Genetic determinants of leukocyte telomere length in children: a neglected and challenging field

Maria G. Stathopoulou[a], Alexandros M. Petrelis[a], Jessica L. Buxton[b,c], Philippe Froguel[d,e], Alexandra I. F. Blakemore[b], Sophie Visvikis-Siest[a,f]*

a. UMR INSERM U1122; Interactions Gène-Environnement en Physiopathologie Cardio-Vasculaire (IGE-PCV), Université de Lorraine, Nancy, F-54000, France

b. Section of Investigative Medicine, Department of Medicine, Imperial College London, London W12 0NN, United Kingdom

c. Centre for Cardiovascular Genetics, UCL Institute of Cardiovascular Science, London WC1E 6JJ, United Kingdom

d. UMR CNRS 8199, Lille Pasteur Institute, Lille, F-59000, France

e. Genomic Medicine, Imperial College London, Hammersmith Hospital, London, W12 0NN, United Kingdom

f. Geriatric Service, Nancy University Hospital, Nancy, F-54511, France

*** Corresponding author:**

Dr. Visvikis-Siest Sophie

Research Director in INSERM

Director of UMR INSERM U1122; Interactions Gène-Environnement en Physiopathologie Cardio-Vasculaire (IGE-PCV), Université de Lorraine

30 Rue Lionnois, 54000, Nancy, France, Tel: +33(0)6.07.60.25.69; fax: +33(0)3.83.32.13.22, E-mail: Sophie.Visvikis-Siest@inserm.fr

ABSTRACT

Background: Telomere length is associated with a large range of human diseases. Genome-wide association studies (GWAS) have identified genetic variants that are associated with leucocyte telomere length (LTL). However, these studies are limited to adult populations. Nevertheless, childhood is a crucial period for the determination of LTL and the assessment of age-specific genetic determinants, although neglected, could be of great importance. Our aim was to provide insights and preliminary results on genetic determinants of LTL in children.

Methods: Healthy children (*n*=322, age range=6.75 - 17 years) with available GWAS data (Illumina Human CNV370-Duo array) were included. The LTL was measured using multiplex quantitative real-time PCR. Linear regression models adjusted for age, gender, parental age at child's birth and body mass index were used to test the associations of LTL with polymorphisms identified in adult GWAS and to perform a discovery-only GWAS.

Results: The previously GWAS-identified variants in adults were not associated with LTL in our paediatric sample. This lack of association was not due to possible interactions with age or gene × gene interactions. Furthermore, a discovery-only GWAS approach demonstrated 6 novel variants that reached the level of suggestive association ($P \leq 5 \times 10^{-5}$) and explain a high percentage of children's LTL.

Conclusions: The study of genetic determinants of LTL in children may identify novel variants not previously identified in adults. Studies in large-scale children populations are needed for the confirmation of these results, possibly through a childhood consortium that could better handle the methodological challenges of LTL genetic epidemiology field.

Key words: childhood, leukocyte telomere length, GWAS, polymorphisms

INTRODUCTION

Telomeres are structures made of tandem repeats of hexanucleotide units with telomeric-interacting proteins that are present at the ends of chromosome and they are responsible for the protection of genomic integrity. The loss of several base pairs from telomeres during each mitotic division leads to telomere shortening which after a critical limit results in telomere dysfunction that triggers cell cycle arrest, apoptosis and/or cellular senescence [1]. Leukocyte telomere length (LTL) has been correlated with a variety of human diseases such as cancer, cardiovascular diseases (CVDs) and CVD risk factors such as obesity, hypertension and diabetes [2].

