Original Article

TP53 common variants and interaction with PPP1R13L and CD3EAP SNPs and lung cancer risk and smoking behavior in a Chinese population

Jiaoyang Yin a,*, Wei Hou b, Ulla Vogel c, Xinxin Li a, Yegang Ma d, Chunhong Wang a, Huiwen Wang a, Zhenxiang Sun a

a Key Laboratory of Environment and Population Health of Liaoning Education Ministry (Shenyang Medical College), Shenyang, Liaoning Province, People's Republic of China
b Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Department of Pathology, Peking University Cancer Hospital & Institute, Beijing, People's Republic of China
c National Research Centre for the Working Environment, DK-2100 Copenhagen, Denmark
d Department of Thoracic Surgery, Liaoning Cancer Hospital, Shenyang, Liaoning Province, People's Republic of China

Abstract

Background: TP53 encodes a tumor suppressor protein containing cell cycle arrest, apoptosis, senescence, DNA repair, or changes in metabolism. The effect of TP53 inactivation is well-known, and genetically determined smaller variations in TP53 activity are related to cancer. Lung cancer causes the highest rates of morbidity and mortality in the world. Epidemiology studies have assessed the association of TP53 single nucleotide polymorphisms with lung cancer.

Methods: We systematically examined the association of five htSNPs (haplotype-tagging single nucleotide polymorphism) (rs12951053, rs1042522, rs8079544, rs12602273 and rs8064946) across the entire TP53 locus and interaction between genes TP53 and PPP1R13L and CD3EAP and smoking-duration related to lung cancer risk in this Chinese study including 544 cases and 550 controls.

Results: No significant associations were observed in analysis of alleles and genotypes with co-dominant, dominant, recessive, and log-additive models after adjustment for smoking status. Haplotype analysis showed that haplotype9 (rs12951053A-rs1042522C-rs8079544C-rs12602273G-rs8064946C) [OR (95% CI) = 0.13 (0.03–0.59), p = 0.0079] was associated with decreased risk of lung cancer after adjusted for smoking-duration. The analysis of smoking-duration within TP53 haplotypes showed that there were more carriers of haplotype1 (AGCCG), 2 (CCCGC) and 4 (CCCCG) in smoking-subgroup of >20 (years) (all p < 0.05). MDR testing analysis identified two significant models (both p < 0.0010) of gene-environment interaction in relation to lung cancer risk in whole study group.

* Corresponding author. Key Laboratory of Environment and Population Health of Liaoning Education Ministry, Shenyang Medical College, 146 Huanghe North St., Shenyang, 110034, Liaoning Province, People's Republic of China.

E-mail address: yinjye@163.com (J. Yin).

Peer review under responsibility of Chang Gung University.

https://doi.org/10.1016/j.bj.2021.01.006

2319-4170/© 2021 Chang Gung University. Publishing services by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Lung cancer is malignant tumors that cause the highest rates of morbidity and mortality in the world [1]. Lung cancer is a complex polygenic disease. Smoking is the most important risk factor for lung cancer. Most patients with lung cancer have developed genetic mutations due to environmental exposure to carcinogens including smoking. Hereditary, genetic, and environmental factors interact in its genesis [2].

The gene tumor protein p53 (TP53, Aliases: BCC7, LFS1, P53, TRP53) (Gene ID: 7157) is located on chromosome 17p13.1 and includes 12 exons. TP53 encodes the tumor suppressor p53 containing transcriptional activation, DNA binding, and oligomerization domains. p53 responds to diverse cellular stresses to regulate expression of target genes, thereby inducing cell cycle arrest, apoptosis, senescence, DNA repair, or changes in metabolism. TP53 is the most commonly mutated gene in human cancers. Approximately half of all human malignancies exhibit TP53 mutations [https://www.ncbi.nlm.nih.gov/gene/7157, [3]]. While the effect of TP53 inactivation is well-known, genetically determined smaller variations in TP53 activity are also related to risk of cancer. Epidemiology studies have assessed the association of TP53 SNPs (single nucleotide polymorphism) with lung cancer [4–12]. However, the published study results are inconsistent [7,13,14].

Two genes governing biological function on Chr19q13.3, PPP1R13L (protein phosphatase 1, regulatory (inhibitor) subunit 13 like) (Gene ID: 10848), one of the most evolutionarily conserved inhibitors of TP53, is related to DNA repair and cell survival and CD3EAP (CD3e molecule, epsilon-associated protein) (Gene ID: 10849) may be related to cell proliferation. SNPs of PPP1R13L rs1970764 and CD3EAP rs967591 and rs735482 have been associated with lung cancer risk among both Caucasian Danes and Chinese in our previous studies [15–19].

