HAL open science

Multi-omic analysis of gametogenesis reveals a novel signature at the promoters and distal enhancers of active genes

Marion Crespo, Annelaure Damont, Melina Blanco, Emmanuelle Lastrucci, Sara El Kennani, Côme Ialy-Radio, Laila El Khattabi, Samuel Terrier, Mathilde Louwagie, Sylvie Kieffer-Jaquinod, et al.

To cite this version:

Marion Crespo, Annelaure Damont, Melina Blanco, Emmanuelle Lastrucci, Sara El Kennani, et al.. Multi-omic analysis of gametogenesis reveals a novel signature at the promoters and distal enhancers of active genes. Nucleic Acids Research, 2020, 48 (8), pp.4115-4138. 10.1093/nar/gkaa163 . hal-02945580

HAL Id: hal-02945580

https://hal.science/hal-02945580v1

Submitted on 22 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Published online 17 March 2020 *Nucleic Acids Research, 2020, Vol. 48, No. 8* **4115–4138**
doi: 10.1093/nar/gkaa163

Multi-omic analysis of gametogenesis reveals a novel signature at the promoters and distal enhancers of active genes

Marion Crespo[1], Annelaure Damont[2], Melina Blanco[3], Emmanuelle Lastrucci[1], Sara El Kennani[1,4], Côme Ialy-Radio[3], Laila El Khattabi[3], Samuel Terrier[2], Mathilde Louwagie[1], Sylvie Kieffer-Jaquinod[1], Anne-Marie Hesse[1], Christophe Bruley[1], Sophie Chantalat[5], Jérôme Govin[1,4], François Fenaille[2], Christophe Battail[6,*], Julie Cocquet[3,*] and Delphine Pflieger[1,7,*]

[1]Univ. Grenoble Alpes, CEA, Inserm, IRIG-BGE, 38000 Grenoble, France, [2]Service de Pharmacologie et d'Immunoanalyse, Laboratoire d'Etude du Métabolisme des Médicaments, CEA, INRA, Université Paris Saclay, MetaboHUB, 91191 Gif-sur-Yvette, France, [3]Institut Cochin, INSERM U1016, CNRS UMR8104, Université de Paris, 75014 Paris, France, [4]CNRS UMR 5309, Inserm U1209, Université Grenoble Alpes, Institute for Advanced Biosciences, 38000 Grenoble, France, [5]Centre National de Recherche en Génomique Humaine (CNRGH), Institut de Biologie François Jacob, CEA, Université Paris-Saclay, 2 rue Gaston Crémieux, CP 5706, 91057 Evry Cedex, France, [6]Univ. Grenoble Alpes, CEA, INSERM, Biosciences and Biotechnology Institute of Grenoble, Biology of Cancer and Infection UMR_S 1036, 38000 Grenoble, France and [7]CNRS, IRIG-BGE, 38000 Grenoble, France

Received September 26, 2019; Revised January 30, 2020; Editorial Decision March 03, 2020; Accepted March 07, 2020

ABSTRACT

Epigenetic regulation of gene expression is tightly controlled by the dynamic modification of histones by chemical groups, the diversity of which has largely expanded over the past decade with the discovery of lysine acylations, catalyzed from acyl-coenzymes A. We investigated the dynamics of lysine acetylation and crotonylation on histones H3 and H4 during mouse spermatogenesis. Lysine crotonylation appeared to be of significant abundance compared to acetylation, particularly on Lys27 of histone H3 (H3K27cr) that accumulates in sperm in a cleaved form of H3. We identified the genomic localization of H3K27cr and studied its effects on transcription compared to the classical active mark H3K27ac at promoters and distal enhancers. The presence of both marks was strongly associated with highest gene expression. Assessment of their co-localization with transcription regulators (SLY, SOX30) and chromatin-binding proteins (BRD4, BRDT, BORIS and CTCF) indicated systematic highest binding when both active marks were present and different selective binding when present alone at chromatin. H3K27cr and H3K27ac finally mark the building of some sperm super-enhancers. This integrated analysis of omics data provides an unprecedented level of understanding of gene expression regulation by H3K27cr in comparison to H3K27ac, and reveals both synergistic and specific actions of each histone modification.

INTRODUCTION

Histone post-translational modifications (PTMs) act as crucial epigenetic regulators in multiple biological processes by modulating chromatin compaction, organizing DNA repair and fine-tuning gene expression. Since its identification as a histone lysine modification in 1963 (1), acetylation of several histone lysine residues has been functionally characterized and shown to activate transcription (2), by binding bromodomain-containing proteins and transcription factors (3). Over the past 12 years, new PTMs that modify lysine residues have been discovered. These modifications, collectively called acylations, possess variable electrostatic and structural features: propionylation and butyrylation bear an additional methyl or ethyl group compared to acetylation (4); crotonylation specifically contains an unsaturated bond, which confers to it a planar configuration (5); malonylation, succinylation and glutarylation end up with a carboxylic acid (6,7), whereas hydroxy-butyrylations

*To whom correspondence should be addressed. Tel: +33 4 38 78 22 65; Fax: +33 4 38 78 50 51; Email: delphine.pflieger@cea.fr
Correspondence may also be addressed to Julie Cocquet. Email: julie.cocquet@inserm.fr
Correspondence may also be addressed to Christophe Battail. Email: christophe.battail@cea.fr

© The Author(s) 2020. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

bear an OH group (8,9). More recently, the landscape of histone lysine PTMs has further broadened with the identification of benzoylation and lactylation (10,11). All these studies have established that histones can be modified by a rich repertoire of acylations, by the reaction between acyl-coenzymes A (acyl-CoAs) and the primary amine on lysine side chain. The epigenetic landscape thus appears to be intricately controlled by the cell metabolic status, and more precisely by the nuclear concentrations of acyl-CoA molecules (12).

One key question that emerged from the discovery of this large palette of PTMs is whether they fulfill redundant functions with acetylation or they are endowed with specific roles, notably in chromatin structure and gene expression control. To address this question, previous works have focused on the identification of enzymes capable of catalyzing acylations, called 'writers'; of enzymes in charge of removing acylations, called 'erasers'; and of the proteins that would preferentially bind non-acetyl acylations compared to acetylation, called 'readers'. The histone acetyltransferase (HAT) p300 was shown to accommodate various acyl-CoA cofactors and thus to catalyze a range of acylations, among which are acetylation, propionylation, butyrylation, crotonylation and hydroxybutyrylations (13–15). Crotonylation can be catalyzed by the acetyltransferase MOF (KAT8) in addition to p300 and CBP (16), while succinylation can be catalyzed by GCN5 (KAT2A) acting in tight collaboration with a nuclear pool of α-ketoglutarate dehydrogenase complex that ensures local production of succinyl-CoA (17). Erasers are globally classified into two families, namely Zn^{2+}-dependent histone deacetylases (HDAC1–11) and NAD^{+}-dependent sirtuin deacetylases (SIRT1–7). While acetylation is removed by HDACs, longer chain acylations are usually removed by diverse sets of Sirtuins: SIRT1-3 erase propionylation and butyrylation, SIRT5 the three acidic acylations, SIRT3 removes β-hydroxybutyrylation at lysine residues not flanked by glycine and HDAC3 catalyzes this removal regardless of the neighboring residues, and SIRT2 ensures de-benzoylation (12,18,19,10). The catalytic removal of crotonylation has been attributed either to SIRT1-3 (20) or to HDAC1-3 (21). Finally, the probable divergence of functions between acetylation and longer chain acylations essentially lies in readers that would preferentially dock onto one type of PTM. Bromodomain-containing proteins have long been described to bind acetylated lysines (22), and their ability to recognize longer chain acylations has been extensively studied. While the majority of human bromodomains only bind acetylated and propionylated peptides, a few also recognize butyrylated and crotonylated lysines (23). Very interestingly, in a short period of time, several studies reported that the double PHD finger (DPF) domains of MOZ and DPF2, and YEATS domains exhibited a strong preference for crotonylated lysines (Kcr) (24–27). More recently, the YEATS domain of GAS41 was demonstrated to recognize succinylated Lys122 from histone H3 (28). Further research is necessary to get the full picture of proteins binding acylations more strongly than acetylation (29) and confer specific roles to them in the context of chromatin.

Lysine crotonylation was originally described in the context of mouse spermatogenesis which is a model system where dramatic changes occur in chromatin (5). During this differentiation process, diploid spermatocytes (SC) undergo meiotic divisions to yield round spermatids (RS). The latter further evolve into elongating and condensing spermatids (EC) whose chromatin gets compacted, while histones are removed from the genome to be eventually replaced by protamines (30). Finally, sperm only contains a few percent of the histones originally present in diploid cells (31). In mouse germinal cells, a genome-wide hyperacetylation at post-meiotic steps has been associated with the eviction of histones and their substitution by protamines (32,33). Besides, a global increase of lysine crotonylation has been described at the latest stages of spermatogenesis (5). Importantly, by ChIP-seq experiments using pan-Kcr antibodies, it was established that histones harboring crotonylation marks and lying around Transcription Start Sites (TSS) are more associated with actively transcribed genes than acetylated histones. Nonetheless, the precise histone lysine residues that contribute to the global Kcr increase are not yet determined. Interestingly, in spite of the evolutionary distance between mouse and yeast, mouse spermatogenesis has functional similarities with yeast sporulation. These two processes share a meiosis step, a dramatic chromatin reorganization accompanied by profound changes in histone PTMs, and end up with the tight compaction of the genome to protect it against environmental cues (34). Dynamic acetylation and phosphorylation occur during this process and a critical role has been attributed to acetylation during the final steps of chromatin organization in spores (35). In particular, hyperacetylation of the N-terminal tail of histone H4 has also been found to occur during yeast sporulation late after meiosis (35). Lysine crotonylation has not yet been studied during this process.

In the present study, we explored in parallel the dynamics of lysine acetylation and crotonylation in histones H3 and H4 in the contexts of mouse spermatogenesis and yeast sporulation. The crotonylation marks H4K8cr, H3K9cr, H3K18cr and H3K27cr, in combination with specific sets of neighboring modified lysine residues, were found to be evolutionary conserved. In mouse, crotonylation levels appeared to be of significant abundance in comparison to acetylation on H4K8 and H3K27. H3K27cr is maintained in sperm where it accumulates in a cleaved form of histone H3. Metabolomic analyses of various acyl-CoA molecules were performed to evaluate a possible correlation between histone modification levels and cellular concentrations of the acyl-donor molecules. Finally, given its similarity to H3K27ac which is described to mark active transcription at promoters and distal enhancers, we obtained the genome-wide localization of H3K27cr in mouse germinal cells and compared its localization with transcription factors/regulators. As a whole, this study reveals that H3K27cr is a hallmark of active genes, located at the promoters of around 7000 genes in round spermatids, particularly of those specific to spermiogenesis, and at ∼9000 putative enhancers. Association of H3K27cr with H3K27ac signifies even stronger gene activation while each mark also exhibits specificities such as concentrating more on specific genomic regions and enhanced binding to certain transcription factors/regulators. Among distal enhancers, super-enhancers concentrate transcription factors and co-

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

activators to induce the expression of distant genes, especially lineage-specific genes. We observed that a combined high level of H3K27ac and H3K27cr exists in super-enhancers determined in SC and RS cells. In addition, some sperm super-enhancers already exist in RS with the presence of both marks, which suggests that H3K27cr contributes to defining at least some super-enhancers.

MATERIALS AND METHODS

Mice

All animals used in the present study were of C57BL/6 background and processed at adult age (between 2- and 6-month-old males). Animal procedures were subjected to local ethical review (Comite d'Ethique pour l'Experimentation Animale, Universite Paris Descartes; registration numbers: CEEA34.JC.114.12; APAFIS 14214-2017072510448522v26).

Germ cell purification by elutriation

Enriched fractions of primary spermatocytes, round spermatids and elongating/condensing spermatids were obtained using two to three mice per experiment (i.e. four to six testes) by centrifugal elutriation as described previously (36). Specifically, freshly dissected testes were chopped in DMEM (GIBCO) and incubated for 30 min at 31°C with 2.5 mg/ml trypsin (GIBCO) and 50 mg/ml DNase I (Sigma). After adding 8% of fetal calf serum (GIBCO), the cells were passed through a 100-μm filter. Cells were then centrifuged at 500 g for 15 min and resuspended in DMEM 0.5% bovine serum albumin (Sigma) with 50 mg/ml DNase I, and cooled on ice. Cells were then separated using a standard chamber with a JE-5.0 rotor in a J-6M/E centrifuge (Beckman). Fractions enriched in elongating/condensing spermatids were collected at 3000 rpm with a flow rate of 16 ml/min (purity ∼90–99%), round spermatids were collected at 3000 rpm with a flow rate of 40 ml/min (purity ∼70–90%) and primary spermatocytes were collected at 2000 rpm with a flow rate of 28 ml/min (purity ∼60–80%). Collected fractions were washed in 1× PBS, then frozen down at −80°C. For ChIP-seq analyses, prior to freezing, cell pellets were crosslinked for 10 min at room temperature in 1× PBS containing 1% formaldehyde. Then glycine was added (125 mM final concentration) and samples were incubated for 5 min at room temperature, before being washed twice with ice-cold 1× PBS and frozen down at −80°C.

Germ cell purification by FACS

Highly enriched fractions of primary spermatocytes (>90%) and round spermatids (∼99%) were collected following a protocol adapted from (37) with the following modifications. Seminiferous tubules were treated with collagenase type I at 120 U/ml (GIBCO) for 30 min at 36°C in 1× HBSS (GIBCO) supplemented with 20 mM HEPES pH 7.2, 1.2 mM $MgSO_4$, 1.3 mM $CaCl_2$, 6.6 mM sodium pyruvate, 0.05% lactate. Then tubules were collected at the surface of a 40 μm cell strainer and incubated in cell dissociation buffer (Invitrogen) supplemented with 10 μg/ml DNase I (Sigma) for 15 min at 36°C. After filtration using a 40 μm cell strainer, the flowthrough was centrifuged for 10 min at 300 g. The cell pellet was resuspended in 1× HBSS supplemented with 20 mM HEPES pH 7.2, 1.2 mM $MgSO_4$, 1.3 mM $CaCl_2$, 6.6 mM sodium pyruvate, 0.05% lactate, 0.5 mM glutamine and 1% fetal bovine serum. Cells were then stained with Hoechst 33342 (Invitrogen) at a concentration of 5 μg/ml for 45 min at 36°C. Prior to FACS analysis, 2 μg/ml propidium iodide (SIGMA) was added. Analysis and cell collection were performed at the Cochin Cytometry and Immunobiology Facility using an ARIA III cell sorter (Becton Dickinson, San Jose, CA, USA) using parameters described in (38).

For germ cell fractions purified by elutriation or FACS, cell purity was assessed for each sample by microscope observation following DAPI (4,6-diamidino-2-phenylindole) staining (VECTASHIELD Mounting Medium with DAPI, Vectorlab, Burlingame, CA, USA) of cells spread onto glass slides and fixed with 4% buffered paraformaldehyde.

Collection of epididymal spermatozoa

Spermatozoa were collected from cauda epididymides as described previously (39). In brief, using small pipette tips, spermatozoa cells were gently squeezed out of cauda. Sperm cell purity was assessed for each sample by microscope observation following DAPI (4,6-diamidino-2-phenylindole) staining (VECTASHIELD Mounting Medium with DAPI, Vectorlab) of sperm cells spread onto glass slides and fixed with 4% buffered paraformaldehyde. All samples used in our analyses contained ⩾99% of spermatozoa.

Immunofluorescence and immunohistochemistry

Testes collected from adult mice were fixed for >5 h in 4% paraformaldehyde, then cut in half and either left in paraformaldehyde or incubated in Bouin reagent overnight. Immunofluorescence and immunohistochemistry were performed on 4-μm testicular sections as previously described (38) or using NovoLink Polymer Detection System (Leica), with modifications as follows. Antigen retrieval was performed by incubating slides for 40 min in 0.01 M sodium citrate solution (pH 6) in a water bath at 96°C. For immunohistochemistry using 3,3′-diaminobenzidine, slides were incubated in peroxidase block (Leica, Wetzlar, Germany) for 30 min. Permeabilization was performed for 10 min with 0.5% Triton X-100. Blocking was performed for 1 h at room temperature in 1× PBS, 0.1% Tween, 1% BSA. Anti-H3K27cr antibody (PTM-526 from PTM BIO) was diluted (at 1/100 for immunofluorescence, 1/300 for immunohistochemistry) in blocking buffer and incubated overnight at 4°C. For immunofluorescence experiments, slides were incubated with Alexa Fluor 488-labeled goat anti-(mouse IgG) (1/500; Life technologies) and Alexa Fluor 594-conjugated peanut agglutinin lectin (1/500; Life technologies) diluted in 1× PBS for 1 h at room temperature. Lectin was used to stain the developing acrosome and determine the stage of testis tubules as described in (40). DAPI (in VECTASHIELD Mounting Medium) was used to stain nuclei. Immunofluorescence pictures were taken with an Olympus BX63 microscope and montage was performed

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

using ImageJ 1.48v (http://imagej.nih.gov/ij/). Immunohistochemistry pictures were taken with Perkin Elmer Lamina slide scanner and analyzed using CaseViewer software.

Acid extraction of histones from mouse germinal cells

Histones were extracted from mouse germinal cells at different steps of spermatogenesis. On an average 5 million cells from each stage, namely spermatocytes, round spermatids and 10 million elongating/condensing cells were used. Briefly, cells were resuspended in 0.2 M sulfuric acid at 4°C before proceeding to six sonication cycles of 5 s at an amplitude of 20% with a break of 5 s between two cycles, using the sonicator (Vibracell 75186 equipped with a 3-mm probe CV18). Lysed cells were incubated on ice for 1h30 to extract histones. The lysates were centrifuged at 14 000 rpm for 10 min at 4°C and the acid extracted histones, contained in the resulting supernatant, were precipitated by incubation with trichloroacetic acid (TCA) at a final concentration of 20% for 45 min. The precipitate was subjected to centrifugation at 18 400 g for 15 min at 4°C. The subsequent histone pellet was washed with cold acetone containing 0.05% HCl, dried at room temperature, and resuspended in 100 μl of SDS-PAGE loading buffer. Extracted histones were separated on a 12% acrylamide gel to evaluate sample quality and histone quantity. For nanoLC–MS/MS analysis, ~10 μg of acid-extracted protein were loaded on a 12% acrylamide gel. After separation, usually four gel slices covering the MW range of histones were cut and then reduced with dithiothreitol, alkylated with iodoacetamide and in-gel digested with 0.1 μg trypsin (V511, Promega) per slice using a Freedom EVO150 robotic platform (Tecan Traging AG, Switzerland).

