Study of five novel non-synonymous polymorphisms in human brain-expressed genes in a Colombian sample

Diego A Ojeda, Diego A Forero

1Laboratory of NeuroPsychiatric Genetics, Biomedical Sciences Research Group, School of Medicine, Universidad Antonio Nariño. Bogotá, Colombia. 2Laboratory of NeuroPsychiatric Genetics, School of Medicine, Universidad Antonio Nariño. Bogotá, Colombia

KEYWORDS
Molecular genetics
Candidate genes
Single nucleotide polymorphisms
Neurogenetics

ABSTRACT
Background: Non-synonymous single nucleotide polymorphisms (nsSNPs) in brain-expressed genes represent interesting candidates for genetic research in neuropsychiatric disorders. Purpose: To study novel nsSNPs in brain-expressed genes in a sample of Colombian subjects. Methods: We applied an approach based on high-throughput sequencing of available genomic data to identify and select novel nsSNPs in brain-expressed genes. We developed novel genotyping assays, based in allele-specific PCR methods, for these nsSNPs and genotyped them in 171 Colombian subjects. Results: Five common nsSNPs (rs6855837; p.Leu395Ile, rs2305160; p.Thr394Ala, rs10503929; p.Met289Thr, rs2270641; p.Thr4Pro and rs3822659; p.Ser735Ala) were studied, located in the CLOCK, NPAS2, NRG1, SLC18A1 and WWC1 genes. We reported allele and genotype frequencies in a sample of South American healthy subjects. There is previous experimental evidence, arising from genome-wide expression and association studies, for the involvement of these genes in several neuropsychiatric disorders and endophenotypes, such as schizophrenia, mood disorders or memory performance. Conclusions: Frequencies for these nsSNPs in the Colombian samples varied in comparison to different HapMap populations. Future study of these nsSNPs in brain-expressed genes, a synaptogenomics approach, will be important for a better understanding of neuropsychiatric diseases and endophenotypes in different populations.

Methods
Bioinformatic analysis of candidate genes
Several dozens of human brain-expressed genes were screened in silico for the presence of nsSNPs, using the BioMart tool. These genes correspond to several well-known functional categories, such as acetylcholine receptors, adaptor proteins, adhesion molecules, calcium channels, dopamine receptors, GABA receptors, glutamate receptors, potassium channels, protein kinases, protein phosphatases, serotonin receptors, signaling proteins, sodium channels, synaptic vesicle and transporters. MAF (>0.01) in HapMap populations was screened to identify nsSNPs.

The catalog of GWAS from the National Human Genome Research Institute (http://www.genome.gov/gwastudies), the public database BioGPS (http://biogps.org), the Stanley Brain Database –SBD- (https://www.stanleygenomics.org) and the HuGeNavigator tool (http://hugenavigator.net/) were used for further functional annotation. The GWAS catalog contains information for available GWAS for diseases and phenotypes, BioGPS contains an extensive set of genome-wide expression data for a wide range of normal human tissues, SBD provides differential expression data for BP and SZ postmortem human brains and the HuGeNavigator contains data for existing meta-analysis of genes and diseases (Tab. 1).

DNA sequences for flanking regions of the nsSNP was obtained from the UCSC genome browser (http://genome.ucsc.edu). Primers were designed with the WASP tool (for allele-specific PCR, -AS-PCR-), Batch Primer3 and Primerer online programs (for Tetra Primer Allele-Specific PCR, -T-ARMS-PCR-). Oligonucleotides were synthesized by Integrated DNA Technologies (Coralville, IA, USA).

Introduction
Common Neuropsychiatric Disorders (NPDs) are a major cause of morbidity and impairments in the quality of life around the world, with very large societal costs in the order of trillions of dollars. Several common NPDs (such as schizophrenia, bipolar disorder, Alzheimer’s disease) have large heritabilities, with a possible shared pathophysiology for some of them, highlighting the importance of the study of genetic factors. In recent years, dozens of genome-wide association studies (GWASs) have been carried out for NPDs. However, very few consistent and strong associations with NPDs have been found.

