Observational Study

Medicine®

OPEN

Genome-wide identification of expression quantitative trait loci for human telomerase

Hanseol Kim, Jihye Ryu, PhD, Chaeyoung Lee, PhD*

Abstract

A genome-wide association study was conducted to identify expression quantitative trait loci (eQTL) for human telomerase.

We tested the genetic associations of nucleotide variants with expression of the genes encoding human telomerase reverse transcriptase (hTERT) and telomerase RNA components (TERC) in lymphoblastoid cell lines derived from 373 Europeans.

Our results revealed 6 eQTLs associated with hTERT ($P < 5 \times 10^{-8}$). One eQTL (rs17755753) was located in the intron 1 of the gene encoding R-spondin-3 (RSPO3), a well-known Wnt signaling regulator. Transcriptome-wide association analysis for these eQTLs revealed their additional associations with the expression of 29 genes ($P < 4.75 \times 10^{-6}$), including prickle planar cell polarity protein 2 (PRICKLE2) gene important for the Wnt signaling pathway. This concurs with previous studies in which significant expressional relationships between hTERT and some genes (β-catenin and Wnt-3a) in the Wnt signaling pathway have been observed.

This study suggested 6 novel eQTLs for hTERT and the association of hTERT with the Wnt signaling pathway. Further studies are needed to understand their underlying mechanisms to improve our understanding of the role of hTERT in cancer.

Abbreviations: EBV = Epstein–Barr virus, eQTL = expression quantitative trait locus, hTERT = human telomerase reverse transcriptase, LD = linkage disequilibrium, PRICKLE2 = prickle planar cell polarity protein 2, RPKM = reads per kilobase per million mapped reads, RSPO3 = R-spondin-3, SHPK = sedoheptulokinase, TERC = telomerase RNA components.

Keywords: expression quantitative trait locus, genome-wide association study, single nucleotide variant, telomerase, Wnt signaling pathway

1. Introduction

Telomeres play key roles in human genome stability through replenishing their G-rich sequences by telomerases.[1] A variety of human cancer cells show high expression levels of telomerase with short telomeres whereas most human somatic cells have low expression levels of telomerase.[2] The telomerases consist of human telomerase reverse transcriptases (hTERT) that regulate telomerase activity and telomerase RNA components (TERC) that are used as template RNAs for lengthening telomeres.[2]

It has been reported that mRNA expression level of hTERT is highly correlated with telomerase activity in tumor cells.[3] TERC are targeted by imetelstat sodium (GRN163L), an inhibitor of telomerase activity in cancer therapy.[4] Nevertheless, genetic factors that regulate the variability of gene expression are not well understood. In particular, knowledge on genetic factors is limited to genetic variability in or near the genes encoding hTERT and TERC. Some nucleotide variants (rs2736108 upstream of hTERT, rs7705526 in intron 2 of hTERT, and rs12696304 downstream of TERC) were associated with telomere length in leukocytes.[5,6] Some intronic variants (rs10069690, rs2242652, and rs7725218) of hTERT have been reported to have association with hTERT expression in prostate cancers.[7] This study aimed to identify nucleotide variants associated with mRNA expression of hTERT and TERC through genome-wide analysis.

2. Material and methods

2.1. Subjects

Expression data of the genes encoding hTERT and TERC in lymphoblastoid cell lines generated from the Geuvadis RNA-sequencing project[8] were used to identify expression quantitative trait loci (eQTLs). Cell lines were derived from 373 Europeans of the following 4 populations: Utah residents with northern and western ancestry (n = 91), Finns (n = 95), British (n = 94), and Toscani (n = 93).[8] We excluded Yoruba population from the project to avoid false positive associations produced by heterogeneous genetic background. Gene expression was calculated as the sum of reads per kilobase per million mapped reads (RPKM) for all transcripts of each gene in each individual.[8] Their corresponding genotypic data were obtained from the 1000 Genomes Project (http://www.1000genomes.org/). Genotypes with minor allele frequency <5% or with missing rate of >5% were removed. After the quality control, genotypes of 5,851,914 SNPs were used for final analysis. Ethical approval was not necessary because we dealt with publically available data.

Editor: Yong Liu.

Funding: This work was funded by the National Research Foundation of Korea, the Ministry of Education, Science, and Technology (Grant No. NRF-2012M3A9D1054705).

