Population Genetics and Pharmacogenetics

Ken K. Ong^{a,b}

^aMedical Research Council Epidemiology Unit, Cambridge, and ^bDepartment of Paediatrics, University of Cambridge, UK

Yearbook 2007 [1] trumpeted the arrival of the genome-wide association study, based on the remarkable technology that can accurately genotype hundreds of thousands of single nucleotide polymorphisms (SNPs) in each individual. Into 2008, genome-wide association studies have successfully spearheaded the escalating discoveries of numerous new genetic variants associated with complex traits such as BMI, height, bone mineral density, fasting glucose concentration, and type 1 and type 2 diabetes. Several of these discoveries are described below and also in the accompanying chapters of this Yearbook. These incontrovertible genetic associations point to several new biological mechanisms and hopefully in time will lead to new understandings of disease pathogenesis and treatments. However, to put the success of these approaches into perspective, these new genetic loci still only explain a small fraction of our genetic predisposition to disease, or our inheritance of body size and shape. While 80% of the inter-individual differences in human height is estimated to be due to genetic factors, the recently published staggering combined list of at least 40 new SNPs convincingly associated with height (see Chapter 4: Growth and Growth Factors, pp 43–62) together probably only explain around 5% its variance. Where are the remaining genetic factors?

One answer is that the heritability of height, or BMI, is the summative result of several hundreds of SNPs and other common genetic variations, each contributing a tiny individual effect size. Demonstrating these small genetic effects and their interactions with the environment will require enormous sized studies, and this is the focus of several major national 'biobank' initiatives.

An intriguing alternative explanation is that heritability is explained by multiple rare variants, which each have a much larger effect on biological function and disease risk than common variants. Different rare variants would contribute the inheritance of height in different pedigrees, i.e. within different families or sub-populations. Emerging evidence for this is described in the final abstract ('The value of rare variants'), while the necessary future sequencing technology is described immediately below ('Mechanism of the year').

Mechanism of the year: your complete genome sequence

The complete genome of an individual by massively parallel DNA sequencing

Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen YJ, Makhijani V, Roth GT, Gomes X, Tartaro K, Niazi F, Turcotte CL, Irzyk GP, Lupski JR, Chinault C, Song XZ, Liu Y, Yuan Y, Nazareth L, Qin X, Muzny DM, Margulies M, Weinstock GM, Gibbs RA, Rothberg JM

Human Genome Sequencing Center, Baylor College of Medicine, Houston, Tex., USA

Nature 2008;452:872–876

Background: There is much optimism for the impact of 'genomic medicine' on disease and drug response. Complete genome sequencing is needed to realize the full potential of genomics for human health. However, the formidable size of the diploid human genome, approximately 6 gigabases, has prevented the routine use of current sequencing methods to decipher complete individual human genomes.

Methods: The authors now report the DNA sequence of a diploid genome of a single individual, James D. Watson, sequenced to 7.4-fold redundancy in 2 months using massively parallel sequencing in picoliter-size reaction vessels.

Results: The authors completed this genome sequence in 2 months at around one-hundredth of the cost of traditional capillary electrophoresis methods. The authors compared this sequence to the reference genome and identified 3.3 million single nucleotide polymorphisms, of which 10,654 are within gene coding sequences and cause amino-acid substitution. In addition, the authors accurately identified small-scale

(2–40,000 base pair (bp)) insertion and deletion polymorphism as well as copy number variation resulting in the large-scale gain or loss of chromosomal segments ranging from 26,000 to 1.5 million bp.

Conclusion(s): Overall, this method sequences the genome of a single individual with good agreement with those recently achieved by traditional methods. In addition to being faster and much less expensive, this sequencing technology avoids the arbitrary loss of genomic sequences inherent in random shotgun sequencing by bacterial cloning because it amplifies DNA in a cell-free system. Therefore, this method acquires a novel human sequence, including novel genes not previously identified by traditional genomic sequencing. This is the first genome sequenced by next-generation technologies and is a pilot for the future challenges of ‘personalized genome sequencing’.

