Title: A Dopamine Transporter Gene Functional Variant Associated with Cocaine Abuse.

Authors: †Camila Guindalini BSc[1,2], †Mark Howard PhD[3], Ronaldo Laranjeira MD PhD[4], David Collier PhD[1,5], Nik Ammar MSc[5], Ian Craig PhD[5], Colin O'Gara MRCPsych[1,6], Tiffany Greenwood PhD[7], John Kelsoe MD[7,8], Phil Asherson MD PhD[5], Robin M. Murray DSc FRCPPsych[1], Adauto Castelo MD PhD[4], John Quinn PhD[3], Homero Vallada MD PhD[2], and *Gerome Breen PhD[1,5].

1. Division of Psychological Medicine, Institute of Psychiatry, King's College London, UK.
2. Department of Psychiatry, University of São Paulo Medical School, Brazil;
3. Department of Physiology, Human Anatomy and Cell Biology, University of Liverpool, UK
4. UNIAD (Unit of Drug and Alcohol Research), Department of Psychiatry, Federal University of São Paulo, Brazil.
5. MRC Social Genetic and Developmental Psychiatry Research Centre, Institute of Psychiatry, King's College London, UK.
6. National Addiction Centre, Institute of Psychiatry, King's College London, UK
7.Department of Psychiatry, University of California San Diego, San Diego, CA, USA.
8. Department of Psychiatry, San Diego VA Healthcare System, La Jolla, CA, USA.

Correspondence to: Gerome Breen, Section of Genetics, Room 222 SGDP Centre, PO 81, MRC Social Genetic and Developmental Psychiatry Research Centre, Institute of Psychiatry, Kings College London. SE5 8AF.
Ph: +44-20-7848-0409. Fax: +44-20-78480575. E-mail: g.breen@iop.kcl.ac.uk

† These authors contributed equally to this work.

Abstract

Background: The dopamine transporter (DAT) mediates the active re-uptake of dopamine from the synapse and is a principal regulator of dopaminergic neurotransmission. Due to the complex pattern of variation and linkage disequilibrium in the DAT gene a more detailed examination of the locus and the identification of robust functional variants is essential for investigation of the role of this gene in disorders related to dopaminergic dysfunction.

Objective: To characterize the influence of functional genetic variants in the dopamine transporter on cocaine abuse.

Design: VNTRs polymorphisms were characterized and genotyped in a Brazilian two samples of current cocaine abusers from the metropolitan region of São Paulo (total n=695) and in a control sample of healthy individuals with no past history of drug abuse (total n=855). Functional effects of the variants on gene expression were analyzed using conventional and novel reporter gene constructs.

Results: A positive genotypic association was found between a 30bp VNTR in intron 8 of the DAT gene and cocaine abuse in both samples (combined Odds Ratio=1.35; 95%CI=1.10-1.67; p=0.0047). Haplotype analyses using other polymorphisms in and near the gene indicate that the Int8 VNTR is the variant responsible for the observed association. The functional analyses demonstrated that the risk allele (allele 3) mediates a differential and increased response (6 fold) to stimulus in reporter constructs in comparison with the other common and "protective" allele 2. ($p<<0.001$)

Discussion: We have demonstrated a robust association between cocaine abuse and a functional VNTR allele in the dopamine transporter, suggesting that this gene may play a role in the aetiology of cocaine abuse. However, the genetic association will require replication in cocaine abuse and examination in other drug abuse phenotypes to clarify the role of this variant in addictive processes.

Introduction

The dopamine transporter (DAT) mediates the active re-uptake of dopamine from the synapse and is a principal regulator of dopaminergic neurotransmission[1]. Dopamine is a key neurotransmitter in brain areas involved in movement and behaviour, particularly reward pathways. The gene encoding DAT (*SLC6A3*) consists of 15 exons spanning 60kb on chromosome 5p15.3[2]. A Pubmed search reveals over one hundred studies that assessed possible associations between variants in the dopamine transporter and psychiatric disorders (http://www.ncbi.nlm.nih.gov; for examples see 3-5). However the majority of these studies have examined only one or two polymorphisms in the gene and have thus only captured part of the genetic information contained within the gene. Recent analyses using a larger selection of markers have shown a complex pattern of variation and linkage disequilibrium in the DAT gene. They have underlined the necessity for more detailed examination of the locus, and the importance of identifying robust functional variants[5,6].

The most widely studied DAT polymorphism is a 40-bp variable number tandem repeat (VNTR) in the 3'untranslated region (UTR) of the gene with repeat copy numbers ranging from three to eleven[2]. Numerous studies have investigated association between alleles and/or genotypes of the 3'UTR-VNTR and clinical phenotypes thought to be related to dysregulation of dopamine transmission, such as attention deficit hyperactivity disorder[8],

Parkinson's disease[9], schizophrenia[10], cocaine-induced paranoia[11], severity of alcohol withdrawal[12] and methamphetamine-induced psychosis[13]. However, the functional effect of this polymorphism is uncertain as the data do not indicate a consistent effect of different alleles and genotypes on gene expression[4,14]. One explanation for these variable results is that the 3'UTR-VNTR is not the sole source of functional variation in the gene and is in linkage disequilibrium with nearby functional polymorphism(s). These may vary in their levels of linkage disequilibrium with the 3'UTR-VNTR in the different case-control and family populations studied.

Our examination of the DAT genomic sequence revealed the 3'UTR-VNTR is not unique. There are ~15 other candidate STRs (simple tandem repeats) and VNTRs in the introns of DAT with 6 or more repeat copies. While SNPs (single nucleotide polymorphisms) are markers of great utility in genetic studies, different alleles of a VNTR represent a much greater physical and chemical change to DNA sequence than different alleles of a SNP. We suggest that these VNTRs may contribute to function in the following ways:

1. They may operate as functional elements binding transcription factors and other proteins that inhibit or promote expression as shown for the serotonin transporter intron 2 VNTR[15] and the pentanucleotide repeat in the PIG3 gene[16].

2. They may affect the efficiency of mRNA splicing as shown for a CA repeat in the eNOS gene[17].

3. They may act as variable spacer elements between functional motifs.

For these reasons we regard VNTRs as very good *a priori* functional candidates. We decided to characterize VNTR polymorphisms in the DAT gene by first investigating a possible relationship between these markers and susceptibility for a specific phenotype and then, secondly, by examination of the functional activity that could be responsible for the development of the disorder.

Initially, we selected for study the 3'UTR-VNTR and a 6 –copy 30bp VNTR located in intron 8 as it lies in a region suggested by Greenwood *et al*[6] to be in weak/variable LD with both the 3' VNTR region and the 5' region including the promoter. This VNTR exists as just one (perfect) copy of the motif in the Chimpanzee genome and is mildly conserved in rat and mouse where this region of intron 8 contains large, presumably polymorphic, GA and CA repeats, respectively (http://genome.ucsc.edu).

