

THE LANCET

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: arcOGEN Consortium and arcOGEN Collaborators. Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. *Lancet* 2012; published online July 3. [http://dx.doi.org/10.1016/S0140-6736\(12\)60681-3](http://dx.doi.org/10.1016/S0140-6736(12)60681-3).

Identification of new susceptibility loci for osteoarthritis -the arcOGEN study

WEBAPPENDIX SUPPLEMENTARY MATERIAL

Eleftheria Zeggini, Kalliope Panoutsopoulou, Lorraine Southam, Nigel W. Rayner, Aaron G. Day-Williams, Margarida C. Lopes, Vesna Boraska, Tonu Esko, Evangelos Evangelou, Albert Hofman, Jeanine J. Houwing-Duistermaat, Thorvaldur Ingvarsson, Ingileif Jonsdottir, Helgi Jonsson, Hanneke J.M. Kerkhof, Margreet Kloppenburg, Steffan D Bos, Massimo Mangino, Sarah Metrustry, P. Eline Slagboom, Gudmar Thorleifsson, Emma V.A. Raine, Madhushika Ratnayake, Michelle Ricketts, Claude Beazley, Hannah Blackburn, Suzannah Bumpstead, Katherine S. Elliott, Sarah E. Hunt, Simon C. Potter, So-Youn Shin, Vijay K. Yadav, Guangju Zhai, Kate Sherburn, Kate Dixon, Elizabeth Arden, Nadim Aslam, Phillippa-Kate Battley, Ian Carluke, Sally Doherty, Andrew Gordon, John Joseph, Richard Keen, Nicola C. Koller, Sheryl Mitchell, Fiona O'Neill, Ellen Paling, Mike R. Reed, Fernando Rivadeneira, Diane Swift, Kirsten Walker, Bridget Watkins, Maggie Wheeler, Fraser Birrell, John P.A. Ioannidis, Ingrid Meulenbelt, Andres Metspalu, Ashok Rai, Donald Salter, Kari Stefansson, Unnur Styrkarsdottir, André G. Uitterlinden, Joyce B.J. van Meurs, arcOGEN Consortium, Kay Chapman, Panos Deloukas, William E.R. Ollier, Gillian A. Wallis, Nigel Arden, Andrew Carr, Michael Doherty, Andrew McCaskie, J. Mark Wilkinson, Stuart H. Ralston, Ana M. Valdes, Tim D. Spector, and John Loughlin.

Address correspondence to:

Dr. Zeggini, Wellcome Trust Sanger Institute, The Morgan Building, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1HH, UK.

eleftheria@sanger.ac.uk

Tel: ++44 1223 496868

Prof. Loughlin, Newcastle University, Institute of Cellular Medicine, 4th Floor Catherine Cookson Building, Newcastle University, Medical School, Framlington Place, Newcastle upon Tyne, NE2 4HH, UK.

john.loughlin@ncl.ac.uk

Tel: ++44 191 2227178

WEBAPPENDIX

SUPPLEMENTARY METHODS

arcOGEN samples

Cases. The arcOGEN case samples were collected and genotyped in two stages. All participants gave written informed consent and the study protocols were reviewed and approved by the appropriate review committees. The stage one samples comprised 3,086 cases from existing DNA collections from five United Kingdom locations within the arcOGEN consortium (London, Nottingham, Oxford, Sheffield, and Southampton). The detailed characteristics of these cases are described elsewhere.¹ Briefly, all were unrelated and of European origin, and all had primary OA of the hip or knee of radiographic Kellgren-Lawrence (KL) grade ≥ 2 , or clinical evidence of disease to a level requiring total joint replacement (TJR). The exclusion criteria included the need for joint replacement due to fracture, secondary OA of any cause, and developmental, vascular, or infective causes of joint disease.

The stage two cases were collected prospectively as part of the arcOGEN study at nine locations across the UK (Edinburgh, London, Newcastle-Upon-Tyne, Nottingham, Oxford, Sheffield, Southampton, Wansbeck, and Worcester). The ascertainment criterion for the majority of cases ($n=4,212$) was primary OA that was severe enough for the individual to require joint replacement of the hip or of the knee; an additional small number ($n=112$) of cases were collected as part of a randomized, placebo controlled trial of vitamin D replacement (VIDEO Study) where the ascertainment criterion was radiographic disease of the knee, with a KL grade ≥ 2 . All cases were unrelated and of European origin. The exclusion criteria included the need for joint replacement due to fracture, secondary OA of any cause, and developmental, vascular, or infective causes of joint disease. The prospective collections were approved by the National Research Ethics Service in the United Kingdom, and all subjects provided written, informed consent prior to inclusion. The arcOGEN stage one and two study sample characteristics are summarized in Table 1.

Controls. The study used two different types of controls: population-based, unrelated UK controls which came from five distinct sources: the 1958 Birth Cohort (58BC) and the UK Blood Donor Service (UKBS) from the Wellcome Trust Case Control Consortium 2 (WTCCC2) study, the 1958 Birth Cohort from the Type 1 Diabetes Genetics Consortium (T1DGC) study, the Avon Longitudinal Study of Parents and Children (ALSPAC) and the People of the British Isles (PoBI) study (ST1); and unrelated, OA-free controls (females only) from the TwinsUK cohort which comprises twins ascertained to study the heritability and genetics of age-related diseases.² These unselected twins were recruited from the general population through national media campaigns in the UK and shown to be comparable to age-matched population singletons in terms of disease-related and lifestyle characteristics.³

Genome-wide genotyping

DNA sample preparation and quality control (QC). Genomic DNA from extant cohorts and blood samples from new cases were shipped by the collection centers to the Centre for Integrated Genomic Medical Research (CIGMR), University of Manchester, for processing. DNA extraction was performed as previously described.¹ 2D bar coded matrix plates and patient data manifests for each set of 96 DNAs were shipped to The Wellcome Trust Sanger Institute, Hinxton. The quality of the DNA and subject identity were validated using a Sequenom iPLEX assay designed to genotype four gender-specific SNPs and 26 SNPs present on the Illumina Human 610-Quad array (Illumina, San Diego, USA). Sample preparation (quantification of DNA concentration and QC criteria) was performed as previously described.¹

Genotyping and allele calling. DNA samples passing QC were rearrayed into 96-well plates for genotyping using Illumina Human 610-Quad BeadChips (Illumina, San Diego, USA), which captures $\sim 90\%$ of common variation in European populations⁴, as previously described.¹ Normalized bead intensity data obtained for each sample were converted into SNP genotypes using the Illuminus genotype calling algorithm.⁵ The same calling algorithm was used for the controls.

Genotype quality control

Sample QC. Sample QC was performed separately for the cases and each control cohort. We used the same method as previously described.¹ Briefly, individuals were excluded if they had a call rate $< 97\%$ and if they had a discrepancy between the supplied gender and the gender that was inferred using X-chromosome heterozygosity rates in PLINK.⁶ We also excluded samples with excess genome-wide homozygosity and heterozygosity (based on autosomal SNPs only) by examining histograms and removing visual outliers. The thresholds we chose for each particular data set were: arcOGEN cases $< 30.5\%$ and $> 34\%$; ALSPAC $< 30.5\%$ and $> 33\%$; PoBI $< 26.8\%$ and $> 28.3\%$; T1DGC $< 31\%$ and $> 32.75\%$; WTCCC2-58BC $< 32.5\%$ and $> 33.74\%$; WTCCC2-UKBS $< 26.6\%$ and $> 28.1\%$; and TwinsUK $< 30.42\%$ and $> 32.67\%$. We removed duplicated/related samples by calculating the genome-wide pair-wise identity by descent (IBD) for each sample; we excluded any sample with a $\pi^{\wedge} > 0.2$.⁶ We examined ethnicity by performing multidimensional scaling (MDS)⁶ and principle component analysis (PCA) using Goldsurfer⁷; we combined autosomal information from each of our populations with 210 founders from the HapMap phase II release 23a data, which included the European (CEU), African (YRI) and Chinese/Japanese (ASN) populations and we excluded samples of non-European descent. There were 7709 cases pre sample QC and SNP QC was performed on 7,420 cases as described below. Subsequently we removed ten cases that failed other phenotype checks. For the ALSPAC, PoBI, WTCCC2-58BC and WTCCC2-UKBS cohorts we excluded individuals identified from additional QC analyses performed by the suppliers of the datasets. The number of samples

passing all QC filters from each population-based cohort is shown in ST1. Sample QC of the female OA-free controls from the TwinsUK dataset has been previously described.¹ We also excluded 47 samples that failed QC when the normalized intensity data for these samples were pooled with data from a further three genotype platforms HumanHap300, 1M-Duo and 1·2M-Duo (Illumina, San Diego, USA) using the following criteria (i) sample call rate <98%; (ii) heterozygosity across all SNPs ≥ 2 SD from the sample mean; (iii) evidence of non-European ancestry, assessed by PCA comparison with HapMap phase III populations; (iv) observed pairwise IBD probabilities suggestive of sample identity errors. We also excluded 154 samples that were part of the TwinsUK replication set. 1,828 OA-free samples passing QC and not overlapping with the TwinsUK replication dataset were used for further analyses.

SNP QC. SNP QC was performed separately for the cases and each control cohort on autosomal SNPs using thresholds that have been previously described.¹ In short, SNPs were excluded if they had a minor allele frequency (MAF) $\geq 5\%$ and a call rate <95%, or a MAF <5% and call rate <99%. All monomorphic SNPs were excluded and also those SNPs with an exact Hardy Weinberg Equilibrium (HWE) $p < 0.0001$. The number of SNPs prior to QC for each dataset was: for the arcOGEN cases: 599,011; for ALSPAC 580,030; for PoBI 1,090,938; for T1DGC 547,458; WTCCC2-58BC 1,115,428; WTCCC2-UKBS 1,115,428. The number of SNPs that passed standard QC in each dataset was: for arcOGEN cases 558,546; for ALSPAC 542,060; for PoBI 965,407; for T1DGC 535,266; for WTCCC2-58BC 1,005,923; for WTCCC2-UKBS 996,623; for TwinsUK 542,037. We extracted 492,811 overlapping SNPs between all population-based control datasets and performed pairwise association tests for all possible pairs. Following inspection of qq plots 866 SNPs with multiplicative $p < 10^{-4}$ were removed from all population-based control datasets prior to pooling them into one common control group. 490,130 SNPs from the common control group overlapped with the clean arcOGEN cases SNPs. Finally we removed 70 SNPs following inspection of intensity plots of the top signals ($p < 10^{-5}$) that were identified in our case/control association analyses and 4,569 SNPs with MAF <1% (estimated separately for each cohort). The resulting OA cases v. population-based controls dataset contained 7,410 cases, 11,009 controls and 485,491 autosomal SNPs (Table 1). From the TwinsUK dataset 532,499 autosomal SNPs overlapped with the QCed SNPs for arcOGEN cases. We also removed 72 SNPs following inspection of intensity plots of the top signals ($p < 10^{-5}$) that were identified in our case/control association analyses and 193 SNPs with MAF <1%. The resulting female OA cases v. OA-free controls from the TwinsUK dataset contained 4,476 cases, 1,828 controls and 532,234 autosomal SNPs (Table 1). Additional QC for differential missingness between cases and controls was performed for the top signals and is described in the section on prioritization of SNPs for replication.

Association analysis of stage 1 and 2 genotype data

Case/control association analyses were carried out using either population-based controls (Table 1 and ST1) or OA-free controls (females only) (Table 1). For the cases v. population-based control analyses 485,491 autosomal SNPs were analyzed under the multiplicative (or log additive) model using PLINK.⁶ Analyses were also carried out stratifying the cases by joint site of OA, gender and also by OA severity (TJR cases only). For the female cases v. female OA-free TwinsUK controls association analyses 532,234 SNPs were analyzed under the multiplicative model. Stratifications were performed according to joint site and severity of OA. The genomic control (GC) value is given as the calibrated inflation factor λ_{1000} .^{8,9}

Principal Component Analysis

To account for population stratification we performed principal components analysis (PCA) using EIGENSTRAT.¹⁰ We repeated all genome-wide analyses by using the first ten PCs as covariates in a logistic regression analysis using PLINK.⁶ In addition, we visualized qqplots and calculated the calibrated inflation factor λ_{1000} in all PC-adjusted and unadjusted analyses.^{8,9}

Prioritization of SNPs for replication

We prioritized 129 SNPs for follow-up by *in silico* replication by the deCODE, EGCUT, GARP, Rotterdam (RSI and RSII cohorts) and TwinsUK studies primarily based on statistical significance. Briefly, we prioritized signals with multiplicative model $p < 10^{-5}$ in the OA v. controls and all other stratified cases v. controls unadjusted analyses and investigated their QC properties. SNPs with differential missingness $p < 10^{-4}$ between cases and controls were not taken forward. Similarly, SNPs with poor cluster plots (plotted separately for the cases and for each control group) were omitted from further analyses. We preferentially selected SNPs with MAF >5% to boost power of replication. Independent SNPs were defined on the basis of $r^2 < 0.4$. We prioritized 100 SNPs, representing 98 independent signals, from the OA v. population-based controls analyses and 29 different SNPs, representing 29 independent signals, from the OA v. OA-free controls analyses (ST4). After obtaining *in silico* replication data from the deCODE, EGCUT, GARP, Rotterdam and TwinsUK studies, we carried out a meta-analysis across all available data and prioritized 26 of the 129 SNPs for *de novo* replication on the basis of their overall statistical significance (ST5).

Imputation and analysis of imputed data

We imputed genotypes for autosomal SNPs that were present in HapMap Phase III but were not present in the genome-wide chip or did not pass direct genotyping QC. In each sample, genotypes were imputed using the directly typed data and phased genotype data from all HapMap III populations.¹¹ Genotypes were imputed using the program IMPUTEv2¹² which determines the probability distribution of missing genotypes conditional on a set of known haplotypes and an estimated fine-scale recombination map. Imputation was based on 485,077 autosomal SNPs with MAF >0.01 (excluding SNPs that demonstrated poor genotype clustering upon manual inspection). Analysis of imputed SNPs was carried out using

SNPTESTv2 appropriately accounting for the probability distributions of imputed genotypes.¹³ We included only SNPs which passed post-imputation QC filters as follows: MAF >0.01 in cases and controls and an imputation information score >0.4. To examine whether any additional signals have arisen following imputation we removed all SNPs with $p \leq 10^{-3}$ within 100Kb of the 129 prioritized SNPs (SF4) (except for rs11177 and rs6976 for which SNPs within 500kb were removed due to extended LD within the associated region).

***In silico* replication samples, genotyping and analysis**

We sought to replicate in as large a sample size as possible and selected five OA cohorts that had previously been subjected to a GWAS for *in silico* replication of promising signals from the arcOGEN GWAS in the first instance: deCODE, EGCUT, GARP, the Rotterdam study (RSI and RSII cohorts) and TwinsUK (Table 1). Summary association statistics (under the log-additive model) were shared for 129 SNPs by EGCUT, RSI, RSII and TwinsUK, and 128 SNPs by deCODE and GARP. The results of these analyses guided the selection of 26 SNPs for further *de novo* replication.

deCODE. The study population comprises patients with OA of the knee and/or hip obtained on the basis of patients' records at hospitals and health care centers in Iceland.¹⁴ All OA patients had undergone joint replacement surgery of the knee or the hip. A clinician reviewed the patient records to verify the diagnosis. Control individuals were recruited as part of various genetic programs at deCODE. All individuals who had chip genotype data that passed quality control were included in the control group provided they had ever been on any OA list. The study was approved by the Data Protection Authority of Iceland and the National Bioethics Committee of Iceland. Informed consent was obtained from all participants. The samples were assayed with the Infinium HumanHap300 or HumanCNV370 SNP chips (Illumina, San Diego, USA) and called with BeadStudio. All of the genotyped SNPs tested in this report passed quality filtering (call rate >96%, MAF >0.01, HWE $p > 1 \times 10^{-6}$ on any of the three chip types used [HumanHap300, HumanHap300-Duo, and HumanCNV370]). SNPs were imputed using the IMPUTE software¹³ and phased haplotypes for the HapMap phase II CEU sample set data release 22. The difference in frequency of each SNP was tested, assuming an additive model, using a logistic regression test. Any samples with a yield <98% were excluded from the analysis.

Estonian Genome Center, University of Tartu (EGCUT). The Estonian cohort is drawn from the population-based biobank of the Estonian Genome Center of University of Tartu (EGCUT). The whole project is conducted according to the Estonian Genes Research Act and all participants have signed the informed consent.¹⁵⁻¹⁷ The current cohort size is over 51,515, from 18 years of age and over, which reflects closely the age distribution of the adult Estonian population. Subjects were recruited randomly when visiting general practitioners (GPs) or hospitals. Each recruit filled out a Computer Assisted Personal interview during 1-2 hours at a doctor's office, including personal data (place of birth, place(s) of living, nationality etc.), genealogical data (family history, three generations), educational and occupational history, and lifestyle data (physical activity, dietary habits, smoking, alcohol consumption, women's health, quality of life). Anthropometric and physiological measurements were also taken. OA was diagnosed by a GP and by an orthopedic surgeon or rheumatologist based on clinical symptoms (pain, stiffness and synovitis) and confirmed by imaging techniques (X-ray, CT scan or MRI). The OA cases for the current study had an ICD10 M16 and/or M17 diagnosis. All diseases are defined according to the ICD10 coding.¹⁸ At the moment of recruitment, the controls did not report diagnosis of osteoarthritis, psoriasis or autoimmune diseases. All the samples were genotyped with Illumina HumanCNV370 or HumanOmniExpress (Illumina, San Diego, USA) according to the Illumina protocol13 in the Estonian Biocenter. Data quality control was performed with PLINK⁶ (SNP call rate >98%; sample call rate >95%; MAF >0.01; HWE $p > 10^{-6}$; cryptic relatedness). Imputation was performed with IMPUTE¹³ and phased haplotypes from HapMap phase II release 22 data. Association analyses were carried out with SNPTEST.¹³ Inflation factors for directly genotyped and imputed data were 1.02 and 1.01 respectively.

Genetics OsteoArthritis and Progression (GARP) Study. The GARP study from Leiden, the Netherlands, consists of 192 sibling pairs concordant for clinical and radiographically confirmed OA at two or more joint sites among hand, spine (cervical or lumbar), knee or hip. As controls, we used a random sample of 1,670 middle aged subjects from the Rotterdam study. Written informed consent was obtained from each subject as approved by the ethical committees of the Leiden University Medical Center and Erasmus Medical Center. In the GARP study, conventional radiographs of the hands (dorso-volar), knees (posterior-anterior (PA) in weight bearing semiflexed and lateral), hips (PA), lumbar (PA and lateral), and cervical spine (anterior-posterior, lateral, and transbuccal) were obtained from all participants. Radiographic characteristics of OA were defined according to KL grade whereas symptoms were defined as pain in the particular joint on most days of the prior month which is in accordance with the American College of Rheumatology (ACR) recommendations and described in detail elsewhere.¹⁹ For the current study cases with radiographic OA (KL grade ≥ 2), knee (N=148), hip (N=106) and knee and/or hip (N=215) were selected. Genotypes of the replication SNPs were filtered from genome wide scan data. For the GARP study the genome wide scan was genotyped using Illumina Infinium HD Human660W-Quad BeadChips (Illumina, San Diego, USA). The controls were genotyped using Illumina Infinium II HumanHap550K Beadchips and Illumina Infinium II HumanHap550-Duo BeadChips (Illumina, San Diego, USA), respectively. Imputation was performed using IMPUTE¹³ and phased haplotypes from HapMap phase II release 22 data. Betas of the association were obtained by performing the 'score' method in an additive model as implemented in SNPTEST,¹³ whereas p-values and standard errors were assessed by applying the score statistic where we adapted the variance by using kinship matrices to adjust for familial dependencies among sibling pairs.²⁰ For SNPs with poor QC properties, the following proxies were provided: rs1764345 for rs1629896 ($r^2=0.71$); rs3827809 for rs1876836 ($r^2=0.95$); rs2739128 for rs10956694 ($r^2=0.84$); rs7319621 for rs4544137 ($r^2=1$). There were no highly correlated SNPs to serve as proxies for rs11725992 so this SNP was removed from the dataset.

The Rotterdam study, RSI and RSII cohorts. The Rotterdam Study is a prospective population-based study on determinants of chronic disabling diseases and consists of three sub-populations.²¹ The Rotterdam Study I (RSI) is the first cohort of 7,983 persons, aged 55 years and over living in Rotterdam in the Netherlands. This cohort was extended in 1999 with 3,011 participants using the same inclusion criteria (the Rotterdam Study II (RSII)). The medical ethics committee of Erasmus University Medical School approved the study and written informed consent was obtained from each participant. For all sub-populations, subjects were scored for the presence of OA using standardized radiographs of the hip and knee. Hip OA cases were defined as a KL grade ≥ 2 (defined as at least definite JSN) or a total hip replacement (THR). Knee OA cases were defined as a KL grade ≥ 2 (defined as at least two osteophytes and possible JSN) or a total knee replacement TKR. Hip and knee replacements due to fracture were excluded. Genotyping of the samples of RSI and RSII with the HumanHap550v3 Genotyping BeadChip (Illumina, San Diego, USA) was carried out at the Genetic Laboratory of the Department of Internal Medicine of Erasmus Medical Center, Rotterdam, the Netherlands. The BeadStudio GenCall algorithm was used for genotype calling. QC procedures were as described previously.²² The following sample exclusion criteria were applied: call rate $< 97.5\%$; gender mismatches with typed X-linked markers; autosomal heterozygosity $> 0.336 \sim \text{FDR} > 0.1\%$, duplicates and/or first or second degree relatives using IBS probabilities $> 97\%$ from PLINK; ethnic outliers using IBS distances $> 3\text{SD}$ from PLINK.⁶ Analysis was restricted to SNPs with a call rate $\geq 98\%$, MAF ≥ 0.01 , and HWE $p \geq 1 \times 10^{-6}$. MACH software was used for imputation and statistical analysis was carried out using MACH2QTL and GRIMP.^{23, 24}

TwinsUK replication. The study participants were white unrelated participants randomly chosen from monozygotic and dizygotic twin pairs from the TwinsUK adult twin registry, a cohort used to study the heritability and genetics of age-related diseases.²⁵ These unselected twins were recruited from the general population through national media campaigns in the UK. In this study, cases had a KL grade ≥ 2 at the hip or the knee, whereas controls had a KL grade < 2 at the hip or the knee. Ethics approval was obtained from the Guy's and St. Thomas' Hospital Ethics Committee. Written informed consent was obtained from every participant. Samples were genotyped with the Infinium HumanHap300 assay (Illumina, San Diego, USA) at the Duke University Genotyping Center (NC USA), Helsinki University (Finland) and the Wellcome Trust Sanger Institute. The Illuminus calling algorithm was used for genotype calling.⁵ Analysis was restricted to SNPs with a call rate $> 90\%$, MAF ≥ 0.01 , and HWE $p \geq 1 \times 10^{-4}$ whilst imputation was performed using IMPUTE.¹³ At imputed loci, all genotypes with posterior probabilities < 0.9 were discarded and the imputed loci were filtered out using usual QC filters.

De novo replication samples, genotyping and analysis

The case set comprised 2,409 subjects collected from arcOGEN stage one or stage two that had not been included in the GWAS. The control set comprised 2,319 individuals from the WTCC2-58BC and WTCC2-UKBS cohorts that were not part of the discovery GWAS. Genotyping was performed using the Sequenom MassARRAY iPLEX Gold assay. For the iPLEX Gold, assays for all SNPs were designed using the eXTEND suite and MassARRAY Assay Design software version 3.1 (Sequenom). Samples were amplified in multiplexed PCR reactions before allele-specific extension. Allelic discrimination was obtained by analysis with a MassARRAY Analyzer Compact mass spectrometer. Genotypes were automatically assigned and manually confirmed using MassARRAY Typer Analyzer software version 4.0 (Sequenom). Gender markers were included in iPLEX assays as a QC metric for confirmation of plate/sample identity. The mean assay call rate across both the case and control datasets was 98.9%, excluding samples with a call rate of $< 80\%$ (97.2% including all samples). For two SNPs that did not pass assay design the following proxies were used: rs13283416 for rs4836732 ($r^2=0.86$); rs1470185 for rs719535 ($r^2=0.96$).

Replication data SNP QC

SNPs were further excluded from the *in silico* replication datasets if their call rate was $< 95\%$, HWE $p < 10^{-4}$ and imputation info score < 0.5 . SNPs were removed from the *de novo* replication dataset if their call rate was $< 80\%$, HWE $p < 10^{-4}$ and if the gender in the manifest was discordant with the gender in the Sequenom iPLEX assay. The number of SNPs passing our additional QC criteria from each dataset is shown in Table 1.

Effective sample size

With the exception of deCODE the effective sample size was calculated by multiplying $N_{\text{eff_cases}}$ by 2 using the following formula $N_{\text{eff_case}} = 2 * N_{\text{case}} * N_{\text{ctrl}} / (N_{\text{case}} + N_{\text{ctrl}})$ where N_{case} is the number of cases and N_{ctrl} is the number of controls.

Meta-analysis

We used a meta-analysis framework to combine results across replication studies and across all data. Combined estimates of ORs for reference alleles were obtained by weighting the logORs of each study by the inverse of their variance using a fixed effects model. We investigated evidence of heterogeneity of ORs using the Cochran's Q and I^2 statistics.^{26, 27} We additionally assessed the combined results using a random effects meta-analysis model. Regional plots of association were plotted using LocusZoom.²⁸

BMI adjustment

We investigated attenuation of the 129 prioritized signals and the established obesity loci by carrying out a logistic regression analysis in the OA v. TwinsUK control dataset using BMI as a covariate. The logistic regression analysis for each SNP was performed for the stratum where the lowest p-value had been observed in the discovery set (ST4 and ST11).

eQTL analysis

We performed an eQTL analysis of index variants at association signals using the GTEx (Genotype-Tissue Expression),²⁹ SCAN³⁰ and Pritchard laboratories (<http://eqtl.uchicago.edu/cgi-bin/gbrowse/eqtl/>)^{29, 31-36} browsers. *In silico* eQTL analysis suggested that six out of the nine replicating SNPs are likely to be trans expression quantitative trait loci (trans eQTLs), as they regulate the expression of distantly located genes.

Gene expression bioinformatics

Gene expression data were assessed through the NCBI GEO (Gene expression Omnibus) database (<http://www.ncbi.nlm.nih.gov/geo/>).

ENCODE regulatory regions

We investigated the presence of enhancer and promoter-associated histone modifications, hypersensitive sites, transcription levels, and transcription factors in the chromosomal region encompassing the association signals by accessing the ENCODE project data (http://www.ensembl.org/Homo_sapiens/encode.html). The rs6976 and rs11177 eQTLs are likely to occur within DNase hypersensitive regions, and rs6976 occurs in a putative transcription factor binding region for C_Fos. The SNP rs10492367 is enriched for DNase hypersensitive sites and contains a binding site for the SP1 transcription factor.

RT-PCR gene expression assay

Joint tissues were obtained from individuals undergoing elective joint replacement of the hip or of the knee for OA. The tissues collected were cartilage, tendon, ligament, meniscus, synovium, fat pad, and osteophyte. We also collected cartilage from patients who did not have clinical OA but who instead had undergone a neck-of-femur (NOF) fracture. The local ethics committee granted ethical approval and informed consent was obtained from each donor. RNA was extracted from the tissues as described previously³⁷ and 1µg of total RNA was then used for each cDNA synthesis using random hexamers, the reverse transcriptase enzyme SuperScript II and the SuperScript kit (Invitrogen). This produced 20µl of cDNA, 0.5µl of which was then used in a 15µl PCR containing 7.5µM each of the gene specific forward and reverse primers (ST13). The PCR products were electrophoresed through 3% w/v agarose gels and stained using ethidium bromide.

Immunohistochemistry and Western Blotting

Cartilage samples were obtained from patients undergoing total knee replacement for osteoarthritis or from normal knee joints obtained from subjects undergoing above knee amputation. Informed consent was obtained in all cases and the studies were approved by NHS Lothian ethical review committee. Immunohistochemical staining for guanine nucleotide-binding protein-like 3 (or nucleostemin) was performed on a touch preparation of fresh human articular cartilage and on cytopins prepared from cultured human articular chondrocytes using a polyclonal goat primary antibody to nucleostemin (R&D Systems®) and a HRP-linked rabbit anti-goat secondary antibody (Dako). The signal was visualised using the DakoEnVision®+ System-HRP (DAB) followed by counterstaining with hematoxylin. Western blots were performed according to standard techniques on protein extracts prepared from short term primary cultures of human articular chondrocytes and probed using a primary monoclonal mouse anti-nucleostemin antibody (Millipore) with detection by a peroxidase-conjugated AffiniPure Donkey Anti-Mouse IgG (Jackson ImmunoResearch Laboratories, Inc.). This antibody detects a major band corresponding to nucleostemin at approximately 62kD. Protein extracts prepared from the JJ012 chondrosarcoma cell line was used as a positive control following treatment with chondroitinase. Beta-actin was used to control for loading in the Western blots. This was detected using a primary rabbit anti-beta-actin antibody (Sigma) and the same secondary antibody as was used for detection of nucleostemin. The signal was quantitated with GeneTool v.3.06.04 software (Syngene) and levels of nucleostemin were corrected for beta-actin.

SUPPLEMENTARY RESULTS

Population stratification

To eliminate possible false positives arising from hidden population structure we first examined multidimensional scaling plots of our samples and the HapMap II populations and removed individuals with non-European ancestry (Supplementary Methods). Large allelic differences between wide-spread geographical locations have been previously observed in the British population for 13 loci,³⁸ none of which overlap with the eight loci reported in this study. To ensure these associations did not arise due to population stratification, we repeated the analysis with adjustment for the first ten principal components (PCs) (Supplementary Methods). The inflation factor λ_{1000} values were very similar in the unadjusted compared with the adjusted analyses (ST2) and remained below 1.015 for all strata. The p-values before and after adjustment were also similar for all signals (ST3) therefore we performed all downstream analyses using unadjusted results.

Imputation results

We imputed genotypes at 907,069 HapMap III SNPs based on 485,077 autosomal directly typed SNPs with MAF >0.01. We included only SNPs that passed post-imputation QC filters in downstream analyses. We inspected qq plots prior to and following the removal of SNPs in the vicinity of the 129 SNPs that were taken forward to replication to examine whether any additional signals have arisen following imputation. Following the removal of the 129 prioritized signals from the imputed data analysis no SNPs with $p < 10^{-5}$ were identified in any of the strata analyzed across the genome (SF3). We then examined the regional plots of association of the eight novel loci to see whether imputation had given rise to any additional

signals. We find that for all of the eight replicating loci imputation has not resulted in the identification of any additional SNPs with substantially lower (i.e. more than an order of magnitude) p-values than the directly typed SNPs that were followed up (SF4).

Effect of phenotype definition in controls

We performed discovery set analyses using two types of controls: population-based controls unselected for OA (n=11,009) and OA-free controls (females only, n=1,828) from TwinsUK. We examined the strength of association of the three previously established OA loci (ST8) and of the novel signals (Table 2) in the two types of analysis. Here, we focus on the stratum where the most significant meta-analysis p-value was observed. Evidence for association with the *GDF5* proxy SNP rs4911494 was stronger when female knee OA cases (n=2,135) were analyzed using OA-free controls (OR 1.10[1.01-1.21], p=0.039) compared to knee OA cases (n=3,498) analyzed using population-based controls (OR 1.01[0.96-1.07], p=0.726) despite the smaller sample size (ST8). Evidence for association with rs3815148 at the chr7 locus was stronger when knee OA cases were analyzed using population-based controls (OR 1.12[1.05-1.19], p=0.0006) compared to the female knee OA cases v. disease-free controls analysis (OR 1.19[1.07-1.32], p=0.002) (ST8). Similarly, evidence for association with the *MCF2L* SNP rs11842874 was stronger when all OA cases were analyzed using population-based controls (OR 1.13[1.05-1.23], p=0.002) compared to the female analysis employing disease-free controls (OR 1.11[0.95-1.29], p=0.186) (ST8). Evidence for association at the novel loci was similar or stronger in the discovery set analyses employing population-based controls (ST9).

We additionally repeated meta-analyses across the discovery and replication sets using OA-free controls only, for the eight replicating loci in Table 2. We find that, for all eight loci, association was more significant in the analysis employing population-based controls. Effect size estimates for the replicating loci are the same or smaller for all variants except one (*CHST11*) in the meta-analyses employing disease-free controls (ST10).

Previously established OA loci

The genome wide significant locus previously reported on chr7 was associated with knee OA (OR 1.12[1.05-1.19], p=0.0006 for knee OA v. population-based controls) (ST8). The OA-associated SNP rs143383 in *GDF5* was not genotyped in our study. Proxy SNP rs4911494 ($r^2=0.93$ with rs143383 based on HapMap CEU) showed weak evidence for association in the analysis employing disease-free controls (OR 1.10[1.01-1.21], p=0.039 for female knee OA v. disease-free controls) but no evidence for association in the analysis employing population-based controls (OR 1.01[0.96-1.07], p=0.7256 for knee OA). Association between rs143383 and hip OA was first described in two independent Japanese populations (OR 1.79[1.53-2.09], p=2x10⁻¹³).³⁹ Evidence for association with knee OA was weaker in datasets from China (p=3x10⁻⁴) and Japan (p=0.002). A large-scale meta-analysis employing 4,791 hip OA cases and 6,006 controls, and 4,367 knee OA cases and 6,291 controls⁴⁰ showed that in samples of European descent there was less compelling evidence for association with hip OA (1.07[1.01-1.14], p=0.034) and more compelling evidence for association with knee OA (1.13[1.06-1.20], p=9x10⁻⁵) but with a much weaker effect size than the East Asian set) (ST8). These differences can be ascribed to allele frequency disparities between ethnic groups. The *GDF5* SNP was eventually found to be genome-wide significantly associated with knee OA in Europeans in a subsequent meta-analysis across a total of 6,861 knee OA cases and 10,103 controls.⁴¹ The arcOGEN knee OA discovery set comprises a much smaller number of knee only OA cases (n=3,498), and this could explain the weaker association observed at the *GDF5* proxy SNP. The rs11842874 variant in the *MCF2L* locus was associated with all OA in the full arcOGEN GWAS (OR 1.13[1.05-1.23], p=2.5x10⁻³) (ST8).

SF1. QQ plots.

Plots A to R show qq plots for arcOGEN cases v. population-based controls. Plots S to X show qq plots for arcOGEN cases v. OA-free controls. The expected p-value is indicated by the red line and the associated 95% confidence intervals are indicated by the black lines either side. The black dots indicate the observed p-values. Because of the large sample size the λ_{1000} is given.^{8,9} A. All OA, $\lambda_{1000}=1.0091$; B. Male OA, $\lambda_{1000}=1.0097$; C. Female OA, $\lambda_{1000}=1.0119$; D. TJR, $\lambda_{1000}=1.01$; E. Male TJR, $\lambda_{1000}=1.0089$; F. Female TJR, $\lambda_{1000}=1.012$; G. Hip OA, $\lambda_{1000}=1.011$; H. Male hip OA, $\lambda_{1000}=1.0117$; I. Female hip OA, $\lambda_{1000}=1.0129$; J. THR, $\lambda_{1000}=1.0118$; K. Male THR, $\lambda_{1000}=1.0121$; L. Female THR, $\lambda_{1000}=1.0131$; M. Knee OA, $\lambda_{1000}=1.013$; N. Male knee OA, $\lambda_{1000}=1.0142$; O. Female knee OA, $\lambda_{1000}=1.0127$; P. TKR, $\lambda_{1000}=1.0133$; Q. Male TKR, $\lambda_{1000}=1.0155$; R. Female TKR, $\lambda_{1000}=1.0145$; S. Female OA, $\lambda_{1000}=1.0096$; T. Female TJR, $\lambda_{1000}=1.0122$; U. Female hip OA, $\lambda_{1000}=1.0149$; V. Female THR, $\lambda_{1000}=1.015$; W. Female knee, $\lambda_{1000}=1.0076$; X. Female TKR, $\lambda_{1000}=1.0164$.

SF2. Manhattan plots.

Plots A to R show Manhattan plots for arcOGEN cases v. population-based controls. Plots S to X show Manhattan plots for arcOGEN cases v. OA-free controls. The red horizontal line indicates genome-wide significance. A. All OA; B. Male OA; C. Female OA; D. TJR; E. Male TJR; F. Female TJR; G. Hip OA; H. Male hip OA; I. Female hip OA; J. THR; K. Male THR; L. Female THR; M. Knee OA; N. Male knee OA; O. Female knee OA; P. TKR; Q. Male TKR; R. Female TKR; S. Female OA; T. Female TJR; U. Female hip OA; V. Female THR; W. Female knee; X. Female TKR.

SF3. QQ plots for directly typed and imputed data.

QQ plots of the all OA v. population-based controls analysis comprising directly typed and HapMap III imputed SNPs. The expected p-value is indicated by the red line and the associated 95% confidence intervals are indicated by the black lines either side. A. All SNPs following post imputation QC filters; B. All SNPs following post imputation QC filters excluding

SNPs in the vicinity of the 129 prioritized SNPs. The same pattern i.e. deflation of the test statistic following removal of the 129 prioritized loci was observed for all stratified analyses that comprise imputed and directly-typed SNPs.

SF4. Regional plots of the replicating signals in analyses comprising directly genotyped and imputed SNPs. Case-control association results ($-\log_{10}$ p-value) for imputed and genotyped SNPs in the discovery set are plotted against genomic position (NCBI build 36) for the stratum where the most significant meta-analysis p-value was observed. The index genotyped SNP is denoted by a purple diamond. The circles indicate association results of imputed SNPs and the squares indicate association results of genotyped SNPs in the region; the color reflects the correlation coefficient (r^2) of each genotyped SNP with the index SNP estimated using the CEU HapMap II panel. Estimated recombination rates (in centimorgans per megabase) are plotted in light blue. The region shown in the plots extends to either 500kb upstream and downstream of the index SNP or until the next recombination hotspot if this lies further than 500kb away. a) chr3 signal centered around rs6976 in TJR; b) chr9 signal centered around rs4836732 in female THR; c) chr6 signal centered around rs9350591 in hip OA; d) chr12 signal centered around rs10492367 in hip OA; e) chr12 signal centered around rs835487 in THR; f) chr3 signal centered around rs12107036 in female TKR; g) chr16 signal centered around rs8044769 in females; h) chr6 signal centered around rs10948172 in males.

SF5. a) Immunohistochemical staining for guanine nucleotide-binding protein-like 3 (nucleostemin) in cytopins of cultured human chondrocytes, showing nucleolar localization of the protein. A higher power view is shown in the inset. b) Immunohistochemical staining for nucleostemin in fresh frozen cartilage showing intranuclear staining (arrows). c) Detection of nucleostemin and beta-actin in cultured chondrocytes from control subjects (n=5) and patients with osteoarthritis of the knee (n=5) by Western blotting and in JJ012 chondrosarcoma cells which were used as a positive control (+ve). d) Quantitation of nucleostemin expression corrected for beta-actin using the data derived from panel c. The p-value shown refers to the difference between the OA and control groups. Values in the graphs are mean + sem. MW = Molecular weight kilodaltons (KDa), NS – nucleostemin; actin – beta actin

SF6. Regional association plots of replicating signals approaching genome-wide significance. Case-control association results ($-\log_{10}$ p-value) for genotyped SNPs in the discovery set are plotted against genomic position (NCBI build 36) for the stratum where the most significant meta-analysis p-value was observed. The index SNP is denoted by a purple diamond in the discovery set and by a purple square in the final meta-analysis. The circles indicate association results of genotyped SNPs in the region; the colour reflects the correlation coefficient (r^2) of each genotyped SNP with the index SNP estimated using the CEU HapMap II panel. Estimated recombination rates (in centimorgans per megabase) are plotted in light blue. The region shown in the plots extends to either 500kb upstream and downstream of the index SNP or until the next recombination hotspot if this lies further than 500kb away. a) chr3 signal centered around rs12107036 in female TKR; b) chr16 signal centered around rs8044769 in females; c) chr6 signal centered around rs10948172 in males.

ST1. Population-based controls used in the arcOGEN GWAS.

ST2. λ_{1000} values for the unadjusted analyses and the analyses adjusted for the first ten PCs.

ST3. Association p-values before and after adjustment for the first ten principal components.

ST4. Association summary statistics and genotype counts for the SNPs prioritized for replication for the stratum where the lowest p-value was observed.

ST5. Association summary statistics for the 129 prioritized SNPs in arcOGEN, replication cohorts and meta-analyses across all strata.

ST6. Association summary statistics for the eight replicating signals in the analyses employing all OA cases.

ST7. Association summary statistics for the replicating signals in analyses employing all cases or TJR-only cases.

ST8. Allele frequencies and association summary statistics for previously established OA loci.

ST9. Association summary statistics for the replicating signals in analyses employing population-based controls or OA-free controls in the discovery set.

ST10. Association summary statistics for the replicating signals in meta-analyses including all controls or OA-free controls only.

ST 11. Association p-values before and after BMI adjustment for the 129 prioritized SNPs.

ST 12. Expression in joint tissues, assessed by reverse-transcription PCR. NOF cartilage was collected from patients undergoing surgery for a neck-of-femur fracture.

