# Ethnic differences in frequencies of gene polymorphisms in the MYCL1 region and modulation of lung cancer patients' survival

Monica Spinola[a], F. Stefania Falvella[a], Antonella Galvan[a], Carmen Pignatiello[a], Vera P. Leoni[a], Ugo Pastorino[b], Rita Paroni[c], Shuqing Chen[d], Vidar Skaug[e], Aage Haugen[e], Tommaso A. Dragani[a,*]

[a] *Department of Experimental Oncology and Laboratories, Istituto Nazionale Tumori, Milan, Italy*
[b] *Unit of Thoracic Surgery, Istituto Nazionale Tumori, Via G. Venezian 1, 20133 Milan, Italy*
[c] *Department of Medicine, Surgery and Dental Science, University of Milan, San Paolo Hospital, Milan, Italy*
[d] *College of Pharmaceutical Sciences, Zhejiang University, Zhejiang, China*
[e] *Section for Toxicology, National Institute of Occupational Health, Oslo, Norway*

**Summary**    Linkage disequilibrium (LD) analysis to refine a region associated with lung cancer progression on chromosome 1p34 identified a 106 kb LD block that includes MYCL1, TRIT1 (tRNA isopentenyltransferase 1) and MFSD2 (major facilitator superfamily domain-containing 2). Case-only association study on SNPs mapping in TRIT1 and MFSD2 indicated that the rare Leu allele (frequency: 0.04) of the TRIT1 Phe202Leu variation predicts short survival as compared to the common Phe/Phe genotype (hazard ratio (HR) = 1.7; 95% CI, 1.03—2.86; $P = 0.039$) in 335 Italian lung adenocarcinoma samples. A replication study in an independent population of 246 Norwegian lung cancer patients confirmed the significant association of the Phe202Leu polymorphism with patients' survival, but the rare allele was associated with better survival rate (HR = 0.5; 95% CI, 0.26—0.91; $P = 0.023$). The rare allele of TRIT1 Phe202Leu SNP was ~seven-fold more frequent in Asian than in Caucasian subjects and three additional SNPs in the TRIT1 and MFSD2 genes showed ethnic differences in allelic frequencies. These results suggest that polymorphisms in the MYCL1 LD region affect lung cancer survival but that the functional element(s) may show population-specific patterns.
© 2006 Elsevier Ireland Ltd. All rights reserved.

## 1. Introduction

A single nucleotide polymorphism (SNP) affecting an *Eco*RI site in the second intron of the MYCL1 (previous symbol: L-myc) gene was found to be associated with patients' outcome in lung cancer with contradictory reports [1—5].

* Corresponding author. Tel.: +39 0223902642; fax: +39 0223903237.
*E-mail address:* tommaso.dragani@istitutotumori.mi.it (T.A. Dragani).

Our recent meta-analysis revealed a significant association between the intronic MYCL1-*Eco*RI variant and lymph node metastasis, distant metastasis, and clinical stage of lung cancer patients [6], supporting the hypothesis that genetic variation in either MYCL1 or a flanking gene determines lung tumour progression. MYCL1 is frequently amplified and over-expressed in small-cell lung cancer but not in a broad range of lung tumour histotypes [7], and no coding polymorphism in this gene was found to be associated with lung tumour prognosis or prognostic factors [8]. Thus, the findings of the meta-analysis are likely explained by linkage disequilibrium (LD) of the MYCL1-*Eco*RI polymorphism with a nearby responsible gene.

Here, we report LD analysis of ~330 kb on human chromosome 1p34 and identification of a ~106 kb region in LD with the MYCL1-*Eco*RI polymorphism which contains two genes that strongly affect tumour phenotypes: TRIT1, which is involved in post-transcriptional modification of tRNA molecules, and MFSD2, which encodes a putative transmembrane protein.

## 2. Materials and methods

### 2.1. Patients and tissue samples

Patients series included 335 lung adenocarcinoma (ADCA) patients enrolled at Istituto Nazionale Tumori (Milan, Italy), 246 non-small cell lung cancer (NSCLC) patients from Norway (described in [9]), and 92 healthy non-cancer subjects recruited from Zhejiang, China. All cases were pathologically documented and personal files were recorded to obtain clinical data (Table 1). Written informed consent was obtained from the participating patients and the use of blood samples and clinical data were in accordance with the ethical committees in the respective country. Genomic DNA was extracted from peripheral blood with: the automatic DNA extractor Extragen 8C (Talent, Trieste, Italy) (Italian samples); phenol/chloroform (Norwegian samples), and the DNA Mini kit (Qiagen, Valencia, CA) (Chinese samples).

