European Journal of Human Genetics (2008) 16, 1–4

www.nature.com/ejhg

NEWS AND COMMENTARY

Popper revisited

GWAS here, last year

Gert-Jan B van Ommen

European Journal of Human Genetics (2008) **16,** 1–2;
doi:10.1038/sj.ejhg.5201970

Often, at the start of a new year, a journal editor tends to prelude on what this year might bring. In this case, however, I would like to reflect on the past year. Indeed, a major development in the field of Human Genetics has been is the arrival of the Genome-Wide Association Studies or GWAS. Many previous years have seen reports of GWAS, and a frequently heard complaint has been, with some exceptions, that the findings were met with irregular or downright controversial replication results. If anything, the year 2007 has been the year of the publication of a large amount of rapidly and widely replicated GWAS. Notably, this has taken place in the field of diabetes, with now over 20 loci identified in total. The largest harvest for 1 year reaped by type 2 diabetes with now in total 11 loci confidently identified, six of which last year by GWAS;[1–7] type 1 coming in second,[8,9] with five loci last year on a total of 11 as well, and additional loci for prostate cancer[10] and rheumatoid arthritis,[11] the latter replicated in parallel by candidate gene analysis.[12] And with bonuses coming along: dissection for body mass index has yielded the first obesity gene, FTO,[2] the prostate cancer risk allele is protective for type 2 diabetes, and some type 2 diabetes risk genes when highly expressed, looks like predisposing for colon cancer when poorly expressed.[6] These findings – and I count on correspondence telling us that we overlooked other examples – supports the concept of *connectivity*: our biology does not respect classical clinical discipline boundaries, and the key switches are rather involved in maintaining balance, with quite different pathology following dearth or excess. The main accelerator of these breakthroughs has been the development of the HapMap resource and the Perlegen SNP set, heirs of the early days SNP consortium. All in the wake of the Human Genome Project, resulting, within half a decade, in robust, very high-density SNP detection platforms of Affymetrix and Illumina. And as each of these is derived from the same HapMap set, tools have already been developed to jointly analyse large study samples typed by either platform by imputation, computing missing values by interpolation (eg, Servin and Stephens[13]).

Similar to positional cloning, GWAS do not need any *a priori* hypothesis of the underlying pathology. However, we are also at the entry of a new era: for the first time in the history of biomedicine, GWAS provide us with a powerful and accurate tool to tackle the complete genome for disease gene search entirely without segregation information.[14] Another crowning event here is undoubtedly the development of the giga-sequencing machines of Roche/454, Solexa-Illumina and recently ABI. These tools allow one to ask and answer genome-wide, yet very refined, molecular questions, as shown recently for the probing of chromatin structure at a genome-wide scale.[15,16] For those of us who have had – or still have – difficult times with grant applications of a non-hypothesis-driven, prospective nature, as they were seen to be out on fishing expeditions, the advent of these exploratory high-throughput approaches is especially welcomed. Indeed, one might maliciously wonder if we are not (temporarily, in this field and pending subsequent functional studies) close to the ultimate consumption date of the Popperian approach of hypothesis-driven research. For was not a main goal of this to unravel the truth in the most efficient, that is, plausible way, faced with a daunting scarcity of collectible data? Well, if it becomes cheaper to just collect *all* data required than to run after a hundred consecutive, plausible, but wrong hypotheses, starting with a hypothesis becomes an economic futility. The hypothesis as a guiding principle is then replaced by a truism: if one does not throw away anything before thoroughly assessing its irrelevance, one will always find what one is looking for■

Gert-Jan B van Ommen is at the Center for Human and Clinical Genetics and Center For Medical Systems Biology, Leiden University Medical Center P.O. Box 9600, 2300 RC Leiden NL. Tel: +31 71 526 9401/9400; Fax: +31 71 526 8285; E-mail: gjvo@lumc.nl

References

1 Sladek R *et al*: A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 2007; **445**: 881–885.
2 Frayling TM *et al*: A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 2007; **316**: 889–894.
3 Saxena R: Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 2007; **316**: 1331–1336.
4 Scott LJ *et al*: A genome-wide association study of type 2 diabetes in finns detects multiple susceptibility variants. *Science* 2007; **316**: 1341–1345.
5 Steinthorsdottir V *et al*: A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat Genet* 2007; **39**: 770–775.
6 Zeggini E *et al*: Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 2007; **316**: 1336–1341.
7 Salonen JT *et al*: Type 2 diabetes whole-genome association study in four populations: the DiaGen consortium. *Am J Hum Genet* 2007; **81**: 338–345.
8 Todd JA *et al*: Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nat Genet* 2007; **39**: 857–864.
9 Hakonarson H *et al*: A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. *Nature* 2007; **448**: 591–594.