GWASs have been based in the common disease/common variant (CDCV) hypothesis. This hypothesis proposes that a single common variant could be a risk factor for a large fraction of patients. Several researchers have proposed that an alternative hypothesis, the common disease/rare variants (CDRV) could lead to the identification of additional risk factors for complex diseases, such as neuropsychiatric disorders. In the context of the CDRV hypothesis, a special interest is focused on functional variants, such as non-synonymous SNPs (nsSNPs), which are a minor fraction of the total number of SNPs. nsSNPs represent the most well understood group of genetic variants of possible functional importance. Alternations in dozens of genes involved in brain plasticity are responsible for increasing or decreasing the risk of NPDs.

In the present study, we screened in silico a large number of brain-expressed genes, identified 5 nsSNPs located in genes of neuropsychiatric relevance, developed novel genotyping assays and reported their frequencies in a sample of Colombian healthy subjects.
Table 1: General information about brain-expressed candidate genes

| Gene Name         | Gene symbol | Position | Size (bp) | Expression in brain | Changes | Meta-Analysis | GWAS | Expression changes | Gene Expression in brain | Expression in brain | Genotype * | PFC | PL | PN | PLB | STN |
|-------------------|-------------|----------|-----------|--------------------|---------|---------------|------|-------------------|-----------------------|--------------------|-------------|-----|----|----|-----|-----|
| CLOCK             | CLOCK       | chr 4: 56,294,068-56,315,663 | 21,596 | PFC, PL, PN         | –       | –             | –    | –                | –                     | PFC, PL, PN         | –            | –  | –  | –  | –   | –   |
| NPAS2             | NPAS2       | Chr 2: 101,436,614-101,613,291 | 176,679 | PFC, PL, PLB        | –       | –             | CFJD, CFS, SZ | ↑    | ↑                   | BP                    | BP, ALS          | –            | –  | –  | –  | –   | –   |
| NRG1              | NRG1        | Chr 8: 31,496,902-32,622,548 | 1,125,739 | PFC, PL, PLB        | –       | –             | –    | –                | –                     | –                  | NAR         | –  | –  | –  | –   | –   |
| SLC18A1           | SLC18A1     | Chr 8: 20,002,366-20,040,717 | 38,352  | PFC, PL, STN        | –       | –             | –    | –                | –                     | HT, THAL, CN      | –            | –  | –  | –  | –   | –   |
| WWC1              | WWC1        | Chr 5: 167,718,656-167,899,308 | 180,653 | HT, THAL, CN        | –       | –             | –    | –                | –                     | –                  | –            | –  | –  | –  | –   | –   |

*Data obtained from Ensembl and UCSC genome browsers. Data about Gene expression, Expression changes, Meta-analysis and Genome-wide association studies (GWAS) were extracted from BP, Bipolar disorder; ALS, Amyotrophic lateral sclerosis; Bi: BioGPS; Stanley Brain Database; HuGe-Navigator and the GWAS catalog (from the NHGRI) respectively. Abbreviations: Amyotrophic lateral sclerosis: ALS; Chronic Fatigue Syndrome: CFS; Creutzfeldt-Jakob disease: CFJD; Episodic and working memory: EwM; Hypothalamus: Hypothalamus; Caudate nucleus: CN; Cerebellum: CLB; Narcolepsy: NAR; Memory Performance: MP; Major Depression: MD; Personality: Pers; Subthalamic nucleus: STN; Thalamus: THAL.

Genotyping of nsSNPs

DNA samples, extracted from peripheral blood using a salting out method, from 171 unrelated Colombian subjects (122 women; mean age = 21.2; SD = 2.92) were genotyped for the 5 nsSNPs. None of the subjects had personal history of neuropsychiatric disorders. Written informed consent was obtained from all subjects and this study was approved by the institutional ethical committee and was conducted according to the Declaration of Helsinki Principles.