Total RNA was extracted with the RNeasy Midi kit (Qiagen) from 20 samples of normal lung tissue and 20 samples of lung tumour tissue of Italian lung ADCA patients who underwent surgical resection at Istituto Nazionale Tumori. Equal amounts (1 μg) of lung normal RNAs and lung tumour RNAs were pooled and the two RNA pools were reverse-transcribed with the SuperScript® First-Strand Synthesis System (Invitrogen, Carlsbad, CA).

### 2.2. Selection of SNPs and genotyping

SNPs mapping in a ~330 kb region centered on MYCL1 were searched in the NCBI public database (http://www.ncbi.nlm.nih.gov/SNP/) and first validated on a subgroup of 50 Italian patients by pyrosequencing analysis on a PSQ96MA system (Biotage AB, Uppsala, Sweden). SNP-containing fragments were PCR-amplified, followed by pyrosequencing analysis according to the manufacturer's instructions using specific primers. Markers confirmed as polymorphic were genotyped on Italian lung ADCA patients to assess pairwise LD between markers. Data on genotypes of 4 MYCL1 SNPs have been reported [5,8].

**Table 1** Characteristics of Italian and Norwegian lung cancer patients

| Subject characteristics | Italians | Norwegians |
|---|---|---|
| Number of subjects | 335 | 246 |
| Median age (years) | 63.0 | 66.0 |
| Gender | | |
| Male | 257 | 174 |
| Female | 78 | 72 |
| Smoker status | | |
| Never | 45 | 16 |
| Ever | 279 | 230 |
| Histology[a] | | |
| ADCA | 335 | 106 |
| SQCC | | 88 |
| LCC | | 42 |
| Others | | 10 |
| Lymph node status[b] | | |
| N0 | 191 | 173 |
| N1 | 111 | 73 |
| Clinical stage | | |
| I | 162 | 160 |
| II | 61 | 44 |
| III | 60 | 42 |
| IV | 27 | 0 |
| Follow-up (months)[c] | 69.1 ($n = 122$) | 74.0 ($n = 90$) |

[a] ADCA, adenocarcinoma; SQCC, squamous cell carcinoma; LCC, large cell carcinoma.
[b] N0, absence of nodal metastasis; N1, presence of nodal metastasis.
[c] Median for patients alive at the end of follow-up.

### 2.3. Analysis of mRNA expression in human lungs

Quantitative real-time PCR amplification was carried out using 0.5 μl of a template cDNA pool from normal or tumour lung tissue, 12.5 μl of 2X QuantiTect SYBR Green PCR Master Mix (Qiagen, Valencia, CA), and 0.3 μM gene-specific intron-spanning primers (MFSD2: 5′-ccttgtttccaggacctcaa-3′ and 5′-aagtaggcgattggctcaga-3′; MYCL1: 5′-cgagagcccaagcgact-3′ and 5′-tgatggatggagatgtggaa-3′). Reactions were run on an ABI GeneAmp 7900 sequence detection system (Applied Biosystems, Foster City, CA). The human HMBS gene (GenBank accession # NM_000190) was amplified with primers 5′-aggatgggcaactgtacctg-3′ and 5′-gcctaccaactgtgggtcat-3′ as a housekeeping control for possible differences in cDNA amounts. Samples were analysed in triplicate in each experiment. Relative differences were calculated according to the comparative $C_t$ method using the normal lung cDNA pool for calibration.

### 2.4. Sequencing analysis of TRIT1 and MFSD2 coding regions

Complete coding regions of TRIT1 and MFSD2 were sequenced from normal DNA of Italian lung ADCA patients

to search for germ-line functional SNPs. All exons were analysed using MassARRAY technology, except for MFSD2 exons 4 and 5, which were directly sequenced. PCR primers were designed in flanking intronic regions or untranslated regions to amplify products not exceeding 600 bp. PCR primers for MassARRAY analysis were modified by addition of a 5'-tail specific for the T7 promoter (forward primer: 5'-cagtaatacgactcactatagggagaaggct-3') or the SP6 promoter (reverse primer: 5'-cgatttaggagacactatagaagagaggct-3'). Analysis by MassARRAY was carried out at Methexis Genomics (Gent, Belgium) as described [10].

