Human Genetics of Face Recognition: Discovery of *MCTP2* Mutations in Humans with Face Blindness (Congenital Prosopagnosia)

Yun Sun[1,2], Weiwei Men[3], Wan Fang[1], Enxing Zhou[1], Wei Yang[1], Zhiqiang Li[4], Hou-Feng Zheng[5], and Yi Rao[1*]

[1]Peking-Tsinghua Center for Life Sciences, PKU-IDG/McGovern Institute for Brain Research, School of Life Sciences, Peking University; Chinese Institute for Brain Research, Beijing; Capital Medical University, Beijing, China.

[2]Institute of Molecular Physiology, Shenzhen Bay Laboratory, Shenzhen, China

[3]Center for MRI Research, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China; Beijing Key Lab for Medical Physics and Engineering, Institute of Heavy Ion Physics, School of Physics, Peking University, Beijing 100871, China.

[4]Bio-X Institute, Ministry of Education Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders, the Collaborative Innovation Center for Brain Science, Shanghai Jiao Tong University, Shanghai 200030, China.

[5]Westlake University School of Life Sciences, Hangzhou, Zhejiang, China.

*Corresponding Author (email: yrao@pku.edu.cn)

SUMMARY

Face recognition is important for both visual and social cognition. While congenital prosopagnosia (CP) or face blindness has been known for seven decades and electrophysiological studies have characterized face specific neurons for half a century, no molecular analyses have been undertaken. Here we report results of research combining classic genetics and modern genomics involving over 7000 subjects. From a large family with 18 CP members, we uncovered a fully co-segregating private mutation in the *MCTP2* gene which encodes a calcium binding transmembrane protein expressed in the central nervous system. More rare mutations in *MCTP2* were detected in CP families screened from a cohort of 2904 individuals and were also associated with CPs in a second cohort of 1928. In a third cohort of 1757, face recognition was different between 14 carriers with a frameshift mutation S80fs in *MCTP2* and 19 non-carrying volunteers. 6 families including one with 10 members showed the S80fs-CP correlation. Functional magnetic resonance imaging (fMRI) indicates that impaired recognition of individual faces by CPs with the *MCTP2* mutations is associated with inability to recognize the same faces in the right fusiform face area (rFFA). Our results have revealed the genetic predisposition of *MCTP2* mutations in CP, 74 years after the initial report of CP. This is the first time a gene required for a higher form of visual social cognition was found in humans.

INTRDUCTION

Although human cognition is fascinating, molecular studies of human cognition are rare. Often, genes were first identified in animals before their studies in humans were undertaken. This approach limits the phenotype to those present in animals, often lower animals because it is difficult to uncover genes by function in non-human primates. Thus, it is difficult to study cognition existing only in humans or those not present in lower animals. However, genetic studies of human diseases have been highly successful (Gusella et al., 1983; Schrott et al., 1972; Tsui et al., 1985). It is unnecessary to assume that human cognition is fundamentally different from diseases in amenibility to genetic studies. Genetics provides a non-invasive approach to study human cognition. Over the last decade, we have carried out several genome-wide association studies (GWAS) of human cognition from memory, social conformity to perceptual switching and top-down control in visual cognition (Chen et al., 2018a; Chen et al., 2018b; Zhu et al., 2019; Zhu et al., 2016; Zhu et al., 2018). In spite of the findings of associated markers, we do not know whether the genes harboring or around the markers are causally linked. Genetically, linkage analyses in large families have succeeded in identifying human disease causing genes. We have undertaken this approach to investigate the molecular genetic basis of face recognition and results from 7004 subjects are reported in this paper.

Face recognition, one of the highest forms of visual cognition, is essential for social cognition (Bruce and Young, 1986; Freiwald et al., 2016; Grill-Spector and Malach, 2004; Tsao and Livingstone, 2008). Neurons responding specifically to the face have been discovered in the interotemporal cortex (IT) (Bruce et al., 1981; Desimone et al., 1984; Gross et al., 1972; Perrett et al., 1987; Perrett et al., 1988; Perrett et al., 1982; Perrett et al., 1984, 1985; Rolls et al., 1985; Saito et al., 1986) and the superior temporal

sulcus (STS) (Bruce et al., 1981) of monkeys for nearly half a century. Electric stimulation in monkeys has functionally implicated neurons in face recognition (Afraz et al., 2006; Moeller et al., 2017). Face specific neurons have also been detected by event-related potentials (Allison et al., 1999; McCarthy et al., 1999; Puce et al., 1999) or direct electrophysiological recordings (Kreiman et al., 2000) in humans. Face responding brain areas have been detected in humans with positron emission tomography (PET) (Haxby et al., 1994; Sergent et al., 1992) and functional magnetic resonance imaging (fMRI) (Barraclough and Perrett, 2011; Clark et al., 1996; Dzhelyova et al., 2011; Kanwisher et al., 1997; Liu et al., 2010; Malach et al., 1995; McCarthy et al., 1997; Puce et al., 1996; Tsao et al., 2003; Tsao et al., 2006; Tsao et al., 2008). Results from trancranial magnetic stimulation are consistent with these regions being functionally important for face recognition (Pitcher et al., 2009; Pitcher et al., 2008; Polk et al., 2007).

Congenital prosopagnosia (CP), also known as developmental prosopagnosia, hereditary prosopagnosia (OMIM 610382), first reported in 1947 (Bodamer, 1947), is a selective impairment of visual learning and recognition of faces, in the absence of any detectable neurological injury (Behrmann and Avidan, 2005; Damasio et al., 1990; Duchaine and Nakayama, 2006b; Gruter et al., 2008; Kress and Daum, 2003; McConachie, 1976; Nunn et al., 2001; Susilo and Duchaine, 2013). Both a questionnaire-based screening method (Kennerknecht et al., 2006; Kennerknecht et al., 2008; Kennerknecht et al., 2007) and behavioral tests (Bowles et al., 2009) have estimated the prevalence of CP at 1.8 to 2.9% in the general population, providing a global estimate of millions of CP individuals (CPs). Familial (De Haan, 1999; Dobel et al., 2007; Duchaine et al., 2007; Galaburda and Duchaine, 2003; Grueter et al., 2007; Johnen et al., 2014; Kennerknecht et al., 2006; Kennerknecht et al., 2007; Lee et al.,

2010; McConachie, 1976; Schmalzl et al., 2008) and twin studies of CP (McKone et al., 2010; Polk et al., 2007; Wilmer et al., 2010; Zhu et al., 2010) suggest that CP and face recognition abilities are highly heritable. In pedigree analysis, a simple autosomal dominant mode of inheritance has often been observed (De Haan, 1999; Duchaine et al., 2007; Grueter et al., 2007; Kennerknecht et al., 2006; Lee et al., 2010; Schmalzl et al., 2008), indicating that mutations in a single gene can lead to difficulties in face recognition.

Beginning with a large pedigree including 18 members with obvious difficulties in face recognition in daily life, we have discovered a CP susceptibility gene, multiple C2 domains, transmembrane 2 (*MCTP2*, NM_018349). More rare mutations in *MCTP2* and correlations in families were detected in people at the lower end with 2 standard deviations (SDs) below the mean by an expanded screening from a cohort of 2904 individuals. In a second cohort of 1928, correlation of the rare alleles in *MCTP2* with the face recognition ability was detected in males by a gene based association analysis. Through a reverse-phenotyping approach, qualitative differences in face recognition were detected between 14 carriers with the same frameshift mutation S80fs in *MCTP2* and 19 non-carrying volunteers from a third cohort of 1757. From the 14 carriers, 6 families including one with 10 members also showed genotype-phenotype correlation. fMRI results indicate that impaired recognition of individual faces by CPs with the *MCTP2* mutations is associated with abnormal responses to repeated faces of the same identity in the rFFA. Our discovery may stimulate further research using genetics and genomics to study higher cognition in humans.

RESULTS

A Three-Generation Pedigree with CP

35 members from a three-generation pedigree (family A) (Table 1, Table S1, Figure 1A), were examined by an interview-based method developed previously (Grueter et al., 2007; Gruter et al., 2008; Kennerknecht et al., 2006) and used by many research groups (Bate et al., 2019; Bate et al., 2014; Carbon et al., 2010; Dinkelacker et al., 2011; Esins et al., 2016; Gruter et al., 2009; Stollhoff et al., 2011; Verfaillie et al., 2014; von Kriegstein et al., 2008; Zhao et al., 2018; Zhao et al., 2016). The mean age at interview was 42.5 ±18.7 (SD) years (range of 15-78 years). 18 family members showed obvious difficulties in face recognition in daily life: 9 individuals (IV:2, IV:4, V:9, V:10, V:11, V15, VI:5, VI:6 and VI:9) had been aware of this deficit in their early life before we contacted them; 4 individuals (IV:6, V:4, V:6 and VI:8) had difficulties in recognizing people for a long time without knowing the reason; while the remaining 5 individuals (IV:10, V:1, V:13, V:19 and VI:7) who ignored this problem in their daily life, adopted strategies relying on cues other than the face (Table 1). There was significantly negative correlation between the spontaneous need of gaze contact and the experience of face recognition difficulty or the use of compensatory strategies ($p < 0.001$, r = -0.658, n = 35). None of the persons showed signs of an autism spectrum disorder or neurodegenerative disorders.

We further investigated face recognition among 25 family members younger than 60 by the Cambridge Face Memory Test-Chinese (CFMT-C) (Bowles et al., 2009; McKone et al., 2012), the matched Cambridge Car Memory Test (CCMT) (Dennett et al., 2012) and the Face Inversion Effect (FIE) Discrimination test (Yovel and Kanwisher, 2005). Consistent with previous studies (Johnen et al., 2014; Schmalzl et al., 2008), there was intra-familial heterogeneity in face recognition. 14 members reported experienced substantial real-world face recognition difficulties, with 10 individuals found to be impaired on at least one of the face behavioral tests with a Z

value below 1.5 SDs. V:1, V:9, V:11, V:13, V:15, V:19 and VI:8 had a score more than two SDs below the normal level on CFMT-C, and V:4 (z = -1.60) and V:6 (z = -1.61), performed noticeably poorly (under 1.5 SDs). In the FIE discrimination test for holistic processing, the face-inversion effect of V:6 (z = -1.62), V:9 (z = -1.62), V:11 (z = -2.56) and VI:5 (z = -1.81) was impaired. V:1 (z = -2.96), V:13 (z = -1.73), V:15 (z = -1.55) and V:19 (z = -1.52) also performed poorly on the ability of discriminating upright faces. There were four individuals (V:10, VI:6, VI:7 and VI:9) who reported everyday difficulties but performed at normal levels on the tests of face recognition. All of them performed well in the CCMT test, indicating that their recognition of other physical stimuli was normal and independent of their abilities to recognize faces. Z scores for each member were listed in Table 1.

A Specific Region on Chromosome 15q Containing a Candidate CP Susceptibility Gene

To position the CP susceptibility gene, we identified 10 members (V:1, V:4, V:6, V:9, V:11, V:13, V:15, V:19, VI:5 and VI:8) as CP cases. They were younger than 60, repeatedly had difficulties in recognizing faces in everyday life via interview and showed impairments in at least one of the face behavioral tests. 4 members (IV:2, IV:4, IV:6 and IV:10) were also CPs with lifelong difficulties in face recognition, though they were over 60 and not suitable for behavioral tests. To be cautious, those with normal performance on behavioral tests but obvious face recognition failure experience (V:10, VI:6 and VI:7) were classified as unclear. Remaining 17 family members (IV:3, IV:5, IV:7, IV:8, IV:9, IV:11, V:3, V:7, V:8, V:12, V:14, V:16, VI:1, VI:2, VI:3, VI:4 and VI:10) were classified as normal for the following genome-wide linkage analysis (named linkage 1) (Table S1).

Parametric linkage analyses with 177,126 SNPs (LD threshold $r^2 < 0.5$) were performed, assuming an autosomal dominant inheritance with 80% penetrance, disease allele frequency at 0.001, and phenocopy rate at 0.05. Results of the genome-wide screen are shown in Figure S1A. The highest LOD (logarithm of odds) score 3.20, which suggested a possible CP-linked gene on chromosome 15, spanned 2.5 megabases (Mb) from rs6497118 (chr15:94266567, 15q26.1) to rs1437588 (chr15:96809782, 15q26.2) (Figure 1B). Keeping the pedigree structure, computer simulations led to an empirical *p* value of 0.017 suggesting that this result was significant at the genome-wide level (Table S2). Four other regions 2p25.1-24.3, 6q16.3, 12p12.1 and14q12 had LOD scores exceeding 1.0 (Figure S1A), but did not reach genome-wide significance by computer simulations (Table S2). Haplotype analysis was used for error elimination during the linkage scan. For the minimum candidate region (MCR), we constructed a genetic haplotype that was shared by all 14 affected individuals but not by the 17 unaffected individuals using markers with LOD scores over 3 (Figure S2A).

It is possible that some affected persons completed face recognition tasks with coping strategies for individual recognition without face recognition, leading to ambiguities in classifications (Dalrymple and Palermo, 2016; Duchaine and Weidenfeld, 2003; Grueter et al., 2007). Thus, we did a second linkage analysis (linkage 2) after adding three individuals (V:10, VI:6 and VI:7), who were set as unclear in the linkage analysis analyzed above (linkage 1), also as CPs (Table S1). Linkage 2 extended results of linkage 1 in further supporting the region between rs6497114 (chr15:94245722) and rs12900387 (chr15:97520230) with LOD scores over 3 (maximum LOD score = 4.27), spanning 3.27 Mb on chromosome 15q26.1-26.2 (Figure 1C, Figure S1B). Two other regions 10q25.1-25.2 and 12p12.1 had LOD scores over 1.0. Keeping the pedigree structure, simulations were performed to assess

empirical *p* values with 1,000 replicates to exclude false positive results due to random chances, and a significant $p < 0.009$ for the maximum LOD score on chr15 existed (Table S2). A genetic haplotype shared by all 17 affected but not by the 17 unaffected individuals using markers with LOD scores over 3, narrowed the MCR to a 1 Mb interval between rs1573086 and kgp9835458 (Figure S2B).

Candidate genes annotated by the newest NCBI database (https://www.ncbi.nlm.nih.gov/genome/gdv/) in the regions of both linkage analyses with LOD scores over 3 are listed in Table S3.

Copy number variations (CNVs) were identified for each member. No correlation was detected between the CNV genotypes and the phenotypes for each linkage analysis (data available upon request).

A Mutation in *MCTP2* Revealed by Whole-Exome Sequencing

To find causative mutations on chromosome 15q26.1-26.2 linked to CP in family A, we employed the whole-exome sequencing (WES) approach with four affected individuals (V:6, V:9, V:11, and V:19, Table S1). After the alignment and variant calling, the mean depth of all samples was 83.55, and 94.6% of the targeted bases were covered at more than 10-fold on average (Table S4). In linkage regions with LOD scores over 3, 100% coding bases were covered at more than 20-fold in at least one of the four individuals. We excluded variants with minor allele frequency (MAF) > 0.05 in multiple databases, including the dbSNP (v150) (Sherry et al., 2001), the 1000 Genomes Project (1000G) (Genomes Project et al., 2010; Genomes Project et al., 2012; Genomes Project et al., 2015), the Exome Aggregation Consortium (ExAC) (Lek et al., 2016) and the Genome Aggregation (genomAD, http://gnomad.broadinstitute.org) databases. In the linkage region with LOD scores over 3, only one variant (c.T2147G;

p.I716S, NM_018349) in *MCTP2* at chr15:94983466, was heterozygous in all the four CPs. This missense mutation (c.T2147G, p.I716S) was private tofamily A, not present in the dbSNP (v150), 1000G, ExAC and genomAD databases, albeit at this location another SNP rs200314451 (c.T2147C; p.I716T) had been reported with a MAF of 2.471e-05 (ExAC_ALL) in the ExAC database, with a MAF of 2e-04 in the 1000G, but not present in genomAD. In our expanded cohorts of 3,600 Chinese samples, we did not find the same mutation.

To answer whether this variant is a strong candidate of CP predisposition, we performed direct Sanger sequencing (primer sequences in Table S5) and co-segregation analysis in the pedigree. We found that the mutation (p.I716S) fully co-transmitted over multiple generations with all 18 family members who had difficulties in recognizing faces in daily life, including those with poor performance on the face behavioral tests and those with normal performance, but not in other family members (Figure 1A). We also performed the whole-genome sequencing (WGS) in two distantly related affected individuals V:19 and VI:5. We filtered our WGS datasets to exclude all variants that have a MAF over 0.05 in the dbSNP, 1000G, Exac and genomAD databases, and filtered for variants within the linkage regions with LOD scores over 3 by all affected individuals. Using this approach, 11 heterozygous variants detected in the linkage region were confirmed by Sanger sequencing (Tables S6 and S7). Notably, only one exonic variant (c.T2147G, p.I716S) occurred within Exon 17 (hg19, NM_018349) of *MCTP2*, as the result from WES. Furthermore, for the structure variations in the MCR detected from the 2 individuals by WGS, we did not find any co-segregating in the family members. Thus, the *MCTP2* mutation encoding p.I716S seems to be the only functional variant shared in the MCR by all patients in this large CP family, and is predicted to be disease causing by SIFT (Ng and Henikoff, 2001).

The full length of the *MCTP2* gene spans 180 kilobases (kb) of genomic DNA, with 22 coding exons, encoding a protein with 878 amino acid residues separated into three C_2 domains and two transmembrane regions (TMRs) (Figure 3A). I716, located in the first TMR, is highly conserved with primate species (Figures 3A and 3B). The expression data of microarray and RNA-seq of two people (H0351.2001, H0351.2002) from the Allen Human Brain Atlas mutually support MCTP2 expression in the right fusiform gyrus which contains the rFFA, the core brain region most consistently and robustly activated by the face selectively.

Additional *MCTP2* Mutations in CPs and CP Families Identified by Expanded Screening

On the basis of the theoretical and experimental considerations, it has been suggested that rare functional alleles are important contributors to the genetics of phenotypes (Cirulli and Goldstein, 2010; Heinzen et al., 2015; Jordan et al., 2010; Pritchard, 2001). The results from the study of a typical CP family A indicate that rare and even private, functional mutations in *MCTP2* could be related to a CP phenotype.

To examine whether the rare functional mutations in *MCTP2* were present in more CPs, we analyzed 2904 individuals by a questionnaire adapted from the 20-item self-report measure for quantifying prosopagnosia (Shah et al., 2015), similar to the interview we used for diagnosing family A. From these we identified 75 individuals who scored below 2 SDs of the mean among the 2904 participants and sequenced their *MCTP2* exons. Five rare heterozygous functional variants including one frameshift and four missense mutations were found in seven individuals (Figures 2A and 3A, Table 2).

The frameshift mutation (c.239delG, p.S80fs) in exon 1 of *MCTP2* was found in 3 of those 75 who reported poor performance in face recognition (individuals 1-1, 1-2

and 1-3) (Figure 2B). Further interviews revealed that all three individuals carrying p.S80fs reported lasting and irritating subjective uncertainties of face recognition, and realized this condition in their early life (Table S8). One relative of individual 1-1 carries p.S80fs and had adopted a strategy to observe other features, which took her more time to recognize a person (Figure 2C, Table S8).

In the family of individual 1-4, those who carry the p.M272L mutation in exon 5, reported their own strategies to recognize others, but still coped poorly with their difficulties (Figures 2D and 2E, Table S8). The p.T374A mutation in exon 8 was detected in individual 1-5 who was aware of having a severe deficit in face recognition (Figure 2F, Table S9). In the family of individual 1-6, the mutation p.V548I in exon 12 was correlated well with the phenotypes (Figures 2G and 2H, Table S8). In the family of individual 1-6, the mutation p.R641Q was found to segregate in all affected children of the family. These mutations are also evolutionarily conserved in animals (Figure 3B).

A Link between Rare Alleles in *MCTP2* and Face Recognition Revealed by a Gene Based Association Analysis

In the next step, we further aimed to evaluate the association of the alleles in *MCTP2* that harbor rare coding variants of moderate or large effects on protein coding with face recognition abilities. We had a second cohort of 1928individuals who took the same 20-item self-report measure as the first cohort of 2904 mentioned above. We resequenced all 22 exons, 5'UTR, 3'UTR and the exon-intron boundaries of *MCTP2* and did a single gene-based burden test for overall, female and male cohorts using the unified optimal sequence kernel association test (SKAT-O) (Lee et al., 2012).

After alignment and variant calling of all samples, 95% of targeted bases were covered at ⩾ 10-fold per individual (Table S4). A total of 53 rare coding and four splice

variants with MAF ⩽ 0.005 were identified and verified by Sanger sequencing (Table S9). Five algorithms (Sift, Polyphen2_HDIV, Polyphen2_HVAR, LRT and Mutation Taster) integrated by ANNOVAR (Adzhubei et al., 2010) were used to predict the functional severity of variants. Primary analyses based on (Purcell et al., 2014) tested 1) nonsynonymous (NS) strict variants: disruptive variants (nonsense, essential splice site and frameshifts) plus missense variants predicted to be damaging by all five algorithms (22 individuals); 2) NSbroad variants: disruptive plus missense variants predicted to be damaging by at least one algorithm (71 individuals) and 3) NSall variants: disruptive plus all missense variants (124 individuals).

Significant association was detected with the disruptive variants plus all missense variants (NSall, p = 0.0266) in the overall cohort based on burden testing (rho = 1 as optimal). In the male cohort, there was a significant enrichment of rare variants associated with face recognition ability in all three categories by both burden testing and optimal testing: NSstrict (p_{burden} = 0.0113, $p_{optical}$ = 0.0212), NSbroad (p_{burden} = 0.0029, $p_{optical}$ = 0.0061) and NSall (p_{burden} = 0.0011, $p_{optical}$ = 0.0024). No association was observed in the male cohort based on SKAT testing (rho = 0 as optimal). When considering the female cohort, no significant aggregation of rare variants was observed for each analysis. Details of the variants contributing to these significant test results are shown in Table 3. Our results suggest a high proportion of causal variants in *MCTP2* exert effects in the same direction and indicate the association between variant burden and face recognition in the male cohort, at least in part, by the effects of *MCTP2*.

A Frameshift Mutation in CP Individuals

Results of pedigree and population studies support correlation between rare alleles in *MCTP2* changing protein coding with the ability of face recognition, implicating

MCTP2 in face recognition. As mentioned above, the frameshift deletion c.239delG (p.S80fs) in the first exon of *MCTP2* which would eliminate most of the MCTP2 protein was present in 3 of 75 individuals reporting poor face recognition from the first cohort of 2904 individuals. In our second cohort, the ExAC and the gnomAD databases in East Asian population, the MAF is 0.0026, 0.0031 and 0.0045 respectively. This frequency lies within the reported range of CP prevalence and is high enough to allow identification of additional carriers consenting to analysis of face recognition through a reverse-phenotyping approach.

We further screened in a third independent cohort of 1757 individuals for this frameshift mutation. We detected 16 individuals carrying this mutation in a heterozygous condition with a MAF of 0.0046. Among them, 14 individuals agreed to be further examined (Table S10). Our interviews documented qualitative differences in the behavior of daily face recognition between the 14 carriers and 19 non-carrying volunteers from the same cohort (Table 4).

Four individuals 3-1, 3-2, 3-3 and 3-4 had clearly realized their difficulties in face recognition before we contacted them. Examination of all available family members of 3-2 and 3-3 showed that this mutation segregated with face recognition deficits in both families (Figure 4A, Table S10).

Individual 3-5 did not report difficulties in recognizing people/faces, but he did feel different from others in the way of recognizing people/faces (Table S10). In the family of 3-5 (Figure 4C, Table S10), the mutation was also identified in M's two relatives (II:2 and III:1), who reported daily failures in face recognition.

For the remaining nine individuals during the interview, they thought they were the same as others or even better than others in face recognition (Table S10). However, eight individuals (from 3-6 to 3-13) developed adaptive behaviors and depended on

explicit learning strategies for recognition, which made it not hard to recognize acquaintances (the whole person, not just the face) in their daily life. When they encountered strangers with few features, or actors especially actresses on the screen, or familiar people out-of-context, the strategies would not always work properly, but they could make adjustments and update the information quickly. The relatives of individual 3-6 who carry this mutation complained difficulties in recognition (Figure 4D, Table S10). For individual 3-8, the strategy to recognize people should be inherited from bothparents (Figure 4E, Table S10). The relative of individual 3-12 did not show prosopagnosia and do not carry the mutation (Figure 4F, Table S10). One individual, 3-14 did not show obvious abnormal face recognition ability during the whole interview.

The inferred haplotypes around the c.239delG (p.S80fs) mutation were estimated by comparison to microsatellites data collected on mother-father-offspring trios (Figure S3) and excluded the possibility of the same founder origin.

Neuroimaging Studies of Family CPs with the *MCTP2* Mutation

With the family A, we carried out fMRI analysis and compared results of three family members with CP and three without CP with normal controls. We found that impaired face recognition in CPs with the *MCTP2* mutation is associated with abnormal responses to individual faces in the fusiform face area (FFA).

The right FFA (rFFA) is the core brain region most consistently and robustly activated by the face selectively, and is involved in detecting the presence of a face and in discriminating individual faces (Grill-Spector et al., 2004; Haxby et al., 2000; Kanwisher et al., 1997; Puce et al., 1996; Rangarajan et al., 2014; Rotshtein et al., 2005; Yovel and Kanwisher, 2004; Zhang et al., 2012).

We functionally localized the rFFA as the region of interest (ROI) in family members and a group of normal controls by contrasting the blood oxygenation level-dependent (BOLD) signal when participants viewed faces versus non-face stimuli. As in normal participants (Figure S4A), there were significant activities in the rFFA of family members with CP (A-VI:5, Figure S4B, A-VI:7, Figure S4C and A-VI:9, Figure S4D) and family members without CP (A-VI:1, Figure S4E, A-VI:2, Figure S4F and AVI:10, Figure S4G). These suggest that the rFFA of the CPs in family A could respond specifically to faces.

Another marker for the special face processing is the Face-Inversion Effect (FIE), to which the rFFA is also sensitive. Assessment of the FIE with the protocol embedded in the adaptation paradigm (Figure 5A, 6A) showed a significantly higher response to upright than to inverted faces in the rFFA in normal subjects (paired t-test $p < 0.00001$, Figure 5B). In addition, in every single control subject there was a higher activation level in response to the upright than the inverted faces: the contrast 'upright face > inverted face' was significant at $p < 0.05$ in 18/20 normal subjects and showed a non-significant trend in the predicted direction in the remaining subjects ($p < 0.09$, $p < 0.14$). Similar to normal controls, the non-CP members of A-VI:1 (Figure 5C), A-VI:2 (Figure 5D), A-VI:10 (Figure 5E) and CP members A-VI:5 (Figure 5F) and A-VI:9 (Figure 5H) showed a higher response to the upright than the inverted faces, except A-VI:7 who showed a lower response to the upright than the inverted faces (Figure 5G). The fMRI-FIE score (the percent signal change under the upright condition versus the inverted condition within the ROI) for each family member was within the range of the normal controls (Figure 7A), except the CP member A-VI:7 who showed a significantly different activation pattern compared with the normal controls ($p = 0.0075$, Figure 7A)

using the Crawford and Howell's modified t test (Crawford and Garthwaite, 2002; Crawford et al., 2010).

