Transactions on Materials, Biotechnology and Life Sciences
ISSN: 2960-1932, eISSN: 2960-2319 | Volume 3
BLSME 2024

Causal analysis on DNA methylation reveals key regulators in aging and senescence

Tingyi Cui *

The High School Affiliated to Renmin University of China, Beijing, China

* Corresponding Author Email: cuitingyi@outlook.com

Abstract. Aging is a multifaceted biological phenomenon influenced by genetic, epigenetic, and environmental factors. Cellular senescence, characterized by an irreversible cell cycle arrest, is a pivotal mechanism contributing to aging and age-related diseases. It is also associated with alterations in DNA methylation, a key epigenetic marker. However, the causal link between senescence-associated DNA methylation changes and aging outcomes remains elusive. In our study, we analyzed three distinct methods to induce cellular senescence and systematically evaluated the resultant DNA methylation profiles. Utilizing Mendelian Randomization, we identified that 4.1% of the methylated loci were significantly associated with aging outcomes and accounted for the majority of the observed variance. Further mediation analysis pinpointed key genes, including *CISD2*, *MARS2*, and *ETFB*, that serve as mediators for these epigenetic effects. Experimental validation using Senescence-Associated-β-Galactosidase (SA-β-gal) staining confirmed that knockdown of these genes triggers cellular senescence. Our findings provide compelling evidence for the causal role of DNA methylation in cellular senescence and its impact on aging.

Keywords: Aging; DNA methylation; Epigenetics; Mendelian Randomization.

1. Introduction

Aging is a complex biological process common to all living organisms and involves multiple factors such as genetics, environment, and lifestyle. Aging increases susceptibility to diseases such as neurodegenerative diseases, cardiovascular diseases, and cancer. It is also associated with other physiological disorders ranging from the molecular to macroscopic levels. However, the cause and effect of aging epigenetics still need to be understood.

Cellular senescence is a permanent halt in the cell cycle that modifies chromatin structure [1]. Cellular senescence was reported to play an important role in aging. It can be triggered by many aging-associated stimuli and result in tissue remodeling and organismal aging [2]. The accumulation of senescent cells can lead to anti-apoptotic phenotypes and senescence-associated secretory phenotype (SASP) [3]. Through SASP, senescent cells can drive normal and pathological aging-associated phenotypes [4]. However, cellular senescence may also provide benefits such as tumor suppressing and tissue repair. SASP helps to maintain tumor suppressive growth arrest of senescent cells [4].

DNA methylation is an epigenetic modification of the cytosine ring by the addition of methyl groups. Thus, the cytosine-guanine (CpG) dinucleotide site is where DNA methylation occurs. DNA methylation plays a crucial role in gene regulation and can influence gene expression by either promoting or inhibiting transcription. Generally, DNA methylation at promoter regions, known as CpG islands, is associated with gene silencing, as it prevents the binding of transcription factors and other regulatory proteins [5]. It is essential for growth and development because it regulates many vital processes, such as genomic imprinting. When dysregulated, DNA methylation can lead to various diseases, including cancer. During aging and senescence, alterations in DNA methylation significantly contribute to cellular dysfunction, activate SASP, and may mediate age-associated conditions.

By selecting genetic variants associated with the exposure but not directly associated with confounding factors, Mendelian randomization (MR) approximates a randomized experiment and

minimizes biases due to confounding variables[6]. MR is a strategy for drawing causal inferences that employ inherent genetic variations as instrumental variables. As genetic variant allocation is dictated by chance during conception, the causal effects derived through MR remain unaffected by environmental variables that might otherwise introduce bias [7,8]. This makes MR a valuable tool to further investigate the relationship between CpG sites causal to life span and sites related to senescence [9]. A previous study suggested that DNA methylation is interconnected with genetic variation, playing a role in gene regulation [10]. The study emphasizes that DNA methylation levels can be influenced by the binding of transcription factors (TFs), which are in turn, affected by both variations in TF abundance and genetic variants at their binding sites. [10]

DNA methylation levels at CpG sites also have a causal relationship to aging. A study indicated that methylation levels in CpG sites located in the genes *ITGA2B, ASPA*, and *PDE4C* can be used to track epigenetic changes, thus estimating the state of aging [11]. DNA methylation level can also be used as a biomarker to predict the cellular senescence state [12].

In our study, the methylation of DNA based on the principles of Mendelian randomization is primarily focused on inferring causality between DNA methylation and aging outcomes. We aim to discover the methylation loci associated with lifespan using Mendelian randomization and integrate them with the methylation changes associated with cellular senescence. As epigenetics and senescence that altogether cause aging are unclear, our comprehensive approach improves the understanding of the mechanisms underlying aging.

2. Method

2.1. Cellular senescence DNA methylation-related data

A variety of human DNA was obtained from diverse cell types, including skin, basal cell carcinoma, keratinocytes, fibroblasts, endothelial cells, keratinocyte stem cells, induced pluripotent stem cells, HUVEC, HCAEC, and neural progenitor cells. These cells were subjected to various conditions and treatments including the use of CCCP (Mitochondrial oxidative phosphorylation uncoupler), bezafibrate, hTERT immortalization, chronic radiation, or rapamycin to induce cellular senescence. Epigenetics data from those samples were generated using an Infinium microarray, a total of 865859 sites. The tibble was filtered to include 96 sites with specific conditions (X-ray, RAS, Replication, None).

2.2. Principal component analysis

Principal component analysis (PCA) is a multivariate method employed to reduce dimensions of highly dimensional datasets [27]. It is performed on the DNA methylation data after using the prcomp function to preprocess the data (centering and scaling). The objective is to extract the core features in the dataset and dimensionally reduce them into fewer orthogonal variables known as principal components. The principal components were plotted by grouping experimental conditions. The Euclidean distance indicates the similarity among data points.

