IDENTIFICATION OF -308 TNF-α PROMOTER SINGLE NUCLEOTIDE POLYMORPHISM IN ACTIVE AND PASSIVE SMOKERS

Gracia A. V. Pollo1), Rooije R.H. Rumende1), Trina E. Tallei1)

1) Department of Biology, Faculty of Mathematics and Natural Sciences, University of Sam Ratulangi
E-mail: gracia_pollo@yahoo.com, rooije.rumende@yahoo.com, trina_tallei@unsrat.ac.id

ABSTRACT

Tumour necrosis factor-α (TNF-α) is one of the pro-inflammatory cytokines that play a role in the inflammatory process, immune system development, and apoptosis. This protein is encoded by the TNF-α gene. Several studies had linked the presence of polymorphisms in the TNF-α promoter region with susceptibility to the onset of chronic obstructive pulmonary disease (COPD) in active and passive smokers. This study aimed to identify and analyze the nucleotide polymorphism in TNF-α gene promoter regions in active and passive smokers in Manado. The method used in this study included blood sampling, DNA extraction, amplification of -308 TNF-α promoter region, sequencing and data analysis. The softwares used for data analysis were Geneious, Multalin, Clustal Omega and DnaSP. DnaSP was used to compute the level of polymorphism based on haplotype statistics. The results did not show single nucleotide polymorphism at position -308 in TNF-α gene promoters in active and passive smokers. This was because at that position, the nucleotides were both in the form of guanosine monophosphate and there were no mutation caused by cigarette smoke exposure.

Keywords: active smoker, passive smoker, single nucleotide polymorphism, TNF-α promoter

BACKGROUND

Smoking activity has become a habit of some people in Indonesia (Kusumawardani et al., 2018). Smoking will produce smoke because of the burning of the cigarettes. Tobacco is the main ingredients of cigarettes and often supplemented with sugar as an additional sweetener (Cheah et al., 2018).

Cigarette smoke that is produced by active smokers will affect the surrounding people and it is inhaled by passive smokers. Passive smokers or second hand smokers are people who do not smoke but exposed continuously to cigarette smoke (Reimondos et al., 2012) for a long period of time.

The exposure of smoke to second hand smoker caused various diseases
(Kementerian Kesehatan Republik Indonesia, 2015), because it affected the human body until the cellular level (Luetragoon et al., 2017). The smoke could cause changes in the DNA sequence, such as in lung tissue (Li et al., 2014) and peripheral blood cells (Luetragoon et al., 2017). The impact of cigarette smoke on peripheral blood cells was in the form of oxidative stress in DNA (Gao et al., 2017) and DNA structure damage (Dobrzynska et al., 2014). These changes could have an impact on gene expression, which in turn will affect cellular interactions (Jamil et al., 2017).

The influence of cigarette smoke at the molecular level of cells could be analyzed in several ways, such as detection of DNA adduct, analysis of oxidative stress DNA, and detection of DNA polymorphism (Phillips and Venitt, 2012). DNA adduct detection was a method detecting certain DNA segments that were covalently bound to the chemical compounds of cigarette smoke (Yun et al., 2018). Oxidative stress analysis was an analysis of the changes in oxidant and antioxidants balance that caused damage of cell structure and even DNA structure (Birben et al., 2012). The detection of single nucleotide polymorphism was the detection of a single nucleotide variation in a specific sequence of genomic DNA (Cui et al., 2018). The polymorphism could be detected using some methods, such as ARMS (Amplification Refractory Mutation System), RLPF (Restriction Length Fragment Polymorphism), DGGE (Denaturing Gradient Gel Electrophoresis), SSCP (Single Strand Conformation Polymorphism), allele discrimination, High-Throughput Assay Chemistry, DNA Arrays, Pirosequencing, and Light Cycler (Koopae and Koshkoiyeh, 2014).

Single nucleotide polymorphism is often conducted because the process is relatively short, accurate, easy and accurate (Cui et al., 2018). On the other hand, the other methods have weaknesses, such as the difficulty in measuring oxidative DNA stress, the effect of individual nutritional status on the results (Kawamura and Muraoka, 2018), the difficulty of identifying unknown DNA adducts and the use of dangerous radioactive phosphate markers (Klaene et al., 2013). Analysis of the effect of smoking activity using SNP detection was carried out on C-reactive protein genes, perforin genes (Luetragoon et al., 2017), and tumor necrosis factor alpha (TNF-α) at position -308 from promoters in Mexico and Taiwan populations (Resendiz et al., 2018; Wu et al., 2018). Polymorphism at position -308 in TNF-α gene promoter was an important marker because of TNF-α was useful in the inflammatory process, development of the immune system, and apoptosis (Gjvestad et al., 2015), so that the emergence of these gene polymorphisms was important to be evaluated in relation with smoking. A research on the identification of -308 TNF-α promoter polymorphism in active and passive smokers in Manado, therefore, was conducted.

