Genes, Brain and Behavior
Official publication of the International Behavioural and Neural Genetics Society

Genes, Brain and Behavior (2012) **11:** 704–711 doi: 10.1111/j.1601-183X.2012.00815.x

Association of Neurexin 3 polymorphisms with smoking behavior

E. Docampo[†], M. Ribasés[‡,§,¶], M. Gratacòs[†], E. Bruguera[¶], C. Cabezas[**], C. Sánchez-Mora[‡], G. Nieva[§,¶], D. Puente[††], J. M. Argimon-Pallàs[‡‡], M. Casas[§,¶,§§], R. Rabionet[*,†] and X. Estivill[†]

[†]Genes and Disease Program, Centre for Genomic Regulation (CRG) and UPF and Centro de Investigación Biomédica en Red en Epidemiología y Salud Pública (CIBERESP), [‡]Psychiatric Genetics Unit, Vall d'Hebron Research Institute (VHIR), [§]Biomedical Network Research Centre on Mental Health (CIBERSAM), [¶]Department of Psychiatry, Hospital Universitari Vall d'Hebron, [**]Subdirecció General de Promoció de la Salut, Direcció General de Salut Pública, Departament de Salut, [††]Institut Universitari d'Investigació en Atenció Primària Jordi Gol, [‡‡]Divisió d'Avaluació, Servei Català de la Salut, and [§§]Department of Psychiatry and Legal Medicine, Universitat Autònoma de Barcelona, Catalonia, Spain
**Corresponding author: R. Rabionet, Genes and Disease Program, Center for Genomic Regulation (CRG-UPF), C/Dr. Aiguader 88, 08003 Barcelona, Spain. E-mail: kelly.rabionet@crg.eu*

The Neurexin 3 gene (*NRXN3)* has been associated with dependence on various addictive substances, as well as with the degree of smoking in schizophrenic patients and impulsivity among tobacco abusers. To further evaluate the role of *NRXN3* in nicotine addiction, we analyzed single nucleotide polymorphisms (SNPs) and a copy number variant (CNV) within the *NRXN3* genomic region. An initial study was carried out on 157 smokers and 595 controls, all of Spanish Caucasian origin. Nicotine dependence was assessed using the Fagerström index and the number of cigarettes smoked per day. The 45 *NRXN3* SNPs genotyped included all the SNPs previously associated with disease, and a previously described deletion within *NRXN3*. This analysis was replicated in 276 additional independent smokers and 568 controls. Case–control association analyses were performed at the allele, genotype and haplotype levels. Allelic and genotypic association tests showed that three *NRXN3* SNPs were associated with a lower risk of being a smoker. The haplotype analysis showed that one block of 16 Kb, consisting of two of the significant SNPs (rs221473 and rs221497), was also associated with lower risk of being a smoker in both the discovery and the replication cohorts, reaching a higher level of significance when the whole sample was considered [odds ratio = 0.57 (0.42–0.77), permuted $P = 0.0075$]. By contrast, the *NRXN3* CNV was not associated with smoking behavior. Taken together, our results confirm a role for *NRXN3* in susceptibility to smoking behavior, and strongly implicate this gene in genetic vulnerability to addictive behaviors.

Keywords: Addiction, association, NRXN3, smoking, SNP

Received 26 Jan 2012, revised 21 Mar 2012 and 22 May 2012, accepted for publication 11 June 2012

According to the 2011 World Health Organization report on the global tobacco epidemic, tobacco kills nearly 6 million people each year, causing hundreds of billions of dollars of economic damage worldwide (WHO 2011). In fact, smoking is the single greatest contributor to preventable ill health and premature death (reviewed in Bierut 2011). The development of nicotine addiction is influenced by environmental and genetic factors, and while environmental factors have a stronger influence on initiation, genetic factors play a more significant role in the transition from regular use to addiction (Bierut 2011; Vink *et al*. 2005).

A genetic component in the development of nicotine addiction has been showed in twin and family studies, showing an estimated heritability of 30–72% (Agrawal *et al*. 2008). Twin studies also showed a hereditary component for a range of diverse smoking-related phenotypes, including age at initiation, intensity and cessation (Heath *et al*. 2002). Similarly, familial studies showed that siblings of habitual-smoking probands are at greater risk of becoming habitual smokers (Bierut *et al*. 1998).

At the molecular level, efforts to untangle the genetic component of nicotine addiction have focused on both the study of candidate genes and on hypothesis-free whole genome analyses. To date, the most robust genetic finding is the link between nicotine addiction and two gene clusters of nicotine receptors: $\alpha 3$, $\alpha 5$ and $\beta 4$ (*CHRNA3*, *CHRNA5* and *CHRNB4)* on chromosome 15; and $\alpha 6$ and $\beta 3$ (*CHRNA*6 and *CHRNB*3) on chromosome 8 (Tagconsortium 2010; Thorgeirsson *et al*. 2008). In spite of the reported strong associations, the contribution of nicotinic receptor genes to susceptibility to nicotine addiction only partially explains the heritability of this condition, suggesting that other genes contribute in an additive or epistatic manner to the development of nicotine addiction.

Nicotine dependence has also been associated with the *Neurexin* genes (Bierut *et al*. 2007; Nussbaum *et al*. 2008). Neurexins are presynaptic cell adhesion proteins that are essential for the development and function of GABAergic and glutamatergic synapses (Craig *et al*. 2006; Sudhof 2008). These synapses participate in key circuits influencing addictive behaviors (Lein *et al*. 2007; Ullrich *et al*. 1995). Two members of the neurexin gene family

© 2012 The Authors
Genes, Brain and Behavior © 2012 Blackwell Publishing Ltd and International Behavioural and Neural Genetics Society

have been associated with nicotine dependence, Neurexin 1 and Neurexin 3 (Bierut *et al*. 2007; Nussbaum *et al*. 2008). Moreover, Neurexin 3 has also been associated with the degree of smoking in schizophrenic patients (Novak *et al*. 2009), impulsivity and substance abuse (Stoltenberg *et al*. 2011), illegal substance abuse (Lachman *et al*. 2007) and alcohol dependence (Hishimoto *et al*. 2007).