The authors have no conflicts of interest to disclose.

Supplemental Digital Content is available for this article.

Department of Bioinformatics and Life Science, Soongsil University, Seoul, Korea.

** Correspondence: Chaeyoung Lee, Department of Bioinformatics and Life Science, Soongsil University, 369 Sangdo-ro, Dongjak-gu, Seoul 06978, Korea (e-mail: clee@ssu.ac.kr).*

Medicine (2016) 95:42(e5209)

Received: 6 July 2016 / Received in final form: 5 September 2016 / Accepted: 24 September 2016

http://dx.doi.org/10.1097/MD.0000000000005209

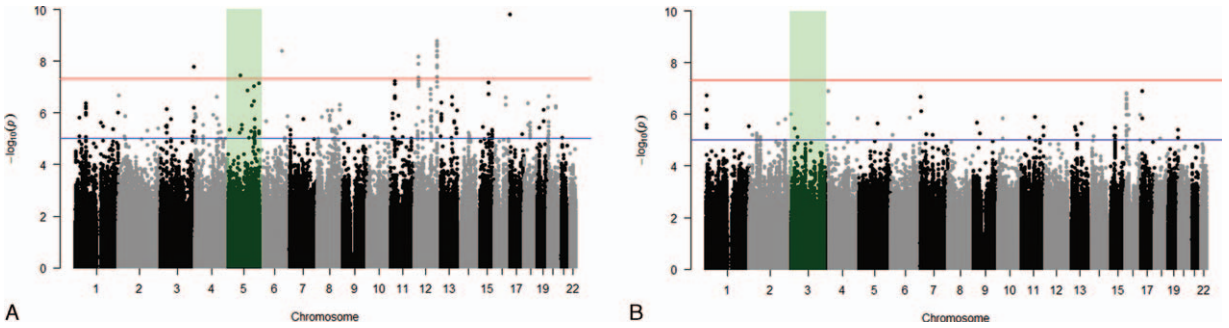

Figure 1. Manhattan plot for genome-wide association between eQTLs and mRNA expression of hTERT (A) or TERC (B). The red line indicates a genome-wide significance threshold ($P=5\times10^{-8}$), and the blue line indicates a threshold for "suggestive" variants ($P=10^{-5}$). The shadow indicates cis regulation. eQTLs = expression quantitative trait loci, hTERT = human telomerase reverse transcriptase, TERC = telomerase RNA components.

Table 1

Genome-wide associations of SNPs with mRNA expression of hTERT gene.

