Trends in Cancer Research
Vol. 13, 2018

Original Communication

Age-related molecular analysis of the *INK4* locus in early onset lung cancer

Thomas Nichols[a], Omid Rouhi[b], and Lela Buckingham[b,*]

[a]College of Health Sciences, [b]Departments of Pathology, Rush University Medical Center, Chicago, IL 60612, USA.

ABSTRACT

Non-small cell lung carcinoma (NSCLC) accounts for about 85% of all lung cancers, with 2% of NSCLC cases under the age of 45. Previous studies have associated the *INK4* locus, a gene cluster encoding three tumor suppressor proteins with lung and other cancers. One of the encoded transcripts, *CDKN2A/p16*, is associated with biological aging. *INK4* locus expression was assessed by qRT-PCR in formalin-fixed lung tumor and non-malignant tissue. Non-malignant expression of *p14*, *p15* and *p16* RNA was similar to that of malignant (Mann-Whitney U = 511, p = 0.440). In tumor tissue, there was low *ANRIL* expression in both age groups (≤50 n = 45, Mean Rank 36.89; >50 n = 33, Mean Rank 43.06; Mann-Whitney U = 625, p = 0.235). *p14* was minimally expressed in non-malignant lung tissue and in both age groups, and only slightly higher in tumor tissue. *p15* demonstrated a similar pattern, while *p16* RNA was more highly expressed in the older group. This trend was reflected in protein expression. The relative levels of *ANRIL* expression were low and quite variable in malignant tissue, the median value being lower in non-malignant than in non-malignant tissue. SNP rs1063192 (C/T), located within the *ANRIL* transcript has been associated with age-related diseases and cancer including glioma and with modulation of *ANRIL* expression. Lack of the rs2811712 G allele was found with increased gene expression normalized to beta2 microglobulin (B2M) in *p16* [*p16*/B2M% = 1.26 (A/AA) vs 0.45 (AG), n = 77] and *p14* [*p14*/B2M% 0.0008 (A/AA) vs 0.008 (AG); n = 13 (p = 0.025)]. These data show measurable differences in *INK4* gene expression in this patient group. The age-related *INK4* locus may contribute to early onset of lung cancer in this patient group. Complex tumorigenic pathways in different histologies as well as in older vs younger tissues complicate clear designation of these effects.

KEYWORDS: *INK4* locus, lung cancer, gene expression, age, p16.

INTRODUCTION

Lung cancer is the leading cause of cancer deaths among men and women. Non small cell lung cancer (NSCLC) is the most frequently encountered type of lung cancer, making it an important target in the health community. NSCLC is further divided into 3 subtypes: Adenocarcinoma, originating from glands, accounting for about 40% of all lung cancer; large-cell carcinoma, originating in transformed epithelium, accounting for about 10 to 15%; and squamous cell carcinoma, with origins as its name suggests, and accounting for 25 to 30%.

Lung cancer is a disease of older individuals. Approximately 68% of individuals diagnosed with lung cancer are 65 or older, less than 4% of diagnoses come from individuals 45 years or younger [1]. Although rare, these early onset cases may offer useful diagnostic value for

*Corresponding author: LelaBArnell@gmail.com

future cases and possibly an insight into the disease process.

The transformation of lung cells into malignancy is the result of genetic and epigenetic aberrations of multiple targets. Therefore, there are many potential treatments for individuals by way of reversing the effects or the processes that lead to cancerous tissue formation. These changes occur in yet not fully described order as evidenced by stage-specific events [2]. There may also be age-specific events. For example, the age-related cyclin dependent kinase inhibitor, *CDKN2B/p16*, is thought to be inactivated late in the events that lead to carcinogenesis [3]. With this, it could be possible to determine how close someone may be to developing this illness and may even provide a temporal target for treatment or diagnosis. The stage of progression at the time of diagnosis may also predict the chances of survival [4].

INK4 Locus

The *INK4a/ARF* locus encodes proteins p16/INK4A, p15/INK4B and p14/ARF and the antisence RNA in the *INK4* locus or *ANRIL* (Figure 1A). The *ARF* (alternative reading frame) gene differs from *p16* by only 1 exon, but is functionally different in that its product regulates the activity of p53 while p16, and p15, regulate pRB. The *ANRIL* gene is antisense to the *p15/CDKN2B-p16/CDKN2A-p14/ARF* gene cluster with its 5' end only about 300 bp from the 3' of *p14/ARF* on the opposite strand. This region plays a role in the progression through the cell cycle and is altered in about 30-40% of human tumors [5]. The *INK4* proteins inhibit cyclin-dependent kinases, which would normally promote cellular growth. *ANRIL* has an association with cell regulation, epigenetic modifying enzymes, and disease pathogenesis, linking methylation and suppression of p16 (Figure 1B). This lncRNA may be the connection between genetic risk factors and disease characterization because of its ability to interact with both PcG proteins and *p16*. *ANRIL* may serve multiple functions due to varying primary, secondary, and tertiary structures interacting with a variety of proteins including the Polycomb Repressive Complexes (PRC) [6].

Expression of the *INK4* locus is regulated directly by the activity of PRC1 and PRC2, and the antagonist to PRC, the Trithoarax Group (TrxG) of proteins which bring about mono-ubiquination of histone H2A at lysine 119 (H2AK119), trimethylation of histone H3 at lysine 27 (H3K27). H3K27me3 is produced by a methyltransferase enzyme of PRC2 called EZH2 [7]. These histone modifications are associated with transcriptional silencing.

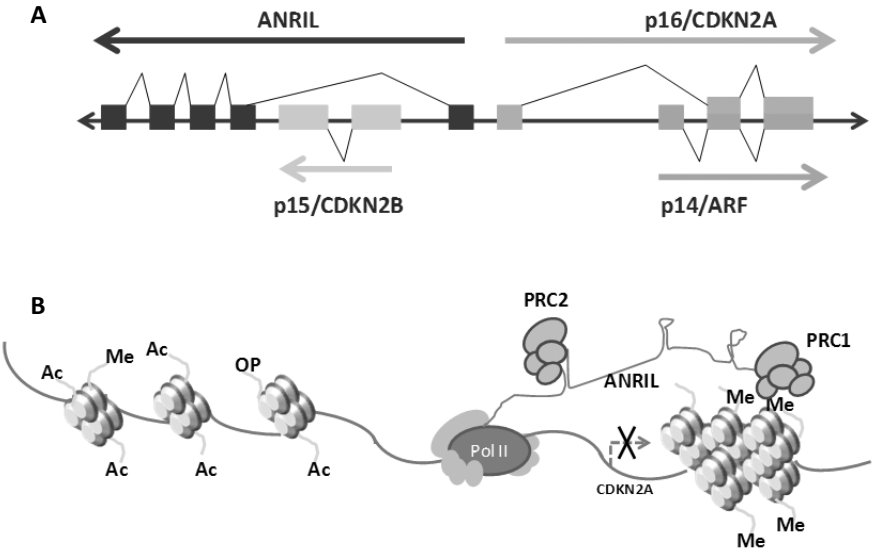

Figure 1. (A) Relative positions of each INK4 locus gene located on 9p21. **(B)** Promoter methylation of p16 *via* PRC1/2 through the assistance of ANRIL [7,13].

