Journal of Molecular Neuroscience (2021) 71:122–136
https://doi.org/10.1007/s12031-020-01633-5

Association of Dopamine Transporter Gene with Heroin Dependence in an Indian Subpopulation from Manipur

Arunkumar Singh Koijam[1] · Aruna Chanu Hijam[1] · Asem Surindro Singh[2] · Preeti Jaiswal[3] · Kanchan Mukhopadhyay[3] · Usha Rajamma[3,4] · Reena Haobam[1]

Received: 20 November 2019 / Accepted: 11 June 2020 / Published online: 18 June 2020

Abstract

Dopamine transporter (DAT) or solute carrier family 6 member 3 (SLC6A3) is a transmembrane protein regulating dopaminergic neurotransmission. It has been implicated in playing important roles in the dopaminergic reward pathways, and thus, *DAT1* is a strong candidate gene for association studies with heroin dependence. A case-control study involving 279 individuals (147 controls and 132 heroin-dependent cases) was conducted. Ten polymorphisms of the *DAT1 (SLC6A3)* gene were analysed for its association with heroin dependence. Following the Hardy-Weinberg equilibrium (HWE) test, genetic association analyses were performed for the study groups. The post hoc statistical power of the study was 0.655 (65.5%). Single-nucleotide polymorphism (SNP) rs246997 was found to be significantly associated with heroin dependence at allelic, genotypic, and haplotypic levels. A significant difference in the distribution of 11R allele and 10R/11R genotype of rs28363170 between heroin-dependent cases and controls was also observed. Nominal significance at degrees of freedom (df) = 5 was also observed for rs28363170. Five bimarker-based haplotype combinations were also found to be associated with heroin dependence. For the first time, 13R allele (7R/13R genotype) and 14R allele (7R/14R genotype) were identified for rs3836790 in the population. The study also reports that the 11R allele and 10R/11R genotype of rs28363170 is associated with protection against heroin dependence. 7R and 6R alleles were also found to be the common alleles of rs3836790 in the study population. The study provides evidence for the association of polymorphisms of *DAT1 (SLC6A3)* with heroin dependence.

Keywords *DAT1/SLC6A3* gene · Polymorphism · Heroin dependence · Genetic association

Introduction

Heroin dependence is a disease characterized by recurring impairment of physiological and psychological conditions, resulting from persistent maladaptive use of the psychoactive substance heroin. It is a complex disease where initiation is primarily environment-dependent, but the progression is dependent upon environmental (Rhee et al. 2003; Robinson and Berridge 2008) and genetic factors (Kobayashi and Schultz 2008). The role of the environmental factors may range widely from the availability of psychoactive substances, prevailing laws in a region, social interaction pattern to the developmental conditions of an individual at home or surrounding. Genetic factors to addiction account for 40–60% of an individual's vulnerability to addiction and can also be influenced by gender, ethnicity, and the developmental stages (National Institute on Drug Abuse 2010).

Dopamine transporter (DAT) is a transmembrane protein responsible for the transportation of dopamine from the synapse into the cytosol. It is a major regulator of dopaminergic neurotransmission (Giros et al. 1996) and thus has been suggested to play an essential role in the dopaminergic reward pathways of the mesocorticolimbic region in substance dependence (Koob and Nestler 1997). Heroin use increases phasic dopamine neurotransmission relative to baseline by increasing dopamine release. The increase in dopamine enhances

✉ Reena Haobam
reena_haobam@rediffmail.com

[1] Department of Biotechnology, Manipur University, Canchipur, Imphal, Manipur 795003, India

[2] Department of Pathology, University of Mississippi Medical Center, 2500 North State Street, Jackson, MS, USA

[3] Manovikas Biomedical Research & Diagnostic Centre, Manovikas Kendra, 482 Madudah, Plot I-24, Sector-J, EM Bypass, Kolkata, West Bengal, India

[4] Centre for Development & Aging Research, Inter University Centre for Biomedical Research & Super Speciality Hospital, MG University Campus, Thalappady, Rubber Board PO, Kottayam, Kerala, India

salience and learning associated with reward and addiction (Sulzer 2011). Impairments in the dopaminergic pathways that regulate the neuronal systems that are associated with self-control, conditioning, stress reactivity, reward sensitivity, and incentive motivation are the major causes of addiction (Volkow et al. 2013). The basolateral amygdala or the nucleus accumbens and the orbital prefrontal cortex (PFC) are richly interconnected, and this neuronal network has been indicated in addiction and conditioned reinforcement. Lesions in this neuronal network have been implicated in impairing the acquisition of cocaine or heroin seeking (Everitt and Robbins 2005). Interactions of the amygdala, hippocampal, and PFC projections in the nucleus accumbens are modulated by mesolimbic dopamine, which in turn can modulate the release of dopamine (Everitt and Robbins 2005). Low dopamine function has been linked to abnormal cravings (Blum et al. 2008).

The human *DAT1* gene (*SLC6A3*) with cytogenetic location 5p15.33 consists of 15 exons and spans approximately 60 kb (Giros et al. 1992; Vandenbergh et al. 1992). The gene exhibits consensus sequences for RNA splicing at each intron-exon junction. Its protein-coding region starts within the exon 2 and ends at the beginning of exon 15 (Bannon et al. 2001). *DAT1* and its variants have been reported to be associated with a responsivity of the reward-related network (Dreher et al. 2009), cocaine-induced paranoia (Gelernter et al. 1994), cocaine addiction (Brewer et al. 2015; Guindalini et al. 2006), alcoholism (Vasconcelos et al. 2015), the severity of alcohol withdrawal (Sander et al. 1997), alcohol withdrawal seizures (Le Strat et al. 2008), smoking (Wetherill et al. 2014), and smoking cessation (Ma et al. 2016) in various populations. This series of evidence suggest that *DAT1* is a strong candidate gene for heroin addiction. Therefore, the present study aims to investigate the association of *DAT1* with heroin abuse in a small homogeneous Indian subpopulation from Manipur, a state in India with a high prevalence of drug dependency. In this study, we have selected eight single-nucleotide polymorphisms (SNPs) and two variable number tandem repeats (VNTRs) of *DAT1* to investigate the association of *DAT1* variants with heroin addiction.

Materials and Methods

Subject Selection

A total of 279 (147 controls and 132 cases) individuals were involved in this study. Participants were recruited with their informed written consent, following the DSM-IV criteria (American Psychiatric Association: Diagnostic and Statiscal Manual of Mental Disorders 1994). The age of heroin-dependent cases ranged from 22 to 57 years. Among the 132 cases, 110 were males, and 22 were females. Out of the total 147 control subjects, aged between 18 and 58 years, 98 were males, and 49 were females. Heroin-dependent subjects were examined and assessed by an expert psychiatrist specialized in substance abuse. All the cases were undergoing oral substitution therapy (OST) with buprenorphine at the time of blood collection. Control participants were selected based on non-substance user criteria provided in DSM-IV. Individuals with other psychiatric diagnoses, such as psychosis, or chronic physical illness such as diabetes or other metabolic disorders were excluded from the study. The study conformed to the Declaration of Helsinki and was approved by the Institutional Human Ethics Committee of Manipur University.

DNA Isolation

Genomic DNA was prepared from WBCs isolated from the peripheral blood of the participants following the salting out procedure (Miller et al. 1988). DNAs were dissolved in TE (Tris EDTA) buffer pH − 8.0 and stored at − 80 °C to be used for downstream genotyping analysis.

Genotyping Analysis

Eight SNPs (rs40184, rs27048, rs37021, rs250683, rs250682, rs427284, rs458609, rs246997) and two VNTR polymorphisms (rs28363170 and rs3836790) located within the 3′ untranslated region (UTR) and intron 8 of the *DAT1* were selected for testing their association with heroin dependence. A diagrammatic representation of *DAT1* is provided in Fig. 1a (NCBI Database n.d.) and the marker details in Table 1. The genomic regions encompassing the selected SNPs and VNTRs for the DNA samples were amplified by PCR (T100™ Thermal Cycler, BIORAD) in a 25 μl reaction mixture containing OneTaq®Quick-Load® 2X Master Mix with standard buffer (New England Biolabs®) and specific oligonucleotide primers (Table 1) purchased from Sigma-Aldrich, India. The specific primers were designed using Primer3 software (Koressaar and Remm 2007). PCR cycling conditions for all the SNPs included initial denaturation at 94 °C for 3 min; 35 cycles of denaturation at 94 °C for 30 s, annealing at X °C (X, 55.2 °C to 60.1 °C depending on the primer) for 30 s; extension at 68 °C for 45 s; and a final extension at 68 °C for 5 min. Cycle conditions for VNTR rs28363170 included initial denaturation at 94 °C for 5 min followed by 35 cycles of denaturation at 94 °C for 20 s, annealing at 73 °C for 30 s; extension at 68 °C for 20 s; and a final extension at 68 °C for 5 min. For VNTR rs3836790, initial denaturation was at 94 °C for 40 s followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 63 °C for 30 s; extension at 68 °C for 40 s; and a final extension at 68 °C for 5 min.

After PCR, the eight SNPs were genotyped by using restriction fragment length polymorphism (RFLP) method, while only PCR was used for genotyping the two VNTRs.