2.3. Differential DNA methylation analysis

We did differential DNA methylation analysis by comparing three conditions: RAS, Replication, and X-ray to identify differentially methylated probes (DMPs) associated with these conditions. A statistical analysis was performed to compare DNA methylation levels between different conditions. A linear model was fitted using the limma package and the empirical Bayes method was performed on the model (https://github.com/cran/limma). The annotation information for the Illumina EPIC array was added using IlluminaHumanMethylationEPICanno.ilm10b2.hg19 (https://bioconductor.org/packages/3.17/data/annotation/html/IlluminaHumanMethylationEPICanno.ilm10b2.hg19.html). The annotation data was processed to extract the gene names associated with

each CpG site. Differential methylation analysis for each condition (RAS, Replication, X-ray) was performed and merged with the gene name annotation tibble for each condition.

2.4. Volcano plot

The significant DNA methylation sites (CGids) for each senescence condition were filtered based on the adjusted p-value threshold of 0.05 (P. Value < 0.05). Correlation between the senescence conditions is calculated and analyzed based on the significant DNA methylation sites to create a volcano plot, mapping logFC to the x-axis, -log10(P. Value) to the y-axis, and using logFC for color and fill aesthetics.

2.5. Compare different senescence conditions

To ensure statistical significance, non-significant sites were filtered out based on the adjusted p-value (adj. P. Val <= 0.05). The significant sites from each dataset were combined to form a unified dataset (507754 observations) containing the common significant sites across all three conditions. Correlation analysis was performed on this merged dataset. The correlation coefficient (r) was calculated to quantify the strength and direction of the relationship between variables. In cases where some correlations were not computable due to missing values, these were imputed as a correlation value of 1 to preserve the integrity of subsequent analyses. To visualize the correlation results, a heatmap was created. Clusters of co-regulated sites that may indicate shared molecular pathways or functional associations among the conditions were identified.

2.6. MeQTL data

DNA methylation data was retrieved from the Genetics of DNA Methylation Consortium (GoDMC). DNA methylation levels were assessed in whole blood samples comprising 36 cohorts which include 27, 750 European subjects. To chart the genetic factors impacting DNA methylation levels, a comprehensive analysis of 420, 509 CpG sites was conducted. MeQTLs with cis-acting effects were identified within a 2 MB radius around the designated CpG site. GoDMC statistics mentioned above can be accessed at http://mqtldb.godmc.org.uk.

2.7. GWAS

A genome-wide association study (GWAS) analysis was performed to investigate the genetic basis of a specific phenotype using the dataset DS_10283_3209 (9085648 SNPs). The dataset was initially obtained in a raw format from the LifeGen Phase 2 study, containing a comprehensive set of genetic variant information. To ensure data quality, variants with low information content (INFO score < 0.8) were filtered out and relevant columns were selected, such as chromosome, position, effect allele, other allele, effect allele frequency (EAF), effect size (beta), standard error (se), p-value (p), and sample size (N). The processed GWAS statistics included 8760737 SNPs. The reference genome (GRCh37) and the dbSNP version (144) is specified to facilitate integration with genomic databases. Variants with INFO scores below 0.8 were excluded from the processed summary statistics to ensure the reliability of the final dataset.

2.8. Mendelian Randomization

When performing the MR analysis, we filtered only to include the genetic variants that are strongly associated with whole blood DNA methylation level (FDR < 0.05) and meQTLs that are in the cis-acting regions. Since the generalized MR method enhances statistical power through the incorporation of partially correlated instruments while accounting for LD structure, we implemented LD clumping. This involved selectively eliminating meQTLs with substantial LD (r2 > 0.3), following the recommendation of Burgess et al [23].

For the primary investigations, we employed three distinct MR techniques: Wald ratio for cases with only one available meQTL, generalized inverse variance weighted (gIVW) when two or more

meQTLs were accessible, and generalized MR-Egger regression (gEgger) when a minimum of three meQTLs were available. Execution of the MR analyses was accomplished using the Mendelian Randomization R package and the TwoSampleMR R package from the GitHub repository (github.com/MRCIEU/TwoSampleMR).

To identify causal CpG sites supported by robust MR evidence, CpG-phenotype pairs exhibiting Padjusted values below 0.05 following Bonferroni correction were selected.

2.9. Identify causal DNA methylation change in senescence

To identify significant sites associated with senescence, data preprocessing was conducted. Only those sites with adjusted p-values (adj. P. Val) less than or equal to 0.05 were retained using filtering procedures. The retained sites were then extracted using the "pull (CGid)" function. Next, all the significant sites identified in the three senescence datasets were combined into a single list. Subsequently, another dataset, "file_raw," was read and filtered based on the criteria of being cis-regulated (cistrans == T) and having a CpG site present in the previously obtained "all_sig_site" list. This filtered dataset was named "cis_list."

2.10. Mediation Analysis

Mediation analysis was assessed for CpG-Aging-GIP1(162716 sites). The Ras, Rep, and Xray data frames (865859 sites) were joined with the "cpg_list" data frame using the "CGid" and "exposure" columns. For each joined table, the "fdr1" and "fdr2" columns are calculated by adjusting the "P.Value" and "pval" columns using the Benjamini-Hochberg (BH) method. The tables are then filtered to retain only rows where "fdr1" and "fdr2" are both less than 0.05. The filtered tables are arranged in descending order based on the absolute value of "b*logFC". Each filtered table included 2536, 1309, and 15 observations, respectively.

The "rep_joined_table", "xray_joined_table", and "ras_joined_table" data frames were created by joining the "mediation" data frame with the "ras_tb", "rep_tb", and "xray_tb" data frames, respectively, using the "Expo_ID" and "CGid" columns.

