DNA methylation patterns of replicative senescence are strand-specific and reflect changes in chromatin conformation

Julia Franzen[1], Theodoros Georgomanolis[2], Anton Selich[3], Reinhard Stöger[4], Lilija Brant[2], Eduardo Fernandez-Rebollo[1], Clara Grezella[1], Alina Ostrowska[1], Matthias Begemann[5], Björn Rath[6], Anthony D. Ho[7], Michael Rothe[3], Axel Schambach[3], Argyris Papantonis[2], Wolfgang Wagner[1]

[1] Helmholtz-Institute for Biomedical Engineering, RWTH Aachen Medical School, 52074 Aachen, Germany

[2] Center for Molecular Medicine, University of Cologne, 50931 Cologne, Germany

[3] Institute of Experimental Hematology, Hannover Medical School, 30625 Hannover, Germany

[4] School of Biosciences, University of Nottingham, Sutton Bonington Campus, Loughborough, Leicestershire, LE12 5RD, United Kingdom

[5] Institute of Human Genetics, Medical Faculty, RWTH Aachen University, 52074, Aachen, Germany

[6] Department for Orthopedics, RWTH Aachen University Medical School, 52074 Aachen, Germany

[7] Internal Medicine Department of Hematology, Oncology and Rheumatology, Heidelberg University Medical Center, 69120 Heidelberg, Germany

Corresponding author: Wolfgang Wagner, M.D., Ph.D., Helmholtz-Institute for Biomedical Engineering, Stem Cell Biology and Cellular Engineering, RWTH Aachen University Medical School, Pauwelsstraße 20, 52074 Aachen, Germany, Phone: +49-241-8088611, Fax: +49-241-8082203, E-mail: wwagner@ukaachen.de

Abstract

Replicative senescence of cells in culture is associated with highly reproducible DNA methylation (DNAm) changes at specific sites in the genome. Thus far, it is largely unclear if these epigenetic modifications are directly regulated, or if they are randomly evoked by other chromatin changes during long-term culture.

We have identified CG dinucleotides (CpGs) that become continuously hyper- or hypo-methylated in the course of culture expansion of mesenchymal stem cells (MSCs) and other cell types. These modifications provide a biomarker for replicative senescence and correlate with the number of passages *in vitro*. During reprogramming into induced pluripotent stem cells (iPSCs) senescence-associated DNAm is reversed simultaneously with pluripotency-associated DNAm changes. Bisulfite barcoded amplicon sequencing (BBA-seq) demonstrated that upon passaging the DNAm patterns of neighboring CpGs become more complex without evidence of continuous pattern development. Notably, BBA-seq of hairpin-linked DNA molecules demonstrated that many CpG dyads are methylated only on the forward or the reverse strand. This hemimethylation was conserved over many passages, indicating that it was not due to insufficient maintenance of DNAm patterns. Circularized chromatin conformation capture (4C) of senescence-associated CpGs revealed reproducible changes during senescence without evidence for preferential interaction between CpGs that become either hyper- or hypomethylated.

Taken together, senescence-associated DNAm fluctuates stochastically at specific sites in the genome. Strand-specific DNAm and reproducible changes in 4C indicate that epigenetic modifications of these CG dyads are not regulated in a targeted manner but rather caused by passage-specific higher order chromatin conformation states.

Introduction

Replicative senescence is a cell intrinsic mechanism that limits the number of cell divisions (Campisi 1997). It is reflected by highly reproducible DNA methylation (DNAm) changes at specific sites in the genome (Bork et al. 2010; Cruickshanks et al. 2013). We have previously demonstrated that DNAm levels at only six cytosine/guanine dinucleotides (CpG sites) can be used to estimate passage numbers and cumulative population doublings (Koch et al. 2012; Franzen et al. 2017). It therefore provides a reliable biomarker for replicative senescence, similar as telomere attrition (Bernadotte et al. 2016) and staining of senescence-associated beta galactosidase (Itahana et al. 2007). However, it is so far unclear how senescence-associated DNAm patterns evolve and why they occur at specific genomic regions.

DNAm patterns on both the forward and the reverse DNA strand (Watson and Crick strands) are generally thought to be identical (Ushijima et al. 2003). DNAm maintenance is particularly mediated by DNA methyltransferase 1 (DNMT1) but also by the *de novo* methyltransferases DNMT3A and DNMT3B (Jones and Liang 2009). Using hairpin-bisulfite PCR (Laird et al. 2004) several studies have demonstrated that, despite the general preference for concordant DNA methylation on both strands, some particular sites are specifically methylated on only one strand (Arand et al. 2015; Choi et al. 2017; Patino-Parrado et al. 2017). More recently, it was shown that such hemimethylation is stably inherited over several passages and that hemimethylated sites are associated with CCCTC-binding factor (CTCF)/cohesin binding sites. Thus, hemimethylation does not only occur transiently after cell division but resembles a stable epigenetic modification related to higher order chromatin structure (Sharif and Koseki 2018; Xu and Corces 2018). Senescence is also associated with changes in chromatin conformation (Chandra et al.

2015; Zirkel et al. 2018). It is therefore conceivable that senescence-associated DNAm changes differ on the complementary DNA strands. In this study, we aimed at better understanding how senescence-associated DNAm patterns evolve during MSC culture expansion.

Results

Tracking replicative senescence by DNA methylation

To identify CpGs that become continuously hyper- or hypomethylated during culture expansion of different cell types, we utilized available DNAm profiles based on the Illumina 450k BeadChip. We compiled 63 DNAm profiles of human primary MSCs (n = 45), fibroblasts (n = 5), and human umbilical vein endothelial cells (n = 13) with precise information on passage numbers (Supplemental Table S1). 15 hypermethylated and 15 hypomethylated sites were filtered by Pearson correlation with passage numbers (R^2 > 0.8) and a linear regression slope (m > 0.02; Supplemental Table S2). For site-specific DNAm analysis we further refined two hyper- and two hypomethylated CpGs that cooperated best for prediction of passage numbers from the Illumina 450k BeadChip datasets. These four CpGs were related to the genes *Arachidonate 12-Lipoxygenase* (*ALOX12,* cg03762994), *Docking Protein 6* (*DOK6,* cg25968937), *Leukotriene C4 Synthase* (*LTC4S,* cg26683398) and *TNNI3 Interacting Kinase* (*TNNI3K,* cg05264232; Fig 1A). Senescence-associated DNAm was then tested by pyrosequencing in MSCs, fibroblasts, and HUVECs at various passages (n = 44). Based on pyrosequencing results, we generated a multivariable model for epigenetic predictions of passage number (R^2=0.81; Fig. 1B and Supplemental Fig. S1A). Ten times 10-fold cross-validation of the training dataset resulted in a R^2=0.84 and a root mean squared error (RMSE) of 3.9 passages. Subsequently, our epigenetic senescence clock was validated on an additional

independent set of samples (n = 42; R^2=0.81; Fig. 1C and Supplemental Fig. S1B). Thus, DNAm analysis at these four CpGs facilitates relative precise estimation of passage numbers, as previously described for six other senescence-associated CpGs (Koch et al. 2012).

Resetting the epigenetic senescence clock is an early event during cell reprogramming into induced pluripotent stem cells

We previously described that senescence-associated DNAm changes are reversed in fully reprogrammed induced pluripotent stem cells (iPSCs) (Koch et al. 2013; Frobel et al. 2014), but the kinetics of this epigenetic rejuvenation were not yet addressed. We therefore utilized DNAm profiles of TRA-1-60 positive cells at various time points after retroviral reprogramming of fibroblasts with *OCT3/4*, *SOX2*, *KLF4* and *c-MYC* (Ohnuki et al. 2014). Particularly CpGs that are hypomethylated at later passages become re-methylated between day 15 and day 20 after reprogramming (Fig. 1D). Notably, these epigenetic changes occur in parallel to the epigenetic modifications at pluripotency-associated genes (Fig. 1E) (Nazor et al. 2012) or age-related CpGs (Supplemental Fig. S2) (Weidner et al. 2014). These results indicate that epigenetic rejuvenation of senescence-associated CpGs might be directly linked to the epigenetic conversion during reprogramming into the pluripotent state (Li et al. 2017).