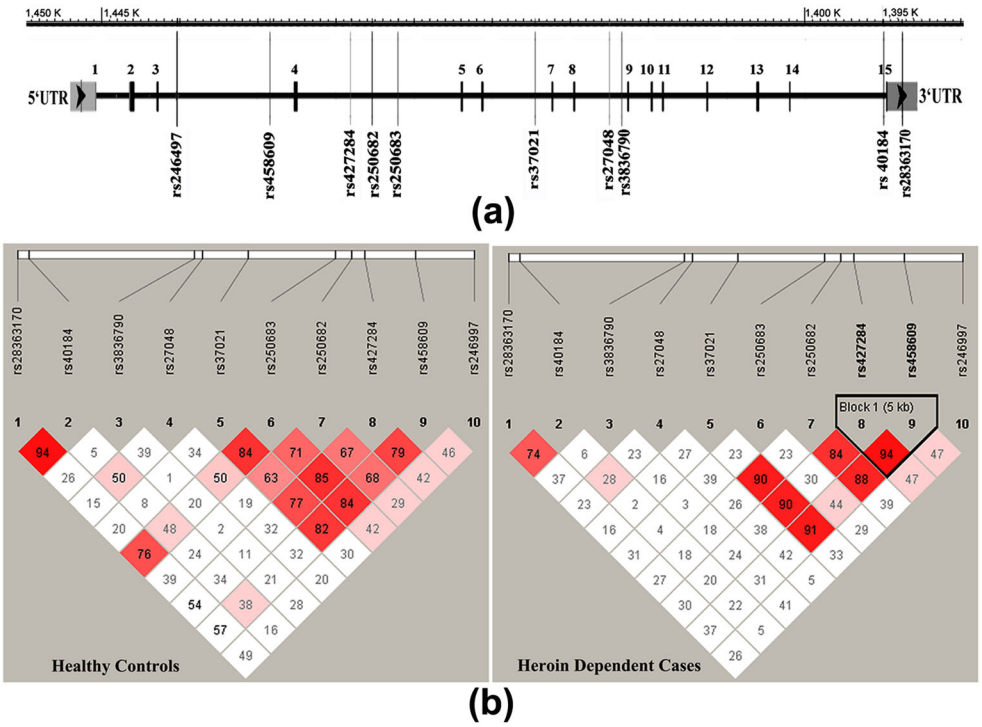

Fig. 1 **a** Diagrammatic representation of the *DAT1* (*SLC6A3*) gene structure. Gene structure was constructed using information gathered from the NCBI gene database. The horizontal black line with vertical bars illustrates the complete gene structure, and the comparative scale in the chromosome is indicated. Dark vertical bars depict exons, and exon numbers are shown above the bar. Locations of the various studied markers are depicted in the gene structure using their respective IDs. Spaces between the exons represent introns. **b** Graphical representation of pairwise linkage disequilibrium (LD) of *DAT1* markers in heroin-dependent cases and controls. Pairwise measures of LD, such as normalized linkage disequilibrium coefficient (D′) and correlation coefficient (r^2), are estimated using Haploview program v4.2. The values in the box represent D′. The intensity of the colour is dependent on the D′/LOD value, and greater intensity of the red colour indicates higher D′ value

Both PCR and RFLP products were analysed by gel electrophoresis. For RFLP, the restriction digestions were performed by the restriction enzymes listed in Table 1, following the manufacturer's protocols (New England Biolabs®).

Statistical Analysis

The statistical power of the present analysis was calculated as post hoc using GAS Power Calculator, assuming that a multiplicative disease model with a disease prevalence of 0.007 (0.7%), the disease allele frequency of 0.35, a significance level of 0.05, and genotype relative risk of 1.5 was 0.655 (65.5%) which is moderate. The Hardy-Weinberg equilibrium (HWE) analysis of cases and controls for each marker was performed using POPGENE version 1.32 (Yeh et al. 1997). Genotypic/allelic frequency calculation and population-based case-control association analyses based on allelic, genotypic, and haplotypic frequencies were performed by using UNPHASED program version 3.1.5 (Dudbridge 2008), and the multiple testing was corrected by 1000 permutations using the same program. The two VNTRs were also analysed based on genotypic and allelic frequencies using SPSS (version 16.0). Odds ratio (OR) and confidence interval (CI) were calculated using OR calculator (http://www.hutchon.net/confidornulhypo.htm). Linkage disequilibrium (LD) between the paired markers of *DAT1* was calculated using HaploView program version 4.2 (Barret et al. 2005). Bimarker-based haplotype association analysis was carried out to identify the unique chromosomal segments that likely harbour disease-specific susceptible haplotypes. For the two VNTRs, the alleles were grouped into two, viz. L (10R and 10+R for rs28363170 and 7R and 7+R for rs3836790) and S (9R and 9-R for rs28363170, and 6R and 6-R for rs3836790), for biallelic representation before the analysis. Bivariate correlation analysis and binary logistic regression analysis were performed using SPSS (version 16.0). Bivariate correlation analysis was performed for different variables, viz. 'years of drug use', 'number of times drug abstained', 'onset', rs28363170 genotypes, and rs246997 genotypes with disease status. Binary logistic regression analysis was performed with heroin use disorder status as a dependent variable and rs28363170-10R/10R, rs28363170-10R/9R, rs28363170-10R/11R, rs246997-TT, rs246997-GT, and rs246997-GG as covariates.

Marker-marker interaction of the *DAT1* gene was carried out using multifactor dimensionality reduction (MDR) 2.0

Table 1 Details of the markers, restriction enzymes, and sequences of the PCR primers used for genotyping analysis of *DAT1* polymorphisms

Markers (SNP/INS ID)	Alleles	MAF			Position	Location in the gene	Enzyme used for genotyping	Oligonucleotide primer sequences	
		Global MAF	Cases	Controls					
rs28363170	3R-13R	0.13	0.11	0.10	1,393,747–1,393,748	3′UTR	–	FP	TGTGGTGTAGGGAACGGCCTGA
								RP	CTTCCTGGAGGTCACGGCTC AAG
rs40184	C/T	0.41 (T)	0.23	0.27	1,394,962	Intron 14	BsaXI	FP	AGTCGCTGGGGTACAATCTG
								RP	ATCCGATGCCTCACTCAAAC
rs3836790	1R-12R	**	0.22	0.26	1,411,740–1,411,741	Intron 8	–	FP	GCATGTGGATGTGTTCTTGC
								RP	GATCAGAGGGCTTTGGTTCA
rs27048	C/T	0.32 (T)	0.19	0.21	1,412,530	Intron 8	PleI	FP	GATGGTGGCTTCGGAGATAA
								RP	CAGAGACCAAACTGCGTTGA
rs37021	A/G	0.38 (A)	0.35	0.36	1,417,239	Intron 6	HaeII	FP	AGCCTCAGGATGAGTTTATGTTG
								RP	CTTCACAACCCCCTGAGTGT
rs250683	T/C	0.42 (C)	0.49 (T)	0.48 (C)	1,426,036	Intron 4	FauI	FP	AGGGACACATCCCACAATTC
								RP	TGGATAGGTCGATAGGTGGA
rs250682	G/C	0.43 (C)	0.46	0.50	1,427,688	Intron 4	BaeI	FP	AGGGACACATCCCACAATTC
								RP	TGGATAGGTCGATAGGTGGA
rs427284	G/C	0.42 (C)	0.48	0.51	1,429,072	Intron 4	BsmAI	FP	CTGAGGTTCCCCTGAAACAG
								RP	GGCATGCCCATTATCTAGGA
rs458609	C/T	0.38 (T)	0.50	0.48	1,434,191	Intron 3	TfiI	FP	GACCCTGGTGGTCTTACGTG
								RP	ATAGCTGCTCATGCCTCTCC
rs246997	G/T	0.45 (T)	0.36	0.50	1,440,165	Intron 3	TfiI	FP	AGGGACACATCCCACAATTC
								RP	TGGATAGGTCGATAGGTGGA

FP forward primer, *RP* reverse primer. **Global MAF not available

beta 8.4 (Hahn et al. 2003). MDR is a genetic model-free, non parametric method for detecting, characterizing, and interpreting nonlinear interactions among discrete genetic and environmental attributes.

Bioinformatics Analysis

The functional protein interaction network of DAT1 protein was retrieved from STRING version 10.5, an online computational prediction tool available at http://string-db.org (Szklarczyk et al. 2017). The confidence score was set at 0.7 and 0.9 to obtain the interaction results. Variant Effect Predictor (VEP) (McLaren et al. 2016) and Gene-Aware Variant INterpretation (GAVIN) (van der Velde et al. 2017) online tools were used to analyse the functional effect of the markers.

Results

Genotypic Distribution of *DAT1* Markers in the Studied Population

The homogeneous distribution of the genotypes in the studied population was evaluated for each *DAT1* marker by HWE analysis. The results showed that the genotypic distributions of all the markers except rs37021 (controls) conformed to HWE equilibrium (Table 2). All the ten markers were subjected to further analysis.

