



**HAL**  
open science

## A genome-wide linkage study of individuals with high scores on NEO personality traits

Cornelia M van Duijn, Najaf Amin, Maaïke Schuur, Elena S Gusareva, Aaron J Isaacs, Yurii Aulchenko, Anatoly V Kirichenko, Irina V Zorkoltseva, Tatiana I Axenovich, Ben Oostra, et al.

### ► To cite this version:

Cornelia M van Duijn, Najaf Amin, Maaïke Schuur, Elena S Gusareva, Aaron J Isaacs, et al.. A genome-wide linkage study of individuals with high scores on NEO personality traits. *Molecular Psychiatry*, 2011, 10.1038/mp.2011.97 . hal-00668053

**HAL Id: hal-00668053**

**<https://hal.science/hal-00668053>**

Submitted on 9 Feb 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **A genome-wide linkage study of individuals with high scores on NEO personality traits**

Najaf Amin<sup>1</sup> PhD, Maaïke Schuur<sup>1</sup> PhD, Elena S. Gusareva<sup>2</sup> PhD, Aaron Isaacs<sup>1,3</sup> PhD,  
Yurii S. Aulchenko<sup>1</sup> PhD, Anatoly V. Kirichenko<sup>4</sup> PhD, Irina V. Zorkoltseva<sup>4</sup> MSc,  
Tatiana I. Axenovich<sup>4,5</sup> PhD, Ben A. Oostra<sup>1,6</sup> PhD, A. Cecile J.W. Janssens<sup>1</sup> PhD,  
Cornelia M. van Duijn<sup>1,3\*</sup> PhD

<sup>1</sup> Department of Epidemiology, Erasmus University Medical Center, Rotterdam, the Netherlands

<sup>2</sup> Department of Molecular and Cellular Immunology, Institute of Molecular Genetics AS CR, v.v.i. Prague, Czech Republic

<sup>3</sup> Centre for Medical Systems Biology, Leiden, the Netherlands

<sup>4</sup> Institute of Cytology & Genetics, Russian Academy of Science, Novosibirsk, Russia

<sup>5</sup> Department of Cytology & Genetics, Novosibirsk State University, Novosibirsk, Russia

<sup>6</sup> Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, the Netherlands

\* Correspondence to: Cornelia M. van Duijn,

Department of Epidemiology, Erasmus University Medical Center Rotterdam,

Dr. Molewaterplein 50, P.O. box 2040, 3000 CA, Rotterdam,

The Netherlands,

E-mail: [c.vanduijn@erasmusmc.nl](mailto:c.vanduijn@erasmusmc.nl)

Tel: +31 10 7043394

Fax: +31 10 7044675

**Running Title:** A genome wide linkage study of personality traits

## **ABSTRACT**

The NEO Five Factor Inventory divides human personality traits into five dimensions: neuroticism, extraversion, openness, conscientiousness and agreeableness. In this study, we sought to identify regions harboring genes with large effects on the five NEO personality traits by performing genome-wide linkage analysis of individuals scoring in the extremes of these traits (> 90<sup>th</sup> percentile). Affected-only linkage analysis was performed using an Illumina 6K linkage array in a family-based study, the Erasmus Rucphen Family study (ERF). We subsequently determined whether distinct, segregating haplotypes found with linkage analysis were associated with the trait of interest in the population. Finally, a dense SNP genotyping array (Illumina 318K) was used to search for copy number variations (CNVs) in the associated regions. In the families with extreme phenotype scores, we found significant evidence of linkage for conscientiousness to 20p13 (rs1434789, LOD=5.86) and suggestive evidence of linkage (LOD > 2.8) for neuroticism to 19q, 21q and 22q, extraversion to 1p, 1q, 9p and 12q, openness to 12q and 19q, and agreeableness to 2p, 6q, 17q and 21q. Further analysis determined haplotypes on 21q22 for neuroticism (p-values = 0.009, 0.007), on 17q24 for agreeableness (marginal p-value = 0.018) and on 20p13 for conscientiousness (marginal p-values = 0.058, 0.038) segregating in families with large contributions to the LOD scores. No evidence for CNVs in any of the associated regions was found. Our findings imply that there may be genes with relatively large effects involved in personality traits, which may be identified with next generation sequencing techniques.

**Key words:** NEO, linkage, family, pedigree, haplotype, CNV

## INTRODUCTION

Research into human personality has always been an integral part of behavioral science. Numerous epidemiological studies show that personality traits are important risk factors for many psychiatric and non-psychiatric disorders (1-9). Personality traits are thought to be complex and determined largely by genetic factors (10). Over the last few decades, various models for the assessment of personality have been developed, including the five factor model that divides human personality into five basic, universal types(11). The NEO Five Factor Inventory (NEO-FFI) assesses these fundamental traits which include (i) neuroticism – a tendency to experience negative emotions, (ii) extraversion – a predisposition towards enthusiasm, positive emotions and action, (iii) openness – the extent of an individual’s intellectual curiosity, (iv) agreeableness – a propensity for cooperation and harmony, and (v) conscientiousness – the inclination to control, regulate and direct impulses (12). Although the five scales are designed to be orthogonal, correlations appear due to self-reporting (13).

To date, genetic research has focused on neuroticism, primarily because of its association with disease and mortality (1-4, 6-9, 14-15). Recent studies suggested that other personality traits might also be important risk factors for different diseases; extraversion is a determinant of bipolar disorder (16), conscientiousness is associated with Alzheimer’s disease (AD) (17), and several studies demonstrated that low scores in novelty seeking, a measure of creativity (18), are associated with a higher risk of developing Parkinson’s disease (19-21). There is increasing interest in the relationship between personality traits and treatment outcomes (22). Additionally, openness and conscientiousness have been associated with higher academic achievement and better work performance(23).

Twin studies provided heritability estimates of the NEO personality traits ranging from 33 to 61% (10, 24-27). Despite the high heritability, the genetics of these traits are not well understood. Candidate gene studies associated neuroticism with SLC6A4 (28-30) and TPH1 (31) and agreeableness, extraversion and conscientiousness with ADH4 and CHRM2 (32). Genome-wide association studies suggested relationships between the CLOCK gene and agreeableness and the MAMDC1 and NKAIN2 genes and neuroticism, however replication efforts were inconsistent (33-36). There have been seven published genome-wide linkage scans for quantitative measures of personality traits, evaluated using either Eysenck's personality questionnaire (EPQ) (37) (measures neuroticism, extraversion and psychoticism) or the Tri-dimensional Personality Questionnaire (TPQ) (18, 38) (assesses novelty seeking, harm avoidance and reward dependence), as continuous outcomes including six exclusively for neuroticism (39-45). These studies provided evidence of linkage between neuroticism/harm avoidance and several different loci (**Table 1**). Of the linked regions 1q, 8p, 11q, 12q and 18q were replicated for neuroticism in various studies (39-40, 42, 44).

Genome-wide association studies (GWAS) and linkage with quantitative outcomes target genes with small to moderate effects. We hypothesized that there may be genes with relatively large effects underlying the NEO personality traits, which may be rare and may have gone undetected in the GWAS, and that the regions containing these might be localized by studying persons scoring in the extremes for these phenotypes. We searched for such regions for all of the NEO-FFI personality traits by performing affected-only linkage analysis and haplotype association analysis. To translate the findings to the full distribution of the scale, we subsequently associated the haplotypes linked to the extremes to the outcomes on a quantitative scale.