Conventional PCR (AS-PCR and T-ARMS-PCR) assays were carried out in Labinet MultiGene 96-well thermal cyclers (Labinet Internation al Inc, Edison, NJ, USA). Amplification reactions were performed in a total volume of 20 μl containing: 2 μl of genomic DNA (~ 50 ng), 2.0 mM MgCl₂, 10 X reaction Buffer, 1 mM of dNTPs, 1 M of Betaine and 0.8 U of Taq polymerase (Bioline, London, United Kingdom) (See Table III for primers concentrations). PCR products were separated in a 2% agarose gel, stained with SYBR Safe (Invitrogen, Carlsbad, CA, USA) and visualized on a UV Transilluminator. Genotyping process for NPAS2 gene was carried out in a CFX96 Real-Time PCR system (BioRad, Hercules, CA, USA). To verify the consistency in the genetic results, a random subsample (10%) of subjects was re-genotyped for each nsSNP. In addition, two different investigators confirmed and validated the results, independently checking all genotypes.

NPAS2-rs2305160 was analyzed using a specific TaqMan SNP Genotyping Assay (Applied Biosystems, Foster City, CA, USA). It was performed as follows: 2 μl (20 ng) genomic DNA, 1 X TaqMan® Genotyping Master Mix (Applied Biosystems), 1 X of TaqMan Pre-Designed SNP Genotyping Assay (C_15976652_10, Applied Biosystems) and water in a total volume of 10 μl. The amplification protocol consisted of 10-min denaturation step at 95°C (1 cycle), 95°C for 15s and 60°C for 90s (50 cycles).

Statistical analysis

PLINK program was used for Hardy-Weinberg equilibrium analysis of genotype and allele frequencies. Allele frequencies for the five nsSNPs were compared using a 2 × 2 contingency table between Colombian subjects and HapMap samples: Utah Residents with Northern and Western European ancestry (CEU); Yoruba in Ibadan, Nigeria (YRI); Han Chinese in Beijing, China (CHB); Japanese in Tokyo, Japan (JPT); and Mexican ancestry from Los Angeles, USA (MXL). Results were considered statistically significant at p < 0.05.

Results

An in silico screening of several dozens of human brain-expressed genes identified 5 nsSNPs (rs6855837; p.Leu395Ile, rs2305160; p.Thr394Ala, rs10503929; p.Met289Thr, rs2270641; p.Thr4Pro, p.Ser735Ala), with higher probabilities of being true positives (MAF > 0.01 in HapMap samples) (Table 2). These nsSNPs are located in genes (CLOCK, NPAS2, NRG1, SLC18A1 and WWC1) with a high relevance for neuropsychiatric disorders and endophenotypes (Table 1), supported by multiple lines of experimental evidence, such as genome-wide expression and association studies and meta-analysis of genetic association studies.

We designed novel cost-effective genotyping assays, based in allele-specific PCR methods, for four of these nsSNPs (Table 3). We genotyped these five nsSNPs in a sample of healthy subjects (CLB). Table 4 shows the allele and genotype frequencies...
Table 2: Details for novel nsSNPs in brain-expressed genes

| Gene   | SNP     | Location     | Reference Allele | Minor Allele | Amino acid change | MAF-CEU | MAF-YRI | MAF-CHB | MAF-JPT | MAF-MEX |
|--------|---------|--------------|------------------|--------------|-------------------|---------|---------|---------|---------|---------|
| CLOCK  | rs6855837 | chr4: 56319244 | G                | T            | p.Leu395Ile       | 0.004   | 0.326   | 0.006   | 0.006   | 0.040   |
| NPAS2  | rs2305160 | chr2: 101591304 | C                | T            | p.Thr394Ala       | 0.296   | 0.027   | 0.220   | 0.198   | 0.330   |
| NRG1   | rs10503929 | chr8: 32613983 | T                | C            | p.Met289Thr       | 0.173   | 0.000   | 0.000   | 0.000   | 0.120   |
| SLC18A1| rs2270641 | chr8: 20038466 | T                | G            | p.Thr4Pro         | 0.376   | 0.106   | 0.256   | 0.333   | 0.286   |
| WWC1   | rs3822659 | chr5: 167858372 | T               | G            | p.Ser735Ala       | 0.035   | 0.350   | 0.207   | 0.186   | 0.070   |

Data obtained from Ensembl, UCSC and HapMap. Populations: CEU: Utah Residents (CEPH) with Northern and Western European ancestry, YRI: Yoruba in Ibadan, Nigeria, CHB: Han Chinese in Beijing, China, JPT: Japanese in Tokyo, Japan, MXL: Mexican Ancestry from Los Angeles USA.

