Promoter methylation of *PGC1A* and *PGC1B* predicts cancer incidence in a veteran cohort

Jacob K Kresovich^{1,2}, Brian T Joyce^{1,2}, Tao Gao¹, Yinan Zheng¹, Zhou Zhang¹, Christopher J Achenbach^{2,3}, Robert L Murphy^{2,3}, Allan C Just⁴, Jincheng Shen⁵, Hushan Yang⁶, Pantel Vokonas⁷, Joel Schwartz⁸, Andrea A Baccarelli⁹ & Lifang Hou^{*,1,2}

¹Center for Population Epigenetics, Robert H Lurie Comprehensive Cancer Center & Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA

²Robert H Lurie Comprehensive Cancer Center, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA

³Center for Global Health, Institute for Public Health & Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA

⁴Department of Preventive Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

⁵Department of Population Health Sciences, University of Utah School of Medicine, Salt Lake City, UT 84108, USA

⁶Division of Population Science, Department of Medical Oncology, Sidney Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, PA 19107, USA

⁷VA Normative Aging Study, Veterans Affairs Boston Healthcare System & the Department of Medicine, Boston University School of Medicine, Boston, MA 02118, USA

⁸Department of Environmental Health, Harvard TH Chan School of Public Health, Harvard University, Boston, MA 02115, USA

⁹Departments of Epidemiology & Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY 10032, USA

*Author for correspondence: l-hou@northwestern.edu

Aim: Previous studies suggest telomere shortening represses *PGC1A* and *PGC1B* expression leading to mitochondrial dysfunction. Methylation of CpG sites within these genes may interact with these factors to affect cancer risk. **Materials & methods:** Among 385 men, we identified 84 incidents of cancers (predominantly prostate and nonmelanoma skin). We examined associations between leukocyte DNA methylation of 41 CpGs from *PGC1A* and *PGC1B* with telomere length, mitochondrial 8-OHdG lesions, mitochondrial abundance and cancer incidence. **Results:** Methylation of five and eight CpG sites were significantly associated with telomere length and mitochondrial abundance at $p < 0.05$. Two CpG sites were independently associated with cancer risk: cg27514608 (*PGC1A*, TSS1500; HR: 1.55, 95% CI: 1.19–2.03, FDR = 0.02), and cg15219393 (*PGC1B*, first exon/5'UTR; HR: 3.71, 95% CI: 1.82–7.58, FDR < 0.01). Associations with cg15219393 were observed primarily among men with shorter leukocyte telomeres. **Conclusion:** *PGC1A* and *PGC1B* methylation may serve as early biomarkers of cancer risk.

First draft submitted: 24 October 2017; Accepted for publication: 7 February 2018; Published online: 11 June 2018

Keywords: cancer biomarkers • cancer epigenetics • DNA methylation

Background

Cancer is an age-associated disease, primarily affecting elderly individuals; nearly half of all malignancies occur in individuals aged over 70 years [1]. Prostate and skin cancers are two of the five most commonly diagnosed tumors among males, with particularly pronounced incidences among older individuals [2]. While carcinogenesis is multifaceted, exploration of biological pathways associated with aging processes may elucidate common events in tumor formation.

Aging and cancer development have common biological foundations [3]. Aging is marked by increases in cellular oxidative stress which may reduce mitochondrial function [4–6] and induce telomere shortening [7–10]. Mitochondrial dysfunction may influence cancer risk by facilitating the generation of reactive oxygen species resulting in increased cellular oxidative stress; the feedback cycle between oxidative stress and mitochondrial dysfunction is believed to lead to cellular alterations that increase cancer risk [11–14]. Telomere shortening is additionally hypothesized to be a marker of aging [15] and may play an important role in the progression to cancer [16–19]. As such, telomere shortening and

mitochondrial dysfunction have been independently observed to be dysregulated during the formation of various tumor types.

Epigenetic regulation of *PGC1A* and *PGC1B* may be a possible link between telomere shortening and mitochondrial dysfunction [20,21]. Telomere reverse transcriptase (*TE RT*) and telomerase RNA component (*TERC*) knockout mice serve as a model to examine the effects of telomere shortening on various physiologic functions. Using this model, researchers observed profound repression of *PGC1A* and *PGC1B* expression leading to subsequent mitochondrial dysfunction; moreover, these functions were restored after enforced *TERC* or *PGC1A* expression [20].

Telomere shortening and mitochondrial dysfunction may interact with epigenetic repression of *PGC1A* and *PGC1B* leading to increased cancer risk. However, previous investigations examining association with cancer incidence remain largely limited to individual components rather than the interrelations of these factors. We aimed to test whether DNA methylation of *PGC1A* and *PGC1B* link telomere shortening with mitochondrial function in humans and if methylation of these genes can serve as a biomarker of future cancer development. We further examined potential modification of *PGC1A* and *PGC1B* methylation associations with cancer development by telomere length and mitochondrial function.

Materials & methods

Sample characteristics

The US Department of Veteran Affairs established the Normative Aging Study (NAS) in 1963 with a total of 2280 healthy men initially enrolled [22]. The NAS is a research component of the Massachusetts Veterans Epidemiology Research and Information Center. Eligibility criteria included veteran status, Boston metro area residence, aged 21–80 years and no history of cancer or other chronic conditions. Participants were recalled every 3–5 years for clinical examinations where they were evaluated for anthropometric measurements using standardized medical examinations and responded to questionnaires about medical history and lifestyle. Information on age, smoking status and body mass index (BMI) were collected at each follow-up examination. Starting in 1999, follow-up examinations included drawing a 7-mL blood sample for epigenetic analyses; of the 829 remaining participants, 802 (97%) consented to these blood donations. We restricted our study to white participants as nonwhite individuals made up fewer than 2% of the total sample. The current study focuses on 541 observations from 385 participants who were cancer free at the first blood draw and had information on telomere length, methylation of *PGC1A* and *PGC1B*, and measures of mitochondrial function. Overall, 230 (60%) participants had one medical visit prior to cancer diagnosis or censorship, 154 (40%) had two medical visits and one participant had three medical visits. A flow chart of the participant selection criteria beginning at first blood draw is shown in Supplementary Figure 1. The NAS was conducted per guidelines laid down in the Declaration of Helsinki and the Institutional Review Board of Harvard University and all participating institutions approved all procedures involving human subjects. All study participants provided informed written consent.

