Conservation of Aging and Cancer Epigenetic Signatures across Human and Mouse

Raúl F. Pérez,[†,1,2,3,4,5] Juan Ramón Tejedor,[†,1,2,3,4,5] Pablo Santamarina-Ojeda,[1,2,3,4,5] Virginia López Martínez,[1,2,3,4,5] Rocío G. Urdinguio,[1,2,3,4,5] Lucía Villamañán,[6] Ana Paula Candiota,[6,7] Noemí Vidal Sarró,[8] Marta Barradas,[9] Pablo Jose Fernandez-Marcos,[9] Manuel Serrano,[10,11,12] Agustín F. Fernández,[*,1,2,3,4,5] and Mario F. Fraga[*,1,2,3,4,5]

[1]Cancer Epigenetics and Nanomedicine Laboratory, Nanomaterials and Nanotechnology Research Center (CINN-CSIC), University of Oviedo, Oviedo, Spain

[2]Institute of Oncology of Asturias (IUOPA), University of Oviedo, Oviedo, Spain

[3]Health Research Institute of Asturias (ISPA), University of Oviedo, Oviedo, Spain

[4]Department of Organisms and Systems Biology (B.O.S.), University of Oviedo, Oviedo, Spain

[5]Rare Diseases CIBER (CIBERER) of the Carlos III Health Institute (ISCIII), Madrid, Spain

[6]Unitat de Bioquímica de Biociències, Departament de Bioquímica i Biologia Molecular, Edifici Cs, Universitat Autònoma de Barcelona, Cerdanyola del Vallès, Spain

[7]Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Madrid, Spain

[8]Servicio Anatomía Patológica, Hospital Universitari de Bellvitge—IDIBELL, Hospitalet de Llobregat, Spain

[9]Metabolic Syndrome Group—BIOPROMET, Madrid Institute for Advanced Studies—IMDEA Food, CEI UAM+CSIC, Madrid, Spain

[10]Tumour Suppression Group, Spanish National Cancer Research Centre (CNIO), Madrid, Spain

[11]Cellular Plasticity and Disease Group, Institute for Research in Biomedicine (IRB Barcelona), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain

[12]Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain

[†]These authors contributed equally to this work.

***Corresponding authors:** E-mails: agustin.fernandez@cinn.es; mffraga@cinn.es.

Associate editor: Li Liu

Abstract

Aging and cancer are two interrelated processes, with aging being a major risk factor for the development of cancer. Parallel epigenetic alterations have been described for both, although differences, especially within the DNA hypomethylation scenario, have also been recently reported. Although many of these observations arise from the use of mouse models, there is a lack of systematic comparisons of human and mouse epigenetic patterns in the context of disease. However, such comparisons are significant as they allow to establish the extent to which some of the observed similarities or differences arise from pre-existing species-specific epigenetic traits. Here, we have used reduced representation bisulfite sequencing to profile the brain methylomes of young and old, tumoral and nontumoral brain samples from human and mouse. We first characterized the baseline epigenomic patterns of the species and subsequently focused on the DNA methylation alterations associated with cancer and aging. Next, we described the functional genomic and epigenomic context associated with the alterations, and finally, we integrated our data to study interspecies DNA methylation levels at orthologous CpG sites. Globally, we found considerable differences between the characteristics of DNA methylation alterations in cancer and aging in both species. Moreover, we describe robust evidence for the conservation of the specific cancer and aging epigenomic signatures in human and mouse. Our observations point toward the preservation of the functional consequences of these alterations at multiple levels of genomic regulation. Finally, our analyses reveal a role for the genomic context in explaining disease- and species-specific epigenetic traits.

Key words: **epigenetics, DNA methylation, histone modification, mouse, human, conservation.**

Article

Introduction

Aging is one of the main risk factors associated with the development of cancer. These two processes have been extensively investigated from the epigenetic perspective and consequently analogous epigenetic alterations have been identified for both, providing explanations for the possible molecular links between them (Aunan et al. 2017; Yu et al. 2020). These alterations mainly relate to local gains in DNA

methylation at CpG-dense regions and global losses at the genomic scale involving repetitive DNA (Baylin and Jones 2016; Sen et al. 2016). Nonetheless, the extent of these parallelisms has not as yet been clearly outlined, particularly within the hypomethylation scenario (Dmitrijeva et al. 2018; Pérez et al. 2018). Moreover, recent whole-genome studies in mouse have also failed to confirm hallmarks such as global hypomethylation with aging in tissues including liver, hippocampus, and stem cells (Sun et al. 2014; Cole et al. 2017; Hahn et al. 2017; Masser et al. 2017; Hadad et al. 2019; Hernando-Herraez et al. 2019). These observations underscore two different, but equally important, issues: first, there is a need to study aging- and cancer-associated epigenetic patterns in a systematic manner to facilitate an integrated comparison between the two processes; second, it is essential to also confirm whether aging- and cancer-associated epigenetic alterations are equivalent in human and mouse models, as the latter are very important experimental proxies of human biology.

Mouse models have been widely used in the characterization of landmark epigenetic mechanisms such as transgenerational inheritance and genomic imprinting (Blewitt and Whitelaw 2013) as well as developmental epigenomic reprogramming (Hanna et al. 2018), and they are often used to characterize alterations in specific disease-associated pathways which sometimes have a translational correspondence in human (Espada and Esteller 2013). To date, reports suggest that DNA methylation patterns preserve global features such as the presence of nonmethylated islands and gene body methylation across mammalian species in general (Long et al. 2013; Schroeder et al. 2015), with other parallelisms extending across more distant vertebrate taxa (Elango and Yi 2008). More directed comparisons involving mouse and human have described a general conservation of their genome-wide epigenetic patterns (Edwards et al. 2010) and have particularly focused on the interspecies conservation of tissue-specific methylation patterns (Kessler et al. 2016; Zhou et al. 2017; Chen et al. 2018) and developmental processes (Lister et al. 2013). Nonetheless, considerable epigenetic differences in proximal species such as primates have been reported (Hernando-Herraez, Heyn, et al. 2015; Mendizabal et al. 2016). So far there have been only limited high-resolution studies that focus on directly assessing specific epigenetic similarities between human and mouse in either aging (Maegawa et al. 2017; Wang et al. 2017) or cancer (Maegawa et al. 2014), and those that exist are often limited by their compilation of data from different studies or the use of different profiling technologies for each species. Some of these issues have arisen from the lack of a large-scale DNA methylation array technology for mouse. Overall, then, there is a lack of back-to-back human–mouse comparisons of DNA methylation at single-base resolution in the context of disease. Even so, this issue is particularly relevant in the study of DNA methylation alterations found independently in both species in order to reliably ascertain the extent to which some of the observed similarities or differences arise from pre-existing species-specific epigenetic traits.

In this study, we sought to perform a systematic, integrated, and intercomparable analysis of the DNA methylation dynamics associated with cancer and aging in both human and mouse. To this end, we profiled the brain methylomes of young and old, tumoral and nontumoral brain samples from human and mouse by reduced representation bisulfite sequencing (RRBS), which provided a robust and interspecies-comparable platform for the analysis at hand (Bock et al. 2010). We first characterized the baseline epigenomic patterns of both species and subsequently focused on the DNA methylation alterations associated with cancer and aging. Next, we described the functional genomic and epigenomic context associated with the alterations, and finally, we integrated our data to study interspecies DNA methylation levels at orthologous CpG sites. Globally, we found considerable differences between the genomic and epigenomic characteristics of DNA methylation alterations in cancer and aging in human and mouse. Furthermore, we describe robust evidence for the conservation of the specific cancer- and aging-associated epigenomic patterns in both species. Our observations point toward the preservation of the functional consequences of these alterations at multiple levels of genomic regulation, including the locations and chromatin context of the alterations, the genetic pathways, and transcription factors (TFs) involved. Additionally, the analysis of orthologous CpG loci allowed us to describe species-common locations of DNA methylation variability and species-common cancer- and aging DNA methylation alterations. Finally, our analyses suggest that the genomic context associated with many CpG sites may help explain the observation of species-discordant DNA methylation as well as cancer and aging-specific DNA methylation alterations.

Results and Discussion

Baseline Epigenomic Patterns Are Conserved between Human and Mouse

To characterize the intricate, conserved epigenetic signatures that may exist in aging and cancer, we studied the genome-wide DNA methylation patterns of young, old, tumoral, and nontumoral brain samples from human and mouse ($n = 24$, 3 per group, fig. 1A). We produced methylation values for an average of 3.94 million and 1.62 million unique CpG sites per sample for human and mouse, respectively, and retained a final total of 1,171,918 and 585,234 CpG sites common to all samples after filtering for coverage (see Materials and Methods and supplementary table 1, Supplementary Material online, for sample and alignment information). We first sought to characterize the global aspects of DNA methylation and its relationship with genomic function in the two species in order to determine the baseline similarities and differences in DNA methylation patterns in our study system. For this initial analysis, we focused on the nontumoral samples because the tumors showed extensive deviation from the rest of the samples using pairwise Pearson correlations (fig. 1B), and our goal was to profile the general species-wise features of DNA methylation in our samples.

Though significant, we did not observe noticeable differences between the species in the distribution of the overall methylation values, with an average methylation of 0.266 for

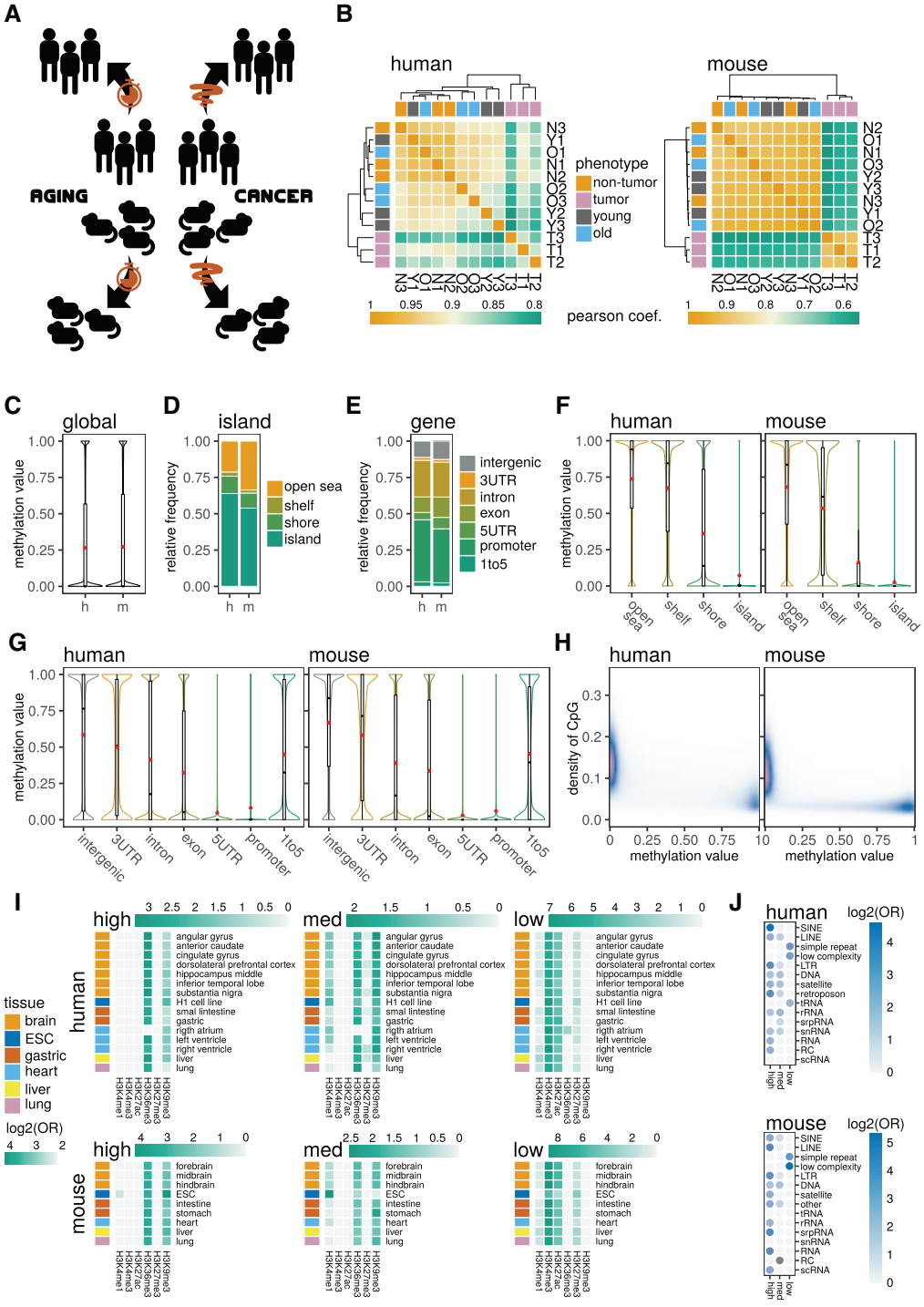

Fig. 1. Baseline epigenomic patterns in human and mouse. (*A*) Schematic of the study design. (*B*) Heatmaps showing the pairwise Pearson correlations between DNA methylation levels of the profiled CpG sites across samples. (*C*) Violin plots of global levels of DNA methylation for the pooled samples (red dots indicate average values; h, human; m, mouse; tumor samples are excluded). (*D* and *E*) Barplots indicating the distribution of the profiled CpG sites in human and mouse across CpG island and gene-region locations. (*F* and *G*) Violin plots showing the DNA methylation levels for the pooled samples with the CpG sites grouped by CpG island locations and gene-region locations (red dots indicate average values). (*H*) Density plots indicating the relationship between DNA methylation values and surrounding density (2-kb bins) in CpG sites for the profiled CpGs. (*I* and *J*) Heatmaps and bubble plots showing the significant overenrichments in log2(odds ratio) of CpG sites classified as high-, medium-, and low-methylation in ChIP-seq peak locations of different histone marks across a panel of tissues or in the locations of repetitive DNA classes across the human and mouse genomes.

human and 0.272 for mouse (fig. 1C, Wilcoxon rank-sum $P = 0.02$). The low levels of methylation observed are to be expected since the RRBS technology targets CpG-dense regions, such as CpG islands, which tend to be unmethylated. On a genome-wide scale, however, most CpG sites are methylated in mammals (Greenberg and Bourc'his 2019), and technologies such as whole-genome bisulfite sequencing (WGBS), which also targets intergenic regions, would detect higher levels of DNA methylation.