"Preprocess" is a function defined to preprocess the joined tables which calculates additional columns such as "b_med" (difference between "b_ivw_tot" and "b_direct_cis") and "mp" (mediation proportion) based on specific calculations. Values of "mp" that exceed 1 were set to 1, and values below 0 were set to 0. The mediation results were obtained by preprocessing the tables and filtering rows where the sign of "b" is different from the sign of "logFC. The mediation results of Rep, Ras, and xray included 645, 126, and 0 sites respectively.

2.11. Cell culture

MRC-5 Human Embryonic Lung Fibroblast Cells were cultured in the National Collection of Authenticated Cell Cultures (Shanghai, China) in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin–streptomycin and incubated at 37°C with 5% CO2 and 70-80% humidity.

2.12. siRNA silencing

The incubated cells were then transfected with short hairpin RNA (shRNA) using Lipofectamine 3000. Briefly, the siRNA and Lipofectamine 3000 were separately incubated. The sequences of siRNAs are as follow tables:

Table 1. Primer bands used for siRNA silencing

Gene	Name	Sequence (5' to 3')
CISD2	CISD2-297-sense	GCUUGAUUAAUCUUAAAAUTT
	CISD2-297-antisense	AUUUUAAGAUUAAUCAAGCTT
	CISD2-368-sense	CUUACUAAAGCAGCUUAUUTT
	CISD2-368-antisense	AAUAAGCUGCUUUAGUAAGTT
	CISD2-470-sense	CCACUAAUACUGAAGAAGATT
	CISD2-470-antisense	UCUUCUUCAGUAUUAGUGGTT
MARS2	MARS2-525-sense	GUCUAUGAAGGUUGGUAUUTT
	MARS2-525-antisense	AAUACCAACCUUCAUAGACTT
	MARS2-745-sense	CAUUUCAUCACGUAGUUCUTT
	MARS2-745-antisense	AGAACUACGUGAUGAAAUGTT
	MARS2-912-sense	CCAAAUGCUGAGUUCAAAUTT
	MARS2-912-antisense	AUUUGAACUCAGCAUUUGGTT
ETFB	ETFB-2406-sense	CCUCAUUUGUAUUUAUGUATT
	ETFB-2406-antisense	UACAUAAAUACAAAUGAGGTT
	ETFB-2489-sense	CCAUAAAUACUGUAAAUUUTT
	ETFB-2489-antisense	AAAUUUACAGUAUUUAUGGTT
	ETFB- 3489-sense	GGAGAUUGGGCGGAUUUGATT
	ETFB-3489-antisense	UCAAAUCCGCCCAAUCUCCTT

The above siRNAs were purchased from General Biosystems (Anhui, China).

To confirm the knockout, we performed quantitative real-time PCR (qRT-PCR) analysis. RNA was extracted using Trizol reagent and subjected to reverse transcription using the HiScript II 1st Strand cDNA Synthesis Kit. The AceQ Universal SYBR qPCR Master Mix is used to detect the expression of the target gene using an Applied Biosystems 7900 instrument.

2.13. SA-β-gal Staining

To induce senescence, we exposed the cells to 10uM of atazanavir for each 20 pmol/ml of siRNA [29]. We then performed Senescence-Associated - β -galactosidase staining to determine the expression of senescence-associated β-galactosidase (SA-β-gal) in cells. The groups were A: siNC, B: siNC+atazanavir, C: siCISD2, D: siCISD2 + atazanavir, E: siMARS2, F: siMARS2 + atazanavir, G: siETFB, H: siETFB+ atazanavir.

3. Results

3.1. Senescent cells induced by different methods show distinct epigenetic profiles

Senescence can be induced by processes such as telomere attrition, unrepaired DNA damage, and overexpression of oncogenes. To investigate whether genome-wide methylation profile is associated with these, we used data from primary human dermal fibroblasts (HDFs) from 14 healthy neonatal donors with four conditions: control, X-irradiated senescence (20 Gy), Ras-transduced senescence and replicative senescence (Fig. 1).