TP53 and PPP1R13L and CD3EAP all belong to pathway of gene expression. TP53 and PPP1R13L share the same 7 pathways such as gene expression, generic transcription pathway, integrated pancreatic cancer pathway, regulation of TP53 activity, regulation of TP53 activity through association with cofactors, transcriptional regulation by TP53 and p53 pathway [https://www.ncbi.nlm.nih.gov/gene/7157, /10848, and /10849, assessed July 2019].

Furthermore, genetic factor of the TP53 hSNPs (haplotype-tagging single nucleotide polymorphism) and interactions of gene-gene and gene-environment related to lung cancer in the same biological pathways will provide important information about carcinogenesis and etiology of the disease. In the present Chinese case-control study of lung cancer, we assessed the association of TP53 hSNPs with lung cancer risk as well as gene-gene and gene-gene-smoking interactions. In addition, we explored potential association between TP53 hSNP haplotypes and smoking-related behaviors.

Conclusion: The present results provide novel evidence that the haplotype of TP53 hSNPs and interaction between genetic variation in TP53 and CD3EAP and smoking-duration may associate with lung cancer risk, and provide additional evidence of association between TP53 hSNP haplotypes and long-term smoking-related behavior.
chemotherapy or radiotherapy for cancer prior to recruitment). Cancer-free controls were selected from the orthopedic wards of Second Affiliated Hospital, Shenyang Medical College, P. R. China. Randomly selected controls were matched to the cases (1:1) by age (±3 years), gender (same) and ethnicity (same). All participants were unrelated ethnic Han Chinese. Stratification criteria were determined as follows: age (10 years an interval), smoking duration (20 years an interval) and histology (3 subgroups). All covariate data were obtained from questionnaires (or medical record) by interview (or extract) of professional doctors.

**htSNP choice in TP53**

We chose htSNPs of TP53 gene from the International HapMap Project (http://www.hapmap.org, HapMap Data Rel 27 Phasel+III, Feb09, on NCBI B36 assembly, dbSNP b26) using the TagSNPs software online and approaches of the algorithm-Tagger-pairwiseTagging on chr17:7512445..7531642, qualified criteria: r2-cut off of 0.8 and MAF (minor allele frequency)-cut off of 0.05 in CHB (Han Chinese in Beijing) samples. Five htSNPs (rs12951053, rs1042522, rs8079544, rs12602273, and rs8064946) were selected across the TP53 gene, representing 95% of the common haplotype diversity. Table 1 shows the information of TP53 five htSNPs and risk SNPs on Chr19q13.3 sub-region (PPP1R13L rs1970764 and CD3EAP rs967591 and rs735482). The genotype data of three risk SNPs on Chr19q13.3 were employed for interaction analyses of gene-gene and gene-environment in current study. The genotype data of three risk SNPs of Chr19q13.3 were previously reported [17,20]. CD3EAP rs735482 was re-genotyped for individuals who genotyping failed in the previous study [17].

**Table 1 Data for TP53 htSNPs selected and SNPs in PPP1R13L and CD3EAP**

| dbSNP ID | Position | Location | Base change | Allele frequency in HapMap HCB | MAF in controls for current study |
|----------|----------|----------|-------------|-------------------------------|-------------------------------|
| Chr17p13.1 |          |          |             |                               |                               |
| TP53     |          |          |             |                               |                               |
| rs12951053 | 7674089  | intron   | A/C         | A0.667/C0.333                 | C: 0.34                       |
| rs1042522 | 7676154  | exon4    | G/C         | G0.511/C0.489                 | C: 0.45                       |
| Codon 72 (R [Arg] [GCC]) ⇒ P [Pro] [CGC] (missense) |          |          |             |                               |                               |
| rs8079544 | 7676374  | intron   | C/T         | C0.878/T0.122                 | T: 0.08                       |
| rs12602273 | 7679695  | intron   | G/G         | G0.678/C0.322                 | G: 0.28                       |
| rs8064946 | 7685993  | intron   | G/C         | G0.622/C0.378                 | C: 0.32                       |
| Chr19q13.3 |          |          |             |                               |                               |
| PPP1R13L |          |          |             |                               |                               |
| rs1970764 | 45387615 | intron   | A/G         | No                            | G: 0.46                       |
| CD3EAP   |          |          |             |                               |                               |
| rs967591 | 45406676 | 5′ UTR   | G/A         | G0.525/A0.475                 | A: 0.39                       |
| rs735482 | 45408744 | exon3    | A/C         | A0.556/C0.444                 | C: 0.45                       |
| Codon 261 (K [Lys] [AAA] ⇒ T [Thr] [ACA]) (missense) |          |          |             |                               |                               |

a Information from NCBI SNP database (GRCh38.p7) and HapMap database.
b Minor allele frequency.
c Han Chinese in Beijing.
d CHB+JPT (Han Chinese in Beijing+ Japanese from 1000 GENOMES).

