Journal of Human Genetics
https://doi.org/10.1038/s10038-018-0418-y

ARTICLE

Nature vs. nurture in human sociality: multi-level genomic analyses of social conformity

Biqing Chen[1,2] · Zijian Zhu[1] · Yingying Wang[1] · Xiaohu Ding[3] · Xiaobo Guo[4] · Mingguang He[3] · Wan Fang[1] · Qin Zhou[5] · Shanbi Zhou[6] · Han Lei[5] · Ailong Huang[5] · Tingmei Chen[5] · Dongsheng Ni[7] · Yuping Gu[7] · Jianing Liu[7] · Yi Rao[1]

Received: 30 October 2017 / Revised: 14 December 2017 / Accepted: 16 January 2018

Abstract

Social conformity is fundamental to human societies and has been studied for more than six decades, but our understanding of its mechanisms remains limited. Individual differences in conformity have been attributed to social and cultural environmental influences, but not to genes. Here we demonstrate a genetic contribution to conformity after analyzing 1,140 twins and single-nucleotide polymorphism (SNP)-based studies of 2,130 young adults. A two-step genome-wide association study (GWAS) revealed replicable associations in 9 genomic loci, and a meta-analysis of three GWAS with a sample size of ~2,600 further confirmed one locus, corresponding to the *NAV3* (Neuron Navigator 3) gene which encodes a protein important for axon outgrowth and guidance. Further multi-level (haplotype, gene, pathway) GWAS strongly associated genes including *NAV3*, *PTPRD* (protein tyrosine phosphatase receptor type D), *ARL10* (ADP ribosylation factor-like GTPase 10), and *CTNND2* (catenin delta 2), with conformity. Magnetic resonance imaging of 64 subjects shows correlation of activation or structural features of brain regions with the SNPs of these genes, supporting their functional significance. Our results suggest potential moderate genetic influence on conformity, implicate several specific genetic elements in conformity and will facilitate further research on cellular and molecular mechanisms underlying human conformity.

Introduction

Conformity is essential to human societies [1–5]. Conformity can be beneficial for individuals (through facilitation of learning by imitating the majority or with help from the majority) and for societies (through social cohesion and stability). Collectively, it is essential to Rousseau's social contract and forms the basis of modern democracy [6]. Individually, it can be used as a frequency-based strategy in social learning [7]. However, conformity also has negative consequences, from suppression of independence and originality, to mass hysteria, violence, dictatorship, and sometimes herding regardless of truth, fairness, or justice [8, 9]. Human individuals may conform to the majority even when the latter is known to be wrong [10]. Social conformity can be informative conformity (using information

These authors contributed equally: Biqing Chen and Zijian Zhu

Electronic supplementary material The online version of this article (https://doi.org/10.1038/s10038-018-0418-y) contains supplementary material, which is available to authorized users.

✉ Yi Rao
yrao@pku.edu.cn

[1] Peking-Tsinghua Center for Life Sciences, PKU-IDG/McGovern Institute for Brain Research, Beijing Advanced Innovation Center for Genomics, Peking University, 100871, Beijing, China

[2] Affiliated Hospital of Nanjing University of Chinese Medicine, Jiangsu Province Hospital of Traditional Chinese Medicine, 210029, Nanjing, China

[3] State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-Sen University, 510060, Guangzhou, China

[4] Department of Statistical Science, School of Mathematics, Sun Yat-Sen University, 510275, Guangzhou, China

[5] College of Laboratory Medicine, Chongqing Medical University, 400016, Chongqing, China

[6] University-Town Hospital Chongqing Medical University, 401331, Chongqing, China

[7] MOE Key Laboratory of Medical Diagnostics, Division of Molecular Nephrology and Creative Training Center for Undergraduates, College of Laboratory Medicine, Chongqing Medical University, 400016, Chongqing, China

Published online: 26 February 2018

Table 1 Summary statistics for conformity phenotypes

Phenotype	Male %	Age (years)	Subjects #	Range	Mean (SD)	Skewness	Kurtosis
Guangzhou Twin Cohort							
CONFP	46	16 ± 3	1,114	0–100%	47.1% (27.8%)	0.05	−0.98
CONFM			964	0–100%	69.1% (25.5%)	−0.76	−0.16
adCONFM			964	0–100%	57.6% (26.5%)	−0.41	−0.62
gCONF			938	0–1.72	1.01 (0.39)	−0.50	−0.41
GWAS Discovery Cohort							
CONFP	12	20 ± 1	793	0–100%	57.3% (28.4%)	−0.29	−0.95
CONFM			793	0–100%	78.0% (21.8%)	−1.11	0.86
adCONFM			793	0–100%	68.8% (24.3%)	−0.80	0.08
gCONF			793	0–1.73	1.17 (0.35)	−0.63	−0.15
Replication Cohort							
CONFP	22	19 ± 2	1,688	0–100%	55.4% (29.6%)	−0.22	−1.07
CONFM			1,686	0–100%	77.9% (22.2%)	−1.28	1.30
adCONFM			1,686	0–100%	68.2% (24.8%)	−0.91	0.32
gCONF			1,686	0–1.73	1.15 (0.36)	−0.68	−0.01
fMRI Cohort							
CONFP*	48	21 ± 2	64	28.6–75.0%	52.2% (10.7%)	0.20	−0.59
All cohorts							
CONFP	28	18 ± 2	3,595	0–100%	53.3% (29.1%)	−0.14	−1.06
CONFM			3,443	0–100%	75.5% (23.4%)	−1.09	0.66
adCONFM			3,443	0–100%	65.4% (25.7%)	−0.73	−0.13
gCONF			3,417	0–1.73	1.12 (0.37)	−0.63	−0.14

Male % the percentage of male subjects in each cohort, *Subjects #* the number of participants in each cohort, *SD* standard deviation, *CONFP* Price Estimation Conformity Test, *CONFM* Memory Conformity Test.

*CONFP measured from the revised price estimation conformity behavioral assay used in the fMRI study.

from the majority to increase the accuracy of individual perception or action) or normative conformity (gaining affiliation with the majority to increase cooperation or trust). Conformity affects both expression and perception: the desire to conform to the majority changes individual's perception [11]. Conformity can change previously established memory that was strong and accurate [12]. Evolutionarily, social conformity has been observed in animals such as rats [13], birds [14], and primates [15, 16]. Developmentally, social conformity occurs in children [17] as early as 7 months of age [18].

Individual differences have been found in human conformity. For example, in the classic Asch's line judgment experiment [3], when asked to choose a line of the same length as the standard line, participants varied from those changing their choices to fit the choice of the majority even when it was obviously wrong to those who held on to their own. Environmental factors have been thought to be the major determinant of individual differences in conformity. Studies of Asch-like line judgment tasks [8] as well as social preferences [19] have shown that residents in collectivist societies are more inclined to conform to the majority than those in individualist societies [20, 21]. Personal differences in conformity have also been attributed to early family experience [22], social anxiety [23], social power [24], and major personality attributes [25].

Functional magnetic resonance imaging (fMRI) has delineated brain regions involved in social conformity [11, 26–29], implicating the posterior medial frontal cortex (pMFC), the caudate, and the ventral striatum. Transcranial magnetic stimulation supports the functional significance of the pMFC because transient down-regulation of the pMFC reduced conformity [30].

The prevalence of conformity suggests the possibility that conformity is genetically influenced, similar to other social behaviors such as helping [31] or aggression [32]. However, only one paper has been published on genetic analysis of conformity, which concluded that there was no genetic influence on conformity [33]. Here we have adopted two behavioral assays to measure conformity. We have performed the first multi-level genomic research on social conformity, including: (1) a twin study to estimate heritability; (2) an estimation of total genetic influence explained by all genomic single-nucleotide polymorphisms (SNPs); (3) a single locus genome-wide association study (GWAS); (4) a gene-level association analysis across the whole