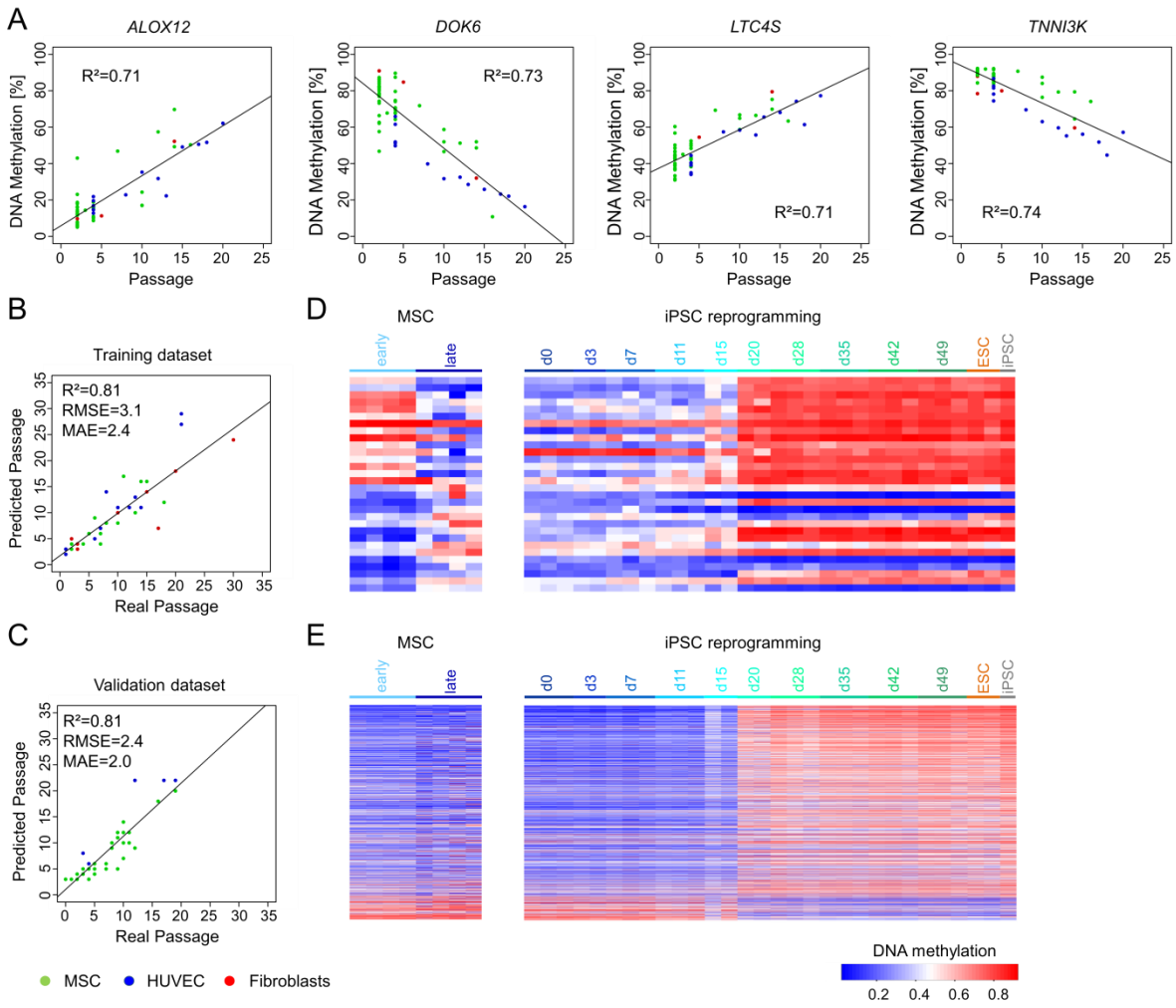

Figure 1: DNA methylation changes during long-term culture and reprogramming into iPSCs.

A) Four senescence-associated CpGs were selected in DNAm profiles of MSCs, fibroblasts, and HUVECs (all 450k BeadChip data). **B)** DNAm at these four CpGs was then analyzed in a training set (n = 44) by pyrosequencing and the results were used to train a multivariable model to estimate passage numbers (RMSE = root mean squared error; MAE = mean average error). **C)** This predictor was validated on pyrosequencing results of an independent set of samples (n = 42). **D)** DNAm changes were analyzed in 30 senescence associated CpG sites using DNAm profiles of MSCs of early and late passage (left; Koch et al. (2013)), and in a datasets that analyzed DNAm changes at various timepoints during reprogramming (right, Ohnuki et al. (2014)). **E)** Reprogramming kinetics of 1,432 pluripotency associated CpG sites (Nazor et al. 2012) in datasets as depicted in D.

Dynamics of DNA methylation patterns at neighboring CpG sites

We previously demonstrated that individual subclones derived from MSC preparations exhibit distinct DNAm patterns at neighboring CpGs (Franzen et al. 2017). Hence, we anticipated that tracking of such patterns over several passages might provide insights into the changing composition of subclones within MSC preparations. To address this question, we used samples from a previously published study (Selich et al. 2016): Umbilical cord derived MSCs from two donors were transduced with lentiviral vectors containing random barcodes and three different fluorescent proteins. Flow cytometry and deep sequencing demonstrated that the diversity of cellular subsets declines and senescent passages became oligoclonal (Fig. 2A and Supplemental Fig. S3A). Genomic DNA was subsequently analyzed via barcoded bisulfite amplicon sequencing (BBA-Seq) to investigate DNAm patterns at the four senescence-associated CpGs mentioned above, as well as the six CpGs of our previous predictor for replicative senescence (associated with the genes *CASR*, *CASP14*, *GRM7*, *KRTAP13.3*, *PRAMEF2* and *SELP*) (Koch et al. 2012). Overall, BBA-Seq facilitated robust measurements of DNAm levels and combined prediction results correlated with the number of passages of umbilical cord derived MSCs (R^2 = 0.87, Fig. 2B). In contrast to pyrosequencing, BBA-Seq facilitates analysis of the sequence of methylated and non-methylated CpGs within individual reads. To this end, we focused on those amplicons that comprised several neighboring CpGs on the BBA-Seq reads (*GRM7*, *CASR*, *LTC4S*, *DOK6* and *ALOX12)*. The DNAm patterns fluctuated over subsequent passages and there was no evidence for continuous development of senescence-associated modifications at these genomic regions. In fact, the modifications seemed to be acquired independently of clonal composition (Fig. 2C). Furthermore, DNAm patterns became more diverse during culture expansion (Fig. 2D and Supplemental Fig S3B), providing additional evidence that development of DNAm patterns in senescence-

associated genomic regions is not related to the oligoclonal composition of MSCs at later passages.

