Progress in Neuropsychopharmacology & Biological Psychiatry 77 (2017) 164–171

Contents lists available at ScienceDirect

Progress in Neuropsychopharmacology & Biological Psychiatry

journal homepage: www.elsevier.com/locate/pnp

Spatial and temporal expression patterns of genes around nine neuroticism-associated loci

Kazutaka Ohi*, Takamitsu Shimada, Toshiki Yasuyama, Kohei Kimura, Takashi Uehara, Yasuhiro Kawasaki

Department of Neuropsychiatry, Kanazawa Medical University, Ishikawa, Japan

ARTICLE INFO

Keywords:
Neuroticism
Genome-wide association study
Temporal-spatial expression
Brain
Glutamate receptor activity

ABSTRACT

Neuroticism is a high-order personality trait. Individuals with higher neuroticism have increased risks of various psychiatric disorders and physical health outcomes. Neuroticism is related to physiological differences in the brain. A recent genome-wide association study identified nine distinct genomic loci that contribute to neuroticism. Brain development and function depend on the precise regulation of gene expression, which is differentially regulated across brain regions and developmental stages. Using multiple publicly available human post-mortem databases, we investigated, in brain and non-brain tissues and across several developmental life stages, the spatial and temporal expression patterns of genes arising from nine neuroticism-associated loci. Functional gene-network analysis for neuroticism-associated genes was performed. The spatial expression analysis revealed that the nearest genes (*GRIK3*, *SRP9*, *KLHL2*, *PTPRD*, *ELAVL2*, *CRHR1* and *CELF4*) from index single-nucleotide polymorphisms (SNPs) at the nine loci were intensively enriched in the brain compared with their representation in non-brain tissues ($p < 1.56 \times 10^{-3}$). The nearest genes associated with the glutamate receptor activity network consisted mainly of *GRIK3* (*FDR* $q = 4.25 \times 10^{-2}$). The temporal expression analysis revealed that the neuroticism-associated genes were divided into three expression patterns: *KLHL2*, *CELF4* and *CRHR1* were preferentially expressed during postnatal stages; *PTPRD*, *ELAVL2* and *MFHAS1* were expressed during prenatal stages; and the other three genes were not expressed during specific life stages. These findings suggest that the glutamate network might be a target for investigating the neurobiological mechanisms underlying susceptibilities to higher neuroticism and several psychiatric disorders and that neuroticism is mediated by genes specifically expressed in the brain during several developmental stages.

1. Introduction

Neuroticism is a high-order personality trait characterized by vulnerability to emotional instability and self-consciousness (Costa and McCrae, 1992) and a tendency toward negative responses to threat, frustration or loss (Lahey, 2009). Neuroticism can be assessed easily via self-report questionnaires, such as the Eysenck Personality Questionnaire (EPQ) and the NEO Five Factor Inventory (NEO-FFI), in which the assessments of neuroticism are highly correlated with one another (Gow et al., 2005). There is growing evidence that neuroticism is a psychological and genetic trait of profound public health significance (Lahey, 2009; Smith et al., 2016). Individuals with higher levels of neuroticism have an increased risk of several types of psychiatric disorders, including major depressive disorder (MDD), schizophrenia (SCZ), bipolar disorder (BIP), anorexia nervosa (AN), anxiety disorder, substance abuse disorder and borderline personality disorder (Barnett et al., 2011; Distel et al., 2009; Kendler and Myers, 2010; Kotov et al., 2010; Ohi et al., 2016b). Higher neuroticism has also been associated with an increased risk of a wide range of physical disorders, including coronary artery disease and diabetes (Goodwin et al., 2006; Jokela et al., 2014). In addition, individuals who show higher neuroticism tend to make greater use of mental health services, regardless of whether they have a mental disorder (ten Have et al., 2005), creating a heavy economic burden (Cuijpers et al., 2010). Therefore, studying neuroticism is not only important for understanding a key dimension of personality but also for illuminating the etiology of a range of psychiatric and physical disorders (Smith et al., 2016).

