Meeting report

Human genetics moves from clinic to bench - and back

Miroslava Ogorelkova* and Xavier Estivill*†*

Addresses: *Genes and Disease Program, Center for Genomic Regulation (CRG) and †Department of Experimental and Life Science, Pompeu Fabra University (UPF), Barcelona Biomedical Research Park, Barcelona, 08003 Spain. †National Center of Genotyping (CEGEN), Barcelona, 08003 Spain.

Correspondence: Miroslava Ogorelkova. E-mail: mira.ogorelkova@crg.es. Xavier Estivill. E-mail: Xavier.estivill@crg.es

Published: 31 August 2005

Genome **Biology** 2005, **6:**343 (doi:10.1186/gb-2005-6-9-343)

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2005/6/9/343

© 2005 BioMed Central Ltd

A report on the European Society of Human Genetics Conference 2005, Prague, Czech Republic, 7-10 May 2005.

As in previous years, this year's European Society of Human Genetics Conference covered a broad and diverse range of topics in clinical and basic human molecular genetics, addressing the interests of clinicians, cytogeneticists, molecular biologists, population geneticists and bioinformaticians. We will focus here on some of the outstanding presentations on single-gene and polygenic diseases, high-throughput technologies and data analysis, the molecular genetics of longevity and ageing, functions of non-coding DNA, segmental duplications, and diagnostic tools and development of therapies.

From single-gene diseases to complex disorders

A number of novel genes involved in the variant pathology of disorders with known single-gene genetic determinants were reported. Regina Bendix-Waltes (Hannover Medical School, Hannover, Germany) reported that a deficiency of the DNArepair protein Rad50, as the result of a hypomorphic germline mutation in the RAD50 gene, caused a variant form of the Nijmegen breakage syndrome, a disease usually caused by mutations in another DNA-repair gene and characterized by chromosomal instability. Anna Benet-Pagès (Institute of Human Genetics, Munich, Germany) reported a homozygous missense mutation in the FGF23 gene as the cause of a variant of familial tumoral calcinosis in which the patient has abnormally low serum phosphate levels rather than the usual elevated serum phosphate. FGF23 encodes a putative circulating fibroblast growth factor that promotes phosphate excretion. Anne Goriely (Weatherall Institute of Molecular Medicine, Oxford, UK) reported that germline mutations in the fibroblast growth factor receptor 2 gene (FGFR2) that lead to amino-acid substitutions conferring gain of function are under positive selection in human spermatogonia. This explains the high prevalence of paternal *de novo* mutations leading to Apert syndrome, which is caused by mutations in *FGFR2*.

Discussions on complex disorders - those resulting from variants in several genes involved in interacting pathways focused on recent progress in understanding inflammatory and immune disorders. Susceptibility to asthma, for example, depends on variation at an unknown number of genetic loci. Among the more than 200 significant genetic associations reported for asthma, less than one third have been replicated, and those are associated with low relative risk. Nevertheless, several new genes shown to be associated with asthma - ADAM33, PHF11, DPP10, DRPR3, and GPRA - should provide new approaches to understanding important aspects of the disease. Juha Kere (Karolinska Institute, Stockholm, Sweden) described the work of his group on GPRA, which encodes the G-protein-coupled receptor for asthma susceptibility, for which no ligand has yet been identified. A distinct distribution of different isoforms of this receptor between bronchial biopsies from healthy and asthmatic individuals has been found; moreover, the analysis of a 70-kb haplotype block in the GPRA gene in a sample of children from the Swedish birth cohort study (BAMSE) has shown an association with sensitization, childhood allergic asthma and allergic rhinoconjunctivitis.

John Todd (University of Cambridge, UK) reviewed progress in understanding the genetic basis of type 1 diabetes (insulindependent diabetes) over the past few years. This type of diabetes arises from the autoimmune destruction of insulin-producing beta cells in the pancreas. Five susceptibility loci for type 1 diabetes have already been identified (HLA class II genes, the insulin gene (INS), CTLA4, CD25 and PTPN22). About 50 genes have been tested for association