In order to ascertain that the RRBS was evaluating a comparable population of CpGs for each species, we annotated the sites to functional genomic locations in terms of their CpG island membership status and genetic elements. For both species, the technology interrogated a majority of CpG sites located at CpG islands (fig. 1D), with a slight increase in open sea CpGs for mouse (Fisher's $P < 0.001$, OR = 1.89). Regarding genetic elements, the distribution of the quantified CpGs was also very similar for both species (fig. 1E), with the largest proportion of sites being located at promoters and gene bodies. These observations suggest that, for human and mouse, the RRBS technology targets populations of CpG sites which have analogous genomic contexts.

Next, we proceeded to characterize the relationship between DNA methylation and CpG context in both human and mouse. With respect to CpG islands, DNA methylation was highest at open sea locations and lowest at islands for both species, with a transition in methylation across shelves and shores (fig. 1F). For gene elements, the relationship with DNA methylation was also parallel between human and mouse, with promoters and 5′-UTR regions showing the lowest values, followed by gene body exons and introns, whereas intergenic regions were those with the highest DNA methylation (fig. 1G). The presence of DNA methylation is inversely correlated to the density in surrounding CpG sites (Chen et al. 2018), and this may help explain some of the patterns found at genomic locations, especially concerning CpG islands and open sea. We therefore characterized the density in the CpG sites associated with each locus by counting the number of neighboring CpG sites in surrounding 2-kb windows. As expected, in general, low methylation values were found associated with locations with higher density in CpG sites, whereas higher methylation values were found at low-density locations, for both species (fig. 1H). Indeed, the different genomic locations (islands and gene regions) displayed different distributions of density in CpG sites (supplementary fig. S1A, Supplementary Material online) and, moreover, the mean methylation value of each element correlated almost perfectly with its average density in CpG sites (supplementary fig. S1B, Supplementary Material online). Collectively, these results confirm that the relationship between DNA methylation and functional genomic elements such as CpG islands and gene regions is conserved between human and mouse and is mostly explained by the density of CpG sites associated with each element (Edwards et al. 2010).

In the regulation of gene expression, epigenetic marks such as DNA methylation are not independent, rather, they act in combination with interrelated marks such as histone modifications to define functional states (Du et al. 2015). Thus, to better study the relationship of DNA methylation concerning its genomic function in mouse and human, we integrated our data with ENCODE and NIH Roadmap Epigenomics ChIP-seq data profiling the histone modifications H3K4me1, H3K4me3, H3K27ac, H3K36me3, H3K27me3, and H3K9me3. We focused on a panel of tissues which was available for the two species, and contained several brain tissue tracks, analyzing a combined total of 144 data sets. We classified CpG sites as high-, medium-, or low-methylation loci (>0.8, [0.2,0.8], and <0.2 average methylation values, respectively) and investigated the colocalization of these sites with the histone marks by overenrichment analyses at FDR <0.05 (see Materials and Methods). With this procedure, we observed practically identical chromatin signatures of methylation status in human and mouse (fig. 1I and supplementary table 2, Supplementary Material online): high methylation sites were associated with active gene-body H3K36me3 and repressive H3K9me3 marks, intermediate methylation also appeared to be related to enhancer/promoter-associated H3Kme1 and low methylation sites were strongly associated with the active promoter and enhancer H3K4me3 and H3K27ac marks. These observations confirm that the relationship between DNA methylation and genomic function is conserved in both species (Roadmap Epigenomics Consortium et al. 2015; Gorkin et al. 2020).

One of the genomic elements to which DNA methylation has been most functionally linked is repetitive DNA (Beisel and Paro 2011). Hence, we mapped the classified CpG sites to RepeatMasker repetitive elements and looked for overenrichment in specific repeat classes. Looking at the most common repeats, we observed, for both human and mouse, that high-methylation loci were associated with SINE, LINE, LTR, and DNA repeats, whereas low-methylation was linked to simple and low-complexity repeats (fig. 1J and supplementary table 3, Supplementary Material online). Expanding the analyses to repetitive element families also revealed parallel patterns in the two species (supplementary fig. S2, Supplementary Material online). These results are also in consonance with existing literature both for mouse and human (Edwards et al. 2010).

Dynamics of DNA Methylation Alterations in Human Aging and Cancer

Once we had confirmed the analogous human and mouse baseline epigenomic patterns in our study system, we proceeded to determine the DNA methylation alterations associated with aging and with cancer for each species independently. Starting with human, we found no general DNA methylation differences associated with cancer or aging by looking at all the sites analyzed by the RRBS, although tumor samples tended to have higher methylation values (fig. 2A). Moreover, PCA clustering suggested differences between the experimental groups as they segregated across the most meaningful dimensions (fig. 2B). We performed a differential methylation analysis with a stringent threshold (FDR <0.05, average change in methylation >0.2) to define differentially methylated CpGs (dmCpGs) in aging and cancer. We identified widespread alterations in both processes, counting

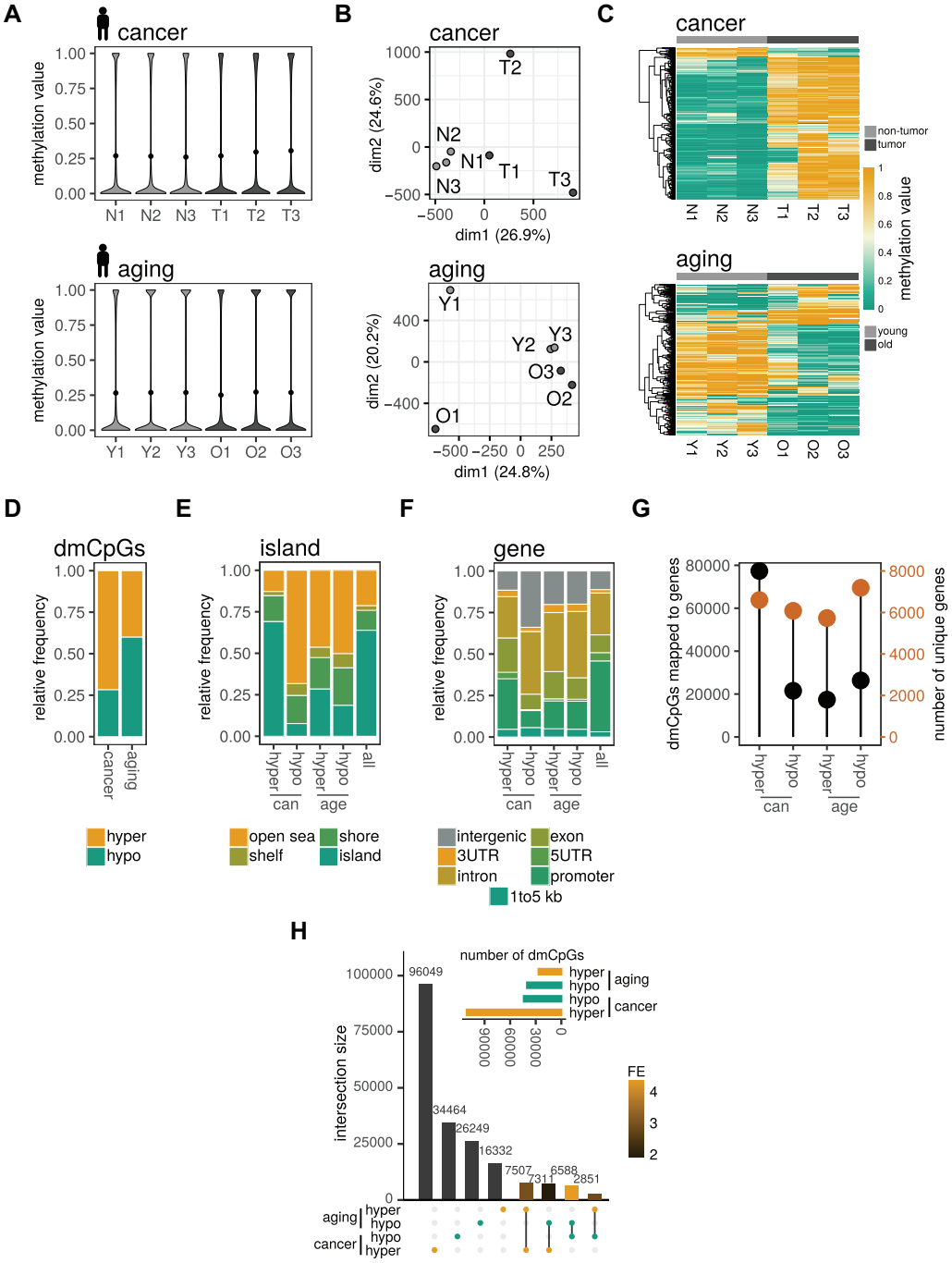

Fig. 2. Cancer and aging DNA methylation alterations in human. (*A*) Violin plots of global levels of DNA methylation of all CpGs across the profiled samples (N, nontumor; T, tumor; Y, young; O, old). (*B*) PCA plots showing the distribution of the samples across the first two principal components, with the percentage of variance each explains. (*C*) Heatmaps showing the methylation values of the top 1,000 dmCpGs found in the aging and cancer comparisons. (*D*) Barplots indicating the relative numbers of hyper- and hypo-dmCpGs (differentially methylated CpGs) in aging and cancer. (*E* and *F*) Barplots indicating the distribution of the dmCpGs across CpG island and gene-region locations, including the distribution of the universe of CpGs profiled by the RRBS ("all"). (*G*) Lollipop plots comparing the number of dmCpGs mapped to genes (left vertical axis, black dots) with the number of different genes that the dmCpGs mapped to (right vertical axis, orange dots). (*H*) UpSet plot describing the dmCpG sets and their intersections, indicating the size of the sets and the fold-enrichment (FE) of the intersections based on their expected overlaps.

154,770 dmCpGs associated with cancer and 66,838 with aging (lists of dmCpGs are available as extended data sets), with hypermethylation dominating the cancer scenario particularly (accounting for 72% of changes, fig. 2C and *D*). Furthermore, performing the cancer differential methylation comparison using all nine nontumoral samples (nontumor, young, and old groups) as the control group yielded comparable results and numbers of dmCpGs (supplementary fig. S3A, Supplementary Material online), indicating that the cancer-associated alterations observed are robust and are

independent of the samples used. When mapping the dmCpGs to CpG island locations (fig. 2E), we observed that hypermethylation occurred for cancer preferentially at CpG islands, with a distribution akin to that of the RRBS background of analyzed sites, whereas hypomethylation and aging-associated changes were, in general, enriched at open sea locations (Fisher's $P < 0.001$, ORs = 8.9, 3.3, and 4.0, respectively). Regarding gene locations, cancer hypomethylation and aging hyper- and hypomethylation were particularly increased at intergenic- and intronic sequences (Fisher's $P < 0.001$, intergenic ORs = 4.5, 2.0, and 2.0, intronic ORs = 1.8, 1.7, and 2.0). In line with this observation, cancer hypermethylation occurred more frequently at CpG-denser regions than the rest of the changes (supplementary fig. S3B, Supplementary Material online) and, additionally, it consisted of the strongest changes (supplementary fig. S3C, Supplementary Material online). These results suggest that there are important differences in the breadth and distribution of hyper- and hypomethylation alterations in cancer, whereas aging-associated alterations appear to be located at similar genomic locations regardless of the direction of change, these locations being comparable to those targeted by cancer-associated hypomethylation.

To expand on the possible functional consequences of the DNA methylation alterations observed, we mapped the dmCpGs to genes and performed pathway enrichment analyses. First, as shown in figure 2F, the largest proportion of dmCpGs mapping to genes was found for cancer hypermethylation. However, of the CpGs that mapped to genes, cancer hypo-dmCpGs and aging dmCpGs, although less numerous, were distributed across similar numbers of genes as cancer hyper-dmCpGs (fig. 2G, χ^2 $P < 0.001$), indicating that the former were more spread-out across different genes. We performed pathway enrichment analyses against several databases which included Gene Ontology, KEGG and Reactome pathways, and the CGP database from MSigDB, which describes empirical pathways mined from biomedical literature. Only cancer hyper-dmCpGs revealed significant enrichment in biological functions, and these were related to Polycomb-target genes, G-protein coupled receptor (GPCR) pathways—with neuropeptide receptors in particular—and cellular development pathways, especially neuronal (supplementary fig. S4, Supplementary Material online, full results are available as extended data sets). Thus, cancer-associated DNA methylation deregulation, particularly hypermethylation, appears to have a more coherent functional impact than aging-associated changes.

In order to investigate common axes of DNA methylation deregulation in both cancer and aging, we intersected the dmCpGs found, observing common CpG sets across all processes (fig. 2H and supplementary table 4, Supplementary Material online). These intersections were all significantly overenriched (SuperExactTest all $P < 0.001$), with common cancer- and aging hypo-dmCpGs displaying the highest fold enrichment, whereas common cancer hyper- and aging hypo-dmCpGs showed the lowest. These findings indicate that, regardless of their direction of change, DNA methylation alterations are readily found at disease-common loci. We performed pathway enrichments on these intersections and only observed enrichments in common CpGs between cancer hypermethylation and aging hyper- and hypomethylation (full results are available as extended data sets). These results suggest that the functional commonalities between DNA methylation deregulation in human cancer and aging mainly pertain to Polycomb-target gene hypermethylation in cancer, whereas these pathways manifest both hyper- and hypomethylation in aging.

