

## Which genome variants matter?

February 8 2007

Findings published today in *Science* will accelerate the search for genes involved in human disease. The report provides a first genome-wide view of how the unique composition of genetic variation within each of us leads to unique patterns of gene activity.

By defining those genetic variants with a biological effect, the results will help prioritise regions of the genome that are investigated for association with disease. This is an important step to understanding links between genes and disease for individuals, and across populations.

The Human Genome Project gave us the instruction manual for building a human. The HapMap and Copy Number Variation (CNV) Projects developed indices of where to find differences in the manuals of different people. One of the challenges for research into variation and disease is that most variants have no consequence for our wellbeing.

The new study gives a global view of the consequences of those differences for gene activity. The work shows that activity of more than 1000 genes is affected by sequence variation and is the first map of human populations that identifies the most important fraction of DNA variation, that which directly affects gene activity.

The research was led by scientists from the Wellcome Trust Sanger Institute, together with colleagues from the University of Cambridge, Hospital for Sick Children/University of Toronto and Harvard Medical School/Brigham and Women's Hospital.



Using the HapMap series of cell samples from four populations, they measured the activity of more than 14,000 genes in cells grown in culture. The cell samples provide a snapshot of genetic activity in one cell type. The activity of each gene was then correlated with genetic variation nearby, as defined by the HapMap, an index of single-base changes (single nucleotide polymorphisms, or SNPs) and the new index of copy number variants (CNVs).

"We've been able to look back into our history and find changes that are older and likely to be shared among populations," explained Dr Manolis Dermitzakis, senior author and Project Leader at the Wellcome Trust Sanger Institute. "But we also find many that are newer and less widespread.

"These are part of our recent evolution and a step along the way to understanding the origin and personal consequences of genetic change, not least for our wellbeing. This is a first generation map of biologically important DNA sequence variation"

The understanding of the genetic basis of gene activity will help medical research to provide individuals with information about their personal predisposition to disease.

The study was a massive undertaking: it included HapMap genotype data on 700,000 SNPs located close to genes, as well as 25,000 sites interrogated for potential structural variation to examine copy-number differences, looking at the activity of 14,000 genes in 210 unrelated individuals.

SNP and CNV variation correlated with altered activity in almost 900 and 240 genes, respectively. The HapMap has been invaluable in detecting variants involved in many diseases and these results suggest that the CNV index will prove similarly useful.



"The remarkable finding was that there is such little overlap in the genes found by using the two indices," commented Dr Matthew Hurles, also a leader of the project at the Wellcome Trust Sanger Institute. "Only about 10% of the activity variants associated with a CNV were also associated with a SNP.

"This suggests that we must include CNV studies in our searches for genetic variation associated with disease or we will be missing a lot of the important genetic effects."

The results show that at least 10-20% of heritable variation in gene activity is due to CNVs. The team found associations that included previously known examples, such as UGT2B17, which has been associated with prostate cancer, proving that the new approach works well.

They also showed for the first time that activity of other genes, located close to UGT2B17, was affected. Finding other effects in this way will enhance the search for critical genes within a region of genetic possibilities.

Some associations were not found in all four populations, two-thirds (CNVs or SNPs) being found in only one population. A gene implicated in Spinal Muscular Atrophy showed an association in three populations, but not in Yoruba from Ibadan, Nigeria. Understanding population differences can help us understand our history.

Variation in copy number can affect gene activity by altering the 'dose' of a gene, by disrupting the active parts of a gene that contain the code for protein, or by disrupting the regulatory regions of the genome that control gene activity - the on/off and dimmer switches in our genome.

"Although the simplest model for a CNV affecting gene activity is where



the variant is a deletion of a gene or part of a gene, we found examples where activity is affected from a distance," commented Barbara Stranger, first author and post-doctoral fellow at the Wellcome Trust Sanger Institute. "This may occur when the CNV reduces the effectiveness of a region that works to switch the genes on or off."

Source: Wellcome Trust Sanger Institute

Citation: Which genome variants matter? (2007, February 8) retrieved 28 April 2024 from <u>https://medicalxpress.com/news/2007-02-genome-variants.html</u>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.