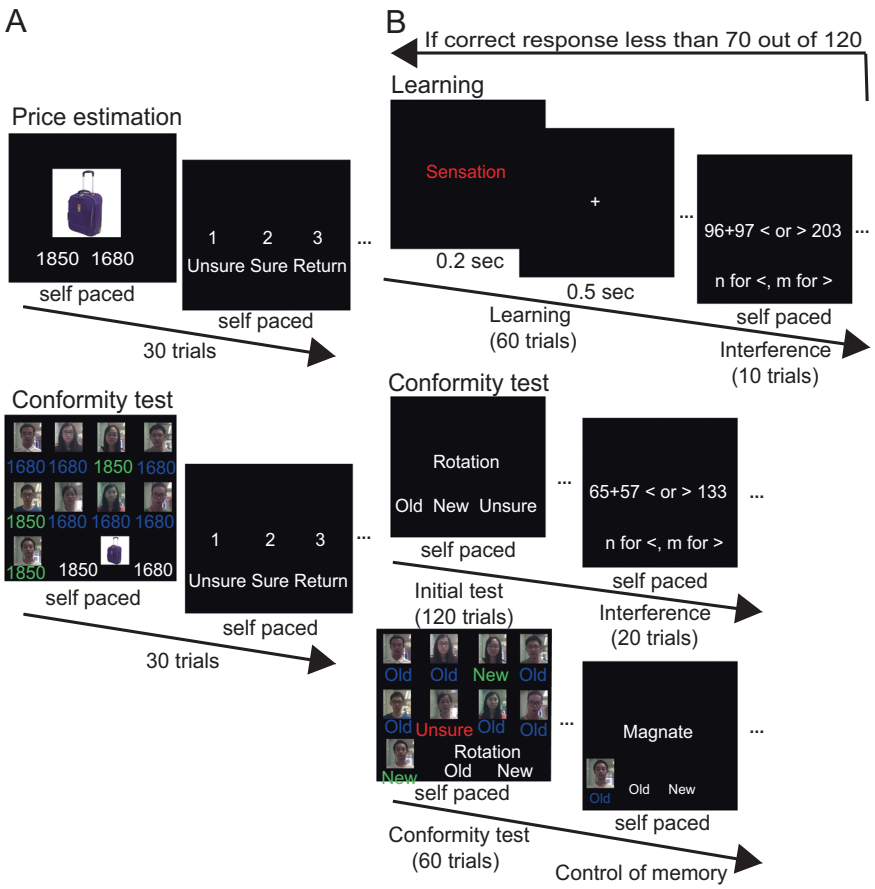

Fig. 1 Behavioral tests. **a** Illustration of procedure for the Price Estimation Conformity Test. Participants chose one out of two possible prices for each commodity, and rated their confidence for the choice. Then all commodities were shown for a second time, when the fabricated answer of each of the eight fictitious "other participants" was presented below the picture of each "participants," along with the participant's original choice in the first phase; **b** Illustration of procedure for the Memory Conformity Test. Sixty words were presented individually to the participant, followed by 10 arithmetic questions for distraction. The participant completed a recognition test by judging whether a word (120 in total) belonged to the series before ("Old") or not ("New"), or by stating their uncertainty about the answer ("Unsure"). The learning–testing process was repeated until the recognition accuracy exceeded 58%, and 20 arithmetic questions were provided to end this phase. The participant was tested again on 70 words to which he/she had given unambiguous answers. Fabricated answers of the eight fictitious "other participants" were presented together with the participant's original answer in the first phase. For one-sixth of the testing trials, answers of the majority were opposite to the participant's previous answers, and therefore used as critical tests of conformity. The last ten trials of the Memory Conformity Test were used as controls for memory, without the display of answers from "other participants." All the tasks were self-paced

genome; (5) a pathway-based association analysis to uncover biological networks; and (6) an MRI study for functional validations by analyzing differences in activation and anatomical features of brain regions between genotypes for each candidate locus.

Materials and methods

Participants

There were three Chinese cohorts (Table 1), a twin sample and two young adult cohorts. Twins came from the Guangzhou Twin Registry, 1,190 participants, aged 10 to 23 years, 54% females. Twin zygosity was confirmed by genotyping 16 polymorphic markers in all the same-sex twins. Four hundred and ninety-nine unrelated individuals (GZ_GWAS), with an average age of 17 years and 53% females, from the twin cohort, were further genotyped genome widely via a chip. The discovery (CQ_GWAS: 820) and replication cohorts (CQ_REP: 1488) consisted of 2,308 college students recruited at the Chongqing Medical University (Chongqing, China), with an average age of 19 years, 85% females, and 95% Han ethnicity. Another replication sample (BJ_REP: 263) and all 66 fMRI subjects were college students recruited from Peking University (Beijing, China), with an average age of 21 years, 52% females, and 98% Han. The institutional review board of Peking University approved the informed consent as well as the experimental protocols. Prior to testing, written

informed consent was obtained from the subjects or their guardians.

Behavioral assays

To minimize systematic error of measurement, two different tests for social conformity were employed in the current study, namely, Price Estimation Conformity Test (phenotype denoted as CONFP) and Memory Conformity Test (phenotype denoted as CONFM). Each test consisted of a judging stage and a conformity testing stage (Fig. 1). A "monetary incentive" condition was adopted in the tests to motivate participants to try their best, in that participants were told that the more accurate they were at the second stage, the higher reward they would receive. The experimental stimuli were programmed with Psychotoolbox in Matlab2010a. Stimuli were presented on an LCD PC Monitor with a resolution of $1{,}024 \times 768$ dpi for behavioral tests and presented through a backset projector with a resolution setting at 800×600 dpi for the fMRI study.

Price Estimation Conformity Test

The first stage is self-paced two-choice judgment of the prices of 60 given objects, represented by pictures, sequentially (Fig. 1a). Participants were instructed to make a guess at the true price and report their confidence in their choice. A "return" option was set to avoid mistakes. At the second stage, participants' own answers at the previous stage as well as those from eight confederates were displayed along with each one's real photos, and they were asked to estimate all the prices again. The computer had manipulated the provided eight answers so that for 20% of the trials, the choice in majority was opposite to the participant's own answers at the first stage (the so-called "conformity trials").

Memory Conformity Test

At the first stage, participants had to memorize a list of sequentially shown words and be tested by a self-paced old/new recognition test (Fig. 1b). This learning–testing cycle was repeated until an accuracy of 58% (70 correct) was achieved, and at last it was followed by an arithmetic distraction task to get rid of working memory effect. At the second stage, the participant was asked to do the recognition test again but with half trials, while answers from eight confederates and his/her own were provided. The last ten trials were used as control for memory, in which only the participant's original answers were provided ("control trials"). One-sixth of the trials were "conformity trials." Responses were also selected so that the participant's original answers were correct in only half of the conformity trials.

Debriefing

After behavioral tests or fMRI experiments, participants were asked two questions to assess their perceptions about the experiment: (1) whether they believed the answers provided at the second stage were real answers from real participants; (2) what their responses were when the majority's answers were different from their previous ones.

Genotyping and quality control

Phenotype data preprocessing

Conformity was calculated as the percentage of changed answers in the "conformity trials." Memory difference was determined by the percentage of changed answers in the "control trials." Considering individual differences in confidence about memory, we derived an adCONFM by subtracting the memory difference from CONFM. We further conducted a principal component analysis (PCA) of all the raw measurements and took the first principal component (PC) as an index of conformity (gCONF). Excluded from the final data analysis were those who were skeptical about the reality and the purpose of the tests; they consisted of 3% of all participants (112 individuals) (Supplementary Table S1). The cleaned samples consisted of 557/482 twins pairs with CONFP/CONFM, 811 individuals for discovery, 1,700 for replication, and 64 for fMRI (Supplementary Table S1). All traits were inverse normal transformed (INT). Individuals with extreme phenotypes outside four standard deviations of the population mean were taken as outliers and removed from subsequent analyses.

Discovery (CQ_GWAS) cohort

DNA was extracted from peripheral blood of 811 individuals using the QuickGene whole blood genome DNA extract system (Kurabo Industries Ltd, Japan), and was genotyped for 894,517 common SNPs using the HumanOmniZhongHua-8 Beadchip v1.1 (Illumina Inc., San Diego, CA, USA). SNPs were included if they met the following criteria: call rate ≥0.95, minor allele frequency (MAF) ≥0.01, and Hardy–Weinberg equilibrium with $p \geq 1 \times 10^{-4}$. Individuals with call rate <0.95 were excluded. Population stratification was examined via PCA using EIGENSTRAT [34]; outliers beyond five standard deviations were excluded automatically with the default mode; the first 10 PCs were extracted. A total of 793

subjects with 830,937 SNPs were included in the final analyses.

Replication cohorts

The replication cohorts of 1,688 individuals (after quality control) was either sequenced via Sequenom iPlex (in Bio Miao Biological Inc., Beijing, China) in 354 individuals (BJ_REP) or genotyped in 1,334 individuals (CQ_REP) on the HumanOmniZhongHua-8 Beadchip v1.2 with the same quality control criteria as before. The fMRI sample was genotyped by sequencing via Sequenom iPlex.

GZ_GWAS cohort

DNA of 499 individuals from the Guangzhou twin cohort was extracted and genotyped on the Affymetrix ASI Axiom 1.0 chip (Thermo Fisher Scientific Inc., Waltham, MA, USA) with 598,317 SNPs, and standard quality control was applied as before, reducing the final set to 500,026 SNPs.

Statistical genetic analysis

Heritability estimation from twin study

For each triple dizygotic (DZ) twin, one individual was discarded randomly, leaving only one pair of DZ twins, resulting in 216 DZ twins. Maximum-likelihood model-fitting analyses were performed via Mx [35] to estimate genetic and environmental components of variance and to test the significance of their contributions to conformity. The genetic–environmental model (Supplementary Figure S1A) was compared to a series of reduced models by likelihood-ratio χ^2 and Akaike Information Criterion.