Dynamics of DNA Methylation Alterations in Mouse Aging and Cancer

Next, we profiled the DNA methylation alterations associated with aging and cancer in mouse. First, regarding the general levels of DNA methylation, we found, in line with the previous observations in human, an even clearer trend of increased DNA methylation in cancer (fig. 3A). The PCA hinted at some differences between the phenotypes in that a slight separation was observed between the groups when looking at the top two dimensions (fig. 3B). The differential methylation analysis (FDR < 0.05, average change in methylation >0.2) uncovered extensive alterations, with 181,027 dmCpGs detected in cancer and 22,564 in aging (lists of dmCpGs are available as extended data sets). When all of the nine nontumoral samples (nontumor, young, and old groups) were used as the control group in the cancer differential methylation comparison, results and numbers of dmCpGs were similar (supplementary fig. S5A, Supplementary Material online), suggesting that the cancer-associated alterations observed are readily detected irrespective of the samples used. Regarding the direction of the alterations, hypermethylation was the predominant phenomenon in both processes (accounting for 68% and 62% of changes, respectively, fig. 3C and D). The genomic-location distributions of the dmCpGs were analogous to those observed in the human system (fig. 3E): cancer hypermethylation occurred at CpG island-associated sites whereas cancer hypomethylation and aging hyper- and hypomethylation changes were much more enriched at open sea locations (Fisher's $P < 0.001$, ORs = 24.7, 6.7, and 10.3). Similarly, with respect to gene locations, these latter alterations were once more enriched at intergenic and intronic sequences (fig. 3F, Fisher's $P < 0.001$, intergenic ORs = 7.5, 2.4, and 3.0, intronic ORs = 1.6, 2.3, and 2.1, respectively). Again, cancer hypermethylation was directly associated with higher CpG-density (supplementary fig. S5B, Supplementary Material online) and accounted for the largest alterations (supplementary fig. S5C, Supplementary Material online). In general, the patterns of disease-associated DNA methylation changes, and especially the contrast between cancer and aging alterations, followed the same trends as those seen for human tissue.

To examine if the parallelisms observed extended to the functional pathways involved, we mapped the dmCpGs to genes and carried out pathway enrichment analyses. Once more, within the dmCpGs mapping to genes, despite cancer hyper-dmCpGs being by far the most numerous, they were concentrated across fewer genes than expected when compared with hypo-dmCpGs and aging changes (fig. 3G, χ^2

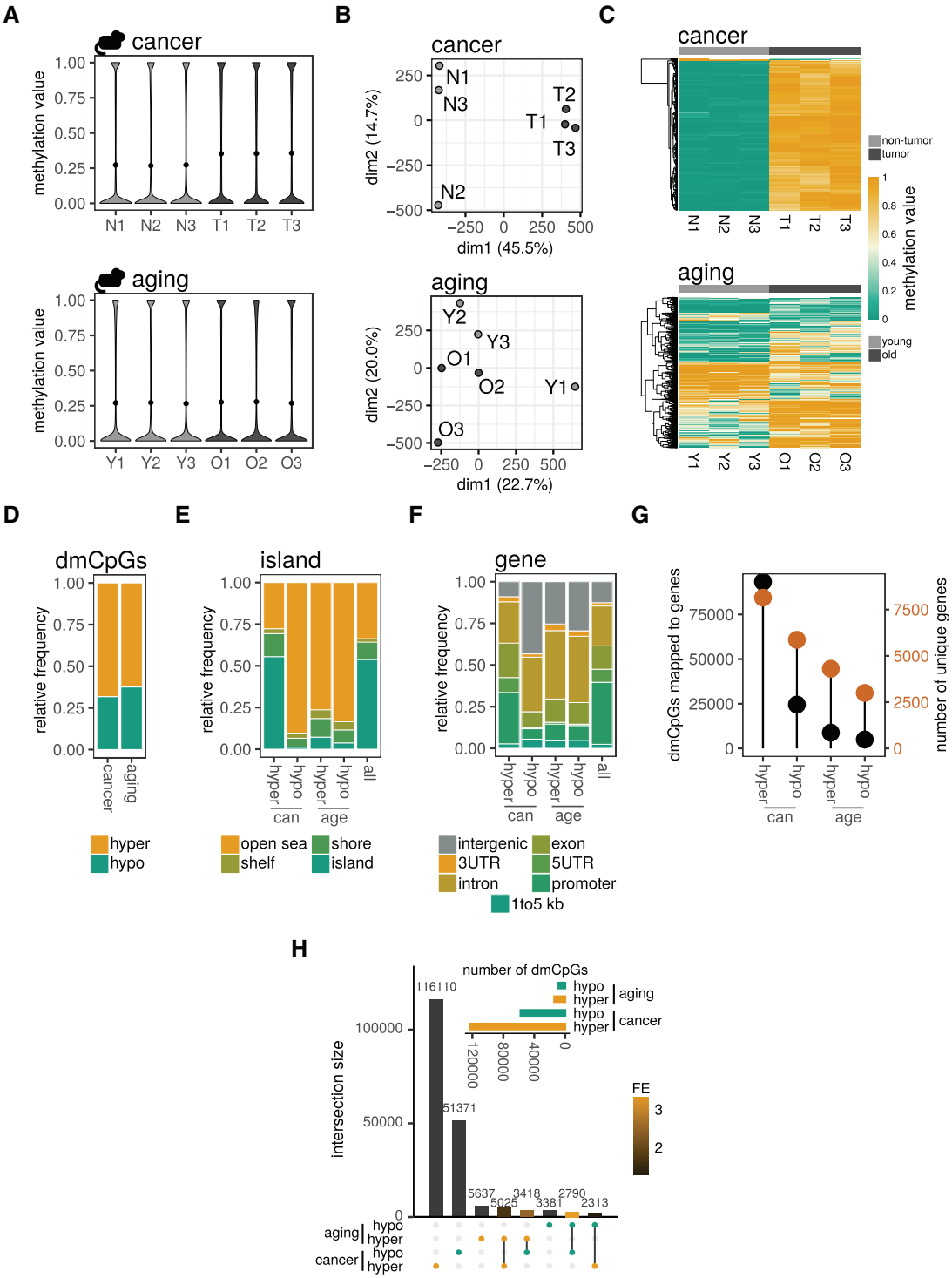

Fig. 3. Cancer and aging DNA methylation alterations in mouse. (*A*) Violin plots of global levels of DNA methylation of all CpGs across the profiled samples (N, nontumor; T, tumor; Y, young; O, old). (*B*) PCA plots showing the distribution of the samples across the first two principal components, with the percentage of variance each explains. (*C*) Heatmaps showing the methylation values of the top 1,000 dmCpGs found in the aging and cancer comparisons. (*D*) Barplots indicating the relative numbers of hyper- and hypo-dmCpGs (differentially methylated CpGs) in aging and cancer. (*E* and *F*) Barplots indicating the distribution of the dmCpGs across CpG island and gene-region locations, including the distribution of the universe of CpGs profiled by the RRBS ("all"). (*G*) Lollipop plots comparing the number of dmCpGs mapped to genes (left axis, black dots) with the number of different genes that the dmCpGs mapped to (right axis, orange dots). (*H*) UpSet plot describing the dmCpG sets and their intersections, indicating the size of the sets and the fold-enrichment (FE) of the intersections based on their expected overlaps.

$P < 0.001$). In this case, the pathway enrichment analyses revealed significant results for all sets of dmCpGs (supplementary fig. S6, Supplementary Material online, full results are available as extended data sets). Cancer hyper-dmCpGs, as with human, were enriched in Polycomb-target genes, GPCRs (Wnt-activated in particular), and cellular development

pathways, including neuronal ones. On the other hand, hypo-dmCpGs were enriched in generic cancer pathways, cell-signaling pathways (including GPCRs), and inflammation terms. Aging hypermethylation was enriched in cell-interaction terms, and hypomethylation in cell-signaling pathways. These results indicate that, in contrast to human processes, aging-associated DNA methylation alterations in mouse may have more directed functional consequences at the gene level.

We integrated the cancer- and aging-associated dmCpGs and found significant overenrichment for all the common sets (fig. 3*H* and supplementary table 4, Supplementary Material online, SuperExactTest all $P < 0.001$). As occurred with human tissue, whereas all the intersections displayed overenrichment, common cancer-, and aging hypo-dmCpGs displayed the highest fold enrichment and cancer hyper- along with aging hypo-dmCpGs the lowest. Next, we performed pathway enrichment analyses of the intersections. This time, we observed no significant enrichments for those containing cancer hyper-dmCpGs, but modest enrichments for inflammation terms did appear, especially for cancer hypo- and aging hyper-dmCpGs (full results are available as extended data sets). Taken together, these results suggest that, in the case of mouse and, in contrast to the human results, the Polycomb-target, developmental gene hypermethylation observed in cancer is not as evident in aging.

The Interspecies Genomic Context of Cancer and Aging DNA Methylation Alterations

In light of the parallelisms observed, we sought to extend the analyses to additional levels of epigenomic regulation by integrating the methylomic data with the previously mentioned histone modification data—involving the histone modifications H3K4me1, H3K4me3, H3K27ac, H3K36me3, H3K27me3, and H3K9me3—across a panel of common healthy tissues in human and mouse (see Materials and Methods). We performed overenrichment analyses (FDR < 0.05) to test the association of the cancer- and aging-related dmCpGs discovered in the two species with the different histone modifications, in order to define chromatin signatures linked to the DNA methylation changes. Using this procedure, we were able to find robust and well-defined similarities between the two species (fig. 4*A* and supplementary table 5, Supplementary Material online). As regards cancer-associated changes, we observed an association of DNA hypermethylation with the repressive H3K27me3 modification and the more active enhancer/promoter-associated H3K4me1. In the case of human, cancer hypermethylation was also strongly enriched at H3K9me3 locations, which are linked to heterochromatic and repressive states. On the other hand, cancer hypomethylated dmCpGs were mainly enriched in the H3K9me3 mark for both species. These cancer-associated signatures have been described across multiple tissues for human (Pérez et al. 2018) whereas the literature for mouse is scant.

When looking at aging-related signatures, we again found comparable patterns between human and mouse. Moreover, for both species, these patterns were markedly different to those found for cancer. We observed enrichments at locations marked with gene-body H3K36me3, heterochromatin H3K9me3, and enhancer/promoter H3K4me1. In this case, even though we observed an increase in H3K4m1 association at aging hypomethylated loci when compared with aging hypermethylation—a finding reported for human in several tissues (Fernández et al. 2015; Pérez et al. 2018)—the global patterns of enrichment found for aging dmCpGs differed moderately from those described in the literature, a result which could have been influenced by the fact that the "classical" chromatin signatures have been mostly derived by using array technology, which targets similar but not identical genomic contexts (supplementary fig. S7, Supplementary Material online). In further support of this idea, a recent study comparing aging-associated DNA methylation changes in liver between human and mouse—using array and RRBS data for each species respectively—reported a high enrichment in H3K36me3-associated loci specifically for mouse aging changes as compared with human (Wang et al. 2017).

Because the histone modification results exposed epigenomic similarities between human and mouse for the methylation changes associated with cancer and aging, we extended our analyses to discover and study chromatin states, which are defined by spatial combinations of the histone marks and reflect functional genomic states (Ernst and Kellis 2010). We trained 18-state models using the six core histone modifications for our selected panel of data sets by using multivariate Hidden Markov Models (see Materials and Methods). The nature of the human states defined was similar to that previously described (Roadmap Epigenomics Consortium et al. 2015) in terms of their histone mark composition and co-occurrence with other genomic elements (supplementary fig. S8*A*, Supplementary Material online). The states found for mouse were, in turn, notably similar to human (supplementary fig. S8*B*, Supplementary Material online), with the exception of some specific ones, such as the one assigned as "EnhG2" (genic enhancer 2). These observations suggest that multilevel epigenomic regulation is functionally parallel between the two species.

We evaluated the overenrichment of the dmCpGs in the different chromatin states (fig. 4*B* and supplementary table 6, Supplementary Material online). Cancer hypermethylation was strongly enriched in bivalent transcription start sites and Polycomb-repressed sites (states 14, 16), and somewhat enriched in transcription start sites and enhancer sites in general (states 2–4, 7–11) for both species. Additionally, and as suggested by the stronger enrichment in H3K9me3 for human as compared with mouse (fig. 4*A*), the chromatin-state analysis revealed that cancer hypermethylation also occurred, specifically for human, both at heterochromatic regions and at zinc-finger (ZNF) genes/repeats (states 12, 13). Indeed, ZNF genes, which are characterized by the high presence of H3K9me3 (Blahnik et al. 2011), displayed the highest overenrichment of all the chromatin states in human cancer hypermethylation. Interestingly, this family shows a great evolutionary expansion in human as compared with other species, including mouse, and its members are posited to play roles in novel transcriptional repression activity