**MATERIAL AND METHODS**

**Place and Time**

Blood sampling and sample preparations were carried out in the Biotechnology Laboratory, Department of Biology, University of Sam Ratulangi. Sequencing was carried out by the provider of 1BASE Malaysia sequencing services. Data analysis was performed in the Biotechnology Laboratory, Department of Biology, University of Sam Ratulangi. This research was conducted from October to November 2018.

**Materials**

The tools used in this research included EDTA tubes, lancing device and its needles, Eppendorf tubes, micropipettes, micropipette tips and spin filters. The used material were alcohol swabs, Qiagen Extraction Kit: Blood and Tissue, a pair of primers, MyTaqTM HS Red Mix Bioline (PCR master mix), gloves, and masks.

The DNA were obtained from peripheral blood samples of two active smokers (GP1 and GP2), two passive smokers (GP3 and GP4) and two non-smokers (GP5 and GP6) who were not exposed to cigarette smoke.

**DNA Extraction**

Blood samples were collected from each individual using lancets (lancing devices). The DNA extraction from blood samples was carried out using Qiagen
Extraction Kit: Blood and Tissue according to the guidelines provided by the manufacturer. The obtained DNA was used for the PCR amplification process.

**Amplification of the TNF-α Gen Promoter Region**

The first step of amplification process was mixing the amplification reactants. The amplification process for each of the DNA sample required 20 μl of the PCR master mix, 15 μl ddH2O, 1.5 μl of each primer, and 2 μl of the DNA sample. All of the reactants was poured in the Eppendorf tube. Each reaction mix was placed into a thermal cycler. The amplification process was carried out using the following temperature setting: of 95ºC for 3 min for initial denaturation, 52ºC for 30 sec for annealing, and 72ºC for 30 sec elongation. The cycle was repeated 34 times (Tallei et al., 2016). The final elongation was set at 72ºC for 5 min. Sample was sequenced bidirectionally at 1BASE Malaysia sequencing service providers twice (forward and reverse) from different directions, according to the available primers. Primers were designed using Primer 3 software (Pollo et al., Unpublished) as follow: forward primer 5'-ATTATGAGTCTCCGGTCC-3' and the reverse primer 5'-GGGTCAGTATGTGAGAGGAGA-3'. The sequenced DNA samples were given additional label after the blood sample label which either _F for samples sequenced by forward direction or _R for samples sequenced by reverse direction.

**Data analysis**

The sequencing results were analysed using procedures from Tallei and Kolondam (2015). The chromatogram was edited using Genious v 5.6. Geneious integrated global alignment process was used in the pairwise alignment to combine sequencing results. Each end of the sequence was cut by about 50 nucleotides to avoid reading errors. Identification of the results of sequencing was done using the Basic Local Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Sequences of TNF-α gene promoter regions together with comparative sequences derived from the NCBI Database (https://www.ncbi.nlm.nih.gov/) were aligned using Multalin V.5.4.1 developed by Corpet (1988) (http://multalin.toulouse.inra.fr/multalin/). All gaps that appeared were considered as missing characters (Zhang et al., 2014). The length of the sequence was then adjusted manually. The final alignment and percentage identity matrix (percent identity matrix) was made using Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/) (Chhabra et al., 2018). Analysis of nucleotide polymorphisms was carried out using DnaSP software. This software computes the level of polymorphism based on statistical tests. The type of statistics in DnaSP used was Haplotype-Based Statistics (Rozas, 2009).

**RESULTS AND DISCUSSION**

**Sequencing Results**

The sequencing results (Table 1) broadly showed that the quality of the sequencing quality were different among the sample groups. Low quality (less than 65%) was found in the samples of active smokers. The passive smoker sample showed quite good sequencing quality in the GP3 sample and was very poor in the GP4 sample, while the non-smoker samples had good sequencing quality (76.2 - 94.3%).

