

Meeting report

## The future is genome-wide

Samuel Deutsch\* and Alexandre Reymond†

Addresses: \*Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva CH-1211, Switzerland. †Center for Integrative Genomics, University of Lausanne, Lausanne CH-1015, Switzerland.

Correspondence: Alexandre Reymond. Email: alexandre.reymond@unil.ch

Published: 24 August 2006

*Genome Biology* 2006, 7:324 (doi:10.1186/gb-2006-7-8-324)

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2006/7/8/324>

© 2006 BioMed Central Ltd

---

A report of the annual meeting of the European Society of Human Genetics, Amsterdam, 6-9 May 2006.

---

More than 1,700 human geneticists from 59 countries congregated in Amsterdam in May for this year's meeting of the European Society of Human Genetics, which mainly focused on the use of post-genome analysis tools to dissect the causes of and mechanisms governing complex traits. Many of the exciting studies presented were based on two technologies: array-based methods for genome-wide genotyping or technologies for high-density comparative genome hybridization (CGH). A highlight of the meeting was the keynote lecture by Nobel laureate Sydney Brenner (Salk Institute for Biological Studies, La Jolla, USA) on 'humanity's genes', which focused on the challenges we face in transforming the information from the human genome into concrete benefits for our societies.

### Whole-genome association studies of complex phenotypes

This year has seen the success of several whole-genome association studies using genotyping for single-nucleotide polymorphisms (SNPs) to identify genes responsible for some common complex phenotypes for both discrete and quantitative traits. A plenary lecture by Kari Stefansson (deCODE Genetics, Reykjavik, Iceland) highlighted the tremendous potential of this approach. Several examples were discussed in which new genes have recently been identified using a combination of linkage and association analysis approaches. One example is a locus on human chromosome 8p12 that confers susceptibility to schizophrenia. Although nucleotide variation around the *NRG1* gene has been known to be associated with schizophrenia for the past 4 years, the mechanism of action of the associated SNPs, located in noncoding regions 5' to the gene, has

remained unclear. Recent evidence strongly suggests that these variants might influence the level of *NRG1* expression. Stefansson suggested that many SNPs involved in the etiology of complex phenotypes are likely to affect gene expression or splicing, and that these variants are under strong selective pressure. A second, more recent, example presented by Stefansson concerns the genetics of myocardial infarction, where two genes involved in the production of the pro-inflammatory molecule leukotriene B<sub>4</sub> have been identified as conferring an increased risk of this disorder. In particular, a haplotype in the leukotriene A<sub>4</sub> hydrolase gene (*LTA4H*) was shown to confer a modest risk in caucasians (relative risk compared with the general population (RR) = 1.16), but a much higher risk in African Americans (RR approximately 3.5). The associated haplotype is likely to confer risk to myocardial infarction through an upregulation of the leukotriene pathway. As in the case of *NGR1* in schizophrenia, quantitative rather than qualitative changes seem to be important.

On the closely related topic of gene expression variation, Harald Göring (Southwest Foundation for Medical Research, San Antonio, USA) discussed the identification of genetic determinants of gene expression, which are excellent candidates for involvement in complex phenotypes. Göring and his collaborators used the Illumina expression system (Sentrix Human-6 gene bead arrays) to study the variation in expression of around 20,000 transcripts in 1,240 individuals of Mexican-American origin. He reported that for most genes, transcript levels vary significantly between individuals, and in virtually all cases the variation in gene expression is genetically controlled (as determined by a significant heritability). Quantitative linkage analysis was performed using the SOLAR computer package, leading to the identification of cis-regulatory loci in 15% of cases. Although similar studies have been performed in the past, this study stands out in terms of the increased statistical power to detect genetic effects.

An interesting example of a genome-wide analysis for a quantitative trait was presented by Arne Pfeufer (Technical University Munich, Germany) in which the loci influencing QT interval (the time the heart takes to recover from a ventricular beat) were identified. Individuals with either too high or too low QT interval are at risk of sudden cardiac death. A whole-genome association study using the Affymetrix 100K chip was carried out taking 100 individuals with extreme QT values at each end of the spectrum. Although no significantly associated SNPs were identified at stage 1 of the study after correction for multiple testing, the top 10 SNPs and additional candidate polymorphisms were genotyped on a second cohort of people. This identified a significant association of SNPs in the *NOS1AP* gene with QT interval. The association was further validated in additional cohorts from Germany and the US. Interestingly, the SNPs with the highest association were located in a conserved noncoding region 5' to the *NOS1AP* gene that has a potential regulatory role.

Another area of exciting research concerns the genetics of infectious diseases in human populations. Adrian Hill (Oxford University, UK) reviewed the major advances in the field, such as the discovery of polymorphisms within the *NRAMP1* gene that affect susceptibility to tuberculosis, and the protective effect of polymorphic deletions within the *CCR5* gene against HIV infection. An interesting recent discovery concerns the *PTPN22* gene encoding a protein phosphatase, and in particular the Arg620 to RP variant. This is a gain-of-function polymorphism that increases the protein's phosphatase activity in T cells and downregulates T-cell responses. This variant has been previously associated with a number of autoimmune diseases such as rheumatoid arthritis and type 1 diabetes. Hill reported that the same genetic variant is also associated with susceptibility to invasive pneumococcal infection. This link between autoimmune disease and susceptibility to infection suggests that other alleles that predispose to autoimmune disorders might play a role in modulating pathogen-host interactions. Strong emphasis was also placed on the role of the Toll-like receptor (TLR) signaling pathway in immune regulation and susceptibility to disease. In particular, coding variants in the gene for TLR2 have recently been shown to be associated to increased susceptibility to tuberculous and lepromatous leprosy in different populations. Other members of this pathway are thus excellent candidates for further study in both mouse and human systems for roles in modulating susceptibility to infectious disease.