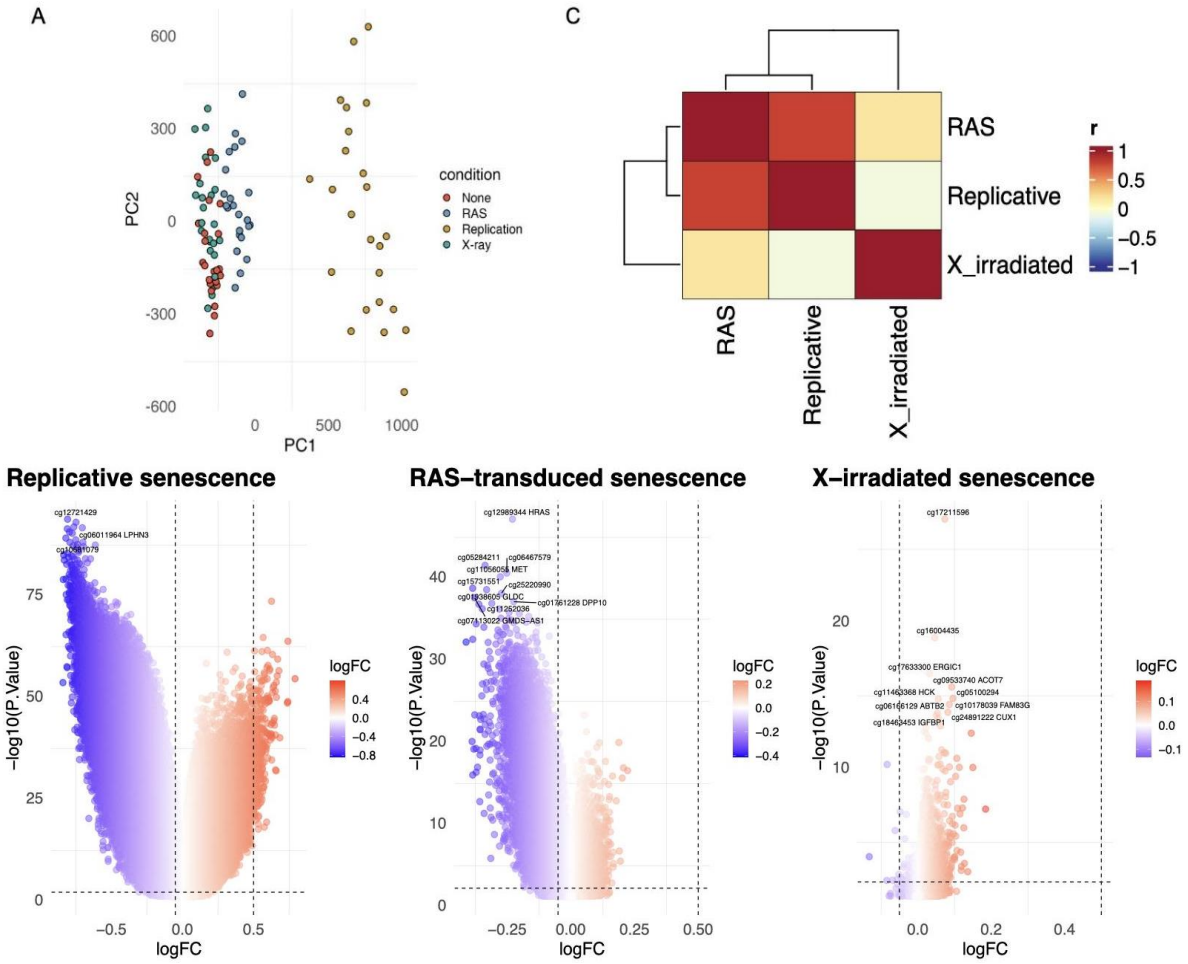

Figure 1. Genome-wide methylation profile of different types of cellular senescence. A. Principal component analysis on ras-transduced, replicative, and x-irradiated cells. Replicative senescence differs significantly as it does not cluster with the other two types of senescence on the plot. B. Differential methylation sites for different types of senescence. The Y-axis shows logFC and X-axis shows −log10 (P. Value). logFC is used for color and fill aesthetics. Dashed lines indicate the adjusted P. Value threshold of 0.05. **C.** This graph shows the correlation between the three types of senescence. The color represents absolute values of corresponding genetic correlations. X-irradiated senescence differs significantly from the other types.

The PCA of the DNA methylation profile shows that methylation pattern of replicative senescence differs significantly from the other three groups (Fig. 1A). The Ras-transduced overexpressed cell also shows a moderate difference compared to the control group, while the X-irradiated cell has a similar global methylation profile with the control group.

We then performed the differential methylation analysis. For Ras-transduced, a total number of 92, 593 sites were differentially methylated, with 21, 067 sites hypermethylated and 71, 526 sites hypomethylated; for replicative senescence, a total number of 499,483 sites differentially methylated, with 217, 521 sites hypermethylated and 281, 942 hypomethylated; for X-irradiated, a total number of 617 sites were differentially methylated, with 611 sites hypermethylated and 6 sites hypomethylated (Fig. 1B). Top hypomethylated sites for replicative senescence are cg12721329, cg06011964, and cg10681079; top hypomethylation sites for Ras-transduced senescence are cg12989344, cg05284211, and cg06467579; top hypermethylation sites for X-irradiated senescence are cg17211596, cg16004435, and cg17633300 (Fig. 1B).

When then analyzed, the correlation of methylation change across three different senescence-induction conditions (Fig. 1C). Consistent with the PCA results, the correlation analysis on differentially methylated sites shows that Ras-transduced and replicative are highly correlated

(Pearson's R = 0.79), while the X-irradiated are only weakly associated with the other types of senescence (for Ras Pearson's R = 0.17, for replicative Pearson's R = -0.10, Fig. 1C). Together our finding suggests that there is a significant difference on methylation profile across different senescence induced methods.

3.2. Mendelian Randomization reveals adaptive and damaging changes during senescence

Cellular senescence plays two distinct roles in aging: (1) protecting against cancer (protective) and (2) disrupting tissue homeostasis (damaging). To uncover protective and damaging methylation change during senescence, the causal effect of each senescence-related methylation change needs to be evaluated. Here, we used 420,509 CpG sites with meQTLs available and performed Mendelian randomization to identify CpG sites causal to aging.

MR is a genetic approach that mimics the principle of randomized trials that is able to assess the causal relationship between exposure and outcome traits. We used Aging-GIP1, a summary-level aging-related trait that represents healthy longevity, as an outcome trait and performed cis-MR on each senescence-related CpG site[6]. Across all three different conditions, we found that 4.1% of differentially methylated sites had a significant causal effect on AGing-GIP1 after adjusting for multiple test settings. For Ras-transduced senescence, 2,536 out of 92,593 differentially methylated sites are shown to have a significant causal effect on Aging-GIP1; for replicative senescence, 21,646 out of 499,483 differentially methylated sites are shown to have a significant causal effect on Aging-GIP1; for X-irradiated senescence, 15 out of 617 differentially methylated sites are shown to have a significant causal effect on Aging-GIP1 (Fig. 2).