## **MATERIAL AND METHODS**

### *Study population*

The study sample consisted of 2657 individuals who participated in the Erasmus Rucphen Family (ERF) study. The ERF cohort was ascertained from a genetically isolated region in the southwest of the Netherlands. The study population descended from 20 related couples that lived in the isolate between 1850 and 1900; genealogical data, available from 1750, captures all individuals in a single 23 generation pedigree including more than 23,000 individuals. Pedigree members 18 years and older were invited to participate. Spouses were invited only for descendants who had children older than 18 years.

### *Personality Assessment*

The five basic personality types (neuroticism, extraversion, openness, agreeableness and conscientiousness) were assessed using the NEO-FFI (12). The short form of this inventory consists of 60 items, 12 for each trait. Answers for each question were given on a 5-point scale and the total scores for each trait ranged from 12 to 60. Because long questionnaires are often repetitious in nature and may cause the respondent to lose interest, the scores were considered invalid if the respondents gave the same response for at least 20 consecutive items ( $n = 114$ ) or if they answered less than 9 questions in total ( $n = 176$ ) on a single scale.

### *Genotyping*

For all participants, genomic DNA was extracted from peripheral venous blood utilizing the salting out method (46). For genome-wide linkage analysis, genotyping was performed using an Illumina 6K linkage panel that includes markers distributed evenly

across the genome (median inter-marker distance = 301 kb). We used 5250 single nucleotide polymorphisms (SNPs) after quality control (call rate > 95%) and the exclusion of X-chromosome SNPs. The genotyping was performed at the Centre National de Génotypage in France according to the manufacturer's protocol. For copy number variation analysis, we used an Illumina 318K SNP array. For this micro-array, genotyping was performed according to the manufacturer's instructions at the Leiden Genome Technology Center at the Leiden University Medical Center.

### Statistical Analyses

#### *Linkage analysis*

We performed affected only linkage analysis by defining the affected as those people who scored above the 90<sup>th</sup> percentile for a given trait. The 90<sup>th</sup> percentile was chosen to maximize power based on our simulation studies evaluating different cut-offs (47). The simulations suggested that, for different sample sizes, power is highest using the 10% cut-off. General characteristics of the samples selected for linkage analysis are given in **Table 2** and the correlation structure of these five traits at the population level is given in **Supplementary Table 1**.

The affected individuals for each trait were linked via a single large pedigree which was later clustered into small ( $\leq 18$  bits) families for analytical purposes using pedigree clustering software(48). We performed multipoint non-parametric linkage (NPL) using the 'all' option in MERLIN(49), which uses the Kong & Cox(50) linear model to evaluate the presence of linkage. We also performed parametric linkage analyses using both dominant and recessive models with complete penetrance and a disease allele frequency of 0.01 using MERLIN. Allele frequencies were estimated from

all genotyped individuals in the pedigrees for each trait. For all five sets of pedigrees, 1249 Mendelian inconsistencies were observed for which we set the genotypes of the individuals from the whole family to missing.

### *Genome-wide simulations*

Genome-wide significance thresholds were determined empirically by performing 500 genome-wide simulations under the null hypothesis of no linkage and using NPL with the ‘all’ option. Marker allele frequencies were set to the ones observed in our data. Pedigree clustering was performed using individuals from the upper 10% of of the distribution for each quantitative trait. For each screen, the highest log of odds (LOD) score was recorded. The cumulative density function of the simulated maximum LOD scores approximates the distribution of the genome-wide type I error rate. Our simulations showed that a LOD score of 4.1 corresponds to a genome-wide type I error rate of 5% and that a LOD of 2.8 corresponds to a genome-wide type I error of 50%. The ‘all’ option was chosen for these simulations since it provides the most conservative threshold keeping in mind the structure of our pedigrees (51).

### *Haplotype construction & Association analysis*

For regions showing significant or suggestive evidence of linkage, haplotypes were constructed for the families with the largest contributions to the LOD score (contributing a LOD score  $\geq 1$ ) in SIMWALK (52-53). Haplotype association analysis was conducted using the “e” option in FBAT (54-55) to test for association in the presence of linkage, once the segregating haplotype was ascertained. For these analyses, we used SNPs from the Illumina 6K panel since the denser chip (Illumina 318K) was not available for all



relatives. Association analyses were run with an additive model when the highest LOD scores were observed with NPL. A recessive or a dominant model was used if the highest LOD score was obtained under one of those models. For binary trait haplotype association analysis, we used the highest scorers (> 90<sup>th</sup> percentile) as cases. In addition, we performed a quantitative trait haplotype association analysis in the total population. Frequencies of the haplotypes ( $f$ ) were estimated from the total population. Bonferroni corrections were applied to the nominal p-values for the number of haplotypes tested per region. To get estimates of the odds ratios (OR) and effect sizes, we performed logistic and linear regression analyses in SPSS for the haplotypes that showed evidence of association in FBAT, using the same model under which association was observed.

#### *CNV analysis*

In the regions that showed evidence of linkage and association, we searched for common copy number variants (CNVs) in individuals who were carrying the haplotype using the CNV partition tool 2.4.4 in Beadstudio. We also used probe intensities (as measured by Log R Ratios) and genotype frequencies of the SNP probes (as measured by B-allele frequencies) to visually detect CNVs in the associated regions (56-57).

## **RESULTS**

We report the Kong & Cox LOD scores for NPL and heterogeneity LOD scores (HLOD) for the parametric linkage analyses. **Table 3** details regions with significant and suggestive linkage in either the parametric or the NPL analyses. **Table 4** provides the results of haplotype association analysis for the identified haplotypes including ORs.

### *Neuroticism*

NPL analysis revealed two regions with suggestive evidence of linkage (LOD > 2.82) on chromosomes 19q13 and 21q22, and the recessive model gave one suggestive locus on chromosome 22q11 (**Figure 1A, Table 3**). None of the LOD scores reached genome-wide significance. We constructed haplotypes for these three regions (data not shown) in the families with large contributions to the LOD scores. For 21q22, we identified two haplotypes (H21<sub>N1</sub> & H21<sub>N2</sub> with frequencies of 0.47 and 0.26). H21<sub>N1</sub> was shared by six of the nine individuals belonging to 3 high LOD score families (**Supplementary Figure 1**), five of which were homozygous carriers. Two of the remaining three were carriers of H21<sub>N2</sub>. H21<sub>N</sub> covered 1.5 centiMorgan (cM) between rs2835574 and rs1016694. Both H21<sub>N1</sub> and H21<sub>N2</sub> showed nominally significant evidence of association in the binary analysis based on the total ERF population (p-value = 0.009, OR = 1.2 and p-value = 0.007, OR = 0.76). Both haplotypes remained significant after correction for multiple testing. H21<sub>N</sub> lies in the Down's syndrome critical region and covers 10 genes including *TTC3*, *DSCR9*, *DSCR3*, *DYRK1A*, *KCNJ6* and *PIGP* (**Table 4**).

### *Extraversion*

Suggestive evidence of linkage was observed for chromosomes 1p31, 1q24, 9p24 and 12q24 in the NPL analysis (**Figure 1B, Table 3**). The LOD score for chromosome 12 (4.01) approached genome-wide significance; a distinct haplotype spanning the region (H12<sub>E</sub>) was observed in three high LOD score families (data not shown). H12<sub>E</sub> covered a 1 cM region between rs1278602 and rs7960480 and was shared by seven out of nine individuals with an extreme phenotype in those families. Three of the seven were

homozygous and 4 were heterozygous for the haplotype. H12<sub>E</sub> was the most frequent haplotype in the region ( $f = 0.36$ ) but was not significant when tested for association (**Table 4**).

