Rare Genetic Variants in Health and Disease:
10,000 Genomes - The UK10K Project

Dawn Muddyman1 on behalf of the UK10K Consortium2

1 Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1HH, UK. Email: dm11@sanger.ac.uk
2 www.uk10k.org

4,000 whole genome samples sequenced
Average 6x depth

>5,500 whole exome samples sequenced
Average 72x depth

2,000 ALSPAC
2,000 TwinsUK

3,000 Neuro disorders
2,113 Schizophrenia
887 Autism

>1,500 Obesity
1,000 SCOOP study
430 Generation Scotland
68 TwinsUK

1,000 Rare disorders
125 Ciliopathies
125 Coloboma
125 Severe insulin resistance
125 Congenital heart disease
125 Familial hypercholesterolaemia
125 Neuromuscular disease
125 Learning disability (FIND study)
125 Thyroid disorders

SNP, indel and larger structural variants associated with phenotypic traits to determine which were linked to disease and which were not

Candidate variants identified in subjects with extreme disease phenotypes

Low frequency variants (down to 0.1% allele frequency) imputed into studies of non-sequenced individuals with GWAS data to increase study power

BAM and VCF files deposited in the European Genome-phenome Archive to create a managed-access data resource

The data generated by UK10K is a valuable resource for research into human genetics, and it is hoped the discovery of rare and low frequency disease-causing variants will lead to further insight into the diagnosis and treatment of disease.