The Texas Medical Center Library

DigitalCommons@TMC

The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access)

The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences

5-2015

SOCIAL-DEMOGRAPHICS, HEALTH BEHAVIORS, TELOMERE LENGTH IN MEXICAN-AMERICAN COHORT AND CIRCULATING DNA METHYLATION IN BLADDER CANCER

Lixia Han

Follow this and additional works at: https://digitalcommons.library.tmc.edu/utgsbs_dissertations

 Part of the Translational Medical Research Commons

Recommended Citation

Han, Lixia, "SOCIAL-DEMOGRAPHICS, HEALTH BEHAVIORS, TELOMERE LENGTH IN MEXICAN-AMERICAN COHORT AND CIRCULATING DNA METHYLATION IN BLADDER CANCER" (2015). *The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access)*. 573.
https://digitalcommons.library.tmc.edu/utgsbs_dissertations/573

This Thesis (MS) is brought to you for free and open access by the The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences at DigitalCommons@TMC. It has been accepted for inclusion in The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact digitalcommons@library.tmc.edu.

The TMC Library Health Sciences Resource Center

SOCIAL-DEMOGRAPHICS, HEALTH BEHAVIORS, TELOMERE LENGTH IN MEXICAN-AMERICAN COHORT AND CIRCULATING DNA METHYLATION IN BLADDER CANCER

Lixia Han, M.S.

APPROVED:

_____
Xifeng Wu, M.D., Ph.D.
Advisory Professor

_____
Oliver Bogler, Ph.D.

_____
Colin P. Dinney, M.D.

_____
Jian Gu, Ph.D.

_____
Alan Wang, Ph.D.

APPROVED:

_____
Dean, The University of Texas
Graduate School of Biomedical Sciences at Houston

SOCIAL-DEMOGRAPHICS, HEALTH BEHAVIORS, TELOMERE LENGTH IN MEXICAN-AMERICAN COHORT AND CIRCULATING DNA METHYLATION IN BLADDER CANCER

A

THESIS

Presented to the Faculty of
The University of Texas
Health Science Center at Houston
and
The University of Texas
MD Anderson Cancer Center
Graduate School of Biomedical Sciences

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

by

Lixia Han, M.S.

Houston, Texas

May, 2015

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Dr. Xifeng Wu, for her guidance and support throughout the development of my thesis, for her unyielding encouragement throughout my graduate education, and for her patience, enthusiasm, and immense knowledge. Her support and patience helped me overcome many crisis situations and finish this dissertation.

I am grateful to my Advisory Committee members. Dr. Jian Gu has been always there to listen and give advices. I am thankful to him for the long discussions that helped me finish the projects and for his revisions of this dissertation. Dr. Oliver Bogler, Dr. Colin P Dinney, and Dr. Alan Wang have all provided valuable suggestions during my thesis work and comments that greatly improved this dissertation.

I have been extremely fortunate to be able to work in an environment full of support and camaraderie. The faculty members, staff, and my fellow trainees in the Department of Epidemiology are all friendly and helpful. Dr. Hua Zhao led me into the Mexican American Cohort study, which was a new area for me, and helped me summarize and interpret the telomere length data. Dr. Yuanqing Ye, Maosheng Huang, M.S., and Qiong Dong, M.S. helped me with the statistical analysis. Dr. David Chang, a senior research scientist, has technical expertise and gave me enormous technical help throughout my thesis work. Dr. Yan Wang, a program manager, gave me many helpful advices. Dr. Julia Mendoza had many long discussions that helped me improve my knowledge. For all the lab members of Dr. Wu – Di Zhang, M.S., Hui Zhang, M.S.,

Sherri Luo, Wei Wang, Huifeng Zhang, Dr. Yilei Gong, Mei Jiang, Yueli Sun, and Anna Garcia – thank you for helping me and for all the helpful suggestions along the way.

Finally, I owe a lot to my family for their love and support. In particular, I must acknowledge my husband, Dr. Gang Chen. Without his love and encouragement, I would not have finished this thesis.

SOCIAL-DEMOGRAPHICS, HEALTH BEHAVIORS, TELOMERE LENGTH IN MEXICAN-AMERICAN COHORT AND CIRCULATING DNA METHYLATION IN BLADDER CANCER

Lixia Han, M.S.

Advisory Professor: Xifeng Wu, M.D., Ph.D.

I. SOCIAL-DEMOGRAPHICS, HEALTH BEHAVIORS, AND TELOMERE LENGTH IN MEXICAN AMERICANS: A COHORT STUDY

Recent studies using a prospective cohort design have suggested that telomere length in peripheral blood leukocytes is not only a potential indicator of cellular aging that has been linked to stressful life experience and health behaviors, but also a prognostic marker for major chronic diseases; however, such study has never been done among adult Mexican Americans. In this current study, we examined cross-sectional associations among social-demographics, lifestyle behaviors, and relative telomere length (RTL) in peripheral blood leukocytes, as well as longitudinal relationships among major chronic diseases, weight gain, and RTL, among 12,792 Mexican Americans aged 20 to 85 years in the Mano-A-Mano, a Mexican American Cohort. As expected, RTL was inversely correlated with age (ρ=-0.15, $p<0.001$). In the multivariate analysis, we found that RTL was positively correlated with levels of education (P=0.021), self-insurance (P=0.041), body mass index (BMI) (P<0.001), and sleeping time per day (P for trend<0.001). In

contrast, RTL was inversely correlated with sitting time per day (P for trend =0.001). In longitudinal analysis, we found that longer RTL was modestly but positively associated with increased risks of diabetes (adjusted hazard ratio (adj.HR) =1.04, 95% conference interval (CI) =1.01-1.08) and hypertension (adj.HR=1.03, 95% CI=1.00-1.08). Similar but marginal association was observed for incident cancer (adj.HR=1.04, 95% CI=0.99-1.10). Further stratified by cancer site, the significant association was observed for lung cancer (adj.HR=1.21, 95% CI=1.00-1.46). In addition, longer RTL was a positive predictor of at least 10% weight gain (adj.HR=1.03, 95% CI=1.00-1.05). In summary, our results in Mexican Americans support the notion that telomere length is a biological mechanism by which social demographics and health behaviors "get under the skin" to affect health.

II. METHYLATION IN SERUM CELL-FREE DNA AND CPG-SNP WITH BLADDER CANCER RISK

Epigenetic alterations are early events of cancers, including bladder cancer. In the present study, we aimed to examine methylation profiles of circulating DNA in serum of bladder cancer patients, in an effort to develop reliable DNA methylation signatures to diagnose the disease. Firstly, we performed whole genome DNA methylation profiling with Illumina Infinium HumanMethylation450 beadchip in 23 participants in a bladder cancer case control study. 396 target cytosine-phosphate- guanine (CpG) sites were identified as hypermethylated in cases compared with the controls. The top 5 candidate CpG sites hypermethylated in cases during the screening phase were further validated by

pyrosequencing, in 100 bladder cancer patients (including 50 non-muscle invasive and 50 muscle invasive bladder cancers) and 50 healthy controls. Successful methylation data were obtained for 4 of the selected CpG sites, located in genes of TTC23, WWOX, ZNF624 and LOC2211122, respectively. Unfortunately, none of these 4 sites exhibited differential methylation between cases and controls in this validation. Interestingly, there was a SNP rs8038732 (G>A) in the methylation site of gene TTC23. The GG genotype showed near complete methylation, G/A heterozygotes half methylation, and AA complete loss of methylation. This SNP is clearly a functional SNP because the G>A transition almost completely knocks out methylation at this site. It would be interesting to see whether this SNP affect bladder cancer risk. However, this SNP was not genotyped in our previously published genome-wide association study (GWAS). But, we found that another SNP, rs1377267 (A>C), on the GWAS chip is in strong linkage disequilibrium (r2=0.85) with rs8038732. Moreover, the SNP rs1377267 significantly associated with the risk of bladder cancer in our GWAS study. Specifically, the heterozygous alleles (AC) was associated with higher bladder cancer risk (OR=1.34, 95% CI=1.09-1.65), whereas the homozygous alleles (CC) were marginally associated with lower risk of bladder cancer (OR=0.78, 95% CI= 0.59-1.03). From this result, we could infer that rs8038732 has similar association with the risk of bladder cancer. Our study suggests that rs8038732 affects the risk of bladder cancer by removing a cytosine-phosphate-guanine (CpG) dinucleotide and knocking out DNA methylation at this site. Loss of methylation may increase the expression of TTC23, which is a potential oncogene.

Table of Contents

List of Figures

List of Tables

Abbreviations

RTL: relative telomere length

OR: odds ratio

HR: hazard ratio

CI: confidence interval

BMI: Body Mass Index

SES: socioeconomic status

NMIBC: Non-muscle invasive bladder cancer

MIBC: muscle invasive bladder cancer

CpG: Cytosine–phosphate–guanine

SNP: single nucleotide polymorphism

GWAS: Genome-wide association study

LD: Linkage disequilibrium

SNAP: SNP annotation and proxy search

Project I: Social-demographics, health behaviors and telomere length in Mexican Americans：a cohort study

Introduction

Telomeres consist of tandem DNA repeats and associated proteins that cap the end of chromosomes and provide protection to eukaryotic chromosomes [1] [2]. As a normal cellular aging process, telomere length decreases with age. However, how quick telomeres shorten varies significantly among the population, and many factors, including genetic, environmental, socio-demographic, cultural and behavioral factors [3-5], may affect the rate of telomere length shortening. More importantly, how long the telomeres are has significant public health and disease implication. Critically short telomeres may lead to chromosomal instability and thus cause diseases [6]. On the other hand, cells with longer telomeres may have prolonged cell life and higher chances of developing DNA damage and chromosomal instability, thus being at high risk of tumorigenesis [7]. Current studies have demonstrated links between telomere length and risk of major chronic diseases, including cancer [8-10], type 2 diabetes [11], and cardiovascular disease [12, 13].

The link between low socioeconomic status (SES) and the development of age-related chronic diseases has been widely observed [14-20]. Several theories have been proposed and the common assumption is that the chronic stress associated with social disadvantage causes wear and tear on the body, which promotes the declined rate of physiological function [20, 21]. Such decline may affect leukocyte telomere length and

further impact biological aging and other age-related chronic diseases. Studying the role of leukocyte telomere length in stress, biological aging, and age-related chronic diseases is particularly relevant to Mexican Americans. Mexican Americans are one of the fastest-growing populations in U.S. [22]. Compared to non-Hispanic whites, Mexican Americans tend to have lower SES, including lower income and homeownership, and less education and access to health care [23]. On the other hand, Mexican Americans are experiencing high burden of age-related chronic diseases, including cardiovascular diseases and type 2 diabetes [24-27]. In addition, they have unique psychological, somatic, and social stress associated with acculturation [28, 29]. Studies have shown that acculturation confers risk for stress, and stress related diseases and poor health behaviors in Mexican Americans [30-32].

Leukocyte telomere length may also be the link between unhealthy behaviors and age-related chronic diseases [3, 33, 34]. Maintaining a healthy lifestyle, including keeping a normal weight, regular physical activity, not smoking, and not heavy drinking, can contribute to decrease the levels of inflammation and oxidative stress [34-38], which helps to prevent chronic diseases [33, 34, 39, 40]. Studying health behaviors, age-related chronic diseases, and leukocyte telomere length is particularly important for Mexican Americans because 3 out of 4 Mexican Americans are either overweight or obese [41, 42]. In addition, they participate in low levels of physical activity [43], and recent surveillance data show only 44.6% met national physical activity recommendations [44].

To date, the relationship among social demographics, lifestyle behaviors, and leukocyte telomere length has not been investigated in any prospective cohort study of adult Mexican Americans. Here, taking advantage of a large ongoing prospective Mexican American Cohort study, we detected relative telomere length (RTL) in peripheral blood leukocytes from 12,792 Mexican American study subjects, and investigated its relationships with social demographics, lifestyle behaviors, and major chronic diseases at baseline, and with the development of cancer, diabetes, and hypertension at follow-up. Additionally, we studied whether RTL could predict prospective weight gain during follow-up.

METHODS AND MATERIALS

Study Population

The samples for the current study were recruited from participants in a large population-based cohort of Mexican origin households recruited from Houston [45]. This Mano a Mano cohort, an ongoing prospective cohort of 1st and 2nd generation Mexican origin immigrant households in Houston, TX, initiated in July 2001 and maintained by the Department of Epidemiology at the University of Texas MD Anderson Cancer Center in Houston, Texas. A detailed description of the sampling and recruitment strategy has been published previously [46]. Briefly, Participants was drawn by random-digit-dialing, block walking in predominantly Mexican American neighborhoods, from community centers and local clinics, or networking through currently enrolled participants [47] [48]. Once the identified eligible households agreed to participate in the study, then the written informed consent would be obtained; the participants were interviewed by well-trained interviewers using a structured questionnaire; the interview last ~45 minutes and was conducted in Spanish or English; the epidemiological data collected by this questionnaire included birthplace, residential history, social-demographic characteristics, lifestyle behaviors, levels of physical activity, personal medical history, family history of chronic disease, acculturation, and occupation exposure [49]. All the participants were followed up through telephone annually; the study was approved by the institutional review board of MD Anderson Cancer Center [46].

Relative Telomere Length (RTL) Assessment

Genomic DNA was isolated from participants' whole blood with the QIAamp Maxi DNA kit from Qiagen (Valencia, CA). Isolated DNA samples were quantified by Quant-iT^{TM} PicoGreen$^{®}$dsDNA Reagent and Kits from Invitrogen (Carlsbad, CA) according to the manufacturer's directions. The RTL was measured with a modified method of real-time quantitative polymerase chain reaction (PCR), which was originally set up by Cawthon [50, 51]. A detailed description of the experimental strategy has been published previously [52-54]. Briefly, the ratio of the telomere repeat copy number (T) to the single gene (human globulin) copy number (S) was detected, which was proportional to the telomere length [54]. The mixture of PCR reaction (14 µl) consisted of 1 x SYBR Green Master Mix (Applied Biosystems, Foster City, CA), 200 nmol/l Tel-1 primer (5'-CGGTTTGTTTGGGTTTGGGTTTGGGTTTGGGTTTGGGTT), 200 nmol/l Tel-2 primer (5'-GGCTTGCCTTACCCTTACCCTTACCCTTACCCTTACCCT), and 5 ng genomic DNA; the PCR reaction mixture (14 µl) for *HGB* gene amplification was similar except the primers- Hgb-1 primer (5'-GCTTCTGACACAACTGTGTTCACTAGC) and Hgb-2 primer (5'-CACCAA CTTCATCCACGTTCACC); the thermal cycling conditions were 1 cycle at 95^{o}C for 10 minutes followed by 40 cycles at 95^{o}C for 15 seconds and at 56^{o}C (for telomere amplification) or 58^{o}C (for *HGB* amplification) for 1 minute; the HT7900 system (Applied Biosystems, CA) was used to perform the Real-time PCR，each sample was detected in duplicates in a 384-well plate; in each assay, negative control, positive control (Roche Applied Science, Pleasanton, CA), calibrator DNA, and a

standard curve were included; the standard deviations for the Ct values were set at 0.25 for all the samples; the standard curve with high coefficient of determination ($R^2 \geq 0.99$) will be selected for further analysis; in addition, the acceptable coefficient of variation were <3% for intra-assay and <5% for the inter assay [53].