### *Openness*

For openness, there were two regions with a LOD > 2.82, one on chromosome 12q24, overlapping with extraversion, and one on 19q13, overlapping with neuroticism (**Figure 1C, Table 3**). This overlap is in line with the correlation structure of the traits (**Supplementary Table 1 and Table 2**). Those scoring higher than the 90<sup>th</sup> percentile of the openness distribution were also significantly different in their means for neuroticism and extraversion when compared with individuals with lower openness scores (**Table 2**). On 12q24, a unique haplotype (H12<sub>O</sub>) was observed in 5 out of 14 individuals with extreme phenotypes from 4 contributing families (data not shown). H12<sub>O</sub> covered a roughly 1 cM region between rs1278602 and rs7960480 and was relatively rare ( $f = 0.05$ ) in the total ERF population. The haplotype differed from the one observed for extraversion in the same region. No evidence of association was observed for H12<sub>O</sub> in either binary or quantitative trait haplotype association analysis (**Table 4**).

### *Agreeableness*

Suggestive evidence of linkage was observed on chromosomes 2p25, 6q27, 17q24, and 21q22 in NPL analysis (**Figure 1D, Table 3**). The 21q22 region was the same as that observed for neuroticism (**Table 3**), in accordance with the correlation structure (**Supplementary Table 1**) and the findings in Table 2, where those who scored high on agreeableness also scored low on neuroticism when compared with the less agreeable

individuals. In contrast to the other NEO traits, where there were only a few families making large contributions to the LOD scores in the linked regions, for agreeableness there was no single family contributing a LOD score  $> 1$ . Instead, there were multiple families each contributing to the LOD score. Construction of haplotypes in the three families with the largest LOD scores revealed haplotypes for chromosomes 6 (H6<sub>A</sub>), 17 (H17<sub>A</sub>) and 21 (H21<sub>A</sub>) (not shown). Of these, only H17<sub>A</sub> (**Supplementary Figure 2**) showed significant evidence of association in the quantitative trait haplotype association analysis ( $p = 0.018$ , the change in agreeableness associated with carriership of 1 extra allele ( $\beta$ ) = 0.31). H17<sub>A</sub> had a frequency of 0.17, covers a 0.3 cM region and is flanked by rs411602 and rs1981647 (**Table 3**). The region covered by the haplotype did not contain any genes, but nearby genes include *KCNJ2*, *KCNJ16* and *MAP2K6*.

### *Conscientiousness*

A significant LOD score was observed for chromosome 20 (HLOD = 5.86, at rs1434789 under the recessive model and 5.28 under the dominant model and a LOD of 4.47 under NPL) (**Figure 1E, Table 3**). Construction of haplotypes in six families revealed two major segregating haplotypes (H20<sub>C1</sub> & H20<sub>C2</sub>, **Supplementary Figure 3**) with frequencies of 0.339 and 0.231. Under a recessive model, H20<sub>C2</sub> showed borderline significance at a nominal level ( $p$ -value = 0.058, OR = 1.4) (**Table 4**). No association was observed under the dominant or additive models. The haplotype covers a half cM region flanked by rs1434789 and rs434609. The region contains 5 genes, all belonging to the beta defensin protein family (DEFB). When tested for association under a recessive model, the top 10% of scorers were all included in only 8 families. To improve statistical power, we extended the association analysis to those scoring in the highest 30% of the

distribution, enlarging the sample to 25 families. In this analysis, the haplotype became marginally significant ( $p = 0.038$ ) but did not survive Bonferroni correction.

### *CNV analysis*

Supplementary Figures 4, 5 and 6 show CNV analysis of the regions 21q22, 17q24 and 20p13 for neuroticism, agreeableness and conscientiousness, respectively, for one carrier of H21<sub>N1</sub>, ten carriers of H17<sub>A</sub> and four carriers of H20<sub>C2</sub>. The CNV partition tool did not detect any CNV in those regions. Visually, the B-allele frequencies revealed a loss of heterozygosity (LoH) for homozygous carriers of the haplotypes in the regions, but the probe intensities did not point towards a deletion. Combining the two analyses suggests that we are dealing with “classical” genomic variation for which individuals are homozygous by descent.

## **DISCUSSION**

In our study of extended families, we found significant evidence of linkage for conscientiousness to 20p13 (rs1434789, LOD = 5.86) and suggestive evidence of linkage for neuroticism to 19q, 21q and 22q, extraversion to 1p, 1q, 9p and 12q, openness to 12q and 19q, and agreeableness to 2p, 6q, 17q and 21q. Haplotype construction and analysis in high LOD score families associated two haplotypes (H21<sub>N1</sub> & H21<sub>N2</sub>) on 21q22 with neuroticism ( $p$ -values = 0.009, 0.007), one on 17q24 with agreeableness (H17<sub>A</sub>,  $p$ -value = 0.018) and one on 20p13 with conscientiousness (H20<sub>C2</sub>,  $p = 0.058$ ). We did not detect evidence of CNVs in these regions. However, these analyses should be interpreted with caution given the low density of the array used.

In this study we aimed to find regions that are overrepresented in individuals with extreme phenotypes. Although it has been argued that this may lead to a reduction of

power as compared to a quantitative trait outcome in which all individuals are used (58), our simulation studies based on variants with large effects suggest that this is a powerful approach (47). When comparing the results discussed here to those derived from an analysis of the traits as quantitative outcomes (**Supplementary Figure 7**), the overlap is limited. This is most likely explained by the fact that the extreme analysis will pick up genes with relatively large effects. The 20p13 and 17q24 regions linked to conscientiousness and agreeableness, however, also provided evidence of linkage (albeit weak) in the quantitative trait analyses.

There are a number of issues to be considered before interpreting the results. We used the short version (NEO-FFI) of NEO personality inventory (NEO-PI-R) to assess the five basic personality traits. The NEO-FFI, although less reliable compared to the NEO-PI-R, is known to tap at least the Five basic personality traits with fidelity and reliability (59). Our study is the first to report linkage analyses of all 5 dimensions of the NEO personality inventory. With haplotype construction and association analyses, we attempted to reduce the size of the regions and find segregating haplotypes that were associated with the traits. It could be argued that the use of SNPs from the denser array for haplotype construction and association might have provided a better insight into the linked regions. Unfortunately, the denser array was available for only a subset of the population. Its use would have led to a loss of power, especially with FBAT, which only uses information from heterozygous parents in complete trios. The family based association analysis is rather robust to low SNP density as the relatives will share large haplotypes. Nevertheless, the construction of haplotypes helped us reduce the size of the linked regions, thus avoiding multiple testing issues in the subsequent association analyses to the extent that we only had to adjust for the number of haplotypes tested.

The family-based design also gave us an opportunity to examine the CNVs that might be segregating in the population. Utilizing a denser array than the 318K chip that we used, however, might have revealed smaller CNVs, in particular deletions. We observed some overlap of the linked regions between different personality traits. The 19q13 locus, for instance, was shared by neuroticism and openness. This is most likely due to the fact that strong correlations were observed between some personality traits. These correlations, although theoretically unexpected for orthogonal scales, was also observed in previous studies (35). It is interesting to note that most of our suggestive linkage peaks come from the NPL analysis but that these are usually followed by somewhat lower scores from dominant and recessive models. This is in accordance with Greenberg et al (60), which suggests that parametric LOD scores are more powerful compared to NPL only when the correct model (which is usually unknown for complex traits) is specified. Misspecification of parameters in parametric linkage analysis has serious negative impact on its power to detect linkage(61).