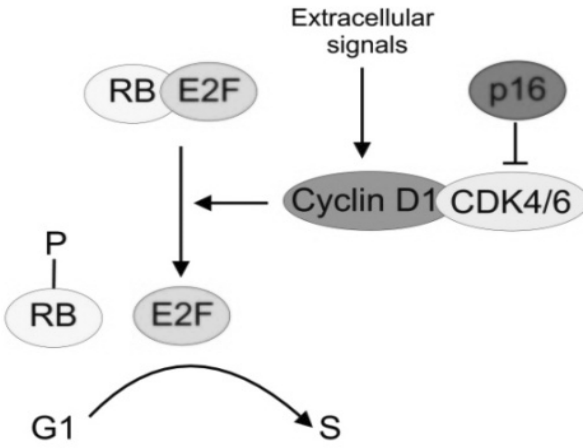

Figure 2. p16 inhibits CDK 4/6 kinase activity that phosphorylates retinoblastoma (pRb), required for movement of the cell division cycle through G1 (Peruala *et al*., 2013).

Immediate actions of p16 involve formation of an inhibitory complex with CDK4 (Figure 2) inhibiting the CDK substrate, cyclin D, halting the progression from G1 to S-phase. Loss of p16 through methylation of histone H3 at lysine 27 (H3K27me) by subunit EZH2 of PRC2 is related to an increased level of inactivated retinoblastoma and cell cycle progression.

Expression of p16 has been considered a marker of physical aging in humans and associated with cell senescence. Nielsen *et al.* (1999) showed a correlation existing between levels of p16 expression and increasing age [8]. A single nucleotide polymorphism (SNP) in the p16 locus (rs2811712) which may affect p16 expression also has significance in physical aging [9]. As cells age, p16 expression increases with loss of the ability to proliferate and senescence which eventually leads to tissue failure.

In lung cancer, the effect of p16 promoter methylation presents a changing threshold. We have observed that patients 60 years of age or less showing hypermethylation of the p16 promoter had a shortened duration of survival, in contrast to patients 60 years of age and older [10]. Because of the tumor suppression role of p16 and its inadvertent tendency to promote cell senescence, we further investigated the role of p16 and the *INK4* locus in lung cancer with young age of onset.

METHODS

Patients and samples

Formalin-fixed paraffin embedded (FFPE) lung tissue samples from 205 lung cancer patients were processed for this study. The tissue blocks were obtained from the Department of Surgical Pathology at Rush University Medical Center in Chicago, IL. Archival samples were chosen at random so long as the selection included an age range suited for this study. The study group included all stages of NSCLC, adenocarcinoma and squamous cell carcinoma. Some patients were treated with chemotherapy following surgery. The majority of patients were smokers. There was no gender bias (Table 1). Tissue sections were cut into 4 µm slices using a microtome. The sections were baked onto glass slides for RNA and DNA isolation.

RNA extraction

RNA was analyzed from malignant and/or non-malignant tissue from formalin-fixed paraffin-embedded (FFPE) tissue samples using the Ambion RecoverAll™ Total Nucleic Acid Isolation kit (Life Technologies, Grand Island NY). To ensure extraction of RNA from only malignant or non-malignant tissue, a reference slide was created. The reference slide consisted of one slide from each patient slide set stained with hematoxylin and eosin (H&E). A resident or attending pathologist was consulted to review each H&E stained slide and to circle the malignant or non-malignant tissue for use as a reference. Tissue was removed from each slide while juxtaposed over the reference slide with a single-use scalpel. Each scalpel was used for malignant or non-malignant tissue from multiple slides of the same patient. The tissue fragments removed from the slides were transferred in microcentrifuge tubes. Some instances required an extra step to successfully transfer the tissue scrapings into centrifugation vessels required for sample processing. This additional step used 10 µL of xylene from a sample's respective centrifugation tube placed onto the slide such that the tissue scrapings became saturated. The same xylene added was immediately removed *via* micropipette to leave behind a mass of tissue that was more easily manipulated. The tissues in the

Table 1. Patient demographics (n = 205[a]).

	≤60 Years (%)[b]	>60 Years (%)[b]	≤50 Years (%)[b]	>50 Years (%)[b]
Total Patients	80	125	25	183
Gender				
Male	21 (31.8)	61 (48.8)	6 (40.0)	76 (43.2)
Female	45 (68.2)	64 (51.2)	9 (60.0)	100 (56.8)
Smoking Status				
Yes	43 (87.8)	80 (80.0)	9 (75.0)	114 (83.2)
No	6 (12.2)	20 (20.0)	3 (25.0)	23 (16.8)
Chemotherapy				
Yes	28 (44.4)	36 (29.8)	8 (57.1)	114 (66.3)
No	35 (55.6)	85 (70.2)	6 (42.9)	58 (33.7)
Histopathology				
Adenocarcinoma	32 (52.4)	62 (39.2)	6 (60.0)	88 (50.0)
SCC	18 (29.5)	39 (24.7)	2 (20.0)	55 (31.2)
Other	11 (18.1)	57 (36.1)	6 (60.0)	33 (19.2)
Stage				
1a	19 (30.2)	35 (28.3)	2 (14.3)	52 (30.0)
1b	26 (41.2)	54 (43.5)	6 (42.9)	74 (42.8)
2a	3 (4.8)	8 (6.4)	1 (7.1)	10 (5.8)
2b	15 (23.8)	27 (21.8)	5 (35.7)	37 (21.4)
Ethnicity				
Caucasian	39 (68.4)	96 (80.7)	6 (60.0)	129 (77.7)
African American	18 (31.6)	20 (16.8)	4 (40.0)	34 (20.5)
Asian	0 (0)	3 (2.5)	0 (0.0)	3 (1.8)