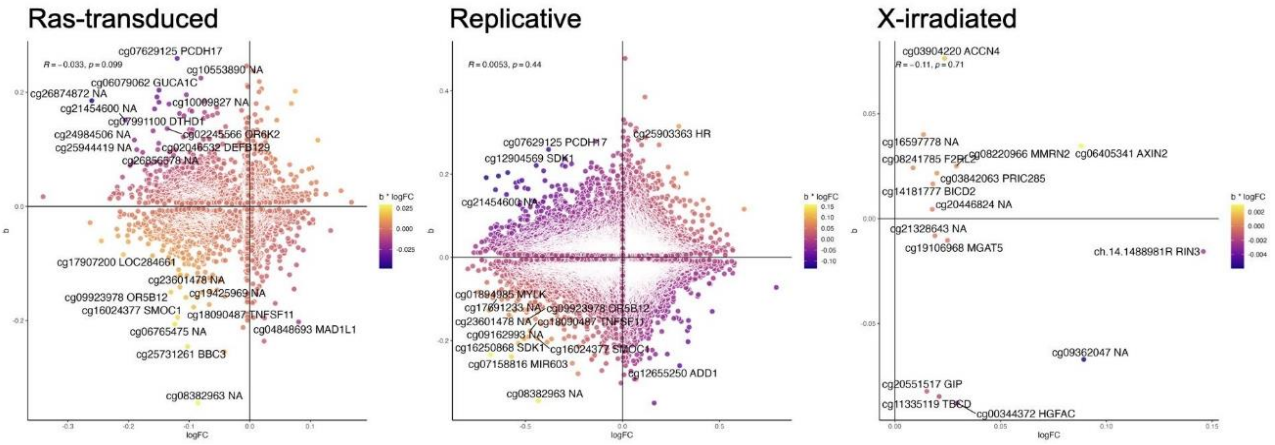

Figure 2. The effect of senescence-related methylation change on aging. The X-axis shows the DNA methylation fold change in log scale (logFC), and the Y-axis shows the causal effect estimate of DNA methylation sites in Aging-GIP1 (b). The color scheme (b*logFC) highlights the expected impact of senescence-related methylation change on aging. The top 30 significant CpG sites based on the absolute value of b*logFC with FDR < 0.05 are annotated.

We then integrate the senescence-related differential methylation information and MR-estimated causal information to infer the damaging and protective role of each methylation site. The role of DNA methylation changes is annotated based on the sign of the product of senescence-related change and causal effect. For instance, if the methylation level of a CpG site increased during senescence and the causal effect estimate is also positive, this CpG site will be annotated as a protective site. Therefore, the methylation sites located in the first and third quadrants are classified as protective, and the methylation sites located in the second and fourth quadrants are classified as damaging. For example, damaging sites such as cg07629125 are associated with chronic obstructive pulmonary disease, and protective sites such as cg25731261 are associated with cancer such as clear cell renal carcinoma and pancreatic ductal adenocarcinoma. [13–15]

3.3. Genetic analysis reveals key mediators of senescence-related methylation sites

To further dissect senescence-related methylation sites, replicative senescence was selected for the mediation analysis. As demonstrated in our PCA, replicative senescence shows the strongest impact on DNA methylome (Fig. 1A). As demonstrated in our differential methylation analysis, replicative senescence included the largest amount of significant hypomethylated and hypermethylated sites (Fig. 1B). For most organisms, replicative senescence is the most common type of senescence experienced compared to X−irradiated senescence and RAS−transduced senescence.

Mediation analysis was conducted to investigate if gene expression is influenced by the effect of top CpG sites. Multivariable MR that dissects significant CpG-phenotype causal effect ($\boldsymbol{\theta}_T$) into direct effects ($\boldsymbol{\theta}_D$) and indirect effects through transcript levels was conducted to estimate the mediation effect. A significant mediation effect was detected in 645 CpG transcriptive pairs (Fig. 3).