These markers were then genotyped in two samples of cocaine abusers from Sao Paulo, Brazil, initially a Caucasian sample, and then a replication sample of mixed African-Caucasian ethnicity. Cocaine binds neurotransmitter transporters thus inhibiting reuptake. Its addictive effects are thought to be principally mediated through blockage of the dopamine transporter, increasing substantially the concentration of extracellular dopamine. This results in elevated stimulation of neurons in brain regions related to reward

and reinforcement behaviour, thus explaining cocaine's pleasurable and addictive effects[18].

Family and twin studies suggest a substantial genetic component in the vulnerability of individuals to become dependent after exposure to cocaine[19-21]. Thus we hypothesize that functional genetic variation in the DAT gene could contribute to an increased susceptibility for cocaine abuse and, consequently, dependence. Our aim was then to confirm associations observed by demonstrating a functional effect of identified risk alleles. Here we report that the 30bp VNTR located in intron 8 of the DAT gene is associated with cocaine abuse in the Brazilian population and further that the risk allele has a differential effect on reporter gene expression.

Methods

Association Study

Subjects

Six hundred and ninety-five cocaine abusers were ascertained in the state of São Paulo, Brazil.[22,23] The study group consisted of drug users who were in treatment from August 1997 to October 1998 in six inpatient and one outpatient clinic located in the metropolitan region of São Paulo City with a catchment of 16.7 million people. These drug treatment centres covered an estimated 75% of all hospitalizations related to drug abuse or illegal drug dependence carried out during the study period (Ministry of Health).

Inclusion criteria were: age 18 years and older, a history of cocaine abuse and under drug treatment at the selected centres. Individuals with a mental or physical disability were excluded. All current cocaine users were then interviewed using a structured interview to collect data on sociodemographic characteristics, sexual behaviours, and drug use profile. All subjects satisfied an ICD10 diagnosis of cocaine dependence[24]. Blood samples were collected from all participants for genetic and other analyses. A total of 64.4% of the participants reported having smoked cocaine (crack) over the previous month and 51.1% had snorted cocaine over the same period. The overall lifetime prevalence for any type of injected drug use (cocaine, heroin and other opiates, amphetamines, and sedatives) was 17.9%.[22]

Controls

Eight hundred and fifty-five healthy controls were recruited from the Blood Transfusion Unit of the Hospital das Clinicas, Faculty of Medicine, University of São Paulo. Each blood donor was screened using a short questionnaire investigating contagious diseases and the use of any kind of drug. Subjects with a past history of drug abuse or with recent use of an illegal drug were excluded. During the act of donation a short interview was conducted and subjects with a lifetime history of a psychiatric disorder requiring admission to hospital or suffering from a psychiatric condition at time of interview were excluded.

Other Details

Ethnic classifications were determined by interviewer monitored self-report. Nine hundred and thirty-two individuals (500 cases, 432 controls; mean age 26.86, SD $\pm$ 7.39 and 95% male; mean age 35.68, SD $\pm$ 11.2 and 62% male, for cases and controls respectively) were classified as Brazilian Caucasian – that is of mixed European background (Italian, Portuguese, German) . Six hundred eighteen individuals were classified as a Mixed Brazilian African/Caucasian group includes (195 cases, 423 controls; mean age 26.24, SD $\pm$ 6.6 and 81%; mean age 36.19, SD $\pm$ 10.6 and 74% male for cases and controls respectively). This group was of mixed European and African background. During the ascertainment of controls, care was taken to ensure the same ethnic classification was used, and that the sampling was from the

same population groups in São Paulo. All the subjects included in this study gave written informed consent and this project was approved by the Ethical Committee of the Federal University of São Paulo and other relevant local ethics committees.

Polymorphisms

In addition to the initial intron 8 VNTR and the 3'UTR-VNTR, we then also selected 2 SNPs, in intron 1 and intron 8 respectively, and a VNTR 10 kb 3' of the DAT gene to help estimate haplotypes and examine those regions for association.

VNTRs

In silico analyses of publicly available human DAT sequence (http://genome.ucsc.edu, July 2003 release, NCBI v.31) revealed a large number of genomic VNTRs in introns. The 40-bp 3' untranslated region VNTR (designated here as 3'*UTR*), and a VNTR in intron 8 (*Int8*) of the gene were selected for the investigation.

Afterwards we selected a novel VNTR (*37bp VNTR*) located ~10kb 3' to the 3'UTR-VNTR. The 37 bp VNTR is described as a 6.1 copies of a 37 bp repetitive element polymorphism and the Int8 as a 6.1 copies of a 30bp element in the Simple Repeat table of the Human Genome in the UCSC "Golden Path" database (http://genome.ucsc.edu).

The three markers were amplified by PCR (polymerase chain reaction) using the following conditions: 5' denaturation at 95 °C; then 35 cycles of 1

min at 95°C, 1 min at 60°C and 1 min at 72°C. The oligonucleotide primers used were:

Int8-VNTRF: 5′-CTTGGGGAAGGAAGGG-3′ and
Int8-VNTR: 5′-TGTGTGCGTGCATGTGG-3′;
3′UTR-VNTR-F: 5′-TGGCACGCACCTGAGAG-3′ and
3′UTR-VNTR-R: 5′-GGCATTGGAGGATGGGG-3′;
37bp-VNTR-F: 5′-CTTGCTTCCGCAGAGCTAAT-3′ and
37bp-VNTR-R: 5′-GGCACTTCCTGAGAGCAGAG-3′.

Amplification products were separated on 3% agarose gel containing ethidium bromide and visualized under UV light. Alleles of the Int8 and 37bp VNTRs were called according to their relative size (smallest = allele 1) whereas the coding system utilising the range from 3 to 11 copies was applied for the 40bp VNTR (3′UTR VNTR).

SNPIn a previous study analysing variation within DAT, Greenwood and colleagues[6] selected a number of SNPs that span the gene from the distal promoter through the 3′ UTR that are suitable for linkage disequilibrium analyses. We decided to genotype two of these SNPs **I1+1036** (rs2963238; C/A) and **I8+ 2086** (rs27048; A/G), in intron 1 and intron 8 respectively, and investigate patterns of LD between these markers and the VNTRs. All genotyping was performed blind to status using an amplifluor assay, and was performed under contract by K-Biosciences (Cambridge, UK; http://www.kbioscience.co.uk/).