[a]All demographic information was not available for all subjects.
[b]Percentages within groups.

tubes were de-waxed by adding 1 mL 100% xylene, centrifuged and dehydrated by adding 1 mL 100% ethanol. The dried ethanol pellet was resuspended into 100 µL digestion buffer (Life Technologies) and 4 µL protease for protein digestion and release of nucleic acid. The samples were incubated for 15 minutes at 50 °C and then 15 minutes at 80 °C. These lysates were then column-purified according to manufacturer's directions. The elution products were analyzed on the NanoDrop Lite spectrophotometer (Thermo Scientific, Waltham MA). Purity was measured as absorbance at a wavelength of 260 nm divided by absorbance at 280 nm. According to the Ambion RecoverAll kit directions, ratios should be in the range of 1.80 to 2.10. All elutions used in this study ranged in absorbance ratios from 1.80 to 2.08. Samples were stored at -20 °C.

DNA extraction

Tissue was distinguished as either malignant or non-malignant using the same protocol as described above for RNA extraction. DNA was obtained from slides carrying 4µm tissue sections de-waxed in 100% xylene (20 minutes in one bath and 5 minutes in another), rehydrated in ethanol alcohol baths of decreasing concentration (5 minutes in each bath of 100%, 95%, and 70%), followed by a final incubation in TE buffer for an additional 5 minutes. Tissue removal was completed by macro-dissection using a pipette and Tris KCl lysis buffer [10 mM Tris, 50 mM KCl, pH (8.0)]. Lysates produced were placed in nuclease-free tubes and incubated at 56 °C over night. Before use, each incubated lysate underwent heat denaturation at 95 °C for 5 min to reduce protease activity. Samples were stored at -20 °C.

Table 2. Primer and probe sequences.

	Allele	Outer/Inner	Forward primer sequence	Reverse primer sequence
rs2811712	A	Inner/Outer	5'- AAT GCC ATC TGA ATA AAC GA -3'	5'- ATG GAA TTT AGA TGT TCA AAC TTT A -3'
	G	Outer/Inner	5'- TAA GGG TAA TTG CTT ATT TCA GA - 3'	5'- TGAACAGAGAGCTTACTA TAT AAT TCT C -3'
rs1063192	C	Outer/Inner	5'- TCA TGA AAA ATT ATC CCT TGA A -3'	5'- CAT TTT CTT TAG TTT CCC TTA ATA TAA G -3'
	T	Inner/Outer	5'- GGA ATC TTT CCT AAT GAC ACC T - 3'	5'- ATA ATG TTT GAG GTT TGC TTT AAA T -3'

Beta 2 microglobulin (B2M)	Probe	5'-/56-FAM/CTGCCGTGT/ZEN/GAACCATGTGACTTTG/3IABkFQ/-3'
	Primer 1	5'-ACCTCCATGATGCTGCTTAC-3'
	Primer 2	5'-GGACTGGTCTTTCTATCTCTTGTAC-3'
ANRIL	Probe	5'-/56-FAM/CCCCAACCC/ZEN/CTAGGCCACA/3IABkFQ/-3'
	Primer 1	5'- TTTTCTCTTCACTACACCAGGG-3
	Primer 2	5'- AACAAGCCACAGACCAGTAC-3'
p16	Probe	5'-/56-FAM/TAGCAGTGT/ZEN/GACTCAAGAGAAGCCAGT/ 3IABkFQ/-3'
	Primer 1	5'-TGAGCTTTGGTTCTGCCATT-3'
	Primer 2	5'-AGCTGTCGACTTCATGACAAG-3'

Gene expression

Extracted RNA was converted to cDNA by reverse-transcriptase PCR (RT-PCR). The components of the reaction mix were as follows: 4 µL 5x RT buffer (Invitrogen-Life Technologies), 2 µL 0.1M DTT, 1 µL RNasin (Promega, Inc., Madison, WI), 1 µL Random Hexamer primer (1/30Xx; Promega), 1uL 10 mM deoxyribonucleotide mix (Thermo), 10 uL sample, and 1 uL of Moloney RT (Invitrogen). The reaction was carried out at 45 °C for 45 minutes with a subsequent enzyme deactivation at 95 °C for 5 minute in a Gene Amp 9700 thermal cycler (Applied Biosystems/Life Technologies).

Expression levels of *ANRIL* lncRNA and *p16* mRNA were determined using qPCR. Primers for qPCR (Prime Time, Integrated DNA Technologies, Coralville, IA; Table 2) were resuspended using 500 µL IX TE buffer (10 mM Tris, 0.1 mM EDTA, ph 8.0) to make a 20X stock solution. The qPCR reaction mix included 1.25 µL primer probe mix, 12.5 µL 2X universal master mix with ROX passive control day (Roche Diagnostics), 4 µL cDNA and 7.25 µL nuclease-free water to 25 µL total volume. Internal controls were beta-2 microglobulin and beta-actin. All qPCR was performed using the ABI PRISM® 7000 Sequence Detection System (Applied Biosystems) using the universal amplification program of 15 minutes at 50 °C followed by 50 cycles of 55 °C for 1 minute and 68 °C for 1 minute.

Immunohistochemistry

Immunohistochemistry was performed on 65 analyzable specimens. Immunohistochemistry (IHC) specimens were 5.0 mm sections of formalin-fixed paraffin-embedded tumor tissue or sections from cytology cell blocks. Immunostaining was performed with mouse monoclonal IgG2a anti p16 antibody (Santa Cruz Biotechnology, Inc.) using a 1:200 dilution as previously described. 17 Staining frequency and intensity of all tumor cells on each slide were estimated on scales of 0 to 4 without knowledge of clinical patient data. Intensity was judged from background and relative density of staining. For frequency, less than 1% positive tumor cells per field

was scored as 0, 0.1%-10% as 1, 11%-35% as 2, 36%-70% as 3, and over 70% as 4. IHC expression was dichotomized into two levels: positive (intensity X frequency > 4) and negative (intensity X frequency < 4).