ACKNOWLEDGEMENTS AND FUNDING

arcOGEN (<http://www.arcogen.org.uk/>) was funded by a special purpose grant from Arthritis Research UK (grant 18030). This study used genotype data from population controls that was generated by the Wellcome Trust Case Control Consortium 2 (<http://www.wtccc.org.uk>), funded by The Wellcome Trust (grant 083948). The population controls were from the 1958 British Birth Cohort collection funded by the Medical Research Council (grant G0000934) and The Wellcome Trust (grant 068545) and from the UK Blood Services Collection of Common Controls funded by The Wellcome Trust. This research utilizes resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. The UK samples used in arcOGEN derive from nine centres; Nottingham, London, Oxford, Sheffield, Southampton, Newcastle, Wansbeck, Worcester and Edinburgh. For Nottingham we acknowledge Arthritis Research UK for funding the collection of the majority of extant cases (stage 1). For London we thank the staff from the TwinsUK unit and the Chingford Study for stage 1 patient ascertainment, we acknowledge financial support from Arthritis Research UK, from the Wellcome Trust, from the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre (BRC) award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London, and from the European Commission framework 7 programme TREAT-OA (grant 200800). For Oxford we acknowledge funding support from the Collisson Foundation, the Botnar Foundation and the Jean Shanks Foundation for stage 1 patient ascertainment and we acknowledge the NIHR for supporting the Biomedical Research Unit (BRU) at the University of Oxford. For Sheffield we acknowledge the NIHR for supporting the Sheffield Musculoskeletal BRU collection site, the South Yorkshire Clinical Research Network for part-funding the Sheffield research nurses and for clerical support, the Royal College of Surgeons of England and the Cavendish Hip Foundation. For Southampton we acknowledge the Wellcome Trust Clinical Research Facility at Southampton General Hospital and we thank Anna Bara, trial manager for the Arthritis Research UK-funded VIDEO study. For Edinburgh we acknowledge the contribution of Gayle Scott and Helen Watters for recruitment and ascertainment of stage 2 cases. For Newcastle we acknowledge the Freeman Hospital arthroplasty surgeons Nigel Brewster, David Deehan, Craig Gerrand, Andy Gray, Munawar Hashmi, James Holland and David Weir for assistance in stage 2 patient ascertainment, we thank Karen Bettinson for her assistance, we acknowledge the support of the UK NIHR BRC for Ageing and Age-related disease award to the Newcastle upon Tyne Hospitals NHS Foundation Trust and the support of the Northumberland, Tyne and Wear CLRN. For Wansbeck we acknowledge the Northumbria arthroplasty consultants Mr Partington, Mr Muller, Mr Emmerson, Mr Prasad, Mr Kramer, Mr Jones, Mr Murty, Mr Asaad, Mr Leitch, Mr Pratt, Mr Townshend and Miss Van Kampen for accrual, and the research nurse Donna Gray for patient ascertainment. For Worcester we acknowledge the support of the West Midlands (south) CLRN. We acknowledge stage 1 and 2 sample management undertaken by the UK DNA Banking Network funded by the Medical Research Council at CIGMR - the Centre for Integrated Genomic Medical Research, University of Manchester. Genotyping was performed at the Wellcome Trust Sanger Institute and we thank Emma Gray, Sarah Edkins and Cordelia Langford for their assistance. We thank Vanessa Cox of the MRC Epidemiology Resource Centre, Southampton, for database and computing services for the arcOGEN stage 2 patient questionnaires. The Leiden University Medical Centre, the Dutch Arthritis Association and Pfizer Inc., Groton, CT, USA support the GARP study, whilst genotypic work was supported by the Netherlands Organization of Scientific Research (MW 904-61-095, 911-03-016, 917 66344 and 911-03-012), Leiden University Medical Centre, and by the "Centre of Medical System Biology" and the "Netherlands Consortium of Healthy Aging" in the framework of the Netherlands Genomics Initiative (NGI). The generation and management of GWAS genotype data for the Rotterdam Study is supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters for their help in creating the GWAS database, and Karol Estrada and Maksim V. Struchalin for their support in creation and analysis of imputed data. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. We would like to thank Karol Estrada, Dr. Fernando Rivadeneira, Dr. Tobias A. Knoch and Anis Abuseiris (Erasmus MC Rotterdam, The Netherlands), for their help in creating GRIMP, and BigGRID, MediGRID, and Services@MediGRID/D-Grid, (funded by the German Bundesministerium fuer Forschung und Technology; grants 01 AK 803 A-H, 01 IG 07015 G) for access to their grid computing resources. EGCUT received financing by FP7 grants (201413, 245536), grant from Estonian Government SF0180142s08, from the EU through the European Regional Development Fund, OPENGENE, in the frame of Centre of Excellence in Genomics and Univ. of Tartu grant SP1GVARENG. We acknowledge EGCUT and Estonian Biocentre personnel, especially Mr V. Soo. The deCODE Genetics OA project is funded in part by the European Commission framework 7 programme TREAT-OA (grant 200800). AGD-W, LS, KP and EZ are supported by the Wellcome Trust (grant 098051), KC is supported by a Botnar Fellowship, KC and NWR are supported by the Wellcome Trust (WT079557MA), JMW is supported by the Higher Education Funding Council for England and MCL is supported by Pfizer. VB is supported by UnityThrough Knowledge

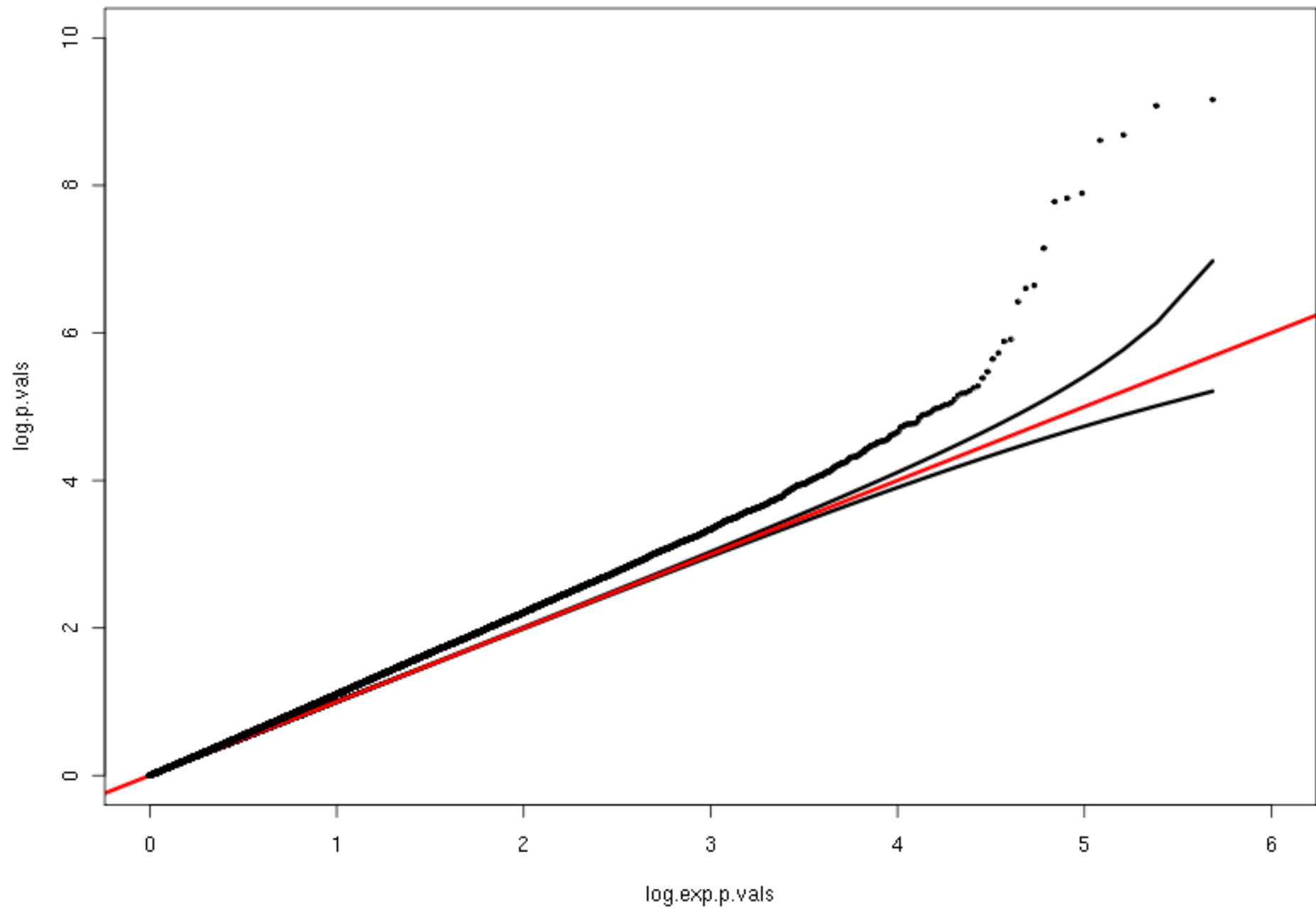
SUPPLEMENTARY REFERENCES

1. Panoutsopoulou K, Southam L, Elliott KS, Wrayner N, Zhai G, Beazley C, et al. Insights into the genetic architecture of osteoarthritis from stage 1 of the arcOGEN study. *Ann Rheum Dis*. 2011; **70**(5): 864-7.
2. TwinsUK. [cited 2010 13th August]; Available from: www.twinsUK.ac.uk
3. Andrew T, Hart DJ, Snieder H, de Lange M, Spector TD, MacGregor AJ. Are twins and singletons comparable? A study of disease-related and lifestyle characteristics in adult women. *Twin Res*. 2001; **4**(6): 464-77.
4. Spencer CC, Su Z, Donnelly P, Marchini J. Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. *PLoS Genet*. 2009; **5**(5): e1000477.
5. Teo YY, Inouye M, Small KS, Gwilliam R, Deloukas P, Kwiatkowski DP, et al. A genotype calling algorithm for the Illumina BeadArray platform. *Bioinformatics*. 2007; **23**(20): 2741-6.
6. 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; **81**(3): 559-75.
7. Pettersson F, Morris AP, Barnes MR, Cardon LR. Goldsurfer2 (Gs2): a comprehensive tool for the analysis and visualization of genome wide association studies. *BMC Bioinformatics*. 2008; **9**: 138.
8. de Bakker PI, Ferreira MA, Jia X, Neale BM, Raychaudhuri S, Voight BF. Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Hum Mol Genet*. 2008; **17**(R2): R122-8.
9. Freedman ML, Reich D, Penney KL, McDonald GJ, Mignault AA, Patterson N, et al. Assessing the impact of population stratification on genetic association studies. *Nat Genet*. 2004; **36**(4): 388-93.
10. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006; **38**(8): 904-9.
11. International HapMap Consortium, Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, et al. Integrating common and rare genetic variation in diverse human populations. *Nature*. 2010; **467**(7311): 52-8.
12. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet*. 2009; **5**(6): e1000529.
13. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet*. 2007; **39**(7): 906-13.
14. Stefansson SE, Jonsson H, Ingvarsson T, Manolescu I, Jonsson HH, Olafsdottir G, et al. Genomewide scan for hand osteoarthritis: a novel mutation in matrilin-3. *Am J Hum Genet*. 2003; **72**(6): 1448-59.
15. EGCUT. [cited 2010 13 August]; Available from: www.geenivaramu.ee
16. Metspalu A. The Estonian Genome Project. *Drug Dev Res*. 2004; **62**(2): 97-101.
17. Nelis M, Esko T, Magi R, Zimprich F, Zimprich A, Toncheva D, et al. Genetic structure of Europeans: a view from the North-East. *PLoS One*. 2009; **4**(5): e5472.
18. International Classification of Diseases. [cited 2010 13th August]; Available from: www.who.int/classifications/icd/en/
19. Riyazi N, Meulenbelt I, Kroon HM, Runday KH, Hellio le Graverand MP, Rosendaal FR, et al. Evidence for familial aggregation of hand, hip, and spine but not knee osteoarthritis in siblings with multiple joint involvement: the GARP study. *Ann Rheum Dis*. 2005; **64**(3): 438-43.
20. Uh HW, Houwing-Duistermaat JJ, Putter H, van Houwelingen HC. Assessment of global phase uncertainty in case-control studies. *BMC Genet*. 2009; **10**: 54.
21. Hofman A, Breteler MM, van Duijn CM, Janssen HL, Krestin GP, Kuipers EJ, et al. The Rotterdam Study: 2010 objectives and design update. *Eur J Epidemiol*. 2009; **24**(9): 553-72.
22. Richards JB, Rivadeneira F, Inouye M, Pastinen TM, Soranzo N, Wilson SG, et al. Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. *Lancet*. 2008; **371**(9623): 1505-12.
23. Li Y, Willer C, Sanna S, Abecasis G. Genotype imputation. *Annu Rev Genomics Hum Genet*. 2009; **10**: 387-406.
24. Estrada K, Abuseiris A, Grosveld FG, Uitterlinden AG, Knoch TA, Rivadeneira F. GRIMP: a web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. *Bioinformatics*. 2009; **25**(20): 2750-2.
25. Spector TD, Williams FM. The UK Adult Twin Registry (TwinsUK). *Twin Res Hum Genet*. 2006; **9**(6): 899-906.
26. Greenland S. Invited commentary: a critical look at some popular meta-analytic methods. *Am J Epidemiol*. 1994; **140**(3): 290-6.
27. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003; **327**(7414): 557-60.
28. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*. 2010; **26**(18): 2336-7.
29. Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, Beazley C, et al. Population genomics of human gene expression. *Nat Genet*. 2007; **39**(10): 1217-24.
30. Gamazon ER, Zhang W, Konkashbaev A, Duan S, Kistner EO, Nicolae DL, et al. SCAN: SNP and copy number annotation. *Bioinformatics*. 2010; **26**(2): 259-62.
31. Montgomery SB, Sammeth M, Gutierrez-Arcelus M, Lach RP, Ingle C, Nisbett J, et al. Transcriptome genetics using second generation sequencing in a Caucasian population. *Nature*. 2010; **464**(7289): 773-7.

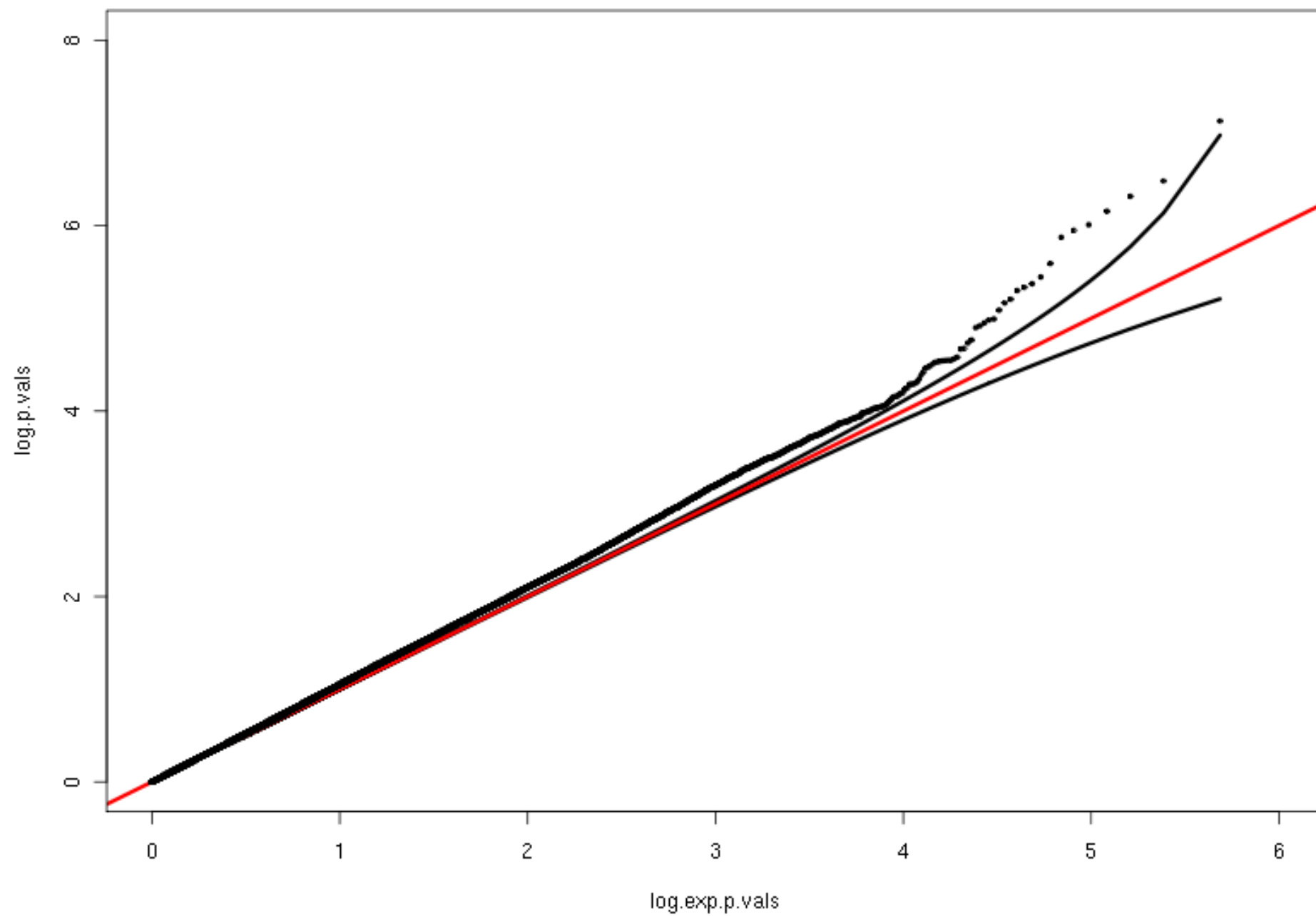
32. Myers AJ, Gibbs JR, Webster JA, Rohrer K, Zhao A, Marlowe L, et al. A survey of genetic human cortical gene expression. *Nat Genet.* 2007; **39**(12): 1494-9.
33. Pickrell JK, Marioni JC, Pai AA, Degner JF, Engelhardt BE, Nkadori E, et al. Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature.* 2010; **464**(7289): 768-72.
34. Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY, et al. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol.* 2008; **6**(5): e107.
35. Veyrieras JB, Kudaravalli S, Kim SY, Dermitzakis ET, Gilad Y, Stephens M, et al. High-resolution mapping of expression-QTLs yields insight into human gene regulation. *PLoS Genet.* 2008; **4**(10): e1000214.
36. Zeller T, Wild P, Szymczak S, Rotival M, Schillert A, Castagne R, et al. Genetics and beyond--the transcriptome of human monocytes and disease susceptibility. *PLoS One.* 2010; **5**(5): e10693.
37. Wilkins JM, Southam L, Price AJ, Mustafa Z, Carr A, Loughlin J. Extreme context specificity in differential allelic expression. *Hum Mol Genet.* 2007; **16**(5): 537-46.
38. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 2007; **447**(7145): 661-78.
39. Miyamoto Y, Mabuchi A, Shi D, Kubo T, Takatori Y, Saito S, et al. A functional polymorphism in the 5' UTR of GDF5 is associated with susceptibility to osteoarthritis. *Nat Genet.* 2007; **39**(4): 529-33.
40. Evangelou E, Valdes AM, Kerkhof HJ, Styrkarsdottir U, Zhu Y, Meulenbelt I, et al. Meta-analysis of genome-wide association studies confirms a susceptibility locus for knee osteoarthritis on chromosome 7q22. *Ann Rheum Dis.* 2011; **70**(2): 349-55.
41. Valdes AM, Evangelou E, Kerkhof HJ, Tamm A, Doherty SA, Kisand K, et al. The GDF5 rs143383 polymorphism is associated with osteoarthritis of the knee with genome-wide statistical significance. *Ann Rheum Dis.* 2011; **70**(5): 873-5.

SF1

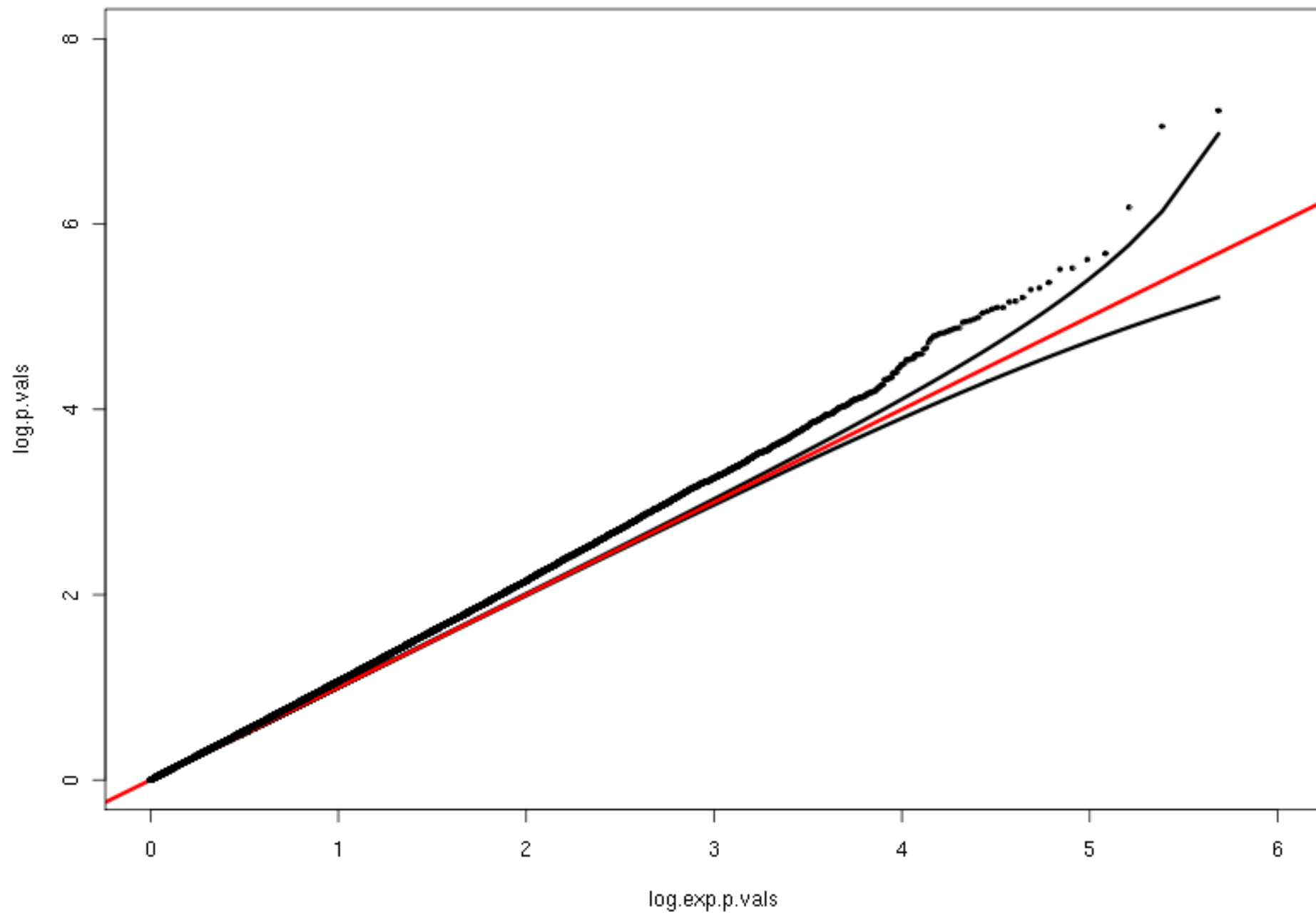
A.



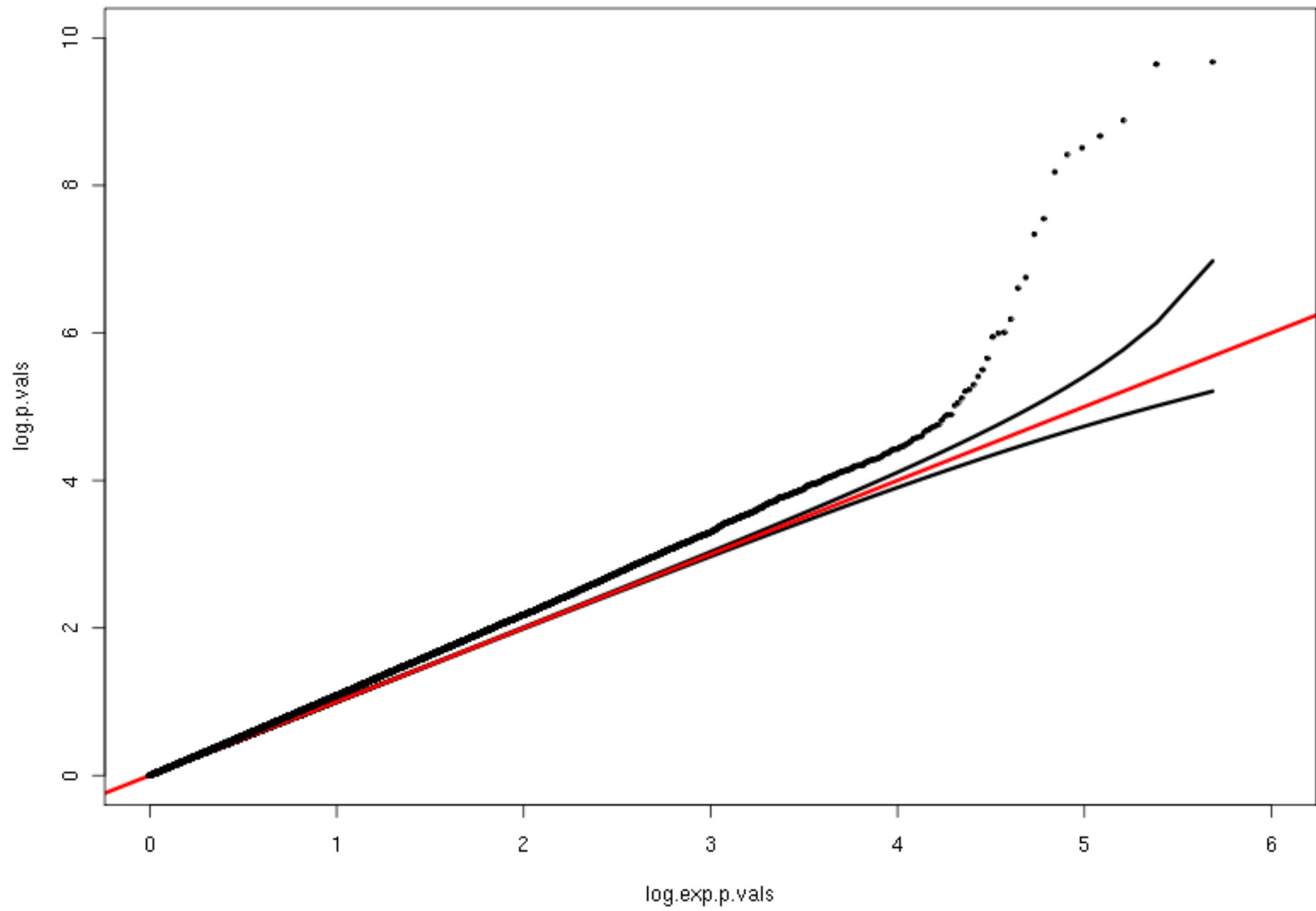
B.



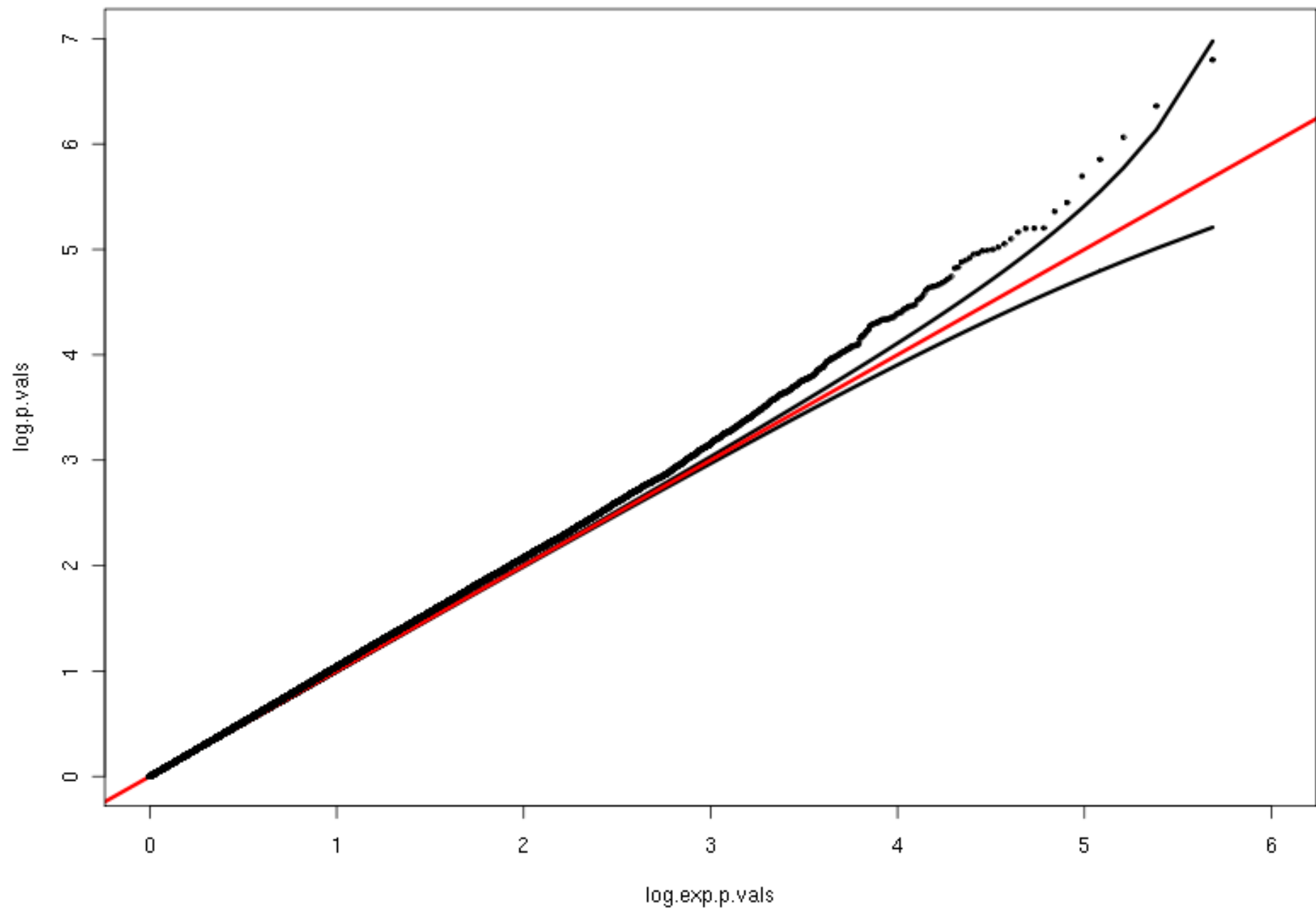
C.



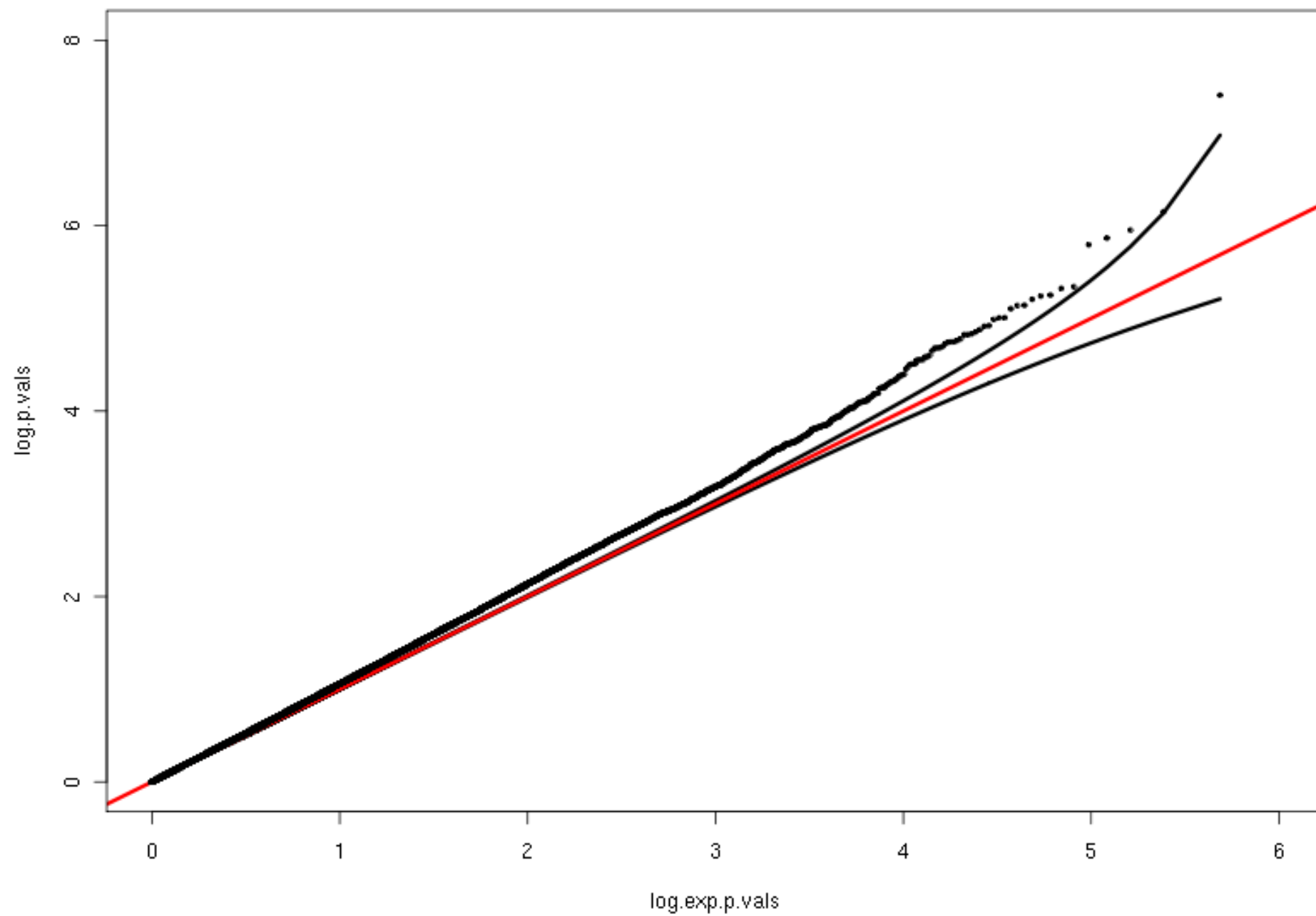
D.



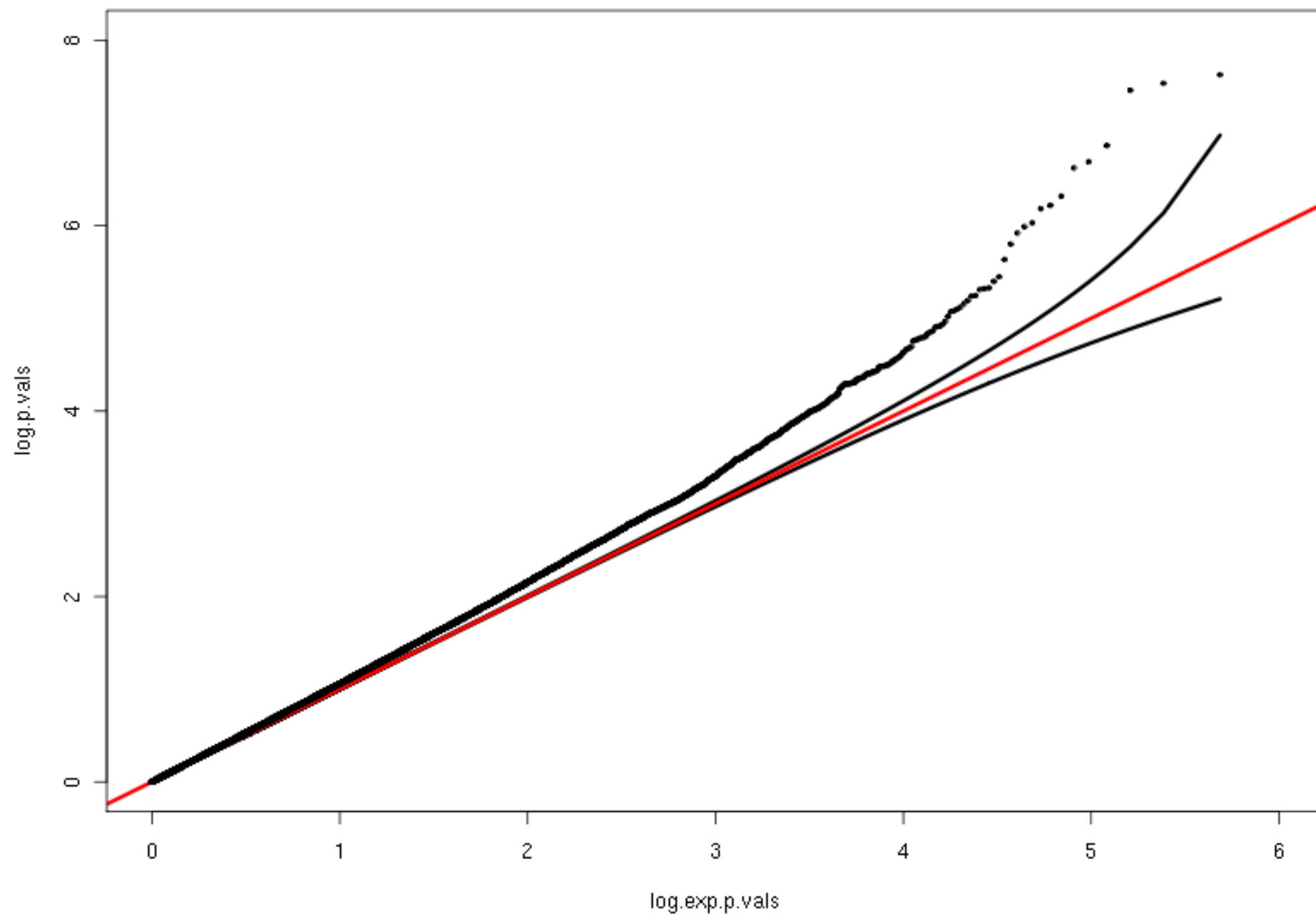
E.



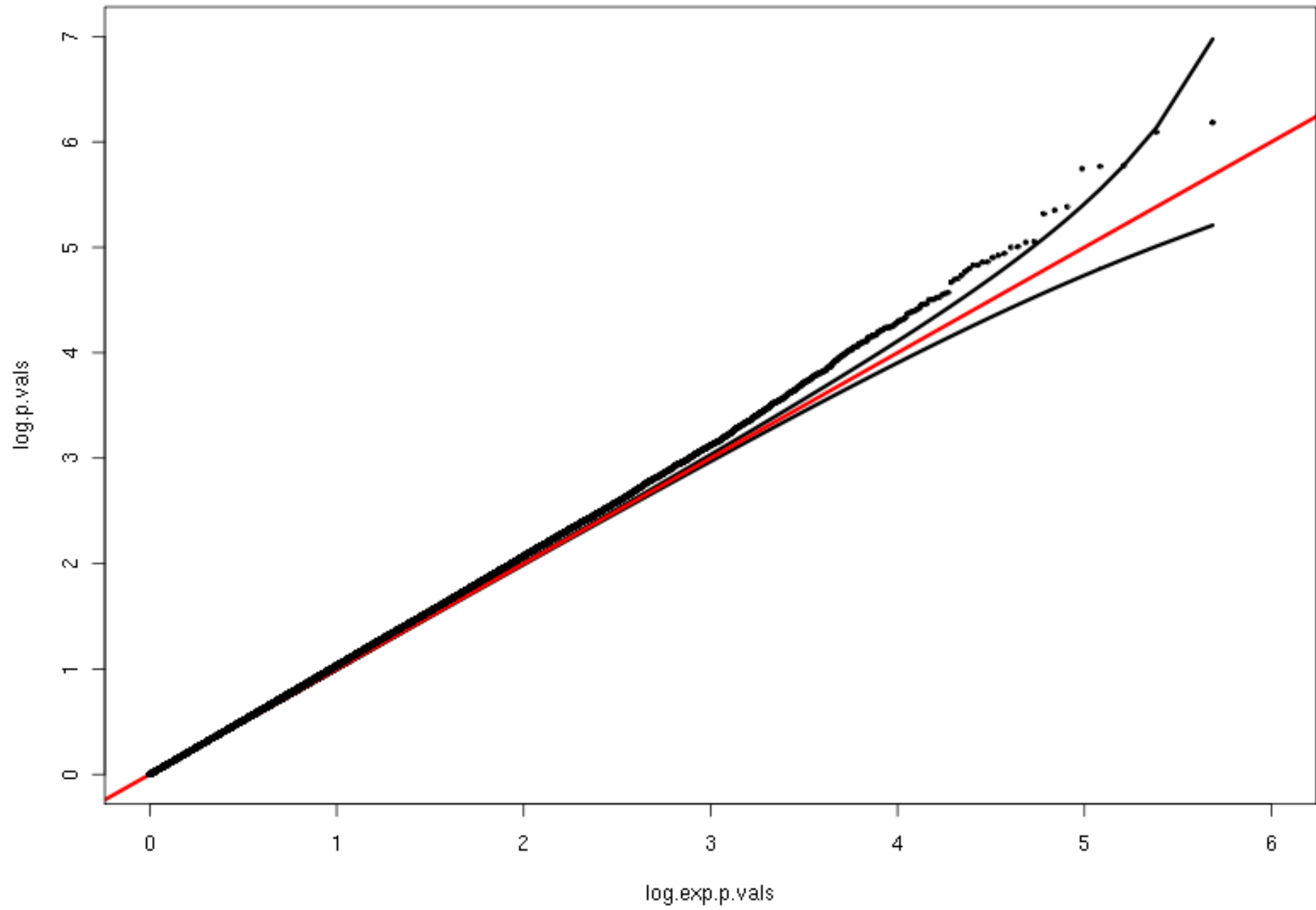
F.