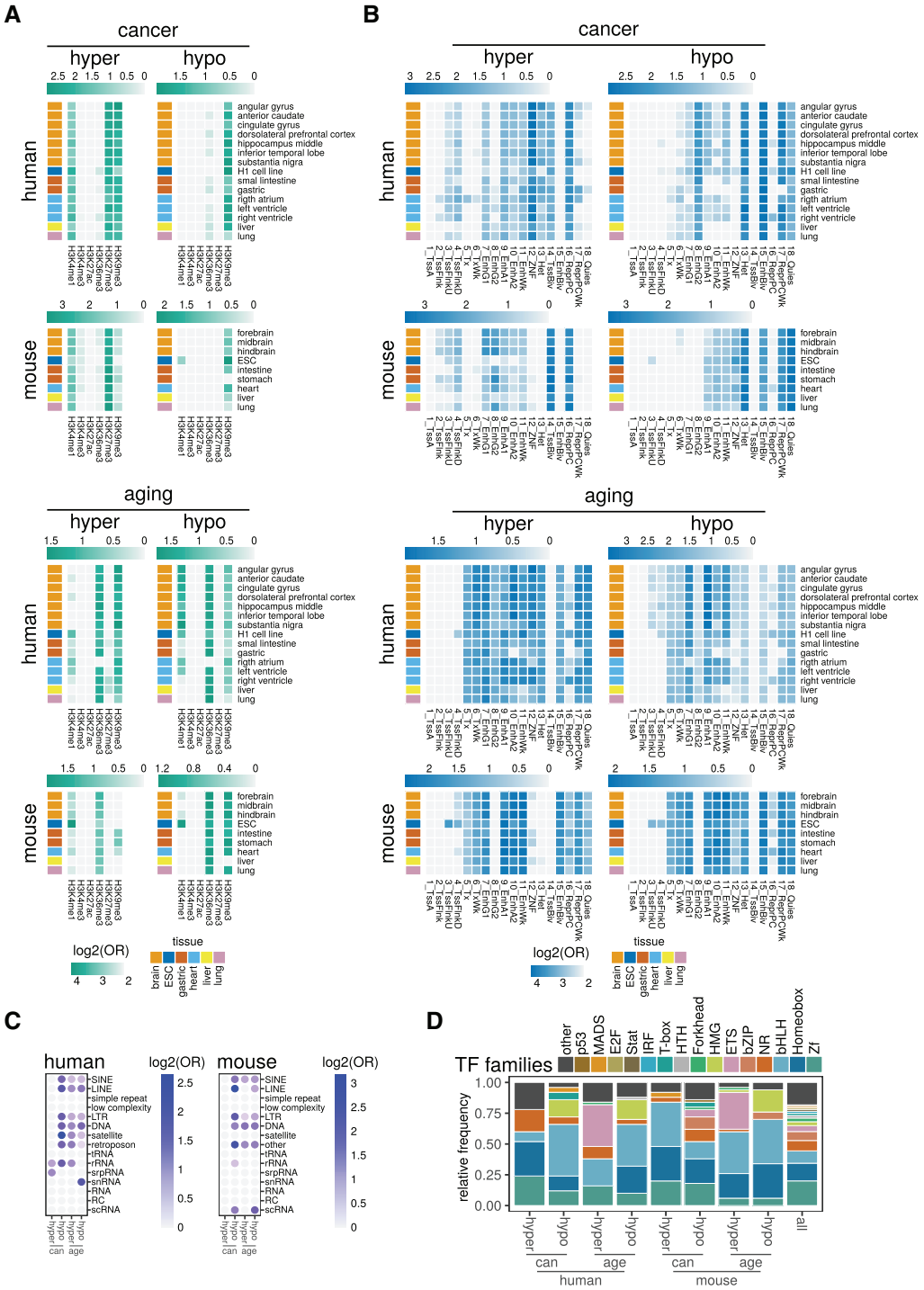

FIG. 4. Chromatin landscape of aging and cancer DNA methylation alterations in human and mouse. (*A* and *B*) Heatmaps showing the significant overenrichments in log2(odds ratio) of hyper- and hypo-dmCpGs in the cancer and aging comparisons, for human and mouse in ChIP-seq peak locations of different histone marks across a panel of tissues or in chromatin states built across the same panel of tissues. (*C*) Bubble plots showing the significant overenrichments in log2(odds ratio) of hyper- and hypo-dmCpGs in the cancer and aging comparisons in the locations of repetitive DNA classes across the genome for human and mouse. (*D*) Barplots indicating the relative numbers of transcription factors, classified by their domains, whose motifs were found to be significantly enriched (HOMER analysis, FDR<0.05) at the genomic locations of the different dmCpG sets, compared with the numbers of all of the transcription factors tested ("all"). Only the top 50 most enriched, significant factors for each set were selected.

(Emerson and Thomas 2009). Thus, the DNA methylation deregulation observed in human cancer could be related to the targeting of human-specific pathways. Regarding cancer hypomethylation, for both species, most of the affected chromatin states (13, 15, 17, 18) involved repressive functions, such as heterochromatin or Polycomb-associated sites, with

some enrichment also being observed in various enhancer associated states. In the case of aging-associated dmCpGs, and much like with the histone signatures, the enrichments were spread out across similar states for hyper- and hypomethylation, and involved gene-body transcription states (5, 6), enhancer states (7–11), and repressive states (15–18). Especially in the case of human, a particular increase in enrichment in enhancer states was observed for aging hypomethylation. The human and mouse profiles were globally similar, with some differences, such as in EnhG2 (state 8) probably being explained by the pre-existing differences in the chromatin models built (supplementary fig. S8, Supplementary Material online). Collectively, our results show that the chromatin context of DNA methylation changes in cancer and aging is considerably different, and that the specific features of each process are robustly recapitulated in human and mouse.

We also looked at the association between DNA methylation alterations and repetitive DNA elements (fig. 4C and supplementary table 7, Supplementary Material online; see supplementary fig. S9, Supplementary Material online, for the analysis expanded to repetitive DNA classes). First, regarding cancer, repetitive loci were mainly associated with hypomethylation across the same large classes of repeats (SINE, LINE, LTR, DNA), as has been well-described previously (Ross et al. 2010), with the main species difference being that satellite repeats showed no changes in DNA methylation for mouse (fig. 4C). In the case of SINE repeats, both the primate-specific Alu family in human and the related B elements (B1, B2, B4) in mouse were targeted by DNA hypomethylation (supplementary fig. S9, Supplementary Material online). Interestingly, we also observed in human the hypermethylation of ribosome-associated elements (rRNA, srpRNA), a finding which has been reported for some types of cancers (Srivastava et al. 2016), although literature on the subject is scarce. Secondly, in the case of aging, both hyper- and hypomethylation were associated with repetitive DNA, with the patterns observed being similar to those of cancer hypomethylation. Although RRBS is mostly limited to examining single-copy sequences, the bidirectionality of the changes found here for aging may help explain observations in the literature that fail to describe any noticeable trend of loss of methylation during aging in the genome-wide levels of brain methylation both in mouse (Hadad et al. 2019) and human (Lister et al. 2013; McKinney et al. 2019). No associations whatsoever were found with simple or low-complexity repeats, which were the main classes which we had previously linked to low levels of basal methylation in nontumoral tissue (fig. 1J). Taken together, these observations suggest that repetitive DNA elements are subject to similar DNA methylation alterations in cancer and aging for both species, with the exception of satellite repeats, which appear to be more affected by these processes in human as compared with mouse.

We next sought to study the possible effects of the DNA methylation changes found at TF-binding sites, because TFs are downstream genomic regulators directly in contact with the genomic structure. To this end, we performed a HOMER analysis to find enrichment in TF motifs associated with dmCpG sites, selecting the top 50 significant (FDR < 0.05) motifs with the highest enrichments found for each comparison (see supplementary table 8, Supplementary Material online, for extended results). First, to obtain a general overview of the types of TFs affected by the DNA methylation changes, we classified our results into TF domain-families (fig. 4D). With this approach, we found both similarities and differences between the processes as well as between species. Cancer hypermethylation dmCpGs were associated with Homeobox TFs for both species (Fisher's $P < 0.01$, ORs $= 2.7$ and 2.7 for human and mouse, respectively), but also associated with NRs in human (Fisher's $P < 0.05$, OR $= 2.8$) and bHLH-domain TFs (Fisher's $P < 0.001$, OR $= 7.8$) in mouse. Regarding cancer hypomethylation, it was particularly enriched at bHLH- and HMG-domain TF motifs in human (Fisher's $P < 0.001$, ORs $= 11.4$ and 8.8, respectively). On the other hand, the patterns of TF sites targeted by aging hyper- and hypomethylation were similar in both species. Aging hypermethylation particularly affected ETS-domain TF motifs (Fisher's $P < 0.001$, ORs $= 48.5$ and 26.9 for human and mouse, respectively) whereas hypomethylation involved bHLH- and HMG-domain TF sites (Fisher's $P < 0.001$, ORs $= 6.9$, 12.0, 7.8, and 16.6 for human bHLH, HMG and mouse bHLH, HMG, respectively).

Subsequently, we focused on the specific TFs involved in the previous observations, by looking at the main sets of shared or unique TFs affected in each comparison (supplementary fig. S10, Supplementary Material online). As suggested by the previous results, the patterns of species-common alterations at TF motifs were especially noticeable for aging-related changes. Aging-associated hypermethylation of ETS-domain TFs, a family strongly linked to oncogenic processes (Sizemore et al. 2017), was the main common epigenetic phenomenon between mouse and human. Several Sox TFs (HMG domains) were commonly targeted by aging hypomethylation in both species, such as Sox2, whose expression has been shown to be deregulated with aging in both human and mouse brain (Carrasco-Garcia 2018). Interestingly, Sox TFs also appeared affected by cancer hypomethylation in human, suggesting possible links between the regulation of aging and of cancer, as has been shown for Sox4 in other tumors (Foronda et al. 2014). In addition, the deregulation of methylation patterns at bHLH-domain TF sites appeared to be a general trend in both cancer and aging processes, appearing at loci related to specific neuronal development and differentiation, including TFs such as NeuroG2, NeuroD1, Olig2, or Atoh1 (Dennis et al. 2019) and also involving Homeobox TFs such as DLX1/2.

To sum up, the analysis of the links between DNA methylation alterations and TF-binding sites revealed that clear functional parallelisms between human and mouse exist in the aging scenario. This is in contrast to our earlier gene pathway enrichment analyses, which showed that most of the affected genetic pathways common in human and mouse involved cancer hypermethylation (of Polycomb-targets and developmental genes). This discrepancy could be explained by aging-associated methylation changes being more functionally important when occurring at genomic sites not

directly associated with gene locations (such as many TF-binding sites). Moreover, as shown in earlier stages of this study (figs. 2F and 3F), aging changes concentrate more at nongenic sites than cancer changes. Thus, the identification of functional targets implicated in DNA methylation alterations in aging may be more effective when focusing on TF-binding site alterations.

Interspecies Integration of Cancer and Aging DNA Methylation Alterations

After establishing the parallel features of DNA methylation alterations in human and mouse, we integrated our CpG-specific data, to be able to directly compare methylation values between the two species at orthologous locations, by lifting the mouse coordinates to the human genome (see Materials and Methods). We recovered interspecies measurements for a sizeable set of 59,100 CpG sites, 35% of which had been classified as dmCpGs at any comparison (see supplementary table 9, Supplementary Material online, for the liftover statistics; the full list of orthologous sites is available as an extended data set). These orthologous sites were overenriched in CpG island locations when compared with the original backgrounds (Fisher's $P < 0.001$, ORs = 7.1 and 6.1 for human and mouse, respectively), and largely maintained their relationship to CpG islands in both species, suggesting that they covered similar functional regions in both genomes (fig. 5A). Correlation analysis of the DNA methylation values of the interspecies lifted CpGs revealed that, excluding cancer, the samples clustered within their species (fig. 5B). Tumors, on the other hand, again showed widespread reconfiguration, clustering independently and, interestingly, mouse tumor CpGs were slightly more correlated to human tumors than to the rest of the mouse samples. Examining only nontumoral samples, methylation was found to be highly correlated between the two species (fig. 5C, left panel; Pearson coef. 0.84, $P < 0.001$). Moreover, when we retained only CpG sites which were lowly variable within each species (SD < 0.1 methylation value) most of the species-discordant CpG sites disappeared (fig. 5C, right panel; Pearson coef. 0.92, $P < 0.001$). Indeed, the subsets of high-variability sites for human and mouse (CpGs with SD $\geq$0.1 methylation value) showed much higher overlap than expected (fig. 5D; Fisher's $P < 0.001$, OR = 16.7), suggesting that locations of methylation variability are conserved between the two species. In order to strictly define species-discordant CpGs (sdCpGs), we performed a differential methylation analysis (FDR < 0.05, average change in methylation >0.2) at the 59,100 loci with interspecies measurements (the list of sites and sdCpGs is available as an extended data set). We identified 1,845 sdCpGs (supplementary fig. S11A, Supplementary Material online), evenly split between those presenting higher and those lower methylation in mouse as compared with human (hyper- and hypo-sdCpGs, respectively, supplementary fig. S11B, Supplementary Material online). Interestingly, when annotating the sdCpGs either to human or mouse genomic contexts, we observed considerable differences in the distribution of the sites, especially according to their CpG island status (supplementary fig. S11C and D, Supplementary Material online). Following this observation, we directly assessed the transitions between the genomic contexts of the sdCpGs in the genome of each species (fig. 5E), finding that, particularly in the case of hyper-sdCpGs, there was a considerable difference in the genomic association to CpG island status between the two species. In this case, CpGs which in human-occupied CpG islands and in mouse-occupied less-dense regions (shores, open sea) had higher methylation levels in mouse (i.e., were classified as hyper-sdCpGs). This is in line with our previous observations (fig. 1F–H) of denser genomic loci being associated with lower methylation values in both species. This finding suggests that interspecies differences in DNA methylation levels may be partly explained by local changes in the functional genomic context of their respective genomes, such as CpG island membership (Hernando-Herraez, Garcia-Perez, et al. 2015).

Next, we focused on the fraction of previously described species-specific cancer- or aging-associated dmCpGs for which we had interspecies measurements (see supplementary table 9, Supplementary Material online, and extended data sets). This allowed us to directly compare cancer- and aging changes at specific loci between human and mouse. We found common and statistically significant intersections between the two species for all sets (fig. 5F and supplementary table 10, Supplementary Material online; SuperExactTest all $P < 0.001$), with the biggest intersection being found for cancer dmCpGs, although the observed size and fold enrichment of the intersections was probably influenced by the differing sizes of the various sets. Differentiating by direction of change revealed that the biggest intersections were dominated by hypermethylation changes in both species (supplementary fig. S12, Supplementary Material online), especially when involving cancer sets. Aging, on the other hand, was limited to much smaller intersections with directionality not playing such an apparent role. Nonetheless, the observed intersections suggest that, added to the previously described similarities in genomic contexts (fig. 4), when we focus on specific orthologous loci, common sites of DNA methylation alteration across the two diseases and the two species are readily detected.

