Insilico Identification of Association between Nicotine Dependence and Pathogenic Missense Mutations of Catechol-O-Methyltransferase Gene

Dr. P Rahmath Meeral[1], Dr. Leelavathi[2*], Dr. Meignana Arumugham Indiran[3], J Vijayashree Priyadharsini[4]

[1] *Department of Public Health Dentistry, Saveetha Dental College and Hospital, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India. Email: prarhmath31@gmail.com, https://orcid.org/0000-0003-1621-5877*
[2] *Department of Public Health Dentistry, Saveetha Dental College and Hospital, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India. Email: leelavathi.sdc@saveetha.com, https://orcid.org/0000-0002-2079-0607*
[3] *Department of Public Health Dentistry, Saveetha Dental College and Hospital, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India. Email: dr.mei.sdc@saveetha.com, https://orcid.org/0000-0002-7519-3877*
[4] *Clinical Genetics Lab Cellular and Molecular Research division, Saveetha Dental College and Hospital, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India. Email: vijayashreej.sdc@saveetha.com, https://orcid.org/0000-0001-7884-5466*

KEYWORDS
Nicotine dependence, tobacco smoking, Pharmacogenomics

ABSTRACT
Introduction: Oral cancers are known to be strongly associated with the usage of tobacco, which is majorly composed of nicotine. Genetic studies evidenced the nicotine addiction as hereditary.
AIM: The aim of the present study is to analyse any genetically functional missense mutations in Catechol-O-methyl-transferase (COMT) gene and derive an association with nicotine dependence that might influence tobacco habit initiation, addiction, and cessation.
Methodology: The study involved an insilico analysis about the association between COMT gene and nicotine dependence. This was done with the help of different databases available in website. Initially, database called Ensembl (https://asia.ensembl.org) was used to pool the genetic data on missense mutations of the human COMT gene. The collected data was then fed into computational tools such as SIFT, PolyPhen, PROVEAN, I-Mutant, MutPred to discover the pathogenic mutations if exists.
Results: Among 78 missense mutations collected, 74 were identified as damaging variants in which 8 variants were found to cause increased protein stability and 66 were found to cause decreased protein stability. Among those 66 missense variants causing decreased protein stability, 25 variants were identified as highly pathogenic while 19 variants were identified as pathogenic.
Conclusion: The in-silico analysis discovered the existence of highly pathogenic mutations in COMT gene which might have a strong association with nicotine dependence that can influence the tobacco habit initiation, addiction, and cessation.

1. Introduction

Globally the premature death due to the tobacco related disease is being increasing every year [1,2]. If these trends continue, then it might to leads to the increase in the annual deaths attributable to tobacco to 10 million by the year 2025 [3]. Tobacco Smoking is influenced by multiple factors including environmental, economic, and genetic factors [4]. A recent study of adult twins showed that a high school education decreased the likelihood of smoking habit while parental smoking and monogynous co-twin who ever smoked increased the likelihood up to 10-fold [5]. This is because the genetic makeup of an individual varies comprising various receptor subtypes, different enzymes for nicotine metabolism that can establish disorders in functions of neurotransmitter field, directly influencing the development of early routine smoking behavior, its intensity, and ability to attain smoking cessation. Hence genetics plays a vital role which makes an individual a dependent consumer of nicotine/ tobacco and a polygenetic hereditary pattern is supposed, where various factors favor the probability for the development of a nicotine or tobacco products addiction.

Numerous pharmacogenetic analyses have concentrated on genes in the dopamine pathway considering the neuroscience of reward [6,7,8,9,10,11], where the dopamine-stimulating effects of cigarettes are partially attributed to the nicotine in tobacco's stimulation of dopamine release in the nucleus accumbens [12,13]. Strong possibilities for influencing smoking behavior characteristics, such as quitting smoking and treatment response, are polymorphisms in dopamine pathway genes. Due to its presence in dopaminergic brain regions and function as a crucial enzyme in the degradation and inactivation of extra neuronally released dopamine, the enzyme catechol O-methyltransferase (COMT) is relevant in many previous pharmacogenetic studies of nicotine dependence and the effectiveness of therapeutic interventions [14,15]. This enzyme was produced by the COMT

gene in the brain region. Numerous studies have attempted to link COMT polymorphisms to various elements of nicotine addiction, but their findings have been inconsistent. The Met/Val functional polymorphism of the catechol-O-methyltransferase gene is one of the variants that is most frequently examined for connections with a variety of addictive behavior features (COMT), however replication has been unreliable for smoking status. With this alteration, the amino acid valine is substituted out for the amino acid methionine in the enzyme's single amino acid structure. This variant occurs at position 158 (written as Val158Met) in the longer form of the enzyme. It appears at position 108 of the shorter form of the enzyme (written as Val108Met). Researchers often shorten this notation to Val108/158Met [16]. Allelic variations among the COMT gene are therefore potentially risky candidates for investigating interindividual differences since the COMT gene has a significant role in dopaminergic circuits important for drug reward. in vulnerability to nicotine dependence. With the usage of both genome-wide linkage and association analysis techniques, major efforts have been made over the past few years to uncover the genes responsible for ND susceptibility. It is anticipated that the discovery of ND susceptibility genes would make it possible to create and customize effective treatment plans and medications for tobacco users as well as prevention methods for those who are at risk. Since there exists a plenty of studies focusing on the genetic mutation/ single nucleotide polymorphism on Val108/158Met allele, the aim of this Insilco study is to discover the existence of any other harmful pathogenic mutations exists in the COMT gene with the help of computational tools in pharmacogenomic field.