Individual differences in neuroticism are relatively stable across the life course (Wray et al., 2007). Twin and adoption studies suggest that approximately 40% of the trait variance for neuroticism is heritable (Keller et al., 2005; Smith et al., 2016; van den Berg et al., 2014; Wray et al., 2007; Yamagata et al., 2006), and 15–37% of the variance for

* Corresponding author at: Department of Neuropsychiatry, Kanazawa Medical University, 1-1 Daigaku, Uchinada, Ishikawa 920-0293, Japan.
E-mail address: ohi@kanazawa-med.ac.jp (K. Ohi).

http://dx.doi.org/10.1016/j.pnpbp.2017.04.019
Received 24 February 2017; Received in revised form 14 April 2017; Accepted 18 April 2017
Available online 19 April 2017

neuroticism is explained by common genetic variants used in genome-wide association studies (GWASs) (de Moor et al., 2015; Smith et al., 2016; van den Berg et al., 2014). Recent GWASs have substantially increased the number of genes that are known to be related to the pathogenesis of neuroticism, although the results of some previous GWASs have been limited by relatively small sample sizes (de Moor et al., 2015; Kim et al., 2015; Okbay et al., 2016; Smith et al., 2016). A recent large meta-analysis of GWASs based on the UK Biobank, the Generation Scotland: Scottish Family Health Study (GS: SFHS) and the Queensland Institute of Medical Research Berghofer Medical Research Institute (QIMR) data from > 108,000 individuals who completed a neuroticism questionnaire and provided DNA for GWAS identified nine distinct neuroticism-associated genomic loci (Smith et al., 2016). Neuroticism is influenced by many genetic variants with small effects, i.e., a polygenic effect. Many psychiatric and physical illnesses also show moderate heritability (Polderman et al., 2015), and twin studies have shown that there are considerable genetic overlaps between neuroticism and these disorders (Hettema et al., 2006; Johnson and Krueger, 2005; Statham et al., 1998). Shared genetic influences between neuroticism and these various psychiatric and physical illnesses, such as MDD, SCZ, AN and coronary artery disease, have been identified (Gale et al., 2016; Harris et al., 2016). Therefore, there is a high level of pleiotropy in associations between neuroticism and psychiatric and physical health illnesses.

The nine neuroticism-associated loci—chromosomes 1p34.3, 1q42.12, 3q13.13, 4q32.3, 8p23.1, 9p24.1, 9p21.3, 17q21.31 and 18q12.2—contain at least 20 genes, including glutamate receptor ionotropic kainate 3 (*GRIK3*) and corticotropin-releasing hormone receptor 1 (*CRHR1*) (Smith et al., 2016). In general, single nucleotide polymorphisms (SNPs) associated with common diseases and phenotypes identified by GWAS are enriched for regulatory regions of the genome (Nicolae et al., 2010), and the SNPs could contribute to disease and phenotype susceptibility by altering the expression of the proxy gene from the SNPs (Fu et al., 2012; Powell et al., 2013). The molecular circuits related to neuroticism may be mediated by the functional downstream convergence of genes in these loci. Brain development and function depend on the precise regulation of gene expression (Rakic, 2009) and gene expression is differentially regulated across brain regions and developmental time points (Kang et al., 2011). Therefore, it is a critical step in the elucidation of the genetic basis of neuroticism to investigate the spatio-temporal dynamics of the transcriptome of the human brain for genes in the loci. Although the molecular and brain mechanisms underlying susceptibility to high neuroticism are unknown, several studies suggest relationships between higher levels of neuroticism and reduced cerebral volumes and attenuated brain functions of the dorsolateral prefrontal cortex (DLPFC) (Bjornebekk et al., 2013; DeYoung et al., 2010; Dima et al., 2015; Kapogiannis et al., 2013; Wright et al., 2006). The DLPFC is a major component of the high-order association cortex engaged in attentional and complex cognitive operations (MacDonald et al., 2000; Miller and Cohen, 2001). For example, the transcriptional activity of the gene-set at the 108 schizophrenia-related loci was shown to be relatively enriched during prenatal life compared with its transcription activity in postnatal life in the DLPFC (Birnbaum et al., 2014; Birnbaum et al., 2015; Ohi et al., 2016a). In addition, we investigated the gene-level temporal and anatomical expression patterns of the 108 schizophrenia-related loci in multiple brain and non-brain tissues and found that some genes in the loci were preferentially expressed in the prefrontal cortex during the fetal or young adult life stages (Ohi et al., 2016a). Considering that high neuroticism indexes a risk constellation that exists prior to the development and onset of any common psychiatric disorders (Jeronimus et al., 2016), the genes in the neuroticism-associated loci may also be preferentially expressed during early life stages.