Statistical Analyses

Genotype and allele frequencies were compared using a χ^2 test, and p-values were assessed using simulation (CLUMP software[25] v.2.2). GENECOUNTING software v.2.0[26,27] was used to estimate haplotype frequencies. Odds ratios and 95% confidence intervals were calculated using Woolf's method (1955)[28], implemented in Microsoft Excel (Microsoft Corporation, CA).

Comparison of Allele Frequencies with CEPH Diversity Panel Data

The Allele Frequency Database (ALFRED[29]; http://alfred.med.yale.edu), was used to access frequencies of intron 8 alleles in populations thought to be of most direct relevance to the current study: (1) the African-American panel and (2) the Europeans Mixed panel of US Caucasians for comparison with the observed allele frequencies in the Brazilian Caucasian and Mixed African/Caucasian groups.

Expression Analyses

Cell Growth

SN4741 is a mouse substantia nigra derived dopaminergic neuronal cell line and was cultured at 33°C, 5% CO2 in D-MEM with Glutamax supplemented with 10% FCS, 1% glucose, and penicillin-streptomycin. Post transfection, cells were incubated in the same conditions[30].

Construction of reporter gene constructs

The plasmids p5′ DAT2, p5′ DAT3, p3′ DAT2 and p3′DAT3 contained alleles of the DAT VNTR found in intron8 of the DAT gene. Primers were designed to amplify the variable region in individuals homozygous for either allele 2 or 3. Forward 5′-GCTTGGGGAAGGAAGGG-3′ and Reverse 5′-TGTGTGCGTGCATGTGG-3′. The PCR product for each fragment was gel purified using the Qiaquick gel extraction kit (Qiagen), and then cloned into the TA cloning plasmid pcRII-TOPO (Invitrogen), validity and directionality being confirmed by DNA sequencing. For p5′ DAT2 and p5′ DAT3 the fragment was then excised by *Acc*65I and *Xho*I and ligated into the MCS of pGL3-P (Promega) introducing the fragment upstream of the start codon of the luciferase gene. For p3′ DAT2 and p3′ DAT3 the fragment was excised with *Bam*HI and **Xho**I and ligated into the *Bam*HI and *Sal*I site of pGL3-P introducing the fragment downstream of the stop codon of the luciferase gene.

To introduce the two common Int8 allelic variants into the intron of the renilla vector phRLsv40 (Promega), we utilised a plasmid supplied by Dr J. Bubb containing an *Asc*I linker within its intronic region. The primers were modified: Forward 5′-TTGGCGCGCCGCTTGGGGAAGGAAGGG-3′ and Reverse 5′-TTGGCGCGCCGTGTGCGTGCATGTGG-3′ to include Asc1 restriction sites enabling direct ligation of the PCR product into phRLsv40. All plasmids were confirmed for validity and directionality by DNA Sequencing.

Transfection

~1x10^4 cells were plated onto 24 well plates 24 hours prior to transfection with Transfast (Promega). Transfections were optimised according to manufacturers' instructions. In brief, reporter plasmid (0.5μg per well) was mixed with serum free cell media (200μl per well) and Transfast was added at a ratio of 2:1 (3μl per well). Cells were then washed twice with PBS, DNA media mixture added for 1 hour, after which 1 ml of relevant cell media containing 10%FBS was added

Reporter gene assay

After 48hour cells were washed twice with PBS, then lysed with Passive Lysis buffer (Promega). After 15 min agitation at room temperature, the cell lysate was centrifuged briefly at 10000g. Supernatants were assayed for reporter gene expression by using the relevant Promega assay system. Supernatants (20µl) was added to 100µl of assay reagent in opaque 96 well plates and the light emission was measured over a given time interval with the Life Sciences Labsystems Luminoskan, model RT. Cells where appropriate were serum deprived, exposed to a potassium evoked depolarisation or a potassium evoked depolarisation combined with Forskolin. Forskolin was used at 10µM final concentration. Depolarsation of SN4741 cells was achieved using final concentrations of 10mM $CaCl_2$ and 40mM KCl. ©Luciferase/renilla results were normalized to total protein concentration, which was measured with the BCA protein assay kit (Pierce) in accordance with the manufacturers instructions. Results are means +/- S.D. of three experiments performed in triplicate using cells of the same or similar passage number.

Results

Genotyping

The analysis was conducted dividing the samples in two groups based on individual's ethnic background: Brazilian Caucasian and Mixed Brazilian African/Caucasian.

The genotype counts and frequencies are indicated in Table 1a, b ,c for the Int8 VNTR, the 3'UTR VNTR, and 37bp VNTR respectively. No significant case-control difference was observed for either the 3'UTR or 37bp VNTRs. However consistent association in both ethnic groups was found with the Int8 VNTR. Evaluating the genotype distribution we observed that the 33 genotype is distinctly increased in the case samples from both ethnic groups whereas the 23 genotype was noticeably decreased.

Subsequently, allele wise odds ratios were calculated by excluding the rare alleles and only considering alleles 2 and 3 of Int8 (as they were >90% of the total alleles observed) using Woolf's method[28]. Genotypic odds ratios were calculated in a similar manner, using a recessive model for the 3 allele and by using 23 as a reference. Using Woolf's method also allowed us to meta-analyse the data to give odds ratios and p-values while allowing for the differing allele and genotype frequencies between the Caucasian and Mixed African/Caucasian samples.

Table 2 summarizes the results. Allele wise analyses did not show any statistical differences between cases and controls in the different groups.

However the recessive model for the int8 VNTR allele 3 showed a significant increase frequency of the 33 genotype in the case sample for both ethnic groups. Moreover when 23 was used as a reference the association became even more evident with an Odds Ratio (OR) =1.40 and 95% Confidence Interval (CI): 1.06-1.83 (p=0.017) for the Caucasian group and an OR=1.68, 95%CI: 1.15-2.45 (p=0.007) for the African/Caucasian sample.

The Meta Analyses finally demonstrated that the 33 genotype is associated with cocaine abuse in the Brazilian sample (for 33 versus 22 and 23: OR=1.35, 95%CI: 1.10-1.67, p=0.0047; for 33 versus 23: OR=1.49, 95%CI: 1.19-1.85; p=0.0004). The *p*-values generated by CLUMP[25] (T1, T4 using 10,000 simulations) for analysis of genotypic and allele wise distributions are also included in table 2 for comparison with the odds ratio analysis.

Genotypes and allele frequencies for the I1+1036 and I8+ 2086 SNPs did not show a significant difference between cases and controls in both groups (data not shown). In addition, both control groups and both case group were in Hardy Weinberg Equilibrium (p>0.05) for the markers tested (as assessed by the HWE program)[31]. Additional clinical variables such as type of cocaine used, age of onset or sex did not show an association/interaction with the markers genotyped (data not shown).