Methylation analysis

Promoter methylation analysis has been described previously [3]. DNA was extracted by proteinase K digestion of tumor cells manually microdissected from paraffin-embedded tissue samples. Bisulfite treatment of DNA was performed using the Qiagen Epitech system according to manufacturer's protocol. The converted DNA was amplified with HotStarTaq® (Qiagen) according to the manufacturer's protocol using modified primers (Table 2). Amplicons were resolved by agarose electrophoresis to confirm proper amplification and quality of product. The detection of the C/T polymorphisms that result from PCR amplification of bisulfite-treated DNA was performed by Pyrosequencing™ (Qiagen) on a PyroMark MD pyrosequencer. For each C/T polymorphism, the relative percent luminescence generated from C (methylated) vs. T (unmethylated) nucleotides at potentially methylated cytosine residues of CG dinucleotides is reported. Non-CG cytosines, which should be 100% converted, are included in each sequence to confirm complete bisulfite conversion. This method was used to detect and quantify the degree of cytosine methylation within each sample promoter being investigated. The p16, -64 to -40 region was analyzed (A of ATG of the start of translation = +1). The output data included site-specific percent methylation for each CpG in the analyzed area as well as overall average percent methylation, defined as the average of the CpG levels within the promoter.

SNP detection

Regions of DNA containing the desired SNP were amplified using sequence-specific PCR (SSP-PCR; Table 2). Each polymorphism pair was investigated with primer sets designed using public software (http://primer1.soton.ac.uk/primer1.html). Each set included one primer for each SNP producing a unique PCR product for that polymorphism (Figure 3). Each SNP was amplified independently in two separate PCR reactions per SNP. The PCR reaction mix is as follows: 15.8 µL H_20, 0.5 µL of 10mM dNTP, 2.5 µL of 10X buffer (Qiagen, Hilden, Germany), 4.0 µL of a primer mix at a concentration of 5 µM/primer, 0.2 µL of Hot Start Taq polymerase, and 2.0 µL of sample. Negative controls included 2 µL of nuclease-free water (Fisher Scientific, Hanover Park, IL) in place of the sample. Sequence-specific PCR was performed using the Gene Amp 9700 thermal cycler. Beginning with activation at 94 °C for 15 minutes, the run proceeded for 45 cycles at 94 °C for 30 seconds; 52 °C for 45 seconds; and 72 °C for 30 seconds, and ended with 72 °C for 7 minutes. All reactions past this time were held at 4 °C. Post PCR products were visualized on 4% agarose-based gel made with GelPilot® LE Agarose (Qiagen). For ease of interpretation, post PCR reactions for each

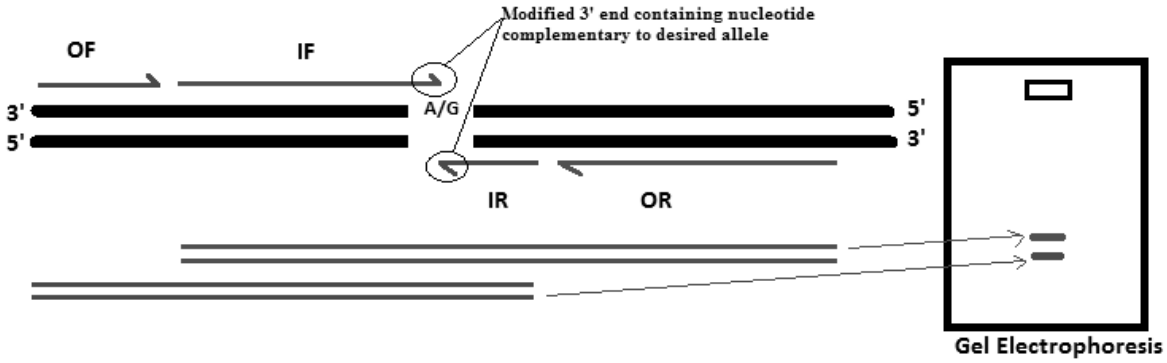

Figure 3. Relative positions of SSP-PCR primers. The size of the PCR product is indicative of the SNP present in the test DNA.

SNP allele of the same patient were run within the same well. Each well was loaded with 3.5 µL of each post-PCR reaction product totaling 14 µL plus an additional 5 µL of gel loading buffer (Promega). Ethidium bromide was added to the gel used to visualize PCR products under UV light.

Statistical analysis

Chi-square analysis was performed to determine the level of significance between age at the time of diagnosis and SNP allele present. Polymorphisms were categorized into groups of hetero/hemizygosity. Since age is a continuous variable patients were dichotomized as young or old at diagnosis using a cut-point of age 45. Statistical analysis of RNA expression levels required the Mann-Whitney to determine significance of association between malignant and non-malignant tissue as well as between age of diagnosis and malignant tissue. All statistical tests were performed using SPSS 16.0 Statistical Software Package.

RESULTS

INK4 locus gene expression and malignancy

Figure 4 shows the mean and standard deviations of normalized *INK4* locus gene expression in tumor and non-malignant lung tissue. Ninety-five