**DNA isolation and genotyping**

Genomic DNA of peripheral blood samples was extracted using the Puregene DNA Isolation Kit or FlexiGene DNA kit 250 (Gentra Systems, Minneapolis, MN, USA or Qiagen, Germany). The status of TP53 rs12951053, rs1042522, rs8079544, rs12602273, and rs8064946 and CD3EAP rs735482 was determined in the study participants using the genotyping assay of ligase detection reaction coupled with polymerase chain reaction (LDR-PCR) as previously published [20,21] in Shanghai Generay Biotechnology Co. Ltd. (P. R. China). The sequences (5′-3′) of primers and probes of TP53 htSNPs and CD3EAP rs735482 are showed in Supplemental Table S1. Each group of LDR probes contained 1 common probe and 2 discriminating probes for the 2 alleles. In brief: performed PCR reactions, completed LDR reactions and sequenced LDR products. The call rate of the genotyping was 93% on average for the five TP53 htSNPs. Repeated genotyping of a subset of the samples yielded 100% identity.

**Statistical analysis**

We conducted tests of general characteristics, allele frequencies, genotype frequencies, Hardy-Weinberg equilibrium, haplotype associations, and LD (pair-wise linkage disequilibrium) employing SPSS® v11.5 (SPSS Inc, Chicago, IL, USA), SNPStats program [22] and SHEsis software online [23]. We performed co-dominant model, dominant model, recessive model and log-additive model for case-control association of each single-locus employing SNPStats program [22]. We applied unconditional logistic regression for measurement of OR, 95% CI (odd ratio, 95% confidence interval) after adjustment for smoking duration. We excluded haplotypes with frequency < 0.01 among both cases and controls from the analysis. We completed the analyses of SNP-SNP and SNP-SNP-smoking duration interactions in relation to lung cancer risk employing MDR (multifactor dimensionality reduction) version 3.0.3. dev. Jar [24]. This software (3.0.3. dev. Jar) is an evolution version which has added permutation testing into the main MDR program. The MDR method is nonparametric and free model. MDR is directly useable to
Table 2 Distribution of selected characteristics in the case-control study population.

| Characteristics | Cases | Controls | p value |
|-----------------|-------|----------|---------|
|                 | n     | %        | n     | %        |         |
| Over all        | 544   |          | 550   |          |         |
| Age (years)     |       |          |       |          |         |
| Mean (±SD)      | 58 (±11) | 0.806<sup>a</sup> | 58 (±11) |          |
| <40             | 29     | 5.3      | 28    | 5.1      |
| 41–50           | 99     | 18.2     | 114   | 20.7     |
| 51–60           | 193    | 35.5     | 189   | 34.4     |
| >60             | 223    | 41.0     | 219   | 39.8     |
| Gender          |       |          |       |          |         |
| Female          | 158    | 29.0     | 161   | 29.3     |
| Male            | 386    | 71.0     | 389   | 70.7     |
| p value         | 0.93<sup>b</sup> |         |       |          |         |
| Smoking duration|       |          |       |          |         |
| Never           | 196    | 36.0     | 294   | 53.5     |
| <20 (years)     | 96     | 17.6     | 91    | 16.5     |
| >20 (years)     | 252    | 46.3     | 165   | 30.0     |
| Family history<sup>c</sup> | | | | |         |
| No              | 463    | 85.1     | 545   | 99.1     |
| Yes             | 81     | 14.9     | 5     | 0.9      |
| p value         | <0.0001<sup>b</sup> |         |       |          |         |
| Histology       |       |          |       |          |         |
| Squamous cell carcinoma | 232 | 42.6 |         |       |
| Adenocarcinoma  | 223    | 41.0     |       |         |
| Other           | 89     | 16.4     |       |         |

<sup>a</sup> For t-test.
<sup>b</sup> For χ² test (two-sided), boldface indicates statistical significance.
<sup>c</sup> Family history of cancer.

Results

This study comprised 544 lung cancer cases and 550 cancer-free controls. The general characteristics of the studied population are summarized in Table 2. There were no statistically significant differences for the distribution of age and gender between case group and control group. However, there were more cases than controls with family history of cancer and cases had longer smoking history (>20 years) than controls (both p < 0.0001).

In previous studies, CD3EAP rs735482 has been associated with lung cancer risk [18,19,17]. We therefore included this SNP in this expanded study population. Table 1 shows the following minor allele frequencies among controls in this population: rs12951053 C: 0.34, rs1042522 C: 0.45, rs8079544 T: 0.08, rs12602273 G: 0.28, and rs8064946 C: 0.32. These data are similar to the frequencies published in the HapMap-CHB of NCBI SNP database. All studied six SNPs were in Hardy-Weinberg equilibrium among controls (data not shown).