Table 1(a). Int8-VNTR Genotype Counts (%)									
	12	13	22	23	25	33	34	35	Total
Caucasian									
Controls	0 (0)	2 (0.5)	28 (6.5)	193 (44.7)	1 (0.2)	201 (46.5)	0 (0)	7 (1.6)	432
Cases	0 (0)	0 (0)	40 (8)	183 (36.6)	1 (0.2)	266 (53.2)	5 (1)	5 (1)	500
African/Caucasian									
Cases	0 (0)	3 (0.7)	48 (11.3)	213 (50.4)	2 (0.5)	152 (35.9)	1 (0.2)	4 (0.9)	423
Case	1 (0.5)	1 (0.5)	31 (15.9)	71 (36.4)	2 (1)	85 (43.6)	1 (0.5)	3 (1.5)	195
Table 1(b) 3'UTR VNTR Genotype Counts (%)									
	14	34	35	44	45	46	55	56	Total
Caucasian									
Controls	0 (0)	1 (0.2)	0 (0)	38 (8.8)	150 (34.9)	3 (0.7)	227 (52.8)	4 (0.9)	430
Cases	0 (0)	2 (0.4)	6 (1.2)	45 (9)	183 (36.6)	2 (0.4)	245 (49)	9 (1.8)	500
African/Caucasian									
Controls	2 (0.5)	5 (1.2)	5 (1.2)	46 (10.9)	126 (29.9)	2 (0.5)	217 (51.4)	6 (1.4)	422
Case	2 (1)	1 (0.5)	3 (1.5)	14 (7.2)	67 (34.5)	3 (1.5)	93 (47.9)	9 (4.6)	194
Table 1(c) 37bp VNTR Genotype Counts (%)									
	12	13	22	23	24	33	34		Total
Caucasian									
Controls	3 (0.7)	2 (0.5)	178 (41.9)	194 (45.6)	5 (1.2)	42 (9.9)	1 (0.2)		425
Cases	5 (1)	0 (0)	240 (48.2)	204 (41)	3 (0.6)	43 (8.6)	3 (0.6)		498
African/Caucasian									
Controls	8 (1.9)	5 (1.2)	178 (42.8)	172 (41.3)	3 (0.7)	48 (11.5)	2 (0.5)		416
Case	3 (1.5)	1 (0.5)	75 (38.7)	92 (47.4)	3 (1.5)	20 (10.3)	0 (0)		194

Table 1: Genotype Number and Frequency (%) of the 3 VNTRs in Healthy Controls and Cocaine Abusers in the Brazilian Caucasian and Mixed African/Caucasian Samples. In Table 1b only the genotype that were represented more than 1% in at least one of the samples are shown. All genotype distributions were in Hardy Weinberg equilibrium.

Table2: Genotype p-values, Odds Ratio analysis of Int8-VNTR alleles, genotypes with 95% Confidence Intervals and p-values.

Sample	Model	Odds Ratio	95% CI	p-value (2-tailed)	Genotype Wise p-values	
Meta-analysis using method of Woolf (1955)	**33 vs (23+22)**	1.35	1.10 - 1.67	0.0047		
	33 vs 23	1.49	1.19 - 1.85	0.0004		
	3 vs 2 allele-wise	1.12	0.95 - 1.31	0.18		
					Clump t1 p-value*	**Clump t4 p-value***
Sample 1 - Brazilian Caucasian	**33 vs (23+22)**	1.31	1.01 - 1.70	0.042	0.017	0.043
	33 vs 23	1.40	1.06 - 1.83	0.017		
	3 vs 2 allele-wise	1.14	0.93 - 1.40	0.216		
Sample 2 - Caucasian-African	**33 vs (23+22)**	1.43	1.01 - 2.03	0.045	0.047	0.019
	33 vs 23	1.68	1.15 - 2.45	0.007	*considering entire distribution	
	3 vs 2 allele-wise	1.08	0.84 - 1.40	0.539		

Table2: Odds Ratio analysis of Int8-VNTR alleles and genotypes with 95% Confidence Intervals and p-values.

A recessive model for allele 3 and a model with the 23 genotype as reference were used in addition to allele wise. Results of Meta-Analyses of both populations using Woolf's method are shown. Genotype wise p-values derived from analysis of data presented in Table 1a using CLUMP are also shown, for comparison.

Comparison of Data with CEPH Diversity Panel Data

In an attempt to verify the influence of genetic stratification in our sample, we decided to compare the allele frequencies of the intron8 VNTR between the Brazilian Caucasian group and Mixed African/Caucasian group with populations from the CEPH diversity panels using the ALFRED database[29] namely:

1. "Europeans, Mixed" - 44 samples collected by K. Kidd from unrelated individuals of European ancestry living in the United States and Canada.

2. "African-American" - This sample is of 91 cell lines (182 chromosomes) from the NIGMS Cell Repository, Corielle Institute for Medical Research and was collected throughout the United States of America. These represent a broad cross-section of African Americans.

Figure 1 summarizes the results of the comparison. The frequency of the alleles in the Brazilian Caucasian group showed to be very similar to the Europeans Mixed population (70% versus 71% for the allele 3, respectively). The observed allelic distribution among the Brazilian Mixed African/Caucasian group was intermediate between the values for "Europeans, Mixed" and African American. It should be noted that our association is primarily genotypic and not allelic in nature and that comparison of genotype frequencies yields similar results for the control groups but less marked similarities for the cases.

