Human Molecular Genetics, 2012, Vol. 21, No. 5 1184–1189
doi:10.1093/hmg/ddr550
Advance Access published on November 24, 2011

Two-stage genome-wide association study identifies variants in *CAMSAP1L1* as susceptibility loci for epilepsy in Chinese

Youling Guo[1], Larry W. Baum[4], Pak Chung Sham[1], Virginia Wong[2], Ping Wing Ng[6], Colin Hiu Tung Lui[7], Ngai Chuen Sin[8], Tak Hong Tsoi[9], Clara S.M. Tang[1], Johnny S.H. Kwan[1], Benjamin H.K. Yip[1], Su-Mei Xiao[3], G. Neil Thomas[10], Yu Lung Lau[2], Wanling Yang[2], Stacey S. Cherny[1,*] and Patrick Kwan[5,*]

[1]Department of Psychiatry and The State Key Laboratory of Brain and Cognitive Sciences, [2]Department of Paediatric and Adolescent Medicine and [3]Department of Medicine and Research Centre of Heart, Brain, Hormone & Healthy Aging, The University of Hong Kong, Hong Kong, China, [4]School of Pharmacy and [5]Division of Neurology, Department of Medicine and Therapeutics, Prince of Wales Hospital, The Chinese University of Hong Kong, Hong Kong, China, [6]Department of Medicine and Geriatrics, United Christian Hospital, Hong Kong, China, [7]Department of Medicine, Tseung Kwan O Hospital, Hong Kong, China, [8]Hospital Authority Head Office, Hong Kong, China, [9]Department of Medicine, Pamela Youde Nethersole Eastern Hospital, Hong Kong, China and [10]Public Health, Epidemiology and Biostatistics, University of Birmingham, Birmingham, UK

Received May 18, 2011; Revised November 11, 2011; Accepted November 18, 2011

In the majority of patients, epilepsy is a complex disorder with multiple susceptibility genes interacting with environmental factors. However, we understand little about its genetic risks. Here, we report the first genome-wide association study (GWAS) to identify common susceptibility variants of epilepsy in Chinese. This two-stage GWAS included a total of 1087 patients and 3444 matched controls. In the combined analysis of the two stages, the strongest signals were observed with two highly correlated variants, rs2292096 [G] [$P = 1.0 \times 10^{-8}$, odds ratio (OR) = 0.63] and rs6660197 [T] ($P = 9.9 \times 10^{-7}$, OR = 0.69), with the former reaching genome-wide significance, on 1q32.1 in the *CAMSAP1L1* gene, which encodes a cytoskeletal protein. We also refined a previously reported association with rs9390754 ($P = 1.7 \times 10^{-5}$) on 6q21 in the *GRIK2* gene, which encodes a glutamate receptor, and identified several other loci in genes involved in neurotransmission or neuronal networking that warrant further investigation. Our results suggest that common genetic variants may increase the susceptibility to epilepsy in Chinese.

INTRODUCTION

Affecting up to 1% of people, epilepsy is the most common serious chronic neurological disorder. Twin studies suggest that epilepsy is highly heritable (1). Although a number of familial epilepsy syndromes are recognized to result from single-gene mutations, in the majority of patients, epilepsy is thought to be a complex disorder with multiple susceptibility genes interacting with various environmental factors that include acquired CNS insults or underlying structural brain abnormalities (e.g. stroke, head trauma, tumor) (2). Despite drug treatment, up to 30% of patients have persistent seizures (3). Discovery of genetic variants predisposing to the development of epilepsy would advance our understanding of epileptogenesis, leading to new drug targets, and facilitate the evaluation of potentially anti-epileptogenic therapies by targeting genetically susceptible individuals following CNS insults.

*To whom correspondence should be addressed at: Department of Psychiatry, The University of Hong Kong, Hong Kong, China. Tel: +852 28199581; Fax: +852 28199550; Email: cherny@hku.hk (S.C.); Department of Medicine and Therapeutics, Prince of Wales Hospital, Hong Kong, China. Tel: +852 26322211; Fax: +852 26321546; Email: patrickkwan@cuhk.edu.hk (P.K.).

There have been many attempts to identify the genetic susceptibility variants for the common epilepsy syndromes using association studies of candidate genes selected either for their role in monogenic epilepsy syndromes or on limited understanding of the pathobiology that underlies epileptogenicity or seizure propagation. Disappointing results of these studies (4) argue for a genome-wide approach without a priori assumptions, which may discover previously unsuspected markers. So far only one genome-wide association study (GWAS) of epilepsy has been reported, which was conducted in European subjects with partial (focal) epilepsy of both known (‘symptomatic’) and unknown (‘cryptogenic’) causes (5). Though the quantile–quantile (Q–Q) plots showed a slight departure from normal expectation, none of the *P*-values in their study reached the genome-wide significance threshold. Because of differences in genetic structures between different ethnic populations, it is possible that some genetic factors influencing susceptibility to epilepsy may also differ. Here, we report the first GWAS of epilepsy in Chinese.