Table 3: Primers and PCR conditions for genotyping assays, based in allele-specific PCR, for novel nsSNPs in brain-expressed genes

| Method       | Allele bp | Primer sequence | Conc (µM) | Primer sequence | Conc (µM) |
|--------------|-----------|-----------------|-----------|-----------------|-----------|
| CLOCK        | AS-PCR    | Wildtype forward | 1.0       | Common reverse  | 1.0       |
| p.Leu395Ile  | 53°C - 35 cycles | 5-TCA GCA GCT TCA GGA GG-3 | 1.0       | 5-TTA AAA CGA ATG ACC AGA CA-3 | 1.0       |
| NRG1         | T-ARMS-PCR | Forward 1       | 1.0       | Reverse 1       | 1.0       |
| p.Met289Thr  | 63°C - 35 cycles | 5-GCA GAG CCT TCG TGA ACG AAA CA AGA C-3 | 1.0       | 5-TCA AGG CCG CTG ACC AAG AGG CTC-3 | 0.2       |
| SLC18A1      | T-ARMS-PCR | Forward 1       | 1.0       | Reverse 1       | 1.0       |
| p.Thr4Pro    | 63°C - 35 cycles | 5-GCA ACC GCT GGG GAG CAT CCA GAA TAGG-3 | 1.0       | 5-CAG TCC GGC CAT CAC CAT GCT CCA GA-3 | 1.0       |
| WWC1         | T-ARMS-PCR | Forward 1       | 0.4       | Reverse 2       | 0.4       |
| p.Ser735Ala  | 63°C - 35 cycles | 5-TCA AGC AGG GTG TAC ACT GCC TGC TGG G-3 | 0.4       | 5-TTT CCT TCG TCT GGA CAG ACT CCC TCC C-3 | 0.4       |

Mismatches in first nucleotide of the 3´ are indicated in bold. AT: Annealing temperature.
Table 4: Allele and genotype frequencies for five novel nsSNPs found in our samples. Minor allele frequencies for the five SNPs ranged between 4.1% (WWC1) and 24.3% (SLC18A1). All markers were in Hardy–Weinberg equilibrium (HWE) (p value >0.05).

We compared MAFs for these five nsSNPs found in the Colombian population with available data for samples from European, African, Asian and Latin American populations (extracted from HapMap). The YRI samples was the only population with significant differences for all genetic markers, with a p value less than 0.01 (see Figure 1). For CHB and JPT samples, we observed significant differences for rs10503929 (NRG1) and rs3822659 (WWC1) SNPs (p<0.002). Finally, the CLB samples showed allelic frequencies similar to those of MEX and CEU samples for all SNPs, with only one exception, rs6855837 SNP (MEX p = 0.039 and CEU p<0.001).

Discussion

Alterations in dozens of genes involved in brain and synaptic plasticity are responsible for a large number of common NPDs.11–13,26 Genome-wide association studies, mainly focused on non-functional SNPs, have found a few consistent and strong associations with NPDs.4 nsSNPs represent the most well understood group of genetic variants of possible functional importance.10 Due to their impact on function of the encoded proteins, nsSNPs are not common in brain-expressed genes.10 We identified five nsSNPs in brain-expressed genes of high relevance for neuropsychiatric disorders: CLOCK, NRG1, NPAS2, SLC18A1 and WWC1. CLOCK is a key regulator of the circadian system, and regulates the transcription of PER1, PER2, PER3, CRY1 and CRY2 genes. NRG1 is known to mediate cell–cell interactions in the nervous system and has been implicated in the etiology of schizophrenia.27 NPAS2 also has a key role in the regulation of circadian mechanisms in the brain.28 SLC18A1 encodes the vesicular amine transporter 1 (also known as VMAT1), important for the functioning of monoaminergic systems.29 WWC1 is involved in long-term potentiation mechanisms and synaptic plasticity.30 There is previous experimental evidence, arising from genome-wide expression and association studies, implicating the involvement of these genes in several neuropsychiatric disorders and endophenotypes, such as schizophrenia, mood disorders or memory performance.27–30 However, a large part of the previously studied variants in these genes were neutral polymorphisms.3 Similar the SNP detection method has wide applications and can be used for diagnosis of other untreated neurodegenerative disorders such as ALS, Parkinson’s and AMD as shown.31–37