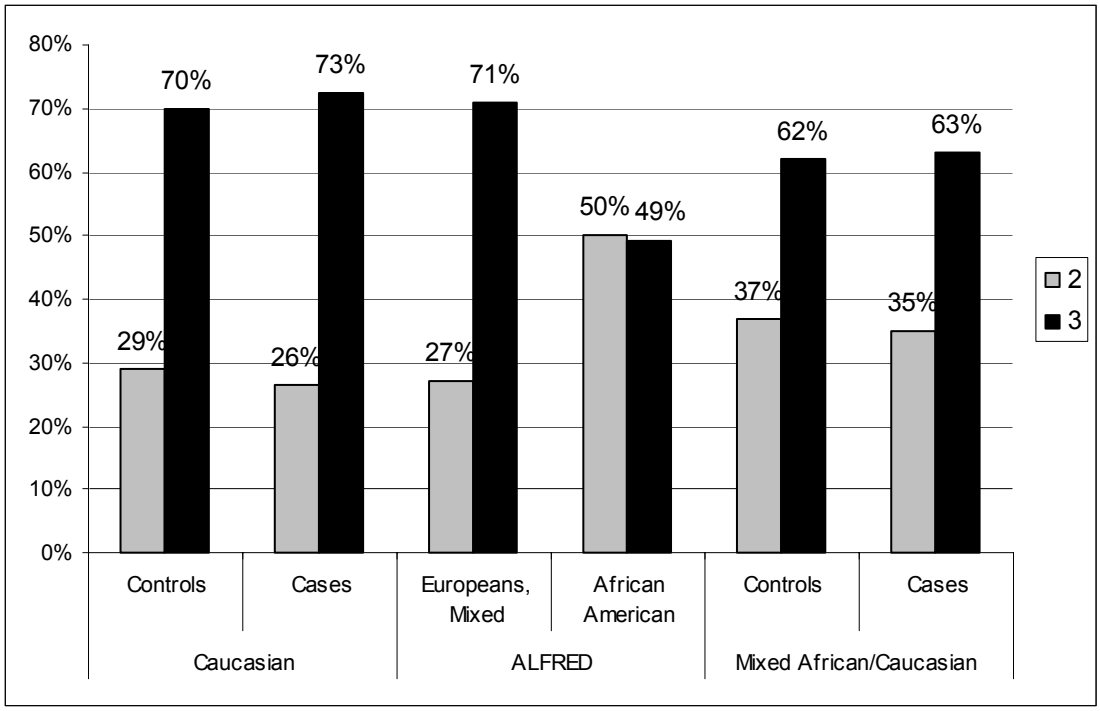

Figure 1. Frequency of Int8-VNTR allele 2 and 3 in cases, controls and reference populations.

Case and control allele frequencies (in % terms) for both samples. For comparison, the allelic distribution in two populations from the ALFRED database, "Europeans, Mixed" and African American, are also shown.

Haplotype and Linkage Disequilibrium Analysis

We analysed haplotypes for all markers involved using Genecounting[26] and Genecounting utilities[27] (data not shown) and did not find significant haplotypic association after permutation analysis. In addition pair-wise LD was calculated using LDpairs program from GC utilities[27] and is given in Table 3 in terms of r (the Cramer's V statistic).

Table3: Pairwise LD Values - Values of r (Cramer's V)

A. Brazilian Caucasian				
	SNP-Int8	Int8-VNTR	3'UTR-VNTR	37bp VNTR
SNP-Int1	0.158	0.146	0.107	0.108
SNP-Int8		0.398	0.147	0.075
Int8-VNTR			0.269	0.146
3'UTR-VNTR				0.236

B. Brazilian African Caucasian				
	SNP-Int8	Int8-VNTR	3'UTR-VNTR	37bp VNTR
SNP-Int1	0.094	0.122	0.192	0.159
SNP-Int8		0.513	0.155	0.074
Int8-VNTR			0.174	0.153
3'UTR-VNTR				0.198

Functional Results

We tested the function of both 3' and intron 8 VNTRs separately in conventional reporter gene constructs (Promega). The intron 8 VNTR common alleles had no activity in a reporter gene assay when cloned in the 3' location; however when the intron 8 VNTR alleles were cloned in the region 5' to the reporter gene both the 2 and 3 copy demonstrated a repressor function, Figure 2. The 3'UTR VNTR was cloned in the 3′ location and displayed no activity (data not shown) although when cloned in the 5′ position it demonstrated enhancer function[14]. Since the location of both the intron 8 and 3' UTR VNTR demonstrated differential function we cloned the intron 8 domain in a more relevant orientation, using a modified renillia vector with a polyclonal site within an intron for testing such intronic VNTR domains as functional units. The intron element when cloned in this domain demonstrated a small but statistically significant increase ($p<0.05$) in reporter gene expression for the 2 allele of the Int8 VNTR.

We then tested the response of the Int8-VNTR-2 and Int8-VNTR-3 alleles to synergistic challenge. Pairing of increased Ca^{2+} and cAMP signals results in robust, synergistic, activation of CRE-mediated transcription. Specifically, the co-administration of Forskolin (10μM) and K^+ (40 mM) triggered a ~6-fold potentiation of reporter expression (see Figure 3). Forskolin (1 - 10 mM) is a membrane permeable activator of adenylyl cyclase. Increasing extracellular potassium will depolarise the membrane

potential of neurons (or any cell). In cultured cells increasing extracellular K+ results leads to an increase in cytosolic Ca^{2+} and results in enhanced (Cre dependant) transcription. Furthermore, the co-activation of Ca^{2+} and cAMP pathways results in a robust synergistic activation of transcription. In this system the Int8-VNTR-3 allele demonstrated a large increase (6-fold) in reporter gene expression in response to synergistic challenge ($p<<0.001$). We have demonstrated that these domains can support differential expression to specific challenges and that Int8 VNTR alleles can differentially mediate response.

Figure.2 Expression of DAT Int8-VNTR alleles in expression constructs

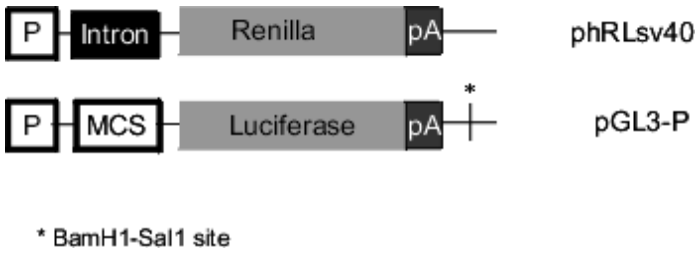

Figure 2(a) A representation of the two reporter plasmids utilized.

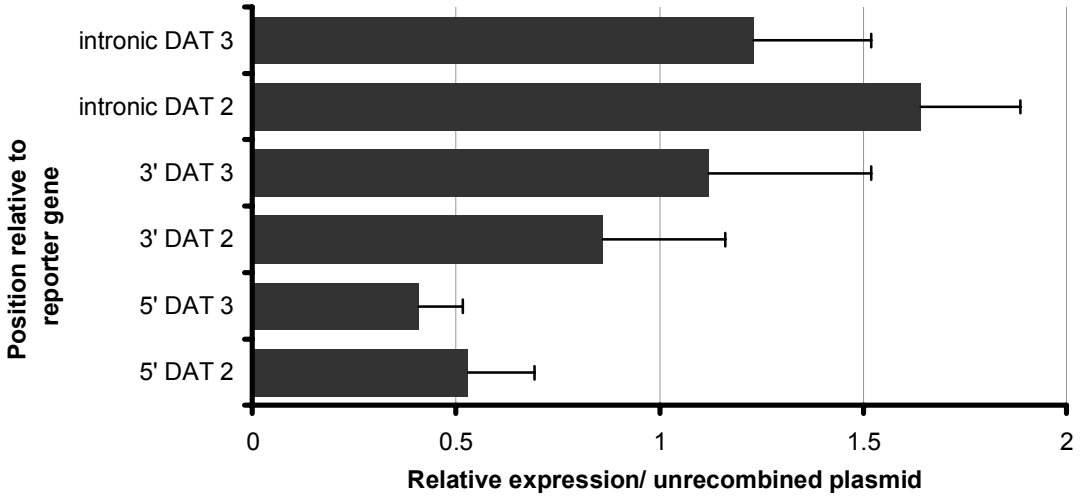

Figure 2(b) Relative expression of DAT Int8-VNTR allele is different reporter construct orientations.