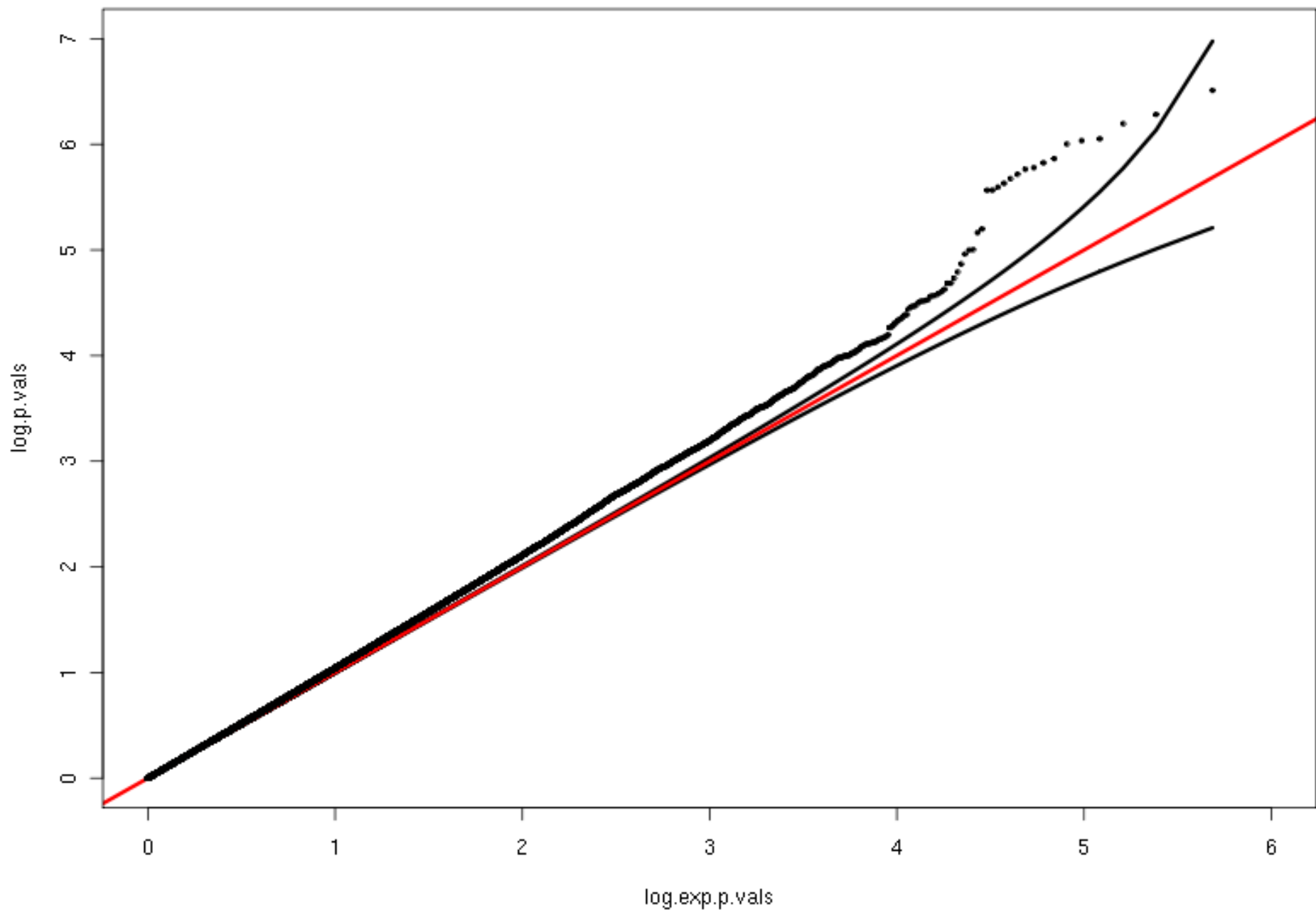
G.



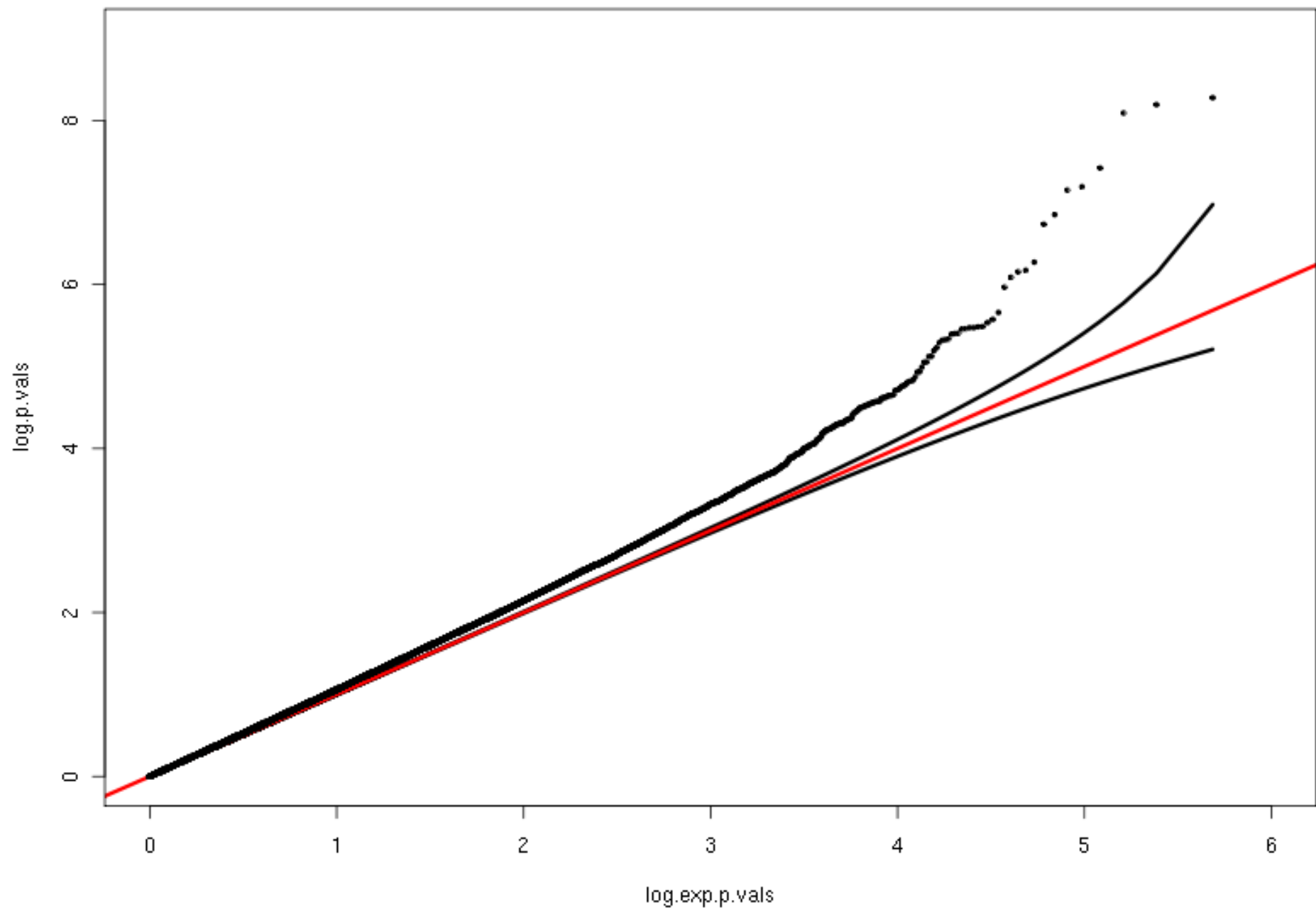
H.



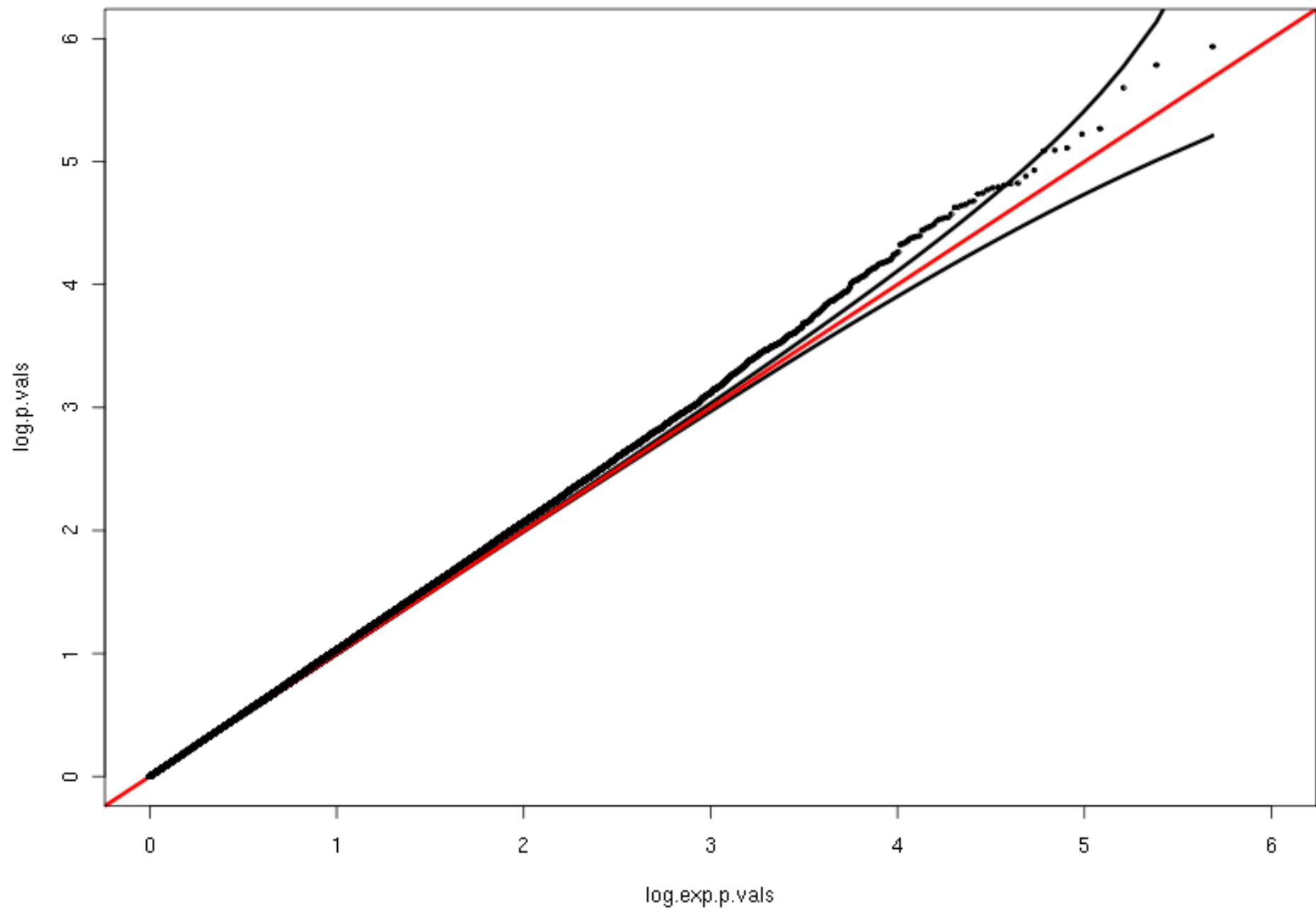
I.



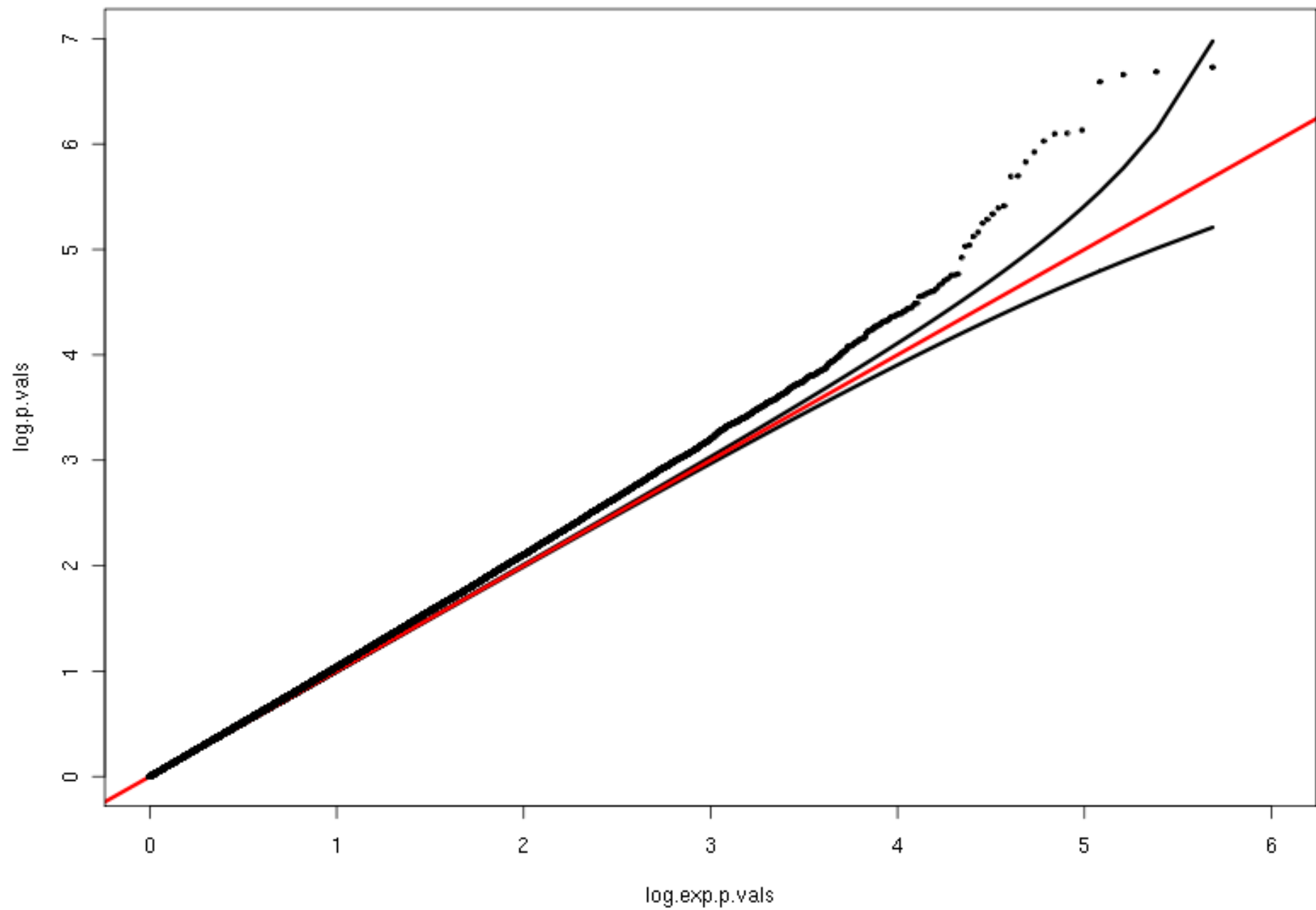
J.



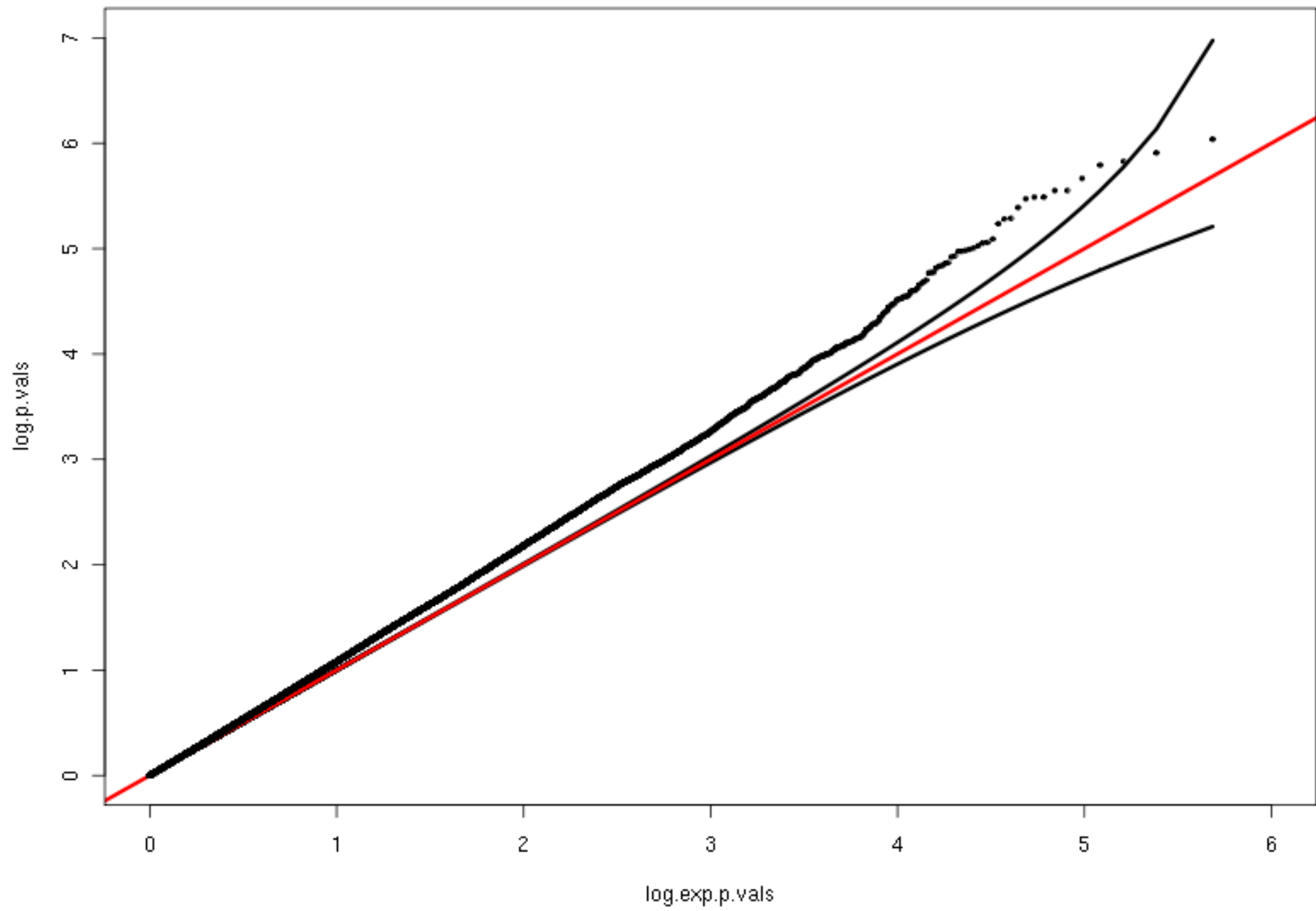
K.



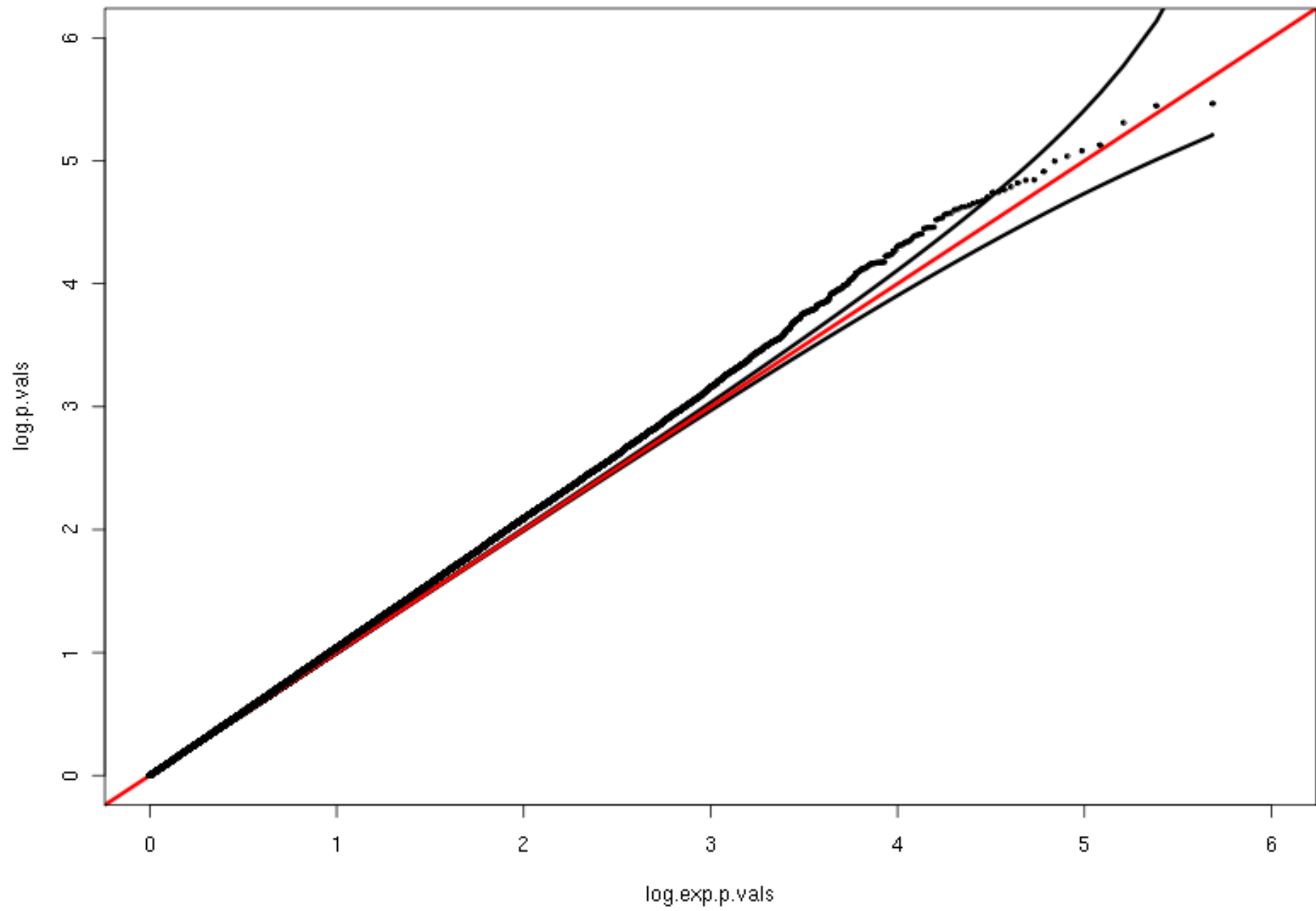
L.



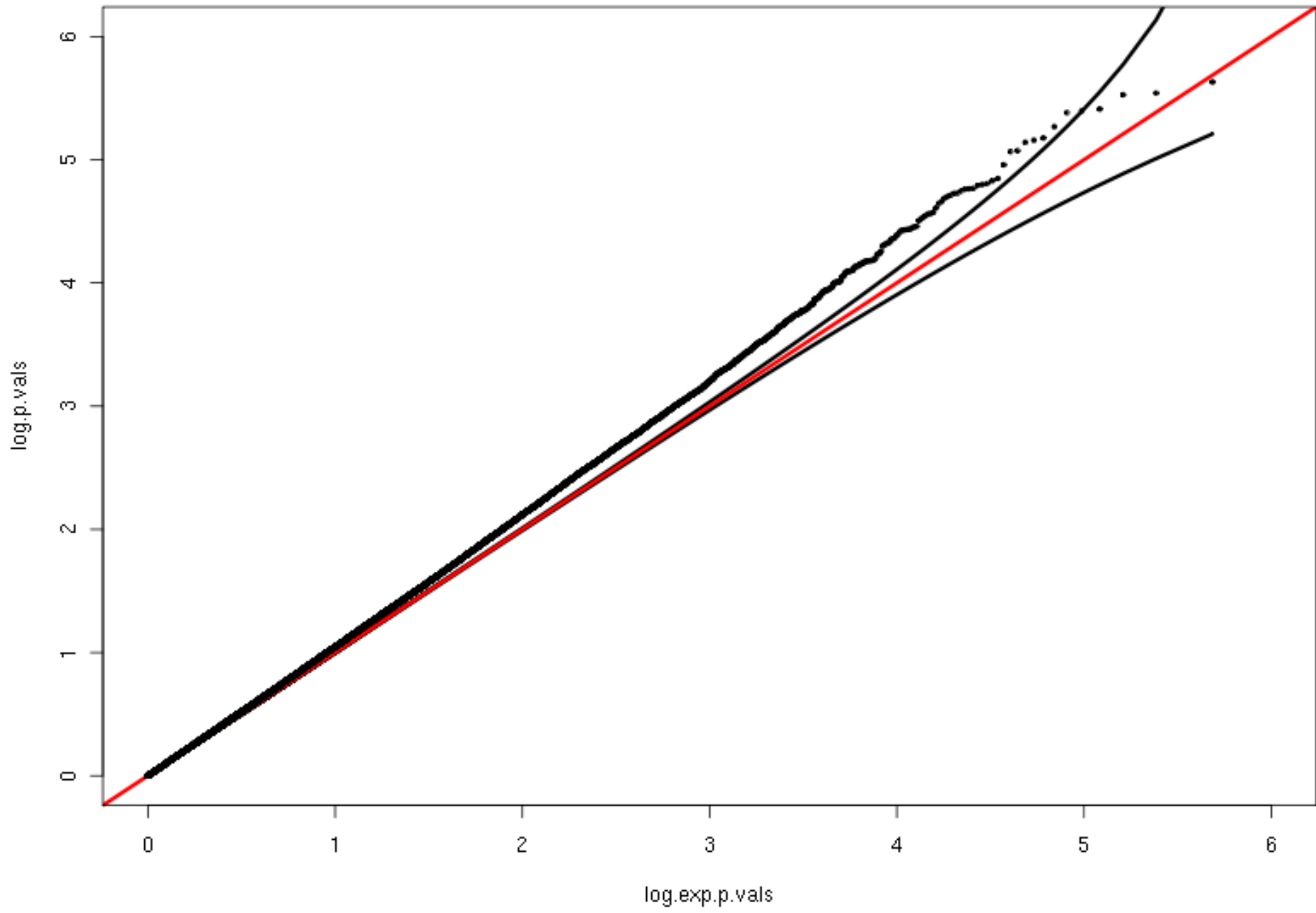
M.



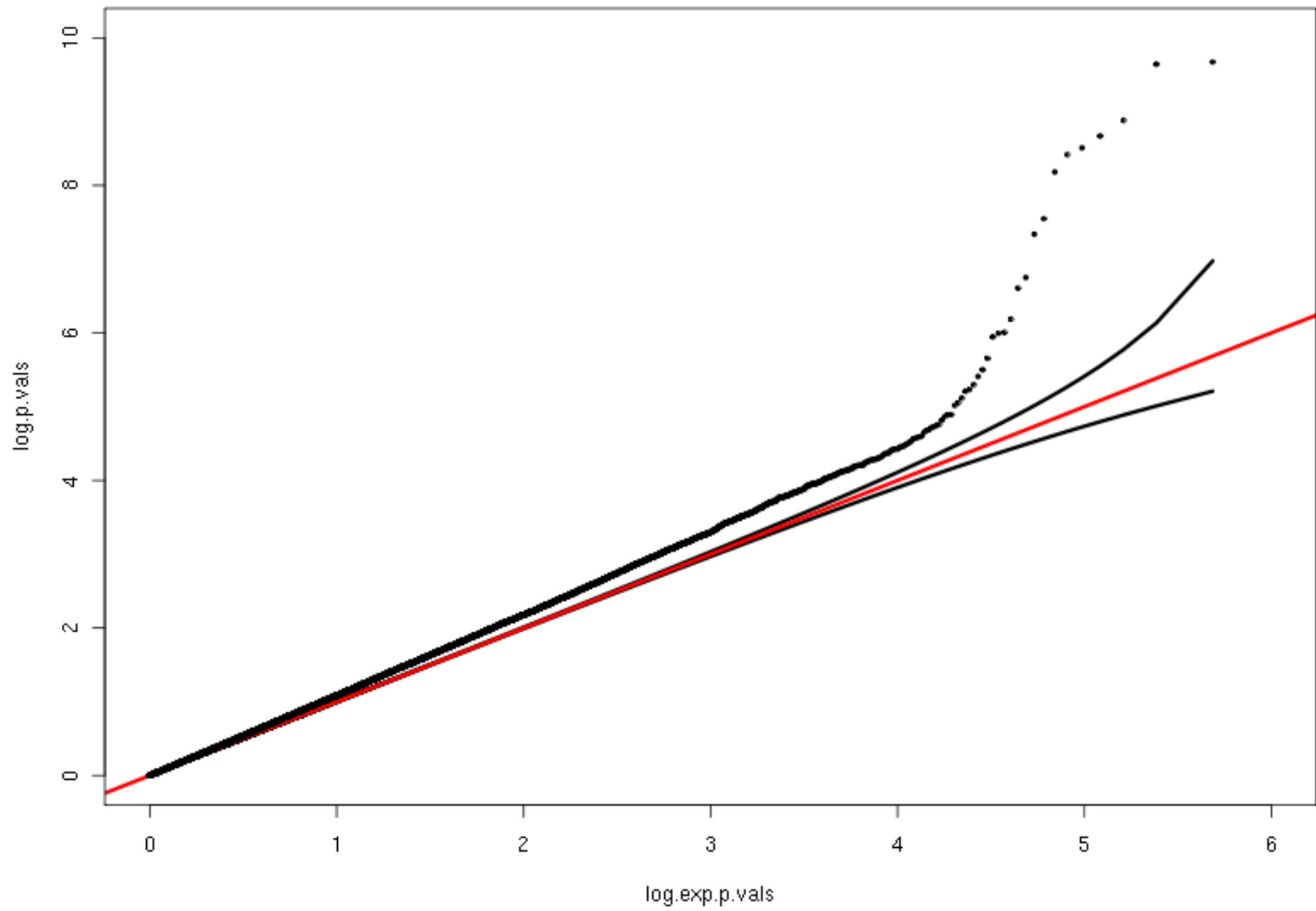
N.



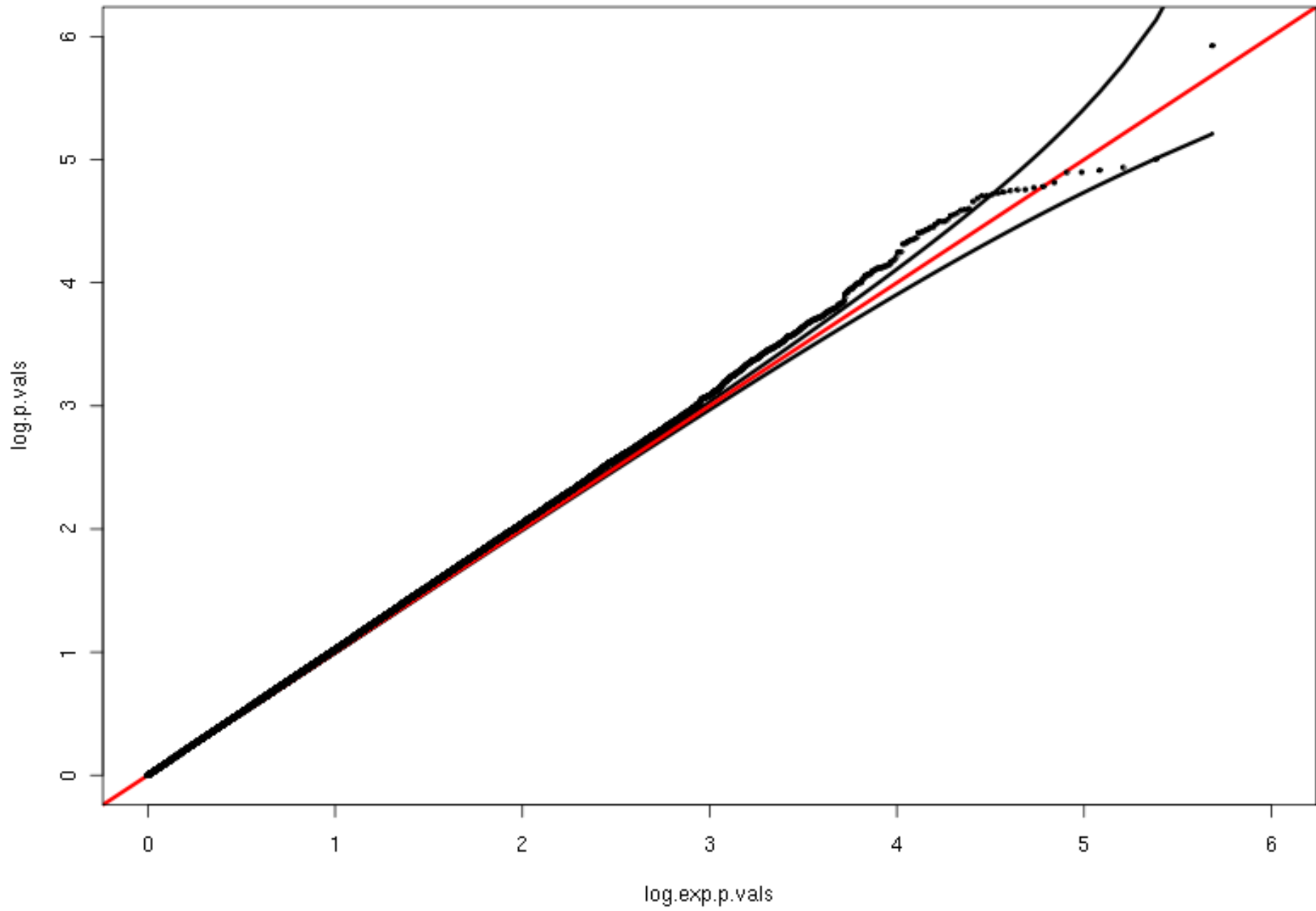
0.



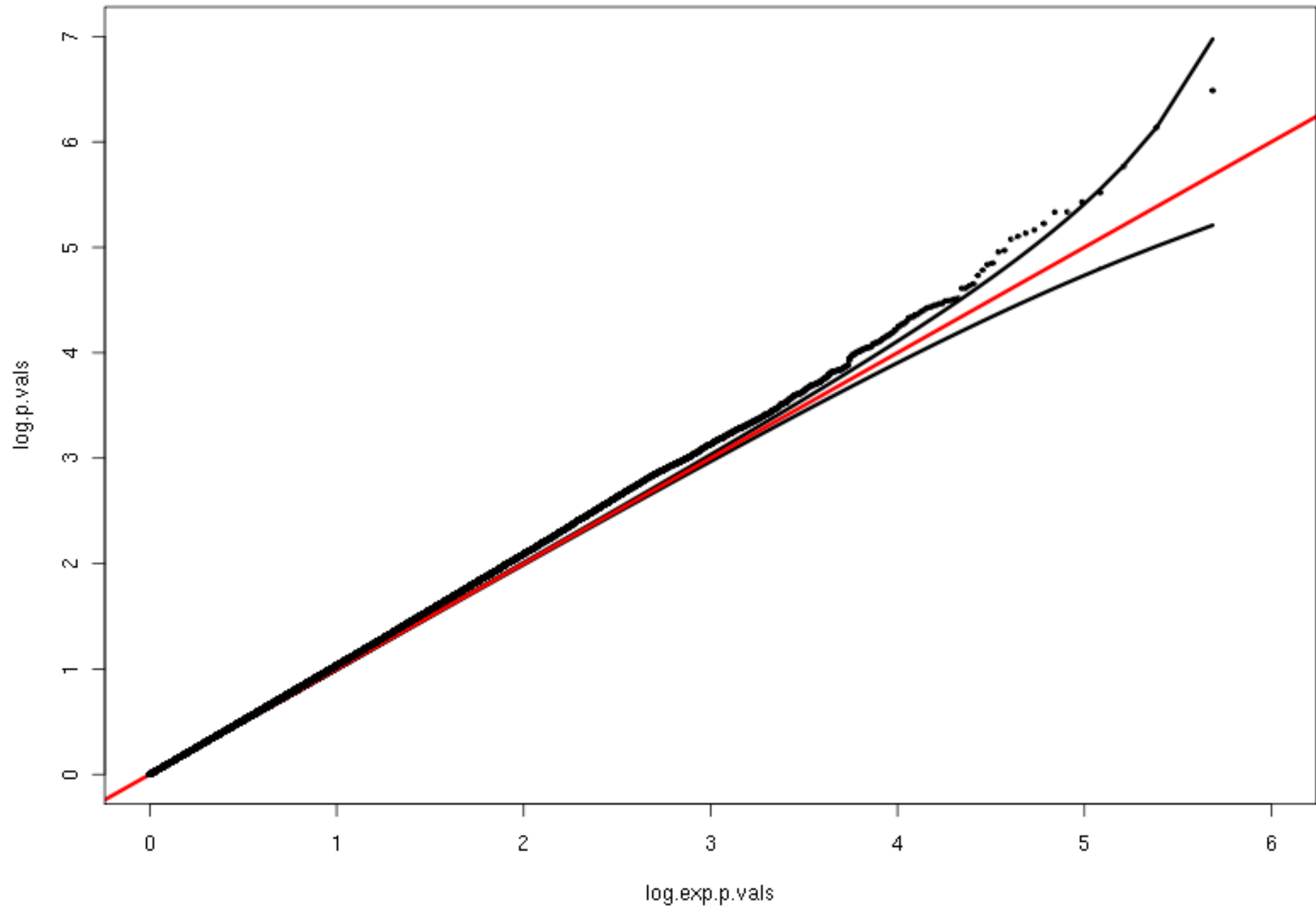
P.



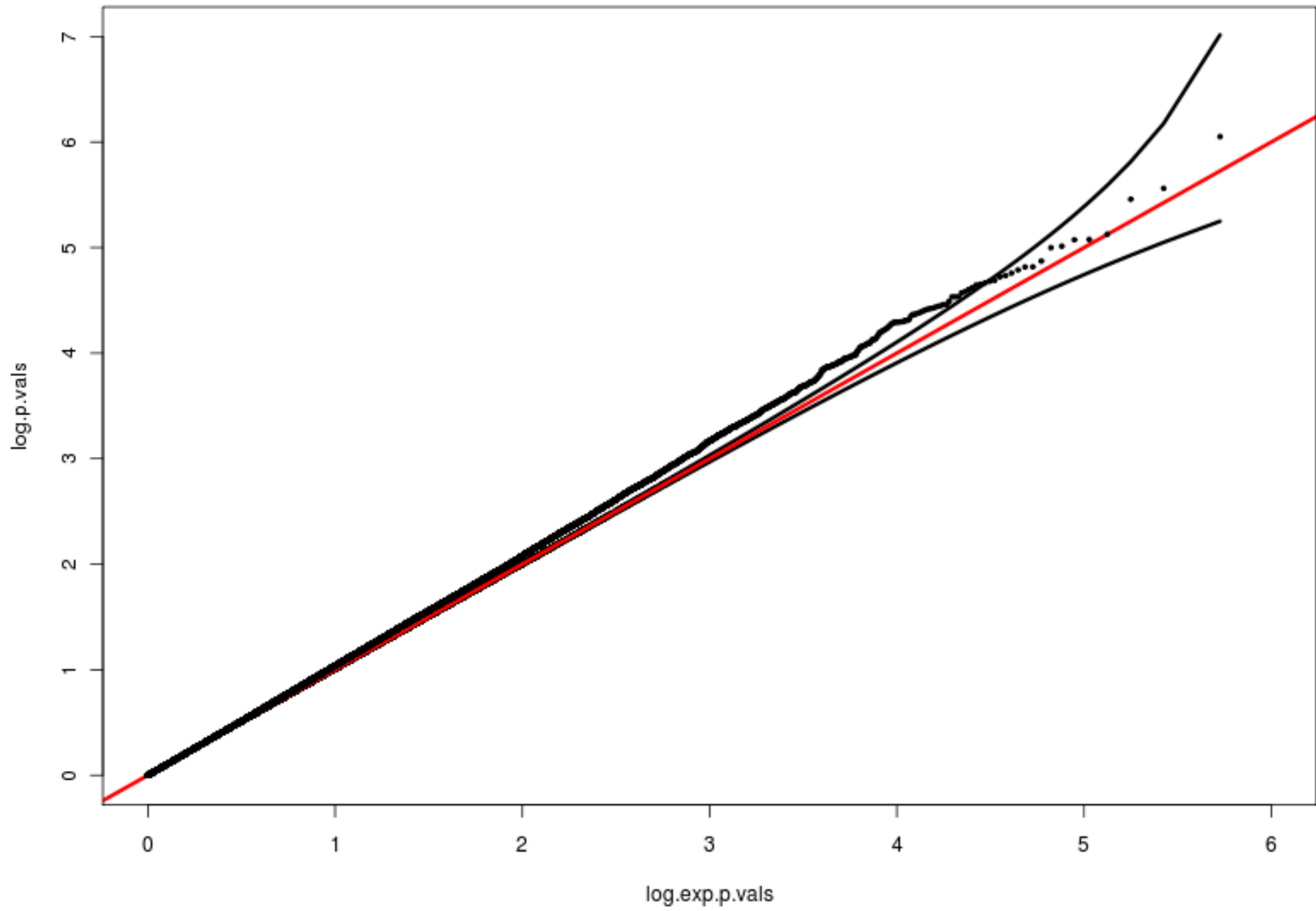
Q.



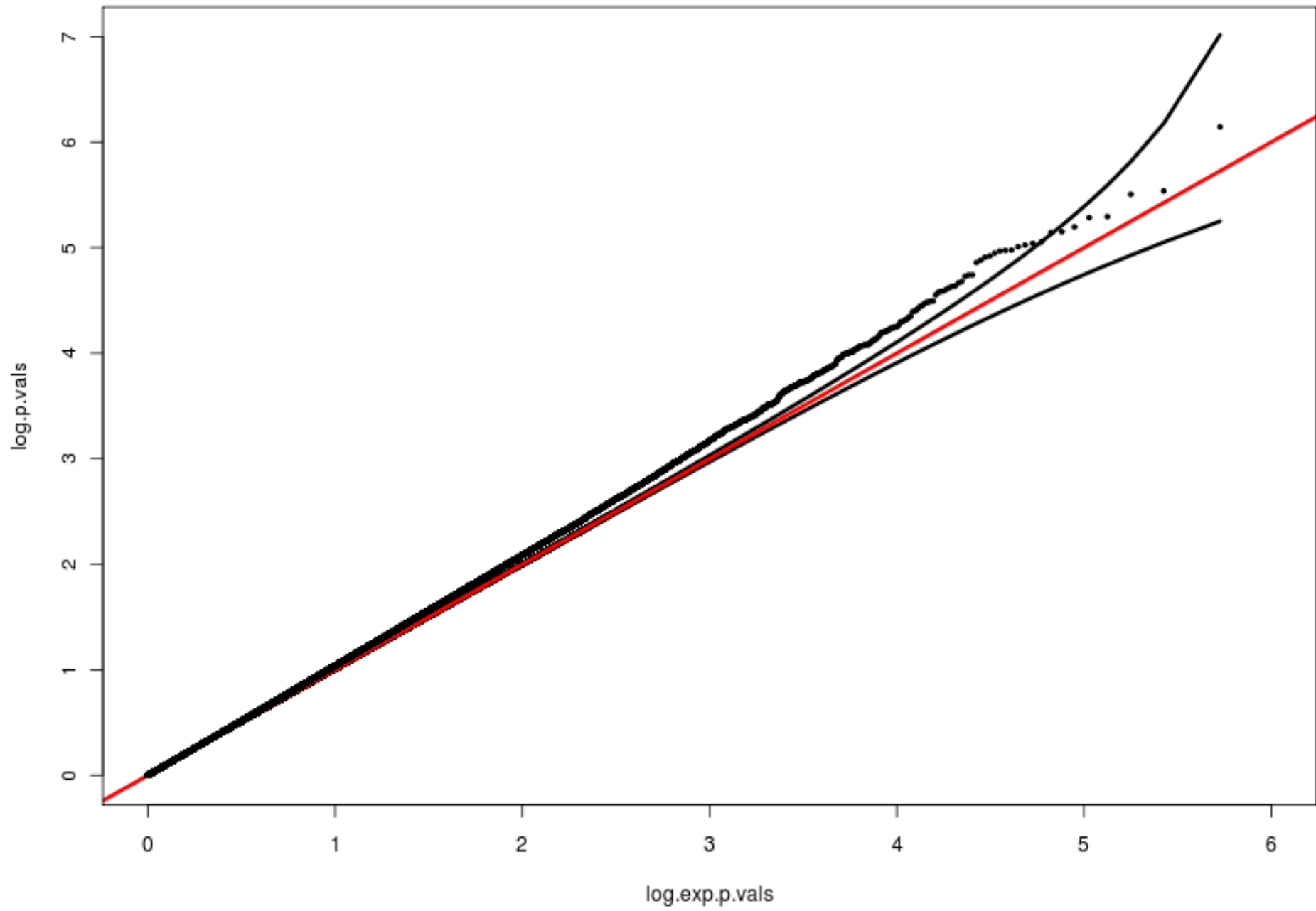
R.