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Sample Type</th>
<th>Sequencing quality (%)</th>
<th>Sequence Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GP1_F</td>
<td>Active Smoker</td>
<td>60.6</td>
<td>611</td>
</tr>
<tr>
<td>GP1_R</td>
<td>Active Smoker</td>
<td>46.9</td>
<td>588</td>
</tr>
<tr>
<td>GP2_F</td>
<td>Active Smoker</td>
<td>64.6</td>
<td>483</td>
</tr>
<tr>
<td>GP2_R</td>
<td>Active Smoker</td>
<td>56.2</td>
<td>473</td>
</tr>
<tr>
<td>GP3_F</td>
<td>Passive Smoker</td>
<td>82.1</td>
<td>821</td>
</tr>
<tr>
<td>GP3_R</td>
<td>Passive Smoker</td>
<td>95.5</td>
<td>801</td>
</tr>
<tr>
<td>GP4_F</td>
<td>Passive Smoker</td>
<td>0.0</td>
<td>0.5</td>
</tr>
<tr>
<td>GP4_R</td>
<td>Passive Smoker</td>
<td>390</td>
<td>803</td>
</tr>
<tr>
<td>GP5_F</td>
<td>Non-Smoker</td>
<td>90.0</td>
<td>848</td>
</tr>
<tr>
<td>GP5_R</td>
<td>Non-Smoker</td>
<td>86.0</td>
<td>850</td>
</tr>
<tr>
<td>GP6_F</td>
<td>Non-Smoker</td>
<td>76.2</td>
<td>821</td>
</tr>
<tr>
<td>GP6_R</td>
<td>Non-Smoker</td>
<td>94.3</td>
<td>803</td>
</tr>
</tbody>
</table>
The compounds contained in the cigarette smoke could form DNA adducts that bind to DNA, such as benzo[a]pyrene, PAH, formaldehyde, acrolein, 4- (methylnitrosamino) -1- (3-pyridyl) -1- butanone (NNK) and N ‘-nitrosonornicotine (NNN) (Phillips and Venitt, 2012). Isopropylene was a DNA adduct that had been studied. This compound could be bound to DNA template and caused the performance of the Taq polymerase enzyme to stop (Cimino et al., 1990). The low amplification and sequencing results in samples of active and passive smokers presumably showed the indirect influence of cigarette smoke. Process of amplification and sequencing could be inhibited by the DNA adduct.

The results of amplification and sequencing had a positive correlation with the intensity of smoking in active smokers and exposure to second hand smoke in passive smokers. This could be seen in the history of the duration of the sample individuals being active and passive smokers. The GP1 samples had a history of smoking for more than 10 years while GP2 for less than 5 years. The GP3 samples had a history of passive smoking for less than 5 years with exposure intensity of 0-25% per day. The GP4 samples was exposed to passive smoking for 6 to 10 years approximately with exposure intensity of 0-25% per day.

The low quality of sequencing of GP4 samples was suspected to be caused by the performance of the Taq polymerase enzyme that could be inhibited by DNA adducts, but it could also be caused by methylation due to DNA adducts and polymorphisms found in the primers attachment sites. The effect of cigarette smoke could lead to an increase in methylation especially in the form of O\(^\text{6}\)-methylguanine DNA adduct due to NNK induction (Hecht, 2002). Methylation found in template DNA could reduce the effectiveness of the PCR process (Kiselev et al., 2015) by the presence of 5-methylcytosine and DNA methylation status. 5-methylcytosine could increase melting temperature (Tm) in DNA (Diele et al., 2010) and DNA methylation status could affect the denaturation level of PCR in human DNA (Bunyan et al., 2011). In the polymerase chain reaction (PCR) cycle, DNA methylation could specifically affect the first cycle of amplification. The efficiency of the first cycle could then affect the efficiency of subsequent cycles (Kiselev et al., 2015).

The quality of the sequencing products in GP4 was low because of the forward primer sequence had 3 polymorphism positions at the last 8 nucleotides, whereas the reverse primer sequence had 1 polymorphism that resulted in a slightly higher quality of 0.5%. The polymorphism found in the primer annealing sequence could decrease primer efficiency, because templates and primers were no longer complement to each other. The template for forward primer sequence had 5 single nucleotide polymorphisms (SNP) sites as follows: 10 (rs1395411507 C / T), 11 (rs202115233 G / T), 13 (rs774555874 C / T), 14 (rs781291243 A / G) and 17 (rs1411897695 G / T) (Pruitt et al., 2007). The template for reverse primer sequence also had several SNPs at the second nucleotides (rs1195109615), 9 (rs372568141 A / T), 10 (rs1245601180 A / G), and 15 (rs1215936409 C / T) (Pruitt et al., 2007).