Statistical Analysis

We used the statistical software package SAS, version 9.4 (SAS, Cary, NC) for analysis. Although the relative telomere length data were not normally distributed, we performed the analysis using both data with and without log transformation. We found there was no significant difference in estimate of the relationships among social-demographics, health behaviors, major chronic diseases, and RTL. Thus, only data without log transformation were presented here. Differences in the distribution of social-demographics, health behaviors, and major chronic diseases were evaluated by Pearson χ^2 test for categorical variables. Spearman correlation was used to assess the relationship between age and RTL. In the univariate analysis, the general linear regression analysis was used to compare RTL by the selected variable. Both age and gender were adjusted as appropriate. Mediation analysis was performed to assess the role of RTL between BMI and chronic diseases at baseline. In the multivariate analysis, the general linear model was used. The social-demographic and health behavior variables with P value less than 0.25 in the univariate analysis were included in the multivariate analysis. For joint effect analysis, we created a social-behavioral score by including 2

SES related variables (education and self-insurance) and 2 lifestyle behaviors (sitting time and sleeping time). For each variable, we created a binary high SES or healthy variable, defined as at least high school education, having self-insurance, sitting less than 2 hours per day, and sleeping more than 6 hours per day, respectively. The social-behavioral score ranged from 0 (low SES and no healthy behaviors) to 4 (high SES and healthy behaviors). Multivariate analysis was used to assess the correlation between the social behavior and RTL. The associations among incident cancer, diabetes, hypertension and RTL were assessed using the multivariable-adjusted Cox proportional-hazard regression model. RTL was analyzed as both a continuous and categorical variable. For categorical variable, cutoff points were set at the quartile values in the non-cancer study population. Adjusted Hazard ratio (HR) and 95% confidence interval (CI) were estimated and co-variants were adjusted as appropriate. Multivariable-adjusted Cox proportional-hazard regression model was also used to examine whether RTL at baseline could predict at least 10% weight gain in the follow-up. All statistical tests were 2-sided, and the value of statistical significance was set as 0.05.

RESULTS

A total of 12,792 MA study subjects were included in the RTL analysis. Among them, 614 were excluded for further analysis due to abnormal RTL value (outside of ±3 standard deviation of age adjusted mean RTL). Thus, a total of 12,178 study subjects were included in further data analysis. Baseline social demographic characteristics, immigration related variables, lifestyle factors, and major chronic diseases are presented in **Table 1**. The mean and median ages of the population were 42 and 39 years old, respectively. The majority of study subjects were women (79.56%) and married (or living together) (77.00%). In this study, we used levels of education, health insurance, home and vehicle ownership to assess SES. Only 40.60% of study subjects had at least high school education. For health insurance, 49.66%, 46.23%, and 68.55% had self, partner, and kids insurances, respectively. In terms of home and vehicle ownership, 45.96% and 78.41% owned a home and a vehicle, respectively. For immigration related variables, the majority was born in Mexico (73.77%) and less acculturated (63.28%). Among Mexico born study subjects, 29.53% arrived in U.S. younger than 20 years old and 46.39% have lived in U.S. for more than 15 years. We included 6 lifestyle behaviors in the analysis, namely cigarette smoking, alcohol consumption, BMI, sedentary behavior, sitting time, and sleeping time. The majority of study subjects were overweight or obese (85.04%), never smokers (71.80%), never alcohol drinkers (66.66%), and sedentary (79.06%). Over a half of study subjects had more than 2 hours sitting time per day (53.42%). The majority of study subjects had 7 to 8 hours sleeping time per day (60.36%). The most

prevalent chronic disease was hypertension (19.48%), followed by diabetes (15.27%), gallbladder disease (8.83%), asthma (3.97%), cancer (2.65%), and kidney disease (1.91%).

The mean and median of RTLs were 0.76 and 0.75, respectively. The range was between 0.21 and 0.99. As expected, a significant inverse relationship between RTL and age was observed (ρ=-0.15, $p<0.001$). For every one year increase in age, there was a 0.002 reduction in RTL. In the univariate analysis, we investigated the relationships among baseline social-demographics, health behaviors, major chronic diseases, and RTL (**Table 1**).

Table 1. Distribution of selected socio-demographics, health behaviors, and major diseases and RTL among 12,178 Mexican American study subjects

Variables	Number of study subjects (%)	Relative telomere length	P value*
Age at enrollment			
20-31	3,116 (25.59)	0.81	
32-39	3,104 (25.49)	0.78	
40-50	2,925 (24.02)	0.75	
>50	3,033 (24.91)	0.72	**<0.001**
Gender			
Men	2,535 (20.44)	0.76	
Women	9,866 (79.56)	0.76	0.324
Marital Status			
Married/living as married	9,365 (77.00)	0.76	
Other	2,797 (23.00)	0.76	0.780
Education			
<High school	7,229 (59.40)	0.76	
At least high school	4,941 (40.60)	0.77	0.069
Self insurance			
Yes	5,498 (49.66)	0.77	
No	5,574 (50.34)	0.76	**0.004**
Partner insurance			
Yes	4,227 (46.23)	0.77	
No	4,917 (53.77)	0.76	0.189
Kids insurance			
Yes	6,502(68.55)	0.77	
No	2,983(31.45)	0.76	0.264
Own home			
Yes	5,025 (45.96)	0.76	
No	5,909 (54.04)	0.76	0.539
Own car			
Yes	8,558 (78.41)	0.76	
No	2,356 (21.59)	0.77	0.618
Birth location			
Mexico	8,973 (73.77)	0.76	

U.S.	3,191 (26.23)	0.76	0.221
Years lived in U.S. (for born in Mexico only)			
<5	1,038 (11.57)	0.75	
5-10	1,841 (20.52)	0.77	
10-15	1,930 (21.51)	0.76	
>15	4,162 (46.39)	0.77	0.255
Age at arrival (for born in Mexico only)			
<20	2,649 (29.53)	0.76	
20-29	3,906 (43.55)	0.77	
≥30	2,415 (26.92)	0.76	0.242
Language acculturation			
High	4,449 (36.72)	0.76	
Low	7,666 (63.28)	0.76	0.308
Cigarettes smoking status			
Current	1,579 (12.99)	0.76	
Former	1,848 (15.20)	0.76	
Never	8,727 (71.80)	0.76	0.882
Alcohol drinking status			
Current	2,770 (22.88)	0.76	
Former	1,267 (10.46)	0.76	
Never	8,071 (66.66)	0.77	0.244
BMI			
Under/Normal weight (<25)	1,791 (14.96)	0.75	
Overweight (25-29.9)	4,072 (34.00)	0.76	
Obese (≥30)	6,112 (51.04)	0.77	**<0.001**
Sedentary behavior			
No	2,380 (20.94)	0.77	
Yes	8,987 (79.06)	0.76	0.052
Sit hours/day			
<2	4,784 (46.58)	0.76	
2-3	2,817 (27.43)	0.74	
>3	2,670 (26.00)	0.75	**<0.001**
Sleep hours/day			
<=6	3,307 (28.96)	0.74	
7-8	6,894 (60.36)	0.75	
≥9	1,220 (10.68)	0.77	**0.001**

Any Major Disease (prevalence)			
Yes	7,298 (60.01)	0.77	
No	4,864 (39.99)	0.75	**0.013**
High blood pressure (prevalence)			
Yes	2,372 (19.48)	0.78	
No	9,805 (80.52)	0.76	**<0.001**
Diabetes (prevalence)			
Yes	1,859 (15.27)	0.77	
No	10,317 (84.73)	0.76	0.401
Gallbladder disease (prevalence)			
Yes	1,075 (8.83)	0.79	
No	11,102 (91.17)	0.76	**0.001**
Asthma (prevalence)			
Yes	483 (3.97)	0.77	
No	11,694 (96.03)	0.76	0.729
All cancer (prevalence)			
Yes	323 (2.65)	0.84	
No	11,855 (97.35)	0.76	**<0.001**
Kidney disease (prevalence)			
Yes	232 (1.91)	0.82	
No	11,945 (98.09)	0.76	**0.014**
* Adjust for age and gender as appropriate			

Among demographic characteristics, the most significant association was observed for age category. As expected, we observed a significant trend of decreasing RTL with the increasing age category from 20-31 to over 50 years old ($p<0.001$). Among SES related variables, study subjects who had less than high school education had marginally shorter RTL than those who had at least high school education ($p=0.069$). Study subjects who had self-insurance had significantly longer RTL than those who did not ($p=0.004$). No

significant association was observed between immigration related variables (including birth location, acculturation, years lived in U.S., and age at arrival) and RTL.

For health behaviors, the most significant association was with BMI. We found a significant positive relationship between BMI (as continuous variable) and RTL ($p<0.001$). In the BMI categorical analysis, after adjusting for age and gender, RTL was shortest among study subjects who were under or normal weight (0.75), median among those who were overweight (0.76), and longest among those who were obese (0.77) ($p<0.001$). When the study subjects were stratified by age categories, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, and 80 or more years old (**Figure 1**), the trend of RTL difference among BMI categories was persistent for most of the age categories. More interestingly, when we further compared the rate of RTL shortening by age among three BMI categories, RTL shortening by age was seemingly quicker among those who were under or normal weight than those who were overweight or obese, although the difference of slope was not statistically significant ($p=0.187$). When we compared RTL difference between those in 20-29 and those in 80 or more years old, the difference was 0.19 for under or normal weight, 0.13 for overweight, and 0.085 for obese groups ($P=0.005$). Significant relationships were also observed for sitting time and sleeping time. Study subjects who had less than 2 hours sitting time per day had the longest RTL (0.76) compared to those who had 2-3 hours (0.74) and more than 3 hours sitting time per day (0.75) ($P<0.001$). RTL was positively associated with hours of sleeping per day. RTL

was longest among those who slept at least 9 hours per day (0.77), median among those who slept 7-8 hours per day (0.75), and shortest among those who slept no more than 6 hours per day (0.74) (P<0.001). In addition, those who were sedentary had marginally shorter RTL than those who were not sedentary (0.76 vs 0.77, P=0.052).

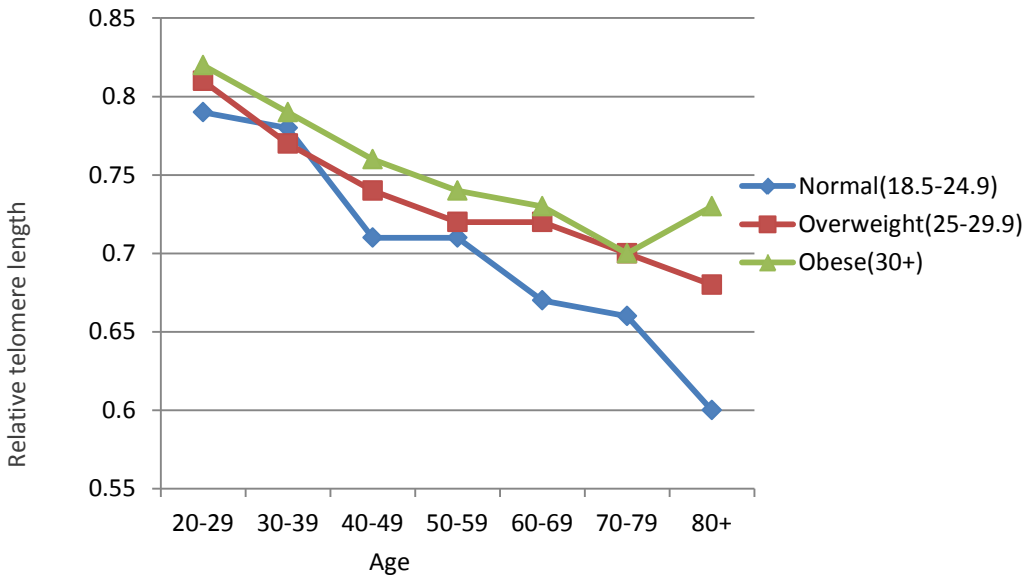

Fig. 1 Association of RTL and BMI in different age categories

For major chronic diseases at baseline, study subjects who had at least one of six major chronic diseases (including hypertension, diabetes, gallbladder disease, asthma, cancer, and kidney disease) had longer RTL than those who have no chronic diseases (RTL: 0.77 vs 0.75, p=0.013). When stratified by type of chronic disease, significant difference was remained for hypertension (P<0.001), gallbladder disease (P<0.001), cancer (P<0.001), and kidney disease (P=0.014). In the mediation analysis, we examined the potential

mediating role of RTL on the association between BMI and major chronic diseases. The significant effect was observed for hypertension and gallbladder disease. About 5.25% of the association between BMI and hypertension, and 5.13% of the association between BMI and gallbladder disease were mediated by RTL (P=0.001 and 0.037, respectively).

Next, we included all social-demographics and health behaviors with p value ≤ 0.25 in the univariate analysis (from **Table 1**) into a multivariate model. The results were presented in **Table 2**. BMI, as a continuous variable, was positively associated with RTL (P<0.001). Compared to study subjects with lower education level, those with at least high school education had significantly longer RTL (P=0.021). Compared to those who slept 6 or less hours per day, those who slept 7-8 and more than 8 hours had significantly longer RTL (P=0.003 and 0.001, respectively). RTL was also longer among those who had self-insurance than those who had no self-insurance (P=0.041). On the other hand, RTL was significantly inversely associated with age (P<0.001), and RTL was shorter among those who sat more than 3 hours per day than those who sat less than 2 hours per day (P<0.001).

Table 2. Multivariate regression analysis of RTL on social-demographics and health behaviors			
Variables	Estimate	Standard Error	P value
age	-0.003	0.0002	**<0.001**
Gender (women vs men)	0.002	0.008	0.804
BMI	0.002	0.0005	**<0.001**
Education (≥ high school vs < high school)	0.013	0.006	**0.021**
Birth place (U.S. vs Mexico)	0.009	0.007	0.197
Ever vs never smokers	-0.008	0.007	0.230
Ever vs never drinkers	0.009	0.007	0.220
Sleep time (≤6 hours/day)			
7-8 hours/day	0.028	0.009	**0.003**
≥9hours/day	0.039	0.010	**0.001**
Sitting time (<2 hours/day)			
2-3 hours/day	-0.003	0.008	0.674
>3 hours/day	-0.026	0.007	**<0.001**
Sedentary lifestyle (yes vs no)	-0.0003	0.007	0.965
Self-insurance (yes vs no)	0.012	0.006	**0.041**
Intercept	0.838	0.021	**<0.001**
R^2	0.026		

To assess the potential joint effect of SES and health behaviors on RTL, we created a social-behavioral score by including 2 SES related variables (education and self-insurance) and 2 lifestyle behaviors (sitting time and sleeping time). For each variable, we created a binary high SES or healthy variable, defined as at least high school education, having self-insurance, sitting less than 2 hours per day, and sleeping more than 6 hours per day, respectively. The social-behavioral score ranged from 0 (low SES and no healthy behaviors) to 4 (high SES and healthy behaviors). We investigated the association between the social-behavioral score and RTL (**Figure 2**). In general, with the score increased from 0 to 4, RTL gradually increased ($p<0.001$). For study subjects with social-behavioral score 0, 1, 2, 3, and 4, their mean RTL was 0.71, 0.73, 0.75, 0.77, and 0.78, respectively.

