

## **Supplementary methods**

### **Subjects entered into different genetic analyses**

For longitudinal heritability analysis we included a maximum of two twins and four siblings (two brothers and two sisters) per family. In total, for the PI-R-ABBR collected in 2002, 5.780 subjects were available, including 2223 monozygotic (MZ) twins (635M/1588F), 1207 dizygotic same sex (DZ-SS) twins (360M/847F), 955 dizygotic opposite sex (DZ-OS) twins (403M/552F) and 1395 siblings (552M/843F). For the PI-R-ABBR collected in 2008, a total of 9.147 subjects were available, including 3731 MZ twins (1017M/2714F), 2035 dizygotic same sex (DZ-SS) twins (623M/1412F), 1683 dizygotic opposite sex (DZ-OS) twins (646M/1037F) and 1698 siblings (629M/1069F).

### **Participants and measures**

The PI-R-ABBR is derived from the Padua Inventory-Revised (PI-R), a widely used self-report inventory on OCS (1;2). The short version of the self-report instrument includes the subscales washing, checking, rumination, precision and impulses and has 12 items, which were chosen by selecting the 2 items per subscale that have had the highest factor loadings in a validation study (3) and an additional item for subscales on rumination and impulses (4;5). Each item is scored on a scale ranging from 0-4, with the OC symptom score per individual calculated as the total sum score (ranging from 0-48).

### **Genotyping and imputation**

After DNA extraction and purification, genotyping was performed on multiple array platforms (AFFYMETRIX PERLEGEN 5.0, ILLUMINA 370, ILLUMINA 660, ILLUMINA OMNI EXPRESS 1M and AFFYMETRIX 6.0). Platform specific software (GENOTYPER, BEADSTUDIO and BIRDSEED) were used for genotype calling. Quality control was performed within and between chip platforms. First, individual

SNP markers were lifted over to build 37 (HG19) of the Human reference genome, using the LIFTOVER tool (6). Second, SNPs that were not mapped at all, SNPs that had ambiguous locations, and SNPs that did not have matching or strand opposite alleles, were removed. Subsequently, the data were strand aligned with the 1000 Genomes phase 1 INTEGRATED RELEASE version 3 ALL panel of March 2012 (7). If a SNP still had mismatching alleles with the imputation reference set, if the allele frequencies differed more than 0.20 with the reference set, if the MAF was <1%, if the HWE P-value was <0.00001 or if the call rate was <95%, they were also removed. In addition, samples were excluded from the data if their expected sex did not match their genotyped sex, if the genotype missing rate was above 10% or if the Plink F inbreeding value was either >0.10 or <-0.10. After quality control, data of the individual chips were merged into a single dataset using the PLINK 1.07 software (8). Within the merged set, identity by descent (IBD) was calculated between all possible pairs of individuals and compared to the expected family structure of the NTR studies. Samples were removed when the data did not match the expected IBD sharing, or, alternatively, when potentially consistent with biographic data, corrections were made to the family structure. For DNA samples that were genotyped on multiple platforms we tested if the overlapping SNPs had a concordance rate above 99.0%. If not, data from these samples were removed. HWE and MAF SNP filters as well as reference allele frequency difference <0.20 checks were re-applied on the merged data. As a final quality step prior to imputation SNPs with C/G and A/T allele combinations were removed if the MAF was between 0.35 and 0.50 to avoid wrong strand alignment. Imputation was performed using the two stage approach. Pre-imputation phasing and imputation of genotype platform specific SNPs was carried out using the MACH software (9). Subsequently, imputation of the reference set was carried out with Minimach. Prior to imputation the person with the highest SNP call rate within an MZ twin pair was selected. Post imputation, the resulting imputed genotypes were duplicated back to the co-twin in the data (10).

## SEM methods

First, familial correlations were estimated. We tested for sex differences in variances and correlations, and for age regression on the OC symptom scores. Next, a bivariate model (with input of a 12 x 12 data matrix; 2 twins, 2 brothers and 2 sisters, with two measures) was fitted to the data to estimate the contribution of A, D and E variance components within and across time points. Parameter estimation was done by raw-data maximum likelihood, and comparison of submodels (model AE, assuming no dominance component) was done by means of likelihood-ratio tests, whose distribution follows a  $\chi^2$  with degrees of freedom equal to the difference in the number of parameters of the two models. If the more restricted model better fits the data, the more general model is dropped. The constrained model was deemed not significantly worse than the previous model if the likelihood-ratio tests yielded a p value > 0.01. For both males and females, and for both PI-R-ABBR collected in 2002 and 2008, the distribution of the data was skewed (skewness<sub>PI-R-ABBR 2002</sub> = 1.100; skewness<sub>PI-R-ABBR 2008</sub> = 1.217). Therefore, data were square root transformed before genetic modeling.

## Estimations of variance explained by common SNPs (GCTA)

From the complete 1000 genomes imputed data transformed to best guess Plink binary format SNPs were selected with a minimum imputation  $R^2$  quality metric of 0.80. In order to avoid explained and artificially increased GRM differences through differences between platforms and between subsamples, additional SNP quality control included evaluation of the SNP platform effect. This was done by defining individuals on a specific platform as cases and the remaining individuals as controls. Allelic association was calculated and SNPs were removed if allele frequencies were significantly different between platforms (p-value < 0.00001). Sex and age were included as covariates. The GRM was calculated per chromosome using GCTA and subsequently the 22 matrixes were merged into a single GRM based on 6.325.240 SNPs that remained in analysis. For the method proposed by Zaitlen et al (11),

a second variance component is included in the model, that only considers closely related individuals for a certain predefined threshold of genetic relatedness (here defined only those with identity by descent values greater than 0.05). At the current sample size of 6881, the power to detect heritability of 0.14 is 86% (<http://spark.rstudio.com/ctgg/gctaPower/>).