Data is presented using unrecombined plasmid as a control in each experiment, the relative expression over the unrecombined plasmid yields a fold increase measure of activity. Cells were left for 48hours post transfection before assay, each experiment was performed three times and in triplicate.

Figure 3. Potassium induced depolarisation (K+IP) and forskolin/K+IP stimulate differential responses in gene expression between the two Int8-VNTR alleles.

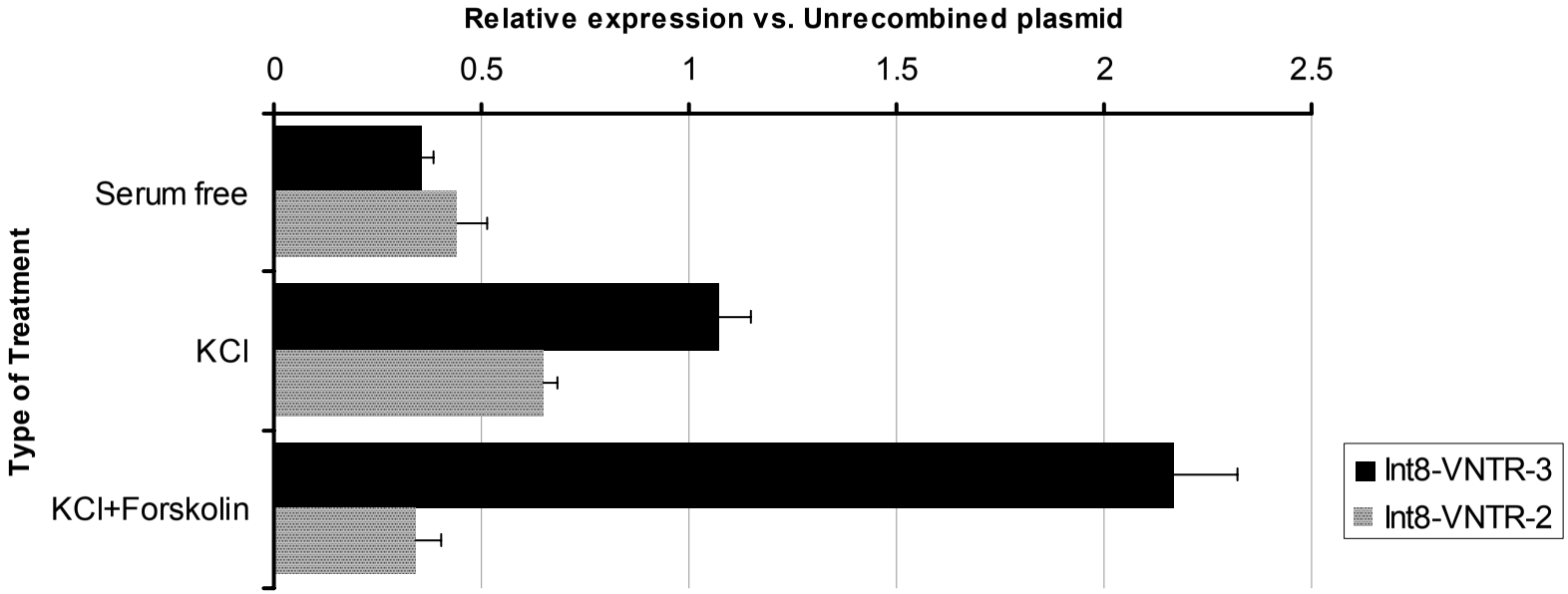

Figure 3. Potassium induced depolarisation (K^+IP) and Forskolin/K^+IP stimulate differential responses in gene expression between the two Int8-VNTR alleles. Data are from experiments performed in triplicate. Forskolin was added to 10μM, and K^+IP was induced using 10mM $CaCl_2$ and 40mM KCl. Data is presented using relative expression over the unrecombined plasmid yielding a fold increase measure of renilla activity.

Discussion

We examined a total of 695 cocaine abusers and 855 controls and identified a positive association with an allele of the 30bp VNTR in intron 8 of the dopamine transporter gene, the Int8 VNTR. Association and haplotype analyses using other polymorphisms in and near the gene (the 3'UTR VNTR, a novel 37bp VNTR and SNPS rs2963238 and rs27048) indicate that the Int8 VNTR is the variant responsible for the observed association, despite significant linkage disequilibrium ($r > 0.4$) existing between it and some of the other variants. Moreover, functional analyses strongly suggest that the risk allele 3 mediates a differential and increased response to stress in reporter constructs in comparison with the other common and "protective" allele 2.

However, the main direction of association observed is genotypic not allelic, suggesting that both alleles have a role to play in regulating expression. Notably the Int8-VNTR genotype was as strongly protective as the 33 genotype was risk-causing. This suggests that if indeed the 3 allele acts to increase risk of cocaine abuse, it does so in a recessive manner. The lack of allele-wise association could be due to this, and a much larger sample may be needed to detect the separate effect of each allele. Unusually, the 22 genotype was neither protective nor significantly risk causing in either sample. This may be due to the smaller number of observations or, alternately, a complex interaction between the two alleles at this marker in regulating gene expression.

Our sample was divided in two ethnic groups for analysis, sample I of patients and controls derived from mixed European ancestry (Italian, Portuguese, German) and another of mixed European African heritage, with association being observed in both groups. Comparison of allele frequencies in our samples with reference populations indicates this division is consistent with observed allele frequencies. Nevertheless, the possibility of stratification means these results should be interpreted cautiously. However, replication in both samples of the same genotypic pattern of association adds weight to the observed association. In addition, our functional evidence shows that the risk allele 3 exhibits a 6-fold differential response to stimulus in comparison to the protective or 2 allele, and is strongly supportive of the association finding.