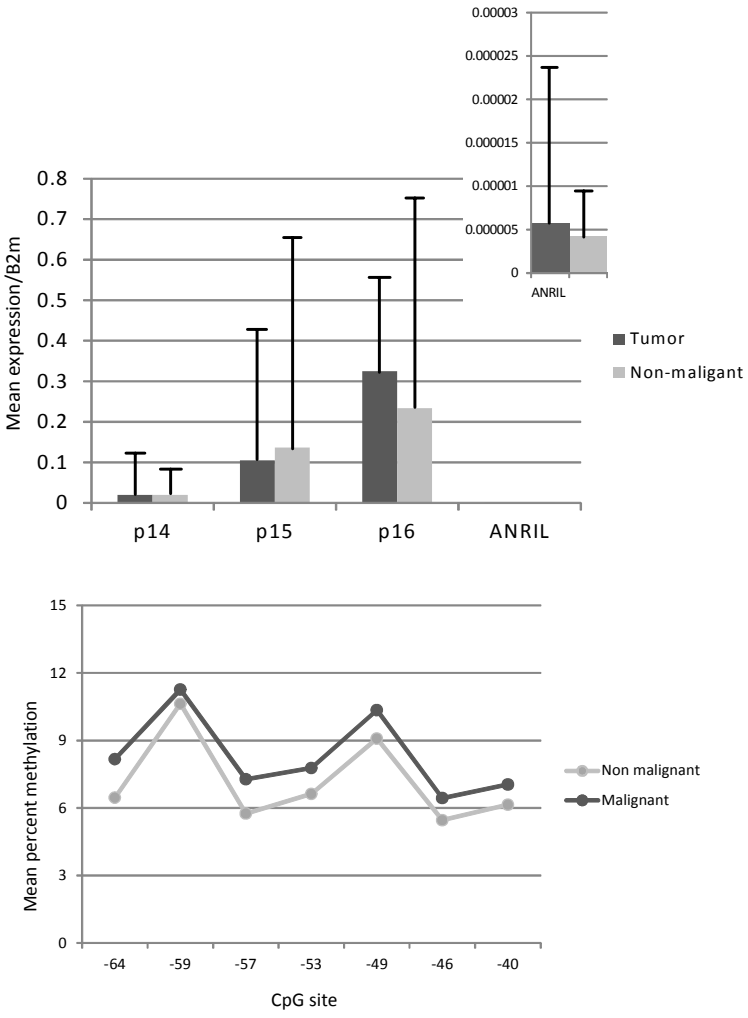

Figure 4. (A) Expression of *p14*, *p15*, *p16* and *ANRIL* normalized to beta 2 microglobulin in non-malignant and tumor tissue. Bars show means with one standard deviation (n = 126). **(B)** DNA methylation levels in non-malignant and malignant tissue at each CpG site in the *p16* promoter region (n = 129).

samples yielded detectable results. A Mann-Whitney test was performed to determine if normalized expression of the *INK4* locus was dysregulated in malignant tissue, compared to non-malignant tissue from the same patients. Non-malignant expression of *p16* RNA was similar to that of malignant ($p = 0.440$). This was also the case with p14 and p15 expression (Figure 4A. No significant difference was seen ($p = 0.761$), and the relative levels of *ANRIL* expression were low and quite variable, the median value being lower in non-malignant than in malignant tissue (malignant median: 0.0069; non-malignant median: 0.0009). If *ANRIL* down regulates p16 expression through histone modification and DNA methylation, the lower but variable *ANRIL* expression in non-malignant tissue compared with malignant tissue supports such a mechanism of regulation.

DNA methylation

Consistent with a previous study on early stage lung cancer, *p16* promoter methylation was not significantly higher in malignant versus non-malignant tissue in this patient group [11], which would result in the transcript expression pattern observed if expression of *p16* being responsive to increased promoter methylation. For the entire patient group, lower protein expression was observed with hypermethylation of all *p16* CpG sites tested (-64 to -40; A of ATG = +1; Figure 4B). In previous studies, hypermethylation at cytosine position -59 had the greatest influence on protein expression as measured by immunohistochemistry (intensity X degree of staining).

Methylation of the p16 promoter area in the *INK* locus was significantly higher in older cases than in younger cases (50 years or younger) in this patient group ($p < 0.001$). Methylation at gene promoters is dynamic, exhibiting both methylation and demethylation activities. Genomic methylation is known to increase with age. To assess genomic demethylation activity in the patient group, an enzyme-linked immunoassay (ELISA) for 5-hydroxymethyl cytosine (5-hmC) was performed. Levels of 5-hmC were consistently higher in younger patient groups using 60 years or 50 years ($p < 0.001$) cutpoints (Figure 5).

INK4 locus activity and age of diagnosis

Patients were dichotomized by age of diagnosis of $\leq$ 50 or >50 years. *ANRIL* was minimally expressed in tumor and normal tissue, regardless of age of diagnosis (Figure 6). In tumor tissue, there was low *ANRIL* expression in the older group and in the younger group ($\leq$50 n = 45, Mean Rank 36.89; >50 n = 33, Mean Rank 43.06; $p = 0.235$). *p14* was minimally expressed in non-malignant lung tissue and in both age groups, and only slightly higher in tumor tissue. *p15* demonstrated a similar pattern, while *p16* RNA was more highly expressed in the older group in both non-malignant and tumor tissue.

To further investigate if patterns of expression might be associated with early-age diagnosis, another Mann-Whitney test was performed comparing tissues from patients 50 years old and younger to those from patients older than 50 years. A relatively smaller mean rank (9.13) was observed in samples from patients 50 years old and younger (n = 12). With a mean rank of 14.82, patients over the age of 50 (n = 35) exhibited relatively higher levels of expression. The difference could be due to the temporal relation previously noted rather than aberrant expression.

This trend was reflected in protein expression. In tumor tissue p16 protein was higher in the older group than in the younger group (mean frequency X intensity 5.71 vs 4.18 in older and younger, respectively, $p = 0.200$). In non-malignant tissue p16 protein was slightly higher in the older group than in the younger group (mean frequency X intensity 2.40 vs 2.00 in older and younger, respectively, $p = 0.871$). This is consistent with previous observations of increased p16 protein expression in older tissues, despite higher p16 promoter methylation in these tissues [8, 10].

Single nucleotide polymorphisms in the *INK4* locus

Genotypic polymorphisms in the *INK4* locus may affect its activity through modifying gene expression or affecting secondary structure and therefore activity of *ANRIL*. The *p16* variant SNP rs2811712 (A/G) G allele has been associated with decreased fragility with advanced age. SNP rs1063192 (C/T) has been found to be associated with age-related diseases and cancer including glioma and with modulation of *ANRIL* expression. SNP rs4977756 is also associated with increased risk of glioma [12]. These SNPs are located in the

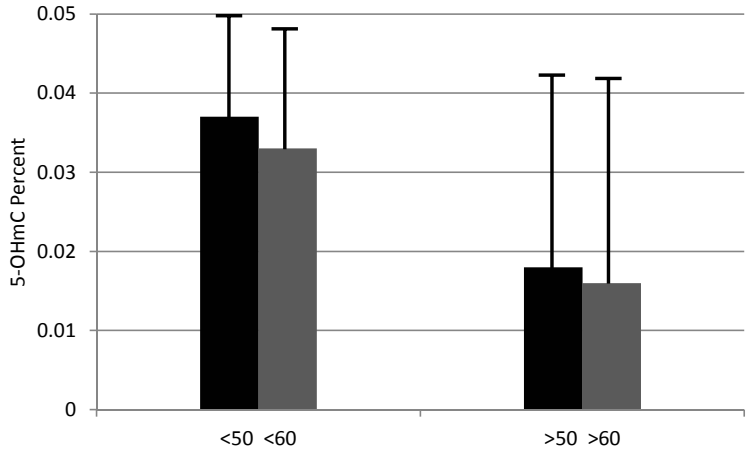

Figure 5. Genomic 5-hmC was assessed in total DNA from tumor tissue by ELISA. 5-hmC levels decreased with age. Using cutpoints of ≤50 or 60 years, there was significantly less 5-hmC detected in the older groups (n = 124).