## **GWAS**

Imputed SNPs were analyzed in PLINK software (8). Genotype imputation uncertainty was accounted for by using allelic dosage (<http://pngu.mgh.harvard.edu/~purcell/plink/>). To correct for familial relatedness robust standard errors provided by the sandwich estimator in Plink (12) were computed. For quality control we filtered on the minor allele frequency (MAF) to ensure a minimum of 5 samples in the minor allele homozygous group, which was calculated as the  $\sqrt{5/\text{effective sample size}} = \sqrt{5/6931} = 0.03$ . Further, we filtered on  $\text{INFO} = 0.4 - 1.1$ , to get rid of SNPs with a bad imputation quality, and calculated a sample based Hardy-Weinberg Equilibrium (HWE) and only included SNPs with a HWE expected  $p\text{-value} > 0.00001$ . At the threshold of  $\alpha = 5.00E-08$ , the required sample size for 80% power to detect a SNP at a minor allele frequency of 0.03, accounting for a proportion of total variance of 0.0025 is 16.950 unrelated individuals (<http://pngu.mgh.harvard.edu/~purcell/gpc/>). This indicates the low power of the study and the need for future Meta-analysis.

## Reference List

- (1) Sanavio E. Obsessions and compulsions: the Padua Inventory. *Behav Res Ther* 1988;26(2):169-77.
- (2) van Oppen P. Obsessions and compulsions: dimensional structure, reliability, convergent and divergent validity of the Padua Inventory. *Behav Res Ther* 1992 Nov;30(6):631-7.
- (3) van Oppen P, Hoekstra RJ, Emmelkamp PM. The structure of obsessive-compulsive symptoms. *Behav Res Ther* 1995 Jan;33(1):15-23.
- (4) Cath DC, van Grootheest DS, Willemsen G, van Oppen P, Boomsma DI. Environmental factors in obsessive-compulsive behavior: evidence from discordant and concordant monozygotic twins. *Behav Genet* 2008 Mar;38(2):108-20.
- (5) den Braber A, van 't Ent D, Cath DC, Wagner J, Boomsma DI, de Geus EJ. Brain activation during cognitive planning in twins discordant or concordant for obsessive-compulsive symptoms. *Brain* 2010 Oct;133(10):3123-40.
- (6) Kuhn RM, Karolchik D, Zweig AS, Wang T, Smith KE, Rosenbloom KR, et al. The UCSC Genome Browser Database: update 2009. *Nucleic Acids Res* 2009 Jan;37(Database issue):D755-D761.
- (7) Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012 Nov 1;491(7422):56-65.
- (8) Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007 Sep;81(3):559-75.
- (9) Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* 2010 Dec;34(8):816-34.
- (10) Nivard MG, Mbarek H, Hottenga JJ, Smit JH, Jansen R, Penninx BW, et al. Further confirmation of the association between anxiety and CTNND2: replication in humans. *Genes Brain Behav* 2014 Feb;13(2):195-201.
- (11) Zaitlen N, Kraft P, Patterson N, Pasaniuc B, Bhatia G, Pollack S, et al. Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. *PLoS Genet* 2013 May;9(5):e1003520.
- (12) Minica CC, Dolan CV, Kampert MM, Boomsma DI, Vink JM. Sandwich corrected standard errors in family-based genome-wide association studies. *Eur J Hum Genet* 2014 Jun 11.

# **Supplementary legends**

## **Figures**

**Supplementary figure 1.** Distribution of OC symptom scores, before transformation of the data, in both PI-R-ABBR collected in 2002 (left), and 2008 (right).

**Supplementary figure 2.** Proportion of variance in height, as measured in the NTR sample, explained by polygenic scores (PRS) obtained from European case-control sample by Stewart et al. 2013, with a range of 15 statistical cutoffs for SNP inclusion in the score (PRS1;  $p < 0.00001$ , PRS2;  $p < 0.0001$ , PRS3;  $p < 0.001$ , PRS4;  $p < 0.01$ , PRS5;  $p < 0.05$ , PRS6;  $p < 0.1$ , PRS7;  $p < 0.2$ , PRS8;  $p < 0.3$ , PRS9;  $p < 0.4$ , PRS10;  $p < 0.5$ , PRS11;  $p < 0.6$ , PRS12;  $p < 0.7$ , PRS13;  $p < 0.8$ , PRS14;  $p < 0.9$ , PRS15;  $p = 1$ ).

**Supplementary figure 3.** Fisher's exact test heatmap plot for clinical OCD  $p$ -values ( $P_2$ ) conditioned on OC symptom scores  $p$ -values ( $P_1$ ).

**Supplementary figure 4.** Quantile--quantile (QQ) plots of observed versus expected  $-\log(P)$  statistics

**Supplementary figure 5.** Regional association plot for the top region identified in the GWAS. The top associated SNP (rs8100480, purple dot at the center) is depicted in a genomic window of 800 Kb (hg19).  $P$ -values are given, as well as linkage disequilibrium (LD) strength ( $r^2$ ; data from the 1000 genomes project European samples) between the sentinel SNP and its flanking markers; the strength of the association is illustrated by the colour of the dots for neighbouring markers (dark blue – low LD, red – high LD). The blue lines indicate the estimated recombination rate (cM per Mb).