RESULTS

In the discovery stage, testing for population stratification using EIGENSOFT and principal components analysis (PCA) found 15 significant principal components ($P < 0.05$) which explained the largest proportion of inter-individual variation. These were controlled for in the subsequent genome-wide association analysis. Supplementary Material, Figure S1 shows the plot of the first two eigenvectors, indicating ancestry difference and individual admixture among the study sample. A Q–Q plot showed an excess of *P*-values expected under the null hypothesis (genomic inflation factor $\lambda = 1.15$), which was reduced ($\lambda = 1.02$) after adjustment for population stratification (Supplementary Material, Fig. S2A). The Manhattan plot showed several genomic regions as potential risk loci (Supplementary Material, Fig. S2B), although none was genome-wide significant.

In the replication stage, among the 80 single-nucleotide polymorphisms (SNPs) analyzed, 7 were nominally significant with one-tailed $P < 0.05$ (Table 1 and Supplementary Material, Table S3), and one was significantly associated with epilepsy after Bonferroni correction (rs2292096 in 1q32.1). In the combined meta-analysis of the discovery and replication sample sets, this SNP in *CAMSAP1L1* achieved genome-wide significance [rs2292096 [G], $P = 1.0 \times 10^{-8}$, odds ratio (OR) = 0.63], and a second, highly correlated SNP ($r^2 = 0.87$) was nearly significant (rs6660197 [T], $P = 9.9 \times 10^{-7}$, OR = 0.69) (Table 1). The OR for the two SNPs in the two stages showed no significant heterogeneity ($P_{het} > 0.1$).

We used SNP imputation to investigate association strength at untyped markers near the significant loci. Considering *CAMSAP1L1*, there was a block of SNPs in tight linkage disequilibrium (LD) ($r^2 > 0.8$) with much stronger significance than any other SNPs at this locus, encompassing a 124 kb region likely to contain the causative variant(s) responsible for the association observed (Fig. 1A). Moderate association was observed for three genotyped SNPs (rs2292096, $P = 6.8 \times 10^{-6}$; rs12742404, $P = 6.8 \times 10^{-6}$; rs6660197,

Table 1. Details of SNPs with one genome-wide significant association and candidate loci with moderate associations in the discovery and replication cohorts, listed by chromosome and position order

Chr	SNP	Position	Gene	Minor allele	Discovery cohort				Replication cohort				Combined cohorts
					Case MAF	Control MAF	OR (95% CI)	*P*-value	Case MAF	Control MAF	OR (95% CI)	*P*-value*	*P*-value
1q32.1[a]	rs6660197	198 970 164	*CAMSAP1L1*	T	0.139	0.200	0.65 (0.53–0.79)	4.83×10^{-5}	0.161	0.203	0.75 (0.60–0.94)	0.00599	9.89×10^{-7}
1q32.1[a]	rs2292096	199 093 392	*CAMSAP1L1*	G	0.120	0.184	0.61 (0.49–0.75)	6.76×10^{-6}	0.139	0.194	0.67 (0.53–0.85)	0.00038	1.04×10^{-8}
2p12[b]	rs4853352	77 993 057	*SNAR-H*	G	0.246	0.185	1.48 (1.25–1.74)	1.54×10^{-6}	0.193	0.177	1.12 (0.89–1.39)	0.169	1.88×10^{-5}
2p12[b]	rs2164851	78 155 723	*SNAR-H*	C	0.181	0.125	1.48 (1.23–1.79)	2.44×10^{-5}	0.258	0.207	1.16 (1.00–1.34)	0.026	5.27×10^{-5}
2p11.2[b]	rs2367575	86 574 352	*KDM3A*	G	0.076	0.042	1.68 (1.27–2.22)	5.41×10^{-5}	0.054	0.037	1.50 (0.97–2.31)	0.033	4.93×10^{-5}
2q34[b]	rs13021324	212 058 490	*ERBB4*	C	0.273	0.331	0.73 (0.63–0.86)	0.000162	0.311	0.346	0.86 (0.72–1.03)	0.048	5.8×10^{-5}
5p13.2[b]	rs6872795	35 713 142	*SPEF2*	A	0.157	0.119	1.47 (1.21–1.79)	0.00013	0.121	0.093	1.33 (1.01–1.75)	0.022	1.81×10^{-5}
6q21[c]	rs9390754	102 071 607	*GRIK2*	G	0.396	0.330	1.39 (1.20–1.60)	1.71×10^{-5}	0.323	0.332	0.96 (0.80–1.15)	0.668	0.0015
6q21[c]	rs4840200	102 433 996	*GRIK2*	T	0.434	0.504	0.73 (0.64–0.85)	4.08×10^{-5}	0.474	0.469	1.02 (0.86–1.21)	0.592	0.0017
7q31.31[b]	rs702416	119 942 506	*KCND2*	C	0.430	0.497	0.76 (0.66–0.87)	0.00012	0.482	0.510	0.89 (0.75–1.06)	0.095	0.00013
21q22.2[b]	rs1980406	40 665 509	*DSCAM*	C	0.068	0.036	1.78 (1.32–2.40)	1.8×10^{-5}	0.049	0.031	1.61 (1.02–2.54)	0.021	1.66×10^{-5}