James Watson and Francis Crick identified the double-helix backbone of DNA in which the bases adenine, thymine, guanine and cytosine fit to encode our genetic makeup. Remarkably, now 55 years later, this paper describes James Watson’s complete DNA sequence of around 3 billion base pairs. Watson’s is not the first complete individual sequence to be published. Craig Venter, who spearheaded the Human Genome Project (HGP), had his sequence reported in 2007. However, this publication is symbolic in achieving this goal at a fraction of the costs and time associated with previous techniques. Compared to around USD 100 million for Venter’s genome, Watson’s code cost less than USD 1 million and was arguably more complete. If your whole genome sequence is now ‘within reach’, what about its interpretation? That is going to be a much larger task. Compared to the HGP consensus sequence, Watson’s showed 10,654 variants that would directly result in an amino acid change. If Watson took his sequence to a genetic counselor, he would identify a handful of mutations that might confer risk on offspring. Furthermore, the ENCODE project recently showed that even the most humble inter-genic sequences often play essential regulatory roles [2]. Piecing together the effects of each of your 3.3 million SNPs and realizing their combined impact on the phenotype, disease risk and drug responses is an enormous future challenge. However, if the HGP gave us the map, then new rapid sequencing technologies provide the highways.

New paradigms: environmental effects on DNA?

Heritable rather than age-related environmental and stochastic factors dominate variation in DNA methylation of the human IGF2/H19 locus

Heijmans BT, Kremer D, Tobi EW, Boomsma DI, Slagboom PE

Molecular Epidemiology Section, Leiden University Medical Centre, The Netherlands

Hum Mol Genet 2007;16:547–554

Background: Epigenetic variation may influence susceptibility to common disease. However, there is currently little known about the extent and causes of epigenetic variation. This study investigated the contribution of heritable influences and the combined effect of environmental and stochastic factors to variation in DNA methylation of the IGF2/H19 locus. Moreover, the study tested whether epigenetic patterns at this locus changed with age, as was previously suggested for global methylation.

Methods: The authors measured methylation of the H19 and IGF2 differentially methylated regions (DMRs) in 196 adolescent and 176 middle-aged twins using a recently developed mass spectrometry-based method.

Results: There was substantial variation in DNA methylation between individuals. The mean methylation of CpGs in the H19 DMR varied between 23 and 30%, and the mean methylation of the IGF2 DMR varied between 43 and 62%. Comparison of differences between monozygotic and dizygotic twins gave estimates of heritability between 20 and 74% for methylation at individual H19 DMR CpG sites, and between 57 and 97% for the IGF2 DMR. For both regions, the average methylation was similar in adolescent and middle-aged individuals and was unrelated to sex and age. Single nucleotide polymorphisms (SNPs) in the IGF2/H19 locus were significantly associated with DNA methylation of the IGF2 DMR ($p = 0.004$). A preliminary analysis suggested an association between H19 DMR methylation and body size ($p < 0.05$).

Conclusion(s): Our study shows that variation in DNA methylation of the IGF2/H19 locus is mainly determined by heritable factors and SNPs in cis. In contrast there was only a limited role for environmental and stochastic factors and age-related degeneration of methylation patterns at this locus.

Epigenetic mechanisms, such as histone acetylation and DNA methylation, regulate whether a gene is expressed or silent. The maternally imprinted *IGF2* gene on chromosome 11p15.5 is one of the best-characterized epigenetically regulated loci. It involves both the *IGF2* region itself, of which the maternal allele is methylated, and the *H19* region, of which the paternal allele is methylated. Demethylation of the imprinting control region leads to *IGF2* overexpression and a ~5-fold increased risk of colorectal neoplasia.

While inherited (imprinted) severe epigenetic mutations in the IGF2/H19 locus cause both the Beckwith-Wiedeman syndrome and the Silver-Russell syndrome, it is unknown whether normal epigenetic patterns in this locus change with age. Furthermore, *IGF2* is a major fetal growth factor and is implicated in body composition and atherosclerosis. There is growing interest in the potential epigenetic effects of environmental factors, such as diet, which may influence long-term disease risks.

The current study showed that while there is marked heterogeneity in the degree of methylation between individuals at the IGF2/H19 locus, most of this variation was due to heritable/genetic factors, while there was remarkably little or no difference in methylation with age between adolescents and middle-aged adults. While they did not have longitudinal DNA samples from the same individuals, the findings imply that methylation patterns, at least at this highly imprinted locus, are very stable during adult life. However, some CpG sites showed only moderate heritability. Therefore, these findings could imply that environmental factors acting during a specific period of plasticity before adolescence, such as antenatal or postnatal dietary factors, could have a long-term impact on methylation patterns and long-term disease risk.