In addition to single nucleotide polymorphisms (SNPs), structural variants involving neurexin genes have been described in autism spectrum disorders (Ching *et al*. 2010; Gauthier *et al*. 2011; Kim *et al*. 2008; Vaags *et al*. 2012; Wisniowiecka-Kowalnik *et al*. 2010) and schizophrenia patients (Kirov *et al*. 2008). These copy number variants (CNVs) are rare, large rearrangements. A common CNV in Neurexin 3 *(NRXN3*) has also been described (Conrad *et al*. 2010). The latter CNV, termed NRXN3_del henceforth, involves a common insertion/deletion polymorphism spanning 8.9 kb of the second intron of the *NRXN3* β isoform, and its association with disease/psychiatric phenotypes has not been investigated. As tobacco addiction is a common mild phenotype, NRXN3_del may be involved in susceptibility to smoking behavior.

In the present study we investigated the role of *NRXN3* in nicotine addiction. We performed a comprehensive genetic study in Spanish smokers and control subjects, which included analysis of tagSNPs and candidate SNPs, and the common CNV located within the *NRXN3* gene.

Materials and methods

Samples

All participants in the smoking group were individuals interested in receiving treatment to give up smoking. The original cohort (TAB_Discovery) included 157 smokers recruited and evaluated at the Tobacco Unit of the Addictive Behavior Unit at the Hospital Universitari Vall d' Hebron (Barcelona, Catalonia, Spain). These subjects were evaluated with the Addiction Research Unit questionnaire (ARU) (West & Russell 1985), the Beck Depression Inventory (BDI) (Beck *et al*. 1961) and the State-Trait Anxiety Inventory (STAI) (Spielberger *et al*., 1970). Subjects smoking regularly that requested for treatment to quit smoking were considered as smokers. The replication sample (TAB_replication) consisted of 276 smokers, recruited at different primary care centers in Catalonia, Spain (ClinicalTrials.gov identifier: NCT00125905). These subjects were evaluated with the Richmond test, to assess their motivation to quit smoking. In these sample sets, subjects smoking regularly and willing to quit smoking were considered as smokers. Nicotine dependence was assessed in both samples using the Fagerström index (Heatherton *et al*. 1991) and through the number of cigarettes smoked per day.

The control sample consisted of 1163 healthy blood donors recruited from the Blood bank (Banc de Sang i de Teixits) at the Hospital Universitari Vall d'Hebron. Blood donors were asked to fill in a questionnaire including the question 'have you ever smoked'; subjects providing a negative answer were considered non-smokers. This control sample was further divided into two random cohorts using the 'Research Randomizer' v.3.0 software (http://www.randomizer.org/): CVH1 ($n = 595$) and CVH2 ($n = 568$). CVH1 and CVH2 were used as controls for the discovery and replication cohorts, respectively. No significant differences in sex or age were detected between the groups, and the demographic and clinical attributes associated with the samples are summarized in Table 1.

All participants, both cases and controls, were of Spanish Caucasian origin and they provided informed consent before enrollment. The project was approved by the ethics committees at all the participating recruitment centers.

Genotyping of NRXN3 variants

We selected 45 SNPs within the genomic region of *NRXN3* using the Haploview software (Barrett *et al*. 2005). These SNPs capture 42% of alleles in *NRXN3*α (1460 kb) and 66% in *NRXN3*β at $r^2 \geq 0.8$, based on the CEU HapMap genotyped SNPs (version 2, release 21). In the selection process, potential functional variants, variants located near splice sites and SNPs previously associated with disease were included as TagSNPs (Heard-Costa *et al*. 2009; Hishimoto *et al*. 2007; Novak *et al*. 2009). A total of 45 SNPs (41 TagSNPs and 4 singleton SNPs) and 1320 samples (CVH1, CVH2 and TAB_Discovery) were genotyped (Table 2) using *VeraCode* technology (Illumina, Inc., San Diego, CA, USA) at the CeGen genotyping facility (Centro Nacional de Genotipado, Genoma España, Barcelona), following the manufacturer's instructions. The assays developed for the VeraCode beads were analyzed using Illumina's BeadXpress Reader System and the data analyzed using BeadStudio v.2.0 software (Illumina, Inc., San Diego, CA, USA). As a quality control, 5% of the genotyped samples were duplicated, and no-template controls and six Hapmap trios were included(NA10840–NA12286–NA12287, NA12766–NA12775–NA12776, NA12818–NA12829–NA12830, NA 12832–NA12842–NA12843, NA12865–NA12874–NA12875, NA128 77–NA12889–NA12890).

We also explored the possible involvement of a previously described *NRXN3* CNV in susceptibility to smoking behavior. This CNV, NRXN3_del, consists of a common insertion/deletion polymorphism spanning 8.9 kb of the second intron of the *NRXN3* β isoform *(Conrad et al. 2010)*. To genotype this structural variant, we included two SNPs located within the deleted region in the Veracode assay (rs12894142 and rs12100748), and we designed a third SNP assay (NRXN3del) with each of its extension probes flanking the breakpoints of the CNV. A combination of the results for these SNPs was used to assess the genotype. In addition, NRXN3_del was genotyped by multiplex polymerase chain reactions (PCRs) with 5′

Table 1: Summary of the demographic and clinical data from the cohorts included in the study

Variable	Controls		Cases	
	CVH1	CVH2	TAB_ Discovery	TAB_ Replication
	($N = 595$)	($N = 568$)	($N = 157$)	($N =$ '276)
Gender, *n* (%)				
Male	216 (36.6)	180 (31.7)	82 (53.6)	118 (43.2)
Female	375 (63.4)	388 (68.3)	69 (46.4)	155 (56.8)
Mean age, yrs (±SD)	53.8 (16.8)	54.1 (17.5)	46.6 (7.7)	45.5 (13.5)
Cigarettes/day (±SD)	–	–	21.9 (9.9)	19.8 (11.4)
Fagerström score (±SD)	–	–	5.12 (2.42)	4.16 (2.58)