## 2.5. Assay for isopentenyladenosine activity of TRIT1 alleles

TRIT1 encodes tRNA isopentenyltransferase, which catalyzes the formation of the modified base $N^6$-isopentenyladenosine ($i^6$A) in position 37 of tRNA molecules that bind codons starting with uridine [11,12]. TRIT1 alleles were tested for functional activity in the yeast strain MT-8, which has defective tRNA-isopentenyltransferase activity and synthesizes tRNAs lacking the isopentenyl-adenosine ($i^6$A) modification, as compared with yeast strain H57, which produces an active tRNA-isopentenyltransferase gene. Yeast transformation and $i^6$A analysis were carried out as described [12]. The TRIT1$^{Phe202}$ allele was obtained by amplification of a cDNA pool of normal lung tissue with primers 5'-caacggaccctacctcttgt-3 and 5'-cacctttccaaaggccact-3' and cloned into the yeast expression vector pYES2.1/V5-His-TOPO (Invitrogen). The TRIT1$^{Leu202}$ allele was generated by in vitro mutagenesis with the QuickChange®II Site-directed Mutagenesis Kit (Stratagene) using the pYES2.1/TRIT1$^{Phe202}$ vector as template and specific primers that allowed substitution of nucleotide 620 (based on the annotated sequence GenBank accession # NM_017646) from thymine to adenine (sense primer 5'-gaatctctcatagtgaattactccatcgtcaacatacgga-3'; antisense primer 5'-tccgtatgttgacgatggagtaattcactatgagagattc-3').

## 2.6. Statistical analysis

Pairwise LD was evaluated calculating $D'$ and $r^2$ as described [13]. LD values were graphed using the Java LINkage disequilibrium plotter (JLIN) (www.genepi.com.au/projects/jlin). Consistency of genotype frequencies with the Hardy—Weinberg equilibrium was tested for each SNP locus using a contingency table of observed versus predicted genotype frequencies [14]. The Kaplan—Meier product-limit method, the log-rank test, and the Cox regression model were used to evaluate the effect of the SNP genotypes on overall survival [15,16]. Median time intervals were estimated by the life-table method. Statistical procedures were carried out with SPSS 10.1 (SPSS Inc., Chicago, IL) software.

## 3. Results

## 3.1. LD extends over 106 kb around MYCL1

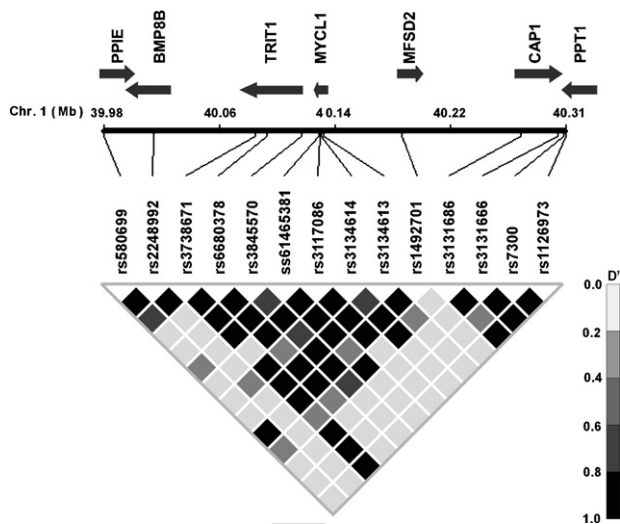The LD pattern of the chromosomal region containing the MYCL1 gene was analysed in an Italian population



**Fig. 1** LD pattern across the MYCL1 chromosomal region. $D'$ values for pairwise LD between each marker are represented in the grey scale, from black ($D' = 1$) to white ($D' = 0$). Spacing of markers and the locations of the genes across the genomic region are shown at the top of the figure. The region from TRIT1 to MFSD2 shows a block of LD.

of lung ADCA patients, using SNPs to identify the genes in LD with the MYCL1-EcoRI polymorphism. A region of ~330 kb spanning from 39.98 to 40.31 Mb and centered on the MYCL1 gene was selected for this analysis and the genes included were: peptidyl-prolyl cis-trans isomerase E (PPIE); bone morphogenetic protein 8 precursor b (BMP8B); tRNA-isopentenyltransferase (TRIT1); major facilitator superfamily domain containing 2 (MFSD2); adenylyl cyclase-associated protein 1 (CAP1); and palmitoyl-protein thioesterase 1 precursor (PPT1) (Fig. 1). 4 MYCL1 SNPs were known [5,8] and 15 SNPs in the other genes were selected and then validated in a subgroup of 50 Italian individuals; 5 SNPs were excluded from the analysis because they were not polymorphic, indicating that the possible rare allele had a frequency lower than 0.01 in the Italian population and was not detected. LD analysis was carried out on a final group of 14 polymorphisms, which included three coding variants: MYCL1-Ser362Thr, TRIT1-Phe202Leu, and TRIT1-Pro11Pro (Table 2).