SNP	Position*	Gene†	Allele‡	MAF	BETA	*P*
rs224514§	17:3,625,779	SHPK (intron 2)	A/G	0.05	0.258	1.50×10^{-10}
rs112953754§	12:132,434,137	LOC105370092 (intron 1)	T/G	0.06	0.216	1.61×10^{-9}
rs76040610	12:132,427,895	LOC105370092 (ncRNA)	A/G	0.06	0.209	2.06×10^{-9}
rs111409387	12:132,436,426	LOC105370092 (intronic variant)	A/C	0.06	0.209	2.16×10^{-9}
rs77898203	12:132,438,006	LOC105370092 (upstream)	T/C	0.06	0.209	2.16×10^{-9}
rs79571493	12:132,438,677	LOC105370092 (upstream)	A/G	0.06	0.209	2.16×10^{-9}
rs113513971	12:132,438,964	LOC105370092 (upstream)	A/G	0.06	0.209	2.16×10^{-9}
rs75873120	12:132,438,995	LOC105370092 (upstream)	A/G	0.06	0.209	2.16×10^{-9}
rs79375673	12:132,439,573	Intergenic	T/C	0.06	0.209	2.16×10^{-9}
rs112669612	12:132,440,762	Intergenic	T/C	0.06	0.209	2.16×10^{-9}
rs76983641	12:132,441,633	Intergenic	A/G	0.06	0.209	2.16×10^{-9}
rs76618659	12:132,441,742	Intergenic	T/C	0.06	0.209	2.16×10^{-9}
rs112091336	12:132,442,010	Intergenic	A/G	0.06	0.209	2.16×10^{-9}
rs79593456	12:132,426,458	Intergenic	T/C	0.06	0.208	2.53×10^{-9}
rs77412187	12:132,434,661	LOC105370092 (intronic variant)	G/C	0.07	0.204	3.90×10^{-9}
rs113852564	12:132,433,798	LOC105370092 (intronic variant)	A/G	0.07	0.204	3.90×10^{-9}
rs76981325	12:132,433,656	LOC105370092 (intronic variant)	A/G	0.07	0.204	3.90×10^{-9}
rs112582060	12:132,431,962	LOC105370092 (intronic variant)	T/C	0.07	0.204	3.90×10^{-9}
rs74998815	12:132,431,325	LOC105370092 (intronic variant)	T/C	0.07	0.204	3.90×10^{-9}
rs117212853	12:132,428,205	LOC105370092 (ncRNA)	A/G	0.07	0.204	3.90×10^{-9}
rs17755753§	6:127,145,296	RSPO3 (intron 1)	C/A	0.07	0.194	3.91×10^{-9}
rs79151326	6:127,153,155	RSPO3 (intron 3)	A/G	0.07	0.194	3.91×10^{-9}
rs111852823	12:132,439,598	Intergenic	C/T	0.07	0.200	5.85×10^{-9}
rs4883600	12:132,438,452	LOC105370092 (upstream)	T/C	0.07	0.200	5.85×10^{-9}
rs113093837	12:132,432,383	LOC105370092 (intronic variant)	T/C	0.07	0.199	6.62×10^{-9}
rs35070061§	12:25,880,566	Intergenic	T/G	0.12	0.159	6.64×10^{-9}
rs117403427	12:25,870,968	Intergenic	T/C	0.11	0.163	1.27×10^{-8}
rs113285167	12:132,430,372	LOC105370092 (intronic variant)	G/T	0.07	0.185	1.41×10^{-8}
rs113633899	12:132,430,331	LOC105370092 (intronic variant)	A/G	0.07	0.185	1.41×10^{-8}
rs2636908§	3:193,958,627	LOC647323 (ncRNA)	C/T	0.07	0.189	1.66×10^{-8}
rs112575769	12:132,443,036	Intergenic	G/T	0.05	0.206	1.78×10^{-8}
rs187444335§	5:71,853,906	Intergenic	C/G	0.05	0.215	3.60×10^{-8}
rs75629604	12:132,438,841	LOC105370092 (upstream)	T/C	0.07	0.179	3.94×10^{-8}
rs73086133	12:25,871,823	Intergenic	A/G	0.13	0.147	4.30×10^{-8}
rs73086129	12:25,867,980	Intergenic	C/T	0.13	0.146	4.30×10^{-8}
rs34704732	12:25,884,449	Intergenic	T/G	0.13	0.146	5.00×10^{-8}
rs73088119	12:25,884,637	Intergenic	C/G	0.13	0.146	5.00×10^{-8}

eQTLs = expression quantitative trait loci, hTERT = human telomerase reverse transcriptase, MAF = minor allele frequency, SNP = single nucleotide polymorphism.

* Chromosome: chromosomal position (bp) from GRCh38.p2.

† The "upstream" in the parentheses indicates that the variant is located within 2 kb upstream of gene. The "ncRNA" indicates that the variant is located in a noncoding RNA gene. The "intronic variant" indicates that the variant is located in introns of LOC105370092, but the intronic number is uncertain because the gene is uncharacterized.

‡ Minor allele/major allele.

§ SNP in bold indicates representative variant with the most significance at each eQTL.

2.2. Statistical methods

Linear regression analysis was performed to discover eQTLs of hTERT and TERC. Multiple testing was employed with significance threshold value of 5×10^{-8}. All the statistical analyses were conducted using PLINK.[9] Linkage disequilibrium (LD) blocks at association signals were constructed using HaploView.[10] The identified eQTLs were analyzed to determine if they were located in transcription factor binding sites using ChIP-seq data from the regulomeDB.[11]

2.3. Transcriptome-wide association analysis of eQTLs for human telomerase reverse transcriptase

Further associations between 10,518 genes and SNPs identified to be associated with the expression of hTERT were examined. For transcriptome-wide association analysis, the significance threshold value was set at 4.75×10^{-6} (= 0.05 divided by the total number of 10,518 genes). Functional enrichment analysis was conducted with the identified genes to examine their functional relevance using the DAVID functional annotation tool.[12]