Purification of histones from sporulating yeast

Histones were obtained at four critical time points of sporulation, namely upon induction of this process by changing the culture medium ($T = 0$), at the time of meiosis ($T = 4$ h), at $T = 10$ h and when sporulation was complete ($T = 48$ h). Sporulation was performed as described in (35). The strain yJG109 in which H3 was N-terminally Flag-tagged has been used to purify histones (35). Yeasts were resuspended in Tris 50 mM pH7.5, EDTA 1 mM, NaCl 300 mM, NP-40 0.5%, glycerol 10%, DTT 1 mM, Complete Protease Inhibitor Cocktail (Roche), Trichostatin A 100 nM and Phosphatase Inhibitor Cocktail (Ref P0044, Sigma) (TENG-300 buffer). They were disrupted in a FastPrep (MP Biomedicals) using glass beads for 45 s at 6.5 m s^{-1}. Lysates were then sonicated three times 30 s with 30 s breaks for a total of ~150 kJ. Soluble extracts were obtained by centrifugation for 15 min at 20 000 g and incubated with anti-Flag M2 resin (A2220, Sigma) for 3 h at 4°C under rotation. Bound proteins were washed (four times with TENG buffer containing NaCl 500 mM, then with a final wash in TENG-300). Histones were eluted in TENG-300 added with 0.5 mg ml^{-1} of M2 Flag peptide (30 min at 4°C). Histone samples were then processed and analyzed in a similar way as those from mouse cells.

Mass spectrometry analysis of mouse germinal cells

The histone tryptic peptides were resuspended in 2.5% acetonitrile (ACN) and 0.05% trifluoroacetic acid (TFA). Peptides were then loaded on a PepMap C18 precolumn (300 μm × 5 mm) and separated on a C18 reversed-phase capillary column (75 μm i.d. × 15 cm ReproSil-Pur C18-AQ, 3 μm particles) using the UltiMate™ 3000 RSLCnano system (Thermo Fisher Scientific) coupled to a Q-Exactive HF mass spectrometer (Thermo Fisher scientific). The sample was firstly washed on the precolumn with 0.1% formic acid for 1.2 min prior to being loaded on the capillary column at a 300 nl/min flow rate. The mobile phases consisted of water with 0.1% formic acid (A) and acetonitrile with 0.08% (v/v) formic acid (B). Peptides were eluted with a gradient consisting of an increase of solvent B from 2.8% to 7.5% for 7.5 min, then from 7.5% to 33.2% over 33.5 min and finally from 33.2% to 48% over 6.5 min.

Mass spectrometry acquisitions were carried out by alternating one full MS scan with Orbitrap detection acquired over the mass range 300–1300 m/z, at a target resolution of 60 000 and with an AGC of 1e6, and data-dependent MS/MS spectra on the 10 most abundant precursor ions detected in MS. The peptides were isolated for fragmentation by higher-energy collisional dissociation (HCD) with a collision energy of 27 using an isolation window of 2 m/z, a target resolution of 60 000 and AGC fixed to 3e6. Dynamic exclusion was applied for 30 s. A first series of LC–MS/MS analyses of histones obtained from a biological triplicate was used for Figure 1C. A second series of analyses were acquired while loading more material of histone H3 onto the C18 column to get most reliable quantification of H3K27-containing peptides shown in Figure 2B.

PTM identification and quantification

Identification of modified peptides was obtained using the manually curated database MS_histoneDB developed in our group (41). MS/MS data interpretation was performed with the program Mascot (http://www.matrixscience.com/), using the following search parameters. The precursor and fragment mass tolerances were 5 ppm and 25 mmu, respectively; enzyme specificity was trypsin; the maximum number of trypsin missed cleavages was set to 5, carbamidomethyl (Cys) was specified as a fixed modification. We were interested in acetylation and crotonylation of Lys residues, methylation and dimethylation of Lys/Arg and trimethylation of Lys, which were indicated as variable PTMs in Mascot, in addition to N-terminal acetylation of proteins. All MS/MS spectra were visually scrutinized to validate peptide identifications and PTM site assignment. Quantification was performed by using the software Proline (http://www.profiproteomics.fr/). Crotonylated and acetylated peptides which contain H4K8, H3K18 and H3K27 and were particularly studied here were further manually quantified by using the Qualbrower tool within Xcalibur (Thermofisher scientific).

Normalization of MS signals to estimate peptide abundances

Normalization was done by dividing the raw MS signals of modified peptides by the MS signals of refer-

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

ence non-modified peptides, namely ISGLIYEETR and DNIQGITKPAIR for histone H4 and STELLIR for histone H3. We further applied a correction factor to compensate for differing ionization efficiencies between variably modified peptides: this factor was calculated from the LC-MS analysis of synthetic peptides (see Supporting file). We then obtained the relative abundance of each modified peptide at a constant histone amount.

We also sought to evaluate the relative level of the PTM combinations including H3K27 by estimating a percentage of each PTM compared to the others. The relative levels of H3K27 acylations was calculated as previously done in (42) and (43), that is by dividing the MS signal of the acylated form of interest by the sum of signals of all modified peptides of same sequence. For example, the relative abundance of KacSAPSTGGVK (K27-K36 peptide) was calculated using the following formula:

$$\text{K27ac}\,(\text{K27} - \text{K36 peptide}) = \frac{\text{Signal K27ac (K27} - \text{K36 peptide)}}{\sum \text{Signals modified K27} - \text{K36 peptides}} \times 100 \quad (1)$$

where $\sum$Signal K27 – K36 peptides = Signal K27me1 + Signal K27me2 + Signal K27me3 + Signal K27ac + Signal K27c.

Antibodies

For Western blotting and ChIP-seq analyses, the primary antibodies used were polyclonal anti-H4 (Active motif, AB_2636967), polyclonal anti-H3 (Abcam, AB_2793771), monoclonal anti-H3K27ac (PTM BIO, PTM-160), monoclonal anti-H3K27cr (PTM BIO, PTM-526) and polyclonal H4K8cr (Abcam, EPR17905(R)).

ChIP-seq analyses

Around 1 million spermatocytes and round spermatids were used for ChIP-seq analyses. The chromatin was prepared by Diagenode using the iDeal ChIP-seq kit for histones (Diagenode Cat# C01010059) and then sheared using the Bioruptor Pico sonication device (Diagenode Cat# B01060001) combined with the Bioruptor Water cooler. ChIP was performed using the IP-Star Compact Automated System (Diagenode Cat# B03000002). The antibody H3K27cr was used at 1 and 2 μg per IP and each cell type was analyzed in two biological replicates. After checking that the two antibody amounts led to very similar ChIP-seq peak profiles, we kept the data obtained with 1 μg to be shown in the results.

Libraries were prepared from input and immuno-purified DNA with 1ng as starting material using MicroPlex Library Preparation Kit v2 (Diagenode Cat# C05010013), then purified using Agencout® AMPPure® XP (Beckman Coulter) and quantified using Qubit™ dsDNA HS Assay Kit (Thermo Fisher Scientific, Q32854). Libraries were pooled and sequenced by Diagenode on an Illumina 3000/4000 HiSeq instrument with 50bp single-end reads at a depth of coverage per sample ranging from 55M to 70M.

Bioinformatic analyses of ChIP-seq data and of 5hmC data

Several published ChIP-seq datasets performed on mouse spermatocytes and round spermatid cells were downloaded from the Sequence Read Archive (SRA) database and re-analyzed using the same bioinformatic procedure (see below). The following ChIP-seq datasets were processed: H3K4me1 and H3K4me3 from GSE49621 (44), pan crotonylation from GSE32663 and GSE69946 (5) (45), H3K27ac from GSE107398 (46), BORIS and CTCF from GSE70764 (47), BRD4 from GSE56526 (48), SLY from PRJNA275694 (39), SOX30 from GSE107644 (49) and H3.1/2 and H3.3 from GSE42629 (50) (Supplementary Table S1). A ChIP-seq dataset on BRDT (51) had been acquired on a SOLiD sequencing platform whereas Illumina sequencers were used for the above listed datasets; we obtained the BED files of already mapped BRDT peaks and then extracted signals around TSS and over putative enhancers as for the other ChIP-seq data.

We obtained the data acquired on the DNA methylation mark 5-hydroxymethylcytosine over SC and RS stages among the several steps of mouse spermatogenesis analyzed in (52). The sequencing data was quality controlled, pre-processed and mapped to the mouse reference genome using the same bioinformatic procedure described for the ChIP-seq data. The 5hmC signal belonging to each exon was quantified in read counts using Bedtools intersect and normalized in BPM. The estimate of the 5hmC signal by gene was then calculated by summing the 5hmC signals associated with each of its exons.

Quality control. The read quality of ChIP-seq datasets was assessed using Fastqc (Andrews S., 2010, http://www.bioinformatics.babraham.ac.uk/projects/fastqc).

Pre-processing. ChIP-seq reads were sampled using Seqtk v20101003 (https://github.com/lh3/seqtk) by taking into account the heterogeneity in duplicated read percentages across samples, as indicated in Supplementary Table S1 in Supporting material, in order to keep similar depths of coverage across samples associated with the same ChIP. Reads were trimmed for low quality bases and Illumina adaptors using Trimmomatic v0.32 (parameters: ILLUMINACLIP:2:30:10, LEADING:30, TRAILING:30, SLIDINGWINDOW:4:30, MINLEN:30) (54).

Mapping. Reads were mapped using Bowtie2 v2.2.9 (default parameters) (55) on the Mus musculus genome assembly GRCm38 (mm10) using Illumina iGenomes reference sequences and annotations (http://emea.support.illumina.com). Reads aligned to multiple genome locations were kept. Read alignments were cleaned and sequence duplicates were removed using the Picard tool suite (addReadGroup, reorder, sort, clean, markDuplicates) (http://broadinstitute.github.io/picard/). Read alignment metrics were calculated using Picard multipleMetrics.

Peak calling. ChIP-seq peaks were called using Macs2 v2.1.1 (parameters: broad, broad-cutoff 0.01) (56). A control ChIP sample was used for peak calling depending on availability and sufficient depth of coverage (Supplementary Table S1).

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

TSS region quantification. Regions of ±500 bp around each gene transcription start site (TSS) were quantified in read counts using Bedtools intersect (53). The read count signal assigned to each TSS region was then normalized in Bins Per Million reads (BPM) (i.e. similar to the Transcripts Per Million reads (TPM) normalization applied on RNA-seq reads). The read counts of each TSS region was divided by its length in kilobases to produce reads per kilobase (RPK). The RPK values per sample were then summed up and divided by 1,000,000 to produce a 'per million' scaling factor. The RPK value assigned to each TSS region was then divided by the 'per million' scaling factor to produce the BPM value.

Enhancer region quantification. Enhancer regions were first identified and delimited using the broad peaks called from the analysis of the H3K4me1 ChIP-seq data, independently on mouse spermatocytes and round spermatid cells, and then merged into a single list. These enhancer regions were then quantified for each ChIP mark using the reads mapped in these regions. The read count signal assigned to each enhancer region was then normalized in BPM, as described in the previous paragraph. The execution of the ChIP-seq data analyses was supervised using the Bpipe workflow manager (57).

Bioinformatic analyses of RNA-seq data

The RNA-seq dataset used to evaluate the effect of H3K27cr mark on transcription was downloaded from (52).

Quality control. The read quality of RNA-seq datasets was assessed using Fastqc.

Pre-processing. Reads were trimmed for low quality bases and Illumina adaptors using Trimmomatic v0.32 (parameters: ILLUMINACLIP:2:30:10, LEADING:30, TRAILING:30, SLIDINGWINDOW:4:30, MINLEN:30).

Mapping and gene quantification. Reads were mapped using Star v2.5.2 (ENCODE standard parameters) (58) on the Mus musculus genome assembly GRCm38 (mm10) using Illumina iGenomes reference sequences and annotations (http://emea.support.illumina.com). Isoform and gene abundances normalized in TPM values were calculated using Rsem v1.3.1 (parameters: estimate-rspd). Read alignments were cleaned and sequence duplicates were removed using the Picard tool suite (addReadGroup, reorder, sort, clean, markDuplicates). Read alignment metrics were calculated using Picard multipleMetrics. The execution of the RNA-seq data analyses was supervised using the Bpipe workflow manager.

Differential gene expression analysis. The identification of differentially expressed genes between mouse spermatocytes and round spermatids was performed using EdgeR (TMM normalization, FDR < 0.01) (59).

Programs used to make figures related to ChIP-seq and RNA-seq data

Files containing TPM and BPM intensity values for RNA-seq and ChIP-seq analyses were handled using the following libraries in R: VennDiagram (https://cran.r-project.org/web/packages/VennDiagram/index.html), ChIPseeker and GenomicFeatures (60,61) to annotate ChIP-seq peaks in terms of genomic regions, ggplot2 (https://ggplot2.tidyverse.org/authors.html) and ggpubr (https://cran.r-project.org/web/packages/ggpubr/index.html) to generate boxplots, ComplexHeatmap (62) to build all heatmaps.

Extraction of acyl-CoA metabolites

On average 5 million cells from each stage, spermatocytes and round spermatids, and 10 millions elongating/condensing spermatids were used. Metabolites from 4 mouse testicles and from 25 mL of yeast cells at each time point were extracted after being metabolically quenched at −80°C. In brief, cells were suspended in 1 ml of pre-cooled (10% (w/v) TCA) and sonicated for 1 min at an amplitude of 20% with a break of 5 s between two cycles of 10 s. Testis tissue was disrupted at an amplitude of 40% and 2 cycles of 1 min were performed. Cell lysates were centrifuged at 15 000 g for 3 min at 4°C to separate the metabolites from protein and cell debris. The supernatant was loaded on SPE cartridges (Oasis HLB 10 mg), washed with 1 ml of water and then eluted with 500 ml of methanol before being dried in a speedvac and kept at −80°C until analysis.

MS analysis of acyl-CoA metabolites

Acetyl-CoA, crotonyl-CoA, as well as propionyl-CoA, butyryl-CoA, 2-hydroxy-butyryl-CoA, malonyl-CoA, succinyl-CoA and glutaryl-CoA standards were purchased from Sigma-Aldrich (Saint Quentin Fallavier, France). External calibration curves were used for the quantification of acyl-CoAs in all metabolic extracts, while acyl-CoAs were also quantified in yeast and testis samples by using the standard addition method. Briefly, dried yeast and testis metabolic extracts were resuspended in 50 µl of 50 mM ammonium formate solution containing two internal standards (acetoacetyl-CoA and $^{13}C_3$-malonyl-CoA both at 0.2 µg/ml – Sigma-Aldrich). Each extract was then split into 3× 10-µl aliquots, and spiked either with 10 µl of a 50 mM ammonium formate or with 10 µl of a 1.0× or a 2.0× solution of a mixture of the seven acyl-CoAs to be quantified (X being the estimated endogenous amount of each acyl-CoA). Dried SC, RS and EC extracts were resuspended in 25 µL of a 50 mM ammonium formate aqueous solution, 10 µl were then withdrawn and mixed with 10 µl of the internal standard solution.

LC–HRMS analyses were performed using an Ultimate 3000 chromatographic system (Thermo Fisher Scientific, Courtaboeuf, France) coupled to a quadrupole-time-of-flight (Q-TOF) mass spectrometer (Impact HD, Bruker Daltonics, Bremen, Germany) equipped with an electrospray ion source and operating in the positive ion mode. Metabolic extracts (15 µl) were loaded and separated on an Atlantis® dC15 5 µm 2.1 × 150 mm column (Waters, Milford, MA, USA) maintained at 30°C. Mobile phases were 50 mM aqueous ammonium formate (A) and acetonitrile (B), at a flow rate of 250 µl/min. The elution consisted of

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

an isocratic step of 1 min at 0% phase B, 0.1 min at 5% B followed by a linear gradient from 5 to 25% of phase B in 5.9 min and then up to 40% B in 1 min. These proportions were kept constant for 3 min before returning to 0% of phase B in 0.1 min and letting the system equilibrate for 8.9 min. HRMS data were acquired between m/z 700 and 1050 at 1.0 Hz in the profile mode and with a resolution set at 40 000 (m/z 1222). Nitrogen was used both as a nebulizer and drying gas. The source conditions were set as follows: end plate offset 700 V, capillary 3800 V, nebulizer 1.7 bar, dry gas 12.0 L/min and dry temperature 240°C. The mass spectrometer was calibrated internally with a sodium formate/acetate solution, which provides a mass measurement accuracy below 2 ppm on average. All raw data were manually treated using Compass QuantAnalysis software (version 2.2, Bruker Daltonics).

RESULTS

Site-specific dynamics of lysine acetylation and crotonylation in histones H3 and H4 during mouse spermatogenesis

To assess the dynamics of crotonylation compared to acetylation at specific histone Lys residues during mouse spermatogenesis, we performed proteomic analysis of histones H3 and H4 extracted from mouse male germ cells (Figure 1A). Histones were purified from three biological replicates of mouse spermatocytes (SC), round spermatids (RS) and elongating/condensing spermatids (EC), then separated on an SDS-PAGE gel and bands covering the molecular weights of histones were cut for in-gel trypsin digestion. We chose not to carry out *in vitro* propionylation of free lysines because the protocol of simple trypsin digestion was shown to yield the highest number of identified crotonylated sites in histones H3 and H4 (5). Overall, crotonylation was detected on H3K9, H3K18, H3K27 and H4K8 residues which were also identified in an acetylated form (Figure 1B). While acetylation at these lysine residues was formerly described in the context of mouse spermatogenesis, crotonylation was not reported on these sites but on H3K37, H4K12, H4K16 and H4K91 within histones H3 and H4 (63). In the latter study, *in vitro* propionylation was performed before trypsin digestion of histones to prevent proteolysis at *in vivo* non-modified lysine residues, which likely favored the detection of different peptides. MS signals of histone tryptic peptides were quantified and normalized to be at a constant histone amount in all samples (see Materials and Methods section). Importantly, peptide sequences H3 K27-K36 from canonical histone H3 and the variant H3.3 differ from each other by Ala31→Ser31. We calculated the relative abundance of each modified form by dividing its MS signal by the sum of signals detected for peptide species of same sequence and bearing different PTM patterns.