These 14 SNPs, which showed frequencies of the rare allele ranging from 0.04 to 0.45 (Table 2), were genotyped in 103 Italian lung ADCA patients. No deviation from the Hardy-Weinberg equilibrium was observed. Pairwise LD analysis showed that MYCL1-EcoRI (rs3134613) maintained strong LD ($D' > 0.8$, Fig. 1) with both its flanking genes, i.e., TRIT1 and MFSD2, whereas LD was no longer displayed beyond these genes in either direction (Fig. 1). Significant LD was detected also between TRIT1 and MFSD2 over a region of 106 kb. Other blocks of strong LD were identified between PPIE and BMP8B over a region of 37 kb and between SNPs in genes CAP1 and PPT1, which are 32 kb away (Fig. 1). These findings restricted the region in LD with MYCL1-EcoRI polymorphism to a block of 106 kb containing the genes TRIT1 and MFSD2 in addition to the proto-oncogene MYCL1.

**Table 2**  SNPs used for analysis of LD in the MYCL1 region

| Gene | SNP accession # | SNP type | Nucleotide change[a] | Amino acid change | Frequency of the rare allele |
|------|-----------------|----------|----------------------|-------------------|------------------------------|
| PPIE | rs580699 | intron | G39979097A | | 0.34 |
| BMP8B | rs2248992 | intron | T40015629C | | 0.25 |
| TRIT1 | rs3738671 | exon5 | A40088475T | Phe202Leu | 0.04 |
| | rs6680378 | intron | T40096684C | | 0.40 |
| | rs3845570 | exon1 | G40121718A | Pro11Pro | 0.40 |
| MYCL1 | ss61465381 | 3′UTR | G40134371C | | 0.24 |
| | rs3117086 | 3′UTR | G40135048C | | 0.17 |
| | rs3134614 | exon2 | C40135641G | Ser362Thr | 0.10 |
| | rs3134613 | intron | A40137390C | | 0.45 |
| MFSD2 | rs1492701 | intron | C40194405G | | 0.19 |
| CAP1 | rs3131686 | intron | A40279809G | | 0.12 |
| | rs3131666 | intron | A40306482T | | 0.19 |
| | rs7300 | 3′UTR | G40310688A | | 0.05 |
| PPT1 | rs1126973 | 3′UTR | T40311663C | | 0.17 |

[a] Contig position is preceded by the common allele and followed by the rare allele of the polymorphism.

## 3.2. TRIT1 and MFSD2 are downregulated in lung ADCAs

We previously showed that TRIT1 mRNA is expressed in normal lung with a complex pattern of alternatively spliced variants and that its expression levels are downregulated ~6- to 14-fold in lung tumour tissue [12]. Similar analysis of MYCL1 and MFSD2 using quantitative real-time PCR on pools of normal and tumour lung tissue revealed no detectable MYCL1 mRNA levels in either normal or ADCA tissue, suggesting that MYCL1 is not a likely lung cancer progression modifier gene, whereas MFSD2 was expressed in normal lung parenchyma and downregulated in lung ADCA tissue. Indeed, analysis of three independent experiments indicated that $\triangle\triangle C_t$ values of MFSD2 in pooled tumour versus normal samples were $3.35 \pm 0.5$ (mean $\pm$ S.E.), indicating a ~10-fold lower transcript abundance in lung ADCAs.

## 3.3. TRIT1 Phe202Leu is associated with patients' survival

To test whether germ-line variants in the candidate genes might underlie the observed differences in lung cancer progression, the entire coding regions of TRIT1 and MFSD2 were screened for functionally relevant coding polymorphisms in addition to those reported in the public databases. The analysis was carried out on two subgroups of Italian lung ADCA patients, each consisting of 12 samples selected for favorable (clinical stage ≥3 and follow-up ≥2 years) and poor (clinical stage ≤2 and follow-up <1 year) prognosis. Thus, direct sequencing of genomic DNA of 24 patients to identify SNPs with a frequency of the rare allele >0.02 confirmed the presence of the known TRIT1 coding polymorphisms rs3845570 (Pro11Pro) and rs3738671 (Phe202Leu) mapping in exon 1 and 5, respectively, but did not identify new coding variations in this gene (Table 2). SNP rs3845570 is a synony-

mous coding SNP (Pro11Pro) in exon 1 of the TRIT1 gene; however, in the shorter lung-specific isoform (AY702933), the rs3845570 SNP is located in the 5′-UTR region, 25-bp upstream the alternative ATG codon [12]. Coding SNPs were not detected in MFSD2.