The aim of the present study was to investigate the spatial and temporal expression patterns in the genes around the nine neuroticism-associated genetic loci. We hypothesized that the nearest protein-coding genes from index SNPs at the nine neuroticism-associated loci were preferentially expressed in brain tissues rather than non-brain tissues and that the nearest genes were expressed during specific developmental stages.

2. Methods

First, the list of the nine index SNPs at nine neuroticism-associated loci was extracted from the original GWAS [Table 2 (Smith et al., 2016)]. The nine nearest protein-coding genes from the index SNPs at nine neuroticism-associated loci and protein-coding genes in the neuroticism-associated region ± 20 kb were extracted using UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly (https://www.genome.ucsc.edu/cgi-bin/hgGateway/).

To investigate whether the genes in the nine loci are specifically expressed in brain tissues, spatial expression of the genes in human tissues were extracted from RNA-Seq datasets of the Genotype-Tissue Expression (GTEx) [GTEx Analysis Release V6p (dbGaP Accession phs000424.v6.p1) http://www.gtexportal.org/home/(Lonsdale et al., 2014)]. The GTEx project, sponsored by the NIH Common Fund, was established to research the correlation between human genetic variation and tissue-specific gene expression by collecting high-quality post-mortem biospecimens in non-diseased (normal) adult individuals. The GTEx contains gene expression profiles from > 50 different non-diseased tissue types from hundreds of donors (http://www.gtexportal.org/home/tissueSummaryPage). Detailed V6p donor information at the GTEx were described in https://gtexportal.org/home/tissueSummaryPage#donorInfo. Briefly, age at death of 544 donors (357 males/187 females) was as follows; age 20–29, n = 44 (8.1%); 30–39, n = 38 (7.0%); 40–49, n = 92 (16.9%); 50–59, n = 188 (34.6%); 60–69, n = 172 (31.6%), and 70–79, n = 10, (1.8%). Most of donors were white (84.3%), followed by African American (13.7%) and Asian (1%). Causes of death were different between ages 20–39 and 60–71; age 20–39, traumatic injury 54.3%, cerebrovascular 16.1%, heart disease 9.9%, liver, renal and respiratory 3.7%, and neurological 3.7%; age 60–71, traumatic injury 5.1%, cerebrovascular 24.7%, heart disease 37.6%, liver, renal and respiratory 16.3%, and neurological 2.3%. The mean expression in brain tissues is calculated from 13 brain tissues: amygdala, anterior cingulate cortex, caudate, cerebellar hemisphere, cerebellum, cortex, frontal cortex, hippocampus, hypothalamus, nucleus accumbens, putamen, spinal cord and substantia nigra. The mean expression in non-brain tissues is calculated from 37 non-neural tissues. The differences in gene expression between brain and non-brain tissues were assessed with a linear regression model using each expression as the dependent variable and tissue type (0 = non-brain tissues; 1 = brain tissues) as the independent variable. The p values for the analysis were adjusted using the Bonferroni correction, and p values $< 1.56 \times 10^{-3}$ (α = 0.05/32 genes) were considered significant.

The spatial expression distributions in the human tissues were confirmed using Gene Enrichment Profiler (GEP) [http://xavierlab2.mgh.harvard.edu/EnrichmentProfiler/(Benita et al., 2010)]. The database was built by collecting publicly available microarray data profiled on Affymetrix U133A chips. Raw CEL files were obtained and normalized as a single experiment. In the GEP database, the expression enrichment of a query gene was calculated based on a reference set obtained from 126 normal tissues and cell types represented by ~650 microarrays on Affymetrix U133A chips (Benita et al., 2010; Ohi et al., 2016a).

A functional gene-network analysis was performed using GeneMANIA software (Multiple Association Network Integration Algorithm; http://www.genemania.org/), which was used for the following: to identify whether there are known interactions between the query genes, and if so, the mechanism for their interaction; to add extra genes strongly connected to the query genes; to search for genetic, physical, pathway and co-expression networks; and to identify func-