3. Results

Genome-wide association analysis revealed 37 eQTLs associated with mRNA expression of hTERT ($P < 5 \times 10^{-8}$; Fig. 1, Table 1). However, no significant eQTL was identified to be associated with mRNA expression of TERC ($P > 5 \times 10^{-8}$; Fig. 1). Some of the identified eQTLs located in chromosomes 6 and 12 turned out to be in strong linkage, and linkage disequilibrium blocks were constructed in Fig. 2. As a result, we found 6 association signals (rs224514, rs112953754, rs17755753, rs35070061, rs2636908, and rs187444335) from the eQTL analysis (Table 1). Among them, the intronic nucleotide variant rs224514 of the gene encoding sedoheptulokinase (SHPK) showed the most significant association signal ($P = 1.50 \times 10^{-10}$). Another intragenic association signal was located in intron 1 of R-spondin-3 (RSPO3) gene (rs17755753, $P = 3.91 \times 10^{-9}$).

Transcriptome-wide association analysis with the 6 eQTLs identified to have association with hTERT expression showed significant associations with mRNA expressions of additional 29 genes ($P < 4.75 \times 10^{-6}$; Table 2). No gene other than hTERT was identified to have significant association with rs35070061 ($P > 4.75 \times 10^{-6}$).

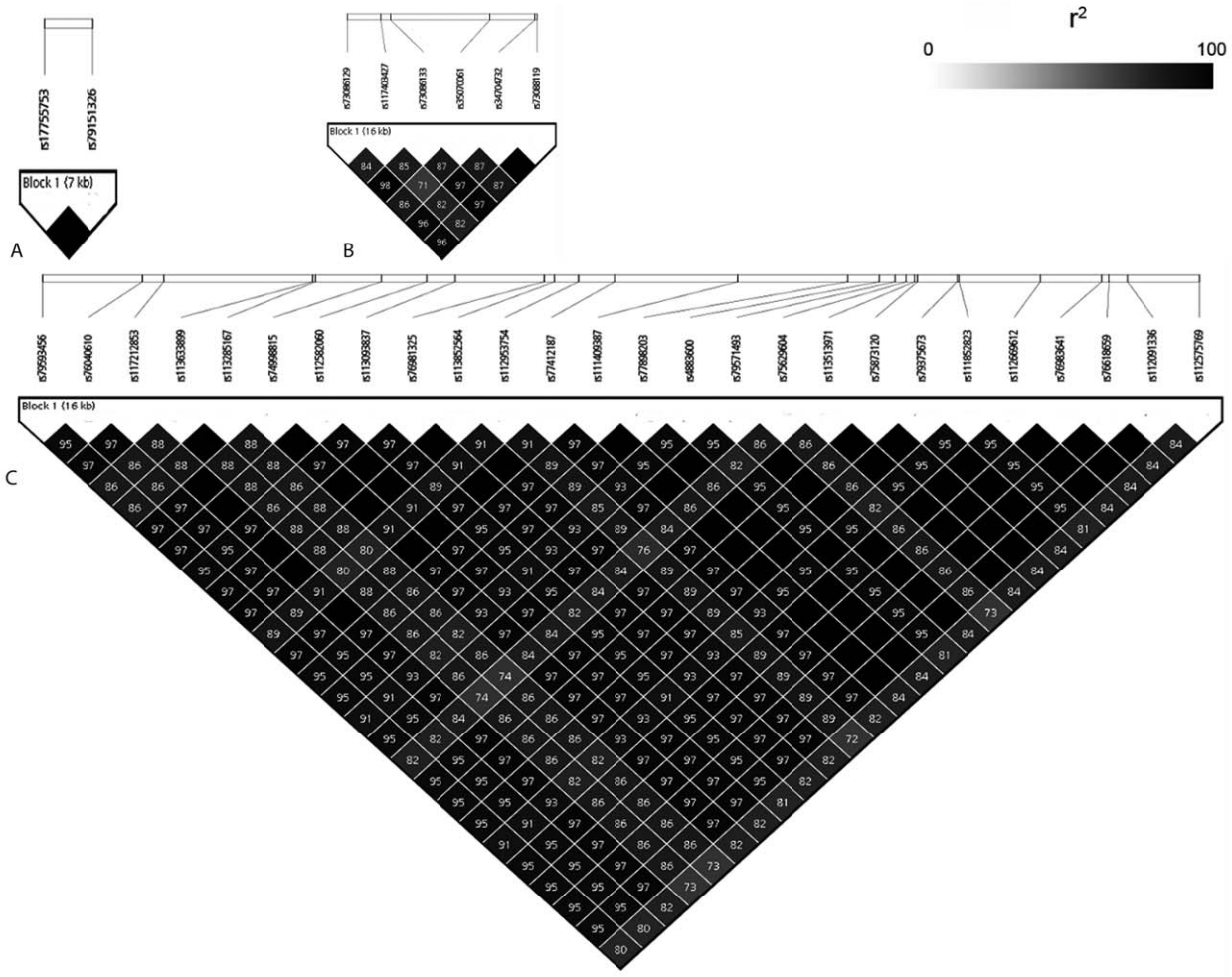

Figure 2. Linkage disequilibrium blocks for signals identified to have association with the expression of hTERT. (A) 6:127,145,296–127,153,155; (B) 12: 25,867,980–25,884,637; (C) 12:132,426,458–132,443,036. hTERT = human telomerase reverse transcriptase.

Table 2

Transcriptome-wide associations of genes with eQTLs identified for hTERT.

eQTL	Gene	Chromosome	BETA	*P*
rs224514	hTERT	5	0.258	1.50×10^{-10}
	LRIG1	3	7.529	6.76×10^{-7}
rs112953754	hTERT	5	0.216	1.61×10^{-9}
	CTNND2	5	0.285	5.95×10^{-8}
rs17755753	ADARB1	21	1.313	1.20×10^{-10}
	hTERT	5	0.194	3.91×10^{-9}
	C2orf89	2	0.263	3.42×10^{-8}