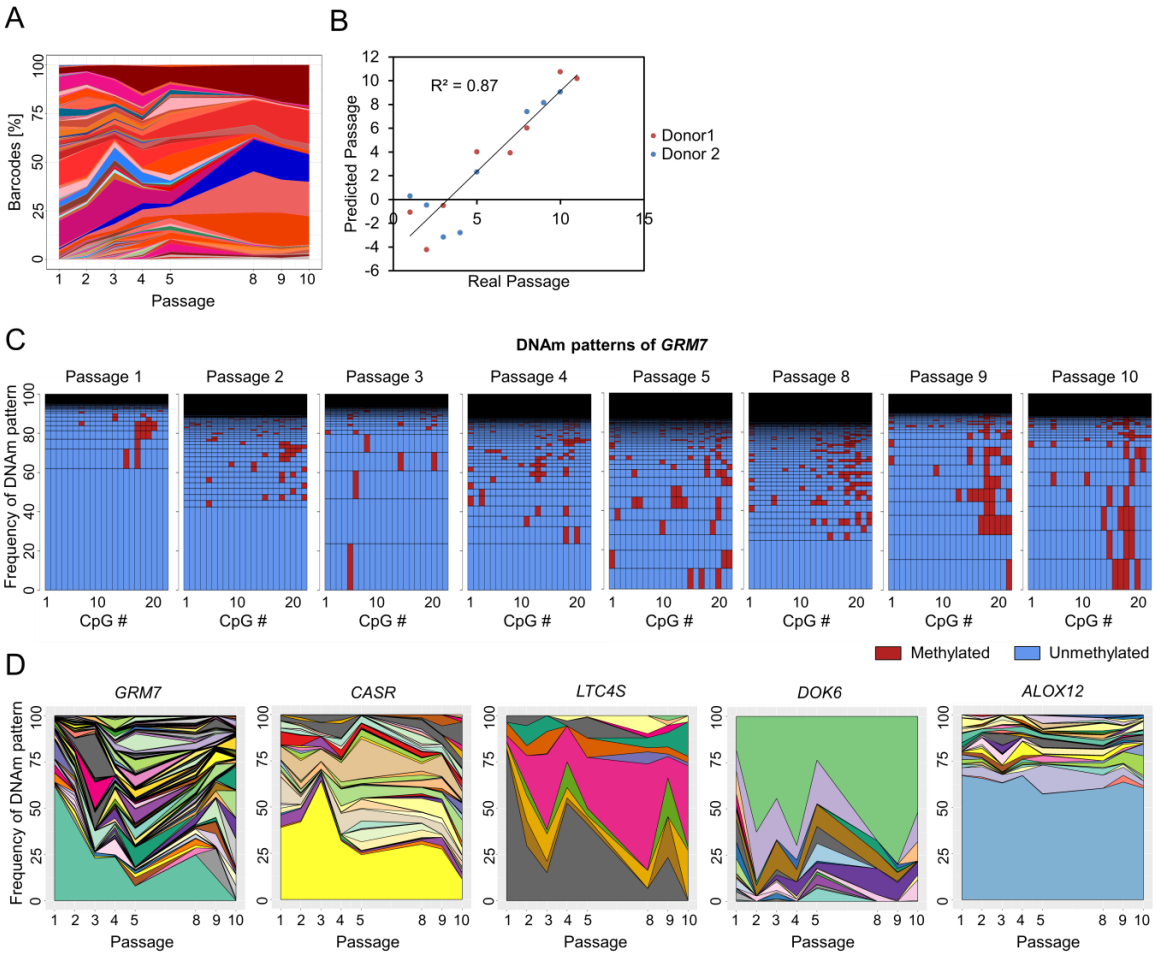

Figure 2: DNA methylation patterns of neighboring CpG sites.

A) Mesenchymal stem cells from umbilical cord were labelled with barcoded-RGB-vectors to correlate DNAm patterns with the composition of cellular subsets (Selich et al. 2016). Deep sequencing analysis of the random barcodes demonstrates that the MSCs at passage 10 are oligoclonal. **B)** These samples were subsequently used for DNAm analysis by barcoded bisulfite amplicon sequencing (BBA-Seq) at ten senescence-associated CpGs. The predicted passage numbers based on these DNAm levels correlated with real passage numbers. **C)** Individual DNAm patterns are exemplarily depicted for 22 neighboring CpGs of *GRM7*. The frequency of DNAm patterns resembles the percentage of corresponding reads in deep sequencing analysis. **D)** Changes in the frequency of different DNAm patterns over the passages are depicted for the neighboring CpGs within the amplicons of *GRM7* (22 CpGs), *CASR* (7 CpGs), *LTC4S* (4 CpGs), *DOK6* (7 CpGs), and *ALOX12* (9 CpGs).

Senescence-associated DNA methylation is often strand-specific

So far, it was unknown if senescence-associated DNAm patterns are identical and stable on the two complementary strands of the DNA. To address this question, we ligated hairpin oligonucleotides to connect the forward and reverse strands of individual DNA molecules (Fig. 3A) (Laird et al. 2004). These hairpins also comprised a unique molecular identifier (UMI) in the loop region to adjust for potential PCR bias (Supplemental Fig. S4A). Eight out of the ten senescence-associated regions encompassed suitable endonuclease restriction sites for targeted hairpin-ligation and could be further analyzed by BBA-Seq with primers specific for these hairpins (*CASR*, *GRM7*, *KRTAP13.3*, *PRAMEF2, SELP*, *DOK6*, *LTC4S*, *TNNI3K*). As a control we considered an additional genomic region that was generally methylated (associated with the genes *C12orf12*) and an imprinting control region in the gene guanine nucleotide binding protein alpha subunit (*GNAS*). Overall, this hairpin BBA-Seq approach revealed similar senescence-associated DNAm levels as conventional BBA-Seq. The accuracy of epigenetic predictions of passage numbers was also similar when reads with the same UMI were only considered once (Supplemental Fig. S4B), indicating that potential PCR bias during amplification does not have major impact on the mean DNAm levels.

Subsequently we compared the DNAm patterns of the two complementary DNA strands. To our surprise most senescence-associated regions revealed preferential hemimethylation at individual CpGs either on the Watson or on the Crick strand. For example, the CpG site next to the hairpin insert of *DOK6* was consistently more methylated on the Crick strand across three different donors in early and late passages (Fig. 3B). We then focused specifically on CpGs that were senescence-associated in the initial Illumina 450k BeadChip analysis. Hemimethylation was consistently observed for the CpGs in *SELP* and *PRAMEF2*, and in early passages for *KRTAP13.3* (Fig. 3C). In addition,

hemimethylation was observed at neighboring CpG sites of *GRM7* and *LTC4S* in late passages (Supplemental Fig. S4C). Notably, the imprinting region of *GNAS* also exhibited hemimethylation, which is in line with recent studies on imprinting regions (Guntrum et al. 2017; Patino-Parrado et al. 2017). The stable hemimethylation at individual CpGs of senescence-associated regions and on specific DNA strands argues against the possibility that this effect is caused by loss of DNAm during cell division – it might rather be evoked by altered states of higher order chromatin conformation (Xu and Corces 2018).

