

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

Xin Zhang<sup>1,2</sup>, Roberta Noberini<sup>3</sup>, Tiziana Bonaldi<sup>3,4</sup>, Jerome Collemare<sup>2,\*</sup>, † and Michael F. Seidl<sup>1,\*</sup>, †

## Abstract

Chemical modifications of DNA and histone proteins impact the organization of chromatin within the nucleus. Changes in these modifications, catalysed 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 MS 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.

## DATA SUMMARY

The predicted proteomes used in this study are publicly available at the JGI (Joint Genome Institute) MycoCosm repository [1]; the species names and abbreviations are listed in Table S1 (available with the online version of this article). 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 (10.5281/zenodo.6586562). The MS results mentioned in Table S6 are deposited in the PRIDE database with the dataset identifier PXD033478.

Received 31 January 2022; Accepted 31 May 2022; Published 21 September 2022

**Author affiliations:** <sup>1</sup>Theoretical Biology & Bioinformatics Group, Department of Biology, Utrecht University, Padualaan 8, 3584 CH Utrecht, the Netherlands; <sup>2</sup>Westerdijk Fungal Biodiversity Institute, Uppsalalaan 8, 3584 CT Utrecht, the Netherlands; <sup>3</sup>Department of Experimental Oncology, IEO, European Institute of Oncology IRCCS, Via Adamello 16, 20139 Milan, Italy; <sup>4</sup>Department of Oncology and Haematology-Oncology, University of Milano, Via Santa Sofia 9/1, 20122 Milano, Italy.

**\*Correspondence:** Jerome Collemare, j.collemare@wi.knaw.nl; Michael F. Seidl, m.f.seidl@uu.nl

**Keywords:** *Aspergilli*; chromatin; DNA methylation; histone modification; H3K27 methylation; mass spectrometry.

**Abbreviations:** AUC, area under the curve; BUSCO, Benchmarking Universal Single-Copy Ortholog; COMPASS, COMplex of Proteins ASSociated with SET1; DNMT, DNA methyltransferase; GNAT, Gcn5-related acetyltransferase; HAT, histone acetyltransferase; HDAC, histone deacetylase; HDMT, histone demethylase; HMM, Hidden Markov Model; HMT, histone methyltransferase; HSP, High-Scoring Pair; JGI, Joint Genome Institute; MS, Mass spectrometry; PAP, presence-absence pattern; PIC, phenyl isocyanate; PRC2, Polycomb Repressive Complex 2; PTM, post-translational modification; RA, relative abundance; RID, Repeat-Induced Point Mutation (RIP)-Defective; RIP, repeat-induced point mutation; SAGA, Spt-Ada-Gcn5 Acetyltransferase; SH-aLRT, Shimodaira-Hasegawa approximate likelihood ratio test; SM, secondary metabolite; XIC, extracted ion chromatogram. †These authors also contributed equally to this work

**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.

000856 © 2022 The Authors



This is an open-access article distributed under the terms of the Creative Commons Attribution License.

### Impact Statement

Uncovering how fungi regulate gene expression and genome organization is an important step to understand how they control biological processes such as growth or biosynthesis of bioactive molecules. Despite the known importance of chromatin modifications in controlling a plethora of biological processes across eukaryotes, they remain particularly poorly understood in filamentous fungi, even in model organisms such as *Aspergilli*. Our systematic evolutionary analysis provides a robust framework for the functional analysis of chromatin modifications in *Aspergillus* species and other fungi. Our results not only implicate candidate enzymes as playing a role in new chromatin modifications, but they also point at species that have experienced duplications or losses of genes encoding enzymes for specific chromatin modifications. We therefore expect that this work will set the stage for future research into the complexity of the histone code and its functional implications on gene regulation and genome organization 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 [2]. Within the nucleus, DNA is arranged as a chromatin fibre formed by repeated nucleosome units consisting of histone protein octamers (two copies each of H2A, H2B, H3 and H4 histones) that are wrapped by 145–147 bp of DNA [3, 4]. 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 [5, 6]. 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 [7–9]. Chromatin dynamics is therefore key for organisms to adapt gene expression patterns in response to a variety of developmental and environmental clues.

In recent decades, several enzymes involved in orchestrating chromatin dynamics have been identified and functionally characterized in diverse eukaryotes [5, 9, 10]. 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 [11]. Chromatin modifiers that catalyse the same type of chemical modification often harbour functionally and evolutionarily conserved catalytic domains. For instance, DNA methylation (5mC) is catalysed by different DNA methyltransferases (DNMTs) which contain a DNA methyltransferase domain, while histone mono-, di- and tri-methylation are nearly exclusively catalysed by proteins containing a SET [*Su (var) 3–9, E(z) and Trithorax*] domain [10]. 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 [12]. Furthermore, these multi-domain proteins, together with other accessory protein subunits, can form multi-protein complexes to find the correct target region within the genome and modify the exact amino acid residue, collectively contributing to the specific functions of these complexes [13, 14]. 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) [15–18]. Evolutionary analyses have suggested that the key enzymes responsible for chromatin modifications have evolved and significantly expanded early in eukaryotic evolution [19]. 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 [19].

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 [20]. 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 [11]. In addition to ancient chromatin modifiers, lineage-specific innovations such as duplications and losses were also 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 [21–23]. 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 [24]. Chromatin dynamics influences gene expression in several *Aspergilli* and regulates growth, sexual development, secondary metabolite biosynthesis and virulence [11, 25–27]. Thus far, very few well-known chromatin modifications and their catalytic protein complexes have been identified and studied in just a few *Aspergillus* species [11]. For example, gene knockout of the H3K9 methyltransferase *Clr4*

homologue in *Aspergillus fumigatus* reduced histone H3K9me1/3 results in reduction of radial growth and conidia production, as well as delayed conidiation [26]. 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 [28]. The SAGA (Spt-Ada-Gcn5 Acetyltransferase) complex, which is responsible for histone acetylation, has been functionally characterized in *Aspergillus nidulans* [29]. Notably, this complex in *A. nidulans* lacks the Sgf11 and Sus1 subunits that have been described in *S. cerevisiae*, suggesting that this complex evolves in distinct trajectories in different fungi [29]. Besides these few examples, we so far lack a complete view of the occurrence and evolution of chromatin modifiers in the genus *Aspergillus*. 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.

## 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>) [1] on 24 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>). Hmmssearch, which is part of the HMMER package (v3.1b2) [30], 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 [31]. TrimAl v1.2 (-gt 0.1) was used to remove positions in the alignments with gaps in more than 90% of the sequences [32]. Maximum-likelihood phylogenetic trees were reconstructed using IQ-TREE v1.6.10 [33, 34] using ultrafast bootstrap and Shimodaira–Hasegawa approximate likelihood ratio test (SH-aLRT) [35] with 1000 pseudo-replicates, and standard model selection to automatically determine the best-fit model. We predicted additional domains in the analysed proteins using Hmmssearch with all Pfam domain profiles that are present in the Pfam database (v28.0) and visualized them along the phylogenetic trees using iTOL [36]. 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 fungi analysed here, we used BUSCO v4.0.1 (Benchmarking Universal Single-Copy Ortholog assessment tool) [37] to identify the occurrence of 758 fungal single-copy genes (Table S2) within the 109 predicted proteomes (Table S1). The identified homologues 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 [33, 34] 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 [35] with 1000 pseudo-replicates separately.

### Chromatin modifier subunits homology search and gene tree construction

We obtained previously studied chromatin modifier complexes from the 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, homologues in the analysed 109 predicted proteomes were detected based on sequence similarity searches. First, BLASTP (v2.2.31+) [38] 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 the query or subject were retained for the following analyses. If we observed fewer than ten hits for the BLASTP search after filtering in Aspergilli, we used PSI-BLAST (Position-Specific Initiated BLAST v2.2.31+) [38] 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 joined.

To account for apparent protein absences due to erroneous gene annotation, we tested these potential losses by using TBLASTN (BLAST v2.2.31+) [38] to query protein sequences against the genome assemblies. Matches with e-value lower than  $10^{-5}$  together with 4500 nt upstream and downstream were extracted, and Exonerate v2.2.0[39] 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[40] 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 phylogenetic trees, the domain organization was identified and visualized as described above. When the domain organization of an *Aspergillus* protein differed from its orthologues, the corresponding gene model was manually checked in the MycoCosm genome browser [1] and corrected. The presence and absence patterns of 83 subunits from 15 different chromatin modifier complexes were visualized using the Seaborn library in Python [41, 42]. 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 MS

Spores of 4- or 5-day-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°C and 200 rpm for 16 h. The sample was centrifuged at 8000 g for 10 min to discard supernatant, and germinating spores were resuspended in 25 ml 0.8 M NaCl and centrifuged again. Ten millilitres of protoplasting solution (20 mg ml<sup>-1</sup> *Trichoderma* lysing enzyme and 5 mg ml<sup>-1</sup> driselease in 0.8 M NaCl) were used to resuspend spores, which were then incubated at 30°C, 100 rpm for 2–3 h. Protoplasts were retrieved by filtering through sterile Mira cloth and were concentrated by centrifugation at 3000 rpm and 4 °C for 5 min. Histone proteins were extracted according to an adapted histone protein extraction protocol [43]. Briefly, 10<sup>8</sup> protoplasts were resuspended in 1 ml Nuclei Isolation Buffer (1×PBS, 0.5 mM PMSE, 5 μM leupeptin, 5 μM aprotinin and 5 mM Na-butyrate) with 1% Triton X-100, followed by vigorous pipetting through a 200 μl pipette several times. This was then centrifuged at 2300 g for 15 min at 4°C, the supernatant was discarded, and the nuclear pellet resuspended in 200 μl Nuclei Isolation Buffer with 0.1% SDS and 250 U Pierce Universal Nuclease and incubated at 37°C for 30 min to digest nucleic acids. The protein concentration was measured using the Bradford Protein Assay [44].

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) [45]. Chemical acylation of lysines (which occurs on unmodified or mono-methylated residues) impairs trypsin cleavage, resulting in proteolytic cleavage at arginine residues only. This treatment allows us to obtain histone peptides of proper length for MS analysis. The samples were then desalted on handmade StageTips columns [46]. Peptide mixtures were separated by reversed-phase chromatography on an EASY-Spray column (Thermo Fisher Scientific), 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 previously described [45].

The acquired RAW data were analysed 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 four 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 labelling was set as a fixed modification [45]. 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 analysed through multiple parallel MaxQuant jobs [47], 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 quantified, either manually or by using a version of the EpiProfile 2.0 software [48] 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 percent 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 [49]. The AUC values are reported (Table S6) and visualized using GraphPad Prism.