	LZTFL1	3	0.990	3.74×10^{-8}
	NLRP3	1	0.202	3.80×10^{-8}
	HHEX	10	0.997	8.89×10^{-8}
	MDK	11	1.593	3.09×10^{-7}
	INPP5F	10	1.662	3.62×10^{-7}
	FN1	2	0.521	9.53×10^{-7}
	PTGFRN	1	0.535	1.13×10^{-6}
	GLT25D2	1	0.210	2.31×10^{-6}
rs35070061	hTERT	5	0.159	6.64×10^{-9}
rs2636908	PRICKLE2	3	0.517	4.11×10^{-11}
	hTERT	5	0.189	1.66×10^{-8}
	ITGA11	15	0.274	2.99×10^{-8}
	LNX1	4	1.660	3.49×10^{-8}
	CHN1	2	0.193	4.22×10^{-8}
	CORO2B	15	0.192	4.89×10^{-8}
	GLT25D2	1	0.237	1.22×10^{-7}
	TSPAN2	1	0.267	1.42×10^{-7}
	FAM9C	X	2.505	1.42×10^{-7}
	MYO16	13	0.072	4.72×10^{-7}
	DIP2C	10	0.692	8.61×10^{-7}
	TFPI	2	0.159	8.98×10^{-7}
	ZNF365	10	0.255	9.00×10^{-7}
	COL1A1	17	0.584	1.05×10^{-6}
	SLIT1	10	0.626	1.10×10^{-6}
	C1QTNF1	17	0.111	1.23×10^{-6}
	PTPN3	9	0.297	1.77×10^{-6}
	CDC14B	9	0.422	2.40×10^{-6}
	AFAP1L2	10	1.886	4.48×10^{-6}
rs187444335	CTNND2	5	0.315	3.22×10^{-8}
	hTERT	5	0.215	3.60×10^{-8}

eQTL = expression quantitative trait locus, hTERT = human telomerase reverse transcriptase, PRICKLE2 = prickle planar cell polarity protein 2.

4. Discussion

The current study identified 6 genetic association signals that might regulate the expression of the gene encoding hTERT and explain partial variability of its expression. In particular, a signal located in the gene of RSPO3 was found. The association of RSPO3 with the expression of hTERT could be supported by previous studies. It has been shown that the R-spondin family can disrupt the inhibition of LRPs caused by DKK1, and thus LRPs activate the Wnt signaling pathway.[13,14] The RSPO3 can enhance angiogenesis and proliferation of human endothelial cells as a Wnt signaling regulator.[15] It has been reported that β-catenin, a critical intracellular signal transducer in the Wnt signaling pathway, can regulate the mRNA expression of hTERT.[16–18] Thus, RSPO3 can induce hTERT through the Wnt signaling pathway.