Hyperacetylation of H4 N-terminus during spermiogenesis co-occurs with methylation of H4R3. A hyperacetylation wave has previously been described at the end of mouse spermatogenesis, particularly on the N-terminal tail of histone H4 that contains the four modifiable residues H4K5, H4K8, H4K12 and H4K16 (35). We observed increases to different degrees from SC to EC cells, depending on the considered combinations of acetylated lysines: the doubly acetylated forms H4-$K_8acK_{12}ac$ and H4-$K_{12}acK_{16}ac$ exhibited a moderate increase, whereas the abundance of triply and quadruply acetylated sequences significantly rose from SC to EC cells (Figure 1C). Interestingly, the most dramatic abundance increase was detected for the fully acetylated N-terminal tail of H4 that additionally contained a methylation on Arg3 (H4-$R_3me1K_5acK_8acK_{12}acK_{16}ac$). Methylation on H4R3 by the Arginine methyltransferase PRMT1 was formerly described to promote hyperacetylation of histone H4 N-terminal tail and to facilitate transcription in other cellular models (64,65). Finally, we were able to reproducibly quantify peptides crotonylated at H4K8 and H3K27, but not at H3K9, probably due to a low modification stoichiometry at this site, nor at H3K18, due to co-elution on the chromatographic column of another peptide of very close mass (Supplementary Figure S1).

Crotonylation dynamics on H4K8. Crotonylation at H4K8 detected within sequence $GGK_8crGLGK_{12}acGGAK_{16}acR$ (Supplementary Figure S1) appeared to increase from SC to EC cells, which was confirmed by Western Blot using an anti-H4K8cr antibody (Figure 2A). The analysis of synthetic peptides $GGK_8GLGK_{12}acGGAK_{16}acR$ acetylated or crotonylated at H4K8 allowed us to estimate their relative ionization efficiencies, and then correcting the signals measured on endogenous peptides to better estimate their relative abundance (Supplementary Figure S2). Our data indicated that histone H4 molecules containing K5un/K8cr/K12ac/K16ac represented about one tenth of molecules bearing K5un/K8ac/K12ac/K16ac, highlighting a non-negligible relative stoichiometry of crotonylation compared to acetylation (Figure 2A). We could only reproducibly detect and quantify H4K8cr together with H4K16ac which sets an active chromatin state (66). This appears coherent with crotonylation being a mark of active transcription (5).

Crotonylation dynamics on H3K27. Proteomic analysis of the gel band at the mass of histone H3 revealed a complex variant-specific pattern of acetylation/crotonylation/methylations on H3K27 in combination with non-modified or dimethylated H3K36. Indeed, mass spectrometry analysis allowed us to distinguish the peptide sequences $K_{27}SAPATGGVK_{36}$ and $K_{27}SAPATGGVK_{36}me2KPHR_{40}$ shared by histones H3.1/H3.2/H3.t from sequences $K_{27}SAPSTGGVK_{36}$ and $K_{27}SAPSTGGVK_{36}me2KPHR_{40}$ belonging to the variant H3.3 (Figure 2B). One prominent feature was that the H3.3 variant was enriched in acetylation and crotonylation on Lys27 compared to methylated forms for both $K_{27}SAPSTGGVK_{36}$ and $K_{27}SAPSTGGVK_{36}me2KPHR_{40}$. This observation was more pronounced for the short peptide $K_{27}SAPSTGGVK_{36}$, with around 80% of the MS signal belonging to acetylated or crotonylated forms in SC, RS and EC cells. Another interesting variant-dependent observation was the increase of crotonylation from RS to EC cells for both peptides $K_{27}SAPSTGGVK_{36}$ and $K_{27}SAPSTGGVK_{36}me2KPHR_{40}$ from H3.3. While crotonylation was not even detected on peptide

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

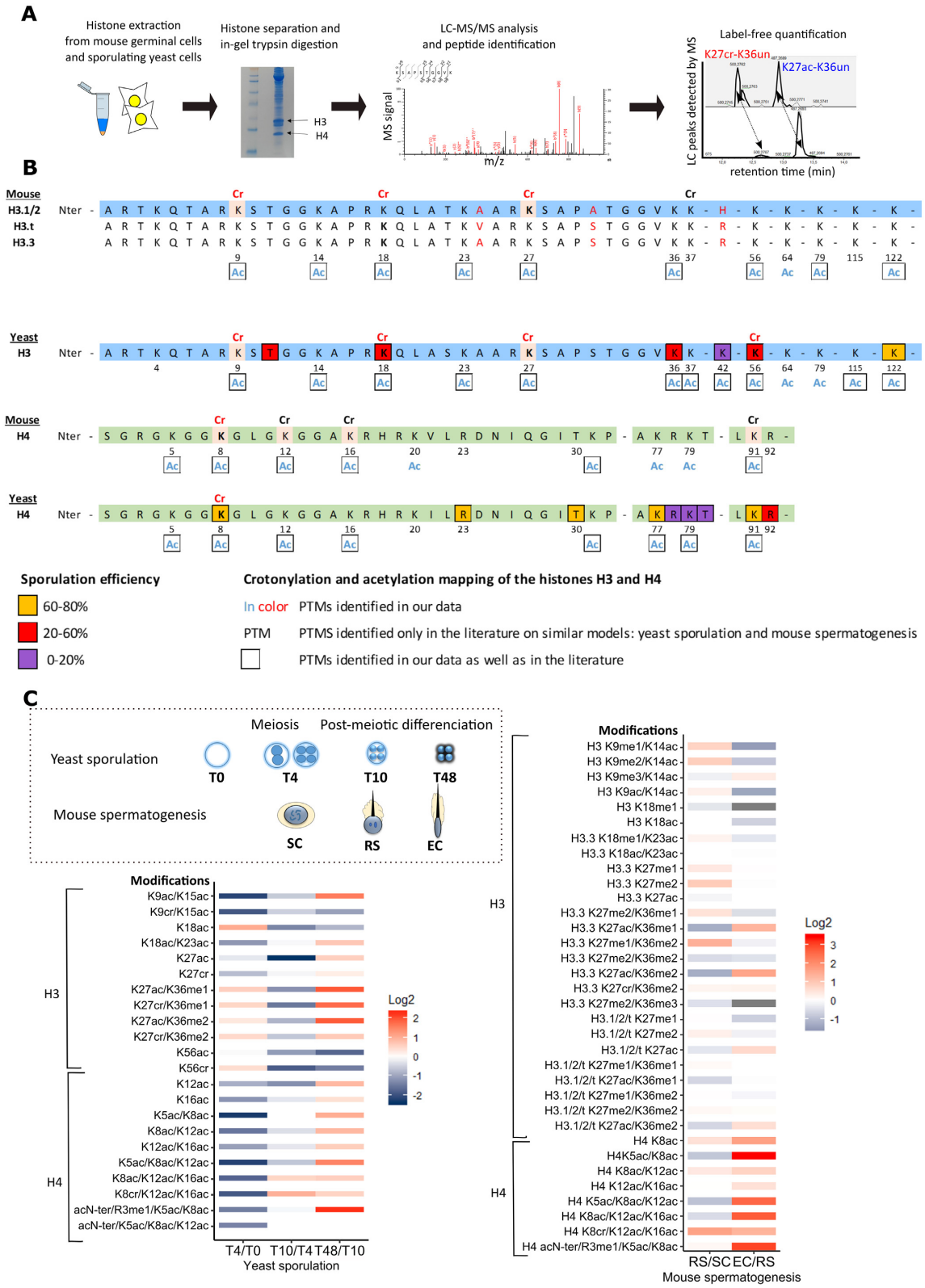

Figure 1. (**A**) Overview of the experimental workflow implemented for the proteomic analysis of histones from mouse germinal cells and sporulating yeast cells. (**B**) Mapping by proteomics of acetylated and crotonylated lysines identified on histones H3 and H4 during yeast sporulation and mouse spermatogenesis. (**C**) Quantitative variations of acetylated and crotonylated peptides during the two processes.

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

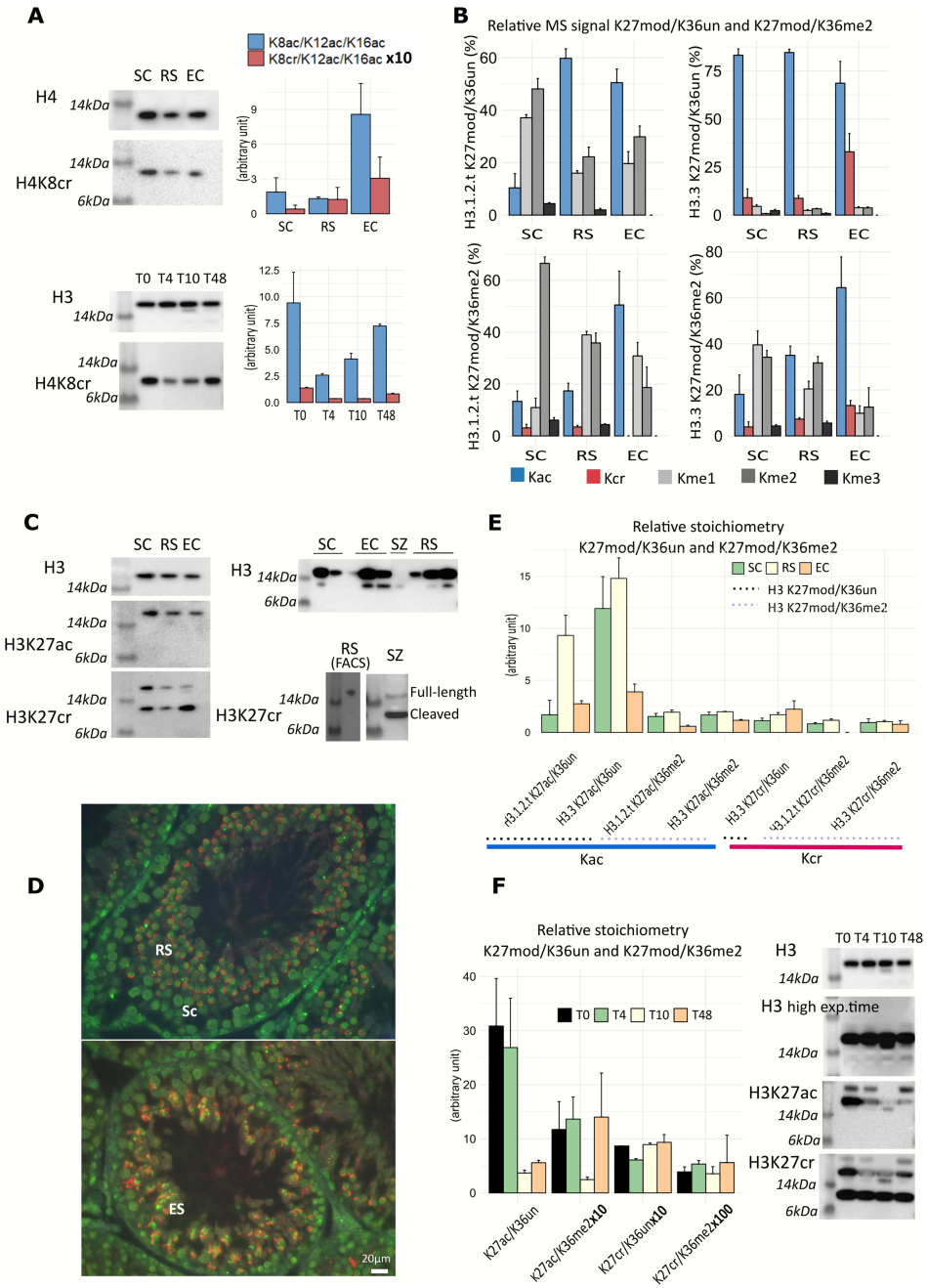

Figure 2. (**A**) Estimation of abundance variations of peptides containing H4K8cr by both proteomic and WB analyses. (**B**) Estimation of abundance variations of peptides K27mod-K36un and K27mod-K36me2 within H3.1/2/t and variant H3.3 between SC, RS and EC cells. The MS signal of each peptide form was divided by the sum of MS signals detected for variably modified forms of same sequence. Bars are not represented for the peptides trimethylated at H3K27 in EC stage, as well as crotonylated at H3K27 in KSAPATGGVK in all stages and in KSAPATGGVKme2KPHR in EC stage, because they were not detected. (**C**) WB of histones extracted from spermatogenic cells and sperm using anti-H3 (used for normalization), anti-H3K27ac and anti-H3K27cr antibodies. In the upper right panel, increasing amounts of material were loaded, to be able to detect the cleaved form of histone H3. The bottom right panel shows the near exclusive detection of full-length H3 with anti-H3K27cr antibody in pure RS cells obtained by FACS, and mostly of cleaved histone H3 in sperm. (**D**) Representative immunodetection images obtained using antibody against H3K27cr (green) on testicular sections from adult wild-type mice. Lectin (red) was used to stain acrosomes in order to determine tubule stage. The top picture shows a stage VIII tubule characterized by acrosomes (lectin red staining) surrounding more than one-third of round spermatid (RS) heads and facing the basal membrane. Sc shows spermatocytes. The bottom picture shows a stage XI tubule stage characterized by asymmetrical acrosomes (lectin red staining) surrounding elongating spermatid (ES) heads. Pictures were taken using the same parameters. Scale bar indicates 20 µm. See also Supplementary Figure S4A and B. (**E**) Relative stoichiometries of peptides starting at K27 modified by acetylation or crotonylation during mouse spermatogenesis. Peptide MS signals were this time normalized by the peptide STELLIR, which is shared by the various sequences of histone H3, to allow better comparison to the observations made by WB. Their MS signals were further normalized by a correction factor to compensate for variable ionization efficiencies between the different peptide sequences. A peptide sequence P1 is said to possess a two times better ionization efficiency than a sequence P2 when one mole of P1 gives rise to a two times higher MS signal than one mole of P2 (**F**) Similarly to (E), relative stoichiometries of peptides starting at K27 modified by acetylation or crotonylation during yeast sporulation obtained by proteomic analysis. WB analyses using antibodies against histone H3, H3K27ac and H3K27cr over sporulation.

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

K_{27}SAPATGGVK_{36} from canonical H3, crotonylation of K_{27}SAPSTGGVK_{36} from H3.3 reached 30% of the MS signal in EC cells. The presence of the active mark crotonylation in H3.3 is in agreement with this variant being preferentially incorporated in active regions of chromatin and bearing marks of active transcription (67,68). Finally, the increase of acetylation and to a lower extent crotonylation on H3K27 was concomitant with the decrease of repressive marks H3K27me1, H3K27me2 and H3K27me3.

Confronting proteomics data on H3K27cr to WB observations. We sought to support our proteomics results with WB using an antibody specific to H3K27cr. Its specificity had been verified by dot blot on peptides containing either H3K23 or H3K27 in crotonylated, acetylated, butyrylated or propionylated forms, and we additionally verified that both sequences from canonical H3 and from H3.3 were detected, whereas a peptide containing H4K8cr was not (Supplementary Figure S3). By WB, a signal of constant intensity between RS and EC cells was detected at the mass corresponding to intact histone H3 (∼16 kDa) (Figure 2C). The antibody also detected a more intense signal at ∼12 kDa, with an intensity strongly increasing from RS to EC cells. Immunofluorescence and histochemistry on testicular sections confirmed the presence of H3K27cr mark in male germ cells, with a stronger signal in elongating spermatids (ES, Figure 2D and Supplementary Figure S4A, B). Inspection of the proteomics data acquired on the gel bands at 12 kDa showed mostly peptides corresponding to histone H4 (since it is roughly the mass of histone H4) but also allowed us to identify histone H3 with peptides spanning the whole protein sequence except residues 1–17 (Supplementary Figure S4C); a signal for peptide K_{27}crSAPSTGGVK_{36}me2KPHR$_{40}$ from H3.3 was detected in this sample. We therefore deduced that the signal at ∼12 kDa detected by the antibody raised against H3K27cr corresponds to a clipped form of histone H3 coming from a cleavage shortly before H3K27. Truncated forms cleaved at Ala21-Thr22 or at Thr22 and Arg26 have recently been described in sperm cells (69). We confirmed by WB the presence of H3K27cr in sperm histones, in large majority in a cleaved form (Figure 2C). Of note, a WB performed on sperm against H4K8cr did not reveal any signal, which confirmed that our anti-H3K27cr antibody did not detect H4K8cr in the lower-mass band. To better assess whether the cleaved form of H3 appears progressively during spermiogenesis, we analyzed by WB the histones extracted from RS cells purified by FACS, being of even higher purity (99%) than RS cells collected by elutriation (Figure 2C). The anti-H3K27cr antibody detected full-length H3 and a faint signal for the cleaved form upon long exposure, which would support the hypothesis that cleavage of histone H3 occurs progressively during spermiogenesis. Besides, a WB obtained in the three cell stages using an antibody raised against H3K27ac only revealed a band at the mass of histone H3 (Figure 2C), whose intensity would decrease a bit from RS to EC stages. We verified by dot blot that the anti-H3K27ac antibody was able to recognize a 17-amino-acid-long peptide starting at H3K27ac (Supplementary Figure S4D). In addition, extraction of the MS signals detected for tryptic peptides K27ac/K36me2 from H3.1/2/t and H3.3 in gel bands at ∼16 kDa and ∼12 kDa indicated that H3K27ac was present at a relative amount of less than 1/25 in the lower-mass band, which was not detected at the exposure times used for revealing WBs (Supplementary Figure S4E). Our results thus hinted at probably different processing during mouse spermiogenesis of histone H3 molecules bearing different PTMs on H3K27, in particular acetylation and crotonylation. These observations are in line with the above report on sperm: an anti-H3K27ac antibody only detected a signal at the mass of intact H3, whereas an antibody raised against H3K27me3 detected the lower-mass band (69).

To draw a better parallel between proteomics and WB results, we also handled our proteomics data acquired on the band of full-length H3 to be at a constant amount of total histone H3. We then normalized the MS data by the peptide STELLIR which is common to all H3 variants. We additionally applied correction factors calculated from synthetic peptides to compensate for variable ionization efficiencies. We then plotted the relative stoichiometry of all the peptides containing H3K27 in an acetylated and a crotonylated forms (Figure 2E). The peptide species acetylated on H3K27 collectively showed a constant amount between SC and RS cells, and then some decrease in EC cells, in agreement with the observation made by WB in Figure 2C. Besides, very interestingly, the relative stoichiometries of crotonylated and acetylated peptides were very high: (i) the abundance of H3.3-K_{27}crSAPSTGGVK_{36} represented about one third of that of H3.3-K_{27}acSAPSTGGVK_{36} and (ii) both K_{27}SAPSTGGVK_{36}me2KPHR$_{40}$ from H3.3 and K_{27}SAPATGGVK_{36}me2KPHR$_{40}$ from canonical H3 were detected at similar stoichiometries to their acetylated counterparts. One can finally note that at EC stage, H3K27cr was only detectable on variant H3.3.