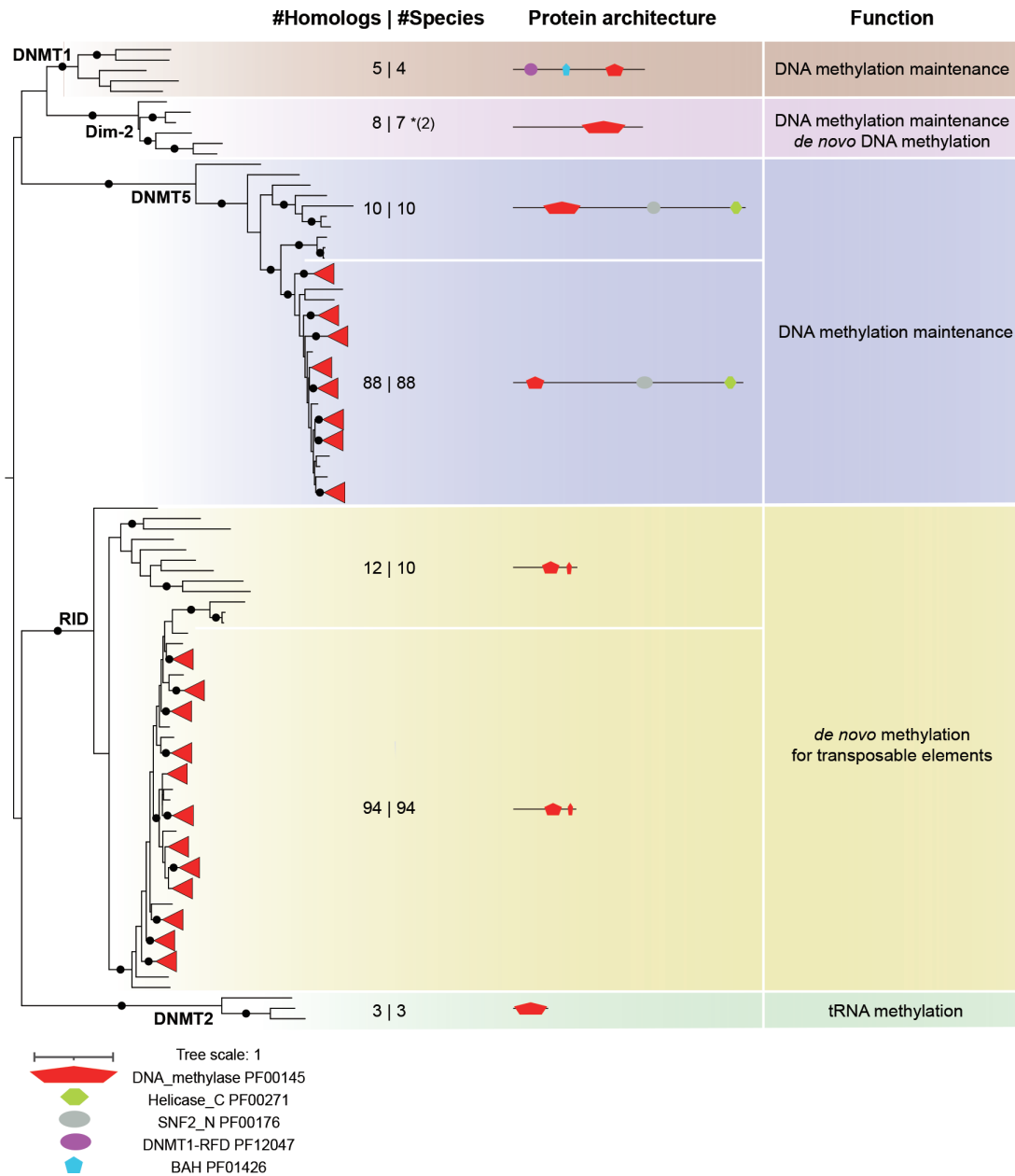
## RESULTS AND DISCUSSION

DNA methylation, histone methylation and histone acetylation are among the most studied chromatin modifications in animals, plants and fungi [50–52]. Protein complexes involved in writing, reading and erasing these modifications have been reported in a few model fungal species to date [16, 22], but the evolution and conservation of these complexes in diverse fungal species have not been largely systematically explored. Besides some research on DNA methyltransferases [21, 23] and SET domain-containing proteins [17] 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 homologues harbouring 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*, *Cremeri*, *Restricti*, *Aspergillus*, *Cervini*, *Clavati*, *Fumigati*, *Candidi*, *Circumdati*, *Janorum*, *Terrei*, *Flavi*, *Ochraceorosei*, *Bispori*, *Cavernicolarum*, *Usti*, *Nidulantes* and *Nigri* [53–59], and 15 other species were used as outgroups (Fig. S1). Two of the outgroup species are Basidiomycota and the remaining species cover each class in the phylum Ascomycota, 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 [60, 61]. We used the identified homologues for each catalytic domain to determine their evolutionary history using maximum-likelihood phylogenies (Figs 1–4, S2 and S3).

### Aspergilli harbour conserved DNA methyltransferases for genome defence and methylation maintenance

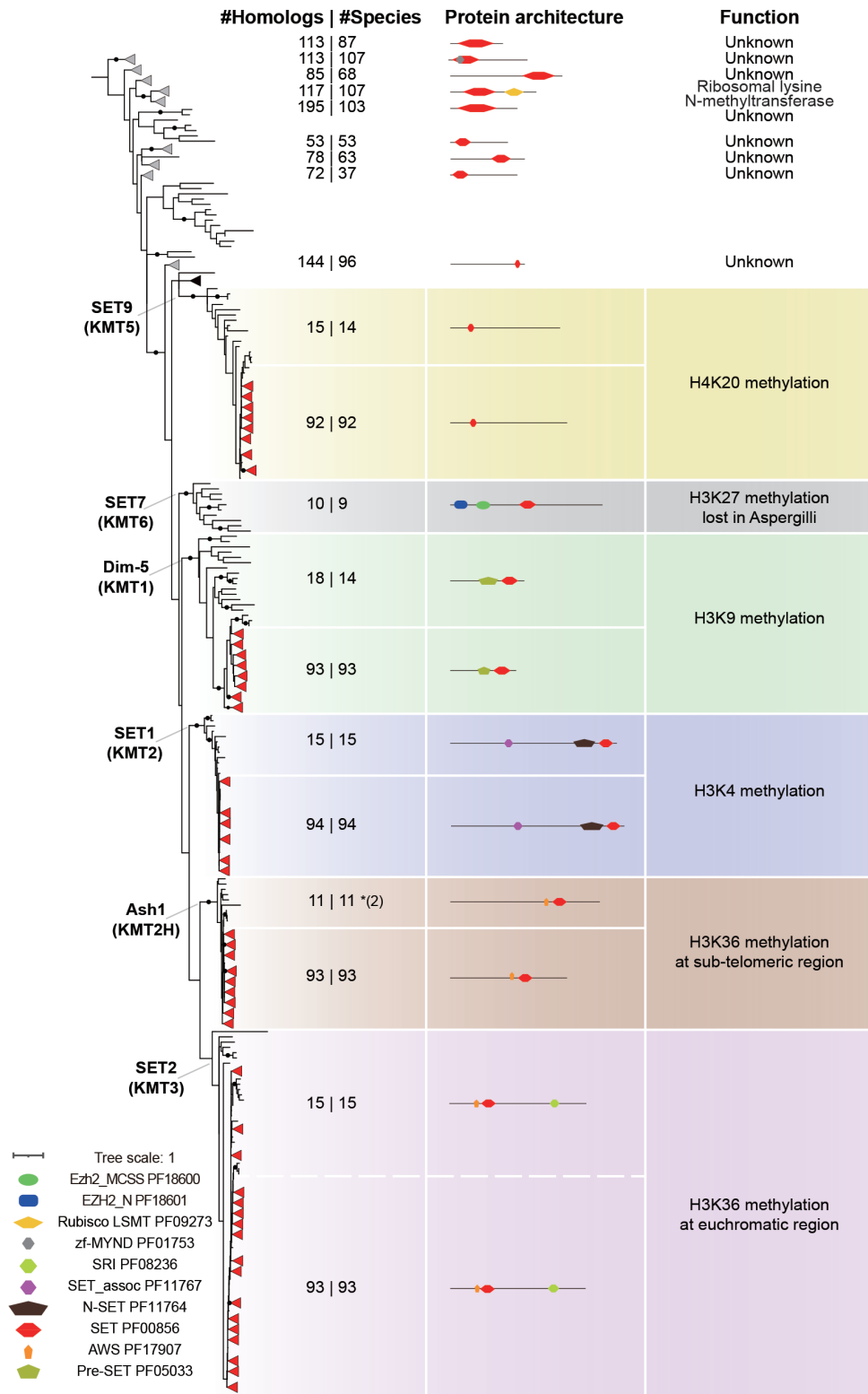
The conserved DNA methyltransferase domain is the only domain known to catalyse the addition of methyl groups to DNA specifically at cytosines (5mC) [21]. Based on previous domain architecture analyses, phylogeny reconstructions and functional experiments, five distinct protein families containing a 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 [21]. Among these five families, DNMT2 seems to function as a tRNA methyltransferase rather than DNA methyltransferase [62]. In line with previous observation [21, 63], 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 catalyse the maintenance of DNA methylation pattern from the parental DNA strand onto synthesized daughter DNA strands, while Dim-2 is also able to catalyse *de novo* methylation [64, 65]. 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 (Fig. 1). Only four outgroup species harbour *DNMT1*, the two Basidiomycota species, *Botrytis cinerea* and *Plectanina melastoma*, which corroborates previous findings [21, 66]. *Dim-2* was identified only in seven Ascomycota outgroup species, and matches in *B. cinerea* and *Zyoseptoria tritici* were included after an additional TBLASTN search in NCBI (Table S6). The original *Dim-2* absence in *B. cinerea* by local Hmsearch 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 [67] and no significant hit was found with Hmsearch or BLASTP searches, but several fragmented hits with TBLASTN searches in NCBI. This loss of *Dim-2* function in *Z. tritici* IPO323 is very recent because other isolates harbour a functional *Dim-2* gene [67]. Its loss leads to increased activity of transposable elements (TEs) and decreased mutation rate, thus influencing genome defence and genome evolution [67]. 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 (Fig. 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), the latter two domains being involved in transcription regulation [68]. The five outgroup species without *DNMT5* are *Saccharomyces cerevisiae*, *Schizosaccharomyces 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 [69–71]. Five out of the six *Aspergillus* species that lost *DNMT5* belong to a single clade within the *Nidulantes* section (*A. undulatus*, *A. olivicola*, *A. filifera*, *A. venezuelensis* and *A. stella-maris*), and the sixth one is *A. robustus* which is assigned to the *Circumdati* section (Fig. S1). In the Basidiomycota fungus *Cryptococcus neoformans*, *DNMT5* maintains DNA methylation using the same cofactors as *DNMT1* (70,72). Thus, it is most likely that *DNMT5* is similarly involved in DNA methylation maintenance in *Aspergilli* as in *Cryptococcus neoformans*. Moreover, DNA methylation maintenance does not appear to be required for the survival of fungi, as shown with the absence of all three of *DNMT1*, *DNMT5* and *Dim-2* not only within the *Nidulantes* section but also in *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* (Fig. 1). However, the absence of these genes could have an impact on the regulation of certain genes such as 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.