Telomere length, *PGC1A* & *PGC1B* methylation & markers of mitochondrial function

Telomere length, DNA methylation of *PGC1A* and *PGC1B*, and mitochondrial functional markers were measured in circulating blood leukocytes.

To determine telomere length, we used quantitative real-time polymerase chain reactions (qPCR) [23]. Relative telomere length was estimated by taking the ratio of the telomere (T) repeat copy number to a single-copy gene (S) copy number (T:S ratio) in a given sample and reported in relative units expressing the ratio between test DNA telomere length and reference pooled telomere length. The reference data were created using DNA from 475 randomly selected participants and used to generate a standard curve used in every T and S qPCR run. A detailed explanation of these methods is described elsewhere [24]. The intra-assay coefficient of variation for the T/S ratio was 8.1%. The average coefficient of variation for the T reaction was 8%, and for the S reaction was 5.6%. Measures were repeated when the coefficient of variation for the T or S reactions was greater than 15% [25].

PGC1A and *PGC1B* methylation was quantified using the Illumina HumanMethylation450 BeadChip. We selected a total of 41 CpG sites located upstream of the gene transcription start site (TSS) through the 3'-untranslated region (UTR) as determined by Illumina annotation (18 sites for *PGC1A* and 23 sites for *PGC1B*). Additional information on chromosomal position and gene region location can be found in Supplementary Table 1. Most of the tested sites were near the promoter region (TSS1500, TSS200 or first exon/5' UTR) as designed by the assay manufacturer (6 of 18 for *PGC1A*, 33%; 7 of 23 for *PGC1B*, 30%). CpG site methylation was converted from β -values to M-values via logit transformation. β -values derived from highly methylated or unmethylated sites

have severe statistical heteroscedasticity due to value constraints between zero and one. A logit transformation of the β -values to M-values reduces this heteroscedasticity. While we recognize the more direct biological interpretation of using β -values, we elected to use M-values in statistical analysis in order to minimize heteroscedasticity in regression models which allows for a more precise and valid measurement of associations [26].

Mitochondrial function was estimated by measuring mitochondrial DNA (mtDNA) copy number abundance (i.e., greater mtDNA abundance is hypothesized to represent impaired mitochondrial function) and mtDNA 8-hydroxy-2'-deoxyguanosine (8-OHdG) lesions caused by oxidative stress [27]. Briefly, mitochondrial abundance was measured using a well-established protocol which examines the ratio of mtDNA copy number to nuclear copy number using qPCR [28–30]. To measure mitochondrial abundance, we used the mtDNA 12S ribosomal ribonucleic acid TaqMan (Applied Biosystems, MA, USA) probe (6FAM-5' TGCCAGCCACCGCG 3'-MGB). The sequences of primers used for amplification in mtDNA were mtF805 (5' CCACGGGAAACAGCAGTGATT 3') and mtR927 (5' CTATTGACTTGGGTTAATCGTGTGA 3'). Samples were run in triplicate. Detailed laboratory methodology for mtDNA copy number quantification can be found elsewhere [27]; the intra- and interassay coefficients of variation were 3.4 and 3.3%.

Mitochondrial DNA 8-OHdG lesions were measured using a qPCR method previously described elsewhere [31,32]. Briefly, when mitochondrial DNA contains 8-OHdG, qPCR efficiency decreases after exogenous digestion with 8-oxoguanine DNA glycosylase. Therefore, the degree of 8-OHdG levels in mtDNA is presented as the ratio of the starting quantity of treated sample over the starting quantity of untreated sample. Prior to qPCR, 10 μ l of mixture with and without 8-oxoguanine DNA glycosylase (New England Biolabs, MA, USA) were added to samples containing 40 ng mtDNA, 1 $\times$ NE buffer (50 mM NaCl, 10 mM Tris-HCl, 10 mM MgCl₂, 1 mM DTT, pH 7.9) and 100 μ g/ml of bovine serum albumin. Samples were incubated at 37°C for 1 h followed by 65°C for 20 min. qPCR was applied for determination of the starting quantity of treated and untreated DNA samples using iQ SYBR Green Supermix (Bio-Rad, CA, USA). The primers used for qPCR for mtDNA were as follows: mtF3212 (5' CACCCAAGAACAGGGTTTGT3') and mtR3319 (5' TGGCCATGGGTATGTTGTTA3'). qPCR conditions were 50°C for 2 min and 95°C for 13 min, followed by 40 cycles of 15 s at 95°C for denaturation and 10 min of annealing at 62°C. qPCR was performed using Bio-Rad CFX384 Touch™ Real-Time PCR Detection System (Bio-Rad). All samples were run in triplicates.

Identification of cancer cases

Cancer diagnosis information was obtained through self-report and confirmed via review of medical records. Of the 385 participants, 84 were diagnosed with cancer after their first blood draw allowing for the assessment of prediagnostic measurements for patients with a subsequent cancer diagnosis. Cancer cases were followed-up for an average of 4.2 years between first blood draw and cancer diagnosis; the participants were generally diagnosed with tumors affecting older individuals, namely skin (33 cases: six melanomas, 27 unspecified; 9% of total sample) and prostate (22; 6%), as well as other tumors from the following systems: respiratory (8; 2%), gastrointestinal (8; 2%) and unclassified (13; 3%). Participants who remained cancer-free for the full duration of the study were censored at 31 December 2013 (average cancer-free follow-up between first blood draw and censorship was 8.7 years).