In our functional work we found good evidence of a differential effect on reporter gene expression of the Int8-VNTR 3 allele when placed in the intron of an expression vector. The data indicates that in a normal or resting state the 3 allele confers slightly less expression of the reporter construct that the 2 allele. However, when challenged with the addition of KCl and Forskolin to the cell culture medium the 3 allele causes a large upregulation of reporter gene expression (6x normal), while the 2 allele does not exhibit any change. The use of K^+ induced depolarisation and Forskolin addition is a well-practised method to "stress" cells and induce responses in gene expression [32-34] . Furthermore, this was observed using a mouse dopaminergic cell line and in 3 replicate experiments.

It is notable that this is the first large scale study of the role of dopamine transporter gene variants in Cocaine abuse both in terms of the sample size examined and the number of polymorphisms genotyped. There have been relatively few genetic association studies of cocaine abuse, only one of which[11] looked at the role of a DAT polymorphism in a small (by more recent standards) sample of cocaine abusers (n=102) through genotyping the 3'UTR-VNTR. Other molecular genetic studies that have examined cocaine abuse as the primary phenotype have principally focussed on the serotonin transporter[35-37], dopamine beta hydroxylase[38] and dopamine receptors[39]. In the previous DAT association study, Gelernter *et al*[11] failed to find an association with cocaine abuse but did suggest a role for DAT 3' UTR-VNTR alleles in cocaine induced paranoia.

The effect of the Int8-VNTR 33 genotype on dopamine transporter expression in the brain is unknown. Given the stress response characteristics of the reporter gene containing the Int8 VNTR in reporter constructs, and the neurochemical and neurotoxic effects of cocaine we hypothesise that people possessing the 33 genotype will exhibit a differential response in terms of DAT gene expression when exposed to cocaine. If, as our data suggest, expression may rise rapidly upon stimulus, this could leave an individual in a dopamine depleted state in the period after an initial dose of cocaine. Conceivably, this could contribute to both the described salience and sensitisation responses observed in abusers[40].

To summarise we have found a robust association between cocaine abuse and a functional VNTR allele in the dopamine transporter. While a false positive finding due to population stratification is a potential concern for this type of study, this was addressed by the implementation of a replication strategy, with confirmation of the association in the two population groups, and also the demonstration of a functional effect for the risk allele. In addition, haplotype analyses in the sample support the idea that the Int8-VNTR is the relevant functional motif in this sample as no other markers and common marker-marker haplotypes exhibit association. However, both the functional and genetic work will require replication in Cocaine abuse. Additional studies of other drug abuse phenotypes than cocaine abuse will be necessary to explore the role of this polymorphism in addictive processes.

Acknowledgements

Gerome Breen is an MRC Bioinformatics Training Fellow. John Kelsoe is a founder of, and holds equity in, a company - Psynomics, Inc. This work was supported by FAPESP (Fundação para o Amparo à Pesquisa do Estado de São Paulo), Grant number 99/04678-2.

References

1. Giros B, Caron MG. Molecular characterization of the dopamine transporter. *Trends Pharmacol Sci*. 1993; 14(2):43-9.

2. Vandenbergh DJ, Persico AM, Hawkins AL, Griffin CA, Li X, Jabs EW, Uhl GR. Human dopamine transporter gene (DAT1) maps to chromosome 5p15.3 and displays a VNTR. *Genomics.* 1992; 14(4):1104-6.

3. Kirley A, Lowe N, Hawi Z, Mullins C, Daly G, Waldman I, McCarron M, O'Donnell D, Fitzgerald M, Gill M. Association of the 480 bp DAT1 allele with methylphenidate response in a sample of Irish children with ADHD. *Am J Med Genet*. 2003; 121B (1):50-4.

4. Martinez D, Gelernter J, Abi-Dargham A, van Dyck CH, Kegeles L, Innis RB, Laruelle M. The variable number of tandem repeats polymorphism of the dopamine transporter gene is not associated with significant change in dopamine transporter phenotype in humans. *Neuropsychopharmacology*. 2001; 24(5):553-60.

5. Ueno S, Nakamura M, Mikami M, Kondoh K, Ishiguro H, Arinami T, Komiyama T, Mitsushio H, Sano A, Tanabe H. Identification of a novel polymorphism of the human dopamine transporter (DAT1) gene and the significant association with alcoholism. *Mol Psychiatry*. 1999; 4(6):552-7.

6. Greenwood TA, Alexander M, Keck PE, McElroy S, Sadovnick AD, Remick RA, Shaw SH, Kelsoe JR. Segmental linkage disequilibrium within the dopamine transporter gene. *Mol Psychiatry.* 2002; 7(2):165-73.

7. Greenwood TA, Kelsoe JR. Promoter and intronic variants affect the transcriptional regulation of the human dopamine transporter gene. *Genomics.* 2003; 82(5):511-20.

8. Chen CK, Chen SL, Mill J, Huang YS, Lin SK, Curran S, Purcell S, Sham P, Asherson P. The dopamine transporter gene is associated with attention deficit hyperactivity disorder in a Taiwanese sample. *Mol Psychiatry*. 2003; 8 (4):393-6.

9. Le Couteur DG, Leighton PW, McCann SJ, Pond S. Association of a polymorphism in the dopamine-transporter gene with Parkinson's disease. *Mov Disord*. 1997; 12(5):760-3.

10. Persico, AM and Macciardi, F. Genotypic association between dopamine transporter gene polymorphisms and schizophrenia. *Am J Med Genet, Neuropsychiatr Genet.* 1997; 74: 53–57.

11. Gelernter, J, Kranzler, HR, Satel, SL and Rao, PA. Genetic association between dopamine transporter protein alleles and cocaine-induced paranoia. *Neuropsychopharmacology.* 1994; 11: 195–200.

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

13. Ujike H, Harano M, Inada T, Yamada M, Komiyama T, Sekine Y, Sora I, Iyo M, Katsu T, Nomura A, Nakata K, Ozaki N. Nine- or fewer repeat alleles

in VNTR polymorphism of the dopamine transporter gene is a strong risk factor for prolonged methamphetamine psychosis. *Pharmacogenomics J.* 2003; 3(4):242-7.

14. Michelhaugh SK, Fiskerstrand C, Lovejoy E, Bannon MJ, Quinn JP. The dopamine transporter gene (SLC6A3) variable number of tandem repeats domain enhances transcription in dopamine neurons. *J Neurochem.* 2001; 79(5):1033-8.

15. K.P. Lesch, U. Balling, J. Gross, K. Strauss, B.L. Wolozin, D.L. Murphy and P. Riederer. Organization of the human serotonin transporter gene. *J. Neural Transm. Gen.* 1994; Sect (95): 157–162.

16. Contente A, Dittmer A, Koch MC, Roth J, Dobbelstein M. A polymorphic microsatellite that mediates induction of PIG3 by p53. *Nat Genet*. 2002; 30(3):315-20.