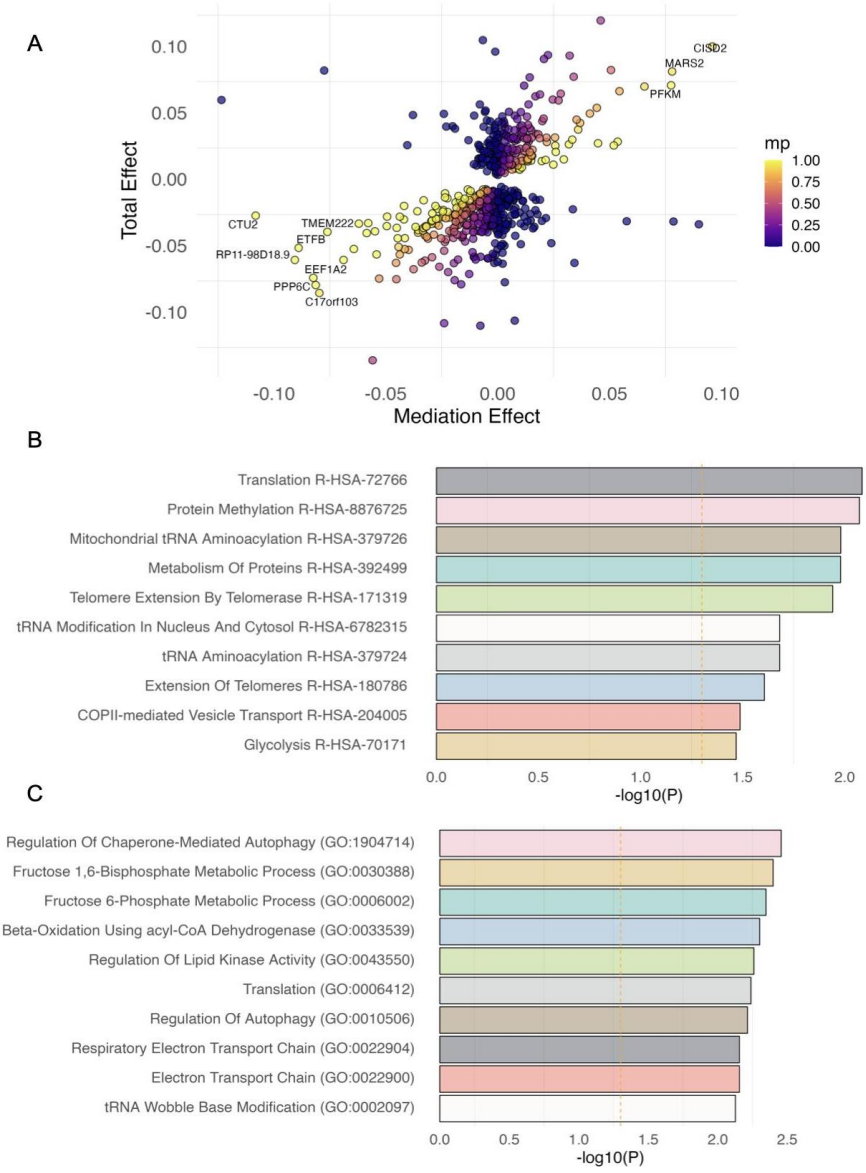

Figure 3. Mediator genes of significant replicative senescence-associated CpG sites. A. The scatter plot indicates MR-estimated causal effects (X-axis) and age-related methylation change (Y-axis) for each significant causal CpG identified in Aging-GIP1. The color scheme represents the mediation proportion (0 being not mediated, 1 being fully mediated). Only CpG sites with adjusted P-values < 0.05 are plotted. The CpG sites with the top 10 largest effect sizes are annotated. B-C. The bar graph shows the enrichment analysis of the top 10 mediator genes of significant replicative senescence-associated CpG sites. The X-axis shows the -log10(P). The pathways were filtered based on a P-value < 0.05. The dashed line represents a threshold for a P-value < 0.05.

We found that 95% of the effect of cg18954925 (chromosome 4q22-24) on Aging-GIP1, which is associated with the disease Wolfram syndrome 2, is mediated through the expression of *CISD2*, 94% of the effect of cg02833180 (chr2:198815064) is mediated through the expression of *MARS2* gene, which is associated with mitochondrial methionyl-tRNA synthetase protein production, and 99% cg17527422 is mediated through the expression of *PFKM* gene, which is associated with enzyme production [16–18] (Fig. 3). We also found that 100% of the effect of cg25728174 (chr19:51814782) is mediated through *ETFB* gene which is associated with electron transfer flavoprotein production and 100% of the effect of cg00620024 (chr9:127952891) is mediated through the *PPP6C* gene [10] (Fig. 3) [19]. Our results suggest that DNA methylation may mediate senescence-related effects at these loci.

We then performed pathway enrichment analysis using the REACTOME and GO to gain insights into the biological processes implicated in our dataset. For REACTOME, the analysis revealed significant enrichment in several key pathways, including translation, protein methylation, Mitochondrial tRNA Animoacylation, Metabolism of Proteins, and Telomere Extension by Telomerease with p-values below 0.05. For instance, the translation pathway was highly enriched with a p-value of 0.008216705, suggesting its potential role in the translation process of protein synthesis. For GO, the analysis revealed significant enrichment in several key pathways, including Regulation of Chaperone-Mediated Autophagy, Fructose 1,6-Bisphosphate Metabolic Process, Fructose 6-Phosphate Metabolic Process, Fatty Acid Beta-Oxidation Using acyl-CoA Dehydrogeerase, and Regulation of Lipid Kinase Activity. The REACTOME results reveal that there exist enrichment changes in telomere extension pathways, which are consistent with current understanding, suggesting that telomere extension pathways are involved in replicative senescence. The GO results reveal that there exist enrichment changes in autophagy pathways, which is supported by studies that show autophagy promotes cellular senescence. [20] These findings provide valuable direction for further investigations into the underlying mechanisms of telomere extension and autophagy.

3.4. Knocking down key mediator genes promotes cellular senescence

We then experimentally validated whether disrupting the key mediator genes of the causal methylation site can affect the cellular senescence process. We knocked down three top mediator genes, namely *CISD2*, *MARS2*, and *PKFM*, in the MRC-5 Human Embryonic Lung Fibroblast Cells using siRNA. We show that the siRNA we designed can successfully knock down the target gene by up to 73% (Fig. 4a). The SA-β-gal staining shows that the cells with key mediating genes knocked down exhibit early cellular senescence phenotype, similar to the positive control with 10uM atazanavir (Fig. 4b). These results suggest that the key mediator genes we identified are essential for cellular maintenance and disrupting them will induce cellular senescence.