TRIT1 and MFSD2 SNPs in Italian lung ADCA patients were tested for an association with clinical parameters. No deviation from the Hardy—Weinberg equilibrium was observed at any marker. No significant association with nodal status, clinical stage or patients' survival was found for the intronic rs6680378 and the exonic rs3845570 of the TRIT1 gene, or for intronic rs1492701 of the MFSD2 gene (not shown). On the other hand, the TRIT1 Phe202Leu variant showed a significant association with patients' survival. Due to its low frequency (0.04), the Leu variant was found only in the heterozygous state and no individual carrying the Leu/Leu genotype was identified. Analysis of survival rates showed a significant association of the Leu202 allele with shorter survival: median survival time was 44.0 months for Phe/Phe individuals and 23.4 months for Phe/Leu patients (P = 0.039, log-rank test; Fig. 2A). Cox proportional hazard analysis of survival, adjusted for sex, smoking, age at diagnosis (in decades), nodal status, and clinical stage showed that patients carrying the Phe/Leu genotype had a poorer survival than patients with the common Phe/Phe genotype (HR, 1.7; 95% CI, 1.03—2.86; P = 0.039; Table 3).

The association study was replicated in an independent Caucasian population from Norway. Unlike the Italian samples, which were all ADCA in histotype, Norwegian samples included squamous cell carcinomas and large cell carcinomas in addition to ADCAs (Table 1). Other parameters were slightly different between the two study populations (Table 1). The Phe202Leu SNP of the TRIT1 gene showed a frequency of the rare Leu allele similar to that observed in the Italian population (i.e., 0.06 versus 0.04) and a significant association with lung cancer patients' survival (Table 3); however, the Leu202 allele, which was associated
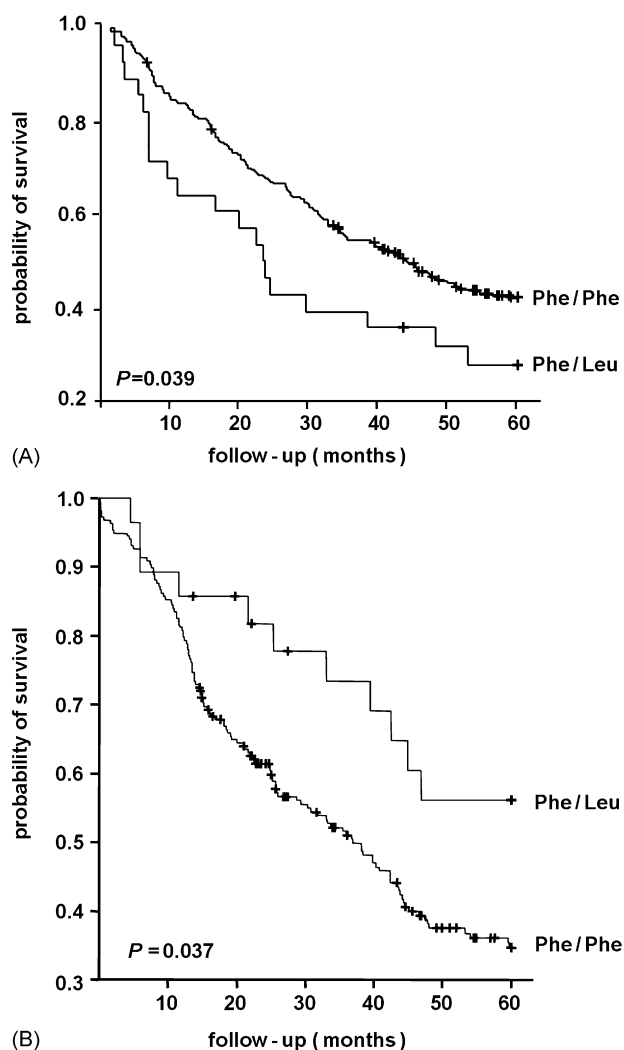
**Fig. 2** Kaplan—Meier survival curves of TRIT1-Phe202Leu polymorphism in (A) Italian or (B) Norwegian lung cancer patients. The curves of patients with the Phe/Phe (Italians, $n = 288$; Norwegians, $n = 218$) or Phe/Leu (Italians, $n = 28$; Norwegians, $n = 28$) genotype are shown. Follow-up is shown truncated at 60 months. Small crosses indicate censored observations. Log-rank analysis indicates a significant difference between curves (Italians, $P = 0.039$; Norwegians, $P = 0.037$).

with poor survival in Italians, showed an opposite trend in Norwegian lung cancer patients. Indeed, median survival times of Phe/Leu and Phe/Phe patients were 96.5 and 36.9 months, respectively ($P = 0.037$, log-rank test; Fig. 2B), and Cox proportional hazard analysis of survival, adjusted for sex, smoking, age at diagnosis (in decades), nodal status, and clinical stage showed that patients carrying the Phe/Leu genotype had a better survival than patients with the common Phe/Phe genotype (HR, 0.5; 95% CI, 0.26—0.91; $P = 0.023$; Table 3).