Table 2: SNPs tested for association

Name	Position	Hwpval	%Geno	MAF	Alleles	*P* value*
rs1004212†	78250978	1	99.6	0.134	G:A	0.1609
rs31431	78510788	1	99.9	0.124	C:A	0.4279
rs17108457	78520530	1	100	0.111	A:T	0.2536
rs2202167	78569377	0.6121	99.6	0.385	C:A	0.5141
rs2202175	78625403	0.2421	99.7	0.394	A:C	0.9386
rs12323794	78631351	0.2865	100	0.365	G:A	0.5657
rs6574495	78669026	1	100	0.396	G:A	0.6821
rs11627269	78784324	1	100	0.332	G:A	0.6109
rs4603484	78925469	0.9725	100	0.432	A:C	0.3402
rs2196443	78946968	0.1608	100	0.194	A:C	0.0369
rs10146997†	79014914	0.047	100	0.179	A:G	0.4977
rs11848580	79038540	0.0332	99.9	0.392	A:G	0.0368
rs1424850	79067756	0.28	100	0.285	T:A	0.0002
rs221415	79115621	0.829	100	0.183	A:G	0.1393
rs7153625	79119014	1	99.7	0.172	G:A	0.4427
rs8022725	79129222	0.8011	100	0.359	G:C	0.2103
rs2543576	79141981	0.433	99.9	0.446	G:A	0.0143
rs221497	79158232	0.3867	100	0.112	G:A	0.0020
rs221473	79174368	1	100	0.187	A:T	0.0009
NRXN3del	79176041	1	65.6	0	G:G	NA
rs12894142	79179213	1.11E-07	83.4	0.027	A:G	NA
rs12100748	79183747	5.00E-04	83.8	0.007	G:A	NA
rs221449	79193929	1	99.5	0.258	G:A	0.1448
rs10130593	79211814	1	100	0.321	C:G	0.4242
rs1030127	79231192	0.8693	100	0.123	A:G	0.5384
rs178377	79247778	0.6315	100	0.147	A:G	0.3535
rs8021767	79263758	1	99.9	0.198	G:A	0.1039
rs5014481	79282614	0.2073	100	0.429	C:A	0.7918
rs8018724	79318557	0.4432	100	0.368	C:A	0.3199
rs2215840†	79339820	0.7087	100	0.371	A:G	0.9289
rs10145867	79339926	0.7731	100	0.083	G:A	0.2061
rs932265	79360177	0.4085	100	0.159	G:A	0.1521
rs1159039	79361345	0.2736	100	0.344	A:G	0.8745
rs760288†	79376682	0.5889	100	0.336	A:G	0.6339
rs2293839	79387212	0.2407	100	0.304	G:A	0.9259
rs8019381†	79390335	0.3627	100	0.162	G:A	0.4028
rs994010	79419355	0.6763	100	0.235	A:G	0.5751
rs9323679	79421543	0.3453	100	0.145	A:G	0.6752

Markers included in the NRXN3-del genotyping appear in bold.
*Genotypic association under the log-additive model corrected by sex and age.
†SNPs analyzed in substance-abuse-related papers.

FAM modification, followed by capillary electrophoresis in a 3730XL automatic sequencer and analysis with the Gene Mapper package (Applied Biosystems, Foster City, CA, USA). In each experiment, we included both negative and positive controls, consisting of three Hapmap samples with validated genotypes for this CNV (Conrad *et al*. 2010): NA06991 (homozygous deleted), NA010831 (heterozygous) and NA06994 (homozygous non-deleted). The observed concordance between the two assays was 100%.

SNPs with a significant association in the discovery samples (rs1424850, rs221473 and rs221497) were subsequently genotyped using the KBiosciences PCR SNP genotyping system (KASPar®, Kbioscience, Hoddesdon, Herts, UK), which uses a competitive allele-specific PCR, following the manufacturer's instructions. Genotyping of the independent smoker cohort (276 additional smokers) was performed at the CeGen genotyping facilities, in the CNIO Node (Centro Nacional de Genotipado, Genoma España). As a quality control, 5% of the genotyped samples had their genotypes confirmed by Sanger sequencing.

Statistical analysis

Quality control and case–control association analyses for both Veracode and Kaspar assays were performed using PLINK software, version 1.07 (http://pngu.mgh.harvard.edu/purcell/plink/) (Purcell *et al*. 2007). SNPs with a low genotyping rate (<95%), not fulfilling the Hardy–Weinberg equilibrium (HWE: $P < 0.05$), or with a minimum allele frequency below 5%, as well as samples with a low genotyping rate (<95%), were excluded from the association analyses. We tested for association for two phenotypes: smoking behavior, and level of addiction based on the Fagerström score. For this, we considered smoking behavior as a dichotomic phenotype (smoker vs. non-smoker), and we used the total Fagerström score, either as a quantitative or a nominal variable: minimally (<4), moderately (4–6) and highly (7–10) dependent. All association tests included sex and age as covariates, and they were corrected for multiple testing based on the Bonferroni correction (35 markers: $P < 0.0014$). Linkage disequilibrium (LD) between polymorphisms and haplotype block structures were evaluated using Haploview software (version 4.1). Regions of strong LD were defined according to the confidence intervals algorithm (Gabriel *et al*. 2002). We studied haplotypes present at a frequency of at least 5%, and estimated the significance of the best result via a permutation procedure (2000 permutations). The SNPassoc R package was used for all analyses (Gonzalez *et al*. 2007).