After the characterization of specific similarities between human and mouse, we turned to examine the influence of sequence conservation in the explanation of the epigenetic patterns found. We integrated our data with UCSC conservation scores for multiple alignments across vertebrate species (see Materials and Methods). First, regarding disease-associated alterations in general (dmCpGs), we observed that, for both species, cancer hypermethylation occurred at more conserved regions than did either cancer hypomethylation or aging changes (fig. 5G). This finding can support the notion that cancer hypermethylation may impact conserved axes of gene regulation across species, something which we had observed through the pathway enrichment analyses. However, these trends may also be partly explained by the genomic distribution of the alterations, because hypomethylation and aging alterations occur more often at intergenic and intronic regions (see figs. 2F and 3F), which we found (using all the profiled CpGs) to be less conserved than other

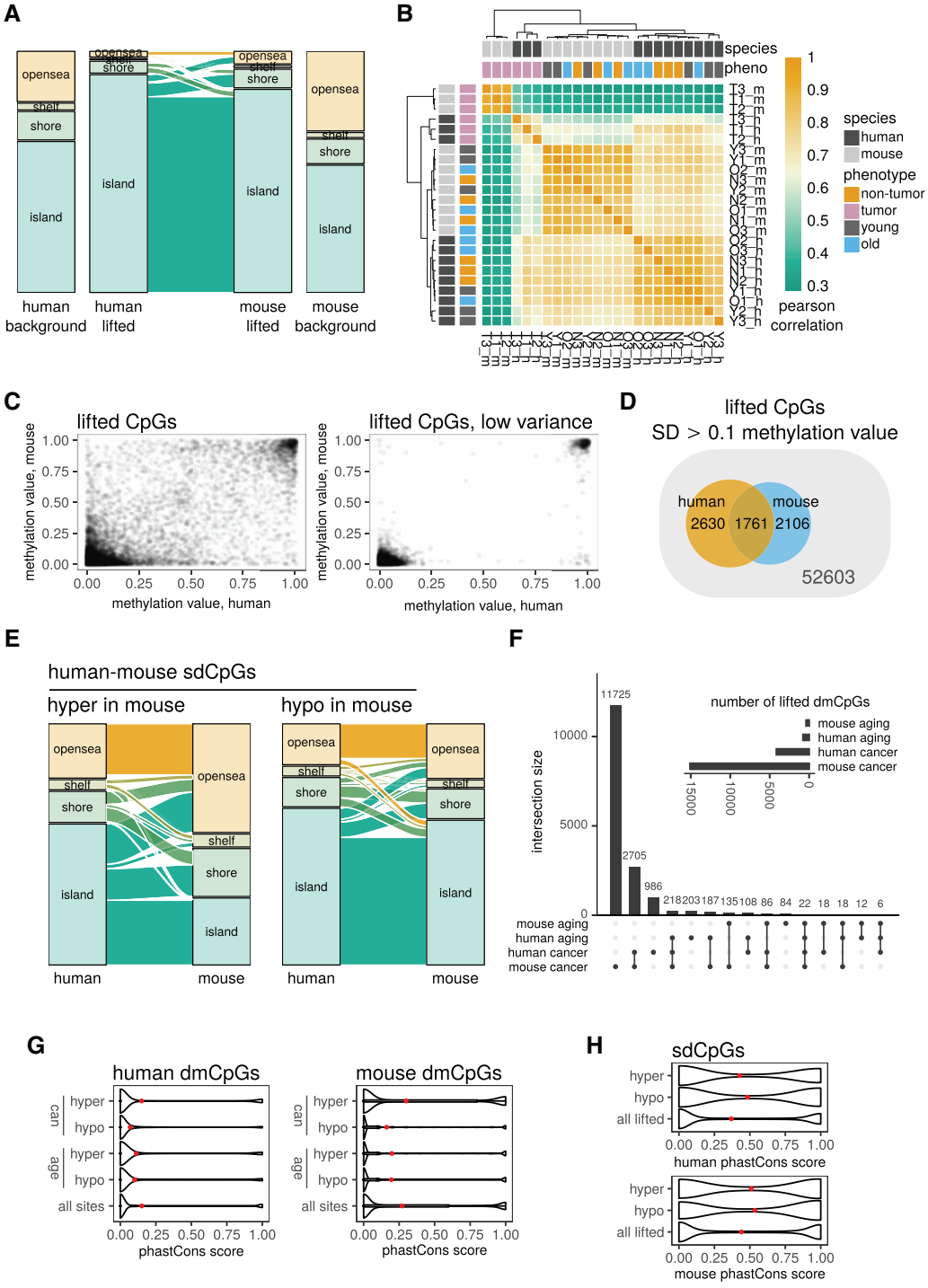

Fig. 5. Interspecies DNA methylation alterations in human and mouse. (*A*) Sankey diagram describing the distribution of changes in the CpG island status of the 59,100 interspecies orthologous CpGs when considering either the human or the mouse genomic annotation. On either side, the original backgrounds of all the profiled CpG sites are shown. (*B*) Heatmap showing the pairwise Pearson correlations of the DNA methylation levels of the interspecies CpGs across all samples. (*C*) Scatter plots showing the correlation between DNA methylation levels at all the profiled interspecies sites (left plot) or excluding intraspecies highly variable CpG sites with SD $\geq$0.1 methylation value (right plot) (tumors excluded). (*D*) Venn diagram showing the intersection of interspecies lifted CpG sites classified as highly variable within each species (SD $\geq$0.1 methylation value) for both mouse and human. (*E*) Sankey diagrams showing the distribution of hyper- and hypo-sdCpGs (species-discordant CpGs) across CpG island locations, taking into account the differences in the genomic context of the human and mouse genome. (*F*) UpSet plot describing the sets of cancer and aging dmCpGs with interspecies measurements, showing their sizes and intersections between human and mouse. (*G*) Violin plots showing the distribution of phastCons conservation scores for the detected aging- and cancer dmCpGs in human and mouse, compared with the profiled RRBS background (red dots indicate average values). (*H*) Violin plots showing the distribution of phastCons scores for the detected interspecies sdCpGs, compared with the 59,100 lifted background, either in human or mouse-centered conservation score (red dots indicate average values).

genetic regions such as exons or promoters, that is, those more impacted by cancer hypermethylation in both species (supplementary fig. S13*A* and *B*, Supplementary Material online). Second, within our 59,100 interspecies loci, we found that sdCpGs were enriched at conserved regions (fig. 5*H*) more frequently than the rest of the interspecies sites. This observation could at first sight appear to challenge the existing literature, which describes the positive correlation between sequence conservation and epigenetic conservation (Zhou et al. 2017), but, again, it may partly be explained by the sdCpGs being enriched at exon locations (see supplementary fig. S11*D*, Supplementary Material online), which are more conserved (see supplementary fig. S13*B*, Supplementary Material online). To address these issues, we applied stratified sampling to compare the CpG sets of interest against randomly sampled background sets of CpGs which had matching distributions of gene locations. With this strategy, we found that most of the trends regarding the difference in conservation scores observed could indeed be partly explained by the genomic context of the CpGs involved, both for cancer- and aging dmCpGs (supplementary fig. S14*A*, Supplementary Material online) and for human- and mouse sdCpGs (supplementary fig. S14*B*, Supplementary Material online). These results suggest that both disease- and species-specific epigenetic traits are influenced by the genomic context surrounding the specific CpG sites involved.

Interspecies Hotspots of Cancer- and Aging Epigenetic Deregulation

Finally, we used the interspecies CpGs to look for common hotspots of cancer- and aging epigenetic deregulation between human and mouse which could have downstream functional implications. We focused on the 22 CpGs which had been identified as dmCpGs in both species and both processes (described in fig. 5*F*). Of these, nine CpG sites presented concordant changes in both species (i.e., being deregulated in the same direction, in cancer and aging, for human and mouse) and were associated with a total of six genes (see supplementary table 12, Supplementary Material online). In order to study the possible functional consequences of the DNA methylation alterations, we obtained TCGA expression data for the LGG–GBM glioma cohorts and analyzed the expression of these candidate genes and their association with survival across glioma patients (see Materials and Methods). Notably, we found evidence of the strong deregulation of five out of the six genes in glioma samples as compared with controls (see below), indicating that our robust approach of looking at orthologous and concordant aging-cancer DNA methylation changes leads to the discovery of functionally relevant targets.

We first identified the *AGAP3* gene as being a target of both cancer-associated and aging-associated deregulation in both human and mouse. *AGAP3* is a component of the neural NMDA receptor complex and AGAP proteins regulate receptor trafficking in neurons (Oku and Huganir 2013). In cancer, this gene showed strong hypermethylation of internal CpG sites which, in both species, colocated with CpG islands upstream and downstream of many alternative transcripts (fig. 6*A*, upper plots), whereas DNA methylation alterations in the same direction, albeit smaller, also appeared with aging in both species (fig. 6*A*, lower plots). The specific location of the DNA methylation changes found hinted at a nonrandom and perhaps functional role for the alterations. Thus, we analyzed *AGAP3* expression in the TCGA LGG-GBM and found a strong decrease in expression in tumors (fig. 6*B*, Wilcoxon rank-sum $P < 0.001$). Moreover, expression levels were strongly associated with overall survival in the cohort (fig. 6*C*, log-rank test $P < 0.001$), with the more lethal GBM subtype having the lowest levels of expression and survival (supplementary fig. S15*A* and *B*, Supplementary Material online). Interestingly, as can be seen in figure 6*A*, the *AGAP3* internal hypermethylation events colocated with the ending and start of several isoforms. We therefore explored the levels of the main *AGAP3* isoforms (supplementary fig. S15*C*, Supplementary Material online) and their association with survival (supplementary fig. S15*D*, Supplementary Material online), discovering that a specific isoform (uc003wjf/ENST00000473312) was mostly explanatory of the effect (supplementary fig. S15*E*, Supplementary Material online, Cox Models log-rank test $P < 0.001$). Moreover, it corresponded to an isoform truncated close to the internal region targeted by the DNA methylation alterations (highlighted in fig. 6*A*). Indeed, studies have described isoform variants of *AGAP3* with biological functions (Oku and Huganir 2013) some of which may have roles in oncogenic processes through alterations in their mRNA expression (Shimizu et al. 2019). These observations exemplify an interspecies DNA methylation deregulation process which, as shown for human, could have a role in *AGAP3* deregulation in glioma, possibly through an effect on isoform switching, with aging-associated methylation changes preluding the cancer-associated alterations.

In addition to *AGAP3*, five other genes were associated with concordant interspecies dmCpGs (fig. 6*D*): *GPRIN1*, *AJM1*, *LHX2*, *CCDC177*, and *TPGS1*. Notably, four of the genes identified, much like with *AGAP3*, were clearly downregulated in tumoral samples as compared with healthy tissue (fig. 6*E*, Wilcoxon rank-sum tests), with the GBM subtype often displaying the strongest downregulation. We also investigated their relationship with survival and found a strong association between gene expression and overall survival for *GPRIN1*, with a more modest effect for *AJM1* and, interestingly, a trend for an association between *CCDC177* expression and survival that was specific to the GBM subtype (supplementary fig. S16, Supplementary Material online). *LHX2* is involved in neural development (Chou and Tole 2019), whereas *GPRIN1* and *TPGS1* play a role in neural cytoskeleton dynamics (Regnard et al. 2003; Nordman and Kabbani 2012) and *AJM1* is related to apical junctions (Köppen et al. 2001). As such, the deregulated genes appear to be predominantly associated with cytoskeletal processes. *LHX2* is known to be oncogenic in several types of cancer (Song et al. 2019) although we have found it to be downregulated in the case of glioma, as has been reported previously (Cheng et al. 2019). Nonetheless, the specific role of *LHX2* and of the rest of the genes in glioma

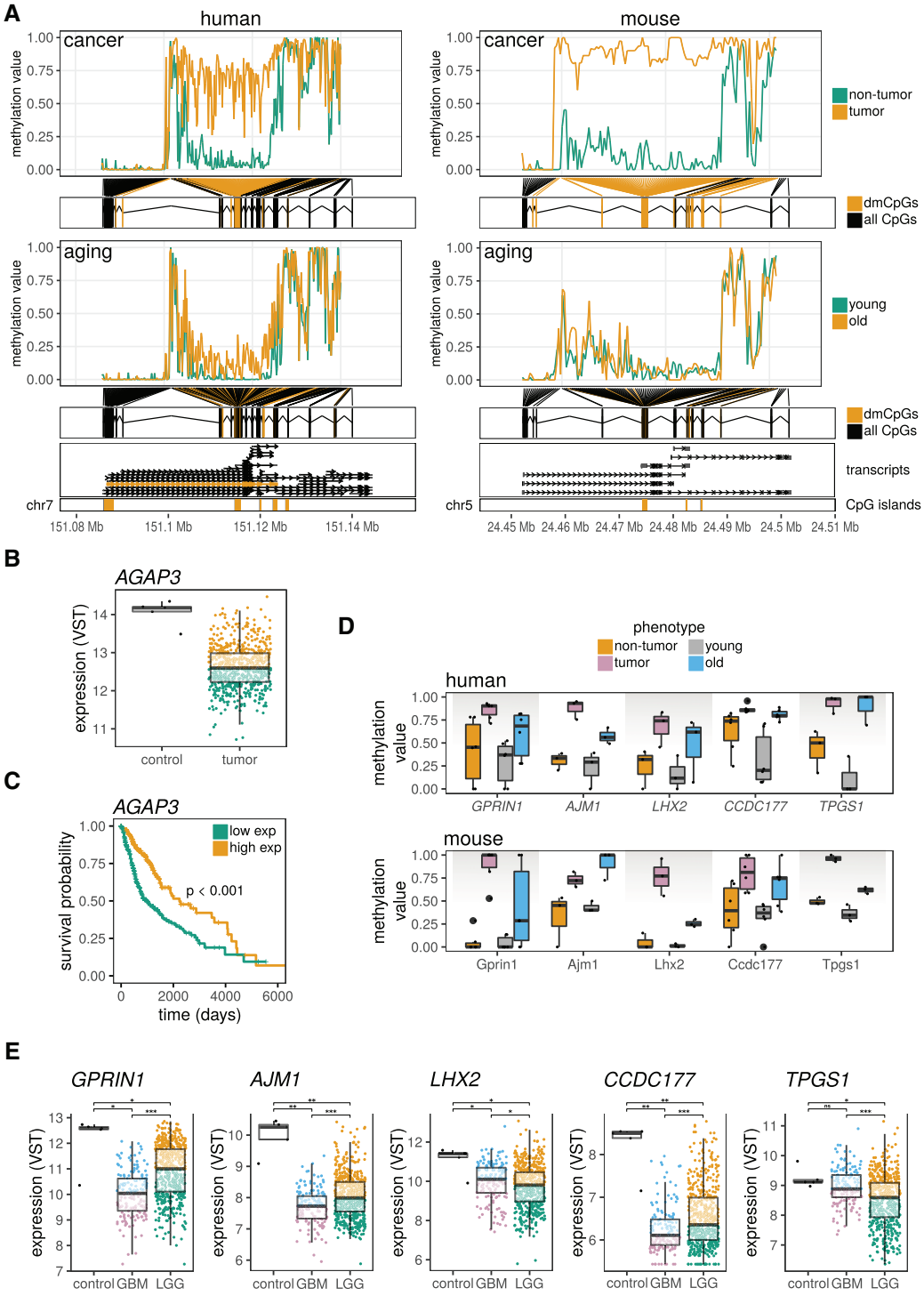

Fig. 6. *AGAP3* is an interspecies cancer- and aging-associated target of DNA methylation alterations. (*A*) Plots describing the average methylation values of tumoral, nontumoral, young, and old samples in human and mouse across the profiled CpG sites mapping to *AGAP3*. Sites marked in yellow represent significant dmCpGs in the aging and cancer comparisons. Below, the transcripts and CpG islands associated with the genomic region described are shown. (*B*) Boxplots showing the expression values of the *AGAP3* gene in control and tumor samples from the TCGA glioma cohort (GBM-LGG). Expression is measured in VST (variance stabilizing transformation) normalized units. High and low-expression samples (above and below the median) are in yellow and green. (*C*) Survival curves showing the association of high and low *AGAP3* expression groups with overall survival across glioma patients. (*D*) Boxplots describing the methylation values of the interspecies dmCpGs that presented concordant changes in cancer and aging across human and mouse and which were associated with specific genes. (*E*) Boxplots showing the expression values (VST normalized units) of the genes associated with the interspecies concordant dmCpGs in control and tumoral samples from the TCGA glioma cohort (GBM-LGG). Glioblastoma (GBM) and lower grade glioma (LGG) samples are shown separately. High- and low-expression samples (above and below the median) are color-coded. Wilcoxon rank-sum tests: **P*<0.05, ***P*<0.001, ****P*<0.0001, "ns," *P*>0.05.

described here, or in cancer as a whole, remains largely unexplored.