Table 2 Case-control association analysis using genotypic distribution and HWE analysis of *DAT1* markers

Marker	Genotypes	Genotype frequency		HWE chi-square (*p* value)		OR[a] (95% CI[b])	LRS[c] (df[d] = 2)	*p* value
		Cases (*n* = 132)	Controls (*n* = 147)	Cases (*n* = 132)	Controls (*n* = 147)			
rs28363170	LL	105 (0.796)	118 (0.803)	0.169 (0.681)	0.195 (0.659)	0.96 (0.53–1.72)	0.461	0.794
	LS	25 (0.189)	28 (0.190)			0.99 (0.55–1.81)		
	SS	2 (0.015)	1 (0.007)			2.19 (0.22–21.26)		
rs40184	CC	80 (0.606)	83 (0.565)	0.399 (0.528)	3.557 (0.059)	1.19 (0.74–1.91)	1.682	0.431
	CT	44 (0.333)	49 (0.333)			1.00 (0.61–1.64)		
	TT	8 (0.061)	15 (0.102)			0.58 (0.25–1.36)		
rs3836790	LL	79 (0.598)	82 (0.558)	0.429 (0.513)	0.053 (0.818)	1.18 (0.73–1.90)	0.770	0.680
	LS	48 (0.364)	55 (0.374)			0.96 (0.59–1.55)		
	SS	5 (0.038)	10 (0.068)			0.55 (0.20–1.57)		
rs27048	CC	87 (0.659)	91 (0.619)	0.038 (0.845)	0.027 (0.869)	1.19 (0.73–1.94)	0.530	0.767
	CT	40 (0.303)	49 (0.333)			0.87 (0.53–1.44)		
	TT	5 (0.038)	7(0.048)			0.79 (0.25–2.51)		
rs37021	GG	56 (0.424)	55 (0.374)	0.001 (0.974)	4.103 (0.043)	1.23 (0.76–1.99)	2.118	0.347
	GA	60 (0.454)	79 (0.537)			0.72 (0.45–1.15)		
	AA	16 (0.121)	13 (0.088)			1.42 (0.66–3.06)		
rs250683	TT	27 (0.205)	36 (0.245)	2.340 (0.126)	1.936 (0.164)	0.79 (0.45–1.39)	0.812	0.666
	TC	75 (0.568)	82 (0.558)			1.04 (0.65–1.67)		
	CC	30 (0.227)	29 (0.197)			1.20 (0.67–2.12)		
rs250682	GG	39 (0.296)	36 (0.245)	0.019 (0.890)	0.042 (0.837)	1.29(0.76–2.19)	1.031	0.597
	GC	65 (0.492)	75 (0.510)			0.93 (0.58–1.49)		
	CC	28 (0.212)	36 (0.245)			0.83 (0.48–1.45)		
rs427284	CC	29 (0.220)	36 (0.245)	0.108 (0.742)	0.290 (0.590)	0.87 (0.50–1.51)	0.522	0.770
	CG	68 (0.515)	77 (0.524)			0.97 (0.60–1.55)		
	GG	35 (0.265)	34 (0.231)			1.20 (0.70–2.06)		
rs458609	CC	35(0.265)	39 (0.265)	0.263 (0.608)	0.003 (0.956)	1.00 (0.59–1.7)	0.166	0.920
	CT	69 (0.523)	74 (0.503)			1.08 (0.68–1.73)		
	TT	28 (0.212)	34 (0.265)			0.90 (0.51–1.57)		
rs246997	TT	14 (0.106)	37 (0.252)	1.267 (0.260)	0.015 (0.901)	0.37 (0.21–0.69)	**12.443**	**0.002**
	TG	67 (0.508)	73 (0.497)			1.02 (0.64–1.63)		
	GG	51 (0.386)	37 (0.252)			**1.86 (1.12–3.08)**		

Best *p* value from 30 tests, 0.001675; adjusted *p* value from 1000 permutation test, 0.04695; standard error, 0.01557; empirical 5% quantile of the best *p* value, 0.002161; [a] Odds ratio; [b] Confidence interval; [c] Likelihood ratio; [d] Degrees of freedom; numbers in bold indicate significant findings

A Biased Genotypic Distribution Between Heroin-Dependent Cases and Controls

Case-control association analysis using the genotypic frequencies demonstrated significant differences between the cases and controls for the marker rs246997. The GG genotype was over-represented in heroin-dependent cases compared with controls ($p = 0.002$; LRS = 12.443), and significance retained after multiple corrections by 1000 permutations (adjusted $p = 0.047$) as shown in Table 2. It may be noted that alleles for 3′ UTR VNTR (rs28363170) and intron 8 VNTR (rs3836790) were grouped into two, for biallelic representation. 'L' allele represents 10R and above for rs28363170, and 7R and above for rs3836790. 'S' allele represents 9R and below for rs28363170, and 6R and below for rs3836790. Therefore, to analyse the individual genotypic frequencies, SPSS was used based on their presence or absence, and a significant difference was observed between heroin-dependent cases and controls for the 10R/11R genotype of rs28363170. The 10R/11R genotype was over-distributed among controls as compared with heroin-dependent cases ($p = 0.007$, chi-square = 7.342, OR = 0.24, df = 1), and further details are provided in Table 3. Overall genotypic analysis by SPSS with df = 5 revealed nominally significant differences between heroin-dependent cases and controls ($p = 0.053$, chi-square = 10.91) (Table 3).

Allelic Distribution of *DAT1* Markers in Heroin-Dependent Cases and Controls

Allelic frequencies of each marker were calculated separately for the cases and controls (Table 4). Allelic frequency-based

Table 3 Population-based analysis of the individual alleles and genotypes of the two VNTR sites using SPSS

Markers	Alleles[s]/ genotypes	Allelic frequency[s]/genotypic frequency		OR[a] (95% CI[b])	Chi-sq. (p value) {df[d] = 1}	Chi-sq. (df[d])	p value
		Cases ($n = 264$)[s] ($n = 132$)	Controls ($n = 294$)[s] ($n = 147$)				
rs28363170	6[s]	8 (0.03)[s]	9 (0.03)[s]	0.99 (0.37–2.60)[s]	0.001 (0.983)[s]	4.70[s] (3)	0.195[s]
	9[s]	21 (0.08)[s]	21 (0.07)[s]	1.12 (0.60–2.11)[s]	0.132 (0.717)[s]		
	10[s]	233 (0.88)[s]	251 (0.85)[s]	1.28 (0.79–2.10)[s]	1.005 (0.316)[s]		
	11[s]	2 (0.01)[s]	**13 (0.05)[s]**	**0.25 (0.09–0.69)[s]**	**7.139 (0.008)[s]**		
	10/10	103 (0.780)	105 (0.714)	1.41 (0.83–2.42)	1.598 (0.206)	10.911 (5)	0.053
	10/11	2 (0.015)	13 (0.089)	**0.24 (0.08–0.67)**	**7.342 (0.007)**		
	10/6	6 (0.046)	9 (0.061)	0.73 (0.26–2.08)	0.340 (0.560)		
	10/9	19 (0.144)	19 (0.129)	1.13 (0.57–2.24)	0.128 (0.721)		
	9/6	2 (0.015)	0 (0)	8.34 (0.52–134.58)	2.243 (0.134)		
	9/9	0 (0)	1(0.007)	0.15 (0.003–7.60)	0.901 (0.342)		
rs3836790	6[s]	58 (0.220)[s]	75 (0.255)[s]	0.79 (0.54–1.17)[s]	0.960 (0.327)[s]	8.269[s] (7)	0.310[s]
	7[s]	198 (0.75)[s]	215 (0.732)[s]	0.70 (0.47–1.03)[s]	0.253 (0.615)[s]		
	8[s]	3 (0.010)[s]	1 (0.003)[s]	3.05 (0.43–21.86)[s]	1.239 (0.266)[s]		
	9[s]	1 (0.004)[s]	0 (0)[s]	8.28 (0.16–419.60)[s]	1.116 (0.291)[s]		
	11[s]	1 (0.004)[s]	0 (0)[s]	8.28 (0.16–419.60)[s]	1.116 (0.291)[s]		
	12[s]	1 (0.004)[s]	3 (0.010)[s]	0.41 (0.06–2.91)[s]	0.805 (0.370)[s]		
	13[s]	1 (0.004)[s]	0 (0)[s]	8.28 (0.16–419.60)[s]	1.116 (0.291)[s]		
	14[s]	1 (0.004)[s]	0 (0)[s]	8.28 (0.16–419.60)[s]	1.116 (0.291)[s]		
	6/6	5 (0.041)	10 (0.064)	0.55 (0.20–1.57)	1.243 (0.265)	8.094 (9)	0.525
	6/7	46 (0.351)	54 (0.357)	0.92 (0.56–1.50)	0.108 (0.743)		
	6/8	2 (0.014)	1 (0.006)	2.19 (0.22–21.26)	0.456 (0.500)		
	7/7	73 (0.552)	79 (0.554)	1.06 (0.66–1.71)	0.068 (0.794)		
	7/8	1 (0.007)	0 (0)	8.28 (0.16–419.61)	1.118 (0.290)		
	7/9	1 (0.007)	0 (0)	8.28 (0.16–419.61)	1.118 (0.290)		
	7/11	1 (0.007)	0 (0)	8.28 (0.16–419.61)	1.118 (0.290)		
	7/12	1 (0.007)	3 (0.019)	0.40 (0.06–2.91)	0.810 (0.368)		
	7/13	1 (0.007)	0 (0)	8.28 (0.16–419.61)	1.118 (0.290)		
	7/14	1 (0.007)	0 (0)	8.28 (0.16–419.61)	1.118 (0.290)		

[s] Allelic data; [a] Odds ratio; [b] Confidence interval; [d] Degrees of freedom. Numbers in bold indicate significant findings

Table 4 Population-based case-control analysis of *DAT1* gene using allelic distribution of the various markers