Assessment of SNP function

We used WGAviewer (Ge *et al*. 2008) and Pupasuite 3.1 (http://pupasuite.bioinfo.cipf.es/) to assess the possible functional consequences of the SNPs associated with smoking behavior. The GENEVAR tool of WGAviewer was used to evaluate the effects of SNPs on gene expression, based on Hapmap genotype and expression data (Stranger *et al*. 2005, 2007). Pupasuite was used to predict the pathogenicity of an SNP based on the disruption of potential transcription factor binding sites, the creation or disruption of splice sites, its location in miRNA sequences or targets, or in promoter or conserved regions.

Results

Discovery cohort analysis

Ten SNPs (4 with a low genotyping rate, 1 not fulfilling HWE and 5 with a low MAF) and 20 samples (with a low genotyping rate) did not fulfill the aforementioned quality control criteria and, thus, a total of 38 SNPs were analyzed in 153 cases (TAB_Discovery) and 583 controls (CVH1) (Fig. 1 and Table 2). No Mendelian errors were detected in the HapMap trios, and 100% concordance was observed with the described genotype in all cases. Allelic and genotypic (log-additive model) association tests showed that three SNPs were significantly associated with a lower risk of being a smoker after Bonferroni correction for multiple testing [rs1424850 [$P = 0.0002$, odds ratio (OR) = 0.55, 95% confidence interval (CI) = 0.40–0.76], rs221497 ($P = 0.0020$, OR = 0.47, 95% CI = 0.28–0.79) and rs221473 ($P = 0.0009$, OR = 0.54, 95% CI = 0.37–0.79); Table 3]. Based on the LD pattern of the region, nine haplotype blocks were defined, and haplotype association analysis showed an association of block 3 with a decreased risk of being a smoker (OR = 0.44, 95% CI = 0.26–0.75; permuted $P = 0.0025$). This block spanned 16 kb and included two of the SNPs identified in the single-SNP analysis (rs221497 and rs221473; Table 4 and Fig. 2). We

Table 3: Summary of the Veracode assay results: SNPs with a statistically significant association

	TAB_Discovery vs. CVH1		TAB_Replication vs. CVH2		All	
	P value*	OR (95% CI)	*P* value*	OR (95% CI)	*P* value*	OR (95% CI)
rs1424850	0.0002†	0.55 (0.40–0.76)	0.59	0.94 (0.73–1.19)	0.0045	0.76 (0.73–0.92)
rs221497	0.0020	0.47 (0.28–0.79)	0.047	0.70 (0.49–1.00)	0.0005†	0.62 (0.46–0.82)
rs221473	0.0009†	0.54 (0.37–0.79)	0.039	0.75 (0.57–0.99)	0.0003†	0.67 (0.54–0.84)

*Genotypic association under the log-additive model corrected by sex and age.
†Significant after Bonferroni correction for multiple testing.