Position is as Genome Build 36.3 in base pairs. Chr, chromosome band; CI, confidence interval. Three types of associations are presented: [a]*CAMSAP1L1*, with SNPs reaching genome-wide significance in the combined analysis; [b]six loci with moderate associations observed, which involve genes that represent candidate genes of epilepsy; [c]a previously reported association, *GRIK2*, which is refined with SNPs attaining significant *P*-values among prime candidate genes in the discovery cohort.
*One-tailed *P*-value in the replication stage.

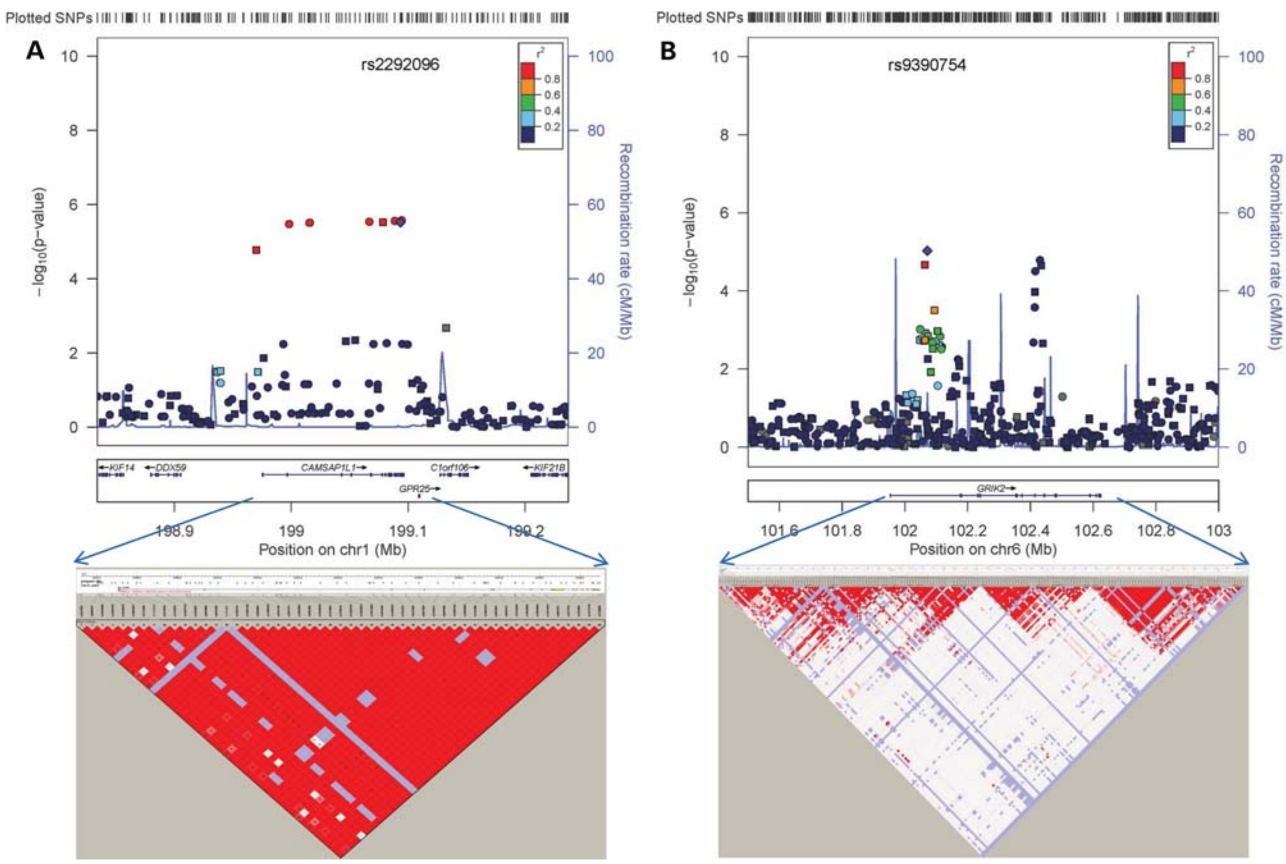

Figure 1. Association and LD plots of regions located on (**A**) 1q32.1 in a 250 kb interval in the *CAMSAP1L1* gene, showing a genome-wide significant association, and on (**B**) 6q21in a 672 kb interval on chromosome 6, refining a previous association signal in the gene *GRIK2*. Discovery cohort association significance in the regions is plotted against the left-hand *y*-axis as −log 10 (*P*-value). Genotyped SNPs were tested using the Cochran–Armitage trend test; imputed SNPs were tested using a regression analysis based on the imputed allelic dosage; both tests were adjusted for population clusters. Genetic coordinates are as NCBI Genome Build 36.3. The diamond denotes the association hit; rectangles, genotyped SNPs; circles, imputed SNPs; color scale, LD with the hit; purple line and right-hand *y*-axis, recombination rate (cM/Mb as per HapMap data in CHB population). LD plots for the highlighted regions are based on HapMap CHB release 22 using the pair-wise correlation coloring scheme of Haploview.

$P = 4.8 \times 10^{-5}$) and six imputed SNPs ($P < 3.5 \times 10^{-6}$) at the locus in the discovery sample set (Fig. 1A).

We also examined the *P*-values from the discovery stage obtained for 1710 SNPs of 194 prime candidate genes previously investigated for possible association with epilepsy (4) (Q–Q plot shown in Supplementary Material, Figure S4). Details of the SNPs and genes examined are provided in Supplementary Material, Table S4. Three SNPs in *GRIK2* on 6q21 were significantly associated, and another nearly so, after controlling the false discovery rate (FDR) for 1710 tests, with rs9390754 attaining $P = 1.7 \times 10^{-5}$; rs7747072, $P = 3.6 \times 10^{-5}$; rs4840200, $P = 4.1 \times 10^{-5}$; and rs9390790, $P = 1.2 \times 10^{-4}$. Association testing of imputed SNPs revealed two regions with *P*-values close to those of the above four genotyped SNPs within *GRIK2*. Two *GRIK2* SNPs (rs9390754 and rs4840200) were genotyped in the replication phase but revealed no significant difference between the minor allele frequencies in cases and controls (Supplementary Material, Table