**Supplementary figure 6.** Quantile-quantile plot for the association analysis.

**Supplementary figure 7.** Manhattan plot for the full gene association analysis.

## Tables

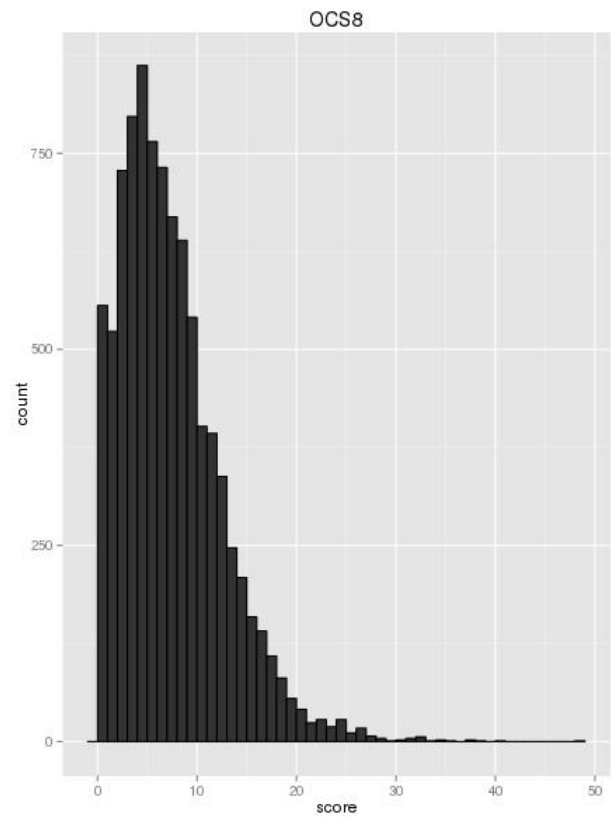
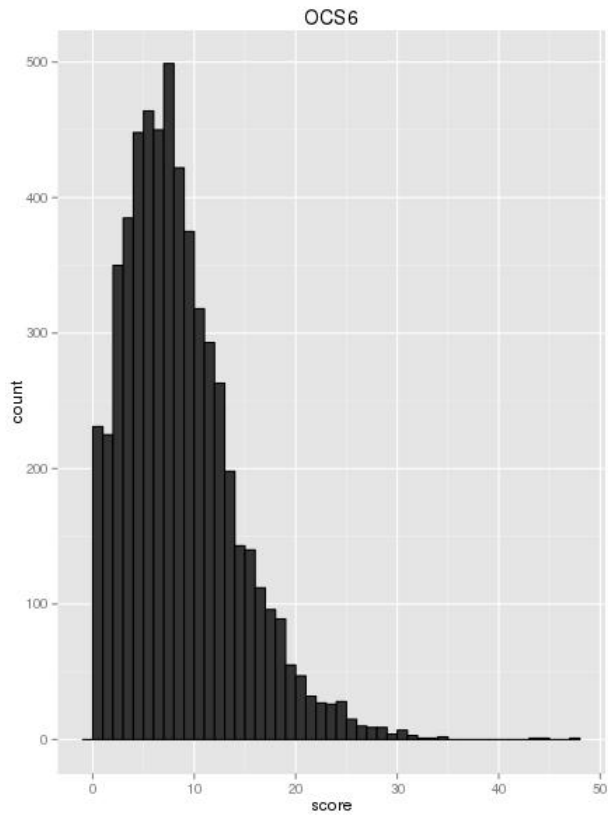
**Supplementary table 1.** Extension of table 3. Listed are all the SNPs with a P-value  $< 10^{-6}$  for the GWAS results. Allele frequency and INFO represent additional information on the frequency of the minor allele and the imputation info score, respectively. Chromosome (Chr) and base pair position (BP), based on hg19 build, are also given. The beta indicates the effect size, and the direction of the association is given by its positive or negative value. The location of each SNP is given in the last column; when located in non-intronic locus, the left and right closest flanking genes are additionally noted. A1 and A2 indicate the effect allele and the non-effect allele, respectively.

**Supplementary table 2.** Abbreviations: SNP, single nucleotide polymorphism; CHR, chromosome; BP, base pair position; IOCDF-GC, The international OCD foundation Genetics Collaborative; MA, minor allele in IOCDF-GC meta-analyses; NA, not applicable. SNPs listed are strongest associated GWAS variants in combined trio–case-control samples reported in Table 1 of Stewart et al. 2013. Directions of effects (IOCDF-GC: DIR and OCS-NTR: DIR) are given based on the A1 in the IOCDF analyses. For the IOCDF-GC GWAS, directions are given for all study samples (EU, AJ, SA, trios).

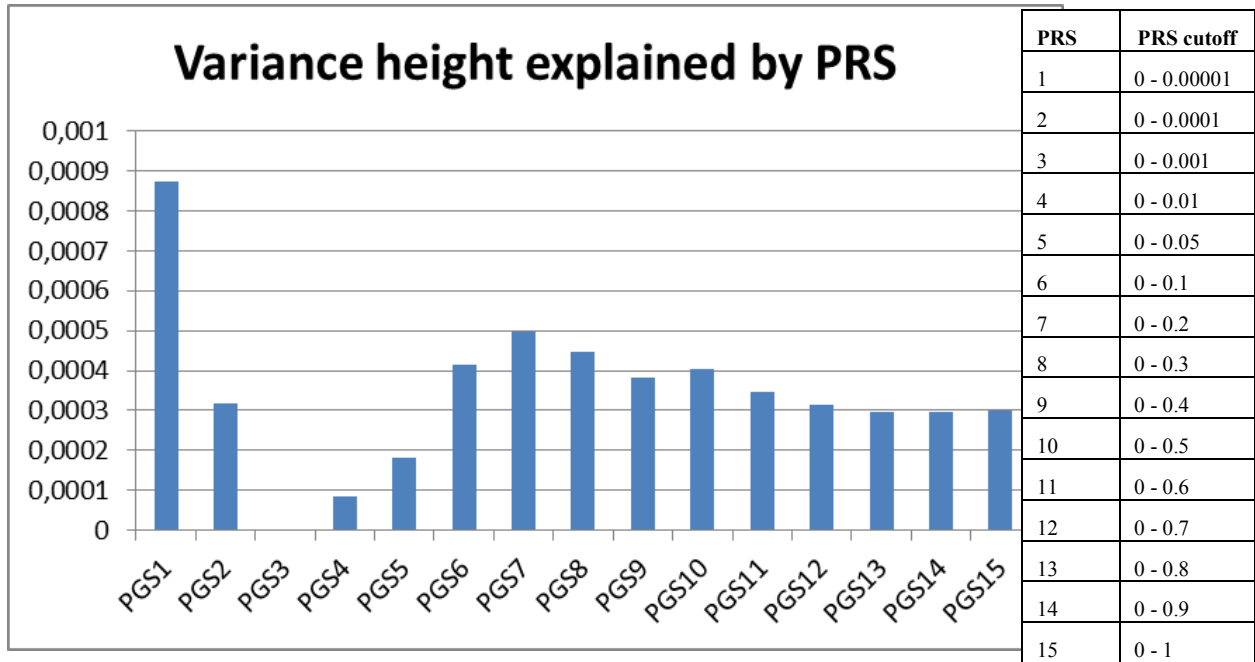
**Supplementary table 3.** Abbreviations: SNP, single nucleotide polymorphism; CHR, chromosome; BP, base pair position; OCGAS, The OCD Collaborative Genetics Association Study; MA, minor allele in OCGAS meta-analyses; NA, not applicable. SNPs listed are strongest associated GWAS variants reported in Table 1 of Mattheisen et al. 2014. Directions of effects (OCGAS: DIR and OCS-NTR: DIR) are given based on the A1 in the OCGAS analyses).



