DOI: https://doi.org/10.54393/pbmj.v8i6.1236

PAKISTAN BIOMEDICAL JOURNAL
https://www.pakistanbmj.com/journal/index.php/pbmj/index
ISSN (E): 2709-2798, (P): 2709-278X
Volume 8, Issue 06 (June 2025)

Original Article

OPEN ACCESS

Association Analysis of *CYP2A6* Gene Variant (rs1801272A>T) with Nicotine Metabolism and Smoking Tendency among Pakistani Youth

Iqra Yasmin[1], Haider Ali[1], Muhammad Rafeh[1], Muhammad Sikandar[1], Abdul Kashif[1], Muhammad Salahuldin[1], Ammad Shafeeq[1] and Rashid Saif[1*]

[1]Department of Biotechnology, Qarshi University, Lahore, Pakistan

ARTICLE INFO

Keywords:
Smoking Genetics, Cytochrome Gene, *CYP2A6* Variants, c.479T>A Locus

How to Cite:
Yasmin, I., Ali, H., Rafeh, M., Sikandar, M., Kashif, A., Salahuddin, M., Shafeeq, A., & Saif, R. (2025). Association Analysis of CYP2A6 Gene Variant (rs1801272A>T) with Nicotine Metabolism and Smoking Tendency among Pakistani Youth: CYP2A6 Gene Variant (rs1801272A>T) with Nicotine Metabolism and Smoking Tendency.Pakistan BioMedical Journal, 8(6),10-15. https://doi.org/10.54393/ pbmj.v8i6.1236

***Corresponding Author:**
Rashid Saif
Department of Biotechnology, Qarshi University, Lahore, Pakistan
rashid.saif37@gmail.com

Received Date: 22nd April, 2025
Revised Date: 2nd June, 2025
Acceptance Date: 10th June, 2025
Published Date: 30th June, 2025

ABSTRACT

Cytochrome P450 2A6 *(CYP2A6)* is a key enzyme in nicotine metabolism, with its genetic variants playing a role in smoking behavior. Particularly, g.40848628A>T is significantly associated with nicotine metabolism and smoking tendency in different populations. **Objectives:** To examine the genetic diversity of this locus and association analysis within smokers and non-smokers cohorts among Pakistani youth. **Methods:** The allele-specific ARMS PCR genotyping technique was applied to examine a total of 100 samples as a case-control study of n=50 from each cohort. **Results:** From the sampled individuals, 92% were found to be homozygous wild-type (AA), 7% were heterozygous (AT), and 1% were homozygous mutant (TT). PLINK software was used for the Chi-square test yielded, χ^2 (1, n=100) =2.91, p=0.088, suggesting a non-significant trend towards association, where alternative allele frequencies were calculated as 0.07 and 0.02 in cases and control cohorts, respectively. Similarly, Hardy-Weinberg Equilibrium (HWE) p=0.1714 indicates genotype frequencies did not significantly deviate from HW expectations and no error or selection in the overall samples. The carriers of the alternative allele have 3.688 times higher odds of being affected by the condition compared to non-carriers with the reference allele. **Conclusions:** It was concluded that future studies with a larger sample size may help to clarify the population structure of the subject locus. Genome-wide association studies using next-generation sequencing may also aid in predicting nicotine metabolism and resistance to smoking cessation in the Pakistani population.