### The coming of age of comparative genomic hybridization

The technique of comparative genomic hybridization (CGH) was developed to detect subtle cytogenetic alterations throughout the genome in a high-throughput manner. Several emerging CGH technologies are now being used to

identify copy number alterations of increasingly smaller regions. Joris Veltman (Radboud University, Nijmegen, The Netherlands) has tested the 'practical' resolution of these new CGH platforms (the Affymetrix 100K SNP array and the Nimblegen 385K CGH array) against that of a homemade 32K array based on bacterial artificial chromosome clones (BACs). Not surprisingly he found that fewer BACs than oligos were necessary for the automatic detection of an imbalance, as these interrogate larger regions; however, with 10 times as many points interrogated, the Nimblegen array reached a resolution of 64 kb, twofold and sixfold better than the resolution of the Affymetrix and BAC arrays, respectively.

Mental retardation and global developmental delay are relatively common disorders, each having a prevalence of 2-3% in the general population, with chromosome abnormalities as the single most common cause. However, the detection rate of chromosome abnormality in surveys of severe mental retardation/global development delay is only 15-40% with traditional techniques. Reasoning that cryptic rearrangements might be uncovered with higher-resolution methods, a number of laboratories have assayed the DNA of patients with idiopathic mental retardation for unbalanced chromosomal rearrangements. Orsetta Zuffardi (University of Pavia, Italy) presented results obtained with 227 DNAs, while Martin Poot (University Medical Centre Utrecht, The Netherlands) and Evan Eichler (University of Washington School of Medicine, Seattle, USA) discussed their analysis of 208 and 291 DNAs, respectively. All found that approximately 12% of the samples showed *de novo* imbalances ranging from 0.1 to 13 Mb in size. Using a 1 Mb resolution homemade BAC array, Bernard Thienpont (University of Leuven, Leuven, Belgium) has similarly analyzed a cohort of patients with a congenital heart defect of unknown cause and either global developmental delay and/or major malformations. He found that 18% of the cohort had genomic rearrangements, of which 11% were relevant *de novo* rearrangements.

Only two recurrent rearrangements were identified by the above laboratories. A deletion at 17q21.3 that could explain as much as 1% of idiopathic mental retardation was identified by Eichler. Interestingly, this region was recently shown to be under positive selection in the European population. A 400 kb duplication in the Xq28 band, which contains the *MECP2* locus, was unveiled by Zuffardi. In a separate study, Guy Froyen (Flanders Interuniversity Institute for Biotechnology, Leuven, Belgium) found a duplication of the same *MECP2*-containing region in a family with severe mental retardation, and subsequently identified three more patients with a similar dosage imbalance. Even though the duplications in different individuals with mental retardation are not of the same length (ranging from 7 to 50 genes), a comparison of the new results with the literature showed that the minimal duplicated region associated with mental retardation includes a set of six genes comprising *MECP2*. The contribution of *MECP2*

and the other genes in the region is currently under investigation. A third recurrent rearrangement in people with mental retardation was recently reported in the literature (a microdeletion on 2q23.1-q23.2 in 3 out of 161 patients).

These reports emphasize the fact that some interstitial aneusomies (that is, the deletions or duplications within chromosome arms) might turn out to be relatively frequent. Zuffardi also presented the results of an analysis of DNAs of patients carrying apparently balanced reciprocal translocations or complex rearrangements, and found that a large proportion of them, 70% and 90%, respectively, also carried previously unidentified insertions and/or deletions. In the case of the translocations, imbalances were detected at the breakpoints and elsewhere in the genome at approximately equal rates. This suggests that care should be exercised when investigating apparently balanced rearrangements, which ideally require a higher-resolution assessment of potential gains or losses of genetic material.

In addition to genomic imbalances that result in pathological phenotypes, a large number of copy-number polymorphisms (CNPs) not associated directly with a phenotype are being identified and studied. For example, Zuffardi reported more than 200 large segments (0.1 to 2 Mb) present in different copy numbers in healthy individuals, while Poot noted copy-number variation in 333 autosomal loci. These loci seemed not to be associated with G-bands (late replicating; A/T-rich DNA) or R-bands (early replicating; G/C-rich DNA), but were often flanked by segmental duplications. Eichler reported similar numbers - 257 new CNPs in 344 normal samples - and demonstrated Mendelian transmission of some of these CNPs, where the number of copies of a locus in an individual was either the sum or the average of the number of copies identified in their parents. The heritability calculated for 25 common CNPs was approximately 100%. In addition, many CNPs are known to be in strong linkage disequilibrium with flanking SNPs, suggesting that they are mostly ancient polymorphisms rather than recurrent events.

Lisenka Vissers (Radboud University, Nijmegen, The Netherlands) focused her attention on 12 known CNPs present in at least 5% of the population, and measured their copy numbers in 309 individuals of Dutch, Finnish, Turkish, Indonesian and Pygmy descent by multiplex ligation-dependent probe amplification (MLPA). A third of these CNPs showed no variation in this sample, suggesting that the CNP databases might be polluted by false positives. Other CNPs showed moderate to extensive variation, and many were identified in all the ethnic groups studied. Eichler noted that his team found few population-specific CNPs. However, Vissers found that the median number of copies was variable in different ethnic groups.

Over the next year we expect to see a dramatic increase in the number and quality of whole-genome association studies,

with the identification of many functional polymorphisms associated with common complex traits. It will be interesting to see how many of these polymorphisms are localized in coding compared with noncoding regions. We anticipate that the next big challenge will be the development of statistical methods able to dissect interactions between different variants as well as environmental effects. Another major hurdle will be understanding the possible role of CNPs in complex phenotypes, and we look forward to hearing about progress in this and other fields at next year's meeting.