**Supplementary figure 1**

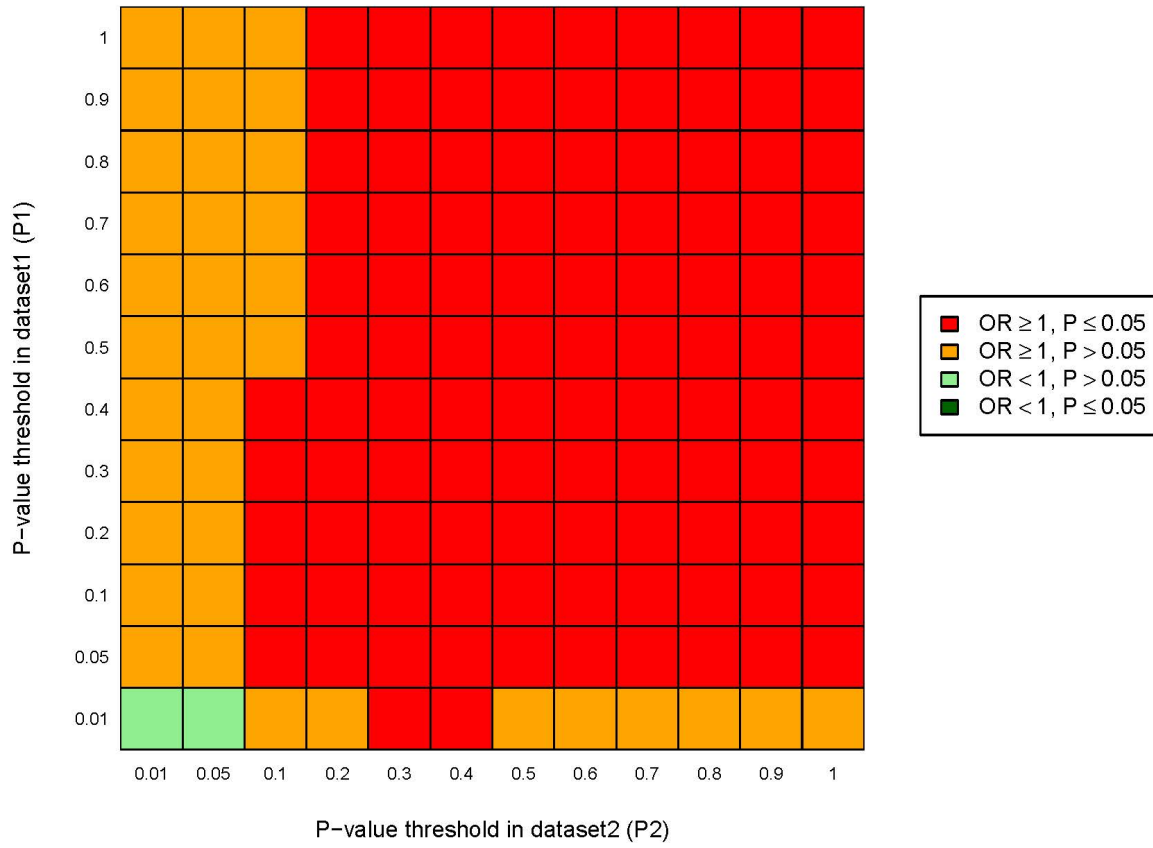


**Supplementary figure 2**

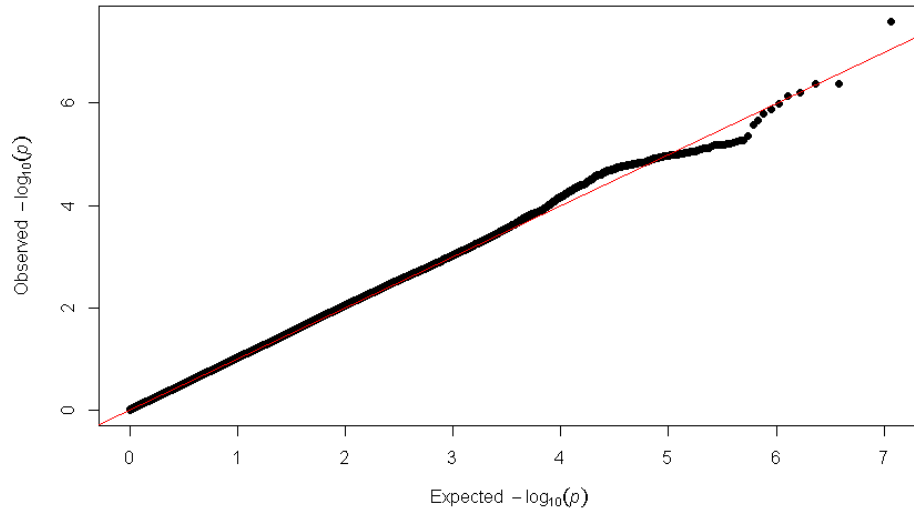


### Supplementary figure 3

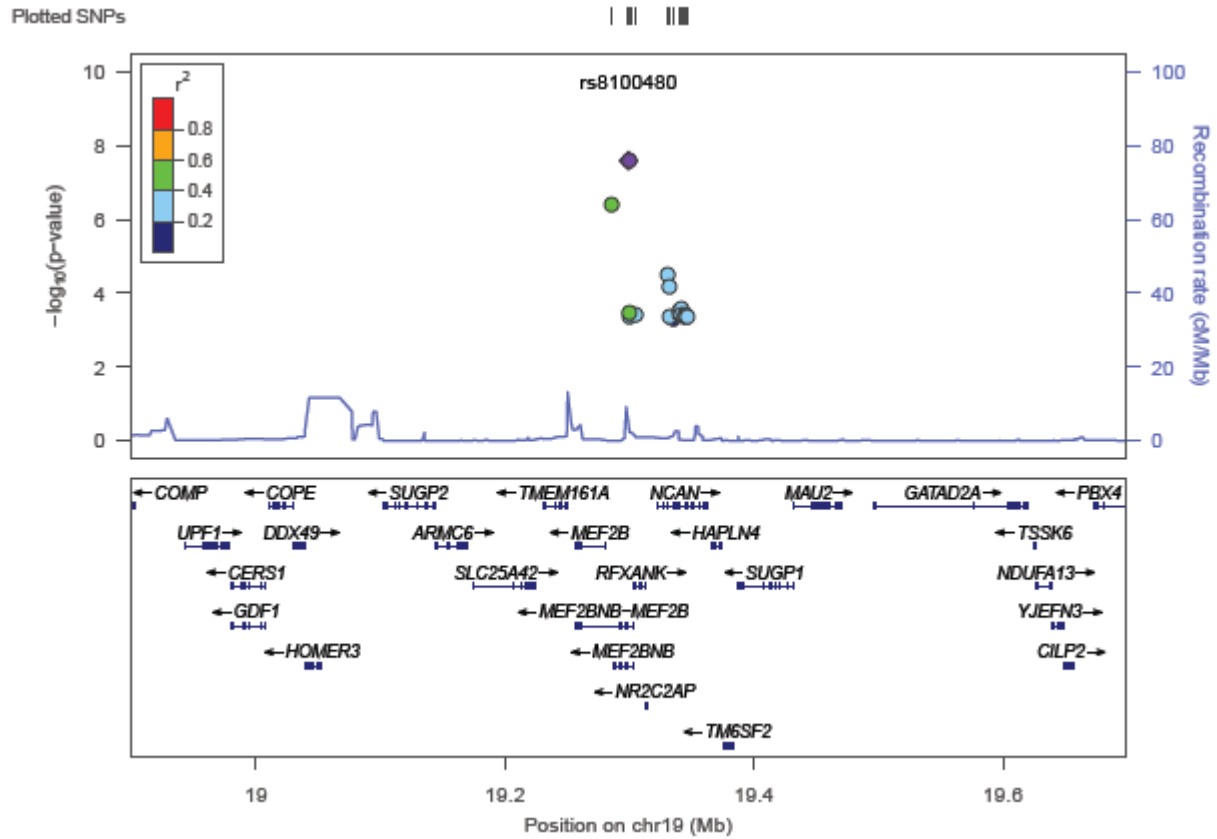
Fisher's tests for association of SNP effect direction between the two datasets