Statistical analysis

We first examined associations between single site *PGC1A* and *PGC1B* methylation and telomere length and markers of mitochondrial function (mitochondrial abundance and mtDNA 8-OHdG lesions). Using mixed-effect linear regression models with random intercepts to account for repeated measures, we treated telomere length as an independent variable and methylation at each CpG site as separate dependent variables to examine associations of telomere length on *PGC1A* and *PGC1B* CpG methylation. We dichotomized telomere length about the median value of all measurements to create equal sized groups representing longer and shorter relative telomere lengths. We then examined relationships between *PGC1A* and *PGC1B* CpG methylation and logit-transformed markers of mitochondrial function by applying mixed-effect linear regression models with random intercepts, treating methylation at each CpG site as an independent variable (in separate models) and logit-transformed mitochondrial abundance or mtDNA 8-OHdG lesions as the dependent variable. We report associations with $p \leq 0.05$, although we additionally calculated the Benjamini-Hochberg false discovery rate (FDR) to examine the effects of multiple testing [33].

To explore relationships between methylation at the 41 measured *PGC1A* and *PGC1B* CpGs and cancer incidence, we employed Cox regression models with discrete time-varying covariates to estimate hazard ratios

Table 1. Sample characteristics at first blood draw by cancer development

Characteristic	Overall sample (n = 385)		No cancer (n = 301)		Developed cancer (n = 84)		p-value [†]
	N	%	n	%	n	%	
Age:							0.17
– <70	137	36	103	34	34	41	
– 70–74	109	28	82	27	27	32	
– 75+	139	36	116	39	23	27	
Body mass index:							0.47
– Under/normal (<25)	74	19	55	18	19	23	
– Overweight (25–29.9)	216	56	168	56	48	57	
– Obese (30+)	95	25	78	26	17	20	
Smoking status:							0.84
– Never	110	29	85	28	25	30	
– Former	256	66	202	67	54	64	
– Current	19	5	14	5	5	6	
Telomere length (median [IQR])	1.21	0.9–1.5	1.21	1.0–1.5	1.20	0.9–1.5	0.51 [‡]
mtDNA 8OH-dG (median [IQR])	0.69	0.6–0.8	0.70	0.6–0.8	0.69	0.6–0.8	0.53 [‡]
mtDNA abundance (median [IQR])	1.01	0.9–1.2	1.01	0.9–1.2	0.99	0.9–1.2	0.53 [‡]

[†]Derived from χ^2 test.[‡]Derived from Wilcoxon-rank sum test (nonparametric test for differences).

(HRs). We considered statistical significance using a traditional cut-point of $FDR(q) \leq 0.05$ to account for multiple testing. We then examined potential interaction of the CpG site-cancer incidence relationship by telomere length and mitochondrial functional markers; interaction was determined by the inclusion of an interaction term in the model using a $p \leq 0.10$ cut-point. We further explored associations between logit-transformed mitochondrial functional markers (mitochondrial abundance and mtDNA 8-OHdG lesions) and cancer incidence overall and stratified by telomere length using the $p \leq 0.05$ threshold. All statistical models were adjusted for age, smoking status and BMI as well as estimated blood cell type composition and processing batch. Due to the small sample size, we were unable to include additional adjustment variables and therefore focused on the most relevant potential confounders. Because cancer is epigenetically heterogeneous, we additionally conducted a sensitivity analyses examining *PGC1A* and *PGC1B* CpG methylation associations with cancer risk stratified by tumor type (skin and prostate only due to sample size constraints). All analyses were conducted using Stata version 14.2 (Stata Corp., TX, USA).

Results

At first blood draw, the sample participants' mean age was 72.3 (standard deviation: 6.7). Participants were generally overweight or obese (81%) and predominately former smokers (66%) (Table 1). There were no significant differences in these characteristics between those who went on to be diagnosed with cancer and those who did not. Moreover, at first blood draw prior to any cancer diagnosis, no significant differences existed in telomere length ($p = 0.51$), mtDNA 8-OHdG lesions ($p = 0.53$) or mitochondrial abundance ($p = 0.53$).

In total, methylation at 13 (32%) distinct *PGC1A* and *PGC1B* CpG sites were associated with either telomere length or mitochondrial abundance at $p \leq 0.05$ (Table 2). No CpG sites were associated with both telomere length and mitochondrial function. Five CpGs were associated with telomere length, although no sites reached statistical significance after accounting for multiple testing (all $q > 0.05$). Generally longer telomere length was associated with lower gene body methylation (*PGC1A*, cg07744449: $\beta = -0.11$, 95% CI: -0.20, -0.01; $q = 0.38$; *PGC1A*, cg24716152: $\beta = -0.07$, 95% CI: -0.13, -0.00; $q = 0.38$) and higher promoter methylation (*PGC1A*, cg20723350: $\beta = 0.07$, 95% CI: 0.01–0.12; $q = 0.38$; *PGC1B*, cg16786458: $\beta = 0.06$, 95% CI: 0.00–0.12; $q = 0.38$; *PGC1B*, cg20563759: $\beta = 0.07$, 95% CI: 0.00–0.14; $q = 0.38$). Only gene body methylation of *PGC1B* was associated with mitochondrial abundance after accounting for multiple testing; higher methylation of these sites was generally associated with greater mitochondrial abundance with one exception (cg26876242: $\beta = 0.08$, 95% CI: 0.04–0.13; $q = 0.01$; cg10500461: $\beta = -0.10$, 95% CI: -0.16, -0.05; $q = 0.01$; cg17645677: $\beta = 0.13$, 95% CI: 0.06–0.20; $q = 0.01$; cg18604199: $\beta = 0.07$, 95% CI: 0.03–0.10; $q = 0.01$). One CpG site from the 3'UTR region of *PGC1A* was additionally associated with increased mitochondrial abundance, but it did not reach significance after

Table 2. CpGs associated with telomere length and mitochondrial abundance at $p < 0.05$.