Cost-effective genotyping assays described in this work for these nsSNPs could be interesting methodological alternatives for researchers in developing and developed countries. These five nsSNPs in brain-expressed genes represent interesting candidates to analyze in future genetic studies of neuropsychiatric disorders and related endophenotypes, a focused synaptogenomics approach.12,13,26

Acknowledgements

This study was supported by grants from Universidad Antonio Nariño (VCTI-UAN).

This article complies with International Committee of Medical Journal Editor’s uniform requirements for manuscript.
Conflict of interest: None.
Competing interests: None, Source of funding: Universidad Antonio Nariño (VCTI-UAN)

Received Date : 29 January 2014; Revised Date : 16 May 2014; Accepted Date : 10 July 2014

References
1. Demyttenaere K, Bruffaerts R, Posada-Villa J, et al. Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys. JAMA. 2004; 291(21): 2581–2590.
2. Uhl GR, Grow RW. The burden of complex genetics in brain disorders. Arch Gen Psychiatry. 2004; 61(3): 223–229.
3. Burmeister M, Mcinnis MG, Zollner S. Psychiatric genetics: progress amid controversy. Nat Rev Genet. 2008; 9(7): 527–540.
4. Cristino AS, Williams SM, Hawi Z, et al. Neurodevelopmental and neuropsychiatric disorders represent an interconnected molecular system. Mol Psychiatry. 2013; 18(3): 294–301.
5. Hindorff LA, Sethupathy P, Junkins HA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci. USA. 2009; 106(23): 9362–9367.
6. Cross-Disorder Group of the Psychiatric Genomics C, Genetic Risk Outcome of Psychosis C. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet. 2013; 381(9875): 1371–1379.
7. Cross-Disorder Group of the Psychiatric Genomics C, Lee SH, Ripke S, et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nat Genet. 2013; 45(9): 984–994.
8. Gu L, Long J, Yan Y, et al. HTR2A-1438A/G polymorphism influences the risk of schizophrenia but not bipolar disorder or major depressive disorder: a meta-analysis. J Neurosci Res. 2013; 91(5): 623–633.
9. Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. Nat Genet. 2008; 40(6): 695–701.
10. Yates CM, Sternberg MJ. The Effects of Non-Synonymous Single Nucleotide Polymorphisms (nsSNPs) on Protein-Protein Interactions. J Mol Biol. 2013; 425(21): 3949–3963.
11. Grant SG. Synaptopathies: diseases of the synaptome. Curr Opin Neurobiol. 2012; 22(3): 522–529.
12. Forero DA, Arboleda GH, Vasquez R, et al. Candidate genes involved in neural plasticity and the risk for attention-deficit hyperactivity disorder: a meta-analysis of 8 common variants. J Psychiatry Neurosci. 2009; 34(5): 361–366.
13. Forero DA, Casadesus G, Perry G, et al. Synaptic dysfunction and oxidative stress in Alzheimer’s disease: emerging mechanisms. J Cell Mol Med. 2006; 10(3): 796–805.
14. Haider S, Ballester B, Smedley D, et al. BioMart Central Portal—unified access to biological data. Nucleic Acids Res. 2009; 37(Web Server issue): W23–27.