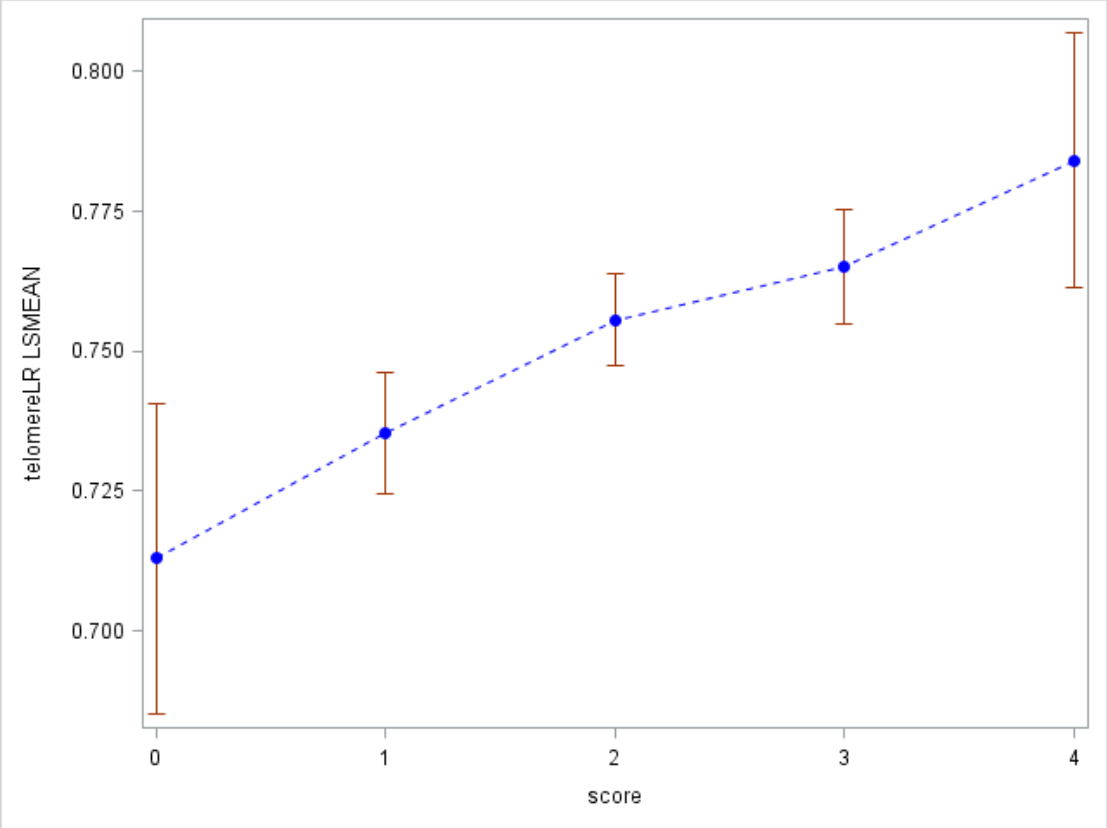

Fig. 2 Association between the social-behavioral score and RTL

To assess whether RTL was a predictor of chronic diseases, we analyzed the relationships between RTL and incident cancer, diabetes, and hypertension during the follow-up. A total of 287 incident cancer cases were observed and confirmed during the follow-up. Overall, RTL was not significant different between incident cancer cases and non-cancer controls (0.77 V.S. 0.76, p=0.555). In the multivariable-adjusted Cox proportional-hazard regression model, using RTL as a continuous variable, we found that longer RTL was marginally associated with increased risk of overall cancer (adjusted HR=1.04, 95%CI=0.99-1.10) (**Table 3**).

Table 3. Risk of incident cancer associated with RTL				
RTL (continuous variables, unit=0.1)	Cancer Cases	Non-cancer controls	Adj. HRs (95% CI)*	P value
All cancer	287	10,340	1.04 (0.99-1.10)	0.111
Breast cancer (women only)	68	8,414	1.01 (0.91-1.12)	0.867
Lung cancer	18	10,605	1.21 (1.00-1.46)	**0.049**
Cervical cancer (women only)	11	8,465	0.96 (0.76-1.23)	0.766
Liver cancer	12	10,611	1.01 (0.78-1.31)	0.960
Prostate cancer (men only)	20	2,120	1.09 (0.86-1.36)	0.898
		All Cancer		
By quartile	N (%)	N (%)	Adj. HRs (95% CI)*	P Value
1st	71 (24.74)	2,615 (25.29)	reference	
2nd	90 (31.36)	2,760 (26.69)	1.15 (0.84-1.58)	0.372
3rd	72 (25.09)	2,468 (23.87)	1.20 (0.86-1.67)	0.279
4th	54 (18.82)	2,497 (24.15)	1.41 (1.00-2.03)	**0.049**
P for trend				0.064

*adjusted for age, obesity status, alcohol drinking, cigarette smoking, and education.

When stratified by top incident cancer site, the only significant association observed was with lung cancer (adjusted HR=1.21, 95% CI=1.00-1.46). In further quartile analysis, compared to individuals who had shortest RTL (1st quartile), those who had longer RTL (2nd, 3rd, and 4th quartiles) had an increased risk of overall cancer. However, the only significant association was observed among those in the 4th quartile with the highest RTL (adjusted HR=1.41, 95% CI=1.00-2.03). The p for trend was 0.064. During the follow-up, a total of 648 and 878 study subjects were diagnosed with diabetes and hypertension. Overall, RTL was significantly positively associated with risks of incident diabetes and hypertension (diabetes: adjusted HR=1.04, 95% CI=1.01-1.08; hypertension: adjusted HR=1.03, 95% CI=1.00-1.06) (**Table 4**).

Table 4. Risk of incident diabetes and high blood pressure associated with RTL				
RTL	Diabetes N (%)	Non-diabetes N (%)	HR(95%CI)	P value
quartile				
1st	152 (23.46)	2,075 (24.82)	ref.	
2nd	211 (32.56)	2,165 (25.89)	1.23 (1.00-1.53)	0.055
3rd	144 (22.22)	2,025 (24.22)	1.00 (0.79-1.26)	0.988
4th	141 (21.76)	2,096 (25.07)	1.31 (1.03-1.67)	**0.026**
P for trend				0.167
Continuous variable (unit=0.1)			1.04(1.01-1.08)	**0.021**
	High blood Pressure N (%)	No high blood Pressure N (%)		
quartile				
1st	227 (25.85)	1,931 (25.14)	ref.	
2nd	257 (29.27)	1,985 (25.84)	0.93(0.77-1.11)	0.422
3rd	217 (24.72)	1,828 (23.80)	0.93(0.77-1.13)	0.474
4th	177 (20.16)	1,937 (25.22)	1.22(1.01-1.50)	**0.044**
P for trend				0.123
Continuous variable (unit=0.1)			1.03(1.00-1.06)	**0.049**

*adjusted for age, sex, obesity status, alcohol drinking, and education.

In further quartile analysis, we found that study subjects who had the longest RTL (4^{th} quartile) had a 1.31-fold increased risk of diabetes (adjusted HR=1.31, 95% CI=1.03-1.67) compared to those who had shortest RTL (1^{st} quartile). Similarly, we observed that study subjects who had highest RTL (4^{th} quartile) had 1.22 fold increased risk of hypertension (adjusted HR=1.22, 95%CI=1.01-1.50) compared to participants who had lowest RTL (1^{st} quartile). Last, in a subset of study subjects with multiple weight data during follow-up, we explored whether RTL could predict weight gain during the follow-up (**Table 5**).

Table 5. Risk of weight gain(>=10%) and RTL		
	HR(95%CI)	P value*
age	0.97 (0.97-0.98)	**<0.001**
Gender: women vs men	1.49 (1.17-1.91)	**0.001**
BMI	0.95 (0.94-0.96)	**<0.001**
marital status: married vs other	0.73 (0.63-0.85)	**<0.001**
acculturation: high vs low	1.20 (1.02-1.41)	**0.026**
Birth place: U.S. vs Mexico	0.83(0.69-1.00)	**0.049**
RTL (0.1 per unit)	1.03 (1.00-1.05)	**0.035**

Using at least 10% weight gain compared to the baseline weight during the follow-up as the event, we performed multivariate regression analysis. The significant variables in the final model included age at baseline, gender, baseline BMI, marriage status, acculturation, birth place, and RTL. RTL was a positive predictor of at least 10% weight gain during the follow-up (adjusted HR=1.03, 95%CI=1.00-1.05). In addition, older age, married (or living together), high BMI, and born in U.S. were associated with less likely to gain at least 10% weight (age: adjusted HR=0.97, 95%CI=0.97-0.98; marriage status: adjusted HR=0.73, 95%CI=0.63-0.85; BMI: adjusted HR=0.95, 95%CI=0.94-0.96; and birth place: adjusted HR=0.83, 95%CI=0.69-1.00, respectively). Being women and higher acculturated were associated with more likely to gain at least 10% weight (gender: adjusted HR=1.49, 95% CI=1.17-1.91; and acculturation: adjusted HR=1.20, 95% CI=1.02-1.41).

Discussion

The associations among social-demographics, health behaviors, and telomere length have rarely been studied in adult Mexican Americans. The only published study was from the National Health and Nutrition Examination Survey (1999-2002), which includes 1,377 Mexican Americans [20]. In that study, the only significant factor correlated with RTL was age. In our current study, with 12,792 Mexican Americans aged 20 to 85 years, in the cross-sectional analysis, we found that lower SES (measured by low education levels and without self-insurance) was significantly associated with shorter RTL. In addition, we observed RTL was positively correlated with BMI, sleeping time, and several major chronic diseases, and negatively correlated with age and sitting time. Then, in the longitudinal analysis, we found longer RTL was a predictor for all cancer, diabetes, hypertension, and weight gain, although the associations were modest to marginal.

Socioeconomic disparities in chronic diseases are well-documented, but little is known about the mechanisms [20]. In the current study, we use indicators of SES including education attainment, health insurance, home and vehicle ownership. We observed an inverse relationship between levels of SES and RTL. Study subjects with less than high school education had shorter RTL than those who had at least high school education (P=0.021). Similarly, study subjects with no self-insurance had shorter RTL than those who had self-insurance (P=0.041). We did not find a significant association

between home and vehicle ownership and RTL. Our results are consistent with several previous studies. For example, Needham et al. found that telomere length was positively associated with education level ($P<0.01$) [20]. However, the significant relationship was only observed among Caucasian American study subjects, not among African and Mexican Americans. The only significant relationship reported among Mexican Americans was from Multi-Ethnic Study of Atherosclerosis (MESA), which evaluated the relationship between SES and telomere length among 963 U.S. adolescents, including 510 Mexican American adolescents [20, 55]. In that study, telomere length was found negatively associated with current home ownership and father's education [20].

Health behavior is a major determinant of chronic disease risk in most of the developed countries [33, 34, 39, 56]. Epidemiological and basic science studies have shown that unhealthy lifestyle including cigarette smoking, heavy drinking, sedentary behavior, and obesity could increase stress levels, decrease immune function, and deteriorate physiological condition [57-62]. In the current study, we found inverse relationships among sitting time, sleeping time, and RTL, and a positive relationship between BMI and RTL. No significant relationship was observed among cigarette smoking, alcohol drinking, sedentary behavior, and RTL. The relationship among sedentary behavior, sitting time, and RTL has been investigated previously [6, 20, 63-70]. The results are still limited and inconsistent. For example, Sjögren et al. found sitting time was associated with longer telomere in sedentary, overweight 68-year-old participants [70]. Using data

from Nurses' Health Study, Du et al. reported that sitting time and sedentary behavior were not associated with leukocyte telomere length [6]. However, they found that longer telomeres are associated with moderate to vigorous amounts of physical activity [6]. Much of the discrepancy among different studies may be due to differences in study population, such as race, age, occupation, and sample size.

One intriguing but puzzling finding from this study is the positive relationship between BMI and telomere length. In a recent systematic review and meta-analysis of 29 existing studies, Muezzinler et al. concluded an inverse association between BMI and RTL in adults [71]. However, in the Genetic Epidemiology Research Study on Adult Health and Aging (GERA) with 100,000 study subjects presented in 2012 American Society of Human Genetic Annual meeting, Schaefer et al. reported that BMI was positively correlated with telomere length in saliva DNAs, which is consistent with our finding. To eliminate the potential age effect, we stratified the study subjects based on their age categories, and we found study subjects in the obese group had steadily longest telomere length and those in the normal weight group had consistently shortest telomere length. More interestingly, the difference of telomere length by BMI status was becoming larger with the age increased. In studying the relationship between RTL and perspective weight gain, we found that long RTL was associated with higher likelihood of gaining at least 10% weight compared to the baseline. Our observation is consistent with the results from the Health ABC study, in which Njajou et al. reported that telomere length was

positively associated with both % body fat and BMI in 7- year follow-up [72]. Clearly, more work is still needed to study the underlying molecular mechanisms between BMI and RTL.

The relationship between RTL and major chronic diseases (both prevalent and incident) was another piece of interesting finding. In cross-sectional analyses, longer RTL was associated with increased prevalence of overall major chronic diseases as well as several individually, including hypertension, gallbladder disease, cancer, and kidney disease. In prospective analyses, longer RTL was associated with modestly to marginally increased risks of diabetes, hypertension, overall cancer, and lung cancer. The relationship between RTL and cancer has been studied extensively in both case control and cohort settings [9, 10, 54, 73-87]. However, the results are inconsistent. One of the exceptions might be lung cancer. In a recent pooled analysis, Seow et al. found that telomere length was positively associated with higher risk of lung cancer [87]. Findings were consistent across the three cohorts, PLCO trial, ATBC trial, and SWHS. Our result is consistent with those findings. Although we do not know the exact biological explanation, upon exposure to carcinogens, cells with longer telomeres may be less likely to enter senescence or apoptosis. Their survival could increase their chance to have extended carcinogen exposure and thereby increase the risk of developing genetic abnormalities.

However, the relationship among diabetes, hypertension, and RTL is puzzling. In a recent meta-analysis study, Zhao et al. found that shortened telomere length was significantly associated with an increased risk of type 2 diabetes (OR=1.29, 95% CI: 1.11-1.14) [11]. Similar association was also observed between shortened telomere length and risk of hypertension [88]. One limitation in our study is that the incident diabetes and hypertension cases were self-reported and we did not separate type 1 from type 2 diabetes. Therefore misclassification may somehow contribute to the discrepancy. In addition, the observed associations in our study are modest at best. Apparently, more studies are needed to further clarify the associations.

There are several limitations in this study. First, the number of incident cancer, diabetes, and hypertension is still small, especially for cancer. This limits our statistical power to detect the associations. Thus, we need to be very cautious to interpret our findings. Second, the analysis of SES and health behavior in relation to RTL is cross-sectional. The findings do not tell us the timing of the effect. Longitudinal data would be necessary to analyze the rate of telomere shortening. Finally, the exposure of stress was not directly assessed [20].

Conclusion

The relationships among SES, health behavior, chronic diseases and telomere length have been studied previously [3, 33]. However, such study has never been comprehensively investigated among adult Mexican Americans. This study fills the gap. Among 12,792 Mexican Americans, we found that RTL is significantly affected by SES and health behavior. In addition, we found longer RTL is a predictor of lung cancer, diabetes, hypertension and weight gain. Together with the results of other studies in different populations and different races, our study suggests that RTL is one mechanism by which social conditions and health behaviors affect human health.

Project II: Methylation in serum cell-free DNA and CpG-SNP with bladder cancer risk

INTRODUCTION

Bladder cancer has the fifth highest incidence rate of all malignancies in the United States, with high recurrence rates [89-91]. In most cases, bladder cancer appears as a non–muscle-invasive urothelial carcinoma [92]. Early detection of bladder cancer is critical for increased chance of bladder preservation and improved overall survival. Biomarkers play an important role in the early detection as they are the early events in the tumorigenesis [93]. The ideal biomarker must be reliable, inexpensive, minimally invasive, specific and sensitive to the disease [93, 94]. We are interested in developing blood-based markers for the early detection of bladder cancer.

Circulating cell free DNA

The presence of tumor DNA in circulating blood (plasma or serum) has been documented dating back to 1977 [95]. Cell-free DNA (cfDNA) was thought to be released from either apoptotic or necrotic cancer cells, from direct secretion or as a byproduct of phagocytosis from macrophages or other scavenger cells [95, 96]. With recent advances in next generation sequencing (NGS) technology, cfDNA has been explored extensively for its potential application to cancer detection [97]. In general, the studies of cfDNA as cancer biomarkers focus on monitoring the presence of methylation, aberrant tumor DNA mutation, microsatellite alternations, and mitochondria DNA in

blood circulation [96, 97]. In this current study, we are going to focus on aberrant methylation in circulating cell free DNAs.