Illumina ID	β	95% CI	p-value	FDR	Gene	Region
Telomere length[†]						
cg20723350	0.07	0.01, 0.12	0.02	0.38	PGC1A	First exon; 5' UTR
cg07744449	-0.11	-0.20, -0.01	0.03	0.38	PGC1A	Body
cg24716152	-0.07	-0.13, -0.00	0.05	0.38	PGC1A	Body
cg16786458	0.06	0.00, 0.12	0.05	0.38	PGC1B	TSS1500
cg20563759	0.07	0.00, 0.14	0.05	0.38	PGC1B	TSS1500
Mitochondrial abundance[‡]						
cg26876242	0.08	0.04, 0.13	<0.01	0.01	PGC1B	Body
cg10500461	-0.10	-0.16, -0.05	<0.01	0.01	PGC1B	Body
cg17645677	0.13	0.06, 0.20	<0.01	0.01	PGC1B	Body
cg18604199	0.07	0.03, 0.10	<0.01	0.01	PGC1B	Body
cg06051582	-0.05	-0.09, -0.01	0.01	0.10	PGC1B	Body
cg07226964	-0.04	-0.08, -0.01	0.03	0.12	PGC1B	Body
cg08928958	0.10	0.02, 0.18	0.02	0.12	PGC1B	Body
cg06772578	0.04	0.00, 0.07	0.03	0.14	PGC1A	3' UTR

No *PGC1A* or *PGC1B* sites were associated with mitochondrial 8-OHdG at $p < 0.05$. All models adjusted for age, BMI, smoking status, blood cell composition and processing batch.

[†] Coefficients derived from models with telomere length serving as the independent variable.

[‡] Coefficients derived from models with individual *PGC1A* and *PGC1B* CpG methylation serving as the independent variables.

FDR: False discovery rate.

accounting for multiple testing (cg06772578: $\beta = 0.04$, 95% CI: 0.00–0.07; $q = 0.14$). Complete results exploring *PGC1A* and *PGC1B* CpG methylation associations with telomere length and mitochondrial functional markers can be found in Supplementary Tables 2–4.

Figures 1 & 2 (panels A & B) depict the genomic, RefSeq and tested CpG site locations for *PGC1A* and *PGC1B*. Panel C of the figures additionally shows the HR point estimates and 99% confidence intervals for the relationships between CpG site methylation and cancer incidence after adjustment for age, BMI, smoking status, blood cell composition and processing batch. We elected to show the 99% confidence interval to visually account for multiple testing. Solid confidence interval denote significance at $q \leq 0.05$. Figure 1 (panel C) shows one CpG site from the promoter proximal region (TSS1500) of *PGC1A* was significantly associated with cancer incidence (cg27514608: HR: 1.55, 95% CI: 1.19–2.03; $q = 0.01$). Figure 2 (panel C) shows that cg15219393, from the promoter proximal region of *PGC1B* (first exon/5' UTR), was significantly associated with cancer incidence (HR: 3.71, 95% CI: 1.82–7.58; $q = 0.003$). Results for all tested CpG loci are shown in Table 3. We generally identified the same strength and direction of associations for cg15219393 and cg27514608 in both prostate (cg15219393: HR: 4.43, 95% CI: 1.39–14.14, $q = 0.18$; cg27514608: HR: 1.72, 95% CI: 1.12–2.62; $q = 0.18$) and skin cancers (cg15219393: HR: 3.93, 95% CI: 0.98–15.74; $q = 0.51$; cg27514608: HR: 2.08, 95% CI: 1.22–3.56; $q = 0.29$), although these sites were not significantly associated with cancer incidence after accounting for multiple testing (Supplementary Table 5).

We further examined potential effect modification by telomere length and mitochondrial function on associations between *PGC1A* and *PGC1B* CpG site methylation and cancer incidence. Interaction was detected by telomere length for the association between cg15219393 methylation and cancer incidence (p -interaction = 0.096). Upon stratification by telomere length, cg15219393 was significantly associated with cancer incidence, but only among those with shorter telomeres (HR: 10.09, 95% CI: 3.56–28.63; $q = 0.003$) (Supplementary Table 6). When stratified by markers of mitochondrial function, we did not detect any significant statistical interaction with mitochondrial 8-OHdG lesions (cg27514608 p -interaction = 0.14, cg15219393 p -interaction = 0.60), although associations between cg27514608 and cancer incidence were mainly observed among those with higher amounts of 8-OHdG lesions (HR: 1.84, 95% CI: 1.26–2.67; $q = 0.04$) (Supplementary Table 7). We additionally detected significant statistical interaction by mitochondrial abundance for associations with cg24716152 and cancer incidence (p -interaction = 0.004). This novel site is located in the gene body of *PGC1A*; methylation of this CpG showed protective associations with cancer incidence among those with lower mitochondrial abundance (HR: 0.22, 95% CI: 0.09–0.56; $q = 0.03$) (Supplementary Table 8).

Figure 1. PGC1A methylation associations with cancer risk. Genomic (A), RefSeq and tested CpG site locations for *PGC1A* (B). Arrows denote direction of transcription (right to left). (C) shows the resulting hazard point estimates and 99% confidence intervals for each individual CpG site. Dotted confidence intervals represent $FDR > 0.05$. FDR: False discovery rate.