10 Gudmundsson J *et al*: Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. *Nat Genet* 2007; **39**: 977–983.
11 Plenge RM *et al*: TRAF1-C5 as a risk locus for rheumatoid arthritis-a genomewide study. *N Engl J Med* 2007; **357**: 1199–1209.
12 Kurreeman FA *et al*: A candidate gene approach identifies the TRAF1/C5 region as a risk factor for rheumatoid arthritis. *PLoS Med* 2007; **4**: e278.
13 Servin B, Stephens M: Imputation-based analysis of association studies: candidate regions and quantitative traits. *PLoS Genet* 2007; **3**: e114.
14 Frayling TM, McCarthy MI: Genetic studies of diabetes following the advent of the genome-wide association study: where do we go from here? *Diabetologia* 2007; **50**: 2229–2233.
15 Robertson G *et al*: Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods* 2007; **4**: 651–657.
16 Mikkelsen TS *et al*: Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 2007; **448**: 553–560.

Diagnostic dilemma's

The congenital disorders of glycosylation are clinical chameleons

David J Coman

European Journal of Human Genetics (2008) **16,** 2–4;
doi:10.1038/sj.ejhg.5201962

The congenital disorders of glycosylation (CDG) are a fascinating and eclectic group of disorders with a multisystem phenotype. The largest impediment to diagnosing CDG is simply considering them in the first instance, which is no small task given their 'clinical mimicry'. Common clinical features associated with CDG involve developmental delay and intellectual impairment, ataxia, seizures, retinopathy, hepatic dysfunction, coagulopathies, failure to thrive, cardiomyopathy and pericardial effusion, hydrops fetalis, endocrine abnormalities, renal dysfunction, skeletal defects, early lethality, and dysmorphic features.[1] While the diagnosis of CDG may not lead to significant alterations in therapeutic management in most instances, an accurate diagnosis is imperative for natural history postulations and genetic counselling.

In CDG, hypoglycosylation of different glycoproteins and occasionally of other glycoconjugates leads to a variety of symptoms that ultimately affect multiple systems. Jaeken described the first CDG patients over 20 years ago, and in the last decade, more than 30 genetic glycosylation disorders mainly encompassing the N-linked and O-linked protein glycosylation pathways have been identified.[2] The CDG have been divided into two groups: type I defects, of which there are 13 subtypes (CDG-Ia to CDG-Im), and type II defects, of which there are 8 subtypes (CDG-IIa to CDG-II/COG8). The CDG are a 'booming field' of the inborn errors of metabolism.

An elegant example of the broadening CDG phenotype is given by Morava *et al*[3], on page XXX of this issue, who define a phenotype of an autosomal recessive cutis laxa (ARCL) associated with a combined N-linked and O-linked glycosylation defect. They describe a fascinating constellation of generalised neonatal cutis laxa that improves with time, joint hypermobility, decreased bone mineralisation, late closure of the anterior fontanels, progressive microcephaly, neuronal migration defects, seizures, hearing loss, transient feed intolerance, developmental and growth delay, and characteristic facial features.

ARCL is a genetically heterogeneous entity divided into two distinct groups, ARCL type I (OMIM 219100) and ARCL type II (OMIM219200). ARCL type I is a very severe disorder of elastic tissue marked by hernaea, visceral involvement, and early lethality. ARCL type II is likely to be a clinically and genetically heterogeneous condition. The cohort described by Morava *et al* displays many similarities to ARCL type II, and are likely to represent a subgroup of this broad umbrella terminology. Dermatological involvement as a major presenting feature is uncommon in CDG, with an ichthyosis-like skin disorder being described in CDG-If (OMIM 609180) and CDG-Im (OMIM 610768). Loose and wrinkled skin has been observed in CDG-IIe (OMIM 608779), but not as prominent as in patients with cutis laxa. CDG-IIe is secondary to a defect in the 7th subunit of the conserved oligomeric golgi complex (COG).[4] Like the patients described by Morava *et al*, deficiencies in COG exhibit abnormal N-linked and O-linked glycosylation. Progeroid Ehlers–Danlos syndrome (PEDS, OMIM 130070) is caused by defective galactosyltransferase-I activity, which results in defects in O-xylosyl-proteoglycans.[5] PEDS is characterised by wrinkled loose skin on the face, fine curly hair, scanty eyebrows and eyelashes, downslanting palpebral fissures, a distinctive facies, developmental delay, skeletal abnormalities, and the classical features of Ehlers–Danlos syndrome.[5] ARCL associated with a combined N-linked and O-linked glycosylation defect displays characteristic dysmorphology, neuroradiological pattern,