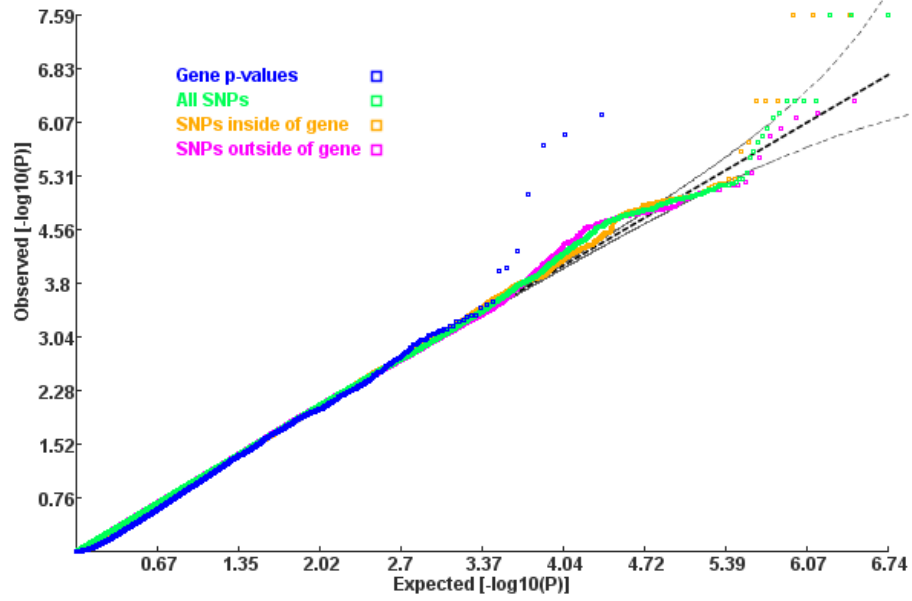
**Supplementary figure 4**



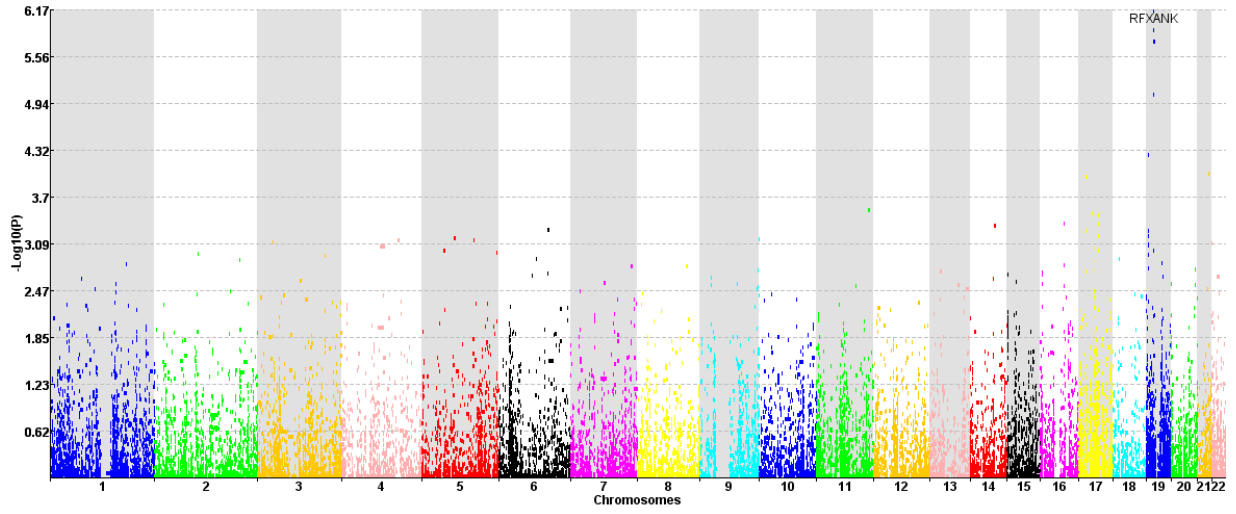
# Supplementary figure 5



## Supplementary figure 6



Supplementary figure 7



**Supplementary table 1.** Top associated variants in NTR-GWAS analysis

| SNP         | CHR | BP(hg19)  | P-value  | A1/<br>A2 | BETA    | Allele<br>frequency | INFO   | Intragenic location      |                           |                            |
|-------------|-----|-----------|----------|-----------|---------|---------------------|--------|--------------------------|---------------------------|----------------------------|
|             |     |           |          |           |         |                     |        | Left gene(kb)            |                           | Right gene(kb)             |
| rs8100480   | 19  | 19299079  | 2.56E-08 | C/T       | 1.4095  | 0.0367              | 0.6877 |                          | <i>MEF2BNB</i>            |                            |
| rs11671119  | 19  | 19286077  | 4.11E-07 | C/T       | 1.2036  | 0.0302              | 0.9816 |                          | <i>MEF2BNB-<br/>MEF2B</i> |                            |
| rs4818048   | 21  | 40908952  | 4.24E-07 | C/G       | 0.8956  | 0.0601              | 0.9606 | <i>LOC729056 (7173)</i>  |                           | <i>B3GALT5(19417)</i>      |
| rs4818050   | 21  | 40910600  | 6.37E-07 | C/T       | 0.8725  | 0.0597              | 0.9789 | <i>LOC729056 (8821)</i>  |                           | <i>B3GALT5(17769)</i>      |
| rs77959192  | 21  | 40911027  | 7.36E-07 | C/A       | 0.8707  | 0.0595              | 0.9782 | <i>LOC729056 (9248)</i>  |                           | <i>B3GALT5(17342)</i>      |
| rs4818049   | 21  | 40910464  | 1.03E-06 | G/A       | 0.8557  | 0.0609              | 0.9734 | <i>LOC729056 (8625)</i>  |                           | <i>B3GALT5(17965)</i>      |
| rs77615161  | 21  | 40911050  | 1.30E-06 | G/A       | 0.8541  | 0.0603              | 0.9714 | <i>LOC729056 (9271)</i>  |                           | <i>B3GALT5(17365)</i>      |
| rs17384439  | 4   | 96424680  | 1.59E-06 | T/C       | 0.8323  | 0.0589              | 0.896  |                          | <i>UNC5C</i>              |                            |
| rs2837096   | 21  | 40978013  | 2.18E-06 | G/A       | 0.747   | 0.0765              | 0.9688 |                          | <i>C21orf88</i>           |                            |
| rs4818052   | 21  | 40912745  | 2.64E-06 | A/G       | 0.8475  | 0.0581              | 0.9723 | <i>LOC729056 (10966)</i> |                           | <i>B3GALT5(15670)</i>      |
| rs77460585  | 5   | 101123995 | 4.45E-06 | G/A       | 0.8668  | 0.0539              | 0.9539 | <i>LOC100420593</i>      |                           | <i>OR7H2P</i>              |
| rs581043    | 3   | 62830115  | 5.26E-06 | C/T       | 0.4358  | 0.3043              | 0.9916 |                          | <i>CADPS</i>              |                            |
| rs999719    | 22  | 34264838  | 5.33E-06 | T/A       | 0.4442  | 0.3063              | 10.007 |                          | <i>LARGE</i>              |                            |
|             | 4   | 96399160  | 5.57E-06 | R/D       | 0.7201  | 0.0682              | 0.9803 |                          |                           |                            |
| rs74276709  | 21  | 40913995  | 5.85E-06 | A/G       | 0.8226  | 0.056               | 0.9907 | <i>LOC729056</i>         |                           | <i>B3GALT5</i>             |
| rs17024030  | 4   | 96399606  | 6.07E-06 | G/A       | 0.7155  | 0.0681              | 0.9851 |                          | <i>UNC5C</i>              |                            |
| rs9520326   | 13  | 107865442 | 6.37E-06 | T/C       | -0.4326 | 0.4301              | 0.9852 |                          | <i>FAM155A</i>            |                            |
| rs79219884  | 21  | 40899981  | 6.40E-06 | A/T       | 0.8467  | 0.0547              | 0.9694 |                          | <i>LOC729056</i>          |                            |
| rs60588302  | 9   | 7900777   | 6.44E-06 | C/T       | 11.278  | 0.0278              | 0.9002 | <i>C9orf123 (100971)</i> |                           | <i>TPRD (413469)</i>       |
| rs11658311  | 17  | 17470526  | 6.50E-06 | C/T       | 0.7719  | 0.0622              | 10.119 |                          | <i>PEMT</i>               |                            |
| rs182187683 | 21  | 40908159  | 6.52E-06 | A/G       | 0.8927  | 0.0507              | 0.9494 | <i>LOC729056 (6380)</i>  |                           | <i>B3GALT5 (20210)</i>     |