There were no significant associations between genotype distributions and lung cancer risk for any of the studied polymorphisms in co-dominant, dominant, recessive, and log-additive models after adjustment for smoking status and log-additive models after adjustment for smoking status and log-additive models after adjustment for smoking status and log-additive models after adjustment for smoking status and log-additive models after adjustment for smoking status. The haplotype distribution of the five TP53 htsSNPs was associated with lung cancer risk (Global haplotype association p-value = 0.0011) and haplotype9 (rs12951053A-rs1042522C-rs8079544C-rs12602273G-rs8064946C) [OR (95% CI) = 1.03 (0.03–0.59), p = 0.0079] was associated with decreased risk of lung cancer after adjusted for smoking duration. The analysis of smoking duration within TP53 haplotypes for 1037 subjects showed that there were more carriers of haplotype1 (AGCCG), 2 (CCCGC) and 4 (CCCCG) in the subgroup of smokers >20 (years) [OR (95% CI) = 1.90 (1.17–3.09), 2.22 (1.47–3.37), 2.65 (1.08–6.51), respectively, all p < 0.05] Table 5]. Combinatorial rare haplotypes consisting of different structures and very low frequencies showed statistical significances in both haplotype analyses [Tables 4 and 5]. MDR testing analysis of TP53, PPP1R13L, CD3EAP and smoking

Table 3 Associations of single htsNP in TP53 and CD3EAP rs735482 with lung cancer risk<sup>b</sup>.

| Gene/rs          | Co-dominant | Dominant | Recessive | Log-additive |
|------------------|-------------|----------|-----------|--------------|
|                  | (AB vs AA)/(BB vs AA)/p | (AB vs BB vs AA)/p | (BB vs AA+AB)/p | - / p         |
| TP53             |             |          |           |              |
| rs12951053 (A>C) |             |          |           |              |
| 509/516          | 0.97 (0.74–1.26) /0.91 (0.59–1.41) | 0.96 (0.74–1.23) /0.73 | 0.93 (0.61–1.40) /0.72 | 0.96 (0.79–1.16) /0.67 |
| rs1042522 (G>C)  |             |          |           |              |
| 489/489          | 1.03 (0.77–1.38) /1.00 (0.69–1.44) | 1.02 (0.77–1.35) /0.89 | 0.98 (0.71–1.34) /0.90 | 1.00 (0.84–1.20) /0.99 |
| rs8079544 (C>T)  |             |          |           |              |
| 509/516          | 1.03 (0.73–1.45) /2.57 (0.23–28.86) /0.72 | 1.05 (0.74–1.47) /0.80 | 2.56 (0.23–28.73) /0.43 | 1.06 (0.76–1.48) /0.72 |
| rs12602273 (C>G) |             |          |           |              |
| 509/516          | 0.94 (0.72–1.23) /0.69 (0.43–1.10) /0.30 | 0.89 (0.69–1.15) /0.37 | 0.70 (0.44–1.12) /0.13 | 0.88 (0.72–1.06) /0.18 |
| rs8064946 (G>C)  |             |          |           |              |
| 509/516          | 0.92 (0.71–1.19) /0.68 (0.44–1.06) /0.23 | 0.87 (0.68–1.12) /0.27 | 0.71 (0.46–1.09) /0.11 | 0.86 (0.71–1.04) /0.12 |
| CD3EAP rs735482 (A>C) |             |          |           |              |
| 522/511          | 1.15 (0.86–1.54) /1.25 (0.88–1.78) /0.43 | 1.18 (0.90–1.55) /0.23 | 1.15 (0.85–1.55) /0.37 | 1.12 (0.94–1.33) /0.20 |

<sup>a</sup> Dominant model: AB (Heterozygote) + BB (Homozygous variant-type) versus AA (Homozygous wild-type), Recessive model: BB versus AA + AB.
<sup>b</sup> Co-dominant model: AB versus AA and BB versus AA, Log-additive model: Analysis of trend where AA is ‘0’, AB is ‘1’ and BB is ‘2’.
<sup>c</sup> OR (95% CI), adjusted for smoking duration.
Table 4 Association of TP53 htSNP haplotypes with lung cancer risk\(^a\).