with the disease, representing a small proportion of the genes that participate in the disorder, but Todd pointed out that most reported positive associations are false positives due to small samples, genotyping problems and population structure, among other things. The HLA locus is the major region associated with type 1 diabetes, representing about 40% of familial clustering, with odds ratios of between 4 and 20 (odds ratios indicate an increased likelihood of getting the disease compared with the general population). Particular variants of the HLA locus are almost necessary for the disease, but it is far from sufficient, with a positive predictive value of about 5%. To identify the other genes involved, thousands of samples and controls will have to be studied. The use of the British Birth Cohort collection in gene-association studies has identified 8,000 cases of type 1 diabetes, allowing Todd's group to confirm the role of CTLA4 in disease susceptibility, with an allele specific 3' untranslated region (UTR) giving an odds ratio of 1.24. Todd also reported the finding that 11% of pediatric cases of type 1 diabetes have autoantibodies to the autoimmune thyroid disease associated antigen thyroid peroxidase, and that the T-cell immunoregulatory CTLA4 gene is involved in only a subgroup of patients.

Volume 6, Issue 9, Article 343

Genome-wide studies utilizing DNA microarrays to study single-nucleotide polymorphisms (SNPs) in a large number of cases and healthy controls have begun, using 10,000-plex SNP assays produced by Parallele and Affymetrix. The work has been organized in several phases, starting with 1,000 cases and 1,000 controls. SNPs with minimum allele frequencies of greater than 10% are chosen to begin with, and the study will then move towards those SNPs showing the most promising associations. One problem the investigators have encountered is the peculiar stratification structure of the British population for several markers, reflecting the Celtic, Scandinavian and Anglo-Saxon origins of the population. In phase I of the study, they also encountered problems due to the lack of normalization of the DNA samples, so that the same protocols and the same labs should be mandatory for these investigations. With regard to the SNPs selected for analysis, Todd pointed out that Tag SNPs (reference SNPs for a given genomic region) are still in their infancy, which means an important loss of potential information from the current studies. A collaborative Consortium for Cases and Controls is also being set up in the UK, with support from the Wellcome Trust, with the aim of studying at least seven complex disorders in the first round of funding.

Chronic inflammatory disorders such as Crohn's disease, eczema, asthma, and psoriasis are triggered by environmental factors acting on the background of susceptibility genes with a polygenic action. Stefan Schreiber (Institute for Clinical Molecular Biology, University Hospital of Schleswig-Holstein, Kiel, Germany) reviewed the current knowledge of the genetics of Crohn's disease, which is a paradigm for inflammatory disorders of the epithelial barriers between the body and the 'outside world'. Inflammatory bowel disease (IBD), of which Crohn's disease is a form, is one of the complex disorders for which we have gained some basic knowledge through genetics and genomics. Three disease genes have been identified for IBD: CARD15 (NOD2), DLG5 and OCTNA12. CARD15, a homolog of CARD4 (NOD1), which is involved in responsiveness to intracellular bacterial lipopolysaccharides, was the first Crohn's disease gene to be identified by positional cloning. Mapping using the two-locus transmission disequilibrium test (TDT; which detects linkage of a major trait to a tightly linked marker, thus avoiding the effect of population stratification in European IBD patients) identified a disease-associated haplotype block in a suggestive linkage region on chromosome 10g, and has identified DLG5 in this region as a likely susceptibility gene for IBD. In contrast to the strong genetic effect of CARD15, where the associated variant confers a relative risk of disease of more than 40 for homozygotes and compound heterozygotes, the risk associated with DLG5 is much smaller (odds ratio of around 1.5). The mechanism by which variants in the third IBD susceptibility gene, for OCTNA12, which encodes an organic cation transporter, contribute to the disease is still unknown. One interesting hypothesis Schreiber put forward was the idea that several diseases that are classified as distinct entities, such as psoriasis, asthma, periodontitis, and sarcoidosis, among others, could all be included in a general category of inflammatory barrier disorders. The genetic dissection of these diseases at the phenotypic level and the use of genomic and metagenomic approaches should mean rapid progress.