INTRODUCTION

Cigarette smoking, a prevalent source of nicotine, poses significant health risks across all age groups, particularly with a concerning surge in popularity among Pakistani adolescents. Nicotine, a proven addictive substance, exhibits varying degrees of dependence in individuals due to the metabolism by the Cytochrome P450 family 2 subfamilies A member 6 *(CYP2A6)* gene product [1, 2]. This primary protein transforms nicotine into its byproduct, cotinine and other metabolites. Divergent levels of *CYP2A6* protein influence nicotine metabolism, impacting an individual's tendency towards smoking. Wild-type *CYP2A6* proteins with normal activity facilitate the standard breakdown of nicotine, mitigating its effects on the brain and discouraging the development of smoking dependence. Conversely, mutated *CYP2A6* proteins with impaired activity hinder nicotine metabolism, allowing its retention in the body, leading to prolonged brain effects with relatively lower cigarette consumption [3, 4]. Despite the well-known health risks associated with smoking, it remains a substantial global health burden. Adult smoking prevalence is 32.6% and 6.5% in men and women, respectively, contributing to ~7.7 million annual deaths

DOI: https://doi.org/10.54393/pbmj.v8i6.1236

worldwide among its total smokers of 1.14 billion [5]. Alarmingly, low and middle-income countries, constituting 80% of total smokers globally, face an increased prevalence of tobacco use [6]. In Pakistan, 13.4% age-standardized prevalence of tobacco use is reported in urban and rural areas with its alarming rates of 16.3% & 11.7% respectively [7]. This escalating trend of tobacco use in Pakistan has prompted the scientific community to delve into its genetic aspects involved in smoking tendency, levels of cigarette consumption, depth of inhalation and smoking cessation ability of the individuals [7-9].The *CYP2A6* gene exhibits numerous variants associated with nicotine metabolism across diverse populations [10].The subject genetic variant located on Chr.19 NC_000019.10 at 19q13.2 locus (NG_008377.1: g.6820T>A, and (ATG start) position 1799T>A).Chromosomal positioning is g.40848628A>T in the (GCF_000001405.40) genome assembly, current variant corresponds to c.479T>A (r.500) in the NM_000762.6 transcript, affecting the protein's p. L160H position situated on exon 3 of total 6907-nucleotide position of *CYP2A6* gene.This variant impacts the protein's function, reflected in its encoded protein ID NP_000753.3 of 494 amino acids. Notably, rs1801272A>T, associated with poor nicotine metabolism, has a global AAF (T=0.0092) as per the 1000Genome database [11].The (A) allele signifies normal *CYP2A6* protein activity, while the (T) allele may alter activity, influencing nicotine metabolism [12].The subject variant was genotyped through allele-specific ARMS-PCR, followed by assessments for (HWE), (χ2), and (OR) statistics.Subsequently, the derived (AAF) will be determined in both cases and controls.

This study aims to find association and correlation insights between the subject variant and nicotine metabolism that may influence smoking tendency among Pakistani youth.

METHODS

Sample inclusion-exclusion criteria and DNA extraction: A comprehensive genotyping study of 100 individuals was carried out, encompassing specimens from both adolescent smokers (n=50) and non-smokers controls (n=50) based on the Rao statistics of sample size calculation. The study was conducted from May 2024 onward. smokers and non-smokers individuals aged 18–30 were included with known smoking history. Sample size was calculated using Rao's statistical formula, based on 80% power, 0.05 alpha, and expected odds ratios derived from previous literature. The primary objective is to investigate the potential association and correlation between the CYP2A6 locus with nicotine metabolism/dependency among Pakistani youth. Blood samples from human male adolescents aged 18–30 years were collected from Diagnostic Zone Lab, Lahore (Ref: DZL#004/25) and Shah Medical Center, Swat, using K_3-EDTA vacutainers and stored at 4°C until subsequent analysis. Through proper channel ethical guidelines were followed for sample collection. Data of the article is available within the manuscript and supplementary file(s). Article medRxiv preprint is also available https://doi.org/10.1101/2025.04.16.25325925. Genomic DNA extraction was performed using column-based kit method (www.favorgen.com), adhering to the manufacturer's instructions for precise and consistent DNA extraction. Primer designing: The allele-specific ARMS primers were designed through specialized software named OligoCalc and validated by NetPrimer. The ARMS-PCR protocol was employed to amplify both wild and mutant-type variants from each sample. Specifically, one ARMS primer was designed for each variant, strategically incorporating a mismatch at the 4nt position from the 3' end of the sequence. Additionally, to uphold PCR accuracy, two internal control (IC) primers were also employed in the experimental setup during optimization. This comprehensive primer design and PCR strategy aim to enhance the precision and reliability of the genetic variant analysis [13]. This comprehensive primer design was seen (Table 1).