Finally, we examined associations between mitochondrial functional markers and cancer incidence overall and stratified by telomere length. Mitochondrial abundance was not independently associated with cancer incidence in either situation, although notably, a higher number of mtDNA 8-OHdG lesions was inversely associated with cancer incidence (HR: 0.28, 95% CI: 0.07–1.02; $p = 0.06$), particularly among those with longer telomeres (p -interaction = 0.087, HR: 0.05, 95% CI: 0.01–0.35; $p = 0.003$) (Supplementary Table 9).

Discussion

This is the first study to examine a possible role of DNA methylation of *PGC1A* and *PGC1B* in relation to telomere length, mitochondrial function and cancer risk. While we identified associations between *PGC1B* gene body methylation and mitochondrial abundance, we did not observe methylation at any CpG site which had overlapping associations with both telomere length and mitochondrial function. We additionally observed that higher methylation at cg15219393 and cg27514608, from the promoter proximal regions of *PGC1B* and *PGC1A*, was associated with increased risk of future cancer diagnosis. Moreover, we showed that telomere length modified the relationship with the *PGC1B* site, such that the effects were mainly observed among those with shorter telomeres.

Telomere shortening has been linked to mitochondrial function through repression of *PGC1A* and *PGC1B* [20,21]. *In vitro* and *in vivo* models found telomere shortening represses expression of these genes and leads to mitochondrial functional decline [34,35]. We found little evidence to support our hypothesis that epigenetic regulation of these genes may be associated with these cellular factors. Transcriptional control of gene expression is a complex process involving many biological mechanisms including heterochromatin, histone modifications and other epigenetic factors as well as activation of transcription factors [36]. It is possible that repression of *PGC1A* and *PGC1B* is primarily controlled by activation of the transcription factor p53, rather than DNA methylation [37]. This potential explanation is supported by the finding that *TER T* and *TERC* knockout mice show activation of p53 and repression of *PGC1A* and *PGC1B* as well as other p53 targets such as *ERRα*, *PPARγ* and *TFAM* [20]. As data on p53 expression were unavailable in our source cohort, we are unable to test this hypothesis, which suggests an important direction for future research.

Figure 2. PGC1B methylation associations with cancer risk. Genomic (A), RefSeq and tested CpG site locations for *PGC1B* (B). Arrows denote direction of transcription (left to right). (C) shows the resulting hazard point estimates and 99% confidence intervals for each individual CpG site. Dotted confidence intervals represent FDR > 0.05. FDR: False discovery rate.

To our knowledge, this is the first study to identify strong, positive associations between *PGC1A* and *PGC1B* proximal promoter methylation and future cancer risk. Previous studies have shown mutations in these genes affect the development of common cancers such as breast, ovarian and colorectal [38–42]. Specifically, prior evidence showed that single nucleotide polymorphisms located in the 5'UTR of *PGC1A* were associated with an increased risk of epithelial ovarian tumor formation [38], while other studies identified cancer associations with single nucleotide polymorphisms in nonpromoter gene regions [39,40] or modification of cancer risk by *PGC1A* or *PGC1B* genotype [41,42]. This study builds upon this literature by showing that epigenetic regulation of the promoter regions of these genes may additionally play a role in cancer development.

We also examined the relationships between methylation of *PGC1A* and *PGC1B* with cancer incidence in the context of telomere length and mitochondrial function. We found proximal promoter methylation of *PGC1B* among participants with shorter telomeres was associated with increased cancer risk. We additionally observed higher methylation at the promoter of *PGC1A* was associated with increased cancer risk among participants with higher mtDNA 8-OHdG lesions. Similarly, gene body methylation of *PGC1A* among participants with lower mitochondrial abundance was associated with decreased cancer risk. These findings align when considering the region of DNA methylation; gene body methylation is commonly associated with increased gene expression while promoter methylation is generally associated with reduced expression [43]. Previous research has identified inconsistent findings regarding the role of telomere length and cancer risk; studies have identified associations with cancer risk for shorter [16–19] as well as longer telomeres [44–47]. These discordant findings could be a result of interactions between telomere length and epigenetic regulation of *PGC1A* and *PGC1B*. Moving forward, researchers investigating the role of telomere length on cancer risk may consider the contexts of *PGC1A* and *PGC1B* gene methylation as our findings suggest these factors modify these relationships.

While this is the first study to identify interactions between telomere length and *PGC1A* and *PGC1B* gene regulation with cancer risk, it is subject to limitations. Most importantly, we did not have *PGC1A* or *PGC1B* expression data available. We therefore cannot confirm promoter DNA methylation is associated with decreased gene expression, although evidence from the ENCODE database and studies examining other genes suggests these

Table 3. Associations between PGC1A and PGC1B CpG sites and cancer incidence.