SNP-based heritability estimation

The discovery and replication cohorts were combined into one sample (2,130 individuals) for SNP-based heritability. The genetic relationships matrix (GRM) was built by estimating genetic relatedness of individual pairs using all autosomal markers with MAF >0.01. We excluded one of each pair of individuals with estimated genetic relatedness >0.025 [36]. The phenotypic variance explained by all SNPs was estimated by the software GCTA version 1.24 [37] using the GREML method, which utilizes the GRM and restricted maximum-likelihood modeling, with the first 20 eigenvectors from PCA included as covariates. Statistical power of estimation by GREML was calculated with the online GCTA-GREML Power Calculator [38]. All estimations have sufficient power over 0.95. Genetic covariance between phenotypes was estimated by bivariate GREML analysis implemented in GCTA.

Genome-wide association study

All GWAS were performed using PLINK [39], assuming a general linear model with a full model assumption, considering both additive effect and dominance component. For each marker, the model with a smaller p value was reported. The following covariates were included in the association tests for CQ_GWAS: gender for CONFM and INT_CONFM; PC2 and PC7 for adCONFM; PC7 for gCONF; PC2 for INT_adCONFM and INT_gCONF. Phenotype permutations and subsequent association testing were conducted with the adaptive permutation program implemented in PLINK. The full set of p values that emerged from association analysis was loaded and visualized in Haploview version 4.2 [40] to generate Manhattan plots. Basic statistical analysis and quantile–quantile (Q–Q) plots were conducted via R version 3.2.1. Quanto Version 1.2 was used for power calculation [41]. A second GWAS was also conducted on the replication cohort, and a joint GWAS of 2,127 was performed on the discovery and replication combined sample, both without covariates.

Candidate selection and replication

The genome-wide significance threshold (type I error considering multiple testing corrections) is 2.5×10^{-8}. Candidate SNPs were identified from CQ_GWAS using the following criteria: (1) $p < 1 \times 10^{-4}$ for both INT and untransformed phenotypes; (2) empirical $p < 1 \times 10^{-4}$ after permutation tests for markers found under an additive assumption; (3) in Hardy–Weinberg equilibrium ($p \geq 0.05$ in the standard asymmetric test implemented in PLINK); (4) within 500 kb around a gene. SNPs selected then served as candidates for all the subsequent analyses, including replication, joint analysis, and meta-analysis. A replication was performed on the CQ_REP and BJ_REP. The number of tested models and phenotypes was taken into consideration in the multiple testing corrections.

Imputation

Imputation of non-genotyped SNPs was performed in all samples with chip data (793 CQ_GWAS, 1,334 CQ_REP, and 499 GZ_GWAS). Genotypes were pre-phased into haplotypes with SHAPEIT [42], and imputation was then performed using IMPUTE2 v2.3.18 [43] with 1000 Genomes haplotype data (Phase I integrated variant set release (SHAPEIT2) in NCBI build 37 (NCBI37)/UCSC *hg19*) as the reference panel, producing 36,820,992 SNPs, 1,384,273 short bi-allelic indels, and 14,017 structural variations. Testing for association at the imputed SNPs was performed with SNPTEST v2.58 [44]. A conditional analysis, focused on the imputed SNPs within 1 Mb around

candidate regions, was performed based on a score method in the frequentist test framework. The association results of imputed genotypes were only used for regional association plots implemented in LocusZoom [45]. Quality control only kept SNPs with information >0.5 and MAF >0.05.

Meta-analysis

A meta-analysis of three GWAS (GZ_GWAS, CQ_GWAS, and CQ_REP, altogether 2,626) was performed using META_v1.6 [46] with inverse-variance method based on a fixed-effects model. GZ_GWAS used imputed genotypes with 4,245,043 SNPs, while CQ_GWAS and CQ_REP used raw genotypes without imputation. Only those SNPs that present in all three GWAS were considered for meta-analysis.

Gene-based and pathway-based analyses

Gene-based tests for association were carried out either using results from primary association tests with FAST-Gates [47] and VEGAS [48] or using raw genotype data with MAGMA [49]. The selection criteria for candidate gene at the discovery stage were $p < 0.01$. FAST-Gates was only adopted for gene-based analysis on the discovery and replication combined sample. The following pathway genomic methods were performed: GSA-SNP [50], MAGMA, and MAGENTA [51]. SNP to gene mapping was based on the NCBI36 (*hg18*) database except for MAGMA, which used NCBI37 (*hg19*). Gene boundaries were set at 20 kb for GSA-SNP, 50 kb for FAST-Gates, VEGAS, and MAGENTA. Reference for pathway analysis was made to Molecular Signatures Database (10,344 pathways) for MAGMA, GO, and KEGG for GSA-SNP, GO, KEGG, REACTOME, BIOCARTA, PANTHER, and INGENUITY (3,216 pathways in total) for MAGENTA; gene sets with 10—200 genes were included.

Functional MRI acquisition and analysis

Behavioral assays

The fMRI experiment took an event-related design. The behavioral assay was modified from the Price Estimation Conformity Test with a few changes to be compatible with fMRI experiments, such as fixed time and irrevocable choice during judgment. There were one run for practice and four runs for experiment; each run consists of seven conformity trials when the majority was conflicted with the subject in answers ("CONF"), nine non-conformity trials when the majority was consistent (denoted as "NON-CONF"), seven half-answer trials when half of the "others" agreed (denoted as "HALF"), and seven no-answer trials when there were only photos of the "others" without their answers (denoted as "NOANS"). Blank screen with white cross-fixation point was set between trials, with a presentation time of a random even number in the range of 2–10 s. Two "NONCONF" trials from each run were randomly chosen and discarded for balance of brain activity signals. Only the conformity testing stage was executed in the scanner. Trial types were randomized in the testing sequence, and the same testing sequence was applied to all subjects.

MRI data acquisition

Imaging was performed on a 3 Tesla Siemens Trio whole-body Magnetom scanner (Siemens AG, Erlangen, Germany) at the Beijing MRI Center for Brain Research. Head movement was restricted with padding. All images were acquired using a 32-channel head matrix coil. Three-dimensional T1-weighted high-resolution anatomical scans were acquired with the following parameters: three-dimensional magnetization-prepared rapid gradient-echo sequence, 176 scans, repetition time (TR) = 2.2 s, echo time (TE) = 3.4 ms, interval time (TI) = 1.1 s, total time (TA) = 421 s, field of vision (FoV) = 256 × 256, 256 × 256 matrix, voxel size = 1 × 1 × 1 mm^3. After performing automatic shimming and acquiring a scout image, we performed four runs of echo-planar imaging (EPI) to maximize the blood oxygenation level-dependent (BOLD) effect associated with neuronal activation (169 scans, TR = 2 s, TE = 30 ms, FoV = 1,152 × 1,152, 64 × 64 matrix, flip angle = 90°, 32 oblique slices without gap, voxel size = 3 × 3 × 3 mm^3, covering the whole cerebrum). The imaging protocol was identical for all subjects.

Image preprocessing

All image analysis was performed with Statistical Parametric Mapping (SPM12; University College London, London, UK). After discarding the first four dummy volumes, functional images were sequentially processed as follows: interpolated to correct for slice timing, realigned to the middle volume, co-registered to structural scans using the mean functional image, unwrapped, spatially normalized to a standard EPI template based on the Montreal Neurological Institute reference brain template (Asia brain), and spatially smoothed with an isotropic 8 mm full-width at half-maximum Gaussian kernel. Structural scans were segmented and spatially normalized to the same template as that used for functional scans.

Region of interest identification and statistical analysis

Contrasts were generated from the design matrix at the individual participant level with a high-pass filter cutoff of

128 s and then entered into a second-level analysis for statistical inference. Data for each condition were convolved with the canonical hemodynamic response using a random-effect general linear model and an unbiased whole-brain contrast was performed to search for regions where BOLD response was greater for all the conformity trials relative to all non-conformity/half trials (denoted as "CONF vs. NONCONF"/"CONF vs. HALF"; whole-brain family wise error-corrected p (P_{FWE}) < 0.05). Clusters showing significant whole-brain activation (cluster-level q_{FDR} < 0.05) were used as region of interests (ROIs) to constrain subsequent genotypic analysis [52]. For completeness, we additionally performed the same ROI analysis as above but included results derived from contrasting "NONCONF vs. HALF," where brain activity is not related with social conformity. Four second-level models, each assuming a specific mode of action for genetic variants, were fitted, regressing the mean BOLD signals in each ROI against genotypes. These second-level models were genotypic mode (analysis of variance (ANOVA): homozygote of minor allele, heterozygote, homozygote of major allele), additive mode (ADD: homozygote of minor allele = 2, heterozygote = 1, homozygote of major allele = 0), dominant mode (DOM: a t test contrasting the minor allele carrier group and the homozygote of major allele group), and recessive mode (REC: a t test contrasting the homozygote of minor allele group and the major allele carrier group). All statistical tests reported are two-sided. A linear regression analysis was further performed between behavioral outputs and ROI activation or genotypes. Allele frequency of all tested SNPs in the fMRI cohort met Hardy–Weinberg equilibrium (by χ^2 test). Results were visualized using xjView toolbox (http://www.alivelearn.net/xjview) and built-in visualization in SPM12.