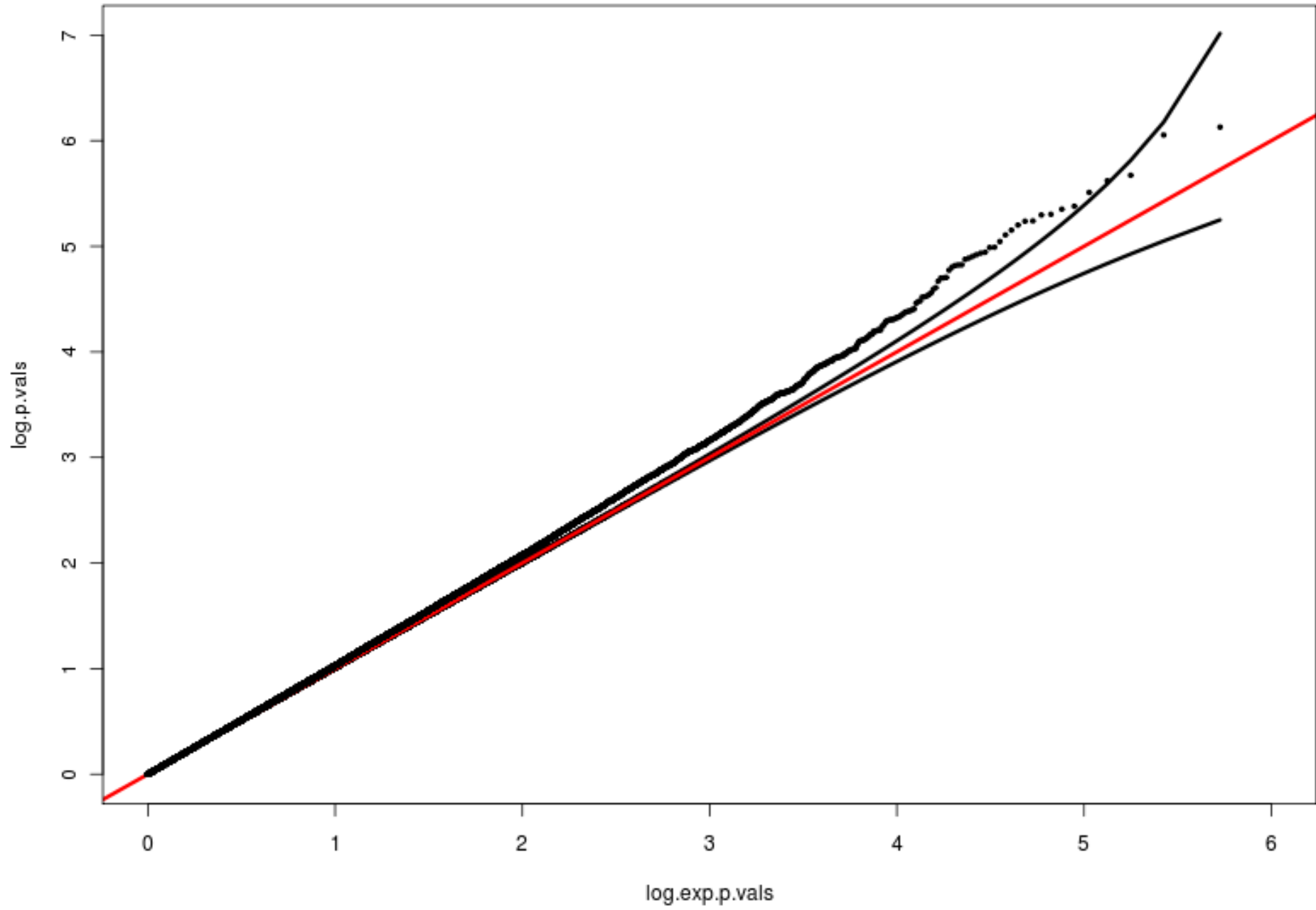
S.



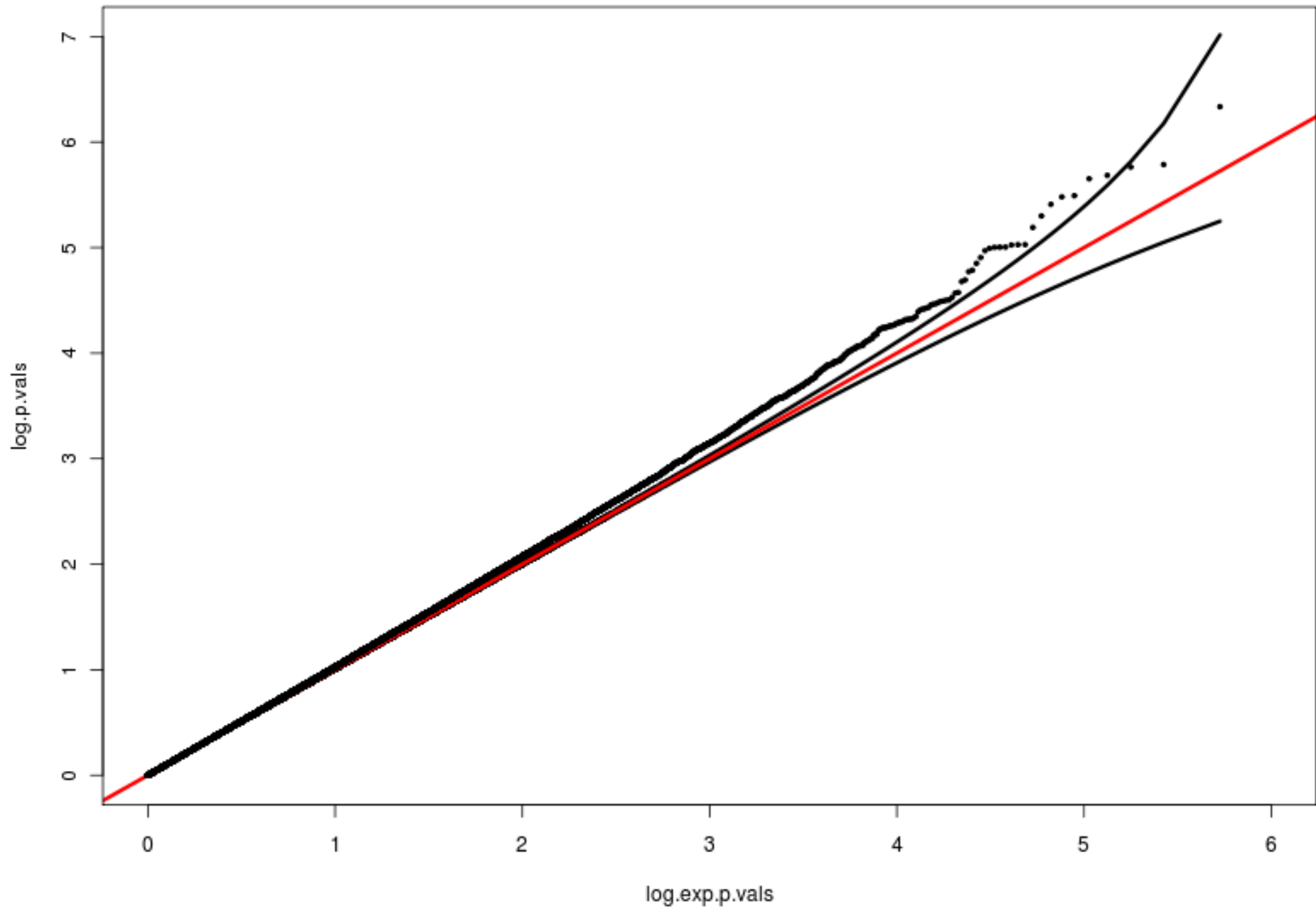
T.



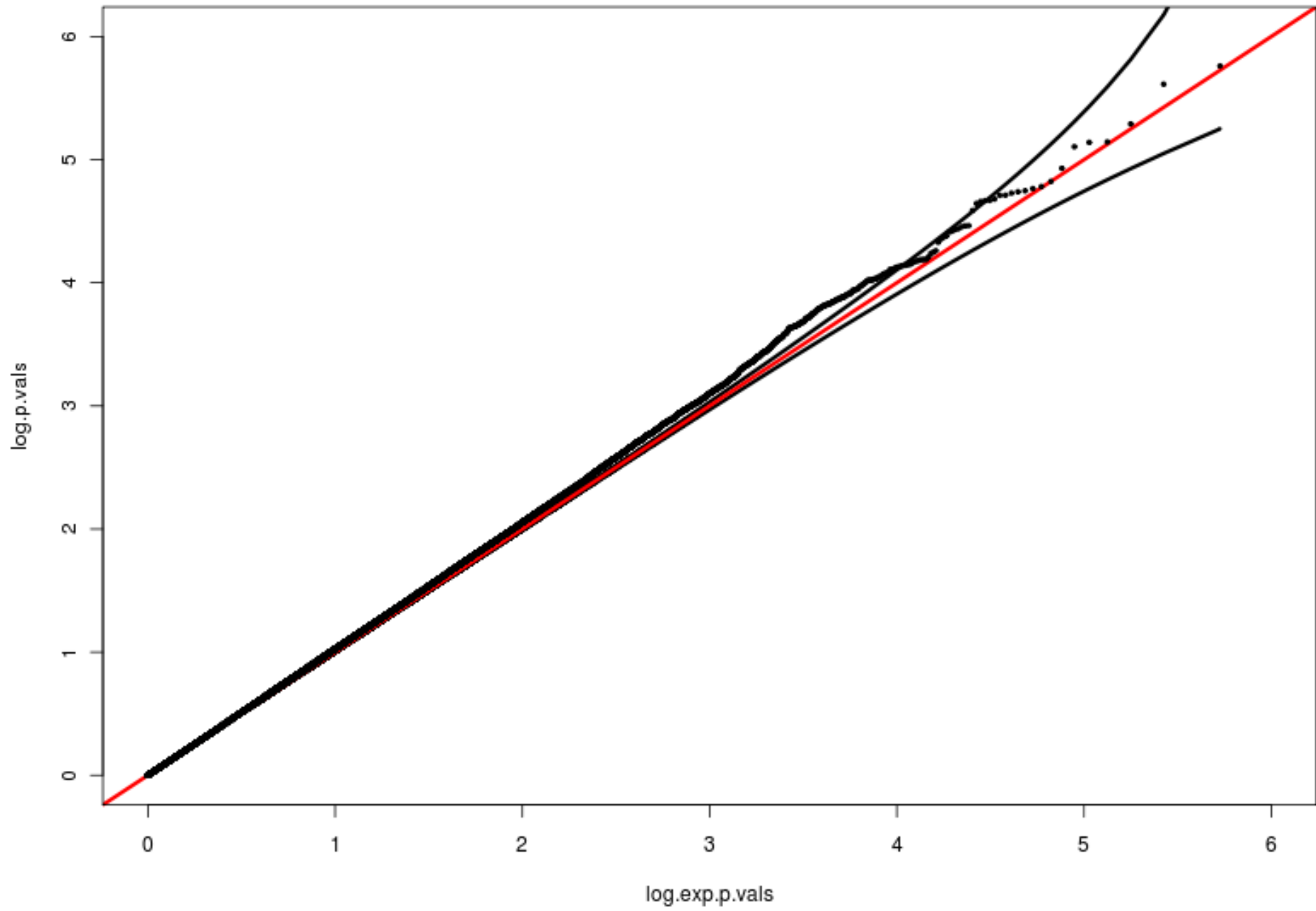
U.



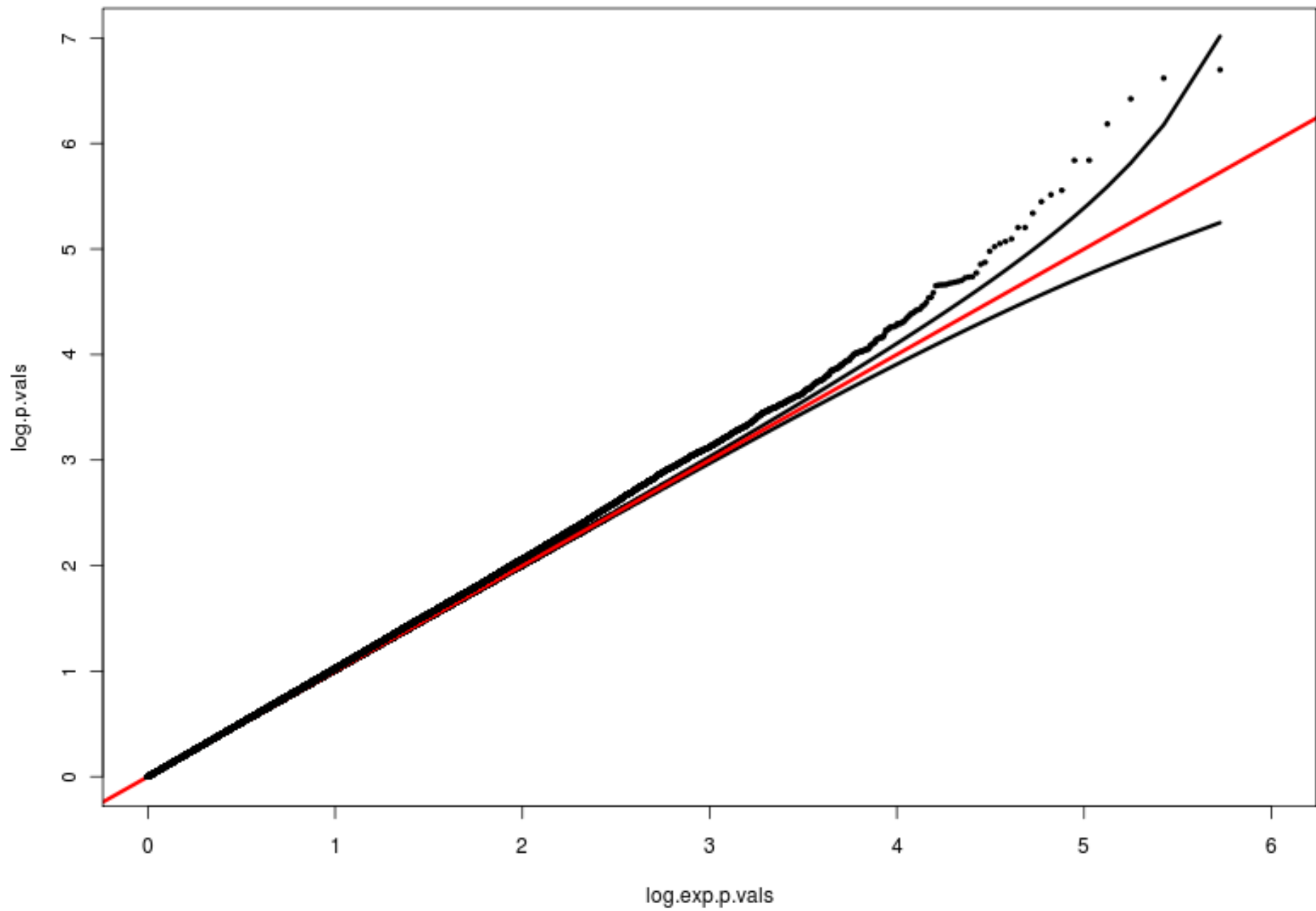
V.



W.

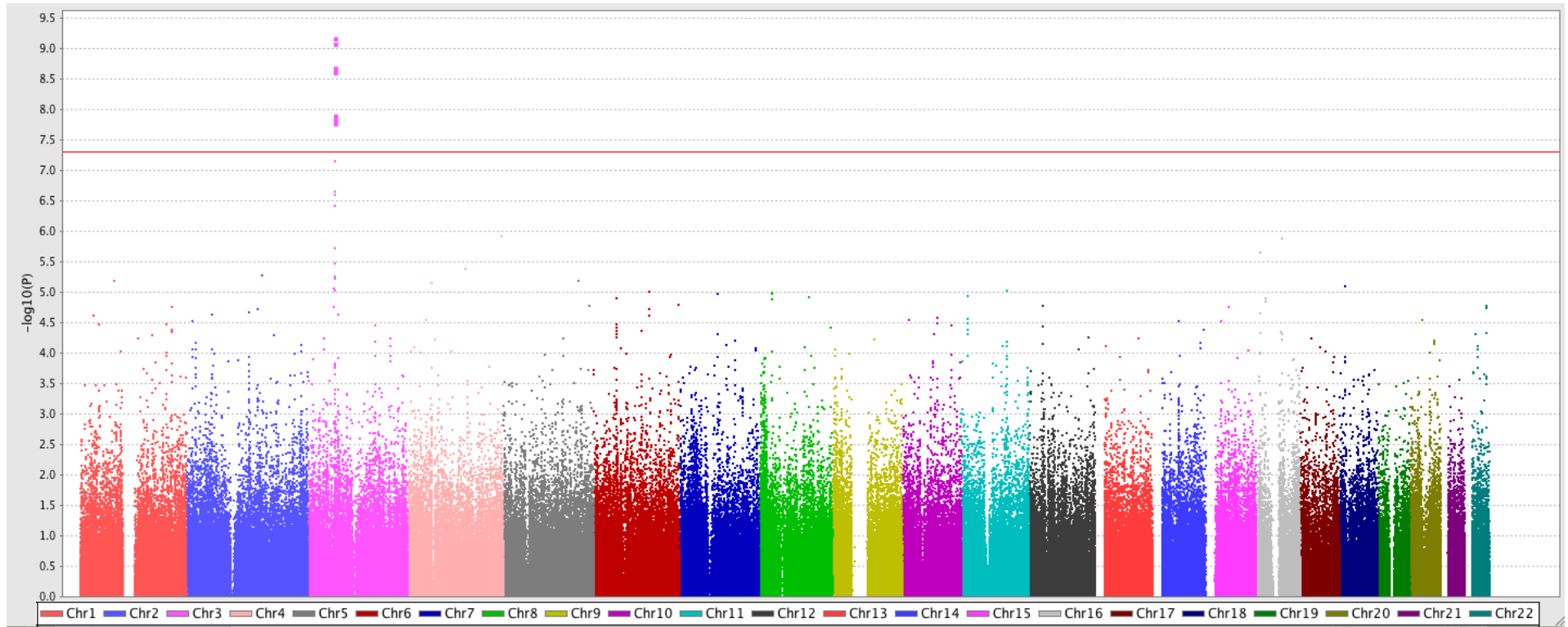


X.

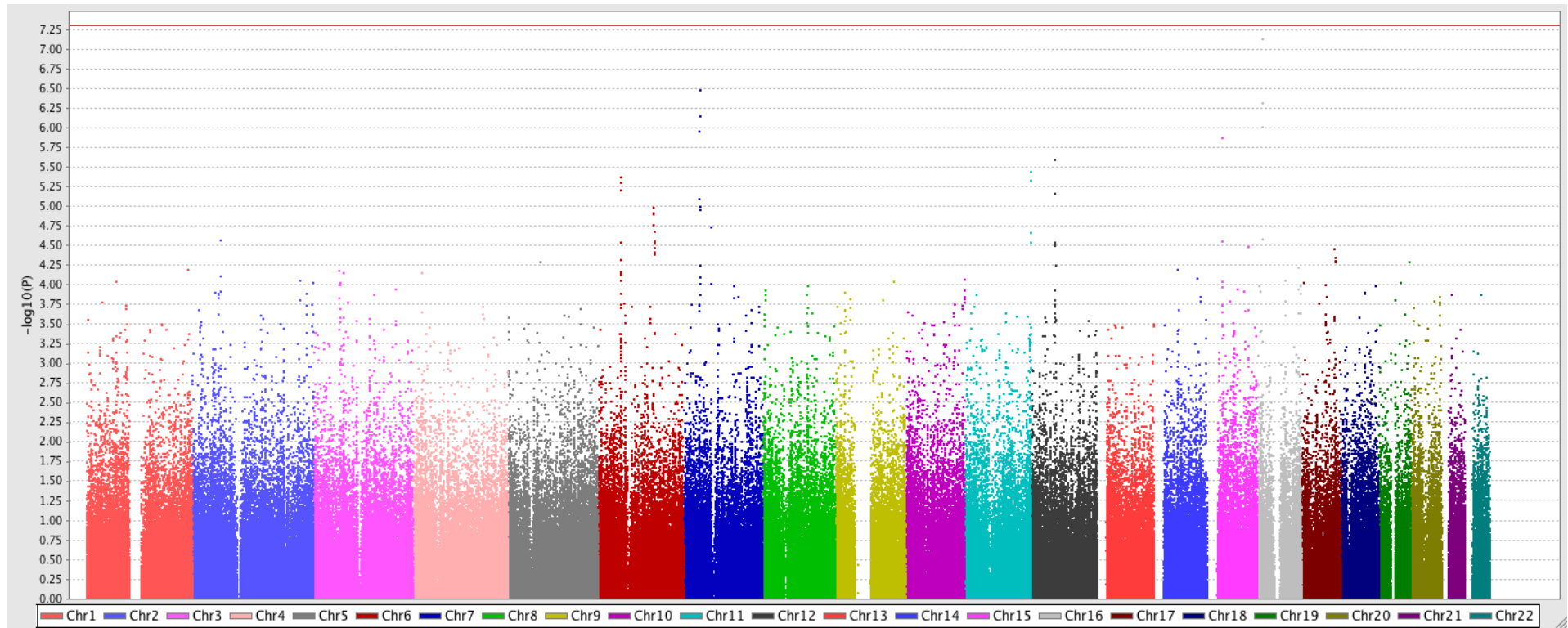


SF2

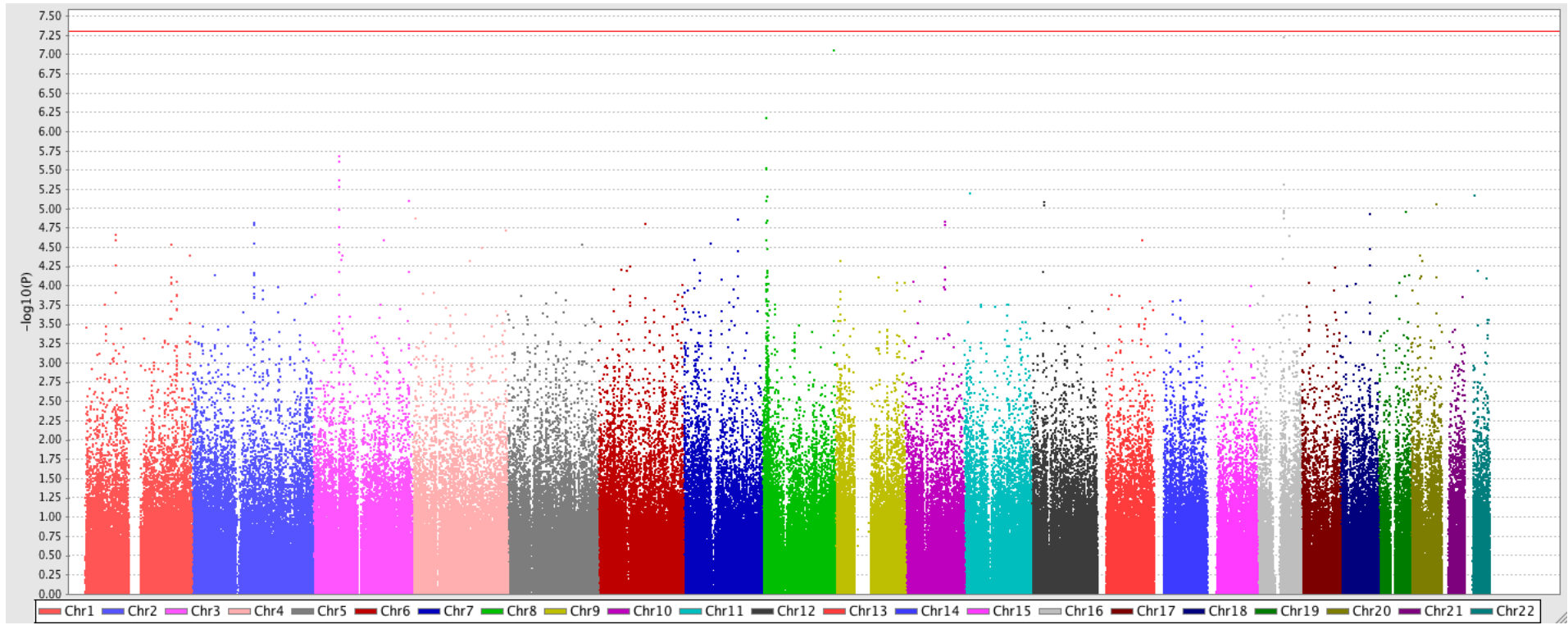
A.



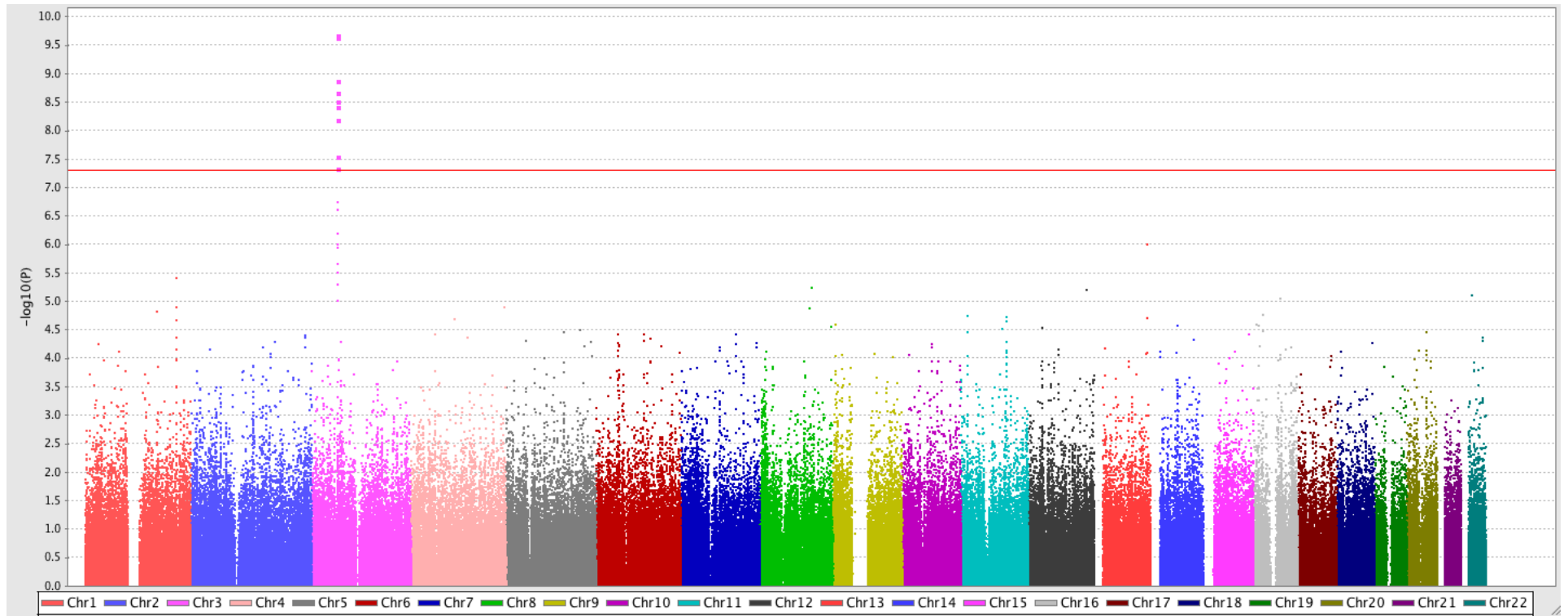
B.



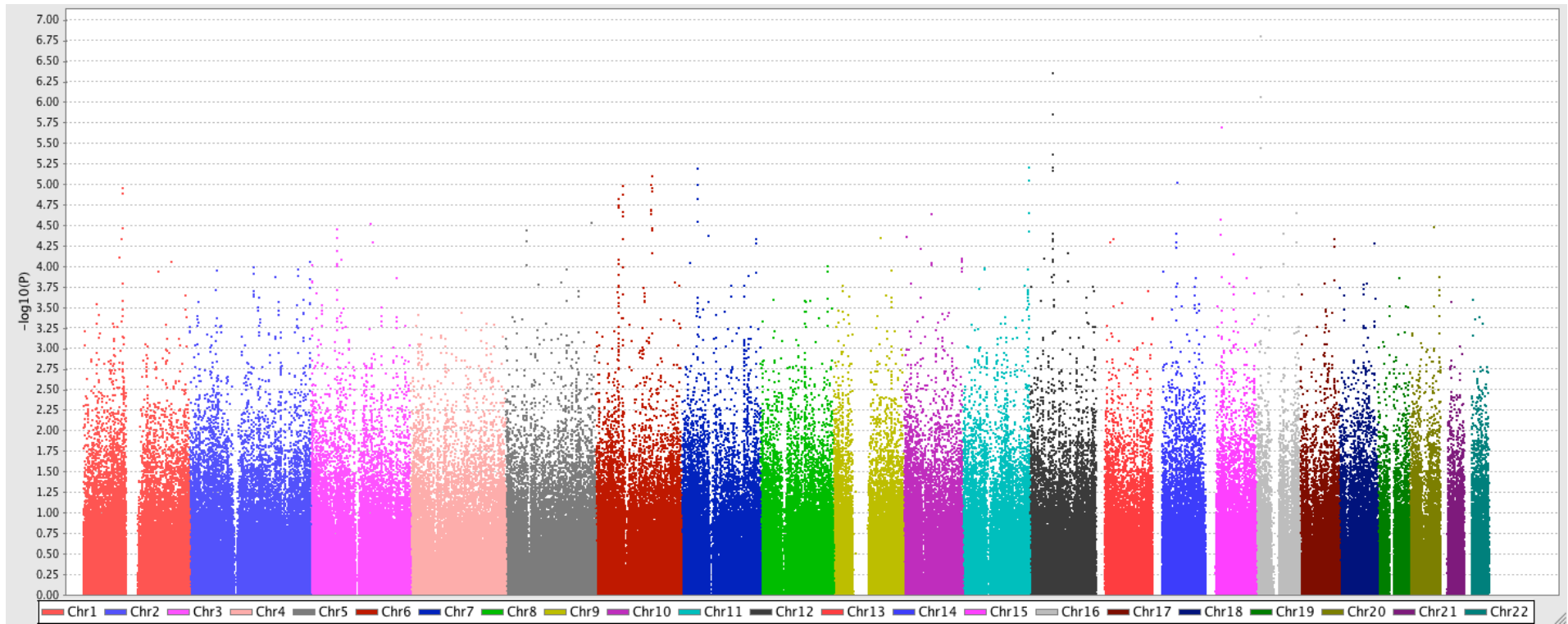
C.



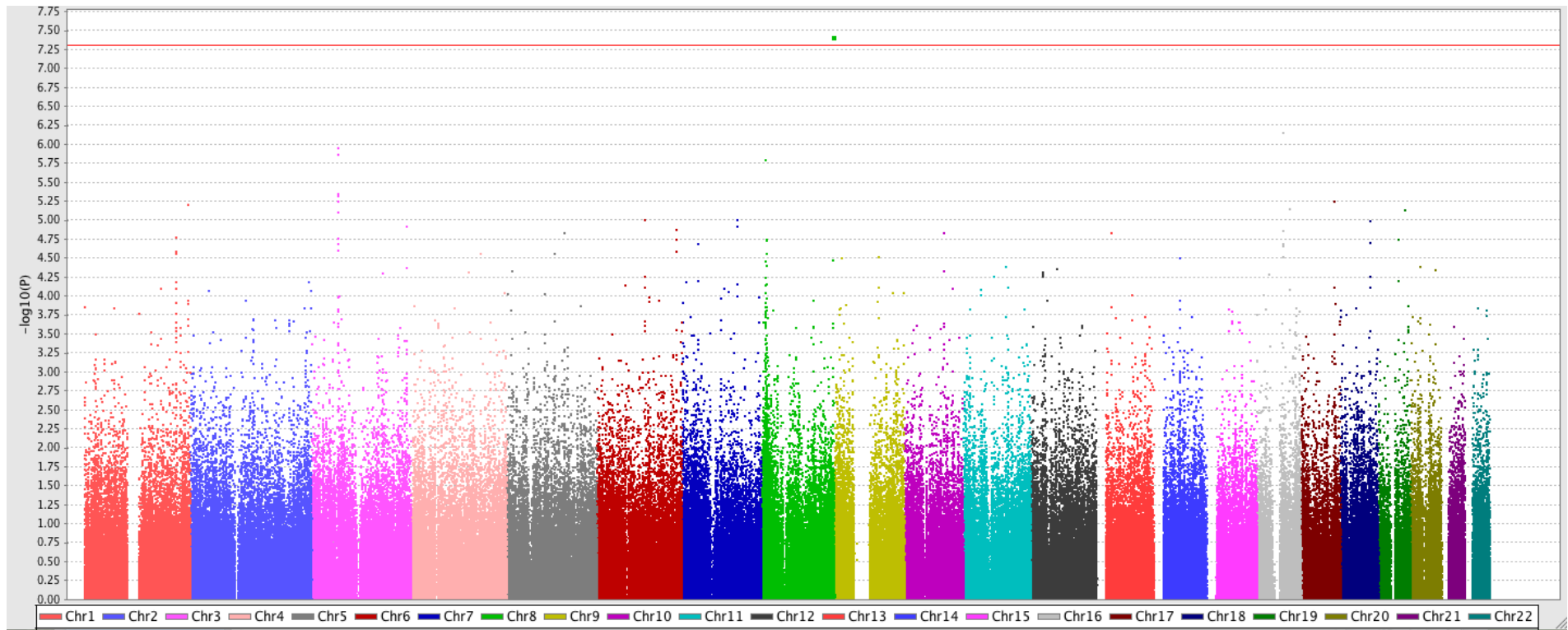
D.



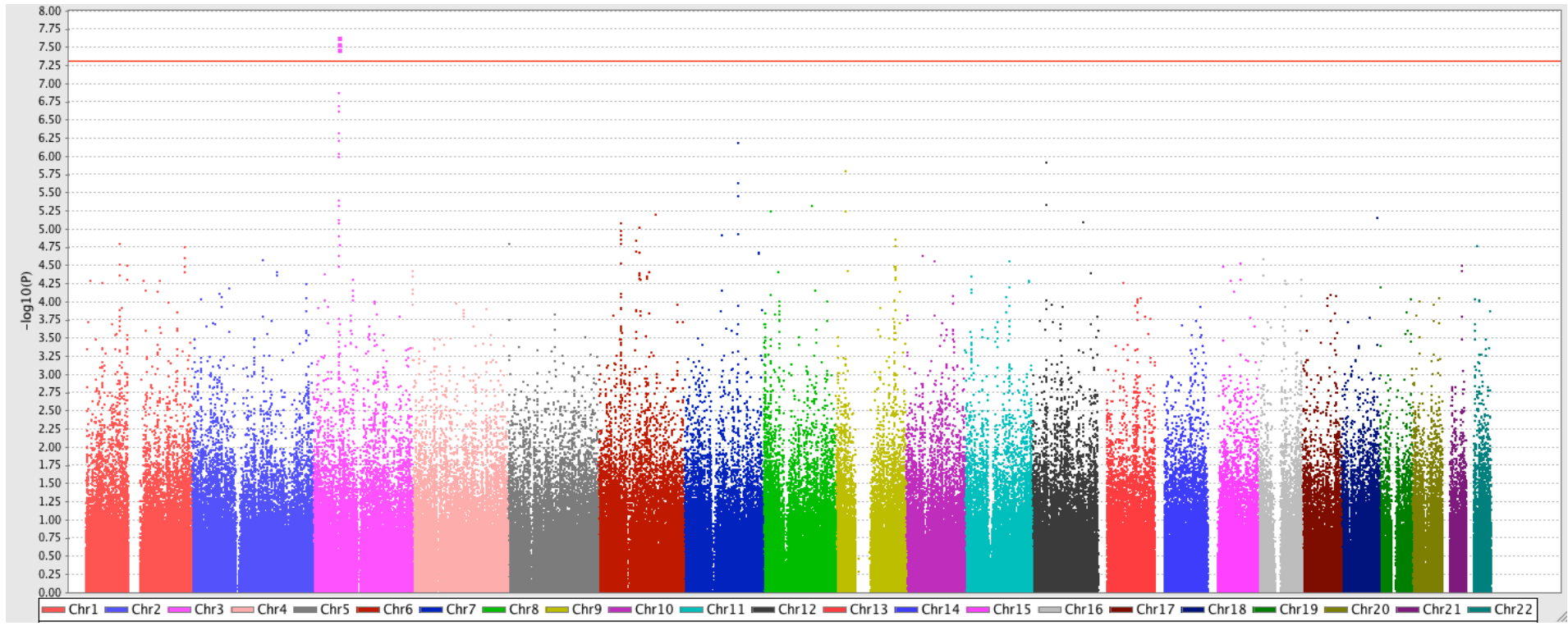
E.



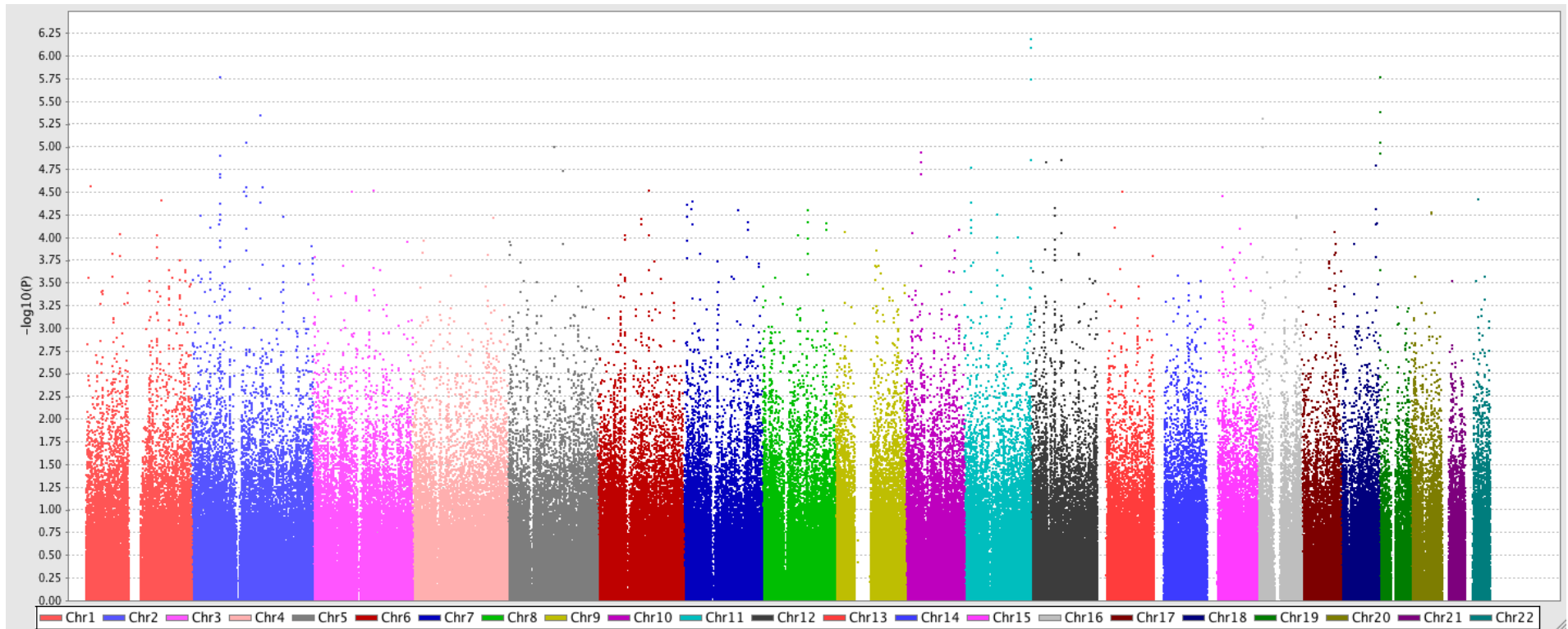
F.