17. Hui J, Reither G, Bindereif A. Novel functional role of CA repeats and hnRNP L in RNA stability. *RNA*. 2003; 9(8):931-6.

18. Ritz MC, Lamb RJ, Goldberg SR, Kuhar MJ. Cocaine receptors on dopamine transporters are related to self-administration of cocaine. *Science*. 1987; 237(4819):1219-23.

19. Merikangas KR, Stolar M, Stevens DE, Goulet J, Preisig MA, Fenton B, Zhang H, O'Malley SS, Rounsaville BJ. Familial transmission of substance use disorders. *Arch Gen Psychiatry*. 1998; 55(11):973-9.

20. Bierut LJ, Dinwiddie SH, Begleiter H, Crowe RR, Hesselbrock V, Nurnberger JI Jr, Porjesz B, Schuckit MA, Reich T. Familial transmission of substance dependence: alcohol, marijuana, cocaine, and habitual smoking: a report from the Collaborative Study on the Genetics of Alcoholism. *Arch Gen Psychiatry*. 1998; 55(11):982-8.

21. Kendler KS, Prescott CA. Cocaine use, abuse and dependence in a population-based sample of female twins. *Br J Psychiatry*. 1998; 173:345-50.

22. Turchi MD, Diaz RS; Martelli CM Turchi; Sabino EC; da Silva WP; Filho OF; Laranjeira RR; Busch MP; Castelo A. *Genetic Diversity and HIV-1 Incidence Estimation Among Cocaine Users in Sao Paulo, Brazil*. JAIDS Journal of Acquired Immune Deficiency Syndromes. 30(5):527-532, August 15, 2002.

23. Messas G, Meira-Lima I; Turchi M; Franco O; Guindalini C, Castelo A, Laranjeira R, Vallada H. Association Study Of Dopamine D2 And D3 Receptor Gene Polymorphisms With Cocaine Dependence. *Psychiatric Genetics* (in press)

24. WHO: The ICD10 classification of mental and behavioural disorders. Diagnostic criteria for research (Geneva: World Health Organization) 1993:.

25. Sham PC, Curtis D. Monte Carlo tests for associations between disease and alleles at highly polymorphic loci. *Ann Hum Genet*. 1995; 59 (Pt 1):97-105.

26. Zhao JH, Lissarrague S, Essioux L, Sham PC: GENECOUNTING: haplotype analysis with missing genotypes: *Bioinformatics.* 2002; 18(12): 1694-5.

27. Zhao, JH. 2LD, GENECOUNTING and HAP: computer programs for linkage disequilibrium analysis. *Bioinformatics.* 2004; 20(8):1325-6.

28. Woolf B: On estimating the relation between blood groups and disease: *Ann Hum Genet*. 1955; 19: 251-253.

29. Howard M. Cann, Claudia de Toma, Lucien Cazes, Marie-Fernande Legrand, Valerie Morel, Laurence Piouffre, Julia Bodmer, Walter F. Bodmer, Batsheva Bonne-Tamir, Anne Cambon-Thomsen, Zhu Chen, Jiayou Chu, Carlo Carcassi, Licinio Contu, Ruofu Du, Laurent Excoffier, Jonathan S. Friedlaender, Helena Groot, David Gurwitz, Rene J. Herrera, Xiaoyi Huang, Judith Kidd, Kenneth K. Kidd, Audre Langaney, Alice A. Lin, S. Qasim Mehdi, Peter Parham, Alberto Piazza, Maria Pia Pistillo, Yaping Qian, Qunfang Shu, Jiujin Xu, S. Zhu, James L. Weber, Henry T. Greely, Marcus W. Feldman, Gilles Thomas, Jean Dausset, and L. Luca Cavalli-Sforza. A Human Genome Diversity Cell Line Panel. *Science.* 2002; 296 (5566): 261b-262b. (in Letters).

30. Son JH, Chun HS, Joh TH, Cho S, Conti B, Lee JW. (1999). Neuroprotection and neuronal differentiation studies using substantia nigra dopaminergic cells derived from transgenic mouse embryos. *J Neurosci*. 1999; **19**(1): 10-20

31. Guo, SW and Thompson, EA. Performing the Exact Test of Hardy-Weinberg Proportion for Multiple Alleles. *Biometrics.* 1992; 48: 361-372.

32. Morrison CF, McAllister J, Dobson SP, Mulderry PK, Quinn JP. An activator element within the preprotachykinin-A promoter. *Mol Cell Neurosci*. 1994; 5(2): 165-75.

33. Morrison CF, McAllister J, Lyons V, Chapman K, Quinn JP. The rat preprotachykinin-A promoter is regulated in PC12 cells by the synergistic action of multiple stimuli. *Neurosci Lett.* 1994; 181(1-2): 117-20.

34. .Walker PD, Andrade R, Quinn JP, Bannon MJ. Real-time analysis of preprotachykinin promoter activity in single cortical neurons. *J Neurochem*. 2000; 75(2): 882-5.

35. Patkar AA, Berrettini WH, Hoehe M, Hill KP, Gottheil E, Thornton CC, Weinstein SP. No association between polymorphisms in the serotonin transporter gene and susceptibility to cocaine dependence among African-American individuals. *Psychiatr Genet*. 2002; 12(3):161-4.

36. Patkar AA, Berrettini WH, Hoehe M, Thornton CC, Gottheil E, Hill K, Weinstein SP. Serotonin transporter polymorphisms and measures of impulsivity, aggression, and sensation seeking among African-American cocaine-dependent individuals. *Psychiatry Res*. 2002; 110(2):103-15.

37. Patkar AA, Berrettini WH, Hoehe M, Hill KP, Sterling RC, Gottheil E, Weinstein SP. Serotonin transporter (5-HTT) gene polymorphisms and susceptibility to cocaine dependence among African-American individuals. *Addict Biol*. 2001; 6(4):337-345.

38. Cubells JF, Kranzler HR, McCance-Katz E, Anderson GM, Malison RT, Price LH, Gelernter J. A haplotype at the DBH locus, associated with low plasma dopamine beta-hydroxylase activity, also associates with cocaine-induced paranoia. *Mol Psychiatry*. 2000; 5(1):56-63.

39. Comings DE, Gonzalez N, Wu S, Saucier G, Johnson P, Verde R, MacMurray JP. Homozygosity at the dopamine DRD3 receptor gene in cocaine dependence. *Mol Psychiatry*. 1999; 4(5):484-487.

40. Volkow ND, Fowler JS, Wang GJ and Swanson JM. Dopamine in drug abuse and addiction: results from imaging studies and treatment implications Molecular Psychiatry (2004) 9, 557–569