Conclusions

In this work, we have performed a systematic analysis of the DNA methylation dynamics associated with cancer and aging in brain tissue in human and mouse. We employed RRBS as a platform which allowed us to profile comparable genomic locations in both species. With this strategy, we first confirmed, in our study system, the parallelisms in baseline DNA methylation patterns between human and mouse. Subsequently, we explored the alterations associated with cancer and aging in both species. Our analyses revealed that there are considerable differences in the genomic and epigenomic features of DNA methylation alterations in cancer and aging, and that the particular characteristics of the alterations found for each process are globally conserved in human and in mouse (see supplementary fig. S17, Supplementary Material online, for a graphical outline of the main findings). We found specific parallelisms at the level of their genomic locations and the magnitude of the changes, the genetic pathways involved, the chromatin context of the alterations, the families of repetitive DNA elements affected, and the alteration of TF regulation sites, among others. Interestingly, we observed that aging-associated changes in general seem to impact more specific functional gene pathways in mouse than in human. Nonetheless, we found that the classical hypermethylation of developmental gene pathways is more shared between cancer and aging in human than in mouse. Intriguingly, epigenetic alteration of satellite repeats appeared to be a feature of both aging and cancer in human, but not in mouse. Significantly, we found very clear interspecies similarities in the aging-associated alterations occurring at the TF-binding sites of specific families, suggesting that the functional impact of aging methylation changes could indirectly occur through TF deregulation rather than through the direct targeting of specific genes. Finally, we derived a large set of orthologous CpG sites with interspecies measurements, which allowed us to observe that sites of epigenetic variability are conserved between the species, that interspecies differences between human and mouse can be partly explained by changes in the local genomic context associated with the CpG sites and that there are common interspecies loci which are altered during both aging and cancer.

The results presented in this manuscript cover several aspects of the epigenetic dynamics in human and mouse, but we should also draw attention to the limitations of our study. First, we recognize that the sample size used in our study design is limited ($n = 3$ per group), which is principally because of the substantial number of groups being analyzed (four phenotypes in two species). However, the interspecies corroboration of the results does indicate that the observed patterns are robust. Second, we have focused on a single tissue (brain). Although this allowed us to perform a comparative study in greater depth, and it is known that epigenetic alterations that arise in cancer or in aging are similar across tissues (Michalak et al. 2019), it will be of value to explore the interspecies conservation of epigenetic signatures in other tissues (Maegawa et al. 2017; Wang et al. 2017), and also other species such as pig or rhesus. Third, we used RRBS to profile analogous genomic locations in human and mouse, a technology which focuses on CpG-dense locations and is more limited when studying intergenic regions. That said, the use of RRBS allowed us to define a large subset of interspecies measurements (59,100 CpG sites) with adequate coverage across all the samples.

From an evolutionary perspective, systematic multispecies comparisons are of great value in the understanding of the etiology of disease. The uncovering of species-common pathways provides robust molecular explanations for the mechanisms at play, whereas interspecies differences may help explain the phenotypic divergences observed across the species. For instance, our observation of the human-specific epigenetic deregulation of zinc-finger genes in cancer may be related to the importance of this family in primates, which have experienced considerable lineage-specific expansion as compared with rodents (Huntley et al. 2006; Emerson and Thomas 2009). The emphasis on species-specific pathways, whereas relevant to evolutionary biology is, conversely, of less value in clinical biology, which seeks the translation of mechanisms to human. In this respect, systematic, genome-wide, multispecies integrative studies such as the one presented here have the potential to answer questions relating to both ends.

Taking a more clinical viewpoint, we can foresee that our results may spark interest in different areas. The mouse is one of the main preclinical models for human disease and, as such, it is essential to have markers for the prediction of treatment success in human (Day et al. 2015). In this vein, in this work, we found that species-common epigenetic changes across diseases were readily associated with genes showing expression alterations in glioma in human. Thus, focusing on conserved epigenetic alterations such as those described here may help narrow the search for functionally relevant targets that are altered in disease. In this sense, current advances in the development of targeted epigenetic editing tools (Liu et al. 2016), combined with the knowledge of epigenetic conservation in mouse and human, may serve as a powerful tool to explore the overall impact of both epigenetic and pathway alterations in multiple biological scenarios.

On the other hand, the set of epigenetically conserved DNA methylation sites could pave the way for the design and development of novel custom high-throughput platforms, either in the context of cost-effective microarray technology or through the use of more sophisticated targeted bisulfite sequencing approaches. This may be of particular interest within the framework of epigenetic profiling during drug screening in preclinical models as alterations of such conserved disease-associated loci could help predict future treatment success in human. For example, glioma is a cancer type with a high prevalence of isocitrate dehydrogenase mutations which lead to important epigenetic alterations (Han et al. 2020). The main chemotherapy treatment in glioma involves temozolomide (Hirst et al. 2013), and its

effectiveness is also associated with various DNA methylation features beyond the well-known *MGMT* promoter methylation (Cheng et al. 2019). Thus, integrating epigenetic conservation knowledge could aid in the identification of drug-sensitive pathways in animal models, which in turn would improve the current standard-of-care treatments as well as also shed light on the mechanistic insights of novel compounds or those rescued from current drug repositioning approaches.

All in all, this work could constitute a baseline for other future studies focusing on the systematic evaluation of disease-associated epigenetic patterns across multiple tissues and species from a multiomic perspective in order to inform the search for inter- and intraspecies patterns of epigenetic regulation.

Materials and Methods

All statistical analyses were carried out using R statistical software (v3.6.2) unless stated otherwise. Software was managed through Bioconda (Grüning et al. 2018). Graphs were produced using R base functions and the *ggplot2* R package (v3.2.1) (Wickham 2016). For the illustration and testing of multiple-set intersections, the *UpSetR* (v1.4.0) and *SuperExactTest* (v1.0.7) R packages were used (Wang et al. 2015; Conway et al. 2017). For the illustration of gene-level DNA methylation profiles, the *ggbio* R package (v1.34.0) was used (Yin et al. 2012).

Sample Description and DNA Extraction

Regarding the aging cohorts, we collected brain cortex tissue from three young (24, 33, 35 years; female, male, male) and three old (65, 69, 70 years; female, male, female) human subjects and three young (10 weeks; all female) and three old (90 weeks; all female) C57BL/6HsdOla/CBA mice. Humans and mice have different lifespans and in consequence their ages are not equally translatable across their different life stages. The young groups were selected to be at least sexually mature adults, a characteristic which is attained at 10 weeks for mice and 20 years for humans (Dutta and Sengupta 2016). Regarding the old groups and considering an average lifespan of 80 years and 24 months for humans and mice, respectively (Dutta and Sengupta 2016), the 90-week-old mice would correspond to 69-year-old humans. Consequently, the age grouping for both species lies within a similar range.

For the cancer cohorts, we collected glioma and paired cortex tissue from three human subjects (71, 31, 44 years; female, female, male), and glioma and paired cortex tissue from three mouse subjects (12–16 weeks; all female) using an in-house murine model of glioma (see supplementary table 1, Supplementary Material online, for additional information on the samples). The murine glioma model consists in the orthotopic allograft of GL261 glioma cells into the brain of C57BL/6 mice as previously described by us (Ferrer-Font et al. 2017).

DNA was extracted by standard phenol–chloroform procedures. Concentration and quality of the DNA were assessed by Qubit fluorometry (Thermo Fisher Scientific) and Fragment Analyzer capillary electrophoresis (Agilent). Because the human cohorts were composed of both male and female individuals, sex chromosomes were not included in the subsequent analyses for any species (see below) in order to avoid sex bias.

Human samples and data from patients included in this study were provided by the following institutions: the Biobank HUB-ICO-IDIBELL (PT17/0015/0024, integrated in the Spanish Biobank Network), the Biobank of Galicia Sur Health Research Institute (IISGS), and the Aragon Health Sciences Institute in the framework of the Biobank of Aragon. All human and mouse samples were processed following standard operation procedures with the appropriate approval of the corresponding Ethical and Scientific Review Boards: Bellvitge University Hospital Ethics Committee (code 07/19, April 11, 2019) for the human aging samples; Biobank IISGS Ethics Committee (code 2019/238, April 28, 2019) and Aragón Clinical Research Ethics Committee (code 14/2019, July 24, 2019) for the human cancer samples; Health Institute Carlos III Ethics Committee (code CBA PA 13_2013, February 26, 2013) for the mouse aging samples; Autonomous University of Barcelona Ethics Commission (code CEEAH-3665, February 5, 2018) for the mouse cancer samples.

Reduced Representation Bisulfite Sequencing

For each sample, 100 ng of genomic DNA were used for library preparation with the Premium Reduced Representation Bisulfite Sequencing Kit (Diagenode). Subsequently, samples were pooled in groups of 6 or 8. PCR clean-up after the final library amplification was performed using Agencourt AMPure XP (Beckman Coulter). The quality of the pools was assessed by Qubit fluorometry (Thermo Fisher Scientific) and Fragment Analyzer capillary electrophoresis (Agilent). Finally, the library pools were sequenced. Human samples were sequenced on a NovaSeq6000 (Illumina) using 50-bp paired-end read sequencing (PE50). Mouse samples were sequenced on a HiSeq3000 (Illumina) using 50-bp single-read sequencing (SR50).

Reduced Representation Bisulfite Sequencing Data Preprocessing

FASTA quality was assessed with FastQC (v0.11.8; https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ last accessed April 29, 2021). Reads were quality- and adapter-trimmed with Trim galore (v0.6.4; https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/ last accessed April 29, 2021) with default parameters under *–rrbs* mode and *–2colour 20* in the case of the human data to account for NovaSeq bias toward identifying high-quality G bases from nonsignal basecalls. Next, the reads were aligned to preindexed and bisulfite converted GRCh38/hg38 and GRCm38/mm10 genomes (primary assemblies) with Bismark (v0.22.2) using default parameters (Bowtie 2 alignment; see supplementary table 1, Supplementary Material online, for additional information on the alignment statistics) (Krueger and Andrews 2011). Finally, methylation counts for cytosines belonging to CpG sites were obtained with Bismark's methylation extractor under default parameters, including *–no_overlap* for paired-end data to avoid scoring overlapping methylation calls twice. For each position, methylation values

were computed as the percentage of methylated cytosines with respect to total cytosines on a scale of 0–1.

Differential Methylation Analyses

Differential methylation analyses were performed on CpG sites, not on individual cytosines, using the following process: 1) the counts for cytosines from both strands belonging to the same CpG site were pooled so as to have one methylation measurement per CpG site; 2) CpG sites mapping to sex chromosomes and alternative haplotypes were filtered out; 3) CpG sites were filtered for low coverage (<10 counts) or very high coverage (>99th percentile); 4) CpG sites common to all human or mouse samples were retained for the analyses. Next, the differential methylation analyses were carried out with the *methylKit* R package (v1.14.2) (Akalin et al. 2012) using the *calculateDiffMeth* function with default parameters and multiple testing adjustment by FDR < 0.05. The calculateDiffMeth function uses logistic regression models to define differentially methylated CpGs between groups. Because of the low number of samples ($n = 3$ per group), a strict threshold of >0.2 change in methylation value was used to define significant changes. No other covariates were included in the models because of the limited information available for other variables. Nonetheless, the inclusion of sex or age variables in the human cancer and aging comparisons led to similar results, with 97–99% of the dmCpGs being the same (data not shown).

Annotation of Sites to Genomic Locations

The profiled CpG sites were annotated to hg38 or mm10 CpG islands and gene locations using the *annotatr* R package (v1.12.1) (Cavalcante and Sartor 2017). Overlapping genetic annotations were collapsed with the following priority: Promoter > 5UTR > 3UTR > Exon > Intron > 1–5 > Intergenic. Finally, CpG sites were mapped to genes and their transcripts through the *TxDb.Hsapiens.UCSC.hg38.knownGene* and *TxDb.Mmusculus.UCSC.mm10.knownGene* R packages (v3.1.0). Annotation for the Illumina Infinium Human Methylation 450 K Beadchip was accessed through the *IlluminaHumanMethylation450kanno.ilmn12.hg19* R package (v0.6.0).