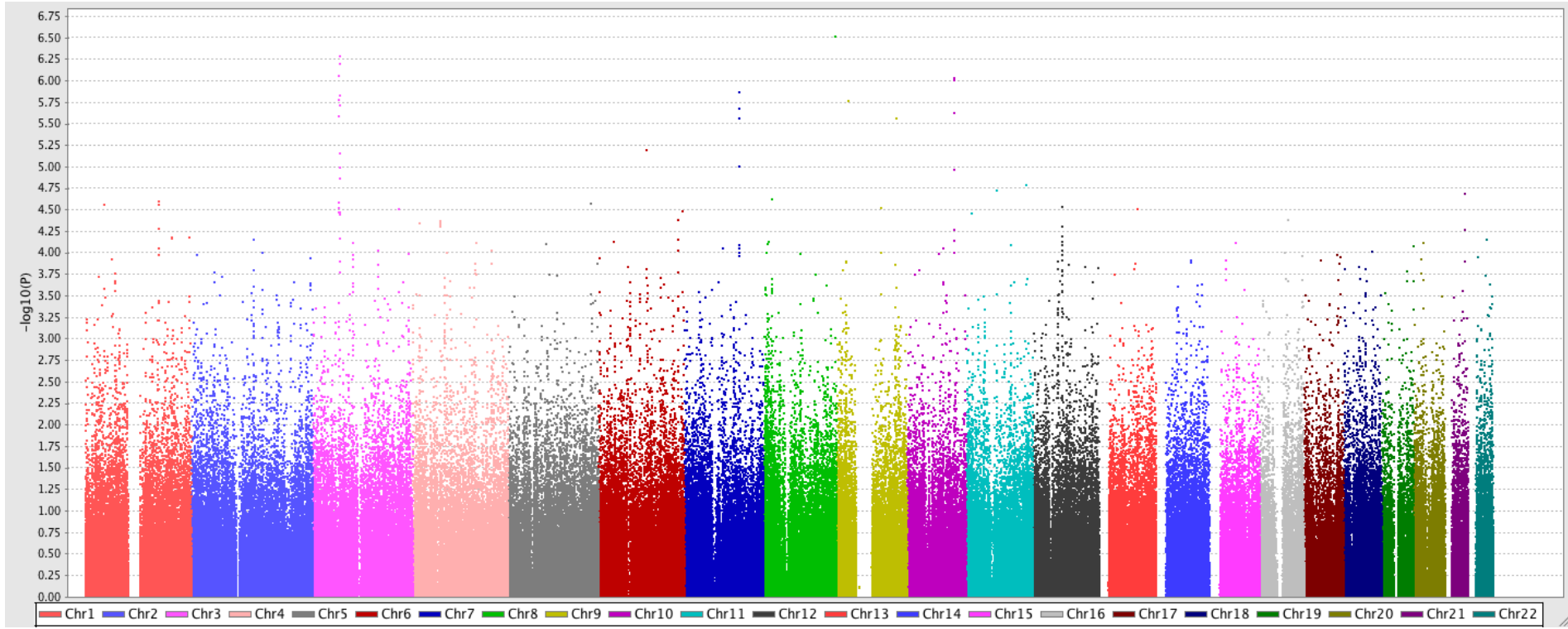
G.



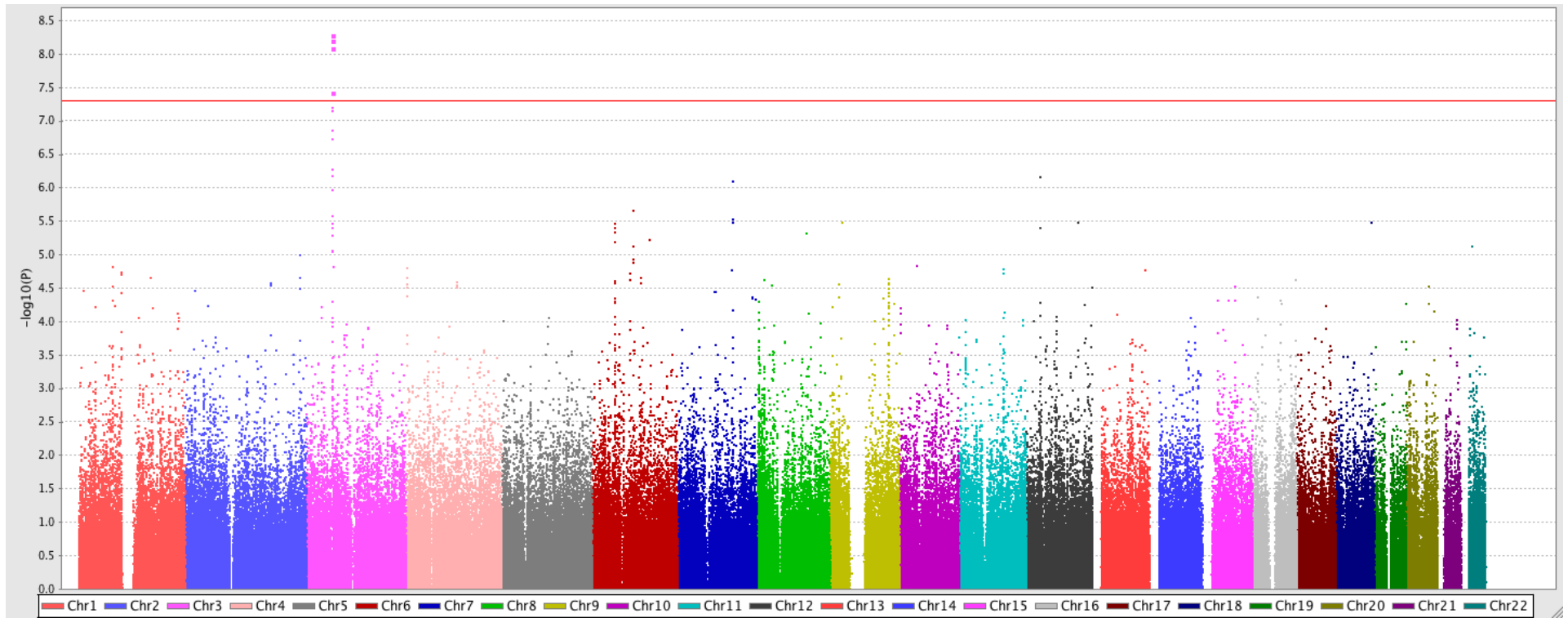
H.



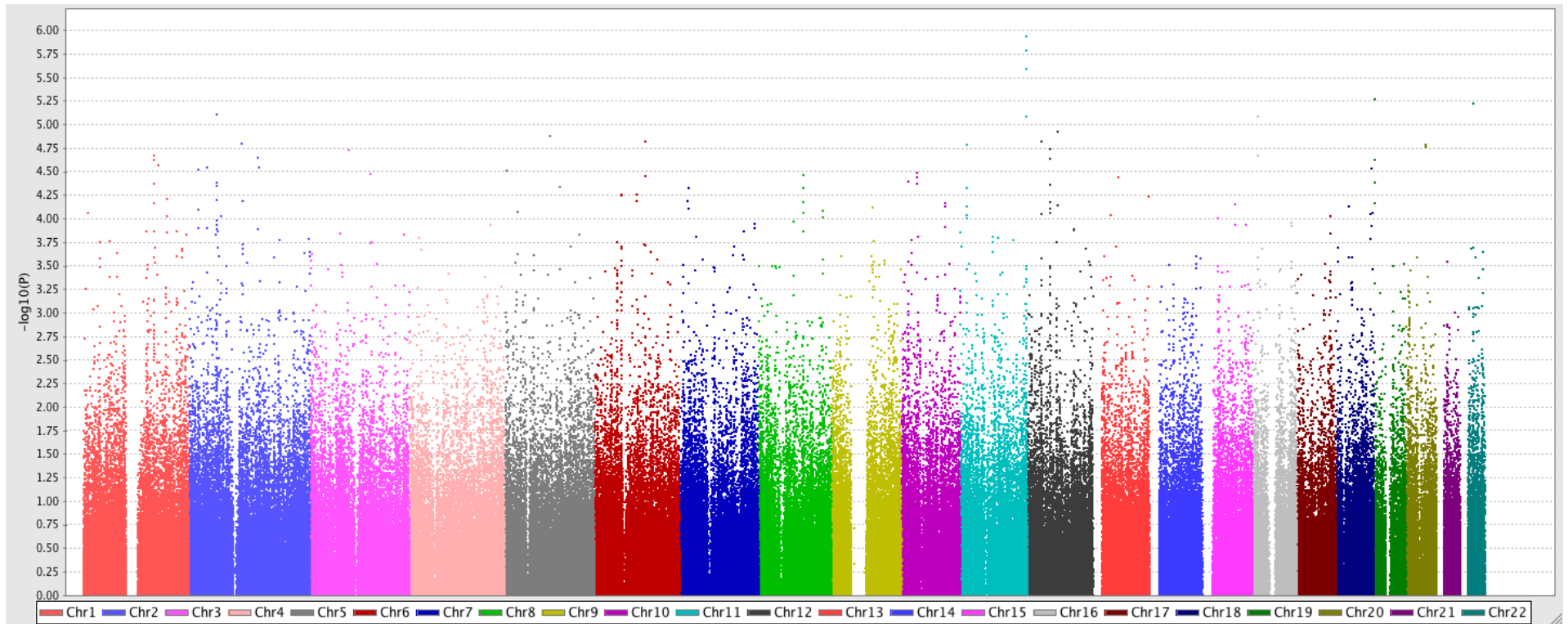
I.



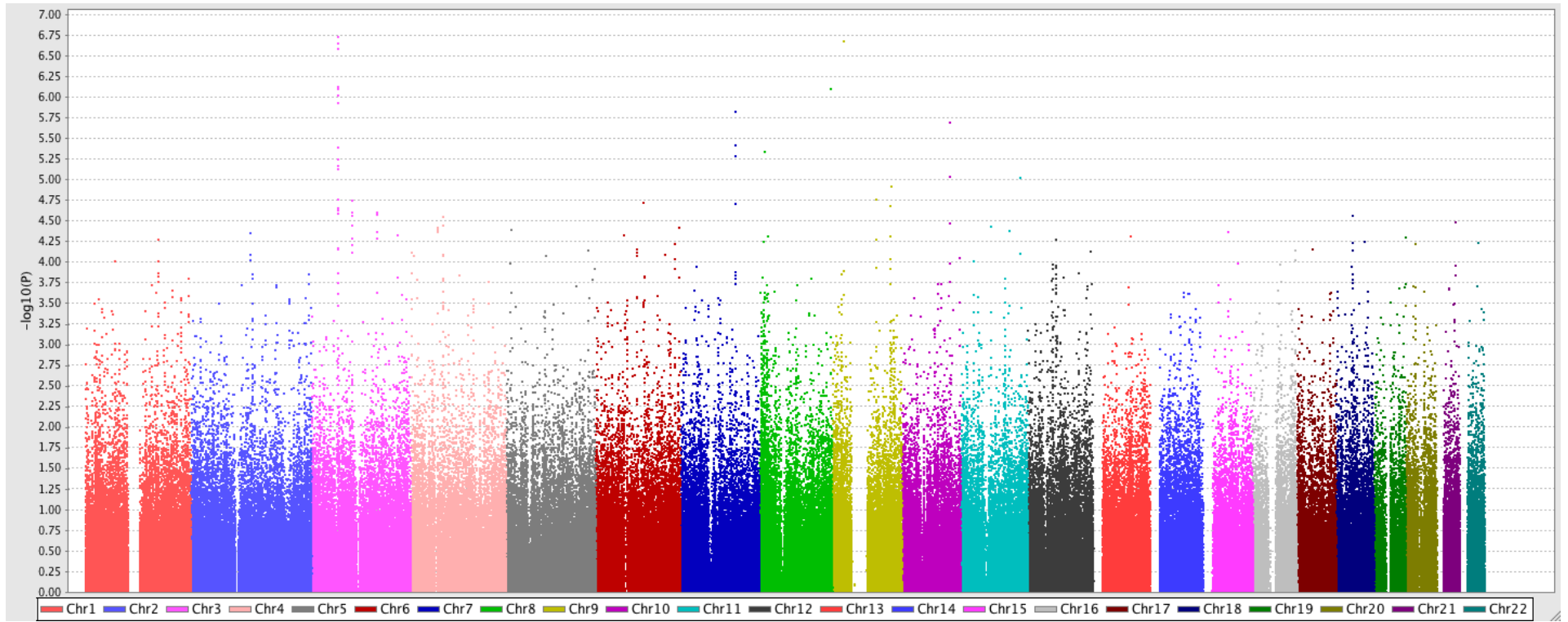
J.



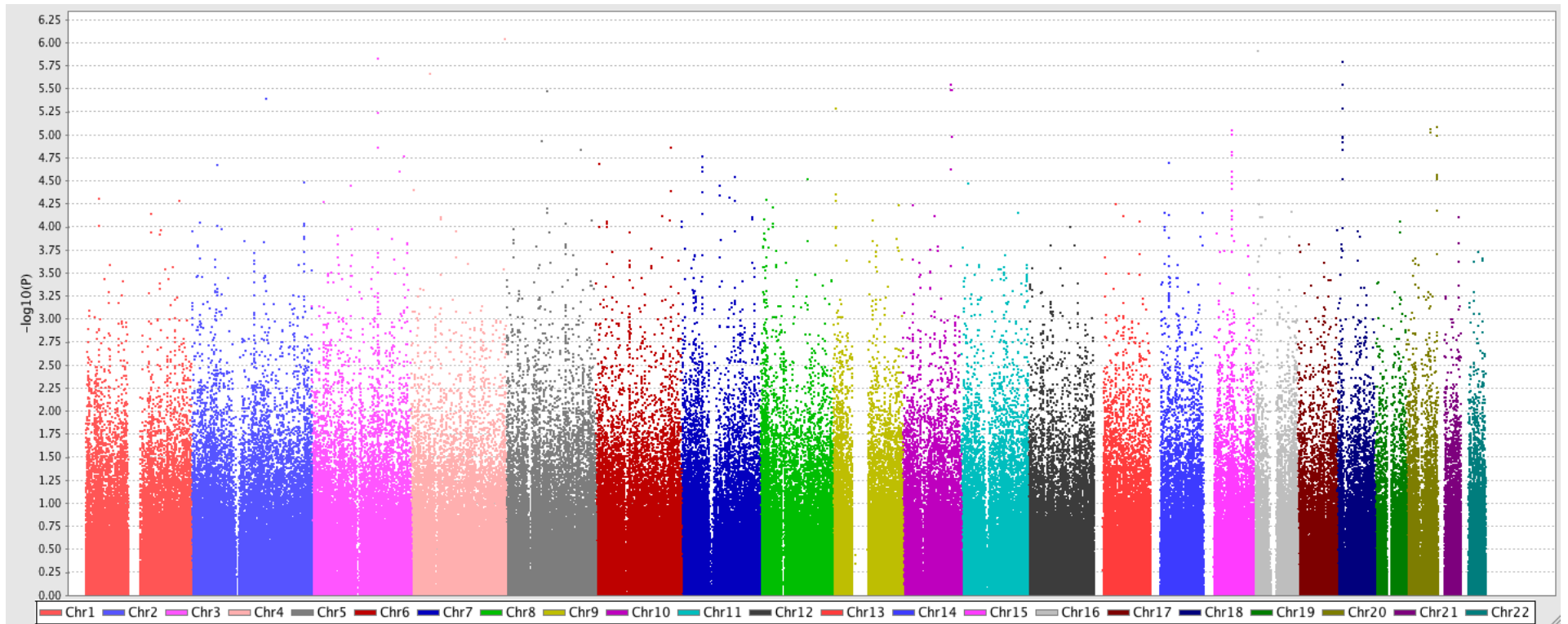
K.



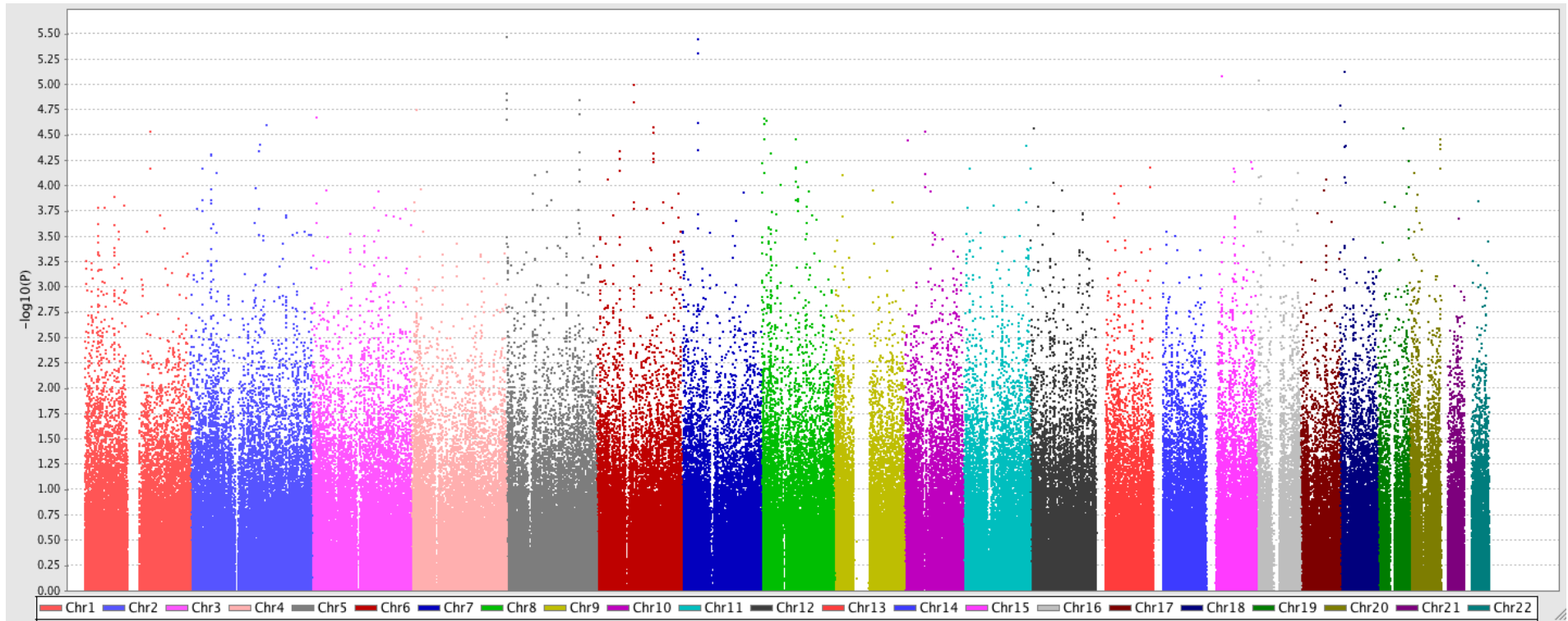
L.



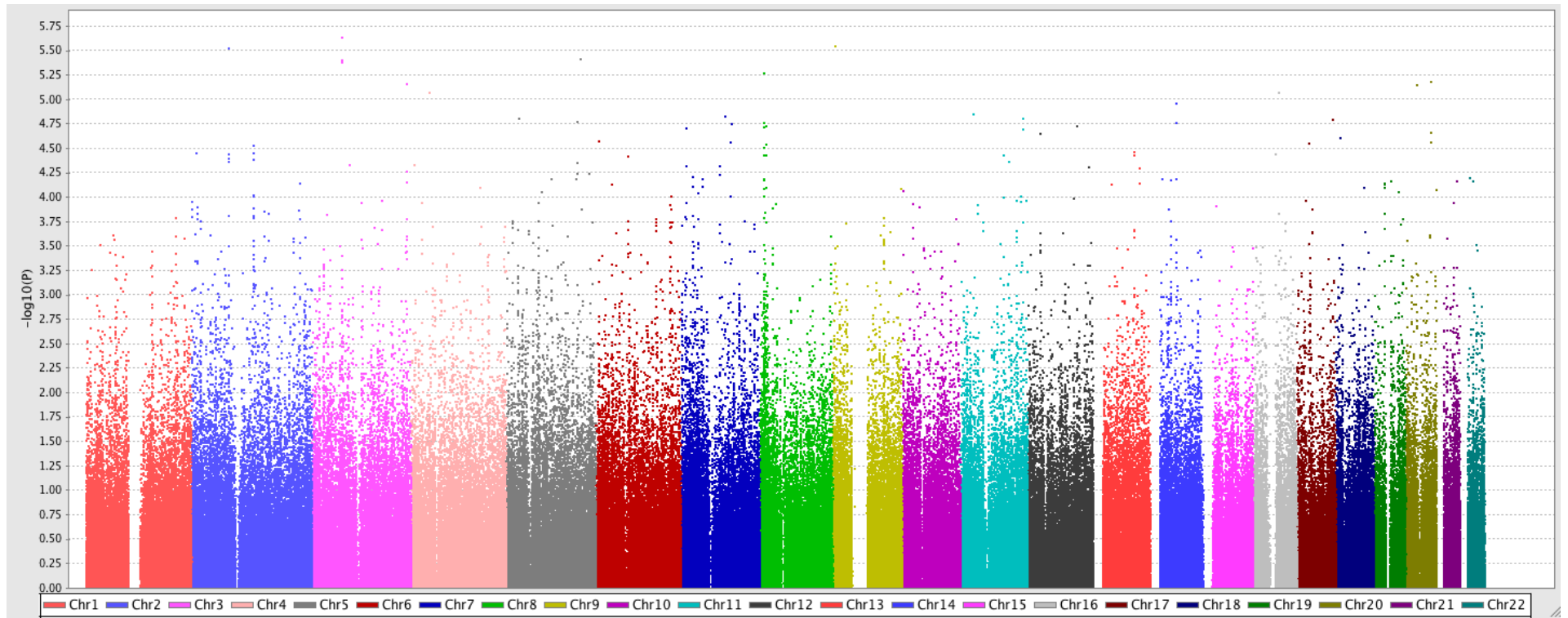
M.



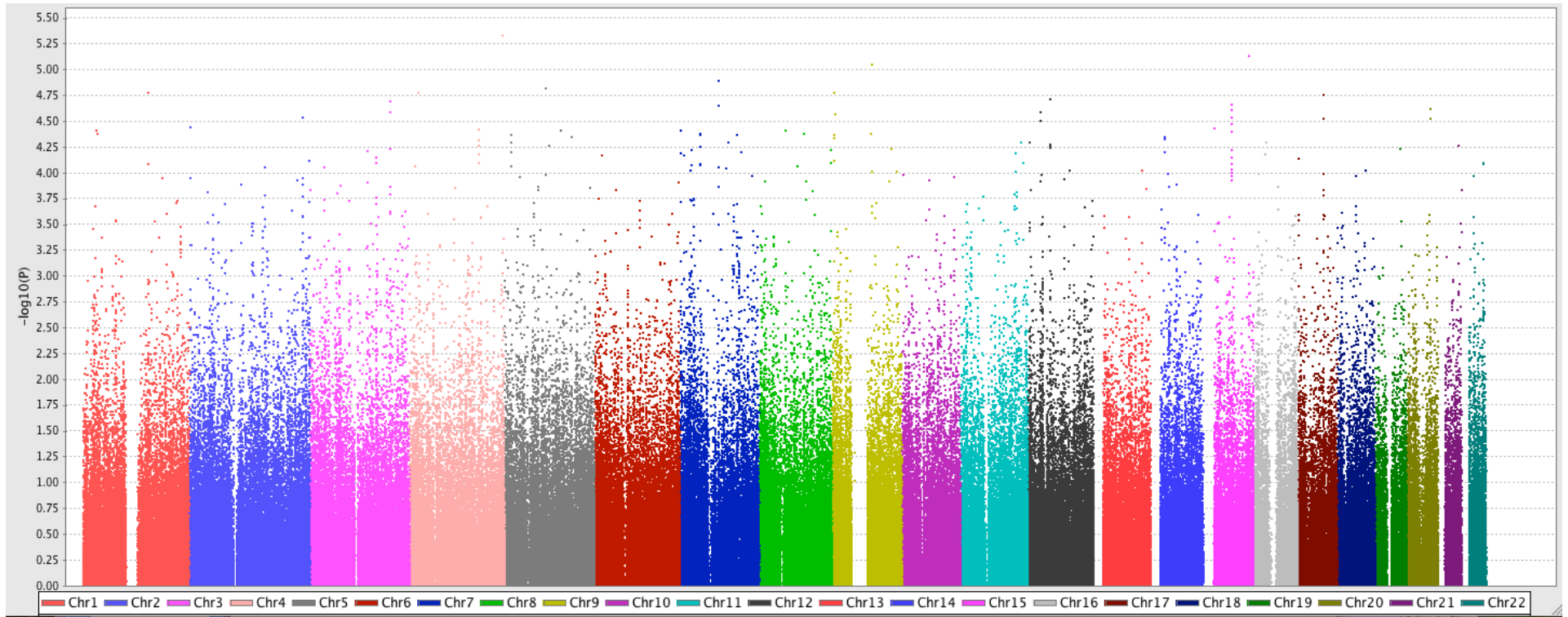
N.



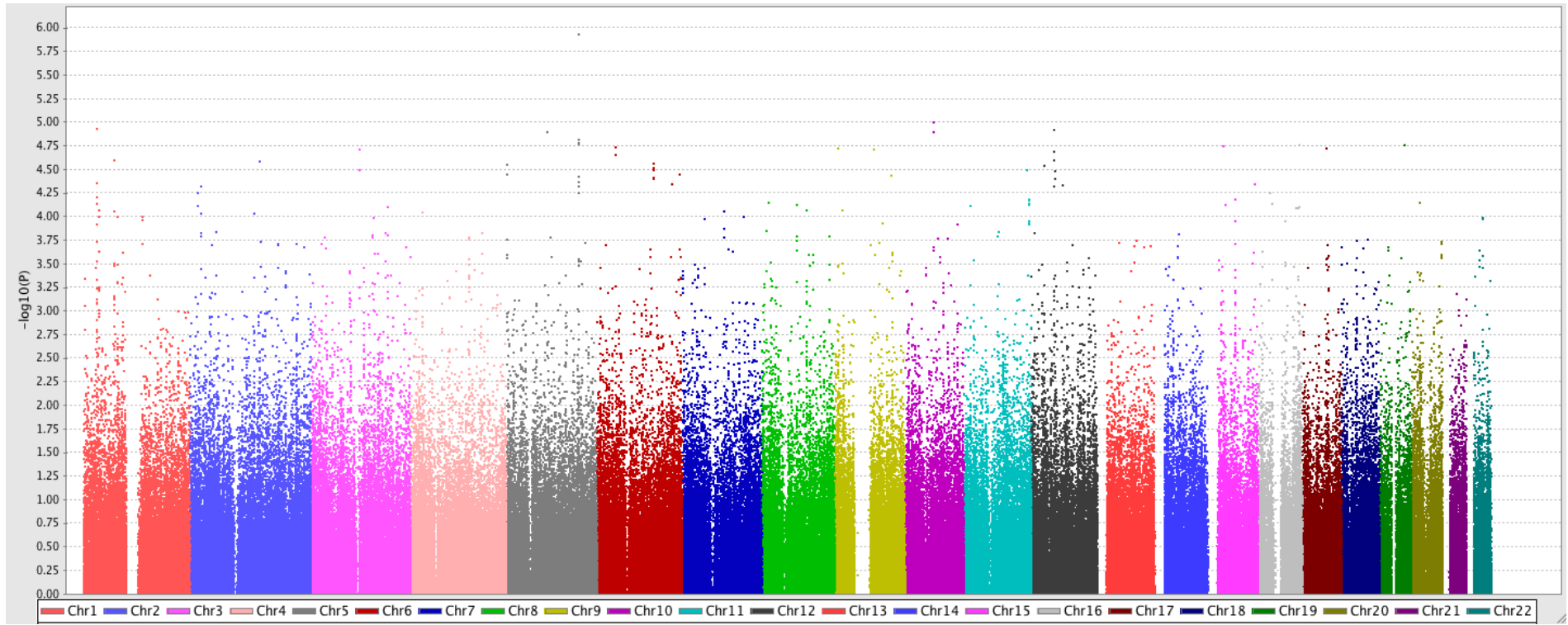
O.



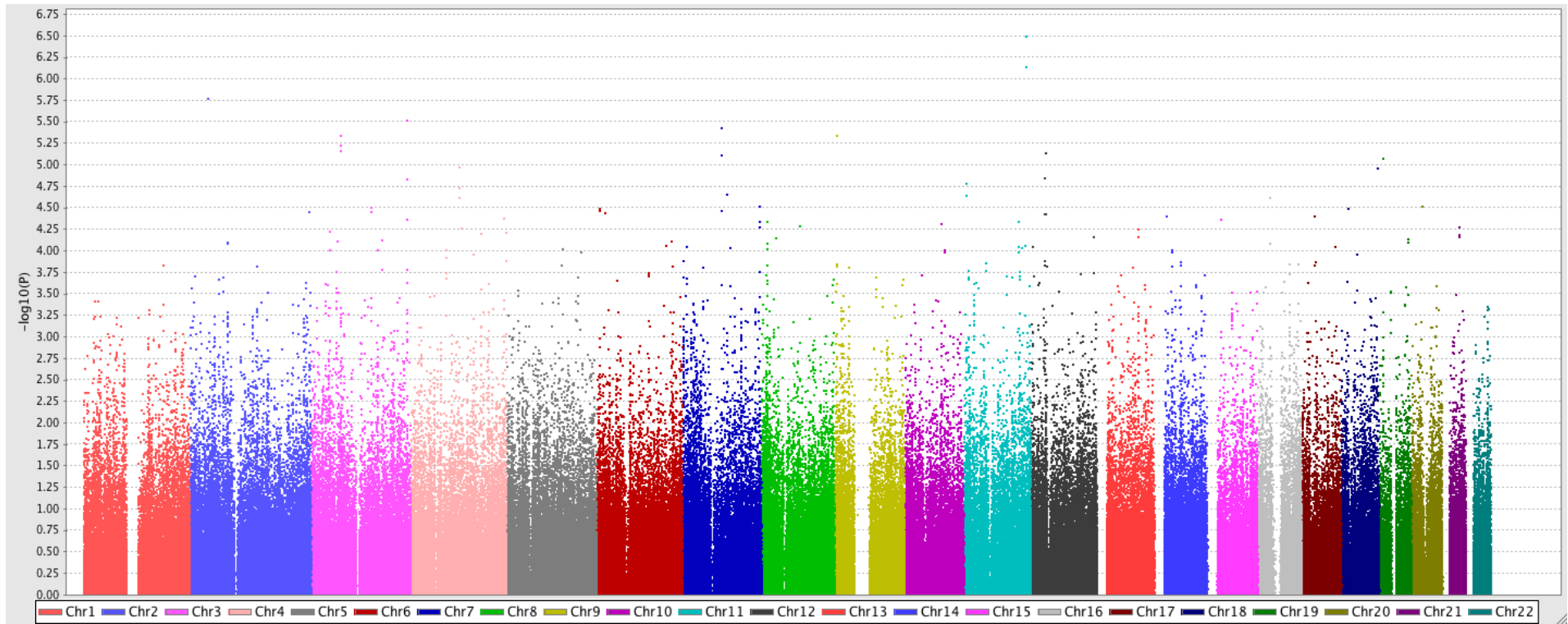
P.



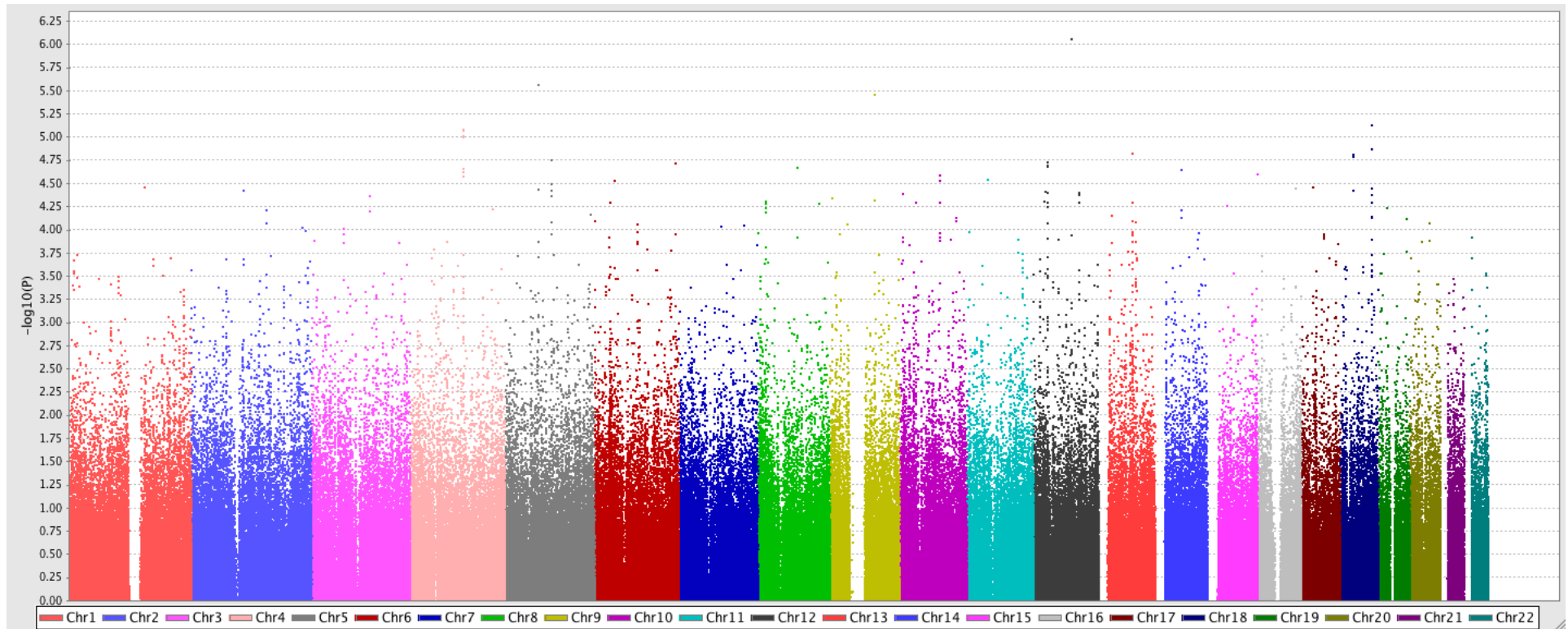
Q.



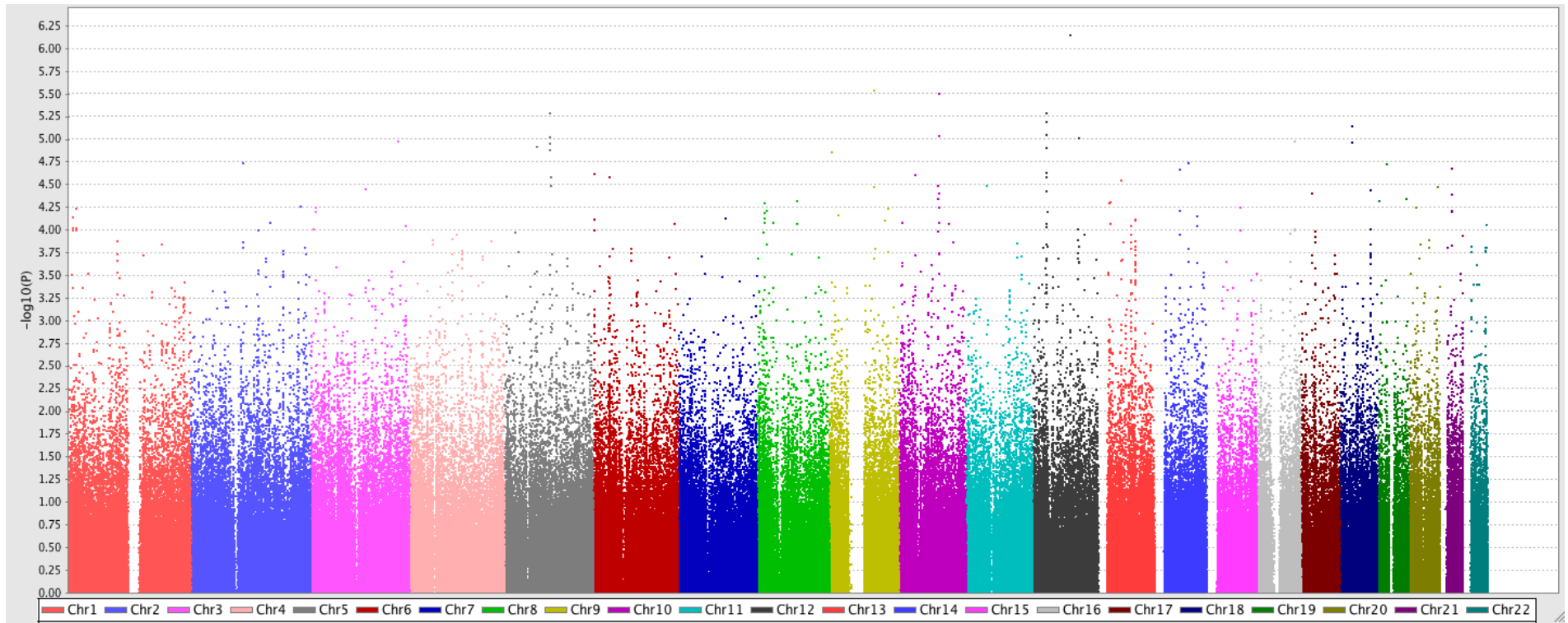
R.



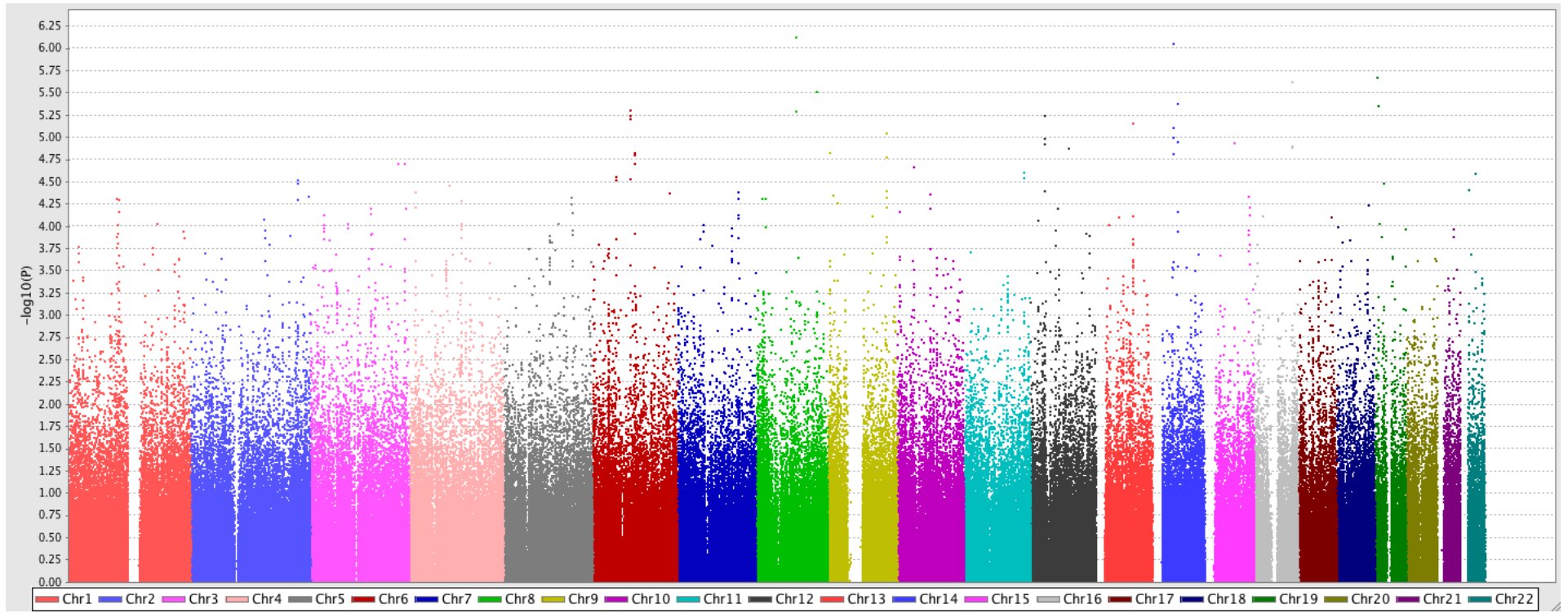
S.



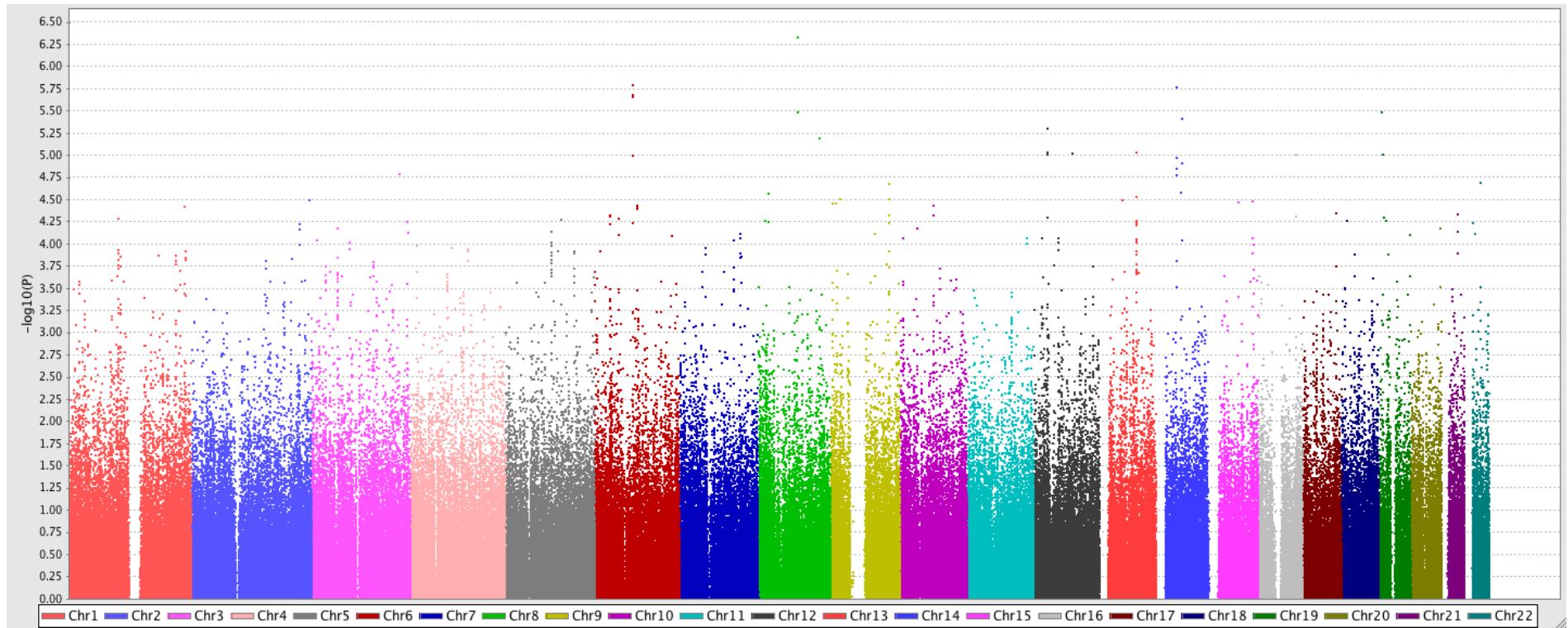
T.