Markers	Allele	Allelic counts in total subjects (frequency)		OR[a] (95% CI[b])	LRS[c]	*p* value (df[d] = 1)
		Cases ($n = 264$)	Controls ($n = 304$)			
rs28363170	L	235 (0.890)	264 (0.898)	0.92 (0.54–1.58)	0.090	0.765
	S	29 (0.110)	30 (0.102)	1.08 (0.63–1.86)		
rs40184	C	204 (0.773)	215 (0.731)	1.25 (0.85–1.83)	1.280	0.258
	T	60 (0.227)	79 (0.269)	0.80 (0.55–1.18)		
rs3836790	L	206 (0.780)	219 (0.745)	1.21 (0.82–1.79)	0.963	0.326
	S	58(0.220)	75 (0.255)	0.82 (0.56–1.22)		
rs27048	C	214 (0.811)	231 (0.786)	1.17 (0.77–1.76)	0.535	0.465
	T	50 (0.189)	63 (0.214)	0.86 (0.57–1.30)		
rs37021	C	172 (0.651)	189 (0.643)	1.04 (0.73–1.47)	0.046	0.831
	T	92 (0.349)	105(0.357)	0.96 (0.68–1.36)		
rs250683	T	129 (0.489)	154 (0.524)	0.87 (0.62–1.21)	0.689	0.407
	C	135 (0.511)	140 (0.476)	1.15 (0.83–1.60)		
rs250682	G	143 (0.542)	147 (0.50)	1.18 (0.85–1.65)	0.968	0.325
	C	121 (0.458)	147(0.50)	0.85 (0.61–1.18)		
rs427284	C	126 (0.477)	149 (0.507)	0.89 (0.64–1.24)	0.485	0.486
	G	138 (0.523)	145 (0.493)	1.13 (0.81–1.57)		
rs458609	C	139(0.526)	152 (0.517)	1.04 (0.75–1.45)	0.050	0.822
	T	125 (0.474)	142 (0.483)	0.96 (0.69–1.34)		
rs246997	T	95 (0.360)	147 (0.50)	0.57 (0.40–0.79)	**11.18**	**0.0008**
	G	169 (0.640)	147 (0.50)	**1.77 (1.26–2.47)**		

Best *p* value from 20 tests, 0.000852; adjusted *p* value from 1000 permutation test, 0.00799; standard error, 0.0028; empirical 5% quantile of the best *p* value, 0.005959; [a] Odds ratio; [b] Confidence interval; [c] Likelihood ratio; [d] Degrees of freedom; numbers in bold indicate significant findings

case-control association analysis showed a significant association of *DAT1* with heroin dependence. The G allele of rs246997 had significantly higher distribution in heroin-dependent cases as compared with controls ($p = 0.0008$; OR = 1.77, LRS = 11.18), and the association remained significant after multiple corrections testing with 1000 permutations (adjusted $p = 0.008$). SPSS analysis with df = 1 depicted a significant difference between heroin-dependent cases and controls for the 11R allele of 3′UTR VNTR (rs28363170). The 11R allele was over-distributed among controls as compared with heroin-dependent cases ($p = 0.008$; chi-square = 7.13; OR = 0.25; 95%CI = 0.09–0.69), but at df = 3, no significant difference was observed as shown in Table 3.

Biased Linkage Disequilibrium of Paired Markers Between Case and Control

Linkage disequilibrium (LD) analysis showed contrasting LD pattern between controls and cases for some marker pairs, and the details of the results are summarized in Table 5 and Fig. 1b. Comparatively, wider LD block starting from rs37021 to rs458609 was observed in controls, whereas it is disrupted in cases by a low LD region between pair of markers involving rs250683 [rs250683-rs250682 ($D' = 0.232$, $r^2 = 0.048$), rs250683-rs427284 ($D' = 0.303$, $r^2 = 0.087$), rs250683-rs458609 ($D' = 0.441$, $r^2 = 0.167$), rs250683-rs246997 ($D' = 0.299$, $r^2 = 0.048$)]. Higher LD profile for the same marker pairs involving rs250683 was observed in controls [rs250683-rs250682 ($D' = 0.711$, $r^2 = 0.460$), rs250683-rs427284 ($D' = 0.850$, $r^2 = 0.675$), rs250683-rs458609 ($D' = 0.840$, $r^2 = 0.687$), rs250683-rs246997 ($D' = 0.428$, $r^2 = 0.167$)]. A high LD forming a haplotype block with $D' = 0.946$ and $r^2 = 0.735$ was observed for the marker pair rs427284-rs458609 in heroin-dependent cases compared with controls ($D' = 0.792$, $r^2 = 0.602$). Comparatively, lower LD was observed in heroin-dependent cases than controls for the marker pairs rs28363170-rs40184 (cases, $D' = 0.746$, $r^2 = 0.234$; controls, $D' = 0.947$, $r^2 = 0.277$) and rs28363170-rs250683 (cases, $D' = 0.311$, $r^2 = 0.012$; controls, $D' = 0.764$, $r^2 = 0.060$). It may be noted again that alleles for rs28363170 and rs3836790 were grouped into two for a biallelic representation as mentioned above.

Case-Control Association Analysis Using Haplotypes

Bimarker-based haplotype analysis demonstrated that out of forty-five bimarker combinations analysed, fourteen bimarker combinations were found to be significant in the bimarker-based haplotypic association analysis (Table 6). Of these, only five bimarker combinations retained their significance after the corrections for multiple testing, viz. rs37021-rs250683, rs250683-rs250682, rs250683-rs427284, rs250683-

Table 5 Linkage disequilibrium analysis between marker pairs in heroin-dependent cases and controls

Markers	rs28363170	rs40184	rs3836790	rs27048	rs37021	rs250683	rs250682	rs427284	rs458609
Total controls (D′/r^2)									
rs40184	0.947/0.277								
rs3836790	0.26/0.022	0.056/ 0.003							
rs27048	0.151/0.009	0.506/0.190	0.398/0.015						
rs37021	0.203/0.003	0.089/0.005	0.013/0.000	0.345/0.018					
rs250683	0.764/0.060	0.481/0.077	0.201/0.013	0.509/0.064	0.841/0.357				
rs250682	0.390/0.017	0.249/0.023	0.027/0.000	0.190/0.010	0.637/0.225	0.711/0.460			
rs427284	0.548/0.033	0.348/0.043	0.116/0.004	0.328/0.029	0.772/0.322	0.850/0.675	0.677/0.446		
rs458609	0.575/0.035	0.388/0.052	0.216/0.015	0.322/0.026	0.823/0.352	0.840/0.687	0.687/0.441	0.792/0.602	
rs246997	0.498/0.028	0.167/0.01	0.281/0.027	0.207/0.012	0.303/0.051	0.428/0.167	0.295/0.087	0.424/0.175	0.464/0.201
Total cases (D′/r^2)									
rs40184	0.746/0.234								
rs3836790	0.377/0.062	0.067/0.004							
rs27048	0.238/0.030	0.288/0.066	0.231/0.004						
rs37021	0.169/0.002	0.027/0.000	0.167/0.015	0.272/0.032					
rs250683	0.311/0.012	0.046/0.001	0.032/0.000	0.393/0.038	0.238/0.032				
rs250682	0.277/0.011	0.187/0.012	0.186/0.011	0.266/0.020	0.908/0.521	0.232/0.048			
rs427284	0.303/0.012	0.201/0.013	0.249/0.019	0.386/0.038	0.902/0.476	0.303/0.087	0.849/0.668		
rs458609	0.373/0.015	0.228/0.014	0.318/0.026	0.423/0.038	0.913/0.401	0.441/0.167	0.884/0.594	0.946/0.735	
rs246997	0.264/0.005	0.057/0.001	0.411/0.027	0.056/0.000	0.338/0.034	0.299/0.048	0.399/0.076	0.471/0.114	0.477/0.142

rs458609, and rs250683-rs246997. Out of the fourteen significant combinations, all marker combinations involving the marker rs246997 showed significance as also demonstrated in allelic and genotypic analyses. It may be noted that rs250683 showed significant association at the haplotypic level and this finding points towards the existence of a combinatorial effect in heroin dependence.

Bimarker-based haplotype analysis demonstrated that haplotype of marker combinations such as L-G rs28363170-rs246997 ($p = 0.008$, OR = 1.72), T-G rs40184-rs246997 ($p = 0.004$, OR = 1.30), L-G rs38336790-rs246997 ($p = 0.004$, OR = 1.72), T-G rs27048-rs246997 ($p = 0.006$, OR = 1.80), G-T and A-C rs37021-rs250683 ($p = 0.0001$, OR = 1.58 and 4.49), G-G rs37021-rs250682 ($p = 0.006$, OR = 1.43), G-G rs37021-rs246997 ($p = 0.007$, OR = 1.64), T-G and C-C rs250683-rs250682 ($p < 0.0001$, OR = 2.54 and 2.83), T-G and C-C rs250683-rs427284 ($p < 0.0001$, OR = 3.56 and 4.45), T-T and C-C rs250683-rs458609 ($p < 0.0001$, OR = 2.89 and 4.14), C-G rs250683-rs246997 ($p = 0.0005$, OR = 2.36), G-G rs250682-rs246997 ($p = 0.002$, OR = 1.82), G-G rs427284-rs246997 ($p = 0.001$, OR = 2.03), and T-G rs458609-rs246997 ($p = 0.004$, OR = 1.81) was significantly over-distributed in heroin-dependent cases compared with controls. On the other hand, haplotype of marker combinations such as L-T rs28363170-rs246997 ($p = 0.008$, OR = 0.55), C-T and T-T rs40184-rs246997 ($p = 0.004$, OR = 0.67 and 0.62), S-T and L-T rs38336790-rs246997 ($p = 0.004$, OR = 0.50 and 0.66), C-T rs27048-rs246997 ($p = 0.006$, OR = 0.60), G-C and A-T rs37021-rs250683 ($p = 0.0001$, OR = 0.74 and 0.56), G-C and A-G rs37021-rs250682 ($p = 0.006$, OR = 0.57 and 0.31), G-T and A-T rs37021-rs246997 ($p = 0.007$, OR = 0.64 and 0.64), T-C rs250683-rs250682 ($p < 0.0001$, OR = 0.52), T-C and C-G rs250683-rs427284 ($p < 0.0001$, OR = 0.50 and 0.66), T-C and C-T rs250683-rs458609 ($p < 0.0001$, OR = 0.62 and 0.68), C-T rs250683-rs246997 ($p = 0.0005$, OR = 0.61), C-T rs250682-rs246997 ($p = 0.002$, OR = 0.53), C-T rs427284-rs246997 ($p = 0.001$, OR = 0.58), and T-T and C-T rs458609-rs246997 ($p = 0.004$, OR = 0.65 and 0.65) was over-represented in controls (Table 6).