DNA Methylation

In general, DNA methylation is thought to be an early event in tumorigenesis and has therefore been proposed as a potential early cancer detection marker [98, 99]. To date, accumulating data have shown that DNA methylation could be used to evaluate cancer risk [100-102], predict patients' prognosis [91, 101-103], and evaluate the response to chemotherapeutic drugs [102, 104]. Mostly, DNA methylation occurs at a cytosine in a CpG dinucleotide, by adding a methyl group to the number 5 carbon of the cytosine pyrimidine ring [105]. Aberrant DNA methylation has been associated with tumorigenesis as a consequence of the alteration it causes in gene expression [98, 106]; with hypermethylation associated with gene s, and hypomethylation with gene activation [107]. Meanwhile, DNA methylation detection is relatively easy due to PCR-based methods, which can compensate for the low content limitation [93], suggesting that methylation of circulating DNA may become a biomarker for cancer diagnosis [93, 94, 105].

CpG-SNP

single-nucleotide polymorphisms in CpG sites are referred as CpG-SNPs. About 34% of the C/T or G/A autosomal SNPs identified by the second-phase HapMap database

are located within a CpG site [105, 108], and some of these CpG SNPs have been related with diseases [105, 109-112]. CpG-SNP would be a potential molecular mechanism through which SNPs influence a phenotype of disease [105].

METHODS AND MATERIALS

Study design

An outline of the study design is shown in **Figure 3**. Firstly, whole-genome serum DNA methylation profiling was performed using the Infinium HumanMethylation450 BeadChip array (Illumina). 23 samples were included (15 cases and 8 controls) in the screening. We found 396 significantly hypermethylated CpG sites in cases and selected the top 5 for validation by pyrosequencing in 150 samples (50 controls, 50 non-muscle invasive and 50 muscle invasive bladder cancer patients).

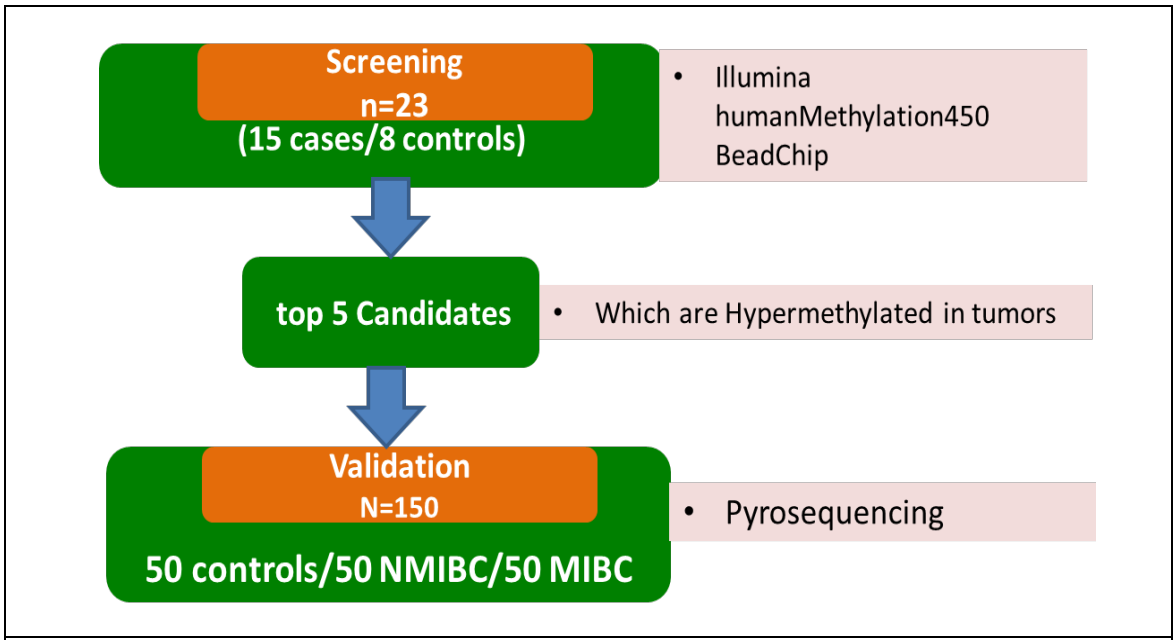

Fig. 3 Flow chart of the study design

Study population and epidemiologic data

The study population has been described elsewhere [52, 113-115]. Briefly, the population consisted of patients with newly diagnosed and histopathologically confirmed

bladder cancer accrued at the University of Texas MD Anderson Cancer Center and Baylor College of Medicine. The Controls were healthy individuals without a prior history of cancer (except nonmelanoma skin cancer) [116], which were selected from a large control pool recruited through a collaboration with the Kelsey-Seybold Clinic, the largest private multi-specialty physician group in Houston [117]. Informed consents were obtained from all the participants before collection of epidemiologic data and blood samples. Epidemiologic and demographic data were collected through in-person interviews by trained MD Anderson interviewers. All subjects were interviewed with a structured questionnaire [118]. All studies were approved by the Institutional Review boards of MD Anderson, Baylor College of Medicine, and Kelsey-Seybold Clinic [119].

Sample collection

All study participants signed informed consent and underwent an in-person interview by trained MD Anderson staff with a structured questionnaire that elicited information on epidemiologic data, including demographic characteristics, medical history and smoking history [120]. Immediately after each interview, a 40-mL peripheral blood sample was drawn. About 10 ml was drawn into a red top serum-separating tube, processed for serum extraction within 2 hours, and then transferred into liquid nitrogen tanks for long-term storage [120].

Circulating cell-free DNA Isolation

Cell-free circulating DNA was extracted from 1 ml of the serum with QIAamp Circulating Nucleic Acid Kit from Qiagen (Hilden, Germany) according to the manufacturer's instructions. The final elution volume was 20 μl and extracted DNA was quantified by NanoDrop ND-1000 spectrometer (Thermo Scientific, DE).

Bisulfite treatment of DNA

For samples used in chip screening, DNA was treated by bisulfite using EZ DNA Methylation Kit (Zymo Research, Orange, CA) following the manufacturer's protocol. In brief, 1 μg of DNA was treated with sodium bisulfite, through which all unmethylated cytosines were converted to uracils, but leaving methyl-cytosines unaltered; after purification, the cfDNA obtained was dissolved in 10 μl of TE buffer and this modified DNA was quantified by NanoDrop ND-1000 spectrometer (Thermo Scientific, DE); the efficiency of cfDNA recovery after bisulfite treatment was around 55% [121]. After bisulfite conversion, each sample was amplified, fragmented by enzymes, and hybridized to the Illumina Infinium HumanMethylation450 BeadChip [122].

For samples analyzed by pyrosequencing, 800 ng of genomic DNA was bisulfite-treated with EZ DNA Methylation-Gold Kit (Zymo Research, Irvine, CA) according to the manufacturer's protocol. To improve the efficiency of DNA elution, the purification columns were stored overnight in M-Elution Buffer at 4^{o}C prior to

centrifugation. The samples were eluted in 40 μl of M-Elution Buffer, and 2 μl (equivalent to 25 ng of bisulfite-modified DNA) was used for each PCR reaction.

Whole-Genome DNA methylation Profiling

Whole-genome serum DNA methylation profiling was performed on 23 serum samples (15 bladder cancer patients and 8 controls) using the Infinium HumanMethylation450 BeadChip (Illumina) according to the manufacturer's standard protocols. This BeadChip covers more than 485,000 methylation sites, located in 99% of RefSeq genes distributed across the promoter, UTR regions, first exon and gene body, and 3'-UTR regions [122, 123]. Samples were bisulfite converted using the EZ DNA methylation kit according to the manufacturer's protocol. Controls on the array were used to track the bisulfite conversion efficiency. The Infinium® HumanMethylation450 BeadChip array (Illumina Inc., San Diego, CA) was used to measure genome-wide CpG methylation levels using beads with specific probes designed to target individual CpG sites on bisulfite-converted genomic DNA [124].

Methylation analysis with pyrosequencing

The subsequent pyrosequencing analysis was done at the DNA Methylation Analysis Core, The University of Texas MD Anderson Cancer Center. Primers of pyrosequencing PCR and sequencing for the individual CpG sites were designed for the forward strand with Pyromark Assay Design SW 1.0 software (Qiagen, Hilden,

Germany), the reverse primer was biotin labeled at the 5'- end [125]. The primer sequences are shown in **Table 6**.

Table 6 Information of Pyrosequencing PCR and sequencing primers

Oligo Name	Sequence 5' to 3'	Sequence to analyze	PCR Product Size (bp)
CG26853071/F[1]	GAGGTTGGTTTAGTGGAAGTAAT	TGGGGYGTTGYGGAAGTTTTTG GGTAGAAA ATGAAA	141
CG26853071/Rbio[2]	[Btn]AACAATTTTCATTTTCTACCCAA AAACTTC		
CG26853071/S[3]	AGTTAGTTTTTGTTAATTTTGTGT		
CG05276469/F[1]	GTTAGGTTGTTTTAGGGATTGG	YGTAGTTTTTTGTYGG	113
CG05276469/Rbio[2]	[Btn]CCCCCCAACTAACCCTACCCA		
CG05276469/S[3]	TTGTTTAGAGGGGTTTAG		
CG01630032/F[1]	GTGTTTGAGATTAAGATTTATTTGGT AAGT	TAGGYGAGTGGAGGGGAGTTT TGTTAGGTA	177
CG01630032/Rbio[2]	[Btn]AACAAAACTCCCCTCCACT		
CG01630032/S[3]	AAAGTAGAAATTAATTTTAATGATT		
CG25576961/F[1]	TTAGTTTATTGTGTAAAATGTGGTTA GAG	GGGYGGGGGAGA	208
CG25576961/Rbio[2]	[Btn]AACCCCATCTACCACCTAAACTC TTACA		
CG25576961/S[3]	GAGGTTAGGATTAGTGG		

1. Forward primer for pyrosequencing PCR
2. Reverse prime with biotin label for pyrosequencing PCR
3. Sequencing primer

A sequencing primer is identified within 1 to 5 base pairs near the CpG site of interest, with an annealing temperature of 40 ± 5 ^{o}C. After that, forward and reverse primers are identified upstream and downstream to the sequencing primer, with a target annealing temperature ranging from 50 ^{o}C to 60 ^{o}C and amplicon product size ranging from 100bp to 200bp. Gradient PCR was used to detect the optimal annealing temperatures for these primers. Controls for high methylation (SssI-treated DNA), low methylation (WGA-amplified DNA), partial methylation (equimolar mixture of SssI-treated and WGA-amplified DNA) and no-DNA template were included in each reaction. For regions that were composed of mostly A and T bases, and additional control of WGA-amplified DNA that was subsequently SssI-treated prior to bisulfite conversion was also included, to account for biased representation of those regions in the whole genome amplified genome.

PCR reactions were performed in a total volume of 15 µl, which was used for the following pyrosequencing reaction (**Table 7, 8**). This method was previously described [126]. In brief, streptavidin-sepharose high-performance beads (GE Healthcare Life Sciences, Piscataway, NJ) were used to purify the PCR product. In addition, co-denaturation of the biotin-labeled PCR products and sequencing primer was performed following the guide of PSQ96 sample preparation. Sequencing was conducted on a PSQ HS 96 system (Biotage AB, Uppsala, Sweden) with the reagents of PyroMark Gold Q96 CDT (Qiagen, Hilden, Germany) according to the manufacturer's

instructions. The methylation level was calculated using the Pyro-Q CpG 1.0.9v software (Biotage AB, Uppsala, Sweden) [126].

Table 7. System of PCR reaction	
PCR Reaction	Volume (μL)
Bisulfite DNA	2
2x Zymo Reaction Buffer	7.5
dNTP Mix (25mM each)	0.15
F primer (10uM)	0.3
R primer (10uM)	0.3
Zymo Taq	0.12
ddH_2O	4.63

Table 8 PCR Cycling conditions		
Step	Temp (^oC)	Duration
1	95	10 m
2	95	30 s
3	53	30s
4	72	30s
5	Repeat (steps 2-4) 44 times (or 49 times, for CG6961)	
6	72	7 m
7	12	infinite

Identification of proxy SNPs

The analysis of proxy SNP was done by SNAP (SNP annotation and proxy) search [127] [30]. The search was based on genotype data released by the 1000 Genomes project and International HapMap Project. Inclusion criteria for LD (linkage disequilibrium) SNPs were set as an r^2 threshold >0.8 from pairwise LD calculations and a distance limit of 500 kb from the query SNP [105].

Genotyping

The method of genotyping has been described before [115]， the primary screen was performed using the Illumina HumanHap 610 chip.

Statistical Analysis

Bead Studio Methylation Module software (Illumina) was used to analyze the different methylation levels among different samples and groups [128]. The STATA software (STATA Corp., College Station, TX) was used for all the statistical analyses. Student's t- test was used to assess differences in DNA methylation levels between cases and controls. Differences in host characteristics between cases and controls were assessed by X^2 test for categorical variables (sex, smoking status, etc.) and the Student's t test for continuous variables (age). Unconditional multivariate logistic regression was performed with adjustment for covariates to calculate odds ratios and 95% confidence intervals (CI). All P values were two-sided, and the significance cut-off point was set as 0.05.

RESULTS

Whole genome DNA methylation profiling

In our study, circulating cell free DNA from serum samples of 23 patients (7 NMIBC, 8 MIBC, and 8 healthy controls) were used for whole-genome methylation profiling with Illumina Infinium HumanMethylation450 beadchip array. The methylation statuses of more than 485,000 CpG sites were assessed. Prior to further analyses, probes failed in any one of the 23 samples and probes in sex chromosomes were removed [128]

. As described before [129, 130], probes with β values being ≤0.2 or ≥0.8 in all samples were also excluded [128]. As a result, we got 275,054 probes for the further analysis, and obtained huge amount of CpG sites differentially methylated in bladder cancer patients compared with the healthy controls. Thus, we tried to narrow down the candidates by selecting those sites that were significantly hypermethylated in cases with $\Delta\beta > 0.2$ ($\Delta\beta = \beta case - \beta control$). As a result, we got 396 candidate CpG sites which are hypermethylated in cases compared with the healthy controls.

Validation of hypermethylated CpG sites by pyrosequencing

Five out of the hypermethylated 396 CpG sites from the microarray (**Table 9**) were selected for validation by bisulfite modification-based pyrosequencing assays in 150 serum samples (50 NMIBC, 50 MIBC, and 50 Controls). The characteristics of participants were summarized in **Table 10**.

Table 9 Five CpG sites are hypermethylated in cases in the screening

Target ID	Gene Symbol	DNA Methylation Level (β value)			P value
		Case (N=15) Mean(SD)	Control(N=8) Mean (SD)	Δβ=Case β -Control β	
cg26853071	LOC221122	0.72 (0.08)	0.40 (0.30)	0.32	**0.005**
cg20823859	OVGP1	0.56(0.17)	0.27 (0.17)	0.29	**0.001**
cg25576961	TTC23	0.38(0.17)	0.11 (0.05)	0.27	**0.001**
Cg05276469	ZNF624	0.59(0.12)	0.38(0.19)	0.21	**0.003**
cg01630032	WWOX	0.54(0.07)	0.34(0.07)	0.20	**0.005**

Table 10. Host characteristics of participants in validation by pyrosequencing

Variables	Overall			Cases		
	Case (n=100)	Control (n=50)	p	NMIBC (n=50)	MIBC (n=50)	p
Age (y)						
Mean(SD)	68.81(8.51)	68.94(8.13)	0.929	68.40(9.11)	69.22(7.92)	0.632
	n (%)	n (%)		n (%)	n (%)	
Gender						
Male	90(90.00)	45(90.00)		45(90.00)	45(90.00)	
Female	10(10.00)	5(10.00)	1	5(10.00)	5(10.00)	1
smoking status						
Never	25(25.00)	20(40.00)		12(24.00)	13(26.00)	
Ever	75(75.00)	30(60.00)	0.059	38(76.00)	37(74.00)	0.817
Death						
Yes				7(14.00)	10(20.00)	
No				43(86.00)	40(80.00)	0.424
Recurrence						
Yes				18(36.00)	1(2.00)	
No				32(64.00)	49(98.00)	**<0.001**
Progression						
Yes				15(30.00)	5(10.00)	
No				35(70.00)	45(90.00)	**0.012**
Treatment						
iBCG				21(42.00)		
mBCG				29(58.00)		
TUR+cystectomy					7(14.00)	
TUR+cystectomy+chemotherapy					14(28.00)	
TUR+chemotherapy					6(12.00)	
TUR					4(8.00)	
Others					19(38.00)	

Successful DNA methylation data were generated for 4 of the 5 CpG sites, located in genes of WWOX, LOC2211122, ZNF624 and TTC23 (**Figure 4**). For these 4 CpG sites, all the data passed the quality controls built into the analysis.