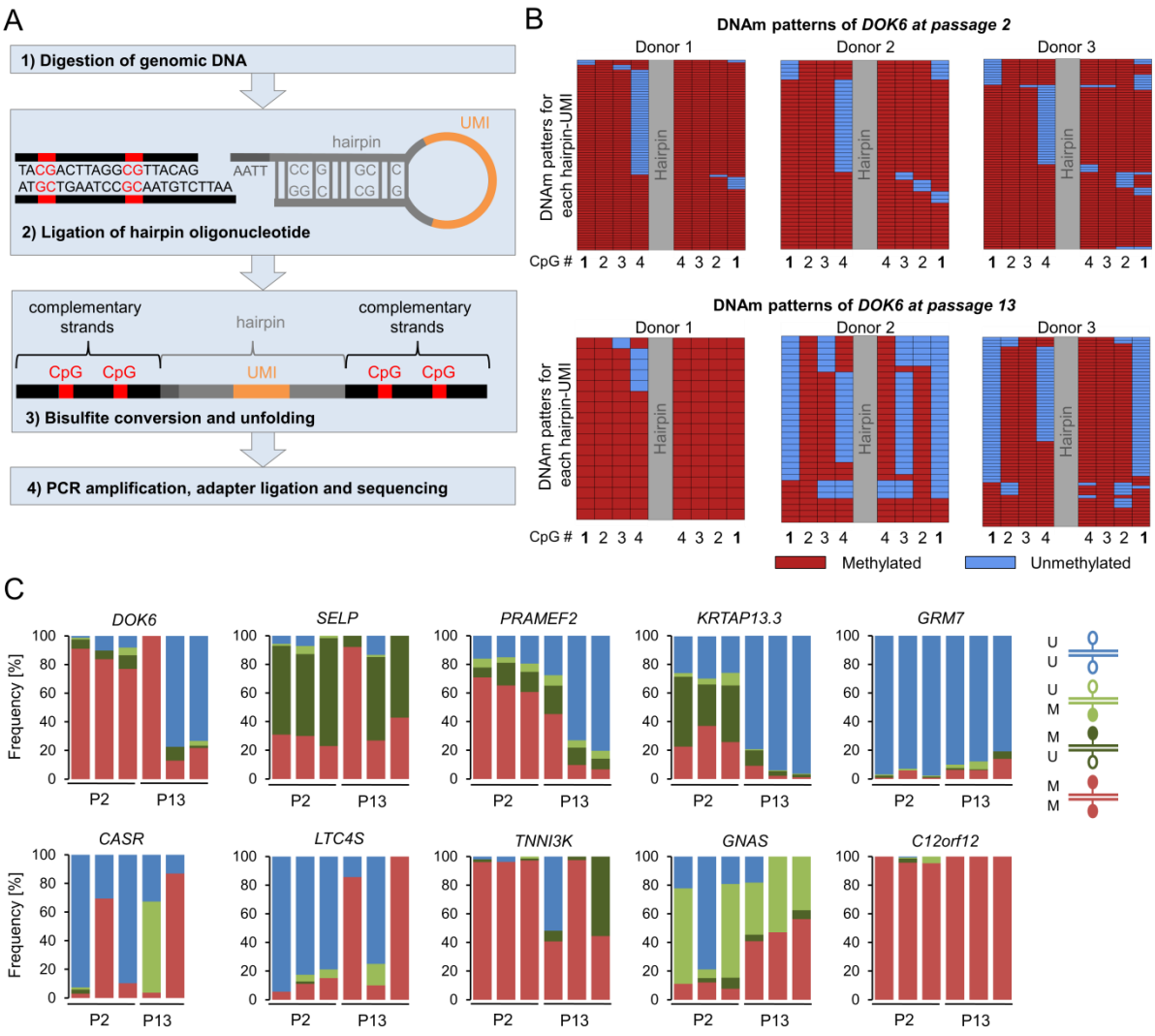

Figure 3: Senescence-associated DNAm is frequently hemimethylated.

A) Schematic presentation of DNAm analysis on complementary DNA strands using hairpin ligation and BBA-Seq (UMI = unique molecular identifier). **B)** DNAm patterns of four neighboring CpGs in *DOK6* are depicted on complementary strands (MSC of three different donors in early and late passages). The CpG site included in the Epigenetic-Senescence-Signature is CpG #1, depicted in bold. Each row represents patterns for an UMI in the hairpin. **C)** Frequency of homo- and hemimethylation at senescence-associated CpG sites that were used for predictions of passage numbers, in an imprinted region (*GNAS*), and at a highly methylated region (*C12orf12*).

Chromatin interactions at senescence-associated genomic regions

Replicative senescence was previously shown to evoke reproducible changes in chromatin organization (Chandra et al. 2015; Zirkel et al. 2018). We explored chromatin interactions and focused specifically on four genomic regions with senescence-associated CpGs (*ALOX12*, *LTC4S*, *CASR* and *KRTAP13.3*) using circular chromatin conformation capture (4C). Two independent MSC preparations at early (P2 and P3) and late (P7 and P9) passages revealed overall reproducible interaction profiles (Fig. 4A). For downstream analyses we only considered highly interacting regions that were categorized as “nearbait” (10 MB around the “bait” locus of interest) or as “cis” (all cis-contacts on the same chromosome). Although trans-chromosomal analysis showed a high number of reproducible interactions (Fig. 4B) we excluded trans-interacting sites from further analysis due to their high background signal, as commonly observed in such studies. The number of interacting sites in nearbait and cis remained similar between early and late passages (Fig. 4C and D). Many of these showed reproducible and significant differences from early to late passages (p value < 0.05), indicating that chromatin interactions at senescence-associated CpGs are modified in conjunction with culture expansion (Fig. 4E and F).

To gain insight into how chromatin interactions are reflected on DNAm we used datasets of MSCs at early passage (n = 5; P2 - P3) and pre-senescent passage (n = 5; P7 – P13; GSE37067) (Koch et al. 2013). The mean DNAm levels at interacting regions was overall similar to the average DNAm of the corresponding chromosome. However, nearbait interactions of *KRTAP13.3*, which becomes hypomethylated upon long-term culture, preferably formed with non-methylated regions both in early and late passages (Fig. 4G). Notably, none of the four senescence-associated genomic regions revealed significant enrichment of interaction with other senescence-associated CpGs (Fig. 4H). Furthermore,

gains and losses in chromatin interactions were not enriched in regions with senescence-associated DNAm changes (Fig. 4I and J).

Hemimethylation was shown to be enriched in CTCF binding sites (Xu and Corces 2018). In fact, nearbait and cis-interacting regions of two senescence-associated sites (*ALOX12* and *LTC4S*) exhibited significant enrichment of binding motifs for CTCF and CCCTC-binding factor like (CTCFL; Supplemental Fig. S5A and B). Similar results were observed using CTCF ChipSeq data of MSCs (Dixon et al. 2015) (Fig. 4K). We and others previously demonstrated that hypomethylation during replicative senescence occurs preferentially at lamina-associated domains (LADs) (Cruickshanks et al. 2013; Shah et al. 2013; Hänzelmann et al. 2015) This is supported by the observation that the hypomethylated region of *KRTAP13.3* was enriched in interactions with LADs, whereas such interactions were depleted for hypermethylated regions (*ALOX12*, *CASR* and *LTC4S*; Fig. 4L). Taken together, senescence-associated DNAm occurs at genomic regions that reveal significant and reproducible changes in chromatin conformation, but it is not synchronously regulated at the interaction sites of different chromatin loops.