To determine the density in surrounding CpGs associated with each site, each locus was expanded to a 2,000-bp window and the number of CpGs present in the surrounding genomic sequence were counted and divided by the number of those possible.

Gene Set Enrichment Analyses

Gene set enrichment analyses were performed with the *goseq* R package (v1.38.0) (2010 P Genome Biol—Young MD). CpGs were mapped to genes, and enrichment in biological functions was carried out with genes that were exclusively hyper- or hypomethylated, using appropriate filtered backgrounds of assayed genes. The bias arising from there being different number of probes per gene was taken into account in the analyses. The GO, KEGG, Reactome, and CGP gene set databases from the Molecular Signature Database (MSigDB v7.0) (Liberzon et al. 2015) were used for the enrichment analyses and were accessed through the *msigdbr* R package (v7.0.1). In order to retain the most important enrichments, only gene sets with FDR < 0.05, OR >2 and containing between 10 and 1,000 genes were represented as statistically significant.

Region Enrichment Analyses

Genomic enrichment analyses were performed with the *LOLA* R package (v1.16.0) (Sheffield and Bock 2016) using custom databases. Sets of CpGs were tested for overenrichment in specific genomic tracks using one-sided Fisher's exact tests (FDR < 0.05) with appropriate filtered backgrounds of the CpGs being used for each case. The genomic regions analyzed included histone marks, chromatin states, and repetitive DNA elements. The histone and chromatin state *LOLA* databases built for the enrichments are available as extended data sets.

Histone ChIP-seq BED tracks for H3K4me1, H3K4me3, H3K36me3, H3K27ac, H3K27me3, and H3K9me3 in 15 healthy human tissues were obtained from LOLA extended universe (databio.org/regiondb) using NIH Roadmap Epigenomics region data corresponding to the hg38 version of the human genome (Roadmap Epigenomics Consortium et al. 2015). Histone BED tracks for the aforementioned marks in nine postnatal mouse tissues mapped to the mm10 version of the mouse genome were obtained from ENCODE (Gorkin et al. 2020).

Chromatin states were built using chromHMM (see below). Consolidated data from the aforementioned six histone tracks (BED tag.align files) and corresponding ChIP-seq input from 15 human tissues were obtained from NIH Roadmap Epigenomics (https://egg2.wustl.edu/roadmap/data/byFileType/alignments/consolidated; last accessed April 29, 2021). Data corresponding to the aforementioned six histone ChIP-seq tracks (BAM files) and the corresponding ChIP-seq input from nine mouse tissues were obtained from ENCODE (https://www.encodeproject.org/; last accessed April 29, 2021).

Repetitive regions identified by RepeatMasker in the hg38 or mm10 genomes (Smit 1996) were obtained with the UCSC Table Browser (Karolchik 2004) and grouped by family or class. Elements with an uncertain classification, labeled with the "?" symbol, were filtered out.

Chromatin State Model Learning

ChromHMM (v1.18) (Ernst and Kellis 2012) was used to generate an 18-state model to predict chromatin states from 15 human and nine mouse tissues. ChIP-seq data sets corresponding to six histone marks (H3K4me1, H3K4me3, H3K36me3, H3K27ac, H3K27me3, and H3K9me3), along with their respective input controls, were obtained from NIH Roadmap Epigenomics and ENCODE for human and mouse samples respectively (see supplementary table 11, Supplementary Material online, for the description of the files). Human BED files were converted from hg19 to hg38 coordinates via UCSC's liftOver tool (v377) (Hinrichs 2006). Resulting files were binarized using the *binarizeBed* function with default parameters using the hg38 chromosome sizes. A model considering 18 states was learned using the

LearnModel function with default parameters, as previously described elsewhere (Roadmap Epigenomics Consortium et al. 2015). In the case of mouse data sets (mm10), BAM files corresponding to replicated samples from different mouse tissues were merged using SAMtools (v1.7) (Li et al. 2009), as previous studies have suggested that chromatin state inferences either from single time series or from replicated data are invariably consistent and highly reproducible. Read duplicates were removed using the *MarkDuplicates* function from Picard tools (broadinstitute.github.io/picard/). The resulting BAM files were sorted using SAMtools and binarized using the *binarizeBam* function from chromHMM. A model considering 18 states was learned using the *LearnModel* function using the same conditions as the human chromHMM pipeline. For both human and mouse analyses, the resulting chromatin states were reordered and renamed to facilitate later interpretation purposes. In addition, enrichment analyses of the different chromatin states in external genomic data sets (CpG Islands, RefSeq Exons, RefSeq genes, RefSeq transcription start sites, RefSeq transcript end sites, and regions within 2,000 bp of a RefSeq TSS) obtained from the UCSC genome browser (hg38 and mm10 genomes) were performed using the *OverlapEnrichment* function from chromHMM.

Motif Enrichment Analyses

Transcription factor motif enrichment analyses were performed with Hypergeometric Optimization of Motif EnRichment (HOMER) software (v4.11.1) (Heinz et al. 2010) for cancer- and aging dmCpGs in human and mouse, using the hg38 or mm10 genomes as reference. HOMER was run under default settings and with the *-keepOverlappingBg* parameter, using as background the filtered universe of CpG sites analyzed by the RRBS for human and mouse, respectively. Significant enrichments for known TFs were first filtered by FDR < 0.05, and subsequently only the top 50 results ordered by enrichment with respect to the background were retained for each comparison (see supplementary table 8, Supplementary Material online, for full results).

Liftover of Orthologous CpGs

Liftover of mm10 to hg38 coordinates was performed with UCSC's liftOver tool (v377) (Hinrichs 2006) by using chain files from https://hgdownload.soe.ucsc.edu/downloads.html last accessed April 29, 2021. Lifted CpG sites which did not reciprocally lift back from hg38 to the original mm10 coordinates were discarded ($\sim$ 2%).

Conservation Analyses

Base-resolution PhastCons tracks for 100-way (hg38) and 60-way (mm10) alignments (Siepel et al. 2005) were accessed via the *GenomicScores* R package (v1.1.10) (Puigdevall and Castelo 2018).

Gene Expression Analyses

Gene and isoform-level gene expression data in the form of RSEM counts and curated survival data were accessed via the Broad Institute Genomic Data Analysis Center (http://gdac.broadinstitute.org/runs/stddata__2016_01_28/data/; last accessed April 29, 2021) and the UCSC Xena (https://xenabrowser.net/datapages/; last accessed April 29, 2021) for the glioma GBMLGG (glioblastoma—lower grade glioma) cohort from the TCGA (The Cancer Genome Atlas) consortium (Brennan et al. 2013; Ceccarelli et al. 2016) (the list of patients and genes analyzed is available as an extended data set). Prior to analysis, data were normalized by the VST (variance stabilizing transformation) procedure using the *DESeq2* R package (v1.26.0) (Anders and Huber 2010). This transformation stabilizes the variance–mean relationship in the data and corrects for differences in library size between the samples. For the isoform analysis, the four most expressed *AGAP3* isoforms were selected. Cox models were analyzed with the *survival* R package (v3.1.11) (Therneau and Grambsch 2000).

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgements

The authors would like to thank Ronnie Lendrum for manuscript editing and the members of the Cancer Epigenetics Laboratory for their positive feedback. Finally, they want to particularly acknowledge the patients and biobanks involved in the study: the Biobank HUB-ICO-IDIBELL (PT17/0015/0024, integrated in the Spanish Biobank Network), the Biobank of Galicia Sur Health Research Institute (IISGS), and the Aragon Health Sciences Institute in the framework of the Biobank of Aragon. Specifically, they thank Laura Arregui Egido (HUB-ICO-IDIBELL), Delia Recalde Frisón (Biobank of Aragon) and Susana Teijeira Bautista, Vanesa Val Varela, and Olga Souto Rodríguez (IISGS) for their contribution in the collection of the samples. This work was supported by the Spanish Association Against Cancer (PROYE18061FERN to M.F.F.), the Asturias Government (PCTI) cofunding 2018-2022/FEDER (IDI/2018/146 to M.F.F.), Fundación General CSIC (0348_CIE_6_E to M.F.F.), and the Health Institute Carlos III (Plan Nacional de I + D+I) cofunding FEDER (PI15/00892 and PI18/01527 to M.F.F and A.F.F.). They also acknowledge support from the Ramón Areces Foundation (CIVP18A3891 to P.J.F.M.), the AECC (SIRTBIO to P.J.F.M.), the MICINN (SAF2017-85766-R to P.J.F.M.), a Ramón y Cajal fellowship (MICINN, RYC-2017-22335 to P.J.F.M.), and the European Commission ATTRACT project (777222 to A.P.C.). J.R.T. is supported by a Juan de la Cierva fellowship from the Spanish Ministry of Science and Innovation (FJCI-2015-26965). R.F.P. and P.S.O. are supported by the Severo Ochoa program (BP17-114 and BP17-165, respectively). R.G.U. is supported by the Centro de Investigación Biomédica en Red de Enfermedades Raras (Health Institute Carlos III). L.V. was supported by the UAB Predoctoral training programme (PIF predoctoral fellowships). They also acknowledge support from the IMDEA Food Institute and IUOPA-ISPA-FINBA (the IUOPA is supported by the Obra Social Cajastur-Liberbank, Spain).

Author Contributions

R.F.P., J.R.T., A.F.F., and M.F.F. conceived, coordinated, and supervised the study. R.F.P. and J.R.T. designed all aspects of the research and contributed equally to this work. R.F.P. and J.R.T. collected the data, performed computational analyses, and wrote the manuscript. A.F.F., P.S.O., V.L.M., and R.G.U. analyzed and interpreted the data. N.V.S. provided clinical samples and biological data for human samples. A.P.C. and L.V. generated the mouse glioma models. M.B., P.J.F.M., and M.S. generated the mouse aging models. All authors revised, read, and approved the final manuscript.

Data Availability

The data underlying this article are available in the article and in its Supplementary Material online. Additionally, extended data sets are available in Zenodo, at https://dx.doi.org/10.5281/zenodo.4573676; last accessed April 29, 2021, principally consisting of scripts used in the analyses, extended results, the matrices of preprocessed methylation values and the lists of orthologous CpG sites, the databases of histones, repetitive DNA elements and chromatin states, information regarding the RNA-seq data analyzed. The sequencing files from the RRBS experiments have been deposited in the European Nucleotide Archive (ENA) under the accession number PRJEB41460 (mouse data) and in the European Genome-Phenome Archive (EGA) under the accession number EGAS00001004851 (human data). Finally, an R Shiny app is available at https://epilabasturias.shinyapps.io/mbe_app/; last accessed April 29, 2021 which describes and facilitates access to the results of the differential methylation analyses.

References

A lin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, Mason CE. 2012. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol*. 13(10):R87.

Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biol*. 11(10):R106.

Aunan JR, Cho WC, Søreide K. 2017. The biology of aging and cancer: a brief overview of shared and divergent molecular hallmarks. *A&D* 8(5):628.

Baylin SB, Jones PA. 2016. Epigenetic determinants of cancer. *Cold Spring Harb Perspect Biol*. 8(9):a019505.

Beisel C, Paro R. 2011. Silencing chromatin: comparing modes and mechanisms. *Nat Rev Genet*. 12(2):123–135.

Blahnik KR, Dou L, Echipare L, Iyengar S, O'Geen H, Sanchez E, Zhao Y, Marra MA, Hirst M, Costello JF, et al. 2011. Characterization of the contradictory chromatin signatures at the 3′ exons of zinc finger genes. *PLoS One* 6:e17121.

Blewitt M, Whitelaw E. 2013. The use of mouse models to study epigenetics. *Cold Spring Harb Perspect Biol*. 5(11):a017939.

Bock C, Tomazou EM, Brinkman A, Müller F, Simmer F, Gu H, Jäger N, Gnirke A, Stunnenberg HG, Meissner A. 2010. Genome-wide mapping of DNA methylation: a quantitative technology comparison. *Nat Biotechnol*. 28(10):1106–1114.

Brennan CW, Verhaak RGW, McKenna A, Campos B, Noushmehr H, Salama SR, Zheng S, Chakravarty D, Sanborn JZ, Berman SH, et al. 2013. The somatic genomic landscape of glioblastoma. *Cell* 155(2):462–477.

Carrasco-Garcia E, Moreno-Cugnon L, Garcia I, Borras C, Revuelta M, Izeta A, Lopez-Lluch G, de Pancorbo MM, Vergara I, Vina J, et al.. 2 9. SOX2 expression diminishes with ageing in several tissues in mice and humans. Mech Ageing Dev. 177:30–36. 10.1016/j.mad.2018.03.008 29574045

Cavalcante RG, Sartor MA. 2017. annotatr: genomic regions in context. *Bioinformatics* 33(15):2381–2383.

Ceccarelli M, Barthel FP, Malta TM, Sabedot TS, Salama SR, Murray BA, Morozova O, Newton Y, Radenbaugh A, Pagnotta SM, et al. 2016. Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. *Cell* 164(3):550–563.

Chen F, Zhang Q, Deng X, Zhang X, Chen C, Lv D, Li Y, Li D, Zhang Y, Li P, et al. 2018. Conflicts of CpG density and DNA methylation are proximally and distally involved in gene regulation in human and mouse tissues. *Epigenetics* 13(7):721–741.

Cheng Q, Huang C, Cao H, Lin J, Gong X, Li J, Chen Y, Tian Z, Fang Z, Huang J. 2019. A novel prognostic signature of transcription factors for the prediction in patients with GBM. *Front Genet*. 10:906.

Chou S-J, Tole S. 2019. Lhx2, an evolutionarily conserved, multifunctional regulator of forebrain development. *Brain Res*. 1705:1–14.

Cole JJ, Robertson NA, Rather MI, Thomson JP, McBryan T, Sproul D, Wang T, Brock C, Clark W, Ideker T, et al. 2017. Diverse interventions that extend mouse lifespan suppress shared age-associated epigenetic changes at critical gene regulatory regions. *Genome Biol*. 18(1):58.

Conway JR, Lex A, Gehlenborg N. 2017. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* 33(18):2938–2940.