When comparing the suggestive regions for neuroticism in our study with those of others, there was no overlap. This might be due to the fact that earlier studies used different instruments to assess neuroticism (or similar traits), such as the EPQ or the TPQ. Nevertheless, there were regions with moderate linkage signals that overlapped with previous studies. These include a locus on 5p, which gave a LOD score of 2.76 under a dominant model in our study for neuroticism, and overlaps with Gillespie et al. (41). Similarly, two loci on chromosome 11, 11p14 (rs1564745, LOD = 1.75) and 11q23 (rs1013582, LOD = 1.95), and one locus on chromosome 10 (rs6580, LOD = 1.26) overlap with previous findings on neuroticism (39-40, 42, 44). For extraversion, the 12q24 locus (LOD = 4.01) overlaps with one identified by Gillespie et al. (p-value =

0.009) (41). This is, perhaps, the most interesting overlap, since with a LOD score of 4.01 this region approximately fulfilled our strict threshold for significant linkage. From regions harboring previously studied candidate genes, modest linkage signals (LOD = 1.16) were observed for the TPH1 region for neuroticism and the ADH4 region for agreeableness (LOD = 1.33) and extraversion (LOD = 0.56). The overlap of our findings with those of others shows that our founder population in the southwestern part of the Netherlands is representative of the general population. This is in line with the findings from our simulation studies that show that variants with frequencies  $> 0.001$  are expected to be conserved in this population (62).

The most interesting region in our study was 20p13 for conscientiousness, which gave significant linkage signals under all three models and contained a haplotype spanning 5 genes belonging to the beta defensin family. Defensins play a role in the immunologic response to invading microorganisms. Other genes in the region include *ZCCHC3*, *SOX12*, and *NRSN2*. This region was previously implicated in spinocerebellar ataxia (63) and neuro-degeneration (64). Despite the high LOD scores observed for this region, the haplotype (H20<sub>C2</sub>) was only borderline associated. An explanation for this may be the reduced number of families available under the recessive model (8 families versus 39 families for the additive model), resulting in a loss of power. This explanation was supported by the fact that a significant p-value (0.038) was observed when the analysis was run using the highest 30% as affected instead of the top10%. No significance was observed under additive or dominant models, which suggests a recessive mode of inheritance for this haplotype and its association with conscientiousness.

The second most interesting region was 21q22 for neuroticism (LOD = 3.73). Although the LOD score is not significant according to our stringent criteria, we found



two haplotypes (H21<sub>N1</sub> & H21<sub>N2</sub>) shared by individuals who had extreme phenotypes spanning the Down syndrome critical region. These haplotypes were significantly associated in the binary trait analysis. This region encompasses 10 genes, including *TTC3*, *DSCR9*, *DSCR3*, *DYRK1A*, *KCNJ6* and *PIGP*, and has been associated with cognitive and mental impairment in individuals with Down's syndrome (65-66). *DYRK1A* has been associated with conscientiousness, which is strongly inversely correlated with neuroticism (35). This is an interesting finding as many of the common physical and cognitive features of Down's syndrome may be present in apparently normal individuals. The haplotypes identified are very frequent (0.26, 0.47) making it very likely that these will also be found in populations outside of the isolate according to the gene dropping simulations and empirical findings(62).

The locus on 17q24 for agreeableness (LOD = 3.32), harbored a shared haplotype, H17<sub>A</sub>, that was observed in individuals with extreme phenotypes in high LOD score families. H17<sub>A</sub> covered an 80 kb region on 17q24 and includes *KCNJ2*, *KCNJ16* and *MAP2K6*. *KCNJ2* has been associated with Andersen syndrome, which is characterized by periodic paralysis, cardiac arrhythmias and dysmorphic features (67) and *MAP2K6* has been associated with Alzheimer's disease (68). H17<sub>A</sub> was associated with the quantitative trait but not the binary trait that provided evidence of linkage. This may be explained by the fact that, unlike other personality traits, for agreeableness there was no single family making a large contribution (> 1) to the LOD scores. Instead, a number of families contributed to the LOD, which points towards the presence of a common variant rather than a rare one.

Of interest is the fact that two of these three regions harbor potassium/magnesium channels that are conserved in most mammalian cells (*KCNJ2*, *KCNJ6*, *KCNJ16*) while

in the other, one of the genes (*SOX12*) is known to interact with a potassium channel (*KCNJ10*). Other interesting regions include 19q13 and 22q11 for neuroticism.

Chromosome 19q13 is a very important region that contains the apolipoprotein E gene (APOE). Both neuroticism and APOE are known to be associated with AD (1, 5, 69). The 22q11 region is of interest since it was previously associated with schizophrenia (70). Since neuroticism is a known risk factor of schizophrenia (9, 16), it is possible that this signal might actually be driven by neurotic aspect of schizophrenia.

In this study, we attempted to identify genes with relatively large effects on personality traits. However, the estimated ORs did not suggest variants with large effects. This may be because the variants tested for association are not those conferring disease susceptibility. Another reason may be the study design, in which we compared those scoring in the top 10% (cases) with those in the bottom 90% (controls) to estimate the odds ratios. Simulations have shown that a comparison of individuals from discontinuous bottom and top percentiles is more powerful compared to dichotomizing the whole sample at a cut-point (71). Nevertheless, these effects are large compared to the low ORs seen in GWAS ( $1.01 < OR < 1.15$ ). Inflation in the ORs is expected if association analysis is performed using the same sample in the region with the highest LOD scores (72). We do not expect a large bias since we used the information from the whole population for haplotype association analysis as opposed to the linkage analysis where we performed affected-only analysis with only top 10% of the phenotypic distribution as affected.

Although GWAS of sufficiently large samples may still yield novel genes with small effects, there is growing interest in whole genome sequencing as a tool to identify loci with large effects. Our study suggests that such loci can be identified for personality traits. Our linkage analysis revealed three interesting regions (20p13 for

conscientiousness, 21q22 for neuroticism and 17q24-25 for agreeableness), which may include genes that exert moderate to large effects on NEO personality traits. The overlap with previous linkage studies, and with associated personality disorders, suggests that the findings in our population can be extrapolated to other populations. We confirm linkage of conscientiousness with 20p13, of neuroticism with 21q22 and of agreeableness with 17q24. The other identified regions need to be followed up with deep sequencing, which may help in capturing the actual mutation responsible for the linkage signals.

Supplementary information is available at *Molecular Psychiatry's* website

### **Acknowledgements:**

The ERF study was further supported by grants from the Netherlands Organisation for Scientific Research (NWO), Erasmus university medical center, the Centre for Medical Systems Biology (CMSB1 and CMSB2) of the Netherlands Genomics Initiative (NGI) and Leiden University Medical Centre (LUMC) for genotyping. We are grateful to all patients and their relatives, general practitioners and neurologists for their contributions and to Petra Veraart for her help in genealogy, Jeannette Vergeer for the supervision of the laboratory work and Peter Snijders for his help in data collection. ESG was supported by The International Human Frontier Science Program Organization.

**Conflict of interest:** The authors declare no conflict of interest.

### **References**

1. Crowe M, Andel R, Pedersen NL, Fratiglioni L, Gatz M. Personality and risk of cognitive impairment 25 years later. *Psychol Aging* 2006; **21**: 573-80.
2. Goodwin RD, Stein MB. Peptic ulcer disease and neuroticism in the United States adult population. *Psychother Psychosom* 2003; **72**: 10-5.
3. Hettema JM, Neale MC, Myers JM, Prescott CA, Kendler KS. A population-based twin study of the relationship between neuroticism and internalizing disorders. *Am J Psychiatry* 2006; **163**: 857-64.