Table 1: Primers Sequence Attributes

ARMS/ IC	Sequence (5'-3')	Tm (°C)	Length (bp)	Product Size (bp)
Reverse Common	GCGTGGTATTCAGCAACG	56.22	24	150
Forward Normal	CGCCAGTGCCCTGGA	55.12	26	
Forward Mutant	CGCCAGTGCCCTGGT	55.26	26	
Forward (IC)	TAACCCACAGCCTCCTACAC	60.50	20	618
Reverse (IC)	TCAGCATCCTCCTCTGGAC	59.50	19	

PCR amplification, each sample, encompassing both wild-type and mutant-type variants were undergoing amplification using a thermocycler.Two distinct PCR reactions were executed for each sample, with each reaction involving the reverse common primer paired separately with forward ARMS primers designed for the wild and mutant-type allele. At the same time, the reverse internal control (IC) and forward internal control (IC) primers were also used to amplify the internal control regions in the process of optimization. The final volume of 14µL, which contained 1µL of 50ng/L genomic DNA as well as the 10mM of each primer, 0.015 IU/µL of Taq polymerase, 2.5 mM each MgCl2 and dNTPs, and 1x Taq buffer with PCR-grade water DEPC treated water was considered as a reaction mixture.To begin the touch-down PCR protocol, an initial denaturation at 94°C was carried out, and 35 cycles were to be repeated with a denaturation of 94°C 30 s, annealing was to begin at 58.4°C with a reduction of 0.5°C/cycle repeated 10 times and the remaining 25 cycles were to be conducted at 54°C 30 s, extension of 72°C 30 s and a final extension of 72°C 07 min [14]. Thereafter, the reaction mixture was transferred to 4°C. This elaborate

DOI: https://doi.org/10.54393/pbmj.v8i6.1236

PCR process seeks to provide a specific and sound amplification of target genomic make-up of DNA (Figure 1).

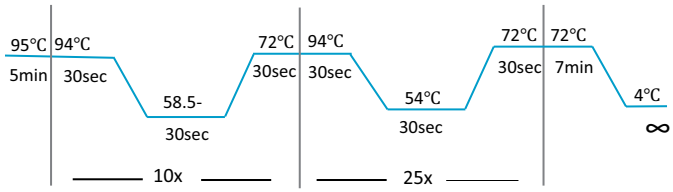

Figure 1: Thermal Cyclic Conditions of ARMS-PCR

The PLINK data analysis toolset was used in computing both observed and expected genotyping frequencies, incorporating considerations for HWE through the application of the equation: p2 + 2pq + q2 = 1. The analysis was extended to Chi-square testing, employing the formula: $\chi 2 = \Sigma(O-E)2/E$, to ascertain the association between the subject variant rs1801272 with nicotine metabolism/dependency and smoking behavior within the sample set. Additionally, p-values and odds-ratios (OR) were undertaken followed by alternative allele frequencies for further enriching the statistical insights using following PLINK commands on command line interface e.g., —assoc for association and correlation tests, —logistic for logistic regression, — hardy for Hardy Weinberg equilibrium, and —model for Dominant (DOM), recessive (REC) models prediction.

RESULTS

CYP2A6 gene variant (rs1801272A>T) showed the variability in Pakistani sampled population, which is in accordance to the previous studies and other world populations (Figure 2).