Illumina ID	HR	95% CI	p-value	FDR	Gene	Region
cg15219393	3.71	1.82–7.58	<0.01	<0.01	PGC1B	First exon; 5'UTR
cg27514608	1.55	1.19–2.03	<0.01	0.02	PGC1A	TSS1500
cg09427718	0.46	0.20–0.79	0.01	0.07	PGC1A	Body
cg18604199	0.53	0.31–0.89	0.02	0.17	PGC1B	Body
cg10955995	0.58	0.34–1.00	0.05	0.34	PGC1A	Body
cg09814231	0.43	0.19–1.00	0.05	0.34	PGC1B	Body
cg16786458	0.60	0.35–1.02	0.06	0.35	PGC1B	TSS1500
cg25226657	0.59	0.32–1.09	0.09	0.48	PGC1B	Body
cg04893087	0.74	0.48–1.13	0.16	0.58	PGC1A	TSS1500
cg07744449	1.33	0.89–2.00	0.17	0.58	PGC1A	Body
cg20723350	0.77	0.53–1.12	0.17	0.58	PGC1A	First exon; 5'UTR
cg17645677	0.48	0.16–1.40	0.18	0.58	PGC1B	Body
cg11270806	1.21	0.91–1.62	0.19	0.58	PGC1A	TSS1500
cg27182172	0.61	0.28–1.30	0.20	0.58	PGC1A	Body
cg27461259	1.18	0.90–1.56	0.23	0.59	PGC1A	TSS1500
cg12691631	0.74	0.45–1.22	0.23	0.59	PGC1A	TSS200
cg06772578	0.75	0.47–1.22	0.25	0.59	PGC1A	3'UTR
cg21184132	0.74	0.40–1.40	0.36	0.81	PGC1B	TSS1500
cg15541091	1.23	0.76–1.98	0.40	0.82	PGC1A	Body
cg05158538	0.81	0.49–1.35	0.42	0.82	PGC1A	Body
cg20563759	0.83	0.52–1.32	0.42	0.82	PGC1B	TSS1500
cg26876242	0.81	0.47–1.39	0.44	0.82	PGC1B	Body
cg20534662	0.82	0.48–1.40	0.47	0.82	PGC1B	Body
cg08550435	1.19	0.71–1.97	0.51	0.82	PGC1A	Body
cg04681472	0.84	0.49–1.44	0.53	0.82	PGC1B	3'UTR
cg24716152	0.83	0.46–1.50	0.54	0.82	PGC1A	Body
cg18547262	1.19	0.66–2.14	0.56	0.82	PGC1B	TSS200
cg15049370	1.18	0.68–2.06	0.56	0.82	PGC1B	Body
cg02896941	0.91	0.64–1.30	0.61	0.82	PGC1A	Body
cg00966557	1.19	0.58–2.44	0.63	0.82	PGC1B	Body
cg07226964	0.88	0.51–1.51	0.63	0.82	PGC1B	Body
cg08928958	0.76	0.23–2.49	0.65	0.82	PGC1B	Body
cg19701879	0.93	0.67–1.29	0.66	0.82	PGC1A	Body
cg15809858	0.87	0.42–1.79	0.70	0.85	PGC1B	Body
cg23459643	0.93	0.59–1.48	0.77	0.90	PGC1B	TSS1500
cg16344026	0.93	0.47–1.84	0.84	0.93	PGC1B	First exon
cg09082664	1.06	0.54–2.09	0.86	0.93	PGC1A	Body
cg06051582	0.96	0.52–1.74	0.88	0.93	PGC1B	Body
cg10500461	0.94	0.41–2.15	0.89	0.93	PGC1B	Body
cg22236469	0.96	0.43–2.11	0.91	0.94	PGC1B	Body
cg13593479	1.01	0.54–1.89	0.98	0.98	PGC1B	Body

associations hold true [48–50]. Another limitation was that our cohort was comprised of a mix of tumor types; while a majority of the tumors were either skin or prostate, a considerable proportion were found in other tissues. As mentioned previously, upon stratification by tumor location, we identified the same strength and direction of association between cg15219393 and cg27514608 methylation with skin and prostate cancer risk, albeit with less precision due to the decreased sample size. This leads us to believe our findings are not specific to one tumor type, rather DNA methylation of these loci is associated with cancer risk in general. Notably, we did not identify any associations between telomere length or markers of mitochondrial function with cancer status at the time of first blood draw. As these factors were measured in circulating blood leukocytes over 4 years prior to any cancer diagnosis,

it is not implausible that we did not observe signals for these factors in our small study population. Similarly, we did not have tumor tissue available; we therefore used blood leukocyte methylation as a surrogate marker of tumor DNA methylation. Blood leukocyte DNA methylation, telomere length and mitochondrial function were used as proxies for tumor characteristics due to the less invasive collection procedures and sample availability. Therefore, at a minimum, blood leukocyte DNA methylation of *PGC1A* and *PGC1B* may serve as a biomarker of cancer risk – additional studies using tumor tissue are needed to examine whether these changes are truly mechanistic. Additionally, due to the limited sample size, we may be overfitting our models such that we were predicting random error; we used the simplest models available that would balance adjusting for confounding and avoiding overfitting. Finally, our study participants were elderly, male veterans; we were therefore unable to examine modification of associations by participant characteristics. Future confirmation of our findings in more diverse populations are required.

Conclusion & future perspective

To our knowledge, we are the first to show that methylation of CpG sites within the promoter region of *PGC1A* and *PGC1B* was associated with cancer incidence. Therefore, blood leukocyte DNA methylation of these genes may serve as a biomarker of cancer risk; collection of blood leukocyte DNA methylation is relatively noninvasive and could be employed on a population level. Moreover, we observed evidence that regulation of *PGC1A* and *PGC1B* is associated with all-cancer incidence suggesting dysregulation of these genes may serve as a common event, early in tumor formation. We also identified interactions between epigenetic regulation of these genes by both telomere length and mitochondrial abundance. These findings together may aid in the development of cancer screening modalities, add to the current knowledge of pan-cancer etiology and facilitate future studies of these mechanisms in the context of cancer risk.