S3). The 6q21 locus between rs9390754 and rs4840200 is marked by three recombination hot spots, which makes it possible that there are combined effects of the two SNPs (Fig. 1B). The two SNPs are weakly correlated ($r^2 = 0.001$, $D' = 0.034$) and are separated by several LD blocks. Multiple logistic regression analysis showed that rs9390754 ($P = 0.0001$) and rs4840200 ($P = 0.0002$) had independent effects. A 2 df likelihood ratio test that accounted for the genotypic additive effects of the two SNPs combined also showed strong evidence of the combined effects ($P = 3.8 \times 10^{-7}$).

Associations with additional SNPs in six loci were replicated with moderate *P*-values, suggesting that other risk variants with a modest effect remain to be identified. These may include variants located on chromosomes 2p12, 2p11.2, 2q34, 5p13.2, 7q31.31 and 21q22.2, which did not reach genome-wide significance but could be considered suggestively associated with epilepsy (Table 1). In the 2p12 locus, eight SNPs lying in a 224 kb region near *SNAR-H* achieved moderate levels of association in the discovery set ($P < 10^{-4}$; Supplementary Material, Fig. S3A). The strongest association in this region was observed with the genotyped marker rs4853352 in the discovery stage ($P = 1.5 \times 10^{-6}$). One marker (rs2164851) in the *SNAR-H* gene also achieved nominal significance in the replication study ($P = 0.026$) and showed a moderate level of significance in the combined analysis ($P = 5.3 \times 10^{-5}$). The regional association plots of these suggestive regions are represented in Supplementary Material, Figure S3A.

DISCUSSION

Studies attempting to identify susceptibility genes have tended to focus on the idiopathic epilepsy syndromes, which usually occur in children/adolescents and are generally drug responsive. However, little attention has been paid to symptomatic epilepsy, which is more often drug resistant, constituting a major unsolved public health burden (3). It has been proposed that causation of epilepsies can be regarded as a biological continuum, and the degree of genetic contribution may differ among different syndromes, being greater in idiopathic than symptomatic syndromes (6). Genetic predisposition to symptomatic epilepsy is not an entirely novel concept, although very few studies have attempted to identify the genetic markers associated with increased risk of epilepsy following CNS insults. This hypothesis is supported by twin studies showing pairwise concordance for symptomatic generalized epilepsy and for partial epilepsy (7). APOE ε4 allele carriers were found to have a 2.4-fold increase in risk of epilepsy following traumatic brain injury (8).