To test whether the rFFA of family CPs could process facial identities, we conducted experiments using an fMRI adaptation paradigm in an event-related design (Yovel and Kanwisher, 2005) with the above-defined ROI (Figure 6A). In group analysis of normal subjects, the expected repetition suppression from fMRI adaptation to facial identity was highly significant in the rFFA (paired t-test $p < 0.0001$, Figure 6B). In addition, in every single control subject there was a higher activation level in response to pairs of up different faces (UDF) than to pairs of up identical faces (UIF): the contrast 'UDF>UIF' was significant at $p < 0.05$ in 18/20 normal subjects and showed a non-significant trend in the predicted direction in the remaining subjects ($p < 0.057$, $p < 0.25$). In each family member without CP (A-VI:1, Figure 6C, A-VI:2, Figure 6D and A-VI:10, Figure 6E), there was a higher activation level in response to pairs of UDF than to pairs of UIF. The contrast 'UDF>UIF' was significant at $p < 0.05$ in each non-CP family member without the *MCTP2* mutation. In contrast, the fMRI signals in the rFFA of the CPs did not show larger response to different faces than identical faces (A-VI:5, Figure 6F, A-VI:7, Figure 6G and A-VI:9, Figure 6H). The region-specific attenuation of neuronal activity due to repetition of identical faces appeared to be reduced in CPs, while different faces yielded similar response in the rFFA in CPs, non-CPs and normal controls. The adaptation score (percent signal change under the upright different condition versus the upright identical condition within the ROI) was lower for CPs than for every individual control subject (Figure 7B) and there was a non-significant trend in the same direction for CPs compared to normal controls (A-VI:4, $p = 0.05407$, A-VI:6, $p = 0.12832$ and A-VI:8, $p = 0.06548$) using the Crawford and Howell's modified t test (Crawford and Garthwaite, 2002; Crawford et al., 2010). These

results show that rFFA responses to identical faces are impaired in CPs with the *MCTP2* mutation, implying a mechanism underlying the difficulty of famliy A CPs in identifying seen-before faces.

In the 3-2 family with the frameshift mutation in the 80th amino acid in *MCTP2*, we observed a normal response to faces in rFFA at the categorical level both in CP III:4 (Figure S6H) and non-CP III:2 (Figure S6I), but a lack of FIE (Figure 5J) and a failure of repetition from fMRI adaptation to facial identity (Figure 6J) also occurred in the rFFA of CP III:4.

DISCUSSION

Our findings indicate that mutations in the *MCTP2* gene contribute to CP in humans. This conclusion is supported by the following: 1) a CP locus at 15q26.1-q26.2 containing *MCTP2* was identified by linkage analysis in family A with autosomal dominant CP (Figure 1, Figure S1); 2) a private mutation c.T2147G (encoding p.I716S) in *MCTP2* was the only functional mutation identified by WES and WGS in the MCR that fully co-segregated with CP in the entire family A (Figures 1); 3) 5 rare functional mutations in *MCTP2* were found in 7 CP individuals from a group of 75 with poor face recognition out of a cohort of 2904 individuals (Figure 2); 4) of those 7 CPs, family members from 4 CPs were available for analysis and showed phenotype-genotype correlations (Figures 2C, 2E, 2H and 2J); 5) 3 out of 7 CP individuals from the cohort with 2904 subjects contained the frameshift deletion c.239delG, p.S80fs in the first exon of *MCTP2* (Figure 2B), a trio of one of these 3 CPs available for testing further supported the correlation between c.239delG, p.S80fs and CP (Figure 2C); 6) 16 individuals from another cohort of 1757 subjects contained the same frameshift deletion mutation c.239delG, p.S80fs, with 14 available for further analysis. Differences in the

behavior of daily face recognition were detected between the 14 carriers and 19 non-carrying volunteers from the same cohort (Table 4); 7) four of 14 had unambiguous CP with two families available for analysis and all supported the correlation between c.239delG, p.S80fs and CP (Figure 4A and 4B), with further support from additional families (Figure 4C, 4D, 4E and 4F) who have developed explicit strategies to recognize people with non-facial clues to overcome their difficulties in face recognition; 8) correlation between the rare alleles in *MCTP2* and face recognition ability was detected in males by a gene based association analysis in a cohort of 1928 individuals. 9) impaired face recognition in CPs with the *MCTP2* mutations was associated with abnormal response to individual faces in the rFFA by neuroimaging studies.

MCTP2 is not the only gene involved in CP. Only 7 of the 75 CPs from the 2904 cohort carried *MCTP2* mutations. Rare alleles in *MCTP2* were not correlated with face recognition ability in females from the cohort of 1928. More linkage studies would be helpful. Higher cognition involves many cells and molecules. Hypothesis-free genomic analysis should be reconsidered with the next generation sequencing method covering both common and rare genetic variants through larger sample sizes, exceeding a million human participants, to find more genetic clues and replicate our *MCTP2* results in human face recognition in the future.

Molecular Studies of Face Recognition

Previous studies have been taken with behavioral, electrophysiological and genetic approaches, but no molecular studies have been undertaken.

In addition to primates (Bruce et al., 1981; Bruce and Young, 1986; Desimone et al., 1984; Freiwald et al., 2016; Grill-Spector et al., 2004; Tsao and Livingstone, 2008), face specific neurons have also been detected in sheep, dogs, chickens, pigeons

(Kendrick et al., 1995; Kendrick et al., 1996; Kendrick and Baldwin, 1987; Pascalis and Kelly, 2009; Ryan and Lea, 1994).

In humans, differential responses to faces occur early in infancy (Goren et al., 1975) and responses to mother's face occur within 4 to 35 days after birth (Bartrip et al., 2001; Bruce et al., 2000; Bushnell et al., 1989; de Haan et al., 2001; Haith et al., 1977; Pascalis et al., 1995; Slater et al., 2000; Slater et al., 1998; Walton et al., 1992).

There are humans defective in recognizing other objects but no face (Feinberg et al., 1994; Germine et al., 2011a; McMullen et al., 2000; Moscovitch et al., 1997), and there are humans who have prosopagnosia, or face blindness, defective in recognizing faces but not other objects (Duchaine and Nakayama, 2005; Duchaine et al., 2006; Farah, 1996; Farah et al., 1995; Henke et al., 1998; Li and Song, 2007; Nunn et al., 2001; Riddoch et al., 2008). Acquired prosopagnosia includes damages by diseases or injuries (Barton et al., 2002; Bouvier and Engel, 2006; Landis et al., 1986; Meadows, 1974; Riddoch et al., 2008; Schiltz et al., 2006; Steeves et al., 2006; Yin, 1969). Several extended CP pedigrees have been reported (Duchaine et al., 2007; Johnen et al., 2014; Schmalzl et al., 2008), but the genes underlying their phenotype remain unclear.

We have performed GWAS of memory, social conformity, visual perceptual switching and visual top-down control in humans (Chen et al., 2018a; Chen et al., 2018b; Zhu et al., 2019; Zhu et al., 2016; Zhu et al., 2018), and understand the limit of GWAS.

This paper represents the first molecular genetic study of face recognition. Our experience with GWAS of human cognition made us determined to use classic linkage analysis, rather than GWAS, to study genes involved in face recognition. An important lesson from our GWAS is that it is very difficult to study causal relationship from GWAS. We therefore persevered in using classic linkage analysis with genomic markers. The availability of family A is particularly helpful in allowing us to find an

MCR with LOD score of 4.27 (Figure 1C). The discovery of *MCTP2* was confirmed in other CP individuals and trios or small families in China (Figures 2 and 4). We have found more families with other unknown genetic reasons and will further analyze them.

The expression data of microarray and RNA-seq of two people (H0351.2001, H0351.2002) from the Allen Human Brain Atlas support that *MCTP2* is expressed in the right fusiform gyrus, covering the face selective region. Although the expression level of *MCTP2* is relatively low, this does not prevent *MCTP2* from playing an important role.

The *MCTP2* gene encodes a protein with three C2 domains and two transmembrane regions (TMRs) with resemblance to proteins involved in synaptic transmission (Shin et al., 2005). Its C2 domains can bind Ca^{2+} (Shin et al., 2005). Many proteins that bear the Ca^{2+}-binding C2 domain are involved in membrane and vesicle trafficking, playing a central role in neuronal transmission (Cho and Stahelin, 2006; Shupliakov and Brodin, 2010). It has also been implicated in lipid droplets formation (Joshi et al., 2018).

The C2 domains and TMRs of MCTPs are evolutionarily conserved from invertebrates such as Caenorhabditis elegans and Drosophila melanogaster, that only have one *MCTP* gene, to mammals that have two genes (*MCTP1* and *MCTP2*). MCTP1 is expressed in the central nervous system and has been implicated in regulating endocytic recycling of specific CNS neurons and synapses (Qiu et al., 2015). In Drosophila, MCTP is involved in stabilizing synaptic transmission and homeostatic plasticity (Genc et al., 2017). The four zebrafish *MCTP* genes are expressed mainly in the nervous and muscular systems. Knocking down *MCTP2b* impaired embryonic development (Espino-Saldaña et al., 2020).

Both *MCTP2* and *MCTP1* had been associated with attention-deficit/hyperactivity

disorder (ADHD), a highly heritable neuropsychiatric disorder (Kweon et al., 2018; Mick et al., 2010). Poelmans et al. found that the gene ontology process "calcium ion binding," which plays an important role in neurite migration, was significantly enriched in the 14 ADHD-associated genes (Poelmans et al., 2011). Genome-wide analyses have identified *MCTP1* single nucleotide polymorphism (SNP) in bipolar diseases (Scott et al., 2009). There was also a report of *MCTP2* SNP in schizophrenia (Djurovic et al., 2009).

Our gene based association analysis suggests a high proportion of causal variants in *MCTP2* exert effects in the same direction and indicates the association between variant burden and face recognition ability in the male cohort, at least in part, by the effects of *MCTP2*. It is likely that the MCTP2 protein functions in the CNS and it will be interesting to investigate whether its role in face recognition is a general feature in neural signaling or a specialized step in specific regions of the brain.

Diagnosis of CP in Linkage Analysis

If both parents exhibit a disability in face recognition which could be caused by different genetic origins, it is more difficult to localize the causing genetic variations (Duchaine et al., 2007; Schmalzl et al., 2008).

For the families with clear inherited models, correct and consistent phenotypical diagnosis for each family member is very important before genetic analysis. Currently, while specific tests have proven useful (Duchaine and Nakayama, 2006a), no single test is of sufficient discriminatory power for the entire spectrum of CP (Duchaine and Nakayama, 2006a; Duchaine and Nakayama, 2004; Duchaine and Weidenfeld, 2003; Grueter et al., 2007; Shah et al., 2015). It is difficult to assess the exact severity of the individual CP cases by behavioral tests, because behavioral adaption, other cognitive

skills, experience with faces or testing forms balance the deficit to an unknown extent, even within a single family with similar genetic and environmental background. Mistaking a true CP as a normal one due to his/her normal performance in behavioral tests would undermine genetic studies.

In our behavioral tests for face recognition in family A, there was also a variable behavioral profile in individuals who experienced everyday difficulties in face recognition. For example, V:10, perfected normally on the CFMT-C (z = -0.06) and excellently on the FIE test (z = 2.90) (Table 1), but he was already aware of his own real-world face recognition difficulties while he attended college, before we contacted him.

Another representative example was VI:7, a distant relative of the index, whose performance was fairly good on the CFMT-C (z =1.26) and FIE (z = 1.48), and would not have been diagnosed as having CP if only according to the available behavioral tests. However, she depended on clues beyond the face when recognizing people (Table 1, more detailed information can be obtained by contacting the corresponding author). Neuroimaging study showed that the activation pattern in her rFFA was indeed defective both in the FIE (Figure 5G, 7A) and the adaptation paradigm (Figure 6G, 7B).

Considering the within and between familial variations among CPs (Dobel et al., 2007; Johnen et al., 2014; Le Grand et al., 2006; Lee et al., 2010; Schmalzl et al., 2008), and the fact that CPs must experience face recognition difficulties on a regular basis in typical situations where recognition is not problematic for individuals without face processing impairments, we adopted a strategy with two step linkage analysis: in linkage 1, we defined family members under 60 as cases of CP (CPs) with clear difficulties in face recognition revealed by the interview and performed poorly (-1.5 SDs) below the average level in at least one of the face behavioral tests; those over 60

not suitable for behavioral tests but with lifelong difficulties in face recognition also as CPs; to be cautious, those with normal performance on behavioral tests but obvious face recognition failure experience as unclear; and the remaining members as normal; in linkage 2, we added people who were set as unclear in linkage 1 but with real-world face recognition defects, also as CPs (Table S1). We detected the same region on chromosome 15q with LOD scores over 3 from both linkage analyses (Figure 1B, 1C, Figure S1A, S1B).

Mutations in the *MCTP2* Gene

The mutation encoding p.I716S we found in *MCTP2* in the linkage region on chromosome 15q fully co-transmitted over several generations with all 18 affected individuals in family A. This mutation is private to family A, not present in public databases or in our in-house data, albeit another rare variant rs200314451 (c.T2147C; p.I716T) has been reported at this location in public databases.

The p.I716S is located at the first TMR of MCTP2 protein (RefGene: NM_018349). The isoleucine at 716 is evolutionally conserved within primates (Figure 3B) which are highly developed in face processing (Freiwald et al., 2016; Hung et al., 2015; Moeller et al., 2008; Tsao et al., 2003; Tsao et al., 2006; Tsao et al., 2008).

Results from the prototypical CP family A indicate that rare, and private, functional mutations in *MCTP2* could be related to the CP phenotype and prompted us to examine more CP individuals and families. This strategy is justified from knowledge of monogenic disorders, where it is typical for different rare variants with large effects of the same gene to segregate with disease in different families. We screened 75 individuals with the most serious self-reported face recognition abnormalities at one end of the population distribution through the questionnaire from a cohort of 2904,

rather than simply verifying them among the people who actively reported their deficient face recognition. Five rare heterozygous variants including one frameshift and four missense mutations were found in seven individuals (Figures 2A and 3A, Table 2), Their difficulties in daily face recognition were confirmed by detailed interviews. All the available family members also showed a clear phenotype-genotype correlation to exclude the observation of *MCTP2* variants by chance. This suggests the significance of rare functional mutations in *MCTP2* in CP. The association of the alleles in *MCTP2* that harbor rare coding variants of moderate or large effect on the protein coding with face recognition ability was further supported in males by a gene based burden analysis from a second cohort of 1928 individuals (Table 3). These results suggest a high proportion of causal variants in *MCTP2* exert effects in the same direction in the male cohort, beyond chance.

The frameshift deletion c.239delG, p.S80fs, in *MCTP2* which was found in three CPs from the 75 top candidate CPs, is likely to eliminate the MCTP2 protein. In our second cohort, the ExAC and the gnomAD databases, the MAFs in the population of East Asian for this mutation were 0.0026, 0.0031 and 0.0045 respectively, which lies within the reported range of CP prevalence and is high enough to attempt the identification of additional carriers consenting to a characterization of their face recognition behavior through a reverse-phenotyping approach. We detected 16 individuals carrying this mutation in a heterozygous condition with a MAF of 0.0046 in a third independent cohort of 1757. Among them, differences in the behavior of daily face recognition were detected between 14 carriers and 19 non-carrying volunteers (Table 4). 4 were clearly aware of their difficulties in face recognition before we contacted them, one felt different from others in recognizing faces, 8 individuals developed adaptive behaviors and used explicit learning strategies for person

recognition, only one did not show obvious abnormal face recognition ability during the whole interview (Table S10). The individual differences in phenotype outcomes may be modulated both on the neurodevelopmental levels and environmental factors with other possible interactions such us more genetic traits, cognitive skills and learned compensatory strategies. Family members with the p.S80fs mutation further confirmed the phenotype-genotype correlation (Table S10, more detailed diagnostic information can be obtained by contacting the corresponding author).

fMRI Analysis

fMRI is helpful both as a diagnostic means for endophenotype and for mechanistic link between genotype and brain activity.

We have focused on the rFFA because of its involvement both in detecting the presence of a face and in discriminating different faces (Grill-Spector et al., 2004; Haxby et al., 2000; Kanwisher et al., 1997; Puce et al., 1996; Rangarajan et al., 2014; Rotshtein et al., 2005; Yovel and Kanwisher, 2004; Zhang et al., 2012).

In the fMRI analysis, the activation of rFFA in Families A and 3-2 without mutations in *MCTP2* was similar to that of the control group, in face detection, upright face recognition, inverted face recognition, different or the same individual face identification. The rFFA of members with the I716S in family A or the S80fs in family 3-2 also showed normal activation during face detection, suggesting these *MCTP2* mutations did not affect face detection. However, in the fMRI adaptation paradigm based on the theory that neural responses to faces are attenuated across repeated presentations of the same stimuli, an effect known as repetition suppression (Grill-Spector et al., 2006; Grill-Spector and Malach, 2001; Henson and Rugg, 2003), the activation of the rFFA in all mutation carriers to paired identical faces was not lower

than that of paired different faces, compared with that of normal people or non-carriers in families. These results indicate, that the mutations in *MCTP2* caused a failure in repetition suppression in the rFFA as well as behavioral difficulties in identifying previously encountered faces. In addition, the rFFA of A-VI:7 and 3-2-III:4 showed a significantly different activation pattern during the presentation of upright faces and inverted faces for the FIE. The rFFA of A-VI:7- was less active to the upright faces than the inverted faces, while the rFFA of 3-2-III:4 was less active to both types of faces.

Three fMRI subjects with the I716S mutation in family A and the one with the S80fs in family 3-2 had difficulties in recognizing people in daily life. In the fMRI study, when they were repeatedly presented with the same face, the repetition-suppression phenomenon in the rFFA was not obvious. It will be helpful if more families with *MCTP2* mutations are examined by fMRI.

We need to further investigate whether CP patients with *MCTP2* mutations have problems in specific adaptive mechanism of facial processing, or problems in other visual stimulation processing in general. Although in some non-face behavioral tests, these CPs performed normally, such as the CCMT test.

Our genetic results provide evidence that *MCTP2* gene mutations underlie CP and our fMRI results suggest that *MCTP2* is involved in neural circuits required for distinguishing different faces.

AUTHOR CONTRIBUTION

YR initiated the idea of genetic analysis of higher cognition in humans, using CP as an example and supervised the project. YR and YS designed the project and wrote the manuscript. YS and WWM carried out fMRI. HFZ, WF and YS collected and coordinated the collection of samples. YS and WF performed CP diagnosis. YS, WF and EXZ did behavioral tests and data analysis. EXZ wrote software programs for behavioral tests. YS and ZQL performed linkage analysis. YS and WY carried out sequencing analysis. YS performed founder origin analysis.

ACKNOWLEDGEMENTS

We are grateful to participants of this project (especially the proband and members of A0001 family), to Dr. Ingo Kennerknecht at the University of Muenster, Germany for sharing the interview and for helpful discussions, to Dr. Jia Liu of Beijing Normal University for sharing the Cambridge Face Memory Test-Chinese version and helpful discussions, to Dr. Bradley Duchaine of Dartmoutth College for sharing the Cambridge Car Memory Test and helpful discussions, to Dr. Jurg Ott, Shihui Han and Taoyu Wu for helpful discussions, to Dr. Pianpian Zhao of Westlake University, Drs. Jing Jiang and Kunjun Mao of Shangrao Medical College, Drs. Liang Chen and Jieyu Chen of Southern Medical University for help with sample collection, to Drs. Zi-Bing Jin and Xue Zhang for critical comments on the manuscript, to the High Performance Computing Platform of Peking/Tsinghua Center for Life Sciences (CLS) for computer support, to the CIBR, the CLS, the Shenzhen Bay Laboratory and the National Natural Science Foundation of China for funding support.

REFERENCES

Abecasis, G.R., Cherny, S.S., Cookson, W.O., and Cardon, L.R. (2002). Merlin[mdash]rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet *30*, 97-101.
Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. Nat Methods *7*, 248-249.
Afraz, S.R., Kiani, R., and Esteky, H. (2006). Microstimulation of inferotemporal cortex influences face categorization. Nature *442*, 692-695.
Allison, T., Puce, A., Spencer, D.D., and McCarthy, G. (1999). Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. Cereb Cortex *9*, 415-430.
Anstey, K.J., Dear, K., Christensen, H., and Jorm, A.F. (2005). Biomarkers, health, lifestyle, and demographic variables as correlates of reaction time performance in early, middle, and late adulthood. Q J Exp Psychol A *58*, 5-21.
Barraclough, N.E., and Perrett, D.I. (2011). From single cells to social perception. Philosophical Transactions of the Royal Society B-Biological Sciences *366*, 1739-1752.
Barton, J.J., Press, D.Z., Keenan, J.P., and O'Connor, M. (2002). Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. Neurology *58*, 71-78.
Bartrip, J., Morton, J., and de Schonen, S. (2001). Responses to mother's face in 3-week to 5-month-old infants. British Journal of Developmental Psychology *19*, 219-232.
Bate, S., Adams, A., Bennetts, R., and Line, H. (2019). Developmental prosopagnosia with concurrent topographical difficulties: A case report and virtual reality training programme. Neuropsychol Rehabil *29*, 1290-1312.
Bate, S., Cook, S.J., Duchaine, B., Tree, J.J., Burns, E.J., and Hodgson, T.L. (2014). Intranasal inhalation of oxytocin improves face processing in developmental prosopagnosia. Cortex *50*, 55-63.
Behrmann, M., and Avidan, G. (2005). Congenital prosopagnosia: face-blind from birth. Trends Cogn Sci *9*, 180-187.
Bodamer, J. (1947). Die Prosop-Agnosie. Archiv für Psychiatrie und Nervenkrankheiten

179, 6-53.
Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114-2120.
Bouvier, S.E., and Engel, S.A. (2006). Behavioral deficits and cortical damage loci in cerebral achromatopsia. Cereb Cortex *16*, 183-191.
Bowles, D.C., McKone, E., Dawel, A., Duchaine, B., Palermo, R., Schmalzl, L., Rivolta, D., Wilson, C.E., and Yovel, G. (2009). Diagnosing prosopagnosia: effects of ageing, sex, and participant-stimulus ethnic match on the Cambridge Face Memory Test and Cambridge Face Perception Test. Cogn Neuropsychol *26*, 423-455.
Bruce, C., Desimone, R., and Gross, C.G. (1981). Visual Properties of Neurons in a Polysensory Area in Superior Temporal Sulcus of the Macaque. Journal of Neurophysiology *46*, 369-384.
Bruce, V., Campbell, R.N., Doherty-Sneddon, G., Import, A., Langton, S., McAuley, S., and Wright, R. (2000). Testing face processing skills in children. British Journal of Developmental Psychology *18*, 319-333.
Bruce, V., and Young, A. (1986). Understanding face recognition. Br J Psychol *77 (Pt 3)*, 305-327.
Burianova, H., Lee, Y., Grady, C.L., and Moscovitch, M. (2013). Age-related dedifferentiation and compensatory changes in the functional network underlying face processing. Neurobiol Aging *34*, 2759-2767.
Bushnell, I.W.R., Sai, F., and Mullin, J.T. (1989). Neonatal Recognition of the Mothers Face. British Journal of Developmental Psychology *7*, 3-15.
Busigny, T., and Rossion, B. (2010). Acquired prosopagnosia abolishes the face inversion effect. Cortex *46*, 965-981.
Carbon, C.C., Gruter, T., Gruter, M., Weber, J.E., and Lueschow, A. (2010). Dissociation of facial attractiveness and distinctiveness processing in congenital prosopagnosia. Vis Cogn *18*, 641-654.
Carbon, C.C., Gruter, T., Weber, J.E., and Lueschow, A. (2007). Faces as objects of non-expertise: processing of thatcherised faces in congenital prosopagnosia. Perception *36*, 1635-1645.
Chen, B.Q., Zhu, Z.J., Na, R., Fang, W., Zhang, W.X., Zhou, Q., Zhou, S.B., Lei, H., Huang, A.L., Chen, T.M.*, et al.* (2018a). Genomic Analyses of Visual Cognition: Perceptual Rivalry and Top-Down Control. Journal of Neuroscience *38*, 9668-9678.

Chen, B.Q., Zhu, Z.J., Wang, Y.Y., Ding, X.H., Guo, X.B., He, M.G., Fang, W., Zhou, Q., Zhou, S.B., Lei, H.*, et al.* (2018b). Nature vs. nurture in human sociality: multi-level genomic analyses of social conformity. Journal of Human Genetics *63*, 605-619.
Cirulli, E.T., and Goldstein, D.B. (2010). Uncovering the roles of rare variants in common disease through whole-genome sequencing. Nat Rev Genet *11*, 415-425.
Clark, V.P., Keil, K., Maisog, J.M., Courtney, S., Ungerleider, L.G., and Haxby, J.V. (1996). Functional magnetic resonance imaging of human visual cortex during face matching: A comparison with positron emission tomography. Neuroimage *4*, 1-15.
Crawford, J.R., and Garthwaite, P.H. (2002). Investigation of the single case in neuropsychology: confidence limits on the abnormality of test scores and test score differences. Neuropsychologia *40*, 1196-1208.
Crawford, J.R., Garthwaite, P.H., and Porter, S. (2010). Point and interval estimates of effect sizes for the case-controls design in neuropsychology: rationale, methods, implementations, and proposed reporting standards. Cogn Neuropsychol *27*, 245-260.
Dalrymple, K.A., and Palermo, R. (2016). Guidelines for studying developmental prosopagnosia in adults and children. Wiley Interdiscip Rev Cogn Sci *7*, 73-87.
Damasio, A.R., Tranel, D., and Damasio, H. (1990). Face agnosia and the neural substrates of memory. Annu Rev Neurosci *13*, 89-109.
De Haan, E.H. (1999). A familial factor in the development of face recognition deficits. J Clin Exp Neuropsychol *21*, 312-315.
de Haan, M., Johnson, M.H., Maurer, D., and Perrett, D.I. (2001). Recognition of individual faces and average face prototypes by 1-and 3-month-old infants. Cognitive Development *16*, 659-678.
Dennett, H.W., McKone, E., Tavashmi, R., Hall, A., Pidcock, M., Edwards, M., and Duchaine, B. (2012). The Cambridge Car Memory Test: a task matched in format to the Cambridge Face Memory Test, with norms, reliability, sex differences, dissociations from face memory, and expertise effects. Behav Res Methods *44*, 587-605.
DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M.*, et al.* (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet *43*, 491-498.
Desimone, R., Albright, T.D., Gross, C.G., and Bruce, C. (1984). Stimulus-Selective Properties of Inferior Temporal Neurons in the Macaque. Journal of Neuroscience *4*, 2051-2062.
Dinkelacker, V., Gruter, M., Klaver, P., Gruter, T., Specht, K., Weis, S., Kennerknecht, I., Elger, C.E., and Fernandez, G. (2011). Congenital prosopagnosia: multistage anatomical and functional deficits in face processing circuitry. J Neurol *258*, 770-782.
Dobel, C., Bolte, J., Aicher, M., and Schweinberger, S.R. (2007). Prosopagnosia without apparent cause: overview and diagnosis of six cases. Cortex *43*, 718-733.
Dricot, L., Sorger, B., Schiltz, C., Goebel, R., and Rossion, B. (2008). The roles of “face” and “non-face” areas during individual face perception: Evidence by fMRI adaptation in a brain-damaged prosopagnosic patient. NeuroImage *40*, 318-332.
Duchaine, B., Germine, L., and Nakayama, K. (2007). Family resemblance: ten family members with prosopagnosia and within-class object agnosia. Cogn Neuropsychol *24*, 419-430.
Duchaine, B., and Nakayama, K. (2005). Dissociations of face and object recognition in developmental prosopagnosia. J Cogn Neurosci *17*, 249-261.
Duchaine, B., and Nakayama, K. (2006a). The Cambridge Face Memory Test: results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. Neuropsychologia *44*, 576-585.
Duchaine, B.C., and Nakayama, K. (2004). Developmental prosopagnosia and the Benton Facial Recognition Test. Neurology *62*, 1219-1220.
Duchaine, B.C., and Nakayama, K. (2006b). Developmental prosopagnosia: a window to content-specific face processing. Curr Opin Neurobiol *16*, 166-173.
Duchaine, B.C., and Weidenfeld, A. (2003). An evaluation of two commonly used tests of unfamiliar face recognition. Neuropsychologia *41*, 713-720.
Duchaine, B.C., Yovel, G., Butterworth, E.J., and Nakayama, K. (2006). Prosopagnosia as an impairment to face-specific mechanisms: Elimination of the alternative hypotheses in a developmental case. Cogn Neuropsychol *23*, 714-747.