2. Material and Methodology

Data collection

Though there exist many genes, the reason for choosing the COMT gene for the present Insilco study is that, the enzyme catechol *O*-methyltransferase influenced by Due to its presence in dopaminergic brain areas involved in the breakdown and inactivation of extra neuronally produced dopamine, the COMT gene is relevant in pharmacogenetic investigations of nicotine dependence and treatment prognosis. To gain a clearer understanding of the COMT gene's relationship to the initiation, addiction, and cessation of tobacco use, it is crucial to identify pathogenic mutations in this gene. The Ensemble database (https://asia.ensembl.org) was used to gather data on human COMT mis-sense mutations [17]. 74 missense mutations were screened as of March 2022 using SIFT, PolyPhen, and PROVEAN, three separate computational techniques. To further assess the data from the three tools and determine the pathogenicity and stability of protein changes, respectively, we employed I-Mutant and MutPred. Each section of the software description is discussed below.

SIFT analysis

To predict the changes that are tolerated and harmful at each point of the query sequence, the Sorting Intolerant from Tolerant programme uses information from various sequence alignments. A tolerance index of 0.05 indicates that a substitution is intolerant or detrimental, while one with a normalized probability greater than 0.05 is tolerated [18].

PolyPhen analysis

PolyPhen-2 (Polymorphism Phenotyping v2) predicts the probable effect of amino acid substitutions on the stability and functionality of human proteins using structural and comparative evolutionary considerations. It determines the risk that a missense mutation will be deleterious using a combination of functional annotation of single nucleotide polymorphisms (SNPs), mapping of coding SNPs to gene transcripts, and data extraction from the genes [19].

PROVEAN analysis

A software tool called PROVEAN (Protein Variation Effect Analyzer) makes predictions about whether an amino acid substitution or indel will affect the biological function of a protein. PROVEAN is helpful for identifying nonsynonymous or indel variants that are anticipated to be functionally significant when filtering sequence variants [20].

I-MUTANT analysis

I-Mutant v3.0 is a programme that uses support vector machines (SVM) to automatically anticipate changes in protein stability brought on by single point mutations. The algorithm bases its predictions on the protein sequence. Three categories of predictions were made: neutral mutation (0.5 DDG 0.5 kcal/mol), substantial drop (0.5 kcal/mol), and decrease. Based on the difference between the unfolding Gibbs free energy change of mutant

and native protein (kcal/mol), I-Mutant 3.0 predicts a free energy change (DDG) [21].

MutPred analysis

MutPred v2 was developed as a standalone and web programme to classify amino acid alterations in people as pathogenic or benign. The wild-type protein sequence in FASTA format is used to find the replacement sites. The likelihood that the mutation will be harmful is reported [22,23].

ExAC data analysis

Exome Aggregation Consortium (ExAC), a team of scientists seeking to make summary data available to the larger scientific community, is gathering exome sequencing data from many large-scale sequencing programmes. Based on consent, consortium approval, exome data quality, and lack of relatedness to other samples, these sequences were extracted for public release. The detected variants documented in the current study were compared to those of reported variants deposited in the ExAC repository using the ExAC genome data [24].

3. Results

There are many computational tools approaching studies for discovering the complex etiology at micro level [25]. Previous study has used the similar tools used in the current study to find out the missense mutation in AMELX gene and Amelogenesis imperfecta [26]. Also, there exists many genetic studies that explores the association of genes with periodontal disease, cancer [27,28,29]. Thus, this study aims to discover the missense mutation in COMT gene with the computational tools and according to the effects determined by three prediction tools (SIFT, PolyPhen, and PROVEAN), the list of missense variants in the transcript (ENST00000380712.7) of the COMT gene was sorted and tabulated. Four SNPs were discovered among the 78 missense variations that were assessed to be harmful, as anticipated by all three of the computational tools mentioned in the methodology section (Table 1, Fig. 1).

