Environment, genes and respiratory health in the SAP ALDIA cohort". 14.02.2006

Principal Investigator:

Nicole M. Probst-Hensch

Main objective

Numerous loci and candidate genes have beeri associated with asthma, accelerated lung function decline, and chronic obstructive pulmonary disease. Only few of the reported genetic associations have been replicated. The interaction of these gene variants with environmental and lifestyle factors has been little explored in the past. Yet, both asthma and COPD, are known to be influenced by a indoor and outdoor air pollution, active and passive smoking, as weIl as occupational exposures. The SAP ALDIA biobank with its detailed characterization of participants and its focus on environmental exposures offers an intemationally unique setting to evaluate the joint influence of environmental exposures and genetic factors on asthma, atopy, bronchial hyperresponsiveness, lung function, and COPD.

Methods

Study population. The population for the proposed study will included aIl SAP ALDIA participants at baseline and follow-up who consented to the genetic protocol and provided a valid DNA sample (N=6200).

Phenotype definition.Subjects having asthma will be defined in two ways. First, subjects reporting an attack of asthma during the last 12 months, or taking asthma medication presently. Second, subjects reporting a physician diagnosis of asthma. Bronchial hyperreactivity is defined as a 20% fall in FEVI from the highest FEVI measured during methacholine challenge. Annuallung function decline is defined as the difference between lung function at follow-up and lung function at baseline, divided by the number of years of follow-up. COPD is defined according to GOLD criteria by an FEVlIFVC ratio <0.70 (based on pre-bronchodilator spirometry).

Selection of Genes. A set of candidate genes has been identified on the basis of reported involvement in asthma, bronchial hyperreactivity, accelerated lung function decline, or COPD. Criteria for selection were reported epidemiological, biological and genetic relevance, minor allele frequency, and likelihood for interaction with environmental exposures (air pollutants, environmental tobacco smoke) collected in the context of the SAPALDIA cohort.

The genes considered included ADAM33, GRPA, CDI4, TLR4, FLAP, LOX-15, NATI, NAT2, SODl, SOD2, SOD3, and CAT.The genes that will be evaluated can be grouped into: - asthma candidate genes identified through genome-wide linkage studies - candidate genes involved in inflammatory/oxidative stress related processes - candidate genes involved in the metabolism of environmental agents

ln the genes selected, a finallist of 60 SNPs will be defined. We will select the minimum subset of SNPs to capture haplotype block diversity using information of the LD profile . obtained on a references set of population samples (Celera, Sequenon, Affymetrix, NCBI).

The inclusion ofhaplotype blocks covering the gene or region ofinterest should reduce the false negative results due to inadequate characterization of the linkage disequilibrium patterns and help to capture the relative risk of mutated genes.

Population stratification markers. A set of 20 SNPs randomly selected across the genome inc1uding the X chromosome and unrelated to health status will be genotyped for the assessment of population-stratification. The markers were selected specifically for European populations. Little is known about differences in the genetic background across different cultural and language regions in S witzerl and. The availability of population-stratification markers will facilitate all future assessments of gene-health associations in SAP ALDIA.

Genotyping. Genotyping will be based on the SNPlex assay developed by ABI; pyrosequencing technology (PSQ96) and Mass Array Technology (MALDI-TOFF) will be used as a genotyping method for some of the variants that cannot be inc1uded in the SNPlex assay. These efficient genotyping methods are not currently available at our own laboratory at the Institute of Medical Genetics in Schwerzenbach. They are implemented, though, at orie of the ECHRS (European Community Respiratory Health Survey) genotyping centre in Barcelona, and assays have alreadybeen set up for genotyping an asthma gene panel in this survey (Principal Investigator: Prof. Dr. M. Kogevinas, IMIM, Barcelona). SAP ALDIA used identical research instruments (questionnaires; spirometry) like the ECRHS study; the SAPALDIA Basel centre is also an ECRHS centre. For;genotyping of the SAPALDIA samples we have access to genotyping at the Fundacion Centro de Regulacion Genomica (Prof. X. Estivill/Dr. R. de Cid) on a fee for service basis. Given the comparability between ECRHS and SAP ALDIA, priority will be given to genes and SNPs genotyped in ECRHS which will allow for validation of each others results as well as for pooling of data, an increasingly important aspect of genetic epidemiology. '

Statistical Analysis. Previous to the main statistical analysis we will evaluate population stratification through the genotyping of 25 population-stratification markers. Indicator " variables for population strata derived from population-stratification markers will be inc1uded as covariates in all models.