Dzhelyova, M.P., Ellison, A., and Atkinson, A.P. (2011). Event-related repetitive TMS reveals distinct, critical roles for right OFA and bilateral posterior STS in judging the sex and trustworthiness of faces. J Cogn Neurosci *23*, 2782-2796.
Esins, J., Schultz, J., Stemper, C., Kennerknecht, I., and Bulthoff, I. (2016). Face Perception and Test Reliabilities in Congenital Prosopagnosia in Seven Tests. Iperception *7*, 2041669515625797.
Farah, M.J. (1996). Is face recognition 'special'? Evidence from neuropsychology. Behav Brain Res *76*, 181-189.
Farah, M.J., Levinson, K.L., and Klein, K.L. (1995). Face Perception and within-Category Discrimination in Prosopagnosia. Neuropsychologia *33*, 661-&.
Feinberg, T.E., Schindler, R.J., Ochoa, E., Kwan, P.C., and Farah, M.J. (1994). Associative Visual Agnosia and Alexia without Prosopagnosia. Cortex *30*, 395-411.
Forshew, T., Murtaza, M., Parkinson, C., Gale, D., Tsui, D.W., Kaper, F., Dawson, S.J., Piskorz, A.M., Jimenez-Linan, M., Bentley, D.*, et al.* (2012). Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. Sci Transl Med *4*, 136ra168.
Freiwald, W., Duchaine, B., and Yovel, G. (2016). Face Processing Systems: From Neurons to Real-World Social Perception. Annu Rev Neurosci *39*, 325-346.
Galaburda, A.M., and Duchaine, B.C. (2003). Developmental disorders of vision. Neurol Clin *21*, 687-707.
Gao, X., Vuong, Q.C., and Rossion, B. (2019). The cortical face network of the prosopagnosic patient PS with fast periodic stimulation in fMRI. Cortex *119*, 528-542.
Genomes Project, C., Abecasis, G.R., Altshuler, D., Auton, A., Brooks, L.D., Durbin, R.M., Gibbs, R.A., Hurles, M.E., and McVean, G.A. (2010). A map of human genome variation from population-scale sequencing. Nature *467*, 1061-1073.
Genomes Project, C., Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., and McVean, G.A. (2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56-65.
Genomes Project, C., Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A.*, et al.* (2015). A global reference for human genetic variation. Nature *526*, 68-74.
Germine, L., Cashdollar, N., Duzel, E., and Duchaine, B. (2011a). A new selective developmental deficit: Impaired object recognition with normal face recognition. Cortex *47*, 598-607.
Germine, L.T., Duchaine, B., and Nakayama, K. (2011b). Where cognitive development and aging meet: face learning ability peaks after age 30. Cognition *118*, 201-210.
Goh, J.O., Suzuki, A., and Park, D.C. (2010). Reduced neural selectivity increases fMRI adaptation with age during face discrimination. Neuroimage *51*, 336-344.
Goren, C.C., Sarty, M., and Wu, P.Y.K. (1975). Visual Following and Pattern-Discrimination of Face-Like Stimuli by Newborn-Infants. Pediatrics *56*, 544-549.
Grill-Spector, K., Henson, R., and Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn Sci *10*, 14-23.
Grill-Spector, K., Knouf, N., and Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. Nat Neurosci *7*, 555-562.
Grill-Spector, K., and Malach, R. (2001). fMR-adaptation: a tool for studying the functional properties of human cortical neurons. Acta Psychol (Amst) *107*, 293-321.
Grill-Spector, K., and Malach, R. (2004). The human visual cortex. Annu Rev Neurosci *27*, 649-677.
Gross, C.G., Rochamir.Ce, and Bender, D.B. (1972). Visual Properties of Neurons in Inferotemporal Cortex of Macaque. Journal of Neurophysiology *35*, 96-&.
Grueter, M., Grueter, T., Bell, V., Horst, J., Laskowski, W., Sperling, K., Halligan, P.W., Ellis, H.D., and Kennerknecht, I. (2007). Hereditary prosopagnosia: the first case series. Cortex *43*, 734-749.
Gruter, T., Gruter, M., Bell, V., and Carbon, C.C. (2009). Visual mental imagery in congenital prosopagnosia. Neurosci Lett *453*, 135-140.
Gruter, T., Gruter, M., and Carbon, C.C. (2008). Neural and genetic foundations of face recognition and prosopagnosia. J Neuropsychol *2*, 79-97.
Gusella, J.F., Wexler, N.S., Conneally, P.M., Naylor, S.L., Anderson, M.A., Tanzi, R.E., Watkins, P.C., Ottina, K., Wallace, M.R., Sakaguchi, A.Y.*, et al.* (1983). A polymorphic DNA marker genetically linked to Huntington's disease. Nature *306*, 234-238.
Haith, M.M., Bergman, T., and Moore, M.J. (1977). Eye contact and face scanning in early infancy. Science *198*, 853-855.

Haxby, J.V., Hoffman, E.A., and Gobbini, M.I. (2000). The distributed human neural system for face perception. Trends Cogn Sci *4*, 223-233.
Haxby, J.V., Horwitz, B., Ungerleider, L.G., Maisog, J.M., Pietrini, P., and Grady, C.L. (1994). The Functional-Organization of Human Extrastriate Cortex - a Pet-Rcbf Study of Selective Attention to Faces and Locations. Journal of Neuroscience *14*, 6336-6353.
Heinzen, E.L., Neale, B.M., Traynelis, S.F., Allen, A.S., and Goldstein, D.B. (2015). The genetics of neuropsychiatric diseases: looking in and beyond the exome. Annu Rev Neurosci *38*, 47-68.
Henke, K., Schweinberger, S.R., Grigo, A., Klos, T., and Sommer, W. (1998). Specificity of face recognition: Recognition of exemplars of non-face objects in prosopagnosia. Cortex *34*, 289-296.
Henson, R.N., and Rugg, M.D. (2003). Neural response suppression, haemodynamic repetition effects, and behavioural priming. Neuropsychologia *41*, 263-270.
Hung, C.C., Yen, C.C., Ciuchta, J.L., Papoti, D., Bock, N.A., Leopold, D.A., and Silva, A.C. (2015). Functional mapping of face-selective regions in the extrastriate visual cortex of the marmoset. J Neurosci *35*, 1160-1172.
Johnen, A., Schmukle, S.C., Huttenbrink, J., Kischka, C., Kennerknecht, I., and Dobel, C. (2014). A family at risk: congenital prosopagnosia, poor face recognition and visuoperceptual deficits within one family. Neuropsychologia *58*, 52-63.
Jordan, D.M., Ramensky, V.E., and Sunyaev, S.R. (2010). Human allelic variation: perspective from protein function, structure, and evolution. Curr Opin Struct Biol *20*, 342-350.
Kanwisher, N., McDermott, J., and Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. J Neurosci *17*, 4302-4311.
Kendrick, K.M., Atkins, K., Hinton, M.R., Broad, K.D., Fabrenys, C., and Keverne, B. (1995). Facial and Vocal Discrimination in Sheep. Animal Behaviour *49*, 1665-1676.
Kendrick, K.M., Atkins, K., Hinton, M.R., Heavens, P., and Keverne, B. (1996). Are faces special for sheep? Evidence from facial and object discrimination learning tests showing effects of inversion and social familiarity. Behavioural Processes *38*, 19-35.
Kendrick, K.M., and Baldwin, B.A. (1987). Cells in Temporal Cortex of Conscious Sheep Can Respond Preferentially to the Sight of Faces. Science *236*, 448-450.
Kennerknecht, I., Grueter, T., Welling, B., Wentzek, S., Horst, J., Edwards, S., and Grueter, M. (2006). First report of prevalence of non-syndromic hereditary prosopagnosia (HPA). Am J Med Genet A *140*, 1617-1622.
Kennerknecht, I., Ho, N.Y., and Wong, V.C. (2008). Prevalence of hereditary prosopagnosia (HPA) in Hong Kong Chinese population. Am J Med Genet A *146A*, 2863-2870.
Kennerknecht, I., Plumpe, N., Edwards, S., and Raman, R. (2007). Hereditary prosopagnosia (HPA): the first report outside the Caucasian population. Journal of Human Genetics *52*, 230-236.
Kreiman, G., Koch, C., and Fried, I. (2000). Category-specific visual responses of single neurons in the human medial temporal lobe. Nature Neuroscience *3*, 946-953.
Kress, T., and Daum, I. (2003). Developmental prosopagnosia: a review. Behav Neurol *14*, 109-121.
Landis, T., Cummings, J.L., Christen, L., Bogen, J.E., and Imhof, H.G. (1986). Are unilateral right posterior cerebral lesions sufficient to cause prosopagnosia? Clinical and radiological findings in six additional patients. Cortex *22*, 243-252.
Le Grand, R., Cooper, P.A., Mondloch, C.J., Lewis, T.L., Sagiv, N., de Gelder, B., and Maurer, D. (2006). What aspects of face processing are impaired in developmental prosopagnosia? Brain Cogn *61*, 139-158.
Lee, S., Wu, M.C., and Lin, X. (2012). Optimal tests for rare variant effects in sequencing association studies. Biostatistics *13*, 762-775.
Lee, Y., Duchaine, B., Wilson, H.R., and Nakayama, K. (2010). Three cases of developmental prosopagnosia from one family: detailed neuropsychological and psychophysical investigation of face processing. Cortex *46*, 949-964.
Lee, Y., Grady, C.L., Habak, C., Wilson, H.R., and Moscovitch, M. (2011). Face processing changes in normal aging revealed by fMRI adaptation. J Cogn Neurosci *23*, 3433-3447.
Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B*., et al.* (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285-291.
Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754-1760.

Li, X.B., and Song, Y.Y. (2007). First congenital prosopagnosia case reported in China. Progress in Natural Science *17*, 165-171.
Liu, J., Harris, A., and Kanwisher, N. (2010). Perception of Face Parts and Face Configurations: An fMRI Study. Journal of Cognitive Neuroscience *22*, 203-211.
Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R., and Tootell, R.B.H. (1995). Object-Related Activity Revealed by Functional Magnetic-Resonance-Imaging in Human Occipital Cortex. Proceedings of the National Academy of Sciences of the United States of America *92*, 8135-8139.
Manichaikul, A., Mychaleckyj, J.C., Rich, S.S., Daly, K., Sale, M., and Chen, W.M. (2010). Robust relationship inference in genome-wide association studies. Bioinformatics *26*, 2867-2873.
McCarthy, G., Puce, A., Belger, A., and Allison, T. (1999). Electrophysiological studies of human face perception. II: Response properties of face-specific potentials generated in occipitotemporal cortex. Cereb Cortex *9*, 431-444.
McCarthy, G., Puce, A., Gore, J.C., and Allison, T. (1997). Face-specific processing in the human fusiform gyrus. J Cogn Neurosci *9*, 605-610.
McConachie, H.R. (1976). Developmental prosopagnosia. A single case report. Cortex *12*, 76-82.
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M.*, et al.* (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res *20*, 1297-1303.
McKone, E., Aimola Davies, A., Fernando, D., Aalders, R., Leung, H., Wickramariyaratne, T., and Platow, M.J. (2010). Asia has the global advantage: Race and visual attention. Vision Res *50*, 1540-1549.
McKone, E., Stokes, S., Liu, J., Cohan, S., Fiorentini, C., Pidcock, M., Yovel, G., Broughton, M., and Pelleg, M. (2012). A robust method of measuring other-race and other-ethnicity effects: the Cambridge Face Memory Test format. PLoS One *7*, e47956.
McMullen, P.A., Fisk, J.D., Phillips, S.J., and Maloney, W.J. (2000). Apperceptive agnosia and face recognition. Neurocase *6*, 403-414.
Meadows, J.C. (1974). The anatomical basis of prosopagnosia. J Neurol Neurosurg Psychiatry *37*, 489-501.
Moeller, S., Crapse, T., Chang, L., and Tsao, D.Y. (2017). The effect of face patch microstimulation on perception of faces and objects. Nature Neuroscience *20*, 743-+.
Moeller, S., Freiwald, W.A., and Tsao, D.Y. (2008). Patches with links: a unified system for processing faces in the macaque temporal lobe. Science *320*, 1355-1359.
Moscovitch, M., Winocur, G., and Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. Journal of Cognitive Neuroscience *9*, 555-604.
Ng, P.C., and Henikoff, S. (2001). Predicting deleterious amino acid substitutions. Genome Res *11*, 863-874.
Nunn, J.A., Postma, P., and Pearson, R. (2001). Developmental prosopagnosia: should it be taken at face value? Neurocase *7*, 15-27.
Park, J., Carp, J., Kennedy, K.M., Rodrigue, K.M., Bischof, G.N., Huang, C.M., Rieck, J.R., Polk, T.A., and Park, D.C. (2012). Neural broadening or neural attenuation? Investigating age-related dedifferentiation in the face network in a large lifespan sample. J Neurosci *32*, 2154-2158.
Pascalis, O., Deschonen, S., Morton, J., Deruelle, C., and Fabregrenet, M. (1995). Mothers Face Recognition by Neonates - a Replication and an Extension. Infant Behavior & Development *18*, 79-85.
Pascalis, O., and Kelly, D.J. (2009). The Origins of Face Processing in Humans: Phylogeny and Ontogeny. Perspectives on Psychological Science *4*, 200-209.
Perrett, D.I., Mistlin, A.J., and Chitty, A.J. (1987). Visual Neurons Responsive to Faces. Trends in Neurosciences *10*, 358-364.
Perrett, D.I., Mistlin, A.J., Chitty, A.J., Smith, P.A.J., Potter, D.D., Broennimann, R., and Harries, M. (1988). Specialized Face Processing and Hemispheric-Asymmetry in Man and Monkey - Evidence from Single Unit and Reaction-Time Studies. Behavioural Brain Research *29*, 245-258.
Perrett, D.I., Rolls, E.T., and Caan, W. (1982). Visual Neurones Responsive to Faces in the Monkey Temporal Cortex. Experimental Brain Research *47*, 329-342.
Perrett, D.I., Smith, P.A.J., Potter, D.D., Mistlin, A.J., Head, A.S., Milner, A.D., and Jeeves, M.A. (1984). Neurons Responsive to Faces in the Temporal Cortex - Studies of Functional-Organization, Sensitivity to Identity and Relation to Perception. Human Neurobiology *3*, 197-208.

Perrett, D.I., Smith, P.A.J., Potter, D.D., Mistlin, A.J., Head, A.S., Milner, A.D., and Jeeves, M.A. (1985). Visual Cells in the Temporal Cortex Sensitive to Face View and Gaze Direction. Proceedings of the Royal Society Series B-Biological Sciences *223*, 293-317.
Pitcher, D., Charles, L., Devlin, J.T., Walsh, V., and Duchaine, B. (2009). Triple Dissociation of Faces, Bodies, and Objects in Extrastriate Cortex. Current Biology *19*, 319-324.
Pitcher, D., Garrido, L., Walsh, V., and Duchaine, B.C. (2008). Transcranial magnetic stimulation disrupts the perception and embodiment of facial expressions. Journal of Neuroscience *28*, 8929-8933.
Polk, T.A., Park, J., Smith, M.R., and Park, D.C. (2007). Nature versus nurture in ventral visual cortex: A functional magnetic resonance Imaging study of twins. Journal of Neuroscience *27*, 13921-13925.
Pritchard, J.K. (2001). Are rare variants responsible for susceptibility to complex diseases? Am J Hum Genet *69*, 124-137.
Puce, A., Allison, T., Asgari, M., Gore, J.C., and McCarthy, G. (1996). Differential Sensitivity of Human Visual Cortex to Faces, Letterstrings, and Textures: A Functional Magnetic Resonance Imaging Study. The Journal of Neuroscience *16*, 5205-5215.
Puce, A., Allison, T., and McCarthy, G. (1999). Electrophysiological studies of human face perception. III: Effects of top-down processing on face-specific potentials. Cereb Cortex *9*, 445-458.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J.*, et al.* (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet *81*, 559-575.
Purcell, S.M., Moran, J.L., Fromer, M., Ruderfer, D., Solovieff, N., Roussos, P., O'Dushlaine, C., Chambert, K., Bergen, S.E., Kahler, A.*, et al.* (2014). A polygenic burden of rare disruptive mutations in schizophrenia. Nature *506*, 185-190.
Rangarajan, V., Hermes, D., Foster, B.L., Weiner, K.S., Jacques, C., Grill-Spector, K., and Parvizi, J. (2014). Electrical stimulation of the left and right human fusiform gyrus causes different effects in conscious face perception. J Neurosci *34*, 12828-12836.
Rausch, T., Zichner, T., Schlattl, A., Stutz, A.M., Benes, V., and Korbel, J.O. (2012). DELLY: structural variant discovery by integrated paired-end and split-read analysis. Bioinformatics *28*, i333-i339.
Riddoch, M.J., Johnston, R.A., Bracewell, R.M., Boutsen, L., and Humphreys, G.W. (2008). Are faces special? A case of pure prosopagnosia. Cognitive Neuropsychology *25*, 3-26.
Righart, R., Andersson, F., Schwartz, S., Mayer, E., and Vuilleumier, P. (2009). Top-Down Activation of Fusiform Cortex without Seeing Faces in Prosopagnosia. Cerebral Cortex *20*, 1878-1890.
Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol *29*, 24-26.
Rolls, E.T., Baylis, G.C., and Leonard, C.M. (1985). Role of Low and High Spatial-Frequencies in the Face-Selective Responses of Neurons in the Cortex in the Superior Temporal Sulcus in the Monkey. Vision Research *25*, 1021-1035.
Rossion, B., Dricot, L., Goebel, R., and Busigny, T. (2011). Holistic Face Categorization in Higher Order Visual Areas of the Normal and Prosopagnosic Brain: Toward a Non-Hierarchical View of Face Perception. *4*.
Rotshtein, P., Henson, R.N., Treves, A., Driver, J., and Dolan, R.J. (2005). Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. Nat Neurosci *8*, 107-113.
Ryan, C.M.E., and Lea, S.E.G. (1994). Images of Conspecifics as Categories to Be Discriminated by Pigeons and Chickens - Slides, Video Tapes, Stuffed Birds and Live Birds. Behavioural Processes *33*, 155-175.
Saito, H., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y., and Iwai, E. (1986). Integration of Direction Signals of Image Motion in the Superior Temporal Sulcus of the Macaque Monkey. Journal of Neuroscience *6*, 145-157.
Schiltz, C., Sorger, B., Caldara, R., Ahmed, F., Mayer, E., Goebel, R., and Rossion, B. (2006). Impaired face discrimination in acquired prosopagnosia is associated with abnormal response to individual faces in the right middle fusiform gyrus. Cereb Cortex *16*, 574-586.
Schmalzl, L., Palermo, R., and Coltheart, M. (2008). Cognitive heterogeneity in genetically based prosopagnosia: a family study. J Neuropsychol *2*, 99-117.
Schrott, H.G., Goldstein, J.L., Hazzard, W.R., McGoodwin, M.M., and Motulsky, A.G. (1972). Familial hypercholesterolemia in a large indred. Evidence for a monogenic mechanism. Ann Intern Med *76*, 711-720.

Sergent, J., Ohta, S., and Macdonald, B. (1992). Functional Neuroanatomy of Face and Object Processing - a Positron Emission Tomography Study. Brain *115*, 15-36.
Shah, P., Gaule, A., Sowden, S., Bird, G., and Cook, R. (2015). The 20-item prosopagnosia index (PI20): a self-report instrument for identifying developmental prosopagnosia. R Soc Open Sci *2*, 140343.
Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M., and Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. Nucleic Acids Res *29*, 308-311.
Slater, A., Bremner, G., Johnson, S.P., Sherwood, P., Hayes, R., and Brown, E. (2000). Newborn Infants' Preference for Attractive Faces: The Role of Internal and External Facial Features. Infancy *1*, 265-274.
Slater, A., Von der Schulenburg, C., Brown, E., Badenoch, M., Butterworth, G., Parsons, S., and Samuels, C. (1998). Newborn infants prefer attractive faces. Infant Behavior & Development *21*, 345-354.
Sobel, E., and Lange, K. (1996). Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. Am J Hum Genet *58*, 1323-1337.
Steeves, J.K., Culham, J.C., Duchaine, B.C., Pratesi, C.C., Valyear, K.F., Schindler, I., Humphrey, G.K., Milner, A.D., and Goodale, M.A. (2006). The fusiform face area is not sufficient for face recognition: evidence from a patient with dense prosopagnosia and no occipital face area. Neuropsychologia *44*, 594-609.
Stollhoff, R., Jost, J., Elze, T., and Kennerknecht, I. (2011). Deficits in long-term recognition memory reveal dissociated subtypes in congenital prosopagnosia. PLoS One *6*, e15702.
Susilo, T., and Duchaine, B. (2013). Advances in developmental prosopagnosia research. Current Opinion in Neurobiology *23*, 423-429.
Thorvaldsdottir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform *14*, 178-192.
Tsao, D.Y., Freiwald, W.A., Knutsen, T.A., Mandeville, J.B., and Tootell, R.B. (2003). Faces and objects in macaque cerebral cortex. Nat Neurosci *6*, 989-995.
Tsao, D.Y., Freiwald, W.A., Tootell, R.B., and Livingstone, M.S. (2006). A cortical region consisting entirely of face-selective cells. Science *311*, 670-674.
Tsao, D.Y., and Livingstone, M.S. (2008). Mechanisms of face perception. Annu Rev Neurosci *31*, 411-437.
Tsao, D.Y., Moeller, S., and Freiwald, W.A. (2008). Comparing face patch systems in macaques and humans. Proc Natl Acad Sci U S A *105*, 19514-19519.
Tsui, L.C., Buchwald, M., Barker, D., Braman, J.C., Knowlton, R., Schumm, J.W., Eiberg, H., Mohr, J., Kennedy, D., Plavsic, N.*, et al.* (1985). Cystic fibrosis locus defined by a genetically linked polymorphic DNA marker. Science *230*, 1054-1057.
Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J.*, et al.* (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics *43*, 11 10 11-33.
Verfaillie, K., Huysegems, S., De Graef, P., and Van Belle, G. (2014). Impaired holistic and analytic face processing in congenital prosopagnosia: Evidence from the eye-contingent mask/window paradigm. Vis Cogn *22*, 503-521.
von Kriegstein, K., Dogan, O., Gruter, M., Giraud, A.L., Kell, C.A., Gruter, T., Kleinschmidt, A., and Kiebel, S.J. (2008). Simulation of talking faces in the human brain improves auditory speech recognition. Proc Natl Acad Sci U S A *105*, 6747-6752.
Walton, G.E., Bower, N.J.A., and Bower, T.G.R. (1992). Recognition of Familiar Faces by Newborns. Infant Behavior & Development *15*, 265-269.
Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S.F., Hakonarson, H., and Bucan, M. (2007). PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res *17*, 1665-1674.
Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res *38*, e164.
Wilmer, J.B., Germine, L., Chabris, C.F., Chatterjee, G., Williams, M., Loken, E., Nakayama, K., and Duchaine, B. (2010). Human face recognition ability is specific and highly heritable. Proceedings of the National Academy of Sciences of the United States of America *107*, 5238-5241.
Yin, R.K. (1969). Looking at upside-down faces. Journal of Experimental Psychology *81*, 141-145.
Yovel, G., and Kanwisher, N. (2004). Face perception: domain specific, not process specific. Neuron *44*, 889-898.

Yovel, G., and Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. Curr Biol *15*, 2256-2262.
Zebrowitz, L., Ward, N., Boshyan, J., Gutchess, A., and Hadjikhani, N. (2016). Dedifferentiated face processing in older adults is linked to lower resting state metabolic activity in fusiform face area. Brain Research *1644*, 22-31.
Zhang, J., Li, X., Song, Y., and Liu, J. (2012). The fusiform face area is engaged in holistic, not parts-based, representation of faces. PLoS One *7*, e40390.
Zhao, Y., Zhen, Z., Liu, X., Song, Y., and Liu, J. (2018). The neural network for face recognition: Insights from an fMRI study on developmental prosopagnosia. Neuroimage *169*, 151-161.
Zhao, Y.F., Li, J.G., Liu, X.Q., Song, Y.Y., Wang, R.S., Yang, Z.T., and Liu, J. (2016). Altered spontaneous neural activity in the occipital face area reflects behavioral deficits in developmental prosopagnosia. Neuropsychologia *89*, 344-355.
Zhu, Q., Song, Y.Y., Hu, S.Y., Li, X.B., Tian, M.Q., Zhen, Z.L., Dong, Q., Kanwisher, N., and Liu, J. (2010). Heritability of the Specific Cognitive Ability of Face Perception. Current Biology *20*, 137-142.
Zhu, X.W., Liu, K.Q., Wang, P.Y., Liu, J.Q., Chen, J.Y., Xu, X.J., Xu, J.J., Qiu, M.C., Sun, Y., Liu, C.*, et al.* (2021). Cohort profile: the Westlake BioBank for Chinese (WBBC) pilot project. BMJ Open *11*, e045564.
Zhu, Z., Chen, B., Na, R., Fang, W., Zhang, W., Zhou, Q., Zhou, S., Lei, H., Huang, A., Chen, T.*, et al.* (2019). Heritability of human visual contour integration-an integrated genomic study. Eur J Hum Genet *27*, 1867-1875.
Zhu, Z., Wang, Y., Cao, Z., Chen, B., Cai, H., Wu, Y., and Rao, Y. (2016). Cue-independent memory impairment by reactivation-coupled interference in human declarative memory. Cognition *155*, 125-134.
Zhu, Z.J., Chen, B.Q., Yan, H.M., Fang, W., Zhou, Q., Zhou, S.B., Lei, H., Huang, A.L., Chen, T.M., Gao, T.M.*, et al.* (2018). Multi-level genomic analyses suggest new genetic variants involved in human memory. Eur J Hum Genet *26*, 1668-1678.