The polymorphism found in the primer sequences affect the DNA strand that was complementary to the primer sequence. This could reduce the stability of the primer annealing and reduce target specificity. The mismatch between the primer sequence and the target DNA could affect the hybridization of the primer sequence to the template and decrease the melting temperature (Tm) by approximately 5 to 18ºC, depending on the position of the mismatch (Prediger, 2008). Primer specificity generally depend on 3’ end sequences. If there was only one mismatch in the middle of the sequence, the PCR amplification product was not too affected. However, if the attachment mismatch was at the 3’ end, especially the incompatibility between A: A, A: G, G: A and C: C, the amplicons formed are greatly decreased, resulting in poor PCR products. The annealing of primers were very good if there were at least 8 nucleotides at the 3’ end that were complementary to the target sequence (Yang et al., 2006). The sequencing quality using forward primer in GP4 was 0%.

**BLAST analysis**

The BLAST analysis results (Table 2) showed that the GP1 sample was a Titin
protein coding DNA sequence (TTN) with an identity level of 99% (forward) and 97% (reverse). GP2, GP3, GP5 and GP6 DNA samples were sequences of amplification targets, namely DNA sequences encoding TNF-α proteins. This was validated with the results of similarity (identity) with samples in the database with accession numbers KF934401.1 and AY274897.1 at 97% and 100%. The GP4 DNA sequence had nothing in common with any sequence, based on BLAST analysis. This was because the quality of GP4 DNA sequencing was so low that it could not be compared to the sequence in the database.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Sample Type</th>
<th>BLAST Analysis Results</th>
<th>Gene</th>
<th>Identity (%)</th>
<th>Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>GP1_F</td>
<td>Active Smoker</td>
<td>Titin</td>
<td>99</td>
<td>NG_011</td>
<td>618.3</td>
</tr>
<tr>
<td>GP1_R</td>
<td>Active Smoker</td>
<td>Titin</td>
<td>97</td>
<td>NG_011</td>
<td>618.3</td>
</tr>
<tr>
<td>GP2_F</td>
<td>Active Smoker</td>
<td>TNF-α</td>
<td>100</td>
<td>KF93444</td>
<td>01.1</td>
</tr>
<tr>
<td>GP2_R</td>
<td>Active Smoker</td>
<td>TNF-α</td>
<td>97</td>
<td>KF93444</td>
<td>01.1</td>
</tr>
<tr>
<td>GP3_F</td>
<td>Passive Smoker</td>
<td>TNF-α</td>
<td>100</td>
<td>KF93444</td>
<td>01.1</td>
</tr>
<tr>
<td>GP3_R</td>
<td>Passive Smoker</td>
<td>TNF-α</td>
<td>100</td>
<td>KF93444</td>
<td>01.1</td>
</tr>
<tr>
<td>GP4_F</td>
<td>Passive Smoker</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GP4_R</td>
<td>Passive Smoker</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GP5_F</td>
<td>Non-Smoker</td>
<td>TNF-α</td>
<td>100</td>
<td>AY2748</td>
<td>97.1</td>
</tr>
<tr>
<td>GP5_R</td>
<td>Non-Smoker</td>
<td>TNF-α</td>
<td>100</td>
<td>AY2748</td>
<td>97.1</td>
</tr>
<tr>
<td>GP6_F</td>
<td>Non-Smoker</td>
<td>TNF-α</td>
<td>100</td>
<td>KF93444</td>
<td>01.1</td>
</tr>
<tr>
<td>GP6_R</td>
<td>Non-Smoker</td>
<td>TNF-α</td>
<td>100</td>
<td>KF93444</td>
<td>01.1</td>
</tr>
</tbody>
</table>