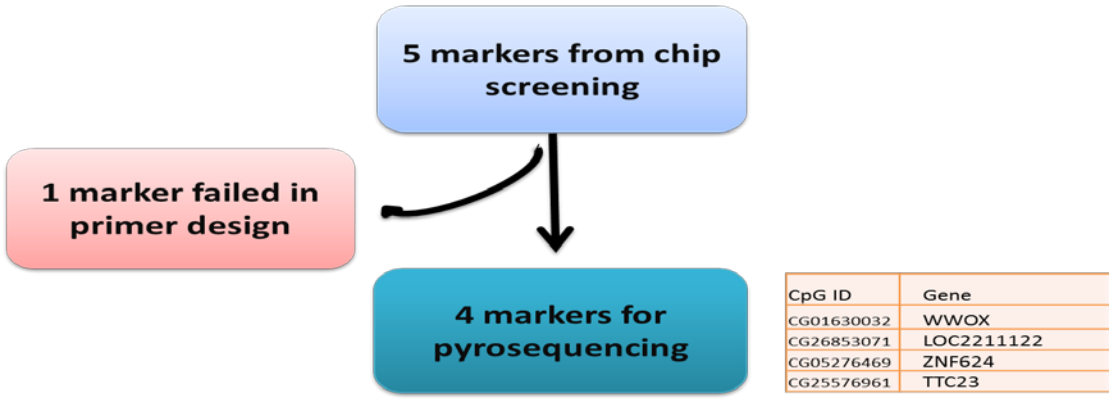

Fig. 4 Successful DNA methylation data were generated for 4 of the 5 CpG sites

There were two methylation sites in gene ZNF624, both of which were hypomethylated (**Figure 5**), so we used the average of methylation levels at the two sites for each sample for analysis.

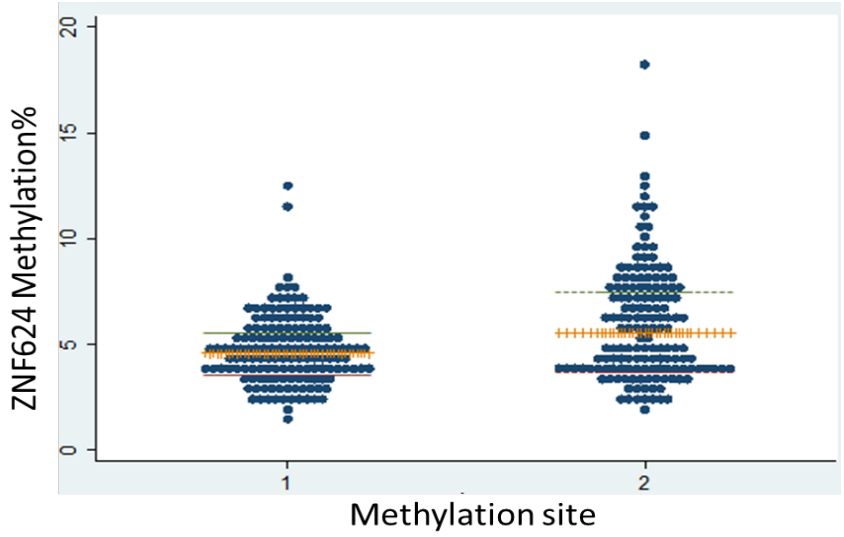

Fig. 5 Distribution of methylation levels of gene ZNF624 in different methylation sites

For gene WWOX, there is a SNP in the methylation site with a minor allele frequency (MAF) less than 0.01, which was ignored. In addition, there is a SNP

rs8038732 (G>A) with an MAF more than 0.42 in the methylation site of gene TTC23. The minor allele (A allele) would cause differential methylation by eliminating the CpG site, so the samples were grouped and analyzed according to the genotype of rs8038732. Unfortunately, none of these 4 sites exhibited differential methylation between cases and controls in our validation (**Table 11, 12)**

Table 11. Association of methylation levels of candidate target genes and bladder cancer

Gene	Methylation %			Methylation %		
	Cases (N=100)	Control(N=50)		NMIBC (N=42)	MIBC (N=42)	
	Mean (SD)	Mean (SD)	P	Mean (SD)	Mean (SD)	P
WWOX	43.31 (22.5)	41.45 (21.27)	0.656	43.24 (21.58)	43.39(23.59)	0.976
ZNF624	5.45 (1.92)	5.34 (1.83)	0.760	5.51 (1.73)	5.4(2.11)	0.789
LOC221112	91.11 (10.72)	91.1 (5.00)	0.995	90.37(14.62)	91.85(4.18)	0.531

Table 12. Association of methylation levels of TTC23 and bladder cancer

Gene	Genotype	Case (n=100)		control (n=50)			NMIBC (n=50)		MIBC(n=50)		
TTC23	rs8038732		Methy%		Methy%			Methy%		Methy%	
	(G>A)	n	Mean(SD)	n	Mean(SD)	p	n	Mean(SD)	n	Mean(SD)	P
	GG	26	91.48(4.20)	14	91.37(3.04)	0.93	16	92.95(2.28)	10	89.14(5.52)	0.055
	GA	53	47.67(6.80)	25	45.88(8.25)	0.313	25	46.92(9.02)	28	48.35(3.97)	0.450
	AA	21	1.74(2.18)	11	2.41(2.47)	0.431	9	2.27(2.88)	12	1.34(1.49)	0.346

CpG-SNP rs8038732 and methylation level of gene TTC23

Interestingly, the methylation level of TTC23 was associated with the genotype of SNP rs8038732 (G>A). As shown in **Figure 6**, the GG genotype showed near complete methylation, G/A heterozygotes half methylation, and AA genotype complete loss of methylation. This SNP is clearly a functional SNP because the G>A transition completely knocks out methylation at this site. It would be interesting to see whether this SNP affect bladder cancer risk.

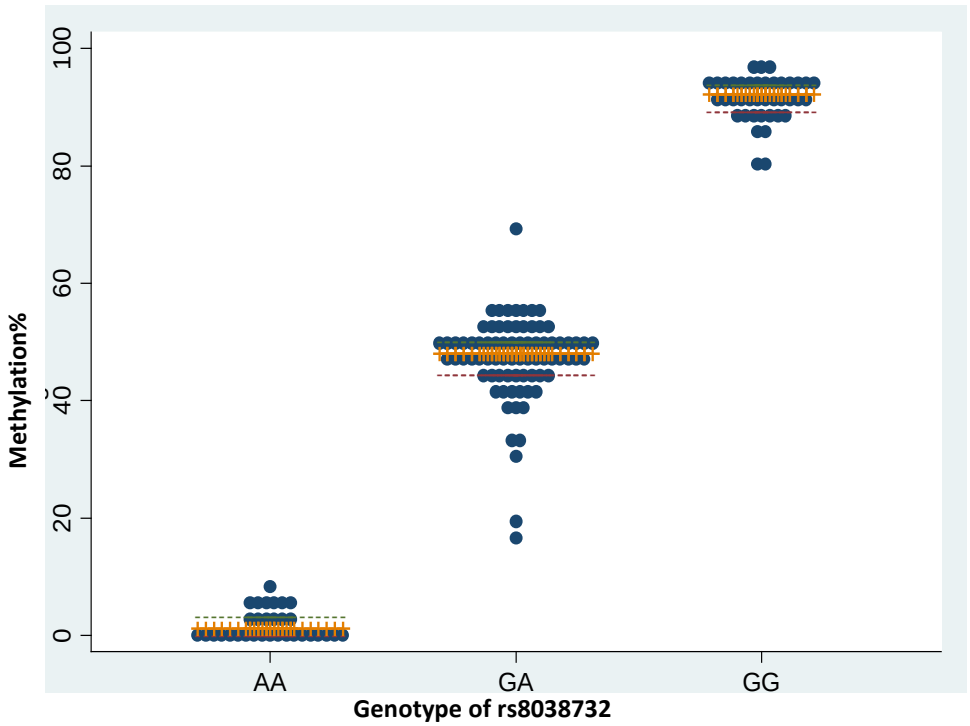

Fig. 6 rs8038732 affects methylation level of TTC23

Association of SNP rs8038732 with bladder cancer risk in GWAS study

SNP rs8038732 was not genotyped in our previously published genome-wide

association study (GWAS) [115]. However, we found that another SNP, rs1377267 (A>C), on the GWAS chip is in strong linkage disequilibrium (r^2=0.85) with rs8038732 (**Table 13**).

Table13 rs8038732 is in high LD with rs1377267

SNP	Linked	Distance	R-Squared	Chromosome	Coordinate_HG18
rs8038732	rs1377267	29908	0.85	chr15	97557412

The SNP rs1377267 was significantly associated with bladder cancer risk in our GWAS study. Specifically, the OR for individuals carrying one copy of the variant allele (C) was 1.34 (95% CI, 1.09–1.65), and for those carrying two copies was 0.78 (95% CI, 0.59–1.03). (**Table 14**)

Table14 Association of rs1377267 and bladder cancer risk

Genotype	Case	Control	OR (95% CI)	P
rs1377267(A>C)				
AA	306	341	1 Ref	
AC	528	426	1.34(1.09-1.65)	**0.005**
CC	135	190	0.78(0.59-1.03)	0.086

The association was also similar in superficial (**Table 15**) and invasive bladder cancer **(Table 16).**

Table 15 Association of rs1377267 and cancer risk in NMIBC patients

Genotype	NMIBC	Control	OR (95% CI)	P
rs1377267 (A>C)				
AA	154	341	Ref	
AC	278	426	1.42(1.11-1.83)	**0.005**
CC	65	190	0.73(0.52-1.05)	0.089

Table 16 Association of rs1377267 and cancer risk in MIBC patients				
Genotype	MIBC	Control	OR (95% CI)	P
rs1377267 (A>C)				
AA	128	341	Ref	
AC	218	426	1.28(0.98-1.69)	0.06
CC	55	190	0.76(0.52-1.11)	0.15

From this result, we could infer that rs8038732 has similar association with the risk of bladder cancer: the heterozygous allele change (GA) increases the bladder cancer risk significantly. The homozygous variant genotype (AA) is protective for bladder cancer risk, although the association is not statistically significant.

DISCUSSION

DNA methylation is one of the most common epigenetic alterations that associate with cancer development and progression [131]. Meanwhile, DNA methylation detection is relatively easy due to PCR based methods, which can compensate for the low content limitation, suggesting that methylation of circulating DNA may be a promising biomarker for cancer detection [93, 94, 128, 132].

In this study, 396 hypermethylated CpG sites were identified by Illumina HumanMethylation450 beadchip array. We selected 5 of the candidate CpG sites for individual validation by pyrosequencing. Successful methylation data were obtained for 4 of the selected CpG sites, located in genes of TTC23, WWOX, ZNF624 and LOC2211122. Unfortunately, none of these 4 sites exhibited differential methylation between cases and controls in this validation. To date, pyrosequencing has been a method with high specificity for DNA methylation analysis [128]. In our study, the pyrosequencing was performed with 25 ng of bisulfite-modified starting DNA, and all the results passed the quality controls. For the whole genome DNA methylation profiling, we started with 1 μg circulating cell free DNA from 1 ml of serum samples. After the bisulfite treatment, ~250 ng bisulfite -converted DNA was loaded to the Illumina HumanMethylation450 beadchip. There might be two reasons that we could not validate the primary screening results. First, the data from Illumina HumanMethylation450 beadchip array are not as accurate as pyrosequencing due to the low amount of input

DNA. Second, the sample size of the primary screening was small (23 samples) and there was high likelihood of false positives for the top statistically significant hits. It is fairly common that in large scale genetic and epigenetic biomarker studies, the validated biomarkers are not among the very top hits during the primary screening. In the future, we need to enlarge the sample size of primary screening and meanwhile select more candidate CpG sites for validation. Nevertheless, our study is the first to use Illumina HumanMethylation450 beadchip to profiling circulating cell free DNA, whereas previous studies have all used genomic DNA from tissue or peripheral blood. [124, 128, 133-135] Our study supports the feasibility of genome-wide CpG methylation profiling and validation of circulating cell-free DNA for the identification of early detection markers.

Interestingly, there was a SNP rs8038732 (G>A) in the methylation site of gene TTC23. The GG genotype showed near complete methylation (Mean of Methylation%=92.6), GA heterozygotes half methylation (Mean of Methylation%=43.4), and AA complete loss of methylation Mean of Methylation%=1.9). This SNP is clearly a functional SNP because the G>A transition completely knocks out methylation at this site. It was reported that altered DNA methylation could suppress tumor suppressor genes or activate oncogenes, thus initiate tumorigenesis and promote the progression of cancer [136]. Therefore, it would be interesting to see whether this SNP affects the risk of bladder cancer. Fortunately, we have data from our previously published GWAS of bladder cancer. Even though this specific SNP was not genotyped the GWAS [115], we

found another SNP rs1377267 (A>C), on the GWAS chip that is in strong linkage disequilibrium (r^2=0.85) with rs8038732. We then analyzed the associations of SNP rs1377267 with bladder cancer risk in the GWAS study of 969 bladder cancer cases and 946 controls. The SNP rs1377267 turned out to be significantly associated with bladder cancer risk. Specifically, the heterozygous alleles (AC) was associated with higher bladder cancer risk (OR=1.34, 95% CI=1.09-1.65), whereas the homozygous alleles (CC) was marginally associated with lower risk of bladder cancer (OR=0.78, 95% CI= 0.59-1.03). Since this SNP is highly linked with our interested SNP, we could infer that rs8038732 has similar association with the risk of bladder cancer.

The SNP rs8038732 is located in the intron region of the TTC23 gene, which was reported as a proto-oncogene. The TTC23 protein was also named Cervical Cancer Proto-Oncogene 8 Protein. A previous paper showed that TTC23 is a functional target gene of the BACH1 transcription factor, which is involved in the oxidative stress response and in cell cycle control [137]. In our study, individuals with the GA genotype had half methylation and higher bladder cancer risk; but individuals with the AA genotype who almost completely lost methylation had reduced risk of bladder cancer. One potential explanation for this observation is that the carcinogenesis is promoted by medium expression of oncogene TTC23, whereas cell death is induced by overexpression of TTC23 (**Figure 7**).