4. Kendler KS, Gatz M, Gardner CO, Pedersen NL. Personality and major depression: a Swedish longitudinal, population-based twin study. *Arch Gen Psychiatry* 2006; **63**: 1113-20.
5. Meins W, Dammast J. Do personality traits predict the occurrence of Alzheimer's disease? *Int J Geriatr Psychiatry* 2000; **15**: 120-4.
6. Robertson DA, Ray J, Diamond I, Edwards JG. Personality profile and affective state of patients with inflammatory bowel disease. *Gut* 1989; **30**: 623-6.
7. Shipley BA, Weiss A, Der G, Taylor MD, Deary IJ. Neuroticism, extraversion, and mortality in the UK Health and Lifestyle Survey: a 21-year prospective cohort study. *Psychosom Med* 2007; **69**: 923-31.
8. Terracciano A, Lockenhoff CE, Crum RM, Bienvenu OJ, Costa PT, Jr. Five-Factor Model personality profiles of drug users. *BMC Psychiatry* 2008; **8**: 22.
9. Van Os J, Jones PB. Neuroticism as a risk factor for schizophrenia. *Psychol Med* 2001; **31**: 1129-34.
10. Jang KL, Livesley WJ, Vernon PA. Heritability of the big five personality dimensions and their facets: a twin study. *J Pers* 1996; **64**: 577-91.
11. McCrae RR. NEO-PI-R data from 36 cultures: Further intercultural comparisons. In: McCrae RR, J A, editors. *The five-factor model of personality across cultures*. New York: Kluwer Academic/Plenum; 2002. p. 105-25.
12. Costa PT, McCrae RR. Revised NEO Personality Inventory (NEO-PI-R) and NEO Five Factor Inventory (NEO-FFI) Professional Manual. *Psychological Assessment Resources* 1992.

13. Biesanz JC, West SG. Towards understanding assessments of the big five: multitrait-multimethod analyses of convergent and discriminant validity across measurement occasion and type of observer. *J Pers* 2004; **72**: 845-76.
14. Winstow DJ, Wakefield JA, Goldsmith WM. The relationship between personality, health symptoms and disease. *Person Individ Diff* 1990; **11**: 717-23.
15. Zbogar, Hema. Neuroticism as a potential risk factor for Alzheimer's disease. *The Journal of Addiction and Mental Health* 2004.
16. Lonqvist JE, Verkasalo M, Haukka J, Nyman K, Tiihonen J, Laaksonen I *et al.* Premorbid personality factors in schizophrenia and bipolar disorder: results from a large cohort study of male conscripts. *J Abnorm Psychol* 2009; **118**: 418-23.
17. Wilson RS, Schneider JA, Arnold SE, Bienias JL, Bennett DA. Conscientiousness and the incidence of Alzheimer disease and mild cognitive impairment. *Arch Gen Psychiatry* 2007; **64**: 1204-12.
18. Cloninger CR, Svrakic DM, Przybeck TR. A psychobiological model of temperament and character. *Arch Gen Psychiatry* 1993; **50**: 975-90.
19. Kaasinen V, Nurmi E, Bergman J, Eskola O, Solin O, Sonninen P *et al.* Personality traits and brain dopaminergic function in Parkinson's disease. *Proc Natl Acad Sci U S A* 2001; **98**: 13272-7.
20. Menza MA, Golbe LI, Cody RA, Forman NE. Dopamine-related personality traits in Parkinson's disease. *Neurology* 1993; **43**: 505-8.
21. Menza M. The personality associated with Parkinson's disease. *Curr Psychiatry Rep* 2000; **2**: 421-6.

22. Canuto A, Giannakopoulos P, Meiler-Mittelau C, Delaloye C, Herrmann FR, Weber K. Personality traits influence clinical outcome in day hospital-treated elderly depressed patients. *Am J Geriatr Psychiatry* 2009; **17**: 335-43.
23. Poropat AE. A meta-analysis of the five-factor model of personality and academic performance. *Psychol Bull* 2009; **135**: 322-38.
24. Waller N. Evaluating the structure of personality. In: Cloninger C, editor. *Personality and psychopathology*. Washington, D.C.: American Psychiatric Association; 1999. p. 155-97.
25. Riemann R, Angleitner A, Strelau J. Genetic and environmental influences on personality: A study of twins reared together using the self- and peer report NEO-FFI scales. *J Personal* 1997; **65**: 449-75.
26. Loehlin JC, McCrae RR, Paul T, Costa J, John OP. Heritabilities of common and measure-specific components of the big five personality factors. *Journal of research in personality* 1998; **32**: 431-53.
27. Jang KL, McCrae RR, Angleitner A, Riemann R, Livesley WJ. Heritability of facet-level traits in a cross-cultural twin sample: support for a hierarchical model of personality. *J Pers Soc Psychol* 1998; **74**: 1556-65.
28. Lesch KP, Bengel D, Heils A, Sabol SZ, Greenberg BD, Petri S *et al*. Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* 1996; **274**: 1527-31.
29. Mazzanti CM, Lappalainen J, Long JC, Bengel D, Naukkarinen H, Eggert M *et al*. Role of the serotonin transporter promoter polymorphism in anxiety-related traits. *Arch Gen Psychiatry* 1998; **55**: 936-40.

30. Sen S, Burmeister M, Ghosh D. Meta-analysis of the association between a serotonin transporter promoter polymorphism (5-HTTLPR) and anxiety-related personality traits. *Am J Med Genet* 2004; **127B**: 85-9.
31. Nash MW, Sugden K, Huezo-Diaz P, Williamson R, Sterne A, Purcell S *et al.* Association analysis of monoamine genes with measures of depression and anxiety in a selected community sample of siblings. *Am J Med Genet B Neuropsychiatr Genet* 2005; **135B**: 33-7.
32. Luo X, Kranzler HR, Zuo L, Zhang H, Wang S, Gelernter J. CHRM2 variation predisposes to personality traits of agreeableness and conscientiousness. *Hum Mol Genet* 2007; **16**: 1557-68.
33. Calboli FC, Tozzi F, Galwey NW, Antoniadou A, Mooser V, Preisig M *et al.* A genome-wide association study of neuroticism in a population-based sample. *PLoS One* 2010; **5**: e11504.
34. Hettema JM, van den Oord EJ, An SS, Kendler KS, Chen X. Follow-up association study of novel neuroticism gene MAMDC1. *Psychiatr Genet* 2009; **19**: 213-4.
35. Terracciano A, Sanna S, Uda M, Deiana B, Usala G, Busonero F *et al.* Genome-wide association scan for five major dimensions of personality. *Mol Psychiatry* 2010; **15**: 647-56.
36. van den Oord EJ, Kuo PH, Hartmann AM, Webb BT, Moller HJ, Hettema JM *et al.* Genomewide association analysis followed by a replication study implicates a novel candidate gene for neuroticism. *Arch Gen Psychiatry* 2008; **65**: 1062-71.
37. Eysenck SBG, Eysenck HJ, Barrett P. A revised version of the psychoticism scale. *Person Individ Diff* 1985; **6**: 21-9.