| Number | Haplotype\(^d\) | Case frequency | Control frequency | OR (95% CI) | p value |
|--------|-----------------|----------------|------------------|-------------|---------|
| 1      | AGCCG           | 0.5071         | 0.4753           | 1.0         | –       |
| 2      | CCGCC           | 0.2237         | 0.2210           | 0.96 (0.77–1.21) | 0.75 |
| 3      | ACTCG           | 0.0754         | 0.0632           | 1.11 (0.77–1.61) | 0.58 |
| 4      | CCGCC           | 0.0572         | 0.0560           | 0.96 (0.63–1.47) | 0.86 |
| 5      | ACCGG           | 0.0432         | 0.0469           | 0.93 (0.59–1.44) | 0.73 |
| 6      | CCCCG           | 0.0404         | 0.0333           | 1.11 (0.66–1.87) | 0.68 |
| 7      | ACGGC           | 0.0160         | 0.0163           | 1.05 (0.47–2.34) | 0.90 |
| 8      | CGCCG           | 0.0128         | 0.0103           | 1.17 (0.42–3.22) | 0.76 |
| 9      | ACCGC           | 0.0003         | 0.0197           | 0.13 (0.03–0.59) | 0.0079 |
| 10     | Rare            | 0.0240         | 0.0580           | 0.36 (0.21–0.64) | 0.0005 |

\(^a\) Adjusted by smoking duration, Global haplotype association p-value = 0.0011.

\(^b\) SNP order: rs12951053-rs1042522-rs8079544-rs12602273-rs8064946.

\(^c\) Boldface means association with decreased risk of lung cancer.

Table 5 Smoking duration within TP53 htSNP haplotypes among 1037 subjects.

| Number | Haplotype\(^a\) | Frequency | OR (95% CI) Never | OR (95% CI) ≤20 (years) | OR (95% CI) >20 (years) |
|--------|-----------------|-----------|-------------------|------------------------|------------------------|
| 1      | AGCCG           | 0.4915    | 1.26 (0.68–2.34)  | 1.90 (1.17–3.09)        | –                      |
| 2      | CCGCC           | 0.2228    | 1.61 (0.90–2.87)  | 2.22 (1.47–3.37)        | –                      |
| 3      | ACTCG           | 0.0694    | 2.18 (0.83–5.72)  | 2.16 (0.99–4.72)        | –                      |
| 4      | CCGCC           | 0.0562    | 1.72 (0.56–5.24)  | 2.65 (1.08–6.51)        | –                      |
| 5      | ACCGG           | 0.0405    | 1.01 (0.27–3.83)  | 1.60 (0.64–4.00)        | –                      |
| 6      | CCCCG           | 0.0367    | 1.64 (0.41–6.58)  | 1.30 (0.44–3.85)        | –                      |
| 7      | AGCGC           | 0.0159    | 2.17 (0.30–15.76) | 7.02 (0.91–53.99)       | –                      |
| 8      | CGCCG           | 0.0119    | 0.77 (0.05–13.12) | 0.84 (0.10–7.07)        | –                      |
| 9      | ACCGC           | 0.0112    | –                  | –                      | –                      |
| 10     | Rare            | 0.0394    | 1.44 (0.22–9.53)  | 4.34 (1.21–15.48)       | –                      |

\(^a\) SNP order: rs12951053-rs1042522-rs8079544-rs12602273-rs8064946.

\(^b\) Boldface indicates statistical significance (p value < 0.05).

Discussion

Studies addressing TP53 SNPs in lung cancer

The previous association studies on TP53 SNPs and lung cancer risk mainly assessed associations of SNP, haplotype/diplotype and gene-gene and gene-gene-environment interactions [4–14] [Table 7].

Variant-homozygote of TP53 rs1042522 was at significantly increased risk of lung squamous cell carcinoma [CC versus GG: OR (95% CI) = 2.2 (1.3–3.9), p = 0.005] in Asian Japanese [6]. TP53 rs1042522 was associated with significantly increased lung cancer risk in the total population [recessive model: CC versus Any G, adjusted OR (95% CI) = 1.57 (1.11–2.21)] and minor-allele carriers (TC or CC) of TP53 rs2078486 were significantly increased lung cancer risk among smokers [adjusted OR (95% CI) = 1.70 (1.08–2.67)] in Asian Chinese [5]. The TP53 rs1042522 C-allele were significantly associated with increased lung cancer risk [GC or CC versus GG: OR (95% CI) = 2.38 (1.38–4.82) and OR (95% CI) = 4.62 (2.31–9.52), respectively] in Asian Bengalese [6]. A study including Caucasians and African Americans reported that among African Americans, carriers of the haplotype rs1042522G-rs895829T-rs2904930A-rs1625895G-rs12951053G had increased risk for lung cancer [OR (95% CI) = 2.32 (1.18–4.57)] and a worsened lung cancer prognosis [HR (hazards ratio) (95% CI) = 2.38 (1.38–4.10)] compared with carriers of the haplotype rs1042522C-rs895829G-rs2904930A-rs1625895G-rs12951053G [7].