| rs186307979 | 21  | 40908160  | 6.53E-06 | G/T       | 0.8926  | 0.0507              | 0.9492 | <i>LOC729056 (6381)</i>  |                           | <i>B3GALT5 (20209)</i>     |
| rs113123495 | 19  | 5587540   | 6.78E-06 | C/T       | 10.582  | 0.0349              | 0.7935 |                          | <i>SAFB2</i>              |                            |
| rs9619381   | 22  | 34262097  | 7.20E-06 | G/T       | 0.4396  | 0.3058              | 0.999  |                          | <i>LARGE</i>              |                            |
| rs10815705  | 9   | 7902402   | 7.58E-06 | T/G       | 11.251  | 0.0275              | 0.9184 | <i>C9orf123 (102596)</i> |                           | <i>PTPRD (411844)</i>      |
| rs75494728  | 1   | 90440967  | 7.69E-06 | G/A       | 0.9241  | 0.1016              | 0.47   | <i>LRRRC8D (38976)</i>   |                           | <i>ZNF326 (19711)</i>      |
| rs142967832 | 21  | 40900470  | 7.72E-06 | T/C       | 0.8391  | 0.0545              | 0.9729 |                          | <i>LOC729056</i>          |                            |
| rs75477292  | 21  | 40900608  | 7.78E-06 | G/A       | 0.8389  | 0.0545              | 0.973  |                          | <i>LOC729056</i>          |                            |
| rs57277948  | 17  | 17481317  | 7.96E-06 | T/C       | 0.856   | 0.062               | 0.8334 |                          | <i>PEMT</i>               |                            |
| rs79874412  | 21  | 40900861  | 8.18E-06 | G/A       | 0.8376  | 0.0544              | 0.9735 |                          | <i>LOC729056</i>          |                            |
| rs117114519 | 21  | 40901172  | 8.30E-06 | T/C       | 0.8372  | 0.0544              | 0.9739 |                          | <i>LOC729056</i>          |                            |
| rs117465668 | 21  | 40901183  | 8.36E-06 | C/G       | 0.8371  | 0.0544              | 0.9739 |                          | <i>LOC729056</i>          |                            |
|             | 17  | 17471374  | 8.74E-06 | R/I       | 0.7628  | 0.0623              | 10.104 |                          | <i>EEF1A1P43</i>          |                            |
| rs7332896   | 13  | 107865698 | 8.82E-06 | C/T       | -0.4269 | 0.4297              | 0.9897 |                          | <i>FAM155A</i>            |                            |
| rs74813064  | 21  | 40901563  | 9.01E-06 | T/C       | 0.8358  | 0.0542              | 0.9745 |                          | <i>LOC729056</i>          |                            |
| rs67366981  | 14  | 77703964  | 9.01E-06 | T/C       | -0.6535 | 0.1253              | 10.126 |                          | <i>TMEM63C</i>            |                            |
| rs67130018  | 14  | 77703280  | 9.00E-06 | C/T       | -0.6815 | 0.1205              | 0.9691 |                          | <i>TMEM63C</i>            |                            |
| rs9904004   | 17  | 17476791  | 9.06E-06 | A/G       | 0.7702  | 0.0618              | 10.096 |                          | <i>PEMT</i>               |                            |
| rs7394339   | 10  | 85120682  | 9.15E-06 | G/A       | -0.6012 | 0.1486              | 0.9544 | <i>MARK2P15 (46475)</i>  |                           | <i>HMG2P8<br/>(720405)</i> |
| rs1997543   | 21  | 40974288  | 9.23E-06 | G/C       | 0.698   | 0.0734              | 10.069 |                          | <i>C21orf88</i>           |                            |
| rs12627674  | 21  | 40901786  | 9.27E-06 | G/C       | 0.8366  | 0.0539              | 0.9745 | <i>LOC729056 (7)</i>     |                           | <i>B3GALT5(26583)</i>      |
| rs8007762   | 14  | 77708465  | 9.62E-06 | A/C       | -0.6602 | 0.1287              | 0.9759 |                          | <i>TMEM63C</i>            |                            |
| rs12627733  | 21  | 40902206  | 9.74E-06 | C/T       | 0.8386  | 0.0534              | 0.9741 | <i>LOC729056 (427)</i>   |                           | <i>B3GALT5 (26163)</i>     |
| rs698737    | 8   | 117222665 | 9.75E-06 | G/C       | 11.281  | 0.033               | 0.8179 | <i>TRPS1 (509366)</i>    |                           | <i>EIF3H (434390)</i>      |
| rs750095    | 13  | 107868758 | 9.86E-06 | G/C       | -0.4258 | 0.3919              | 0.992  |                          | <i>FAM155A</i>            |                            |
| rs17798894  | 3   | 9061714   | 9.90E-06 | T/G       | 0.4585  | 0.2303              | 10.133 |                          | <i>SRGAP3</i>             |                            |