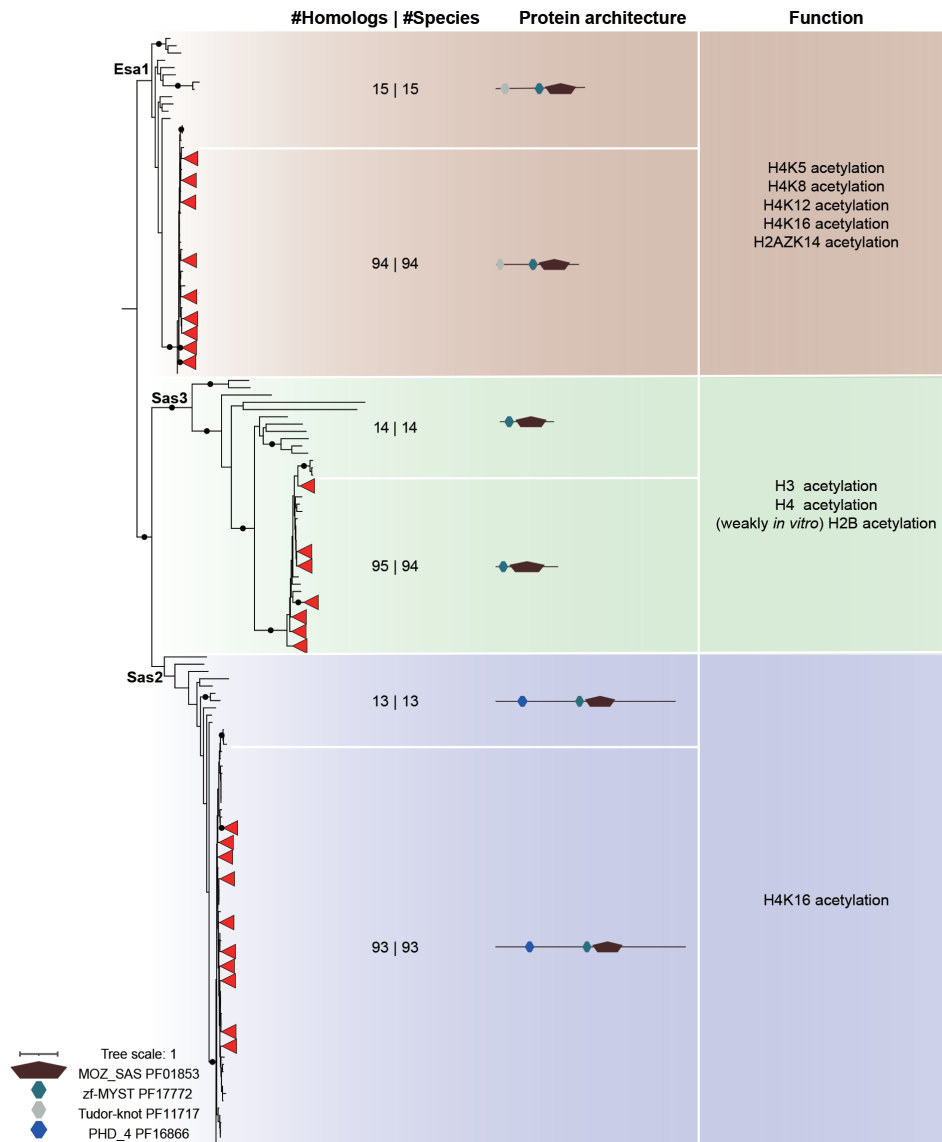


**Fig. 1.** Conservation of DNA methyltransferase families in the genus *Aspergillus* and outgroup species. The maximum-likelihood phylogeny of DNA methyltransferase domain (PF00145) found in 94 *Aspergilli* and 15 outgroup species was determined using IQ-TREE [37, 38]. A black dot on the branch indicates 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 an asterisk 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>).

RID is a more recently acquired single-domain DNMT that specifically occurs in fungi and is involved in a fungal-specific genome defence mechanism called RIP (repeat-induced point mutation) [21]. 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 catalyse 5mC has not yet been demonstrated [72]. Here, we identified RID in all analysed *Aspergilli* species as well as in 10 outgroup species. A RID-like homologue knock-out mutant in the model filamentous Ascomycota *Podospora anserina* shows blocked sexual development, as gamete formation and fertilization can happen but dikaryotic cells cannot develop normally [73]. Interestingly, the knock-out mutant of *DmtA* (homologue of *RID* in *A. flavus*) shows altered conidiation, sclerotial production, aflatoxin biosynthesis and virulence [27], suggesting that RID might have additional functions next to genome defence. Additionally, fungi lost the gene for



**Fig. 2.** Conservation of histone methyltransferase families in the genus *Aspergillus* and outgroup species. The maximum-likelihood phylogeny of SET domains (PF00856) found in 94 *Aspergilli* and 15 outgroup species was determined using IQ-TREE [37, 38]. A black dot on a 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 an asterisk 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>).



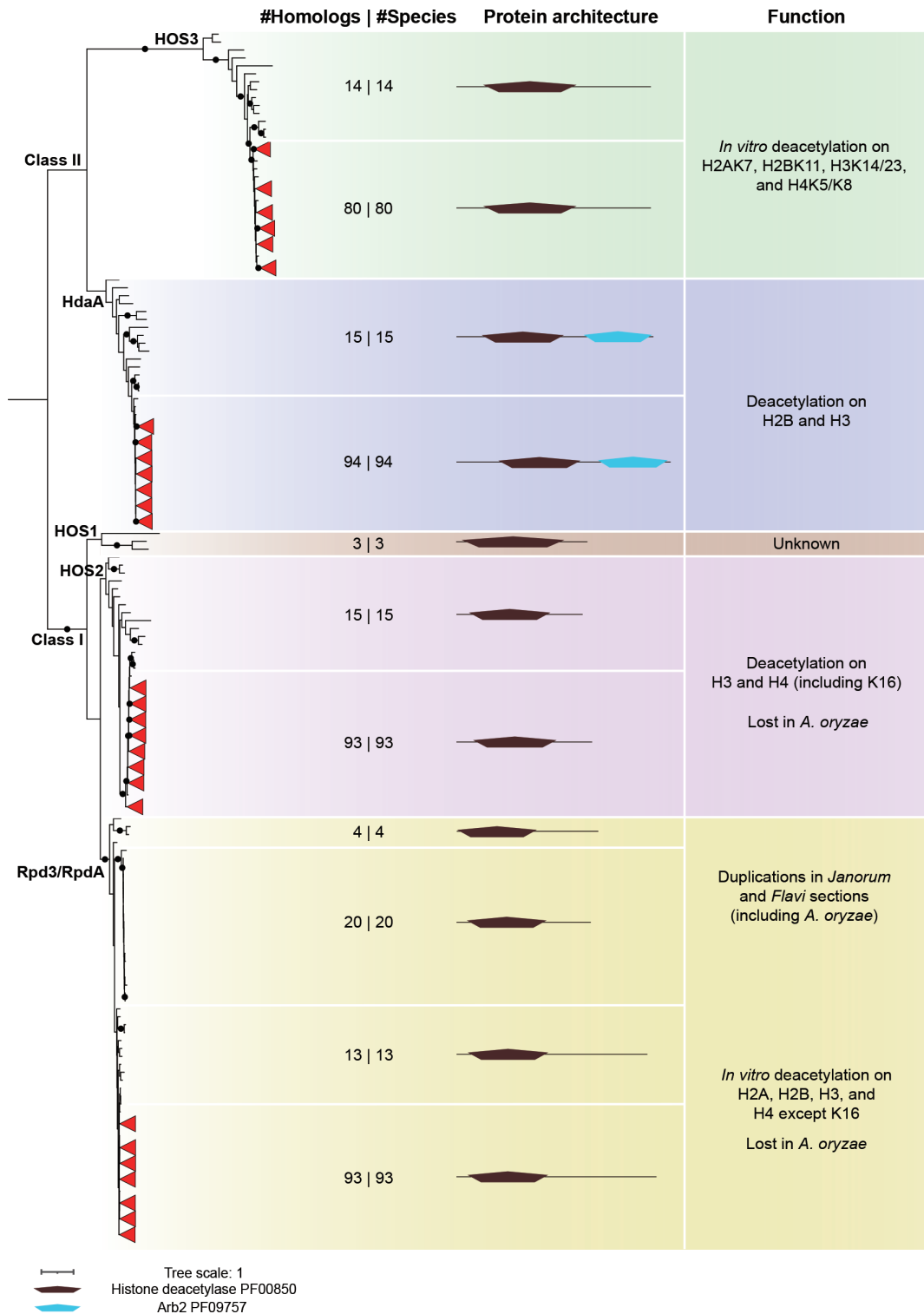
**Fig. 3.** Conservation of MOZ/SAS FAMILY domain-containing histone acetyltransferases in the genus *Aspergillus* genus and outgroup species. The maximum-likelihood phylogeny of MOZ/SAS FAMILY domains (PF01853) found in 94 *Aspergilli* and 15 outgroup species was determined using IQ-TREE [37, 38]. The A black dot on the a 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>).

*de novo* DNA methyltransferase, *DNMT3*, around 50–150 million years ago [74, 75], and only *DNMT5* and *RID* were consistently found in *Aspergilli* (Fig. 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 bisulphite sequencing in *A. flavus* has uncovered very little genome-wide DNA methylation [76]. However, systematic analyses of DNA methylation patterns and experimental evidence for the roles of DNMTs in *Aspergilli* are still largely missing.