New genetic associations: puberty and sexual development

Male infertility and variation in CAG repeat length in the androgen receptor gene: a meta-analysis

Davis-Dao CA, Tuazon ED, Sokol RZ, Cortessis VK

Department of Preventive Medicine, University of Southern California, Los Angeles, Calif., USA

J Clin Endocrinol Metab 2007;92:4319–4326

Background: Many studies have looked at the association between male infertility and trinucleotide repeat polymorphisms in the androgen receptor (AR) gene, but no comprehensive meta-analysis has been conducted. This study summarized the published data on associations between AR CAG and GGC repeat lengths and male infertility.

Methods: The authors searched for reports published before October 2006 using Medline, PubMed, and Web of Science. Inclusion criteria for selected studies were: a case group with infertility based on measured semen parameters, a control group of known or presumed fertile men, and measurement of AR CAG and/or GGC repeat lengths among cases and controls. 39 reports met these criteria, and 33 were included in the meta-analysis. One investigator extracted data on sample size, mean and SD of trinucleotide repeat length, and study characteristics.

Results: Cases had significantly longer AR CAG repeat length than controls. Estimates of the standardized mean difference (95% confidence interval) were 0.19 (0.09–0.29) for all 33 studies, and 0.31 (0.14–0.47) for a subset of 13 studies that used more stringent case and control selection criteria. Publication date appeared to be a significant source of variation between studies.

Conclusion(s): This meta-analysis provides support for an association between increased androgen receptor CAG length and idiopathic male infertility, suggesting that even subtle disruptions in the androgen axis may compromise male fertility.

Spermatogenesis is regulated by androgens in a largely paracrine fashion. Leydig cells of the adult testis secrete testosterone, which likely acts on androgen receptors (AR) expressed by the Sertoli cells.

The results of this large meta-analysis indicate that AR variants (longer CAG repeat lengths), which confer reduced AR activity in response to ligand binding and therefore limited Sertoli cell function, may contribute to spermatogenetic deficits in men with idiopathic infertility. The literature to date on this question has been inconclusive. Based on the observed mean differences in CAG length between cases and controls, the authors calculate that a single study would require 3,533 cases and 3,533 controls to have sufficient power. None of the 33 individual studies even approached such size, and this indicates the large collaborative efforts needed to confidently demonstrate, or refute, common genetic associations. Recent studies have also reported associations between shorter CAG repeat length and susceptibility to polycystic ovary syndrome in US adults [3] and premature adrenarche in Finnish children [4]. If confirmed, those findings would support the etiological role of increased androgen receptor activity in those conditions and therefore the rationale for androgen receptor blockade therapy.

Risk factors for hypospadias in the estrogen receptor 2 gene

Beleza-Meireles A, Kockum I, Lundberg F, Soderhall C, Nordenskjöld A

Department of Molecular Medicine and Surgery, Karolinska University Hospital, Stockholm, Sweden

J Clin Endocrinol Metab 2007;92:3712–3718

Background: The etiology of hypospadias remains elusive. Estrogen receptor-2 (ESR2) is the predominantly expressed ESR in the developing male urethra, suggested a role of estrogens in the otherwise androgen-dependent male genital differentiation. Longer (CA)_n variants in the ESR2 gene have been associated with lower androgen levels as well as hypospadias. The aim of this study was to systematically explore the role of ESR2 gene variants on the risk of hypospadias.

Methods: The authors genotyped 4 haplotype-tagging single nucleotide polymorphisms (SNPs; rs2987983, rs1887994, rs1256040, and rs1256062), the (CA)_n polymorphism, and 2 additional promoter SNPs (rs10483774 and rs1271572), mapping to a transcription factor binding region in 354 Swedish boys with non-syndromic hypospadias and 380 healthy controls.

Results: Hypospadias was associated with longer variants of the (CA)_n polymorphism in intron 6 ($p < 0.002$) and also with SNPs rs2987983 ($p < 0.01$) and rs10483774 ($p < 0.01$), which are located in the putative promoter region, in an area of intense transcription factor binding. The two regions were in low-linkage disequilibrium, however logistic regression analysis indicated that these two risk effects were not independent.

Conclusion(s): This study provides evidence for the involvement of variants in the ESR2 gene in hypospadias and supports the concept that estrogens and their receptors are important for the developing male urethra.