Day C-P, Merlino G, Van Dyke T. 2015. Preclinical mouse cancer models: a maze of opportunities and challenges. *Cell* 163(1):39–53.

Dennis DJ, Han S, Schuurmans C. 2019. bHLH transcription factors in neural development, disease, and reprogramming. *Brain Res*. 1705:48–65.

Dmitrijeva M, Ossowski S, Serrano L, Schaefer MH. 2018. Tissue-specific DNA methylation loss during ageing and carcinogenesis is linked to chromosome structure, replication timing and cell division rates. *Nucleic Acids Res*. 46(14):7022–7039.

Du J, Johnson LM, Jacobsen SE, Patel DJ. 2015. DNA methylation pathways and their crosstalk with histone methylation. *Nat Rev Mol Cell Biol*. 16(9):519–532.

Dutta S, Sengupta P. 2016. Men and mice: relating their ages. *Life Sci*. 152:244–248.

Edwards JR, O'Donnell AH, Rollins RA, Peckham HE, Lee C, Milekic MH, Chanrion B, Fu Y, Su T, Hibshoosh H, et al. 2010. Chromatin and sequence features that define the fine and gross structure of genomic methylation patterns. *Genome Res*. 20(7):972–980.

Elango N, Yi SV. 2008. DNA methylation and structural and functional bimodality of vertebrate promoters. *Mol Biol Evol*. 25(8):1602–1608.

Emerson RO, Thomas JH. 2009. Adaptive evolution in zinc finger transcription factors. *PLoS Genet*. 5(1):e1000325.

Ernst J, Kellis M. 2010. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat Biotechnol*. 28(8):817–825.

Ernst J, Kellis M. 2012. ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods*. 9(3):215–216.

Espada J, Esteller M. 2013. Mouse models in epigenetics: insights in development and disease. *Brief Funct Genomics*. 12(3):279–287.

Fernández AF, Bayón GF, Urdinguio RG, Toraño EG, García MG, Carella A, Petrus-Reurer S, Ferrero C, Martinez-Camblor P, Cubillo I, et al. 2015. H3K4me1 marks DNA regions hypomethylated during aging in human stem and differentiated cells. *Genome Res*. 25(1):27–40.

Ferrer-Font L, Villamañan L, Arias-Ramos N, Vilardell J, Plana M, Ruzzene M, Pinna L, Itarte E, Arús C, Candiota A. 2017. Targeting protein kinase CK2: evaluating CX-4945 potential for GL261 glioblastoma therapy in immunocompetent mice. *Pharmaceuticals* 10(4):24.

Foronda M, Martínez P, Schoeftner S, Gómez-López G, Schneider R, Flores JM, Pisano DG, Blasco MA. 2014. Sox4 links tumor suppression

to accelerated aging in mice by modulating stem cell activation. *Cell Rep*. 8(2):487–500.
Gorkin DU, Barozzi I, Zhao Y, Zhang Y, Huang H, Lee AY, Li B, Chiou J, Wildberg A, Ding B, et al. 2020. An atlas of dynamic chromatin landscapes in mouse fetal development. *Nature* 583(7818):744–751.
Greenberg MVC, Bourc'his D. 2019. The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol*. 20(10):590–607.
Grüning B, Dale R, Sjödin A, Chapman BA, Rowe J, Tomkins-Tinch CH, Valieris R, Köster J. 2018. Bioconda: sustainable and comprehensive software distribution for the life sciences. *Nat Methods*. 15(7):475–476.
Hadad N, Masser DR, Blanco-Berdugo L, Stanford DR, Freeman WM. 2019. Early-life DNA methylation profiles are indicative of age-related transcriptome changes. *Epigenet Chromatin*. 12(1):58.
Hahn O, Grönke S, Stubbs TM, Ficz G, Hendrich O, Krueger F, Andrews S, Zhang Q, Wakelam MJ, Beyer A, et al. 2017. Dietary restriction protects from age-associated DNA methylation and induces epigenetic reprogramming of lipid metabolism. *Genome Biol*. 18(1):56.
Han S, Liu Y, Cai SJ, Qian M, Ding J, Larion M, Gilbert MR, Yang C. 2020. IDH mutation in glioma: molecular mechanisms and potential therapeutic targets. *Br J Cancer*. 122(11):1580–1589.
Hanna CW, Demond H, Kelsey G. 2018. Epigenetic regulation in development: is the mouse a good model for the human? *Hum Reprod Update*. 24(5):556–576.
Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 38(4):576–589.
Hernando-Herraez I, Evano B, Stubbs T, Commere P-H, Jan Bonder M, Clark S, Andrews S, Tajbakhsh S, Reik W. 2019. Ageing affects DNA methylation drift and transcriptional cell-to-cell variability in mouse muscle stem cells. *Nat Commun*. 10(1):4361.
Hernando-Herraez I, Garcia-Perez R, Sharp AJ, Marques-Bonet T. 2015. DNA methylation: insights into human evolution. *PLoS Genet*. 11(12):e1005661.
Hernando-Herraez I, Heyn H, Fernandez-Callejo M, Vidal E, Fernandez-Bellon H, Prado-Martinez J, Sharp AJ, Esteller M, Marques-Bonet T. 2015. The interplay between DNA methylation and sequence divergence in recent human evolution. *Nucleic Acids Res*. 43(17):8204–8214.
Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H, Diekhans M, Furey TS, Harte RA, Hsu F, Hillman-Jackson J. 2006. The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res*. 34(Database issue):D590–D598.
Hirst TC, Vesterinen HM, Sena ES, Egan KJ, Macleod MR, Whittle IR. 2013. Systematic review and meta-analysis of temozolomide in animal models of glioma: was clinical efficacy predicted? *Br J Cancer*. 108(1):64–71.
Huntley S, Baggott DM, Hamilton AT, Tran-Gyamfi M, Yang S, Kim J, Gordon L, Branscomb E, Stubbs L. 2006. A comprehensive catalog of human KRAB-associated zinc finger genes: insights into the evolutionary history of a large family of transcriptional repressors. *Genome Res*. 16(5):669–677.
Karolchik D. 2004. The UCSC Table Browser data retrieval tool. *Nucleic Acids Res*. 32(90001):493D–496.
Kessler NJ, Van Baak TE, Baker MS, Laritsky E, Coarfa C, Waterland RA. 2016. CpG methylation differences between neurons and glia are highly conserved from mouse to human. *Hum Mol Genet*. 25(2):223–232.
Köppen M, Simske JS, Sims PA, Firestein BL, Hall DH, Radice AD, Rongo C, Hardin JD. 2001. Cooperative regulation of AJM-1 controls junctional integrity in *Caenorhabditis elegans* epithelia. *Nat Cell Biol*. 3(11):983–991.
Krueger F, Andrews SR. 2011. Bismark: a flexible aligner and methylation caller for bisulfite-Seq applications. *Bioinformatics* 27(11):1571–1572.
Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25(16):2078–2079.
Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. 2015. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 1(6):417–425.
Lister R, Mukamel EA, Nery JR, Urich M, Puddifoot CA, Johnson ND, Lucero J, Huang Y, Dwork AJ, Schultz MD, et al. 2013. Global epigenomic reconfiguration during mammalian brain development. *Science* 341(6146):1237905–1237905.
Liu XS, Wu H, Ji X, Stelzer Y, Wu X, Czauderna S, Shu J, Dadon D, Young RA, Jaenisch R. 2016. Editing DNA methylation in the mammalian genome. *Cell* 167(1):233–247.e17.
Long HK, Sims D, Heger A, Blackledge NP, Kutter C, Wright ML, Grützner F, Odom DT, Patient R, Ponting CP, et al. 2013. Epigenetic conservation at gene regulatory elements revealed by non-methylated DNA profiling in seven vertebrates. *eLife* 2:e00348.
Maegawa S, Gough SM, Watanabe-Okochi N, Lu Y, Zhang N, Castoro RJ, Estecio MRH, Jelinek J, Liang S, Kitamura T, et al. 2014. Age-related epigenetic drift in the pathogenesis of MDS and AML. *Genome Res*. 24(4):580–591.
Maegawa S, Lu Y, Tahara T, Lee JT, Madzo J, Liang S, Jelinek J, Colman RJ, Issa J-PJ. 2017. Caloric restriction delays age-related methylation drift. *Nat Commun*. 8(1):539.
Masser DR, Hadad N, Porter HL, Mangold CA, Unnikrishnan A, Ford MM, Giles CB, Georgescu C, Dozmorov MG, Wren JD, et al. 2017. Sexually divergent DNA methylation patterns with hippocampal aging. *Aging Cell* 16(6):1342–1352.
McKinney BC, Lin C-W, Rahman T, Oh H, Lewis DA, Tseng G, Sibille E. 2019. DNA methylation in the human frontal cortex reveals a putative mechanism for age-by-disease interactions. *Transl Psychiatry*. 9(1):39.
Mendizabal I, Shi L, Keller TE, Konopka G, Preuss TM, Hsieh T-F, Hu E, Zhang Z, Su B, Yi SV. 2016. Comparative methylome analyses identify epigenetic regulatory loci of human brain evolution. *Mol Biol Evol*. 33(11):2947–2959.
Michalak EM, Burr ML, Bannister AJ, Dawson MA. 2019. The roles of DNA, RNA and histone methylation in ageing and cancer. *Nat Rev Mol Cell Biol*. 20(10):573–589.
Nordman JC, Kabbani N. 2012. An interaction between α7 nicotinic receptors and a G-protein pathway complex regulates neurite growth in neural cells. *J Cell Sci*. 125(22):5502–5513.
Oku Y, Huganir RL. 2013. AGAP3 and Arf6 regulate trafficking of AMPA receptors and synaptic plasticity. *J Neurosci*. 33(31):12586–12598.
Pérez RF, Tejedor JR, Bayón GF, Fernández AF, Fraga MF. 2018. Distinct chromatin signatures of DNA hypomethylation in aging and cancer. *Aging Cell* 17(3):e12744.
Puigdevall P, Castelo R. 2018. GenomicScores: seamless access to genomewide position-specific scores from R and Bioconductor. *Bioinformatics* 34(18):3208–3210.
Regnard C, Fesquet D, Janke C, Boucher D, Desbruyéres E, Koulakoff A, Insina C, Travo P, Eddé B. 2003. Characterisation of PGs1, a subunit of a protein complex co-purifying with tubulin polyglutamylase. *J Cell Sci*. 116(Pt 20):4181–4190.
Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, Ziller MJ, et al. 2015. Integrative analysis of 111 reference human epigenomes. *Nature* 518(7539):317–330.
Ross JP, Rand KN, Molloy PL. 2010. Hypomethylation of repeated DNA sequences in cancer. *Epigenomics* 2(2):245–269.
Schroeder DI, Jayashankar K, Douglas KC, Thirkill TL, York D, Dickinson PJ, Williams LE, Samollow PB, Ross PJ, Bannasch DL, et al. 2015. Early developmental and evolutionary origins of gene body DNA methylation patterns in mammalian placentas. *PLoS Genet*. 11:e1005442.
Sen P, Shah PP, Nativio R, Berger SL. 2016. Epigenetic mechanisms of longevity and aging. *Cell* 166(4):822–839.
Sheffield NC, Bock C. 2016. LOLA: enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor. *Bioinformatics* 32(4):587–589.

Shimizu D, Masuda T, Sato K, Tsuruda Y, Otsu H, Kuroda Y, Eguchi H, Kodera Y, Mimori K. 2019. CRMP5-associated GTPase (CRAG) is a candidate driver gene for colorectal cancer carcinogenesis. *Anticancer Res*. 39(1):99–106.

Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth J, Hillier LW, Richards S, et al. 2005. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res*. 15(8):1034–1050.

Sizemore GM, Pitarresi JR, Balakrishnan S, Ostrowski MC. 2017. The ETS family of oncogenic transcription factors in solid tumours. *Nat Rev Cancer*. 17(6):337–351.

Smit AF. 1996. The origin of interspersed repeats in the human genome. *Curr Opin Genet Dev*. 6(6):743–748.

Song H, Liu J, Wu X, Zhou Y, Chen X, Chen J, Deng K, Mao C, Huang S, Liu Z. 2019. LHX2 promotes malignancy and inhibits autophagy via mTOR in osteosarcoma and is negatively regulated by miR-129-5p. *Aging* 11(21):9794–9810.

Srivastava R, Srivastava R, Ahn SH. 2016. The epigenetic pathways to ribosomal DNA silencing. *Microbiol Mol Biol Rev*. 80(3):545–563.

Sun D, Luo M, Jeong M, Rodriguez B, Xia Z, Hannah R, Wang H, Le T, Faull KF, Chen R, et al. 2014. Epigenomic profiling of young and aged HSCs reveals concerted changes during aging that reinforce self-renewal. *Cell Stem Cell* 14(5):673–688.

Therneau TM, Grambsch PM. 2000. Modeling survival data: extending the Cox model. New York: Springer-Verlag.

Wang M, Zhao Y, Zhang B. 2015. Efficient test and visualization of multi-set intersections. *Sci Rep*. 5(1):16923.

Wang T, Tsui B, Kreisberg JF, Robertson NA, Gross AM, Yu MK, Carter H, Brown-Borg HM, Adams PD, Ideker T. 2017. Epigenetic aging signatures in mice livers are slowed by dwarfism, calorie restriction and rapamycin treatment. *Genome Biol*. 18(1):57.

Wickham H. 2016. ggplot2: elegant graphics for data analysis. Springer International Publishing. ISBN 978-0-387-98140-6

Yin T, Cook D, Lawrence M. 2012. ggbio: an R package for extending the grammar of graphics for genomic data. *Genome Biol*. 13(8):R77.

Yu M, Hazelton WD, Luebeck GE, Grady WM. 2020. Epigenetic aging: more than just a clock when it comes to cancer. *Cancer Res*. 80(3):367–374.

Zhou J, Sears RL, Xing X, Zhang B, Li D, Rockweiler NB, Jang HS, Choudhary MNK, Lee HJ, Lowdon RF, et al. 2017. Tissue-specific DNA methylation is conserved across human, mouse, and rat, and driven by primary sequence conservation. *BMC Genomics* 18(1):724.