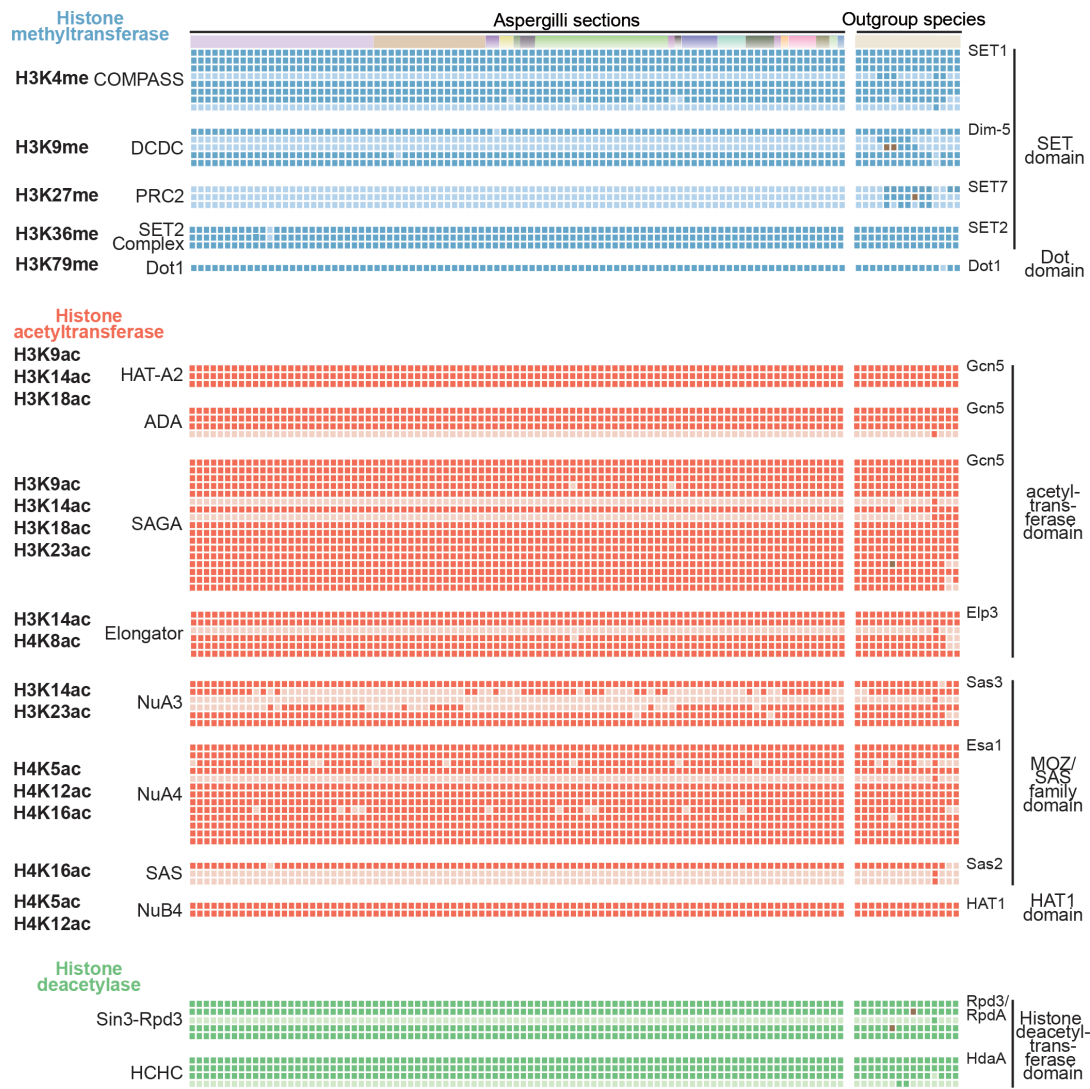
### **Aspergilli lack *SET7* (*KMT6*), but *SET*-domain proteins are abundant and other histone methyltransferases are conserved**

Histone methylation is typically catalysed by SET domain-containing proteins [17, 77]. We identified a total of 1548 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 homologues of the known HMTs *SET9* (*KMT5*), *Dim-5* (*KMT1*), *SET1* (*KMT2*), *Ash1* (*KMT2H*) and *SET2* (*KMT3*), respectively





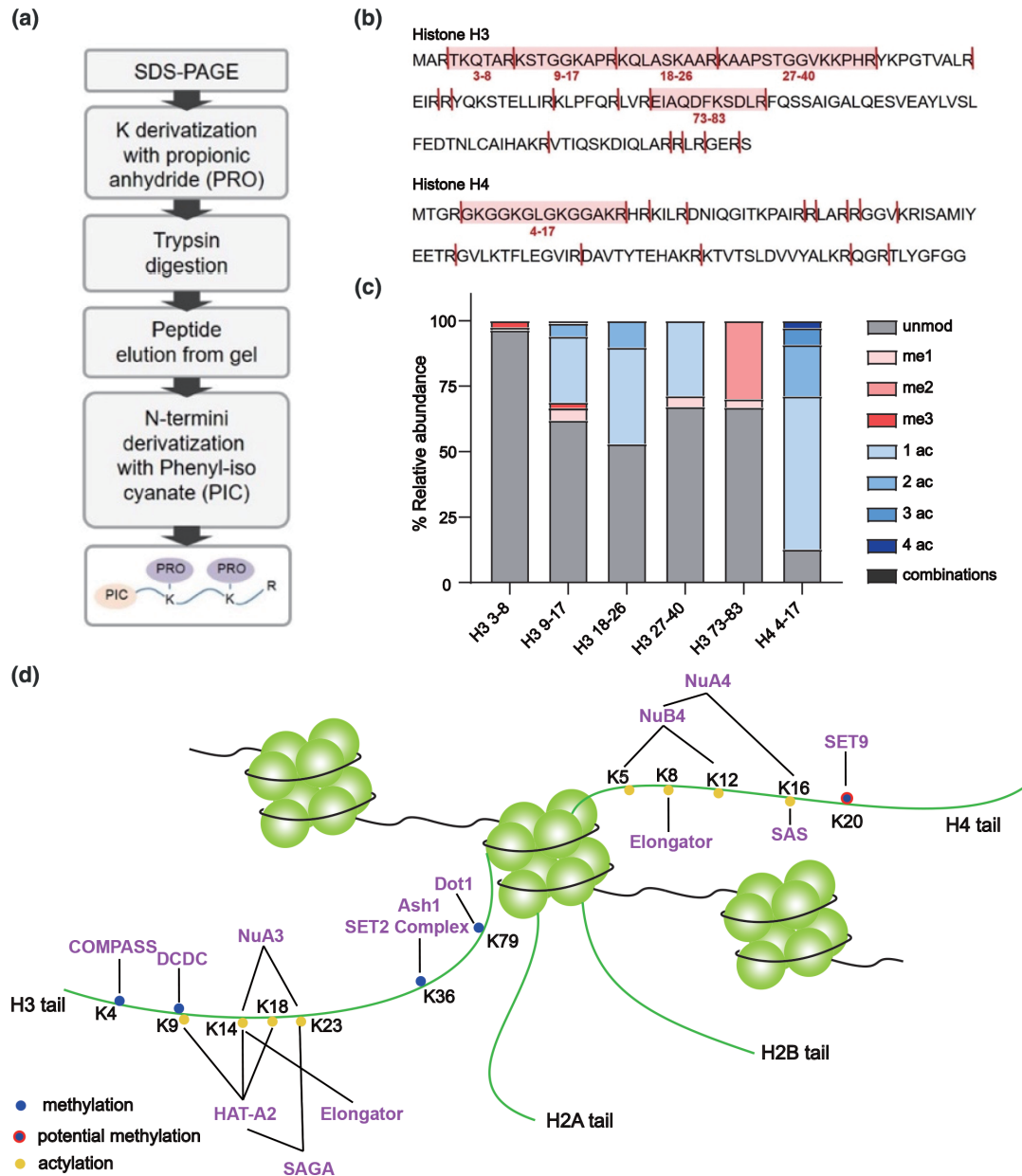
**Fig. 4.** Conservation of histone deacetylase domain-containing proteins in the genus *Aspergillus* and outgroup species. The maximum-likelihood phylogeny of histone deacetylase domains (PF00850) found in 94 *Aspergillus* and 15 outgroup species was determined using IQ-TREE [37, 38]. The black dot on the branch indicates ultrafast bootstrap values over 95 and SH-aLRT values over 80. Leaves from *Aspergillus* 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>).



**Fig. 5.** Presence-absence patterns of subunits that constitute characterized histone modification complexes. The top row with distinct colour 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 (Fig. S1) without *Fusarium oxysporum* and *Aspergillus kawachii* since they were excluded in the latter analysis. Each row represents a subunit in the histone modification complexes; the order agrees with the summary list (Table S5). 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 colour) or absence (light colour) of a subunit (Table S4). The brown boxes indicate additional hits found with TBLASTN searches on NCBI (Table S6).

(Fig. 2). The remaining homologues 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 (Fig. 2). Enzymes in this group contain a SET domain together with a Rubisco LSMT substrate-binding domain (PF09273). While the homologue in plants transfers a methyl group to the large subunit of the Rubisco holoenzyme complex [78, 79], the homologue in *Saccharomyces cerevisiae* has been characterized as a ribosomal lysine N-methyltransferase [80].

The well-described HMTs, SET1, Dim-5, SET2, Ash1 and SET9, all contain distinct accessory domains and catalyse methylation on lysine residues at different positions of H3 or H4 tails [17]. SET1 is composed of the SET domain, the N-SET domain (PF11764) as well as the SET\_assoc domain (PF11767), and is responsible for H3K4me<sub>1/2/3</sub> in *Saccharomyces cerevisiae* [81]. SET1 occurs in a single copy in all species (Fig. 2), indicating that this enzyme and its catalysed 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* [82]. Dim-5 was found in all species included in this study except *Saccharomyces cerevisiae* and *A. calidoustus*, among which *Saccharomyces cerevisiae* is reported to lack H3K9me [82, 83]. How such a loss has



**Fig. 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 S6). (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 phylogenetic analysis are labeled as solid dots with a red outline.

impacted gene expression in *A. calidoustus* in comparison to its closest relative *A. carlsbadensis* deserves further investigation. Both SET2 and Ash1 catalyse H3K36me in *N. crassa* and share the SET domain and the AWS domain (PF17907) (Fig. 2), yet SET2 harbours one additional SRI domain (PF08236) that interacts with RNA polymerase II and links H3K36me with transcript elongation [84–86]. In *N. crassa*, Ash1 was shown to contribute to H3K36me at sub-telomeric regions, while SET2 acts primarily

at euchromatic regions [87, 88]. 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 (Fig. 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 *A. 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 [88]. SET9 homologues harbour only a SET domain, and were shown to catalyse H4K20me in *Schizosaccharomyces pombe* [89]. This group is well conserved in most of the species except in *Saccharomyces cerevisiae* [90] and in the two closely related *Aspergilli*, *A. novoparasiticus* and *A. arachidicola* that both belong to section *Flavi* (Figs 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 catalyse methylation on H3K27, a PTM that is deposited at facultative heterochromatin [16]. Even though PRC2 is considered to be a key chromatin modifier complex, its catalytic subunit SET7 was lost in many fungal species [17]. SET7 has previously been reported to be absent in *A. nidulans* and *A. fumigatus* [15–17], and we here demonstrate that SET7 is absent from all analysed *Aspergilli* as well as from *Penicillium* (Fig. 2), suggesting that this enzyme was probably lost at or before the divergence of the last common ancestor of these two fungal genera. Thus, in line with previous reports [17, 25, 91, 92], our results indicate that methylation on H3K4, H3K9, H3K36 and H4K20 are probably catalysed by conserved enzymes in most *Aspergilli*, while H3K27 methylation is absent. The loss of this methylation is usually reported to trigger serious phenotypes in humans [93] and mice [94]. Loss of H3K27me3 in fungi typically leads to gene expression changes without severe growth defects [16, 95], and thus further investigating how *Aspergilli* and other fungi are able to cope with such a loss will probably 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. In particular, four uncharacterized SET domain-containing groups with 144, 195, 113, and 113 members are conserved in *Aspergilli* and thus probably play important roles in different cellular processes.