Structural imaging analysis

Brain structural scans were analyzed using default procedures implemented in the FreeSurfer software package (http://freesurfer-software.org). Sixteen morphometric features (e.g., the volume and the curvature) of the superior frontal gyrus (SFG) and the medial frontal gyrus (MeFG), were modeled against genotypes with SNP-wise one-factor analysis of covariance.

Gene expression and brain activity correlation analysis

A gene differential search was done by contrasting the conformity-defined ROI (the SFG) to the whole brain using the online Allen Human Brain Atlas (http://human.brain-map.org). The correlation between gene expression and brain functional activation across the whole brain was calculated in Pearson's correlations by pairing β-value (from previous fMRI) of each MRI coordinate point with gene expression data at the proximal point, using the Student's t distribution with df = 839.

Results

Behavioral measures of conformity

Social conformity can be measured by the act of changing one's behavior to be in agreement with that of others. Here we have developed two behavioral assays modified from well-established conformity experiments, a Price Estimation Conformity Test and a Memory Conformity Test to ensure consistency of conformity measurement. Conformity was found to vary between individuals (0–100%). The average conformity was 75.5% for CONFM and 53.3% for CONFP (Table 1). The within-test reliability was high (for CONFP: Spearman's correlation coefficient $r = 0.68$, 95% confidence interval (CI) = 0.66–0.70, $p < 2.2 \times 10^{-16}$). Social conformity measured in CONFM was consistent with that in CONFP, showing both phenotypic correlation ($r = 0.42$, 95% CI: 0.39–0.45, $p < 2.2 \times 10^{-16}$) and genetic correlation ($r = 1.00$, standard error (SE) = 0.55). Thus, measures in our study were internally consistent. Analysis showed homogeneity of sampling for each cohort, with no significant effects ($p > 0.05$) on conformity exerted by demographic factors such as gender, age, ethnicity, or birthplace. Individual differences in confidence in one's own choices could not account for the individual differences measured in our assays as the correlation between one's conformity score and the percentage of "unsure" was very low ($r = -0.089$, $n = 2{,}448$). Besides, it was of low possibility that the studied inter-individual differences were mainly caused by the inner differences in changing one's opinion irrespective of others' opinions, because very high correlations were found between the measured conformity score and the same score subtracted by the amount of answer change under non-conformity condition ($r = 0.98$, $n = 2{,}230$).

Heritability estimation from twin studies

The final twin cohort consisted of 340 pairs of monozygotic (MZ), and 216 pairs of DZ adolescent twins (54% female) with an average age of 16 years. The MZ twin correlations of CONFP and CONFM (INT) were $r(338) = 0.36$ ($p < 0.001$) and $r(286) = 0.37$ ($p < 0.001$), respectively, while the corresponding DZ twin correlations were $r(214) = 0.24$ ($p < 0.001$) and $r(184) = 0.14$ ($p = 0.065$), respectively (Supplementary Table S2). There was gender difference in CONFP in the twin sample, thus we separate males and females in the analysis of CONFP. Correlation was higher for MZ than DZ twins, indicating genetic influences on

Table 2 Estimation of heritability using twin sample

Phenotype	χ^2	AIC	A	C/D	E	h^2
CONFP-female	3.623	−8.377	0.034	0.298	0.542	0.06
CONFP-male	5.836	−6.164	0.244	0.045	0.670	0.25
CONFM	0.429	−11.571	0.155	0.175	0.558	0.37

χ^2 is a parameter for testing a model's goodness of fit. For CONFP/CONFM, C/D effect was estimated

AIC Akaike Information Criterion, a statistical criterion for model selection, a model with smaller AIC being better, *A* additive genetic effect, *C* common environment effect, *D* dominant/non-additive genetic effect, *E* specific environmental effect, h^2 heritability, *CONFP* Price Estimation Conformity Test, *CONFM* Memory Conformity Test

conformity (Supplementary Figure S1). The pattern of twin correlation for CONFP fitted the classic ACE model, while that for CONFM fitted the ADE model. For CONFP, males and females differed in genetic contributions with heritability estimations of 0.25 (male) and 0.06 (female); for CONFM, the general heritability estimation was 0.37 (Table 2). These results indicate moderate genetic effects on conformity, in contrast to the recent conclusion of zero heritability by other researchers [33], though there might be gender differences.

SNP-based heritability estimation

Two thousand one hundred and thirty college students were retained after exclusion of related samples. Results showed that about 17.8–37.6% of the phenotypic variance in conformity could be explained by all autosomal SNPs (Supplementary Table S3). These results were consistent with those from our classic twin study and support potentially moderate genetic contributions to conformity.

Genome-wide survey of single-marker associations

Among the 830,937 tested markers, 107–317 were identified with $p < 10^{-4}$ in CQ_GWAS (Supplementary Table S4). The number of nominally significant ($p < 0.05$) markers was higher than that expected by chance, as illustrated by the Q–Q plots (Fig. 2a, b, Supplementary Figure S2B). The genomic inflation factor λ was 1.00 for all association tests, indicating the significance reported herein was not affected by population stratification. The association results across the whole genome are shown in "Manhattan" plots (Fig. 2c, d, Supplementary Figure S2C–H). Three of the tested markers, located in the gene *MIRLET7BHG* (MicroRNA Let-7b Host Gene), reached genome-wide significance ($p = (2.5–8.4) \times 10^{-9}$, Fig. 2d); however, they were not significant in the replication cohort ($p > 0.05$, Supplementary Table S4); nevertheless, conditional association analysis of imputed genotypes around this locus in the combined sample (2,130 participants) revealed two imputed SNPs with genome-wide significance (rs112585135 and rs118122886, see Supplemental Table S5). Further gene-based analysis showed the gene *MIRLET7BHG* was associated with several conformity-related phenotypes in the combined sample ($p =$ 0.013–0.035, by FAST_Gates, Supplementary Table S10). Additional GWAS was performed on the CQ_REP cohort (1,340 individuals) and revealed four SNPs (rs10892909, rs197581, rs78739647, rs6625475) that reached genome-wide significance (Supplementary Table S6), and these SNPs also reached $p < 10^{-3}$ in the joint GWAS. Two hundred and sixty-nine markers, corresponding to 190 loci, passed the selection rule and were analyzed further (Supplementary Table S4), and these markers accounted for half of the total heritability in conformity. Among them, 33 reached $p < 0.05$ in the replication GWAS and the joint GWAS. Twelve SNPs achieved false discovery rate-corrected significance level ($P_{FDR} < 0.05$) when analyzed jointly (Table 3).

The most significant locus in the joint analysis of the 12 SNPs was rs2381801 (effect size $\beta_{discovery} = -0.16$, see Table 3), which is located in the intron of protein tyrosine phosphatase receptor type D (*PTPRD*). Individuals carrying more minor alleles of this locus tend to be less conformed to the majority. The second most significant locus corresponded to three SNPs (rs770122, rs1479010, rs2619056) on 12q21.2, locating in the intron of the gene *NAV3* (neuron navigator 3). For each of the three candidate SNPs, carrying one minor allele would reduce conformity by 4% (Table 3). Regional association plots (Fig. 2e) displayed more association signals within this region in the joint GWAS (Supplementary Table S5); in addition, one imputed SNP rs61936251, in the intron of *NAV3*, was associated with gCONF with a genome-wide significant signal in the discovery GWAS and the joint GWAS ($p_{discovery} = 1.3 \times 10^{-8}$, $p_{replication} = 1.2 \times 10^{-4}$, $p_{joint} = 2.3 \times 10^{-9}$). *NAV3* was found at gene level in the joint analysis by FAST-Gates with a *p*-value of 0.0013.

A meta-analysis of the three GWAS in our study with a sample size of ~2,600 was further performed on the selected SNPs (Supplementary Table S7). Of the 111 common SNPs analyzed, 36 reached Bonferroni-corrected combined $p < 0.05$, 10 reached genome-wide significance; of the 36 top candidates, five SNPs, rs7709420 ($p_{meta} = 4.3 \times 10^{-25}$), rs244515 ($p_{meta} = 4.0 \times 10^{-20}$), rs1479010 ($p_{meta} = 2.0 \times 10^{-7}$), rs7589342 ($p_{meta} = 1.7 \times 10^{-5}$), and rs12053259 ($p_{meta} = 4.1 \times 10^{-5}$), were significant in the joint GWAS and two of the three independent cohorts. It is notable that rs1479010 is in the intron of *NAV3*, further supporting its association with social conformity.