LTL is a highly heritable trait, as indicated by previous studies (40-80%) [3, 4] and verified in the most recent meta-analysis by Broer et al (70%) [5], and a number of genome-wide association studies (GWAS) have assessed its biological determinants [6-12]. However, none of the studies included or referred to paediatric populations (<18 years). The rate of LTL attrition in infancy and early childhood is increased compared to adults, possibly due to increased growth rates [13]. A plateau is observed between 4 years old and early adulthood, followed by a gradual decrease [14]. Therefore, the study of genetic determinants of LTL in the age range of 4-18 years will likely identify the physiological factors that explain the high inter-individual variability of this trait. Furthermore, it is known that LTL attrition in adulthood is related to environmental factors such as obesity, smoking, inflammation and oxidative stress [15]. In healthy children, many of these environmental factors are absent, thus increasing our confidence that novel genetic determinants of LTL, non-modified by environment, may arise from studies in similar populations compared to adults. To our knowledge, cohorts of healthy children with GWAS data and LTL measurements are limited. Therefore, our aim was to test the association of the GWAS-identified polymorphisms from

adult studies in a paediatric cohort, in order to assess their possible effect on this age group, and to seek for novel genetic determinants through a preliminary GWAS approach.

METHODS

Subjects

A group of unrelated children (age <18 years) was selected from the STANISLAS Family Study, a 10-year longitudinal cohort of 1,006 European families free of chronic disease (cardiovascular or cancer). All participants and their parents gave a written informed consent. Protocols were approved by the local ethics committees for the protection of subjects for biomedical research [Comité Consultatif de Protection des Personnes dans la Recherche Biomédicale (CCPPRB)]. For the present study, the inclusion criteria were based on normal weight [defined as body mass index (BMI) <97th percentile for age and sex according to French reference cohort [16]] and complete availability of GWAS data and LTL measurements. The selection of normal weight children only was made with the purpose to avoid the effect of obesity and obesity-related factors to LTL [17]. The final sample size consisted of 322 children (age: 11.90±2.26 years, 50.6% males).

Genotyping

The genotyping was performed using the Illumina Human CNV370-Duo array, as previously described [18].

LTL measurements

The relative LTL was measured using monochrome multiplex quantitative real-time PCR (MM-qPCR) [19]. This approach is used in human epidemiological studies, since large numbers of samples can be measured in a high-throughput and cost-effective way that uses minimal amounts of DNA. The telomere measurement values obtained using qPCR have been shown to correlate with the Southern blotting method, which is regarded as the 'gold standard' method, and are less expensive [20, 21]. LTL measurements were carried out using the MM-

qPCR method with minor modifications, as described in detail elsewhere [17]. Briefly, all test samples were measured in duplicate, and five serial dilutions of the reference sample (leukocyte DNA from an adult female) spanning 5–50 ng were run in triplicate on each plate. The telomere amplicon signal (T) to single-copy gene control amplicon (S) ratios for each sample were calculated as T, the amount of reference DNA that matched the experimental sample for copy number of the telomere template, divided by S, that matched the copy number of the single-copy gene template. The mean coefficient of variation for the T/S measurements of duplicate samples was 4.5%. and the overall mean T/S ratio was 1.63±0.52. All T/S values were log-transformed before statistical analyses to ensure normal distribution of the data. As reported previously, and as expected, T/S measurements were negatively correlated with age (R= -0.17, *P*=0.002).

Statistical analyses

We tested the previous GWAS-identified single nucleotide polymorphisms (SNPs) from adults in our children cohort using a candidate-gene approach. We used a list of 14 SNPs summarized by the GWAS investigators HuGENavigator and the NHGRI Catalog of published GWAS (http://www.genome.gov/gwasstudies) for the telomere length trait and 7 SNPs from the most recent GWAS meta-analysis [6] (table 1). Among these, 10 were genotyped in our GWAS array and for the rest, proxy SNPs were used (table 1). In total, 19 polymorphisms were tested using linear regression models for the assessment of the associations with log-transformed T/S ratios, under an additive genetic model using PLINK, in a significance level of $P \leq 2.6 \times 10^{-3}$. The gene × environment interactions were tested using similar linear regression models with the addition of the interaction term in a significance level of $P \leq 2.6 \times 10^{-3}$. The epistatic interactions were tested using the BiForce toolbox, in a significance level of $P \leq 2.9 \times 10^{-4}$. The genome-wide associations were tested

using linear regression models under an additive genetic model using PLINK ($P = 5 \times 10^{-8}$). Conditional analyses in order to test the independence of association between SNPs and LTL were also assessed. All models were adjusted for age, gender, parental age at child's birth and BMI. Parental age has been shown to be inversely associated with LTL in children and was used as a covariate in the models [22]. The levels of significance were calculated using the Bonferroni formula ($P=\alpha/n$, where n is the number of tests for each stage). Power calculations were performed using QUANTO version 1.2.4.