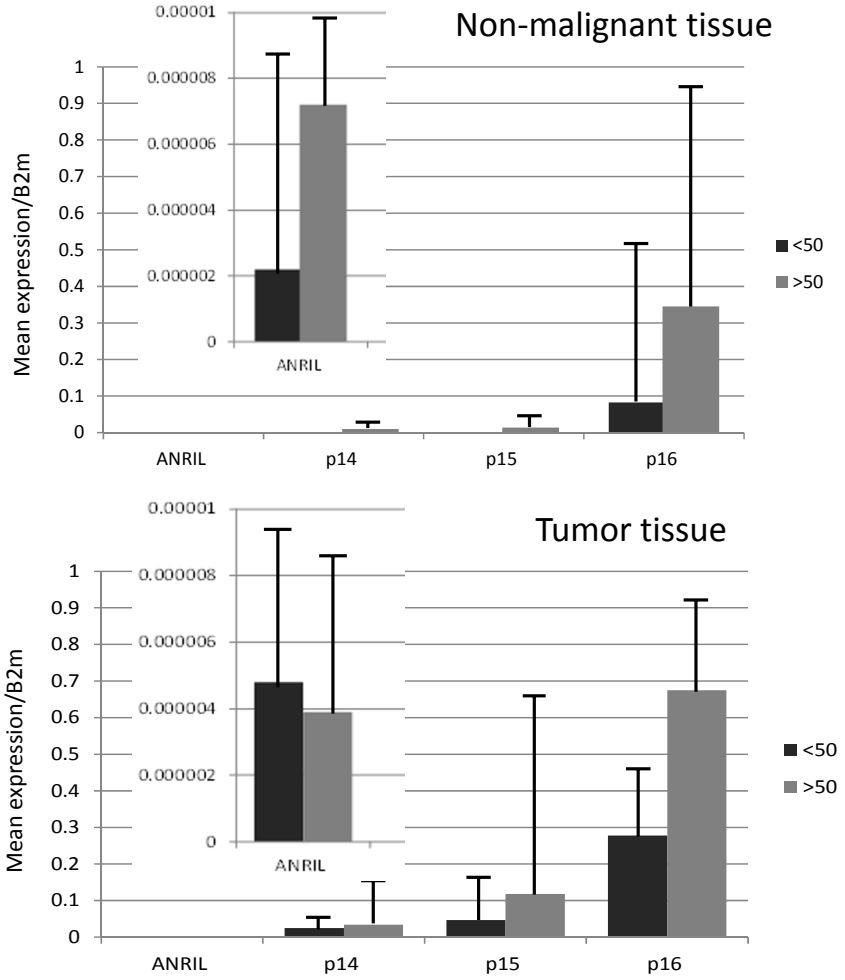

Figure 6. Mean normalized expression of *ANRIL*, *p14*, *p15* and *p16* in dichotomized age groups ≤50 and >50 in non-malignant and tumor tissue (n = 126).

CDK2NB region of *ANRIL* (Figure 1a). Allele frequencies are shown in Table 3.

SNP vs age of diagnosis

Previous studies reported an association of SNP alleles located in the *INK4* locus with age-related conditions (fragility, malignancy). To determine any influence of these variants in the *INK4* locus and age at the time of diagnosis in this lung cancer group, chi-square analyses for the association between age of diagnosis ($\leq$ 50 or >50) and SNP rs2811712 (A/G) was performed. The results demonstrated a tendency toward decreased frequency of the G allele in younger patients (p - 0.092; Figure 7). 12.1% of the younger lung patients were hemizygous/homozygous A, missing the G allele (n = 58), compared to 6.2% of older patients (n = 64).

In contrast to rs2811712, chi-square analyses of rs1063192 and rs4977756 showed no such associations. For rs1063192, 20.0% of the younger lung patients were hemizygous/ homozygous T, missing the C allele (n = 15), compared to 19.0% of older patients (n = 58). With respect to rs4977756, 9.5% of the younger lung patients were hemizygous/homozygous A (n = 2), compared to 3.7% of older patients. Heterozygosity in rs4977756 was observed in 61.9% of younger patients and 68.4% of older patient. Kaplan Meier

Table 3. Allele frequencies of INK4 polymorphisms in the current study (top panel). Bottom panel: observed range of allele frequencies in world populations (n = 112).

Polymorphism	**Allele**	**(%)**	**Allele**	**(%)**	**Allele**	**(%)**
rs2811712	AA	87.5	AG	12.5	GG	0.0
rs1063192	CC	6.9	CT	62.7	TT	29.4
rs4977756	AA	8.6	AG	69.6	GG	21.7
Population Frequencies[a]						
Polymorphism	**Allele**	**(%)**	**Allele**	**(%)**	**Allele**	**(%)**
rs2811712	AA	79	AG	14	GG	7
rs1063192	CC	21	CT	48	TT	31
rs4977756	AA	34	AG	48	GG	18

[a]dbSNP (http://www.ncbi.nlm.nih.gov/SNP/)

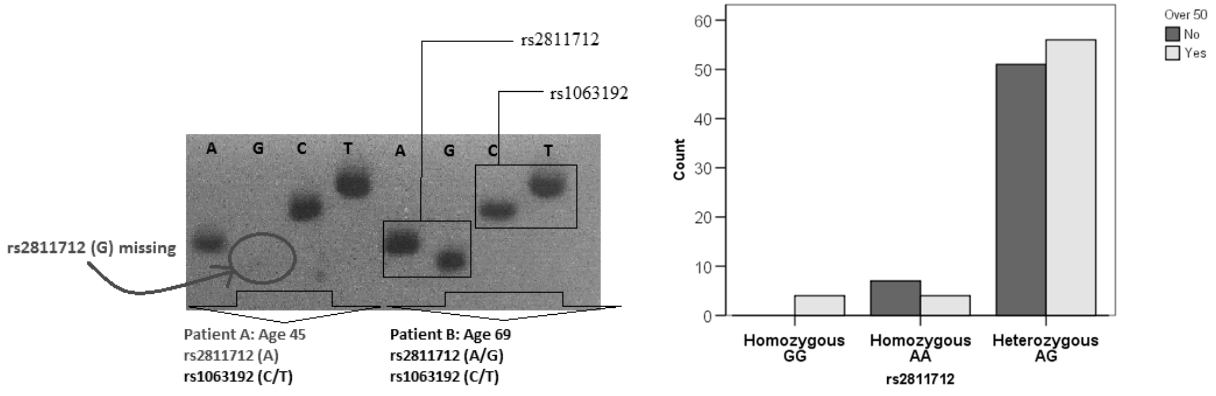

Figure 7. Example of missing rs2811712 (G) in a younger lung cancer patient and rs2811712 (A/G) in a patient 45 years of age (left). Homozygous AA (or hemizygous A) was more frequently observed in the younger patient group as shown in the graph on the right. Another SNP, re1063192, did not show this tendency.