Standard descriptive statistics will be performed for the case and control population, inc1uding allele frequencies, genotype frequencies and Hardy-Weinberg measure. Haplotype frequencies will be measured within each group of subjects using an ExpectationMaximization algorithm and Bayesian methods. Variability will be examined investigating differences in pairwise linkage disequilibrium and haplotype patterns between groups. Statistical packages used inc1ude Haploview and Phase.

The statistical analysis of genotype effects will be done intwo phases, separately for the different phenotypes. ln a first phase main effects of environmental and genetic factors will be evaluated. ln second phase, gene-envirohment and gene-gene interactions will be evaluated. To investigate SNP-disease associations we will use SNP allelic, genotypic and extended haplotypes data, in contingency tables andusing unconditionallogistic regression with odds ratios (OR) and 95% confidence intervals (95% CI) adjusted for age, sex, centre, and possible population-stratification marker. The statistical package STATA8.0 will be used. Given the large number of comparisoris that will be done and to minimise false positives, a Bayesian type approach will be followed based on the observed p-value, the power of the test and on the prior probability that the association is real. '

Additive and multiplicative relative risk models will be used to assess additive and multiplicative interactions, respectively. When the genotype and exposure are independent in the population, multiplicative interactions can be evaluated among cases only. This approach has great power than the conventional case-control approach, assuming that the independence assumption is correct. Therefore, we will use both, the conventional case-control and the caseonly approaches to assess the multiplicative interactions. The validity of the assumptions of independence between the genotype and exposure of interest will be evaluated in the control population.

Scientific Interest of the Study

The incidence ofboth, asthma and COPD is increasing worldwide. Understanding of disease etiology is a fundamental prerequisite for the development of novel preventive and therapeutic approaches. Large studies with detailed characteriziation of subjects and associated biobanks are needed for the investigation of gene-environment interactions. SAP ALDIA is one of the few international studies with substantial research potential in this area. The collaboration with ECRHS, which is comparable in design and instruments used, is essential from a genetic epidemiology perspective, because it offers the opportunity for replication of findings and pooling of data. An according collaboration is ongoing for the assessment of the association between genetic variants in TNF -alpha and asthma, where a SAP ALDIA/ECHRS abstract has been submitted for the A TS annual meeting 2008 and a manuscript is in preparation.

The project is based on the application of state-of-the-art methodology applied in genetics and epidemiology. The application of new high-throughput methods for genotyping within the context of a large epidemiological study will allow for availability of a large number of candidate SNPs.

Timetable'

January-March 2007:	Aliquoting and shipping of standardised DNA for the Barcelona laboratory
January 2007:	Finalize selection of candidate gene and SNPs
February-AprI 2007:	Design of assays (where not already available)
F ebruary- August 2007:	Finalize genotyping of SAPALDIA samples
September-December 2007:	Statistical analysis of results
January-March 2008:	Prepare scientific publications

Budget

Salary costs	no additional salary costs are needed; personnel for aliquoting of DNA, statistical analysis and publication writing is provided within the SNF-Grant 3347CO-I08796 (PI: Prof. Th. Rochat).
Genotyping	40'000.- Euro, fee for service for genotyping 85 SNPs total: 40'000 CHF from Nested Project Money 25'000 CHF from own funds (N. Probst)

TravellShipping of DNA	8'000.- CHF (shipping DNA; scientific meetings Barcelona group; attendance A TS annual meeting 2008)
Total Costs:	73 '000 CHF
SNF Nested Project Money	**48’000 CHF**