Correlation Analysis

Bivariate correlation analysis among different variables demonstrated significant positive correlation between rs246997-GG ($p = < 0.001$, Pearson correlation = 0.217) with heroin dependence. Significant negative correlations were observed for heroin dependence with rs246997-TT ($p = 0.001$, Pearson correlation = − 0.203) and disease status with rs28363170-10R/11R ($p = 0.007$, Pearson correlation = −0.162). No significant correlations were observed for the variables ‘years of drug use’, ‘number of times drug abstained’, and ‘onset’ with rs28363170 genotypes or rs246997 genotypes.

Table 6 Case-control association analysis using haplotypic distribution of bimarker combinations of *DAT1* for heroin dependence

Marker pairs	Haplotypes	Cases (freq)	Controls (freq)	OR (95% CI)	LRS (*p* value)
rs28363170-rs246997	S-T	7.68 (0.03)	7.53 (0.03)	1.14 (0.41–3.16)	11.70 (0.008)
	S-G	21.32 (0.08)	22.46 (0.08)	1.06 (0.57–1.97)	
	L-T	87.32 (0.33)	139.5 (0.47)	0.55 (0.39–0.77)	
	L-G	147.7(0.56)	124.5 (0.42)	1.72 (1.23–2.40)	
rs40184-rs246997	C-T	20.36 (0.08)	32.9 (0.11)	0.67 (0.38–1.18)	13.36 (0.004)
	C-G	39.64(0.15)	46.1 (0.16)	0.95 (0.60–1.51)	
	T-T	74.64 (0.28)	114.1 (0.39)	0.62 (0.44–0.89)	
	T-G	129.4 (0.49)	100.9 (0.34)	1.30 (1.31–2.57)	
rs3836790-rs246997	S-T	12.3 (0.05)	26.96 (0.09)	0.50 (0.26–0.96)	13.53 (0.004)
	S-G	45.7 (0.17)	48.04 (0.16)	1.07 (0.69–1.67)	
	L-T	82.7 (0.31)	120 (0.41)	0.66 (0.47–0.94)	
	L-G	123.3 (0.47)	98.96 (0.34)	1.72 (1.23–2.42)	
rs27048-rs246997	C-T	78.02 (0.30)	122 (0.42)	0.60 (0.42–0.84)	12.50 (0.006)
	C-G	136 (0.52)	109 (0.37)	1.80 (1.28–2.51)	
	T-T	16.98 (0.06)	24.99 (0.08)	0.74 (0.40–1.39)	
	T-G	33.02 (0.13)	38.01 (0.13)	0.96 (0.59–1.58)	
rs37021-rs250683	G-T	72.86 (0.28)	56.94 (0.19)	1.58 (1.07–2.35)	20.96 (0.0001)*
	G-C	99.14 (0.38)	132.1 (0.45)	0.74 (0.53–1.03)	
	A-T	56.14 (0.21)	97.06 (0.33)	0.56 (0.38–0.81)	
	A-C	35.86 (0.14)	7.94 (0.03)	4.49 (2.42–8.33)	
rs37021-rs250682	G-C	33.59 (0.13)	61.08 (0.21)	0.57 (0.36–0.88)	12.38 (0.006)
	G-G	138.4 (0.52)	127.9 (0.44)	1.43 (1.02–1.99)	
	A-C	87.41 (0.33)	85.92 (0.29)	1.20 (0.84–1.72)	
	A-G	4.59 (0.02)	19.08 (0.06)	0.31 (0.14–0.71)	
rs37021-rs246997	G-T	73.07 (0.28)	110.4 (0.38)	0.64 (0.45–0.91)	12.05 (0.007)
	G-G	98.93 (0.37)	78.57 (0.27)	1.64 (1.15–2.34)	
	A-T	21.93 (0.08)	36.57 (0.12)	0.64 (0.37–1.11)	
	A-G	70.07 (0.27)	68.43 (0.23)	1.19 (0.81–1.75)	
rs250683-rs250682	T-C	73.5 (0.28)	126.8 (0.43)	0.52 (0.36–0.73)	23.93 (2.582e-005)*
	T-G	55.5 (0.21)	27.2 (0.09)	2.54 (1.59–4.05)	
	C-C	47.5 (0.18)	20.2 (0.07)	2.83 (1.70–4.71)	
	C-G	87.5 (0.33)	119.8 (0.41)	0.72 (0.51–1.02)	
rs250683-rs427284	T-C	81.07 (0.31)	138.4 (0.47)	0.50 (0.36–0.71)	41.49 (5.141e-005)*
	T-G	47.93 (0.18)	15.64 (0.05)	3.56 (2.11–6.00)	
	C-C	44.93 (0.17)	10.64 (0.04)	4.45 (2.55–7.74)	
	C-G	90.07 (0.34)	129.4 (0.44)	0.66 (0.47–0.93)	
rs250683-rs458609	T-T	34.15 (0.13)	13.57 (0.04)	2.89 (1.60–5.24)	32.66 (3.80e-007)*
	T-C	94.85 (0.36)	140.4 (0.48)	0.62 (0.44–0.86)	
	C-T	90.85 (0.34)	128.4 (0.44)	0.68 (0.48–0.95)	
	C-C	44.15 (0.17)	11.57 (0.04)	4.14 (2.38–7.20)	
rs250683-rs246997	T-T	32.54 (0.12)	47.04 (0.16)	0.74 (0.46–1.19)	17.55 (0.0005)*
	T-G	96.46 (0.37)	107 (0.36)	1.01 (0.71–1.42)	
	C-T	62.46 (0.24)	99.96 (0.34)	0.61 (0.42–0.87)	
	C-G	72.54 (0.27)	40.04 (0.14)	2.36 (1.56–3.57)	
rs250682-rs246997	C-T	26.17 (0.10)	51.84 (0.18)	0.53 (0.33–0.85)	15.11 (0.002)
	C-G	94.83 (0.36)	95.16 (0.32)	1.17 (0.83–1.66)	
	G-T	68.83 (0.26)	95.16 (0.32)	0.74 (0.51–1.06)	
	G-G	74.14 (0.28)	51.84 (0.18)	1.82 (1.22–2.70)	
rs427284-rs246997	C-T	23.99 (0.09)	43.77 (0.15)	0.58 (0.35–0.97)	15.89 (0.001)
	C-G	102 (0.39)	105.2 (0.36)	1.13 (0.80–1.59)	
	G-T	71.01(0.27)	103.2 (0.35)	0.68 (0.48–0.98)	

Table 6 (continued)

Marker pairs	Haplotypes	Cases (freq)	Controls (freq)	OR (95% CI)	LRS (*p* value)
	G-G	66.99 (0.25)	41.77 (0.14)	2.03 (1.34–3.09)	
rs458609-rs246997	T-T	68.85(0.26)	103.9 (0.35)	0.65 (0.45–0.93)	13.08 (0.004)
	T-G	56.15 (0.21)	38.07 (0.13)	1.81 (1.16–2.82)	
	C-T	26.15 (0.10)	43.07 (0.15)	0.65 (0.39–1.07)	
	C-G	112.8 (0.43)	108.9 (0.37)	1.27 (0.90–1.78)	

Best *p* value from 180 tests: 2.169e-005

Adjusted *p* value from permutation test, 0.000999; standard error, 0.000999

Empirical 5% quantile of the best *p* value, 0.0005534

*Denotes significance after multiple testing correction with1000 permutations

Regression Analysis

Binary logistic regression analysis with heroin use disorder status as dependent variable and rs28363170-10R/10R, rs28363170-10R/9R, rs28363170-10R/11R, rs246997-TT, rs246997-GT, and rs246997-GG as covariates revealed significant odds ratio for rs246997-GG (OR = 2.576, 95% CI = 1.49–4.45, $p = 0.001$), rs28363170-10R/11R (OR = 0.176, 95% CI = 0.04–0.81, $p = 0.025$), and rs246997-TT (OR = 0.392, 95% CI = 0.19–0.80, $p = 0.010$). The Hosmer-Lemeshow test of goodness of fit at df = 1 was not significant ($p = 0.887$) depicting the fitness of proposed model. ln(odds) of the derived logistic regression equation was − 0.302 with $p = 0.041$ and odds ratio of 0.74.