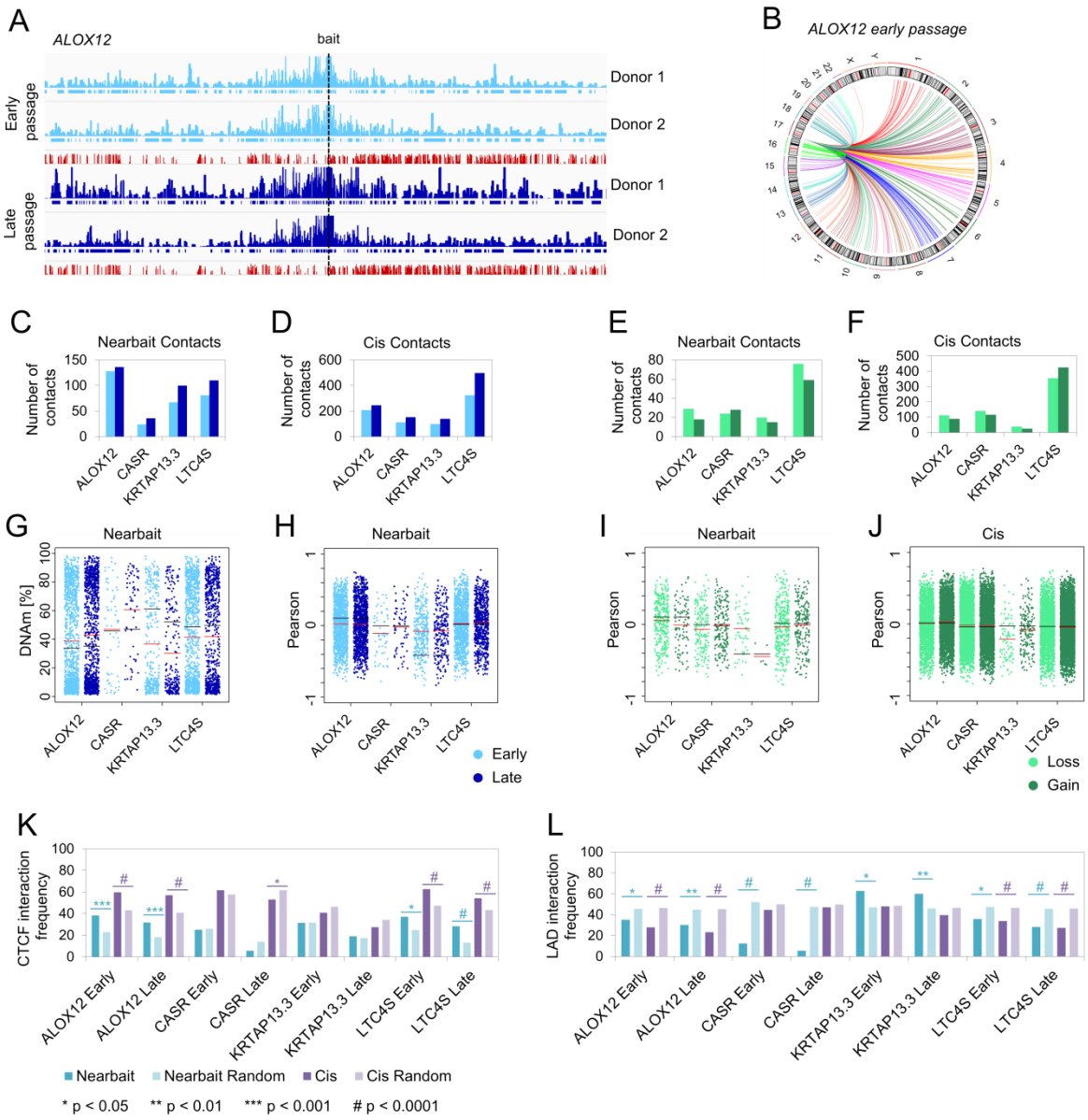

Figure 4: Circular chromatin conformation capture (4C) of senescence-associated CpGs.

A) Integrative Genomics Viewer (IGV) overview of the bait region of *ALOX12* in MSCs of two donors at early (light blue) and late (dark blue) passage. Sequencing peaks are presented as normalized counts. High interacting regions called by 4Cker tool are indicated by horizontal bars beneath the peaks. Mean methylation levels of CpGs on the Illumina 450k BeadChip are depicted for MSCs at early and late passages (GSE37067; red) (Koch et al. 2013). **B)** Circos-plot of highly interacting regions across different chromosomes (trans) is exemplarily depicted for *ALOX12* (interactions are reproducible in two MSC preparations at early passage). **C)**

Numbers of highly interacting regions within the nearbait regions of the four senescence-associated CpGs were similar at early and late passages. **D)** Numbers of cis-interacting regions (within the same chromosome). **E**, **F)** Number of differential contacts (p-value < 0.05) between early and late passages of all four CpG sites in nearbait and cis locations. **G)** DNAm levels of all CpG sites within high interacting regions in early and late passages. Red lines indicate mean DNAm levels of all CpG sites within interacting regions, black lines depict mean DNAm of all CpG sites on the 450k BeadChip within the nearbait region. **H)** The association of DNAm with passage numbers (Pearson correlation) is presented for all CpG sites within high interacting regions. Mean correlations are depicted for CpG sites at highly interacting regions (red) and all CpGs within the nearbait regions (black). **I, J)** Association of DNAm with passage numbers (Pearson correlation) is presented for all CpG sites at differential contact regions within nearbait and cis (red and black lines in analogy to H). **K,** Enrichment of CTCF ChipSeq peaks (Dixon et al. 2015) and **L)** lamina-associated domains (LADs) within high interacting sites of the four baits compared to random background regions. Significance was tested by Fisher's exact test (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; # $p < 0.0001$).

Discussion

Senescence is not just a process observed in cell culture. There is growing evidence that this phenomenon is also relevant in organismal aging and tumorigenesis (Childs et al. 2015). Depletion of senescent cells might prolong healthy aging, and this raises high hopes for development of senolytic drugs (Jeon et al. 2017; Kirkland et al. 2017; Grezella et al. 2018). In fact, replicative senescence – in analogy to aging – is associated with highly reproducible DNAm changes at specific sites in the genome (Cruickshanks et al. 2013; Lowe et al. 2015). Yet, epigenetic modifications upon aging and upon senescence are distinct (Lowe et al. 2016) and therefore epigenetic senescence predictions are applicable to cells from young and elderly donors.

The highly reproducible nature of senescence-associated DNAm suggests that this process is tightly controlled (Xie et al. 2018). Nevertheless, our results indicate that it is possibly not directly mediated by regulatory protein complexes for the following reasons: 1) resetting of senescence-associated DNAm during reprogramming into iPSCs occurs synchronously with the epigenetic changes in pluripotency genes, which is directly linked to the epigenetic transition itself; 2) BBA-seq analysis demonstrated that senescence-associated DNAm patterns do not develop in an additive manner at neighboring CpGs, as would be expected if it was governed by a protein complex (Lövkvist et al. 2016); and 3) senescence-associated DNAm occurs in regions with highly reproducible changes in chromatin interactions. However, there was no evidence that senescence-associated DNAm is co-regulated at the interaction sites of different chromatin loops, as anticipated for regulation via epigenetic modifying protein complexes.

It has been suggested that at least senescence-associated hypomethylation can be attributed to insufficient maintenance function of DNMT1 (Nadejda et al. 2002; Cruickshanks et al. 2013). In fact, we often observed methylation at only one of the complementary DNA strands – but this hemimethylation was consistently observed on specific strands and during culture expansion. Thus, hemimethylation at senescence-associated regions is not due to insufficient maintenance of DNA methylation. Our results provide additional evidence that hemimethylation is a stable epigenetic modification state, as recently proposed (Sharif and Koseki 2018; Xu and Corces 2018). It is still unclear how strand-specific hemimethylation is established and maintained. We speculate, that it is not directly regulated by specific protein complexes. Hemimethylation might be randomly acquired due to some favorable chromatin conformation, such as bending of DNA or interaction with histone complexes – and then stabilize the later *vice versa*. Alternatively, interaction with DNA binding proteins might restrict access to methyltransferases on

individual DNA strands. This hypothesis might be supported by enrichment of CTCF and CTCFL binding motifs, which mediate chromatin looping (Xu and Corces 2018).

Replicative senescence is evidently associated with major changes in chromatin structure. The nucleus becomes much larger, while chromatin volume decreases due to extensive reorganization of hetero- and euchromatin conformation (Chandra et al. 2015; Criscione et al. 2016). Nuclear depletion of HMGB2 and its induction of CTCF clustering are early events on the path to senescence, which disturb the chromosomal 3D organization (Zirkel et al. 2018). Our 4C analysis supports the notion that the 3D chromatin structure undergoes highly reproducible changes during culture expansion. Senescence-associated DNAm may therefore be indirectly mediated by the histone code, insulators, chromatin loops and the overarching nuclear structure.

It is generally anticipated that maintenance of the DNAm pattern and *de novo* methylation are tightly controlled by specific epigenetic modifiers. Our results indicate that at least senescence-associated DNAm changes may rather be indirectly controlled by changes in chromatin conformation that impact on the probability of methylation and demethylation events. In analogy, age-associated DNAm changes might be similar modified, but this warrants further analysis. The functional relevance of DNAm changes is yet unclear. It has been demonstrated that senescence-associated DNAm is only partly associated with gene expression changes of corresponding genes (Koch et al. 2013; Sidler et al. 2014; Sakaki et al. 2017). Yet, it appears to be an oversimplification to consider only transcriptional regulation as functionally relevant. Chromatin conformation, loop structures, histone modifications, and lamina-associated domains may favor modulation and generation of specific DNAm patterns – and on the other hand DNAm and hemimethylation may stabilize such chromatin structures. To understand the underlying mechanism that drive molecular

changes in replicative senescence it may therefore be relevant to take different levels of chromatin organization into account.