Figure 1. Mapping of a Novel Autosomal Dominant CP Locus to the Short Arm of Chromosome 15 (15q26.2).

(a) The pedigree plot of family A with genotypes for c.T2147G, p.I716S, NM_018349 and phenotypes. Crosses, DNA samples not available; diagonal slashes, dead members; filled symbols, CP candidates in linkage 1; half-filled symbols, added as CP candidates in linkage 2; an arrow head, the index of family A.

(b) Graphical representation of parametric linkage results of linkage 1 (LOD score = 3.20) in family A with diagnosis based only on interviews as well as behavioral tests, assuming a rare dominant model.

(c) Graphical representation of parametric linkage results of linkage 2 (LOD score = 4.27) in family A with diagnosis based only on interviews but not behavioral tests, assuming a rare dominant model.

Figure 2. Mutations Detection in *MCTP2* and Co-segregation in CP.

(a) The distribution of the 20-item questionnaire scores in 743 males (the blue approximate normal curve) and 2161 females (the red approximate normal curve). Dashed lines indicate the cut-off of 2 SDs poorer than the mean. Sticks with different colors indicate different individuals with different mutations.

(b) A mutation c.239delG, leading to p.S80fs in individuals 1-1, 1-2, 1-3 with CP. All three of them were aware of difficulties with face recognition early in their life (Table S8).

(c) Correlation of the (c.239delG, p.S80fs) genotype with the CP phenotype in the family of 1-1, with 3 members tested.

(d) A mutation c. 814A>T, leading to p. M272L in individual 1-4.

(e) Correlation of the genotype (c.814A>T, p.M272L) with the CP phenotype in the family of 1-4.

(f) A mutation c.1120A>G, leading to p.T374A in individual 1-5.

(g) A mutation c.1642G>A, leading to p.V548I in individual 1-6.

(h) Correlation of the (c.1642G>A, p.V548I) genotype with the CP phenotype in the family of 1-6, with 6 members tested (Table S8).

(i) A mutation c.1922G>A, leading to p.R641Q in individual 1-7.

(j) Correlation of the (c.1922G>A, p.R641Q) genotype with the CP phenotype in the family of 1-7, with 5 members tested (Table S8).

Figure 3. The Positions in the MCTP2 Protein Corresponding to *MCTP2* Mutations in CPs

(a) The domain plot of the MCTP2 protein and locations of alterations found in CPs. The blue rectangles represent C_2 domains, the yellow hexagons represent transmembrane segments; the orange lines indicate putative Ca^{2+} binding sites; the red triangles indicate amino acid changes.

(b) Evolutionary conservation of amino acid residues in the MCTP2 protein in different species. Amino acids corresponding to mutations found in CP subjects are denoted in different colors. Green: identical to those in the human MCTP2, blue: non-identical to those in human MCTP2, red: non-identical amino acids but the same as the mutated in human MCTP2. Mutated amino acid residues are indicated by triangles on the top.

Figure 4. Co-segregation of c.239delG, p.S80fs in *MCTP2* in CP.

16 frameshift deletion mutation (c.239delG, p.S80fs) were detected in a cohort of 1757 students. 14 of them were available for further analysis. More members from 6 families were available (Table S10).

(a) Correlation of the (c.239delG, p.S80fs) genotype with the CP phenotype in the family of 3-2, with 10 family members tested.

(b) Correlation of the (c.239delG, p.S80fs) genotype with the CP phenotype in the family of 3-3, with 4 family members tested.

(c) Correlation of the (c.239delG, p.S80fs) genotype with the CP phenotype in the family of 3-5 with 6 family members tested.

(d) Correlation of the (c.239delG, p.S80fs) genotype with the CP phenotype in the family of 3-6 with 4 family members tested.

(e) Correlation of the (c.239delG, p.S80fs) genotype with the CP phenotype in the family of 3-8, with 3 members tested.

(f) The pedigree plot of family 3-12 with genotype and phenotype.

Figure 5. rFFA Activation in Face Sensitivity

(a) Experimental paradigm. Gray-scale photographs of 24 pairs of upright or inverted faces were presented sequentially in a random order during the experiment. Cartoon faces are used as examples. The average percent signal change (±SD for controls) from baseline fixation is plotted for the upright and the inverted face conditions. Bold response to pairs of upright different faces or inverted different faces in the rFFA of controls (n=20) (b); non-CP members A-VI:1 (c), A-VI:2 (d), A-VI:10 (e) and CP members from family A with the I716S mutation in *MCTP2*: A-VI:5 (f), A-VI:7 (g),

and A-VI:9 (h); non-CP 3-2-III:2 (i) and CP 3-2-III:4 (j) from family 3-2 with the S80fs mutation in *MCTP2*.

Figure 6. rFFA Activation in Face Discrimination

Experimental paradigm as Figure 5A. The average percent signal change (±SD for controls) from baseline fixation is plotted for the different and identical upright face conditions. Bold response to pairs of upright different faces or identical faces in the rFFA of controls (n=20) (b); non-CP members A-VI:1 (c), A-VI:2 (D), A-VI:10 (E) and CP members A-VI:5 (F), A-VI:7 (G), and A-VI:9 (H) from family A with the I716S mutation in *MCTP2*; non-CP 3-2-III:2 (I) and CP3-2-III:4 (J) from family 3-2 with the S80fs mutation in *MCTP2*.

Figure 7. Comparison between CP and Control Individuals in the FIE and Adaptation Experiments

(a) The FIE score (difference in percent signal change between responses to upright different faces and inverted different faces) is plotted for CP individuals (black bars) and each individual control subject (white bars) in an increasing order.

(b) The adaptation score (difference in percent signal change between responses to upright different faces and upright identical faces) is plotted for CP individuals (black bars) and each individual control subject (white bars) in an increasing order.

Table 1. Diagnosis of CP in Family A

Members with dark grey background indicate CP candidates in linkage 1. Members with light grey background were added as CP candidates in linkage 2. Face processing z scores for members younger than 60 years of age in family A are coded as described in the method such that negative values indicate performances poorer than the mean. Red indicates below -2 SDs. Blue indicates below -1.5 SDs. Whether there is an obvious key manifestation for CP is expressed by yes or no. More detailed diagnostic information can be obtained by contacting the corresponding author.

Table 2. *MCTP2* Mutations Found in Sporadic CP Candidates Screened by a 20-Item Questionnaire

NS: nonsynonymous single nucleotide variant; FD: frameshift_deletion; AAF_EAS: alternative allele frequency in EAS (East Asian) population; SIFT: predicts whether an amino acid substitution affects protein function as D (Deleterious) or T (tolerated) by SIFT algorithm; Polyphen2_HDIV, Polyphen2_HVAR: predicts possible impact of an amino acid substitution on the structure and function of a human protein using straightforward physical and comparative considerations as D (Probably damaging), P (possibly damaging) or B (benign) by PolyPhen 2 (Polymorphism Phenotyping v2); LRT: uses comparative genomics to predict whether variants disrupt highly conserved amino acids as D (deleterious), N (neutral) or U (unknown) by LRT (Likelihood Ratio Test); Mutation Taster: predicts an alteration as one of four possible types: D (disease causing - i.e. probably deleterious), A (disease causing automatic - i.e. known to be deleterious), N (polymorphism - i.e. probably harmless), P (polymorphism automatic - i.e. known to be harmless). These abbreviations also apply to the following tables. * Start Position Based on the *MCTP2* RefSeq (hg19, NM_018349); # dbSNP build 150, & RefGene: NM_018349.

Table 3. Gene Based Association Analysis for Rare Variants in *MCTP2*

Table 4. Differences in Daily Face Recognition between Target Group and Controls Regarding the p.S80fs Mutation.

Figure S1. Genome-wide Screen Results of CP-susceptibility Loci. Allele-sharing multipoint LOD scores are shown. (a) for linkage 1, and (b) for linkage 2.

Figure S2. Haplotype Analysis for MCR on Chromosome 15q

(A) 315 markers with LOD scores over 3 in linkage 1 were used in haplotype analysis. Haplotypes are color-coded, with red indicating co-segregating with CP.

(B) 400 markers with LOD scores over 3 in linkage 2 were used in haplotype analysis. Haplotypes are color-coded, with red indicating co-segregating with CP.

Figure S3. Inferred Haplotypes around the c.239delG (p.S80fs) Mutation

These were estimated by comparison to microsatellites data collected on mother-father-offspring trios. Haplotypes coded in black contain the c.239delG (p.S80fs) mutation. The haplotype on the left is paternal and the haplotype on the right is maternal.

Figure S4. Brain fMRI from the Localization Experiment: Identifying the FFA ROI.

Statistical parametric maps by contrasting the response to faces versus non-face stimuli, showing significant face-specific bilateral or unilateral activity in the FFA. Color scale represents statistical values of a one level t test comparing the fMRI signal while subjects viewed blocks of faces versus blocks of non-face stimuli.

(A) rFFA responses in normal controls, $p_{FWE} < 0.05$.

(B) rFFA responses in A-VI:5 CP member with the I716S mutation, $p_{FWE} < 0.05$.

(C) rFFA responses in A-VI:7 CP member with the I716S mutation, $p < 0.05$.

(D) rFFA responses in A-VI:9 CP member with the I716S mutation, $p_{FWE} < 0.05$.

(E) rFFA responses in A-VI:1 non-CP member, $p_{FWE} < 0.05$.

(F) rFFA responses in A-VI:2 non-CP member, $p_{FWE} < 0.05$.

(G) rFFA responses in A-VI:10 non-CP member, $p < 0.001$.

(H) rFFA responses in 3-2-III:4 S80fs CP member, $p < 0.0001$.

(G) rFFA responses in 3-2-III:2 non-CP member, $p_{FWE} < 0.05$.

Figure S5. Scatterplots of Age against Total Correct for (a) the Cambridge Face Memory Test-Chinese (CFMT-C), and (b) the Cambridge Car Memory Test (CCMT)

Male control individuals are shown as blue circles; female control individuals are shown as red circles. The curves show the best fitting second-order polynomial relating age to the relevant performance measure with different genders, blue for male, red for female.

Figure S6. Scatterplots of Age against Total Correct for (a) the Upright Face, and (b) the Inverted Face.

Control individuals are shown as blue circles. The curves show the best fitting second-order polynomial relating age to the relevant performance measure.

Table S1. Summary of Phenotypic Analysis and Genetic Analysis of Each Member of Family A

Table S2. Summary of Computer Simulations of Linkage Analysis

Table S3. Candidate Genes Annotated by the NCBI Database (NC_000015.9) in Regions of Linkage Analyses with LOD Scores over 3

Table S4. the Summary Analysis of the Sequencing Data

Table S5. PCR Primer Sequences for Each Exon and Mutation Site of *MCTP2*

L, the left primer; R, the right primer; seq, the primer for sequencing.

Table S6. 11 Heterozygous Variants with MAF under 0.05 Shared with the Affected Individuals of Family A by WGS

Table S7. PCR Primer Sequences for 11 Heterozygous Variants (MAF under 0.05) Shared by the Affected Individuals of Family A by WGS

Table S8. Most Discriminative Symptoms for the Diagnosis of Additional CP Individuals and Families

Member codes with dark grey background indicate CP candidates who showed daily hallmarks of CP. The index for each family is shown in bold and underlined. Whether there is an obvious key manifestation for CP is expressed by yes or no. More detailed diagnostic information can be obtained by contacting the corresponding author.

Table S9. Rare Variants (MAF<0.005) in *MCTP2* from a Second Cohort of 1928 Students

Table S10. Symptoms for the Diagnosis in Persons with c.239delG, p.S80fs

Member codes with dark grey background indicate CP candidates who showed daily symptoms of CP. The index for each family is shown in bold and underlined. Whether there is an obvious key manifestation for CP is expressed by yes or no. More detailed diagnostic information can be obtained by contacting the corresponding author.

SUBJECT DETAILS

Participants

This study has been approved by the ethics committee at Peking University. Written informed consent to participate in the research study and to have the results of this research work published was obtained from participants or their legal representatives prior to any tests. Genomic DNA was isolated from whole blood (BD Vacutainer 3.2% Sodium Citrate (1:9)) using the Gentra Puregene Blood Kit (QIAGEN) or saliva using

the GeneFiX DNA Saliva Collector and Isolation Kit.

Family A

The proband V:11in family A was ascertained by self-referring. Initially, 35 available family individuals (18 CPs, including 9 males and 9 females, aged from 16 to 72; 17 non-CPs, including 9 males and 8 females, aged from 15 to 78) were assessed based on a standardized semi-structured interview (Carbon et al., 2007; Grueter et al., 2007; Kennerknecht et al., 2006). 25 family members younger than 60 years of age performed the CFMT-C (McKone et al., 2012), the matched CCMT (Dennett et al., 2012) and the FIE Discrimination Test (Yovel and Kanwisher, 2005). 34 individuals (not including VI:9) were included in the linkage study. Four individuals (V:6, V:9, V:11, and V:19) were selected for the WES. Two individuals (V:19 and VI:5) were selected for the WGS. This information is listed in Table S1.

Additional CP Individuals with Mutations in *MCTP2*

To look for more CP individuals and examine whether rare functional mutations in *MCTP2* are associated with CP, 2904 individuals (average age of 19.25 $\pm$ 1.30, 2161 females, 743 males) from the Westlake BioBank for Chinese (WBBC) pilot project (Zhu et al., 2021), were screened with a questionnaire adapted from the 20-item self-report measure for quantifying CP traits. 78 individuals scored worse than the mean by 2 SDs. 75 of them provided DNA samples. 44 individuals including 7 individuals with *MCTP2* mutations agreed to an interview for CP diagnosis and family members were contacted for availability of further studies.

Samples for Gene Based Association Analysis

The second cohort of 1928 individuals (average age of 18.51 ± 0.93, 1085 females, 843 males) used in the association of the rare functional alleles in *MCTP2* with face recognition ability were also from the WBBC project (Zhu et al., 2021). The coding sequences of the *MCTP2* gene for each person was analyzed by tagged-amplicon deep sequencing.

Individuals with the c.239delG, p.S80fs Mutation in *MCTP2*

A third cohort of 1757 individuals (average age of 19.13 ± 1.07, 1295 females, 462 males) were sequenced for the presence of the c.239delG, p.S80fs mutation in the exons of *MCTP2*. Individuals and family members carrying this mutation were further contacted and assessed by the standardized semi-structured interview. The coding sequences of the *MCTP2* gene for each person was analyzed by tagged-amplicon deep sequencing. The differences in the behavior of daily face recognition between the 14 carriers and 19 non-carrying volunteers from the same cohort were analyzed by independent sample t-test.

Control Samples for Behavioral Tests

338 normal participants (average age of 42.68 ± 16.53, 164 females, 174 males) were tested. These participants were unrelated. Participants were not selected for face recognition ability except for known history of major brain injury, or other major disorders likely to affect face recognition (e.g., Alzheimer's disease), representing a random sample of the community. They were tested on a battery of tests including the CFMT-C, the CCMT, the FIE Test and the Cambridge Face Perception Test - Chinese (CFPT-C).

fMRI Subjects

Age-related dedifferentiation and compensatory changes in the functional network underlying face processing have been found in studies (Burianova et al., 2013; Germine et al., 2011b; Goh et al., 2010; Le Grand et al., 2006; Lee et al., 2011; Park et al., 2012; Zebrowitz et al., 2016). Therefore, in order to exclude age influence on fMRI results, we mainly studied family members younger than 30 years old. Family members who participated in the fMRI experiment were on a voluntary basis. In family A, 3 CP individuals (VI:5, VI:7 and VI:9) and 3 non-CP individuals (VI:1, VI:2 and VI:10) took part in fMRI; in family 3-2, III:4 (CP) and III:2 (non-CP) took part in fMRI.

21 non-CP students (average age = 23.63 $\pm$ 3.71, 6 females, 15 males) were recruited for fMRI analysis.

CP referred to those with obvious face recognition problems in daily life diagnosed by the structured interview.

All participants reported normal or corrected to normal vision, no history of neurological or psychiatric conditions and all were right handed. Anatomical volumes (i.e., structural MRIs) had been routinely checked. One of the normal students was excluded from neuroimaging studies because the maximum head rotation was over 1.5 degree or the maximum translation was over 2 mm during localization.

We adopted the methods of studying AP (acquired prosopagnosia) cases that the severity of sporadic cases could be reported by comparison with the control population (Busigny and Rossion, 2010; Dricot et al., 2008; Gao et al., 2019; Righart et al., 2009; Rossion et al., 2011; Schiltz et al., 2006). In fMRI analysis, each family member was compared individually to a small sample of normal controls by a modified t-test (Crawford and Garthwaite, 2002; Crawford and Howell, 1998).

METHODS DETAILS

Structured Interview for CP Diagnosis

The diagnosis of CP was based on a standardized semi-structured interview (Table 1, S8 and S10) (Grueter et al., 2007; Kennerknecht et al., 2006), which documented a variety of impressive qualitative differences between CP and non-CP, and had been validated with objective face recognition tests in previous studies (Carbon et al., 2007; Grueter et al., 2007). These criteria were employed in recent literature on CP (Bate et al., 2019; Bate et al., 2014; Carbon et al., 2010; Dinkelacker et al., 2011; Esins et al., 2016; Gruter et al., 2009; Stollhoff et al., 2011; Verfaillie et al., 2014; von Kriegstein et al., 2008; Zhao et al., 2018; Zhao et al., 2016).

The interviewer asked questions in a semi-structured interview format with three or four questions about each diagnostic item. Interviews were held to embed the questions into conversations and questions about the same diagnostic items not asked sequentially. Interviews included a medical history in order to exclude conditions which might cause or mimic prosopagnosia. Affected people present a lack of confidence with face recognition. Symptoms include lasting and irritating subjective uncertainties of face recognition, failure to recognize familiar people out of context or in crowded places, overlooking familiar people and confusing strangers with familiar people, face recognition/learning time longer than socially accepted, development of adaptive behavior of critical situations and strategies for visual recognition of people and time of onset.

Consistent with the interview results, some individuals were aware of their CP before we contacted them. Some individuals who could identify people via non-facial features such as voice, gait and general appearance and manner, or were unaware of face recognition problems, but had developed obvious compensatory strategies to cope

with difficulties were also diagnosed as CPs.

CP Questionnaire

To effectively screen CP candidates from big samples, we adapted a Chinese 20-item version from the 20-item self-report measure for quantifying prosopagnosic traits (Shah et al., 2015), which asks about tangible experiences. This 20-item Questionnaire was included in a set of questionnaires for many research purposes and filled out by two cohorts of individuals online. Invalid questionnaires were dropped due to no distinction between forward and reverse questions. 2904 valid questionnaires for the first cohort and 1928 valid questionnaires for the second cohort were collected. The internal reliability measured by Cronbach's α was 0.828 and 0.902 respectively. In the first cohort, 343 individuals finished the questionnaire for a second time several weeks later and the correlation for the first and second results across each individual was very high by Pearson correlation coefficient analysis, $r = 0.081$, $p < 0.001$. There was a significant difference of the score distribution between the female and male participants ($p < 0.001$, two tailed t test), so the candidate CPs were screened and the gene based association study were carried with different genders respectively. An individual with a score beyond 2 SDs of the mean of the controls was defined as a CP.

Stimuli and Procedures of Behavioral Tests

All behavioral tests were adapted and integrated into a whole set using the Hyper Text Markup Language. All participants were tested individually. The tests were run on a desktop PC with screen resolution 1024 × 768, refresh rate 85 Hz. Participants were seated at a viewing distance of approximately 50 cm from the screen. All participants were tested wearing their usual optical correction. Participants were asked to confirm

that they could focus without seeing blur on the computer screen. No participant reported any difficulty with focus at these distances. Gray-scale adult Chinese faces were used.

Cambridge Face Memory Test-Chinese (CFMT-C)

The Chinese face version of CFMT (McKone et al., 2012) was kindly provided by Professor Jia Liu of Beijing Normal University which we integrated with other behavioral tests into a whole set, and performed according to the standard procedure including the practice phase, the "Learn" phase with 18 trials, the "Novel Images" phase with 30 trials and the final "Novel Images with Noise" with 24 trials. All faces were Chinese male, shown without hair or facial blemishes and with neutral expressions. Participants were instructed to press the key corresponding to the number 1, 2 and 3 below faces. The test includes a total of 72 trials. Scores were reported as percent correct across the full test.

Cambridge Car Memory Test (CCMT)

CCMT is a test similar in the experimental design as CFMT, with stimuli replaced by whole cars (Dennett et al., 2012). We used CCMT as a control of CFMT to test for potential general object recognition deficits and to quantify the individual ability of performing such kind of tests. The original version of CCMT was kindly provided by Professor Bradley Duchaine of Dartmouth College, USA, which we integrated with other behavioral tests into a whole set, and performed according to the standard procedure (Dennett et al., 2012) including the practice phase, the "Learn" phase with 18 trials, the "Novel Images" phase with 30 trials and the final "Novel Images with Noise" with 24 trials. The stimuli were the modified computer-generated images of

actual car makes and models from the original CCMT. Participants were instructed to press the key corresponding to the numbers 1, 2 and 3 below cars. The test included a total of 72 trials. Scores were reported as percent correct across the full test.

Cambridge Face Perception Test-Chinese (CFPT-C)

We developed a version of CFPT using Chinese faces as stimuli (CFPT-C), according to the standard procedure (Duchaine et al., 2007) by morphing six different individuals with the target face, containing 88%, 76%, 64%, 52%, 40%, and 28% of the target face in turn. The Chinese faces were male, shown without hair or facial blemishes and with neutral expressions. All faces were photographs of Chinese students at Peking University, with written consent forms collected before photographing. Each individual was photographed in the same range of views and lighting conditions. Eight upright and eight inverted trials were intermixed, with the upright trial occurring first half the time. Participants had one minute to arrange six morphed faces according to their similarity to a target face by clicking on a face and indicating where that face should be moved by clicking in the area between two faces. Scores were computed according to the previous paper (Duchaine et al., 2007). The internal reliability measured by Cronbach's α in our sample of 170 individuals was 0.705 for upright faces, and 0.483 for inverted faces. However, after our pilot study, the CFPT-C was excluded from further study, because strategies such as just comparing partial facial features were used in normal people as in CPs and inconsistent results were reported by others (Anstey et al., 2005; Bowles et al., 2009; Duchaine et al., 2007).

Face Inversion Effect (FIE) Discrimination Test

The stimuli consisted of 20 gray-scale individual face images, cropped using the same

4X4.5cm oval window to remove cues from the hairline and face contour. All faces were photographs of Chinese students at Peking University, with written consent forms collected before photographing. Photographs were not repeatedly used in different tests. Pairs of face stimuli were presented sequentially either upright or inverted in a randomized order. The first face stimulus was presented in the upper-left quadrant of the screen for 0.5 s. After an interstimulus interval of 0.5 s, the second stimulus was presented in the lower-right quadrant for 0.5 s. The next trial will not begin until a same or different response was made by pressing one of two keys to respond by the participant. 80 trials were conducted in this test, half of them with upright faces and half with inverted faces. In each condition (upright or inverted), the chance with the two identical faces was 50%. A few practice trials were presented before the beginning of the experiment. Scores were reported as percent correct for each condition. For the measure of FIE, the difference in performance level between upright and inverted faces, an FIE index was calculated by entering the correction of performance for upright and for inverted faces in the following formula: FIE = (upright –inverted) / (upright + inverted).

Data Analysis for Behavioral Tests

Due to possible effects of ageing and sex difference on the scores of CFMT-C, CCMT, upright faces and inverted faces in FIE and the FIE index, we used multiple stepwise regression analyses to identify the covariates (sex, age) specific to each trait from 338 participants. The results of CFMT-C and CCMT showed noticeable age-related decline and sex differences. The scores of upright faces and inverted faces in FIE correlated with ageing but not sex, while the FIE index was not affected by age or sex among the control samples.

Figure S1 shows a scatterplot of age against CFMT-C score (a) and CCMT score (b) including all participants (aged 15–83 years) in different sex. We fitted age-related curves to the CFMT-C and CCMT scores with different sex, as shown $\text{CFMT-C}_{female} = -0.0087\text{age}^2 + 0.4954\text{age} + 71.508$,

$\text{CFMT-C}_{male} = -0.0069\text{age}^2 + 0.3159\text{age} + 69.506$,

$\text{CCMT}_{female} = -0.0073\text{age}^2 + 0.5034\text{age} + 50.433$,

$\text{CCMT}_{male} = -0.0087\text{age}^2 + 0.4954\text{age} + 71.508$.

Figure S2 shows a scatterplot of age against upright face score (a) and inverted face score (b) in FIE including all participants (aged 15–83 years). We also fitted age-related curves to the upright face and inverted face scores, as shown $\text{Upright} = -0.0051\text{age}^2 + 0.2159\text{age} + 79.821$,

$\text{Inverted} = -0.0016\text{age}^2 + 0.0173\text{age} + 63.813$.

Examining function curves suggested that the behavior performance remained stable across early middle age, but began to decline noticeably at approximately 50 years of age. This is consistent with previous studies. In terms of the validity of the behavioral scores, only individuals with ages under 60 (average age = 36.51 $\pm$ 12.34, 130 females, 138 males) were included to calculate the best estimate of CP cut-offs.

Significant covariates for each behavioral score were considered for the adjustment of behavioral performance. We used the fit-and-residual procedure to calculate the standard residual of each participant (z score, a reflection of the behavioral performance) for CP family members as described previously (Bowles et al., 2009). The regression function describing the relationship between age, sex and CFMT was: CFMT score = -0.212 age + 6.188 sex + 73.724. The difference between the participant's CFMT score and the estimated score from the function given his/her age and sex (i.e. the residual)

was calculated. The standard deviation of the controls' residuals of CFMT score was: 11.189. Z score was calculated by dividing the participant's residual by the standard deviation of the controls' residuals.

The regression function describing the relationship between age, sex and CCMT was: CCMT score = -0.134 age - 2.968 sex + 69.444. The difference between the participant's CCMT score and the estimated score from the function given his/her age and sex (i.e. the residual) was calculated. The standard deviation of all the participants' residuals for CCMT score was: 10.252. The z score was calculated by dividing the participant's residual by the standard deviation of the controls' residuals. The regression function describing the relationship between age and upright faces was: Upright score = -0.170 age + 87.641. The difference between the participant's upright faces score and the estimated score from the function given his/her age (i.e. the residual) was calculated. The standard deviation of all the participants' residuals for upright faces score was: 8.862. The z score was calculated by dividing the participant's residual by the standard deviation of the controls' residuals. Participants with z score below 2 SDs were considered performing poorly in these tests.

For the FIE index, there was no correlation with age or sex. Normalized z score was calculated for each subject by subtracting the mean of the control sample and dividing by the control samples' SD, and was considered as performing poorly if below 2 SDs.

GENETIC ANALYSES

Genome-wide Genotyping for Linkage Analysis

DNA samples were genotyped using Infinium Human OmniZhongHua BeadChips (Illumina), and normalized bead intensity data obtained for each sample were converted

into SNP genotypes using Genome Studio. SNPs were then selected according to the following parameters: genotyping rate > 95%, minor allelic frequency > 1% and no significant deviation from Hardy–Weinberg proportions ($p > 0.001$) by PLINK software (Purcell et al., 2007). Gender corresponding to each DNA sample was checked by analysis of X chromosome heterozygosity using PLINK. The initial Mendelian inheritance in family A0001 was analyzed by PLINK and KING toolset (Manichaikul et al., 2010).