Summary points

- Aging and cancer development have common biological foundations such as telomere shortening and reductions in mitochondrial function.
- Epigenetic regulation of *PGC1A* and *PGC1B* is hypothesized to be a possible link between these factors.
- Epigenetic regulation of *PGC1A* and *PGC1B* and telomere length may additionally interact in predicting cancer risk.
- We used mixed-effect linear regression models to examine whether methylation of *PGC1A* and *PGC1B* CpGs linked telomere shortening with mitochondrial dysfunction.
- We used Cox regression models with discrete time-varying covariates and to examine associations between *PGC1A* and *PGC1B* CpG methylation and cancer incidence; we additionally examined interactions with telomere length and mitochondrial function.
- Five and eight distinct *PGC1A* and *PGC1B* CpG sites were associated with telomere length and mitochondrial abundance, although we did not identify any overlapping sites.
- cg27514608 from the TSS1500 of *PGC1A*, and cg15219393 from the first exon/5' UTR of *PGC1B* were significantly associated with cancer incidence.
- Upon stratification by telomere length, cg15219393 was significantly associated with cancer incidence among those with shorter telomeres.
- This is the first study to examine a possible role of DNA methylation in *PGC1A* and *PGC1B* in relation to telomere length, mitochondrial function and cancer risk.
- Blood leukocyte DNA methylation of these genes may serve as a biomarker of future cancer risk.

Financial & competing interests disclosure

The Epidemiology Research and Information Center of US Department of Veterans Affairs (NIEHS R01-ES015172) support the Normative Aging Study, which is a research component of the Massachusetts Veterans Epidemiology Research and Information Center (MAVERIC). L Hou, C Achenbach and R Murphy received support from the National Cancer Institute (U54-CA221205). L Hou received additional support from the Northwestern University Robert H Lurie Comprehensive Cancer Center Rosenberg Research Fund. A Baccarelli and J Schwartz received additional support from the National Institute of Environmental Health Sciences (NIEHS R01-ES021733 and NIEHS P30-ES00002). J Kresovich received additional support from the National Cancer Institute Cancer Education and Career Development Program (NIH R25-CA057699). The authors have no other relevant affiliations or financial

involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

1. Navarrete-Reyes AP, Soto-Pérez-de-Celis E, Hurria A. Cancer and aging: a complex biological association. *Rev. Invest. Clin.* 68(1), 17–24 (2016).
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. *CA Cancer J. Clin.* 67(1), 7–30 (2017).
3. Finkel T, Serrano M, Blasco MA. The common biology of cancer and ageing. *Nature* 448(7155), 767–774 (2007).
4. Kong Y, Trabucco SE, Zhang H. Oxidative stress, mitochondrial dysfunction and the mitochondria theory of aging. *Interdiscip. Top. Gerontol.* 39, 86–107 (2014).
5. Wang CH, Wu SB, Wu YT, Wei YH. Oxidative stress response elicited by mitochondrial dysfunction: implication in the pathophysiology of aging. *Exp. Biol. Med. (Maywood)* 238(5), 450–460 (2013).
6. Mikhed Y, Daiber A, Steven S. Mitochondrial oxidative stress, mitochondrial DNA damage and their role in age-related vascular dysfunction. *Int. J. Mol. Sci.* 16(7), 15918–15953 (2015).
7. Demissie S, Levy D, Benjamin EJ *et al.* Insulin resistance, oxidative stress, hypertension, and leukocyte telomere length in men from the Framingham Heart study. *Aging Cell* 5(4), 325–330 (2006).
8. von Zglinicki T. Oxidative stress shortens telomeres. *Trends Biochem. Sci.* 27(7), 339–344 (2002).
9. Kawanishi S, Oikawa S. Mechanism of telomere shortening by oxidative stress. *Ann. NY Acad. Sci.* 1019, 278–284 (2004).
10. Fouquerel E, Lormand J, Bose A *et al.* Oxidative guanine base damage regulates human telomerase activity. *Nat. Struct. Mol. Biol.* 23(12), 1092–1100 (2016).
11. De Luca A, Fiorillo M, Peiris-Pagès M *et al.* Mitochondrial biogenesis is required for the anchorage-independent survival and propagation of stem-like cancer cells. *Oncotarget* 6(17), 14777–14795 (2015).
12. Shen J, Gopalakrishnan V, Lee JE, Fang S, Zhao H. Mitochondrial DNA copy number in peripheral blood and melanoma risk. *PLoS ONE* 10(6), e0131649 (2015).
13. He Y, Gong Y, Gu J, Lee JJ, Lippman SM, Wu X. Increased leukocyte mitochondrial DNA copy number is associated with oral premalignant lesions: an epidemiology study. *Carcinogenesis* 35(8), 1760–1764 (2014).
14. Xu E, Sun W, Gu J, Chow WH, Ajani JA, Wu X. Association of mitochondrial DNA copy number in peripheral blood leukocytes with risk of esophageal adenocarcinoma. *Carcinogenesis* 34(11), 2521–2524 (2013).
15. Sanders JL, Newman AB. Telomere length in epidemiology: a biomarker of aging, age-related disease, both, or neither? *Epidemiol. Rev.* 35, 112–131 (2013).
16. Du J, Zhu X, Xie C *et al.* Telomere length, genetic variants and gastric cancer risk in a Chinese population. *Carcinogenesis* 36(9), 963–970 (2015).
17. Hurwitz LM, Heaphy CM, Joshu CE *et al.* Telomere length as a risk factor for hereditary prostate cancer. *Prostate* 74(4), 359–364 (2014).
18. Shen J, Terry MB, Gurvich I, Liao Y, Senie RT, Santella RM. Short telomere length and breast cancer risk: a study in sister sets. *Cancer Res.* 67(11), 5538–5544 (2007).
19. Willeit P, Willeit J, Mayr A *et al.* Telomere length and risk of incident cancer and cancer mortality. *JAMA* 304(1), 69–75 (2010).
20. Sahin E, Colla S, Liesa M *et al.* Telomere dysfunction induces metabolic and mitochondrial compromise. *Nature* 470(7334), 359–365 (2011).
21. Finkel T. Telomeres and mitochondrial function. *Circ. Res.* 108(8), 903–904 (2011).
22. Bell B, Rose C, Damon A. The normative aging study: an interdisciplinary and longitudinal study of health and aging. *Int. J. Aging Hum. Dev.* 3(1), 5–17 (1972).
23. Cawthon RM. Telomere measurement by quantitative PCR. *Nucleic Acids Res.* 30(10), e47 (2002).
24. Hou L, Savage SA, Blaser MJ *et al.* Telomere length in peripheral leukocyte DNA and gastric cancer risk. *Cancer Epidemiol. Biomarkers Prev.* 18(11), 3103–3109 (2009).
25. Hou L, Joyce BT, Gao T *et al.* Blood telomere length attrition and cancer development in the normative aging study cohort. *EBioMedicine* 2(6), 591–596 (2015).
26. Du P, Zhang X, Huang CC *et al.* Comparison of β -value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* 11, 587 (2010).
27. Zhong J, Cayir A, Trevisi L *et al.* Traffic-related air pollution, blood pressure, and adaptive response of mitochondrial abundance. *Circulation* 133(4), 378–387 (2016).
28. Malik AN, Czajka A. Is mitochondrial DNA content a potential biomarker of mitochondrial dysfunction? *Mitochondrion* 13(5), 481–492 (2013).