Variant C-allele of TP53 rs1042522 was significantly associated with increased risk of lung squamous cell carcinoma [CC+GC versus GG: OR (95%) = 1.65 (1.10–2.47), p = 0.016], the risk was markedly increased in heavy smokers with lung squamous cell carcinoma [CC versus GG: OR (95%) = 2.80 (1.19–6.58), p = 0.019] and combined effect of TP53...
rs1042522 C-allele and $P21/\text{CDKN1A}$ (cyclin dependent kinase inhibitor 1 A) rs1801270 CC-genotype was most pronounced in heavy smokers with lung squamous cell carcinoma [TP53 rs1042522$^{CC} + CG/P21$ rs1801270$^{CC}$ versus TP53 rs1042522$^{CC}/P21$ rs1801270$^{AA}/AC}$; OR (95% CI) = 3.84 (1.46–10.1), $p = 0.007$] in Caucasians Germans [8]. The TP53 rs1042522 was significantly associated with increased risk of lung adenocarcinoma [CC versus GG: adjusted OR (95% CI) = 1.55, (1.17–2.06)] and gene-gene interaction was found for the combination of TP53 rs1042522$^{CC}$ and MDM2 (MDM2 proto-oncogene) rs2279744$^{GG}$ genotypes [adjusted OR (95% CI) = 2.66 (1.54–4.60)] related to risk of lung adenocarcinoma in Asian-Chinese female non-smokers [9]. TP53 rs1042522 was associated with risk of NSCLC (non-small-cell lung cancer), both independently [dominant model: OR (95% CI) = 1.809 (1.159–2.825), $p < 0.05$; recessive model: OR (95% CI) = 1.933 (1.096–3.409), $p < 0.05$] and in combination with miR-502-binding site SNP (rs16917496) in the 3’ UTR of SET8 (set domain-containing
Table 7 Results of TP53 single nucleotide polymorphisms and risk of lung cancer from epidemiological studies.a.

| Lung cancerb | Reference | SNP | Location/Populationc | Cases/Controls | Comparisond | OR (95% CI) | P valuee |
|--------------|-----------|-----|----------------------|----------------|-------------|-------------|-----------|
| LC | Sakiyama et al. [4] | rs1042522 | Japan/Hospital-based case-control | 1002/685 | CC vs. GG/SQC | 2.2 (1.3–3.9) | 0.005 |
| LC | Li et al. [5] | rs1042522 | China/Hospital-based case-control | 399/466 | CC vs. Any G | 1.57 (1.11–2.21) | – |
| | | | | rs2078486 | | | |
| LC | Mostaid et al. [6] | rs1042522 | Bangladesh/Population-based case-control | 106/116 | TC vs. TT/Smoker | 1.70 (1.08–2.67) | – |
| LC | Mechanic et al. [7] | rs1042522 | USA/Hospital-based Case-control/AFA | 120/204 | Haplotype with C vs. G | 2.32 (1.18–4.57) | – |
| | | | | rs1042522C- rs9895829T- rs2909430A- rs1625895G- rs12951053G vs. G-T-A-G-T | | |
| LC | Popanda et al. [8] | rs1042522 | Germany/Hospital-based case-control | 405/404 | CC vs. GG/SQC | 1.65 (1.10–2.47) | 0.016 |
| | | | | | CC versus GG/HS | 2.80 (1.19–6.58) | 0.019 |
| ADC | Ren et al. [9] | rs1042522 | China/Hospital-based case-control/FNS | 764/983 | CC vs. GG | 3.84 (1.46–10.1) | 0.007 |
| NSCLC | Yang et al. [10] | rs1042522 | China/Hospital-based case-control | 164/199 | Combination genotypes with CC | 2.66 (1.54–4.60) | <0.001 |
| | | | | | Dominant model | 1.809 (1.159–2.825) | <0.05 |
| | | | | | Recessive model | 1.933 (1.096–3.409) | <0.05 |
| | | | | | Combination genotypes with GG | 3.032 (1.580–5.816) | – |
| | | | | | Dipotype with CC vs. GG+GC | 3.68 (1.43–9.45) | – |
| LC | Myneni et al. [11] | rs1042522 | China/Population-based case-control | 399/466 | Combination genotypes with CC | 2.5 (1.2–5.0) | – |
| LC | Chua et al. [12] | rs1042522 | Singapore/Hospital-based case-control | 126/162 | Combination genotypes with C | 2.5 (1.2–5.0) | – |
| LC | Mechanic et al. [7] | rs1042522 | USA/Hospital-based case-control/CA | 323/343 | MDM2 rs2279744TT vs. TP53 rs1042522GGCC + MDM2 rs2279744TT | 1.23 (0.86–1.76) | – |
| | | | | | AB or BB or AB+BB vs. AA: | 0.87 (0.41–1.84)/1.18 (0.84–1.66), 1.48 (0.78–2.82)/not determined | – |
| | | | | | | 1.48 (0.78–2.82), 1.17 (0.77–1.78)/ | – |
| | | | | | | 1.08 (0.31–3.76)/1.16 (0.77–1.74), 1.12 (0.74–1.68)/0.93 (0.25–3.41)/ | – |
| | | | | | | 1.10 (0.74–1.64), 0.91 (0.56–1.49)/ | – |
| | | | | | | 1.97 (0.19–20.6)/0.94 (0.58–1.52) | – |
| LC | Guan et al. [13] | rs7837822 | USA/Hospital-based case-control/NHW 1014/1076 | 323/343 | AC vs. AA | 0.84 (0.51–1.37) | 0.379 |

