Application number/Title: 24299 - COPD follow up – access to genotype and phenotype information and DNA samples

Applicant PI: Professor Kari Stefansson

Application Institution: deCODE genetics ehf, Sturlugata 8, Reykjavik IS101, Iceland

Lead Collaborators: 1) Professor Ian Hall

Collaborating Institutions and Addresses: 1) University of Nottingham, Therapeutics & Mol Medicine, D Floor South Block, Queen's Medical Centre, Nottingham, NG7 2UH, United Kingdom

Funding body: deCODE genetics ehf

Keywords provided by the Applicant PI to describe the research project:

COPD, Emphysema, Bronchitis, WGS, association, rare

Application Lay Summary:

1a: deCODE's ongoing genetic study includes 6,000 COPD cases and 300,000 controls. Many sequence variants reported to associate with COPD or lung function replicate in our dataset. We have many novel variants associating with COPD, requiring follow up in a large dataset like the UKBiLEVE cohort, in particular rare variants identified by whole genome sequencing of the Icelanders. Our aim is to use existing genotypes or where necessary to undertake further genotyping of samples from UKBiLEVE subjects to confirm associations found in the Icelandic population, thereby establishing novel, robust associations between sequence variants and COPD or lung function.
1b: Identifying novel variants in the genome that associate with COPD or lung function will contribute to increased understanding of the disease and may provide new targets for development of improved treatments for the disease.

1c: Existing genotypes for UKBiLEVE subjects will be analysed for selected variants to confirm novel associations with COPD or its subphenotypes, emphysema and chronic bronchitis, or quantitative lung function traits, already

identified in the deCODE sample set. DNA from UKBiLEVE subjects will be genotyped for rare variants for which no genotypes exist in UK Biobank. To identify UKBiLEVE subjects with COPD, emphysema, chronic bronchitis or asthma, information on self-reported or HES diagnosis of those diseases would be needed.

1d: Genotype information for 48931 individuals from UKBiLEVE and DNA samples from a subset of 19017 individuals, defined as COPD cases or controls in the study, are requested. In addition, we request genotype information and DNA samples for with self reported, or registered diagnosis of, COPD, emphysema or chronic bronchitis for up to 7428 individuals, fewer if the selected groups overlap. In total, genotypes for 56359 and DNA for 26445 individuals are requested.

PROJECT EXTENSION APPROVED 26/10/2017
"Given the large range of physiological, environmental and genetic factors that may lead to the diagnosis of COPD, it is important to get as complete as possible picture of the respiratory health of the study subjects. This includes knowledge of the overlap of COPD with other respiratory diseases. deCODE has through collaboration with Landspitali, the National University Hospital of Iceland and major specialist clinics in Iceland received quality controlled hospital and specialist diagnosis data for all major respiratory/lung diseases, including asthma and related allergic and atopic conditions known to share etiology, pathology and/or genetic risk with asthma (conjunctivitis, atopic dermatitis, urticaria, anaphylaxis and angioedema), tuberculosis, sarcoidosis, bronchiolitis, idiopathic pulmonary fibrosis, cystic fibrosis, sleep apnoea, lung viral and bacterial infections (pneumonia) and lung cancer. We periodically perform GWAS analyses on these patient groups (sample size varies between 500 to 15,000 patients) against more than 200,000 controls, and seek other genetic datasets to follow up our discoveries in those studies.
We would therefore like to request an extension of scope of our application, to include follow up of deCODE discovery using the UK Biobank GWAS data and to perform GWAS meta-analyses with our Icelandic data on hospital and specialist clinics diagnoses of the full range of respiratory diseases. Specifically, we refer to hospital spell and episode data with episode/admission start and end date associated with as the main or secondary diagnoses any ICD10 codes starting with J (respiratory diseases) and other relevant allergic diseases: H10 (conjunctivitis), L20 (atopic dermatitis), L50 (urticaria) and T78 (anaphylaxis/angioedema), and in addition other diseases of the lung, codes starting with A15 (tuberculosis), C34 (lung cancer, with SNOMED codes or other histological information if available), D86 (sarcoidosis), E84 (cystic fibrosis), G47 (sleep apnoea).
In case the month and year of birth data already applied for would only cover the case and control groups defined in the main application, we apply for month and year of birth for all UK biobank participants.
With access to this data from UK Biobank we hope to be able to uncover new genetic associations with respiratory/lung diseases and linked conditions, both with each one in its own right and in the context of COPD development and risk."