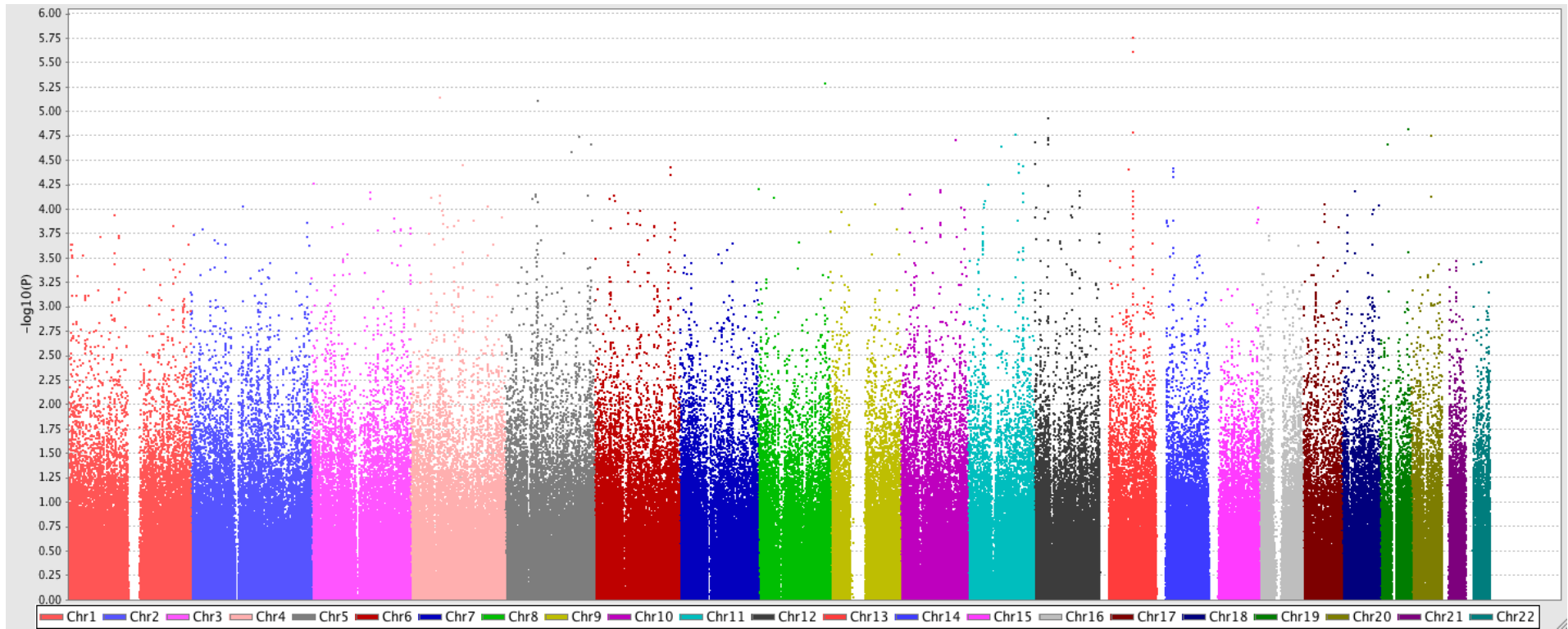
U.



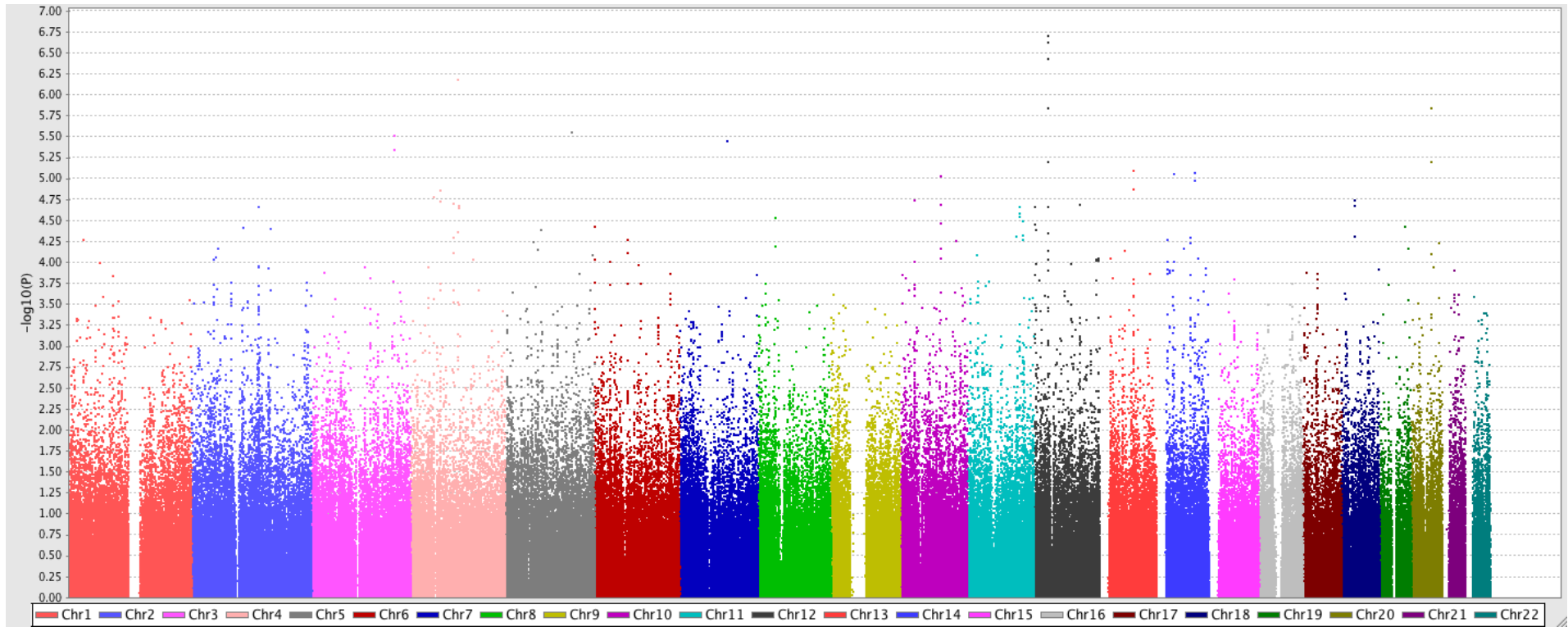
V.



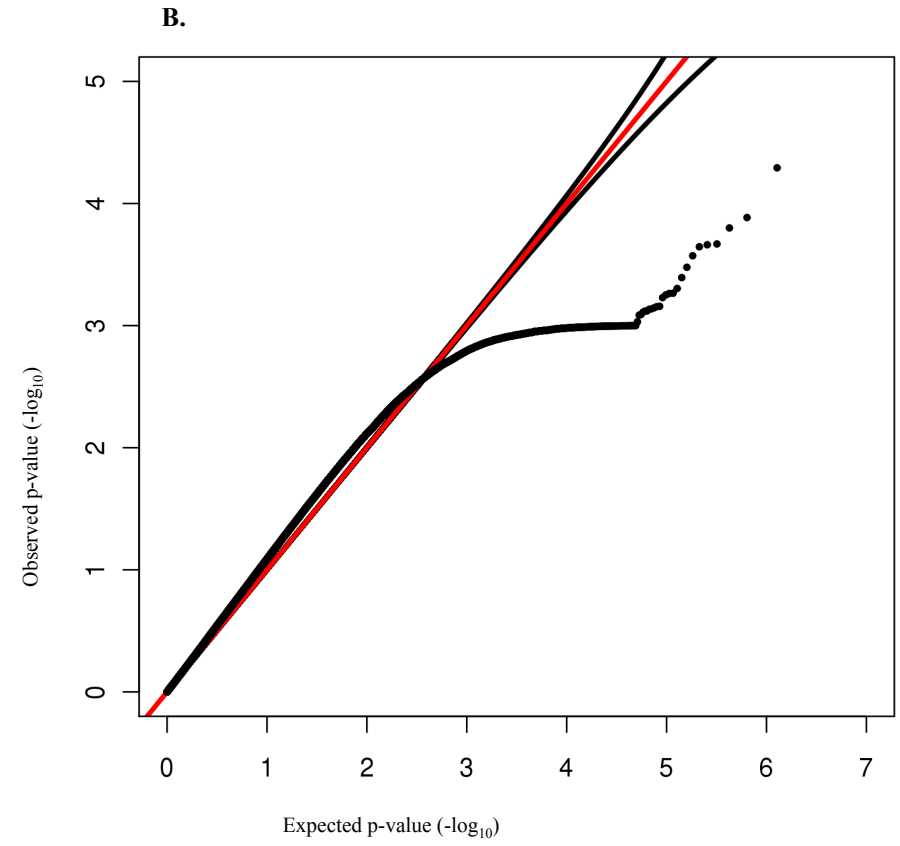
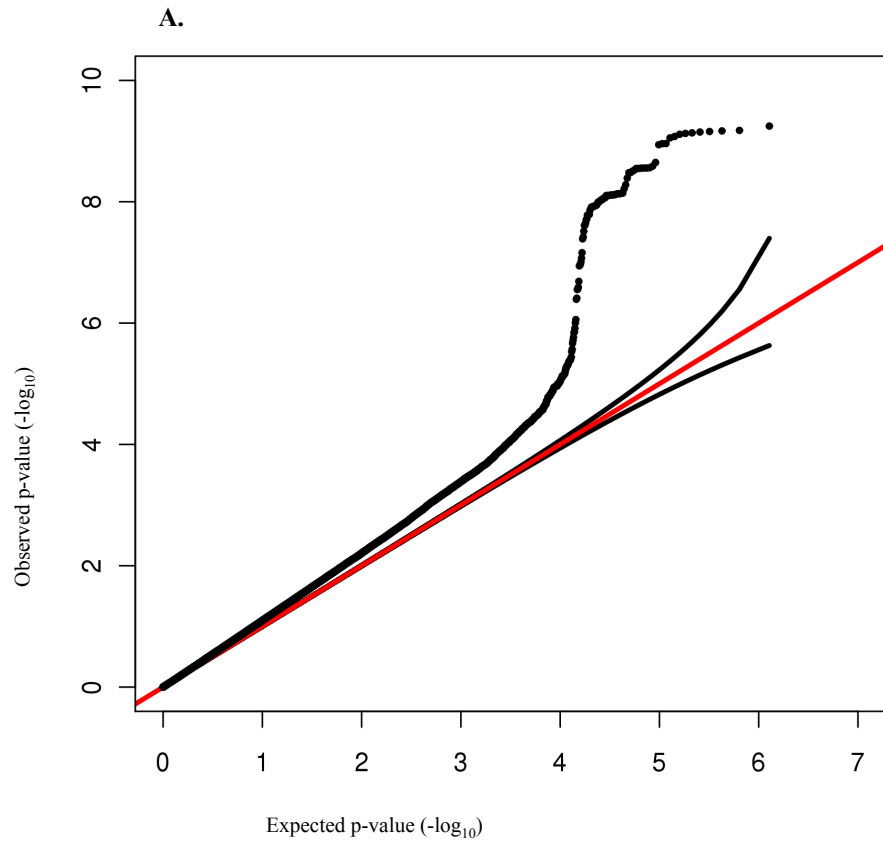
W.



X.

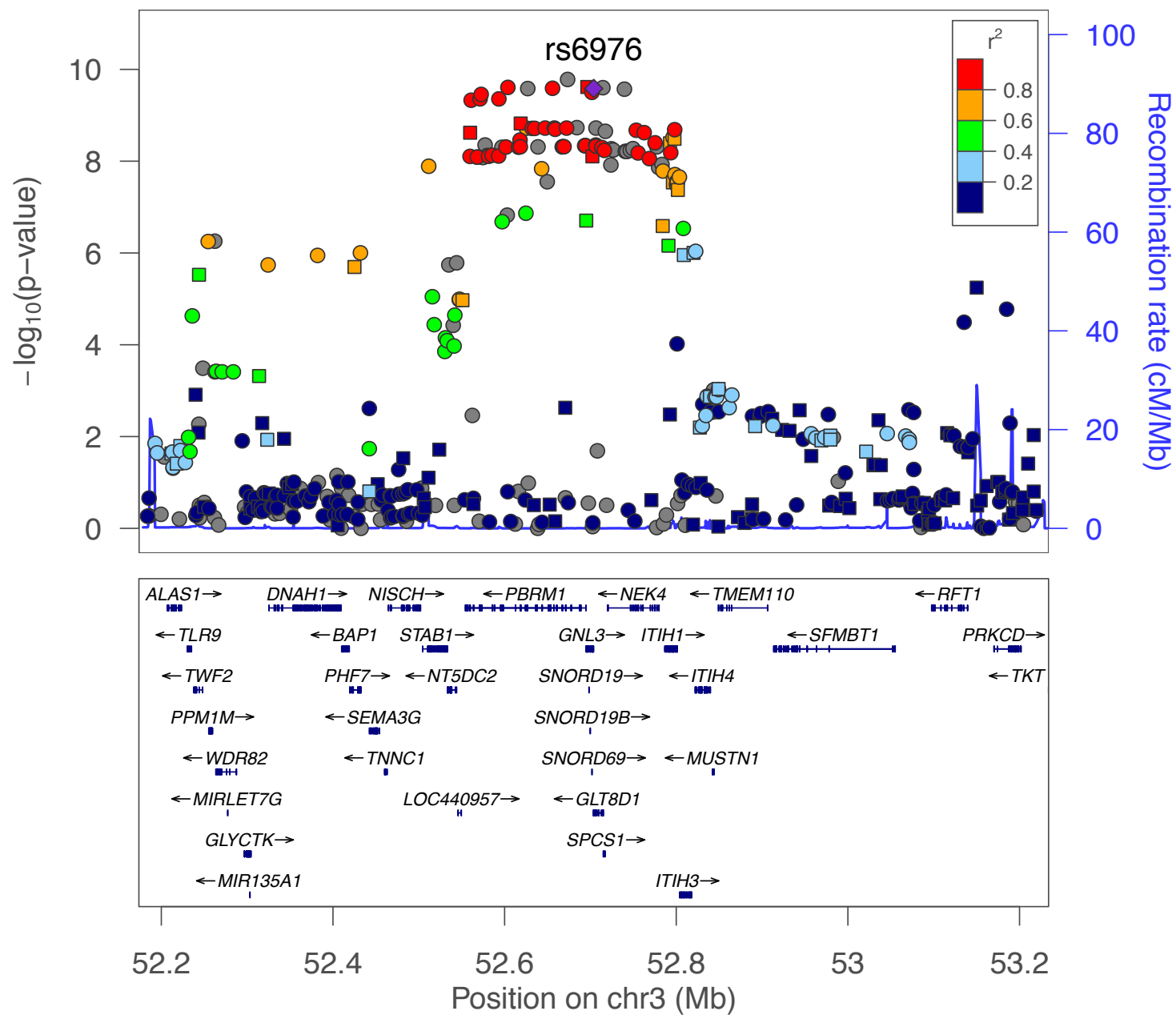


SF3



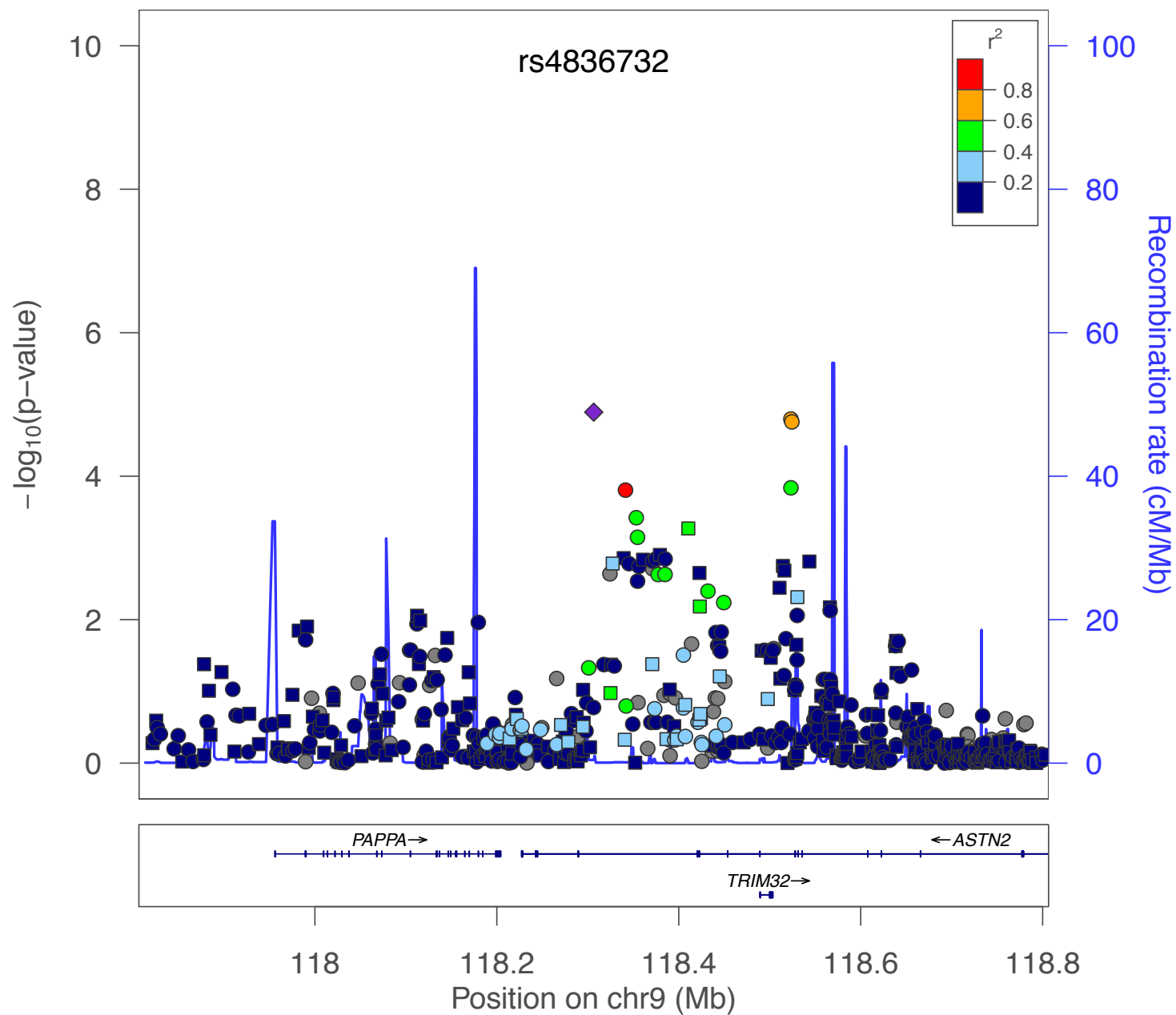
SF4 a)

Plotted SNPs



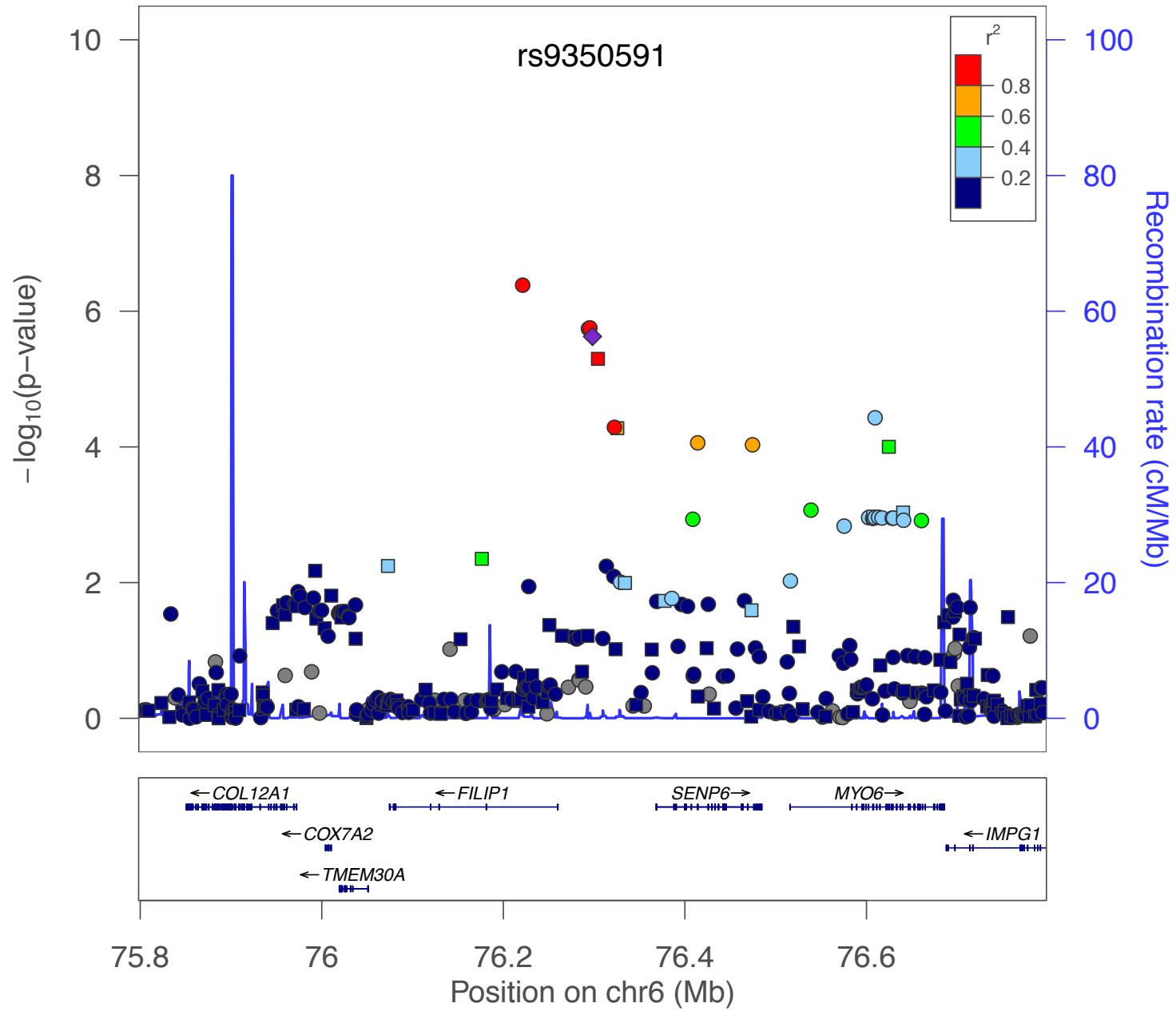
b)

Plotted SNPs



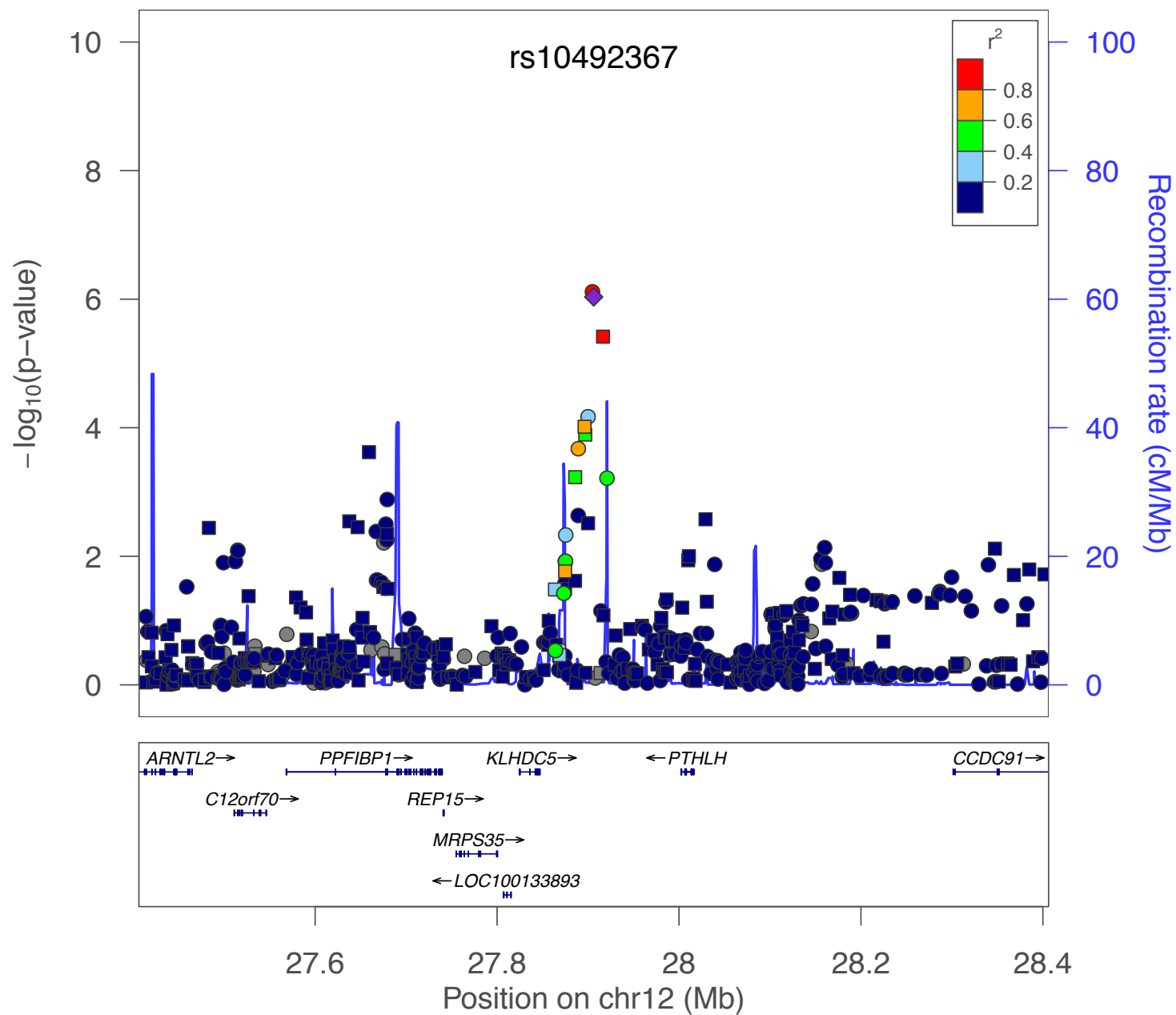
c)

Plotted SNPs



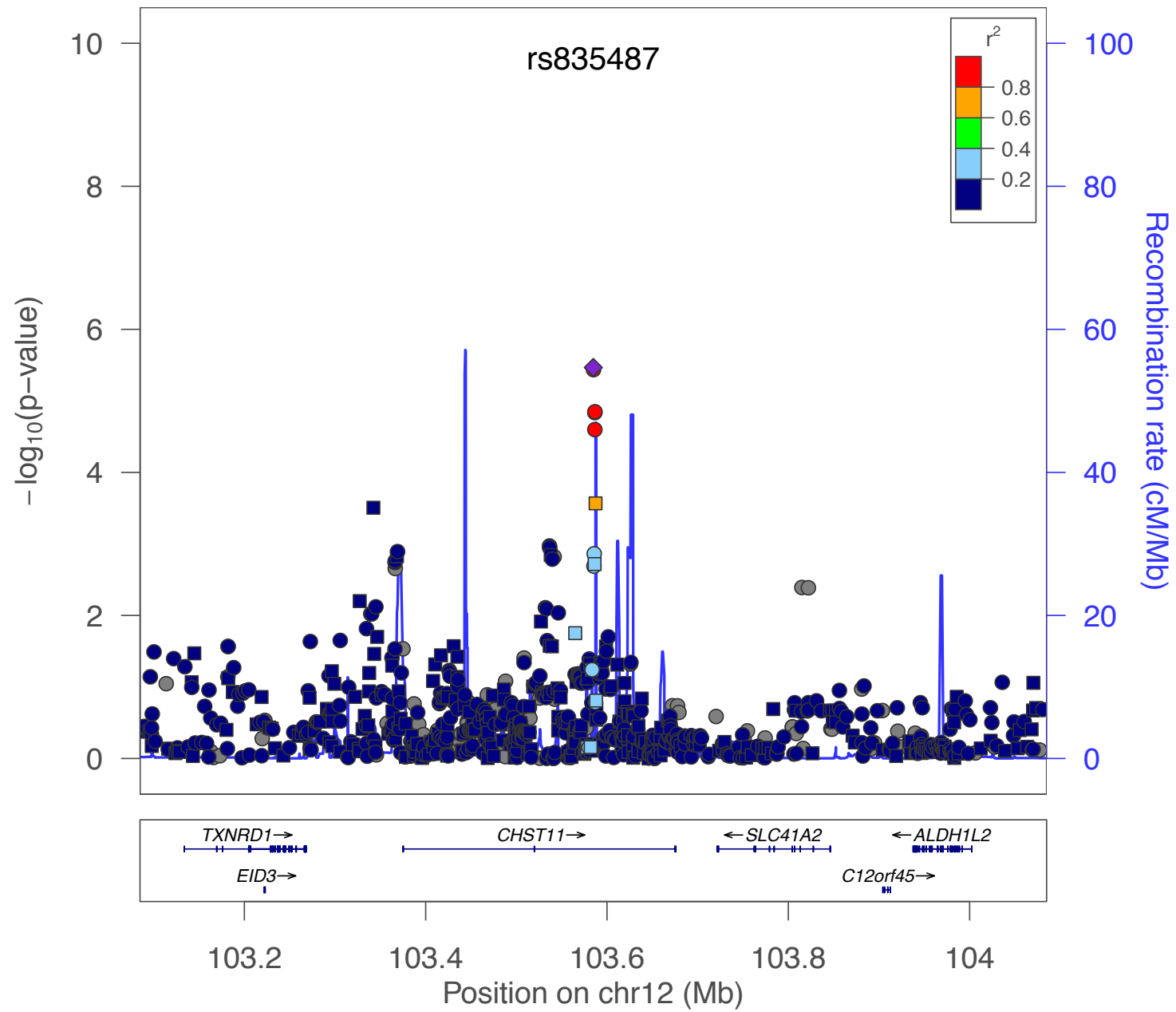
d)

Plotted SNPs



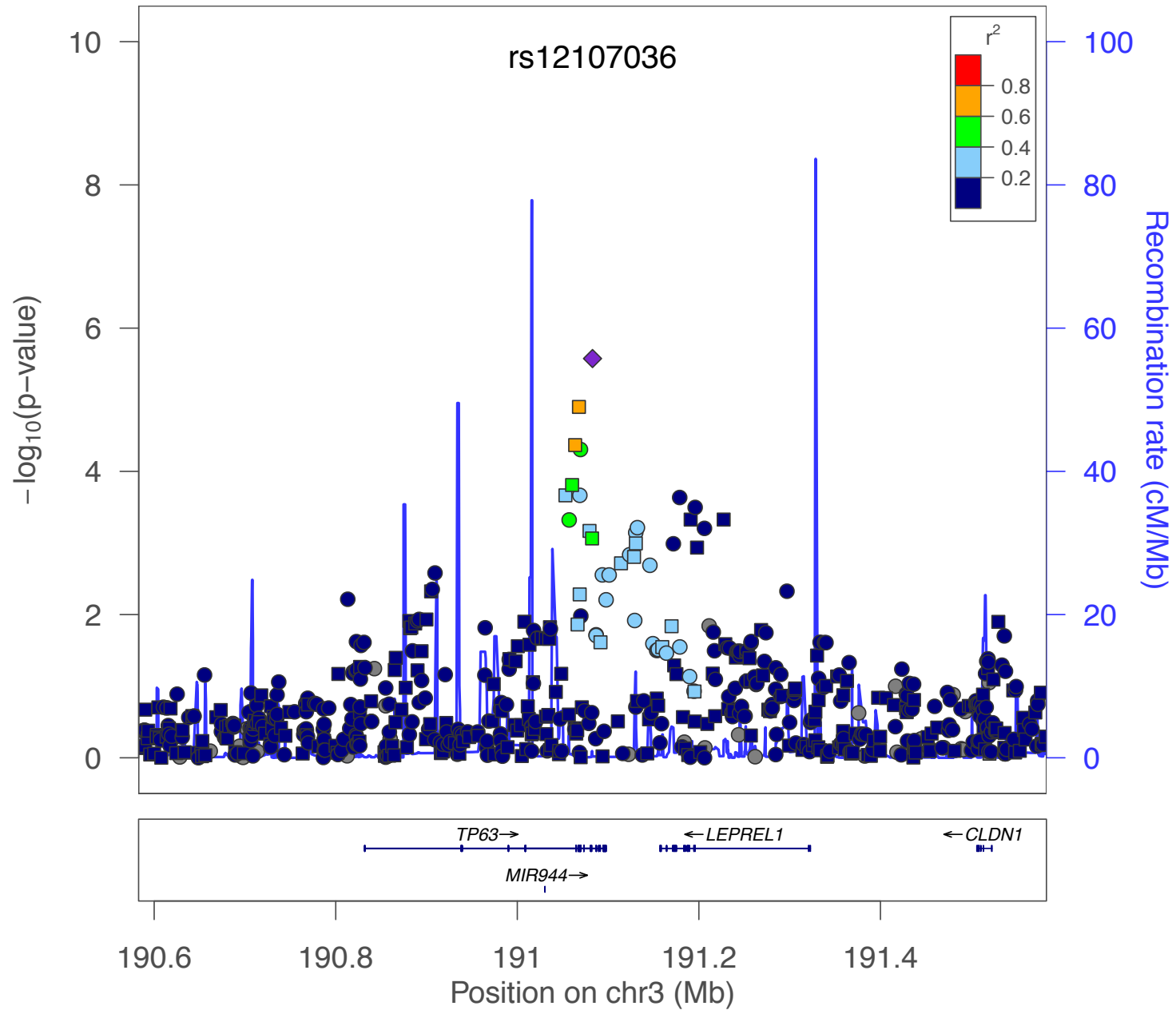
e)

Plotted SNPs



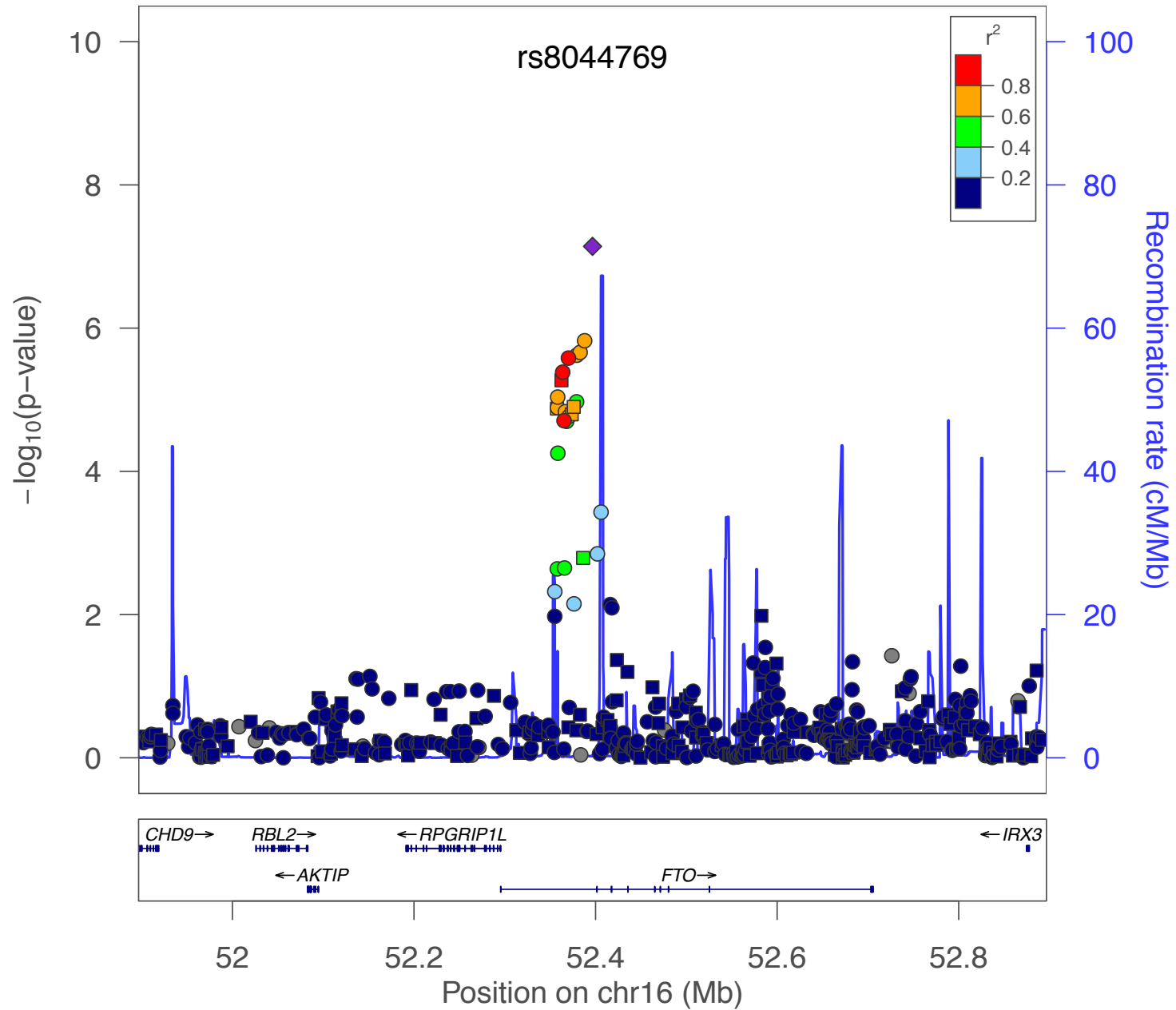
f)

Plotted SNPs



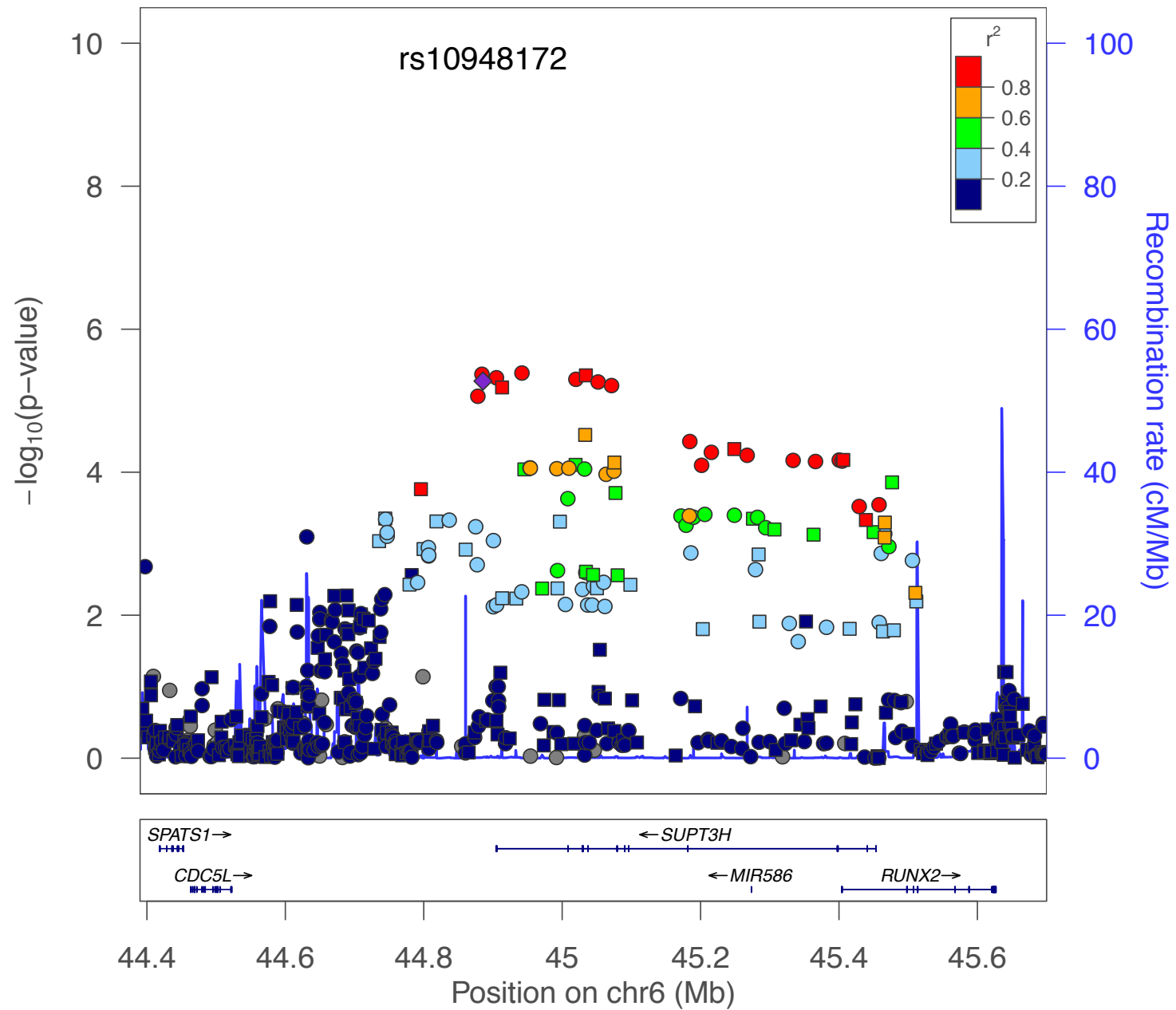
g)