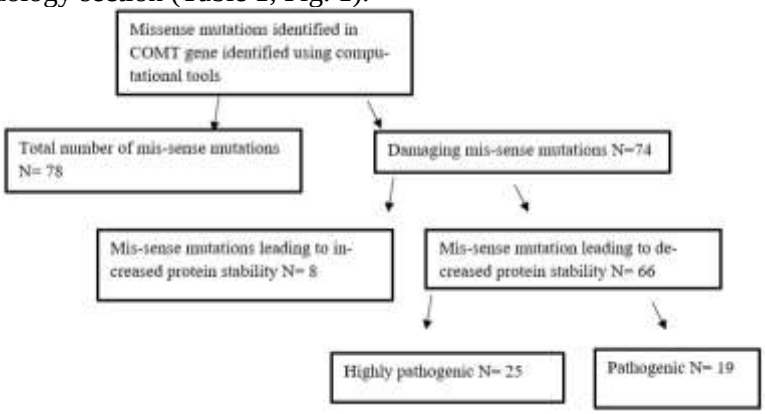

Figure 1: Flowchart of missense pathogenic genes identification.

Based on the standard free energy change at 25 °C and pH 7.0, the structural stability of the protein was evaluated (Table 2), and it was discovered that this change increased or decreased depending on the protein stability free energy (DDG) change. There is enhanced stability when DDG>0 and decreased stability when DDG 0. About 66 variants were found by I-Mutant Suit to have lower stability, whereas 8 variants had higher stability (Table 3).

Table 1: List of missense variants in the transcript (ENST00000361682.11) of COMT gene sorted based on their effects as assessed by three prediction tools (SIFT, PolyPhen and PROVEAN)

c S.no	Variant ID	AA	AA coord	sift_class	SIFT	polyphen_class	polyphen_sort	PRO-pred	PRO-Score
1	rs1243773629	E/K	56	deleterious	0.02	probably damaging	973	Deleterious	-2.596
2	rs1480698560	R/C	58	deleterious	0	probably damaging	979	Deleterious	-6.476
3	rs766469681	R/H	58	deleterious	0	probably damaging	954	Deleterious	-4.212
4	rs1326159112	V/M	63	deleterious	0.03	probably damaging	991	Deleterious	-2.679
5	rs746764639	V/A	63	deleterious	0	probably damaging	992	Deleterious	-3.793
6	rs1464634867	A/T	67	deleterious	0	probably damaging	988	Deleterious	-3.763