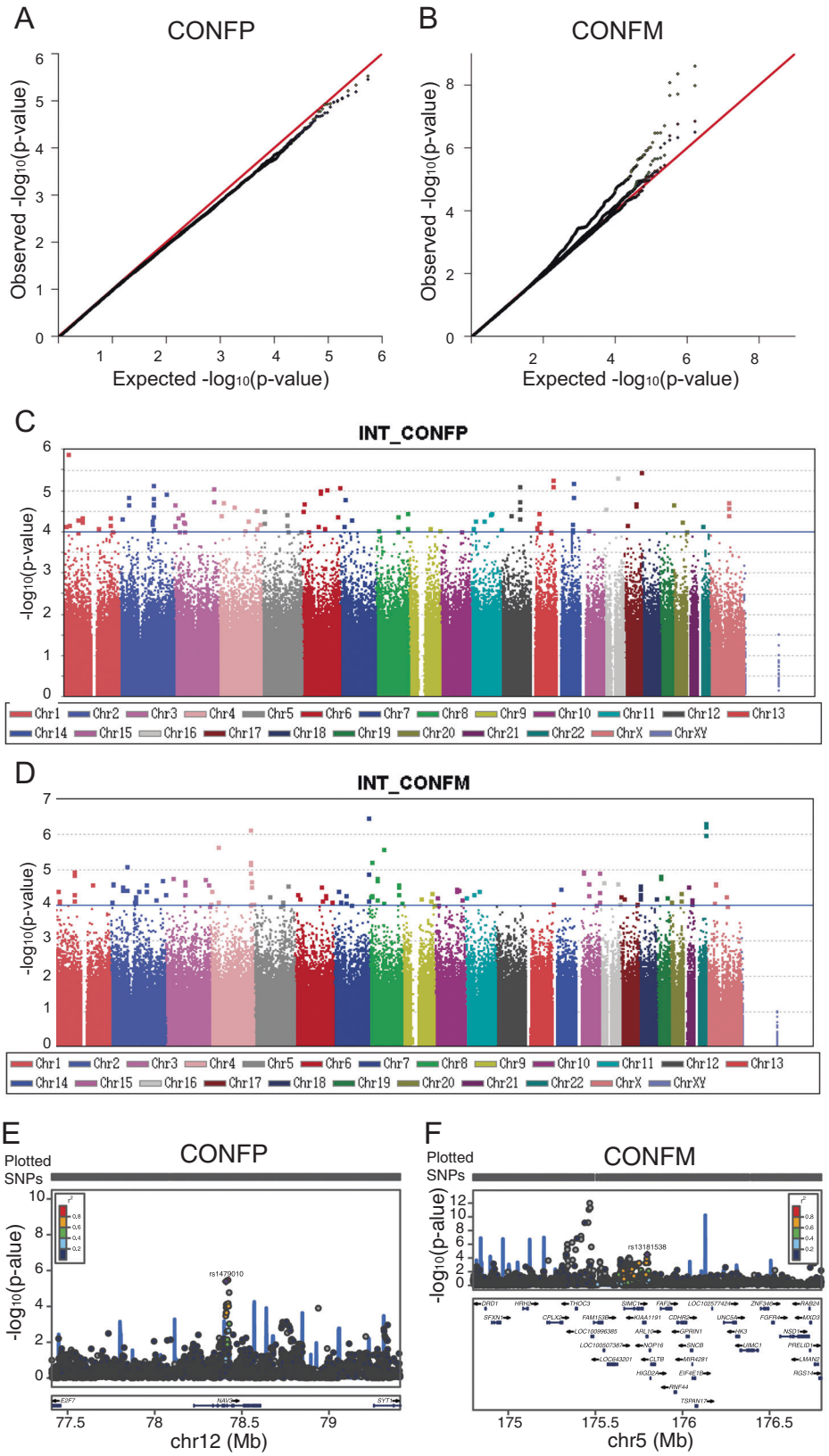

Fig. 2 Plots of phenotype–genotype association results across the genome. Quantile–quantile (Q–Q) plots of GWAS results for **a** CONFP and **b** CONFM. Observed *p*-values relative to expected were plotted based on *p*-values calculated using linear regression and including significant principal components as covariates; the red line indicates the null hypothesis of no association; yellow and green dots represent association results from untransformed raw data; blue and red dots represent association results from inverse normally transformed data; yellow and blue dots are results of adCONFM; green and red dots are results of CONFM. Figures were generated with the R package. Manhattan plots of GWAS results ($-\log_{10}$ *p*-values) are shown in chromosomal order for individually genotyped quality control-positive SNPs that were tested for linear regression with **c** CONFP and **d** CONFM in the discovery sample. Association tests considered both additive and dominant models; the *p*-values plotted were the smaller ones under these two genetic models. The blue line indicates the suggestive significance level (10^{-4}). Chromosomes are shown in different colors for clarity. Regional association plots for top successfully replicated SNPs, **e** rs770122, rs1479010, and rs2619056 (the gene *NAV3*), **f** rs13170785 (the gene *ARL10*). Genome Build is *hg*19/1000 Genomes Nov 2014 ASN. Data for association plots come from the joint analysis of imputed genotypes from 2,130 individuals. Figures were plotted with the web-based LocusZoom program

Table 3 Top loci in association with conformity

SNP	Proxy gene	β_{dis}	P_{dis}	β_{rep}	P_{rep}	β_{join}	P_{join}	Genotype (mean±SD)		
rs2381801	*PTPRD* intron	−0.16	4.3×10^{-6}	−0.07	0.013	−0.10	4.8×10^{-6}	G/G	G/C	C/C
							N=	457	1,007	654
							INT_adCONFM	-0.27 ± 0.83	-0.11 ± 0.86	-0.03 ± 0.86
rs2619056	*NAV3* intron	−0.16	1.2×10^{-5}	−0.06	0.018	−0.10	8.1×10^{-6}	A/A	A/G	G/G
							N=	390	999	729
							CONFP	0.51 ± 0.29	0.57 ± 0.29	0.60 ± 0.29
rs2056708	*CCDC146* intron	−0.14	7.6×10^{-5}	−0.06	0.034	−0.09	1.8×10^{-5}	A/A	A/G	G/G
							N=	530	1,027	554
							INT_gCONF	-0.02 ± 0.83	-0.21 ± 0.85	-0.08 ± 0.91
rs11257080	*PFKFB3* intron	−0.26	1.8×10^{-5}	−0.10	0.044	−0.17	2.1×10^{-5}	A/A	A/G	G/G
							N=	33	418	1,667
							gCONF	0.90 ± 0.38	1.18 ± 0.35	1.17 ± 0.35
rs10062113	*CTNND2* 81 kb upstream	0.17	2.1×10^{-5}	0.06	0.027	0.10	2.5×10^{-5}	A/A	A/G	G/G
							N=	196	894	1,025
							CONFP	0.48 ± 0.28	0.59 ± 0.29	0.57 ± 0.30
rs13170785	*ARL10* intron	0.15	3.9×10^{-5}	0.06	0.032	0.09	4.7×10^{-5}	G/G	G/A	A/A
							N=	364	1,017	739
							INT_CONFM	-0.20 ± 0.87	-0.05 ± 0.81	-0.20 ± 0.87
rs56324903	*IGSF21* 44 kb downstream	0.36	1.6×10^{-5}	0.12	0.038	0.19	5.8×10^{-5}	A/A	A/G	G/G
							N=	21	403	1,694
							gCONF	0.90 ± 0.45	1.20 ± 0.34	1.16 ± 0.35
rs3929673	*PRDM2* 201 kb downstream	0.15	1.8×10^{-5}	0.06	0.038	0.09	6.0×10^{-5}	C/C	C/A	A/A
							N=	437	1,069	612
							INT_CONFM	-0.20 ± 0.83	-0.06 ± 0.83	-0.21 ± 0.87
rs2614595	*LINC01098* 498 kb downstream	−0.16	1.6×10^{-5}	−0.06	0.038	−0.09	1.1×10^{-4}	A/A	A/G	G/G
							N=	251	951	905
							CONFM	0.74 ± 0.24	0.79 ± 0.22	0.79 ± 0.21

Behavioral outputs shown for each locus are the phenotypes with the most significant associated signals in the joint analysis. Values here are mean ± standard deviation (SD). For each locus, only the SNP with the most significant signals is shown. $\beta_{dis/rep/joint}$ β values (effect size) in the discovery/replication/joint analysis, $P_{dis/rep/joint}$ p-values in the discovery/replication/joint analysis, *N* the sample size under the particular genotype, *INT* inverse normal transformation, *PTPRD* protein tyrosine phosphatase receptor type D, *NAV3* neuron navigator 3, *CCDC146* coiled-coil domain containing 146, *PFKFB3* 6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3, *CTNND2* catenin delta 2, *ARL10* ADP ribosylation factor-like GTPase 10, *IGSF21* immunoglobin superfamily member 21, *PRDM2* PR/SET domain 2, *CONFP* Price Estimation Conformity Test, *CONFM* Memory Conformity Test, *SNP* single-nucleotide polymorphism