In this two-stage GWAS of epilepsy in Chinese, combined analysis showed the strongest signals with two highly correlated variants, namely rs2292096 and rs6660197 on 1q32.1 in the *CAMSAP1L1* gene, with the former reaching genome-wide significance. Imputation revealed a block of SNPs likely containing the causative variant(s), although none of the SNPs examined is in the translated regions of the gene. CAMSAP1L1 (CAMSAP1-like 1; also called CAMSAP1L2 or KIAA1078) is a calmodulin-regulated spectrin-associated protein (CAMSAP) that belongs to a novel family of cytoskeletal proteins of little-known function. CAMSAP1 has been reported to be expressed in neurons and astrocytes in the mammalian nervous system, where it is suggested to interact with intermediate filaments (9). Further work has identified a unique structural domain common to the three members of the protein family that is able to inhibit neurite extension, most likely by blocking microtubule function (10).

Using results of the discovery stage, we also refined a previously reported association with rs9390754 in the *GRIK2* gene on 6q21. *GRIK2*, also called *GluR6*, is the glutamate receptor 6 gene, one of the high-ranking candidate genes for epilepsy. Knockout mice deficient in the kainate-selective *GRIK2* subunit of the kainate receptor had reduced susceptibility to kainate-induced seizures (11). Alteration in *GRIK2* mRNA editing in neocortical tissue reflects an adaptive reaction to ongoing seizure activity and may play a role in pathological processes which contribute to seizure maintenance (12). Forced overexpression of the *GRIK2* kainate receptor within the hippocampus induced seizures (13).

Weaker signals were observed in other loci, some of which involved genes that could also represent candidate genes of epilepsy (Table 1). These include *KCND2*, which encodes the voltage-gated potassium channel Kv4.2, a key component of the A-type potassium currents in the CNS that critically regulate action potential back propagation and the induction of specific forms of synaptic plasticity (14). Kv channels are increasingly recognized to play an important role in the pathogenesis of epilepsy. Specifically, Kv4.2 knockout mice demonstrated increased susceptibility to seizure induction (15). Dynamic alterations in Kv4.2 channel expression and localization were observed in a variety of focal lesions associated with refractory epilepsy in humans (16). Another potential locus was found in *ERBB4*, which is recognized as a schizophrenia-susceptibility gene but has also been reported to be mutated in a case of early myoclonic encephalopathy (17). It encodes ErbB4, which is a member of the type I receptor tyrosine kinase subfamily that promotes synapse formation of GABA-containing interneurons in the hippocampus (18).

Other suggestive signals were found in genes previously unsuspected but which could plausibly be associated with epilepsy based on limited knowledge of their functions. They include *SNAR-H*, a member of the small NF90-associated RNA family expressed in many human tissues, with minor distribution in the brain (19). *LRRTM4*, ~400 kb from *SNAR-H*, is abundant in the dentate gyrus and helps regulate cell–cell contact (20), which could plausibly be involved in neuronal connectivity in epilepsy. *DSCAM*, on 21q22.2, is suggested to play an important role in neural circuitry development by allowing neurite 'self avoidance', which refers to the tendency of branches from the same neuron to selectively avoid one another (21). *DSCAM* knockout mice have dysregulated central respiratory function because of impaired neural synchroneity; whether the abnormality extends to other parts of the brain is unclear (22). *KDM3A* (also called *JMJD1A* or *JHDM2A*) on 2p11.2 is a histone demethylase that may reprogram neural stem cells (23). *SPEF2* (also known as *KPL2*) is preferentially expressed in tissues that contain axonemal structures such as the brain, lung and testis, and as with *CAMSAP1L1*, has a calponin homology domain (24).

The present study was limited by a small number of cases, so that only associations with relatively large effect could be detected. Given that the associations found in the present study were not detected (Supplementary Material, Table S2) in a previous GWAS of focal epilepsy in patients of European ancestry (5), it is possible that they are unique to the Chinese population. Allele frequencies in controls in Europeans (0.118 for both *CAMSAP1L1* SNPs; see Supplementary Material, Table S2) are much lower than those in Chinese controls (0.200 and 0.184) and even lower than in Chinese cases (0.139 and 0.120). Still, the frequencies in European cases are slightly lower, in the same direction of effect as found in the present Chinese sample. The discrepant results between the two GWASs might also be due to differences in epilepsy phenotypes. The European study, employing a heterogeneous sample, likely only investigated genetic factors shared across partial epilepsies, disregarding the type of epilepsy (idiopathic, cryptogenic or symptomatic). For instance, 27% of patients in the European study had hippocampal sclerosis and the cause of epilepsy was unknown in 41% (5). In comparison, the most common structural-metabolic cause in the present study was stroke (17% of patients), and only patients with symptomatic epilepsy were included in the discovery stage, although the replication stage also included patients with cryptogenic epilepsy, which may have a different pathogenesis. Our Hong Kong cases appear to be recruited from a more homogeneous and narrowly defined group of patients, and it is possible that syndrome-specific common genetic causes do exist and were detected in the present study.