**Supplementary table 2.** Follow-up of IOCDF-GC GWAS results

| MARKERINFO       |          |                  | IOCDF-GC        |            |             | OCS-NTR       |          | Intragenic location |              |            |
|------------------|----------|------------------|-----------------|------------|-------------|---------------|----------|---------------------|--------------|------------|
| SNP              | CHR      | BP (hg19)        | P-Value         | A1/A2      | DIR         | P-Value       | DIR      | Left gene           |              | Right gene |
| rs11898020       | 2        | 144282078        | 2.56E-04        | A/G        | ++-         | 0.3457        | -        |                     | ARHGAP1<br>5 |            |
| rs10165908       | 2        | 158315629        | 0.0169          | C/T        | +++         | 0.227         | -        | CYTIP               |              | ACVR1C     |
| rs1838733        | 5        | 58533392         | 3.82E-05        | T/C        | ----        | 0.4355        | -        |                     | PDE4D        |            |
| rs26728          | 5        | 106946056        | 1.01E-04        | T/C        | ---         | 0.2454        | -        |                     | EFNA5        |            |
| <b>rs4868342</b> | <b>5</b> | <b>173504522</b> | <b>3.20E-05</b> | <b>C/T</b> | <b>++++</b> | <b>0.0049</b> | <b>+</b> |                     | <b>HMP19</b> |            |
| rs9499708        | 6        | 104445367        | 2.96E-06        | C/T        | --+         | NA            | NA       | GRIK2               |              | HACE1      |
| rs2205748        | 6        | 104462555        | 8.52E-06        | G/A        | ++++        | NA            | NA       | GRIK2               |              | HACE1      |
| rs182320         | 6        | 130073291        | 2.25E-05        | C/T        | ++++        | 0.8009        | +        | ARHGAP18            |              | C6orf191   |
| rs6531002        | 8        | 12722703         | 0.0067          | T/C        | ---         | NA            | NA       | LONRF1              |              | KIAA1456   |
| rs11611761       | 12       | 33025612         | 0.115           | A/C        | --+         | 0.0746        | +        |                     | PKP2         |            |
| rs297941         | 12       | 50319086         | 4.99E-07        | G/A        | ----        | NA            | NA       | FAIM2               |              | AQP2       |
| rs9652236        | 13       | 72688774         | 5.14E-06        | T/G        | ++++        | 0.4476        | -        | DACH1               |              | MZT1       |
| rs11081062       | 18       | 3662879          | 2.92E-05        | T/C        | +0++        | 0.6318        | +        |                     | DLGAP1       |            |
| rs11663827       | 18       | 3663631          | 2.31E-05        | A/G        | ++++        | 0.5986        | +        |                     | DLGAP1       |            |
| rs485186         | 19       | 49207206         | 9.94E-06        | G/A        | ++++        | 0.3297        | -        |                     | FUT2         |            |
| rs6131295        | 20       | 11996267         | 3.63E-05        | G/A        | ---+        | 0.7904        | -        | BTBD3               |              | SPTLC3     |