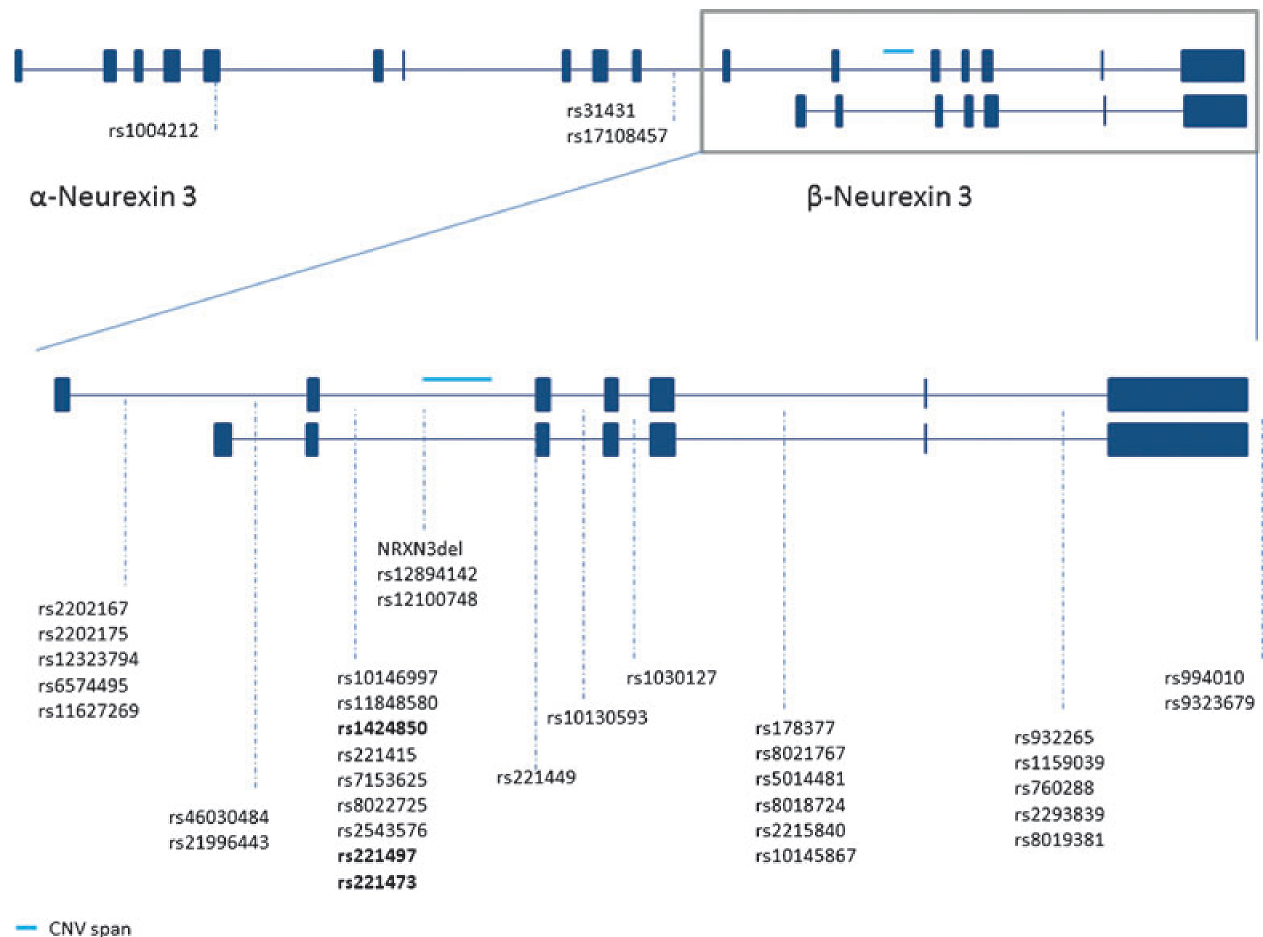

Figure 1: SNPs location. Location of the 38 SNPs (including the CNV assay) along the α and β *NRXN3* genomic region. SNPs exhibiting association with nicotine addiction appear in bold.

also tested for any association with the level of addiction, as determined by the Fagerström index, although no significant relationships were observed for any of the SNPs evaluated.

Association analysis for NRXN3_del was performed separately from the SNP analysis. After quality control, 154 cases and 580 controls were included in the analysis, and this deleted allele was observed at a similar frequency in both the cases and controls (cases: 0.42; controls: 0.40). Hence, there was no association between the CNV polymorphism and smoking behavior ($P = 0.3$, genotypic association; $P = 0.46$, allelic association Fisher test).

Replication

To replicate these findings, the 3 SNPs exhibiting an association were genotyped in an independent cohort of 276 individuals addicted to nicotine (TAB_replication) and in an independent control set (CVH2). We found a nominal association for rs221473 ($P = 0.039$, OR = 0.75, 95% CI = 0.57–0.99) and

Table 4: Haplotype association

	TAB_Discovery vs. CVH1		TAB_Replication vs. CVH2		All	
	P value*	OR (95% CI)	*P* value*	OR (95% CI)	*P* value*	OR (95% CI)
Haplotype[†]	0.0031	0.46 (0.27–0.77)	0.037	0.68 (0.47–0.98)	0.0003	0.57 (0.42–0.77)
Permuted haplotype[‡]	0.017	–	0.23	–	0.0075	–

*Genotypic association under the log-additive model corrected by sex and age.
[†]rs221497(A)-rs221473(T): freq 0.11.
[‡]global *P* value for 2000 permutations.

rs221497 ($P = 0.047$, OR = 0.70, 95% CI = 0.49–1.00), and their composite haplotype. When considering both cohorts, the level of significance increased and haplotype association persisted after permutation (OR = 0.57; 95% CI = 0.42–0.77, permuted $P = 0.0075$; Tables 3 and 4).

Evaluation of SNP function

The three SNPs significantly associated with smoking behavior were located in an intronic region, and *in silico* evaluation of their functional implications using Genevar showed no correlation between the SNP genotypes and levels of expression. Evaluation of possible functional effects using Pupasuite 3.1 showed that rs221497 and rs221449 were highly conserved as compared with the mouse genome, although no potential functional consequences were showed.

Discussion

In the present study, we performed a high-density SNP association study of the link between *NRXN3* and smoking behavior in two independent samples of Spanish smokers. Our findings show a significant association between two *NRXN3* SNPs and smoking behavior, and a nominal association for a third SNP. To the best of our knowledge, this is the first study to assess genetic variation in the *NRXN3* gene by analyzing tagSNPs, additional putative functional variants and structural variation.

NRXN3 is one of the three members of the neurexin gene family. Neurexins are presynaptic cell adhesion proteins that are essential for the development and function of synapses, and that are implicated in neurotransmitter release. *NRXN3* is located on chromosome 14 and the two existing promoters drive the expression of its α and β isoforms. Together with five canonical sites of alternative splicing, these promoters give rise to over 1000 isoforms (Rowen *et al*. 2002; Tabuchi & Sudhof 2002). *NRXN3* is expressed in key circuits that have been implicated in addictive behaviors, including glutamatergic neurons projecting from the prefrontal cortex to the striatum and GABAergic neurons (Lein *et al*. 2007; Ullrich *et al*. 1995). Hence, *NRXN3* is a strong candidate for substance abuse susceptibility.

Genetic variants in *NRXN3* have been associated with various substance dependence phenotypes (Bierut 2007; Hishimoto 2007; Liu 2005; Novak 2009; Stoltenberg 2011), including the degree of smoking in schizophrenic patients (Novak *et al*. 2009) and nicotine dependence (Bierut *et al*. 2007). The convergence of associations between *NRXN3* and different phenotypes suggests that genetic factors contribute to developing dependence on various classes of drugs (Bierut 2011). Nevertheless, there is only a partial overlap in the SNPs analyzed in each of the studies, various SNPs have been analyzed in only one study, and the phenotypes tested in the different studies, although related, are not the same. It is possible that different SNPs in the same gene are associated with the different substance dependence phenotypes. Several of the SNPs included in this work have also been analyzed in other studies (Table 2), although most of them did not show any association. rs1004212 had been associated with heavy smoking in schizophrenic patients and has shown nominal association with alcohol problems in men (Stoltenberg 2011), and rs8019381 had been associated with alcohol dependence (Hishimoto 2007). However, we did not find these SNPs associated with smoking behavior. Several explanations can be offered for the conflicting SNPs identified in different studies (including ours). First, as mentioned above, although related, the phenotypes investigated in the different studies are not the same, thus allowing for different SNP associations. Also, previous studies were performed on American (Bierut *et al*. 2007; Stoltenberg *et al*. 2011), Canadian (Novak *et al*. 2009) and Australian (Bierut *et al*. 2007) populations, indicating that the failure to replicate these data may also be due to genetic heterogeneity among populations. Alternatively, the observed differences may reflect distinct gene $\times$ environment interactions (Ioannidis *et al*. 2007). Finally, it is possible that an untyped causative variant in LD with the different signals identified in the *NRXN3* gene may exist.