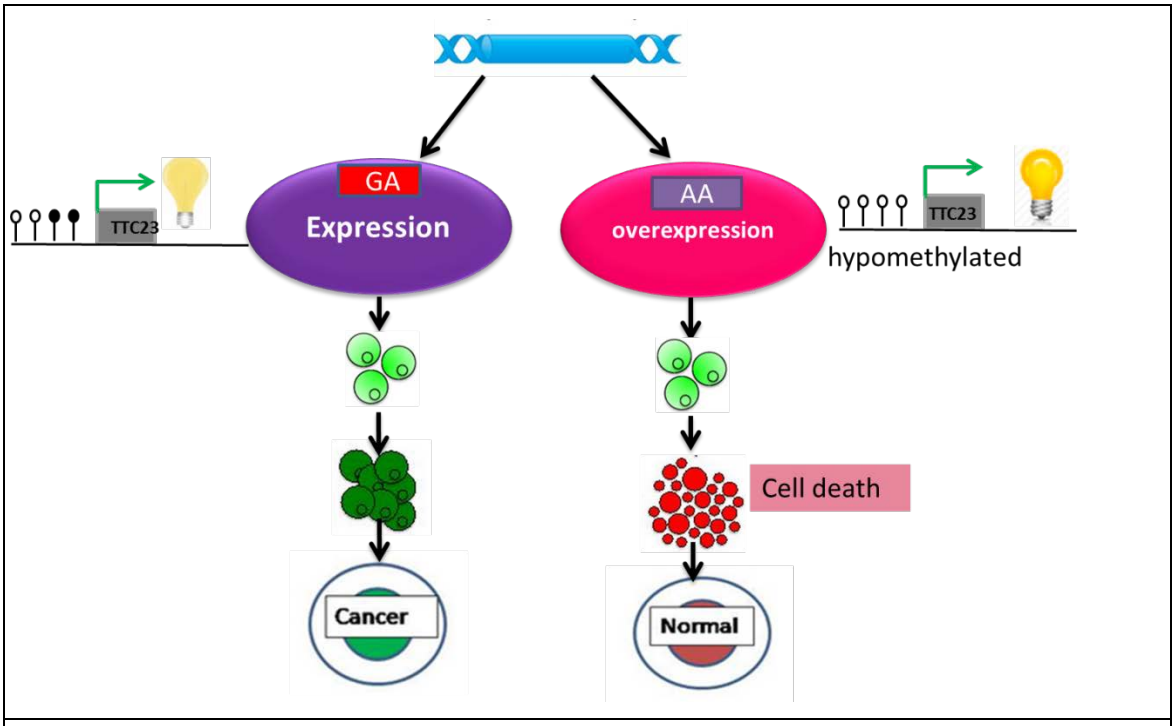

Fig. 7 Potential mechanism: overexpression of TTC23 could induce cell death

In general, oncogene is believed to promote carcinogenesis through increasing cell proliferation and survival. However, there has been evidence showing that oncogenes have anti-survival characteristics and contribute to tumor suppression. For instance, Ras was one of the first identified human proto-oncogenes. Ras is a membrane-anchored protein, which plays an important role in regulating cell differentiation, growth and apoptosis [138-145]. A study published in 1983 found the expression of Ras induced growth arrest in fibroblasts, which was one of the first indications that Ras may have dual functions in cell growth [145, 146]. Subsequent studies have shown that whether Ras signaling plays either a pro-apoptotic or a pro-growth role depends on the balance between the selectively deactivated or activated

downstream pathways [145]. The active overexpression of Ras was also reported to induce autophagic cell death [147]. Therefore, our observation that the TTC23 may be a risk or protective factor for bladder cancer etiology is consistent with these previous observations.

CONCLUSION

This study shows the feasibility of performing genome-wide CpG methylation profiling of ctDNA as a tool to screening biomarkers for bladder cancer early detection. A serendipitous observation was that a SNP rs8038732 in a CpG site of TTC23 is significantly associated with bladder cancer risk. This CpG-SNP affects local DNA methylation which, in turn, may affect gene expression and increase bladder cancer risk. Our data suggest that the abrogation of CpG dinucleotide, which is the site of DNA methylation, is a potential mechanism through which rs8038732 affects the risk of bladder cancer.

REFERENCES

1. Blackburn, E.H., *The telomere and telomerase: nucleic acid-protein complexes acting in a telomere homeostasis system. A review.* Biochemistry (Mosc), 1997. **62**(11): p. 1196-201.
2. Duggan, C., R. Risques, C. Alfano, D. Prunkard, I. Imayama, S. Holte, K. Baumgartner, R. Baumgartner, L. Bernstein, R. Ballard-Barbash, P. Rabinovitch, and A. McTiernan, *Change in peripheral blood leukocyte telomere length and mortality in breast cancer survivors.* J Natl Cancer Inst, 2014. **106**(4): p. dju035.
3. Starkweather, A.R., A.A. Alhaeeri, A. Montpetit, J. Brumelle, K. Filler, M. Montpetit, L. Mohanraj, D.E. Lyon, and C.K. Jackson-Cook, *An integrative review of factors associated with telomere length and implications for biobehavioral research.* Nurs Res, 2014. **63**(1): p. 36-50.
4. Eisenberg, D.T., *An evolutionary review of human telomere biology: the thrifty telomere hypothesis and notes on potential adaptive paternal effects.* Am J Hum Biol, 2011. **23**(2): p. 149-67.
5. Kurenova, E.V. and J.M. Mason, *Telomere functions. A review.* Biochemistry (Mosc), 1997. **62**(11): p. 1242-53.
6. Du, M., J. Prescott, P. Kraft, J. Han, E. Giovannucci, S.E. Hankinson, and I. De Vivo, *Physical activity, sedentary behavior, and leukocyte telomere length in women.* Am J Epidemiol, 2012. **175**(5): p. 414-22.
7. Lan, Q., R. Cawthon, M. Shen, S.J. Weinstein, J. Virtamo, U. Lim, H.D. Hosgood,

3rd, D. Albanes, and N. Rothman, *A prospective study of telomere length measured by monochrome multiplex quantitative PCR and risk of non-Hodgkin lymphoma.* Clin Cancer Res, 2009. **15**(23): p. 7429-33.

8. Hou, L., X. Zhang, A.J. Gawron, and J. Liu, *Surrogate tissue telomere length and cancer risk: shorter or longer?* Cancer Lett, 2012. **319**(2): p. 130-5.

9. Ma, H., Z. Zhou, S. Wei, Z. Liu, K.A. Pooley, A.M. Dunning, U. Svenson, G. Roos, H.D. Hosgood, 3rd, M. Shen, and Q. Wei, *Shortened telomere length is associated with increased risk of cancer: a meta-analysis.* PLoS One, 2011. **6**(6): p. e20466.

10. Wentzensen, I.M., L. Mirabello, R.M. Pfeiffer, and S.A. Savage, *The association of telomere length and cancer: a meta-analysis.* Cancer Epidemiol Biomarkers Prev, 2011. **20**(6): p. 1238-50.

11. Zhao, J., K. Miao, H. Wang, H. Ding, and D.W. Wang, *Association between telomere length and type 2 diabetes mellitus: a meta-analysis.* PLoS One, 2013. **8**(11): p. e79993.

12. Haycock, P.C., E.E. Heydon, S. Kaptoge, A.S. Butterworth, A. Thompson, and P. Willeit, *Leucocyte telomere length and risk of cardiovascular disease: systematic review and meta-analysis.* BMJ, 2014. **349**: p. g4227.

13. Cabarcas, S.M., E.M. Hurt, and W.L. Farrar, *Defining the molecular nexus of cancer, type 2 diabetes and cardiovascular disease.* Curr Mol Med, 2010. **10**(8): p. 744-55.

14. Quinlan, J., M.T. Tu, E.V. Langlois, M. Kapoor, D. Ziegler, H. Fahmi, and M.V. Zunzunegui, *Protocol for a systematic review of the association between chronic stress during the life course and telomere length.* Syst Rev, 2014. **3**: p. 40.

15. Skinner, E.H., M. Foster, G. Mitchell, M. Haynes, M. O'Flaherty, and T.P. Haines, *Effect of health insurance on the utilisation of allied health services by people with chronic disease: a systematic review and meta-analysis.* Aust J Prim Health, 2014. **20**(1): p. 9-19.

16. Amaral, S. and R. Patzer, *Disparities, race/ethnicity and access to pediatric kidney transplantation.* Curr Opin Nephrol Hypertens, 2013. **22**(3): p. 336-43.

17. Thomson, H., S. Thomas, E. Sellstrom, and M. Petticrew, *Housing improvements for health and associated socio-economic outcomes.* Cochrane Database Syst Rev, 2013. **2**: p. CD008657.

18. Chen, E. and G.E. Miller, *Socioeconomic status and health: mediating and moderating factors.* Annu Rev Clin Psychol, 2013. **9**: p. 723-49.

19. Sewell, J.L. and F.S. Velayos, *Systematic review: The role of race and socioeconomic factors on IBD healthcare delivery and effectiveness.* Inflamm Bowel Dis, 2013. **19**(3): p. 627-43.

20. Needham, B.L., N. Adler, S. Gregorich, D. Rehkopf, J. Lin, E.H. Blackburn, and E.S. Epel, *Socioeconomic status, health behavior, and leukocyte telomere length in the National Health and Nutrition Examination Survey, 1999-2002.* Soc Sci Med, 2013. **85**: p. 1-8.

21. Geronimus, A.T., M. Hicken, D. Keene, and J. Bound, *"Weathering" and age patterns of allostatic load scores among blacks and whites in the United States.* Am J Public Health, 2006. **96**(5): p. 826-33.

22. Koebnick, C., A.M. Langer-Gould, M.K. Gould, C.R. Chao, R.L. Iyer, N. Smith, W. Chen, and S.J. Jacobsen, *Sociodemographic characteristics of members of a large, integrated health care system: comparison with US Census Bureau data.* Perm J, 2012. **16**(3): p. 37-41.

23. Morales, L.S., M. Lara, R.S. Kington, R.O. Valdez, and J.J. Escarce, *Socioeconomic, cultural, and behavioral factors affecting Hispanic health outcomes.* J Health Care Poor Underserved, 2002. **13**(4): p. 477-503.

24. Caballero, A.E., *Understanding the Hispanic/Latino patient.* Am J Med, 2011. **124**(10 Suppl): p. S10-5.

25. Al Snih, S., M.N. Fisher, M.A. Raji, K.S. Markides, G.V. Ostir, and J.S. Goodwin, *Diabetes mellitus and incidence of lower body disability among older Mexican Americans.* J Gerontol A Biol Sci Med Sci, 2005. **60**(9): p. 1152-6.

26. Ventura, H., I.L. Pina, and C.J. Lavie, *Hypertension and antihypertensive therapy in Hispanics and Mexican Americans living in the United States.* Postgrad Med, 2011. **123**(6): p. 46-57.

27. Centers for Disease, C. and Prevention, *Hypertension-related mortality among Hispanic subpopulations--United States, 1995-2002.* MMWR Morb Mortal Wkly Rep, 2006. **55**(7): p. 177-80.

28. Schneider, M.G. and D.A. Chiriboga, *Associations of stress and depressive symptoms with cancer in older Mexican Americans.* Ethn Dis, 2005. **15**(4): p. 698-704.

29. Sullivan, M.M. and R. Rehm, *Mental health of undocumented Mexican immigrants: a review of the literature.* ANS Adv Nurs Sci, 2005. **28**(3): p. 240-51.

30. Marsiglia, F.F., J.M. Booth, A. Baldwin, and S. Ayers, *Acculturation and Life Satisfaction Among Immigrant Mexican Adults.* Adv Soc Work, 2013. **14**(1): p. 49-64.

31. MacKenzie, J.J., T.W. Smith, and B.N. Uchino, *Cardiovascular reactivity during stressful speaking tasks in Mexican-American women: effects of language use and interaction partner ethnicity.* J Behav Med, 2013. **36**(6): p. 621-31.

32. Gallo, L.C., S. Shivpuri, P. Gonzalez, A.L. Fortmann, K.E. de los Monteros, S.C. Roesch, G.A. Talavera, and K.A. Matthews, *Socioeconomic status and stress in Mexican-American women: a multi-method perspective.* J Behav Med, 2013. **36**(4): p. 379-88.

33. Shalev, I., S. Entringer, P.D. Wadhwa, O.M. Wolkowitz, E. Puterman, J. Lin, and E.S. Epel, *Stress and telomere biology: a lifespan perspective.* Psychoneuroendocrinology, 2013. **38**(9): p. 1835-42.

34. Babizhayev, M.A. and Y.E. Yegorov, *Smoking and health: association between telomere length and factors impacting on human disease, quality of life and life span in a large population-based cohort under the effect of smoking duration.*

Fundam Clin Pharmacol, 2011. **25**(4): p. 425-42.

35. Dorresteijn, J.A., F.L. Visseren, and W. Spiering, *Mechanisms linking obesity to hypertension.* Obes Rev, 2012. **13**(1): p. 17-26.

36. Hybertson, B.M., B. Gao, S.K. Bose, and J.M. McCord, *Oxidative stress in health and disease: the therapeutic potential of Nrf2 activation.* Mol Aspects Med, 2011. **32**(4-6): p. 234-46.

37. Teixeira-Lemos, E., S. Nunes, F. Teixeira, and F. Reis, *Regular physical exercise training assists in preventing type 2 diabetes development: focus on its antioxidant and anti-inflammatory properties.* Cardiovasc Diabetol, 2011. **10**: p. 12.

38. Ramos, L.F., A. Shintani, T.A. Ikizler, and J. Himmelfarb, *Oxidative stress and inflammation are associated with adiposity in moderate to severe CKD.* J Am Soc Nephrol, 2008. **19**(3): p. 593-9.

39. Russo, P., M. Fini, and A. Cesario, *Editorial: disease control and active and healthy ageing: new paradigms of therapeutic strategy.* Curr Pharm Des, 2014. **20**(38): p. 5919-20.

40. Green, J. and R. Jester, *Health-related quality of life and chronic venous leg ulceration: Part 2.* Br J Community Nurs, 2010. **15**(3): p. S4-6, S8, S10, passim.

41. Rossen, L.M., *Neighbourhood economic deprivation explains racial/ethnic disparities in overweight and obesity among children and adolescents in the U.S.A.* J Epidemiol Community Health, 2014. **68**(2): p. 123-9.

42. Gallo, L.C., A.L. Fortmann, K.E. de Los Monteros, P.J. Mills, E. Barrett-Connor, S.C. Roesch, and K.A. Matthews, *Individual and neighborhood socioeconomic status and inflammation in Mexican American women: what is the role of obesity?* Psychosom Med, 2012. **74**(5): p. 535-42.

43. Neighbors, C.J., D.X. Marquez, and B.H. Marcus, *Leisure-time physical activity disparities among Hispanic subgroups in the United States.* Am J Public Health, 2008. **98**(8): p. 1460-4.

44. Prevention, C.f.D.C.a., *Prevalence and Trends Data: Nationwide (States and DC) - 2011 Physical Activity* 2011.

45. Archer, N.P., A.V. Wilkinson, N. Ranjit, J. Wang, H. Zhao, A.C. Swann, and S. Shete, *Genetic, psychosocial, and demographic factors associated with social disinhibition in Mexican-origin youth.* Brain Behav, 2014. **4**(4): p. 521-30.

46. Chow, W.H., M. Chrisman, R.D. C, Y. Ye, H. Gomez, Q. Dong, C.E. Anderson, S. Chang, S. Strom, H. Zhao, and X. Wu, *Cohort Profile: The Mexican American Mano a Mano Cohort.* Int J Epidemiol, 2015.

47. de Heer, H.D., A.V. Wilkinson, L.L. Strong, M.L. Bondy, and L.M. Koehly, *Sitting time and health outcomes among Mexican origin adults: obesity as a mediator.* BMC Public Health, 2012. **12**: p. 896.

48. Wilkinson, A.V., M.R. Spitz, S.S. Strom, A.V. Prokhorov, C.H. Barcenas, Y. Cao, K.C. Saunders, and M.L. Bondy, *Effects of nativity, age at migration, and acculturation on smoking among adult Houston residents of Mexican descent.* Am

J Public Health, 2005. **95**(6): p. 1043-9.

49. Barcenas, C.H., A.V. Wilkinson, S.S. Strom, Y. Cao, K.C. Saunders, S. Mahabir, M.A. Hernandez-Valero, M.R. Forman, M.R. Spitz, and M.L. Bondy, *Birthplace, years of residence in the United States, and obesity among Mexican-American adults.* Obesity (Silver Spring), 2007. **15**(4): p. 1043-52.