In conclusion, this GWAS identified common genetic variants that may increase the susceptibility to epilepsy in Chinese. *CAMSAP1L1* and the other genes where suggestive loci were found might be considered candidate genes because of their known or potential role in neurotransmission, neuronal networking and connectivity. These findings lend support to the concept that epileptogenesis may result from a distributed hyperexcitable circuitry rather than a homogenous epileptogenic focus (25). We suggest that a GWAS in larger cohorts of Chinese subjects should be performed to confirm the findings, and studies should be conducted to explore the mechanisms for these associations.

MATERIALS AND METHODS

Study participants

Epilepsy patients of Han Chinese ethnicity aged between 2 and 91 years were recruited from neurology clinics of five regional hospitals in Hong Kong covering a combined catchment population of approximately 3 million. Exclusion criteria included significant psychiatric comorbidity, history of pseudoseizures, alcohol or illicit drug abuse, and presence of progressive or degenerative neurological or systemic disorders. Syndromic classification was adapted from the revised international organization of phenotypes in epilepsy (26). The study included a total of 1087 Chinese epilepsy patients and 3444 ethnically matched controls. The discovery stage included 504 patients with symptomatic focal epilepsy and 2947 ethnically matched controls. All epilepsy patients were recruited in Hong Kong. Non-epilepsy controls were recruited from two sources: subjects recruited for other studies conducted in the University of Hong Kong, genotyped with the same platform ($n = 1947$), and healthy individuals recruited in Taiwan ($n = 1000$). All subjects in the replication stage were recruited in Hong Kong. They consisted of 583 patients with either symptomatic or cryptogenic focal epilepsy and 497 controls who were healthy blood donors kindly contributed by the Hong Kong Red Cross. Supplementary Material, Table S1 provides the clinical characteristics of the cases included in the data analysis after quality control filters. The study was approved by ethics committees of the participating hospitals, and all patients or their legal guardians gave written informed consent.

Genotyping and quality control

Genotyping of the discovery cohorts was performed using the Illumina platforms at deCODE Genetics, Iceland (http://www.decode.com). Cases and Hong Kong controls were genotyped with the HumanHap 610-Quad BeadChip, and Taiwan controls were genotyped using the HumanHap 550-Duo BeadChip. The results were then merged using PLINK (http://pngu.mgh.harvard.edu/purcell/plink). Common SNPs typed in both groups were identified by filtering against the HumanHap 550K Quad chip. A total of 88 subjects (16 cases, 72 controls) were excluded according to the following criteria: (i) genotyping call rate <95% ($n = 7$); (ii) very strong positive or negative autosomal heterozygosity ($n = 15$); (iii) related or duplicated individuals ($n = 25$); (iv) sex discrepancies ($n = 21$); and (v) outliers in a plot of multidimensional scaling analysis ($n = 20$). Nearly half a million common autosomal SNPs ($n = 461\,024$) passed the quality control thresholds of >95% call rate, >1% minor allele frequency (MAF) and Hardy–Weinberg equilibrium (HWE, $P \geq 0.0001$). The total genotyping call rate in included individuals was 99.83% for cases and 99.85% for controls.

In the replication stage, SNPs with the lowest *P*-values from the discovery stage were selected as follows. Genotyping assays were designed for a Sequenom MassARRAY iPlex System at the Hong Kong University Genome Research Centre, Hong Kong (http://genome.hku.hk). SNPs <100 kb from other selected SNPs with lower *P*-values were not initially selected. SNPs that could not be pooled together for genotyping within three pools were replaced with other SNPs within 100 kb or, if none, other SNPs with the next lowest *P*-values. The process was repeated until no more SNPs could be grouped into three pools, at which point there were 82 SNPs. These were genotyped. After quality control measures, 29 subjects (12 cases, 17 controls) were excluded owing to low call rates. In addition, results of two SNPs were excluded, one due to a call rate <95% and one due to violation of HWE ($P < 0.0001$) in the controls, resulting in 80 SNPs for analysis.

Statistical analysis

In the discovery stage, EIGENSOFT (http://helix.nih.gov/Applications/eigensoft.html) (27) and PCA were used to control for population stratification. For the correction of population structure, we excluded a subset of SNPs ($n = 363\,904$) in approximate LD with each other ($r^2 > 0.2$) before running PCA, and then obtained correlation matrices among remaining SNPs. Genome-wide association analysis was performed for the full set of SNPs, using ancestry-adjusted genotypes and phenotypes, controlling for the 15 significant components. Associations with $P < 5 \times 10^{-8}$ were considered genome-wide significant, as is generally accepted. Q–Q plots were constructed by contrasting uncorrected and corrected experimental *P*-value distributions to the expected uniform 0–1 distribution.