Marker-Marker Interaction Analysis

To investigate the interaction between markers of the *DAT1* gene, an MDR test was performed. The interaction of all the markers in the study is presented in the dendrogram generated through MDR (Fig. 2a). The strength of the synergistic interactions between the studied markers along with their values is demonstrated by the entropy-based interaction graph (Fig. 2b). Marker rs250683 has shown a maximum synergistic effect with most of the studied markers. The estimated information gain (IG) for some pairwise combination (interaction effect) was found to be higher than that of individual markers, as seen in the case of rs250683 (individual entropy value = 0.21) with rs427284 (individual entropy value = 0.13,

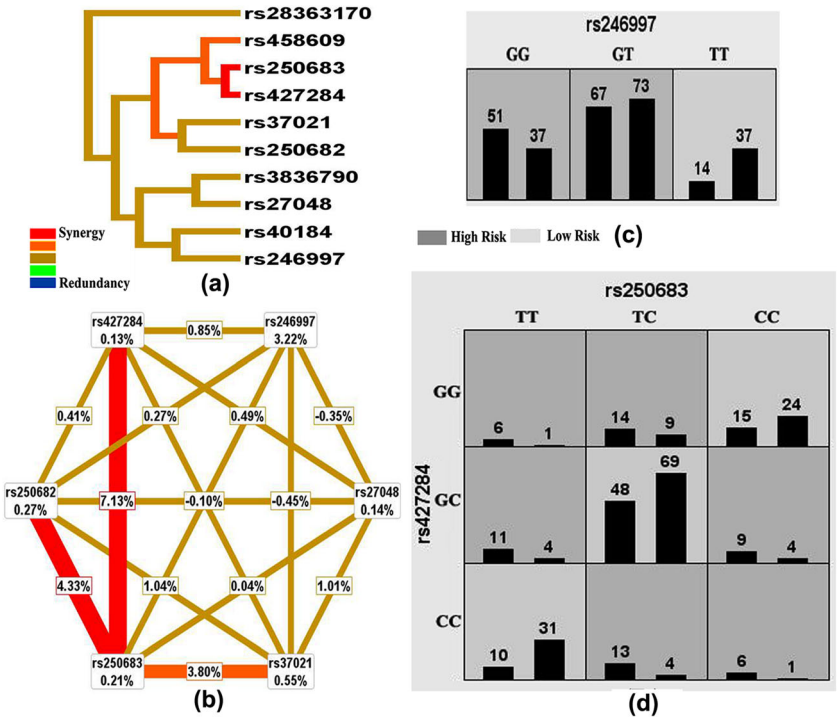

Fig. 2 MDR generated individual and interactive data. **a** Dendrogram showing interactions of all the studied markers of the *DAT1* gene. Colour variation indicates the strength of interaction, and the increase in interaction is presented from blue (redundant) towards red (synergy). **b** Circle graph showing different entropy levels of individual and interacting markers for heroin dependence. The strong interaction is indicated when the entropy value is higher in the combined set than the individual markers. **c** Graphical view showing high-risk and low-risk genotype for heroin dependence in single-marker model. **d** Graphical view showing high-risk and low-risk genotype for heroin dependence in two marker interaction model

synergistic entropy value = 7.13), rs250682 (individual entropy value = 0.27, synergistic entropy value = 4.33), and rs37021 (individual entropy value = 0.55, synergistic entropy value = 3.80) (Fig. 2b). Marker rs246997 has shown a high individual entropy value of 3.22 but no effective synergism with any of the studied markers. Further, the best interaction models were generated based on maximum cross-validation consistency (CVC) and the lowest prediction error (PE). The result provided four best models which include a highest single-marker effect of rs246997 ($p = 0.0017$, OR = 2.84, CVC = 10/10, PE = 0.427), highest interactive effects of two marker combination, viz. rs250683-rs427284 ($p = < 0.0001$, OR = 4.36, CVC = 10/10, PE = 0.355), three marker combination, viz. rs37021-rs250683-rs427284 ($p = < 0.0001$, OR = 4.97, CVC = 5/10, PE = 0.322), and four marker combination, viz. rs27048-rs37021-rs250683-rs250682 ($p = < 0.0001$, OR = 8.94, CVC = 4/10, PE = 0.272), for heroin dependence. Genotype GG of rs246997 was more prevalent among cases than the controls (Fig. 2c). The most effective marker interaction was found to be rs250683-rs427284, where TT-GG, TT-GC, TC-GG, TC-CC, CC-GC, and CC-CC genotypes interact more among cases compared with controls (Fig. 2d).

Bioinformatics Analysis

Variant Effect Predictor (VEP) and Gene-Aware Variant INterpretation (GAVIN) online tools were used to check the predicted effects of the markers. Both VEP and GAVIN have scored the impact of all the studied ten markers as modifier (usually non-coding variants or variants affecting non-coding genes, where predictions are difficult, or there is no evidence of impact). DAT protein interaction network depicting all known, predicted, and other interactions generated through the online search tool 'STRING' is provided in Fig. 3. The maximum DAT interaction score was observed with DRD2 (0.971). The remaining nine highest scoring DAT-interacting proteins were SNCA (0.968), CDH1 (0.908), FLOT2 (0.900), CAMK2A (0.870), SLC618A2 (0.860), TH (0.840), CAMK2G (0.837), CAMK2B (0.804), and PRKCA (0.804).

Discussion

The major results of the present study are as follows: (i) this is the first study to report the association of marker rs246997 of DAT1 gene with heroin dependence at allelic, genotypic, and haplotypic levels; (ii) association of five bimarker-based haplotype combination of DAT1 gene, viz. rs37021-rs250683, rs250683-rs250682, rs250683-rs427284, rs250683-rs458609, and rs250683-rs246997, with heroin dependence even after multiple testing correction; (iii) association of 11R allele and 10R/11R genotype of 3′UTR VNTR (rs28363170) of DAT1 gene with protection against heroin dependence, which is being reported for the first time; (iv) the finding of 7R and 6R alleles as the common alleles of intron 8 VNTR (rs3836790) of DAT1 gene in this study population indicates ethnicity-linked variation in allele frequency of this marker; and (v) this is for the first time that 13R allele (7R/13R genotype) and 14R allele (7R/14R genotype) are identified for intron 8 VNTR (rs3836790).

In the present study, the marker rs246997 of *DAT1* was found to be significantly associated with heroin dependence even after correction for multiple testing with 1000 permutation, and it has not been cited or reported earlier for its association with heroin dependence in other similar studies. MDR also demonstrated the highest single-marker effect of rs246997 with a cross-validation consistency of 10/10 and a p value of 0.027. Correlation analysis and regression analysis have also demonstrated the association of this marker with heroin dependence. However, no correlation of rs246997

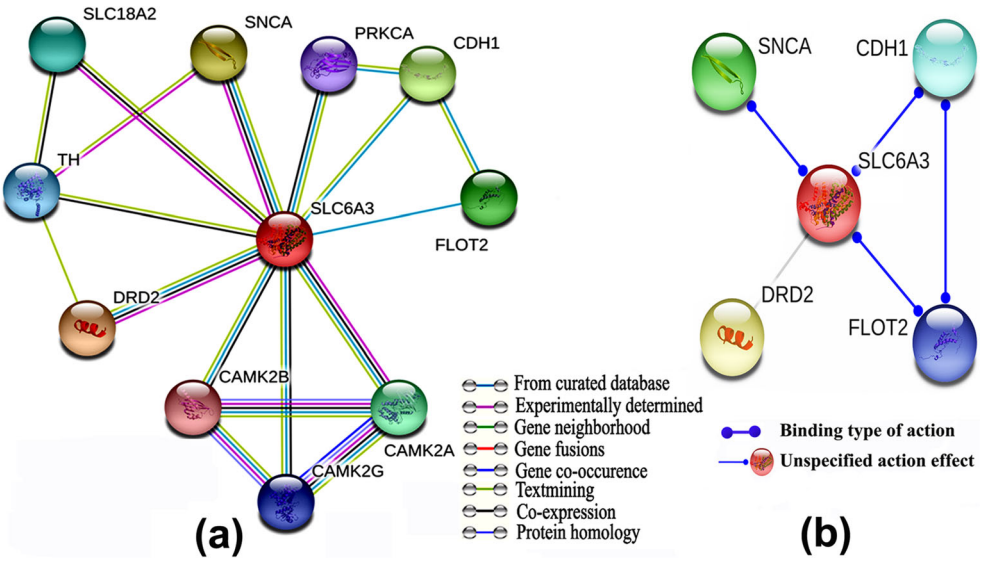

Fig. 3 DAT1/SLC6A3 protein interaction model generated through the online search tool 'STRING'. **a** DAT1/SLC6A3 and other protein interactions, showing evidence from various sources at a set confidence score of 0.7. **b** DAT1/SLC6A3 and other protein interactions, depicting molecular action at a set confidence score of 0.9

genotypes with possible determinants of heroin dependence, such as years of drug use, age at onset, and the number of times drug abstained, was observed. This could be because of the small sample size of the study. Moreover, the assessment and analysis of these possible determinants of heroin dependence were based on individual responses to a questionnaire. A previous study (Koijam et al. 2020) with the same population had demonstrated a significant association of an opioid receptor mu 1 (*OPRM1*) variant with the number of times drug abstained.

Since the VNTRs have more than two alleles and three genotypes, possible differential role of alleles and genotypes as a group and as an individual entity was checked using UNPHASED and SPSS. Analysis using UNPHASED by the grouping of all repeats into higher repeat alleles and lower repeat alleles failed to reveal any possible effect on phenotype. However, when analysed individually with df = 1 by SPSS, a significant difference between heroin-dependent cases and controls for 11R allele and 10R/11R genotype of 3′UTR VNTR (rs28363170) was demonstrated. When overall individual alleles and genotypes of 3′UTR VNTR (rs28363170) were analysed with df > 1, nominal significance was observed at the genotypic level ($p = 0.053$, df = 5) but not at allelic level.