Data Analysis for Linkage Study

Parametric linkage analysis was performed with the Merlin programs (Abecasis et al., 2002) , assuming autosomal dominant inheritance with 80% penetrance (for linkage 1, 90% for linkage 2), disease allele frequency 0.001, and phenocopy rate 0.05. Because the presence of linkage disequilibrium (LD) might inflate linkage statistics, LD maps were constructed with the PLINK tool within family members with the LD thresholds ($r^2 < 0.5$). Following data QC, 177,126 ($r^2 < 0.5$) informative SNPs were selected for linkage analysis with Merlin (Abecasis et al., 2002). Because the Merlin software could analyze a large set of polymorphic markers in a limited number of family members, family A was split into 2 sub-families with common ancestors. 1000 simulation analyses were performed to test for the statistical power of the significant results under the null hypothesis of no linkage, with simulated data, while maintaining the pedigree structure, allele frequencies, and recombination fraction. By convention, parametric linkage results were reported in LOD scores. A LOD score of > 2.5 was taken as being highly suggestive of linkage while a LOD score of > 3.0 was taken as convincing evidence of linkage. Haplotype construction was performed using the SimWalk2 program (Sobel and Lange, 1996).

Generation of CNV Calls

Genotyping and signal intensity data (genotype, GType; Log R ratio, LRR and B allele frequency, BAF) were exported from the GenomeStudio software 2011.1 (San Diego, CA, USA). The subsequent CNV calling analyses were performed using PennCNV (v.2011 Jun16) (Wang et al., 2007) according to the manual. Minimum number of markers of 3 were set as threshold for all CNV calling. The samples met the following QC measures were thought with good quality for further analysis: (1) within the standard deviation (SD) of the normalized intensity (LRR < 0.35); (2) |GC base pair wave factor (GCWF)| < 0.05; (3) with CNV call count < 70.

Whole Exome Sequencing (WES)

WES was performed by the Next-Generation Sequencing Center of Biomedical Pioneering Innovation Center, PKU. Exome capture was performed using Agilent SureSelect Human All Exon V4 capture kit. 100 bp paired-end sequencing was done on Illumina HiSeq 2500.

Data Analysis for WES

Sequences were aligned to human genome build hg19 using BWA (Li and Durbin, 2009). Duplicate reads were marked using Picard (http://broadinstitute.github.io/picard/). Genome Analysis Toolkit (GATK) (DePristo et al., 2011; McKenna et al., 2010; Van der Auwera et al., 2013) was used for base quality score recalibration (BQSR) and local realignment around indels to refine alignment artifacts around putative insertions or deletions. Variant discovery was performed in two steps: variant calling with GATK HaplotypeCaller followed by joint

genotyping using GATK GenotypeGVCFs. The resulting variant call set was refined using Variant Quality Score Recalibration (VQSR) as implemented in GATK VariantRecalibrator according to the GATK Best Practices recommendations (DePristo et al., 2011; Van der Auwera et al., 2013). The VQSR scores were used to define low quality variants for downstream processing. If the coverage in MCR was under 20-fold depth in all the 4 CP individuals investigated by WES, we performed direct Sanger sequencing to verify these regions. All variants were annotated to RefSeq hg19 and ANNOVAR (Wang et al., 2010) were used to add alternative allele frequencies, variant effect predictions and functional annotations.

Whole Genome Sequencing (WGS)

WGS was also performed by the Next-Generation Sequencing Center of Biomedical Pioneering Innovation Center, PKU. Sequencing libraries were built with NEBNext Ultra DNA Library Prep Kit for Illumina. 150 bp paired-end sequencing was done on Illumina HiSeq 4000.

Data Analysis for WGS

FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was applied to perform quality checks on various aspects of sequencing quality. Low quality bases were removed by Trimmomatic software (Bolger et al., 2014) with the parameters HEADCROP:5, LEADING:25, TRAILING:25, SLIDINGWINDOW:4:15 MINLEN:35. After that, clean sequences were aligned to human genome build hg19 using BWA (Li and Durbin, 2009). Duplicate reads were marked using Picard (http://broadinstitute.github.io/picard/). SNPs and INDELs were called by GATK HaplotypeCaller (McKenna et al., 2010). DELLY (Rausch et al., 2012) and

BreakDancer (http://www.nature.com/nmeth/journal/v6/n9/abs/nmeth.1363.html) were used to discover structural variants. BAM files locally realigned around INDELs were used to determine the average coverage, using GATK DepthOfCoverage package and default settings. If the coverage in MCR was under 20-fold depth in all the 2 CP individuals investigated by WGS, we performed direct Sanger sequencing to verify these regions. All variants were annotated to RefSeq hg19 and ANNOVAR(Wang et al., 2010) were used to add alternative allele frequencies, variant effect predictions and functional annotations.

Tagged-Amplicon Deep Sequencing

Target-specific primers for the coding sequences of the *MCTP2* gene (NM_018349) were designed with universal primer sequences (termed CS1 and CS2) appended at the 5'-end as described previously (Forshew et al., 2012).

A preamplification step was performed with a pool of the target-specific primer pairs to preserve representation of all alleles in the template material. The coding sequences of *MCTP2* gene were amplified using a 6μl PCR reaction mixture, containing 3μl KAPA2G Robust HotStart ReadyMix (2X), 1μl primers mix (500nM) and 2μl DNA templates (10ng/μl). PCR conditions were as follows: initial denaturation for 2 min at 95℃; 45 cycles of 30s denaturation at 95℃, 30s annealing at 56℃ and 1min extension at 72℃; and a final extension step for 5min at 72℃. Following PCR amplification, 1.5μl Shrimp Alkaline Phosphatase (SAP, affymetrix)/Exonuclease I (Exo I, BioLabs) mix was added to 2.5μl PCR products and incubated for 60min at 37℃, then for 20min at 80℃. SAP/Exo I mix comprised 600μl SAP (1U/μl), 240μl SAP buffer, 150μl Exo I (20,000U/ml), 200μl Exo I buffer and 3000μl ddH_2O. The

barcode primers (Fluidigm) consisted of the PE1 and PE2 sequences for Illumina cluster generation, a 10-bp barcode, and the CS1 and CS2 adaptors, used in pairs: PE1-CS1 with PE2-BC-CS2, and PE1-CS2 with PE2-BC-CS1. For each sample, 1 μl of the 100-fold diluted PCR products was added to each of two PCR plates containing 9μl pre-sample mix containing 0.4μl 1U/μl KAPA HiFi HotStart DNA Polymerase, 4μl 5× KAPA HiFi Buffer, 120μM each of dNTPs and 4μl ddH_2O. In the first plate, 10μl primers containing PE1-CS1 and PE2-BC-CS2 were added to each well. In the second plate, 10μl of primers containing PE1-CS2 + PE2-BC-CS1 were added to each well. The corresponding wells in both plates contained primers with the same barcode sequences. Reaction products in plates were amplified for 12 cycles: 95 ℃ 10 min, 12 cycles of 95 ℃ 15 s, 60 ℃ 30 s, 72 ℃ 3 min, 1 cycle of 72 ℃ for 3 min. After PCR products were barcoded, they were analyzed by gel electrophoresis to ensure expected insert size was obtained. They were then pooled together with equal volume and purified using size selected AMPure XP beads.

The library was quantified by Agilent BioAnalyzer and sequenced using the Illumina Hiseq2500 platforms with paired-end reads of 150bp according to the manufacturer's instructions. Custom sequencing primers targeted to the CS1 and CS2 were used to target the paired reads and 10-base indexing (barcode) read according to recommendations from Fluidigm.

Data Analysis for Tagged-Amplicon Deep Sequencing

The raw paired 150 bp-long reads were mapped to the human reference genome (build hg19) using BWA. GATK was then used to perform local realignment and recalibrate base quality scores, producing a BAM file for each individual. The raw single

nucleotide variants (SNVs) were filtered based on the variant quality score recalibration module in GATK. All variants were annotated to RefSeq hg19 and ANNOVAR(Wang et al., 2010) were used to add alternative allele frequencies, variant effect predictions and functional annotations. We manually inspected each mutation using the Integrative Genomics Viewer (Robinson et al., 2011; Thorvaldsdottir et al., 2013) to rule out false positive findings. If the coverage for each exon was under 500-fold depth in each individual, we performed direct Sanger sequencing to verify these regions.

Sequence Validation

Mutations were amplified by PCR and validated by direct Sanger DNA sequencing. All reactions were 100% successfully validated. Primer sequences for PCR amplification and Sanger sequencing are included in Table S5, S7.

Founder Origin Testing

We genotyped nine multi-allelic microsatellite markers (CHLC.ATA22D04, AFMB077YD5, GATA128A02, AFM072YB11, AFM357TD9/D15S1038, CHLC.GATA73F01, GATA161C02, AFM217ZG1, AFM309VG9) around the c.239delG (p.S80fs) mutation of *MCTP2* from 94255659 bp to 96210769 bp on chromosome 15 in families 1-1, 3-6, 3-2, 3-8, 3-5 and 3-3. The inferred haplotypes were estimated by comparison to these microsatellites data collected on mother-father-offspring trios to test the possibility of the same founder origin.

NEUROIMAGING

Stimuli.

In the localizer experiment, images of faces, non-face objects (e.g., chairs, food, and

tools), and texture patterns (scrambled faces) were presented at the center of the screen, subtended 6.2 °×6.2 °. In the adaptation paradigm, the stimuli were gray-scale images of young Chinese men (hair cropped with neutral expressions), the same as those used in the FIE behavioral test. The stimuli were presented by a MRI compatible projector system (SA-9900, The Shenzhen Sinorad Medical Electronics Co., Ltd, China, http://www.sinorad.com), with a spatial resolution of 1,024×768 and a refresh rate of 60 Hz.

Experimental Design and Procedures.

Each participant completed the same single scan session consisting of one functional localizer run and five runs for the adaptation paradigm. During the experiment, participants lay on their back in the scanner, using earplugs to reduce noise and sponges to hold their heads in place to reduce head movement. Participants viewed the stimuli presented on a translucent screen visible via a mirror mounted to the head coil at a distance of 60 cm. In the localizer run which lasted 360 seconds (s), with a 12 s dummy at the beginning of the run, images appeared at a rate of 2 Hz in blocks of 12 s, interleaved with 12 s blank blocks. There were five blocks for each type of images in the run. Each image was presented for 300 microseconds (ms), followed by a 200 ms blank interval. Subjects performed a one-back task during scanning to ensure maintenance of attention to the stimuli. In the adaptation paradigm, pairs of face or house stimuli were presented sequentially either in the upright or the inverted manner in a randomized order that was optimized for the extraction of the hemodynamic response in an event-related fast presentation design. Each trial lasted 2000 ms. The first and second stimuli were presented for 250 ms, with an interval of 500 ms and followed by 1 s fixation. Blank screen with white cross fixation point was set between trials, with a presentation time of a random even number in the range of 0-10 s to

optimize the efficiency of the event-related fMRI design by optseq2 (http://surfer.nmr.mgh.harvard.edu/optseq). Each run lasted 332 s, containing eight conditions (upright same face, upright different face, inverted same face, inverted different face, upright same house, upright different house, inverted same house and inverted different house). Five runs were included, each containing 12 trials for each condition. Subjects made a same/different response on each trial. The purpose we used an event-related design instead of block design of the adaptation paradigm, was to mimic the FIE behavioral test used to explore the ability of face recognition in CPs based on computers. Here we focus only on the face stimuli.

fMRI Scanning

The fMRI data were collected in a 3T GE MR 750 scanner, with an 8-channel phase-array head coil (GE Healthcare, Waukesha, WI) at Peking University Center for MRI Research.

The gradient-echo echo planar imaging (EPI) sequence was employed for BOLD images acquisition, and the imaging parameters were as below: repetition time (TR) = 2000 ms, echo time (TE) = 30 ms, field of view (FOV) = 224mm × 224mm, matrix = 64 × 64, flip angle = 90°, slice thickness = 3.5 mm with 0.7 mm spacing, voxel size = 3.5 × 3.5 × (3.5 + 0.7) mm, 33 oblique slices covering the whole brain).

The structure images were acquired by a 3D inversion recovery-prepped T1-weighted sequence (fSPGR, sagittal acquisition, TR = 6.65ms, TE = 2.92ms, TI = 450ms, flip angle = 12°, FOV = 256mm × 256 mm， matrix = 256 × 256, 192 continue slices with 1mm slice thickness, voxel size = 1mm × 1mm × 1mm).

Data Analysis for Neuroimaging Studies

Preprocessing and data analysis were performed with SPM12 (Wellcome Trust Centre for Neuroimaging, London; http://www.fil.ion.ucl.ac.uk/spm/software/spm12/). Functional images were sequentially processed in accordance with the standard SPM approach as follows: interpolated to correct for slice timing, realigned to the middle volume, co-registered to structural scans using the mean functional image, spatially normalized to a standard echo-planar image (EPI) template based on the Montreal Neurological Institute reference brain template (MNI152, Asia brain), and spatially smoothed with an isotropic 8 mm full width at half-maximum Gaussian kernel. For anatomical reference, the statistical maps computed were overlaid to the 3D T1-weighted scans. First, face-selective regions were localized in each subject by BOLD signals. By comparing the "faces" condition with the "non-face" condition in the localizer experiment at the first-level analysis, we assessed this face-selective ROI for every participant by building a 5-mm radius sphere surrounding the coordinate of the maximum activation in rFFA at a threshold $p < 0.05$, the family-wise error (FWE) correcting for multiple testing, with MARSBAR followed by visual confirmation of their anatomical location. For participants who lacked a face-selective area with this criterion, we explored liberal uncorrected significance thresholds (as liberal as $p < 0.05$), to avoid missing effects that might be apparent at less stringent uncorrected thresholds. Because faces are processed more dominantly in the right hemisphere and because the ROIs in this study were localized more consistently in the right hemisphere, we chose to restrict our ROI-based analyses to regions in the right hemisphere (unless noted otherwise). Second, the above-defined ROI were tested for FIE using the contrast (upright different face > inverted different face) and adaptation to facial identity using the contrast (up different faces > up identical faces) by BOLD signals at the first-level analysis. Third, the time course of percent signal change from baseline fixation was

extracted from the ROI for each condition in each individual with MARSBAR, and was plotted for each condition in controls (±SD) and family members by an in-house Matlab programme. Fourth, in order to directly compare each CP or non-CP with the control subjects, the percent signal change in the ROI was computed for each condition. Three data points around the peak of the hemodynamic response defined individually were averaged to estimate the percent signal change. Fifth, the percent signal change was used to compute both the FIE (upright different face – inverted different face) and adaptation scores (upright different face – upright identical face) for each subject, allowing a comparison between both the CP and non-CP family members with the control group in each experiment by means of Z-scores and the modified t-test score (Crawford and Garthwaite, 2002; Crawford et al., 2010).

Results were visualized using xjView toolbox (http://www.alivelearn.net/xjview) in addition to built-in visualization in SPM12.

Quantification and Statistical Analysis

Where applicable, statistical parameters including sample size, precision measures (standard error or standard deviation) and statistical significance are reported in the figures and corresponding legends. *P*-values of less than 0.05 are significant.

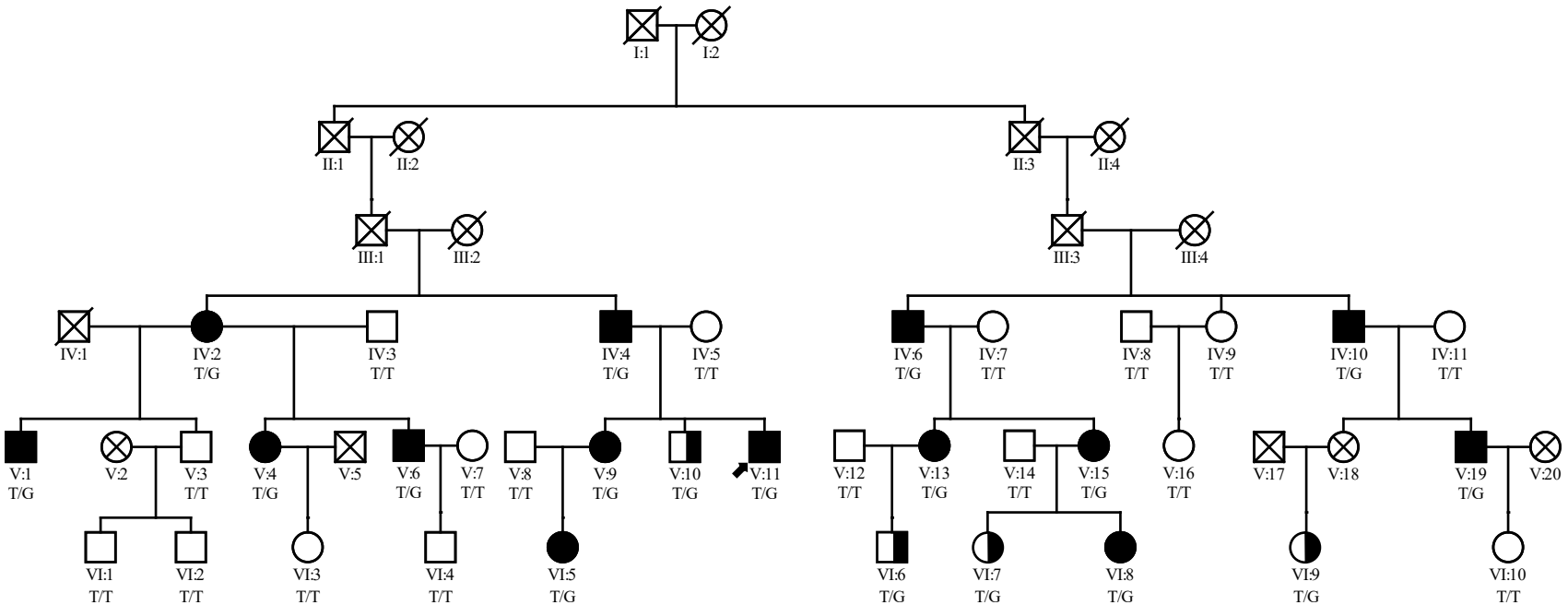
Data and Software Availability

Further information and requests for raw data, analysis details, and DNA samples may be directed to and fulfilled in compliance with IRB protocol requirements by Yi Rao (yrao@pku.edu.cn). Whole exome/genome sequencing data are not deposited in a public repository because the genetic sharing plans of the IRB approved consent forms used in this study do not permit the deposition of such data into public or controlled-

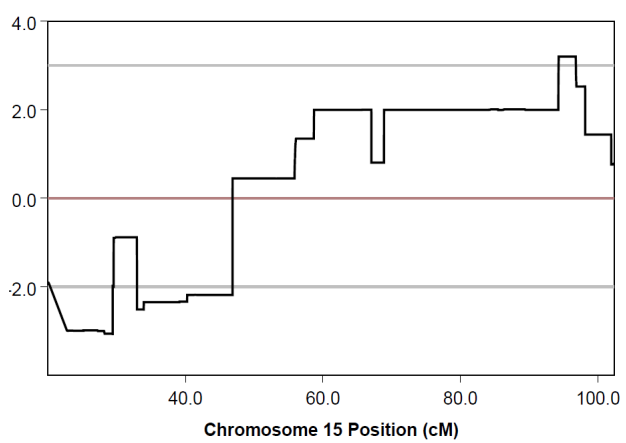
access databases.

Figure 1

A


I:1
I:2
II:1
II:2
II:3
II:4
III:1
III:2
III:3
III:4
IV:1
IV:2 T/G
IV:3 T/T
IV:4 T/G
IV:5 T/T
IV:6 T/G
IV:7 T/T
IV:8 T/T
IV:9 T/T
IV:10 T/G
IV:11 T/T
V:1 T/G
V:2
V:3 T/T
V:4 T/G
V:5
V:6 T/G
V:7 T/T
V:8 T/T
V:9 T/G
V:10 T/G
V:11 T/G
V:12 T/T
V:13 T/G
V:14 T/T
V:15 T/G
V:16 T/T
V:17
V:18
V:19 T/G
V:20
VI:1 T/T
VI:2 T/T
VI:3 T/T
VI:4 T/T
VI:5 T/G
VI:6 T/G
VI:7 T/G
VI:8 T/G
VI:9 T/G
VI:10 T/T

B


4.0
2.0
0.0
-2.0
40.0
60.0
80.0
100.0
Chromosome 15 Position (cM)

C

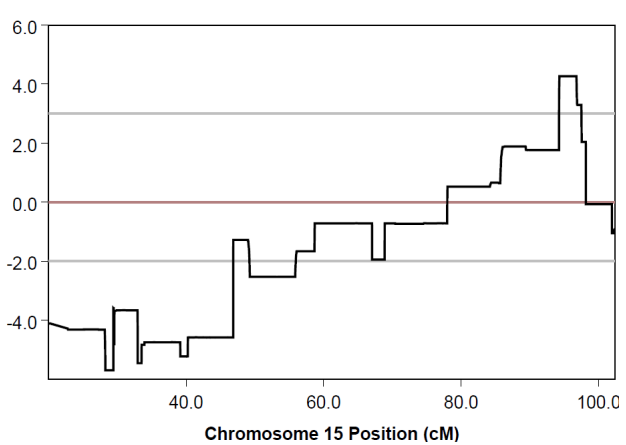

6.0
4.0
2.0
0.0
-2.0
-4.0
40.0
60.0
80.0
100.0
Chromosome 15 Position (cM)

Figure 2

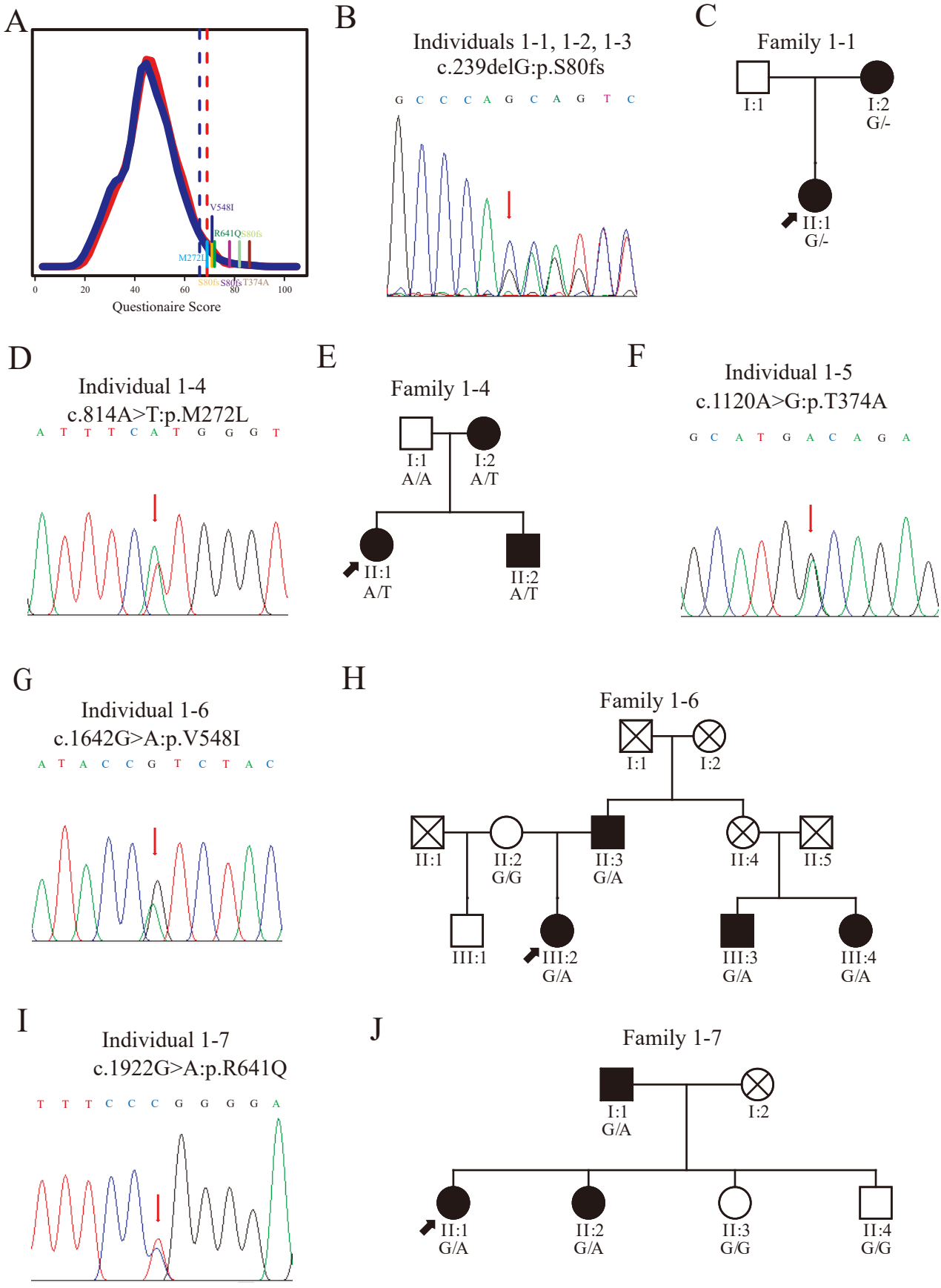
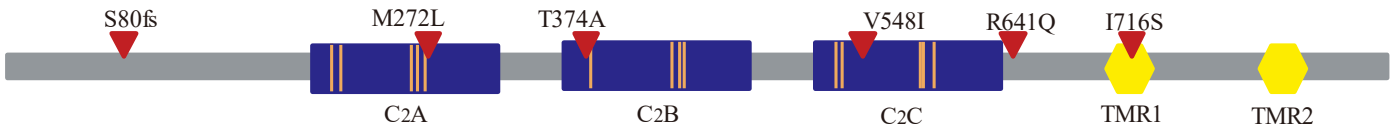

A
V548I
R641Q
S80fs
M272L
S80fs
S80fs
T374A
0
20
40
60
80
100
Questionaire Score
B
Individuals 1-1, 1-2, 1-3
c.239delG:p.S80fs
G C C C A G C A G T C
C
Family 1-1
I:1
I:2
G/-
II:1
G/-
D
Individual 1-4
c.814A>T:p.M272L
A T T T C A T G G G T
E
Family 1-4
I:1
A/A
I:2
A/T
II:1
A/T
II:2
A/T
F
Individual 1-5
c.1120A>G:p.T374A
G C A T G A C A G A
G
Individual 1-6
c.1642G>A:p.V548I
A T A C C G T C T A C
H
Family 1-6
I:1
I:2
II:1
II:2
G/G
II:3
G/A
II:4
II:5
III:1
III:2
G/A
III:3
G/A
III:4
G/A
I
Individual 1-7
c.1922G>A:p.R641Q
T T T C C C G G G G A
J
Family 1-7
I:1
G/A
I:2
II:1
G/A
II:2
G/A
II:3
G/G
II:4
G/G