38. Cloninger CR, Przybeck TR, Svrakic DM. The Tridimensional Personality Questionnaire: U.S. normative data. *Psychol Rep* 1991; **69**: 1047-57.
39. Cloninger CR, Van Eerdewegh P, Goate A, Edenberg HJ, Blangero J, Hesselbrock V *et al.* Anxiety proneness linked to epistatic loci in genome scan of human personality traits. *Am J Med Genet* 1998; **81**: 313-7.
40. Fullerton J, Cubin M, Tiwari H, Wang C, Bomhra A, Davidson S *et al.* Linkage analysis of extremely discordant and concordant sibling pairs identifies quantitative-trait loci that influence variation in the human personality trait neuroticism. *Am J Hum Genet* 2003; **72**: 879-90.
41. Gillespie NA, Zhu G, Evans DM, Medland SE, Wright MJ, Martin NG. A genome-wide scan for Eysenckian personality dimensions in adolescent twin sibships: psychoticism, extraversion, neuroticism, and lie. *J Pers* 2008; **76**: 1415-46.
42. Kuo PH, Neale MC, Riley BP, Patterson DG, Walsh D, Prescott CA *et al.* A genome-wide linkage analysis for the personality trait neuroticism in the Irish affected sib-pair study of alcohol dependence. *Am J Med Genet B Neuropsychiatr Genet* 2007; **144B**: 463-8.
43. Nash MW, Huezio-Diaz P, Williamson RJ, Sterne A, Purcell S, Hoda F *et al.* Genome-wide linkage analysis of a composite index of neuroticism and mood-related scales in extreme selected sibships. *Hum Mol Genet* 2004; **13**: 2173-82.
44. Neale BM, Sullivan PF, Kendler KS. A genome scan of neuroticism in nicotine dependent smokers. *Am J Med Genet B Neuropsychiatr Genet* 2005; **132**: 65-9.
45. Wray NR, Middeldorp CM, Birley AJ, Gordon SD, Sullivan PF, Visscher PM *et al.* Genome-wide linkage analysis of multiple measures of neuroticism of 2 large cohorts from Australia and the Netherlands. *Arch Gen Psychiatry* 2008; **65**: 649-58.

46. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988; **16**: 1215.
47. Van Gestel S, Houwing-Duistermaat JJ, Adolfsson R, van Duijn CM, Van Broeckhoven C. Power of selective genotyping in genetic association analyses of quantitative traits. *Behav Genet* 2000; **30**: 141-6.
48. Liu F, Kirichenko A, Axenovich TI, Duijn CMv, Aulchenko YS. An approach for cutting large and complex pedigrees for linkage analysis. *European Journal of Human Genetics* 2008; **in press**.
49. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002; **30**: 97-101.
50. Kong A, Cox NJ. Allele-sharing models: LOD scores and accurate linkage tests. *Am J Hum Genet* 1997; **61**: 1179-88.
51. McPeck MS. Optimal allele-sharing statistics for genetic mapping using affected relatives. *Genet Epidemiol* 1999; **16**: 225-49.
52. Sobel E, Lange K. Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. *Am J Hum Genet* 1996; **58**: 1323-37.
53. Sobel E, Papp JC, Lange K. Detection and integration of genotyping errors in statistical genetics. *Am J Hum Genet* 2002; **70**: 496-508.
54. Horvath S, Xu X, Laird NM. The family based association test method: strategies for studying general genotype--phenotype associations. *Eur J Hum Genet* 2001; **9**: 301-6.
55. Laird NM, Horvath S, Xu X. Implementing a unified approach to family-based tests of association. *Genet Epidemiol* 2000; **19 Suppl 1**: S36-42.

56. Conlin LK, Thiel BD, Bonnemann CG, Medne L, Ernst LM, Zackai EH *et al.* Mechanisms of mosaicism, chimerism and uniparental disomy identified by single nucleotide polymorphism array analysis. *Hum Mol Genet* 2010; **19**: 1263-75.
57. Yau C, Holmes CC. CNV discovery using SNP genotyping arrays. *Cytogenet Genome Res* 2008; **123**: 307-12.
58. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet* 1998; **62**: 1198-211.
59. Rosellini AJ, Brown TA. The NEO Five-Factor Inventory: latent structure and relationships with dimensions of anxiety and depressive disorders in a large clinical sample. *Assessment* 2011; **18**: 27-38.
60. Durner M, Vieland VJ, Greenberg DA. Further evidence for the increased power of LOD scores compared with nonparametric methods. *Am J Hum Genet* 1999; **64**: 281-9.
61. Clerget-Darpoux F, Bonaiti-Pellie C, Hochez J. Effects of misspecifying genetic parameters in lod score analysis. *Biometrics* 1986; **42**: 393-9.
62. Pardo LM, MacKay I, Oostra B, van Duijn CM, Aulchenko YS. The effect of genetic drift in a young genetically isolated population. *Ann Hum Genet* 2005; **69**: 288-95.
63. Verbeek DS, van de Warrenburg BP, Wesseling P, Pearson PL, Kremer HP, Sinke RJ. Mapping of the SCA23 locus involved in autosomal dominant cerebellar ataxia to chromosome region 20p13-12.3. *Brain* 2004; **127**: 2551-7.
64. Zhou B, Westaway SK, Levinson B, Johnson MA, Gitschier J, Hayflick SJ. A novel pantothenate kinase gene (PANK2) is defective in Hallervorden-Spatz syndrome. *Nat Genet* 2001; **28**: 345-9.

65. Altafaj X, Dierssen M, Baamonde C, Marti E, Visa J, Guimera J *et al.* Neurodevelopmental delay, motor abnormalities and cognitive deficits in transgenic mice overexpressing Dyrk1A (minibrain), a murine model of Down's syndrome. *Hum Mol Genet* 2001; **10**: 1915-23.
66. Capone GT. Down syndrome: advances in molecular biology and the neurosciences. *J Dev Behav Pediatr* 2001; **22**: 40-59.
67. Andelfinger G, Tapper AR, Welch RC, Vanoye CG, George AL, Jr., Benson DW. KCNJ2 mutation results in Andersen syndrome with sex-specific cardiac and skeletal muscle phenotypes. *Am J Hum Genet* 2002; **71**: 663-8.
68. Zhu X, Rottkamp CA, Hartzler A, Sun Z, Takeda A, Boux H *et al.* Activation of MKK6, an upstream activator of p38, in Alzheimer's disease. *Journal of neurochemistry* 2001; **79**: 311-8.
69. van der Flier WM, Schoonenboom SN, Pijnenburg YA, Fox NC, Scheltens P. The effect of APOE genotype on clinical phenotype in Alzheimer disease. *Neurology* 2006; **67**: 526-7.
70. Karayiorgou M, Morris MA, Morrow B, Shprintzen RJ, Goldberg R, Borrow J *et al.* Schizophrenia susceptibility associated with interstitial deletions of chromosome 22q11. *Proc Natl Acad Sci U S A* 1995; **92**: 7612-6.
71. Deng HW, Chen WM, Recker RR. QTL fine mapping by measuring and testing for Hardy-Weinberg and linkage disequilibrium at a series of linked marker loci in extreme samples of populations. *Am J Hum Genet* 2000; **66**: 1027-45.
72. Goring HH, Terwilliger JD, Blangero J. Large upward bias in estimation of locus-specific effects from genomewide scans. *Am J Hum Genet* 2001; **69**: 1357-69.



## Figure legends

### Figure 1

Genome wide linkage results from dominant (green), recessive (red) models, and non-parametric linkage (NPL) analysis (blue). The X-axis depicts the whole autosomal genome divided into 22 chromosomes. The Y-axis shows the heterogeneity LOD scores (HLOD) for dominant and recessive models and LOD scores from NPL analysis. (A) Neuroticism, (B) Extraversion, (C) Openness, (D) Agreeableness and (E) Conscientiousness.