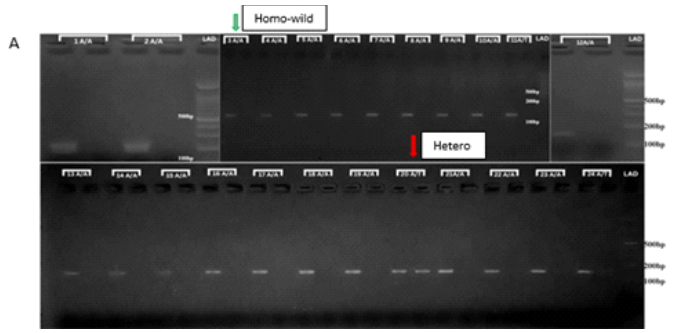

Figure 2: Gel Picture of ARMS-PCR Amplification of Targeted Variant within Cases A. Green and Red Arrows Shows Homozygous-Wild Type and Heterozygous Samples

The study depicts Chr.19 locus g.40848628 is variable and under-selection (Figure 3).

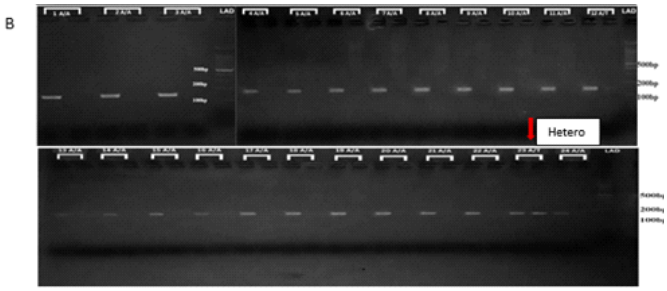

Figure 3: Gel Picture of ARMS-PCR Amplification of Targeted Variant within Controls (B), Red Arrow Shows Heterozygous Samples

A total of 100 samples were genotyped, (smokers=50) and (non-smokers=50). After experimental and statistical analyses, it was concluded that within smoker's cohort, there are 05 heterozygous (AT), 44 homozygous-wild (AA) and 01 sample was observed as homozygous-mutant. Similarly, within non-smokers, 02 and 48 individuals are heterozygous, homozygous-wild respectively and none of the sample was observed as homozygous-mutant. Remaining samples gel pictures are provided in Supplementary Figure 1. The overall genotypic frequencies in our sampled population are 0.92 as homozygous-wild, 0.07 reflected as the heterozygous and 0.01 as the homozygous-mutant. Thereafter, the analysis *(HWE)* of Chi-square was also taken out to check whether our sampled population is in his principle or not and the following results of 42 (1, N=100), 0.1714, which manifests that our sampled population is in accordance with the *HWE* equilibrium as the p-value is above the set threshold confidence interval of 0.05 so accepting our null-hypothesis of observing *HWE*, means no selection and population is randomly bred. In addition, there was also an alternative allele frequency observed as 0.07and 0.02 in our cases and control cohorts with a 2.90, $\chi 2$ statistics value and p=0.088 that indicates insignificant relationship of screened variant with nicotine metabolism and smoking tendency in youth of Pakistan. Similarly, an odds-ratio (OR) of 3.688 was also shown indicating relative risk of subject phenotype is higher and indicating that the prevalence of the odds/mutant variant is about 4 rid higher in cases compared to controls (Table 2).

Table 2: Plink Association of *CYP2A6* Gene Variant (rs1801272A>T) with Nicotine Metabolism and Smoking Tendency among Pakistani Youth

Sample	Chr.Pos.	cDNA variant NM_000762.6	Protein variant NP_000753.3	Genotypic Frequencies			Alternative Allele Frequencies		p-value (OR)
				Homo-wild (AA)%	Hetero- (AT)%	Homo-mutant (TT)%	Case	Contr.	
100	19:40848628	c.479A>Tr.500	p. L160H	92	7	1	0.07	0.02	0.088 (3.68)