RESULTS AND COMMENTS

Among the 19 polymorphisms, previously associated with LTL in adults, none was significantly associated with LTL in children. A possible explanation could be the effect of gene × age and gene × gene interactions that could mask the associations in childhood. However, no statistically significant interactions were observed either. Therefore, it seems that the genetic determinants of adults' LTL do not affect the trait in childhood. It is, thus, possible that the identified variants in adults may be related to telomere maintenance during life, whereas, the genetic component of childhood LTL reflects the telomere length at birth. Hence, these findings support the notion that non-identified genetic variants determine LTL in children. Although the study was not sufficiently powered for the identification of low frequency polymorphisms with weak effects on LTL, it has an adequate power for more common variants and medium effects. Therefore, we would be expect to observe some significant results if these associations truly existed in children. Nevertheless, the interactions results should be considered with caution due to power issues.

In order to further assess this idea, we also applied a GWAS approach in our population; however, due to the absence of other childhood studies with GWAS and LTL data we could only test our hypothesis in a GWAS discovery-only cohort.

In our discovery cohort, we did not identify any genome-wide significant associations. However, six SNPs reached the levels of suggestive association ($P \leq 5 \times 10^{-5}$) and could be tested in replication studies and meta-analyses. Conditional analysis revealed that each of these SNPs was independently associated with log T/S ratio ($P \leq 4.6 \times 10^{-4}$) (table 2). These SNPs have not been identified by GWAS in adult populations and this finding could support the hypothesis of differences in LTL genetic determinants between children and adults. Furthermore, these polymorphisms explain up to 23.9% of the variability of the phenotype in our sample (r^2 calculated from linear regression model which includes only the 6 SNPs).

Similarly to our study, previous GWAS on LTL in adults did not demonstrate genome-wide significant associations in their discovery cohorts [7, 8, 11, 12], indicating methodological challenges for the study of this trait. Our sample size, as well as the lack of a replication population, suggests a lack of statistical power. The study was sufficiently powered (>80%) to identify common variants (minor allele frequency, MAF >20%), with a relatively large effect (beta $\geq \pm 0.08$) in a genome-wide significance level. However, these effect sizes are not the rule in LTL studies as indicated by the previous GWAS findings in adults (ranging from 0.04-0.09 in the most recent study of Codd et al [6]). This is a limitation of the study of LTL through GWAS and these possible associations should be replicated in larger GWAS on paediatric populations before any conclusions could be considered reliable. Furthermore, due to the limited longitudinal data regarding LTL attrition rate in childhood, we cannot exclude the presence of genetic variability in the age range of the group used in this study.

LTL is a trait related to a large range of diseases and the identification of its genetic background is significant for the understanding of its dynamics and physiology and for the diseases prevention. The regulation of its levels during childhood is crucial and given its high heritable component, the determination of its genetic regulators in younger ages is of high significance. In this report, we have shown by a candidate-gene approach that the GWAS-

identified markers in adults are not associated with LTL in children, and that this result is not due to gene × age or gene × gene interactions. This observation was also verified by a preliminary GWAS study. Our results, although of limited statistical power, suggest a differentiated genetic regulation of LTL between adults and children, as the six identified SNPs are novel and have not been demonstrated in previous adult studies. Nevertheless, these findings need to be re-assessed and replicated in larger childhood cohorts.

Finally, we would like to stress the need for large consortia of investigators researching children with LTL data, which will likely be required to investigate this neglected and very challenging field of telomere biology research.

ACKNOWLEDGEMENTS

Samples and data are part of the biological resource center 'Interactions Gène-Environnement en Physiopathologie Cardio-Vasculaire' (IGE-PCV) in Nancy, France.

The study was funded by the "BioIntelligence" project. The Section of Investigative Medicine at Imperial College London is funded by grants from the MRC, BBSRC, NIHR, an Integrative Mammalian Biology (IMB) Capacity Building Award, an FP7- HEALTH- 2009-241592 EuroCHIP grant and is supported by the NIHR Imperial Biomedical Research Centre Funding Scheme. The work of JB and AIFB was also funded by The Wellcome Trust, MRC and Diabetes UK.