29. Malik AN, Parsade CK, Ajaz S *et al*. Altered circulating mitochondrial DNA and increased inflammation in patients with diabetic retinopathy. *Diabetes Res. Clin. Pract.* 110(3), 257–265 (2015).
30. Andreu AL, Martinez R, Marti R, García-Arumí E. Quantification of mitochondrial DNA copy number: pre-analytical factors. *Mitochondrion* 9(4), 242–246 (2009).
31. Lin CS, Wang LS, Tsai CM, Wei YH. Low copy number and low oxidative damage of mitochondrial DNA are associated with tumor progression in lung cancer tissues after neoadjuvant chemotherapy. *Interact. Cardiovasc. Thorac. Surg.* 7(6), 954–958 (2008).
32. Lin CS, Wang LS, Chou TY *et al*. Cigarette smoking and hOGG1 Ser326Cys polymorphism are associated with 8-OHdG accumulation on mitochondrial DNA in thoracic esophageal squamous cell carcinoma. *Ann. Surg. Oncol.* 20(Suppl. 3), S379–S388 (2013).
33. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat. Soc. Series B Stat. Methodol.* 57, 289–300 (1995).
34. Moslehi J, DePinho RA, Sahin E. Telomeres and mitochondria in the aging heart. *Circ. Res.* 110(9), 1226–1237 (2012).
35. Sahin E, DePinho RA. Axis of ageing: telomeres, p53 and mitochondria. *Nat. Rev. Mol. Cell Biol.* 13(6), 397–404 (2012).
36. Grewal SI, Moazed D. Heterochromatin and epigenetic control of gene expression. *Science* 301(5634), 798–802 (2003).
37. Berkers CR, Maddocks OD, Cheung EC, Mor I, Vousden KH. Metabolic regulation by p53 family members. *Cell Metab.* 18(5), 617–633 (2013).
38. Permuth-Wey J, Chen YA, Tsai YY *et al*. Inherited variants in mitochondrial biogenesis genes may influence epithelial ovarian cancer risk. *Cancer Epidemiol. Biomarkers Prev.* 20(6), 1131–1145 (2011).
39. Wirttenberger M, Tchatchou S, Hemminki K *et al*. Associations of genetic variants in the estrogen receptor coactivators PPARGC1A, PPARGC1B and EP300 with familial breast cancer. *Carcinogenesis* 27(11), 2201–2208 (2006).
40. Cho YA, Lee J, Oh JH *et al*. Genetic variation in PPARGC1A may affect the role of diet-associated inflammation in colorectal carcinogenesis. *Oncotarget* 8(5), 8550–8558 (2017).
41. Martínez-Nava GA, Burguete-García AI, López-Carrillo L, Hernández-Ramírez RU, Madrid-Marina V, Cebrián ME. PPARγ and PPARGC1B polymorphisms modify the association between phthalate metabolites and breast cancer risk. *Biomarkers* 18(6), 493–501 (2013).
42. Li Y, Wedrén S, Li G *et al*. Genetic variation of ESR1 and its co-activator PPARGC1B is synergistic in augmenting the risk of estrogen receptor-positive breast cancer. *Breast Cancer Res.* 13(1), R10 (2011).
43. Lou S, Lee HM, Qin H *et al*. Whole-genome bisulfite sequencing of multiple individuals reveals complementary roles of promoter and gene body methylation in transcriptional regulation. *Genome Biol.* 15(7), 408 (2014).
44. Haycock PC, Burgess S, Nounu A *et al*. Association between telomere length and risk of cancer and non-neoplastic diseases: a Mendelian randomization study. *JAMA Oncol.* 3(5), 636–651 (2017).
45. Julin B, Shui I, Heaphy CM *et al*. Circulating leukocyte telomere length and risk of overall and aggressive prostate cancer. *Br. J. Cancer* 112(4), 769–776 (2015).
46. Svenson U, Roos G, Wikström P. Long leukocyte telomere length in prostate cancer patients at diagnosis is associated with poor metastasis-free and cancer-specific survival. *Tumour Biol.* 39(2), 1010428317692236 (2017).
47. Caini S, Raimondi S, Johansson H *et al*. Telomere length and the risk of cutaneous melanoma and non-melanoma skin cancer: a review of the literature and meta-analysis. *J. Dermatol. Sci.* 80(3), 168–174 (2015).
48. Myers RM, Stamatoyannopoulos J, Snyder M *et al*. A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol.* 9(4), e1001046 (2011).
49. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* 13(7), 484–492 (2012).
50. Rauscher GH, Kresovich JK, Poulin M *et al*. Exploring DNA methylation changes in promoter, intragenic, and intergenic regions as early and late events in breast cancer formation. *BMC Cancer* 15, 816 (2015).