## 3.4. Phe202Leu TRIT1 alleles have similar biochemical activity

The biochemical activity of the two TRIT1 alleles was compared in a yeast complementation assay. Amplification of TRIT1 from a pool of human normal lung parenchyma produced only the full-length transcript encoding the common Phe variant, confirming the low frequency of the Leu variant. The rare transcript was obtained by *in vitro* mutagenesis and vectors expressing TRIT1 variants were transformed in the MT-8 yeast strain, which synthesizes tRNAs devoid of isopentenyladenosine ($i^6A$). HPLC and mass spectrometry analysis of tRNAs extracted from control and transformed yeast strains revealed no significant differences between the two alleles. Indeed, both variants were functionally active and produced similar amounts of $i^6A$ ($1.43 \pm 0.41$ and $1.46 \pm 0.36$ pmol/$\mu$g tRNA for the Phe and Leu allele, respectively), which were comparable to $i^6A$ levels measured in the wild-type H57 yeast strain ($1.81 \pm 0.40$ pmol/$\mu$g tRNA).

## 3.5. Ethnic differences in allele frequency at TRIT1 and MFSD2 genes

Genotyping of the Phe202Leu variation in a small group ($n = 92$) of Chinese non-cancer controls revealed a higher frequency of the rare Leu202 allele (0.28) as compared to the Italian (0.04) or Norwegian (0.06) samples. To verify and extend these findings on ethnic differences, we have extracted the HapMap (http://www.hapmap.org/) genotype data of SNPs located in TRIT1, MYCL1, or MFSD2 genes. The frequency of the Leu202 allele was 0.03 in Caucasian (CEPH, $n = 90$) and 0.32 in Asian (JPT + CHB, $n = 90$) subjects (Table 4), thus our findings confirm HapMap data.

Comparison of allele frequencies between Caucasian and Asian subjects showed significant differences ($-\log P > 5.0$, $\chi^2$-test) at SNPs rs3768308, rs3738671 (Phe202Leu), and

**Table 3** Association between TRIT1 Phe202Leu polymorphism and survival of Italian or Norwegian lung cancer patients

| Genotype | Number of subjects | Events[a] | Hazard ratio[b] | 95% CI | $P$ |
|---|---|---|---|---|---|
| Italians |  |  |  |  |  |
| Phe/Phe | 288 | 163 | 1 |  |  |
| Phe/Leu | 28 | 20 | 1.7 | 1.03—2.86 | 0.039 |
| Norwegians |  |  |  |  |  |
| Phe/Phe | 218 | 129 | 1 |  |  |
| Phe/Leu | 28 | 11 | 0.5 | 0.26—0.91 | 0.023 |

[a] Number of deaths within 60 months of follow-up.
[b] Adjusted for sex, smoking, age at diagnosis (in decades), $N$ status (N0 or $N > 0$), and clinical stage (stage 1 or stage > 1).

**Table 4**  Type and allele frequency of SNPs in TRIT1 or MFSD2 genes by ethnicity[a]

| Gene | SNP | Chromosome position (kb)[b] | Gene location | Alleles[c] | Frequency of the rare allele | | −log P |
|------|-----|------------------------------|---------------|-----------|-------------|-------|--------|
| | | | | | Caucasians | Asians | |
| TRIT1 | rs7315 | 40079 | 3′-UTR | A/G | G (0.42) | A (0.34) | 3.2 |
| TRIT1 | rs3768308 | 40084 | intron | T/C | C (0.02) | C (0.31) | 10.9 |
| TRIT1 | rs3738671 | 40088 | Phe202Leu | T/A | A (0.03) | A (0.31) | 10.6 |
| TRIT1 | rs2294814 | 40092 | intron | A/C | C (0.49) | A (0.38) | 1.5 |
| TRIT1 | rs12086005 | 40095 | intron | C/A | A (0.00) | A (0.02) | 1.1 |
| TRIT1 | rs12028138 | 40097 | intron | T/C | C (0.00) | C (0.20) | 7.8 |
| MYCL1 | rs3134614 | 40136 | Ser362Thr | G/C | C (0.11) | C (0.00) | 4.3 |
| MYCL1 | rs3134613 | 40137 | intron (*Eco*RI) | T/G | G (0.46) | G (0.45) | 0.0 |
| MFSD2 | rs12072037 | 40193 | 5′-region | C/A | A (0.01) | C (0.49) | 23.0 |
| MFSD2 | rs4360512 | 40195 | intron | C/T | T (0.00) | T (0.03) | 1.7 |
| MFSD2 | rs12747592 | 40200 | intron | C/T | T (0.04) | T (0.00) | 2.1 |
| MFSD2 | rs3103778 | 40206 | intron | G/A | A (0.37) | A (0.49) | 1.4 |

[a] Genotypes were extracted from the HapMap data (http://www.hapmap.org/).
[b] According to the Ensembl database (http://www.ensembl.org/Homo_sapiens/index.html).
[c] Common/rare in Caucasians.

rs12028138 of TRIT1 gene, and rs12072037 of MFSD2 gene. On the other hand, MYCL1-*Eco*RI SNP (rs3134613) did not show different allele frequencies between Asian and Caucasian subjects (Table 4).