In the replication stage, case–control analysis was performed using the Cochran–Armitage trend test as implemented in PLINK (28). Meta-analysis of the discovery and replication cohorts was then performed using an inverse-variance-weighted method under a fixed-effects model as implemented in PLINK (28). Homogeneity for SNP effect across the studies was tested using the Cochran *Q* test (29).

Imputation analyses were performed with IMPUTE v2 (https://mathgen.stats.ox.ac.uk/impute/impute_v2.html) (30), taking data from CHB + CHD individuals from HapMap 3 as the reference set of haplotypes. We analyzed only regions surrounding significantly or marginally associated SNPs that were either genotyped or could be imputed with relatively high calling confidence (>90%). Association analysis of imputed SNPs was performed assuming an underlying additive model and including the first 15 EIGENSOFT eigenvectors as covariates, which accounted for uncertainty in prediction of the imputed data by weighing genotypes by their posterior probabilities.

The combined analysis of the two-stage study had >80% power to identify SNPs conferring genotypic relative risks of 1.5–4.5 with minor allele frequencies of 0.01–0.5 at $\alpha = 5 \times 10^{-8}$. Power calculations were performed using CaTS (www.sph.umich.edu/csg/abecasis/CaTS/).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

We are grateful to all participants who contributed to this study. We also thank Professor Annie W.C. Kung, Department of Medicine, The University of Hong Kong, for contributing to the control data in the discovery stage.

Conflict of Interest statement. The authors declare no competing financial interests.

FUNDING

The study was supported by the Research Grants Council of the Hong Kong Special Administrative Region, China (project numbers HKU7623/08M, HKU7747/07M and CUHK4466/06M).