Population admixture in our sample may represent a potential limitation to our analysis, although no significant population effects have been observed in previous studies on Spanish populations (Julia *et al*. 2008; Ribases *et al*. 2009). Moreover, all samples in the present study were recruited in the same geographical region (Catalonia) and, thus, the west-to-east trend reported previously is unlikely to influence our data (Julia *et al*. 2008). Indeed, population admixture was excluded from part of the control cohort in an independent study (Ribases *et al*. 2009) and, hence, the influence of population admixture in the present work is probably negligible.

In our two-step study, the association observed was evident in both the discovery and replication cohorts, although

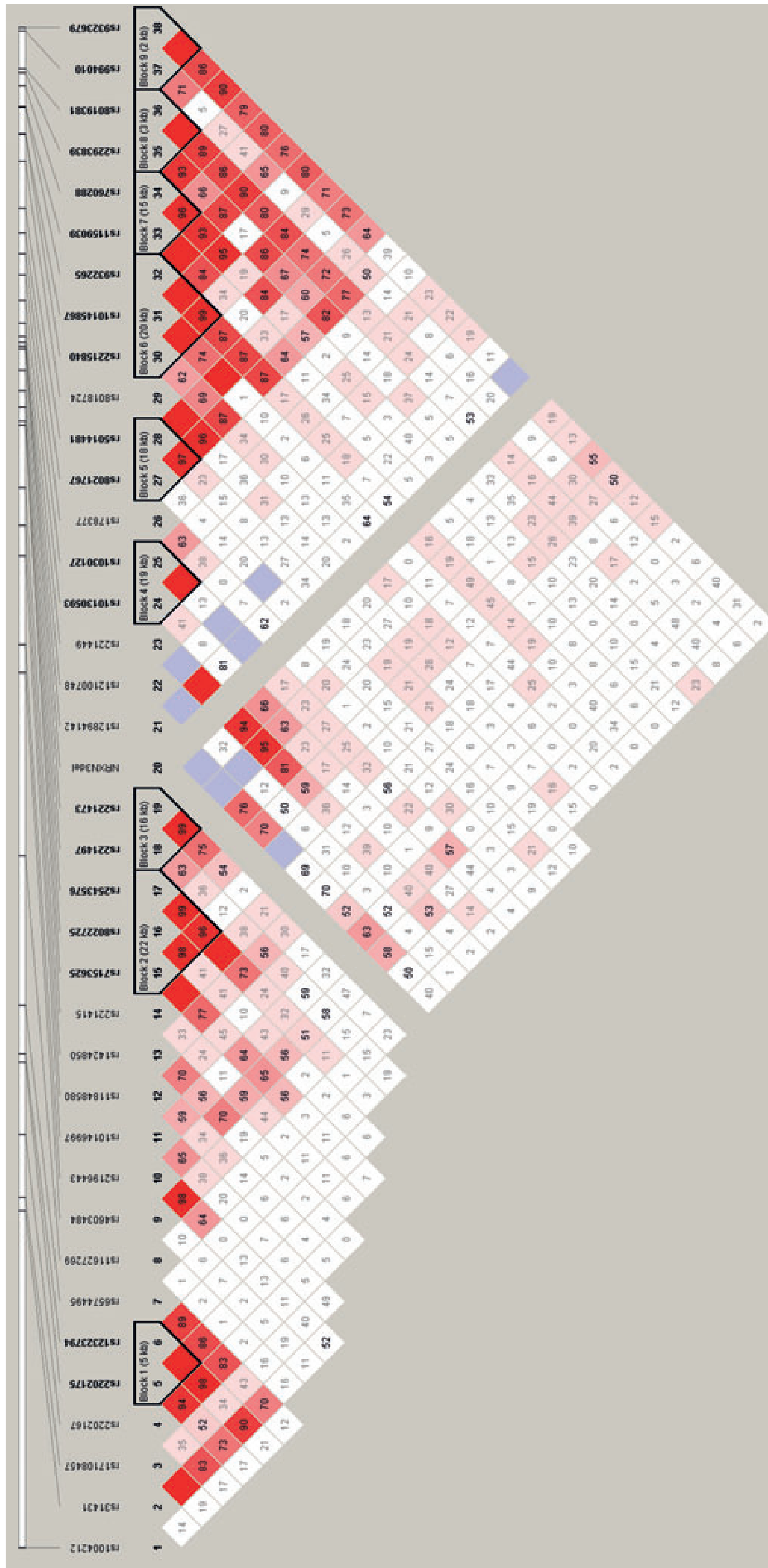

Figure 2: Linkage disequilibrium plot. LD of the 38 (35 + 3 assessing NRXN3_DEL) included in the analysis, as assessed using Haploview software. The D′ score value is provided inside the boxes.

the level of significance was lower in the latter of the two cohorts. This is most probably due to clinical heterogeneity between these patient cohorts, as the Fagerström index and the number of cigarettes smoked per day were significantly higher in the discovery cohort. Nonetheless, global analysis of both cohorts together showed a significant protective pooled effect.

Although we detected an association with status as a smoker, we observed no association with the Fagerström index (Table S1), suggesting that although related, these two phenotypes may have distinct genetic bases. Alternatively, a lack of statistical power may explain the failure to detect association with the Fagerström index, given the smaller effective sample size once the sample is divided into the different Fagerström categories. In fact, we did observe an association in two separate smoker cohorts, albeit in small sample sets. Performing the replication study in a larger cohort from a different population would provide strong support for the present findings, and may even show an association with the different categories defined by the Fagerström index.