## 4. Discussion

The MYCL1-*Eco*RI polymorphism was the first germ-line variation reported to be associated with cancer patients' prognosis. Although contrasting results have been described on the effects of this SNP, a meta-analysis pointed to its significant role in the modulation of cancer prognostic parameters, in particular lymph node status and clinical stage of lung cancer patients [6]. This finding is most likely explained by LD of this SNP with a nearby functional polymorphism that affects lung cancer progression.

Our case-only association study on Italian lung ADCA patients identified a 106 kb LD block including the proto-oncogene MYCL1 and the two flanking genes TRIT1 and MFSD2, suggesting the candidacy of these genes as likely lung cancer progression modifiers. Our LD results are in agreement with those reported by the HapMap project that also shows a significant LD extending between the TRIT1 and MFSD2 genes in Caucasian (36 informative SNPs) as well as in Asian (37 informative SNPs) subjects (not shown).

MYCL1 is reported to be preferentially amplified in small cell lung cancer but not in other lung cancer histotypes [7] and, in fact, we detected no MYCL1 expression in either normal lung or ADCA tissue, indicating that it cannot explain the reported prognostic effect. Conversely, TRIT1 and MFSD2 are expressed in normal lung tissue and downregulated in lung ADCA samples, suggesting they might play a role as lung tumour suppressor genes. Consistent with this hypothesis, we found frequent allelic losses in lung ADCA samples in this region [12]. TRIT1 catalyzes post-transcriptional modification of tRNA molecules and might affect cancer progression

through deregulation of the synthesis of essential proteins. As for MFSD2, no information is currently available on its role except its protein product contains several transmembrane domains similar to those of other carbohydrates transporters.

TRIT1 non-synonymous Phe202Leu SNP was significantly associated with worse survival in Italian lung cancer patients, and the adjusted logistic regression analysis showed that the effect was independent from lymph node status or clinical stage. Further analysis of the functional role of the TRIT1 Phe202Leu SNP in modulating survival of lung cancer patients was conducted in an independent group, i.e., Norwegian lung cancer patients whose clinical and follow-up data were available. The rare Leu202 allele showed a similar frequency in Italian and Norwegian patients, but it was significantly associated with better survival in the Norwegian series (Fig. 2B, Table 3), unlike the Italian population. In the Norwegian series, association of Phe202Leu SNP in ADCA patients only showed the same trend without reaching statistical significance. Also, a combined analysis of both populations (ADCA only or all histotypes) did not show significant associations, as expected by the opposite trends of the Phe202Leu SNP in the two populations. Because the small number of Norwegian adenocarcinomas ($n = 106$), we cannot rule out the possibility that a larger replication series could confirm the findings on Italians.

As a further attempt, we generated haplotypes by the combination of TRIT1 and MFSD2 SNPs. TRIT1 rs3845570 (Pro11Pro) and MFSD2 rs1492701 SNPs were chosen because of the relatively high frequency of their rare allele, that would allow to generate haplotype groups containing a reasonable high number of patients for association analysis. Indeed, each of the four possible haplotype groups contained from 33 to 177 patients; each haplotype group was compared with the group of patients not carrying that specific haplotype for association with lung cancer survival, however, no significant associations were detected.

Whereas false-positive random results cannot be ruled out entirely to explain these contrasting observations, results reflecting a true effect would support findings of the previous meta-analysis on MYCL1-*Eco*RI polymorphism pointing to the major role of this chromosomal region in the control of lung cancer progression. However, contradictory results of the association studies in the two different populations argue against the identification of TRIT1 Phe202Leu as the responsible polymorphism, a notion also suggested by yeast complementation studies indicating comparable biochemical activity of the two alleles. A similar result was observed with the intronic KRAS polymorphism whose *Rsa*I A2 allele showed a significant association with disease survival of lung ADCA patients from two unrelated populations, but with an opposite trend [17]. Those findings are consistent with a role for genetic elements of those regions in the control of lung cancer progression and patients' survival, although the functional elements remain to be identified.