### Histone acetyltransferases are conserved in *Aspergilli*

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 [96–98]. We focused here only 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 [96]. The MOZ/SAS FAMILY domain is the catalytic domain of MYST-related HATs [99]. As previously reported [100], the three characterized MYST-related HATs, Sas3, Sas2 and Esa1, belong to distinct monophyletic clades, and these paralogues are nearly fully conserved in the fungi analysed here (Fig. 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* [101]. 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 *Saccharomyces cerevisiae*, only H3K14ac can be detected [102]. Sas3 occurs in most of the species but is lost in *Schizosaccharomyces pombe*, as previously observed [103, 104], and is duplicated in *A. carbonarius* (Fig. 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 (Fig. 3). Esa1 is present as a single copy in all species included in this study, and its functions are known to contribute to the acetylation of K5, K8, K12 and K16 on histone H4, as well as on H2AZK14 in *Saccharomyces cerevisiae* [105–108]. 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 *Saccharomyces cerevisiae* and *Candida albicans* [105, 109, 110]. Although Esa1 and Sas2 catalyse H4K16ac, it was shown in *Candida albicans* that they are differentially recruited at different stages of development [105]. Because both Esa1 and Sas2 are well conserved in the genus *Aspergillus*, we hypothesize that these two HATs may also have distinct roles during development.

GNATs have an acetyltransferase domain as their catalytic core, and 2604 acetyltransferase domain-containing proteins were found in our search (Fig. S2). The genes for two well-characterized GNATs, Gcn5 and Elp3, occur in a single copy in every species included in our study (Fig. S2). Gcn5 harbours a Bromodomain (PF00439) in addition to the Acetyltransferase domain and is a crucial subunit of distinct HAT complexes such as the HAT-A2 complex, the SAGA complex and the ADA complex in *Saccharomyces cerevisiae* [96, 98]. All these complexes contain two coactivator proteins, Ada2 and Ada3, that directly bind to Gcn5 [111] and the recruitment of different other subunits determines the location where they catalyse acetylation [96]. Elp3 is characterized by the Acetyltransferase domain, Radical SAM (PF04055) domain and Radical SAM C-terminal domain (PF16199), and has been shown to catalyse H3K14ac and H4K8ac in *Saccharomyces cerevisiae* [112–114]. Elp3 also exhibits a dual enzymatic activity as it catalyses demethylation in mouse zygotes [115], which is due to the Radical SAM domain that catalyses the demethylation of methylated lysyl residues on histones [116].



Next to Gcn5 and Elp3, we have identified 2386 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 [117] and can also function as a histone crotonyltransferase to regulate gene transcription *in vitro* [118]. It is expected that other less common histone modifications, i.e. histone lysine butyrylation, propionylation or malonylation, may also be catalysed by histone acetyltransferases because these histone modifications all use acyl-Coenzyme A to complete the reactions [119–121]. Further research is needed to determine whether these less common histone modifications are catalysed 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 genus *Aspergillus*, 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 such as butyrylation, succinylation and malonylation.

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

Five families of HDACs can be grouped into two classes based on sequence similarity: class I (HOS1, HOS2 and Rpd3/RpdA) and class II (HdaA and HOS3) [122, 123]. They all harbour the Histone deacetylase catalytic domain, and our phylogenetic tree corroborates the HDAC classification into two distinct classes (Fig. 4). *HdaA* is found as a single copy in each species included in this study (Fig. 4). It harbours the Histone deacetylase domain and C-terminal Arb2 domain (PF09757) that serves as an anchor to target the centromeric heterochromatin region [124, 125]. *HdaA* deacetylates lysine in histones H3 and H2B, but not H4 or H2A in *Saccharomyces cerevisiae* [126, 127], and the  $\Delta hdaA$  knock-out strains show increased or reduced production of different secondary metabolites in *A. nidulans* [128], *A. niger* [128] and *A. fumigatus* [129]. By contrast, *HOS3* is found in only 80 *Aspergillus* species, as it is lost in nine species (*A. nomiae*, *A. pseudocaelatus*, *A. tamarii*, *A. pseudotamarii*, *A. caelatus*, *A. sergii*, *A. novoparasiticus*, *A. minisclerotigenes*, *A. arachidicola*) in the *Flavi* section, three species (*A. desertorum*, *A. stercorarius* and *A. multicolor*) in the *Nidulantes* section, and two species (*A. heteromorphus* and *A. carbonarius*) in the *Nigri* section. *HOS3* is well conserved in all outgroup species except *Schizosaccharomyces pombe* (Fig. 4) [130]. *HOS3* has been shown to catalyse deacetylation with distinct specificity for histone H2AK7, H2BK11, H3K14 and H3K23, as well as H4K5 and H4K8 in *Saccharomyces cerevisiae* cell extracts [131]. Notably, *HOS1* was only detected in three outgroup species, *Saccharomyces cerevisiae* and both Basidiomycota (Fig. 4). The specificity of *Hos1* in catalysing deacetylation on histone proteins is not yet known [132], but it can catalyse the deacetylation of the Smc3 subunit of cohesin and influence the regulation of chromosome segregation during mitosis in *S. cerevisiae* [133]. *HOS2* is found in 15 outgroup species and most *Aspergilli* except *A. oryzae* (Fig. 4). It is known to deacetylate lysine residues in H3 and H4 histone tails [134]. 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 [128].

Within class I HDACs, the Rpd3/RpdA clade is divided into two sub-clades (Fig. 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 (Fig. 4). Our phylogenetic reconstruction suggests a duplication event that has occurred in the ancestor of all *Aspergilli* giving rise to a paralogue that was retained in two sections only and could explain the *A. oryzae* loss in the conserved sub-clade. The two paralogues from the same species vary in length as they share a conserved N-terminal catalytic domain but have a different C terminus (Fig. 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* [135] that is catalysed by *HOS2* [134]. *Saccharomyces cerevisiae* lacking *Rpd3* is sensitive to high osmolarity and shows compromised expression of osmotic stress genes [136]. Deletion of *RpdA* in *A. nidulans* is lethal [137], while only heterokaryon disruptants could be obtained in *A. oryzae* [138], 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 paralogue was retained in *A. oryzae*, the presence of two paralogues in sections *Flavi* and *Janorum* may provide functional redundancy or diversification, which remains to be determined.

Lastly, we reconstructed the phylogeny of histone demethylases harbouring the Jumonji C (JmjC) domain (PF02373) (Fig. 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 [91]. The first five groups occur in most of the outgroup and *Aspergillus* species, but *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 catalyse the demethylation of histone proteins, e.g. KDM5/JARID for H3K4me2/3 and KDM4A/JHDM3/JMJD2 for H3K9me2/3 and H3K36me2/3 [139]. Histone demethylase activity has not been detected for DMM-1 [140] or Epe1 [141] in *N. crassa*, while KDM2 can demethylate H3K36me1/2 in mouse [142] and JMJD6/PKDM11 is an arginine demethylase in *Arabidopsis* [143]. Interestingly, the human KDM6B/JMJD3 is responsible for the demethylation of H3K27me2/3 [144]. 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 catalyse the demethylation of H3K27 [144, 145], we cannot unambiguously correlate the absence of *KDM6B/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 catalysed by the conserved domains found in HMTs, HATs and HDACs (Figs 2–4 and S2), but these chromatin-modifying enzymes often do not act 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 [13, 14]. 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 *Saccharomyces 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* (Fig. 5 and Table S4). Ten subunits show PAPs in both *Aspergilli* and outgroup species, while nine subunits appear specific to *Saccharomyces 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 (Fig. 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 [146]. 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 (Fig. 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* [147]. 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 *A. indicus* may suggest altered H3K9 methylation in this species. Consistent with the absence of *SET7* (Fig. 2), all components of the PRC2 complex are lost in *Aspergilli* (Fig. 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 [72], and cause the redistribution of H3K27me3 [15]. In *Aspergillus* species, although HP1 and the writer for H3K9me3, DCDC complex, are well conserved (Fig. 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 HAT complexes that differ by their distinct catalytic subunits (Fig. 5). The HAT-A2, ADA, SAGA and Elongator complexes use an acetyltransferase domain-containing protein as catalytic subunit [29, 148]. The SAS (Something About Silencing), NuA4 (Nucleosomal Acetyltransferase of histone H4), and NuA3

complexes use a MOZ/SAS FAMILY domain-containing protein [149]. The NuB4 (Nuclear Hat1p-containing type B histone acetyltransferase) complex uses the histone acetyltransferase HAT1 domain-containing protein [150]. 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 catalysing histone acetylation, and other subunits in these complexes guide them to specific locations and link them to the transcriptional machinery [151]. Besides the catalytic module, the ADA complex also harbours the Ahc1 subunit that is found in *Saccharomyces cerevisiae* only [148, 152]. Twelve additional subunits of the SAGA complex are well conserved in *Aspergilli* and outgroup species (Fig. 5). The SAGA complex also contains a deubiquitylase (DUB) module, for which the composition differs between *Aspergilli* and other species. *Sgf11* is unique to *Saccharomyces cerevisiae* and *Sus1* was probably lost in the ancestor of Pezizomycotina because we can only identify *Sus1* orthologues in both Basidiomycota and the yeasts (Fig. 5). The Elongator complex appears more variable as four subunits are mostly conserved, but *Elp6* is specific to *Saccharomyces cerevisiae*, and *Hpa2* and *Hpa3* show PAPs in both *Aspergilli* and outgroup species.