A study by Hou and Li (2009) on 3′UTR VNTR (rs28363170) of the DAT1 gene in Han Chinese population found no association of the DAT1 gene with heroin dependence. Vereczkei et al. (2013) reported a lack of association of DAT1 gene variants 3′UTR VNTR (rs28363170) and intron 8 VNTR (rs3836790) with heroin dependence in the Hungarian population. Another study by Yeh et al. (2010) on Han Chinese involving 15 variants of the DAT1 gene, including rs27048, 3′ UTR VNTR (rs28363170), and intron 8 VNTR (rs3836790), reported absence of association with heroin dependence. The non-association of rs27048 and intron 8 VNTR (rs3836790) with heroin dependence was replicated in the present study. The non-association results involving 3′UTR VNTR (rs28363170) from these studies were based on analysis after the dualization or grouping method of VNTR, which is similar to our result under similar analysis method. However, the present study analysed 3′UTR VNTR (rs28363170) further at individual allele and individual genotype level with df = 1 and df > 1 rather than a limited binary or grouping approach, based on which association of 3′UTR VNTR (rs28363170) was observed as described above. The contradicting results generated from different studies in different populations may be because of differences in the study design, variants studied, analysis methods, sample size, differential allele frequencies, ethnic differences, and population-specific epigenetic interactions with other genes. Further, gene polymorphisms involved in the vulnerability or association with heroin dependence may be population-specific, as demonstrated by Bhaskar et al. (2012).

The participants in the present study are from Manipur of North East India only and are traced to East Asian origin (Saraswathy et al. 2009). In this study, utmost care has been taken in the selection of the subjects. They belonged to an ethnically homogeneous and small population, which is also different from the rest of the Indian population. Controls with appropriate size from a similar background were included for proper representation. Analysis based on population stratification was not performed due to the very low frequency of ethnically different groups and overall small sample size. Population admixtures and gene flow in many generations may play a role in the actual inference of the study as well.

The present study reports for the first time the association of 11R allele and 10R/11R genotype of 3′UTR VNTR (rs28363170) with protection against heroin dependence. Future studies are suggested to look into 11R allele and 10R/11R genotype of 3′UTR VNTR towards association with heroin dependence. The study has also highlighted a differential distribution of the alleles of intron 8 VNTR (rs3836790) in the population, indicating ethnicity-oriented changes in allele frequency. The 7R and 6R alleles of this marker were found to be more common in the population, and lower repeat alleles were undetectable. The commonly reported alleles of the marker are 5R and 6R alleles. Similar allelic prevalence was reported by Kim et al. (2017) in the Korean population and by Zhou et al. (2014) in the African American population, where 6R of intron 8 VNTR was the minor allele as against 5R allele (Guindalini et al. 2006). In the case of intron 8 VNTR (rs3836790), alleles such as 6R, 7R, 8R, 9R, 11R, 12R, 13R, and 14R were detected in the study in heroin-dependent cases; though, most of the alleles showed very low frequencies. However, 11R, 13R, and 14R alleles were absent in the controls in the same study group. 13R (7R/13R genotype) and 14R (7R/14R genotype) alleles of intron 8 VNTR (rs3836790) were also detected in two heroin-dependent cases for the first time. To date, studies have reported the occurrence of only up to 12R alleles for intron 8 VNTR (Gadow et al. 2014).

VEP and GAVIN online prediction tools scored the impact of all the ten markers as modifiers. It may be noted that the mere absence of evidence of impact or difficulty in predictions for a variant cannot rule its involvement in a particular phenotype or regulation out. The 3′UTR VNTR and intron 8 VNTR have also been scored as modifiers, whereas they have been suggested to regulate or modulate DAT expression (Pinsonneault et al. 2011). DAT-interacting proteins provided by STRING like DRD2, SNCA, CAMK2A, and CAMK2B have reports of association with substance dependence (Easton et al. 2014; Janeczek et al. 2014; Lehrmann et al. 2006; Spronk et al. 2016). DAT may also act in combination with these proteins to provide an overall effect on the phenotype of heroin dependence.

Some of the limitations of the present study include relatively smaller sample size and less number of variants screened, which could be extended with higher sample size

and more variants. Further, DAT expression levels in brain were not included, and some of the possible determinants of heroin dependence such as years of drug use, age at onset, number of times drug abstained, etc. used in the study were solely based on individual responses to questionnaire.

Such association studies help in the identification of genetic risk factors to a particular disease. The identified genetic risk factors may not necessarily affect expression of the genes. However, the risk factors may vary from population to population due to population-specific variations. Studies of such gene polymorphisms give better understanding of the phenotypic effects that they may induce. One of the extensively studied polymorphisms is the SNP (118A>G) in the *OPRM1* (mu opioid receptor) gene whose product is involved in reward, addictive behaviour, and pain regulation. The SNP (118A>G) changes a single amino acid, Asn 40 to Asp in the receptor protein (Beyer et al. 2004), leading to reduced protein translation and availability of the receptor likely stimulating opioid addiction (Ahmed et al. 2018). Therefore, gene polymorphisms may also affect the functions of genes, which in turn may affect the pathophysiology of drug dependence. Further, the extension of such heroin dependence studies towards a metabonomic and clinical perspective (Xie et al. 2015) is highly encouraging. In the present study, individuals carrying the G allele, GG genotype at rs246997, have been identified as having a higher risk for susceptibility to drug dependence. Therefore, the study suggests prioritizing individuals with G allele, GG genotype at rs246997, during prevention, intervention, and clinical treatment of heroin dependence. Further, the 11R allele at 3′UTR VNTR (rs28363170) has been identified as a protective allele for heroin dependency. This result suggests that this marker could be used for stratification of patients with lower risk. The findings of the present study, in concurrence with other similar studies (Dang et al. 2014; Levran et al. 2015), provide strategies for the development of novel prevention, patient stratification, targeted intervention, and personalized treatment for heroin dependence. However, the strategies stemming from such association studies have an inherent problem of the inability for generalized usage across all human populations, owing to the complex nature of the disease such as ethnic susceptibility differences, individual differences, epistatic interactions, and involvement of multiple genes and other environmental conditions. Thus, as with any other complex disease, there is a constant need for understanding all the possible factors and mechanisms underlying heroin dependency.

Conclusion

The present study provides evidence that the *DAT1* gene is a susceptibility gene for heroin dependence in a subpopulation from India, where such phenotype is highly prevalent. The presence of 'G' allele at rs246997 in an individual is a risk factor that increases the probability of getting heroin-dependent. Furthermore, the presence of 11R at 3′UTR VNTR (rs28363170) confers protection from heroin dependency. Identification of a novel marker and low-frequency allele of a commonly studied polymorphism as effectors for heroin dependence opens up further avenues for research to understand the possible mechanisms through which it modifies the phenotype to reinforce its association with heroin dependence. Although the variants identified in this study are suggested to produce low relative risk towards vulnerability to heroin dependency, it may uncover novel mechanisms of addictive behaviour.

Acknowledgements We are thankful to the Social Awareness Service Organization (SASO), Imphal, and all the participants in the study.

Authors' Contributions Reena Haobam, Usha Rajamma, and Kanchan Mukhopadhyay conceived, designed, and supervised the work; Arunkumar Singh Koijam performed research work; Aruna Chanu Hijam and Preeti Jaiswal assisted in research work; Arunkumar Singh Koijam and Asem Surindro Singh analysed data; Arunkumar Singh Koijam and Reena Haobam prepared the manuscript for publication. All the authors read and agreed to the manuscript.

Funding Information The present study is supported by the Department of Biotechnology (DBT), Govt. of India grant to Reena Haobam and Usha Rajamma vide BT/251/NE/TBP/2011 dated 14.5.2012. Arunkumar Singh Koijam worked for this project as Junior Research Fellow.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in the study were per the ethical standards of the Institutional Human Ethics Committee, Manipur University, and with the 1964 Helsinki declaration and its later amendments.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

Ahmed M, Haq I, Faisal M, Waseem D, Taqi MM (2018) Implication of *OPRM1* A118G polymorphism in opioids addicts in Pakistan: in vitro and in silico analysis. J Mol Neurosci 65:472–479

American Psychiatric Association: Diagnostic and Statiscal Manual of Mental Disorders (1994) 4th edn. Washington, DC

Bannon MJ, Michelhaugh SK, Wang J, Sacchetti P (2001) The human dopamine transporter gene: gene organization, transcriptional regulation, and potential involvement in neuropsychiatric disorders. Eur Neuropsychopharmacol 11:449–455

Barret JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263–265

Beyer A, Koch T, Schroder H, Schulz S, Hollt V (2004) Effect of the A118G polymorphism on binding affinity, potency and agonist-mediated endocytosis, desensitization, and resensitization of the human mu-opioid receptor. J Neurochem 89:553–560

Bhaskar LVKS, Thangaraj K, Wasnik S, Singh L, Rao VR, Dopamine transporter (DAT1) VNTR (2012) Polymorphism and alcoholism in two culturally different populations of South India. Am J Addict 21: 343–347

Blum K, Chen AL, Chen TJ, Braverman ER, Reinking J, Blum SH, Cassel K, Downs BW, Waite RL, Williams L, Prihoda TJ, Kerner MM, Palomo T, Comings DE, Tung H, Rhoades P, Oscar-Berman M (2008) Activation instead of blocking mesolimbic dopaminergic reward circuitry is a preferred modality in the long term treatment of reward deficiency syndrome (RDS): a commentary. Theor Biol Med Model 5:24. https://doi.org/10.1186/1742-4682-5-24

Brewer AJ, Nielsen DA, Spellicy CJ, Hamon SC, Gingrich J, Thompson-Lake DGY, Nielsen EM, Mahoney JJ, Kosten TR, Newton TF, De La Garza R (2015) Genetic variation of the dopamine transporter (DAT1) influences the acute subjective responses to cocaine in volunteers with cocaine use disorders. Pharmacogenet Genomics 25: 296–304

Dang W, Zhang Q, Zhu Y-S, Lu X-Y (2014) The evidence for the contribution of the autism susceptibility candidate 2 (AUTS2) gene in heroin dependence susceptibility. J Mol Neurosci 54:811–819