Association of MYCL1-*Eco*RI with lung cancer patients' survival was observed in all of four studies of Asians, but in none of three studies on Caucasians [6], including a study in an Italian [5] and a study in a Norwegian series [4]. As frequencies of TRIT1 Phe202Leu SNP and of additional SNPs in TRIT1 and MFSD2 genes are significantly different among Caucasian and Asian subjects (Table 4), these results raise the possibility of population-specific allele frequency patterns in this chromosomal region, which might involve also the polymorphism(s) controlling lung cancer progression, and they may suggest that Caucasian series may have a low power to detect the effect on survival of the functional polymorphism in the MYCL1 region.

The LD between TRIT1 and MFSD2 and the tumour suppressor effects exerted by TRIT1 on lung cancer cell lines *in vitro* [12] points to their candidacy as lung cancer progression modifier genes. Polymorphic genetic elements of the region containing the two genes may control cancer patients' survival by fine modulation of their expression and/or of their isoform pattern. Further studies to identify regulatory SNPs in this region and to genotype these SNPs in large lung cancer series, including Asian patients, are needed to address this point.

## Conflict of interest statement

None declared.

## Acknowledgements

## References

[1] Kawashima K, Shikama H, Imoto K, Izawa M, Naruke T, Okabayashi K, et al. Close correlation between restriction fragment length polymorphism of the L-MYC gene and metastasis of human lung cancer to the lymph nodes and other organs. Proc Natl Acad Sci USA 1988;85:2353—6.

[2] Shih CM, Kuo YY, Wang YC, Jian SL, Hsu YT, Wu HY, et al. Association of L-myc polymorphism with lung cancer susceptibility and prognosis in relation to age-selected controls and stratified cases. Lung Cancer 2002;36:125—32.

[3] Ge H, Lam WK, Lee J, Wong MP, Yew WW, Lung ML. Analysis of L-myc and GSTM1 genotypes in Chinese non-small cell lung carcinoma patients. Lung Cancer 1996;15:355—66.

[4] Tefre T, Borresen AL, Aamdal S, Brogger A. Studies of the L-myc DNA polymorphism and relation to metastasis in Norwegian lung cancer patients. Br J Cancer 1990;61:809—12.

[5] Spinola M, Nomoto T, Manenti G, Falvella FS, Brega Massone PP, Conti B, et al. Linkage disequilibrium pattern in the L-myc gene in Italian and Japanese non-small-cell lung-cancer patients. Int J Cancer 2001;95:329—31.

[6] Spinola M, Pedotti P, Dragani TA, Taioli E. Meta-analysis suggests association of L-myc EcoRI polymorphism with cancer prognosis. Clin Cancer Res 2004;10:4769—75.

[7] Makela TP, Saksela K, Alitalo K. Amplification and rearrangement of L-myc in human small-cell lung cancer. Mutat Res 1992;276:307—15.

[8] Spinola M, Conti B, Ravagnani F, Fabbri A, Incarbone M, Cataldo I, et al. A new polymorphism [Ser362Thr] of the L-myc gene is not associated with lung adenocarcinoma risk and prognosis. Eur J Cancer Prev 2004;13:87—9.

[9] Zienolddiny S, Campa D, Lind H, Ryberg D, Skaug V, Stangeland L, et al. Polymorphisms of DNA repair genes and risk of non-small cell lung cancer. Carcinogenesis 2006;27:560—7.

[10] Stanssens P, Zabeau M, Meersseman G, Remes G, Gansemans Y, Storm N, et al. High-throughput MALDI-TOF discovery of genomic sequence polymorphisms. Genome Res 2004;14:126—33.

[11] Gillman EC, Slusher LB, Martin NC, Hopper AK. MOD5 translation initiation sites determine *N*6-isopentenyladenosine modification of mitochondrial and cytoplasmic tRNA. Mol Cell Biol 1991;11:2382—90.

[12] Spinola M, Galvan A, Pignatiello C, Conti B, Pastorino U, Nicander B, et al. Identification and functional characterization of the candidate tumor suppressor gene TRIT1 in human lung cancer. Oncogene 2005;24:5502—9.

[13] Devlin B, Risch N. A comparison of linkage disequilibrium measures for fine-scale mapping. Genomics 1995;29:311—22.

[14] Weir BS. Genetic Data Analysis 2: Methods for Discrete Population Genetic Data. Sunderlands: Sinauer Associates, Inc.; 1996.

[15] Kaplan EL, Meier P. Non-parametric estimation from incomplete observations. J Am Stat Assoc 1958;3:457—81.

[16] Peto R, Peto J. Asymptotically efficient rank invariant test procedures. J R Stat Soc [A] 1972;135:185—206.

[17] Dragani TA, Hirohashi S, Juji T, Kawajiri K, Kihara M, Ono-Kihara M, et al. Population-based mapping of pulmonary adenoma susceptibility 1 [PAS1] locus. Cancer Res 2000;60:5017—20.