Table 1: A summary of published genome wide linkage studies for neuroticism

Study	Questionnaire/trait	Chrom	Region (cM)	Markers	Lod score	P-value
Gillespie et al. 2008	JEPQ/Neuroticism	5	15	-	1.52 <sup>a</sup>	0.008
		10	105-125	-	1.79 <sup>a</sup>	0.004
		12	110	-	1.58 <sup>a</sup>	0.007
		15	100	-	1.79 <sup>a</sup>	0.004
		16	90-95	-	2.35 <sup>a</sup>	0.001
		19	30	-	1.91 <sup>a</sup>	0.003
Wray et al. 2008	EPQ/NEO/Neuroticism	2	112	D2S1790	1.6	-
		5	191	D5S211	2.2	-
		6	75	D6S2410	1.5	-
		8	45	D8S1771	1.6	-
		8	134	D8S592	1.6	-
		10	5	D10S1412	2.0	-
		10	175	D10S212	1.7	-
		14	103	D14S1434	2.6	-
		15	17	GTTTT001	1.8	-
		18	117	D18S61	1.9	-
Kuo et al. 2007	EPQ/Neuroticism	1	49	D1S470	1.77	-
		1	83	D1S476	1.38	-
		2	157	D2S349	1.07	-
		9	162	D9S1826	1.02	-
		11	43	D11S4080	2.06	-
		12	175	D12S1638	2.13	-
		15	124	D15S1014	1.03	-
		18	91	D18S68	1.39	-
Neale et al. 2005	EPQ/Neuroticism	18	115	D18S1009	1.10	-
		1	137	AMY	1.91 <sup>a</sup>	0.003
		3	104.5	D3S2406	1.39 <sup>a</sup>	0.0113
		6	147.3	D6S1003	1.39 <sup>a</sup>	0.0112
		11	132	D11S4150	1.41 <sup>a</sup>	0.0108
Nash et al. 2004	EPQ/Neuroticism	12	45.5	D12S1042	2.22 <sup>a</sup>	0.0014
		1	80	D1S2890	1.6	0.62
Fullerton et al. 2003	EPQ/Neuroticism	6	47	D6S1610	2.7	0.07
		1	126	D1S2868	3.25 <sup>a</sup>	0.00011
Cloninger et al. 1998	TPQ/Harm-avoidance	4	176	D4S1539	3.15 <sup>a</sup>	0.00014
		7	42	D7S516	3.18 <sup>a</sup>	0.00013
		8	8	D8S277	2.29 <sup>a</sup>	0.00117
		11	99	D11S898	3.00 <sup>a</sup>	0.00020
		12	105	D12S346	3.95 <sup>a</sup>	0.00002
		13	64	D13S153	3.12 <sup>a</sup>	0.00015
		8*	17	D8S1106	3.2	0.00006
11	194	D11S1327	1.6	0.003		
18	109		1.6	0.004		

EPQ = Eysenck personality questionnaire, JEPQ = Junior Eysenck personality questionnaire, TPQ = Tri-dimensional personality questionnaire

\* region confirmed by Zohar et al. 2003 in an independent candidate region study

<sup>a</sup> LOD scores not reported in the actual studies but calculated for comparison in this study using the conversion  $\chi^2 / 2 * \ln(10)$

Table 2: General characteristics of the individuals selected for the linkage analysis

	Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness
<b>Number</b>	221	176	215	227	178
<b>Age</b>	47.5(14.2)	41.7(14.8)	42.1(13.5)	49.9(14.0)	50.0(14.0)
<b>% women</b>	70	54	61	78	56
<b>Range in scale</b>	12-59	16-60	13-52	20-60	23-60
<b>Threshold score at 90<sup>th</sup> Percentile</b>	42	48	40	51	54
<b>Scores</b>					
<b>Neuroticism</b>	45.9(3.8)	25.5(7.4)**	30.1(8.4)*	27.3(8.9)**	26.8(7.2)**
<b>Extraversion</b>	33.5(7.1)**	51.4(2.3)	42.3(5.9)**	42.8(6.9)**	45.1(6.5)**
<b>Openness</b>	32.4(6.1)*	34.5(6.8)**	43.8(2.7)	31.8(6.4)*	32.1(6.5)*
<b>Agreeableness</b>	42.4(6.4)**	46.5(5.9)**	43.8(6.3)	53.6(2.5)	47.1(6.6)**
<b>Conscientiousness</b>	42.3(6.9)**	51.6(5.4)**	46.1(7.0)	50.5(5.5)**	56.8(1.7)

\* Difference between means of highest 10% and lowest 90% at  $p < 0.05$ , \*\* at  $p < 0.001$   
 Values are means (standard deviation), unless otherwise indicated



Table 3: Regions in suggestive or significant linkage with NEO-FFI traits

start SNP	end SNP	Cytogenetic location	Position (start)	Position (end)	Dominant (HLOD) <sup>†</sup>	Recessive (HLOD) <sup>†</sup>	Non-parametric (LOD) <sup>∞</sup>	Genome-wide p-value*
<b>Neuroticism</b>								
rs964795	rs3499	19q13.4	63029177	63785296	2.50	<i>2.94</i>	<i>3.73</i>	0.11
rs1012959	rs1016694	21q22	36983492	38156688	0.15	0.51	<i>3.42</i>	0.21
rs713816	rs4444	22q11	24491639	29529888	2.60	<i>3.07</i>	2.50	0.37
<b>Extraversion</b>								
rs437749	rs1413527	1p31	82965819	88142058	0.14	0.43	<i>3.05</i>	0.38
rs913257	rs1004959	1q24-25	167253034	175983330	2.88	1.43	1.31	0.49
rs1532310	rs1412256	9p24	592986	1454067	0.43	0.10	<i>2.96</i>	0.44
rs1278602	rs7960480	12q24.3	132109288	132388516	1.94	2.06	<i>4.01</i>	0.06
<b>Openness</b>								
rs1278602	rs7960480	12q24.3	132109288	132388516	1.87	2.21	<i>3.71</i>	0.12
rs964795	rs3499	19q13.4	63029177	63785296	2.28	<i>3.00</i>	2.78	0.41
<b>Agreeableness</b>								
rs168293	rs6432244	2p25	8674900	12079927	0.51	1.43	<i>3.11</i>	0.36
rs727619	rs8770	6q27	170623826	170804163	1.69	0.86	<i>3.54</i>	0.16
rs411602	rs759563	17q24-25	66051172	67882483	0.92	0.63	<i>3.32</i>	0.25
rs2834380	rs2836301	21q22	34413370	38599459	1.99	0.80	<i>3.64</i>	0.14
<b>Conscientiousness</b>								
rs1434789	rs434609	20p13	85900	184992	<b>5.25</b>	<b>5.86</b>	<b>4.47</b>	0.03

Significant regions in bold

Suggestive regions in italic

† heterogeneity log of odds

∞ Kong & Cox log of odds

\* Empirical genome-wide p-value estimated as a proportion of simulated scans resulting in NPL (LOD)  $\geq$  the observed (500 simulations)

Table 4. The association of the haplotypes identified in families showing evidence of linkage to the outcome of interest

Trait	chr	Haplotype Name	Haplotype Configuration	Flanking SNPs	Position	$f^1$	p-value		OR (CI) Effect (SE) <sup>2</sup>	Genes
							Binary	quantitative		
<b>Neuroticism</b>										
	21	H21 <sub>N1</sub>	1111111111	rs2835574- rs1016694	37377194- 38156688	0.473	0.009*	0.888	1.2(0.97,1.46)	TTC3, DSCR9,DSCR3, DYRK1A, KCNJ6
	21	H21 <sub>N2</sub>	2222222222	rs2835574- rs1016694	37377194- 38156688	0.261	0.007*	0.520	0.76(0.57,0.99)	TTC3, DSCR9,DSCR3, DYRK1A, KCNJ6
<b>Extraversion</b>										
	12	H12 <sub>E</sub>	1111112	rs1278602- rs7960480	132109288- 132388516	0.361	0.696	0.780		
<b>Openness</b>										
	12	H12 <sub>O</sub>	1121111	rs1278602- rs7960480	132109288- 132388516	0.052	0.309	0.638		
<b>Agreeableness</b>										
	6	H6 <sub>A</sub>	12222	rs909472- rs8770	170731815- 170804163	0.598	0.843	0.616		
	17	H17 <sub>A</sub>	111	rs411602- rs1981647	66051172- 66132788	0.168	0.395	0.018*	0.31(0.23)	KCNJ2, KCNJ16, MAP2K6,
	21	H21 <sub>A1</sub>	22222222	rs2835574- rs2835653	37377194- 37480506	0.521	0.366	0.943		
	21	H21 <sub>A2</sub>	11111111	rs2835574- rs2835653	37377194- 37480506	0.255	0.349	0.689		
<b>Conscientiousness</b>										
	20	H20 <sub>C1</sub>	2221222	rs1434789- rs434609	85900- 184992	0.231	0.058	0.152	1.4 (0.96,1.99)	DEFB, ZCCHC3, SOX12, NRSN2
	20	H20 <sub>C2</sub>	1112111	rs1434789- rs434609	85900- 184992	0.339	0.408	0.094		