Figure 3

A



B

	S80fs	M272L	T374A	V548I	R641Q	I716S
Human	TSVPSSLST	TTSDFMGSAFV	SGGSMTEMFVQ	LQTHTVYKNLN	RTFTPREKRFV	LALLLIFVYNF
Chimp	TSVPSSLST	TTSDFMGSAFV	SGGSMTEMFVQ	LQTHTVYKNLN	RTFTPREKRFV	LALLLIFVYNF
Gorilla	TSVPSSLST	TTSDFMGSAFV	SGGSMTEMFVQ	LQTHTVYKNLN	RTFTPREKRFV	LALLLIFVYNF
Orangutan	TSVPSSLST	TTSDFMGSAFV	SGGSMTEMFVQ	LQTHTVYKNLN	RTFTPREKRFV	LALLLIFIYNF
Gibbon	TSVPSSLST	TKSDFMGSAFV	SGGSMTEMFVQ	LQTHTVYKNLN	RTFTPREKRFV	LALLLIFVYNF
Rhesus	TSVPSSLST	TTSDFMGSAFV	SGGSMTEMFVQ	LQTHTVYKNLN	RTFTPREKRFV	LALLLMFVYNF
Baboon	TSVPSSLST	TTSDFMGSAFV	SGGSMTEMFVQ	LQTHTVYKNLN	RTFTPREKRFV	LALLLIFVYNF
Marmoset	TSVPSSLST	TTSDFMGSAFV	SGGSMTELFVQ	LQTHTVYKNLN	RTFTPREKRFV	LALLLIFIYNF
Mouse	ISVPNSLST	TKSDFMGSAFV	SGGNMTEMFVQ	LQTHTIYKNLN	RTFTPREKRFV	LALLLLFLYNF
Rat	ISVPNSLST	TKSDFMGSAFV	SGGNMTEMFVQ	LQTHTIYKNLN	RTFSPREKRFV	LAMLLLFLYNF
Cow	TSVPSSLST	TISDFMGSAFV	SGGSVAEMFVQ	LQTHTIYKNLN	RTFTPREKRFV	LALLLLFVYNS
Dog	TSVPSSLST	TTSDFMGSAFV	SGGSMTEMFVQ	LQTHTIYKNLN	RTFTPREKRFV	LALLLLFIYNF
Chicken	SGYPSGAPR	TSSDFMGSAFV	PFGGLAEVFIL	LQTHTVYKNLN	RTFKPREKRFT	LALLVLFAYNF

Figure 4

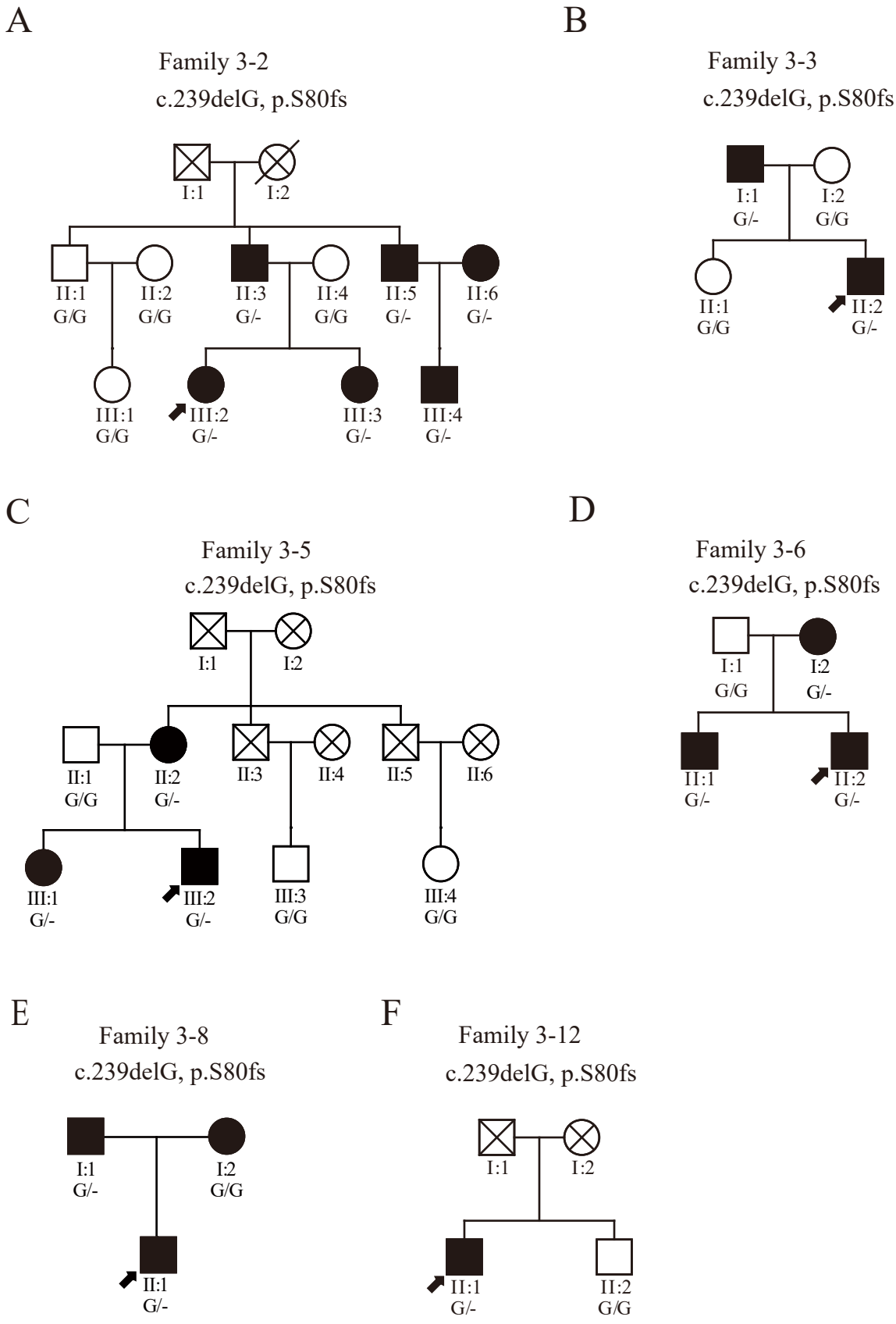

A
Family 3-2
c.239delG, p.S80fs
I:1
I:2
II:1
G/G
II:2
G/G
II:3
G/-
II:4
G/G
II:5
G/-
II:6
G/-
III:1
G/G
III:2
G/-
III:3
G/-
III:4
G/-
B
Family 3-3
c.239delG, p.S80fs
I:1
G/-
I:2
G/G
II:1
G/G
II:2
G/-
C
Family 3-5
c.239delG, p.S80fs
I:1
I:2
II:1
G/G
II:2
G/-
II:3
II:4
II:5
II:6
III:1
G/-
III:2
G/-
III:3
G/G
III:4
G/G
D
Family 3-6
c.239delG, p.S80fs
I:1
G/G
I:2
G/-
II:1
G/-
II:2
G/-
E
Family 3-8
c.239delG, p.S80fs
I:1
G/-
I:2
G/G
II:1
G/-
F
Family 3-12
c.239delG, p.S80fs
I:1
I:2
II:1
G/-
II:2
G/G

Figure 5

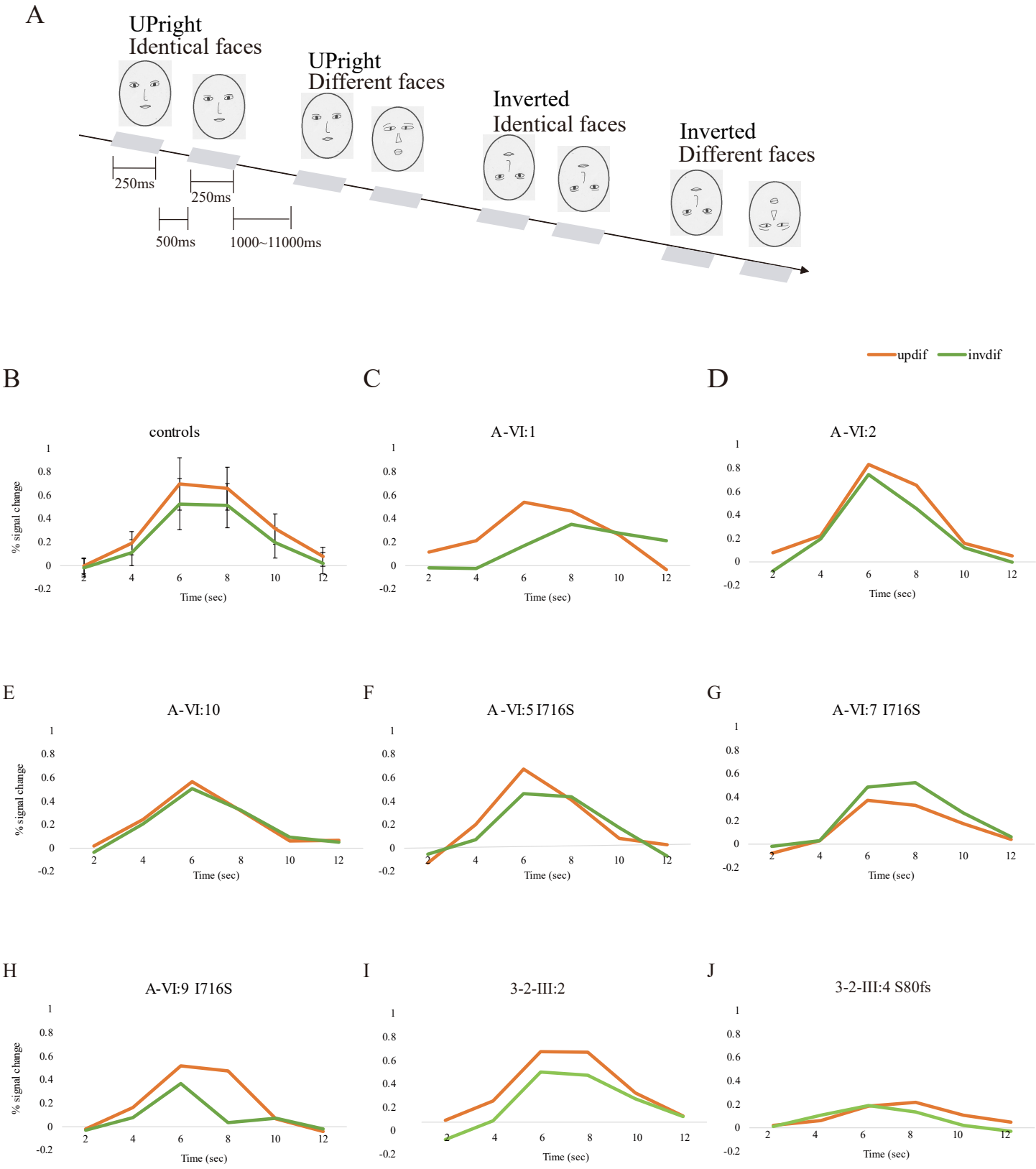

A
UPright
Identical faces
UPright
Different faces
Inverted
Identical faces
Inverted
Different faces
250ms
250ms
500ms
1000~11000ms
updif
invdif
B
controls
C
A-VI:1
D
A-VI:2
E
A-VI:10
F
A-VI:5 I716S
G
A-VI:7 I716S
H
A-VI:9 I716S
I
3-2-III:2
J
3-2-III:4 S80fs
% signal change
Time (sec)

Figure 6

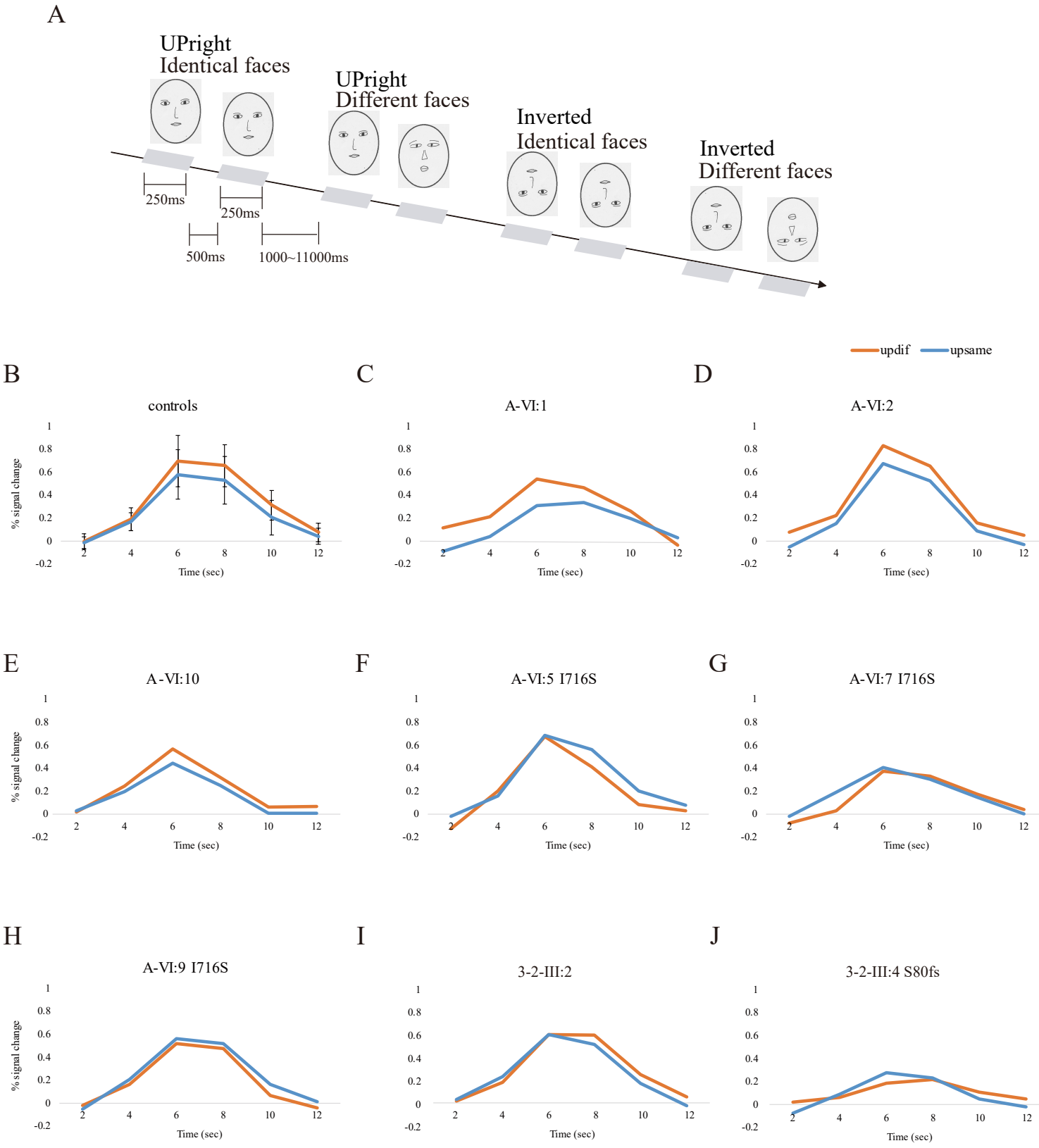
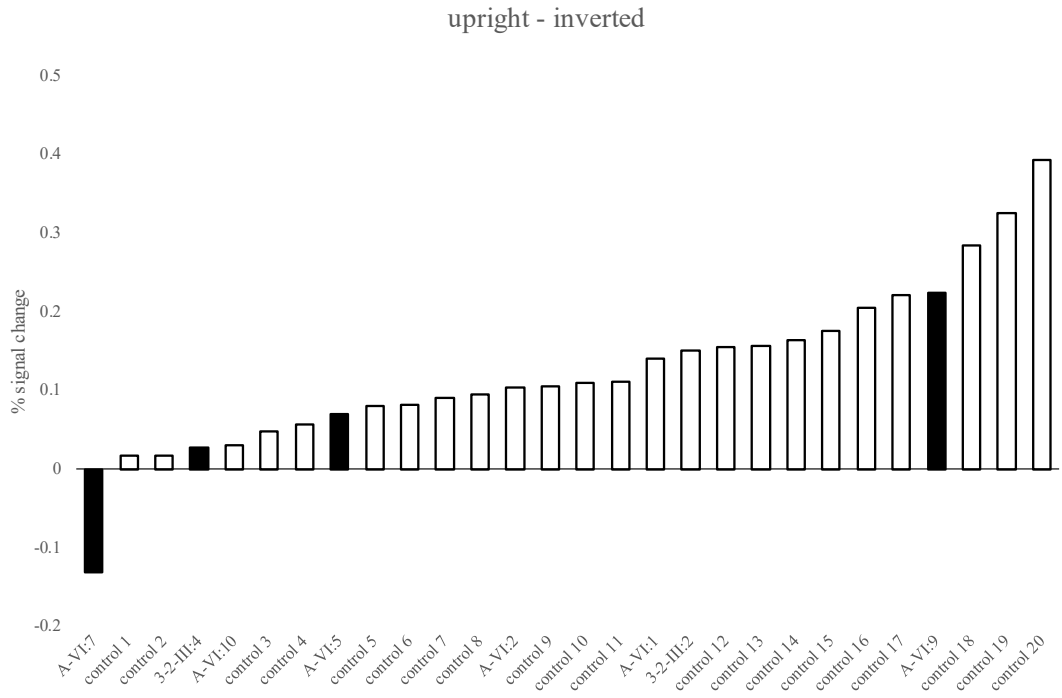

A
UPright
Identical faces
UPright
Different faces
Inverted
Identical faces
Inverted
Different faces
250ms
250ms
500ms
1000~11000ms
updif
upsame
B
controls
C
A-VI:1
D
A-VI:2
E
A-VI:10
F
A-VI:5 I716S
G
A-VI:7 I716S
H
A-VI:9 I716S
I
3-2-III:2
J
3-2-III:4 S80fs
% signal change
Time (sec)

Figure 7

A


upright - inverted
% signal change
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
A-VI:7
control 1
control 2
3-2-III:4
A-VI:10
control 3
control 4
A-VI:5
control 5
control 6
control 7
control 8
A-VI:2
control 9
control 10
control 11
A-VI:1
3-2-III:2
control 12
control 13
control 14
control 15
control 16
control 17
A-VI:9
control 18
control 19
control 20

B

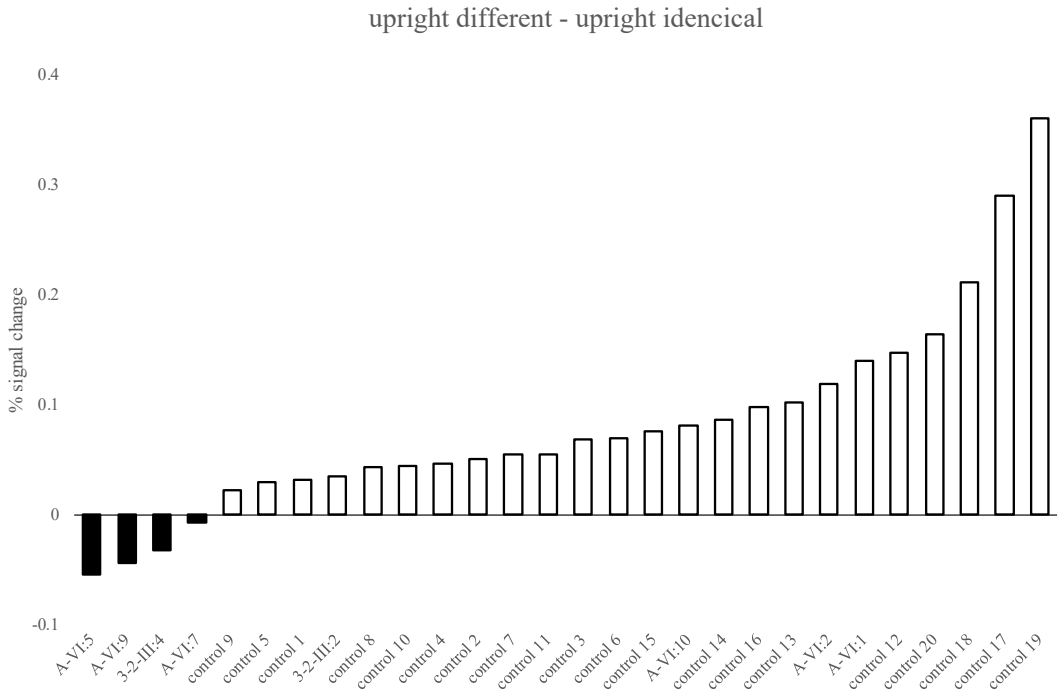

upright different - upright idencical
% signal change
0.4
0.3
0.2
0.1
0
-0.1
A-VI:5
A-VI:9
3-2-III:4
A-VI:7
control 9
control 5
control 1
3-2-III:2
control 8
control 10
control 4
control 2
control 7
control 11
control 3
control 6
control 15
A-VI:10
control 14
control 16
control 13
A-VI:2
A-VI:1
control 12
control 20
control 18
control 17
control 19

Table 1. Diagnosis of CP in Family A

Family Member	Key Manifestations									Facultative symptoms		Memory		Discrimination	
	Lasting and irritating subjective uncertainty of face recognition	Face recognition deficit especially in crowded places or out-of-context encounters	False negative and false positive face recognition events	Face recognition time longer than socially accepted	Face learning time longer than socially accepted	Development of adaptive behavior	Use of explicit learning strategies for visual person recognition	First Time to Realize*	Affected first degree relatives	Images of close friends' faces	Gaze contact necessary in conversation	CFMT	CCMT	FIE	Upright
IV:2	yes			yes		yes		in childhood	yes	blur face	yes	-	-	-	-
IV:3	no			no		no		-	yes	clear face	yes	-	-	-	-
IV:4	yes			yes		yes		in childhood	yes	blur face	yes	-	-	-	-
IV:5	no			no		no		-	yes	clear face	yes	-	-	-	-
IV:6	yes			yes		yes		after contacted	yes	blur face	no	-	-	-	-
IV:7	no			no		no		-	yes	clear face	yes	-	-	-	-
IV:8	no			no		no		-	no	clear face	yes	-	-	-	-
IV:9	no			no		no		-	yes	clear face	yes	-	-	-	-
M61	yes			yes		yes		not realize by himself	yes	whole body & action	no	-	-	-	-
IV:11	no			no		no		-	yes	clear face	yes	-	-	-	-
V:1	yes			yes		yes		not realize by himself	yes	body & wearings	diliberately need	-2.06	-0.11	-1.33	-2.96
V:3	no			no		no		-	yes	clear face	yes	-0.98	0.67	-0.05	0.94
V:4	yes			yes		yes		after contacted	yes	blur face	no	-1.60	0.56	-1.41	-0.94
V:6	yes			yes		yes		after contacted	yes	local face	no	-1.61	0.35	-1.62	0.11
V:7	no			no		no		-	no	clear face	yes	-0.25	0.63	0.48	0.45
V:8	no			no		no		-	yes	clear face	yes	-0.16	1.90	-0.35	1.33
V:9	yes			yes		yes		in middle school	yes	only facial expression	yes	-2.67	-0.09	-1.62	-0.34
V:10	yes			yes		yes		at university	yes	contour	diliberately need	-0.06	0.40	2.90	1.31
V:11	yes			yes		yes		in middle school	yes	blur face	diliberately need	-2.09	0.37	-2.56	-0.71
V:12	no			no		no		-	yes	clear face	yes	-1.11	-0.91	0.06	0.25
V:13	yes			yes		yes		not realize by herself	yes	whole body	no	-3.27	0.06	-0.51	-1.73

V:14	no	no	no	-	yes	clear face	yes	1.13	0.46	-0.40	0.26
V:15	yes	yes	yes	from the age of 30	yes	blur face	yes	-2.03	-0.74	-0.65	-1.55
V:16	no	no	no	-	no	clear face	yes	-0.52	-0.82	-0.27	-0.70
V:19	yes	yes	yes	not realize by himself	yes	blur face	yes	-2.79	-0.55	-1.15	-1.52
VI:1	no	no	no	-	no	clear face	yes	-0.13	0.72	0.30	-0.14
VI:2	no	no	no	-	no	clear face	yes	0.84	0.17	-0.58	1.25
VI:3	no	no	no	-	yes	clear face	yes	0.48	-0.66	1.16	0.35
VI:4	no	no	no	-	yes	clear face	yes	0.71	-0.89	0.24	-0.01
VI:5	yes	yes	yes	in middle school	yes	what she images is not the same as she sees, "a kind of feeling"	no	-1.37	-0.39	-1.81	-0.50
VI:6	yes	yes	yes	at university	yes	face contour	yes	-0.56	0.41	-0.36	-0.48
VI:7	yes	-	yes	not realize by herself	yes	blur face	diliberately need	1.26	1.31	1.48	0.95
VI:8	yes	yes	yes	after contacted	yes	no face	yes	-2.05	-0.05	0.05	-1.40
VI:9	yes	yes	yes	in middle school	unknown	what she images is not the same as she sees, "a kind of feeling"	diliberately need	1.01	2.52	-0.71	0.67
VI:10	no	no	no	-	yes	clear face	yes	-0.53	-0.86	0.07	0.89

Members with dark grey background indicate CP candidates in linkage 1. Members with light grey background were added as CP candidates in linkage 2. Face processing z scores for members younger than 60 years of age in Family A are coded as described in the method such that negative values indicate performances poorer than the mean. Red indicates below -2 SDs. Blue indicates below -1.5 SDs. Whether there is an obvious key manifestation for CP is expressed by yes or no. More detailed diagnostic information can be obtained by contacting the corresponding author.

Table 2. *MCTP2* Mutations Found in Sporadic CP Candidates Screened by a 20-Item Questionnaire

Patients	Chr15 Position*	RefSNP ID#	Type of Variants	Amino Acid Change &	AAF_EAS				Non-Synonymous Variants Annotation				
					1000G	ExAC	gnomAD		SIFT	Polyphen2 HDIV	Polyphen2 HVAR	LRT	Mutation Taster
							exome	exome					
1-1 1-2 1-3	94841733	rs574533536	FD	exon1:c.239delG:p.S80fs	0.005	0.0031	0.0036	0.0045	.	.	.	.	.
1-4	94883461	rs773393905	NS	exon5:c.A814T:p.M272L	.	0.0001	.	.	T	P	P	D	D
1-5	94899480	rs764700909	NS	exon8:c.A1120G:p.T374A	.	0.0002	0.0002	.	T	B	B	N	D
1-6	94927310	rs182595113	NS	exon12:c.G1642A:p.V548I	0.003	0.0008	0.0013	0.0006	T	B	B	N	N
1-7	94943181	rs749217960	NS	exon15:c.G1922A:p.R641Q	.	.	0.0000544	.	D	P	B	D	D

NS: nonsynonymous single nucleotide variant; FD: frameshift_deletion; AAF_EAS: alternative allele frequency in EAS (East Asian); SIFT: predicts whether an amino acid substitution affects protein function as D (Deleterious) or T (tolerated) by SIFT algorithm; Polyphen2_HDIV, Polyphen2_HVAR: predicts possible impact of an amino acid substitution on the structure and function of a human protein using straightforward physical and comparative considerations as D (Probably damaging), P (possibly damaging) or B (benign) by PolyPhen 2 (Polymorphism Phenotyping v2); LRT: uses comparative genomics to predict whether variants disrupt highly conserved amino acids as D (deleterious), N (neutral) or U (unknown) by LRT (Likelihood Ratio Test); Mutation Taster: predicts an alteration as one of four possible types: D (disease causing - i.e. probably deleterious), A (disease causing automatic - i.e. known to be deleterious), N (polymorphism - i.e. probably harmless), P (polymorphism automatic - i.e. known to be harmless). These abbreviations also apply to the following tables.
* Start Position Based on the MCTP2 RefSeq (hg19, NM_018349); # dbSNP build 150, & RefGene: NM_018349.