The two *NRXN3* SNPs associated with smoking behavior are located in an intronic region and no specific effect was predicted for either SNP. Although this intronic region is highly conserved between species, suggesting that it fulfills an important role, it is possible that the SNPs identified merely tag an as-yet-unidentified functional variant. Next-generation sequencing technologies may help identify additional functional variants within the gene that could have been tagged by the associated SNPs. Targeted sequencing of *NRXN3* in larger nicotine addiction cohorts will be an important next step to determine its role in substance abuse. To better understand the functional involvement of these and other SNPs in NRXN3, it will be of interest to correlate these variants with the expression of different splice variants. Other functional assays, such as testing the response to nicotine exposure of neuronal cell lines carrying the different genotypes, will also help elucidate the functional consequences of these polymorphisms.

Despite finding no association between the NRXN3_del variant and being a smoker, we have confidently identified an INDEL by genotyping an internal SNP in combination with a specifically designed assay. The concordance of this genotyping technique with more traditional multiplex PCR was 100%, suggesting that it constitutes a useful tool to genotype common CNVs with known and defined breakpoints in a high-throughput manner.

In summary, we describe new *NRXN3* SNPs that are associated with susceptibility to being a smoker in agreement with previous studies implicating neurexins in the development of addiction.

References

Agrawal, A., Lynskey, M.T., Pergadia, M.L., Bucholz, K.K., Heath, A.C., Martin, N.G. & Madden, P.A. (2008) Early cannabis use and DSM-IV nicotine dependence: a twin study. *Addiction* **103**, 1896–1904.

Barrett, J.C., Fry, B., Maller, J. & Daly, M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265.

Beck, A.T., Ward, C.H., Mendelson, M., Mock, J. & Erbaugh, J. (1961) An inventory for measuring depression. *Arch Gen Psychiatry* **4**, 561–571.

Bierut, L.J. (2011) Genetic vulnerability and susceptibility to substance dependence. *Neuron* **69**, 618–627.

Bierut, L.J., Dinwiddie, S.H., Begleiter, H., Crowe, R.R., Hesselbrock, V., Nurnberger, J.I. Jr, Porjesz, B., Schuckit, M.A. & Reich, T. (1998) 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* **55**, 982–988.

Bierut, L.J., Madden, P.A., Breslau, N. *et al*. (2007) Novel genes identified in a high-density genome wide association study for nicotine dependence. *Hum Mol Genet* **16**, 24–35.

Conrad, D.F., Pinto, D., Redon, R. *et al*. (2010) Origins and functional impact of copy number variation in the human genome. *Nature* **464**, 704–712.

Ching, M.S., Shen, Y., Tan, W.H. *et al*. (2010) Deletions of NRXN1 (neurexin-1) predispose to a wide spectrum of developmental disorders. *Am J Med Genet B Neuropsychiatr Genet* **153B**, 937–947.

Craig, A.M., Graf, E.R. & Linhoff, M.W. (2006) How to build a central synapse: clues from cell culture. *Trends Neurosci* **29**, 8–20.

Gabriel, S.B., Schaffner, S.F., Nguyen, H., Moore, J.M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A., Faggart, M., Liu-Cordero, S.N., Rotimi, C., Adeyemo, A., Cooper, R., Ward, R., Lander, E.S., Daly, M.J. & Altshuler, D. (2002) The structure of haplotype blocks in the human genome. *Science* **296**, 2225–2229.

Gauthier, J., Siddiqui, T.J., Huashan, P. *et al*. (2011) Truncating mutations in NRXN2 and NRXN1 in autism spectrum disorders and schizophrenia. *Hum Genet* **130**, 563–573.

Ge, D., Zhang, K., Need, A.C., Martin, O., Fellay, J., Urban, T.J., Telenti, A. & Goldstein, D.B. (2008) WGAViewer: software for genomic annotation of whole genome association studies. *Genome Res* **18**, 640–643.

Gonzalez, J.R., Armengol, L., Sole, X., Guino, E., Mercader, J.M., Estivill, X. & Moreno, V. (2007) SNPassoc: an R package to perform whole genome association studies. *Bioinformatics* **23**, 644–645.

Heard-Costa, N.L., Zillikens, M.C., Monda, K.L. *et al*. (2009) NRXN3 is a novel locus for waist circumference: a genome-wide association study from the CHARGE Consortium. *PLoS Genet* **5**, e1000539.

Heath, A.C., Martin, N.G., Lynskey, M.T., Todorov, A.A. & Madden, P.A. (2002) Estimating two-stage models for genetic influences on alcohol, tobacco or drug use initiation and dependence vulnerability in twin and family data. *Twin Res* **5,** 113–124.

Heatherton, T.F., Kozlowski, L.T., Frecker, R.C. & Fagerstrom, K.O. (1991) The Fagerstrom Test for Nicotine Dependence: a revision of the Fagerstrom Tolerance Questionnaire. *Br J Addict* **86**, 1119–1127.

Hishimoto, A., Liu, Q.R., Drgon, T., Pletnikova, O., Walther, D., Zhu, X.G., Troncoso, J.C. & Uhl, G.R. (2007) Neurexin 3 polymorphisms are associated with alcohol dependence and altered expression of specific isoforms. *Hum Mol Genet* **16**, 2880–2891.

Ioannidis, J.P., Patsopoulos, N.A. & Evangelou, E. (2007) Heterogeneity in meta-analyses of genome-wide association investigations. *PLoS One* **2**, e841.

Julia, A., Ballina, J., Canete, J.D., Balsa, A., Tornero-Molina, J., Naranjo, A., Alperi-Lopez, M., Erra, A., Pascual-Salcedo, D., Barcelo, P., Camps, J. & Marsal, S. (2008) Genome-wide association study of rheumatoid arthritis in the Spanish population: KLF12 as a risk locus for rheumatoid arthritis susceptibility. *Arthritis Rheum* **58**, 2275–2286.

Kim, H.G., Kishikawa, S., Higgins, A.W. *et al*. (2008) Disruption of neurexin 1 associated with autism spectrum disorder. *Am J Hum Genet* **82**, 199–207.

Kirov, G., Gumus, D., Chen, W., Norton, N., Georgieva, L., Sari, M., O'Donovan, M.C., Erdogan, F., Owen, M.J., Ropers, H.H. & Ullmann, R. (2008) Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. *Hum Mol Genet* **17**, 458–465.