http://genomebiology.com/2005/6/9/343

There were several other presentations on polygenic disorders. Emma Banfield (Wellcome Trust Centre for Human Genetics, Oxford, UK) reported a multivariate linkage analysis of original data from a whole-genome scan for genetic associations with specific language impairment. Using this approach, she and her colleagues detected three new possible quantitative trait loci involved in these conditions on chromosomes 4, 5, and 10, and investigated the relationship between the phenotypes influenced by the previously identified loci SLI1 and SLI2. Silvia Paracchini (University of Oxford, UK) reported that a previously identified 70 kb risk haplotype for dyslexia on chromosome 6p22 is associated with a reduction in transcription of KIAA0319 gene of about 40%. To analyze the allele-specific level of expression of genes within the candidate region, Paracchini and her colleagues tested lymphoblastoid and neuroblastoma cell lines heterozygous for the risk haplotype on the Sequenom MassARRAY platform for measuring gene expression. Cisca Wijmenga (University Medical Center Utrecht, The Netherlands) reported novel loci for celiac disease on 6q, 9p and 19p, identified by the sib-pair approach. The 6q locus might represent a more common autoimmune locus. And finally, one of us (X.E.) described data from our group on the association of anorexia and bulimia with variants in the genes for brain-derived neurotrophic factor (BDNF) and its receptor (NTRK2), defining risk and protective haplotypes for

NTRK2, as well as high BDNF levels in the blood of patients with eating disorders.

High-throughput technology and data analysis

On the high-throughput front, a novel array-based wholegenome genotyping assay that effectively enables unlimited SNP genotyping from a single sample preparation was reported by Kevin Gunderson (Illumina, San Diego, USA). This is accomplished by hybridizing the product of a singletube whole-genome amplification reaction to arrays of 50-mer probes and conducting an array-based allele-specific primer extension assay. Using Illumina's bead-based high-density array platform BeadChip, over 100,000 assays for SNPs located in exons can be combined on a single chip. Siv Fokstuen (University Hospital of Geneva, Switzerland) described the applicability of high-throughput resequencing for the diagnosis of high-prevalence disorders. The authors have developed a CustomSeq Resequencing Array (Affymetrix) with 30,000 probes that enables the rapid molecular diagnosis of hypertrophic cardiomyopathy. The array comprises all coding exons, splice-site junctions and known promoters of 12 genes known to be mutated in the disease; more than 90% of all mutations reported to date were detected. The complexity of the data obtained from genome-wide studies has meant the continual development of new methods for analyzing it. Geraldine Clarke (Wellcome Trust Centre for Human Genetics, Oxford, UK) presented a model-free multipoint method based on dense sequence-polymorphism data from parent-offspring trios (two parents and one child) to estimate recombination rates between adjacent markers. With dense maps of markers on trios, the effects of linkage disequilibrium and linkage can be separated, allowing estimation of recombination rates in a model-free setting which is the basis for the presented multipoint method.

Salma Kotti and her colleagues (Inserm U535, Villejuif, France) reported the use of TDT to analyze a suggested interaction between the HLA-DRB1 and CTLA4 genes that are involved in predisposition to multiple sclerosis. The effect of each gene was confirmed, but there was no evidence of any interaction between them. When the homogeneity transmission test was applied (that is, the transmission rate of one locus from heterozygous parents is compared between positive or negative for the other candidate locus) however, a significant difference in transmission was observed, suggesting an interaction. Kotti explained this disagreement by showing that, unlike the homogeneity transmission test, the TDT (involving two loci) is not affected by the presence of population stratification. Thus, she suggested, the homogeneity transmission test may sometimes lead to a false conclusion of interaction.

Molecular genetics of longevity and ageing

Identifying genes controlling human longevity is of increasing interest in the light of the general increase in life span in developed countries. In his talk on methods of identifying genes for human longevity, Rudi Westendorp (Leiden University Medical Hospital, The Netherlands) suggested that although 25% of the variation of human life span is explained by genetic factors, and evolutionarily conserved mechanisms are prime candidates for influencing longevity, ageing is under unintended genetic control. There will be positive selection of haplotypes associated with fertility rather than longevity, as evolutionary theory predicts that a genetic predisposition to fertility has effects on survival, and many animal models for longevity show decreased fertility. He pointed out that the search for genetic loci that explain the variability in human longevity has peculiar characteristics. Linkage studies are complicated, as some individuals are too young to determine their phenotypic status while others, whose status is defined, have already died. Association studies suffer from a lack of equilibrium, as an agematched control group is unavailable and use of a younger control group could introduce a bias towards selection for fertility over two generations. Westendorp presented two alternative approaches: investigating linkage in an affected sib-pair design, and assessment of association in a classic prospective follow-up study.