REFERENCES

1. Kjeldsen, M.J., Kyvik, K.O., Christensen, K. and Friis, M.L. (2001) Genetic and environmental factors in epilepsy: a population-based study of 11900 Danish twin pairs. *Epilepsy Res.*, **44**, 167–178.
2. Crino, P.B. (2007) Gene expression, genetics, and genomics in epilepsy: some answers, more questions. *Epilepsia*, **48**(Suppl. 2), 42–50.
3. Kwan, P. and Brodie, M.J. (2000) Early identification of refractory epilepsy. *N. Engl. J. Med.*, **342**, 314–319.
4. Cavalleri, G.L., Weale, M.E., Shianna, K.V., Singh, R., Lynch, J.M., Grinton, B., Szoeke, C., Murphy, K., Kinirons, P., O'Rourke, D. *et al.* (2007) Multicentre search for genetic susceptibility loci in sporadic epilepsy syndrome and seizure types: a case-control study. *Lancet Neurol.*, **6**, 970–980.
5. Kasperaviciute, D., Catarino, C.B., Heinzen, E.L., Depondt, C., Cavalleri, G.L., Caboclo, L.O., Tate, S.K., Jamnadas-Khoda, J., Chinthapalli, K., Clayton, L.M. *et al.* (2010) Common genetic variation and susceptibility to partial epilepsies: a genome-wide association study. *Brain*, **133**, 2136–2147.
6. Helbig, I., Scheffer, I.E., Mulley, J.C. and Berkovic, S.F. (2008) Navigating the channels and beyond: unravelling the genetics of the epilepsies. *Lancet Neurol.*, **7**, 231–245.
7. Berkovic, S.F., Howell, R.A., Hay, D.A. and Hopper, J.L. (1998) Epilepsies in twins: genetics of the major epilepsy syndromes. *Ann. Neurol.*, **43**, 435–445.
8. Diaz-Arrastia, R., Gong, Y., Fair, S., Scott, K.D., Garcia, M.C., Carlile, M.C., Agostini, M.A. and Van Ness, P.C. (2003) Increased risk of late posttraumatic seizures associated with inheritance of APOE epsilon4 allele. *Arch. Neurol.*, **60**, 818–822.
9. Yamamoto, M., Yoshimura, K., Kitada, M., Nakahara, J., Seiwa, C., Ueki, T., Shimoda, Y., Ishige, A., Watanabe, K. and Asou, H. (2009) A new monoclonal antibody, A3B10, specific for astrocyte-lineage cells recognizes calmodulin-regulated spectrin-associated protein 1 (Camsap1). *J. Neurosci. Res.*, **87**, 503–513.
10. Baines, A.J., Bignone, P.A., King, M.D., Maggs, A.M., Bennett, P.M., Pinder, J.C. and Phillips, G.W. (2009) The CKK domain (DUF1781) binds microtubules and defines the CAMSAP/ssp4 family of animal proteins. *Mol. Biol. Evol.*, **26**, 2005–2014.
11. Mulle, C., Sailer, A., Perez-Otano, I., Dickinson-Anson, H., Castillo, P.E., Bureau, I., Maron, C., Gage, F.H., Mann, J.R., Bettler, B. *et al.* (1998) Altered synaptic physiology and reduced susceptibility to kainate-induced seizures in GluR6-deficient mice. *Nature*, **392**, 601–605.
12. Kortenbruck, G., Berger, E., Speckmann, E.J. and Musshoff, U. (2001) RNA editing at the Q/R site for the glutamate receptor subunits GLUR2, GLUR5, and GLUR6 in hippocampus and temporal cortex from epileptic patients. *Neurobiol. Dis.*, **8**, 459–468.
13. Telfeian, A.E., Federoff, H.J., Leone, P., During, M.J. and Williamson, A. (2000) Overexpression of GluR6 in rat hippocampus produces seizures and spontaneous nonsynaptic bursting *in vitro*. *Neurobiol. Dis.*, **7**, 362–374.
14. Birnbaum, S.G., Varga, A.W., Yuan, L.L., Anderson, A.E., Sweatt, J.D. and Schrader, L.A. (2004) Structure and function of Kv4-family transient potassium channels. *Physiol. Rev.*, **84**, 803–833.
15. Barnwell, L.F., Lugo, J.N., Lee, W.L., Willis, S.E., Gertz, S.J., Hrachovy, R.A. and Anderson, A.E. (2009) Kv4.2 knockout mice demonstrate increased susceptibility to convulsant stimulation. *Epilepsia*, **50**, 1741–1751.
16. Aronica, E., Boer, K., Doorn, K.J., Zurolo, E., Spliet, W.G., van Rijen, P.C., Baayen, J.C., Gorter, J.A. and Jeromin, A. (2009) Expression and localization of voltage dependent potassium channel Kv4.2 in epilepsy associated focal lesions. *Neurobiol. Dis.*, **36**, 81–95.
17. Backx, L., Ceulemans, B., Vermeesch, J.R., Devriendt, K. and Van Esch, H. (2009) Early myoclonic encephalopathy caused by a disruption of the neuregulin-1 receptor ErbB4. *Eur. J. Hum. Genet.*, **17**, 378–382.
18. Fazzari, P., Paternain, A.V., Valiente, M., Pla, R., Lujan, R., Lloyd, K., Lerma, J., Marin, O. and Rico, B. (2010) Control of cortical GABA circuitry development by Nrg1 and ErbB4 signalling. *Nature*, **464**, 1376–1380.
19. Parrott, A.M. and Mathews, M.B. (2007) Novel rapidly evolving hominid RNAs bind nuclear factor 90 and display tissue-restricted distribution. *Nucleic Acids Res.*, **35**, 6249–6258.
20. Lauren, J., Airaksinen, M.S., Saarma, M. and Timmusk, T. (2003) A novel gene family encoding leucine-rich repeat transmembrane proteins differentially expressed in the nervous system. *Genomics*, **81**, 411–421.
21. Fuerst, P.G., Bruce, F., Tian, M., Wei, W., Elstrott, J., Feller, M.B., Erskine, L., Singer, J.H. and Burgess, R.W. (2009) DSCAM and DSCAML1 function in self-avoidance in multiple cell types in the developing mouse retina. *Neuron*, **64**, 484–497.
22. Amano, K., Fujii, M., Arata, S., Tojima, T., Ogawa, M., Morita, N., Shimohata, A., Furuichi, T., Itohara, S., Kamiguchi, H. *et al.* (2009) DSCAM deficiency causes loss of pre-inspiratory neuron synchroneity and perinatal death. *J. Neurosci.*, **29**, 2984–2996.
23. Ma, D.K., Chiang, C.H., Ponnusamy, K., Ming, G.L. and Song, H. (2008) G9a and Jhdm2a regulate embryonic stem cell fusion-induced reprogramming of adult neural stem cells. *Stem Cells*, **26**, 2131–2141.
24. Ostrowski, L.E., Andrews, K., Potdar, P., Matsuura, H., Jetten, A. and Nettesheim, P. (1999) Cloning and characterization of KPL2, a novel gene induced during ciliogenesis of tracheal epithelial cells. *Am. J. Respir. Cell. Mol. Biol.*, **20**, 675–683.
25. Dichter, M.A. (2009) Emerging concepts in the pathogenesis of epilepsy and epileptogenesis. *Arch. Neurol.*, **66**, 443–447.
26. Berg, A.T., Berkovic, S.F., Brodie, M.J., Buchhalter, J., Cross, J.H., van Emde Boas, W., Engel, J., French, J., Glauser, T.A., Mathern, G.W. *et al.* (2010) Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology, 2005–2009. *Epilepsia*, **51**, 676–685.
27. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A. and Reich, D. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.*, **38**, 904–909.
28. 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.
29. Higgins, J.P. and Thompson, S.G. (2002) Quantifying heterogeneity in a meta-analysis. *Stat. Med.*, **21**, 1539–1558.
30. Howie, B.N., Donnelly, P. and Marchini, J. (2009) A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.*, **5**, e1000529.