50. Cawthon, R.M., *Telomere length measurement by a novel monochrome multiplex quantitative PCR method.* Nucleic Acids Res, 2009. **37**(3): p. e21.

51. Li, Z., J. Tang, H. Li, S. Chen, Y. He, Y. Liao, Z. Wei, G. Wan, X. Xiang, K. Xia, and X. Chen, *Shorter telomere length in peripheral blood leukocytes is associated with childhood autism.* Sci Rep, 2014. **4**: p. 7073.

52. Gu, J., M. Chen, S. Shete, C.I. Amos, A. Kamat, Y. Ye, J. Lin, C.P. Dinney, and X. Wu, *A genome-wide association study identifies a locus on chromosome 14q21 as a predictor of leukocyte telomere length and as a marker of susceptibility for bladder cancer.* Cancer Prev Res (Phila), 2011. **4**(4): p. 514-21.

53. Xie, H., X. Wu, S. Wang, D. Chang, R.E. Pollock, D. Lev, and J. Gu, *Long telomeres in peripheral blood leukocytes are associated with an increased risk of soft tissue sarcoma.* Cancer, 2013. **119**(10): p. 1885-91.

54. Bau, D.T., S.M. Lippman, E. Xu, Y. Gong, J.J. Lee, X. Wu, and J. Gu, *Short telomere lengths in peripheral blood leukocytes are associated with an increased risk of oral premalignant lesion and oral squamous cell carcinoma.* Cancer, 2013. **119**(24): p. 4277-83.

55. Carroll, J.E., A.V. Diez Roux, A.L. Fitzpatrick, and T. Seeman, *Low social support is associated with shorter leukocyte telomere length in late life: multi-ethnic study of atherosclerosis.* Psychosom Med, 2013. **75**(2): p. 171-7.

56. Epel, E.S., *Psychological and metabolic stress: a recipe for accelerated cellular aging?* Hormones (Athens), 2009. **8**(1): p. 7-22.

57. Nunn, A.V., G.W. Guy, J.S. Brodie, and J.D. Bell, *Inflammatory modulation of exercise salience: using hormesis to return to a healthy lifestyle.* Nutr Metab (Lond), 2010. **7**: p. 87.

58. Ji, L.L., J.R. Dickman, C. Kang, and R. Koenig, *Exercise-induced hormesis may help healthy aging.* Dose Response, 2010. **8**(1): p. 73-9.

59. Radak, Z., H.Y. Chung, E. Koltai, A.W. Taylor, and S. Goto, *Exercise, oxidative stress and hormesis.* Ageing Res Rev, 2008. **7**(1): p. 34-42.

60. Gomez-Pinilla, F., *The influences of diet and exercise on mental health through hormesis.* Ageing Res Rev, 2008. **7**(1): p. 49-62.

61. Ji, L.L., M.C. Gomez-Cabrera, and J. Vina, *Exercise and hormesis: activation of cellular antioxidant signaling pathway.* Ann N Y Acad Sci, 2006. **1067**: p. 425-35.

62. Radak, Z., H.Y. Chung, and S. Goto, *Exercise and hormesis: oxidative stress-related adaptation for successful aging.* Biogerontology, 2005. **6**(1): p. 71-5.

63. Kim, J.H., J.H. Ko, D.C. Lee, I. Lim, and H. Bang, *Habitual physical exercise has beneficial effects on telomere length in postmenopausal women.* Menopause, 2012.

19(10): p. 1109-15.

64. Puterman, E., J. Lin, E. Blackburn, A. O'Donovan, N. Adler, and E. Epel, *The power of exercise: buffering the effect of chronic stress on telomere length.* PLoS One, 2010. **5**(5): p. e10837.

65. Cherkas, L.F., J.L. Hunkin, B.S. Kato, J.B. Richards, J.P. Gardner, G.L. Surdulescu, M. Kimura, X. Lu, T.D. Spector, and A. Aviv, *The association between physical activity in leisure time and leukocyte telomere length.* Arch Intern Med, 2008. **168**(2): p. 154-8.

66. Rae, D.E., A. Vignaud, G.S. Butler-Browne, L.E. Thornell, C. Sinclair-Smith, E.W. Derman, M.I. Lambert, and M. Collins, *Skeletal muscle telomere length in healthy, experienced, endurance runners.* Eur J Appl Physiol, 2010. **109**(2): p. 323-30.

67. LaRocca, T.J., D.R. Seals, and G.L. Pierce, *Leukocyte telomere length is preserved with aging in endurance exercise-trained adults and related to maximal aerobic capacity.* Mech Ageing Dev, 2010. **131**(2): p. 165-7.

68. Cassidy, A., I. De Vivo, Y. Liu, J. Han, J. Prescott, D.J. Hunter, and E.B. Rimm, *Associations between diet, lifestyle factors, and telomere length in women.* Am J Clin Nutr, 2010. **91**(5): p. 1273-80.

69. Woo, J., N. Tang, and J. Leung, *No association between physical activity and telomere length in an elderly Chinese population 65 years and older.* Arch Intern Med, 2008. **168**(19): p. 2163-4.

70. Sjogren, P., R. Fisher, L. Kallings, U. Svenson, G. Roos, and M.L. Hellenius,

Stand up for health--avoiding sedentary behaviour might lengthen your telomeres: secondary outcomes from a physical activity RCT in older people. Br J Sports Med, 2014. **48**(19): p. 1407-9.

71. Muezzinler, A., A.K. Zaineddin, and H. Brenner, *Body mass index and leukocyte telomere length in adults: a systematic review and meta-analysis.* Obes Rev, 2014. **15**(3): p. 192-201.

72. Njajou, O.T., R.M. Cawthon, E.H. Blackburn, T.B. Harris, R. Li, J.L. Sanders, A.B. Newman, M. Nalls, S.R. Cummings, and W.C. Hsueh, *Shorter telomeres are associated with obesity and weight gain in the elderly.* Int J Obes (Lond), 2012. **36**(9): p. 1176-9.

73. Kotsopoulos, J., J. Prescott, I. De Vivo, I. Fan, J. McLaughlin, B. Rosen, H. Risch, P. Sun, and S.A. Narod, *Telomere length and mortality following a diagnosis of ovarian cancer.* Cancer Epidemiol Biomarkers Prev, 2014. **23**(11): p. 2603-6.

74. Wu, X., H. Zhao, R. Suk, and D.C. Christiani, *Genetic susceptibility to tobacco-related cancer.* Oncogene, 2004. **23**(38): p. 6500-23.

75. Burke, L.S., P.L. Hyland, R.M. Pfeiffer, J. Prescott, W. Wheeler, L. Mirabello, S.A. Savage, L. Burdette, M. Yeager, S. Chanock, I. De Vivo, M.A. Tucker, A.M. Goldstein, and X.R. Yang, *Telomere length and the risk of cutaneous malignant melanoma in melanoma-prone families with and without CDKN2A mutations.* PLoS One, 2013. **8**(8): p. e71121.

76. Gramatges, M.M., Q. Liu, Y. Yasui, M.F. Okcu, J.P. Neglia, L.C. Strong, G.T.

Armstrong, L.L. Robison, and S. Bhatia, *Telomere content and risk of second malignant neoplasm in survivors of childhood cancer: a report from the Childhood Cancer Survivor Study.* Clin Cancer Res, 2014. **20**(4): p. 904-11.

77. Heaphy, C.M., G.S. Yoon, S.B. Peskoe, C.E. Joshu, T.K. Lee, E. Giovannucci, L.A. Mucci, S.A. Kenfield, M.J. Stampfer, J.L. Hicks, A.M. De Marzo, E.A. Platz, and A.K. Meeker, *Prostate cancer cell telomere length variability and stromal cell telomere length as prognostic markers for metastasis and death.* Cancer Discov, 2013. **3**(10): p. 1130-41.

78. Kim, S., D.P. Sandler, G. Carswell, L.A. De Roo, C.G. Parks, R. Cawthon, C.R. Weinberg, and J.A. Taylor, *Telomere length in peripheral blood and breast cancer risk in a prospective case-cohort analysis: results from the Sister Study.* Cancer Causes Control, 2011. **22**(7): p. 1061-6.

79. Lan, Q., R. Cawthon, Y. Gao, W. Hu, H.D. Hosgood, 3rd, F. Barone-Adesi, B.T. Ji, B. Bassig, W.H. Chow, X. Shu, Q. Cai, Y. Xiang, S. Berndt, C. Kim, S. Chanock, W. Zheng, and N. Rothman, *Longer telomere length in peripheral white blood cells is associated with risk of lung cancer and the rs2736100 (CLPTM1L-TERT) polymorphism in a prospective cohort study among women in China.* PLoS One, 2013. **8**(3): p. e59230.

80. Mirabello, L., M. Garcia-Closas, R. Cawthon, J. Lissowska, L.A. Brinton, B. Peplonska, M.E. Sherman, and S.A. Savage, *Leukocyte telomere length in a population-based case-control study of ovarian cancer: a pilot study.* Cancer

Causes Control, 2010. **21**(1): p. 77-82.

81. Qin, Q., J. Sun, J. Yin, L. Liu, J. Chen, Y. Zhang, T. Li, Y. Shi, S. Wei, and S. Nie, *Telomere length in peripheral blood leukocytes is associated with risk of colorectal cancer in Chinese population.* PLoS One, 2014. **9**(2): p. e88135.

82. Sanchez-Espiridion, B., M. Chen, J.Y. Chang, C. Lu, D.W. Chang, J.A. Roth, X. Wu, and J. Gu, *Telomere length in peripheral blood leukocytes and lung cancer risk: a large case-control study in Caucasians.* Cancer Res, 2014. **74**(9): p. 2476-86.

83. Shen, M., R. Cawthon, N. Rothman, S.J. Weinstein, J. Virtamo, H.D. Hosgood, 3rd, W. Hu, U. Lim, D. Albanes, and Q. Lan, *A prospective study of telomere length measured by monochrome multiplex quantitative PCR and risk of lung cancer.* Lung Cancer, 2011. **73**(2): p. 133-7.

84. Varadi, V., A. Brendle, A. Brandt, R. Johansson, K. Enquist, R. Henriksson, U. Svenson, B. Tavelin, G. Roos, K. Hemminki, P. Lenner, and A. Forsti, *Polymorphisms in telomere-associated genes, breast cancer susceptibility and prognosis.* Eur J Cancer, 2009. **45**(17): p. 3008-16.

85. Weischer, M., B.G. Nordestgaard, R.M. Cawthon, J.J. Freiberg, A. Tybjaerg-Hansen, and S.E. Bojesen, *Short telomere length, cancer survival, and cancer risk in 47102 individuals.* J Natl Cancer Inst, 2013. **105**(7): p. 459-68.

86. Wu, X., C.I. Amos, Y. Zhu, H. Zhao, B.H. Grossman, J.W. Shay, S. Luo, W.K. Hong, and M.R. Spitz, *Telomere dysfunction: a potential cancer predisposition*

factor. J Natl Cancer Inst, 2003. **95**(16): p. 1211-8.

87. Seow, W.J., R.M. Cawthon, M.P. Purdue, W. Hu, Y.T. Gao, W.Y. Huang, S.J. Weinstein, B.T. Ji, J. Virtamo, H.D. Hosgood, 3rd, B.A. Bassig, X.O. Shu, Q. Cai, Y.B. Xiang, S. Min, W.H. Chow, S.I. Berndt, C. Kim, U. Lim, D. Albanes, N.E. Caporaso, S. Chanock, W. Zheng, N. Rothman, and Q. Lan, *Telomere length in white blood cell DNA and lung cancer: a pooled analysis of three prospective cohorts.* Cancer Res, 2014. **74**(15): p. 4090-8.

88. Butt, H.Z., G. Atturu, N.J. London, R.D. Sayers, and M.J. Bown, *Telomere length dynamics in vascular disease: a review.* Eur J Vasc Endovasc Surg, 2010. **40**(1): p. 17-26.

89. Burger, M., J.W. Catto, G. Dalbagni, H.B. Grossman, H. Herr, P. Karakiewicz, W. Kassouf, L.A. Kiemeney, C. La Vecchia, S. Shariat, and Y. Lotan, *Epidemiology and risk factors of urothelial bladder cancer.* Eur Urol, 2013. **63**(2): p. 234-41.

90. Burger, M., W. Oosterlinck, B. Konety, S. Chang, S. Gudjonsson, R. Pruthi, M. Soloway, E. Solsona, P. Sved, M. Babjuk, M.A. Brausi, C. Cheng, E. Comperat, C. Dinney, W. Otto, J. Shah, J. Thurof, and J.A. Witjes, *ICUD-EAU International Consultation on Bladder Cancer 2012: Non-muscle-invasive urothelial carcinoma of the bladder.* Eur Urol, 2013. **63**(1): p. 36-44.

91. Babjuk, M., M. Burger, R. Zigeuner, S.F. Shariat, B.W. van Rhijn, E. Comperat, R.J. Sylvester, E. Kaasinen, A. Bohle, J. Palou Redorta, and M. Roupret, *EAU guidelines on non-muscle-invasive urothelial carcinoma of the bladder: update*

2013. Eur Urol, 2013. **64**(4): p. 639-53.

92. Blick, C.G., S.A. Nazir, S. Mallett, B.W. Turney, N.N. Onwu, I.S. Roberts, J.P. Crew, and N.C. Cowan, *Evaluation of diagnostic strategies for bladder cancer using computed tomography (CT) urography, flexible cystoscopy and voided urine cytology: results for 778 patients from a hospital haematuria clinic.* BJU Int, 2012. **110**(1): p. 84-94.

93. Xiao, Y.F., X. Yong, Y.H. Fan, M.H. Lu, S.M. Yang, and C.J. Hu, *microRNA detection in feces, sputum, pleural effusion and urine: novel tools for cancer screening (Review).* Oncol Rep, 2013. **30**(2): p. 535-44.

94. Weber, J.A., D.H. Baxter, S. Zhang, D.Y. Huang, K.H. Huang, M.J. Lee, D.J. Galas, and K. Wang, *The microRNA spectrum in 12 body fluids.* Clin Chem, 2010. **56**(11): p. 1733-41.

95. Leon, S.A., B. Shapiro, D.M. Sklaroff, and M.J. Yaros, *Free DNA in the serum of cancer patients and the effect of therapy.* Cancer Res, 1977. **37**(3): p. 646-50.

96. Schwarzenbach, H., D.S. Hoon, and K. Pantel, *Cell-free nucleic acids as biomarkers in cancer patients.* Nat Rev Cancer, 2011. **11**(6): p. 426-37.

97. Kaiser, J., *Medicine. Keeping tabs on tumor DNA.* Science, 2010. **327**(5969): p. 1074.

98. Sandoval, J. and M. Esteller, *Cancer epigenomics: beyond genomics.* Curr Opin Genet Dev, 2012. **22**(1): p. 50-5.

99. Jones, P.A. and S.B. Baylin, *The fundamental role of epigenetic events in cancer.*

Nat Rev Genet, 2002. **3**(6): p. 415-28.

100. Peng, D., A. Belkhiri, T. Hu, R. Chaturvedi, M. Asim, K.T. Wilson, A. Zaika, and W. El-Rifai, *Glutathione peroxidase 7 protects against oxidative DNA damage in oesophageal cells.* Gut, 2012. **61**(9): p. 1250-60.

101. Lu, X.X., J.L. Yu, L.S. Ying, J. Han, S. Wang, Q.M. Yu, X.B. Wang, X.H. Fang, and Z.Q. Ling, *Stepwise cumulation of RUNX3 methylation mediated by Helicobacter pylori infection contributes to gastric carcinoma progression.* Cancer, 2012. **118**(22): p. 5507-17.