Martin Holzenberger (Hôpital Saint-Antoine, Paris, France) discussed signaling by insulin-like growth factor 1 (IGF-1) signaling and longevity. The insulin/IGF-1 signaling pathway has been identified as a major evolutionarily conserved player in life-span regulation. He described work from his lab showing that complete suppression of the mouse IGF-1 receptor (IGF-1R) is not compatible with life, but heterozygous knockout IGF-1R+/- mice have a significantly increased life span, by 33% in females and 16% in males, and are highly resistant to oxidative stress. IGF-1R+/- mice have normal energy metabolism, nutrient uptake and physical activity, and no change is seen in their fertility and reproduction. Analysis of their rate of growth suggests that significant changes in response to IGF-1R deficiency occur during short time periods. Holzenberger also pointed out a link between aging and longevity and growth hormone signaling. Deficiency of growth hormone in mice leads to a significantly increased life span associated with reduced fertility or infertility and a reduction in size and weight, suggesting that the somatotrophin-growth factor pathway is connected to the regulation of fertility, reproduction and ageing. Both growth hormone and IGF-1 are promoters of somatic growth and development; at the level of IGF-1R, however, fertility, growth and longevity phenotypes dissociate.

A number of genetically determined syndromes can be associated with premature or accelerated aging, or progeria. Nicholas Levy (Faculté de Médecine Timone, Marseille, France) discussed the involvement of lamins, which are nuclear proteins of the intermediate filament family in these syndromes. Alterations in the expression levels, functions and distribution of lamins A and C characterize a set of therapy.

abnormal aging syndromes, including lipodystrophy-atypical Werner syndromes (LIRLLC and WC), mandibuloacral dysplasia (MAD), Hutchinson-Gilford progeria syndrome (HGPS) and restrictive dermopathy (RD). These disorders range from those with mild effects to those that reduce lifespan or are even fatal in newborns. Mutations in the *LMNA* gene, which encodes both lamin A and lamin C, can be a primary cause of most of the laminopathies; defects in ZMPSTE24, a metalloproteinase involved in the processing of lamin A precursors, can cause MAD and RD. Levy reported that reduction of dominant-negatively acting unprocessed prelamin A *in vitro* and *in vivo* can reverse the cellular pathological phenotype, thus providing the hope of a targeted molecular

Functions of noncoding DNA

Manolis Dermitzakis (The Wellcome Trust Sanger Institute, Hinxton, UK) presented preliminary results of a wholegenome association study aimed at identifying functionally variable regulatory regions that are likely to contribute to complex phenotypes and disorders. Dermitakis and his colleagues have surveyed expression levels for around 700 genes (350 in regions defined by Encyclopedia of DNA Elements (ENCODE) project, 250 on chromosome 21 and 200 on chromosome 20) in immortalized lymphoblastoid cell lines from 60 unrelated humans from the Centre d'Etude du Polymorphism Humain (CEPH) pedigrees, and used the SNP genotypes of the same individuals, publicly available from the haplotype-mapping project HapMap, to perform association analysis. Dermitzakis reported data on 30 ENCODE genes and 120 genes on chromosome 20. Large numbers of genes with significant inter-individual variation in expression were identified, and a strong association bias was found between individual genes and specific SNPs. Both cis and trans effects on expression variation were found. The overall cis-effect distance from the target gene can be up to 4 Mb, suggesting that the regulatory landscape may be different from that hypothesized previously.

Ultraconserved elements in the human genome were the subject of a talk by David Haussler (Howard Hughes Medical Institute, University of California, Santa Cruz, USA). A comparative analysis of human, mouse, rat and chicken genomes revealed 481 such highly conserved elements of length 200 bp or greater that are totally unchanged in human, mouse and rat and have, on average, 96% identity with a species as distant as the chicken. About half of these sequences overlap with mRNA-coding sequences, especially those of genes involved in RNA processing. The remainder neither code for protein nor appear in the UTRs of known genes, and often appear in clusters within approximately one Mb regions that surround genes for transcription factors involved in embryonic development. There is experimental evidence that some of these elements act as distal enhancers located many hundreds of kilobases away from the transcription-factor gene