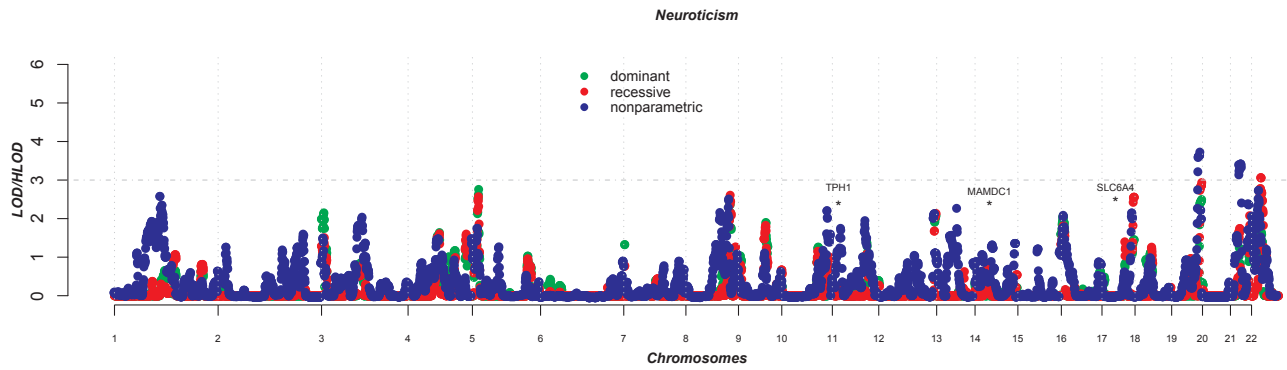
<sup>1</sup> frequency of the haplotype in the total population

<sup>2</sup> Odds ratio calculated using simple logistic regression in SPSS if association was observed under binary model, effect size from a linear regression analysis is given when association was observed under quantitative model

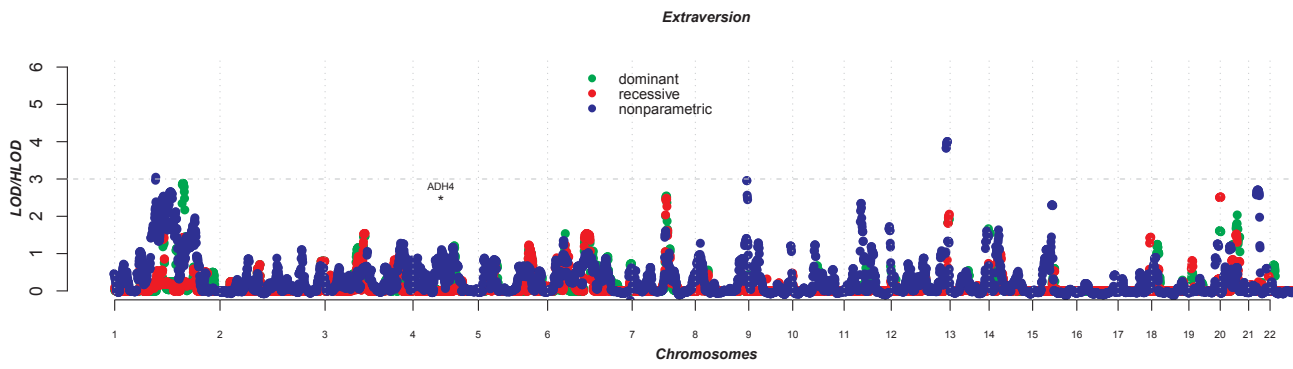
\* significant after Bonferroni adjustment

Figure 1

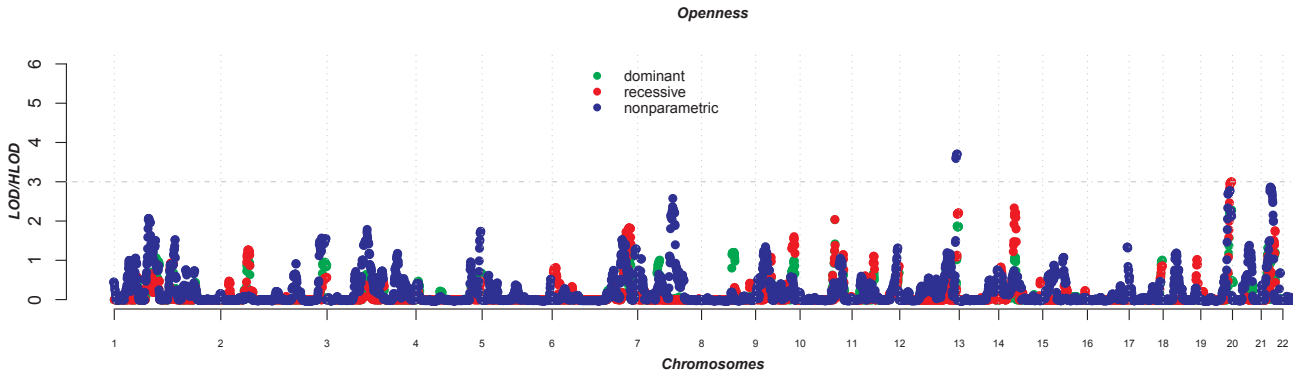
A



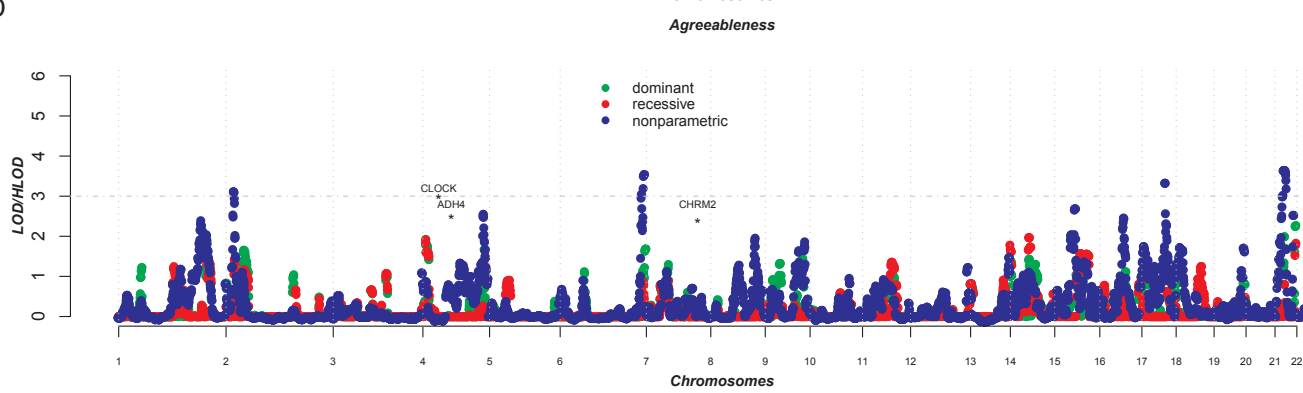
B



C



D



E