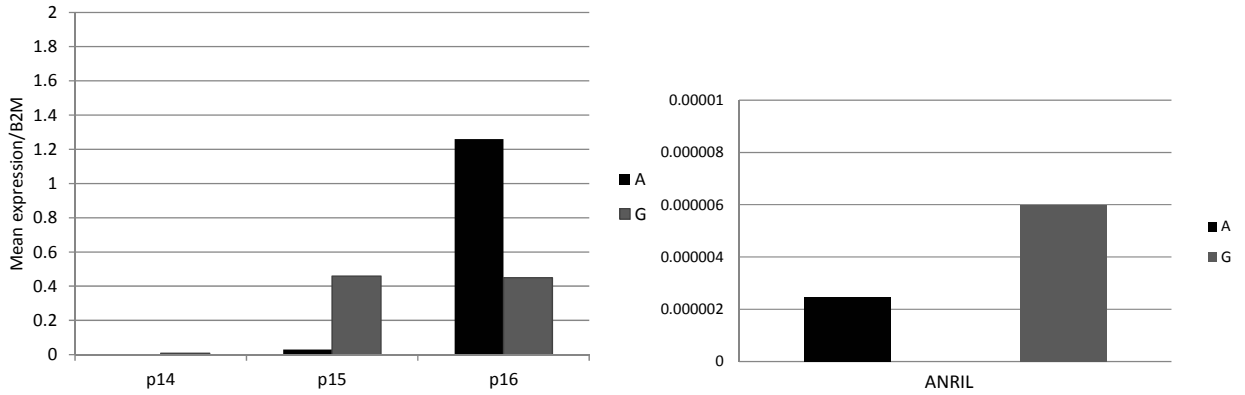

Figure 8. *INK4* gene expression in different rs2811712 genotypes (n = 77).

analysis of time to tumor recurrence and overall survival did not reveal significant associations with the SNP genotypes analyzed.

Somatic versus inherited polymorphism

Polymorphisms in matched non-malignant tissue were compared to their respective malignant tissue. Nineteen direct comparisons were made between malignant tissue and paired non-malignant tissue. The rs2811712 allelic genotypes of matched non-malignant tissue were consistent with the tumor genotypes in all cases except two (89% concordance). Both exceptions were heterozygous in malignant tissue and hemi/homozygous in non-malignant tissue, one missing the G allele and one missing the A allele. SNP rs4977756 allelic genotypes of matched non-malignant tissue were also consistent with the tumor genotypes in all cases except one (95% concordance). The exception was hemi/homozygous in malignant tissue and heterozygous in non-malignant tissue, missing the G allele. These data are consistent with a germline origin of the SNPs genotype; however, additional analyses would confirm whether the observed variants were germline or somatic.

SNP rs2811712: Somatic mutation versus *INK4* locus gene expression

To determine if the effect of the SNP genotype would be manifested through changes in gene expression, the effects of rs2811712 on expression of genes in the *INK4* locus was analyzed. Although there were relative differences in average expression of p14, p16 and ANRIL between AA and AG genotypes (Figure 8), lack of the G allele was found with increased p16 [p16/B2M% = 1.26 (A/AA) vs 0.45 (AG), n = 77] and p14 [p14/B2M% 0.0008 (A/AA) vs 0.008 (AG); n = 13 (p = 0.025)] gene expression, but decreased expression of p15 [p15/B2M% = 0.03 (A/AA) vs 0.46 (AG), n = 22] and ANRIL [ANRIL/B2M% = 0.00000247 (A/AA) vs 0.00000598 (AG); n = 78].

DISCUSSION

Cyclin dependent kinase 2A (CDKN2A/p16) protein acts to arrest the cell in the G1 phase of the cell cycle. Damaged cells under arrest eventually become senescent rather than replicating into possibly malignant clones. This is important with advancing age as the possibility of carcinogenesis increases. Previous studies have shown that *p16* expression increases with age resulting in protection against tumor formation. This study addresses genetic and epigenetic regulatory mechanisms in the *p16* (*INK4*) gene locus in lung cancer.

Age-related levels of *p16* expression are well-supported, and previous studies have shown an inverse relationship between the lncRNA, *ANRIL* and *p16* expression [13]. This relationship is consistent with the scaffold-guide functions of *ANRIL* and the silencing methylation activity exhibited by associated PcG proteins. Further, *ANRIL* has been implicated in the progression of cancer through MAPK signalling pathways [14]. Assuming this putative mechanism of *p16* regulation through *ANRIL* is correct, it follows

that *ANRIL* transcripts would be more highly expressed in the younger lung tissue samples. This elevated expression would lead to increased *p16* promoter methylation, *p16* silencing, and progression through the cell cycle; a mechanism proposed to be diminished with age to protect the body from carcinogenesis. Although the *p16* age-related findings on methylation [10] and expression in tumors are consistent with this model, the *ANRIL* expression levels were too low to measure definitive differences with the methods used. The observations do not address, however, the overall mechanism accounting for these differences, but they may hint as to why patients of early-age diagnosis tend to suffer from a more aggressive cancer; more *p16* inhibition means less cell cycle control, resulting in a higher rate of division. The age-related differences observed in this study may simply be a result of age-related phenotypes already in place.