There is increasing evidence for a major role of estrogens in male reproductive tract development. In mouse models, estrogen exposure inhibits cell proliferation in the developing urethra in a dose-dependent manner and can also induce hypospadias. In humans, deleterious gene mutations in the androgen pathway explain only a small proportion of non-syndromic hypospadias. The current study shows that common genetic variation in estrogen activity in humans may influence male urethral tract development. These findings also support the hypothesis that the xenoestrogens, the environmental endocrine disrupters, could have important functional consequences on male reproductive tract abnormalities at the estrogen receptor (ESR) level rather than as anti-androgens. In other fields, consistent associations with common variants in the ESR1 and ESR2 genes have helped to demonstrate the role of estrogen activity in childhood growth, bone mineral density and body composition [5] as well as cognitive function and breast cancer risks in later life [6].

Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes

Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, de Bakker PI, Abecasis GR, Almgren P, Andersen G, Ardlie K, Bostrom KB, Bergman RN, Bonnycastle LL, Borch-Johnsen K, Burtt NP, Chen H, Chines PS, Daly MJ, Deodhar P, Ding CJ, Doney AS, Duren WL, Elliott KS, Erdos MR, Frayling TM, Freathy RM, Gianniny L, Grallert H, Grarup N, Groves CJ, Guiducci C, Hansen T, Herder C, Hitman GA, Hughes TE, Isomaa B, Jackson AU, Jorgensen T, Kong A, Kubalanza K, Kuruvilla FG, Kuusisto J, Langenberg C, Lango H, Lauritzen T, Li Y, Lindgren CM, Lyssenko V, Marville AF, Meisinger C, Midtjell K, Mohlke KL, Morken MA, Morris AD, Narisu N, Nilsson P, Owen KR, Palmer CN, Payne F, Perry JR, Pettersen E, Platou C, Prokopenko I, Qi L, Qin L, Rayner NW, Rees M, Roix JJ, Sandbaek A, Shields B, Sjogren M, Steinthorsdottir V, Stringham HM, Swift AJ, Thorleifsson G, Thorsteinsdottir U, Timpson NJ, Tuomi T, Tuomilehto J, Walker M, Watanabe RM, Weedon MN, Willer CJ, Illig T, Hveem K, Hu FB, Laakso M, Stefansson K, Pedersen O, Wareham NJ, Barroso I, Hattersley AT, Collins FS, Groop L, McCarthy MI, Boehnke M, Altshuler D Wellcome Trust Centre for Human Genetics, University of Oxford, UK
Nat Genet 2008;40:638–645

Background: Genome-wide association (GWA) studies have identified multiple loci at which common variants modestly but reproducibly influence risk of type 2 diabetes (T2D). Established associations to common and rare variants explain only a small proportion of the heritability of T2D. Larger sample sizes are needed to overcome the limited power of previously published analyses to identify variants with modest effects.

Methods: The authors carried out a meta-analysis of three T2D GWA scans comprising approximately 2.2 million SNPs (directly genotyped and imputed) in 10,128 individuals of European descent. Replication testing was then performed in an independent sample with an effective sample size of up to 53,975.

Results: We detected at least six previously unknown loci with robust evidence for association, including the JAZF1 ($p = 5.0 \times 10^{-14}$), CDC123-CAMK1D ($p = 1.2 \times 10^{-10}$), TSPAN8-LGR5 ($p = 1.1 \times 10^{-9}$), THADA ($p = 1.1 \times 10^{-9}$), ADAMTS9 ($p = 1.2 \times 10^{-8}$) and NOTCH2 ($p = 4.1 \times 10^{-8}$) gene regions.

Conclusion(s): Our results illustrate the value of large discovery and follow-up samples for gaining further insights into the inherited basis of T2D.

Before this study, GWA studies had identified 10 common genetic variants with validated associations with T2D (excluding FTO, which has its primary effect through obesity). Apart from TCF7L2 with an odds ratio (OR) per risk allele of 1.35, the 9 other variants conferred ORs between 1.1 and 1.2. The authors had already shown that large numbers were needed. Indeed, in the original 'separate' back-to-back publications of these three studies last year the findings were highly dependent on a limited pooling of the 'top hits' from each study. The current report is based on a meta-analysis of the complete GWA data from each study (stage 1). Promising SNPs at previously unknown loci were then selected for replication in large independent samples.