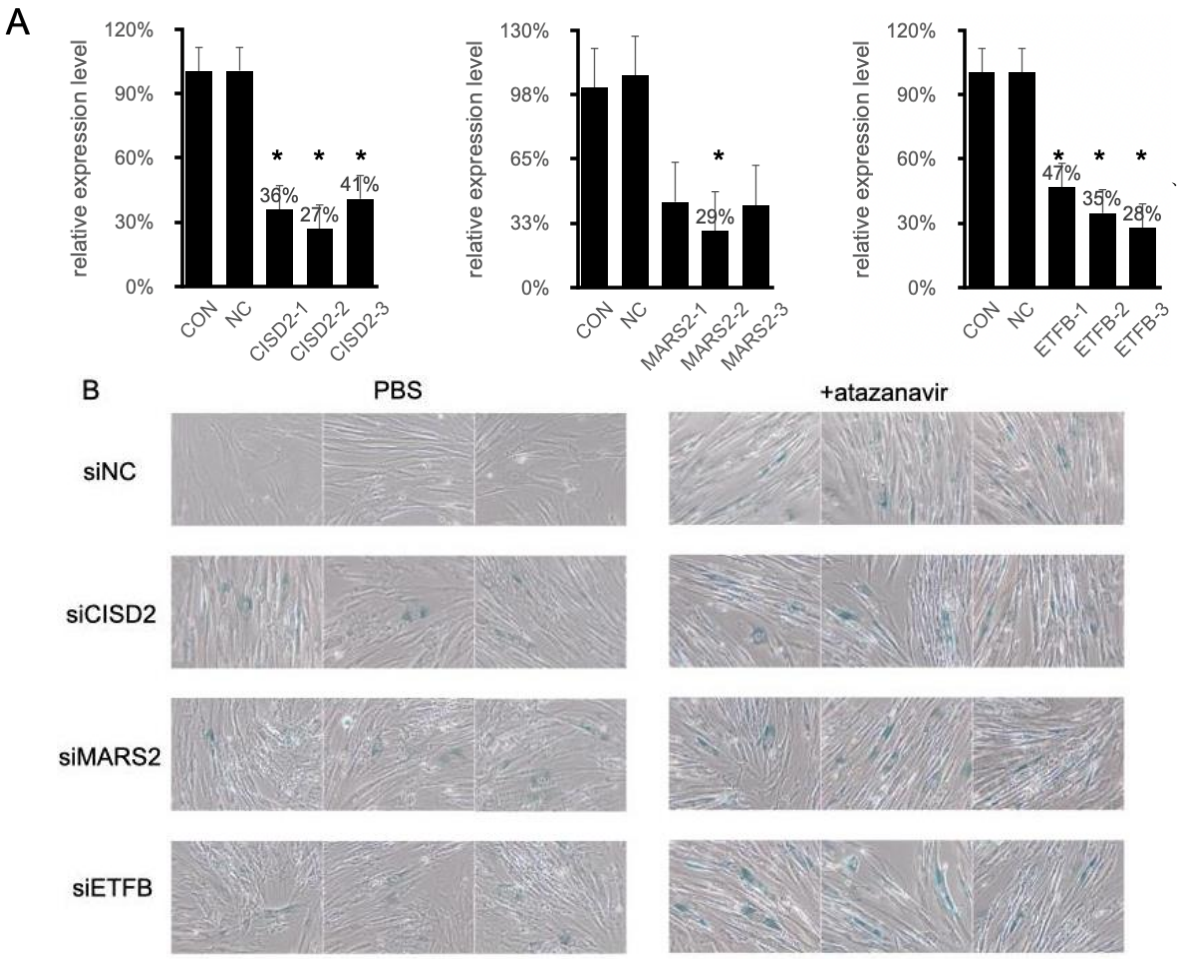

Figure 4. Knocking down key mediator genes induces early cellular senescence A. The Bar graph shows the relative expression change of CISD2, MARS2, and ETFB after knocking down in human fibroblast cells. Error bars indicate SD (n=2). B. Senescence-associated-β-galactosidase (SA-β-gal) staining experiment. The experiment shows that knocking down genes with protective CpG sites induces cellular senescence.

4. Discussion and Conclusion

During senescence, DNA methylation alterations are a key feature, and these changes are distinct from those associated with malignant transformation [21]. Replicative senescence is associated with reproducible DNA hypermethylation changes at specific sites in the genome [22, 23]. Epigenetic alterations to DNA and chromatin are crucial in mediating the induction and maintenance of senescence [24]. During aging, CpG sites are differentially methylated. Some CpG sites are protective, meaning that methylation at these sites is associated with decreased risk of disease, while other CpG sites are damaging, meaning that methylation at these sites is associated with increased risk of disease. It is crucial that protective and damaging CpG sites are distinguished as they could have many implications in the research and medical field. By identifying protective CpG sites, researchers can develop therapies that increase methylation at these sites to decrease the risk of disease. By identifying damaging CpG sites, researchers can develop diagnostic tools that detect methylation at these sites as an early warning sign of disease. Researchers can gain insights into the specific genes and pathways that are involved in disease development by studying the effects of methylation at different CpG sites.

The *CISD2* gene is associated with Wolfram syndrome, a progressive neurodegenerative disease characterized by diabetes insipidus, diabetes mellitus, optic atrophy, and deafness [25]. While genetic studies in Drosophila and humans suggest that *CISD2* might function in concert with *PPT1, CLN3*, and several other genes under normal or possibly pathological states, the function of *CISD2* has yet

to be fully resolved[25]. *CISD2* could serve as a target to elucidate the correlation between deafness and aging.

While we utilized the Illumina 450K methylation array's DNA meQTLs, which included 420, 509 CpG sites, there still exists a number of unmeasured CpG sites. The large number of unmeasured CpG sites can limit the accuracy due to the high correlation between nearby CpG sites and methylation patterns. Also, CpG sites were primarily chosen from a large meQTL study in whole blood, which included 36 cohorts and 27, 750 European subjects. As a result, the causal CpG sites identified are primarily valid in blood while up to 73% of cis-meQTLs are shared across tissues, suggesting that the identified causal CpG sites may also act in other tissues to affect lifespan and health span [26]. Future large-scale meQTL studies across tissues may facilitate the identification of tissue-specific epigenetic effects on aging.

Acknowledgements

I would like to express my sincere gratitude to the following individuals who have contributed to the success of this article: Dr. Jianjun Wu, for providing valuable insights and support throughout the experiment and writing process, Dr. Albert Ying, for sharing his expertise and knowledge on the topic. Additionally, I would like to acknowledge all the individuals who have encouraged me to pursue this research, especially my dad, my mom, and my sister.

References

[1] Cellular Senescence - an overview | ScienceDirect Topics. https://www.sciencedirect.com/topics/medicine-and-dentistry/cellular-senescence.