Plotted SNPs

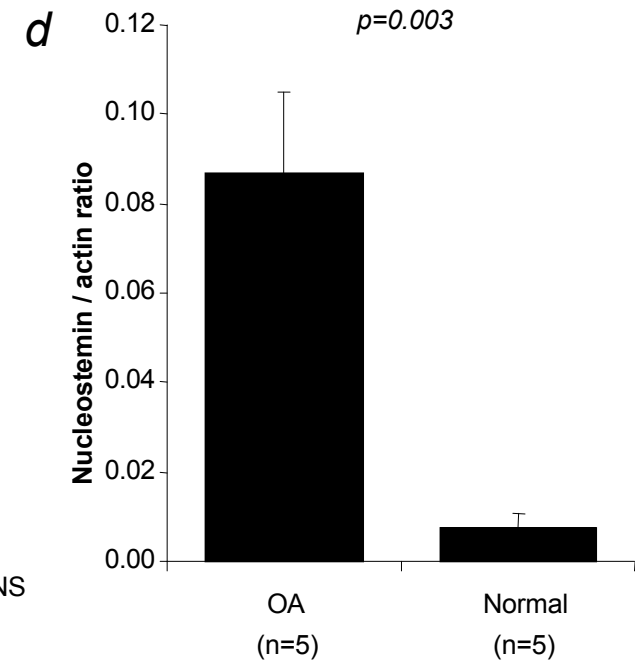
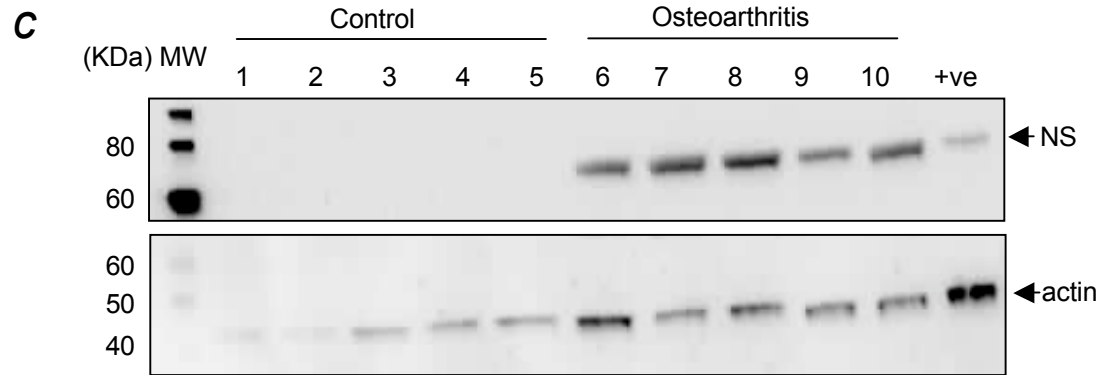
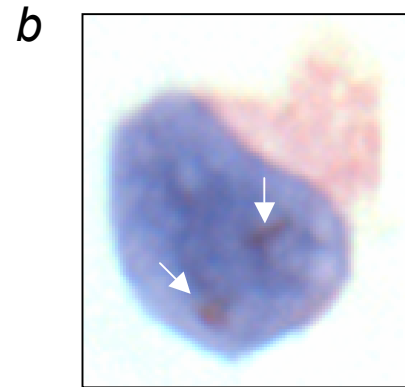
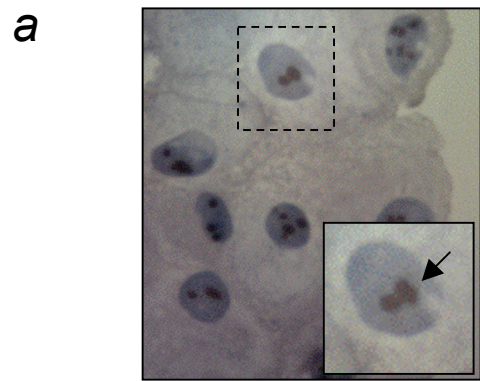


h)

Plotted SNPs 

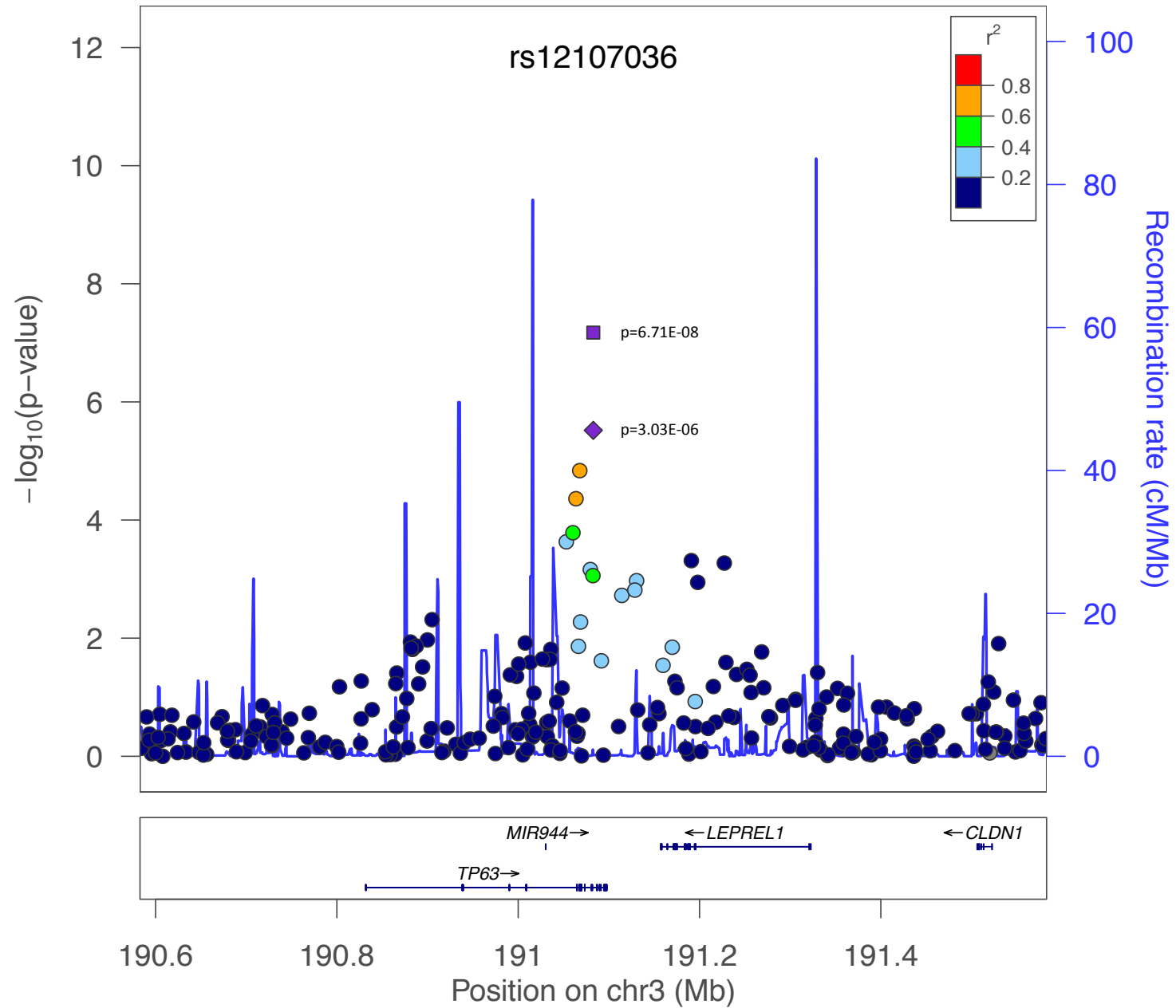


SF5



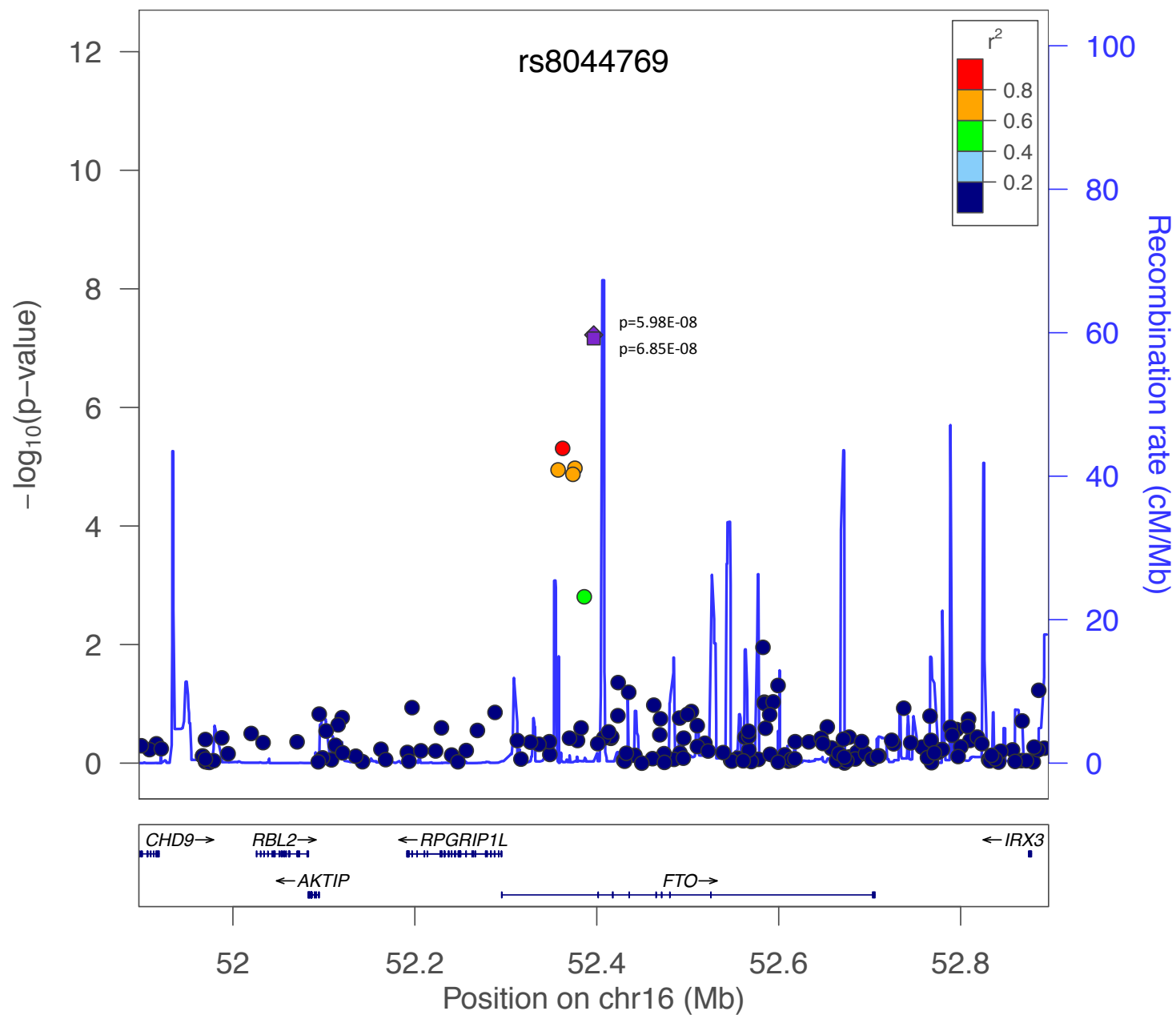
SF6 a)

Plotted SNPs



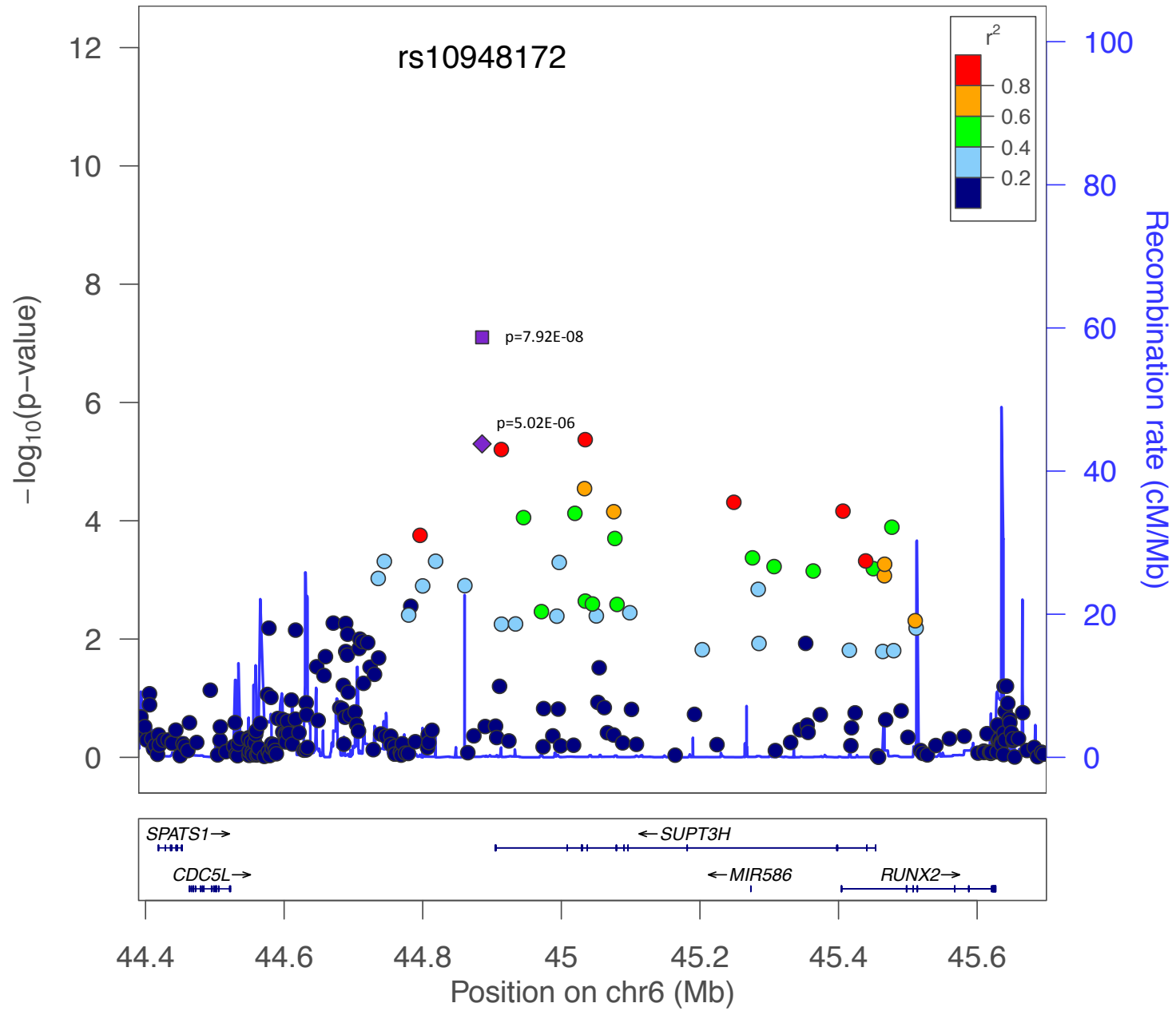
b)

Plotted SNPs



c)

Plotted SNPs



ST1. Population-based controls used in the arcOGEN GWAS.

Cohort	Platform	N samples pre QC	N samples post QC
WTCCC2-UKBS	Illumina Human1-2M-DuoCustom	2737	2501
WTCCC2-58BC	Illumina Human1-2M-DuoCustom	2930	2699
T1DGC-58BC	Illumina HumanHap550-2	2596	2530
ALSPAC	Illumina Human670-QuadCustom	772	743
PoBI	Illumina Human1-2M-DuoCustom	2912	2536
All		11947	11009

ST2. λ_{1000} values for the unadjusted analyses and the analyses adjusted for the first 10 principal components (PCs).

Stratum	λ_{1000} unadjusted	λ_{1000} adjusted
All OA	1.0091	1.0082
Male OA	1.0097	1.0091
Female OA	1.0119	1.0101
TJR	1.01	1.0095
Male TJR	1.0089	1.0092
Female TJR	1.012	1.011
Hip OA	1.011	1.0101
Male hip OA	1.0117	1.0121
Female hip OA	1.0129	1.0119
THR	1.0118	1.0111
Male THR	1.0121	1.0126
Female THR	1.0131	1.0116
Knee OA	1.013	1.0104
Male knee OA	1.0142	1.0119
Female knee OA	1.0127	1.0107
TKR	1.0133	1.013
Male TKR	1.0155	1.0149
Female TKR	1.0145	1.0145

ST3. Association p-values before and after adjustment for the first 10 principal components.

Chr	SNP	Position	P unadjusted	P adjusted	P adj/P unadj
3	rs69763 ¹	52703844	2.27E-10	5.72E-10	2.52
3	rs111773 ¹	52696345	2.12E-10	6.02E-10	2.84
9	rs4836732	118306516	1.19E-05	1.42E-05	1.19
6	rs9350591	76298247	2.49E-06	4.63E-06	1.86
12	rs10492367	27906237	1.20E-06	9.76E-07	0.81
12	rs835487	103584897	3.26E-06	3.00E-06	0.92
3	rs12107036	191082854	3.03E-06	3.98E-06	1.31
16	rs8044769	52396636	5.98E-08	1.27E-07	2.12
6	rs10948172	44885669	5.02E-06	4.47E-06	0.89

¹ Both SNPs represent the same signal, $r^2=1$.

ST4. Association summary statistics and genotype counts for the SNPs prioritised for replication for the stratum where the lowest p-value was observed.

Chr	SNP	Position	EA ¹	NEA ²	EAF ³ cases	EAF controls	OR	OR 95L	OR 95U	P	Stratum	Genotype counts ⁴ cases	Genotype counts controls
1	rs1854169	102046864	C	T	0.192	0.2113	0.89	0.84	0.93	6.53E-06	All	296/2251/4855	505/3642/6859
1	rs1417066	217740451	G	A	0.4628	0.4364	1.11	1.06	1.16	3.90E-06	TJR	1234/2902/1666	2084/5432/3482
1	rs1629896	223752551	G	A	0.09447	0.07824	1.23	1.13	1.34	4.82E-06	Knee**	53/677/3414	65/1592/9347
1	rs987179	242876559	A	G	0.2759	0.3078	0.86	0.8	0.92	6.23E-06	TJR Female	265/1325/1772	500/2390/2617
2	rs1024816	37454712	G	A	0.1045	0.07534	1.43	1.24	1.66	1.72E-06	TKR Female	9/236/970	39/753/4723
2	rs3791679	55950396	G	A	0.1944	0.2379	0.77	0.7	0.86	1.71E-06	Hip Male	46/426/860	330/1954/3210
2	rs10779958	74609286	C	A	0.1677	0.1379	1.26	1.14	1.39	2.97E-06	Knee Female	58/600/1477	113/1295/4106
2	rs6754683	108522707	A	G	0.3791	0.4265	0.82	0.75	0.9	8.84E-06	Hip Male	183/644/505	981/2723/1789
2	rs4622728	125523313	T	C	0.4467	0.4849	0.86	0.8	0.92	6.22E-06	Knee** Female	530/1211/801	1300/2746/1466
2	rs1432237	137524977	G	A	0.1672	0.2068	0.77	0.69	0.86	4.44E-06	Hip Male	36/373/922	211/1849/3432
2	rs12618428	151441225	G	A	0.1128	0.1341	0.82	0.76	0.89	7.72E-07	Knee**	60/814/3267	192/2566/8243
2	rs6716254	175250720	C	T	0.2846	0.3207	0.84	0.78	0.91	8.05E-06	Hip** Female	173/985/1180	539/2459/2517
3	rs6793234	8801396	G	A	0.1305	0.1023	1.32	1.17	1.48	6.86E-06	Knee** Male	26/366/1210	63/998/4430
3	rs11177	52696345	A	G	0.4105	0.375	1.16	1.11	1.22	2.12E-10	TJR	998/2768/2037	1547/5154/4297
3	rs6976	52703844	T	C	0.4103	0.3748	1.16	1.11	1.22	2.27E-10	TJR	998/2763/2039	1546/5157/4301
3	rs10510816	59630570	A	C	0.4243	0.4667	0.84	0.78	0.9	2.34E-06	Knee Female	376/1059/699	1188/2769/1555
3	rs1355782	132970356	C	T	0.3435	0.375	0.87	0.83	0.92	3.83E-07	Knee**	490/1866/1787	1553/5150/4305
3	rs16851066	142325511	A	G	0.07696	0.05761	1.36	1.19	1.56	5.37E-06	Hip** Female	9/342/1988	16/603/4892
3	rs9867979*	166345704	A	G	0.113	0.07772	1.51	1.27	1.8	3.06E-06	TKR Female	16/243/958	15/254/1558
3	rs12107036	191082854	A	G	0.4269	0.479	0.81	0.74	0.89	3.03E-06	TKR Female	210/619/388	1255/2772/1486
4	rs16837352	5555175	T	G	0.1647	0.1434	1.18	1.1	1.26	3.78E-06	Knee**	108/1148/2886	230/2697/8079
4	rs17578878	37577120	T	C	0.09631	0.1161	0.81	0.75	0.88	9.87E-07	Knee**	30/738/3375	140/2276/8592
4	rs11725992	48082608	C	T	0.06502	0.0774	0.83	0.76	0.9	7.01E-06	All	31/901/6473	69/1565/9367
4	rs6842739*	60172105	A	G	0.2697	0.3156	0.8	0.73	0.88	7.19E-06	Knee Female	146/858/1128	178/798/852
4	rs2626053*	96474111	A	G	0.1487	0.1986	0.71	0.61	0.81	6.54E-07	TKR Female	31/300/886	73/580/1175
4	rs2636726*	106528673	A	G	0.4801	0.4365	1.19	1.1	1.29	8.44E-06	All Female	1068/2160/1246	350/894/582

4	rs12500935	114760230	A	G	0.1068	0.09227	1.18	1.1	1.26	4.10E-06	All	73/1437/5899	95/1841/9070
4	rs1566347	186969560	T	C	0.3124	0.3443	0.87	0.82	0.92	9.15E-07	Knee	344/1497/1656	1278/5021/4706
5	rs4957048	636442	A	G	0.185	0.2261	0.78	0.7	0.86	3.42E-06	Knee Male	52/400/910	260/1963/3269
5	rs1083523*	66314669	C	T	0.4896	0.4436	1.2	1.11	1.3	2.75E-06	All Female	1102/2175/1195	349/923/555
5	rs2610424	71292415	C	T	0.4661	0.4372	1.12	1.07	1.18	6.62E-06	Knee**	908/2045/1189	2108/5408/3490
5	rs457008	81110411	T	C	0.4574	0.4893	0.88	0.83	0.93	3.37E-06	Knee	721/1757/1019	2638/5483/2874
5	rs10036746	91780688	A	G	0.1152	0.08752	1.36	1.19	1.56	9.86E-06	Hip Male	22/263/1047	56/849/4585
5	rs11135394*	93212891	C	T	0.2138	0.2519	0.81	0.73	0.89	9.46E-06	TJR Female	154/1129/2078	121/679/1028
5	rs12515798*	134333096	T	C	0.2136	0.166	1.37	1.2	1.56	2.78E-06	TKR Female	50/420/747	53/501/1274
5	rs10515550	144792886	T	C	0.4126	0.4567	0.84	0.77	0.9	9.93E-06	Knee** male	270/781/550	1189/2639/1665
5	rs13167773	148567182	T	C	0.1104	0.1384	0.77	0.69	0.86	3.87E-06	Knee Female	22/427/1685	113/1299/4097
5	rs10051783	167755230	T	C	0.1649	0.1379	1.23	1.13	1.35	6.94E-06	Knee** Female	64/710/1767	96/1329/4088
6	rs10948172	44885669	G	A	0.3265	0.2925	1.17	1.1	1.26	5.02E-06	All Male	337/1239/1354	452/2308/2731
6	rs9350591	76298247	T	C	0.1337	0.1136	1.2	1.11	1.3	2.49E-06	Hip**	61/923/2924	143/2215/8647
6	rs569731*	76414313	A	G	0.1158	0.1545	0.72	0.63	0.82	1.64E-06	THR Female	18/373/1375	40/484/1301
6	rs10943837	82727239	T	C	0.3708	0.3382	1.15	1.09	1.22	2.19E-06	THR	435/1384/1220	1260/4925/4821
6	rs2493984	95440213	G	A	0.3472	0.3093	1.19	1.11	1.28	3.25E-06	Hip** Female	292/1041/1007	541/2328/2643
6	rs7755798	109743847	T	C	0.1813	0.1636	1.13	1.07	1.2	9.87E-06	All	264/2156/4982	307/2975/7686
6	rs2027532	112741889	C	T	0.2581	0.2191	1.24	1.13	1.36	3.72E-06	Knee** male	101/625/876	247/1914/3333
6	rs6931833	114078284	A	G	0.2239	0.2522	0.86	0.8	0.92	5.92E-06	THR	166/1028/1843	679/4194/6135
7	rs17627827	16731521	G	A	0.04488	0.06843	0.64	0.53	0.77	1.71E-06	Hip** Male	3/135/1433	25/701/4761
7	rs6461679	22976588	T	C	0.3501	0.3886	0.85	0.79	0.91	2.67E-06	Knee** Female	304/1171/1066	828/2629/2056
7	rs7785659	32427540	T	C	0.1761	0.136	1.36	1.22	1.51	1.37E-08	Knee** male	46/472/1083	98/1296/4092
7	rs7805536	32540095	C	T	0.1951	0.1544	1.33	1.2	1.47	4.38E-08	Knee** male	53/519/1030	123/1450/3919
7	rs1859572	76870933	T	C	0.2393	0.2132	1.16	1.09	1.23	1.74E-06	Hip**	223/1426/2262	491/3709/6799
7	rs6466265	77998184	G	A	0.2288	0.1878	1.28	1.16	1.43	3.72E-06	TKR Female	52/453/712	222/1626/3664
7	rs1488517*	95446689	C	T	0.3127	0.2582	1.31	1.17	1.46	3.56E-06	TKR Female	114/533/570	121/702/1005
7	rs2966417	109944782	A	G	0.2036	0.2329	0.84	0.79	0.9	6.57E-07	Hip	125/1079/2060	594/3937/6471

8	rs4841020	8638760	G	A	0.1516	0.1814	0.81	0.74	0.88	3.04E-06	Knee** Female	64/642/1834	192/1617/3705
8	rs1876836	8719166	T	C	0.224	0.2542	0.85	0.79	0.9	6.59E-07	All Female	227/1551/2698	372/2059/3082
8	rs4841067	8784655	G	A	0.13	0.1593	0.79	0.71	0.87	5.40E-06	Knee Female	39/477/1619	149/1458/3903
8	rs609792	9857300	T	C	0.2465	0.2099	1.23	1.13	1.35	4.61E-06	THR Female	112/648/1009	238/1838/3437
8	rs6997710	10247861	G	A	0.1061	0.08732	1.24	1.13	1.36	6.92E-06	All Female	49/852/3574	34/895/4585
8	rs1421259	15607043	A	G	0.3792	0.3486	1.14	1.08	1.21	5.74E-06	Hip	487/1503/1276	1317/5032/4646
8	rs2979715*	80724814	C	T	0.236	0.1874	1.34	1.2	1.5	4.63E-07	THR Female	85/663/1017	71/542/1212
8	rs2467753*	80759808	G	T	0.188	0.1481	1.33	1.19	1.5	1.54E-06	Hip** female	74/732/1535	44/453/1330
8	rs2289496	98357340	C	A	0.3571	0.3894	0.87	0.82	0.92	4.78E-06	THR	385/1400/1253	1658/5253/4093
8	rs16870112	104191422	A	C	0.09592	0.08131	1.2	1.11	1.3	5.85E-06	TJR	59/995/4748	90/1609/9302
8	rs10956114*	124156008	G	A	0.333	0.3846	0.8	0.73	0.88	3.09E-06	Hip Female	207/874/853	246/914/668
8	rs4072286	142805713	G	T	0.3905	0.3495	1.19	1.12	1.27	3.92E-08	TJR Female	504/1616/1240	683/2488/2342
9	rs2146423	4647040	G	A	0.4101	0.4519	0.84	0.78	0.91	2.88E-06	Knee Female	375/1001/759	1149/2686/1679
9	rs945442	18917768	C	A	0.2505	0.2221	1.17	1.1	1.25	1.58E-06	Hip	213/1209/1841	559/3770/6674
9	rs3780296	79327254	C	T	0.4344	0.3981	1.16	1.09	1.24	8.97E-06	TKR	431/1018/715	1747/5270/3989
9	rs4072357*	89688030	T	C	0.4071	0.4546	0.82	0.76	0.89	2.90E-06	TJR Female	546/1642/1170	379/904/545
9	rs4836732	118306516	C	T	0.5081	0.467	1.18	1.1	1.26	2.48E-06	Hip** Female	602/1169/564	1209/2710/1571
9	rs719535*	118516091	T	C	0.2653	0.2213	1.27	1.14	1.41	9.02E-06	Hip Female	153/720/1061	91/627/1110
9	rs2900277	131434091	G	T	0.2193	0.2438	0.87	0.82	0.93	8.44E-06	Knee**	194/1429/2519	676/4013/6316
10	rs2474714*	33535989	G	A	0.4022	0.451	0.82	0.75	0.89	7.56E-06	Hip** Female	365/1153/823	365/918/544
10	rs9299558	56603927	A	C	0.436	0.491	0.8	0.73	0.88	9.92E-06	TKR Male	178/468/299	1314/2757/1413
10	rs1250552*	80728033	G	A	0.4984	0.4505	1.21	1.12	1.31	3.13E-06	TJR Female	819/1711/830	369/909/550
10	rs787640	95114663	G	A	0.3064	0.2654	1.22	1.13	1.33	9.23E-07	Hip Female	168/849/917	391/2145/2979
10	rs11188412	97391986	A	G	0.05532	0.07134	0.76	0.68	0.85	3.24E-06	Knee	7/373/3118	54/1462/9487
11	rs4757420	12371653	T	G	0.1543	0.1783	0.84	0.78	0.91	6.20E-06	All Female	116/1149/3211	163/1639/3709
11	rs7119797	61922811	T	C	0.4559	0.4957	0.85	0.8	0.91	5.06E-06	Hip** Female	464/1204/670	1352/2761/1399
11	rs17159801	82496946	A	G	0.1332	0.1131	1.21	1.12	1.3	2.34E-06	Hip**	68/906/2938	145/2198/8656
11	rs2691829	90240224	T	C	0.4096	0.3866	1.1	1.06	1.15	9.38E-06	All	1238/3593/2577	1680/5149/4176

11	rs6589848	120229228	T	C	0.1668	0.2005	0.8	0.72	0.88	9.36E-06	THR Female	48/494/1227	222/1766/3523
11	rs7941193	125876053	A	G	0.215	0.1712	1.33	1.19	1.48	3.26E-07	TKR Female	66/391/759	151/1585/3776
11	rs7122854	132673733	G	A	0.205	0.1647	1.31	1.18	1.46	8.08E-07	Hip Male	55/436/841	142/1524/3823
12	rs7305794	24860894	T	C	0.2926	0.2642	1.15	1.08	1.23	8.26E-06	All Female	396/1825/2251	379/2153/2977
12	rs10492367	27906237	T	G	0.2157	0.1873	1.19	1.11	1.28	7.00E-07	THR	117/1076/1843	366/3392/7249
12	rs6487684	28534480	T	C	0.2646	0.2223	1.26	1.14	1.39	7.35E-06	TKR Female	85/474/658	273/1905/3334
12	rs1034762	46675910	A	C	0.1701	0.1392	1.27	1.16	1.39	4.36E-07	TJR Male	64/701/1672	106/1317/4071
12	rs11107957*	76956186	C	A	0.2882	0.2429	1.26	1.15	1.38	7.19E-07	TJR Female	279/1380/1703	130/628/1070
12	rs1404866	83888610	T	G	0.2581	0.23	1.17	1.1	1.24	3.04E-07	Knee**	292/1551/2293	565/3932/6508
12	rs11107219*	92770487	G	A	0.08326	0.05935	1.44	1.22	1.69	9.81E-06	TJR Female	18/524/2821	6/205/1617
12	rs835487	103584897	G	A	0.3768	0.3446	1.15	1.08	1.22	3.26E-06	THR	420/1450/1169	1330/4924/4751
12	rs11064722	118157566	T	C	0.2148	0.194	1.14	1.08	1.2	6.16E-06	TJR	296/1901/3605	424/3421/7155
13	rs4544137*	67072419	T	C	0.04925	0.07467	0.64	0.53	0.77	2.44E-06	Knee Female	1/208/1923	6/261/1561
13	rs7994526*	67263831	G	A	0.05298	0.0796	0.65	0.54	0.77	1.74E-06	Knee Female	5/216/1912	8/275/1545
13	rs2036717*	75645031	C	T	0.1116	0.08096	1.43	1.22	1.67	7.04E-06	Hip Female	20/391/1520	10/276/1542
13	rs9520058	105855421	G	T	0.344	0.3709	0.89	0.85	0.93	1.01E-06	TJR	681/2629/2491	1533/5100/4374
14	rs8015303*	34316223	G	A	0.08628	0.05689	1.57	1.28	1.91	8.89E-06	TKR Female	6/198/1013	6/196/1626
14	rs1627411*	42241078	T	C	0.1529	0.1958	0.74	0.66	0.84	8.83E-07	Hip Female	41/509/1383	77/562/1189
14	rs7155791	49487980	T	C	0.1486	0.1229	1.25	1.13	1.37	9.52E-06	TJR Male	57/611/1772	93/1164/4236
14	rs1998094*	51653209	C	T	0.1343	0.09907	1.41	1.23	1.62	7.91E-07	Hip** Female	52/525/1764	18/326/1483
14	rs10144429*	77722272	A	G	0.3439	0.2899	1.28	1.15	1.43	8.48E-06	TKR Female	119/599/499	149/762/917
15	rs11634255	30874764	G	A	0.4742	0.4354	1.17	1.1	1.25	1.34E-06	All Male	663/1455/814	1015/2752/1725
15	rs4646626	56043419	C	T	0.485	0.4546	1.13	1.07	1.19	8.86E-06	Knee	810/1770/915	2303/5399/3303
15	rs8032755	91403060	A	G	0.3863	0.4231	0.86	0.8	0.92	7.40E-06	TKR	313/1046/805	1951/5408/3644
16	rs2908646	4971845	A	G	0.5275	0.4801	1.21	1.11	1.32	9.17E-06	Knee Male	386/666/311	1277/2720/1496
16	rs12596722	9324789	C	A	0.3517	0.3205	1.15	1.09	1.22	1.23E-06	Knee	455/1548/1491	1121/4813/5073
16	rs12923310	9695438	T	G	0.3274	0.369	0.83	0.78	0.89	7.45E-08	All Male	313/1293/1325	721/2611/2160
16	rs8044769	52396636	T	C	0.4592	0.4977	0.86	0.81	0.91	5.98E-08	All Female	969/2171/1334	1368/2753/1393

16	rs9937815	54888333	G	A	0.3451	0.3177	1.13	1.07	1.2	8.75E-06	Hip**	473/1752/1684	1119/4753/5131
16	rs7404629	64970230	A	G	0.1958	0.2243	0.84	0.78	0.91	7.23E-06	TJR Female	124/1068/2168	302/1869/3342
16	rs16947129*	76709247	C	T	0.2006	0.1587	1.33	1.18	1.5	2.39E-06	Hip Female	75/625/1232	49/482/1296
17	rs4078062	16919559	A	G	0.1354	0.1111	1.25	1.13	1.39	9.41E-06	Knee** Female	48/592/1901	65/1095/4354
17	rs8080960	67298538	G	A	0.4533	0.4185	1.15	1.08	1.23	5.75E-06	TJR Female	712/1624/1026	963/2688/1861
18	rs11665347	10468812	A	G	0.2915	0.3133	0.9	0.86	0.94	7.97E-06	All	627/3064/3716	1092/4711/5200
18	rs12454396*	62299667	T	C	0.335	0.3769	0.83	0.77	0.9	7.47E-06	All Female	514/1968/1989	268/842/718
18	rs17089382	70369681	T	C	0.1267	0.1056	1.23	1.13	1.34	3.28E-06	THR	59/652/2327	135/2054/8813
19	rs12974139	2135005	C	T	0.3795	0.4305	0.81	0.74	0.88	1.69E-06	Hip Male	186/639/507	1030/2670/1793
19	rs1610093*	4505774	C	T	0.06437	0.09382	0.66	0.56	0.79	2.13E-06	Hip Female	7/235/1692	10/323/1495
19	rs10426443	8017982	G	A	0.1464	0.1845	0.76	0.67	0.86	8.37E-06	TKR Female	32/292/892	179/1677/3658
19	rs10426377	53784046	A	C	0.2706	0.2405	1.17	1.09	1.26	7.29E-06	TJR Female	232/1355/1774	336/1978/3196
20	rs1360400	22909890	A	G	0.4655	0.4254	1.18	1.1	1.26	7.22E-06	Knee Female	458/1068/605	1012/2667/1835
20	rs172981*	40543116	C	T	0.06409	0.09929	0.62	0.51	0.76	1.45E-06	TKR Female	5/146/1066	18/327/1483
20	rs17785895	47866466	G	A	0.1016	0.1211	0.82	0.75	0.9	9.41E-06	Knee	35/640/2820	162/2340/8499
20	rs4925370	60300986	G	A	0.4866	0.456	1.13	1.07	1.19	8.11E-06	Knee	821/1760/915	2303/5428/3271
22	rs3747081	19682120	A	G	0.1223	0.1022	1.22	1.12	1.34	6.81E-06	All Female	70/955/3451	57/1013/4444
22	rs7287616	22259852	T	C	0.2307	0.259	0.86	0.8	0.92	7.52E-06	THR	169/1064/1805	726/4246/6030
22	rs5762347	26489672	A	G	0.3661	0.3193	1.23	1.13	1.35	5.98E-06	THR Male	174/582/514	548/2411/2532

¹ EA Effect allele.

² NEA Non effect allele.

The bold allele is the risk allele.

³ EAF Effect allele frequency.

⁴ Genotype counts: 11/12/22, where 1 is the EA and 2 is the NEA.

* Signals prioritised from the OA vs TwinsUK analyses.

** Indicates analyses where a subset of samples with OA at the hip and the knee are also included in the arcOGEN discovery set and in the UK replication set.

Meta areOEN_insilco_denovo_random

Table with 12 columns: OR, L95, U95, p, q_p, i2. Contains 100 rows of numerical data.

Meta insilco_denovo_fixed

Table with 12 columns: OR, L95, U95, p, q_p, i2. Contains 100 rows of numerical data.

Meta insilco_denovo_random

Table with 12 columns: OR, L95, U95, p, q_p, i2. Contains 100 rows of numerical data.

Meta TRJ_areOEN_insilco_denovo_fixed

Table with 12 columns: OR, L95, U95, p, q_p, i2. Contains 100 rows of numerical data.

Meta TRJ_areOEN_insilco_denovo_random

Table with 12 columns: OR, L95, U95, p, q_p, i2. Contains 100 rows of numerical data.

Meta TRJ_insilco_denovo_fixed

Table with 12 columns: OR, L95, U95, p, q_p, i2. Contains 100 rows of numerical data.

Meta TRJ_insilco_denovo_random

Table with 12 columns: OR, L95, U95, p, q_p, i2. Contains 100 rows of numerical data.

Meta TRJ_insilco_denovo_random

Table with 12 columns: OR, L95, U95, p, q_p, i2. Contains 100 rows of numerical data.

Meta TRJ_insilco_denovo_random

Table with 12 columns: OR, L95, U95, p, q_p, i2. Contains 100 rows of numerical data.

Meta TRJ_insilco_denovo_random

Table with 12 columns: OR, L95, U95, p, q_p, i2. Contains 100 rows of numerical data.

Table with columns: Meta arCGEN insulino denovo random, Meta insulino denovo fixed, Meta insulino denovo random, Meta TIR arCGEN insulino denovo fixed, Meta TIR arCGEN insulino denovo random, Meta TIR insulino denovo fixed. Each column contains a grid of numerical data points.