While this was an obvious next step for these three studies, the resulting successful discovery of a further six new T2D loci has provided much encouragement to genetic epidemiologists who are still searching to explain the majority of the heritability of complex diseases. Increasing sample sizes beyond those initially planned by pooling data from existing studies lead to an amplification of true positive signals. Consortia of groups working on other diseases and traits are also pooling their GWA data with similar successful results. More laborious resequencing and genotyping is needed to pinpoint the specific causal variants at each loci and should enhance the effect sizes and increase the proportion of heritability explained.

A polymorphism within the G6PC2 gene is associated with fasting plasma glucose levels

Bouatia-Naji N, Rocheleau G, Van Lommel L, Lemaire K, Schuit F, Cavalcanti-Proenca C, Marchand M, Hartikainen AL, Sovio U, De Graeve F, Rung J, Vaxillaire M, Tichet J, Marre M, Balkau B, Weill J, Elliott P, Jarvelin MR, Meyre D, Polychronakos C, Dina C, Sladek R, Froguel P

CNRS UMR 8090 Institute of Biology, Pasteur Institute of Lille, Lille, France

Science 2008;320:1085–1088

Background: Several studies have shown that healthy individuals with high-normal fasting plasma glucose (FPG) levels have an increased risk of mortality. The aim of this study was to identify common genetic determinants that contribute to inter-individual variation in FPG.

Methods: The authors genotyped 392,935 single nucleotide polymorphisms (SNPs) in 654 normoglycemic subjects for association with FPG. The most strongly associated SNP was then selected for replication in 9,353 subjects.

Results: SNP rs560887 showed the strongest association with FPG ($p = 4 \times 10^{-7}$). SNP rs560887 maps to intron 3 of the G6PC2 gene, which encodes glucose-6-phosphatase catalytic subunit-related protein (also known as IGRP), a protein selectively expressed in pancreatic islets. The replication study confirmed the association between rs560887 and FPG ($\beta = -0.06 \text{ mmol/l per A-allele}$; $p = 3 \times 10^{-8}$), combined data from three populations ($p = 4 \times 10^{-23}$) and with pancreatic β -cell function (Homa-B model, combined $p = 3 \times 10^{-13}$); however it was not associated with type 2 diabetes risk.

Conclusion(s): The authors speculate that G6PC2/IGRP regulates FPG by altering the set-point for glucose-stimulated insulin secretion in pancreatic β -cells.

This novel genome-wide association study reports a new genetic variant that is consistently associated with fasting glucose levels, but is not associated with clinical cases of type 2 diabetes (T2D). G6PC2/IGRP is proposed to be part of the same glucose phosphorylation pathway that includes the glucokinase gene (GCK) and IGRP may antagonize glucokinase activity in β -cells. It may be highly relevant that severe GCK mutations are responsible for the monogenic form of diabetes, MODY2, which has a phenotype of mild, asymptomatic, non-progressive hyperglycemia and is unassociated with typical diabetic complications. Similarly G6pc2-null mice show a mild reduction in glucose levels but no differences in body weight, insulin sensitivity or major changes in other metabolic parameters.

There has been long debate over the concordance, or rather the relative discordance, between the fasting versus stimulated glucose concentration criteria used to diagnose T2D and to indicate later cardiovascular disease risk. Several investigators have suggested that the two measurements are regulated by different processes, relating to fasting versus glucose-stimulated insulin secretory responses. Almost all of the expanding list of T2D associated genetic loci appear to influence β -cell insulin secretory responses to oral or intravenous glucose [7]. Comparison between these genetic loci for fasting versus stimulated glucose and insulin levels could help to establish the degree of overlap between their underlying biological mechanisms, and to distinguish if any or both have causal effects on cardiovascular disease risk.

Common variants near MC4R are associated with fat mass, weight and risk of obesity