Altogether, these results indicated that histone H3 crotonylated at residue K27 was maintained on the genome until sperm, mostly within variant H3.3 and essentially in a cleaved form devoid of its N-terminal part. In addition, at the EC stage, the relative stoichiometries between crotonylated and acetylated H3K27 was around 1:3 on full-length histone molecules non-modified on H3K36 and about 1:1 on molecules bearing H3K36me2. Finally, when considering all crotonylated sites identified here, both H3K27cr and H4K8cr appear to contribute to the general wave of crotonylation described toward the end of spermatogenesis (5).

Assessing conservation of crotonylation profiles during evolution: patterns observed during yeast sporulation

In spite of the evolutionary distance between the two organisms, yeast sporulation and mouse spermatogenesis share similarities, including a meiosis step and a post-meiosis differentiation that culminates with extensive chromatin restructuring and compaction. At the histone PTM level, an increase in phosphorylation at H3T11 and H4S1 was reported in the two processes at the time of meiosis, and hyperacetylation of the N-terminal tail of H4 was shown to be conserved at the end of post-meiotic differentiation, which allows recruitment of the bromodomain-containing

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

proteins Bdf1 in yeast and BRDT in mouse (35). We then wanted to assess as to what extent this parallel applied to crotonylation sites and their dynamics. Histones were obtained from yeast cells at four critical time points of sporulation, namely upon induction of this process by changing the culture medium ($T = 0$), at the time of meiosis ($T = 4$ h), after meiosis ($T = 10$ h) and in fully mature spores, after the final post-meiotic differentiation ($T = 48$ h). Histone samples were then processed and analyzed in a similar way as those from mouse cells (Figure 1A).

First, an abundance increase was observed for acetylated peptides from the H4 N-terminal tail in the course of yeast sporulation, with the highest change being observed for the sequence quadruply acetylated at K5/K8/K12/K16 in association with R3me1 (Figure 1C), which matched observations made during mouse spermatogenesis. Besides, crotonylation was detected at H3K9, H3K18, H3K27 and H4K8, as was the case in mouse (Figure 1B). More precisely, H3K27cr was detected in association with non-modified, mono- and dimethylated H3K36, showing a rather constant stoichiometry during sporulation with a relative stoichiometry cr/ac of 1/10 for K_{27}modSAPSTGGVK_{36} and 1/100 for K_{27}cmodSAPSTGGVK_{36}me2KPHR_{40} (Figure 2F). WB analysis of this mark again detected a signal at the mass of H3 and a signal at about the mass of H4, corresponding to a cleaved form of H3 (Figure 2F). Knowing that the FLAG-tag used to perform nucleosome purification from yeast cells is located on the N-terminus of H3, it is possible that this cleaved form associates in the same nucleosome with a non-cleaved histone H3 molecule. Besides, H4K8cr was again detected in the sequence GGK_{8}crGLGK_{12}acGGAK_{16}acR and exhibited a drop from T0 to T4 and then some increase at the end of sporulation. Its stoichiometry never exceeded about 1/100 of the same sequence acetylated at H4K8 (Figure 2A). Finally, we again detected the pair of peptides H3K18cr-K23un and H3K18ac-K23un, yet reliable quantification was hampered due to co-elution of peptides of very close mass to the peptides of interest. Altogether, these results indicated that the crotonylation sites H4K8, H3K18 and H3K27 detected by proteomics were conserved during evolution. The relative cr/ac stoichiometries were about ten times higher in mouse spermatogenesis than in yeast sporulation. Some increase for H4K8cr was observable at the end of both processes and abundance variations were similar for H3K27cr in both organisms when considering the variant H3.3 in mouse.

Metabolomic analyses of acyl-CoA molecules in mouse spermatogenesis and yeast sporulation

While lysine acetylation is catalyzed by Histone Acetyl-Transferases p300/CBP from acetyl-CoA, crotonyl-CoA is the donor allowing lysine crotonylation by the same enzymes and may also be added non-enzymatically (70). We wanted to investigate whether the relative cellular amounts of the two metabolites and their possible variations during the studied processes could explain the observed relative stoichiometries of histone lysine acetylation and crotonylation. A panel of acyl-CoA molecules were analyzed in SC, RS, EC cells and in spermatozoa, in total mouse testis and in yeast cells at the four time points covering sporulation. It should be noted that the protocol used to collect fractions enriched in spermatogenic cells might perturb metabolism. Besides, histone acylation levels are correlated to nuclear concentrations in acyl-CoA metabolites, therefore the metabolomic analyses should ideally be performed on nuclear extracts (13,70). To our knowledge, however, no such analysis on a nuclear fraction has yet been successfully carried out (71,72). Having these limitations in mind, we quantified acetyl-CoA, crotonyl-CoA, as well as propionyl-CoA, butyryl-CoA, 2-hydroxy-butyryl-CoA, malonyl-CoA and glutaryl-CoA which are the donors for the corresponding acylations on yeast and mammalian histones (13,70) (Supplementary Figures S5A and B). Succinyl-CoA was also readily detected, yet its precise quantification was hampered by its noticeable chemical instability over successive LC-MS analyses.

Acetyl-CoA and crotonyl-CoA as well as butyryl-CoA and malonyl-CoA were successfully quantified in mouse spermatogenic cells (Supplementary Figure S5C, top panel). The cellular amount of all these metabolites decreased in the course of spermiogenesis. The relative cellular concentration of crotonyl-CoA/acetyl-CoA was about 1:100, to be compared to the relative stoichiometries of acetylation and crotonylation on histone lysine residues, which varied between 1:20 on H4K8 and about 1:3 on H3K27. In yeast (Supplementary Figure S5D), whereas acetyl-CoA was easily quantified, crotonyl-CoA was below the detection limit, leading to a probable ratio of acetyl-CoA/crotonyl-CoA well >100. This is in agreement with the fact that the observed relative stoichiometries of crotonylation as compared to acetylation were lower in yeast than in mouse cells. Beyond the two acyl-CoA being central in this study, other metabolites could be quantified. In particular, propionyl-CoA, butyryl-CoA, malonyl-CoA and glutaryl-CoA represented significant amount of total acyl-CoA in yeast cells and/or total testis extract (Supplementary Figure S5D). They might point to a significant modification level of cellular proteins and of histones; in fact butyrylation has already been functionally characterized on H4K5/K8 during mouse spermatogenesis (73).

The observation of an increased crotonylation of H3K27 and H4K8 residues during mouse spermatogenesis would at first sight appear contradictory with the decreased cellular amount of the donor metabolite. However, SC are much bigger cells than RS (i.e. at least 4 times bigger since SC are 4N and RS are 1N), and the amount of histones to be modified also decreases substantially from SC to EC, typically by a factor of 6 to 8 as estimated from our proteomics experiments. Representing the variations of each metabolite divided by the sum of all quantified metabolites indicated more stable amounts of acetyl-CoA and crotonyl-CoA (Supplementary Figure S5C bottom panel). Besides, our analysis of available RNA-seq data acquired on SC, RS and EC cells (52) suggest that p300/CBP, which was shown to catalyze acetylation and crotonylation (13), was progressively less expressed (74) (Supplementary Figure S5E). Interestingly a recent work provides a clear demonstration that the activity and specificity of p300/CBP is enhanced by the p300/CBP activator NUT, a factor that is specifically expressed in late developing male germ cells

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

(33). In addition, HDAC1/2/3 that catalyze deacetylation and also decrotonylation (21), and SIRT1/2/3 that are also able to remove crotonylation (20) similarly became less transcribed. Besides, ACSS2 (Acyl-CoA Synthetase Short Chain Family Member 2) which produces crotonyl-CoA from crotonate also exhibited decreased expression (75). Assuming a good correlation between mRNA and protein abundances, this data combined with our results do not allow concluding whether an increased or a decreased crotonylation is favored. One more actor contributing to balancing crotonylation levels was recently described: the protein chromodomain Y-like (CDYL), which is a crotonyl-CoA hydratase converting crotonyl-CoA into β-hydroxybutyryl-CoA and negatively impacts the crotonylation level of H3K27 and H4K8, as well as of H3K9 and H2BK12 (71). More precisely, CDYL protein amounts decrease in elongating/condensing spermatids, which would favor the increased crotonylation of H3K27 and H4K8. CDYL was also suggested to play a role in the expression of sex-chromosome-linked genes in post-meiotic cells by impacting Kcr levels at gene promoters.

Selection of a mark for genome-wide localization by ChIP-seq

Given the increased crotonylation of H4K8 and H3K27 during mouse spermatogenesis, it was tempting to speculate that these marks could contribute to the active transcription of specific genes in RS compared to SC cells, and particularly of genes on the X and Y chromosomes escaping from Meiotic Sex Chromosome Inactivation (MSCI). H4K8 looked particularly interesting because it was formerly studied in mouse spermatogenesis when modified by acetylation (45), 2-OH-isobutyrylation (8) and butyrylation (73). We sought to perform ChIP-seq analysis of H4K8cr in SC and RS cells, yet did not obtain reproducibly detected peaks. Second, to our knowledge, genome localization of H3K27cr has not been obtained in any system yet, even though an attempt was performed in cultured NCI-H1299 cells derived from lung cancer (76). In our proteomic analyses, H3K27cr was mostly detected within variant H3.3, possibly in association with H3K36me2. This combination rendered H3K27cr very attractive for several reasons. First, variant H3.3 is critical for proper spermatogenesis: its deletion leads to distorted chromatin reorganization and male infertility (77). Second, H3K27ac is described to be an active mark both at promoters and at distal enhancers, where it occurs in association with H3K4me1. Third, the combination H3 K4me1/K27ac/K36me2 has recently been suggested to mark super-enhancers that particularly concentrate transcription factors and co-activators to induce the expression of distant genes (78). We therefore sought to decipher the roles of H3K27cr at promoters and at distal enhancers to better understand its putatively complementary roles to those of H3K27ac.

Localization of H3K27cr on the genome

Interest of keeping ChIP-seq reads mapped at multiple positions on the genome. ChIP-seq analyses were performed on two biological replicates of SC and RS cells using the antibody raised against H3K27cr. Chromatin from EC cells is rarely analyzed by ChIP-seq, because of its peculiarly compacted structure and its heterogeneous nature, which renders the nucleosome extraction for ChIP-seq analysis difficult (5,73,69). Given that chromosomes X, Y, 5 and 14 contain multicopy genes expressed during spermiogenesis (79), in order not to under-estimate the ChIP-seq signals associated with these evolutionary conserved sequences, we kept reads mapped at both unique and multiple positions on the genome. This mapping strategy allowed us to gain a total of 1105 genes in the four pooled datasets when compared to the conservation of only uniquely mapped reads (Supplementary Figure S6). Among these genes, 150 and 170 were located on the X and Y chromosomes, respectively. In particular, we were able to quantify with the H3K27cr signals 16 out of the 22 TSS located in amplicons described on the X chromosome (79).

Positioning of H3K27cr with respect to H3K27ac, to Kcr and to variant H3.3: H3K27cr is notably present on distal enhancers and on the Y chromosome in RS cells. Very reproducible peak intensities were detected for the H3K27cr mark in both RS and SC cells (Supplementary Figure S7). We verified that the majority of our ChIP-seq peaks were included in those arising from an antibody raised against Kcr, by considering the data from publication (5) (Supplementary Figures S8 and S9, for all peaks and those located at promoters). We similarly compared the distribution of H3K27cr peaks with that of H3K27ac obtained in (46) (Supplementary Figures S8 and S9). The two marks overlapped more significantly at promoters, by 77% to 82% in RS cells and 75% to 93% in SC cells. We also compared the peak overlap observable between the two biological replicates of ChIP-seq analyses performed on each of the two marks H3K27cr and H3K27ac. This allowed observing that ChIP-seq signals obtained on the two distinct marks were more divergent (Supplementary Figures S8 and S9).

Figure 3A shows a browser view of the ChIP-seq signals detected for H3K27ac, H3K27cr and H3K4me1 in SC and RS cells over key genes expressed during spermiogenesis, namely genes coding for the transition protein *Tnp2* and for two protamines *Prm1* and *Prm3*. Both H3K27ac and H3K27cr are present at the promoters of these genes, and H3K27cr is well visible over a probable distal enhancer marked by H3K4me1. The global distribution of H3K27cr peaks on promoters (defined as ±500 bp windows around the TSS), introns and distal intergenic regions was obtained, and compared to the distributions of Kcr and H3K27ac peaks (Figure 3B). The distribution of H3K27cr peaks on functional genomic regions appears to be intermediate between those of Kcr and of H3K27ac. In particular, whereas about two thirds of H3K27ac peaks lie at promoters, less than half of H3K27cr peaks do so, but a higher proportion of them localize at distal intergenic regions. At promoters, H3K27cr appeared more enriched downstream of TSS both in SC and RS cells, similarly to H3K27ac and H3K4me3 that signs active transcription (Figure 3C). At putative distal enhancers predicted from H3K4me1 mark, H3K27cr overlapped with H3K27ac and cr/ac signal ratios were much above those observed at promoters (Figure 3C).

Our proteomic analyses indicated the preferential presence of H3K27cr, and to some extent of H3K27ac, on

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

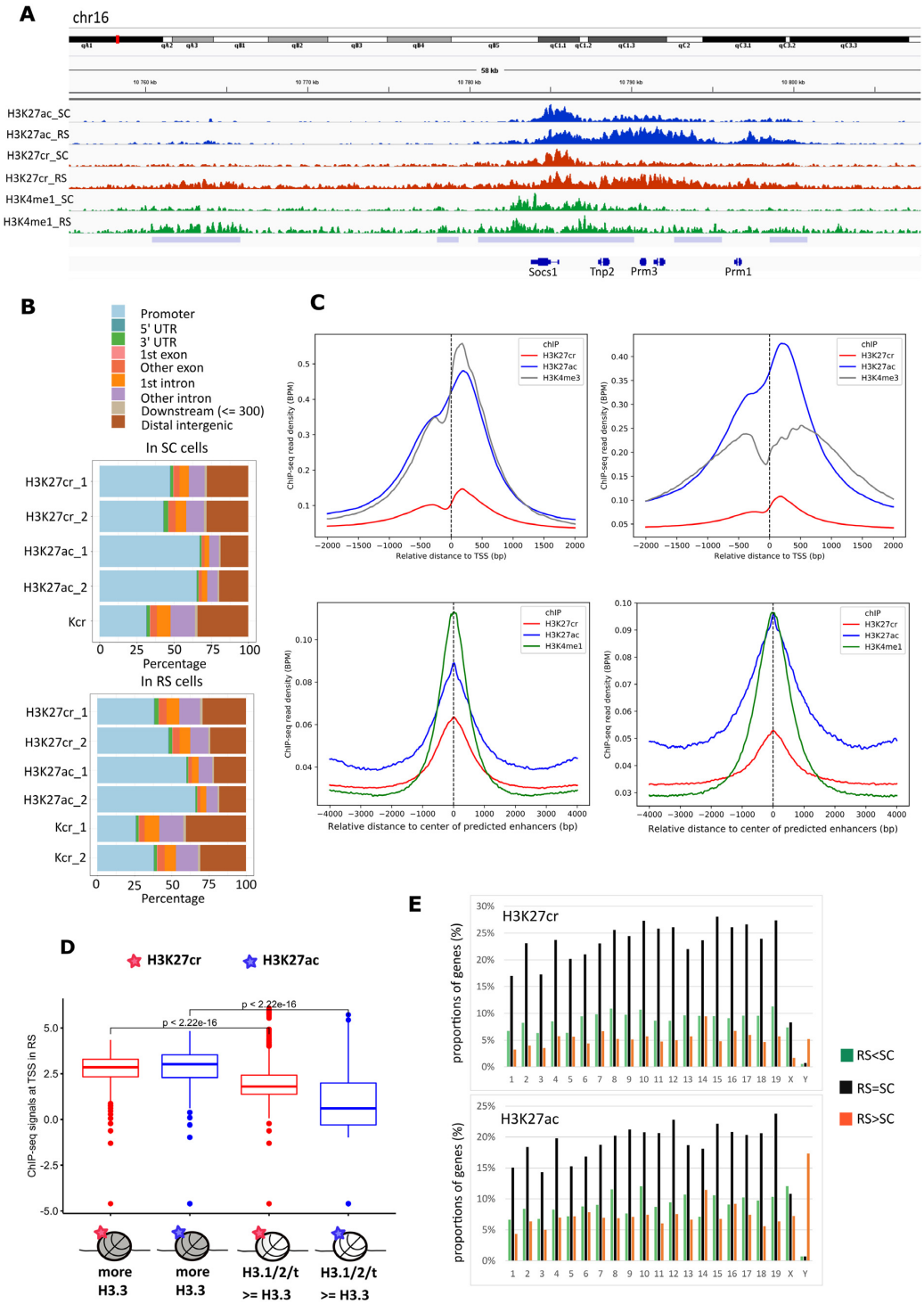

Figure 3. (**A**) Genome browser view showing the ChIP-seq signals detected for H3K27cr, H3K27ac and H3K4me1 in SC and RS cells around genes coding for the transition protein Tnp2 and protamines Prm1 and Prm3. Intense marking by H3K4me1 and H3K27cr in RS cells indicates a probable active distal enhancer on the left side. (**B**) Distribution in SC and RS cells of ChIP-seq peaks detected for the mark H3K27cr in terms of various genomic region types including promoters (defined as +/− 500 bp windows around TSS) and distal intergenic regions. (**C**) Curves showing the mean reads density of the indicated histone modifications around all known TSS, or around enhancers predicted by H3K4me1. Reads densities are calculated within a 20 bp sliding window and normalized in BPM. (**D**) H3K27ac and H3K27cr co-localize more with H3.3 than with H3.1/2/t. The ChIP-seq signals detected around TSS for H3.3 and for H3.1/2/t were split into two groups: one group in which ChIP-seq signals are higher for H3.3 than for H3.1/2/t (signal(H3.3) > 1.5× signal(H.1/2/t), indicated on the figure by 'more H3.3') and the other group corresponding to the remaining signals (signal(H3.3) ≤ 1.5 × signal(H3.1/2/t), indicated on the figure by 'H3.1.2.t ≥ H3.3'). The ChIP-seq signals detected for H3K27ac and H3K27cr in these two groups are represented in boxplots. (**E**) Distribution over chromosomes of H3K27cr and H3K27ac peaks at promoters. The proportions of genes (%) bearing H3K27cr (top panel) or H3K27ac (bottom panel) are indicated for each chromosome. Three sub-categories 'RS<SC', 'RS = SC' and 'RS > SC' were defined by considering that RS/SC ratios of ChIP-seq signals (in BPM per start position) were unequal when they were <0.66 or >1.5.