(continued on next page)
### Table 7 (continued)

| Lung cancer | SNP | Location/Population | Case/Controls | Comparison | OR (95% CI) | P value | P value
|-------------|-----|---------------------|---------------|------------|-------------|---------|---------|
| LC Zhang et al. [14] | rs1042522 | China/Hospital-based | 640/650 CG or GG or CG | GG vs. CC | 1.02 (0.79–1.31) | 0.882 | 0.99 (0.72–1.37) | 0.963 | 1.00 (0.80–1.29) | 0.924 |
| LC Yin et al. | rs1042522 | China/Hospital-based | 544/550 Haplotype with C vs. G | 0.13 (0.03–0.59) | 0.0079 |

### Main findings, implications and strengths of current study

In the present study, we report no association with lung cancer risk for the individual TP53 htSNPs (including TP53 rs1042522) [Table 3]. This is in agreement with a previously reported regard to TP53 rs1042522 in Asian-Chinese Han population [14]. TP53 five htSNPs were in stronger pair-wise LD for our study population (Supplementary Table S2). Haplotype analysis could increase the estimated effect. Haplotype encompassing rs1042522 and other 4 htSNPs of TP53 showed association evidence. Haplotype 9 (rs12951053A-rs1042522C-rs8079544T-rs16202273G-rs8064946C) with 2% frequency in the controls was associated with lowered risk of lung cancer [Table 4]. This significant observation is not consistent with previously significant associated findings in an African-Americans population [7]. The difference is that the haplotype encompassing rs1042522C was protective in current Chinese population, while the haplotype encompassing rs1042522C was risky in African Americans. The polymorphisms included in the haplotypes studied differed between the studies and only rs1042522 and rs12951053 were included in both haplotypes in the two studies. There were statistically significant differences of the two alleles frequencies in control groups among current Chinese and African Americans for rs1042522 (C = 0.45 and C = 0.55, this was in inversion for minor allele and major allele, χ² = 5.733, p = 0.017) and rs12951053 (C = 0.34 and C = 0.1, χ² = 38.512, p < 0.001). Thus the observed discrepancy may result from differences of SNPs or allele frequencies composing haplotype or differences of LD status and haplotype frequency in...
In addition, the analysis of smoking duration within TP53 haplotypes among 1037 subjects exhibited carriers with haplotype1 (AGGCCG), haplotype2 (CCCCGG) and haplotype4 (CCCCCG) were over-represented in smoking subgroup of >20 (years). This showed that the three haplotypes played coincident roles with respect to smoking duration. It suggested that three haplotypes (AGGCCG, CCCCCG and CCCCCG) consisting of TP53 htSNPs (htSNPs order: rs12951053-rs1042522-rs8079544-rs12602273-rs8064946) may be a potentially genetic predisposing factor for behavior of long-term smoking.

We assessed the possible functionality of the studied polymorphisms using the web tool: SNPinfo [32]. This analysis indicated that TP53 rs12951053 (Regulatory Potential Score = 0.058167), rs1042522 (rsSNP: Yes, Polyphen: benign, Regulatory Potential Score = 0.31032, Conservation Score = 0.002), rs8079544 (Regulatory Potential Score = 0.204487) and rs8064946 (Transcription Factor Binding Sites: Yes, Regulatory Potential Score = 0.118648) may all be biologically functional, whereas rs12602273 was not. Rs1042522 was the most important functional htSNP, and lead to a non-conservative Arg to Pro amino acid substitution.

**Limitations**

With current genotypes we had 88%, 79%, 70%, 90% and 89% and 82% chance of detecting OR = 1.5 at 0.05 significant level and two sided test under dominant model for TP53 rs12951053, rs1042522, rs8079544, rs12602273 and rs8064946 and CD3EAP rs735482, respectively. Further studies with larger sample sizes are warranted. The matching concerning age, gender and ethnicity between cases and controls was insufficient to exclude potential confounding factors such as smoking in this study.

**Conclusion**

In conclusion, the present results provide novel evidence that the haplotype of TP53 htSNPs and interaction between genetic variation in TP53 and CD3EAP and smoking-duration may associate with lung cancer risk, and provide additional evidence of association between TP53 htSNP haplotypes and long-term smoking-related behavior.