c S.no	Variant ID	AA	AA coord	sift_class	SIFT	polyphen_class	polyphen_sort	PRO-pred	PRO-Score
7	rs757163626	A/V	67	deleterious	0.03	probably damaging	994	Deleterious	-3.88
8	rs149909767	G/R	70	deleterious	0.01	probably damaging	1001	Deleterious	-7.859
9	rs149909767	G/W	70	deleterious	0	probably damaging	1001	Deleterious	-7.859
10	rs999200429	V/M	75	deleterious	0	probably damaging	980	Deleterious	-2.81
11	rs999200429	V/L	75	deleterious	0	probably damaging	952	Deleterious	-2.71
12	rs773373201	V/A	75	deleterious	0.04	probably damaging	965	Deleterious	-3.8
13	rs760899191	L/Q	76	deleterious	0	probably damaging	955	Deleterious	-4.84
14	rs753856814	I/T	79	deleterious	0	probably damaging	979	Deleterious	-3.71
15	rs148620887	C/R	83	deleterious	0	probably damaging	998	Deleterious	-8.51
16	rs373611092	M/V	90	deleterious	0.01	possibly damaging	904	Deleterious	-3.42
17	rs1255952971	M/T	90	deleterious	0.01	probably damaging	994	Deleterious	-4.87
18	rs774847933	N/D	91	deleterious	0.01	probably damaging	957	Deleterious	-4.43
19	rs755825980	N/K	91	deleterious	0	probably damaging	993	Deleterious	-5.09
20	rs76452330	D/Y	94	deleterious	0	probably damaging	992	Deleterious	-7
21	rs1601527545	K/E	96	deleterious	0	probably damaging	1001	Deleterious	-3.97
22	rs1173091650	G/D	97	deleterious	0	probably damaging	1001	Deleterious	-6.46
23	rs749349703	L/P	115	deleterious	0	probably damaging	1001	Deleterious	-6.44
24	rs1364785401	Y/S	118	deleterious	0	probably damaging	953	Deleterious	-7.59
25	rs1174983685	G/D	120	deleterious	0	probably damaging	1000	Deleterious	-6.71
26	rs766020740	A/D	123	deleterious	0	probably damaging	977	Deleterious	-4.17
27	rs199710929	R/C	125	deleterious	0	probably damaging	1001	Deleterious	-6.83
28	rs759305975	R/H	125	deleterious	0.01	probably damaging	1001	Deleterious	-4.26
29	rs764846399	R/C	128	deleterious	0	probably damaging	995	Deleterious	-6.58
30	rs752384128	R/H	128	deleterious	0	probably damaging	987	Deleterious	-4.07
31	rs757861911	L/P	130	deleterious	0	probably damaging	1000	Deleterious	-6.07
32	rs1286744830	T/P	138	deleterious	0.03	probably damaging	996	Deleterious	-5.05
33	rs575273220	E/K	140	deleterious	0	probably damaging	1000	Deleterious	-3.8
34	rs145228139	P/R	143	deleterious	0	probably damaging	911	Deleterious	-6.01
35	rs4986871	A/V	146	deleterious	0.01	probably damaging	938	Deleterious	-3.82
36	rs756499908	A/T	156	deleterious	0	probably damaging	992	Deleterious	-3.81
37	rs1445081098	G/A	157	deleterious	0	probably damaging	1001	Deleterious	-5.78
38	rs1410929509	G/E	167	deleterious	0	probably damaging	1000	Deleterious	-6.19
39	rs1410929509	G/A	167	deleterious	0.01	probably damaging	944	Deleterious	-4.47
40	rs768611995	I/N	173	deleterious	0	probably damaging	1001	Deleterious	-6.06
41	rs138056183	P/S	174	deleterious	0.03	probably damaging	940	Deleterious	-6.14
42	rs1395452462	P/R	174	deleterious	0.04	probably damaging	959	Deleterious	-7.01
43	rs1404781926	D/H	183	deleterious	0	probably damaging	953	Deleterious	-2.78
44	rs777226221	L/P	185	deleterious	0	probably damaging	999	Deleterious	-4.85
45	rs1601530510	D/V	186	deleterious	0	probably damaging	1001	Deleterious	-8.38
46	rs765568181	F/L	189	deleterious	0	probably damaging	1001	Deleterious	-5.58
47	rs1181546591	D/N	191	deleterious	0	probably damaging	989	Deleterious	-4.65
48	rs1410147025	H/R	192	deleterious	0	probably damaging	1000	Deleterious	-7.31
49	rs142545956	H/Q	192	deleterious	0	probably damaging	1000	Deleterious	-7.28
50	rs1287503209	D/H	200	deleterious	0	probably damaging	1001	Deleterious	-6.51
51	rs1193284272	L/F	203	deleterious	0	probably damaging	994	Deleterious	-3.46
52	rs1252489480	L/W	204	deleterious	0	probably damaging	1001	Deleterious	-5.3
53	rs374941634	L/F	204	deleterious	0.01	probably damaging	998	Deleterious	-3.42
54	rs567483468	L/P	210	deleterious	0	probably damaging	1000	Deleterious	-5.84
55	rs144463570	R/W	211	deleterious	0	probably damaging	985	Deleterious	-5.36
56	rs747375116	G/R	213	deleterious	0	probably damaging	1001	Deleterious	-7.11
57	rs747375116	G/W	213	deleterious	0	probably damaging	1001	Deleterious	-7.11
58	rs750159038	G/E	213	deleterious	0	probably damaging	1001	Deleterious	-7.11
59	rs376273380	A/D	218	deleterious	0	probably damaging	999	Deleterious	-4.88
60	rs775712233	V/L	221	deleterious	0	probably damaging	929	Deleterious	-2.26
61	rs1318512993	V/E	221	deleterious	0	probably damaging	1001	Deleterious	-4.92
62	rs758187540	G/S	225	deleterious	0	probably damaging	999	Deleterious	-4.68

c S.no	Variant ID	AA	AA coord	sift_class	SIFT	polyphen_class	polyphen _sort	PRO-pred	PRO-Score
63	rs1211653120	P/R	227	deleterious	0	probably damaging	959	Deleterious	-7.49
64	rs773064787	D/Y	228	deleterious	0	probably damaging	964	Deleterious	-4.12
65	rs760480469	L/P	230	deleterious	0	probably damaging	1000	Deleterious	-5.35
66	rs999446749	V/M	233	deleterious	0.02	probably damaging	997	Deleterious	-2.26
67	rs139449932	R/C	234	deleterious	0	probably damaging	999	Deleterious	-6.09
68	rs200150695	R/H	234	deleterious	0	probably damaging	992	Deleterious	-3.97
69	rs1311075732	L/P	248	deleterious	0	probably damaging	994	Deleterious	-4.53
70	rs201407028	G/R	256	deleterious	0	probably damaging	987	Deleterious	-4.83
71	rs754537148	E/K	258	deleterious	0	probably damaging	1000	Deleterious	-2.7
72	rs1327994695	K/T	259	deleterious	0	probably damaging	994	Deleterious	-3.88
73	rs1030230368	G/D	264	deleterious	0	probably damaging	1001	Deleterious	-3.71
74	rs1030230368	G/V	264	deleterious	0	probably damaging	1001	Deleterious	-4.47

Table 2: Protein structural stability assessment based on standard free energy change at 25°C with pH 7.0