Understanding how lncRNA molecules are regulated may also be of help in determining their true nature. One study found that *ANRIL* expression is induced by E2F1 in the ataxia telangiectasia mutated-dependent (ATM-dependent) pathway, just after DNA repair has been initiated [15]. This time-specific up-regulation, resulting in *INK4* protein repression, is in place to keep cells from dividing before DNA damage is fixed. Conversely, the same study by Wan *et al*. also noted that less expression of *ANRIL* led to an increase in apoptosis, in the presence of DNA damage. A negative feedback loop was proposed as a form of regulating the activity of E2F by *p14/ARF* [16]. The promoter regions of *p14/ARF* and *ANRIL* overlap and are co-regulated by E2F1 [17]. If INK4 proteins are to be expressed to arrest the cell cycle, then *ANRIL* must remain inactive. But, as E2F acts as a transcription factor on p14 it may also act on *ANRIL*. The negative feedback of *p14/ARF* on E2F is a way of keeping the *INK4* locus proteins expressed while down-regulating their regulatory mechanism of repression. These observations may partially explain the low *ANRIL* expression in non-malignant tissue.

Similar to the functional variety of proteins, lncRNAs can be classified into four groups: scaffolding, decoys, guides, or signals [18]. *ANRIL* has two known functions: guidance, to mediate site-specific methylation by PcG proteins; and a scaffold to bind PcG proteins (Figure 1B). Just as proteins may be affected by changes in sequence, lncRNAs have the same potential. Aside from intentional regulatory mechanisms, lncRNAs may be affected by polymorphisms depending on their genomic location [19]. While Cunnington noted that SNP rs2811712 was significantly associated with *CDKN2B* (*p15*) expression, and not *ANRIL* expression, he did not mention the functionality of *ANRIL* related to investigated SNPs. A SNP may not affect expression levels, but still affect the overall structure and, therefore, function of the lncRNA. *ANRIL* displays complex secondary structure including circular or linear shapes [20]. It seems plausible that the effects of polymorphisms associated with different diseases found in non-coding regions of DNA manifest themselves as altering secondary structure and function. Comparing the minor allele frequency for rs2811712 (G) in various ethnic groups to the occurrence of NSCLC in each group, reveals a similarity. The incidence and deaths from lung cancer seems to correlate with the presence of the minor (G) allele. Because decreased frailty is associated with rs2811712 (G) and with increased cellular proliferation [21] there may be an increased chance of tumorigenesis with the minor (G) allele.

CONCLUSION

The results presented support a role of epigenetic factors in lung cancers arising at younger ages. Higher expression of *ANRIL* (which was observed to be increased in malignant tissue) in normal tissue of younger patients along with lower expression of tumor p16 support altered epigenetic activities in the younger cancer cases. An effect of the age-related SNP tested on *ANRIL* sequence structure and/or activity was not observed; however, this does not rule out other possible effects of other sequence changes or multiple changes.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest. This work was funded through philanthropic donations to Rush University and the Department of Pathology.

REFERENCES

1. American Cancer Society: Cancer Facts & Figures 2015.
2. Liang, J., Lv, J. and Liu, Z. 2015, Tumour Biology, 36, 6391-6399.
3. Okamoto, A., Hussain, S. P., Hagiwara, K., Spillare, E. A., Rusin, M. R., Demetrick, D. J., Serrano, M., Hannon, G. J., Shiseki, M., Zariwala, M., Zariwala, M., Xiong, Y., Beach, D. H., Yokota, J. and Harris, C. C. 1995, Cancer Research, 55, 1448-1451.
4. Alzahouri, K., Martinet, Y., Briançon, S. and Guillemin, F. 2006, European Journal of Cancer Care, 15,348-354.
5. Kotake, Y., Nakagawa, T., Kitagawa, K., Suzuki, S., Liu, N., Kitagawa, M. and Xiong, Y. 2011, Oncogene, 30, 1956-1962.
6. Li, C. and Chen, Y. 2016, International Journal of Biochemistry and Cell Biology, 45, 1895-1920.
7. Yap, K., Li, S., Muñoz-Cabello, A. M., Raguz, S., Zeng, L., Mujtaba, S., Gil, J., Walsh, M. J. and Zhou, M. M. 2010, Molecular Cell, 38, 662-674.
8. Nielsen, G., Stemmer-Rachamimov, A. O., Shaw, J., Roy, J. E., Koh, J. and Louis, D. N. 1999, Laboratory Investigation, 79, 1137-1143.
9. Melzer, D., Frayling, T. M., Murray, A., Hurst, A. J., Harries, L. W., Song, H., Khaw, K., Luben, R., Surtees, P. G., Bandinelli, S. S., Corsi, A. M., Ferrucci, L., Guralnik, J. M., Wallace, R. B., Hattersley, A. T. and Pharoah, P. D. 2007, Mechanisms of Aging and Development, 128, 370-377.
10. Bradly, D., Gattuso, P., Pool, M., Basu, S., Liptay, M., Bonomi, P. and Buckingham, L. 2012, Diagnostic Molecular Pathology, 21, 207-213.
11. Buckingham, L., Faber, L. P., Kim, A., Liptay, M., Barger, C., Basu, S., Fidler, M., Walters, K., Bonomi, P. and Coon, J. 2010, International Journal of Cancer, 126, 1630-1639.
12. Lu, H., Yang, Y., Wang, J., Liu, Y., Huang, M., Sun, X. and Ke, Y. 2015, International Journal of Clinical Experimental Medicine, 8, 17480-17488.
13. Pasmant, E., Sabbagh, A., Vidaud, M. and Bièche, I. 2011, FASEB Journal, 25, 444-448.
14. Xu, X., Wan, X. and Zhang, Z. 2016, International Journal of Clinical and Experimental Pathology, 9, 10803-10809.
15. Wan, G., Mathur, R., Hu, X., Liu, Y., Zhang, X., Peng, G. and Lu, X. 2013, Cell Signal, 25,1086-1095.
16. Mason, S., Loughran, O. and La Thangue, N. B. 2002, Oncogene, 21, 4220-4230.
17. Sato, K., Nakagawa, H. and Tajima, A. 2010, Oncology Reports, 24, 701-707.
18. Da Sacco, L., Baldassarre, A. and Masotti, A. 2012, International Journal of Molecular Science, 131, 97-114.
19. Cunnington, M., Santibanez Koref, M., Mayosi, B. M., Burn, J. and Keavney, B. 2010, PLoS Genetics, 6, e1000899.
20. Burd, C., Jeck, W. R., Liu, Y., Sanoff, H. K., Wang, Z. and Sharpless, N. E. 2010, PLoS Genetics, 6, e1001233.
21. Sharpless, N. and DePinho, R. A. 2007, Nature Reviews in Molecular and Cellular Biology, 8,703-713.