Lachman, H.M., Fann, C.S., Bartzis, M., Evgrafov, O.V., Rosenthal, R.N., Nunes, E.V., Miner, C., Santana, M., Gaffney, J., Riddick, A., Hsu, C.L. & Knowles, J.A. (2007) Genomewide suggestive linkage of opioid dependence to chromosome 14q. *Hum Mol Genet* **16**, 1327–1334.

Lein, E.S., Hawrylycz, M.J., Ao, N. *et al*. (2007) Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176.

Liu, Q.R., Drgon, T., Walther, D., Johnson, C., Poleskaya, O., Hess, J. & Uhl, G.R. (2005) Pooled association genome scanning: validation and use to identify addiction vulnerability loci in two samples. *Proc Natl Acad Sci U S A* **102**, 11864–11869.

Novak, G., Boukhadra, J., Shaikh, S.A., Kennedy, J.L. & Le Foll, B. (2009) Association of a polymorphism in the NRXN3 gene with the degree of smoking in schizophrenia: a preliminary study. *World J Biol Psychiatry* **10**, 929–935.

Nussbaum, J., Xu, Q., Payne, T.J., Ma, J.Z., Huang, W., Gelernter, J. & Li, M.D. (2008) Significant association of the neurexin-1 gene (NRXN1) with nicotine dependence in European- and African-American smokers. *Hum Mol Genet* **17**, 1569–1577.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. & Sham, P.C. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559–575.

Ribases, M., Bosch, R., Hervas, A. *et al*. (2009) Case-control study of six genes asymmetrically expressed in the two cerebral hemispheres: association of BAIAP2 with attention-deficit/hyperactivity disorder. *Biol Psychiatry* **66**, 926–934.

Rowen, L., Young, J., Birditt, B., Kaur, A., Madan, A., Philipps, D.L., Qin, S., Minx, P., Wilson, R.K., Hood, L. & Graveley, B.R. (2002) Analysis of the human neurexin genes: alternative splicing and the generation of protein diversity. *Genomics* **79**, 587–597.

Spielberger, C.D., Gorsuch, R. L. & Lushene, R.E. (1970) *Manual for the State-Trait Inventory. Consulting Psychology Press, Palo Alto*.

Stoltenberg, S.F., Lehmann, M.K., Christ, C.C., Hersrud, S.L. & Davies, G.E. (2011) Associations among types of impulsivity, substance use problems and Neurexin-3 polymorphisms. *Drug Alcohol Depend* **119**, e31–e38.

Stranger, B.E., Forrest, M.S., Clark, A.G., Minichiello, M.J., Deutsch, S., Lyle, R., Hunt, S., Kahl, B., Antonarakis, S.E., Tavare, S., Deloukas, P. & Dermitzakis, E.T. (2005) Genome-wide associations of gene expression variation in humans. *PLoS Genet* **1**, e78.

Stranger, B.E., Forrest, M.S., Dunning, M., Ingle, C.E., Beazley, C., Thorne, N., Redon, R., Bird, C.P., de Grassi, A., Lee, C., Tyler-Smith, C., Carter, N., Scherer, S.W., Tavare, S., Deloukas, P., Hurles, M.E. & Dermitzakis, E.T. (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* **315**, 848–853.

Sudhof, T.C. (2008) Neuroligins and neurexins link synaptic function to cognitive disease. *Nature* **455**, 903–911.

Tabuchi, K. & Sudhof, T.C. (2002) Structure and evolution of neurexin genes: insight into the mechanism of alternative splicing. *Genomics* **79**, 849–859.

TAGconsortium (2010) Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet* **42**, 441–447.

Thorgeirsson, T.E., Geller, F., Sulem, P. *et al*. (2008) A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature* **452**, 638–642.

Ullrich, B., Ushkaryov, Y.A. & Sudhof, T.C. (1995) Cartography of neurexins: more than 1000 isoforms generated by alternative splicing and expressed in distinct subsets of neurons. *Neuron* **14**, 497–507.

Vaags, A.K., Lionel, A.C., Sato, D. *et al*. (2012) Rare deletions at the Neurexin 3 locus in autism spectrum disorder. *Am J Hum Genet* **90**, 133–141.

Vink, J.M., Willemsen, G. & Boomsma, D.I. (2005) Heritability of smoking initiation and nicotine dependence. *Behav Genet* **35**, 397–406.

West, R.J. & Russell, M.A. (1985) Pre-abstinence smoke intake and smoking motivation as predictors of severity of cigarette withdrawal symptoms. *Psychopharmacology (Berlin)* **87**, 334–336.

WHO (2011) 2011 World Health Organization report on the global tobacco epidemic.

Wisniowiecka-Kowalnik, B., Nesteruk, M., Peters, S.U., Xia, Z., Cooper, M.L., Savage, S., Amato, R.S., Bader, P., Browning, M.F., Haun, C.L., Duda, A.W. 3rd, Cheung, S.W. & Stankiewicz, P. (2010) Intragenic rearrangements in NRXN1 in three families with autism spectrum disorder, developmental delay, and speech delay. *Am J Med Genet B Neuropsychiatr Genet* **153B**, 983–993.

Acknowledgments

This work was supported by the Spanish Ministry of Science and Innovation (SAF2008-00357), the European Commission (ENGAGE_201413), the Fundació de la Marató de TV3 (041810), and the Generalitat de Catalunya. E.D. is supported by the Spanish Ministry of Health (FIS-PFIS, FI 08/00148). R.R. received a Ramon y Cajal grant from the Spanish Ministry of Health (RyC04_Raquel Rabionet) and M.R. is supported by a Miguel de Servet contract from the ''Instituto de Salud Carlos III, ''Ministerio de Ciencia e Innovación'', Spain.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1: Association of the SNPs with the Fagerström score.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.