S No	Variant ID	AA	AA Position	Stability	DDG (kcal/mol)
1	rs1243773629	E/K	56	decreased	-0.15
2	rs1480698560	R/C	58	decreased	-0.59
3	rs766469681	R/H	58	decreased	-1.03
4	rs1326159112	V/M	63	decreased	-0.13
5	rs746764639	V/A	63	decreased	-0.87
6	rs1464634867	A/T	67	decreased	-1.04
7	rs757163626	A/V	67	increased	0.29*
8	rs149909767	G/R	70	decreased	-0.5
9	rs149909767	G/W	70	decreased	-0.81
10	rs999200429	V/M	75	increased	0.17*
11	rs999200429	V/L	75	increased	0.41*
12	rs773373201	V/A	75	decreased	-0.36
13	rs760899191	L/Q	76	decreased	-1.67
14	rs753856814	I/T	79	decreased	-3.5
15	rs148620887	C/R	83	decreased	-0.5
16	rs373611092	M/V	90	decreased	-1.48
17	rs1255952971	M/T	90	increased	0.08*
18	rs774847933	N/D	91	decreased	-0.56
19	rs755825980	N/K	91	decreased	-0.38
20	rs76452330	D/Y	94	decreased	-0.32
21	rs1601527545	K/E	96	decreased	-0.45
22	rs1173091650	G/D	97	decreased	-1.58
23	rs749349703	L/P	115	decreased	-0.43
24	rs1364785401	Y/S	118	decreased	-2.05
25	rs1174983685	G/D	120	decreased	-1.24
26	rs766020740	A/D	123	decreased	-0.38
27	rs199710929	R/C	125	decreased	-0.78
28	rs759305975	R/H	125	decreased	-0.78
29	rs764846399	R/C	128	decreased	-0.9
30	rs752384128	R/H	128	decreased	-0.9
31	rs757861911	L/P	130	decreased	-0.86
32	rs1286744830	T/P	138	decreased	-1.24

S No	Variant ID	AA	AA Position	Stability	DDG (kcal/mol)
33	rs575273220	E/K	140	decreased	-0.76
34	rs145228139	P/R	143	decreased	-0.45
35	rs4986871	A/V	146	decreased	-0.97
36	rs756499908	A/T	156	decreased	-1.13
37	rs1445081098	G/A	157	decreased	-0.22
38	rs1410929509	G/E	167	increased	0.61*
39	rs1410929509	G/A	167	decreased	-0.33
40	rs768611995	I/N	173	decreased	-1.35
41	rs138056183	P/S	174	decreased	-1.35
42	rs1395452462	P/R	174	decreased	-0.85
43	rs1404781926	D/H	183	decreased	-0.76
44	rs777226221	L/P	185	decreased	-1.96
45	rs1601530510	D/V	186	decreased	-0.31
46	rs765568181	F/L	189	decreased	-3.05
47	rs1181546591	D/N	191	decreased	-1.43
48	rs1410147025	H/R	192	decreased	-0.33
49	rs142545956	H/Q	192	decreased	-1.25
50	rs1287503209	D/H	200	increased	0*
51	rs1193284272	L/F	203	increased	0.21*
52	rs1252489480	L/W	204	increased	0.04*
53	rs374941634	L/F	204	decreased	-0.18
54	rs567483468	L/P	210	decreased	-0.34
55	rs144463570	R/W	211	increased	0.16*
56	rs747375116	G/R	213	decreased	-1.71
57	rs747375116	G/W	213	decreased	-0.13
58	rs750159038	G/E	213	decreased	-0.19
59	rs376273380	A/D	218	decreased	-0.19
60	rs775712233	V/L	221	decreased	-2.45
61	rs1318512993	V/E	221	decreased	-1.58
62	rs758187540	G/S	225	decreased	-1.58
63	rs1211653120	P/R	227	decreased	-1.41
64	rs773064787	D/Y	228	decreased	-1.22
65	rs760480469	L/P	230	decreased	-1.27
66	rs999446749	V/M	233	decreased	-1.41
67	rs139449932	R/C	234	decreased	-0.87
68	rs200150695	R/H	234	decreased	-1.18
69	rs1311075732	L/P	248	decreased	-0.99
70	rs201407028	G/R	256	decreased	-1.56
71	rs754537148	E/K	258	decreased	-1.07
72	rs1327994695	K/T	259	decreased	-1.41
73	rs1030230368	G/D	264	decreased	-1.85
74	rs1030230368	G/V	264	decreased	-1.77

Table 3: Prediction of pathogenicity of proteins with decreased stability as assessed by MutPred tool

