Appendix e-1

Supplementary information

Cohort ancestry

Our cohorts consisted of individuals of European and Han Chinese ethnicity; subjects were assigned to the two broad ancestry groups according to results of genotype-based principal component analysis (PCA) using the programme GCTA.[1] The principal components plot of PC1 versus PC2 is shown in Supplementary Figure e-2 with study subjects overlaid on HapMap CEU, CHB and YRI ancestral populations.

Genotyping quality control

BeadStudio (Illumina; version 2.0) was used to call genotypes, normalize signal intensity data and establish the log R ratio and B allele frequency at every SNP. Quality control of genotypes was performed separately for each cohort using PLINK 1.9,[2] as follows: all markers with very high (> 0.9) rate of missing genotypes were removed. Genotypes for 2,766 ethnicity-sensitive SNPs common to all SNP arrays were extracted and STRUCTURE 2.2 was used to derive European and Asian ancestry probabilities, with Hapmap Yoruban, Nigeria (YRI) and Han Chinese in Beijing, China (CHB); and Utah residents with ancestry from northern and western Europe (CEU) samples as respective reference populations.[3] Subjects with less than 90% European ancestry in all cohorts were removed, except for the Hong Kong cohort where we excluded samples with less than 90% Asian ancestry. Hardy-Weinberg equilibrium (HWE) was calculated per ethnic group and in the case of the EpiPGX cohort, HWE was checked separately per recruitment site. A HWE cut-off threshold of $p < 1 \times 10^{-6}$ per marker was applied. All samples with < 0.98 genotype rate, all markers with < 0.95 genotype rate and all markers below < 0.01 minor allele frequency were removed. Next, a subset of markers independent of each other with respect to linkage disequilibrium (LD) was created using a window size of 100 markers shifting by 25 markers at a time and removing one half of every SNP pair with genotypic $r^2 > 0.1$. Using this subset of markers, heterozygosity (HET) and identity by state (IBS) was calculated in order to remove all samples with outlying HET values (> 5 standard deviations from the median of the whole sample) and one half of all sample pairs > 0.9 IBS. Subjects were removed if sex determined from genotype did not match reported gender. Array-specific maps retrieved at the website of Will Rayner at the Wellcome Trust (http://www.well.ox.ac.uk/~wrayner/strand/) were used to update all marker positions and chromosome numbers to the Genome Reference Consortium Human Build 37 (GRCh37) and all A/T and C/G markers were removed to avoid strand issues. Genotypes were split up according to chromosome arms (and in the case of chromosome X, it was split additionally into pseudo-autosomal regions (PAR) and non-PAR) and created phased haplotypes using SHAPEIT v2 with recommended effective size setting (11,418), and using 1000 Genomes phase 1 integrated (v3) map files as the reference map.[4] Following haplotype phasing, genotypes were imputed into our dataset using IMPUTE v2.3.0 with recommended effective population size settings (20,000) and 1000 Genomes phase 1 integrated (v3) genotypes as reference.[5] The haplotype phasing and imputation was also performed in separate batches for each cohort. Additionally, HLA alleles were imputed on to genotype data using SNP2HLA from the HapMap CEU reference panel for Europeans and an in-house Han Chinese reference panel for the Hong Kong population.[6,7] Post-imputation quality control filters were applied to remove imputed variants with imputation score <

0.9, call rate < 0.95 in either cases or controls and a minimum minor allele frequency of 2% across all samples to reduce spurious associations.

Study Power

We estimated that our meta-analyses had 80% power to detect a marker with allele frequency >2% and an alpha level of 1.25×10^{-8} with relative risks (approximated to odds ratio) ≥3, ≥4.5, ≥5 or ≥7 for all MPE, carbamazepine-MPE, lamotrigine-MPE and phenytoin-MPE respectively. For our European-specific analyses, we estimated that we had 80% power to detect relative risks ≥3.5, ≥6, ≥5.5 or ≥9.5 for all MPE, carbamazepine-MPE, lamotrigine-MPE and phenytoin-MPE respectively. For our Han Chinese population-specific analyses, we estimated that we had 80% power to detect relative risks ≥5, ≥7, ≥32 or ≥15 for all MPE, carbamazepine-MPE, lamotrigine-MPE and phenytoin-MPE respectively (see Supplementary Figure e-2).