Methods

Identification of senescence-associated CpGs

We compiled 64 available DNAm datasets of untreated primary cells with reliable information on passage numbers (all Illumina 450k Methylation BeadChip; Supplemental Table S1). For further analysis we excluded CpGs from X and Y chromosomes and performed k-nearest neighbor imputation and quantile normalization using the R packages impute and lumi (Du et al. 2008), respectively. An outlier test using the R package car was performed resulting in exclusion of one sample (GSM1004625). Subsequently, 15 hyper- and 15 hypomethylated CpG sites were selected based on Pearson correlation of DNAm levels (beta values) and passage numbers ($R^2 > 0.8$) and slope of regression ($m > 0.02$; Supplemental Table S2). Using the R package leaps we calculated the best multivariable linear regression model including two hypo- and two hypermethylated CpG sites.

Analysis of DNAm changes during reprogramming into iPSCs

The kinetics of senescence-associated DNAm changes during reprogramming into iPSCs were investigated using the dataset of Ohnuki *et al.* (GSE45848; 450k BeadChip) (Ohnuki et al. 2014). For comparison we used DNAm profiles of MSCs at early and late passage (GSE37067) (Koch et al. 2013). We either focused on the above mentioned 15 hypo- and 15 hypermethylated CpGs or on 2,116 (hypo) and 1,702 (hyper) senescence-associated CpGs as identified in our previous work (results not demonstrated) (Koch et al. 2013). Furthermore, we focused on 1,432 pluripotency-associated CpGs (Nazor et al. 2012) and

99 age-associated CpGs (Weidner et al. 2014). Heatmaps were produced with R package gplots.

Cell culture

Mesenchymal stem cells were isolated from the bone marrow of donors undergoing orthopedic surgery (n = 11; BM-MSC RWTH Aachen) (Fernandez-Rebollo et al. 2017), from bone marrow aspirates of allogeneic hematopoietic stem cell donors (n = 8; BM-MSC University of Heidelberg) (Horn et al. 2008), from subcutaneous adipose tissue lipoaspirates (n = 3, AT-MSC RWTH Aachen) (Cholewa et al. 2011) and from umbilical cord pieces (n = 2; UCP-MSC; University of Hannover) (Selich et al. 2016). Fibroblasts were isolated from dermis (n = 4; RWTH Aachen) (Koch et al. 2012). Human umbilical vein endothelial cells were isolated from umbilical cords of healthy donors after cesarean sections (HUVECS; n = 4; RWTH Aachen) (Baudin et al. 2007) or obtained from Lonza (n = 3, Basel, Switzerland) (Franzen et al. 2017). All samples were taken after informed and written consent and the study was specifically approved by the ethics committees of the corresponding Universities. All cell preparations were thoroughly characterized (including morphology, immunophenotype, and three lineage *in vitro* differentiation potential) and culture conditions were used as described in detail in our previous work (Horn et al. 2008; Cholewa et al. 2011; Koch et al. 2012; Selich et al. 2016; Fernandez-Rebollo et al. 2017; Franzen et al. 2017). In addition, HUVECS were cultured on 0.1% gelatin in M199 medium (Thermo Fisher, Waltham, USA) supplemented with 20% fetal calf serum (FCS) (Gibco Thermo Fisher), 1 % penicillin/streptomycin (PAA)**,** 0.1 % heparin (5000 IU/ml, Ratiopharm) and 50 µg/ml endothelial cell growth supplement (ECGS) (Sigma-Aldrich, St. Louis, USA)**.** For long-term culture all cells were passaged at approximately 90% confluency and reseeded at 10,000 cells/cm^2.

Pyrosequencing

Genomic DNA was isolated with the NucleoSpin Tissue kit (Macherey&Nagel, Düren, Germany) and bisulfite converted using the EZ DNA Methylation kit (Zymo Research, Irvine, CA, USA). Pyrosequencing was performed on a PyroMark ID System (Biotage, Uppsala, Sweden). Primers for pyrosequencing were designed with the PSQ assay design software (Biotage; Supplemental Table S3). DNAm levels were determined with the Pyro-Q-CpG Software (Biotage). To train the Epigenetic-Senescence-Signature on pyrosequencing data we divided the pyrosequencing samples into a training and validation set (Supplemental Tables S8 and S9). The multivariable linear regression model based on DNAm levels (β-values) at the four CpGs in α = *ALOX12* (cg03762994), β = *DOK6* (cg25968937), γ = *LTC4S* (cg26683398) and δ = *TNNI3K* (cg05264232) was as follows:

Predicted passage = 39.0341 – 10.9266 α - 0.4219 β + 5.8979 γ - 38.889 δ

Finally, we used the R package caret (Kuhn 2008) to perform a 10-fold cross-validation on the training dataset.

Lentiviral Barcode-RGB marking

Primary MSCs in umbilical cord tissue were transduced with three different lentiviral vectors containing the fluorescent proteins mCherry (red), Venus (green) or Cerulean (blue) and a barcode of 16 random and 15 vector-backbone-specific nucleotides, as previously published (Selich et al. 2016). Clonal dynamics were assessed at each passage by flow cytometry using a BD LSR II flow cytometer (BD Biosciences, Heidelberg, Germany) and by deep sequencing of PCR amplified barcodes using ion torrent sequencing. Area plots were produced with the R package ggplot2.

Barcoded-bisulfite-amplicon sequencing

Bisulfite converted DNA was used for a nested PCR using the PyroMark PCR kit (Qiagen; primers are provided in Supplemental Table S4). The second PCR added barcoded Illumina adapters that allowed to distinguish donors and passages, as described before (Franzen et al. 2017). Amplicons were pooled and sequenced on an Illumina MiSeq lane with the v2 nano reagents (Illumina) in 250 PE mode. Bisulfite converted sequencing data was analyzed using TrimGalore, Bismark (Krueger and Andrews 2011) and bowtie2 (Langmead and Salzberg 2012). Further pattern analysis and visualization was performed with custom perl and R scripts or with R package ggplot2 for area plots.

Analysis of hemimethylation

Hemimethylation analysis was modified from a protocol by Laird et al. (Laird et al. 2004). Genomic DNA (4 µg, three donors in passages 2 and 13) was digested with restriction enzymes that cut close to our CpG of interest: AccI (*CASR*, *TNNI3K*, *C12orf12*), DdeI (*KRTAP13.3*, *LTC4S*) and CviQI (*DOK6*, *GRM7*, *PRAMEF2*, *SELP*, *GNAS*). Hairpin linkers (Supplemental Table S5) were denatured at 95°C and subsequently folded by slow cooling to room temperature. Ligation was performed with 4,000 U ligase and 3.3 µM hairpin linker DNA over night at 16° C. The ligated DNA was denatured with 0.3 M NaOH at 42°C for 15 min and 99°C for 2 min before a 0.4 g/ml sodium bisulfite and 1 mg/ml hydroquinone solution was added. Bisulfite conversion was performed at 55°C overnight with 10 intervening denaturation steps at 99°C to prevent renaturation of hairpin structures. The regions of interest were subsequently amplified by PCR using the PyroMark PCR kit (Qiagen; primers listed in Supplemental Table S6). Illumina adapters were ligated using the GeneRead DNA Library I Core Kit (Qiagen) and GeneRead DNA I Amp Kit (Qiagen). Final library cleaning was performed with the Select-a-size DNA Clean & Concentrator kit (Zymo Research). MiSeq v2 nano reagents (Illumina) were used for

library dilution to 4 nM and 20% PhiX were spiked in to increase sequencing diversity. Sequencing was performed on an Illumina MiSeq in 250 PE mode and analyzed as described above.