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

the variant H3.3 (Figure 2B and E). Making use of previously published ChIP-seq data on canonical H3 and H3.3 acquired in RS (50), we observed that a strong signal was concomitantly detected for H3K27ac and H3K27cr within nucleosomes mostly containing H3.3 (Figure 3D). In contrast, promoters occupied by balanced amounts of H3.1/2/t and H3.3 or mostly occupied by H3.1/2/t exhibited lower H3K27cr and H3K27ac. Our observation that the two active histone marks are present more on H3.3 is in agreement with the transcription-coupled eviction of canonical histones described in RS: Erkek *et al.* indeed described a more pronounced removal of H3.1/H3.2 from the TSS of highly expressed genes (50).

In the course of spermatogenesis, genes coded by the sex chromosomes undergo transcriptional repression, a process named MSCI for Meiotic Sex Chromosome Inactivation (80). After meiosis, despite persistence of a repressive chromatin environment characterized by H3K9me3 mark, a significant proportion of genes are re-activated or *de novo* expressed in RS. A finely tuned orchestration of histone PTMs is at the basis of this phenomenon (45,81,82). In the original paper describing crotonylation in mouse spermatogenesis, Tan *et al*. highlighted that a very large fraction of Kcr signals of higher intensity in RS than in SC cells was enriched on the X chromosome, and that this scenario correlated with post-meiotic gene expression activation (5). This observation was later confirmed by Namekawa's team (45) who found a higher intensity of crotonylation in X-linked genes escaping repression. Additionally, Liu *et al.* showed that this specific crotonylation is essential for X-linked genes reactivation (71). Plotting the distribution of H3K27cr ChIP-seq peak intensities on the chromosomes did not exhibit such an enrichment on the X chromosome (Figure 3E). Our data however showed a particular trend for the Y chromosome, with 29 out of 36 marked genes exhibiting an RS/SC ratio of H3K27cr signal >1.5. A similar observation was made for H3K27ac, with 97 out of 105 genes showing RS/SC ratios >1.5 for this mark (Figure 3E). All 29 genes exhibiting an increased H3K27cr marking also showed increased H3K27ac. It is worth noting that those genes are highly similar and belong to the same family (*Sly*, *Ssty*, also called multicopy genes). Hence, this observation suggests that site-specific crotonylation and acetylation on lysine 27 from H3 could be involved in regulating the activity of a family of genes located on the Y-chromosome.

H3K27cr is a mark of active promoters

H3K27cr and H3K27ac often evolve in a synchronized manner from SC to RS at promoters, but also mark preferentially specific subsets of genes. Most of the genes exhibiting increased Kcr in RS compared to SC were originally shown to be more expressed post meiosis (5,45). We then evaluated whether a similar trend was observed for the more specific mark H3K27cr. We quantified ChIP-seq signals for H3K27cr and H3K27ac within ±500 bp windows around TSS. An increased transcription level was indeed associated with an increase of H3K27cr mark between the two cellular stages; a similar trend was observed for H3K27ac (Figure 4A and B). However, it should be kept in mind that H3K27cr and H3K27ac often coexist at the same TSS, with ~80% of peaks of one mark co-localizing with peaks of the other mark (Supplementary Figure S9).

To assess further the effects of H3K27ac and H3K27cr on transcription, we classified genes according to their RS/SC ratios of H3K27cr and H3K27ac signals at promoters and annotated their corresponding gene expression ratio (Figure 4C). We identified several gene clusters associated with different trends. Cluster C1 contains genes with the highest gene expression increases from SC to RS cells, together with the highest gain of both marks at their TSS. This cluster is notably enriched in genes involved in spermatogenesis, spermatid differentiation and sperm motility. In contrast, cluster C5 corresponds to genes that globally show decreasing expression and whose promoters become less marked by both histone PTMs from SC to RS stages. From this list of genes, GO terms related to meiosis were enriched, which is in agreement with the fact that RS are post-meiotic cells. Rho and Ras signal transduction and chromatin remodeling were also enriched in this cluster, in line with previous reports (83,45). Yet some gene TSS harbor an inverted evolution of H3K27cr and H3K27ac marking from SC to RS cells, leading to variable mRNA expression outcomes (clusters C3 and C4). In this case, one hypothesis is that mRNA expression might be regulated by the respective amounts of each mark at the gene promoter.

We then provided to the program STRING (84) the list of genes the promoter of which only harbored detectable H3K27cr peaks together with H3K4me3, to search for enriched molecular functions, pathways and protein complexes (Supplementary Figures S10). Both in SC and RS cells, the Takusan gene family was significantly overrepresented, with a total of 97 members out of 209 recorded in STRING (IGV screenshots on three genes are shown in Supplementary Figure S10E). This family is composed of multi-copy genes located on chromosomes 5 and 14. We quantified the signal intensities (BPM) for H3K27cr and H3K27ac at Takusan gene promoters. The normalization in BPM brings the summed ChIP-seq signals over all promoters to one million for each mark, which erases any information of H3K27ac and H3K27cr relative stoichiometries, yet allows observing the accumulation of one mark at some specific genomic regions. We observed that at both SC and RS stages, the sum of normalized signals measured for H3K27cr at Takusan gene promoters was more than three times higher than for H3K27ac (Supplementary Table S2). In addition, 55 out of 58 genes exhibiting a significant expression variation between SC and RS cells (P-value < 0.01) were more transcribed in RS cells, which usually correlated with RS/SC H3K27cr intensity ratios above 1 (Supplementary Table S2). Beside Takusan genes, 72 genes coding for proteins of the cytoskeleton were also pinpointed in RS cells as bearing specifically H3K27cr peaks. The expression of such genes needs to be precisely orchestrated during spermiogenesis, given the dramatic changes of cellular shape until differentiation into spermatozoa. Among them were six members of the axonemal dynein complex, including Dnah2, Dnah12 and Dnah17 that are implicated in sperm flagellar assembly. We next performed a similar analysis with genes that only presented H3K27ac peaks at their promoters (Supplementary Figure S10). Several of these genes coded for Spindlin/spermiogenesis-specific proteins

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

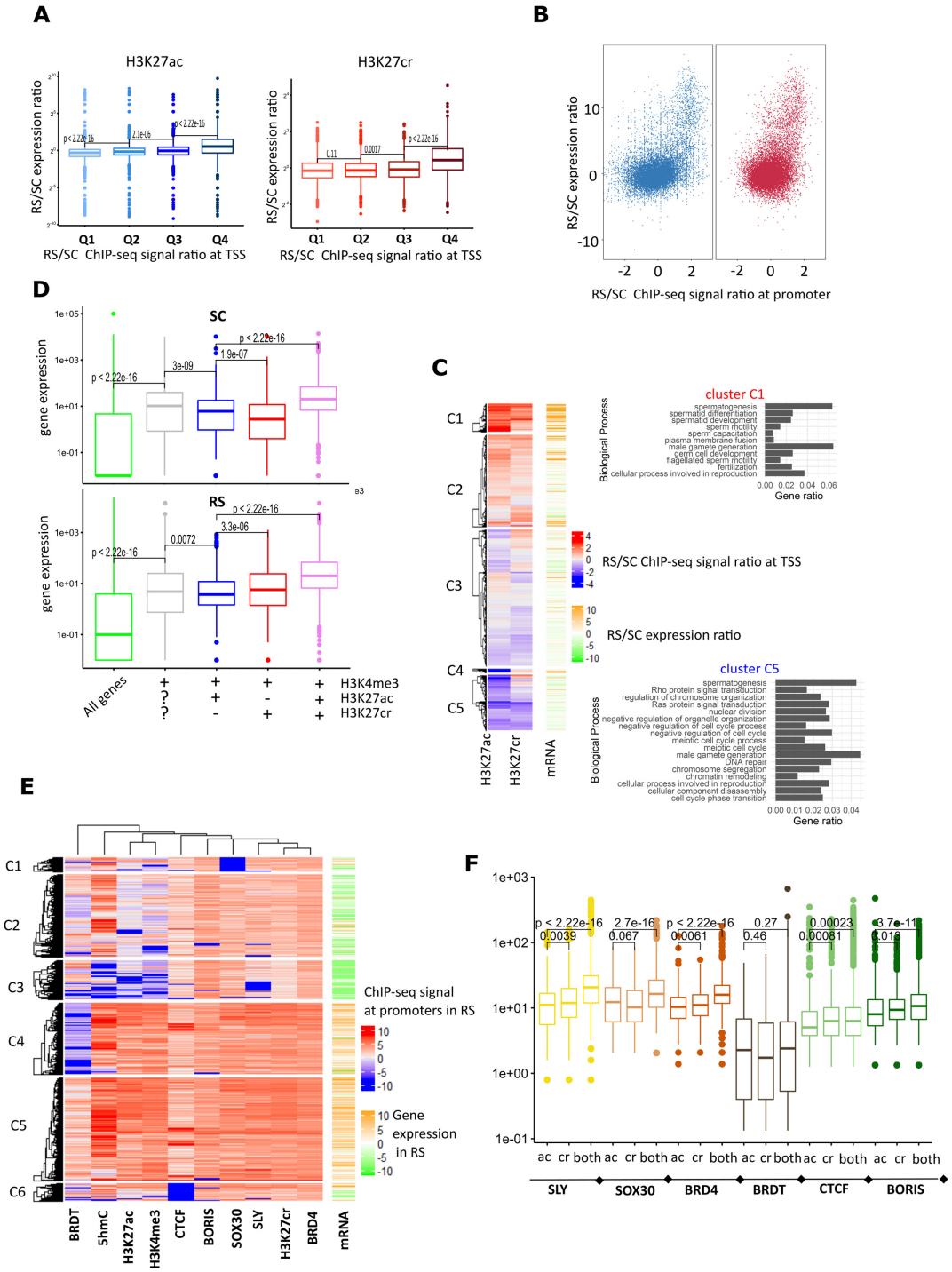

Figure 4. All gene expression and ChIP-seq signal values are expressed in $\log_2$(TPM + 0.01) and log2(BPM+0.01), respectively. (**A**) ChIP-seq signals detected at promoters for H3K27ac and H3K27cr marks were organized into quartiles of increasing RS/SC ratios. The distributions of RS/SC mRNA abundance ratios corresponding to the four sets of genes were plotted as boxplots for each mark. (**B**) Evaluation of the correlation between RS/SC mRNA ratios and RS/SC fold change of ChIP-seq signal for H3K27ac and H3K27cr marks considering all quantified mRNA individually. (**C**) Heatmap matching the RS/SC ratios of ChIP-seq signals detected at promoters for H3K27ac and H3K27cr with the RS/SC mRNA abundance ratios of the corresponding genes. GO terms were extracted from clusters C1 and C5 using ClusterProfiler (110) within ProteoRE with *P*-values below 10^{-3}. (**D**) Distribution of gene expression levels in SC (upper panel) and RS (lower panel) when considering successively: all genes, genes with H3K4me3 mark at their promoter, genes bearing additionally H3K27ac mark or H3K27cr mark or both active marks. (**E**) Heatmap matching the presence in RS cells of the marks H3K27ac and H3K27cr with a series of transcription factors/regulators at gene promoters, with 5hmC signals over exons of the corresponding genes and with the expression levels of the genes. (**F**) Distributions of ChIP-seq signals detected for SLY, SOX30, BRD4, BRDT, CTCF and BORIS at gene promoters marked solely by H3K27ac peaks ('ac' label), or solely by H3K27cr peaks ('cr' label), or bearing both peaks ('both' label). Only the subset of genes exhibiting increased expression from SC to RS were considered here; results for all genes and for the quartiles of most upregulated genes are shown in Supplementary Figure S11.

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

which are located on the Y chromosome. Among them, the mRNA of 13 genes were reliably quantified and were all *de novo* expressed in RS cells. Interestingly, for all 13 genes, both histone marks increased from SC to RS, by a factor of 6.5 for H3K27ac and of 3.0 for H3K27cr on average. This gene family thus belongs to cluster 1 of Figure 4C, and contributes to the RS/SC increase of both marks particularly observed on the Y chromosome (Figure 3E).

We finally plotted the distribution of transcript abundance in SC and RS cells separately, depending on whether a peak had been detected at their TSS for H3K27ac and for H3K27cr. We performed this assessment when the former PTMs occurred together with the active mark H3K4me3 (Figure 4D). We observed that in both SC and RS cells (i) genes marked by H3K27ac but not by H3K27cr were less expressed than the whole group of actively transcribed genes characterized by the presence of H3K4me3 and (ii) the highest gene expression levels were reached when both marks co-existed at the same TSS. Besides, in SC cells H3K27ac alone would more strongly promote transcription than H3K27cr alone, whereas the opposite trend prevailed in RS cells (Figure 4D). These results suggested the existence of different transcriptional factors/regulators in SC and RS.

Combined H3K27cr and H3K27ac co-localize with higher amounts of transcription factors and regulators, yet each mark exhibits selective association. We compared the co-localization at promoters (i.e. ±500 bp from the TSS) of H3K27cr and H3K27ac with transcription factors (TFs) and transcription regulators involved in spermiogenesis. ChIP-seq data in RS are available for BRDT (51), an essential regulator of gene expression during meiosis prophase (85,86), for the TF SOX30 (49), for the transcription regulator SLY (39) and for the two closely related chromatin-binding factors CTCF and BORIS that contribute to shaping the 3D structure of chromatin by bridging distant regions (47,87). We also included the bromodomain-containing protein BRD4 which was shown to be highly associated with the promoter of genes expressed in RS; BRD4 participates in the regulation of transcription elongation by RNA polymerase II by its binding to acetylated histones, particularly to the poly-acetylated N-terminus of histone H4, and to Mediator (48,88). Besides, we had formerly verified that SLY ChIP-seq peaks largely overlapped with Kcr peaks (39), which encouraged us to scrutinize more specifically the overlap with H3K27cr. We finally also included the signals measured over exons for 5-hydroxymethylcytosine (5hmC), which was shown to correlate with gene expression levels during mouse spermatogenesis (52). The signals quantified around TSS for H3K27cr, H3K27ac, SLY, SOX30, BRD4, BRDT, CTCF and BORIS, and over exons for 5hmC, resulted in a heatmap containing two balanced clusters of genes that are globally highly expressed in RS and whose promoters homogenously harbor all of the considered marks, except BRDT which is of either high or low level (C4 and C5 in Figure 4E). This observation is in line with the fact that only about half of the genes requiring this protein for their activation bear it at their TSS (51). Two other clusters, C2 and C3, are characterized by genes of low expression level, whose promoters accordingly harbor little H3K4me3 and H3K27ac. Nonetheless, these gene promoters also bear a significant amount of H3K27cr, together with SLY, SOX30 and BRD4, and for a large fraction of them, also bear BORIS and CTCF.

To more quantitatively assess whether some transcription factors and regulators would preferentially associate with H3K27cr rather than with H3K27ac, we plotted the distribution of their signals detected at promoters bearing only H3K27ac peaks or only H3K27cr peaks or peaks for both marks (Figure 4F and Supplementary Figure S11). More precisely, we considered successively (i) all the genes under the control of promoters matching the former criteria, or (ii) those genes exhibiting an increased expression from SC to RS (Figure 4F), and finally (iii) the quartile of most induced genes (highest RS/SC mRNA ratios). The highest amount of each chromatin binder was always observed at promoters bearing simultaneously H3K27ac and H3K27cr, in agreement with the observations made on Figure 4E. SLY was globally more associated with H3K27cr than with H3K27ac, except when considering the quartiles of most induced genes (Supplementary Figure S11). By contrast, SOX30 appeared more associated with H3K27ac when considering these subsets of most induced genes. BRD4, BORIS and CTCF were more associated with H3K27cr than to H3K27ac when focusing on upregulated genes (Figure 4F and Supplementary Figure S11). Finally, BRDT exhibited the least signals, in agreement with its favored presence in intergenic regions (51). We inquired further on SOX30 by making use of the RNA-seq and ChIP-seq data published on precisely fractionated spermatogenic cells (49). This report established a list of genes exhibiting a reduced expression upon deletion of SOX30. We compared them to our lists of genes whose promoters were only attributed H3K27ac peaks, or H3K27cr peaks or the combination of both. The vast majority of overlapping genes bore both active marks, in agreement with their combination correlating with favored TF attachment (Supplementary Table S3). As for X-encoded genes expressed in RS (whether re-activated after MSCI or *de novo* expressed in RS) (49), our analysis showed that most of these genes were predominantly controlled by H3K27ac (Supplementary Table S3).

H3K27cr is a mark of active distal enhancers

In our proteomic analyses, H3K27cr was largely detected in association with H3K36me2 (Figure 2E). The latter mark was described to be rarely present at promoters (89) and was suggested to mark enhancers and even more super-enhancers (78) that concentrate TFs and other transcription co-regulators, to define the expression of nearby cell-type-specific genes and thus lead to cell identity determination (78,90,91). Enhancers are distant regions from gene promoters and are characterized by the presence of H3K4me1; among them, active enhancers are classically defined as bearing H3K27ac, which is causative to enhancer activity (92,93). We then inquired whether H3K27cr might also mark active enhancers and super-enhancers.

H3K27cr and H3K27ac often co-exist at enhancers. We looked for H3K4me1 peaks present at >3000 bp from TSS to localize putative distal enhancers in SC and RS cells,

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

and searched for H3K27ac and H3K27cr peaks overlapping with them. As a whole, 2.7 and 1.7 times more enhancers bore H3K27cr compared to H3K27ac in SC and RS cells, respectively (Figure 5A). A total of 1067 enhancers bore only peaks for H3K27ac, 3531 enhancers only for H3K27cr and 5467 harbored peaks for both marks. We observed that H3K27ac was detected with higher signal intensity when H3K27cr was also present (two first boxplots of Figure 5B), and reciprocally for H3K27cr signal (two last boxplots of Figure 5B; see also Supplementary Figure S12 for SC cells). In other words, the presence of both acetylation and crotonylation on H3K27 was favored in a subset of enhancers.