[2] López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The Hallmarks of Aging. *Cell* 153, 1194 – 1217 (2013).

[3] Arisa Kita, Yuki Saito, Norihiro Miura, Maki Miyajima, Sena Yamamoto, Tsukasa Sato, Takatoshi Yotsuyanagi, Mineko Fujimiya & Takako S. Chikenji. Altered regulation of mesenchymal cell senescence in adipose tissue promotes pathological changes associated with diabetic wound healing | Communications Biology. https://www.nature.com/articles/s42003-022 - 03266 - 3.

[4] Campisi, J. Aging, Cellular Senescence, and Cancer. *Annu. Rev. Physiol.* 75, 685 – 705 (2013).

[5] Jin, B., Li, Y. & Robertson, K. D. DNA Methylation. *Genes Cancer* 2, 607 – 617 (2011).

[6] Timmers, P. R. H. J. *et al.* Mendelian randomization of genetically independent aging phenotypes identifies LPA and VCAM1 as biological targets for human aging. *Nat. Aging* 2, 19–30 (2022).

[7] Davies, N. M., Holmes, M. V. & Davey Smith, G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ* k601 (2018) doi:10.1136/bmj. k601.

[8] Burgess, S. & Labrecque, J. A. Mendelian randomization with a binary exposure variable: interpretation and presentation of causal estimates. *Eur. J. Epidemiol.* 33, 947 – 952 (2018).

[9] Ying, K. *et al.* Causal Epigenetic Age Uncouples Damage and Adaptation. 2022.10.07.511382 Preprint at https://doi.org/10.1101/2022.10.07.511382 (2022).

[10] Gutierrez-Arcelus, M. *et al.* Passive and active DNA methylation and the interplay with genetic variation in gene regulation. *eLife* 2, e00523 (2013).

[11] Weidner, C. I. *et al.* Aging of blood can be tracked by DNA methylation changes at just three CpG sites. *Genome Biol.* 15, R24 (2014).

[12] Carmen M. Koch, Sylvia Joussen, Anne Schellenberg, Qiong Lin, Martin Zenke, Wolfgang Wagner. Monitoring of cellular senescence by DNA-methylation at specific CpG sites - Koch - 2012 - Aging Cell - Wiley Online Library. https: //onlinelibrary.wiley.com/doi/full/10.1111/j.1474-9726.2011.00784.x.

[13] Hillary, R. F. *et al.* Blood-based epigenome-wide analyses on the prevalence and incidence of nineteen common disease states. 2023.01.10.23284387 Preprint at https://doi.org/10.1101/2023.01.10.23284387 (2023).

[14] Nones, K. *et al.* Genome-wide DNA methylation patterns in pancreatic ductal adenocarcinoma reveal epigenetic deregulation of SLIT-ROBO, ITGA2 and MET signaling. *Int. J. Cancer* 135, 1110 – 1118 (2014).

[15] Wozniak, M. B. *et al.* Integrative genome-wide gene expression profiling of clear cell renal cell carcinoma in Czech Republic and in the United States. *PloS One* 8, e57886 (2013).

[16] Shen, Z.-Q. *et al.* CISD2 maintains cellular homeostasis. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* 1868, 118954 (2021).

[17] Zhou, Z., Sun, B., Yu, D. & Bian, M. Roles of tRNA metabolism in aging and lifespan. *Cell Death Dis.* 12, 548 (2021).

[18] PFKM gene: MedlinePlus Genetics. https://medlineplus.gov/genetics/gene/pfkm/.

[19] ETFB gene: MedlinePlus Genetics. https://medlineplus.gov/genetics/gene/etfb/.

[20] Kwon, Y., Kim, J. W., Jeoung, J. A., Kim, M.-S. & Kang, C. Autophagy Is Pro-Senescence When Seen in Close-Up, but Anti-Senescence in Long-Shot. *Mol. Cells* 40, 607 – 612 (2017).

[21] Xie, W. *et al.* Abstract B20: Malignant transformation initiates a stochastic DNA methylation alteration pattern distinct from that in senescence. *Cancer Res.* 76, B20–B20 (2016).

[22] Hänzelmann, S. *et al.* Replicative senescence is associated with nuclear reorganization and with DNA methylation at specific transcription factor binding sites. *Clin. Epigenetics* 7, 19 (2015).

[23] 23. Franzen, J. *et al.* DNA Methylation Changes Upon Senescence are Strand-Specific and Reflect Chromatin Conformation. *SSRN Electron. J.* (2019) doi:10.2139/ssrn.3312726.

[24] Crouch, J., Shvedova, M., Thanapaul, R. J. R. S., Botchkarev, V. & Roh, D. Epigenetic Regulation of Cellular Senescence. *Cells* 11, 672 (2022).

[25] Jones, M. A. *et al.* Genetic studies in Drosophila and humans support a model for the concerted function of CISD2, PPT1 and CLN3 in disease. *Biol. Open* 3, 342 – 352 (2014).

[26] Lin, D. *et al.* Characterization of cross-tissue genetic-epigenetic effects and their patterns in schizophrenia. *Genome Med.* 10, 13 (2018).

[27] Abdi, H. & Williams, L. J. Principal component analysis. *WIREs Comput. Stat.* 2, 433 – 459 (2010).

[28] Burgess, S., Zuber, V., Valdes-Marquez, E., Sun, B. B. & Hopewell, J. C. Mendelian randomization with fine-mapped genetic data: Choosing from large numbers of correlated instrumental variables. *Genet. Epidemiol.* 41, 714 – 725 (2017).

[29] A proteomic atlas of senescence-associated secretomes for aging biomarker development - PubMed. https://pubmed.ncbi.nlm.nih.gov/31945054/.