The NuA4, SAS and NuB4 complexes are well conserved in *Aspergilli* as well as in the outgroup species (Fig. 5). These complexes appear slightly different in *Saccharomyces 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 catalyse H4K16ac with opposite effects during cell growth and morphogenesis in *Candida albicans* [105]. 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 *Saccharomyces cerevisiae*. *Pdp3* and *Nto1* seem to be lost in a complementary manner, suggesting a possible functional redundancy (Fig. 5). Finally, two histone deacetylase complexes, Sin3-Rpd3 and HCHC, are fully conserved (Fig. 5). The Sin3-Rpd3 complex is slightly different in *Saccharomyces cerevisiae* with the specific *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 MS 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 (Figs 2–5), we extracted histone proteins from *A. nidulans* and performed unbiased MS analyses (Fig. 6a, b). H3 and H4 peptides with PTMs corresponding to H3K4me, H3K9me, H3K36me and H3K79me were detected (Fig. 6c), which we anticipated being catalysed by SET1, Dim-5, Ash1 and SET2 and Dot1, respectively. H4K20me, which is catalysed by SET9 in *Schizosaccharomyces pombe* [89], was not detected (Fig. 6c), probably 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 (Fig. 6c). We did not observe any peptide that would indicate the presence of H3K27me in *A. nidulans*, while we did detect H3K27ac (Fig. 6c and Table S6). 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* [17, 153, 154]. Thus, the MS 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.

## CONCLUSION

Changes in chromatin organization play an important role in regulating the expression of genes [130]. The fungal genus *Aspergillus* is well known for its ability to thrive in a broad range of different environmental niches and the huge potential for secondary metabolite biosynthesis [11, 155]. 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 MS. 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 such as histone lysine succinylation, crotonylation, butyrylation, propionylation and malonylation, are all probably catalysed by HATs and HDACs as they all use Coenzyme A (CoA) to complete the reactions [119–121], but may require specific, yet to be discovered subunits. In conclusion, our study provides a detailed overview of the evolutionary routes of the 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*.

### 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.

### Acknowledgement

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. In particular, we thank 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 the *A. nidulans* strain. We thank Alessandro Vai for support with analysis of MS data for histone PTM identification on *Aspergillus* core histone sequences.

### Author 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.

### Conflicts of interest

The authors declare that there are no conflicts of interest.

### References

1. Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, et al. MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res* 2014;42:D699–704.
2. Labrador M, Corces VG. Setting the boundaries of chromatin domains and nuclear organization. *Cell* 2002;111:151–154.
3. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 1997;389:251–260.
4. Luger K. Structure and dynamic behavior of nucleosomes. *Curr Opin Genet Dev* 2003;13:127–135.
5. Jenuwein T, Allis CD. Translating the histone code. *Science* 2001;293:1074–1080.
6. Peterson CL, Lanier M-A. Histones and histone modifications. *Curr Biol* 2004;14:R546–51.
7. Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, et al. Active genes are tri-methylated at K4 of histone H3. *Nature* 2002;419:407–411.

8. Grewal SIS, Jia S. Heterochromatin revisited. *Nat Rev Genet* 2007;8:35–46.
9. Berger SL. The complex language of chromatin regulation during transcription. *Nature* 2007;447:407–412.
10. Xiao B, Wilson JR, Gamblin SJ. SET domains and histone methylation. *Curr Opin Struct Biol* 2003;13:699–705.
11. Collemare J, Seidl MF. Chromatin-dependent regulation of secondary metabolite biosynthesis in fungi: is the picture complete? *FEMS Microbiol Rev* 2019;43:591–607.
12. Helmlinger D, Papai G, Devys D, Tora L. What do the structures of GCN5-containing complexes teach us about their function? *Biochim Biophys Acta Gene Regul Mech* 2021;1864:194614.
13. Hajheidari M, Koncz C, Bucher M. Chromatin evolution—key innovations underpinning Morphological complexity. *Front Plant Sci* 2019;10:1–12. ;
14. Gates LA, Foulds CE, O'Malley BW. Histone Marks in the 'Driver's Seat': functional roles in steering the transcription cycle. *Trends Biochem Sci* 2017;42:977–989.
15. Jamieson K, Wiles ET, McNaught KJ, Sidoli S, Leggett N, et al. Loss of HP1 causes depletion of H3K27me3 from facultative heterochromatin and gain of H3K27me2 at constitutive heterochromatin. *Genome Res* 2016;26:97–107.
16. Jamieson K, Rountree MR, Lewis ZA, Stajich JE, Selker EU. Regional control of histone H3 lysine 27 methylation in *Neurospora*. *Proc Natl Acad Sci U S A* 2013;110:6027–6032.
17. Freitag M. Histone methylation by SET domain proteins in Fungi. *Annu Rev Microbiol* 2017;71:413–439.
18. Kamei M, Ameri AJ, Ferraro AR, Bar-Peled Y, Zhao F, et al. Imitation switch is required for normal chromatin structure and gene repression in PRC2 target domains. *Proc Natl Acad Sci U S A* 2021;118:e2010003118.
19. Weiner AKM, Cerón-Romero MA, Yan Y, Katz LA. Phylogenomics of the Epigenetic toolkit reveals punctate retention of genes across *Eukaryotes*. *Genome Biol Evol* 2020;12:2196–2210.
20. Choi J, Kim S-H. A genome tree of life for the fungi kingdom. *Proc Natl Acad Sci U S A* 2017;114:9391–9396.
21. Bewick AJ, Hofmeister BT, Powers RA, Mondo SJ, Grigoriev IV, et al. Diversity of cytosine methylation across the fungal tree of life. *Nat Ecol Evol* 2019;3:479–490.
22. Kouzminova E, Selker EU. Dim-2 encodes a DNA methyltransferase responsible for all known cytosine methylation in *Neurospora*. *EMBO J* 2001;20:4309–4323.
23. Nai YS, Huang YC, Yen MR, Chen PY. Diversity of fungal DNA Methyltransferases and their association with DNA Methylation patterns. *Front Microbiol* 2020;11:616922.
24. Samson RA, Visagie CM, Houbbraken J, Hong S-B, Hubka V, et al. Phylogeny, identification and nomenclature of the genus *Aspergillus*. *Stud Mycol* 2014;78:141–173.
25. Cánovas D, Marcos AT, Gacek A, Ramos MS, Gutiérrez G, et al. The histone acetyltransferase GcnE (GCN5) plays a central role in the regulation of *Aspergillus* asexual development. *Genetics* 2014;197:1175–1189.
26. Palmer JM, Perrin RM, Dagenais TRT, Keller NP. H3K9 methylation regulates growth and development in *Aspergillus fumigatus*. *Eukaryot Cell* 2008;7:2052–2060.
27. Yang K, Liang L, Ran F, Liu Y, Li Z, et al. The DmtA methyltransferase contributes to *Aspergillus flavus* conidiation, sclerotial production, aflatoxin biosynthesis and virulence. *Sci Rep* 2016;6:1–13.
28. Palmer JM, Bok JW, Lee S, Dagenais TRT, Andes DR, et al. 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.
29. Georgakopoulos P, Lockington RA, Kelly JM. The Spt-Ada-Gcn5 Acetyltransferase (SAGA) complex in *Aspergillus nidulans*. *PLoS ONE* 2013;8:e65221.
30. Eddy SR. Accelerated profile HMM searches. *PLoS Comput Biol* 2011;7:e1002195.
31. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 2013;30:772–780.
32. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. TrimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 2009;25:1972–1973.
33. Chernomor O, von Haeseler A, Minh BQ. Terrace aware data structure for Phylogenomic inference from supermatrices. *Syst Biol* 2016;65:997–1008.
34. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015;32:268–274.
35. Anisimova M, Gil M, Dufayard J-F, Dessimoz C, Gascuel O. Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. *Syst Biol* 2011;60:685–699.
36. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 2019;47:W256–W259.
37. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 2015;31:3210–3212.
38. Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, et al. BLAST: a more efficient report with usability improvements. *Nucleic Acids Res* 2013;41:W29–33.
39. Slater GSC, Birney E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* 2005;6:31.
40. Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* 2000;16:276–277.
41. Hunter JD. Matplotlib: a 2D graphics environment. *Comput Sci Eng* 2011;9:90–95.
42. Waskom ML. Seaborn: statistical data visualization. *JOSS* 2011;6:3021.
43. Noberini R, Restellini C, Savoia EO, Bonaldi T. Enrichment of histones from patient samples for mass spectrometry-based analysis of post-translational modifications. *Methods* 2020;184:19–28.
44. He F. Bradford protein assay. *BIO Protoc* 2011;1.
45. Noberini R, Savoia EO, Brandini S, Greco F, Marra F, et al. Spatial epi-proteomics enabled by histone post-translational modification analysis from low-abundance clinical samples. *Clin Epigenetics* 2021;13:1–16.
46. Prudhomme N, Gianetto-Hill C, Pastora R, Cheung W-F, Allen-Vercoe E, et al. Quantitative proteomic profiling of shake flask versus bioreactor growth reveals distinct responses of *Agrobacterium tumefaciens* for preparation in molecular pharming. *Can J Microbiol* 2021;67:75–84.
47. Bremang M, Cuomo A, Agresta AM, Stugiewicz M, Spadotto V, et al. Mass spectrometry-based identification and characterisation of lysine and arginine methylation in the human proteome. *Mol Biosyst* 2013;9:2231–2247.
48. Yuan Z-F, Sidoli S, Marchione DM, Simithy J, Janssen KA, et al. EpiProfile 2.0: a computational platform for processing Epi-Proteomics mass spectrometry data. *J Proteome Res* 2018;17:2533–2541.
49. Pesavento JJ, Mizzen CA, Kelleher NL. Quantitative analysis of modified proteins and their positional isomers by tandem mass spectrometry: human histone H4. *Anal Chem* 2006;78:4271–4280.
50. Wade PA, Pruss D, Wolffe AP. Histone acetylation: chromatin in action. *Trends Biochem Sci* 1997;22:128–132.
51. Binda O. On your histone mark, SET, methylate! *Epigenetics* 2013;8:457–463.
52. Wolffe AP, Matzke MA. Epigenetics: regulation through repression. *Science* 1999;286:481–486.