Table 3. Gene Based Association Analysis for Rare Variants in *MCTP2*

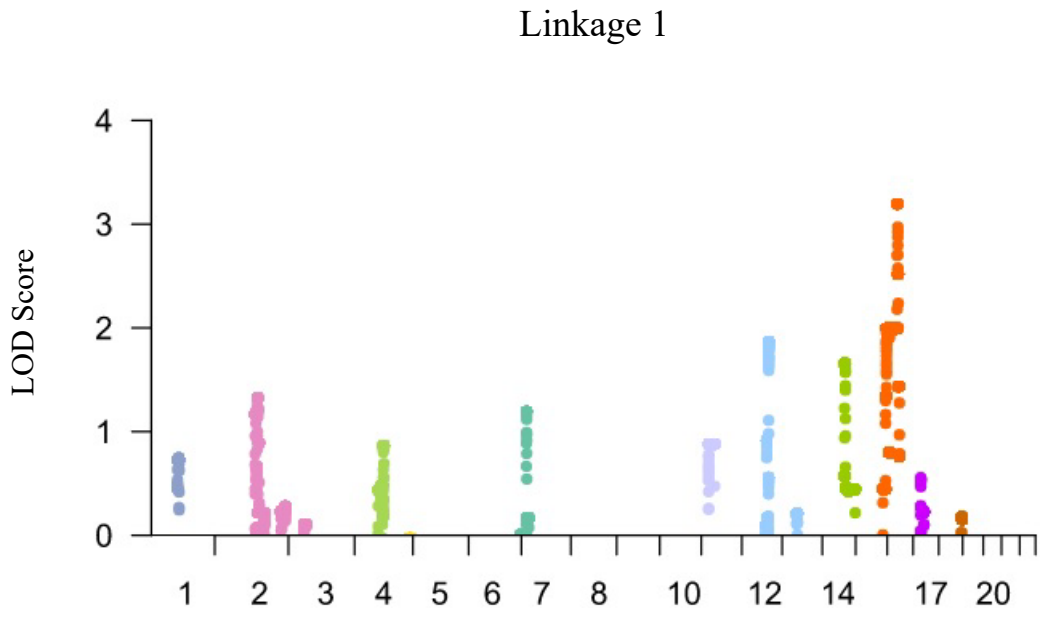
		All	Male	Female
Nsstrict	Burden	0.1286	**0.0113**	0.8746
	SKAT	0.9560	0.7612	0.9009
	SKAT-O	0.2249	**0.0212**	1.0000
Nsbroad	Burden	0.0603	**0.0029**	0.9518
	SKAT	0.9914	0.7611	0.9276
	SKAT-O	0.1141	**0.0061**	1.0000
Nsall	Burden	**0.0266**	**0.0011**	0.9229
	SKAT	0.9368	0.7997	0.9744
	SKAT-O	0.0499	**0.0024**	1.0000

Table 4. Differences in Daily Face Recognition between Target Group and Controls Regarding the p.S80fs Mutation.

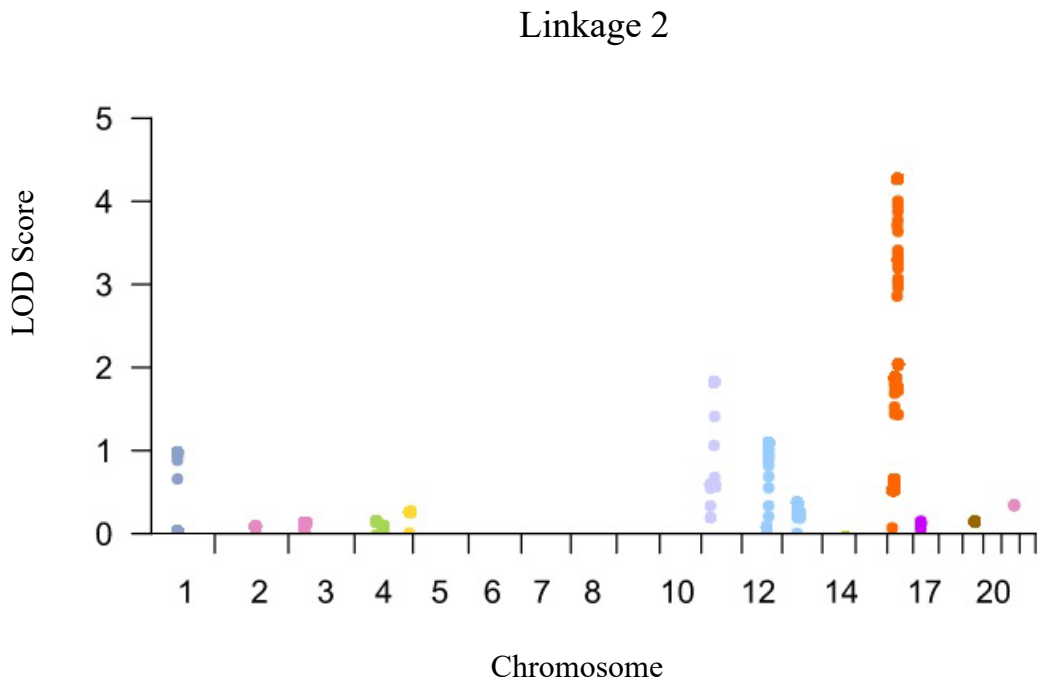
Manifestations of poor face recognition	Target group N=14 with the p.S80fs mutation (6 males/8 females)	Controls N=19 without the p.S80fs mutation (5 males/14 females)	Fisher's exact test, p value, two-tailed
Lasting and irritating subjective uncertainty of face recognition	9/14	1/19	<0.001
Face recognition deficit especially in crowded places or out-of-context encounters	12/14	4/19	<0.001
False negative and false positive face recognition events	10/14	4/19	0.006
Face recognition time longer than socially accepted	11/14	3/19	<0.001
Face learning time longer than socially accepted	11/14	3/19	<0.001
Development of adaptive behavior	12/14	5/19	0.01
Use of explicit learning strategies for visual person recognition	13/14	4/19	<0.001
Self-evaluation of impaired visual recognition of faces	5/14	1/19	0.062

Supplementray Figure 1

A

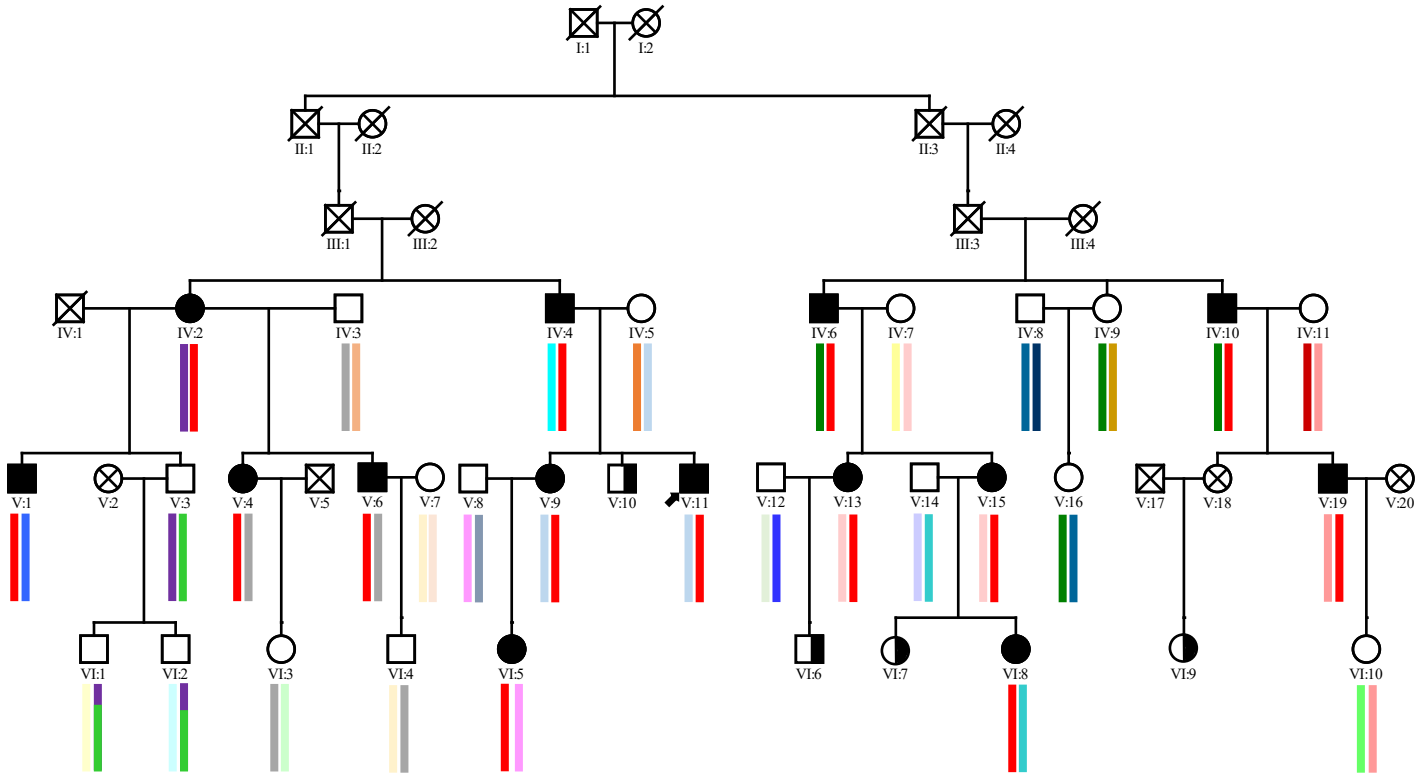


B

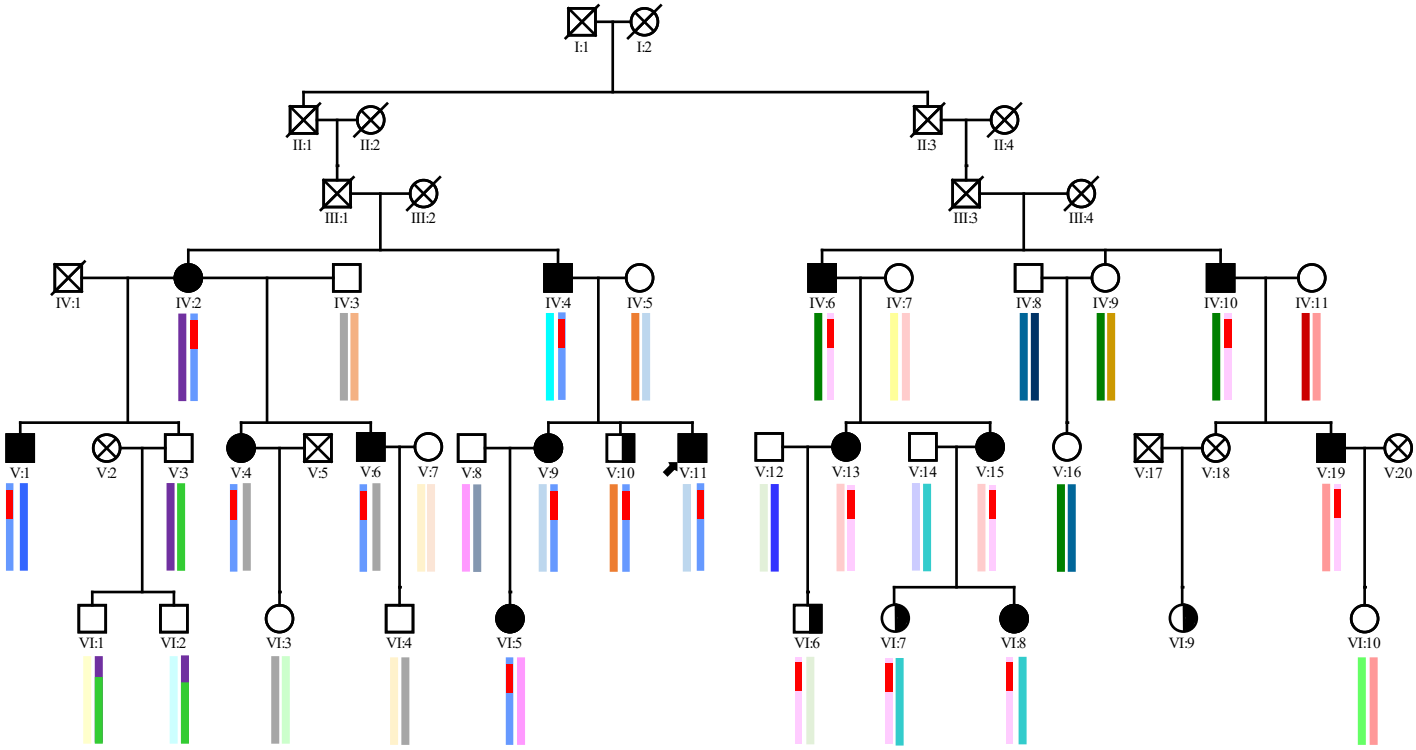


Supplementary Figure 2

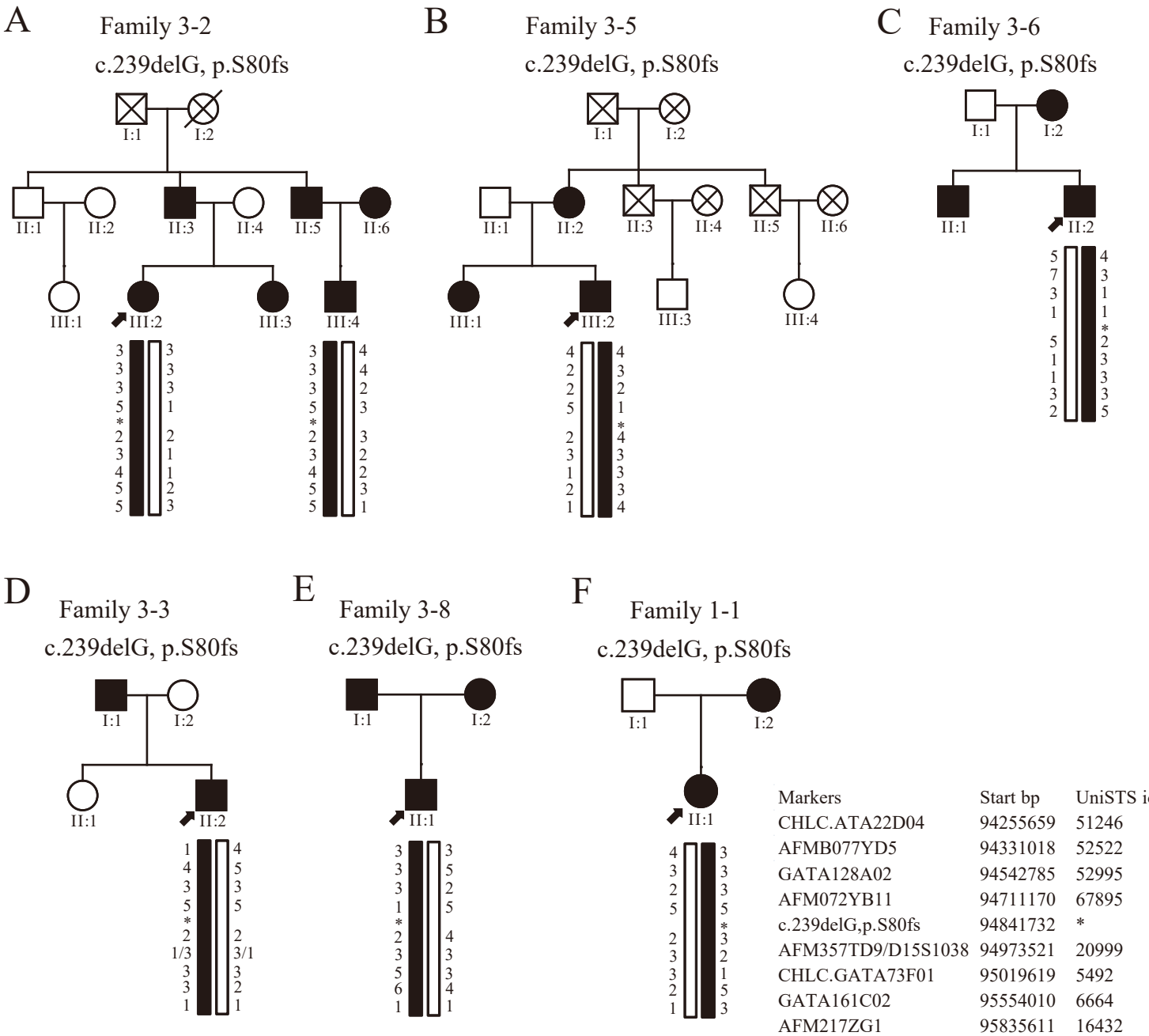
A



B

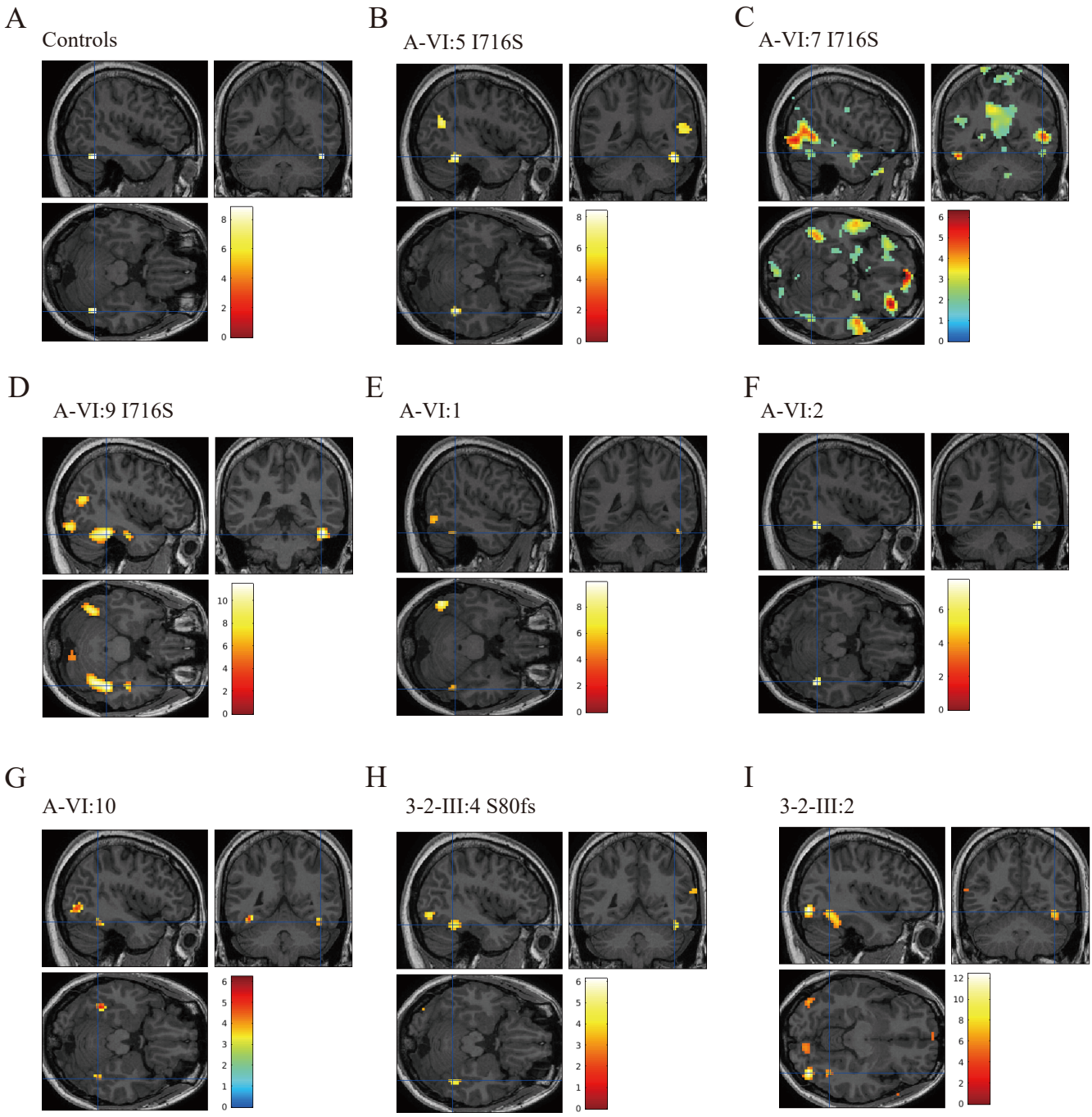


Supplementary Figure 3



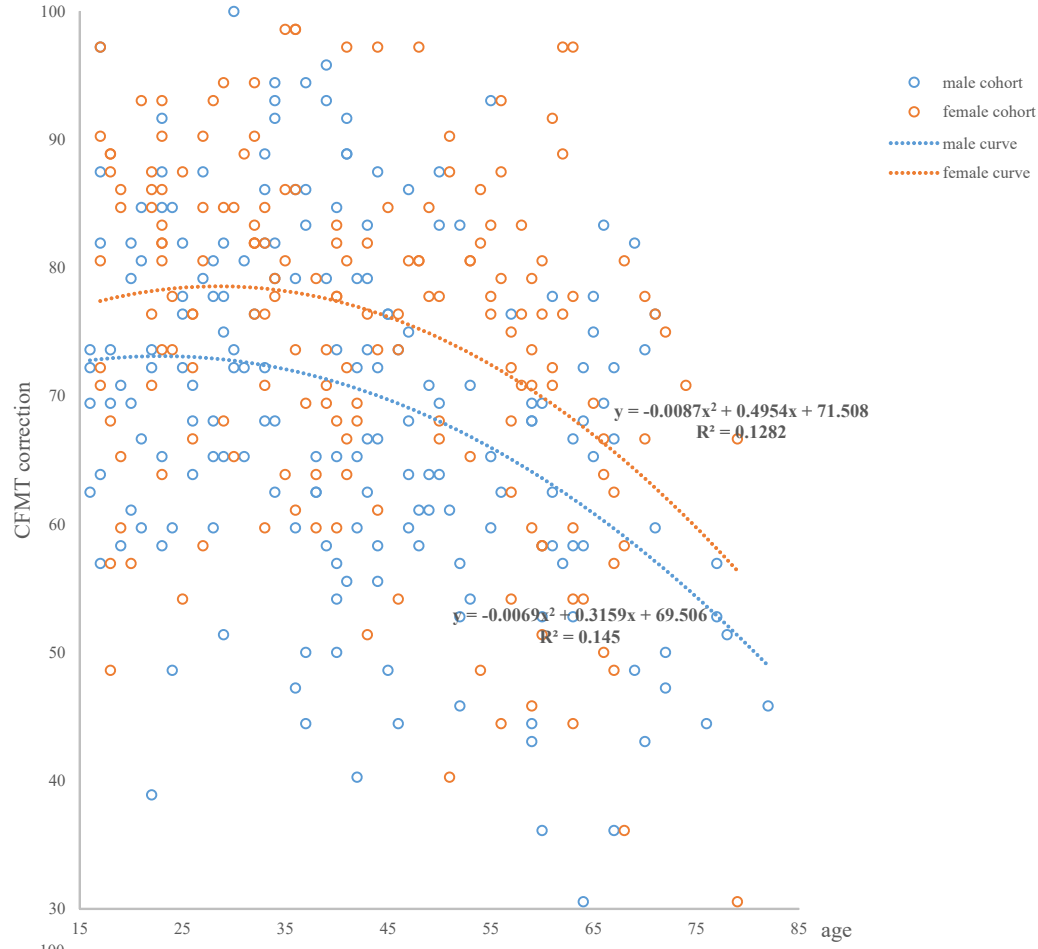
Markers	Start bp	UniSTS id
CHLC.ATA22D04	94255659	51246
AFMB077YD5	94331018	52522
GATA128A02	94542785	52995
AFM072YB11	94711170	67895
c.239delG,p.S80fs	94841732	*
AFM357TD9/D15S1038	94973521	20999
CHLC.GATA73F01	95019619	5492
GATA161C02	95554010	6664
AFM217ZG1	95835611	16432
AFM309VG9	96210769	10812

Supplementary Figure 4



Supplementary Figure 5

A



B

Supplementary Figure 6

A

B

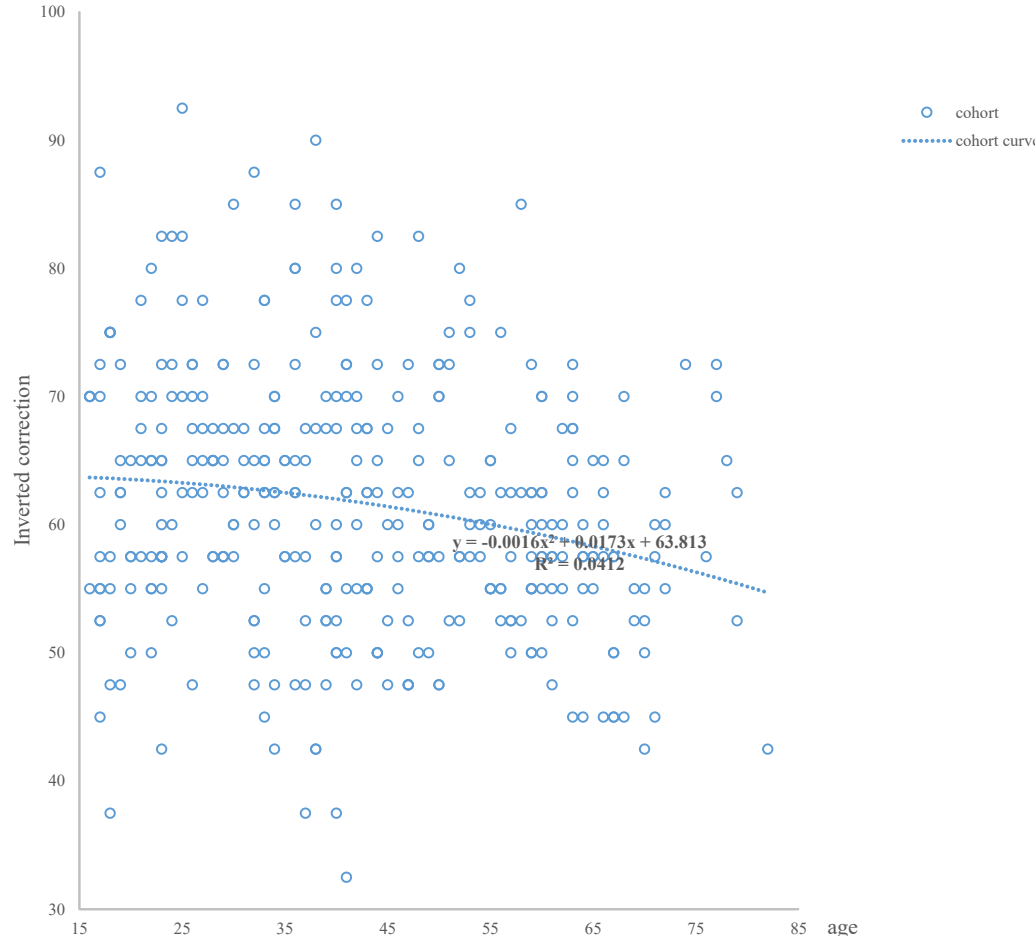


Table S1. Summary of Phenotypic Analysis and Genetic Analysis of Each Member of Family A