102. Ling, Z.Q., P. Lv, X.X. Lu, J.L. Yu, J. Han, L.S. Ying, X. Zhu, W.Y. Zhu, X.H. Fang, S. Wang, and Y.C. Wu, *Circulating Methylated DNA Indicates Poor Prognosis for Gastric Cancer.* PLoS One, 2013. **8**(6): p. e67195.

103. Yu, Q.M., X.B. Wang, J. Luo, S. Wang, X.H. Fang, J.L. Yu, and Z.Q. Ling, *CDH1 methylation in preoperative peritoneal washes is an independent prognostic factor for gastric cancer.* J Surg Oncol, 2012. **106**(6): p. 765-71.

104. Charlet, J., M. Schnekenburger, K.W. Brown, and M. Diederich, *DNA demethylation increases sensitivity of neuroblastoma cells to chemotherapeutic drugs.* Biochem Pharmacol, 2012. **83**(7): p. 858-65.

105. Dayeh, T.A., A.H. Olsson, P. Volkov, P. Almgren, T. Ronn, and C. Ling, *Identification of CpG-SNPs associated with type 2 diabetes and differential DNA methylation in human pancreatic islets.* Diabetologia, 2013. **56**(5): p. 1036-46.

106. Feinberg, A.P. and B. Vogelstein, *Hypomethylation distinguishes genes of some*

human cancers from their normal counterparts. Nature, 1983. **301**(5895): p. 89-92.

107. Bock, C., J. Walter, M. Paulsen, and T. Lengauer, *CpG island mapping by epigenome prediction.* PLoS Comput Biol, 2007. **3**(6): p. e110.

108. Sigurdsson, M.I., A.V. Smith, H.T. Bjornsson, and J.J. Jonsson, *HapMap methylation-associated SNPs, markers of germline DNA methylation, positively correlate with regional levels of human meiotic recombination.* Genome Res, 2009. **19**(4): p. 581-9.

109. Ling, C., P. Poulsen, S. Simonsson, T. Ronn, J. Holmkvist, P. Almgren, P. Hagert, E. Nilsson, A.G. Mabey, P. Nilsson, A. Vaag, and L. Groop, *Genetic and epigenetic factors are associated with expression of respiratory chain component NDUFB6 in human skeletal muscle.* J Clin Invest, 2007. **117**(11): p. 3427-35.

110. Taqi, M.M., I. Bazov, H. Watanabe, D. Sheedy, C. Harper, K. Alkass, H. Druid, P. Wentzel, F. Nyberg, T. Yakovleva, and G. Bakalkin, *Prodynorphin CpG-SNPs associated with alcohol dependence: elevated methylation in the brain of human alcoholics.* Addict Biol, 2011. **16**(3): p. 499-509.

111. Harlid, S., M.I. Ivarsson, S. Butt, S. Hussain, E. Grzybowska, J.E. Eyfjord, P. Lenner, A. Forsti, K. Hemminki, J. Manjer, J. Dillner, and J. Carlson, *A candidate CpG SNP approach identifies a breast cancer associated ESR1-SNP.* Int J Cancer, 2011. **129**(7): p. 1689-98.

112. Mill, J. and A. Petronis, *Molecular studies of major depressive disorder: the*

epigenetic perspective. Mol Psychiatry, 2007. **12**(9): p. 799-814.

113. Wu, X., J. Gu, H.B. Grossman, C.I. Amos, C. Etzel, M. Huang, Q. Zhang, R.E. Millikan, S. Lerner, C.P. Dinney, and M.R. Spitz, *Bladder cancer predisposition: a multigenic approach to DNA-repair and cell-cycle-control genes.* Am J Hum Genet, 2006. **78**(3): p. 464-79.

114. Hudmon, K.S., S.E. Honn, H. Jiang, R.M. Chamberlain, W. Xiang, G. Ferry, W. Gosbee, W.K. Hong, and M.R. Spitz, *Identifying and recruiting healthy control subjects from a managed care organization: a methodology for molecular epidemiological case-control studies of cancer.* Cancer Epidemiol Biomarkers Prev, 1997. **6**(8): p. 565-71.

115. Wu, X., Y. Ye, L.A. Kiemeney, P. Sulem, T. Rafnar, G. Matullo, D. Seminara, T. Yoshida, N. Saeki, A.S. Andrew, C.P. Dinney, B. Czerniak, Z.F. Zhang, A.E. Kiltie, D.T. Bishop, P. Vineis, S. Porru, F. Buntinx, E. Kellen, M.P. Zeegers, R. Kumar, P. Rudnai, E. Gurzau, K. Koppova, J.I. Mayordomo, M. Sanchez, B. Saez, A. Lindblom, P. de Verdier, G. Steineck, G.B. Mills, A. Schned, S. Guarrera, S. Polidoro, S.C. Chang, J. Lin, D.W. Chang, K.S. Hale, T. Majewski, H.B. Grossman, S. Thorlacius, U. Thorsteinsdottir, K.K. Aben, J.A. Witjes, K. Stefansson, C.I. Amos, M.R. Karagas, and J. Gu, *Genetic variation in the prostate stem cell antigen gene PSCA confers susceptibility to urinary bladder cancer.* Nat Genet, 2009. **41**(9): p. 991-5.

116. Ye, Y., S.M. Lippman, J.J. Lee, M. Chen, M.L. Frazier, M.R. Spitz, and X. Wu,

Genetic variations in cell-cycle pathway and the risk of oral premalignant lesions. Cancer, 2008. **113**(9): p. 2488-95.

117. Cassidy, A., W. Wang, X. Wu, and J. Lin, *Risk of urinary bladder cancer: a case-control analysis of industry and occupation.* BMC Cancer, 2009. **9**: p. 443.

118. Yang, H., C.P. Dinney, Y. Ye, Y. Zhu, H.B. Grossman, and X. Wu, *Evaluation of genetic variants in microRNA-related genes and risk of bladder cancer.* Cancer Res, 2008. **68**(7): p. 2530-7.

119. Lin, J., M.R. Forman, J. Wang, H.B. Grossman, M. Chen, C.P. Dinney, E.T. Hawk, and X. Wu, *Intake of red meat and heterocyclic amines, metabolic pathway genes and bladder cancer risk.* Int J Cancer, 2012. **131**(8): p. 1892-903.

120. Wang, Y., J. Gu, J.A. Roth, M.A. Hildebrandt, S.M. Lippman, Y. Ye, J.D. Minna, and X. Wu, *Pathway-based serum microRNA profiling and survival in patients with advanced stage non-small cell lung cancer.* Cancer Res, 2013. **73**(15): p. 4801-9.

121. Martinez-Galan, J., B. Torres-Torres, M.I. Nunez, J. Lopez-Penalver, R. Del Moral, J.M. Ruiz De Almodovar, S. Menjon, A. Concha, C. Chamorro, S. Rios, and J.R. Delgado, *ESR1 gene promoter region methylation in free circulating DNA and its correlation with estrogen receptor protein expression in tumor tissue in breast cancer patients.* BMC Cancer, 2014. **14**: p. 59.

122. Eriksson, A., M.J. Williams, S. Voisin, I. Hansson, A. Krishnan, G. Philippot, O. Yamskova, F.M. Herisson, R. Dnyansagar, G. Moschonis, Y. Manios, G.P.

Chrousos, P.K. Olszewski, R. Frediksson, and H.B. Schioth, *Implication of coronin 7 in body weight regulation in humans, mice and flies.* BMC Neurosci, 2015. **16**(1): p. 13.

123. Lokk, K., V. Modhukur, B. Rajashekar, K. Martens, R. Magi, R. Kolde, M. Koltsina, T.K. Nilsson, J. Vilo, A. Salumets, and N. Tonisson, *DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns.* Genome Biol, 2014. **15**(4): p. r54.

124. Bibikova, M., B. Barnes, C. Tsan, V. Ho, B. Klotzle, J.M. Le, D. Delano, L. Zhang, G.P. Schroth, K.L. Gunderson, J.B. Fan, and R. Shen, *High density DNA methylation array with single CpG site resolution.* Genomics, 2011. **98**(4): p. 288-95.

125. Krohn, A., T. Ahrens, A. Yalcin, T. Plones, J. Wehrle, S. Taromi, S. Wollner, M. Follo, T. Brabletz, S.A. Mani, R. Claus, B. Hackanson, and M. Burger, *Tumor cell heterogeneity in Small Cell Lung Cancer (SCLC): phenotypical and functional differences associated with Epithelial-Mesenchymal Transition (EMT) and DNA methylation changes.* PLoS One, 2014. **9**(6): p. e100249.

126. Estecio, M.R., S. Maddipoti, C. Bueso-Ramos, C.D. DiNardo, H. Yang, Y. Wei, K. Kondo, Z. Fang, W. Stevenson, K.S. Chang, S.A. Pierce, Z. Bohannan, G. Borthakur, H. Kantarjian, and G. Garcia-Manero, *RUNX3 promoter hypermethylation is frequent in leukaemia cell lines and associated with acute myeloid leukaemia inv(16) subtype.* Br J Haematol, 2015. **169**(3): p. 344-51.

127. Johnson, A.D., R.E. Handsaker, S.L. Pulit, M.M. Nizzari, C.J. O'Donnell, and P.I. de Bakker, *SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap.* Bioinformatics, 2008. **24**(24): p. 2938-9.

128. Xu, E., J. Gu, E.T. Hawk, K.K. Wang, M. Lai, M. Huang, J. Ajani, and X. Wu, *Genome-wide methylation analysis shows similar patterns in Barrett's esophagus and esophageal adenocarcinoma.* Carcinogenesis, 2013. **34**(12): p. 2750-6.

129. Laczmanska, I., P. Karpinski, M. Bebenek, T. Sedziak, D. Ramsey, E. Szmida, and M.M. Sasiadek, *Protein tyrosine phosphatase receptor-like genes are frequently hypermethylated in sporadic colorectal cancer.* J Hum Genet, 2013. **58**(1): p. 11-5.

130. Fang, W.J., Y. Zheng, L.M. Wu, Q.H. Ke, H. Shen, Y. Yuan, and S.S. Zheng, *Genome-wide analysis of aberrant DNA methylation for identification of potential biomarkers in colorectal cancer patients.* Asian Pac J Cancer Prev, 2012. **13**(5): p. 1917-21.

131. Ponomaryova, A.A., E.Y. Rykova, N.V. Cherdyntseva, T.E. Skvortsova, A.Y. Dobrodeev, A.A. Zav'yalov, L.O. Bryzgalov, S.A. Tuzikov, V.V. Vlassov, and P.P. Laktionov, *Potentialities of aberrantly methylated circulating DNA for diagnostics and post-treatment follow-up of lung cancer patients.* Lung Cancer, 2013. **81**(3): p. 397-403.

132. Greenberg, A.K., B. Rimal, K. Felner, S. Zafar, J. Hung, E. Eylers, B. Phalan, M. Zhang, J.D. Goldberg, B. Crawford, W.N. Rom, D. Naidich, and S. Merali,

S-adenosylmethionine as a biomarker for the early detection of lung cancer. Chest, 2007. **132**(4): p. 1247-52.

133. Ong, M.L. and J.D. Holbrook, *Novel region discovery method for Infinium 450K DNA methylation data reveals changes associated with aging in muscle and neuronal pathways.* Aging Cell, 2014. **13**(1): p. 142-55.

134. Joubert, B.R., S.E. Haberg, R.M. Nilsen, X. Wang, S.E. Vollset, S.K. Murphy, Z. Huang, C. Hoyo, O. Midttun, L.A. Cupul-Uicab, P.M. Ueland, M.C. Wu, W. Nystad, D.A. Bell, S.D. Peddada, and S.J. London, *450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy.* Environ Health Perspect, 2012. **120**(10): p. 1425-31.

135. Bady, P., D. Sciuscio, A.C. Diserens, J. Bloch, M.J. van den Bent, C. Marosi, P.Y. Dietrich, M. Weller, L. Mariani, F.L. Heppner, D.R. McDonald, D. Lacombe, R. Stupp, M. Delorenzi, and M.E. Hegi, *MGMT methylation analysis of glioblastoma on the Infinium methylation BeadChip identifies two distinct CpG regions associated with gene silencing and outcome, yielding a prediction model for comparisons across datasets, tumor grades, and CIMP-status.* Acta Neuropathol, 2012. **124**(4): p. 547-60.

136. Li, L., J.Y. Choi, K.M. Lee, H. Sung, S.K. Park, I. Oze, K.F. Pan, W.C. You, Y.X. Chen, J.Y. Fang, K. Matsuo, W.H. Kim, Y. Yuasa, and D. Kang, *DNA methylation in peripheral blood: a potential biomarker for cancer molecular epidemiology.* J Epidemiol, 2012. **22**(5): p. 384-94.

137. Warnatz, H.J., D. Schmidt, T. Manke, I. Piccini, M. Sultan, T. Borodina, D. Balzereit, W. Wruck, A. Soldatov, M. Vingron, H. Lehrach, and M.L. Yaspo, *The BTB and CNC homology 1 (BACH1) target genes are involved in the oxidative stress response and in control of the cell cycle.* J Biol Chem, 2011. **286**(26): p. 23521-32.

138. Karnoub, A.E. and R.A. Weinberg, *Ras oncogenes: split personalities.* Nat Rev Mol Cell Biol, 2008. **9**(7): p. 517-31.

139. Olson, M.F. and R. Marais, *Ras protein signalling.* Semin Immunol, 2000. **12**(1): p. 63-73.

140. Sasaki, A.T. and R.A. Firtel, *Regulation of chemotaxis by the orchestrated activation of Ras, PI3K, and TOR.* Eur J Cell Biol, 2006. **85**(9-10): p. 873-95.

141. Pemberton, L.F. and B.M. Paschal, *Mechanisms of receptor-mediated nuclear import and nuclear export.* Traffic, 2005. **6**(3): p. 187-98.

142. Krieglstein, K., S. Richter, L. Farkas, N. Schuster, N. Dunker, R.W. Oppenheim, and K. Unsicker, *Reduction of endogenous transforming growth factors beta prevents ontogenetic neuron death.* Nat Neurosci, 2000. **3**(11): p. 1085-90.

143. Subramaniam, S., J. Strelau, and K. Unsicker, *Growth differentiation factor-15 prevents low potassium-induced cell death of cerebellar granule neurons by differential regulation of Akt and ERK pathways.* J Biol Chem, 2003. **278**(11): p. 8904-12.

144. Borasio, G.D., J. John, A. Wittinghofer, Y.A. Barde, M. Sendtner, and R.

Heumann, *ras p21 protein promotes survival and fiber outgrowth of cultured embryonic neurons.* Neuron, 1989. **2**(1): p. 1087-96.

145. Ehrkamp, A., C. Herrmann, R. Stoll, and R. Heumann, *Ras and rheb signaling in survival and cell death.* Cancers (Basel), 2013. **5**(2): p. 639-61.

146. Newbold, R.F. and R.W. Overell, *Fibroblast immortality is a prerequisite for transformation by EJ c-Ha-ras oncogene.* Nature, 1983. **304**(5927): p. 648-51.

147. Denton, D., S. Nicolson, and S. Kumar, *Cell death by autophagy: facts and apparent artefacts.* Cell Death Differ, 2012. **19**(1): p. 87-95.

VITA

Lixia Han was born in Shandong, China on June, 9th, 1984, the daughter of Mingjing Han and Yurong Han. After completing her Bachelor of Science degree in Biology at Heze University at Shangdong, China in 2008, she entered the Capital Medical University in Beijing and received a Master of Science degree in Cell Biology in August, 2011. She then worked as an independent technician at Beijing Tongren Hospital for one year. In September of 2009 she entered the Ph.D. program of The University of Texas Graduate School of Biomedical Sciences (GSBS) at Houston. She performed her M.S. thesis research in Dr. Xifeng Wu's lab at the University of Texas MD Anderson Cancer Center.

Permanent address:

3626 Rose Water Ct

Manvel, Texas 77578