Dreher JC, Kohna P, Kolachana B, Weinberger DR, Berman KF (2009) Variation in dopamine genes influences responsivity of the human reward system. Proc Natl Acad Sci U S A 106:617–622

Dudbridge F (2008) Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. Hum Hered 66:87–98

Easton AC, Lourdusamy A, Havranek M, Mizuno K, Solati J, Golub Y, Clarke T-K, Vallada H, Laranjeira R, Desrivières S, Moll GH, Mössner R, Kornhuber J, Schumann G, Giese KP, Fernandes C, Quednow BB, Müller CP (2014) αCaMKII controls the establishment of cocaine's reinforcing effects in mice and humans. Transl Psychiatry 4:e457. https://doi.org/10.1038/tp.2014.97

Everitt BJ, Robbins TW (2005) Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. Nat Neurosci 8: 1481–1489

Gadow KD, Pinsonneault J, Perlman G, Sadee W (2014) Association of dopamine gene variants, emotion dysregulation and ADHD in autism spectrum disorder. Res Dev Disabil 35:1658–1665

Gelernter J, Kranzler HR, Satel SL, Rao PA (1994) Genetic association between dopamine transporter protein alleles and cocaine-induced paranoia. Neuropsychopharmacology 191:195–200

Giros B, el Mestikawy S, Godinot N, Zheng K, Han H, Yang-Feng T, Caron MG (1992) Cloning, pharmacological characterization, and chromosome assignment of the human dopamine transporter. Mol Pharmacol 42:383–390

Giros B, Jaber M, Jones SR, Wightman RM, Caron MG (1996) Hyperlocomotion and indifference to cocaine and amphetamine in mice lacking the dopamine transporter. Nature 379:606–612

Guindalini C, Howard M, Haddley K, Laranjeira R, Collier D, Ammar N, Craig I, O'Gara C, Bubb VJ, Greenwood T, Kelsoe J, Asherson P, Murray RM, Castelo A, Quinn JP, Vallada H, Breen G (2006) A dopamine transporter gene functional variant associated with cocaine abuse in a Brazilian sample. Proc Natl Acad Sci U S A 103: 4552–4557

Hahn LW, Ritchie MD, Moore JH (2003) Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions. Bioinformatics 19:376–382

Hou QF, Li SB (2009) Potential association of DRD2 and DAT1 genetic variation with heroin dependence. Neurosci Lett 464:127–130

Janeczek P, MacKay RK, Lea RA, Dodd PR, Lewohl JM (2014) Reduced expression of α-synuclein in alcoholic brain: influence of SNCA-Rep1 genotype. Addict Biol 19:509–515

Kim WT, Lee SR, Roh YG, Kim S, Choi YH, Mun MH, Jeong MS, Koh SS, Leem SH (2017) Characterization of VNTRs within the entire region of SLC6A3 and its association with hypertension. DNA Cell Biol 36:1–10

Kobayashi S, Schultz W (2008) Influence of reward delays on responses of dopamine neurons. J Neurosci 28:7837–7846

Koijam AS, Chakraborty B, Mukhopadhyay K, Rajamma U, Haobam R (2020) A single nucleotide polymorphism in *OPRM1* (rs483481) and risk for heroin use disorder. J Addict Dis 38:214–222. https://doi.org/10.1080/10550887.2020.1740070

Koob GF, Nestler EJ (1997) The neurobiology of drug addiction. J Neuropsychiatry Clin Neurosci 9:482–497

Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3. Bioinformatics 23:1289–1291

Le Strat Y, Ramoz N, Pickering P, Burger V, Boni C, Aubin H, Ade's J, Batel P, Gorwood P (2008) The 3′ part of the dopamine transporter gene DAT1/SLC6A3 is associated with withdrawal seizures in patients with alcohol dependence. Alcohol Clin Exp Res 32:1–9

Lehrmann E, Colantuoni C, Deep-Soboslay A, Becker KG, Lowe R, Huestis MA, Hyde TM, Kleinman JE, Freed WJ (2006) Transcriptional changes common to human cocaine, cannabis and phencyclidine abuse. PLoS One 1(1):e114. https://doi.org/10.1371/journal.pone.0000114

Levran O, Randesi M, daRosa JC, Ott J, Rotrosen J, Adelson M, Kreek MJ (2015) Overlapping dopaminergic pathway genetic susceptibility to heroin and cocaine addictions in African Americans. Ann Hum Genet 79:188–198

Ma Y, Yuan W, Cui W, Li MD (2016) Meta-analysis reveals significant association of 3′-UTR VNTR in SLC6A3 with smoking cessation in Caucasian populations. Pharmacogenomics J 16:10–17

McLaren W, Laurent G, Hunt SE, Riat HS, Ritchie GRS, Thormann A, Flicek P, Cunningham F (2016) The Ensembl Variant Effect Predictor. Genome Biol 17:122. https://doi.org/10.1186/s13059-016-0974-4

Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16:1215

National Institute on Drug Abuse (2010) Drugs, brains and behavior. The Science of Addiction. NIH Publication No. 10–5605. National Institute of Health, Bethesda

NCBI Database. https://www.ncbi.nlm.nih.gov/gene/6531

Pinsonneault JK, Han DD, Burdick KE, Kataki M, Bertolino A, Malhotra AK, Gu HH, Sadee W (2011) Dopamine transporter gene variant affecting expression in human brain is associated with bipolar disorder. Neuropsychopharmacology 36(8):1644–1655

Rhee SH, Hewitt JK, Young SE, Corley RP, Crowley TJ, Stallings MC (2003) Genetic and environmental influences on substance initiation, use and problem use in adolescents. Arch Gen Psychiatry 60: 1256–1264

Robinson TE, Berridge KC (2008) The incentive sensitization theory of addiction: some current issues. Philos Trans R Soc Lond Ser B Biol Sci 363:3137–3146

Sander T, Harms H, Podschus J, Finckh U, Nickel B, Rolfs A, Rommelspacher H, Schimidt LG (1997) Allelic association of a dopamine transporter gene polymorphism in alcohol dependence with withdrawal seizures or delirium. Biol Psychiatry 41:299–304

Saraswathy KN, Kiranmala N, Murry B, Sinha E, Saksena D, Kaur H, Sachdeva MP, Kalla AK (2009) A genomic insight into the diversity among tribal and non-tribal population groups of Manipur. Biochem Genet 47:694–706

Spronk DB, Van der Schaaf ME, Cools R, De Bruijn ERA, Franke B, van Wel JHP, Ramaekers JG, Verkes RJ (2016) Acute effects of cocaine and cannabis on reversal learning as a function of COMT and DRD2 genotype. Psychopharmacology 233:199–211

Sulzer D (2011) How addictive drugs disrupt presynaptic dopamine neurotransmission. Neuron 69:628–649

Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ, von Mering C (2017) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res 45:D362–D368. https://doi.org/10.1093/nar/gkw937

van der Velde KJ, de Boer EN, van Diemen CC, Sikkema-Raddatz B, Abbott KM, Knopperts A, Franke L, Sijmons RH, de Koning TJ, Wijmenga C, Sinke RJ, Swertz MA (2017) GAVIN: gene-aware variant INterpretation for medical sequencing. Genome Biol 18:6. https://doi.org/10.1186/s13059-016-1141-7

Vandenbergh DJ, Persico AM, Hawkins AL, Griffin CA, Li X, Jabs EW, Uhl GR (1992) Human dopamine transporter gene (DAT1) maps to chromosome 5p15.3 and displays a VNTR. Genomics 14:1104–1106

Vasconcelos ACCG, Neto ESR, Pinto GR, Yoshioka FKN, Motta FJN, Vasconcelos DFP, Canalle R (2015) Association study of the SLC6A3 VNTR (DAT) and DRD2/ANKK1 Taq1A polymorphisms with alcohol dependence in a population from northeastern Brazil. Alcohol Clin Exp Res 39:205–211

Vereczkei A, Demetrovics Z, Szekely A, Sarkozy P, Antal P, Szilagyi A, Sasvari-Szekely M, Barta C (2013) Multivariate analysis of dopaminergic gene variants as risk factors of heroin dependence. PLoS One 8:e66592. https://doi.org/10.1371/journal.pone.0066592

Volkow ND, Wang GJ, Tomasi D, Baler RD (2013) Obesity and addiction: neurobiological overlaps. Obes Rev 14:2–18. https://doi.org/10.1111/j.1467-789X.2012.01031.x

Wetherill RR, Jagannathan K, Lohoff FW, Ehrman R, O'Brien CP, Childress AR, Franklin TR (2014) Neural correlates of attentional bias for smoking cues: modulation by variance in the dopamine transporter gene. Addict Biol 19:294–304

Xie P, Wang T, Yin G, Yan Y, Xiao L, Li Q, Bi K (2015) Metabonomic study of biochemical changes in human hair of heroin abusers by liquid chromatography coupled with ion trap-time of flight mass spectrometry. J Mol Neurosci 58:93–101

Yeh FC, Yang R-C, Boyle TB, Ye Z, Mao JX (1997) POPGENE, the user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Canada, p 10

Yeh YW, Lu RB, Tao PL, Shih MC, Lin WW, Huang SY (2010) Neither single-marker nor haplotype analyses support an association between the dopamine transporter gene and heroin dependence in Han Chinese. Genes, Brain Behav 9:638–647

Zhou Y, Michelhaugh SK, Schmidt CJ, Liu JS, Bannon MJ, Lin Z (2014) Ventral midbrain correlation between genetic variation and expression of the dopamine transporter gene in cocaine-abusing versus non-abusing subjects. Addict Biol 19:122–131