Gene-level and pathway-level association analyses

Gene-based tests can be more powerful than single-marker level association because of the intermingling of weak signals in a gene with random noise in GWAS and the relatively smaller number of tested genes. We adopted three software programs, MAGMA (Supplementary Table S8), VEGAS (Supplementary Table S9), and FAST-Gates (Supplementary Table S10) for gene-based analysis. MAGMA found one gene, *TMEM173* (transmembrane protein 173), reaching genome-wide significance with a p-value of 3.1×10^{-6} and it was replicated with $p < 0.05$ at the replication stage and $P_{FDR} < 0.05$ in the joint analysis for all memory-based conformity phenotypes. Eleven genes were discovered by both two software programs and replicated by at least one program (Supplementary Table S11). Among them, *ASPH* (aspartate beta-hydroxylase) displayed suggestive significance at the discovery stage via the SNP-level GWAS. Haplotype-based analysis gave similar results (see Supplementary Text and Table S12 for details).

It is likely that social conformity is a complex phenotype so that multiple genes in the genome, each with a modest effect, contribute to phenotypic associations. Thus, we carried out association analysis at pathway level, taking advantage of three software programs, MAGMA (Supplementary Table S13), GSA-SNP (Supplementary Table S14), and MAGENTA (Supplementary Table S15). One gene set discovered by MAGMA, chr5q31, reached genome-wide significance ($p < 4.8 \times 10^{-6}$). 57 gene sets were associated

Table 4 Candidate loci with significant genotypic difference (FDR-corrected) in the brain activity of social conformity-related brain regions

Contrast	SNP	Gene	Model	P_{beh}	ROI	Brain region	*t*/*F*	*P*
CONF vs. NONCONF	rs13181538	*ARL10*	DOM	0.008	[63, −28, −10]	Middle temporal gyrus	3	0.0019
			ADD	0.025	[−54, −61, 32]	Supramarginal gyrus	2.94	0.0023
	rs2448226	*NAV3*	ANOVA	0.941	[−15, 5, 11]	Extranuclear	6.77	0.0022
	rs11133644	*CTNND2*	ANOVA	0.134	[−42, 47, −7]	Middle frontal gyrus	6.93	0.0019
CONF vs. HALF					[−30, 17, −13]	Inferior frontal gyrus	7.28	0.0015
					[−21, 2, 56]	Subgyral	7.86	0.0009

CONF vs. NONCONF/CONF vs. HALF, contrasting "CONF" condition with "NONCONF"/"HALF" condition

MODEL genetic model being tested, *SNP* single-nucleotide polymorphism, *ADD* additive, *DOM* dominant, *ANOVA* analysis of variance, P_{beh} *p*-values of associations between genotypes and behavioral outputs in the fMRI sample, *ROI* ROI names in terms of the coordinates of their peak voxel, based on MNI templates, *t*/*F* *t*-test or *F*-test statistic; *P* uncorrected *p*-values of ROI analysis, *CONFP* Price Estimation Conformity Test, *CONFM* Memory Conformity Test, *MNI* Montreal Neurological Institute, *ROI* region of interest

with conformity with $p < 0.05$ in both the discovery and replication cohorts by MAGMA; the gene *NAV3* was included in one of these gene sets. The three programs discovered 18 pathways that were candidates by at least two programs and were further successfully replicated by at least one program ($p < 0.05$, Supplementary Table S16). The 18 pathways include the gene *TMEM173* whose association with conformity was revealed by haplotype-level and gene-level GWAS.

Functional validation of candidate loci via brain imaging

To further investigate the functional significance of loci identified by genetic association, fMRI was conducted. The fMRI assay was revised from the Price Estimation Conformity Test whose behavioral score was closer to normal distribution than that from memory-based behavioral assays. All 64 fMRI subjects showed conformity to some degree and their average conforming tendency was similar to that in the discovery and replication cohorts (Table 1).

A one-sample *F* test at the group level on the brain activation patterns of all subjects, by contrasting the "conformity" condition to the "non-conformity" condition, detected 26 clusters showing significant correlations ($P_{FWE} < 0.05$); a contrast of the "conformity" condition to the "half" condition produced a very similar brain activation map with 30 significant clusters at the whole brain level (Supplementary Figure S3 and Supplementary Table S17). The most significantly activated clusters (the contrast "CONF vs. NONCONF": centering at [−3, 32, 35], $F = 145.49$; the contrast "CONF vs. HALF": centering at [0, 47, 20], $F = 153.81$) across the whole brain are located in the Frontal_Sup_Medial_L (in Automated Anatomical Labeling human brain atlas) of the MeFG. Most significantly activated voxels ($P_{FWE} < 0.05$) were in the SFG, the MeFG, and the middle frontal gyrus (MidFG) (Supplementary Figure S3D, E and Supplementary Table S18). The ROIs were defined as the clusters significantly correlated with conformity-specific conditions (cluster-level $q_{FDR} < 0.05$; 20 ROIs for the contrast "CONF vs. NONCONF," 18 ROIs for the contrast "CONF vs. HALF"); the subsequent genotypic differential brain activation analysis was confined to these ROIs. Significant differences ($p < 0.05$) in brain activation within these ROIs between genotypes were found for all the 10 tested SNPs (all of which were candidate loci in our GWAS); three SNPs (rs2448226, rs11133644, and rs13181538), linked to the genes *NAV3*, *CTNND2* (catenin delta 2), and *ARL10* (ADP ribosylation factor-like GTPase 10), remained significant after Bonferroni correction for multiple testing (Table 4); notably, nominal significance in differential activations between genotypes in both two groups of ROIs were found for seven SNPs linked to the gene for *CTNND2*, *PTPRD*, *ARL10*, and *NAV3* (Supplementary Table S19). Five *NAV3* intron SNPs within 135 kb around the loci shown the most significant SNP-level association signals were tested and all of them showed genotypic differences with $p < 0.05$ in brain activation in at least one ROI; one *NAV3* SNP rs2448226 was associated with conformity-related brain activity in the most significantly activated ROI; significant ($P_{FWE} < 0.05$) genotypic differences at the whole brain level were found in the brain activity within the clusters located in the conformity-related regions for two *NAV3* SNPs (Supplementary Table S20). These results above further support the involvement of *NAV3* in social conformity. Most of the tested genotypes were associated with anatomical features of the conformity-related brain regions, for example, *NAV3* SNPs were significantly associated with the curvature of the MidFG and *CTNND2* SNP was weakly correlated with the gray matter volume of the MidFG and SFG ($p < 0.05$) (Supplementary Table S21). Furthermore, significant ($p < 0.05$) differences between genotypes of rs1566622 (near *CTNND2*) and rs13181538 (near *ARL10*) in behavioral performance were detected in the fMRI cohort. However,

no significant link was found between activities in these defined ROIs and the individual level of conformity across subjects.

Gene expression analysis was performed with human brain microarray data from the Allen Brain Atlas (http://human.brain-map.org). Of the top candidate genes, *CTNND2*, *ARL10*, and *NAV3*, showed significantly correlated ($P_{\text{Bonferroni-corrected}} < 0.05$) expression patterns with fMRI brain activities across the whole brain (Supplementary Table S22).

Discussion

The two behavioral assays used in the research were modified from well-established ones [26, 27, 53]. Memory conformity has been extensively studied [12, 23, 53, 54]. The two behavioral assays together could reduce potential unknown confounding factors brought on by each assay. A combined factor from PCA, gCONF, can explain 73% of all the phenotypic variances, supporting that gCONF represents the major part of one's conforming tendency. It also supports the idea that memory is not a confounder for memory-based conformity.

The heritability found in CONFM is indicative of heredity in conformity rather than memory because: (1) the correlation between the percentage of answers changed under conformity condition and that under memory control condition was low ($r = 0.04$, $p = 0.01$), indicating memory performance could not explain the major variance in conformity; (2) after adjustment by memory performance (adCONFM), conformity still achieved a similar heritability as that derived from unadjusted one; (3) CONFM was correlated to CONFP ($r = 0.42$).