Distal enhancers bearing combined H3K27ac and H3K27cr are associated with higher concentrations of transcription factors and regulators. We observed above that promoters bearing H3K27ac and H3K27cr concomitantly were associated with the highest amount of TFs and chromatin-binding proteins. We then asked whether a similar rule would apply within distal enhancers. Focusing on RS cells in which SLY is specifically expressed, we localized SLY, SOX30, BRD4, BRDT, BORIS and CTCF at the putative enhancers bearing solely H3K27ac peaks, solely H3K27cr peaks, or peaks for both marks (Figure 5C). Those enhancers bearing simultaneously H3K27ac and H3K27cr were more systematically associated with all the chromatin-binding proteins, except SOX30 that was equally present on enhancers exhibiting only H3K27ac peaks (Supplementary Table S4 for Mann-Whitney test results). CTCF, together with cohesin, have long been shown to bridge enhancers and their target promoters by forming chromatin loop anchors (94,95,90). A high CTCF signal was harbored by a minority of enhancers; their genomic coordinates were provided to the program GREAT (96) to study the ontology of neighbor genes. The subset of 23 enhancers solely bearing H3K27ac did not allow extracting any GO term. By contrast, the 318 enhancers bearing only H3K27cr led to highlighting 'positive regulation of MAP kinase activity' (most notably by MAP3K7, FGFR1, FGFR4, VEGFA, FL1, KDR and NTF3) (Supplementary Figure S13). In addition, the cluster of 486 enhancers marked by all transcription regulators and bearing both H3K27ac and H3K27cr allowed extracting genes involved in the regulation of nuclear division and in Rac protein signal transduction (Supplementary Figure S13). The identification of genes involved in MAPK signaling, in particular MAP3K7 and FGFR1 (97,98), is in line with previous reports in the context of spermatogenesis.

Because several enhancers often share the same nearest gene the expression of which they may regulate, we asked whether the one bearing the maximum signal for both H3K27ac and H3K27cr might concentrate more chromatin-binding proteins. Following this gene-centric reasoning, we counted that about 70% of enhancers bore the maximum signal for both H3K27ac and H3K27cr in SC and RS cells. Because active enhancers can get transcribed, we evaluated the presence of H3K4me3. As much as 73–78% of enhancers exhibiting simultaneously the maximum of H3K27cr and H3K27ac also bore a maximum signal for H3K4me3 in SC and RS cells. We assessed the co-concentration of BRD4, SOX30, SLY, BORIS and BRDT on the enhancers bearing the maximum of H3K27ac and H3K27cr. BRD4 was tested first because this protein was described to induce gene expression in RS (48) and determined to be enriched at super-enhancers in human ESCs (88); SOX30 was next required due to its established role in the proper progression of mouse spermatogenesis (49,99–101). This progressive skimming led to the selection of 4678 enhancers, of which 99% were maintained when requiring a maximum signal for CTCF (Supplementary Figure S14). This final set represented 48% of the 9742 initially considered enhancers bearing a maximum signal for both H3K27ac and H3K27cr in RS cells. By contrast, among the 4094 and 4069 enhancers decorated with either a maximum signal for H3K27ac or for H3K27cr, only between 25% and 29% overlapped with the maximum of any single chromatin binder, with the exception of SLY binding 39% of H3K27cr-marked enhancers and of BRDT binding 34% of H3K27ac-marked enhancers. Requesting having the maximum signal for all chromatin binders led to the selection of <0.5% of the enhancers bearing solely the maximum for H3K27ac or H3K27cr. These assessments indicated that the combined strong marking of enhancers by H3K27ac and H3K27cr is most often associated with the accumulation of transcription regulators and of CTCF.

Enhancers marked by H3K27cr may be more associated with highly expressed genes than those solely marked by H3K27ac. Assuming that a given enhancer modulates the expression of the closest gene (92,93,95), we plotted the distribution of gene expression levels with regard to the presence of each active mark on the associated enhancer(s). Considering the four scenarios of genes depicted in Figure 5D, we observed that genes under the control of H3K27cr-modified enhancers, alone or in combination with H3K27ac, were significantly more expressed than those under the control of enhancers solely modified by H3K27ac.

Assessing a possible synergy between marks at promoters and distal enhancers to induce the expression of spermiogenesis-related genes. The combined increase of both marks at promoters was observed to contribute to highest gene expression induction (Figure 4D). We asked whether genes regulated by the two active marks at their promoters also matched doubly marked enhancers. We considered the various enhancers sharing the same nearest gene and selected the one harboring the maximum signal of each histone mark. We then handled data in a stringent way, by requesting that a maximum of signal for H3K27ac and H3K27cr be present at the considered enhancers at both SC and RS stages. We assessed the correlation between gene expression induction and an increase between SC and RS stages of H3K27ac and H3K27cr marks at both promoters and enhancers. This evaluation led to the heatmap of Figure 5E, in which cluster C2 is mostly populated by over-expressed genes (Supplementary Table S5). The program STRING extracted from this cluster the GO terms 'sperm part', 'acrosomal vesicle' and 36 Takusan genes (Supplementary Figure S15). Clusters with overall stable or decreased marking at enhancers (cluster C4) or at promoters (cluster C3) were both associated with a lack of induced gene expression, which indicates that gene expression from SC to RS

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

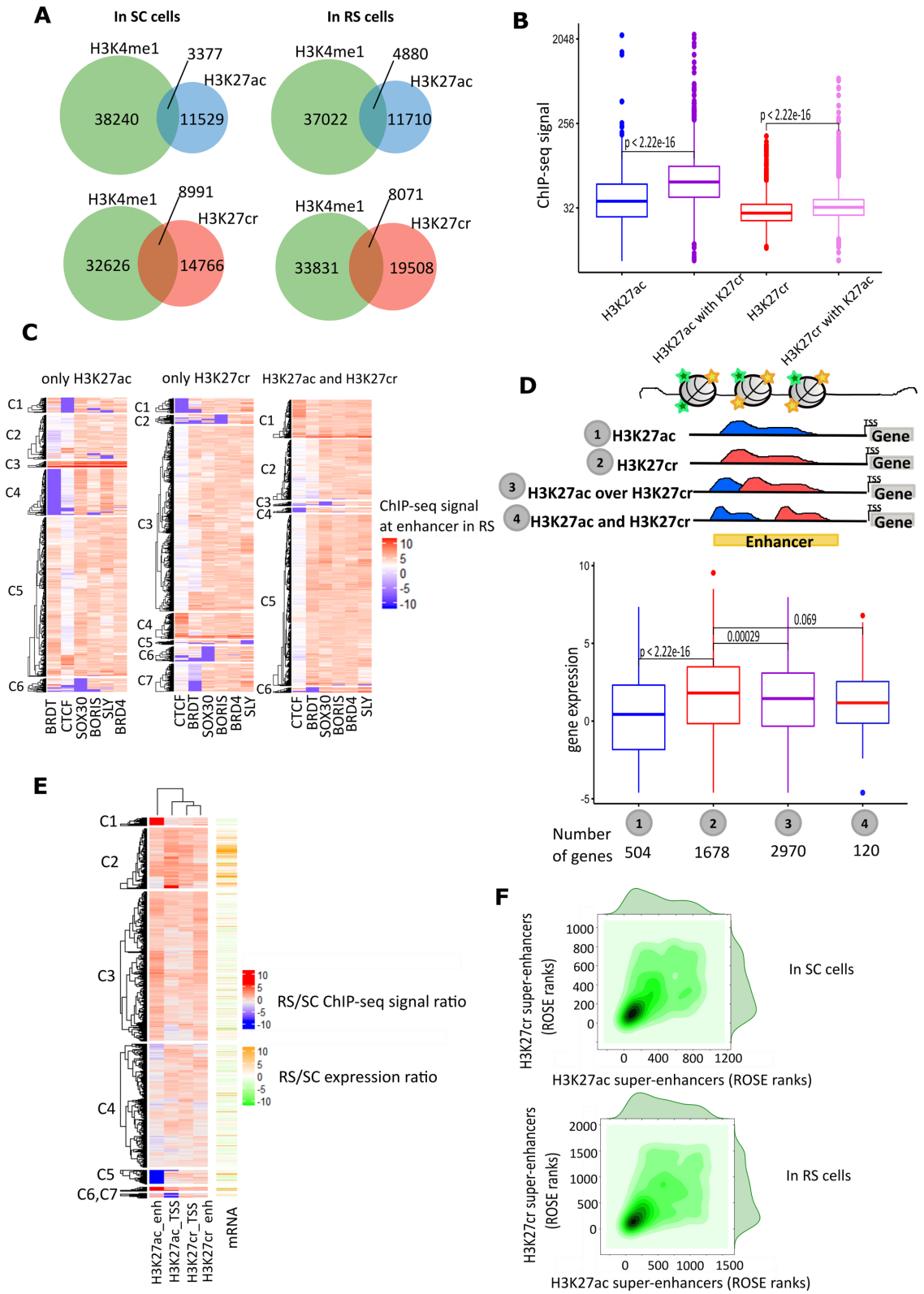

Figure 5. (**A**) Overlaps of H3K4me1 peaks localized at >3000 bp from gene TSS with H3K27ac peaks and H3K27cr peaks. (**B**) Distribution of ChIP-seq signals for H3K27ac and for H3K27cr on enhancers bearing only one mark or both concomitantly in RS cells; 'H3K27ac with H3K27cr' labels the distribution of H3K27ac signals at enhancers where H3K27cr peaks are also detected, and *vice versa*. (**C**) Heatmaps obtained in RS stage showing the signals detected for chromatin-binding proteins on enhancers only marked by H3K27ac peaks (left), or by H3K27cr peaks (middle), or by peaks for both modifications (right). (**D**) Distribution of gene expression levels depending whether they are associated solely to enhancers marked by H3K27ac, or by H3K27cr, or to both kinds of enhancers being either overlapping or separated. (**E**) Heatmap showing the correlation between RS/SC mRNA ratios and the RS/SC ratios of H3K27ac and H3K27cr marks at both promoters and distal enhancers. (**F**) Correlation between the ROSE ranks of super-enhancers determined using H3K27ac or H3K27cr in SC and RS cells.

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

stage is regulated by histone marks at both promoters and enhancers. In conclusion, a combined increase of H3K27cr and H3K27ac at both promoters and distal enhancers induces the expression of genes that ensure the proper differentiation of round spermatids into final sperm, notably genes involved in sperm motility and fertilization.

First-rank super-enhancers bear both H3K27ac and H3K27cr and some sperm super-enhancers are pre-defined in RS cells. Finally, we used the program ROSE (rank ordering of super-enhancers) to define super-enhancers among distal enhancers using the classically considered mark H3K27ac or using H3K27cr (91,102). In SC cells, 930 and 871 super-enhancers were determined from H3K27ac and H3K27cr, respectively, whereas 1296 and 1625 were so in RS cells (Supplementary Figure S16A). About half of super-enhancers determined from H3K27cr thus appear *de novo* in RS cells (Supplementary Figure S16B). This is in agreement with histone crotonylation increasing toward the end of spermiogenesis. As many as 962 super-enhancers bore both marks in RS cells and 562 did so in SC cells. A representation in 2D density plots of the ROSE ranks of super-enhancers determined for both marks indicated that the subset of super-enhancers bearing the most intense marking for H3K27ac also bore strongest H3K27cr in both SC and RS (Figure 5F). To assess whether stage-specific super-enhancers might be associated with genes exhibiting dramatic expression changes from SC to RS stage, we looked for (i) genes matching RS-specific super-enhancers and showing a log2(RS/SC) expression ratio above 3, and (ii) genes matching SC-specific super-enhancers and showing a $\log_2$(RS/SC) expression ratio below –3. A total of 115 and 136 genes corresponded to (i) while considering super-enhancers determined from H3K27ac and from H3K27cr, respectively (Supplementary Table S6). We established that there was a significant association between the specific presence of a super-enhancer at RS stage and the highly increased expression of the closest gene (Chi-squared test, P-value = 0.002 for H3K27ac and P-value = 0.017 for H3K27cr). Besides, 14 genes corresponded to (ii) when considering H3K27ac to define super-enhancers; a significant association between SC-specific super-enhancers and strongly decreased gene expression could be established (Chi-squared test, P-value = 0.0006, and Fisher exact test, P-value = 0.0008). Given that only 77 super-enhancers determined from H3K27cr were SC-specific, no association analysis with gene expression could be performed. Supplementary Table S6 contains Lyzl4, Pmfbp1 and Ropn1 that allowed highlighting the term 'flagellum' using the program STRING. Interestingly, the two first proteins are related to super-enhancers marked by both H3K27ac and H3K27cr. Among genes under the probable control of H3K27cr-labeled super-enhancers lies H2afb1, a histone variant that is incorporated in nucleosomes during late spermatogenesis and facilitates replacement of histones by protamines. Altogether, this evaluation indicated that super-enhancers defined by H3K27ac were associated with the stage-selective expression of genes in both SC and RS cells, while super-enhancers defined by H3K27cr were only associated with the expression of RS-selective genes. Finally, we compared the genomic coordinates of the 645 sperm super-enhancers established in (31) to those of super-enhancers determined here. In SC, sperm super-enhancers overlapped with 86, 92 and 75 super-enhancers established from H3K27ac, from H3K27cr or bearing both marks, respectively. In RS, these figures increased to 104, 138 and 95 super-enhancers (Supplementary Figure S16C). A fraction of sperm super-enhancers thus appears to be already built at the RS stage, and notably bears H3K27cr.

DISCUSSION

The recent discovery of various chemical structures possibly modifying histone lysine residues has opened up new avenues for better deciphering the intricate mechanisms of gene expression regulation. Among these modifications collectively called acylations, crotonylation was originally described in mouse spermatogenesis, and its abundance on histones was shown to increase during this differentiation process (5). More precisely, an increase from meiotic (SC) to post-meiotic cells (RS) of the global level of Kcr at promoters correlates with increased gene expression, particularly on the X chromosome. We mapped by proteomics the sites of histones H3 and H4 that bear crotonylation and observed a maximum stoichiometry of this mark at H3K27 and H4K8 at the end of spermiogenesis (EC stage). This result could not be easily explained by metabolomics analyses which did not indicate obvious changes of the quantifiable acyl-CoA metabolites in whole cells, including crotonyl-CoA, from SC to sperm. Ideally, metabolomics analyses of acyl-CoA should be performed on nuclear extracts, but both cell and subcellular fractionations are likely to perturb metabolism. Our data are however in agreement with the reported decreased amount of Chromodomain Y-like protein (CDYL), which converts crotonyl-CoA to β-hydroxybutyryl-CoA and thus negatively regulates crotonylation levels (71). Additional elements are still needed to grasp the regulation of histone lysine crotonylation, including the fraction of cellular enzymes being located in the nucleus, the enzymatic activities and the site-specific regulation (e.g. relative cr/ac stoichiometry is maximal at H3K27). Metabolomic analyses also highlighted relatively high concentrations of propionyl-CoA, malonyl-CoA and glutaryl-CoA in both yeast and mouse testis, which might lead to significant histone PTM levels and would warrant further investigation.

Western Blot analysis using an antibody against H3K27cr revealed the presence of this mark on a truncated form of H3, whose abundance surpassed the one on full-length histone H3 in sperm. First, these observations indicate that histones crotonylated on H3K27 are maintained in spermatozoa despite the near complete replacement of histones by protamines. In (71), the authors observed that overexpression of CDYL (and thus decreased level of crotonylated histones) led to a higher proportion of soluble transition proteins and protamines, suggesting that crotonylation would favor histone replacement by protamines. Our results on H3K27cr do not fit with this proposed mechanism, which may be relevant for other crotonylation sites, including H4K8cr that we did not detect in sperm. The modification H3K27cr might rather be compared to H4K5/K8buty: the bromodomain-

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

containing protein BRDT binds and removes histone H4 hyperacetylated at its N-terminus but binds more weakly H4 butyrylated at K5/K8, which slows down its eviction from the genome (73). Second, the detection of a cleaved histone form bearing H3K27cr completes the panel of clipped histones devoid of their N-terminal part observable in sperm, which include H3 methylated at Lys27 but interestingly enough, barely detectable H3K27ac (69). Further work is required to determine whether different roles are attributable to full-length and cleaved histone H3 crotonylated on Lys27, which crotonylated sites may actually favor replacement of histones by protamines, and which histone PTM readers are at work in this process. Some YEATS-domain proteins might play a role in histone removal (25–27).

Proteomic analysis revealed the co-occurrence of several PTMs on a given histone region and discriminating between histone variants differing by a few amino acids. H3.1/H3.2/H3.t and H3.3 are thus distinguishable by Ala31>Ser31 and we observed a complex interplay involving these two groups of H3 sequences, with H3K27 being acetylated or crotonylated in combination with H3K36 being non-modified or dimethylated. Most interestingly, we estimated that the relative stoichiometry between H3.3-K27cr/K36un and H3.3-K27ac/K36un was about 1/3 whereas H3.3-K27cr/K36me2 and H3.3-K27ac/K36me2 were of similar abundances. These ratios are well above the relative concentrations of crotonyl-CoA to acetyl-CoA measured by metabolomics, and may indicate the favored addition of crotonyl onto H3K27 by yet to determine mechanisms.

We obtained the genomic localization of H3K27cr by ChIP-seq analysis of SC and RS cells and compared it to H3K27ac ChIP-seq data and transcription dynamics between SC and RS. Both active marks exhibited a significant overlap, yet a larger fraction of H3K27cr was located on distal intergenic regions when the majority of H3K27ac resided at promoters. Acetylation and crotonylation being two active marks, they may co-exist on the same nucleosome. To address such a question, a very elegant strategy combining MNase digestion, ChIP against one mark and quantitative proteomic analysis was reported by Danny Reinberg's lab (103). Yet implementing it here would require taking into account all the active marks possibly modifying H3K27 (including propionylation, butyrylation, succinylation, 2-hydroxyisobutyrylation and beta-hydroxybutyrylation), testing the anti-H3K27ac antibody for its ability to equally purify symmetric and assymetric nucleosomes, and obtaining a stringently prepared sample of mononucleosomes, because di/multinucleosomes would significantly impact quantitative results. Besides, it would also be interesting to estimate the open chromatin status of regions bearing H3K27ac and H3K27cr either alone or together.