Subject cDNA variant c.479A>T alters the lucine amino acid (aa) to histidine at the 160th position, this aa is responsible for slower nicotine metabolism, which makes an individual more prone to quit smoking easily as compared to those with normal *CYP2A6* activity. Further, functional genetics studies are still needed to confirm and validate this postulated hypothesis. Another statistical test named Cochran-Armitage Trend test was also conducted to evaluate the association between subject variant and our selected phenotype, assessing whether the frequency of

DOI: https://doi.org/10.54393/pbmj.v8i6.1236

minor allele differs between smokers and non-smokers individuals The test yielded a chi-square value of2.453 (df=1) and a p-value=0.1173, indicating no significant association, which suggests that the genotype distribution at this variant does not show a meaningful trend between allele dosage and disease status in our sampled population. Similarly, dominant (DOM) and recessive (REC) models testing were also used to explore the potential genetic association of our subject variant with the nicotine metabolism. However, both models yielded NA values for chi-square and p-values, indicating lack of variation in genotype distribution to perform statistical testing. This suggests that neither a dominant nor a recessive model pattern could be established for this variant in the given dataset, because only (1+5) = 6 vs. 44 genotypes were observed in affected and (0+2) = 2 vs. 48 in unaffected population respectively. The logistic regression analysis was also applied under the additive genetic settings which effectively models binary traits while adjusting covariates, suggesting that for each additional copy of the affected allele, the odds of being affected increase by a factor of 3.16. The additive model assumes a linear effect of each additional risk allele on subject phenotype susceptibility. However, the p-value=0.1455 is not significant, indicating no strong association due to limited sample size, low allele frequency or random variation.

DISCUSSION

The current study examined the association between the *CYP2A6* gene variant rs1801272A>T with nicotine metabolism and smoking behavior among Pakistani youth. The statistics revealed no significant association with nicotine metabolism and smoking tendency among our sampled population. Current results showed overall l92% of the sampled population was homozygous-wild (AA), 7% was heterozygous (AT), 1% was homozygous-mutant (TT) with p-value=0.088 and OR=3.68. As individuals with the "T" allele tend to metabolize nicotine more slowly, leading to lower cigarette consumption and potentially have reduced risk of nicotine addiction [15, 16]. Prior research around the globe has reported genetic associations on *CYP2A6* polymorphisms in nicotine metabolism.One of the European study indicated that the rs1801272(A) allele frequency is around 5% in CEU and IBS populations, with a reported p-value=0.81 [17]. In the Mexican population, the allele frequency was found to be <1%, leading to its exclusion from statistical analysis due to low prevalence [17].An Egyptian study reported association between this variant and nicotine metabolism, with statistical significance [18].In contrast, Japanese populations exhibit a higher frequency of loss of function of the *CYP2A6* protein, correlating with slower nicotine metabolism and altered smoking tendencies, with a reported p-value=0.34 [19].The current study genotypic distribution showed that among cases, only five individuals were heterozygous (AT), while the majority (44/50) were homozygous wild-type (AA), a one sample was observed as homo-mutant (TT). While in controls, two individuals carried the heterozygous variant, with the rest (48/50) being homozygous wild. This suggests that the rs1801272T variant is rare in the Pakistani population, limiting its influence on smoking behaviour [1]. Nicotine metabolism is primarily mediated by *CYP2A6*, which converts nicotine into cotinine. Genetic variations in *CYP2A6* may influence smoking initiation, dependency and cessation activities.Studies indicate that slow metabolizers, due to reduced function of the *CYP2A6* protein, tend to smoke less and experience lower nicotine dependence [20]. As far as the clinical relevance and pharmacogenomics are concerned, this variant may influence the effectiveness of the nicotine replacement therapy (NRT) for smoking cessation [19, 21]. However, rs1801272A>T alone is not a major determinant of smoking behaviour in Pakistani youth.

CONCLUSIONS

It was concluded that this pilot study suggested the variability of the subject *CYP2A6* gene locus rs1801272A>T and seems not significantly associated with nicotine metabolism and smoking tendency among Pakistani youth with p-value=0.088 and OR=3.68. Further, large sample size single locus genotyping or GWA studies may be conducted to evaluate the smoking genetics in the Pakistani population.