Loos RJ, Lindgren CM, Li S, Wheeler E, Zhao JH, Prokopenko I, Inouye M, Freathy RM, Attwood AP, Beckmann JS, Berndt SI, Jacobs KB, Chanock SJ, Hayes RB, Bergmann S, Bennett AJ, Bingham SA, Bochud M, Brown M, Cauchi S, Connell JM, Cooper C, Smith GD, Day I, Dina C, De S, Dermizakis ET, Doney AS, Elliott KS, Elliott P, Evans DM, Sadaf Farooqi I, Froguel P, Ghorji J, Groves CJ, Gwilliam R, Hadley D, Hall AS, Hattersley AT, Hebebrand J, Heid IM, Lamina C, Gieger C, Illig T, Meitinger T, Wichmann HE, Herrera B, Hinney A, Hunt SE, Jarvelin MR, Johnson T, Jolley JD, Karpe F, Keniry A, Khaw KT, Luben RN, Mangino M, Marchini J, McArdle WL, McGinnis R, Meyre D, Munroe PB, Morris AD, Ness AR, Neville MJ, Nica AC, Ong KK, O'Rahilly S, Owen KR, Palmer CN, Papadakis K, Potter S, Pouta A, Qi L, Kraft P, Hankinson SE, Hunter DJ, Hu FB, Randall JC, Rayner NW, Ring SM, Sandhu MS, Scherag A, Sims MA, Song K, Soranzo N, Speliotes EK, Lyon HN, Voight BF, Ridderstrale M, Groop L, Syddall HE, Teichmann SA, Timpson NJ, Tobias JH, Uda M, Scheet P, Sanna S, Abecasis GR, Albal G, Nagaraja R, Schlessinger D, Ganz Vogel CI, Wallace C, Waterworth DM, Weedon MN, Willer CJ, Jackson AU, Tuomilehto J, Collins FS, Boehnke M, Mohlke KL, Wraight VL, Yuan X, Zeggini E, Hirschhorn JN, Strachan DP, Ouwehand WH, Caulfield MJ, Samani NJ, Frayling TM, Vollenweider P, Waeber G, Mooser V, Deloukas P, McCarthy MI, Wareham NJ, Barroso I
MRC Epidemiology Unit, Institute of Metabolic Science, Cambridge, UK
Nat Genet 2008;40:768–775

Background: Recently, the FTO gene was identified as the first locus harboring common variants with an unequivocal impact on obesity predisposition and fat mass at the population level. This study aimed to identify common variants influencing body mass index (BMI).

Methods: The authors analyzed genome-wide association data from 16,876 individuals of European descent. Replication studies were performed in a further 60,352 adults and 5,988 children, as well as in several groups of obese cases.

Results: After the previously reported variants in FTO, the strongest association signal (rs17782313, $p = 2.9 \times 10^{-6}$) mapped 188 kb downstream of MC4R (melanocortin-4 receptor). The rs17782313 association with BMI was confirmed by independent replication in adults (per allele effect = 0.05 Z-score units; $p = 2.8 \times 10^{-15}$) and in children aged 7–11 (0.13 Z-score units; $p = 1.5 \times 10^{-8}$). In case-control analyses ($n = 10,583$), the odds for severe childhood obesity reached 1.30 ($p = 8.0 \times 10^{-11}$). Overtransmission of the risk allele to obese offspring was observed in 660 families ($p = 2.4 \times 10^{-4}$).

Conclusion(s): The SNP location and patterns of phenotypic associations are consistent with effects mediated through altered MC4R function. These findings establish that common variants near MC4R influence fat mass, weight and obesity risk at the population level and reinforce the need for large-scale data integration to identify variants influencing complex biomedical traits.

This study reports only the second common genetic variant that is incontrovertibly associated with BMI and obesity risk. The authors purport that the effect might be due to altered function at the MC4R gene, in which rare mutations are the leading cause of monogenic severe childhood-onset obesity. The variant is located far downstream of MC4R and future major efforts will be required to identify the causal variant(s) and to determine the mechanism of the hypothesized long range regulation of MC4R. Consistent with the rare MC4R monogenic cases, the effect of this common variant (minor allele frequency 24%) on BMI standard deviation scores was twice as strong in children as in adults. Furthermore, as for FTO, the effects of this variant appeared to be specific for childhood body fat rather than lean body mass.

The other major finding of this study is 'what it did not find'. This combined collaborative analysis of GWA data is one of the largest to date (as indicated by the number of authors). Yet there was little evidence for any other common variant that even approached the effect sizes of the novel MC4R variant and FTO, around 0.05–0.06 fractions of one standard deviation of adult BMI, which is equivalent to a difference of 1.5 kg in adult body weight per risk allele. Further meta-analyses of most of the available worldwide GWA data for BMI and obesity are in progress and will likely identify several more novel common variants with likely even smaller effect sizes. However, the genetic basis for obesity risk in the general population remains largely elusive. It is intriguing that the majority of the monogenic and common variants appear to have a greater impact in children and even in infants than in adults [8, 9]. New obesity variant discovery projects in children would perhaps yield greater fruit.