Most scientists regard conformity as mainly determined by environmental factors, such as culture [20], early family experience [22], and social power [24]. A recent study has attempted to study genetic contributions to conformity, but came up with a zero heritability conclusion [33]. A major caveat in that study is the behavioral paradigm used to measure conformity: they defined conformity as the choice of a pen of the "majority" color. In this paradigm, there was neither informational nor normative conflict with the majority of humans, thus lacking the usual necessary conditions for conformity [2, 19]. Furthermore, the authors only relied on subjects' self-report of non-conformity to validate assay reliability. However, we have found poor correlations between self-reported conformity and actual behavior readouts ($r_{\text{CONFP}} = 0.41$, $r_{\text{CONFM}} = 0.24$, $p < 0.001$), which means that variance in subjective report could only explain less than 20% of the variance in behavior. The discrepancy between self-report and actual behavior indicates that individuals were either unaware of, or reluctant to admit, their real tendency in social behaviors. It could explain the smaller genetic effects on self-reported conformity found by two twin studies (16% and 22%) [55, 56] comparing to ours (~37%). Several other cognitive behaviors, such as prosocial behavior [57], aggression [32], general intelligence [58], and personality [59] have already been found to be highly heritable, for example, 42% of social responsibility [57] and 56% of altruism [32] can be explained by genetic effects. However, most previous studies about heritability were based on questionnaires rather than behavioral experiments. Our finding of poor correlation between immediate after-test self-report and actual behavior suggests inaccuracies of self-report.

Combining association analysis at multiple levels, we discovered two genes, namely, *NAV3* and *CTNND2*, in strong associations with social conformity tendency (Supplementary Table S23). They displayed replicable association signals with $P_{\text{FDR}} < 0.05$ at single-marker GWAS and $p < 0.05$ at gene-level and pathway-level analyses; their SNPs showed genotypic differences in fMRI signal and anatomical features of conformity-related brain regions; and their mRNAs and proteins are enriched in the conformity-related brain regions with more than 1.3-folds change of gene expression in the SFG (data from Allen Human Brain Atlas). Furthermore, whole brain gene expression patterns of both two genes were significantly correlated with brain activation map during conformity behavior ($p < 0.05$), providing more evidence for their involvement in human conformity behavior. *NAV3*'s family member *NAV2* was also found by gene-level and pathway-level analyses with $p < 0.05$. NAV3 protein, expressed predominantly in the nervous system and enriched in synaptic regions, is an evolutionally conserved protein with functional importance in axon outgrowth and guidance conserved from *Caenorhabditis elegans* to mammals [60, 61]. *CTNND2* plays an important role in cortical development and differentiation, synapse formation and growth [62]. Previous reports have shown associations between *CTNND2* and psychiatric disorders such as anxiety [63] and schizophrenia [64].

Previous studies have implicated the dopaminergic pathway in social conformity [65–67]. At the discovery stage of our analysis, SNPs in most of dopamine-related genes showed nominally significant ($p < 0.05$) signals of association (Supplementary Table S24). Among them, four SNPs in *DRD2* (dopamine receptor D2) were replicated with $p < 0.05$ in the replication cohort and in the joint analysis; *COMT* (catechol-*O*-methyltransferase), *DBH* (dopamine β-hydroxylase), and *DDC* (Dopa decarboxylase) genes have conformity-related SNPs with nominal significance in the joint analysis. Gene-level analysis in the combined cohorts found that most genes in the dopamine pathway showed significant signals of association ($p < 0.05$) with conformity-related phenotypes

(Supplementary Table S25); of these genes, the association signals of *DRD2* remained significant ($p < 0.05$) after Bonferroni correction of multiple testing. Our results support the involvement of D_2 receptor in social conformity [65–67]. Oxytocin was also implicated in social conformity [68, 69]. Analysis of *OXT* and *OXTR* (genes encoding oxytocin and its receptor) in our cohorts showed positive associations by at least one gene-level analysis.

Our fMRI showed conformity-related brain activation in the superior medial prefrontal area (Supplementary Figure S3), similar to that in previous fMRI studies of conformity [26, 70, 71]. Other researchers have reported that these brain areas were activated during cognitive control when faced with conflicts [72], involved in social and emotional information processing [73]. However, the current study could not distinguish social conflict from simple decision conflict, which requests more delicate experimental design in the future.

Social conformity has been reported to vary among cultures [8, 14, 20, 21]. Whether the uncovered genes in this paper contribute similarly to conformity in other cultures is unknown. Furthermore, allele frequency can differ between populations, so those genetic components that affect conformity in the studied population may not exert similar influence on other diverse populations. Considering the majority of our samples are Han Chinese but there was possibly slight population stratification, the generalizability of this research remained to be studied. Clearly, environmental factors are important, as further supported by our study.

The power is enough to find very common variants with large effect size, but to discover more rare variants with small effect size needs much larger samples which requires more labs in collaboration to achieve.

In summary, this paper is the first to uncover genetic contributions to human social conformity; it has identified specific genes including *NAV3* to be involved in conformity with genomic association approach and functional validation via brain imaging. These results support the use of genetic analysis in studying human cognition.

Significance statement

Social conformity is important both for individuals and societies. It has long been thought to be influenced only by environment but not by genes. This is the first systematic genetic study of social conformity. Through twin studies and multi-level genomic analyses, we have revealed potential heritable contribution to social conformity and uncovered several genetic loci highly associated with individual differences in conformity. Further brain imaging study provides functional validation for the results.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Project 31421003); the Beijing Advanced Innovation Center for Genomics at Peking University; the Peking-Tsinghua Center for Life Sciences; the Applied Development Program from the Science and Technology Committee of Chongqing (Grant number cstc2014yykfB10003, cstc2015shms-ztzx10006); and the Program of Mass Creativities Workshops from the Science and Technology Committee of Chongqing. We are grateful to Dr. Chen Wu, Dr. Jurg Ott, and Dr. Houfeng Zheng for comments on the manuscript, and to Zhangyan Guan and Huizhen Yang for help with DNA preparation.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Sherif M. The Psychology of Social Norms. New York, NY, USA: Harper & Brothers; 1936.
2. Asch SE. in Effects of group pressure upon the modification and distortion of judgements. (ed.) Guetzkow H. Groups, leadership and men. Pittsburgh, PA, USA: Carnegie Press; 1951. p. 177–190.
3. Asch SE. Studies of independence and conformity: 1. A minority of one against a unanimous majority. Psychol Monogr. 1956;70:1–70.
4. Crutchfield RS. Conformity and character. Am Psychol. 1955;10:191–8.
5. Cialdini RB, Goldstein NJ. Social influence: compliance and conformity. Annu Rev Psychol. 2004;55:591–621.
6. Grofman B, Feld SL. Rousseaus general will—a condorcetian perspective. Am Polit Sci Rev. 1988;82:567–76.
7. Kendal J, Giraldeau L, Laland K. The evolution of social learning rules: payoff-biased and frequency-dependent biased transmission. J Theor Biol. 2009;260:210–9.
8. Bond R, Smith PB. Culture and conformity: a meta-analysis of studies using Asch's (1952b, 1956) Line judgment task. Psychol Bull. 1996;119:111–37.
9. Raafat RM, Chater N, Frith C. Herding in humans. Trends Cogn Sci. 2009;13:420–8.
10. Stein R. The pull of the group: conscious conflict and the involuntary tendency towards conformity. Conscious Cogn. 2013;22:788–94.
11. Berns GS, Chappelow J, Zink CF, Pagnoni G, Martin-Skurski ME, Richards J. Neurobiological correlates of social conformity and independence during mental rotation. Biol Psychiatry. 2005;58:245–53.
12. Edelson MG, Sharot T, Dolan RJ, Dudai Y. Following the crowd: brain substrates of long-term memory conformity. Science. 2011;333:108–11.
13. Galef B, Whiskin E. 'Conformity' in Norway rats? Anim Behav. 2008;75:2035–9.
14. Claidiere N, Whiten A. Integrating the study of conformity and culture in humans and nonhuman animals. Psychol Bull. 2012;138:126–45.
15. Whiten A, Horner V, de Waal FB. Conformity to cultural norms of tool use in chimpanzees. Nature. 2005;437:737–40.
16. van de Waal E, Borgeaud C, Whiten A. Potent social learning and conformity shape a wild primate's foraging decisions. Science. 2013;340:483–5.
17. Corriveau KH, Harris PL. Preschoolers (sometimes) defer to the majority in making simple perceptual judgments. Dev Psychol. 2010;46:437–45.

