

## **INFORMATION AND GUIDELINES FOR ACCESS TO DATA FROM THE WELLCOME TRUST CASE-CONTROL CONSORTIUM**

The Wellcome Trust case-control Consortium Data Access Committee will consider applications for access to genotype data generated as a result of the Consortium activities. Access to data will be granted to qualified investigator for appropriate use. A qualified investigator means a scientific researcher who is employed, or a student enrolled at, or legitimately affiliated with an academic, non-profit or government institution, or a commercial company. The decision as to who is or is not a qualified investigator is the primary responsibility of the Consortium Data Access Committee.

Access is conditional upon availability of data and agreement to abide by policies related to publication, data/sample disposal, ethical approval and confidentiality.

Access to data will be governed by the provisions laid out in the associated informed consent and the original research ethics committee approval for each case collection or control group.

The Consortium Data Access Committee is concerned only with access to the core, anonymised, genotype data and samples generated by this study. The only phenotypic information held by the Consortium is that which is implied by membership of a particular case or control group. The Committee will not consider requests for more detailed phenotypic information that is held by the principal investigators for the individual case collections. Access to this data would be by arrangement with the relevant principal investigator.

The Committee will consider applications that include collaborators, but each Institution must sign a separate Data Access Agreement. Should you wish to share the data with additional collaborators not previously approved, they must make a separate application for access to the Data.

Members of the Consortium will have early access to genotype data relating to their own collections, to control data, and to preliminary analyses, but will be bound by the same agreements as external users concerning the use of these data.

The Consortium Data Access Committee will keep confidential records of all applications, but may supply Consortium investigators with summary data of access requests to aid investigators in obtaining further funding.

### **1. Data Sets Available**

**a)** Genotype data is available from the following collections using the Affymetrix 500K SNP chip:

- 1,500 samples from the 1958 British Birth Cohort controls
- 1,500 samples from the UK Blood Service controls
- 2,000 samples from the following disease cases:
  - type 1 diabetes
  - type 2 diabetes
  - bipolar disorder
  - rheumatoid arthritis
  - coronary heart disease
  - hypertension
  - inflammatory bowel disease
- quantile normalised signal file data for control and disease samples listed above

**b)** Genotype data is available from the following collections tested on a custom Illumina chip containing approximately 15,000 non synonymous SNPs:

- 1,500 samples from the 1958 British Birth Cohort controls
- 1,000 samples from the following disease cases:
  - multiple sclerosis

- autoimmune thyroid disease
- ankylosing spondylitis
- breast cancer

Individual-level genotypes for all SNPs typed in the course of this study will be accompanied by:

- Chromosome position and strand information
- age at onset or collection (in 10 year age bands)
- broad geographical region of residence
- sex

The ethnicity of the participants in the 1958 British Birth Cohort was recorded by a school medical officer at ages 11 and 16, and self-declared by cohort members at age 33 and 43. The DNA samples issued to the WTCCC were restricted to those 97% of the cohort who are Caucasian. The samples tested are therefore estimated to be 99.8% Caucasoid.

The ethnicity of the UK Blood Service control group was self-declared as 99% Caucasian.

The genotype data will be available as:

- anonymised genotype data analysed using the BRLMM and CHIAMO algorithms
- filtered with quality control checks performed

The Data, along with the associated information, will be available on a password-protected website at the Wellcome Trust Sanger Institute.

**The data available from the 1958 British Birth Cohort, or any part thereof, cannot be used for the creation of products for sale, or for any commercial purpose.**

## 2. Availability of DNA samples

The mechanism for distribution of the DNA samples from the 1500 UK Blood Service Control group has yet to be determined. Once finalised, details will be posted on the WTCCC website.

Access to UK Blood Service Control Group samples will be conditional upon the investigator signing an agreement to submit genotype results obtained from the samples to the Consortium for incorporation into the central database. This will enhance the scientific value of the resource and promote further genetic research.

DNA samples from the 1958 British Birth Cohort can be obtained from the Genetics Advisory Group, a sub-group of the 1958 Oversight Committee which can be found at: [www.b58cgene.sgul.ac.uk/application.php](http://www.b58cgene.sgul.ac.uk/application.php)

Note that the disease DNA samples used in the WTCCC are not available through this mechanism. Some samples may be available through the specific disease investigators and/or their research sponsors, and you should contact them directly to explore the options. The contact details for principal investigators can be found at: [www.wtccc.org.uk](http://www.wtccc.org.uk)

## 3. Application Procedure

The Consortium Data Access Committee will accept applications for access to data at any time. Applications for access to the data should be submitted using the form "Application for Access to Genotype Data".

The Consortium Data Access Committee will consider straightforward applications on a rolling basis and aim to provide a decision within 2 months of receipt. Applications which require further consideration will be referred to the next available Access Committee meeting for discussion.

You will be informed if your application has been referred and you should receive a decision within 2 weeks following the meeting at which your application is considered.

**Applicants must address all questions on the form and include a reference list.** Application forms will be returned if submitted with incomplete or unclear information.

Completed application forms should be submitted electronically to the following email address: [cdac@wellcome.ac.uk](mailto:cdac@wellcome.ac.uk)

Any queries regarding access procedures or completion of the forms should be sent to this address.

The application should be accompanied by a signed Data Access Agreement. This must include a signature from the Head of Institute or person with the appropriate institutional or administrative authority. A signed hard copy of the agreement should be posted or faxed to:

Dr A Duncanson  
Molecules, Genes and Cells  
The Wellcome Trust  
215 Euston Road  
London NW1 2BE  
Fax: 020 7611 8352

Successful applicants for access to data will be designated as an "Approved User" and will receive a username and password to enable access to the WTCCC database.