Circular Chromatin Conformation Capture (4C)

Ten million cells from two MSC preparations at early (P2 and 3) and late (P7 and 9) passages were cross-linked with 4% paraformaldehyde (Electron Microscopy Sciences, Hatfield, PA, USA) for 10 min., harvested into ice cold PBS with 0.125 M glycine, and frozen under a protease inhibitor cocktail (Roche, Basel, Suisse). 4C-seq was performed as described before (Stadhouders et al. 2013) using *ApoI* as the primary and *DpnII* as the secondary restriction enzyme. Bait-specific primers for the circularized inverse PCR are listed in Supplemental Table S7. Amplicons were sequenced on a HiSeq2500 platform (Illumina), mapped to the reference genome (hg19) and analyzed with the Hidden-Markov-Model-based tool *4Cker* (Raviram et al. 2016). This tool corrects for increasing signal noise in *trans* chromosomal interactions and far-cis chromosomal interactions by adaptive window sizes. We used the k^{th}-next-neighbor adaptive window sizes of k=5 for nearbait (10 MB around the “bait” region of interest) and cis interaction analysis and k=20 for *trans* chromosomal interactions. We focused particularly on the high-interacting reads that were called in both replicates. Circos-plots were generated with the R package RCircos (Zhang et al. 2013). For comparison we used DNAm profiles of MSCs at early (n = 5) and late passages (n = 5; GSE37067) (Koch et al. 2013). CTCF and CTCFL enrichment in interacting regions was tested with the RGT motif enrichment tool (http://www.regulatory-genomics.org/motif-analysis/introduction). CTCF ChipSeq data of MSCs (Dixon et al. 2015) was analyzed using TrimGalore, bowtie2 (Langmead and Salzberg 2012) and MACS2 (Zhang et al. 2008). CTCF ChipSeq peaks and enrichment of interactions with

lamina-associated domains (LADs) (Guelen et al. 2008) were analyzed in comparison to randomly chosen regions by Fisher's exact test using R stats.

Additional information

Acknowledgements

We would like to thank all donors for their valuable support of our research.

Authors' contributions

JF contributed to experimental design, analysis and writing of the manuscript; TG, LB and AP performed 4C experiments; AnS, MR and AxS performed barcoded-RGB-vector experiments; RS helped to establish hairpin BBA-Seq; EFR, CG and AO contributed to long term cell culture and cellular characterization; MB helped to sequence data; BR and ADH contributed important material. WW contributed to experimental design, data analysis and writing of the manuscript.

Competing interests

Wolfgang Wagner is cofounder of Cygenia GmbH (www.cygenia.com), which can provide service for the Epigenetic-Senescence-Signature to other scientists. All other authors do not have a conflict of interest to declare.

Funding

This work was particularly supported by the Else Kröner-Fresenius-Stiftung (AP and WW: 2014_A193), by the Deutsche Forschungsgemeinschaft (WW: WA 1706/8-1; AP: UoC Advancer Research grant of the DFG Excellence Initiative, MR: RO 5102/1-1), by the German Ministry of Education and Research (WW: OBELICS, 01KU1402B), by the Interdisciplinary Center for Clinical Research within the faculty of Medicine at the RWTH Aachen University (WW: IZKF O3-3), and by CMMC core funding (AP).

References

Arand J, Wossidlo M, Lepikhov K, Peat JR, Reik W, Walter J. 2015. Selective impairment of methylation maintenance is the major cause of DNA methylation reprogramming in the early embryo. *Epigenetics & Chromatin* **8**: 1.

Baudin B, Bruneel A, Bosselut N, Vaubourdolle M. 2007. A protocol for isolation and culture of human umbilical vein endothelial cells. *Nature protocols* **2**: 481-485.

Bernadotte A, Mikhelson VM, Spivak IM. 2016. Markers of cellular senescence. Telomere shortening as a marker of cellular senescence. *Aging (Albany NY)* **8**: 3-11.

Bork S, Pfister S, Witt H, Horn P, Korn B, Ho AD, Wagner W. 2010. DNA methylation pattern changes upon long-term culture and aging of human mesenchymal stromal cells. *Aging cell* **9**: 54-63.

Campisi J. 1997. The biology of replicative senescence. *European journal of cancer (Oxford, England : 1990)* **33**: 703-709.

Chandra T, Ewels Philip A, Schoenfelder S, Furlan-Magaril M, Wingett Steven W, Kirschner K, Thuret J-Y, Andrews S, Fraser P, Reik W. 2015. Global Reorganization of the Nuclear Landscape in Senescent Cells. *Cell Reports* **10**: 471-483.

Childs BG, Durik M, Baker DJ, van Deursen JM. 2015. Cellular senescence in aging and age-related disease: from mechanisms to therapy. *Nature medicine* **21**: 1424-1435.

Choi M, Genereux DP, Goodson J, Al-Azzawi H, Allain SQ, Simon N, Palasek S, Ware CB, Cavanaugh C, Miller DG et al. 2017. Epigenetic memory via concordant DNA methylation is inversely correlated to developmental potential of mammalian cells. *PLoS genetics* **13**: e1007060.

Cholewa D, Stiehl T, Schellenberg A, Bokermann G, Joussen S, Koch C, Walenda T, Pallua N, Marciniak-Czochra A, Suschek CV et al. 2011. Expansion of Adipose Mesenchymal Stromal Cells is Affected by Human Platelet Lysate and Plating Density. *Cell transplantation* **20**: 1409-1422.

Criscione SW, De Cecco M, Siranosian B, Zhang Y, Kreiling JA, Sedivy JM, Neretti N. 2016. Reorganization of chromosome architecture in replicative cellular senescence. *Science Advances* **2**: e1500882.

Cruickshanks HA, McBryan T, Nelson DM, Vanderkraats ND, Shah PP, van Tuyn J, Singh Rai T, Brock C, Donahue G, Dunican DS et al. 2013. Senescent cells harbour features of the cancer epigenome. *Nature cell biology* **15**: 1495-1506.

Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W et al. 2015. Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**: 331-336.

Du P, Kibbe WA, Lin SM. 2008. lumi: a pipeline for processing Illumina microarray. *Bioinformatics (Oxford, England)* **24**: 1547-1548.

Fernandez-Rebollo E, Mentrup B, Ebert R, Franzen J, Abagnale G, Sieben T, Ostrowska A, Hoffmann P, Roux P-F, Rath B et al. 2017. Human Platelet Lysate versus Fetal Calf Serum: These Supplements Do Not Select for Different Mesenchymal Stromal Cells. *Scientific Reports* **7**: 5132.

Franzen J, Zirkel A, Blake J, Rath B, Benes V, Papantonis A, Wagner W. 2017. Senescence-associated DNA methylation is stochastically acquired in subpopulations of mesenchymal stem cells. *Aging cell* **16**: 183-191.

Frobel J, Hemeda H, Lenz M, Abagnale G, Joussen S, Denecke B, Šarić T, Zenke M, Wagner W. 2014. Epigenetic Rejuvenation of Mesenchymal Stromal Cells Derived from Induced Pluripotent Stem Cells. *Stem Cell Reports* **3**: 414-422.

Grezella C, Fernandez-Rebollo E, Franzen J, Ventura Ferreira MS, Beier F, Wagner W. 2018. Effects of senolytic drugs on human mesenchymal stromal cells. **9**: 108.

Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W et al. 2008. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**: 948-951.

Guntrum M, Vlasova E, Davis TL. 2017. Asymmetric DNA methylation of CpG dyads is a feature of secondary DMRs associated with the Dlk1/Gtl2 imprinting cluster in mouse. *Epigenetics & Chromatin* **10**: 31.

Hänzelmann S, Beier F, Gusmao EG, Koch CM, Hummel S, Charapitsa I, Joussen S, Benes V, Brümmendorf TH, Reid G et al. 2015. Replicative senescence is associated with nuclear reorganization and with DNA methylation at specific transcription factor binding sites. *Clinical Epigenetics* **7**: 19.