that they regulate. Haussler presented an experimental functional analysis of POLA/ARX ultraconserved elements that cluster on the X chromosome. Eight of these most highly conserved elements in the human genome were tested for their ability to act as enhancers and for expression of novel transcripts. Five drive brain-specific expression during mouse embryonic development, as shown using a β-galactosidase reporter. The expression pattern of four of these is a subset of the aristaless related homeobox (ARX) gene-expression pattern in the developing mouse brain. Notably, in humans ARX is associated with spasticity and intellectual disability, X-linked West syndrome, Partington syndrome, non-syndromic mental retardation and X-linked lissencephaly with abnormal genitalia. Two of the eight ultraconserved elements code for small ubiquitously expressed RNAs. Haussler suggested that the ultraconserved elements are likely to have multiple functions: some acting as nodes of enhancing or repressing binding sites, others encoding novel functional RNAs. Other possible functions could be regulation of chromatin or higher-order chromosome structures and regulation of mRNA transcripts.

Damian Labuda (Université de Montréal, Canada) discussed the contribution of gene conversion to genetic diversity. He described screening for DNA sequence variability in promoter regions, regions arbitrarily defined as 2 kb segments directly upstream of the first exon, in 40 individuals of African, Middle-Eastern, European and East-Asian descent. The variability of the promoter regions was found to be similar to that of other genomic segments. There was, however, a great variance in diversity indices among loci, which might be at least partially due to selection. Of note, larger than expected average amounts of recombination were estimated, and this was not only characteristic of promoter regions but was found throughout the genes. The greater than expected incidence of recombination apparently reflects gene conversion rather than the presence of recombination hot spots. Labuda suggested that gene conversion, by disturbing long-range linkage disequilibria and locally changing the redistribution of functional variants among haplotypes, may profoundly affect the outcome of linkage studies and thus disease mapping.

Segmental duplications and large copy-number variants within the human genome

A hot topic in genomic variability are the large segments of our genome that have undergone recent evolution and show variability in their copy number and gene content. Charles Lee (Brigham and Women's Hospital, Boston, USA) reported on the analysis of large-scale copy-number polymorphisms in the human genome using array-based comparative genomic hybridization (array-CGH). The microarrays contained probes based on bacterial artificial chromosome (BAC) human genomic clones spaced at 1 Mb intervals within the genome. Pairwise analysis of genomic

DNA from 39 unrelated and apparently healthy individuals identified more than 200 loci that contain large-size copynumber polymorphisms, some involving hundreds of kilobases of DNA. These loci are scattered throughout the genome and the polymorphism explains previously observed locus-specific variation. An average of 12 large-size polymorphisms were detected per individual using pooled control DNA from ten unrelated, healthy individuals. More than 50% of these polymorphic regions overlap known genes and approximately 25% of the identified loci map to regions previously thought to contain segmental duplications. Interestingly, some 10% of the loci reside within the 100 kb of gaps in the current presentation of the human genome. Together, these large-size copy-number variations may represent as much as a tenfold greater amount of genetic variation than SNPs in humans. A searchable database that will provide an updated catalog of these variations for accurate interpretation of whole-genome-directed arrayCGH assays in research and clinical settings has been established [http://projects.tcag.ca/variation/]. The previously unappreciated large-scale genomic heterogeneity argues for a more dynamic picture of the structure of the human genome. Further studies are likely to yield evidence as to whether these regions are associated with disease-linked rearrangements or account for genetic differences in susceptibility to diseases or reaction to specific environmental stimuli.

Bert De Vries (Human Genetics Nijmegen, Radboud University, Nijmegen, The Netherlands) reported the use of wholegenome tiling-path array-CGH to detect submicroscopic chromosome alterations at 100 kb resolution in patients with mental retardation and/or congenital anomalies. DNA copynumber alterations were detected in most patients, most variation being inherited from parents and corresponding to large-scale copy-number variants of the type described above. Some patients had de novo alterations - deletions or duplications - varying in size between 300 kb and 10 Mb. Submicroscopic alterations were identified in around 14% of the cases. Stefan Kirsch (University of Heidelberg, Germany) described an analysis of interchromosomal segmental duplications of the pericentromeric region on the human Y chromosome. He and his colleagues have identified and analyzed a euchromatic island within the pericentromeric repeats of the human Y chromosome. This 450 kb island was not detected and is not contained within the published Y chromosomal sequence. The entire 450 kb interval is almost completely duplicated and consists predominantly of interchromosomal duplication rather than the intrachromosomal duplications that are usually prevalent on the Y chromosome. The duplicated region contains eight putative genes with open reading frames, involving members of the DUX homeobox gene family.