**Supplementary table 3.** Follow-up of OCGAS-GWAS results

| MARKERINFO |         |           | OCGAS    |       |     | OCS-NTR |         | Intragenic location |              |
|------------|---------|-----------|----------|-------|-----|---------|---------|---------------------|--------------|
| SNP        | CH<br>R | BP (hg19) | P-Value  | A1/A2 | DIR | P-Value | DI<br>R | Left gene           | Right gene   |
| rs4401971  | 9       | 11890045  | 4.13E-07 | A/G   | -   | 0.1437  | +       | LOC646114           | LOC100049717 |
| rs6876547  | 5       | 25572301  | 1.76E-06 | G/T   | +   | 0.0825  | +       | CDH10               | MSNL1        |
| rs13437953 | 7       | 54313171  | 9.69E-06 | A/C   | -   | 0.6638  | +       | FLJ45974            | LOC222005    |
| rs6452234  | 5       | 24922789  | 1.13E-05 | A/G   | -   | NA      | NA      | CDH10               | MSNL1        |
| rs10149510 | 14      | 24039559  | 1.49E-05 | T/C   | +   | NA      | NA      |                     | JPH4         |
| rs7462051  | 8       | 80266553  | 1.69E-05 | A/G   | +   | 0.6567  | +       | IL7                 | LOC100128963 |
| rs1544352  | 16      | 19713882  | 1.94E-05 | C/T   | +   | NA      | NA      | C16orf62            | C16orf88     |
| rs1532154  | 4       | 157279663 | 2.22E-05 | C/T   | -   | 0.4974  | -       | FTHP2               | hCG_1814936  |
| rs509876   | 11      | 79811268  | 2.29E-05 | G/A   | +   | 0.1965  | -       | LOC646112           | LOC729790    |
| rs10392    | 20      | 37550935  | 2.87E-05 | A/G   | -   | 0.0484  | -       |                     | PPP1R16<br>B |
| rs2821204  | 9       | 11683901  | 2.89E-05 | G/A   | +   | 0.8537  | +       | LOC646114           | LOC100049717 |
| rs10882583 | 10      | 97109019  | 3.56E-05 | A/G   | -   | 0.1149  | -       |                     | SORBS1       |
| rs2278144  | 8       | 25634077  | 3.94E-05 | T/C   | -   | 0.7461  | -       | CDCA2               | EBF2         |
| rs8026755  | 15      | 37247722  | 4.03E-05 | A/C   | -   | 0.7806  | +       |                     | MEIS2        |
| rs838209   | 3       | 176352763 | 4.14E-05 | C/T   | -   | 0.1190  | -       | LOC730168           | TBL1XR1      |
| rs17735629 | 5       | 168653502 | 5.03E-05 | C/T   | +   | 0.6468  | +       |                     | SLIT3        |
| rs7003102  | 8       | 126303493 | 5.14E-05 | G/A   | +   | 0.0598  | -       |                     | NSMCE2       |
| rs12547343 | 8       | 26833198  | 5.38E-05 | A/C   | -   | 0.4138  | -       | LOC100127897        | LOC100132229 |
| rs7593878  | 2       | 44358504  | 5.44E-05 | C/A   | -   | 0.9034  | +       | LRPPRC              | PPM1B        |
| rs16867406 | 2       | 182189405 | 6.10E-05 | T/C   | +   | 0.8464  | -       | LOC729026           | LOC100127923 |
| rs9541148  | 13      | 35095884  | 6.73E-05 | C/T   | -   | NA      | NA      | LOC100130499        | LOC100129452 |
| rs1671253  | 18      | 50040561  | 7.04E-05 | T/C   | -   | 0.9944  | +       |                     | DCC          |
| rs973714   | 6       | 133578733 | 7.44E-05 | A/G   | +   | 0.9983  | +       |                     | EYA4         |
| rs3902042  | 10      | 17065067  | 7.71E-05 | T/C   | +   | NA      | NA      |                     | CUBN         |
| rs8120171  | 20      | 50712059  | 7.85E-05 | T/C   | -   | NA      | NA      |                     | ZFP64        |
| rs16851014 | 4       | 75216365  | 7.89E-05 | A/G   | -   | 0.2436  | -       | EPGN                | EREG         |
| rs9845643  | 3       | 179531062 | 8.58E-05 | G/A   | -   | 0.4279  | -       |                     | PEX5L        |
| rs11265908 | 9       | 92077132  | 8.60E-05 | C/T   | -   | 0.3481  | -       |                     | SEMA4D       |
| rs16997877 | 20      | 16533162  | 8.68E-05 | C/T   | +   | 0.7431  | -       |                     | KIF16B       |
| rs2183738  | 9       | 71211200  | 8.73E-05 | G/A   | +   | 0.5192  | +       | C9orf71             | LOC347097    |
| rs12096987 | 1       | 185201061 | 8.87E-05 | C/T   | +   | 0.6838  | +       |                     | C1orf26      |
| rs7005206  | 8       | 130620813 | 9.13E-05 | A/G   | -   | 0.7417  | +       | LOC100129525        | LOC100130376 |