References

1. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011; **88**(1): 76-82.
2. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**(3): 559-75.
3. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics* 2000; **155**(2): 945-59.
4. Delaneau O, Zagury JF. Haplotype inference. *Methods Mol Biol* 2012; **888**: 177-96.
5. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* 2009; **5**(6): e1000529.
6. Jia X, Han B, Onengut-Gumuscu S, et al. Imputing amino acid polymorphisms in human leukocyte antigens. *PLoS One* 2013; **8**(6): e64683.
7. Gui H, Kwok M, Baum L, Sham P, Kwan P, Cherny SS. SNP-based HLA allele tagging, imputation, and association with antiepileptic drug-induced cutaneous reactions in Hong Kong Han Chinese. *The Pharmacogenomics Journal (under review)* 2016.
8. McCormack M, Alfirevic A, Bourgeois S, et al. HLA-A*3101 and carbamazepine-induced hypersensitivity reactions in Europeans. *N Engl J Med* 2011; **364**(12): 1134-43.
9. Amstutz U, Ross CJ, Castro-Pastrana LI, et al. HLA-A 31:01 and HLA-B 15:02 as genetic markers for carbamazepine hypersensitivity in children. *Clin Pharmacol Ther* 2013; **94**(1): 142-9.
10. Kim SH, Lee KW, Song WJ, et al. Carbamazepine-induced severe cutaneous adverse reactions and HLA genotypes in Koreans. *Epilepsy Res* 2011; **97**(1-2): 190-7.
11. Ozeki T, Mushiroda T, Yowang A, et al. Genome-wide association study identifies HLA-A*3101 allele as a genetic risk factor for carbamazepine-induced cutaneous adverse drug reactions in Japanese population. *Hum Mol Genet* 2011; **20**(5): 1034-41.
12. Chung WH, Hung SI, Hong HS, et al. Medical genetics: a marker for Stevens-Johnson syndrome. *Nature* 2004; **428**(6982): 486.
13. Cheung YK, Cheng SH, Chan EJ, Lo SV, Ng MH, Kwan P. HLA-B alleles associated with severe cutaneous reactions to antiepileptic drugs in Han Chinese. *Epilepsia* 2013; **54**(7): 1307-14.
14. Tassaneeyakul W, Tiamkao S, Jantararoungtong T, et al. Association between HLA-B*1502 and carbamazepine-induced severe cutaneous adverse drug reactions in a Thai population. *Epilepsia* 2010; **51**(5): 926-30.
15. Li X, Yu K, Mei S, et al. HLA-B*1502 increases the risk of phenytoin or lamotrigine induced Stevens-Johnson Syndrome/toxic epidermal necrolysis: evidence from a meta-analysis of nine case-control studies. *Drug Res (Stuttg)* 2015; **65**(2): 107-11.
16. Chung WH, Chang WC, Lee YS, et al. Genetic variants associated with phenytoin-related severe cutaneous adverse reactions. *JAMA* 2014; **312**(5): 525-34.
17. Tassaneeyakul W, Prabmeechai N, Sukasem C, et al. Associations between HLA class I and cytochrome P450 2C9 genetic polymorphisms and phenytoin-related severe cutaneous adverse reactions in a Thai population. *Pharmacogenetics and genomics* 2016; **26**(5): 225-34.
18. Moon J, Park HK, Chu K, et al. The HLA-A*2402/Cw*0102 haplotype is associated with lamotrigine-induced maculopapular eruption in the Korean population. *Epilepsia* 2015; **56**(10): e161-7.
19. Chen CB, Hsiao YH, Wu T, et al. Risk and association of HLA with oxcarbazepine-induced cutaneous adverse reactions in Asians. *Neurology* 2017; **88**(1): 78-86.
20. Lv YD, Min FL, Liao WP, et al. The association between oxcarbazepine-induced maculopapular eruption and HLA-B alleles in a northern Han Chinese population. *BMC Neurol* 2013; **13**: 75.