**Conflicts of interest**

The authors have no conflicts of interest relevant to this article.

**Acknowledgements**

The Human Genetic Resource Administration of China, Ministry of Science and Technology of the People’s Republic of China (Beijing, P. R. China) approved this study (Approval No. 2001015). The National Natural Science Foundation of China (Grant No. 30571016 and No. 81072384) supported this project.

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at [https://doi.org/10.1016/j.bj.2021.01.006](https://doi.org/10.1016/j.bj.2021.01.006).
REFERENCES

[1] Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2013;156:E359–86.

[2] Algranti E, Buschinelli JT, De Capitani EM. Occupational lung cancer. J Bras Pneumol 2010;36:784–94.

[3] Lin HY, Huang CH, Wu WJ, Chang LC, Lung FW. TP53 codon 72 gene polymorphism paradox in associated with various carcinoma incidences, invasiveness and chemotherapy responses. Int J Biomed Sci 2008;4:248–54.

[4] Sakiyama T, Kohno T, Mimaki S, Ohita T, Yanagitani N, Sobue T, et al. Association of amino acid substitution polymorphisms in DNA repair genes TP53, POLI, REV1 and LG4 with lung cancer risk. Int J Cancer 2005;114:730–7.

[5] Li Y, Chang SC, Niu R, Liu L, Crabtree-Ide CR, Zhao B, et al. P53 Arg72Pro and MDM2 SNP309 polymorphisms cooperate to increase lung adenocarcinoma risk in Chinese female non-smokers: a case control study. Asian Pac J Cancer Prev 2013;14:5415–20.

[6] Mostaid MS, Ahmed MU, Islam MS, Bin Sayeed MS, Hasnat A. Lung cancer risk in relation to TP53 codon 47 and codon 72 polymorphism in Bangladeshi population. Tumour Biol 2014;35:10309–17.

[7] Mechanic LE, Bowman ED, Welsh JA, Khan MA, Hagiwara N, Enoweld L, et al. Common genetic variation in TP53 is associated with lung cancer risk and prognosis in African Americans and somatic mutations in lung tumors. Cancer Epidemiol Biomarkers Prev 2007;16:214–22.

[8] Popanda O, Edler L, Waas P, Schattenberg T, Butkiewicz D, Muley T, et al. Elevated risk of squamous-cell carcinoma of the lung in heavy smokers carrying the variant alleles of the TP53 Arg72Pro and p21 Ser31Arg polymorphisms. Lung Cancer 2007;55:25–34.

[9] Ren YW, Yin ZH, Wan Y, Guan P, Wu W, Li XL, et al. P53 Arg72Pro and MDM2 SNP309 polymorphisms cooperate to increase lung adenocarcinoma risk in Chinese female non-smokers: a case control study. Asian Pac J Cancer Prev 2013;14:5415–20.

[10] Lang S, Guo H, Wei B, Zhu S, Cai Y, Jiang P, et al. Association of miR-502-binding site single nucleotide polymorphism in the 3′-untranslated region of SET8 and TP53 codon 72 polymorphism with non-small cell lung cancer in Chinese population. Acta Biochim Biophys Sin (Shanghai) 2014;46:149–54.

[11] Myreni AA, Chang SC, Niu R, Liu L, Hao B, Shi J, et al. Ataxia telangiectasia-mutated (ATM) polymorphisms and risk of lung cancer in a Chinese population. Front Publ Health 2017;5:102.

[12] Chua HW, Ng D, Choo S, Lum SS, Li H, Soh LY, et al. Effect of MDM2 SNP309 and p53 codon 72 polymorphisms on lung cancer risk and survival among non-smoking Chinese women in Singapore. BMC Cancer 2010;10:88.

[13] Guan X, Wang LE, Liu Z, Sturgis EM, Wei Q. Association between a rare novel TP53 variant (rs78378222) and melanoma, squamous cell carcinoma of head and neck and lung cancer susceptibility in non-Hispanic Whites. J Cell Mol Med 2013;17:873–83.

[14] Zhang H, Hao S, Zhao J, Zhou B, Ren Y, Yan Y, et al. Common genetic variants in 53BP1 associated with nonsmall-cell lung cancer risk in Han Chinese. Arch Med Res 2014;45:84–9.

[15] Vogel U, Laros I, Jacobsen NR, Thomsen BL, Bak H, Olsen A, et al. Two regions in chromosome 19q13.2-3 are associated with risk of lung cancer. Mutat Res 2004;546:65–74.

[16] Yin J, Vogel U, Ma Y, Wang H, Wang C, Liang D, et al. A specific dipolotype defined by PPP1R13L rs1970764, CD3EAP rs967591 and ERCC1 rs11615 and lung cancer risk in a Chinese population. Lung Cancer 2012;76:286–91.