53. James TY, Stajich JE, Hittinger CT, Rokas A. Toward a fully resolved fungal tree of life. *Annu Rev Microbiol* 2020;74:291–313.
54. Sun BD, Houbraken J, Frisvad JC, Jiang XZ, Chen AJ, et al. New species in *Aspergillus* section Usti and an overview of *Aspergillus* section *Cavernicularum*. *Int J Syst Evol Microbiol* 2020;70:5401–5416. ;
55. Chen AJ, Frisvad JC, Sun BD, Varga J, Kocsubé S, et al. *Aspergillus* section *Nidulantes* (formerly *Emericella*): polyphasic taxonomy, chemistry and biology. *Stud Mycol* 2016;84:1–118.
56. Steenwyk JL, Shen X-X, Lind AL, Goldman GH, Rokas A. A robust phylogenomic time tree for biotechnologically and medically important fungi in the genera *Aspergillus* and *Penicillium*. *mBio* 2019;10:1–25.
57. Vesth TC, Nybo JL, Theobald S, Frisvad JC, Larsen TO, et al. Investigation of inter- and intraspecies variation through genome sequencing of *Aspergillus* section Nigri. *Nat Genet* 2018;50:1688–1695.
58. Kjærboelling I, Vesth T, Frisvad JC, Nybo JL, Theobald S, et al. A comparative genomics study of 23 *Aspergillus* species from section Flavi. *Nat Commun* 2020;11:1106.
59. Souza SC, Pereira VM, Moreira SI, Costa SS, Moreira GM, et al. *Aspergillus trisporus*: a new jani section species from Brazilian soil. *CREAM* 2019;9:175–186.
60. Grunstein M, Gasser SM. Epigenetics in *Saccharomyces cerevisiae*. *Cold Spring Harb Perspect Biol* 2013;5:a017491.
61. Nielsen JC, Grijseels S, Prigent S, Ji B, Dainat J, et al. Global analysis of biosynthetic gene clusters reveals vast potential of secondary metabolite production in *Penicillium* species. *Nat Microbiol* 2017;2:April.
62. Ko Y-H, So K-K, Chun J, Kim D-H, Kronstad JW. Distinct roles of two DNA Methyltransferases from *Cryphonectria parasitica* in fungal virulence, responses to hypovirus infection, and viral clearance. *mBio* 2017;12:1–16. ;
63. Jurkowski TP, Jeltsch A, Lyko F. On the evolutionary origin of Eukaryotic DNA Methyltransferases and Dnmt2. *PLoS ONE* 2017;6:e28104.
64. Bestor T. Structure of mammalian DNA methyltransferase as deduced from the inferred amino acid sequence and direct studies of the protein. *Biochemical Society Transactions* 2017;16:944–947.
65. Okano M, Xie S, Li E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat Genet* 1998;19:219–220.
66. Zhang Z, He C, Chen Y, Li B, Tian S. DNA Methyltransferases regulate pathogenicity of *Botrytis cinerea* to horticultural crops. *J Fungi (Basel)* 2021;7:659.
67. Möller M, Habig M, Lorrain C, Feurtey A, Hauelsen J, et al. Recent loss of the Dim2 DNA methyltransferase decreases mutation rate in repeats and changes evolutionary trajectory in a fungal pathogen. *PLoS Genet* 2021;17:1–27.
68. Delmas V, Stokes DG, Perry RP. A mammalian DNA-binding protein that contains a chromodomain and an SNF2/SWI2-like helicase domain. *Proc Natl Acad Sci U S A* 1993;90:2414–2418.
69. Capuano F, Müllereder M, Kok R, Blom HJ, Ralser M. Cytosine DNA methylation is found in *Drosophila melanogaster* but absent in *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and other yeast species. *Anal Chem* 2014;86:3697–3702.
70. Huff JT, Zilberman D. Dnmt1-independent CG methylation contributes to nucleosome positioning in diverse eukaryotes. *Cell* 2014;156:1286–1297.
71. Proffitt JH, Davie JR, Swinton D, Hattman S. 5-Methylcytosine is not detectable in *Saccharomyces cerevisiae* DNA. *Mol Cell Biol* 1984;4:985–988.
72. Rountree MR, Selker EU. DNA methylation and the formation of heterochromatin in *Neurospora crassa*. *Heredity (Edinb)* 2010;105:38–44.
73. Grognet P, Timpano H, Carlier F, Ait-Benkhalil J, Berteaux-Lecellier V, et al. A RID-like putative cytosine methyltransferase homologue controls sexual development in the fungus *Podospora anserina*. *PLoS Genet* 2019;15:e1008086.
74. Catania S, Dumesic PA, Pimentel H, Nasif A, Stoddard CI, et al. Evolutionary persistence of DNA Methylation for millions of years after ancient loss of a *De Novo* Methyltransferase. *Cell* 2020;180:263–277.
75. Madhani HD. Unbelievable but true: epigenetics and chromatin in fungi. *Trends Genet* 2021;37:12–20.
76. Liu S-Y, Lin J-Q, Wu H-L, Wang C-C, Huang S-J, et al. Bisulfite sequencing reveals that *Aspergillus flavus* holds a hollow in DNA Methylation. *PLoS ONE* 2012;7:e30349.
77. Spellmon N, Holcomb J, Trescott L, Sirinupong N, Yang Z. Structure and function of SET and MYND domain-containing proteins. *Int J Mol Sci* 2015;16:1406–1428.
78. Raunser S, Magnani R, Huang Z, Houtz RL, Trievel RC, et al. Rubisco in complex with Rubisco large subunit methyltransferase. *Proc Natl Acad Sci U S A* 2009;106:3160–3165.
79. Trievel RC, Flynn EM, Houtz RL, Hurley JH. Mechanism of multiple lysine methylation by the SET domain enzyme Rubisco LSM1. *Nat Struct Biol* 2003;10:545–552.
80. Porras Y, Whitelegge JP, Clarke S. Yeast ribosomal/cytochrome c SET domain methyltransferase subfamily: identification of Rpl23ab methylation sites and recognition motifs. *J Biol Chem* 2007;282:12368–12376.
81. Roguev A, Schaft D, Shevchenko A, Pijnappel WW, Wilms M, et al. The *Saccharomyces cerevisiae* Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *EMBO J* 2001;20:7137–7148.
82. Zhang X, Tamaru H, Khan SI, Horton JR, Keefe LJ, et al. Structure of the *Neurospora* SET domain protein DIM-5, a Histone H3 Lysine Methyltransferase. *Cell* 2002;111:117–127.
83. Marina DB, Shankar S, Natarajan P, Finn KJ, Madhani HD. A conserved ncRNA-binding protein recruits silencing factors to heterochromatin through an RNAi-independent mechanism. *Genes Dev* 2013;27:1851–1856.
84. Strahl BD, Grant PA, Briggs SD, Sun Z-W, Bone JR, et al. Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Mol Cell Biol* 2002;22:1298–1306.
85. Xiao T, Hall H, Kizer KO, Shibata Y, Hall MC, et al. Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. *Genes Dev* 2003;17:654–663.
86. Li M, Phatnani HP, Guan Z, Sage H, Greenleaf AL, et al. Solution structure of the Set2-Rpb1 interacting domain of human Set2 and its interaction with the hyperphosphorylated C-terminal domain of Rpb1. *Proc Natl Acad Sci U S A* 2005;102:17636–17641.
87. Bicocca VT, Ormsby T, Adhvariy KK, Honda S, Selker EU. ASH1-catalyzed H3K36 methylation drives gene repression and marks H3K27me2/3-competent chromatin. *Elife* 2018;7:1–19.
88. Janevska S, Baumann L, Sieber CMK, Münsterkötter M, Ulrich J, et al. 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* 2018;208:153–171.
89. Wang Y, Reddy B, Thompson J, Wang H, Noma K, et al. Regulation of Set9-mediated H4K20 methylation by a PWWP domain protein. *Mol Cell* 2009;33:428–437.
90. Veerappan CS, Avramova Z, Moriyama EN. Evolution of SET-domain protein families in the unicellular and multicellular *Ascomycota* fungi. *BMC Evol Biol* 2008;8:1–20.
91. Gacek-Matthews A, Noble LM, Gruber C, Berger H, Sulyok M, et al. KdmA, a histone H3 demethylase with bipartite function, differentially regulates primary and secondary metabolism in *Aspergillus nidulans*. *Mol Microbiol* 2015;96:839–860.
92. Bachleitner S, Sørensen JL, Gacek-Matthews A, Sulyok M, Studt L, et al. Evidence of a Demethylase-independent role for