18. Powell LJ, Spelke ES. Preverbal infants expect members of social groups to act alike. Proc Natl Acad Sci USA. 2013;110: E3965–3972.
19. Kim H, Markus HR. Deviance or uniqueness, harmony or conformity? A cultural analysis. J Pers Soc Psychol. 1999;77:785–800.
20. Triandis HC. in Cross-cultural studies of individualism and collectivism. (ed.) Berman JJ. Nebraska symposium on motivation. Lincoln, NE, USA: University of Nebraska Press; 1989. p. 41–133.
21. Takano Y, Sogon S. Are Japanese more collectivistic than Americans? Examining conformity in in-groups and the reference-group effect. J Cross Cult Psychol. 2008;39:237–50.
22. Kasser T, Koestner R, Lekes N. Early family experiences and adult values: a 26-year, prospective longitudinal study. Pers Soc Psychol Bull. 2002;28:826–35.
23. Wright DB, London K, Waechter M. Social anxiety moderates memory conformity in adolescents. Appl Cogn Psychol. 2010;24:1034–45.
24. Galinsky AD, Magee JC, Gruenfeld DH, Whitson JA, Liljenquist KA. Power reduces the press of the situation: implications for creativity, conformity, and dissonance. J Pers Soc Psychol. 2008;95:1450–66.
25. de Young CG, Peterson JB, Higgins DM. Higher-order factors of the Big Five predict conformity: are there neuroses of health? Pers Indiv Differ. 2002;33:533–52.
26. Klucharev V, Hytonen K, Rijpkema M, Smidts A, Fernandez G. Reinforcement learning signal predicts social conformity. Neuron. 2009;61:140–51.
27. Campbell-Meiklejohn DK, Bach DR, Roepstorff A, Dolan RJ, Frith CD. How the opinion of others affects our valuation of objects. Curr Biol. 2010;20:1165–70.
28. Izuma K, Adolphs R. Social manipulation of preference in the human brain. Neuron. 2013;78:563–73.
29. Stallen M, Smidts A, Sanfey AG. Peer influence: neural mechanisms underlying in-group conformity. Front Hum Neurosci. 2013;7:1–7.
30. Klucharev V, Munneke MAM, Smidts A, Fernandez G. Down regulation of the posterior medial frontal cortex prevents social conformity. J Neurosci. 2011;31:11934–40.
31. Constantino JN, Todd RD. Genetic structure of reciprocal social behavior. Am J Psychiatry. 2000;157:2043–5.
32. Rushton JP, Fulker DW, Neale MC, Nias DKB, Eysenck HJ. Altuism and aggression: the heritability of individual differences. J Pers Soc Psychol. 1986;50:1192–8.
33. Li XT, Zhang JD, Huang Y, Xu M, Liu J. Nurtured to follow the crowd: a twin study on conformity. Chin Sci Bull. 2013;58:1175–80.
34. 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;38:904–9.
35. Neale MC, Boker SM, Xie G, Maes HH. Mx: Statistical modeling. Richmond, VA, USA: Department of Psychiatry, Virginia Commonwealth University; 2003.
36. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. Nat Genet. 2010;42:565–9.
37. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet. 2011;88:76–82.
38. Visscher PM, Hemani G, Vinkhuyzen AA, Chen GB, Lee SH, Wray NR, et al. Statistical power to detect genetic (co)variance of complex traits using SNP data in unrelated samples. PLoS Genet. 2014;10:e1004269.
39. 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;81:559–75.
40. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005;21:263–5.
41. Gauderman WJ, Morrison JM. QUANTO 1.1: a computer program for power and sample size calculations for genetic-epidemiology studies. 2006. http://hydra.usc.edu/gxe.
42. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nat Genet. 2012;44:955–9.
43. 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:e1000529.
.44. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies via imputation of genotypes. Nat Genet. 2007;39:906–13.
45. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010;26:2336–7.
46. Liu JZ, Tozzi F, Waterworth DM, Pillai SG, Muglia P, Middleton L, et al. Meta-analysis and imputation refines the association of 15q25 with smoking quantity. Nat Genet. 2010;42:436–40.
47. Chanda P, Huang H, Arking DE, Bader JS. Fast association tests for genes with FAST. PLoS ONE. 2013;8:e68585.
48. Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM, et al. A versatile gene-based test for genome-wide association studies. Am J Hum Genet. 2010;87:139–45.
49. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: Generalized gene-set analysis of GWAS data. PLoS Comput Biol. 2015;11:e1004219.
50. Nam D, Kim J, Kim SY, Kim S. GSA-SNP: a general approach for gene set analysis of polymorphisms. Nucleic Acids Res. 2010;38:W749–W754.
51. Segrè AV, DIAGRAM Consortium, MAGIC investigators, Groop, L., Mootha VK, Daly MJ, et al. Common inherited variation in mitochondrial genes is not enriched for associations with Type 2 diabetes or related glycemic traits. PLoS Genet. 2010;6: e1001058.
52. Brett M, Jean-Luc A, Valabregue R, Jean-Baptiste P. Region of interest analysis using an SPM toolbox. Presented at the 8th International Conference on Functional Mapping of the Human Brain, 2–6 June 2002, Sendai, Japan. Available in CD-ROM in *Neuroimage* 2002;**16** (abstract).
53. Reysen M. The effects of conformity on recognition judgements. Memory. 2005;13:87–94.
54. Horry R, Palmer MA, Sexton M, Brewer N. Memory conformity for confidently recognized items: the power of social influence on memory reports. J Exp Soc Psychol. 2012;48:783–6.
55. Agrawal A, Dick DM, Bucholz KK, Madden PAF, Cooper ML, Sher KJ, et al. Drinking expectancies and motives: a genetic study of young adult women. Addiction. 2008;103:194–204.
56. Kristjansson SD, Agrawal A, Littlefield AK, Pergadia ML, Lessov-Schlaggar CN, Sartor CE, et al. Drinking motives in female smokers: factor structure, alcohol dependence, and genetic influences. Alcohol Clin Exp Res. 2011;35:345–54.
57. Rushton JP. Genetic and environmental contributions to pro-social attitudes: a twin study of social responsibility. Proc Biol Sci. 2004;271:2583–5.
58. Jacobs N, Gestel SV, Derom C, Thiery E, Vernon P, Derom R, et al. Heritability estimates of intelligence in twins: effect of chorion type. Behav Genet. 2001;31:209–17.
59. Jang KL, Livesley WJ, Vernon PA. Heritability of the big five personality dimensions and their facets: a twin study. J Pers. 1996;64:577–91.

60. Maes T, Barcelo A, Buesa C. Neuron navigator: a human gene family with homology to unc-53, a cell guidance gene from *Caenorhabditis elegans*. Genomics. 2002;80:21–30.
61. Peeters PJ, Baker AK, Goris I, Daneels G, Verhasselt P, Luyten WH, et al. Sensory deficits in mice hypomorphic for a mammalian homologue of unc-53. Brain Res Dev Brain Res. 2004;150:89–101.
62. Kosik KS, Donahue CP, Israely I, Liu X, Ochiishi T. Delta-catenin at the synaptic-adherens junction. Trends Cell Biol. 2005;15:172–8.
63. Nivard MG, Mbarek H, Hottenga JJ, Smit JH, Jansen R, Penninx BW, et al. Further confirmation of the association between anxiety and CTNND2: replication in humans. Genes Brain Behav. 2014;13:195–201.
64. Vrijenhoek T, Buizer-Voskamp JE, van der Stelt I, Strengman E, Genetic Risk and Outcome in Psychosis (GROUP) Consortium, Sabatti CG, et al. Recurrent CNVs disrupt three candidate genes in schizophrenia patients. Am J Hum Genet. 2008;83:504–10.
65. Cervenka S, Gustavsson JP, Halldin C, Farde L. Association between striatal and extrastriatal dopamine D2-receptor binding and social desirability. Neuroimage. 2010;50:323–8.
66. Egerton A, Rees E, Bose SK, Lappin J, Stokes PR, Turkheimer FE, et al. Truth, lies or self-deception? Striatal D(2/3) receptor availability predicts individual differences in social conformity. Neuroimage. 2010;53:777–81.
67. Campbell-Meiklejohn DK, Simonsen A, Jensen M, Wohlert V, Gjerloff T, Scheel-Kruger J, et al. Modulation of social influence by methylphenidate. Neuropsychopharmacology. 2012;37:1517–25.
68. Stallen M, de Dreu CK, Shalvi S, Smidts A, Sanfey AG. The herding hormone oxytocin stimulates in-group conformity. Psychol Sci. 2012;23:1288–92.
69. Huang Y, Kendrick KM, Zheng H, Yu R. Oxytocin enhances implicit social conformity to both in-group and out-group opinions. Psychoneuroendocrinology. 2015;60:114–9.
70. Wei Z, Zhao Z, Zheng Y. Neural mechanisms underlying social conformity in an ultimatum game. Front Hum Neurosci. 2013;7:896.
71. Edelson MG, Dudai Y, Dolan RJ, Sharot T. Brain substrates of recovery from misleading influence. J Neurosci. 2014;34:7744–53.
72. Milham MP, Banich MT. Anterior cingulate cortex: an fMRI analysis of conflict specificity and functional differentiation. Hum Brain Mapp. 2005;25:328–35.
73. Rudie JD, Shehzad Z, Hernandez LM, Colich NL, Bookheimer SY, Iacoboni M, et al. Reduced functional integration and segregation of distributed neural systems underlying social and emotional information processing in autism spectrum disorders. Cereb Cortex. 2012;22:1025–37.