Diagnostic tools and development of therapies

Mutations in the *BRCA1* gene are implicated in an increased risk of breast and ovarian cancer, and in 2001 the European

Patent Organization (EPO) granted three patents on BRCA1, in the face of strong opposition from European scientists. Gert Matthijs (Center of Human Genetics, Catholic University of Leuven, Belgium) discussed the history of this opposition, and its successful outcome. The patentees involved the US company Myriad Genetics, based in Salt Lake City, Utah, and the University of Utah. The patents gave Myriad Genetics a monopoly on diagnostic testing of BRCA1 and the right to levy a license fee, which evoked strong opposition from European geneticists, supported by several European governments. The first patent, which claimed "a method for diagnosing the predisposition for breast and ovarian cancer" by comparing the patient's BRCA1 sequence with a reference sequence, was revoked in May 2004. In January 2005, the two other patents were maintained in modified form after a final hearing. They no longer include a method for diagnostics, but only relate to a probe for BRCA1 and a probe for a common 185delAG mutation found in Ashkenazi populations, respectively. In practice, the patents will no longer interfere with breast and ovarian cancer diagnostics in Europe. The successful attack was based on errors contained in the DNA sequence as it was disclosed in the patent application in 1994. The process of opposing and reconsidering the patents granted by EPO has been extremely costly for the European community, and Matthijs suggested that patents should be restricted to technologies and innovative discoveries that improve diagnostics and treatment, and that no natural DNA sequences and variants should be subject to patenting.

Between 40% and 70% of people who receive treatment with a particular drug do not respond satisfactorily. The adverse effects of drugs are also a high cost to society, estimated at several million euros and hundreds of thousands of deaths every year worldwide. Magnus Ingelman-Sundberg (Karolinska Institute, Stockholm, Sweden) reviewed the variability between individuals in drug action and response. Variability can arise from the fact that many of the proteins involved in drug transport and metabolism are polymorphic; this has both qualitative and quantitative effects, leading, for example, to the enhancement of drug metabolism and rapid loss of the drug from the body. Ingelman-Sundberg cited several examples in which drugs are rendered ineffective as a result of ultra-rapid metabolism caused by multiple copies of particular genes or by a high level of allelic variability. The best current explanation for variability in the copy number of some genes is related to their selection to respond to environmental stress. The cytochrome P450 isoenzymes CYP2C19 and CYP2D6, for example, are involved in the metabolism of many widely used drugs, and polymorphisms in these enzymes give rise to inter-individual and interethnic-group variability in the metabolism of several therapeutic drugs, including some antidepressants. People who carry two null alleles of either gene are known as poor metabolizers (PMs), while those who carry more than two copies of a functional CYP2D6 gene are ultrarapid metabolizers (UMs); UMs are over-represented in the non-responder

http://genomebiology.com/2005/6/9/343

group compared with the control population. As more than 40 million Europeans carry multiple copies of the CYP2D6 gene with differences depending on the geographic location, this might largely explain the frequent lack of response to antidepressants and other drugs in the European population. Another source of variability in drug response is alternative splicing of several of the CYP3A genes, which leads to variability between individuals from different ethnic backgrounds. Exploration of copy number, splicing alleles and other specific variants should facilitate the understanding of the consequences for drug response between and within populations.

Recent advances in human genetics, providing molecular, biochemical and cellular understanding of diseases have raised considerable hopes for a better future for patients with genetic diseases. A stimulating presentation by Arnold Munnich (Hôpital des Enfants Malades Necker, Paris, France) showed the range of approaches that can now be explored to overcome specific molecular defects. These include protein engineering, enzyme replacement therapy, rectification of splicing defects, and re-expressing embryonic genes, as in the successful re-expression of fetal haemoglobin by hydroxyurea stimulation to treat thalassemias and sicklecell anemia. Similarly, Munnich showed that even if it is not possible to correct or replace the underlying defective gene, therapeutic strategies based on current knowledge should be able to ameliorate particular aspects of a disorder. We can look forward to further progress in this direction at future year's European Conferences.