One of the 6 association signals covered 26 nucleotide variants in strong linkage within LOC105370092. Some of these variants were corresponding to a transcription factor binding sites uncovered by ChIP-Seq (RegulomeDB; Supplementary Table 1, http://links.lww.com/MD/B366), suggesting that these variants might regulate the expression of the uncharacterized ncRNA gene. Significant associations of these variants with the expression of CTNND2 were found ($P < 4.75 \times 10^{-6}$; Table 2), indicating that the ncRNA might influence the expression of hTERT and CTNND2. The RegulomeDB showed that rs113285167 was found in lymphoblastoid cells as binding sites of transcription factor 12 (TCF12) and CCCTC-binding factor (CTCF) (Supplementary Table 1, http://links.lww.com/MD/B366), which have opposite functions as a transcription factor and a repressor, respectively. The functionally contrasting proteins could bind the locus even in the same cell line (GM12878; ENCODE http://genome.ucsc.edu/ENCODE), but relative binding affinity remained unknown in these studies, especially by alleles. Further research on specific underlying mechanism on their interaction with binding sites is required to understand how gene expression is affected by alleles.

Transcriptome-wide association analysis revealed that the 6 signals associated with hTERT expression were further associated with the expression of many other genes. The most significant association was observed between prickle planar cell polarity protein 2 (PRICKLE2) and rs2636908. Since the PRICKLE2 gene and this SNP were both located in chromosome 3, a regulatory element for expression of PRICKLE2 might include this SNP or other SNPs strongly linked to it. The complex of PRICKLE2 with Vangl2 can regulate the Wnt/planar cell polarity (PCP) pathway that changes cytoskeleton.[19] Genes identified with eQTL of rs2636908 were enriched with cytoskeleton (Supplementary Table 2, http://links.lww.com/MD/B366). Acute withdrawals of TERT in mouse have triggered a rapid change in the expression of genes with the functions in the cytoskeleton,[20] suggesting that hTERT might be involved in cytoskeletal changes associated with the PCP pathway.

The current study dealt with expression of telomerase genes in lymphoblastoid cell lines transformed by Epstein–Barr virus (EBV).[8] The lymphoblastoid cell lines may be suitable for identifying eQTLs of telomerase genes since they demonstrated strong telomerase activity to be immortalized.[21–23] In EBV-infected B cell, expression of ectopic hTERT simultaneously increased with expression of basic leucine zipper ATF-like transcription factor (BATF) which maintains EBV latent status, and hTERT silenced by shRNA can induce B-cell death through transition into lytic cycle of EBV.[24] Thus, lymphoblastoid cells transformed by EBV can be pertinent to hTERT expression studies. Also, we suspect that the Wnt signaling pathway can be instrumental in EBV-transformed lymphoblastoid cells. This is because β-catenins are expressed in lymphoblastoid cells, and moreover, EBV-transformed lymphoblastoid cells increased the expression of β-catenin more than EBV-negative cells.[25–28]

Any eQTLs for TERC, another important component of human telomerase, were not found in this study ($P > 5 \times 10^{-8}$). However, many false negatives are suspected because we employed a conservative significance threshold. Some "suggestive" eQTLs for TERC are presented in Fig. 1 as potential false negatives ($10^{-5} > P > 5 \times 10^{-8}$).

The current GWAS identified 6 novel eQTLs for hTERT. Some of these eQTLs were involved in the Wnt signaling pathway critical to the production of tumors. Functional studies are needed in order to understand the underlying mechanisms of the Wnt signaling pathway influenced by hTERT. This will provide some evidence to unveil the reason why hTERT expression is associated with tumor.

References

[1] Blackburn EH, Epel ES, Lin J. Human telomere biology: a contributory and interactive factor in aging, disease risks, and protection. Science 2015;350:1193–8.

[2] Shay JW, Zou Y, Hiyama E, et al. Telomerase and cancer. Hum Mol Genet 2001;10:677–85.

[3] Yan P, Coindre JM, Benhattar J, et al. Telomerase activity and human telomerase reverse transcriptase mRNA expression in soft tissue tumors correlation with grade, histology, and proliferative activity. Cancer Res 1999;59:3166–70.

[4] Akiyama M, Hideshima T, Shammas MA, et al. Effects of oligonucleotide N3′→ P5′ thio-phosphoramidate (GRN163) targeting telomerase RNA in human multiple myeloma cells. Cancer Res 2003;63:6187–94.