Horn P, Bork S, Diehlmann A, Walenda T, Eckstein V, Ho A, Wagner W. 2008. Isolation of human mesenchymal stromal cells is more efficient by red blood cell lysis. *Cytotherapy* **10**: 676-685.

Itahana K, Campisi J, Dimri GP. 2007. Methods to detect biomarkers of cellular senescence: the senescence-associated beta-galactosidase assay. *Methods in molecular biology (Clifton, NJ)* **371**: 21-31.

Jeon OH, Kim C, Laberge R-M, Demaria M, Rathod S, Vasserot AP, Chung JW, Kim DH, Poon Y, David N et al. 2017. Local clearance of senescent cells attenuates the development of post-traumatic osteoarthritis and creates a pro-regenerative environment. *Nature Medicine* **23**: 775.

Jones PA, Liang G. 2009. Rethinking how DNA Methylation Patterns are Maintained. *Nature reviews Genetics* **10**: 805-811.

Kirkland JL, Tchkonia T, Zhu Y, Niedernhofer LJ, Robbins Paul D. 2017. The Clinical Potential of Senolytic Drugs. *Journal of the American Geriatrics Society* **65**: 2297-2301.

Koch C, Reck K, Shao K, Lin Q, Joussen S, Ziegler P, Walenda G, Drescher W, Opalka B, May T et al. 2013. Pluripotent Stem Cells Escape From Senescence-Associated DNA Methylation Changes. *Genome Res* **23**: 248-259.

Koch CM, Joussen S, Schellenberg A, Lin Q, Zenke M, Wagner W. 2012. Monitoring of cellular senescence by DNA-methylation at specific CpG sites. *Aging cell* **11**: 366-369.

Krueger F, Andrews SR. 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics (Oxford, England)* **27**: 1571-1572.

Kuhn M. 2008. Building Predictive Models in R Using the caret Package. *2008* **28**: 26.

Laird CD, Pleasant ND, Clark AD, Sneeden JL, Hassan KM, Manley NC, Vary JC, Jr., Morgan T, Hansen RS, Stoger R. 2004. Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary strands of individual DNA molecules. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 204-209.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**: 357-359.

Li D, Liu J, Yang X, Zhou C, Guo J, Wu C, Qin Y, Guo L, He J, Yu S et al. 2017. Chromatin Accessibility Dynamics during iPSC Reprogramming. *Cell stem cell* **21**: 819-833.e816.

Lövkvist C, Dodd IB, Sneppen K, Haerter JO. 2016. DNA methylation in human epigenomes depends on local topology of CpG sites. *Nucleic Acids Research* **44**: 5123-5132.

Lowe D, Horvath S, Raj K. 2016. Epigenetic clock analyses of cellular senescence and ageing. *Oncotarget* **7**: 8524-8531.

Lowe R, Overhoff MG, Ramagopalan SV, Garbe JC, Koh J, Stampfer MR, Beach DH, Rakyan VK, Bishop CL. 2015. The senescent methylome and its relationship with cancer, ageing and germline genetic variation in humans. *Genome Biology* **16**: 194.

Nadejda L, F. HJ, G. AL, C. PJ, Sabita S, Trygve T. 2002. Differential maintenance and de novo methylating activity by three DNA methyltransferases in aging and immortalized fibroblasts. *Journal of Cellular Biochemistry* **84**: 324-334.

Nazor KL, Altun G, Lynch C, Tran H, Harness JV, Slavin I, Garitaonandia I, Muller FJ, Wang YC, Boscolo FS et al. 2012. Recurrent variations in DNA methylation in human pluripotent stem cells and their differentiated derivatives. *Cell stem cell* **10**: 620-634.

Ohnuki M, Tanabe K, Sutou K, Teramoto I, Sawamura Y, Narita M, Nakamura M, Tokunaga Y, Nakamura M, Watanabe A et al. 2014. Dynamic regulation of human endogenous retroviruses mediates factor-induced reprogramming and differentiation potential. *Proceedings of the National Academy of Sciences of the United States of America* **111**: 12426-12431.

Patino-Parrado I, Gomez-Jimenez A, Lopez-Sanchez N, Frade JM. 2017. Strand-specific CpG hemimethylation, a novel epigenetic modification functional for genomic imprinting. *Nucleic Acids Res* **45**: 8822-8834.

Raviram R, Rocha PP, Muller CL, Miraldi ER, Badri S, Fu Y, Swanzey E, Proudhon C, Snetkova V, Bonneau R et al. 2016. 4C-ker: A Method to Reproducibly Identify Genome-Wide Interactions Captured by 4C-Seq Experiments. *PLoS computational biology* **12**: e1004780.

Sakaki M, Ebihara Y, Okamura K, Nakabayashi K, Igarashi A, Matsumoto K, Hata K, Kobayashi Y, Maehara K. 2017. Potential roles of DNA methylation in the initiation and establishment of replicative senescence revealed by array-based methylome and transcriptome analyses. *PLoS ONE* **12**: e0171431.

Selich A, Daudert J, Hass R, Philipp F, von Kaisenberg C, Paul G, Cornils K, Fehse B, Rittinghausen S, Schambach A et al. 2016. Massive Clonal Selection and Transiently Contributing Clones During Expansion of Mesenchymal Stem Cell Cultures Revealed by Lentiviral RGB-Barcode Technology. *Stem Cells Translational Medicine* **5**: 591-601.

Shah PP, Donahue G, Otte GL, Capell BC, Nelson DM, Cao K, Aggarwala V, Cruickshanks HA, Rai TS, McBryan T et al. 2013. Lamin B1 depletion in senescent cells triggers large-scale changes in gene expression and the chromatin landscape. *Genes & Development* **27**: 1787-1799.

Sharif J, Koseki H. 2018. Hemimethylation: DNA's lasting odd couple. *Science* **359**: 1102-1103.

Sidler C, Woycicki R, Kovalchuk I, Kovalchuk O. 2014. WI-38 senescence is associated with global and site-specific hypomethylation. *Aging (Albany NY)* **6**: 564-574.

Stadhouders R, Kolovos P, Brouwer R, Zuin J, van den Heuvel A, Kockx C, Palstra RJ, Wendt KS, Grosveld F, van Ijcken W et al. 2013. Multiplexed chromosome conformation capture sequencing for rapid genome-scale high-resolution detection of long-range chromatin interactions. *Nature protocols* **8**: 509-524.

Ushijima T, Watanabe N, Okochi E, Kaneda A, Sugimura T, Miyamoto K. 2003. Fidelity of the methylation pattern and its variation in the genome. *Genome research* **13**: 868-874.

Weidner CI, Lin Q, Koch CM, Eisele L, Beier F, Ziegler P, Bauerschlag DO, Jöckel K-H, Erbel R, Mühleisen TW et al. 2014. Aging of blood can be tracked by DNA methylation changes at just three CpG sites. *Genome Biology* **15**: R24-R24.

Xie W, Kagiampakis I, Pan L, Zhang YW, Murphy L, Tao Y, Kong X, Kang B, Xia L, Carvalho FLF et al. 2018. DNA Methylation Patterns Separate Senescence from Transformation Potential and Indicate Cancer Risk. *Cancer cell* **33**: 309-321.e305.

Xu C, Corces VG. 2018. Nascent DNA methylome mapping reveals inheritance of hemimethylation at CTCF/cohesin sites. *Science* **359**: 1166-1170.

Zhang H, Meltzer P, Davis S. 2013. RCircos: an R package for Circos 2D track plots. *BMC Bioinformatics* **14**: 244-244.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W et al. 2008. Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* **9**: R137.

Zirkel A, Nikolic M, Sofiadis K, Mallm J-P, Brackley CA, Gothe H, Drechsel O, Becker C, Altmüller J, Josipovic N et al. 2018. HMGB2 Loss upon Senescence Entry Disrupts Genomic Organization and Induces CTCF Clustering across Cell Types. *Molecular Cell* **70**: 730-744.e736.