The funding sources had no involvement in the study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication. All authors declare that there are no conflicts of interest.

REFERENCES

1. Armanios M. Telomeres and age-related disease: how telomere biology informs clinical paradigms. *J Clin Invest*. 2013; 123:996-1002.
2. Calado RT, Young NS. Telomere diseases. *N Engl J Med*. 2009; 361:2353-2365.
3. Okuda K, Bardeguez A, Gardner JP, Rodriguez P, Ganesh V, Kimura M, et al. Telomere length in the newborn. *Pediatr Res*. 2002; 52:377-381.
4. Slagboom PE, Droog S, Boomsma DI. Genetic determination of telomere size in humans: a twin study of three age groups. *Am J Hum Genet*. 1994; 55:876-882.
5. Broer L, Codd V, Nyholt DR, Deelen J, Mangino M, Willemsen G, et al. Meta-analysis of telomere length in 19 713 subjects reveals high heritability, stronger maternal inheritance and a paternal age effect. *Eur J Hum Genet*. 2013; 21:1163-1168.
6. Codd V, Nelson CP, Albrecht E, Mangino M, Deelen J, Buxton JL, et al. Identification of seven loci affecting mean telomere length and their association with disease. *Nat Genet*. 2013; 45:422-427, 427e421-422.
7. Gu J, Chen M, Shete S, Amos CI, Kamat A, Ye Y, et al. 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; 4:514-521.
8. Codd V, Mangino M, van der Harst P, Braund PS, Kaiser M, Beveridge AJ, et al. Common variants near TERC are associated with mean telomere length. *Nat Genet*. 2010; 42:197-199.
9. Levy D, Neuhausen SL, Hunt SC, Kimura M, Hwang SJ, Chen W, et al. Genome-wide association identifies OBFC1 as a locus involved in human leukocyte telomere biology. *Proc Natl Acad Sci U S A*. 2010; 107:9293-9298.
10. Mangino M, Hwang SJ, Spector TD, Hunt SC, Kimura M, Fitzpatrick AL, et al. Genome-wide meta-analysis points to CTC1 and ZNF676 as genes regulating telomere homeostasis in humans. *Hum Mol Genet*. 2012; 21:5385-5394.
11. Mangino M, Richards JB, Soranzo N, Zhai G, Aviv A, Valdes AM, et al. A genome-wide association study identifies a novel locus on chromosome 18q12.2 influencing white cell telomere length. *J Med Genet*. 2009; 46:451-454.
12. Prescott J, Kraft P, Chasman DI, Savage SA, Mirabello L, Berndt SI, et al. Genome-wide association study of relative telomere length. *PLoS One*. 2011; 6:e19635.
13. Zeichner SL, Palumbo P, Feng Y, Xiao X, Gee D, Sleasman J, et al. Rapid telomere shortening in children. *Blood*. 1999; 93:2824-2830.
14. Frenck RW, Jr., Blackburn EH, Shannon KM. The rate of telomere sequence loss in human leukocytes varies with age. *Proc Natl Acad Sci U S A*. 1998; 95:5607-5610.
15. Aviv A. The epidemiology of human telomeres: faults and promises. *J Gerontol A Biol Sci Med Sci*. 2008; 63:979-983.
16. Rolland-Cachera MF, Cole TJ, Sempe M, Tichet J, Rossignol C, Charraud A. Body Mass Index variations: centiles from birth to 87 years. *Eur J Clin Nutr*. 1991; 45:13-21.
17. Buxton JL, Walters RG, Visvikis-Siest S, Meyre D, Froguel P, Blakemore AI. Childhood obesity is associated with shorter leukocyte telomere length. *J Clin Endocrinol Metab*. 2011; 96:1500-1505.
18. Froguel P, Ndiaye NC, Bonnefond A, Bouatia-Naji N, Dechaume A, Siest G, et al. A genome-wide association study identifies rs2000999 as a strong genetic determinant of circulating haptoglobin levels. *PLoS One*. 2012; 7:e32327.