S.No	Variant ID	AA	AA coord	MUT-score
1	rs1243773629	E/K	56	0.689
2	rs149909767	G/W	70	0.532
3	rs760899191	L/Q	76	0.699
4	rs753856814	I/T	79	0.617
5	rs148620887	C/R	83	0.723
6	rs373611092	M/V	90	0.717
7	rs774847933	N/D	91	0.527
8	rs755825980	N/K	91	0.659
9	rs76452330	D/Y	94	0.781**
10	rs1173091650	G/D	97	0.902**
11	rs749349703	L/P	115	0.91**
12	rs1286744830	T/P	138	0.898**
13	rs575273220	E/K	140	0.903**
14	rs145228139	P/R	143	0.768**
15	rs4986871	A/V	146	0.596
16	rs756499908	A/T	156	0.735
17	rs1445081098	G/A	157	0.848**
18	rs1410929509	G/A	167	0.78**
19	rs768611995	I/N	173	0.812**
20	rs1404781926	D/H	183	0.56
21	rs1601530510	D/V	186	0.924**
22	rs765568181	F/L	189	0.871**
23	rs1181546591	D/N	191	0.861**
24	rs1410147025	H/R	192	0.842**
25	rs142545956	H/Q	192	0.779**
26	rs374941634	L/F	204	0.502
27	rs567483468	L/P	210	0.933**
28	rs747375116	G/R	213	0.921**
29	rs747375116	G/W	213	0.923**
30	rs750159038	G/E	213	0.928**
31	rs376273380	A/D	218	0.928**
32	rs775712233	V/L	221	0.76**
33	rs1318512993	V/E	221	0.924**
34	rs758187540	G/S	225	0.867**
35	rs1211653120	P/R	227	0.722
36	rs773064787	D/Y	228	0.719
37	rs760480469	L/P	230	0.714
38	rs139449932	R/C	234	0.594
39	rs1311075732	L/P	248	0.895**
40	rs201407028	G/R	256	0.706
41	rs754537148	E/K	258	0.707
42	rs1327994695	K/T	259	0.609
43	rs1030230368	G/D	264	0.883**
44	rs1030230368	G/V	264	0.887**

** Increased pathogenicity as predicted by MutPred [Score > 0.5 – pathogenic]

4. Discussion

Nicotine was been shown by many research as the main ingredient of tobacco smoking products, stimulating the dopamine release from neurons of the ventral tegmental area, which is responsible for an action believed to be responsible for its rewarding effect consequences [30,13]. As a result, it is one of the primary psychoactive components in tobacco that fuels the dangerous tobacco use habit [31], which raises morbidity and mortality rates globally [32] Additionally, compared to reliance on any other substance of abuse, nicotine dependence is more common [33].

According to numerous prior research, nicotine dependence is a complicated illness caused by the interactions

of many different genes, each of which has a minor influence. These interactions are expected to impart additive genetic contributions. Twin and adoption studies over a long period of time have shown that the heredity of liability for nicotine dependency is at least 50%. [34].

It has been demonstrated that a functional polymorphism in the catechol-O-methyltransferase enzyme (COMT) gene influences executive cognition and the physiology of the prefrontal cortex in humans. This effect is likely caused by changes in prefrontal dopamine signaling, and it may therefore be the cause of poor prefrontal function [35], which may lead to tobacco addiction. Though there exists a lot of literature and research studies which evidence about the association between individual's genetic makeup and smoking behavioral pattern, many of the studies has focused on the single nucleotide polymorphisms (SNP) on the Val108/158Met allele located on the exon 3 of the COMT gene (rs4680) [36]. Low gene activity caused by the Met (A) allele causes a 3- to 4-fold reduction in COMT enzyme activity. It is hypothesized that because of the reduced inactivation caused by the lower enzyme activity, substantially more dopamine will be produced [37].

Women who were homozygous for the low-activity Met allele were significantly more likely to have given up smoking, according to research by Colilla et al. [11]. Other studies have revealed that the Val allele may be a risk allele for Caucasian Americans [38], Croatian Caucasian men [39], Plains Indian women [40], and Japanese men [41] to start or continue smoking. The COMT Met/Val poly-morphism is substantially related with quitting smoking, according to a study by Omidvar et al. The risk allele known as Met is the one that makes it less likely for both men and women to give up smoking. This current Insilco study was conducted to discover if there exists any other SNP in the alleles of the COMT gene other than the Val allele with the help of various genetic software tools.

This software tools can filter the large number of variants of gene in terms of amino acids substitution, protein stability etc., For the identification of potentially harmful missense mutations, numerous computational methods have been invented which are also known as variant effect predictors. However, because different tools use distinct predictive traits, they frequently disagree with one another. Performance can be enhanced by ensemble approaches, which integrate the findings of numerous individual predictors [42]. Hence in this study we used multiple computational tools such as SIFT, PolyPhen, PROVEAN, MutPred and I Mutant to identify pathogenic missense mutations in the COMT gene. These methods provide feature information as input to machine learning algorithms for phenotypic prediction using sequence information from homologues, structural information, such as accessible surface area, and changes in amino acid characteristics. They are dependent on sets of mutation/phenotype relationship data that already exist [43]. Finally, the output of the present study was that, from the 74 missense mutations, 66 variants were found to have decreased protein stability in which 25 variants were highly pathogenic and 19 were pathogenic.

5. Conclusion

Thus, the present Insilco study had shed a light on variants of COMT genes other than the Val/Met allele that could have a strong association with tobacco habit initiation, addiction, and cessation which on confirmation with future genetic studies, leads to the appropriate strategy to achieve prevention of nicotine dependence and excellent prognosis of nicotine therapy.