Our proteomic analyses indicated the preferential presence of H3K27cr, and to some extent of H3K27ac, in the variant H3.3. Comparison with ChIP-seq data acquired for H3.1/2 and H3.3 in RS cells (50) confirmed the preferential association of H3K27ac and H3K27cr with H3.3 observed by proteomics. The original publication describing Kcr (5) highlighted that increased crotonylation levels from SC to RS cells were associated with gene expression activation, and that a large fraction of these genes resided on the X chromosome; this group of genes thus overcoming the repressive chromatin environment following meiotic sex chromosome inactivation (MSCI) (5). The preferential association of Kcr with the transcriptional start sites of sex-linked genes that are activated in post-meiotic cells was also later confirmed by the Namekawa's group (45). The mark H3K27cr did not exhibit a particular enrichment on the X chromosome, and X-encoded genes expressed in post-meiotic cells (49) appeared to be mostly enriched in H3K27ac. Crotonylated lysine residues other than K27 must therefore be associated with post-meiotic expression of X-encoded genes (5,45,71). However, a large number of multi-copy genes located on the Y-chromosome exhibited a simultaneous increase of both H3K27cr and H3K27ac at their promoters, among which were Spindlin/spermiogenesis-specific proteins, suggesting that both crotonylation and acetylation on H3K27 could be involved in the regulation of these particular genes. Beyond sex chromosomes, we observed that H3K27cr and H3K27ac evolved from SC to RS in a synchronized manner at a subset of gene promoters, with an increase of both marks associated with highest expression of genes involved in spermatid differentiation, sperm motility and fertilization, while a decrease of both marks was associated with gene expression shutdown. R. Schneider and colleagues identified H3K14pr and H3K14bu as novel marks inducing active transcription and reported a synergistic effect of these modifications together with the well-known active mark H3K9ac (104). The work by Goudarzi *et al*. (73) also demonstrated that high levels of both acetylation and butyrylation on H4 K5 and K8 in spermatogenic cells are correlated with high levels of gene expression; their data also suggest that these two modifications could actually represent alternated modifications of the same residues, reflecting a characteristic feature of dynamic chromatin regions resulting from the combined action of HATs and HDACs. Our data suggest that H3K27ac and H3K27cr also have a synergistic effect on gene expression and that H3K27 could also be a residue involved in controlling the dynamics of specific chromatin regions. However, differences were highlighted between the two marks: in SC cells, H3K27ac appeared to be associated with more expressed genes than H3K27cr, while the opposite was observed in RS cells. These results suggested that both marks might be differently associated with transcriptional factors/regulators, which may depend on the expression level of those factors in different cellular contexts. Indeed some factors, such as SLY, are only expressed in RS cells.

In RS, we observed that the presence of both marks at promoters was associated with highest concentrations of transcription factors and regulators, namely SOX30, SLY, BORIS, CTCF and BRD4, with highest 5hmC marking over exons and highest gene expression levels. Yet when considering promoters exhibiting a detectable peak only for one mark on H3K27, genes induced from SC to RS showed a favored overlap of SOX30 with H3K27ac and a preferred correlation of SLY, BORIS, CTCF and BRD4 with H3K27cr. The fact that the three latter general transcription regulators are more associated with H3K27cr is

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

in line with lysine crotonylation being a more efficient mark of active transcription than lysine acetylation. Importantly, BRD4 is able to bind to H3K27ac (105) but not to butyrylated peptides (106), which renders it unlikely to bind H3K27cr. BRD4 was shown to be associated with histone H4 acetylated at its N-terminus (48). One hypothesis would be that H3K27cr-containing histone H3 preferentially coexists with acetylated histone H4, either in the same nucleosome or in adjacent ones. Despite its overlap with active epigenetic marks and enrichment at the promoters of highly expressed genes, SLY has been found to induce gene repression rather than gene activation, at least on sex chromosome genes and some autosomal genes (44). In line with these observations, we found here that SLY enrichment is reduced at the most upregulated genes which are H3K27cr-positive (Supplementary Figure S11B). Future research on SLY and on SLX, its X-linked homolog which appears to have an opposite role on gene expression, will be needed to better understand their mechanism and their interplay with histone crotonylation, in particular at H3K27. Finally, it would be interesting to extend the parallel drawn between high levels of H3K27ac and H3K27cr at promoters and of 5-hydroxymethylcytosine over exons of the downstream genes to the recently published data on DNA methylation and on the repressive mark H3K9me3 (107). This would provide an even more complete picture of the respective contributions on gene expression regulation of histone and DNA modifications.

In our proteomic analyses, H3K27cr was largely detected in association with H3K36me2, previously found to mark enhancers and super-enhancers (78). Active enhancers, which contribute to gene expression by interacting with their target promoters, are currently identified by the presence of H3K27ac. Even though H3K27ac is also known to accumulate at super-enhancers, additional characteristics are needed to locate and define these genomic elements with certainty. We observed that distal enhancers bearing simultaneously high levels of H3K27ac and H3K27cr colocalized most frequently with all the considered transcription factors and regulators: SOX30, SLY, BRD4, BRDT, BORIS and CTCF. In addition, a combined increase of H3K27cr and H3K27ac at both promoters and distal enhancers was associated with the expression of genes essential for the differentiation of round spermatids into final sperm, notably genes involved in sperm motility and fertilization. The determination of super-enhancers by considering H3K27ac or H3K27cr revealed a significant overlap of first-rank super-enhancers of both marks. In addition, both H3K27ac and H3K27cr define super-enhancers controlling the expression of RS-selective genes. Finally, of the 645 super-enhancers established in sperm (31), about 100 overlapped with doubly marked super-enhancers of RS cells, which indicates that at least some sperm super-enhancers are already present in post-meiotic cells.

Overall, our work suggests a prevalent role of H3K27cr in gene expression regulation both at promoters and at distal enhancers in the context of mouse spermatogenesis. It would be interesting to further dig into the regulation mechanisms controlled by H3K27cr in other biological contexts, such as in the brain. Indeed, in meiotic germ cells (SC), H3K27cr was observed to be enriched at the promoters of genes involved in postsynaptic specialization (Supplementary Figure S10B). In line with this clue, a recent publication reported the regulation of H3K27ac and H3K27cr levels by the long non-coding RNA NEAT1, and their impact on the expression of endocytosis-related genes in Alzheimer disease (108). Interestingly, in this context, the two marks appeared to have opposite effects, as H3K27cr correlated with transcriptional repression.

DATA AVAILABILITY

ChIP-seq data produced and used in this study were deposited on ArrayExpress under accession number E-MTAB-8328 and in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB34135. Proteomics data were deposited on the PRIDE archive under the accession number PXD015380 (Project DOI: 10.6019/PXD015380) (109).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

M.C. and D.P. wish to thank the staff of EDyP Service who contributed excellent technical support for proteomic analyses. We also want to thank staff of the animal house, cytometry (CYBIO) and histology (HistIM) core facilities from Cochin Institute (INSERM U1016, CNRS UMR8104, Université Paris Descartes). D.P. and M.C. greatly appreciated fruitful discussions with Saadi Khochbin and Sophie Rousseaux, and are grateful to them for their critical reading of this manuscript. We thank Nicolas Wiart for his support on the use of the high-performance computing clusters at CNRGH. We wish to thank Marie Arlotto for preparing yeast histones and yeast cells for metabolomic analyses, Sandrine Miesch-Fremy for preparing mice testis, Marie Courçon for technical advice in biochemistry, Yves Vandenbrouck for advice on the use of ProteoRE, and Naganand Rayapuram and Myriam Ferro for critical reading of this manuscript.

FUNDING

University Grenoble Alpes (UGA) by PhD funding (to M.C.); Agence Nationale de la Recherche [ANR-14-CE19-0014-01 to D.P.]; proteomics platform supported by ProFI [ANR-10-INBS-08]; GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS [ANR-17-EURE-0003]; INSERM and ANR [ANR-17-CE12-0004-01 to J.C.]; MetaboHUB infrastructure [ANR-11-INBS-0010]; CEA by a PhD fellowship (to S.E.K.); Fond d'Intervention of the University Grenoble Alpes (to J.G.); French National Research Agency [ANR-11-PDOC-0011 EpiGam to J.G., ANR-10-LABX-49-01 GRAL to J.G. and S.E.K.]; European Union FP7 Marie Curie Action 'Career Integration Grant' [304003 to J.G.]. Funding for open access charge: [ANR-10-INBS-08].

Conflict of interest statement. None declared.

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

REFERENCES

1. Phillips,D.M. (1963) The presence of acetyl groups of histones. *Biochem. J.*, **87**, 258–263.
2. Allfrey,V.G., Faulkner,R. and Mirsky,A.E. (1964) Acetylation and methylation of histones and their possible role in the regulation of rna synthesis. *Proc. Natl. Acad. Sci. U.S.A.*, **51**, 786–794.
3. Wu,S.-Y. and Chiang,C.-M. (2007) The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation. *J. Biol. Chem.*, **282**, 13141–13145.
4. Chen,Y., Sprung,R., Tang,Y., Ball,H., Sangras,B., Kim,S.C., Falck,J.R., Peng,J., Gu,W. and Zhao,Y. (2007) Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Mol. Cell Proteomics MCP*, **6**, 812–819.
5. Tan,M., Luo,H., Lee,S., Jin,F., Yang,J.S., Montellier,E., Buchou,T., Cheng,Z., Rousseaux,S., Rajagopal,N. *et al.* (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell*, **146**, 1016–1028.
6. Xie,Z., Dai,J., Dai,L., Tan,M., Cheng,Z., Wu,Y., Boeke,J.D. and Zhao,Y. (2012) Lysine succinylation and lysine malonylation in histones. *Mol. Cell Proteomics MCP*, **11**, 100–107.
7. Hirschey,M.D. and Zhao,Y. (2015) Metabolic regulation by lysine malonylation, succinylation, and glutarylation. *Mol. Cell Proteomics MCP*, **14**, 2308–2315.
8. Dai,L., Peng,C., Montellier,E., Lu,Z., Chen,Y., Ishii,H., Debernardi,A., Buchou,T., Rousseaux,S., Jin,F. *et al.* (2014) Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark. *Nat. Chem. Biol.*, **10**, 365–370.
9. Xie,Z., Zhang,D., Chung,D., Tang,Z., Huang,H., Dai,L., Qi,S., Li,J., Colak,G., Chen,Y. *et al.* (2016) Metabolic regulation of gene expression by histone lysine β-Hydroxybutyrylation. *Mol. Cell*, **62**, 194–206.
10. Huang,H., Zhang,D., Wang,Y., Perez-Neut,M., Han,Z., Zheng,Y.G., Hao,Q. and Zhao,Y. (2018) Lysine benzoylation is a histone mark regulated by SIRT2. *Nat. Commun.*, **9**, 3374.
11. Zhang,D., Tang,Z., Huang,H., Zhou,G., Cui,C., Weng,Y., Liu,W., Kim,S., Lee,S., Perez-Neut,M. *et al.* (2019) Metabolic regulation of gene expression by histone lactylation. *Nature*, **574**, 575–580.
12. Sabari,B.R., Zhang,D., Allis,C.D. and Zhao,Y. (2017) Metabolic regulation of gene expression through histone acylations. *Nat. Rev. Mol. Cell Biol.*, **18**, 90–101.
13. Sabari,B.R., Tang,Z., Huang,H., Yong-Gonzalez,V., Molina,H., Kong,H.E., Dai,L., Shimada,M., Cross,J.R., Zhao,Y. *et al.* (2018) Intracellular crotonyl-CoA stimulates transcription through p300-catalyzed histone crotonylation. *Mol. Cell*, **69**, 533.
14. Kaczmarska,Z., Ortega,E., Goudarzi,A., Huang,H., Kim,S., Márquez,J.A., Zhao,Y., Khochbin,S. and Panne,D. (2017) Structure of p300 in complex with acyl-CoA variants. *Nat. Chem. Biol.*, **13**, 21–29.
15. Huang,H., Tang,S., Ji,M., Tang,Z., Shimada,M., Liu,X., Qi,S., Locasale,J.W., Roeder,R.G., Zhao,Y. *et al.* (2018) p300-mediated lysine 2-hydroxyisobutyrylation regulates glycolysis. *Mol. Cell*, **70**, 663–678.
16. Liu,X., Wei,W., Liu,Y., Yang,X., Wu,J., Zhang,Y., Zhang,Q., Shi,T., Du,J.X., Zhao,Y. *et al.* (2017) MOF as an evolutionarily conserved histone crotonyltransferase and transcriptional activation by histone acetyltransferase-deficient and crotonyltransferase-competent CBP/p300. *Cell Discov.*, **3**, 17016.
17. Wang,Y., Guo,Y.R., Liu,K., Yin,Z., Liu,R., Xia,Y., Tan,L., Yang,P., Lee,J.-H., Li,X.-J. *et al.* (2017) KAT2A coupled with the α-KGDH complex acts as a histone H3 succinyltransferase. *Nature*, **552**, 273–277.
18. Zhao,S., Zhang,X. and Li,H. (2018) Beyond histone acetylation-writing and erasing histone acylations. *Curr. Opin. Struct. Biol.*, **53**, 169–177.
19. Zhang,X., Cao,R., Niu,J., Yang,S., Ma,H., Zhao,S. and Li,H. (2019) Molecular basis for hierarchical histone de-β-hydroxybutyrylation by SIRT3. *Cell Discov.*, **5**, 35.
20. Bao,X., Wang,Y., Li,X., Li,X.-M., Liu,Z., Yang,T., Wong,C.F., Zhang,J., Hao,Q. and Li,X.D. (2014) Identification of 'erasers' for lysine crotonylated histone marks using a chemical proteomics approach. *eLife*, **3**, e02999.
21. Wei,W., Liu,X., Chen,J., Gao,S., Lu,L., Zhang,H., Ding,G., Wang,Z., Chen,Z., Shi,T. *et al.* (2017) Class I histone deacetylases are major histone decrotonylases: evidence for critical and broad function of histone crotonylation in transcription. *Cell Res.*, **27**, 898–915.
22. Filippakopoulos,P., Picaud,S., Mangos,M., Keates,T., Lambert,J.-P., Barsyte-Lovejoy,D., Felletar,I., Volkmer,R., Müller,S., Pawson,T. *et al.* (2012) Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell*, **149**, 214–231.
23. Flynn,E.M., Huang,O.W., Poy,F., Oppikofer,M., Bellon,S.F., Tang,Y. and Cochran,A.G. (2015) A subset of human bromodomains recognizes butyryllysine and crotonyllysine histone peptide modifications. *Struct. Lond. Engl.*, **23**, 1801–1814.
24. Xiong,X., Panchenko,T., Yang,S., Zhao,S., Yan,P., Zhang,W., Xie,W., Li,Y., Zhao,Y., Allis,C.D. *et al.* (2016) Selective recognition of histone crotonylation by double PHD fingers of MOZ and DPF2. *Nat. Chem. Biol.*, **12**, 1111–1118.
25. Andrews,F.H., Shinsky,S.A., Shanle,E.K., Bridgers,J.B., Gest,A., Tsun,I.K., Krajewski,K., Shi,X., Strahl,B.D. and Kutateladze,T.G. (2016) The Taf14 YEATS domain is a reader of histone crotonylation. *Nat. Chem. Biol.*, **12**, 396–398.
26. Li,Y., Sabari,B.R., Panchenko,T., Wen,H., Zhao,D., Guan,H., Wan,L., Huang,H., Tang,Z., Zhao,Y. *et al.* (2016) Molecular coupling of histone crotonylation and active transcription by AF9 YEATS domain. *Mol. Cell*, **62**, 181–193.
27. Zhao,D., Guan,H., Zhao,S., Mi,W., Wen,H., Li,Y., Zhao,Y., Allis,C.D., Shi,X. and Li,H. (2016) YEATS2 is a selective histone crotonylation reader. *Cell Res.*, **26**, 629–632.
28. Wang,Y., Jin,J., Chung,M.W.H., Feng,L., Sun,H. and Hao,Q. (2018) Identification of the YEATS domain of GAS41 as a pH-dependent reader of histone succinylation. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, 2365–2370.
29. Zhai,G., Dong,H., Guo,Z., Feng,W., Jin,J., Zhang,T., Chen,C., Chen,P., Tian,S., Bai,X. *et al.* (2018) An efficient approach for selective enrichment of histone modification readers using self-assembled multivalent photoaffinity peptide probes. *Anal. Chem.*, **90**, 11385–11392.
30. Barral,S., Morozumi,Y., Tanaka,H., Montellier,E., Govin,J., de Dieuleveult,M., Charbonnier,G., Couté,Y., Puthier,D., Buchou,T. *et al.* (2017) Histone variant H2A.L.2 guides transition protein-dependent protamine assembly in male germ cells. *Mol. Cell*, **66**, 89–101.
31. Jung,Y.H., Sauria,M.E.G., Lyu,X., Cheema,M.S., Ausio,J., Taylor,J. and Corces,V.G. (2017) Chromatin states in mouse sperm correlate with embryonic and adult regulatory landscapes. *Cell Rep.*, **18**, 1366–1382.
32. Goudarzi,A., Shiota,H., Rousseaux,S. and Khochbin,S. (2014) Genome-scale acetylation-dependent histone eviction during spermatogenesis. *J. Mol. Biol.*, **426**, 3342–3349.
33. Shiota,H., Barral,S., Buchou,T., Tan,M., Couté,Y., Charbonnier,G., Reynoird,N., Boussouar,F., Gérard,M., Zhu,M. *et al.* (2018) Nut directs p300-dependent, genome-wide H4 hyperacetylation in male germ cells. *Cell Rep.*, **24**, 3477–3487.
34. Govin,J. and Berger,S.L. (2009) Genome reprogramming during sporulation. *Int. J. Dev. Biol.*, **53**, 425–432.
35. Govin,J., Dorsey,J., Gaucher,J., Rousseaux,S., Khochbin,S. and Berger,S.L. (2010) Systematic screen reveals new functional dynamics of histones H3 and H4 during gametogenesis. *Genes Dev.*, **24**, 1772–1786.
36. Cocquet,J., Ellis,P.J.I., Yamauchi,Y., Mahadevaiah,S.K., Affara,N.A., Ward,M.A. and Burgoyne,P.S. (2009) The multicopy gene Sly represses the sex chromosomes in the male mouse germline after meiosis. *PLoS Biol.*, **7**, e1000244.
37. Bastos,H., Lassalle,B., Chicheportiche,A., Riou,L., Testart,J., Allemand,I. and Fouchet,P. (2005) Flow cytometric characterization of viable meiotic and postmeiotic cells by Hoechst 33342 in mouse spermatogenesis. *Cytom. Part J. Int. Soc. Anal. Cytol.*, **65**, 40–49.
38. Comptour,A., Moretti,C., Serrentino,M.-E., Auer,J., Ialy-Radio,C., Ward,M.A., Touré,A., Vaiman,D. and Cocquet,J. (2014) SSTY proteins co-localize with the post-meiotic sex chromatin and interact with regulators of its expression. *FEBS J.*, **281**, 1571–1584.
39. Moretti,C., Serrentino,M.-E., Ialy-Radio,C., Delessard,M., Soboleva,T.A., Tores,F., Leduc,M., Nitschké,P., Drevet,J.R., Tremethick,D.J. *et al.* (2017) SLY regulates genes involved in chromatin remodeling and interacts with TBL1XR1 during sperm differentiation. *Cell Death Differ.*, **24**, 1029–1044.