Evidence for a strong genetic influence on childhood adiposity despite the force of the obesogenic environment

Wardle J, Carnell S, Haworth CM, Plomin R

Department of Epidemiology and Public Health, Health Behaviour Research Centre, University College London, London, UK

Am J Clin Nutr 2008;87:398–404

Background: Body mass index (BMI) is highly heritable. However, most of the studies showing this were carried out in cohorts born before the ‘obesity epidemic’. This study aimed to quantify the genetic and environmental influences on BMI in children growing up during the dramatic rises in pediatric obesity.

Methods: The authors analyzed BMI and waist circumference (WC) data at age 8–11 years in a UK cohort of 5,092 twin pairs born in 1994–1996. Quantitative genetic model-fitting was used to assess heritability. Bivariate quantitative genetic model-fitting was used to assess covariance between BMI and WC.

Results: These data confirmed substantial heritability for BMI and WC (77% for both). Although the genetic influence on WC was largely common to BMI (60%), there was also a significant independent genetic effect (40%). For both BMI and WC, there was a very modest shared-environment effect, and the remaining environmental variance was unshared.

Conclusion(s): Genetic influences on BMI and WC are high in children born since the onset of the pediatric obesity epidemic. Targeting the family may be vital for obesity prevention in the earliest years, but longer-term weight control will require a combination of individual engagement and society-wide efforts to modify the environment, especially for children at high genetic risk.

Low physical activity accentuates the effect of the FTO rs9939609 polymorphism on body fat accumulation

Andreassen CH, Stender-Petersen KL, Mogensen MS, Torekov SS, Wegner L, Andersen G, Nielsen AL, Albrechtsen A, Borch-Johnsen K, Rasmussen SS, Clausen JO, Sandbaek A, Lauritzen T, Hansen L, Jorgensen T, Pedersen O, Hansen T
Steno Diabetes Center, Gentofte, Denmark

Diabetes 2008;57:95–101

Background: Several studies have shown that common genetic variation in the FTO gene is associated with BMI and obesity. The present study examined the effect of FTO variation on metabolic traits including obesity, type 2 diabetes, and related quantitative phenotypes.

Methods: The FTO rs9939609 polymorphism was genotyped in a total of 17,508 Danes from five different study groups.

Results: In studies of 3,856 type 2 diabetic case subjects and 4,861 normal glucose-tolerant control subjects, the minor A-allele of rs9939609 was associated with type 2 diabetes (odds ratio 1.13, 95% CI 1.06–1.20, $p = 9 \times 10^{-5}$). This association was completely explained by BMI (odds ratio 1.06, 95% CI 0.97–1.16, $p = 0.2$). Among 17,162 middle-aged Danes, the A-allele associated with overweight (odds ratio 1.19, 95% CI 1.13–1.24, $p = 1 \times 10^{-12}$) and obesity (odds ratio 1.27, 95% CI 1.20–1.34, $p = 2 \times 10^{-16}$). Furthermore, obesity-related quantitative traits such as body weight, waist circumference, fat mass, and fasting serum leptin levels were significantly elevated in A-allele carriers. An interaction between the FTO rs9939609 genotype and physical activity ($p = 0.007$) was found, where the genotype effect on BMI was greater (1.95 kg/m^2) among physically inactive individuals than light-medium (0.69 kg/m^2) or very active individuals (0.47 kg/m^2) for homozygous risk A-allele carriers compared with homozygous T-allele carriers.

Conclusion(s): This study confirms that variation in FTO confers increased body fat mass and hence increased risk of type 2 diabetes. Furthermore, low physical activity seems to accentuate the effect of FTO rs9939609 on body fat accumulation.

What role can genetic factors have in the modern obesity epidemic? The marked recent increases in obesity in children and adults must surely be due to changes in our lifestyles and environment rather than our genes. Furthermore, the absence of major common genetic factors that explain a significant

proportion of obesity risk have questioned our previous high estimates of the heritability of BMI. However, these two papers suggest that the impact of genetic factors on obesity risk may actually have recently increased. The first paper by Wardle et al. reports a relatively high estimate of heritability of BMI in children born in the mid-1990s and living through the UK boom in childhood obesity rates. The current figure of 77% is towards the higher end of the range of 50–90% from a review of studies published up to 1997. There are a number of potential biases in twins studies which could inflate the estimate of heritability, so the true value is likely lower than 77%, however these biases do not affect the relative comparison of the different estimates over time.