Authors Contribution

Conceptualization: RS
Methodology: IY, HA, MY, MS, AK, MS, AS, RS
Formal analysis: RS
Writing review and editing: IY, HA

All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

Source of Funding

The authors gratefully acknowledge the support of Qarshi University, Lahore, Pakistan, whose funding made this pilot-scale study possible, their commitment to advancing research and innovation is deeply appreciated.

DOI: https://doi.org/10.54393/pbmj.v8i6.1236

REFERENCES

[1] Tomaz PR, Kajita MS, Santos JR, Scholz J, Abe TO, Gaya PV *et al*. Cytochrome P450 2A6 and 2B6 polymorphisms and smoking cessation success in patients treated with varenicline. European journal of Clinical Pharmacology.2019Nov;75:1541-5.doi: 10.1007/s00228-019-02731-z.

[2] Yuan JM, Nelson HH, Carmella SG, Wang R, Kuriger-Laber J, Jin A *et al. CYP2A6* genetic polymorphisms and biomarkers of tobacco smoke constituents in relation to risk of lung cancer in the Singapore Chinese Health Study.Carcinogenesis.2017Apr; 38(4): 411-8. doi: 10.1093/carcin/bgx012.

[3] Verde Z, Santiago C, Rodríguez González-Moro JM, de Lucas Ramos P, López Martín S, Bandrés F *et al.* 'Smoking genes': a genetic association study.PLOS One.2011Oct;6(10):e26668.doi:10.1371/journal.pone .0026668.

[4] Jones SK, Wolf BJ, Froeliger B, Wallace K, Carpenter MJ, Alberg AJ. Nicotine metabolism predicted by *CYP2A6* genotypes in relation to smoking cessation: a systematic review. Nicotine and Tobacco Research. 2022 May; 24(5): 633-42. doi: 10.1093/ntr/ntab175.

[5] Reitsma MB, Kendrick PJ, Ababneh E, Abbafati C, Abbasi-Kangevari M, Abdoli A *et al.* Spatial, temporal, and demographic patterns in prevalence of smoking tobacco use and attributable disease burden in 204 countries and territories, 1990–2019: a systematic analysis from the Global Burden of Disease Study 2019.The Lancet.2021Jun;397(10292):2337-60.

[6] Schoedel KA, Hoffmann EB, Rao Y, Sellers EM, Tyndale RF.Ethnic variation in *CYP2A6* and association of genetically slow nicotine metabolism and smoking in adult Caucasians. Pharmacogenetics and Genomics.2004Sep;14(9):615-26.doi:10.1097 /00008571-200409000-00006.

[7] Siddiqi K, Siddiqui F, Boeckmann M, Islam Z, Khan A, Dobbie F *et al.* Attitudes of smokers towards tobacco control policies: findings from the Studying Tobacco users of Pakistan (STOP) survey.Tobacco Control. 2022Jan;31(1):112-6.doi:10.1136/tobaccocontrol-2020-055995.

[8] Al Koudsi N, O'Loughlin J, Rodriguez D, Audrain-McGovern J, Tyndale RF. The genetic aspects of nicotine metabolism and their impact on adolescent nicotine dependence.Journal of Pediatric Biochemistry.2010 Jun; 1(02): 105-23. doi: 10.1055/s-0036-1586369.

[9] Pehlivan S, Uysal MA, Cagatay T, Nursal AF, Cinar CK, Erkan F *et al. CYP2A6* gene variants may explain smoking status in a Turkish cohort.Psychiatry and Clinical Psychopharmacology.2019Jul;29(3):340-5. doi: 10.1080/24750573.2018.1547177.