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

40. Ahmed,E.A. and de Rooij,D.G. (2009) Staging of mouse seminiferous tubule cross-sections. *Methods Mol. Biol. Clifton, NJ*, **558**, 263–277.
41. El Kennani,S., Adrait,A., Shaytan,A.K., Khochbin,S., Bruley,C., Panchenko,A.R., Landsman,D., Pflieger,D. and Govin,J. (2017) MS_HistoneDB, a manually curated resource for proteomic analysis of human and mouse histones. *Epigenet. Chromatin*, **10**, 2.
42. Lin,S., Wein,S., Gonzales-Cope,M., Otte,G.L., Yuan,Z.-F., Afjehi-Sadat,L., Maile,T., Berger,S.L., Rush,J., Lill,J.R. *et al.* (2014) Stable-isotope-labeled histone peptide library for histone post-translational modification and variant quantification by mass spectrometry. *Mol. Cell. Proteomics MCP*, **13**, 2450–2466.
43. Yu,Y., Chen,J., Gao,Y., Gao,J., Liao,R., Wang,Y., Oyang,C., Li,E., Zeng,C., Zhou,S. *et al.* (2016) Quantitative profiling of combinational K27/K36 modifications on histone H3 variants in mouse organs. *J. Proteome Res.*, **15**, 1070–1079.
44. Hammoud,S.S., Low,D.H.P., Yi,C., Carrell,D.T., Guccione,E. and Cairns,B.R. (2014) Chromatin and transcription transitions of mammalian adult germline stem cells and spermatogenesis. *Cell Stem Cell*, **15**, 239–253.
45. Sin,H.-S., Kartashov,A.V., Hasegawa,K., Barski,A. and Namekawa,S.H. (2015) Poised chromatin and bivalent domains facilitate the mitosis-to-meiosis transition in the male germline. *BMC Biol.*, **13**, 53.
46. Adams,S.R., Maezawa,S., Alavattam,K.G., Abe,H., Sakashita,A., Shroder,M., Broering,T.J., Sroga Rios,J., Thomas,M.A., Lin,X. *et al.* (2018) RNF8 and SCML2 cooperate to regulate ubiquitination and H3K27 acetylation for escape gene activation on the sex chromosomes. *PLoS Genet.*, **14**, e1007233.
47. Pugacheva,E.M., Rivero-Hinojosa,S., Espinoza,C.A., Méndez-Catalá,C.F., Kang,S., Suzuki,T., Kosaka-Suzuki,N., Robinson,S., Nagarajan,V., Ye,Z. *et al.* (2015) Comparative analyses of CTCF and BORIS occupancies uncover two distinct classes of CTCF binding genomic regions. *Genome Biol.*, **16**, 161.
48. Bryant,J.M., Donahue,G., Wang,X., Meyer-Ficca,M., Luense,L.J., Weller,A.H., Bartolomei,M.S., Blobel,G.A., Meyer,R.G., Garcia,B.A. *et al.* (2015) Characterization of BRD4 during mammalian postmeiotic sperm development. *Mol. Cell Biol.*, **35**, 1433–1448.
49. Chen,Y., Zheng,Y., Gao,Y., Lin,Z., Yang,S., Wang,T., Wang,Q., Xie,N., Hua,R., Liu,M. *et al.* (2018) Single-cell RNA-seq uncovers dynamic processes and critical regulators in mouse spermatogenesis. *Cell Res.*, **28**, 879–896.
50. Erkek,S., Hisano,M., Liang,C.-Y., Gill,M., Murr,R., Dieker,J., Schübeler,D., van der Vlag,J., Stadler,M.B. and Peters,A.H.F.M. (2013) Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa. *Nat. Struct. Mol. Biol.*, **20**, 868–875.
51. Gaucher,J., Boussouar,F., Montellier,E., Curtet,S., Buchou,T., Bertrand,S., Hery,P., Jounier,S., Depaux,A., Vitte,A.-L. *et al.* (2012) Bromodomain-dependent stage-specific male genome programming by Brdt. *EMBO J.*, **31**, 3809–3820.
52. Gan,H., Wen,L., Liao,S., Lin,X., Ma,T., Liu,J., Song,C.-X., Wang,M., He,C., Han,C. *et al.* (2013) Dynamics of 5-hydroxymethylcytosine during mouse spermatogenesis. *Nat. Commun.*, **4**, 1995.
53. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinforma. Oxf. Engl.*, **26**, 841–842.
54. Bolger,A.M., Lohse,M. and Usadel,B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinforma. Oxf. Engl.*, **30**, 2114–2120.
55. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
56. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
57. Sadedin,S.P., Pope,B. and Oshlack,A. (2012) Bpipe: a tool for running and managing bioinformatics pipelines. *Bioinforma. Oxf. Engl.*, **28**, 1525–1526.
58. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinforma. Oxf. Engl.*, **29**, 15–21.
59. Robinson,M.D., McCarthy,D.J. and Smyth,G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinforma. Oxf. Engl.*, **26**, 139–140.
60. Yu,G., Wang,L.-G. and He,Q.-Y. (2015) ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinforma. Oxf. Engl.*, **31**, 2382–2383.
61. Lawrence,M., Huber,W., Pagès,H., Aboyoun,P., Carlson,M., Gentleman,R., Morgan,M.T. and Carey,V.J. (2013) Software for computing and annotating genomic ranges. *PLoS Comput. Biol.*, **9**, e1003118.
62. Gu,Z., Eils,R. and Schlesner,M. (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinforma. Oxf. Engl.*, **32**, 2847–2849.
63. Luense,L.J., Wang,X., Schon,S.B., Weller,A.H., Lin Shiao,E., Bryant,J.M., Bartolomei,M.S., Coutifaris,C., Garcia,B.A. and Berger,S.L. (2016) Comprehensive analysis of histone post-translational modifications in mouse and human male germ cells. *Epigenet. Chromatin*, **9**, 24.
64. Wang,H., Huang,Z.Q., Xia,L., Feng,Q., Erdjument-Bromage,H., Strahl,B.D., Briggs,S.D., Allis,C.D., Wong,J., Tempst,P. *et al.* (2001) Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. *Science*, **293**, 853–857.
65. Strahl,B.D., Briggs,S.D., Brame,C.J., Caldwell,J.A., Koh,S.S., Ma,H., Cook,R.G., Shabanowitz,J., Hunt,D.F., Stallcup,M.R. *et al.* (2001) Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. *Curr. Biol. CB*, **11**, 996–1000.
66. Taylor,G.C.A., Eskeland,R., Hekimoglu-Balkan,B., Pradeepa,M.M. and Bickmore,W.A. (2013) H4K16 acetylation marks active genes and enhancers of embryonic stem cells, but does not alter chromatin compaction. *Genome Res.*, **23**, 2053–2065.
67. McKittrick,E., Gafken,P.R., Ahmad,K. and Henikoff,S. (2004) Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 1525–1530.
68. Szenker,E., Ray-Gallet,D. and Almouzni,G. (2011) The double face of the histone variant H3.3. *Cell Res.*, **21**, 421–434.
69. Yamaguchi,K., Hada,M., Fukuda,Y., Inoue,E., Makino,Y., Katou,Y., Shirahige,K. and Okada,Y. (2018) Re-evaluating the localization of sperm-retained histones revealed the modification-dependent accumulation in specific genome regions. *Cell Rep.*, **23**, 3920–3932.
70. Simithy,J., Sidoli,S., Yuan,Z.-F., Coradin,M., Bhanu,N.V., Marchione,D.M., Klein,B.J., Bazilevsky,G.A., McCullough,C.E., Magin,R.S. *et al.* (2017) Characterization of histone acylations links chromatin modifications with metabolism. *Nat. Commun.*, **8**, 1141.
71. Liu,S., Yu,H., Liu,Y., Liu,X., Zhang,Y., Bu,C., Yuan,S., Chen,Z., Xie,G., Li,W. *et al.* (2017) Chromodomain protein CDYL acts as a crotonyl-CoA hydratase to regulate histone crotonylation and spermatogenesis. *Mol. Cell*, **67**, 853–866.
72. Sivanand,S., Viney,I. and Wellen,K.E. (2018) Spatiotemporal control of acetyl-CoA metabolism in chromatin regulation. *Trends Biochem. Sci.*, **43**, 61–74.
73. Goudarzi,A., Zhang,D., Huang,H., Barral,S., Kwon,O.K., Qi,S., Tang,Z., Buchou,T., Vitte,A.-L., He,T. *et al.* (2016) Dynamic competing histone H4 K5K8 acetylation and butyrylation are hallmarks of highly active gene promoters. *Mol. Cell*, **62**, 169–180.
74. Boussouar,F., Goudarzi,A., Buchou,T., Shiota,H., Barral,S., Debernardi,A., Guardiola,P., Brindle,P., Martinez,G., Arnoult,C. *et al.* (2014) A specific CBP/p300-dependent gene expression programme drives the metabolic remodelling in late stages of spermatogenesis. *Andrology*, **2**, 351–359.
75. Jiang,G., Nguyen,D., Archin,N.M., Yukl,S.A., Méndez-Lagares,G., Tang,Y., Elsheikh,M.M., Thompson,G.R., Hartigan-O'Connor,D.J., Margolis,D.M. *et al.* (2018) HIV latency is reversed by ACSS2-driven histone crotonylation. *J. Clin. Invest.*, **128**, 1190–1198.
76. Mi,W., Guan,H., Lyu,J., Zhao,D., Xi,Y., Jiang,S., Andrews,F.H., Wang,X., Gagea,M., Wen,H. *et al.* (2017) YEATS2 links histone acetylation to tumorigenesis of non-small cell lung cancer. *Nat. Commun.*, **8**, 1088.

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020

77. Yuen,B.T.K., Bush,K.M., Barrilleaux,B.L., Cotterman,R. and Knoepfler,P.S. (2014) Histone H3.3 regulates dynamic chromatin states during spermatogenesis. *Dev. Camb. Engl.*, **141**, 3483–3494.
78. Soldi,M., Mari,T., Nicosia,L., Musiani,D., Sigismondo,G., Cuomo,A., Pavesi,G. and Bonaldi,T. (2017) Chromatin proteomics reveals novel combinatorial histone modification signatures that mark distinct subpopulations of macrophage enhancers. *Nucleic Acids Res.*, **45**, 12195–12213.
79. Mueller,J.L., Mahadevaiah,S.K., Park,P.J., Warburton,P.E., Page,D.C. and Turner,J.M.A. (2008) The mouse X chromosome is enriched for multicopy testis genes showing postmeiotic expression. *Nat. Genet.*, **40**, 794–799.
80. Turner,J.M.A. (2007) Meiotic sex chromosome inactivation. *Dev. Camb. Engl.*, **134**, 1823–1831.
81. Moretti,C., Vaiman,D., Tores,F. and Cocquet,J. (2016) Expression and epigenomic landscape of the sex chromosomes in mouse post-meiotic male germ cells. *Epigenet. Chromatin*, **9**, 47.
82. Ernst,C., Eling,N., Martinez-Jimenez,C.P., Marioni,J.C. and Odom,D.T. (2019) Staged developmental mapping and X chromosome transcriptional dynamics during mouse spermatogenesis. *Nat. Commun.*, **10**, 1251.
83. Wegner,S.H., Yu,X., Pacheco Shubin,S., Griffith,W.C. and Faustman,E.M. (2015) Stage-specific signaling pathways during murine testis development and spermatogenesis: A pathway-based analysis to quantify developmental dynamics. *Reprod. Toxicol.*, **51**, 31–39.
84. Szklarczyk,D., Gable,A.L., Lyon,D., Junge,A., Wyder,S., Huerta-Cepas,J., Simonovic,M., Doncheva,N.T., Morris,J.H., Bork,P. *et al.* (2019) STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.*, **47**, D607–D613.
85. Shang,E., Nickerson,H.D., Wen,D., Wang,X. and Wolgemuth,D.J. (2007) The first bromodomain of Brdt, a testis-specific member of the BET sub-family of double-bromodomain-containing proteins, is essential for male germ cell differentiation. *Dev. Camb. Engl.*, **134**, 3507–3515.
86. Manterola,M., Brown,T.M., Oh,M.Y., Garyn,C., Gonzalez,B.J. and Wolgemuth,D.J. (2018) BRDT is an essential epigenetic regulator for proper chromatin organization, silencing of sex chromosomes and crossover formation in male meiosis. *PLoS Genet.*, **14**, e1007209.
87. Rivero-Hinojosa,S., Kang,S., Lobanenkov,V.V. and Zentner,G.E. (2017) Testis-specific transcriptional regulators selectively occupy BORIS-bound CTCF target regions in mouse male germ cells. *Sci. Rep.*, **7**, 41279.
88. Hnisz,D., Abraham,B.J., Lee,T.I., Lau,A., Saint-André,V., Sigova,A.A., Hoke,H.A. and Young,R.A. (2013) Super-enhancers in the control of cell identity and disease. *Cell*, **155**, 934–947.
89. Zhang,T., Cooper,S. and Brockdorff,N. (2015) The interplay of histone modifications - writers that read. *EMBO Rep.*, **16**, 1467–1481.
90. Schoenfelder,S. and Fraser,P. (2019) Long-range enhancer-promoter contacts in gene expression control. *Nat. Rev. Genet.*, **20**, 437–455.
91. Whyte,W.A., Orlando,D.A., Hnisz,D., Abraham,B.J., Lin,C.Y., Kagey,M.H., Rahl,P.B., Lee,T.I. and Young,R.A. (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*, **153**, 307–319.
92. Creyghton,M.P., Cheng,A.W., Welstead,G.G., Kooistra,T., Carey,B.W., Steine,E.J., Hanna,J., Lodato,M.A., Frampton,G.M., Sharp,P.A. *et al.* (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 21931–21936.
93. Raisner,R., Kharbanda,S., Jin,L., Jeng,E., Chan,E., Merchant,M., Haverty,P.M., Bainer,R., Cheung,T., Arnott,D. *et al.* (2018) Enhancer activity requires CBP/P300 bromodomain-dependent histone H3K27 acetylation. *Cell Rep.*, **24**, 1722–1729.
94. Ing-Simmons,E., Seitan,V.C., Faure,A.J., Flicek,P., Carroll,T., Dekker,J., Fisher,A.G., Lenhard,B. and Merkenschlager,M. (2015) Spatial enhancer clustering and regulation of enhancer-proximal genes by cohesin. *Genome Res.*, **25**, 504–513.
95. Dowen,J.M., Fan,Z.P., Hnisz,D., Ren,G., Abraham,B.J., Zhang,L.N., Weintraub,A.S., Schujiers,J., Lee,T.I., Zhao,K. *et al.* (2014) Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell*, **159**, 374–387.
96. McLean,C.Y., Bristor,D., Hiller,M., Clarke,S.L., Schaar,B.T., Lowe,C.B., Wenger,A.M. and Bejerano,G. (2010) GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.*, **28**, 495–501.
97. Li,M.W.M., Mruk,D.D. and Cheng,C.Y. (2009) Mitogen-activated protein kinases in male reproductive function. *Trends Mol. Med.*, **15**, 159–168.
98. Cotton,L., Gibbs,G.M., Sanchez-Partida,L.G., Morrison,J.R., de Kretser,D.M. and O'Bryan,M.K. (2006) FGFR-1 signaling is involved in spermiogenesis and sperm capacitation. *J. Cell Sci.*, **119**, 75–84.
99. Feng,C.-W.A., Spiller,C., Merriner,D.J., O'Bryan,M.K., Bowles,J. and Koopman,P. (2017) SOX30 is required for male fertility in mice. *Sci. Rep.*, **7**, 17619.
100. Zhang,D., Xie,D., Lin,X., Ma,L., Chen,J., Zhang,D., Wang,Y., Duo,S., Feng,Y., Zheng,C. *et al.* (2018) The transcription factor SOX30 is a key regulator of mouse spermiogenesis. *Development*, **145**, dev164723.
101. Bai,S., Fu,K., Yin,H., Cui,Y., Yue,Q., Li,W., Cheng,L., Tan,H., Liu,X., Guo,Y. *et al.* (2018) Sox30 initiates transcription of haploid genes during late meiosis and spermiogenesis in mouse testes. *Development*, **145**, dev164855.
102. Lovén,J., Hoke,H.A., Lin,C.Y., Lau,A., Orlando,D.A., Vakoc,C.R., Bradner,J.E., Lee,T.I. and Young,R.A. (2013) Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell*, **153**, 320–334.
103. Voigt,P., LeRoy,G., Drury,W.J., Zee,B.M., Son,J., Beck,D.B., Young,N.L., Garcia,B.A. and Reinberg,D. (2012) Asymmetrically modified nucleosomes. *Cell*, **151**, 181–193.
104. Kebede,A.F., Nieborak,A., Shahidian,L.Z., Le Gras,S., Richter,F., Gómez,D.A., Baltissen,M.P., Meszaros,G., Magliarelli,H.de F., Taudt,A. *et al.* (2017) Histone propionylation is a mark of active chromatin. *Nat. Struct. Mol. Biol.*, **24**, 1048–1056.
105. Ji,X., Dadon,D.B., Abraham,B.J., Lee,T.I., Jaenisch,R., Bradner,J.E. and Young,R.A. (2015) Chromatin proteomic profiling reveals novel proteins associated with histone-marked genomic regions. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 3841–3846.
106. Vollmuth,F. and Geyer,M. (2010) Interaction of propionylated and butyrylated histone H3 lysine marks with Brd4 bromodomains. *Angew. Chem. Int. Ed Engl.*, **49**, 6768–6772.
107. Liu,Y., Zhang,Y., Yin,J., Gao,Y., Li,Y., Bai,D., He,W., Li,X., Zhang,P., Li,R. *et al.* (2019) Distinct H3K9me3 and DNA methylation modifications during mouse spermatogenesis. *J. Biol. Chem.*, **294**, 18714–18725.
108. Wang,Z., Zhao,Y., Xu,N., Zhang,S., Wang,S., Mao,Y., Zhu,Y., Li,B., Jiang,Y., Tan,Y. *et al.* (2019) NEAT1 regulates neuroglial cell mediating Aβ clearance via the epigenetic regulation of endocytosis-related genes expression. *Cell. Mol. Life Sci. CMLS*, **76**, 3005–3018.
109. Perez-Riverol,Y., Csordas,A., Bai,J., Bernal-Llinares,M., Hewapathirana,S., Kundu,D.J., Inuganti,A., Griss,J., Mayer,G., Eisenacher,M. *et al.* (2019) The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.*, **47**, D442–D450.
110. Yu,G., Wang,L.-G., Han,Y. and He,Q.-Y. (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics J. Integr. Biol.*, **16**, 284–287.

Downloaded from https://academic.oup.com/nar/article/48/8/4115/5809165 by CEA user on 22 September 2020