Meta insidulo denovo fixed						Meta TJR aROGEN insidulo denovo fixed						Meta TJR insidulo denovo random						Meta TJR insidulo denovo fixed											
OR	L95	U95	p	q.p	z2	OR	L95	U95	p	q.p	z2	OR	L95	U95	p	q.p	z2	OR	L95	U95	p	q.p	z2						
0.95	0.91	1.00	2.59E-01	3.08E-01	0.166	0.91	0.85	0.97	1.02E-02	1.27E-01	0.570	0.95	0.88	1.03	7.85E-01	0.570	0.91	0.85	1.03	7.85E-01	0.570	0.95	0.88	1.03	7.85E-01	0.570			
1.06	1.00	1.13	5.43E-02	1.17E-01	0.000	1.07	1.02	1.13	6.14E-03	8.67E-01	0.000	1.06	1.08	1.14	1.37E-01	9.45E-01	0.000	1.06	1.08	1.14	1.37E-01	9.45E-01	0.000	1.06	1.08	1.14	1.37E-01	9.45E-01	0.000
0.94	0.82	1.08	3.76E-01	5.15E-01	0.494	0.94	0.74	1.19	5.87E-01	1.15E-01	0.494	1.08	0.98	1.19	1.40E-01	1.35E-01	0.552	1.06	0.90	1.25	4.79E-01	1.35E-01	0.552	1.06	0.90	1.15	6.69E-01	1.00E+00	na
1.05	0.97	1.14	2.34E-01	2.68E-01	0.000	1.11	1.05	1.18	8.60E-03	9.30E-01	0.000	1.05	1.11	1.18	6.04E-01	0.91	1.12	1.05	1.12	8.80E-01	0.91	1.12	8.80E-01	0.91	1.12	8.80E-01	0.91	1.12	8.80E-01
0.99	0.85	1.15	8.69E-01	5.33E-01	0.000	0.99	0.85	1.15	8.69E-01	5.33E-01	0.000	1.09	0.98	1.20	1.21E-01	4.30E-01	0.000	1.09	0.98	1.20	1.21E-01	4.30E-01	0.000	1.01	0.81	1.25	9.51E-01	1.00E+00	0.000
1.01	0.92	1.11	7.82E-01	3.02E-01	0.178	1.01	0.90	1.11	9.82E-01	3.02E-01	0.178	0.91	0.85	0.98	1.02E-01	2.85E-03	0.888	0.95	0.76	1.19	6.45E-01	2.85E-03	0.888	0.77	0.94	1.21	2.98E-01	1.00E+00	na
1.02	0.92	1.13	6.81E-01	1.08E-01	0.552	1.03	0.87	1.22	7.63E-01	1.08E-01	0.552	1.06	0.98	1.14	1.29E-01	9.98E-01	0.000	1.06	0.93	1.21	3.76E-01	9.98E-01	0.000	1.06	0.93	1.21	3.76E-01	9.98E-01	0.000
0.92	0.81	1.09	5.18E-01	2.45E-01	0.000	0.92	0.81	1.09	5.18E-01	2.45E-01	0.000	0.92	0.81	1.09	5.18E-01	2.45E-01	0.000	0.92	0.81	1.09	5.18E-01	2.45E-01	0.000	0.92	0.81	1.09	5.18E-01	2.45E-01	0.000
0.98	0.91	1.06	6.55E-01	7.15E-02	0.572	0.98	0.86	1.13	8.02E-01	7.15E-02	0.572	0.97	0.91	1.02	2.33E-01	7.10E-01	0.000	0.97	0.91	1.02	2.33E-01	7.10E-01	0.000	0.98	0.89	1.08	7.14E-01	1.00E+00	na
1.05	0.96	1.15	3.07E-01	9.63E-01	0.000	1.05	0.96	1.15	3.07E-01	9.63E-01	0.000	0.91	0.85	0.98	9.80E-03	1.06E-02	0.847	0.94	0.77	1.14	5.22E-01	1.06E-02	0.847	1.04	0.92	1.18	5.20E-01	1.00E+00	na
1.11	1.00	1.24	6.08E-02	8.74E-01	0.000	1.11	1.00	1.24	6.08E-02	8.74E-01	0.000	0.94	0.87	1.03	1.70E-01	1.73E-03	0.898	0.98	0.75	1.30	9.06E-01	1.73E-03	0.898	1.14	0.99	1.31	7.84E-02	1.00E+00	na
1.14	1.02	1.28	2.55E-02	4.17E-01	0.000	1.14	1.02	1.28	2.55E-02	4.17E-01	0.000	1.15	1.05	1.25	1.93E-01	1.17E-01	0.147	1.15	1.05	1.26	4.07E-03	2.79E-01	0.147	1.22	1.06	1.41	6.44E-03	1.00E+00	na
1.08	1.02	1.15	1.06E-02	5.47E-01	0.000	1.08	1.02	1.15	1.06E-02	5.47E-01	0.000	1.12	1.06	1.17	2.56E-03	2.72E-01	0.140	1.11	1.05	1.18	2.05E-04	3.12E-01	0.140	1.07	0.99	1.15	9.19E-02	9.67E-01	0.000
1.04	0.92	1.16	5.47E-01	2.75E-01	0.000	1.04	0.92	1.16	5.47E-01	2.75E-01	0.000	1.12	1.06	1.17	2.56E-03	2.72E-01	0.140	1.11	1.05	1.18	2.05E-04	3.12E-01	0.140	1.07	0.99	1.15	9.19E-02	9.67E-01	0.000
1.01	0.94	1.09	7.79E-01	4.46E-01	0.000	1.01	0.94	1.09	7.79E-01	4.46E-01	0.000	1.00	0.95	1.06	8.85E-01	4.18E-01	0.000	1.00	0.95	1.06	8.85E-01	4.18E-01	0.000	0.97	0.88	1.07	5.55E-01	1.00E+00	na
0.90	0.83	0.97	5.07E-03	3.96E-01	0.000	0.90	0.83	0.97	5.07E-03	3.96E-01	0.000	0.92	0.87	0.97	4.69E-03	4.94E-01	0.000	0.92	0.87	0.97	4.69E-03	4.94E-01	0.000	0.95	0.86	1.05	2.91E-01	1.00E+00	na
1.04	0.96	1.26	2.34E-01	6.59E-01	0.000	1.09	0.94	1.26	2.34E-01	6.59E-01	0.000	1.04	0.93	1.17	4.46E-01	8.46E-01	0.000	1.04	0.93	1.17	4.46E-01	8.46E-01	0.000	1.06	0.88	1.28	5.45E-01	1.00E+00	na
1.00	0.94	1.07	9.17E-01	3.28E-01	0.135	1.01	0.94	1.08	8.82E-01	3.28E-01	0.135	1.00	0.95	1.06	8.85E-01	9.42E-01	0.000	1.00	0.94	1.09	7.32E-01	9.42E-01	0.000	1.01	0.94	1.09	7.32E-01	9.42E-01	0.000
0.95	0.89	1.01	1.02E-01	0.74E-01	0.241	0.95	0.80	1.11	9.74E-01	2.11E-02	0.241	0.95	0.90	1.01	8.07E-02	6.10E-02	0.715	0.96	0.86	1.07	4.46E-01	6.10E-02	0.715	0.91	0.85	0.98	1.17E-02	1.00E+00	na
1.03	0.92	1.16	5.56E-01	4.70E-01	0.000	1.03	0.92	1.16	5.56E-01	4.70E-01	0.000	0.94	0.86	1.03	1.83E-01	1.24E-01	0.578	0.95	0.83	1.10	5.19E-01	1.24E-01	0.578	1.04	0.89	1.21	6.39E-01	1.00E+00	na
1.08	0.74	0.99	3.70E-02	6.41E-02	0.587	1.08	0.69	1.19	4.60E-01	6.41E-02	0.587	0.83	0.74	0.92	7.49E-04	6.95E-01	0.000	0.83	0.74	0.92	7.49E-04	6.95E-01	0.000	0.85	0.71	1.03	1.01E-01	1.00E+00	na
1.08	0.96	1.21	2.18E-01	7.28E-01	0.000	1.08	0.96	1.21	2.18E-01	7.28E-01	0.000	1.09	1.00	1.19	6.35E-02	8.45E-01	0.000	1.09	1.00	1.19	6.35E-02	8.45E-01	0.000	1.08	0.92	1.25	3.40E-01	1.00E+00	na
0.99	0.91	1.07	7.54E-01	6.28E-01	0.000	0.99	0.91	1.07	7.54E-01	6.28E-01	0.000	0.94	0.89	1.00	4.06E-02	3.07E-01	0.040	0.94	0.89	1.00	4.06E-02	3.07E-01	0.040	0.98	0.89	1.09	7.20E-01	1.00E+00	na
0.92	0.86	0.99	3.50E-02	1.06E-01	0.476	0.93	0.83	1.04	2.04E-01	1.06E-01	0.476	0.90	0.85	0.96	1.11E-03	4.50E-01	0.000	0.90	0.85	0.96	1.11E-03	4.50E-01	0.000	0.93	0.85	1.02	1.03E-01	3.12E-01	0.022
1.04	0.92	1.12	2.81E-01	1.80E-01	0.000	1.04	0.92	1.12	2.81E-01	1.80E-01	0.000	1.12	1.06	1.19	1.06E-02	6.62E-01	0.000	1.12	1.06	1.19	1.06E-02	6.62E-01	0.000	1.11	1.03	1.23	3.96E-02	1.00E+00	na
1.03	0.95	1.11	4.74E-01	5.50E-01	0.000	1.03	0.95	1.11	4.74E-01	5.50E-01	0.000	0.97	0.92	1.02	2.50E-01	3.94E-01	0.000	0.97	0.92	1.02	2.50E-01	3.94E-01	0.000	1.00	0.91	1.11	9.63E-01	1.00E+00	na
1.03	0.91	1.17	6.59E-01	3.59E-02	0.699	1.05	0.82	1.35	6.87E-01	3.59E-02	0.699	1.10	1.00	1.21	3.93E-02	2.52E-03	0.890	1.04	0.76	1.42	8.03E-01	2.52E-03	0.890	0.88	0.74	1.05	1.53E-01	1.00E+00	na
1.01	0.93	1.08	8.77E-01	4.04E-01	0.000	1.01	0.93	1.08	8.77E-01	4.04E-01	0.000	0.94	0.89	1.00	3.57E-02	3.72E-02	0.770	0.96	0.85	1.08	4.98E-01	3.72E-02	0.770	1.03	0.93	1.13	6.07E-01	1.00E+00	na
0.97	0.89	1.05	4.72E-01	4.98E-01	0.000	0.97	0.89	1.05	4.72E-01	4.98E-01	0.000	0.94	0.89	1.00	3.57E-02	3.72E-02	0.770	0.96	0.85	1.08	4.98E-01	3.72E-02	0.770	1.03	0.93	1.13	6.07E-01	1.00E+00	na
1.00	0.90	1.11	9.89E-01	2.03E-01	0.349	0.99	0.86	1.15	9.11E-01	2.03E-01	0.349	1.01	0.94	1.10	7.12E-01	4.33E-01	0.000	1.06	0.93	1.22	3.93E-01	4.33E-01	0.000	1.06	0.93	1.22	3.93E-01	1.00E+00	0.000
1.11	1.04	1.19	2.48E-03	8.59E-01	0.000	1.11	1.04	1.19	2.48E-03	8.59E-01	0.000	1.15	1.09	1.22	1.97E-07	3.38E-01	0.848	1.15	1.09	1.22	1.97E-07	3.38E-01	0.848	1.13	1.04	1.23	3.17E-03	1.90E-01	0.417
1.10	1.00	1.21	4.67E-02	3.22E-01	0.144	1.10	1.00	1.21	4.67E-02	3.22E-01	0.144	1.10	1.00	1.21	4.67E-02	3.22E-01	0.144	1.10	1.00	1.21	4.67E-02	3.22E-01	0.144	1.04	0.93	1.17	4.84E-02	4.04E-01	0.000
0.97	0.90	1.05	4.03E-01	4.04E-01	0.000	0.97	0.90	1.05	4.03E-01	4.04E-01	0.000	1.01	0.95	1.07	7.61E-01	1.44E-01	0.532	1.00	0.91	1.10	9.97E-01	1.44E-01	0.532	1.05	0.85	1.05	3.02E-01	1.00E+00	na
1.01	0.95	1.08	7.02E-01	7.29E-01	0.000	1.02	0.97	1.08	4.44E-01	7.29E-01	0.000	1.02	0.97	1.08	4.44E-01	7.29E-01	0.000	1.02	0.97	1.08	4.44E-01	7.29E-01	0.000	1.00	0.92	1.08	9.81E-01	2.23E-01	0.328
1.00	0.91	1.10	9.86E-01	5.31E-01	0.000	1.00	0.91	1.10	9.86E-01	5.31E-01	0.000	1.17	1.08	1.26	4.97E-05	6.64E-02	0.703	1.14	0.98	1.32	8.67E-02	6.64E-02	0.703	1.04	0.91	1.20	5.35E-01	1.00E+00	na
1.00	0.94	1.15	9.17E-01	5.54E-01	0.000	1.01	0.97	1.12	8.54E-01	5.54E-01	0.000	1.12	1.07	1.15	2.62E-02	1.24E-01	0.587	1.14	0.97	1.26	9.62E-02	1.24E-01	0.587	1.05	0.94	1.18	3.96E-01	1.00E+00	na
1.08	1.00	1.18	6.22E-02	5.40E-01	0.000	1.08	1.00	1.18	6.22E-02	5.40E-01	0.000	0.98	0.92	1.04	4.79E-01	1.93E-03	0.896	1.01	0.82	1.24	9.95E-01	1.93E-03	0.896	1.12	1.01	1.25	3.74E-02	1.00E+00	na
0.97	0.78	1.22	8.07E-01	4.50E-01	0.000	0.78	0.65	0.86	9.67E-05	1.00E+00	na	0.75	0.65	0.86	9.67E-05	1.00E+00	na	0.75	0.65	0.86	9.67E-05	1.00E+00	na	0.95	0.86	1.05	3.39E-01	1.00E+00	na
0.97	0.90	1.05	5.14E-01	6.33E-01	0.000	0.97	0.90																						

ST6. Association summary statistics for the eight replicating signals in the analyses employing all OA cases.

						Discovery			Replication		Discovery and replication	
SNP	Chr	Position	Nearest gene(s)	EA ¹	Stratum	EAF ²	OR [95%CI]	P	OR [95%CI]	P	OR [95%CI]	P
rs6976 ³	3	52703844	<i>GLT8D1</i>	T	All OA	0.38	1.14 [1.10-1.19]	8.34x10 ⁻¹⁰	1.05 [1.01-1.09]	1.75x10 ⁻⁰²	1.09 [1.06-1.12]	6.56x10 ⁻⁰⁹
rs11177 ³	3	52696345	<i>GNL3</i>	A	All OA	0.38	1.14 [1.10-1.19]	6.87x10 ⁻¹⁰	1.05 [1.01-1.09]	1.65x10 ⁻⁰²	1.09 [1.06-1.12]	5.13x10 ⁻⁰⁹
rs4836732	9	118306516	<i>ASTN2</i>	C	All OA	0.47	1.05 [1.01-1.10]	2.14x10 ⁻⁰²	1.02 [0.98-1.07]	2.81x10 ⁻⁰¹	1.04 [1.01-1.07]	1.56x10 ⁻⁰²
rs9350591	6	76298247	<i>FILIP1; SENP6</i>	T	All OA	0.11	1.10 [1.03-1.18]	2.75x10 ⁻⁰³	1.07 [1.01-1.14]	2.07x10 ⁻⁰²	1.09 [1.04-1.13]	2.78x10 ⁻⁰⁴
rs10492367	12	27906237	<i>KLHDC5; PTHLH</i>	T	All OA	0.19	1.12 [1.06-1.18]	1.68x10 ⁻⁰⁵	1.02 [0.97-1.07]	4.22x10 ⁻⁰¹	1.06 [1.03-1.10]	9.02x10 ⁻⁰⁴
rs835487	12	103584897	<i>CHST11</i>	G	All OA	0.35	1.07 [1.02-1.12]	3.33x10 ⁻⁰³	1.04 [1.00-1.08]	4.32x10 ⁻⁰²	1.05 [1.02-1.08]	6.22x10 ⁻⁰⁴
rs12107036	3	191082854	<i>TP63</i>	G	All OA	0.53	1.08 [1.04-1.13]	2.38x10 ⁻⁰⁴	1.01 [0.98-1.05]	4.54x10 ⁻⁰¹	1.05 [1.02-1.08]	2.15x10 ⁻⁰³
rs8044769	16	52396636	<i>FTO</i>	C	All OA	0.51	1.11 [1.06-1.16]	1.30x10 ⁻⁰⁶	1.03 [0.99-1.07]	1.08x10 ⁻⁰¹	1.07 [1.04-1.10]	3.56x10 ⁻⁰⁶
rs10948172	6	44885669	<i>SUPT3H; CDC5L</i>	G	All OA	0.29	1.10 [1.05-1.15]	3.31x10 ⁻⁰⁵	1.06 [1.02-1.11]	5.87x10 ⁻⁰³	1.08 [1.05-1.12]	6.14x10 ⁻⁰⁷

¹EA, Effect allele.

²EAF, Effect allele frequency in controls.

³Both SNPs represent the same signal, $r^2=1$.

ST7. Association summary statistics for the replicating signals in analyses employing all cases or TJR-only cases.

SNP	CHR	BP	EA1	Stratum	TJR arcOGEN					all arcOGEN					TJR meta-analysis including arcOGEN					All meta-analysis including arcOGEN				
					OR	OR_95L	OR_95U	p	N cases/controls	OR	OR_95L	OR_95U	p	N cases/controls	OR	OR_95L	OR_95U	p	N cases/controls	OR	OR_95L	OR_95U	p	N cases/controls
rs6976†	3	52703844	T	All	1.16	1.11	1.22	2.27E-10	5804/11009	1.14	1.10	1.19	8.34E-10	7410/11009	1.12	1.08	1.16	7.24E-11	9852/44815	1.09	1.06	1.12	6.56E-09	14883/53947
rs11177†	3	52696345	A	All	1.16	1.11	1.22	2.12E-10	5804/11009	1.14	1.10	1.19	6.87E-10	7410/11009	1.12	1.08	1.16	1.25E-10	9852/44815	1.09	1.06	1.12	5.13E-09	14883/53947
rs4836732	9	118306516	C	Hip female	1.19	1.10	1.28	1.19E-05	1769/5515	1.19	1.11	1.28	2.72E-06	1934/5515	1.20	1.13	1.27	6.11E-10	3145/23058	1.16	1.10	1.22	1.24E-08	4528/28031
rs9350591	6	76298247	T	Hip**	1.20	1.11	1.31	1.94E-05	3039/11009	1.20	1.11	1.30	2.49E-06	3912/11009	1.18	1.11	1.26	1.59E-07	5673/44815	1.18	1.12	1.25	2.42E-09	7714/53947
rs10492367	12	27906237	T	Hip	1.19	1.11	1.28	7.00E-07	3039/11009	1.18	1.11	1.27	1.20E-06	3266/11009	1.16	1.10	1.22	1.61E-08	5429/44815	1.14	1.09	1.20	1.48E-08	6861/53947
rs835487	12	103584897	G	Hip	1.15	1.08	1.22	3.26E-06	3039/11009	1.14	1.08	1.21	7.95E-06	3266/11009	1.13	1.09	1.18	1.64E-08	5429/44815	1.11	1.06	1.15	2.79E-07	6861/53947
rs12107036	3	191082854	G	Knee female	1.23	1.13	1.35	3.03E-06	1217/5515	1.18	1.10	1.26	6.94E-06	2135/5515	1.21	1.13	1.29	6.71E-08	2151/23058	1.10	1.05	1.15	1.20E-04	4961/28031
rs8044769	16	52396636	C	Female	1.17	1.10	1.24	7.09E-07	3363/5515	1.17	1.10	1.23	5.98E-08	4476/5515	1.14	1.08	1.19	1.12E-07	5725/28031	1.11	1.07	1.15	6.85E-08	9266/28031
rs10948172	6	44885669	G	Male	1.17	1.09	1.26	1.51E-05	2441/5494	1.17	1.10	1.26	5.02E-06	2934/5494	1.15	1.09	1.22	1.97E-07	4127/21757	1.14	1.09	1.20	7.92E-08	5518/24066

†EA, Effect allele.

† Both SNPs represent the same signal, $r^2=1$.

** Indicates analyses where samples with OA at the hip and the knee are also included in the arcOGEN discovery set and in the UK replication set.

ST8. Allele frequencies and association summary statistics for previously established OA loci.

	Type of study	Ethnicity	EA ¹	EAF ² cases	EAF controls	Freq HapMap	P	OR [95% CI]
rs143383								
Miyamoto hip OA v. controls	Discovery	Japanese	T	0.811	0.73	0.756	0.002	1.59 [1.18-2.15]
Miyamoto hip OA v. controls	Discovery and replication	Japanese	T	0.836	0.74	0.756	2x10 ⁻¹³	1.79 [1.53-2.09]
Miyamoto knee OA v. controls	Discovery	Japanese	T	0.788	0.741	0.756	0.002	1.30 [1.10-1.53]
Miyamoto knee OA v. controls	Discovery	Chinese	T	0.784	0.702	0.779	3x10 ⁻⁰⁴	1.54 [1.22-1.95]
Evangelou hip OA v. controls	Meta-analysis	European	A	N/A	N/A	0.667	0.034	1.07 [1.01-1.14]
Evangelou knee OA v. controls	Meta-analysis	European	A	N/A	N/A	0.667	9x10 ⁻⁰⁵	1.13 [1.06-1.20]
Valdes knee OA v. controls	Meta-analysis	European and Asian	T	N/A	N/A	N/A	6x10 ⁻¹¹	1.17 [1.12-1.23]
Valdes knee OA v. controls	Meta-analysis	European	T	N/A	N/A	0.667	1x10 ⁻⁰⁸	1.16 [1.10-1.22]
rs4911494 (r²=0.93 with rs143383)								
All OA v. population-based controls	Discovery	European	A	0.635	0.634	0.650	0.742	1.01 [0.96-1.05]
Hip OA v. population-based controls	Discovery	European	A	0.635	0.634	0.650	0.918	1.00 [0.95-1.06]
Knee OA v. population-based controls	Discovery	European	A	0.636	0.634	0.650	0.726	1.01 [0.96-1.07]
Female OA v. female TwinsUK controls	Discovery	European	A	0.638	0.617	0.650	0.027	1.09 [1.01-1.18]
Hip female OA v. female TwinsUK controls	Discovery	European	A	0.634	0.617	0.650	0.131	1.07 [0.98-1.18]
Knee female OA v. TwinsUK	Discovery	European	A	0.640	0.617	0.650	0.039	1.10 [1.01-1.21]
rs3815148								
Kerkhof hip female OA v. controls	Discovery	European	C	N/A	N/A	0.233	0.810	1.03
Kerkhof knee female OA v. controls	Discovery	European	C	N/A	N/A	0.233	7x10 ⁻⁰⁵	1.42
Kerkhof knee female and/or hand OA v. controls	Discovery	European	C	N/A	N/A	0.233	7x10 ⁻⁰⁵	1.32 [1.15-1.51]
Kerkhof knee and/or hand OA v. controls	Meta-analysis	European	C	N/A	N/A	0.233	8x10 ⁻⁰⁸	1.14 [1.09-1.19]
All OA v. population-based controls	Discovery	European	C	0.235	0.228	0.233	0.098	1.04 [0.99-1.10]
Hip OA v. population-based controls	Discovery	European	C	0.228	0.228	0.233	0.987	1.00 [0.94-1.07]
Knee OA v. population-based controls	Discovery	European	C	0.247	0.228	0.233	6x10 ⁻⁰⁴	1.12 [1.05-1.19]
Female OA v. female TwinsUK controls	Discovery	European	C	0.235	0.216	0.233	0.022	1.12 [1.02-1.22]
Hip female OA v. female TwinsUK controls	Discovery	European	C	0.226	0.216	0.233	0.277	1.06 [0.95-1.19]
Knee female OA v. female TwinsUK controls	Discovery	European	C	0.246	0.216	0.233	0.002	1.19 [1.07-1.32]
rs11842874								

Day-Williams knee and/or hip OA v. controls	Discovery	European	A	0.940	0.920	0.946	1.67x10 ⁻⁰⁵	1.32 [1.16-1.50]
Day-Williams knee and/or hip OA v. controls	Meta-analysis	European	A	0.933	0.921	0.946	2.07x10 ⁻⁰⁸	1.17 [1.11-1.23]
All OA v. population-based controls	Discovery	European	A	0.933	0.925	0.946	2.45x10 ⁻⁰³	1.13 [1.05-1.23]
Hip OA v. population-based controls	Discovery	European	A	0.931	0.925	0.946	0.079	1.10 [0.99-1.23]
Knee OA v. population-based controls	Discovery	European	A	0.935	0.925	0.946	4.52x10 ⁻⁰³	1.16 [1.04-1.30]
Female OA v. female TwinsUK controls	Discovery	European	A	0.936	0.930	0.946	0.186	1.11 [0.95-1.29]
Hip female OA v. female TwinsUK controls	Discovery	European	A	0.937	0.930	0.946	0.236	1.12 [0.93-1.34]
Knee female OA v. female TwinsUK controls	Discovery	European	A	0.937	0.930	0.946	0.226	1.12 [0.94-1.33]

¹EA Effect allele.

²EAF Effect allele frequency.

ST9. Association summary statistics for the replicating signals in analyses employing population-based controls or OA-free controls in the discovery set.

			Cases v. population-based controls			Female cases v. female OA-free controls		
SNP	EA ¹	Stratum	OR [95%CI]	P	N cases/controls	OR [95%CI]	P	N cases/controls
rs6976	T	TJR	1.16 [1.11-1.22]	2.3x10 ⁻¹⁰	5804/11009	1.14 [1.05-1.23]	2.5x10 ⁻⁰³	3363/1828
rs4836732	C	THR	1.13 [1.07-1.20]	2.3x10 ⁻⁰⁵	3039/11009	1.19 [1.08-1.30]	3.5x10 ⁻⁰⁴	1769/1828
rs9350591	T	Hip**	1.20 [1.11-1.30]	2.5x10 ⁻⁰⁶	3912/11009	1.22 [1.07-1.39]	3.8x10 ⁻⁰³	2341/1828
rs10492367	T	Hip	1.18 [1.11-1.27]	1.2x10 ⁻⁰⁶	3266/11009	1.31 [1.16-1.47]	5.8x10 ⁻⁰⁶	1934/1828
rs835487	G	THR	1.15 [1.08-1.22]	3.3x10 ⁻⁰⁶	3039/11009	1.14 [1.03-1.25]	8.9x10 ⁻⁰³	1769/1828
rs12107036	G	TKR	1.11 [1.04-1.19]	1.2x10 ⁻⁰³	2164/11009	1.17 [1.05-1.29]	3.4x10 ⁻⁰³	1217/1828
rs8044769	C	All	1.11 [1.06-1.16]	1.3x10 ⁻⁰⁶	7410/11009	1.15 [1.06-1.24]	4.3x10 ⁻⁰⁴	4476/1828
rs10948172	G	All	1.10 [1.05-1.15]	3.3x10 ⁻⁰⁵	7410/11009	1.03 [0.95-1.12]	4.9x10 ⁻⁰¹	4476/1828

¹EA, Effect allele.

** Indicates analyses where a subset of samples with OA at the hip and the knee are also included in the arcOGEN discovery set and in the UK replication set.

ST10. Association summary statistics for the replicating signals in meta-analyses including all controls or OA-free controls only.

						Meta-analysis including all controls			Meta-analysis including OA-free controls		
SNP	Chr	Position	Nearest gene(s)	EA ¹	Stratum	OR [95%CI]	P	N cases/controls	OR [95%CI]	P	N cases/controls
rs6976 ²	3	52703844	<i>GLT8D1</i>	T	TJR	1.12 [1.08-1.16]	7.24x10 ⁻¹¹	9852/44815	1.10 [1.04-1.16]	2.90x10 ⁻⁰⁴	5394/33315
rs11177 ²	3	52696345	<i>GNL3</i>	A	TJR	1.12 [1.08-1.16]	1.25x10 ⁻¹⁰	9852/44815	1.10 [1.05-1.16]	1.60x10 ⁻⁰⁴	5394/33315
rs4836732	9	118306516	<i>ASTN2</i>	C	THR-female	1.20 [1.13-1.27]	6.11x10 ⁻¹⁰	3145/23058	1.19 [1.11-1.28]	4.39x10 ⁻⁰⁷	2540/19037
rs9350591	6	76298247	<i>FILIP1; SENP6</i>	T	Hip ^{**}	1.18 [1.12-1.25]	2.42x10 ⁻⁰⁹	7714/53947	1.16 [1.07-1.25]	1.60x10 ⁻⁰⁴	4734/38246
rs10492367	12	27906237	<i>KLHDC5; PTHLH</i>	T	Hip	1.14 [1.09-1.20]	1.48x10 ⁻⁰⁸	6861/53947	1.12 [1.05-1.19]	5.03x10 ⁻⁰⁴	4327/38246
rs835487	12	103584897	<i>CHST11</i>	G	THR	1.13 [1.09-1.18]	1.64x10 ⁻⁰⁸	5429/44815	1.15 [1.08-1.22]	5.44x10 ⁻⁰⁶	3164/ 33315
rs12107036	3	191082854	<i>TP63</i>	G	TKR-female	1.21 [1.13-1.29]	6.71x10 ⁻⁰⁸	2151/23058	1.14 [1.06-1.24]	1.01x10 ⁻⁰³	1702/ 19037
rs8044769	16	52396636	<i>FTO</i>	C	Female	1.11 [1.07-1.15]	6.85x10 ⁻⁰⁸	9266/28031	1.11 [1.06-1.16]	1.14x10 ⁻⁰⁵	7484/ 21659
rs10948172	6	44885669	<i>SUPT3H; CDC5L</i>	G	Male	1.14 [1.09-1.20]	7.92x10 ⁻⁰⁸	5518/24066	1.09 [1.00-1.18]	4.37x10 ⁻⁰²	1628/ 16587

¹EA, Effect allele.

²Both SNPs represent the same signal, r²=1.

^{**}Indicates analyses where a subset of samples with OA at the hip and the knee are also included in the arcOGEN discovery set and in the UK replication set.

ST11. Association p-values before and after BMI adjustment for the 129 prioritized SNPs.

Chr	SNP	Position	P_unadjusted	P_BMI_adjusted
1	rs1854169	102046864	0.1304	0.203
1	rs1417066	217740451	0.05318	0.1015
1	rs1629896	223752551	0.4974	0.5562
1	rs987179	242876559	0.02743	0.1414
2	rs1024816	37454712	0.004964	0.007703
2	rs3791679	55950396	0.868	0.8224
2	rs10779958	74609286	0.03986	0.3123
2	rs6754683	108522707	0.4807	0.4278
2	rs4622728	125523313	0.02417	0.232
2	rs1432237	137524977	0.933	0.8576
2	rs12618428	151441225	0.4088	0.617
2	rs6716254	175250720	0.01225	0.01121
3	rs6793234	8801396	0.6399	0.8961
3	rs11177	52696345	0.01815	0.01162
3	rs6976	52703844	0.01876	0.01185
3	rs10510816	59630570	0.01787	0.01275
3	rs1355782	132970356	0.05988	0.06085
3	rs16851066	142325511	0.04054	0.06121
3	rs9867979*	166345704	0.000007714	0.0001797
3	rs12107036	191082854	0.004629	0.008156
4	rs16837352	5555175	0.05	0.02485
4	rs17578878	37577120	0.02794	0.02284
4	rs11725992	48082608	0.4048	0.6794
4	rs6842739*	60172105	0.00007574	0.0000738
4	rs2626053*	96474111	0.00001253	0.001522
4	rs2636726*	106528673	0.00005061	0.0002791
4	rs12500935	114760230	0.2216	0.3102
4	rs1566347	186969560	0.5776	0.546
5	rs4957048	636442	0.8404	0.8016
5	rs1083523*	66314669	0.00001819	0.0002898
5	rs2610424	71292415	0.005757	0.01183
5	rs457008	81110411	0.3066	0.2176
5	rs10036746	91780688	0.7919	0.5747
5	rs11135394*	93212891	4.841E-07	0.00000169
5	rs12515798*	134333096	0.0000109	0.0001752
5	rs10515550	144792886	0.6042	0.3032
5	rs13167773	148567182	0.00008868	0.0000545
5	rs10051783	167755230	0.0004402	0.0009038
6	rs10948172	44885669	0.8535	0.6911
6	rs9350591	76298247	0.01322	0.02511
6	rs569731*	76414313	0.0008494	0.0001632
6	rs10943837	82727239	0.003018	0.008799
6	rs2493984	95440213	0.006584	0.005263
6	rs7755798	109743847	0.01978	0.03172
6	rs2027532	112741889	0.1551	0.1733
6	rs6931833	114078284	0.8543	0.8936
7	rs17627827	16731521	0.5464	0.4206
7	rs6461679	22976588	0.0006546	0.001775

7	rs7785659	32427540	0.1505	0.353
7	rs7805536	32540095	0.1665	0.1955
7	rs1859572	76870933	0.2269	0.5753
7	rs6466265	77998184	0.0002065	0.001685
7	rs1488517*	95446689	0.00008932	0.00345
7	rs2966417	109944782	0.0006126	0.004868
8	rs4841020	8638760	0.03563	0.01004
8	rs1876836	8719166	0.1023	0.1061
8	rs4841067	8784655	0.02033	0.00298
8	rs609792	9857300	0.007901	0.006625
8	rs6997710	10247861	0.03274	0.03251
8	rs1421259	15607043	0.02156	0.04276
8	rs2979715*	80724814	1.464E-08	4.27E-08
8	rs2467753*	80759808	0.00001013	0.0000147
8	rs2289496	98357340	0.4036	0.2598
8	rs16870112	104191422	0.1047	0.1939
8	rs10956114*	124156008	0.000006739	0.00000354
8	rs4072286	142805713	0.01148	0.0165
9	rs2146423	4647040	0.001239	0.0002433
9	rs945442	18917768	0.005024	0.002601
9	rs3780296	79327254	0.5945	0.934
9	rs4072357*	89688030	0.000005694	0.0000292
9	rs4836732	118306516	0.00000692	0.00000288
9	rs719535*	118516091	0.00002235	0.0000401
9	rs2900277	131434091	0.002476	0.01879
10	rs2474714*	33535989	0.0008353	0.001965
10	rs9299558	56603927	0.228	0.3012
10	rs1250552*	80728033	0.000002497	0.0001308
10	rs787640	95114663	0.02506	0.04294
10	rs11188412	97391986	0.06331	0.1534
11	rs4757420	12371653	0.5509	0.8094
11	rs7119797	61922811	0.01811	0.008324
11	rs17159801	82496946	0.001266	0.0005232
11	rs2691829	90240224	0.5601	0.3535
11	rs6589848	120229228	0.001544	0.005266
11	rs7941193	125876053	0.0004613	0.002248
11	rs7122854	132673733	0.09471	0.06992
12	rs7305794	24860894	0.01265	0.05208
12	rs10492367	27906237	0.00002596	0.0000123
12	rs6487684	28534480	0.00001242	0.0000223
12	rs1034762	46675910	0.8002	0.9124
12	rs11107957*	76956186	0.000003129	0.0004511
12	rs1404866	83888610	0.0003551	0.0001185
12	rs11107219*	92770487	0.000005263	0.0000545
12	rs835487	103584897	0.008516	0.01282
12	rs11064722	118157566	0.7795	0.7302
13	rs4544137*	67072419	0.00001705	0.0001463
13	rs7994526*	67263831	0.00002614	0.0004633
13	rs2036717*	75645031	0.000476	0.0006804
13	rs9520058	105855421	0.02588	0.09082

14	rs8015303*	34316223	0-0001215	0-00000176
14	rs1627411*	42241078	0-00002329	0-0000378
14	rs7155791	49487980	0-0881	0-4414
14	rs1998094*	51653209	0-00003446	0-00000214
14	rs10144429*	77722272	0-00000357	0-0001263
15	rs11634255	30874764	0-688	0-9063
15	rs4646626	56043419	0-2945	0-3997
15	rs8032755	91403060	0-0461	0-04206
16	rs2908646	4971845	0-178	0-328
16	rs12596722	9324789	0-0001885	0-001024
16	rs12923310	9695438	0-437	0-2658
16	rs8044769	52396636	0-004069	0-1706
16	rs9937815	54888333	0-6869	0-61
16	rs7404629	64970230	0-05379	0-04675
16	rs16947129*	76709247	0-000001504	0-0000114
17	rs4078062	16919559	0-4242	0-2574
17	rs8080960	67298538	0-00114	0-0000718
18	rs11665347	10468812	0-02049	0-02802
18	rs12454396*	62299667	0-00004345	0-0000237
18	rs17089382	70369681	0-01933	0-009003
19	rs12974139	2135005	0-2036	0-2152
19	rs1610093*	4505774	0-000002781	0-000031
19	rs10426443	8017982	0-004553	0-005486
19	rs10426377	53784046	0-008796	0-01162
20	rs1360400	22909890	0-002412	0-0004732
20	rs172981*	40543116	0-000006284	0-000077
20	rs17785895	47866466	0-3062	0-1136
20	rs4925370	60300986	0-7025	0-9269
22	rs3747081	19682120	0-2335	0-3887
22	rs7287616	22259852	0-1077	0-04072
22	rs5762347	26489672	0-752	0-9592

* Signals prioritised from the OA vs TwinsUK analyses.

ST12. Expression in joint tissues, assessed by reverse-transcription PCR. NOF cartilage was collected from patients undergoing surgery for a neck-of-femur fracture.

Chr	Gene	Tissue							
		OA cartilage	NOF cartilage	Tendon	Ligament	Meniscus	Synovium	Fat pad	Osteophyte
3p21.1	<i>STAB1</i>	X ¹	X	Y ²	Y	Y	Y	Y	Y
	<i>NT5DC2</i>	Y	Y	Y	Y	Y	Y	Y	Y
	<i>PBRM1</i>	Y	Y	Y	Y	Y	Y	Y	Y
	<i>GNL3</i>	Y	Y	Y	Y	Y	Y	Y	Y
	<i>GLT8D1</i>	Y	Y	Y	Y	Y	Y	Y	Y
	<i>SPCS1</i>	Y	Y	Y	Y	Y	Y	Y	Y
	<i>NEK4</i>	X	X	Y	Y	Y	Y	Y	Y
	<i>ITIH1</i>	X	X	Y	X	Y	Y	X	X
	<i>ITIH3</i>	X	X	Y	Y	Y	Y	Y	Y
	<i>ITIH4</i>	X	X	Y	Y	Y	Y	Y	Y
	<i>MUSTN1</i>	Y	Y	Y	Y	Y	Y	Y	Y
	<i>TMEM110</i>	Y	Y	Y	Y	Y	Y	Y	Y
3q28	<i>TP63</i>	X	X	Y	X	X	X	Y	Y
6p21.1	<i>SUPT3H</i>	Y	Y	Y	Y	Y	Y	Y	Y
	<i>RUNX2</i>	Y	Y	Y	Y	Y	Y	Y	Y
6q13-q14.1	<i>COL12A1</i>	Y	Y	Y	Y	Y	Y	Y	Y
	<i>COX7A2</i>	Y	X	Y	Y	Y	Y	Y	Y
	<i>TMEM30A</i>	Y	Y	Y	Y	Y	Y	Y	Y
	<i>FILIP1</i>	Y	Y	Y	Y	Y	Y	Y	Y
	<i>SENP6</i>	Y	Y	Y	Y	Y	Y	Y	Y
	<i>MYO6</i>	Y	Y	Y	Y	Y	Y	Y	Y
	<i>IMPG1</i> ³								
9q33.1	<i>PAPPA</i>	Y	Y	Y	X	Y	Y	Y	Y
	<i>ASTN2</i>	X	Y	Y	Y	Y	Y	Y	Y
	<i>TRIM32</i>	Y	Y	Y	Y	Y	Y	Y	Y
12p11.22	<i>KLHDC5</i>	Y	Y	Y	Y	Y	Y	Y	Y
	<i>PTHLH</i>	Y	Y	Y	Y	Y	Y	Y	Y
12q23.3	<i>CHST11</i>	Y	Y	Y	Y	Y	Y	Y	
16q12.2	<i>FTO</i>	Y	Y	Y	Y	Y	Y	Y	

¹ X, gene is not expressed; where a gene is marked as not showing expression in a particular tissue, this is the result of studying at least three independent tissue samples for that gene.

² Y, gene is expressed.

³ Amplification not possible in our study.

ST13. Oligonucleotide primers used to assess expression of the genes in joint tissue RNA.

Chr	Gene	Forward primer (5'-3')	Location	Reverse primer (5'-3')	Location	PCR product size (bp)
3p21·1	<i>STAB1</i>	cctggcaggctcagcttcg	Exon 1	cggcagtcctgggttatctg	Exon 2	168
	<i>NT5DC2</i>	gaaaactcgaatgagaaggc	Exon 10	ccagatcactatagaggtgg	Exon 11	159
	<i>PBRM1</i>	taggagttgtcggataacc	Exon 10	agaactcgtttagatggc	Exon 11	243
	<i>GNL3</i>	gcatgacctgccataagcgg	Exon 2	gcctgtcaagtttctgctgc	Exon 4	208
	<i>GLT8D1</i>	atcttggctctggctgttc	Exon 4	gcgagttgtgtgctgaatgc	Exon 5	249
	<i>SPCS1</i>	gccagaagctagctgaacag	Exon 2	attcttgaacaggaaccac	Exon 4	201
	<i>NEK4</i>	acagtgtttctggagaggc	Exon 6	taagatgtcaatattacgc	Exon 7	234
	<i>ITIH1</i>	tgccacggacttcagtgg	Exon 19	caaagtgccaccgtcgtcc	Exon 20	173
	<i>ITIH3</i>	tgtataacctgggctttggc	Exon 11	agccatcgtagaagtctgg	Exon 12	231
	<i>ITIH4</i>	cccaagcaatccgtggagg	Exon 11	gcctctccatgaagttgtgg	Exon 12	233
	<i>MUSTN1</i>	cgcttcagagaccagcgc	Exon 1	ctctcgatgacctggtagg	Exon 2	174
	<i>TMEM110</i>	tttcaaatgtatacctagc	Exon 3	agccagggactcccactgc	Exon 4	160
3q28	<i>TP63</i>	tgtccttccagcagtcgagc	Exon 4	tcaccacctccgtgacgtgc	Exon 5	189
6p21·1	<i>SUPT3H</i>	gcagcctgtggtttgaggc	Exon 9	aagccatcccattcctgcgg	Exon 10	133
	<i>RUNX2</i>	catcccagatgagagtagg	Exon 6	cgctcatctggctcaggtagg	Exon 7	174
6q13-q14·1	<i>COL12A1</i>	ggtgctgcagggattgaagg	Exon 13	ccgcgaaccattcctggagc	Exon 14	198
	<i>COX7A2</i>	gcatacgcaagactcggagg	Exon 2	ggaagcagtgcttatctgcc	Exon 3	154
	<i>TMEM30A</i>	ttctatcaaaaccatcgtcg	Exon 3	atgctgttggcaatagctcc	Exon 4	155
	<i>FILIP1</i>	cgtatacaggttagagaacg	Exon 4	tctcttagttattgagtcg	Exon 5	177
	<i>SENP6</i>	agaatacccacctcatgtcc	Exon 7	tcctcttaattcaggctcc	Exon 8	128
	<i>MYO6</i>	agtgttctctcagaagagg	Exon 3	tcaaagatggattcactgc	Exon 5	165
	<i>IMPG1¹</i>					
9q33·1	<i>PAPPA</i>	gccatgttgacctccactgc	Exon 7	tagttatgacccaatactgg	Exon 8	178
	<i>ASTN2</i>	gagccccacaccagtctgc	Exon 16	ggcgcctgaggtcacactgc	Exon 17	166
	<i>TRIM32</i>	tccggcgggtgactcgtcgg	5' UTR	cttcccggaggcatccagg	Exon 2	164
12p11·22	<i>KLHDC5</i>	tcattggaggatacactacc	Exon 2	gtcctcgtctgctgctcc	Exon 3	200
	<i>PTHLH</i>	tcgcggtgttctctgctgagc	Exon 3	agatggtgaaggaagaatcg	Exon 4	157
12q23·3	<i>CHST11</i>	ggaatccctttggtgtggac	Exon 3	cagtagatgagctcgtgtcc	Exon 4	235
16q12·2	<i>FTO</i>	gactgccatcctgcctcgc	Exon 8	gtcaaacggcagaggcatcg	Exon 9	160

¹ Amplification not possible in our study.