[10] Perez-Paramo YX, Watson CJ, Chen G, Thomas CE, Adams-Haduch J, Wang R *et al.* Impact of genetic variants in the nicotine metabolism pathway on nicotine metabolite levels in smokers.Cancer Epidemiology, Biomarkers and Prevention. 2023 Jan; 32(1): 54-65. doi: 10.1158/1055-9965.EPI-22-0868.

[11] Consortium GP, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM. A global reference for human genetic variation.Nature.2015Oct;526(7571):68-74.

[12] Chenoweth MJ, Lerman C, Knight J, Tyndale RF. Influence of *CYP2A6* genetic variation, nicotine dependence severity, and treatment on smoking cessation success. Nicotine and Tobacco Research. 2023 Jun; 25(6): 1207-11. doi: 10.1093/ntr/ntac268.

[13] Amjad G, Saif R, Hussain MG. Genetic Association of CYP1A2 Gene Variant (rs762551) with Caffeine Induced-Hypertension Susceptibility and Cytochrome P450 1A2 Protein Analyses.bioRxiv. 2023 Oct: 2023-10. doi: 10.1101/2023.10.07.561327.

[14] Khalid A, Ashraf H, Asim H, Ayman M, Saif R. Association Analysis of LCORL Genetic Variant (rs657074013) with Wither-Height in Pakistani Goats: Association of LCORL Variant (rs657074013) with Goat Height.Futuristic Biotechnology.2023Sep:46-50. doi: 10.54393/fbt.v3i02.46.

[15] Murphy SE. Nicotine metabolism and smoking: ethnic differences in the role of P450 2A6.Chemical Research in Toxicology.2017Jan;30(1):410-9.doi: 10.1021/acs.chemrestox.6b00387.

[16] Bergen AW, Michel M, Nishita D, Krasnow R, Javitz HS, Conneely KN *et al.* Drug metabolizing enzyme and transporter gene variation, nicotine metabolism, prospective abstinence, and cigarette consumption. PLOS One.2015Jul;10(7):e0126113.doi:10.1371/journal .pone.0126113.

[17] Pérez-Rubio G, López-Flores LA, Ramírez-Venegas A, Noé-Díaz V, García-Gómez L, Ambrocio-Ortiz E *et al.* Genetic polymorphisms in *CYP2A6* are associated with a risk of cigarette smoking and predispose to smoking at younger ages. Gene.2017Sep;628:205-10. doi:10.1016/j.gene.2017.07.051.

[18] Langlois AW, Chenoweth MJ, Twesigomwe D, Scantamburlo G, Whirl-Carrillo M, Sangkuhl K *et al.* PharmVar GeneFocus: *CYP2A6*.Clinical Pharmacology & Therapeutics.2024Oct;116(4):948-62.doi: 10.1002/cpt.3387.

[19] Ezzeldin N, El-Lebedy D, Darwish A, El Bastawisy A, Abd Elaziz SH, Hassan MM *et al.* Association of genetic polymorphisms *CYP2A6** 2 rs1801272 and CYP2A6* 9 rs28399433 with tobacco-induced lung Cancer: case-control study in an Egyptian population. BioMed Central Cancer. 2018 Dec; 18: 1-9. doi: 10.1186/s12885-018-4342-5.

DOI: https://doi.org/10.54393/pbmj.v8i6.1236

[20] Giratallah H, Chenoweth MJ, Pouget JG, El-Boraie A, Alsaafin A, Lerman C *et al. CYP2A6* associates with respiratory disease risk and younger age of diagnosis: a phenome-wide association Mendelian Randomization study. Human Molecular Genetics. 2024 Jan; 33(2): 198-210. doi: 10.1093/hmg/ddad172.

[21] Perez-Paramo YX and Lazarus P. Pharmacogenetics factors influencing smoking cessation success; the importance of nicotine metabolism. Expert Opinion on Drug Metabolism and Toxicology. 2021Mar;17(3): 333-49. doi: 10.1080/17425255.2021.1863948.