19. Cawthon RM. Telomere length measurement by a novel monochrome multiplex quantitative PCR method. *Nucleic Acids Res*. 2009; 37:e21.
20. Aubert G, Hills M, Lansdorp PM. Telomere length measurement-caveats and a critical assessment of the available technologies and tools. *Mutat Res*. 2012; 730:59-67.
21. Aviv A, Hunt SC, Lin J, Cao X, Kimura M, Blackburn E. Impartial comparative analysis of measurement of leukocyte telomere length/DNA content by Southern blots and qPCR. *Nucleic Acids Res*. 2011; 39:e134.
22. Aviv A, Susser E. Leukocyte telomere length and the father's age enigma: implications for population health and for life course. *Int J Epidemiol*. 2013; 42:457-462.

TABLES

Table 1. GWAS-identified polymorphisms in adult populations and related proxies polymorphisms tested in the children cohort

Polymorphism	**Closest genes**	**Chr[a]**	**Reference**	**Genotyped or proxy polymorphism[b]**	**Rsquared**
rs2162440	*BRUNOL4, PIK3C3*	18	Mangino et al [11]	Yes	
rs12696304	*TERC*	3	Codd et al [8]	rs1997392	1
			Prescott et al J [12]		1
rs4452212	*CXCR4*	2	Levy et al [9]	rs10221893	1
rs2736428	*SLC44A4*	6		Yes	
rs1975174	*ZNF676*	19		rs11668269	1
rs4387287	*OBFC1*	10		NA[c]	
rs6028466	*DHX35*	20	Gu e al [7]	Yes	
rs398652	*PELI2*	14		Yes	
rs654128	*KPNA5*	6		Yes	
rs621559	*WDR65*	1		Yes	

rs3027234	*CTC1*	17	Mangino et al [10]	rs8075078	0.862
rs412658	*ZNF676*	19		rs10419926	1
rs1317082	*TERC*	3		rs10936599	1
rs9419958	*OBFC1*	10		Yes	
rs10936599	*TERC*	3	Codd et al [6]	Yes	
rs2736100	*TERT*	5		Yes	
rs7675998	*NAF1*	4		rs11100479	0.953
rs9420907	*OBFC1*	10		Yes	
rs8105767	*ZNF208*	19		rs7257051	0.853
rs755017	*RTEL1*	20		rs2281929	1
rs11125529	*ACYP2*	2		rs11890390	0.932

[a] **Chr,** chromosome, [b]**Yes,** for polymorphisms genotyped in the Illumina Human CNV370-Duo array or **rs number** for proxy polymorphisms genotyped in the Illumina Human CNV370-Duo array, [c]**NA,** not genotyped and no available proxy polymorphism in the Illumina Human CNV370-Duo array

Table 2. Conditional analysis of the SNPs with suggestive GWAS association ($P \leq 5 \times 10^{-5}$)

Chr[a]	SNP[b]	Minor allele	MAF[c]	P_{GWAS}	Location		Conditional analysis	
					Function	Closest genes	β (SE)[d]	*P*
2	rs10496920	G	0.18	5.00×10^{-5}	Intergenic	*LRP1B/LOC100129955*	0.05 (0.01)	4.60×10^{-4}
4	rs528983	G	0.11	2.93×10^{-5}	Intron	*NDST4*	-0.07 (0.01)	7.88×10^{-6}
6	rs594119	T	0.19	3.11×10^{-5}	Intron	*NKAIN2*	-0.05 (0.01)	2.19×10^{-5}
8	rs12678295	G	0.40	3.53×10^{-5}	Intergenic	*MYOM2/CSMD1*	-0.04 (0.01)	1.92×10^{-4}
21	rs2300383	G	0.47	1.88×10^{-5}	Intron	*ITSN1*	-0.04 (0.01)	7.42×10^{-6}
22	rs11703393	G	0.27	5.00×10^{-5}	Intron	*PARVB*	-0.04 (0.01)	1.69×10^{-4}

[a]**Chr,** chromosome, [b]**SNP,** single nucleotide polymorphism, [c]**MAF**, minor allele frequency, [d]β, effect size; **SE**, standard error; β coefficient refers to log T/S ratio