The incompatibility of GP1 DNA samples with the TNF-α sequence was not caused by the primer sequence incompatibility because it was designed for TNF-α promoter region sequences. The analysis results showed identical GP1 samples of 99% and 97% with sequences. This was validated by analysing the forward and reverse primer sequences using the BLAST analysis. The results of the analysis did not show any similarity between the primer sequence and the TTN sequence. The amplification and sequencing errors could be caused by the influence of cigarette smoke which resulted in DNA adduct (Du et al., 2018) that interfered with and affected the quality of amplification and sequencing. Compounds in cigarette smoke generally induced the appearance of DNA adducts were nitrosamines 4- (methylNitosamino) -1- (3-pyrimidyl) -1-butane (NNK) and N'-nitrosonornicotine (NNN) (Carlson et al., 2016). Several types of DNA adducts could cause miscoding during replication. One example of miscoding due to DNA adduct was the error of reading the polymerase DNA due to the presence of O6-methylguanine due to NNK induction. This resulted in thymidine insertion at the time of replication because deoxyguanine was read as deoxyadenosine (Hecht, 2002). The existence of a miscoding mechanism due to the DNA adduct implied that the primer used could bind with and amplified the Titin coding sequence due to an error in reading the primer attachment sequence and an error in the amplification stage by the polymerase protein.

The query sequences included GP2_F, GP3_F, GP3_R, GP5_F, GP5_R, GP6_F, and GP6_R and the subject sequences were AY274897.1 and KF934401.1. The GP1_F and GP1_R samples were not used in alignment because it was not a TNF-α sequence. Even though GP2_R sequence was positive a TNF-α coding sequence, however the fragment length was only 473 bp so it did not cover the region of -308.

**Polymorphism detection**

One of the important polymorphisms in TNF-α gene expression regulation was located in -308 (Zhou et al., 2011). Based on Multalin alignment, there was no polymorphism at the position of -308 TNF-α gene promoter in all sample sequences (active and passive smokers, non-smokers and subject samples). The nitrogen base in the nucleotide found in this position was guanine, although there are variations in polymorphism in the form of adenine (Chen et al., 2018). This does not show that although position -308 had polymorphism in other populations, there was no influence of cigarette smoke which causes nucleotide
Pollo, Rumende dan Tallei: Identification of -308 TNF-α

Figure 1. Allele in -308 of Promoter TNF-α

changes in the sample, especially at position -308 (Fig. 1).

The number of alleles and genotypes at -308 could be compared with other populations. GG genotype was more than AA and AG in the Brazilian population (Borges et al., 2018). In the Chinese population, the G allele was more common than the A allele, and the genotypes found in the population more are GG and the least AA genotypes were found (Chen et al., 2018). Different results were shown in the North Indian population, where the A allele was more found (Kumari et al., 2018).

Nucleotide at position -308 was very important in the regulation of transcription. Because of its location between DNA fragments -316 and -161, it could form complexes with nuclear proteins on TEP macrophages (Collart et al., 1990). The four sample sequences used in this research had G alleles which are common alleles in the population and had lower TNF-α regulation than A allele (Yi et al., 2018). If TNF-α was present in a small size, then this protein could act locally as an inflammatory mediator in autocrine and paracrine leukocytes and also in endothelial cells, determining the expression of membrane receptors for leukocyte migration, acting as an angiogenic factor such as fibroblast growth factor and could improve apoptosis in several cell types (Cereda et al., 2012).

The nucleotide sequences of GP2_F, GP3_F, GP3_R, GP5_F, GP5_R, GP6_F, and GP6_R used in this polymorphism analysis using DnaSP are sequences in area -308, which located at positions 31,574,961 to 31,575,338 (377 bp). The haplotype diversity of the seven sequences was 0.476 with a standard deviation 0.171. The analyzed polymorphism was at position -238, because there was no polymorphism at position -308G so there was no difference between the seven samples used.

The G allele at -308 TNF-α promoter found in samples, showed a negative association with COPD. It was because the G allele regulated TNF-α gene expression at a normal level, meanwhile, the A allele regulated the gene expression at a higher level (Yi et al., 2018). The up-regulation of TNF-α could mediate the pathophysiological changes associated to COPD (Mukhopadhyay et al., 2006).

CONCLUSION

Single nucleotide polymorphisms at position -308 from TNF-α gene promoter between active and passive smokers were not found. This was because by that position, the nucleotides are both in the form of guanosine monophosphate and there was no mutation due to the influence of exposure to cigarette smoke.

REFERENCES


Kumari, R., S. Kumar, M. Ahmad, R. Singh, S. Kant, A. Pradhan, S. Chandra, S. Kumar. 2018. Promoter Variants of TNF-α rs1800629 and IL-10 rs1800871 are independently Associated with the Susceptibility of Coronary Artery Disease in North Indian. *Cytokine* 110: 131-136.