[5] Codd V, Mangino M, van der Harst P, et al. Common variants near TERC are associated with mean telomere length. Nat Genet 2010;42: 197–9.

[6] Bojesen SE, Pooley KA, Johnatty SE, et al. Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. Nat Genet 2013;45:371–84.

[7] Kote-Jarai Z, Saunders EJ, Leongamornlert DA, et al. Fine-mapping identifies multiple prostate cancer risk loci at 5p15, one of which associates with TERT expression. Hum Mol Genet 2013;22:2520–8.

[8] Lappalainen T, Sammeth M, Friedländer MR, et al. Transcriptome and genome sequencing uncovers functional variation in humans. Nature 2013;501:506–11.

[9] Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–75.

[10] Barrett JC, Fry B, Maller JDMJ, et al. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21: 263–5.

[11] Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 2012;22:1790–7.

[12] Huang DW, Sherman BT, Tan Q, et al. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res 2007;35: W169–75.

[13] De Lau WB, Snel B, Clevers HC. The R-spondin protein family. Genome Biol 2012;13:242.

[14] Aoki M, Mieda M, Ikeda T, et al. R-spondin3 is required for mouse placental development. Dev Biol 2007;301:218–26.

[15] Kazanskaya O, Ohkawara B, Heroult M, et al. The Wnt signaling regulator R-spondin 3 promotes angioblast and vascular development. Development 2008;135:3655–64.

[16] Park JI, Venteicher AS, Hong JY, et al. Telomerase modulates Wnt signalling by association with target gene chromatin. Nature 2009;460: 66–72.

[17] Zhang Y, Toh L, Lau P, et al. Human telomerase reverse transcriptase (hTERT) is a novel target of the Wnt/β-catenin pathway in human cancer. J Biol Chem 2012;287:32494–511.

[18] Hoffmeyer K, Raggioli A, Rudloff S, et al. Wnt/β-catenin signaling regulates telomerase in stem cells and cancer cells. Science 2012;336: 1549–54.

[19] Nagaoka T, Ohashi R, Inutsuka A, et al. The Wnt/planar cell polarity pathway component Vangl2 induces synapse formation through direct control of N-cadherin. Cell Rep 2014;6:916–27.

[20] Choi J, Southworth LK, Sarin KY, et al. TERT promotes epithelial proliferation through transcriptional control of a Myc-and Wnt-related developmental program. PLoS Genet 2008;4:e10.

[21] Sugimoto M, Ide T, Goto M, et al. Reconsideration of senescence, immortalization and telomere maintenance of Epstein–Barr virus-transformed human B-lymphoblastoid cell lines. Mech Ageing Dev 1999;107:51–60.

[22] Sugimoto M, Tahara H, Ide T, et al. Steps involved in immortalization and tumorigenesis in human B-lymphoblastoid cell lines transformed by Epstein–Barr virus. Cancer Res 2004;64:3361–4.

[23] Takahashi T, Kawabe T, Okazaki Y, et al. In vitro establishment of tumorigenic human B-lymphoblastoid cell lines transformed by Epstein–-Barr virus. DNA Cell Biol 2003;22:727–35.

[24] Dolcetti R, Giunco S, Dal Col J, et al. Epstein–Barr virus and telomerase: from cell immortalization to therapy. Infect Agent Cancer 2014;9:8.

[25] Shackelford J, Maier C, Pagano JS. Epstein–Barr virus activates β-catenin in type III latently infected B lymphocyte lines: Association with deubiquitinating enzymes. Proc Natl Acad Sci 2003;100:15572–6.

[26] Morrison JA, Klingelhutz AJ, Raab-Traub N. Epstein-Barr virus latent membrane protein 2A activates β-catenin signaling in epithelial cells. J Virol 2003;77:12276–84.

[27] Morrison JA, Raab-Traub N. Roles of the ITAM and PY motifs of Epstein–Barr virus latent membrane protein 2A in the inhibition of epithelial cell differentiation and activation of β-catenin signaling. J Virol 2005;79:2375–82.

[28] Everly DN, Kusano S, Raab-Traub N. Accumulation of cytoplasmic β-catenin and nuclear glycogen synthase kinase 3β in Epstein–Barr virus-infected cells. J Virol 2004;78:11648–55.