Fig. 1: Radial plot of minor allele frequencies for 5 novel nsSNPs in six different populations. Colombian population from Bogotá (CLB); Utah Residents with Northern and Western European ancestry (CEU); Yoruba in Ibadan, Nigeria (YRI); Han Chinese in Beijing, China (CHB); Japanese in Tokyo, Japan (JPT); Mexican Ancestry from Los Angeles USA (MXL).
15. International Hapmap C. The International HapMap Project. Nature. 2003; 426(6968): 789–796.
16. Wu C, Macleod I, Su Al. BioGPS and MyGene.info: organizing online, gene-centric information. Nucleic Acids Res. 2013; 41(Database issue): D561–565.
17. Higgs BW, Elashoff M, Richman S, et al. An online database for brain disease research. BMC Genomics. 2006; 7: 70.
18. Yu W, Gwinn M, Clyne M, et al. A navigator for human genome epidemiology. Nat Genet. 2008; 40(2): 124–125.
19. Meyer LR, Zweig AS, Hinrichs AS, et al. The UCSC Genome Browser database: extensions and updates 2013. Nucleic Acids Res. 2013; 41(Database issue): D64–69.
20. Wangkumhang P, Chaichoompu K, Ngamphiw C, et al. WASP: a Web-based Allele-Specific PCR assay designing tool for detecting SNPs and mutations. BMC Genomics. 2007; 8(275).
21. Ye S, Dhillon S, Ke X, et al. An efficient procedure for genotyping single nucleotide polymorphisms. Nucleic Acids Res. 2001; 29(17): E88–88.
22. You FM, Huo N, Gu YQ, et al. BatchPrimer3: a high throughput web application for PCR and sequencing primer design. BMC Bioinformatics. 2008; 9(253).
23. Ojeda DA, Perea CS, Nino CL, et al. A novel association of two non-synonymous polymorphisms in PER2 and PER3 genes with specific diurnal preference subscales. Neurosci Lett. 2013; 553C (52-56).
24. Beskow LM, Burke W, Merz JF, et al. Informed consent for population-based research involving genetics. JAMA. 2001; 286(18): 2315–2321.
25. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81(3): 559–575.
26. Forero DA, Van Der Ven K, Callaerts P, et al. miRNA genes and the brain: implications for psychiatric disorders. Hum Mutat. 2010; 31(11): 1195–1204.
27. Buonanno A. The neuregulin signaling pathway and schizophrenia: from genes to synapses and neural circuits. Brain Res Bull. 2010; 83(3-4): 122–131.
28. Buhr ED, Takahashi JS. Molecular components of the Mammalian circadian clock. Handb Exp Pharmacol. 2013; (217): 3–27.
29. Lawal HO, Krantz DE. SLC18: Vesicular neurotransmitter transporters for monoamines and acetylcholine. Mol Aspects Med. 2013; 34(2-3): 360–372.
30. Makuch L, Volk L, Anggono V, et al. Regulation of AMPA receptor function by the human memory-associated gene KIBRA. Neuron. 2011; 71(6): 1022–1029.
31. Gupta PK, Prabhakar S, Abburi C, et al. Vascular endothelial growth factor-A and CCL2 genes are upregulated in peripheral blood monocytes in Indian Amyotrophic lateral sclerosis patients. Journal of Neuroinflammation. 2011; 8: 114.
32. Gupta PK, Prabhakar S, Sharma S, et al. VEGF-a and CCL2 in Indian ALS patients. Journal of Neuroinflammation. 2011; 8:47.
33. Anand A, Gupta PK, Sharma NK et al. sVEGR1 as a novel marker of ALS in north Indian ALS population. European Journal of Neurology. 2012; 19(5): 788–92.
34. Vinish M, Anand A, Prabhakar S. Altered oxidative stress levels in Indian Parkinson’s disease patients with PARK2 mutations. Acta Biochem Polonica. 2011; 58, 2: 165–169.
35. Sharma NK, Gupta A, Prabhakar S, et al. CC Chemokine receptor-3 as new target for age related macular degeneration. Gene. 2013; 523: 106–111.
36. Anand A, Sharma NK, Prabhakar S, et al. Single Nucleotide Polymorphisms in MCP-1 and its Receptor are associated with the risk of Age Related Macular Degeneration. PLoS ONE. 2012; 7(11): doi:10.1371/journal.pone.0049905.
37. Sharma NK, Gupta A, Prabhakar S, et al. Single nucleotied polymorphism and serum levels of VEGFR2 are associated with age related macular degeneration. Curr Neurovasc Res. 2012; 9(4): 256–65.