Conflicts:

No conflicts of interest.

Funding:

Nil.

Reference

[1] Doll R, Peto R, Wheatley K, Gray R, Sutherland I. Mortality in relation to smoking: 40 years' observations on male British doctors. Bmj. 1994 Oct 8;309(6959):901-11. DOI: 10.1136/bmj.309.6959.901

[2] Neubauer S, Welte R, Beiche A, Koenig HH, Buesch K, Leidl R. Mortality, morbidity and costs attributable to smoking in Germany: update and a 10-year comparison. Tobacco control. 2006 Dec 1;15(6):464-71.3. DOI:10.1136/tc.2006.016030

[3] Zhang J, OU JX, BAI CX. Tobacco smoking in China: prevalence, disease burden, challenges and future strategies. Respirology. 2011 Nov;16(8):1165-72. PMID: 21910781

[4] Benowitz NL. Nicotine addiction. Primary Care: Clinics in Office Practice. 1999 Sep 1;26(3):611-31. DOI: https://doi.org/10.1016/S0095-4543(05)70120-2

[5] Hamilton AS, Lessov-Schlaggar CN, Cockburn MG, Unger JB, Cozen W, Mack TM. Gender differences in determinants of smoking initiation and persistence in California twins. Cancer Epidemiology Biomarkers & Prevention. 2006 Jun 1;15(6):1189-97.

[6] Robinson TE, Berridge KC. The neural basis of drug craving: an incentive-sensitization theory of addiction. Brain Res Brain Res Rev. 1993;18:247–91. DOI: https://doi.org/10.1158/1055-9965.EPI-05-0675

[7] Johnstone EC, Yudkin PL, Hey K, Roberts SJ, Welch SJ, Murphy MF, Griffiths SE, Walton RT. Genetic variation in dopaminergic pathways and short-term effectiveness of the nicotine patch. Pharmacogenetics and Genomics. 2004 Feb 1;14(2):83-90.PMID: 15077009

[8] Lerman C, Jepson C, Wileyto EP, Epstein LH, Rukstalis M, Patterson F, Kaufmann V, Restine S, Hawk L, Niaura R, Berrettini W. Role of functional genetic variation in the dopamine D2 receptor (DRD2) in response to bupropion and nicotine replacement therapy for tobacco dependence: results of two randomized clinical trials. Neuropsychopharmacology. 2006 Jan;31(1):231-42. PMID: 16123753.

[9] Yudkin P, Munafò M, Hey K, Roberts S, Welch S, Johnstone E, Murphy M, Griffiths S, Walton R. Effectiveness of nicotine patches in relation to genotype in women versus men: randomised controlled trial. Bmj. 2004 Apr 22;328(7446):989-90. PMID: 15033882.

[10] Berrettini WH, Wileyto EP, Epstein L, Restine S, Hawk L, Shields P, Niaura R, Lerman C. Catechol-O-methyltransferase (COMT) gene variants predict response to bupropion therapy for tobacco dependence. Biological psychiatry. 2007 Jan 1;61(1):111-8. PMID:16876132.

[11] Colilla S, Lerman C, Shields PG, Jepson C, Rukstalis M, Berlin J, DeMichele A, Bunin G, Strom BL, Rebbeck TR. Association of catechol-O-methyltransferase with smoking cessation in two independent studies of women. Pharmacogenetics and genomics. 2005 Jun;15(6):393.PMID: 15900212.

[12] Balfour DJ. The neurobiology of tobacco dependence: a preclinical perspective on the role of the dopamine projections to the nucleus. Nicotine & Tobacco Research. 2004 Dec 1;6(6):899-912. PMID: 15801566.

[13] Pontieri FE, Tanda G, Orzi F, Chiara GD. Effects of nicotine on the nucleus accumbens and similarity to those of addictive drugs. Nature. 1996 Jul 18;382(6588):255-7. DOI: 10.1038/382255a0

[14] Chen J, Lipska BK, Halim N, Ma QD, Matsumoto M, Melhem S, Kolachana BS, Hyde TM, Herman MM, Apud J, Egan MF. Functional analysis of genetic variation in catechol-O-methyltransferase (COMT): effects on mRNA, protein, and enzyme activity in postmortem human brain. The American Journal of Human Genetics. 2004 Nov 1;75(5):807-21. PMID: 15457404.

[15] Akil M, Kolachana BS, Rothmond DA, Hyde TM, Weinberger DR, Kleinman JE. Catechol-O-methyltransferase genotype and dopamine regulation in the human brain. Journal of Neuroscience. 2003 Mar 15;23(6):2008-13.

[16] https://medlineplus.gov/genetics/gene/comt/.

[17] Sarah EH, McLaren W, Laurent G et al. Ensembl variation resources. Database 2018;2018:119.