- the H3K4-specific Histone Demethylases in *Aspergillus nidulans* and *Fusarium graminearum* secondary metabolism. *Front Microbiol* 2019;10:1759.
93. Sneeringer CJ, Scott MP, Kuntz KW, Knutson SK, Pollock RM, et al. Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. *Proc Natl Acad Sci U S A* 2010;107:20980–20985.
  94. Tao Y, Neppl RL, Huang Z-P, Chen J, Tang R-H, et al. The histone methyltransferase Set7/9 promotes myoblast differentiation and myofibril assembly. *J Cell Biol* 2011;194:551–565.
  95. Kramer HM, Seidl MF, Thomma B, Cook DE. Local rather than Global H3K27me3 dynamics are associated with differential gene expression in *Verticillium dahliae*. *mBio* 2022:e0356621.
  96. Carrozza MJ, Utley RT, Workman JL, Côté J. The diverse functions of histone acetyltransferase complexes. *Trends Genet* 2003;19:321–329.
  97. Torchia J, Glass C, Rosenfeld MG. Co-activators and co-repressors in the integration of transcriptional responses. *Curr Opin Cell Biol* 1998;10:373–383.
  98. Sterner DE, Berger SL. Acetylation of histones and transcription-related factors. *Microbiol Mol Biol Rev* 2000;64:435–459.
  99. Roth SY, Denu JM, Allis CD. Histone acetyltransferases. *Annu Rev Biochem* 2001;70:81–120.
  100. Kawahara T, Siegel TN, Ingram AK, Alsford S, Cross GAM, et al. Two essential MYST-family proteins display distinct roles in histone H4K10 acetylation and telomeric silencing in trypanosomes. *Mol Microbiol* 2008;69:1054–1068.
  101. Takechi S, Nakayama T. Sas3 is a histone acetyltransferase and requires a zinc finger motif. *Biochem Biophys Res Commun* 1999;266:405–410.
  102. Taverna SD, Ilin S, Rogers RS, Tanny JC, Lavender H, et al. 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. *Mol Cell* 2006;24:785–796.
  103. Gómez EB, Espinosa JM, Forsburg SL. Schizosaccharomyces pombe mst2+ encodes a MYST family histone acetyltransferase that negatively regulates telomere silencing. *Mol Cell Biol* 2005;25:8887–8903.
  104. Nugent RL, Johnsson A, Fleharty B, Gogol M, Xue-Franzén Y, et al. Expression profiling of *S. pombe* acetyltransferase mutants identifies redundant pathways of gene regulation. *BMC Genomics* 2010;11:59.
  105. Kim J, Park S, Lee JS. Epigenetic control of oxidative stresses by Histone Acetyltransferases in *Candida albicans*. *J Microbiol Biotechnol* 2018;28:181–189.
  106. Millar CB, Xu F, Zhang K, Grunstein M. Acetylation of H2AZ Lys 14 is associated with genome-wide gene activity in yeast. *Genes Dev* 2006;20:711–722.
  107. Searle NE, Torres-Machorro AL, Pillus L. Chromatin regulation by the NuA4 Acetyltransferase complex is mediated by essential interactions between enhancer of Polycomb (Epl1) and Esa1. *Genetics* 2017;205:1125–1137.
  108. Allard S, Utley RT, Savard J, Clarke A, Grant P, et al. NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p. *EMBO J* 1999;18:5108–5119.
  109. Wang X, Chang P, Ding J, Chen J. Distinct and redundant roles of the two MYST histone acetyltransferases Esa1 and Sas2 in cell growth and morphogenesis of *Candida albicans*. *Eukaryot Cell* 2013;12:438–449.
  110. Meijnsing SH, Ehrenhofer-Murray AE. The silencing complex SAS-I links histone acetylation to the assembly of repressed chromatin by CAF-I and Asf1 in *Saccharomyces cerevisiae*. *Genes Dev* 2001;15:3169–3182.
  111. Balasubramanian R, Pray-Grant MG, Selleck W, Grant PA, Tan S. Role of the Ada2 and Ada3 transcriptional coactivators in histone acetylation. *J Biol Chem* 2002;277:7989–7995.
  112. Winkler GS, Kristjuhan A, Erdjument-Bromage H, Tempst P, Svejstrup JQ. Elongator is a histone H3 and H4 acetyltransferase important for normal histone acetylation levels *in vivo*. *Proc Natl Acad Sci U S A* 2002;99:3517–3522.
  113. Krogan NJ, Greenblatt JF. Characterization of a six-subunit holo-elongator complex required for the regulated expression of a group of genes in *Saccharomyces cerevisiae*. *Mol Cell Biol* 2001;21:8203–8212.
  114. Wittschieben BO, Otero G, de Bizemont T, Fellows J, Erdjument-Bromage H, et al. A novel histone acetyltransferase is an integral subunit of elongating RNA polymerase II holoenzyme. *Mol Cell* 1999;4:123–128.
  115. Okada Y, Yamagata K, Hong K, Wakayama T, Zhang Y. A role for the elongator complex in zygotic paternal genome demethylation. *Nature* 2010;463:554–558.
  116. Booker SJ, Grove TL. Mechanistic and functional versatility of radical SAM enzymes. *F1000 Biol Rep* 2010;2:52.
  117. Yuan Y, Yuan H, Yang G, Yun H, Zhao M, et al. IFN- $\alpha$  confers epigenetic regulation of HBV cccDNA minichromosome by modulating GCN5-mediated succinylation of histone H3K79 to clear HBV cccDNA. *Clin Epigenetics* 2020;12:1–16.
  118. Kollenstart L, de Groot AJL, Janssen GMC, Cheng X, Vreeken K, et al. Gcn5 and Esa1 function as histone crotonyltransferases to regulate crotonylation-dependent transcription. *J Biol Chem* 2019;294:20122–20134.
  119. Jones R. Identification and verification of lysine propionylation and butyrylation in yeast core histones using ptmap software. *Bone* 2014;23:1–7.
  120. Liu B, Lin Y, Darwanto A, Song X, Xu G, et al. Identification and characterization of propionylation at histone H3 lysine 23 in mammalian cells. *J Biol Chem* 2009;284:32288–32295.
  121. Papanicolaou KN, O'Rourke B, Foster DB. Metabolism leaves its mark on the powerhouse: recent progress in post-translational modifications of lysine in mitochondria. *Front Physiol* 2014;5:301.
  122. Ekwall K. Genome-wide analysis of HDAC function. *Trends Genet* 2005;21:608–615.
  123. Bernstein BE, Tong JK, Schreiber SL. Genomewide studies of histone deacetylase function in yeast. *Proc Natl Acad Sci U S A* 2000;97:13708–13713.
  124. Honda S, Bicocca VT, Gessaman JD, Rountree MR, Yokoyama A, et al. Dual chromatin recognition by the histone deacetylase complex HCHC is required for proper DNA methylation in *Neurospora crassa*. *Proc Natl Acad Sci U S A* 2016;113:E6135–E6144.
  125. Job G, Brugger C, Xu T, Lowe BR, Pfister Y, et al. SHREC Silences Heterochromatin via Distinct Remodeling and Deacetylation Modules. *Mol Cell* 2016;62:207–221.
  126. Wu J, Suka N, Carlson M, Grunstein M. TUP1 utilizes histone H3/H2B-specific HDA1 deacetylase to repress gene activity in yeast. *Mol Cell* 2001;7:117–126.
  127. Kurdistani SK, Grunstein M. Histone acetylation and deacetylation in yeast. *Nat Rev Mol Cell Biol* 2003;4:276–284.
  128. Li X, Pan L, Wang B, Pan L. The Histone Deacetylases HosA and HdaA affect the phenotype and transcriptomic and metabolic profiles of *Aspergillus niger*. *Toxins (Basel)* 2019;11:E520.
  129. Lee I, Oh J-H, Shwab EK, Dagenais TRT, Andes D, et al. HdaA, a class 2 histone deacetylase of *Aspergillus fumigatus*, affects germination and secondary metabolite production. *Fungal Genet Biol* 2009;46:782–790.
  130. Brosch G, Loidl P, Graessle S. Histone modifications and chromatin dynamics: a focus on filamentous fungi. *FEMS Microbiol Rev* 2008;32:409–439.
  131. Carmen AA, Griffin PR, Calaycay JR, Rundlett SE, Suka Y, et al. Yeast HOS3 forms a novel trichostatin A-insensitive homodimer



- with intrinsic histone deacetylase activity. *Proc Natl Acad Sci U S A* 1999;96:12356–12361.
132. Pérez-Martínez ME, Benet M, Alepuz P, Tordera V. Nut1/Hos1 and Sas2/Rpd3 control the H3 acetylation of two different sets of osmotic stress-induced genes. *Epigenetics* 2020;15:251–271.
  133. Li S, Yue Z, Tanaka TU. Smc3 deacetylation by Hos1 facilitates efficient dissolution of sister chromatid cohesion during early anaphase. *Mol Cell* 2017;68:605–614.
  134. Wang A, Kurdistani SK, Grunstein M. Requirement of Hos2 histone deacetylase for gene activity in yeast. *Science* 2002;298:1412–1414.
  135. Suka N, Suka Y, Carmen AA, Wu J, Grunstein M. Highly specific antibodies determine histone acetylation site usage in yeast heterochromatin and euchromatin. *Mol Cell* 2001;8:473–479.
  136. De Nadal E, Zapater M, Alepuz PM, Sumoy L, Mas G, et al. The MAPK Hog1 recruits Rpd3 histone deacetylase to activate osmoresponsive genes. *Nature* 2004;427:370–374.
  137. Tribus M, Bauer I, Galehr J, Rieser G, Trojer P, et al. A novel motif in fungal class 1 histone deacetylases is essential for growth and development of *Aspergillus*. *Mol Biol Cell* 2010;21:345–353.
  138. Kawauchi M, Nishiura M, Iwashita K. Fungus-specific sirtuin HstD coordinates secondary metabolism and development through control of LaeA. *Eukaryot Cell* 2013;12:1087–1096.
  139. Kloze RJ, Yamane K, Bae Y, Zhang D, Erdjument-Bromage H, et al. The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature* 2006;442:312–316.
  140. Tamaru H. Confining euchromatin/heterochromatin territory: jumonji crosses the line. *Genes Dev* 2010;24:1465–1478.
  141. Isaac S, Walfridsson J, Zohar T, Lazar D, Kahan T, et al. Interaction of epe1 with the heterochromatin assembly pathway in *Schizosaccharomyces pombe*. *Genetics* 2007;175:1549–1560.
  142. Turberfield AH, Kondo T, Nakayama M, Koseki Y, King HW, et al. KDM2 proteins constrain transcription from CpG island gene promoters independently of their histone demethylase activity. *Nucleic Acids Res* 2019;47:9005–9023.
  143. Qian S, Wang Y, Ma H, Zhang L. Expansion and Functional Divergence of Jumonji C-Containing Histone Demethylases: Significance of Duplications in Ancestral Angiosperms and Vertebrates. *Plant Physiol* 2015;168:1321–1337.
  144. Hong SH, Cho YW, Yu LR, Yu H, Veenstra TD, et al. Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases. *Proc Natl Acad Sci U S A* 2007;104:18439–18444.
  145. Agger K, Cloos PAC, Christensen J, Pasini D, Rose S, et al. UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* 2007;449:731–734.
  146. Dehé P-M, Dichtl B, Schaft D, Roguev A, Pamblanco M, et al. Protein interactions within the Set1 complex and their roles in the regulation of histone 3 lysine 4 methylation. *J Biol Chem* 2006;281:35404–35412.
  147. Lewis ZA, Adhvaryu KK, Honda S, Shiver AL, Knip M, et al. DNA Methylation and normal chromosome behavior in *Neurospora* depend on five components of a histone Methyltransferase complex, DCDC. *PLoS Genet* 2010;6:e1001196.
  148. Eberharther A, Sterner DE, SchieltzD, HassanA, YatesJR, et al. The ADA complex is a distinct histone acetyltransferase complex in *Saccharomyces cerevisiae*. *Mol Cell Biol* 1999;19:6621–6631.
  149. Doyon Y, Côté J. The highly conserved and multifunctional NuA4 HAT complex. *Curr Opin Genet Dev* 2004;14:147–154.
  150. Verzijlbergen KF, van Welsem T, Sie D, Lenstra TL, Turner DJ, et al. A barcode screen for epigenetic regulators reveals a role for the NuB4/HAT-B Histone Acetyltransferase complex in Histone turnover. *PLoS Genet* 2011;7:e1002284.
  151. Sendra R, Tse C, Hansen JC. The yeast histone acetyltransferase A2 complex, but not free Gcn5p, binds stably to nucleosomal arrays. *J Biol Chem* 2000;275:24928–24934.
  152. Grant PA, Duggan L, Côté J, Roberts SM, Brownell JE, et al. Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an ada complex and the SAGA(spt/ada) complex. *Genes Dev* 1997;11:1640–1650.
  153. Erlendson AA, Friedman S, Freitag M, Heitman J, Stukenbrock EH. A matter of scale and dimensions: chromatin of chromosome landmarks in the fungi. *Microbiol Spectr* 2017;5:4–5.
  154. Buscaino A. Chromatin-Mediated Regulation of Genome Plasticity in Human Fungal Pathogens. *Genes (Basel)* 2019;10:E855.
  155. Steenwyk JL, Mead ME, Knowles SL, Raja HA, Roberts CD, et al. Variation among biosynthetic gene clusters, secondary metabolite profiles, and cards of virulence across *Aspergillus* species. *Genetics* 2020;216:481–497.

### Five reasons to publish your next article with a Microbiology Society journal

1. When you submit to our journals, you are supporting Society activities for your community.
2. Experience a fair, transparent process and critical, constructive review.
3. If you are at a Publish and Read institution, you'll enjoy the benefits of Open Access across our journal portfolio.
4. Author feedback says our Editors are 'thorough and fair' and 'patient and caring'.
5. Increase your reach and impact and share your research more widely.

Find out more and submit your article at [microbiologyresearch.org](https://microbiologyresearch.org).