**Zooming in on functional variants by combined transcriptome and genome sequencing**

Peter A.C. ’t Hoen1, Michael Sammeth2,3, Marc R. Friedländer2, Jean Monlong3, Manuel A Rivas4, Natalja Kurbatova5, Mar Gonzalez-Porta5 Matthias Barann6, Thomas Wieland7, Jonas Almlöf8, The Geuvadis Consortium, Ralf Sudbrak9, Philip Rosenstiel6, Roderic Guigó3, Ivo G. Gut2, Xavier Estivill3, Emmanouil T Dermitzakis10, Tuuli Lappalainen10

1 Department of Human Genetics, Leiden University, the Netherlands

2 Centro Nacional d'Anàlisi Genòmica, Barcelona, Spain

3 Center for Genomic Regulation, Barcelona, Spain

4 Wellcome Trust Centre for Human Genetics, Univeristy of Oxford, UK

5 EMBL-EBI, Hinxton, UK

6 Institute of Clinical Molecular Biology, Christian-Albrechts-University, Kiel, Germany

7 Institute of Human Genetics, Helmholtz Zentrum München, Germany

8 Department of Medical Sciences, Uppsala University, Uppsala, Sweden

9 Max Planck Institute for Molecular Genetics, Berlin, Germany

10 Department of Genetic Medicine and Development, University of Geneva, Switzerland

Understanding functional effects of genetic variants is one of the biggest challenges in human genomics. Towards this goal, we sequenced mRNA and small RNA from lymphoblastoid cell lines of 465 individuals from 5 populations of the 1000 Genomes Project. Sequencing was distributed over seven different centers. The variation between laboratories appeared to be considerably smaller than the already limited biological variation. We found extremely widespread regulatory variation, with genetic variants associating to expression levels of 7825 genes, transcript structure of 639 genes, as well as expression of 60 miRNAs. Allele-specific transcription analysis allowed us to characterize also rare regulatory effects and underlying genetic variants. Importantly, genome sequencing data combined with functional annotation of the genome allowed us to infer putative causal regulatory variants for 57-75% of eQTL genes. We employed this to predict causal variants for 91 disease-associated loci, and also characterize transcriptome effects of 1688 loss-of-function variants, thus linking our discoveries to genetic associations to human disease. Altogether, the integration of genome sequencing and functional genomics data in this study takes us towards understanding and predicting molecular genomic effects of causal functional variants in the human genome.