Use of multinomial regression model to identify loci underlying diseases with variable age of onset

Sarnowski C (1,2), Dizier MH (1,3), Ahmed I (2,4), Margaritte-Jeannin P (1,3), Lathrop M (5,6), Demenais F (1,3), Bouzigon E (1,3) and the EGEA cooperative group
(1) INSERM, UMR 946, Paris, France
(2) Univ. Paris Sud, Paris, France
(3) Univ. Paris Diderot, Paris, France
(4) CESP, INSERM, UMRS 1018, Villejuif, France
(5) McGill Univ., Montréal, Canada
(6) CEA/CNG, Evry, France

Asthma is a heterogeneous disease and age of onset is one of the simplest features that can be used to differentiate asthma phenotypes. To characterize the genetic factors influencing asthma age-of-onset, we conducted a GWAS using a multinomial regression model applied to 750 asthmatics categorized according to their age-of-onset and 1,085 non-asthmatics from the French EGEA study with HapMap2 imputed data. Asthmatics were split into four specific age-of-onset sub-phenotypes: A) age-of-onset ≤ 4yrs (early-onset), B) 5-12yrs (before puberty), C) 13-20yrs (between puberty and adulthood) and D) > 20yrs (adult-onset). First, we applied an association test allowing heterogeneity of SNP effect between sub-phenotypes (Morris et al. Genet Epidemiol 2010) and detected 60 SNPs with P-value ≤10^{-5}.

Then, we tested whether these SNPs had a heterogeneous effect among the four sub-phenotypes. We identified 53 SNPs located in 16 regions with an interclass heterogeneity P-value ≤10^{-3}. Among these regions, six had intra-class association P-values ≤10^{-5}. We confirmed the specific association between 17q12-q21 genetic variants and early-onset asthma (P=10^{-6}) (Bouzigon et al. N Engl J Med 2008). We also detected five new regions with SNP effect restricted to one asthma age-of-onset sub-phenotype: 9q34 with phenotype A (P=5x10^{-6}), 3q25 with phenotype B (P=2x10^{-7}), 1p13-p12, 3p22 and 3q27-q28 with phenotype C (P≤3x10^{-7}). This analysis will be extended to GABRIEL Asthma consortium datasets. Thus, taking into account the age of onset in a multinomial regression framework can be a powerful approach to identify new loci underlying complex diseases.