Family Member	obvious difficulty in face recognition in daily life by interview	impaired on at least one of the face behavioral tests	Linkage 1	Linkage 2	Array	WES	WGS	fMRI
			cases defined by interview and tests	cases defined by interview				
IV:2	V	-	CP	CP	V	-	-	-
IV:3		-	normal	normal	V	-	-	-
IV:4	V	-	CP	CP	V	-	-	-
IV:5		-	normal	normal	V	-	-	-
IV:6	V	-	CP	CP	V	-	-	-
IV:7		-	normal	normal	V	-	-	-
IV:8		-	normal	normal	V	-	-	-
IV:9		-	normal	normal	V	-	-	-
M61	V	-	CP	CP	V	-	-	-
IV:11		-	normal	normal	V	-	-	-
V:1		V	CP	CP	V	-	-	-
V:3			normal	normal	V	-	-	-
V:4	V	V	CP	CP	V	-	-	-
V:6	V	V	CP	CP	V	V	-	-
V:7			normal	normal	V	-	-	-
V:8			normal	normal	V	-	-	-
V:9	V	V	CP	CP	V	V	-	-
V:10	V		unclear	CP	V	-	-	-
V:11	V	V	CP	CP	V	V	-	-
V:12			normal	normal	V	-	-	-
V:13	V	V	CP	CP	V	-	-	-
V:14			normal	normal	V	-	-	-
V:15	V	V	CP	CP	V	-	-	-
V:16			normal	normal	V	-	-	-
V:19	V	V	CP	CP	V	V	V	-
VI:1			normal	normal	V	-	-	V
VI:2			normal	normal	V	-	-	V
VI:3			normal	normal	V	-	-	-
VI:4			normal	normal	V	-	-	-
VI:5	V	V	CP	CP	V	-	V	V
VI:6	V		unclear	CP	V	-	-	-
VI:7	V		unclear	CP	V	-	-	V
VI:8	V	V	CP	CP	V	-	-	-
VI:9	V		-	-	-	-	-	V
VI:10			normal	normal	V	-	-	V

Table S2. Summary of Somputer Computer Simulations of Linkage Analysis

	independent linked regions		observed maximum LOD score	empirical p value
Linkage 1	chr2	18150570-19775673	1.325	0.932
	chr6	102275330-103291012	1.192	0.942
	chr12	21989321-22891967	1.867	0.591
	chr14	26341293-29161732	1.657	0.659
	chr15	94266567-96809782	3.196	**0.017**
Linkage 2	chr10	110316481-112555472	1.826	0.528
	chr12	22015114-22325528	1.095	0.854
	chr15	94245722-97520230	4.270	**0.009**

Table S3. Candidate Genes Annotated by the NCBI Database (NC_000015.9) in Regions of Linkage Analyses with LOD Scores over 3

Start	Stop	Gene symbol	Strand	NCBI Gene ID	Name
94399789	94421585	LINC02207	plus	101927153	long intergenic non-protein coding RNA 2207
94439288	94443828	LINC01579	minus	283682	long intergenic non-protein coding RNA 1579
94443930	94527379	LINC01580	plus	101927129	long intergenic non-protein coding RNA 1580
94448632	94651167	LINC01581	minus	101927112	long intergenic non-protein coding RNA 1581
94711497	94712177	H3P40	plus	115482723	H3 histone pseudogene 40
94774790	95027181	MCTP2	plus	55784	multiple C2 and transmembrane domain containing 2
94816725	94817231	RPL26P5	plus	654386	ribosomal protein L26 pseudogene 5
95398592	95400295	LOC440311	plus	440311	NOP53 ribosome biogenesis factor pseudogene
95822517	95870329	LINC01197	minus	400456	long intergenic non-protein coding RNA 1197
95976322	96051076	LINC00924	plus	145820	long intergenic non-protein coding RNA 924
96289033	96289220	RNU2-3P	plus	26854	RNA, U2 small nuclear 3, pseudogene
96625281	96671070	LOC105369212	minus	105369212	uncharacterized LOC105369212
96670589	96870590	NR2F2-AS1	minus	644192	NR2F2 antisense RNA 1
96869275	96883492	NR2F2	plus	7026	nuclear receptor subfamily 2 group F member 2
96876490	96876536	MIR1469	plus	100302258	microRNA 1469
96931442	96932281	TUBAP12	minus	100289303	tubulin alpha pseudogene 12
96984599	96985595	PGAM1P12	minus	390643	phosphoglycerate mutase 1 pseudogene 12
96991356	96991760	RPL31P55	plus	100128250	ribosomal protein L31 pseudogene 55
97029812	97030121	RN7SKP254	plus	106481820	RN7SK pseudogene 254
97300082	97301822	FAM149B1P1	plus	388181	family with sequence similarity 149 member B1 pseudogene 1
97315235	97326567	SPATA8-AS1	minus	100652749	SPATA8 antisense RNA 1 (head to head)
97326665	97328845	SPATA8	plus	145946	spermatogenesis associated 8
97385299	97385614	RN7SKP181	plus	106479176	RN7SK pseudogene 181

Table S4. the Summary Analysis of the Sequencing Data

Terms	Exome Sequence of Family A	Whole Genome Sequence of Family A	Candidate Region Sequence of the second cohort of 1928 medical students	Candidate Region Sequence of the third cohort of 1757 medical students
Sample number	4	2	1928	1757
Reads_mapped (%)	99.86	99.86	99.91	98.99
Target_Efficiency (%)	83.48	-	72.48	82.57
Mean Depth	83.55	42.66	10255	4501.25
>=10Coverage (%)	94.6	95.55	95.44	92.24
>=20Coverage (%)	89.2	90.05	94.40	88.58
>=40Coverage (%)	77.1	56.7	93.13	83.84
>=100Coverage (%)	33.3	-	90.57	76.29

Table S5. PCR Primer Sequences for Each Exon and Mutation Site of *MCTP2*

Primer ID		Sequence (5' to 3')
Exon1/c.239delG:p.S80fs	L	TGGTGCCAACCTTATCATCTCAT
	R	GAGCCCAGCCCTATTCTCACC
	seq	ATTGCAGTTTTCAGTAGAGG
Exon2	L/seq	TATCCAGTGACTTAGGCGGCA
	R	ATTGTCTCTGTCCAGTGCTCTCC
Exon3	L/seq	CACTTCTTCCTCCTGTTGCCTCT
	R	TACACCCACCATAACAGACCCAG
Exon4	L/seq	GGTGGGCTTCTCCCTAAGACTAC
	R	CAGTTTTAACCCAGATGTTGTAGCA
Exon5/c.814A>T:p.M272L	L/seq	CATAGTAGGCAAAATGGGCAAGA
	R	TTAACACGATCACTCCCATGTCA
Exon6	L/seq	CCAACTGTTGAATTAAACAATGCC
	R	ATATTCTCCTCACCCACCCCAC
Exon7	L/seq	ATGAGAGGGTTTTGCTTGGAGA
	R	TAAAATGCCACCACTAAAACAAGA
Exon8/c.1120A>G:p.T374A	L/seq	AAGGAGTCCCACCAAAGAGCG
	R	TGGCGTGAAGACTAGCAAGCA
Exon9	L/seq	CTGGGGGGCAGCTAAGATAGTAG
	R	GTAGTGCCCTCAGCCCTAACAGA
Exon10	L/seq	AGGCTGGGGGTGAGCATTAT
	R	TTTCAAAATAGCAAAACCAACTCTT
Exon11	L/seq	TATCTATGCCATCACCAGGGG
	R	TTTTGCTTTGCCAGGAGATTG
Exon12/c.1642G>A:p.V548I	L/seq	AAAGCAGTCTTGCCTTCCCTC
	R	TATGGGATGGGTCGCTCAGA
Exon13	L/seq	TGGATGTCTGAATCTTCTATGATGG
	R	AGTAACATCAGGGATATGAGAACTCG
Exon14	L/seq	ATCCACACACCAGTTTCAGCC
	R	GCACTTGTTCAAACCACCAGACT
Exon15/c.1922G>A:p.R641Q	L	GTCTGGTGGTTTGAACAAGTGCT
	R/seq	GTGAAATGCCAACACATCCCA
Exon16	L/seq	GCCCAAGGTGAACTCCAATCC
	R	ATTATGGGCGGTCACTGGCAC
Exon17/c.T2147G:p.I716S	L/seq	ATTGGGCAGCTAAGGGCATAC
	R	CCACAGAGCCAGCAAAGGTT
Exon18	L/seq	CGTAGCAGTTGGTTCATTTGTCAG
	R	AGGGGAAGGCTAATCTTGGCT
Exon19	L	GGGGAAGTGAAGTGCTTGGTT
	R/seq	TCTTGCCTCCCTTTGTCGC
Exon20	L	GTGGTGTCTCTTGCCTGTAGTCC
	R/seq	CCTCTCTGTTCAGGCACTGGTTA
Exon21	L	CCCGAGTCAGGCGAGTTTT
	R/seq	GCATCAGTGCCCAGTAGCCA
Exon22	L/seq	CGAAGGTAGAAACCCACGCA
	R	GGGCTGACCCTCTACGCAA

Note: L, the left primer; R, the right primer; seq, the primer for sequencing.

Table S6. 11 Heterozygous Variants with MAF under 0.05 Shared with the Affected Individuals of Family A by WGS

Position	REF	ALT	snp150	Func.refGene	Gene.refGene	AAChange.refGene	1000g2015aug_eas	genoMAD_EAS
15:94279037	C	T	.	intergenic	LOC105370980,LINC02207	.	.	.
15:94300722	G	A	rs77009857	intergenic	LOC105370980,LINC02207	.	0.0099	0.0135
15:94377882	A	G	rs79380132	intergenic	LOC105370980,LINC02207	.	0.005	0.0077
15:94437129	C	G	.	intergenic	LINC02207,LINC01579	.	.	.
15:94763672	T	C	rs1378593	intergenic	LINC01581,MCTP2	.	0.0288	0.0244
15:94765715	G	A	rs62017614	intergenic	LINC01581,MCTP2	.	0.0288	0.0232
15:94766118	A	G	rs62017615	intergenic	LINC01581,MCTP2	.	0.0268	0.0237
15:94880697	G	T	rs139334340	intronic	MCTP2	.	0.0218	0.0257
15:94983466	T	G	.	exonic nonsynonymous	MCTP2	NM_018349:exon17: c.T2147G:p.I716S	.	.
15:95030981	C	T	rs75162624	intergenic	MCTP2,LOC440311	.	0.0268	0.0282
15:95034599	A	T	rs1121393	intergenic	MCTP2,LOC440311	.	0.0387	0.0359

Table S7. PCR Primer Sequences for 11 Heterozygous Variants (MAF under 0.05) Shared by the Affected Individuals of Family A by WGS

Primer ID		Sequence (5' to 3')
15:94279037	L/seq	CAGACAGGCTGACACTTAGGGAA
	R	CTGAAAGCAAGTGACACCAGAGAG
15:94300722	L/seq	GTGGTTTTCTCAGTTCAGCAGTG
	R	TATTTCGGTCAGCACCTGTTCTC
15:94377882	L/seq	ATCCCAGGCATTACTCCCATT
	R	TTATTCCTCCCACAACTCGTTTC
15:94437129	L/seq	CTTCTGGCTCTGGAGACCATAAT
	R	ACACCATCCTGCTCCATCCTT
15:94763672	L/seq	GTAACTTTGGGGGGTGAAGATAA
	R	ACTTACCATGCAAACCTGTATATCC
15:94765715	L/seq	GCACGGCTCAGGAATAGGAAC
	R	TTTATGTCCCCACCTCACCCT
15:94766118	L/seq	GCACGGCTCAGGAATAGGAAC
	R	TTTATGTCCCCACCTCACCCT
15:94880697	L/seq	TGTCAAGATGCCTGGGGATTA
	R	TGAGAACATAGGAGTAGAAGCCCAT
15:94983466	L/seq	ATTGGGCAGCTAAGGGCATAC
	R	CCACAGAGCCAGCAAAGGTT
15:95030981	L/seq	AACCAGAACCAAAGGAGTGCC
	R	AATGTCTGTTAGGTGTGTTTGTTCAA
15:95034599	L/seq	AAGTGGGCAAAGAAAATGAACAG
	R	TAGTTCCATCTATGTTGCTGCTGA

Note: L, the left primer; R, the right primer; seq, the primer for sequencing.

Table S8. Most Discriminative Symptoms for the Diagnosis of Additional CP Individuals and Families

ID	Family member	Mutation	Key Manifestations									Facultative symptoms	
			Lasting and irritating subjective uncertainty of face recognition	Face recognition deficit especially in crowded places or out-of-context encounters	False negative and false positive face recognition events	Face recognition time longer than socially accepted	Face learning time longer than socially accepted	Development of adaptive behavior	Use of explicit learning strategies for visual person recognition	First time to realize*	Affected first degree relatives	Images of close friends' faces*	Gaze contact necessary in conversation
1-1	I:1	-	no			no		no		-	yes	clear face	yes
	I:2	p.S80fs	yes			yes		yes		-	yes	clear face, but a desk with things on it is more clear than a face	yes
	II:1 (1-1)	p.S80fs	yes			yes		yes		in middle school	yes	local face	not necessarily
1-2	1-2	p.S80fs	yes			yes		yes		in middle school	not clear	clear face, but a desk with things on it is more clear than a face	yes
1-3	1-3	p.S80fs	yes			yes		yes		in childhood	not clear	blur face	no
1-4	I:1	-	no			no		no		-	yes	clear face	yes
	I:2	p.M272L	yes			yes		yes		-	yes	clear face	no
	II:1 (1-4)	p.M272L	yes			yes		yes		have no idea about others, think she is the same as others	yes	clear face	diliberately need
	II:2	p.M272L	yes			yes		yes		-	yes	local face	diliberately need

1-5	1-5	p.T374A	yes	yes	yes	in middle school	not clear	blur face	no
1-6	II:2	-	no	no	no	-	yes	clear face	yes
	II:3	p.V548I	yes	yes	yes	in recent years	yes	blur face	no
	III:1 M31	-	no	no	no	-	yes	clear face	yes
	III:2 (1-6)	p.V548I	yes	yes	yes	at university	yes	clear face	yes
	III:3	p.V548I	yes	-	yes	have no idea about others, think he is the same as his sister III:4	yes	clear face	diliberately need
	III:4	p.V548I	yes	yes	yes	after starting job carrer	yes	only beard	yes
1-7	I:1	p.R641Q	yes	yes	yes	-	yes	clear face	no
	II:1 (1-7)	p.R641Q	yes	yes	yes	in childhood	yes	what she images is not the same as she sees	yes
	II:2	p.R641Q	yes	yes	yes	-	yes	clear face	yes
	II:3	-	no	no	no	-	yes	clear face	yes
	II:4	-	no	no	no	-	yes	clear face	yes

Member codes with dark grey background indicate CP candidates who showed daily hallmarks of CP. The index for each family is shown in bold and underlined. Whether there is an obvious key manifestation for CP is expressed by yes or no. More detailed diagnostic information can be obtained by contacting the corresponding author.

Table S9. Rare Variants (MAF<0.005) in *MCTP2* from a Second Cohort of 1928 Students

Chr15 Position*	RefSNP ID#	AA Change.refGene	AAF				Non-Synonymous Variants Annotation				
		NM_018349	1000G2015Aug_EAS	ExAC_EAS	gnomAD_EAS	our cohort of 1928 individuals	SIFT	Polyphen2 HDIV	Polyphen2 HVAR	LRT	Mutation Taster
94841499	rs776199245	exon1:c.A5T:p.D2V	.	0.0004	0.0006	0.0003	D	D	D	D	D
94841634	rs149133063	exon1:c.G140A:p.R47H	.	0.0001	0	0.0003	D	D	P	D	D
94841676	.	exon1:c.T182A:p.L61X	.	.	.	0.0003	.	.	.	D	A
94841690	rs141454907	exon1:c.C196T:p.R66W	.	0.0002	0.0006	0.0003	D	D	B	N	N
94841732	rs574533536	exon1:c.239delG:p.S80Tfs*17	0.005	0.0031	0.0045	0.0026	.	.	.	.	.
94841776	rs376018109	exon1:c.C282G:p.S94R	.	0.0009	0.0039	0.0010	D	P	B	N	D
94841843	rs748435445	exon1:c.349_350insAAACAGA:p.D119Efs*13	.	.	.	0.0003	.	.	.	.	.
94841861	rs200962853	exon1:c.G367A:p.A123T	0.004	0.0045	0.0045	0.0075	T	B	B	N	N
94841867	rs377631883	exon1:c.G373A:p.A125T	.	0.0001	0	0.0005	T	B	B	N	N
94841889	rs199860245	exon1:c.G395A:p.R132Q	0.002	0.0011	0.0019	0.0023	T	B	B	N	N
94841934	rs764405023	exon1:c.G440T:p.G147V	.	0.0005	0.0013	0.0016	D	P	B	N	N
94857525	rs749409827	exon2:c.C480A:p.S160R	.	0.0001	.	0.0005	D	P	B	N	N
94857550	rs768863067	exon2:c.T505C:p.S169P	.	0.0005	.	0.0003	D	D	D	D	D
94857572	rs772412690	exon2:c.C527G:p.S176C	.	0.0004	.	0.0013	T	B	B	N	D
94858867	rs750195259	exon3:c.637+1->T	.	0.0016	0.0006	0.0013	.	.	.	.	.
94882518	rs756489641	exon4:c.638-1G>C	.	0.0001	.	0.0005	.	.	.	.	D
94882549	.	exon4:c.T668C:p.L223P	.	.	.	0.0003	D	D	D	D	D
94883462	.	exon5:c.T815C:p.M272T	.	.	.	0.0003	D	D	D	D	D
94884094	rs199933047	exon6:c.A910C:p.M304L	0.001	.	0.0006	0.0003	T	P	B	D	D
94884106	rs145490381	exon6:c.G922C:p.V308L	0.002	.	0.0006	0.0003	T	B	B	D	D
94899388	rs77311454	exon8:c.G1028A:p.R343Q	.	0.0001	0	0.0005	D	D	P	D	D
94899411	rs762149227	exon8:c.A1051G:p.N351D	.	0.0001	.	0.0005	D	D	P	D	D
94899432	rs199572394	exon8:c.A1072G:p.I358V	.	.	.	0.0003	T	B	B	D	D
94899475	.	exon8:c.G1115A:p.S372N	.	.	.	0.0003	T	B	B	N	D
94899480	rs764700909	exon8:c.A1120G:p.T374A	.	0.0002	.	0.0005	T	B	B	N	D
94899488	.	exon8:c.G1128A:p.M376I	.	.	.	0.0003	T	B	B	N	D
94899522	.	exon8:c.A1162G:p.K388E	.	.	.	0.0003	D	D	D	D	D
94901803	rs199614826	exon9:c.G1263A:p.W421X	0.001	0.0006	.	0.0005	.	.	.	D	A
94901805	rs761523056	exon9:c.G1265A:p.G422E	.	0.0002	.	0.0003	T	B	B	N	D
94901834	rs747502432	exon9:c.1295delT:p.L432Rfs*13	.	0.0006	.	0.0026	.	.	.	.	.
94901842	rs746093287	exon9:c.1301+1G>A	.	0	.	0.0005	.	.	.	.	D
94910856	.	exon10:c.C1324T:p.L442F	.	.	.	0.0003	T	D	P	D	D
94910913	.	exon10:c.C1381T:p.L461F	.	.	.	0.0005	D	D	D	D	N

94910933	.	exon10:c.1402dupA:p.T468Nfs*10	.	.	.	0.0003	.	.	.	.	.
94911020	rs2289010	exon10:c.T1488A:p.Y496X	0.001	0.0009	0.0006	0.0013	.	.	.	D	A
94913325	.	exon11:c.A1498T:p.N500Y	.	.	.	0.0003	D	D	P	D	D
94913349	rs772110471	exon11:c.G1522A:p.V508I	.	0.0001	.	0.0005	T	B	B	N	N
94927287	.	exon12:c.A1619G:p.N540S	.	.	.	0.0003	D	B	B	D	D
94927292	rs374061496	exon12:c.C1624T:p.R542X	0.004	0.0024	0.0032	0.0023	.	.	.	D	A
94927310	rs182595113	exon12:c.G1642A:p.V548I	0.003	0.0008	0.0006	0.0003	T	B	B	N	N
94927336	.	exon12:c.G1668A:p.W556X	.	.	.	0.0003	.	.	.	D	A
94927338	.	exon12:c.1671delC:p.N557Kfs*17	.	.	.	0.0003	.	.	.	.	.
94928755	rs200626556	exon13:c.1788+1G	0.001	0.0005	0	0.0010	.	.	.	.	D
94942206	rs138009671	exon14:c.C1805T:p.P602L	.	0.0001	0.0006	0.0008	T	B	B	N	D
94942260	rs749725540	exon14:c.T1859C:p.I620T	.	0.0002	.	0.0003	D	D	D	D	D
94943181	rs749217960	exon15:c.G1922A:p.R641Q	.	0	.	0.0003	D	B	B	D	D
94943208	rs137927917	exon15:c.G1949A:p.R650H	.	.	0	0.0003	T	D	P	D	D
94945175	.	exon16:c.T2012C:p.I671T	.	.	.	0.0003	D	B	B	D	D
94945235	rs576073379	exon16:c.C2072T:p.T691I	0.001	0.001	0.0013	0.0003	T	B	B	N	N
94945247	rs370153540	exon16:c.C2084T:p.A695V	.	0.001	0.0006	0.0008	T	B	B	N	N
94983525	rs751758425	exon17:c.C2206T:p.Q736X	.	0.0001	.	0.0005	.	.	.	N	A
94986160	rs767710235	exon18:c.A2221G:p.I741V	.	0.0002	0.0006	0.0005	T	B	B	N	N
94986169	.	exon18:c.G2230A:p.E744K	.	.	.	0.0003	D	B	B	D	D
95019951	rs889873034	exon21:c.C2497T:p.R833X	.	.	0.0006	0.0008	.	.	.	D	A
95019952	rs187187495	exon21:c.G2498A:p.R833Q	.	0.0001	0	0.0003	D	D	D	D	D
95019969	rs140817032	exon21:c.G2515A:p.D839N	.	0	0	0.0003	T	D	D	D	D
95022248	rs746937797	exon22:c.G2622T:p.K874N	.	0.0002	.	0.0005	D	D	D	N	D

Note:* Start Position Based on the MCTP2 RefSeq (hg19, NM_018349);# dbSNP build 150

Table S10. Symptoms for the Diagnosis in Persons with c.239delG, p.S80fs

ID	Family member	Mutation	Key Manifestations									Facultative symptoms	
			Lasting and irritating subjective uncertainty of face recognition	Face recognition deficit especially in crowded places or out-of-context encounters	False negative and false positive face recognition events	Face recognition time longer than socially accepted	Face learning time longer than socially accepted	Development of adaptive behavior	Use of explicit learning strategies for visual person recognition	First time to realize*	Affected first degree relatives	Images of close friends' faces*	Gaze contact necessary in conversation
3-1	3-1	p.S80fs	yes			yes		yes		in childhood	-	clear face	no
3-2	II:1	-	no			no		no		-	yes	clear face	yes
	II:2	-	no			no		no		-	yes	clear face	yes
	II:3	p.S80fs	yes			yes		yes		after contacted	yes	clear face	occasionally
	II:4	-	no			no		no		-	no	clear face	yes
	II:5	p.S80fs	yes			yes		yes		after contacted	yes	clear face	yes
	II:6	p.S80fs	yes			yes		yes		before contacted	yes	blur face	no
	III:1	-	no			no		no		-	no	clear face	yes
	III:2 (3-2)	p.S80fs	yes			yes		yes		in childhood/ When she was a child, it was strange that she would always meet 'new people' at her family reunion. Actually, they were her relatives met several times every year	yes	local face	yes

	III:3	p.S80fs	yes	yes	yes	at university/joked by her roommates as 'faceblind'	yes	clear face	no
	III:4	p.S80fs	yes	yes	yes	after contacted	yes	local face	no
3-3	I:1	p.S80fs	yes	yes	yes	think the same as others	yes	clear face	yes
	I:2	-	no	no	no	-	yes	clear face	yes
	II:1	-	no	no	no	-	yes	clear face	yes
	II:2 (3-3)	p.S80fs	yes	yes	yes	at university after take a face behavioral test	yes	local face	no
3-4	3-4	p.S80fs	yes	yes	yes	at university	-	local face	seldom
3-5	II:1	-	no	no	no	no	yes	clear face	yes
	II:2	p.S80fs	yes	yes	yes	before we contacted her	yes	clear face	yes
	III:1	p.S80fs	yes	yes	yes	think the same as others	yes	clear face	yes
	III:2 (3-5)	p.S80fs	yes	yes	yes	different from others in the way of recognizing faces	yes	clear face	yes
	III:3	-	no	no	no	-	-	clear face	yes
	III:4	-	no	no	no	-	-	clear face	yes
3-6	I:1	-	no	no	no	-	yes	clear face	yes
	I:2	p.S80fs	yes	yes	yes	think the same as others	yes	clear face	no
	II:1	p.S80fs	yes	yes	yes	think the same as others	yes	clear face	yes
	II:2 (3-6)	p.S80fs	yes	yes	yes	think the same as others	yes	clear face	yes
3-7	3-7	p.S80fs	yes	yes	yes	think the same as others	-	blur face	yes
	I:1	p.S80fs	yes	yes	yes	think the same as others	-	clear face	yes

3-8	I:2	-	yes	no	yes	think better than others	-	clear face	yes
	II:1 (3-8)	p.S80fs	yes	yes	yes	think the same as others	yes	local face	yes
3-9	3-9	p.S80fs	yes	yes	yes	think the same as others	-	clear face	diliberately need
3-10	3-10	p.S80fs	yes	yes	yes	think better than others	-	clear face	yes
3-11	3-11	p.S80fs	yes	yes	yes	think the same as others	-	what she images is not the same as she looks, "a kind of sense"	yes
3-12	**II:1 (3-12)**	p.S80fs	yes	yes	yes	think better than others	-	clear face	yes
	II:2	-	no	no	no	-	yes	clear face	yes
3-13	3-13	p.S80fs	yes	yes	yes	think better than others because pay special attention to faces	-	clear face	yes
3-14	3-17	p.S80fs	no	no	no	-	-	clear face	yes

Member codes with dark grey background indicate CP candidates who showed daily hallmarks of CP. The index for each family is shown in bold and underlined. Whether there is an obvious key manifestation for CP is expressed by yes or no. More detailed diagnostic information can be obtained by contacting the corresponding author.