[18] Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. Nat Protco 2016;11:1-9.

[19] Adzhubei IA, Schmidt S, Peshkin L et al. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248-249.

[20] Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics 2015; 31:2745-2747.

[21] Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res 2005;33:W306-31.

[22] Li B, Krishnan VG, Mort ME et al. Automated inference of molecular mechanisms of disease from amino acid substitutions. Bioinformatics 2009;25:2744-2750.

[23] Pejaver V, Urresti J, Lugo-Martinez J et al. Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. Nat. commun. 2020;11(1):5918.

[24] Lek M, Karczewski KJ, Minikel EV et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536(7616):285-91.

[25] Jayaseelan VP, Arumugam P. A Computational Data Mining Strategy to Identify the Common Genetic Markers of Temporomandibular Joint Disorders and Osteoarthritis. Global Medical Genetics. 2022 Mar 9;9(02):159-65.

[26] Shivani N, Smiline-Girija AS, Paramasivam A, Vijayashree-Priyadharsini J. Computational approach towards identification of pathogenic missense mutations in AMELX gene and their possible association with amelogenesis imperfecta. Molecular biology research communications. 2020 Jun;9(2):63

[27] Paramasivam A, George R, Priyadharsini JV. Genomic and transcriptomic alterations in m6A regulatory genes are associated with tumorigenesis and poor prognosis in head and neck squamous cell carcinoma. American journal of cancer research. 2021;11(7):3688.

[28] Murthykumar K, Varghese S, Jayaseelan VP. Association of SRXN1 receptor gene polymorphism with susceptibility to periodontitis. Contemporary Clinical Dentistry. 2022 Oct;13(4):363.

[29] Ramasubramanian A, Paramasivam A, Ramani P. FASTK family of genes linked to cancer. Bioinformation. 2022;18(3):206.

[30] Nisell M, Nomikos GG, Svensson TH. Systemic nicotine-induced dopamine release in the rat nucleus accumbens is regulated by nicotinic receptors in the ventral tegmental area. Synapse. 1994;16(1):36-44.

[31] Royal College of Physicians of London 2000 Nicotine addiction in Britain: a report of the Tobacco Advisory Group of the Royal College of Physicians. London, UK: Royal College of Physicians of London.

[32] Murray, C. J. & Lopez, A. D. Alternative projections of mortality and disability by cause 1990–2020: global burden of disease study. Lancet1997; 349:1498–1504.

[33] Anthony, J. C., Warner, L. A. & Kessler, R. C. Comparative epidemiology of dependence on tobacco, alcohol, controlled substances, and inhalants: basic findings from the National Comorbidity Survey. Exp. Clin. Psychopharmacol. 1994; 2: 244–268.

[34] Li MD. The genetics of nicotine dependence. Curr Psychiatry Rep. 2006;8(2):158-64.

[35] Akil M, Kolachana BS, Rothmond DA, Hyde TM, Weinberger DR, Kleinman JE. Catechol-O-methyltransferase genotype and dopamine regulation in the human brain. J.Neurosci. 2003;23(6):2008-13.

[36] Witte AV, Flöel A. Effects of COMT polymorphisms on brain function and behavior in health and disease. Brain Res. Bull. 2012;88(5):418-28.

[37] Johnstone EC, Elliot KM, David SP, Murphy MF, Walton RT, Munafò MR. Association of COMT Val108/158Met genotype with smoking cessation in a nicotine replacement therapy randomized trial. Cancer Epidemiol Biomarkers Pre. 2007;16(6):1065-9.

[38] Shiels MS, Huang HY, Hoffman SC et al. A community-based study of cigarette smoking behavior in relation to variation in three genes involved in dopamine metabolism: Catechol-O-methyltransferase (COMT), dopamine beta-hydroxylase (DBH) and monoamine oxidase-A (MAO-A). Prev. Med. 2008;47(1):116-22.

[39] Nedic G, Nikolac M, Borovecki F, Hajnsek S, Muck-Seler D, Pivac N. Association study of a functional catechol-o-methyltransferase polymorphism and smoking in healthy Caucasian subjects. Neurosci Lett 2010; 473:216–219.

[40] Enoch MA, Waheed JF, Harris CR, Albaugh B, Goldman D. Sex differences in the influence of COMT Val158Met on alcoholism and smoking in plains American Indians. Alcohol Clin Exp Res 2006; 30:399–406.

[41] Tochigi M, Suzuki K, Kato C et al. Association study of monoamine oxidase and catechol-Omethyltransferase genes with smoking behavior. Pharmacogenet Genomics 2007; 17:867–872.

[42] Omidvar M, Stolk L, Uitterlinden AG, Hofman A, Van Duijn CM, Tiemeier H. The effect of catechol-O-methyltransferase Met/Val functional polymorphism on smoking cessation: retrospective and prospective analyses in a cohort study. Pharmacogenomics J. 2009;19(1):45-51.

[43] Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet. 2016;99(4):877–85.