The second paper by Andreassen et al. sheds some light on the current role of genetic factors. The novel finding in that study was the interaction between FTO and physical activity. 34% of those middle-aged Danes were self-reportedly inactive, and the effect of the obesity risk allele was significantly greater in this group than among more active Danes. Therefore certain genetically predisposed individuals may be more prone to the obesogenic effects of our slothful and plentiful western environments. The Pima Indians are a case in point as they are considered to have a marked genetic susceptibility to changes from traditional to western environments [10]. The good news is that even in the most genetically prone populations obesity should be preventable. Further identification of specific genetic factors and their interactions may one day indicate individual risk pathways to obesity as well as individualized interventions.

Concepts revised: the d3-growth hormone receptor polymorphism

The presence of the d3-growth hormone receptor polymorphism is negatively associated with fetal growth but positively associated with postnatal growth in healthy subjects

Jensen RB, Vielwerth S, Larsen T, Greisen G, Leffers H, Juul A

University Department of Growth and Reproduction, Rigshospitalet, Copenhagen, Denmark

J Clin Endocrinol Metab 2007;92:2758–2763

Background: A common exon 3 deletion (d3) polymorphism in the growth hormone receptor (GHR) gene has been associated with increased growth response in GH-treated patients. No study has examined its association with prenatal growth.

Methods: 115 healthy adolescents were divided into those born small for gestational age (SGA) and appropriate for gestational age. Identification of intrauterine growth restriction was based on prospective data on third-trimester fetal growth velocity (FGV, measured by serial ultrasonography). Isoforms of the d3-GHR gene (fl/fl, d3/fl, and d3/d3) were determined.

Results: The prevalence of the d3-GHR isoforms (overall 50%) differed between the groups ($p = 0.006$). The highest prevalence (88%) was in the group born SGA with verified intrauterine growth restriction. Among SGA individuals, d3-GHR alleles were associated with slower third-trimester FGV ($p = 0.05$). In the entire cohort, carriers of the d3-GHR allele were taller (height SD score: 0.34 vs. -0.10 ; $p = 0.017$) and showed greater change in height SD score from birth to adolescence (0.57 vs. -0.02 SD score; $p = 0.005$) compared with carriers of the full-length GHR allele.

Conclusion(s): This study showed an increased spontaneous postnatal growth velocity in the carriers of the d3-GHR allele. Interestingly, the d3-GHR allele appeared to have the opposite effect on prenatal growth in the SGA group.

The common d3-GHR variant of the GH-receptor has been shown to increase GHR sensitivity in vitro. The findings of this study that the d3-GHR variant confers taller stature and faster spontaneous growth are therefore consistent with this biological effect. In addition, these individuals had unique data on fetal growth velocity from serial ultrasound measurements. Intriguingly, the authors found that, in contrast to postnatal growth, the d3-GHR allele was associated with slower fetal growth rates (see Concepts revised: genetic influences on birth weight).

The clinical role of the d3-GHR polymorphism has been a long-standing question in previous editions of this *Yearbook* chapter. d3-GHR alleles have been reported to confer increased growth response to

GH therapy in various childhood conditions. However, studies have been conflicting. This confusion is little clarified by other publications on the d3-GHR in the last 12 months which report: increased short-term growth response to GH in GHD children [11] and in preterm-born children [12], and greater metabolic responses to GH in GHD adults [13]. In contrast, there was no association with growth responses in short SGA [14] or short normal birth weight children [15], and there were no apparent associations with longer-term outcomes [11, 13]. However, the sample sizes in all these studies ranged from 60 to 180. While such sample sizes represent large clinical collections, they are small relative to those needed to robustly confirm or refute effects of common genetic variants. It would be very helpful if in future each new study could include an updated meta-analysis of studies in that population. Furthermore, the biological relevance of the d3-GHR to spontaneous childhood growth needs to be confirmed in much larger childhood populations. Common wisdom is that fetal growth is independent of fetal GH levels, and hence, GHR variation should have little if any impact on fetal growth. This common wisdom may have to be reconsidered.

Concepts revised: the Barker hypothesis – genes or environment

Genetic and shared environmental factors do not confound the association between birth weight and hypertension: a study among Swedish twins

Bergvall N, Iliadou A, Johansson S, de Faire U, Kramer MS, Pawitan Y, Pedersen NL, Lichtenstein P, Cnattingius S
Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
Circulation 2007;115:2931–2938

Background: Studies have found associations between low birth weight and increased risks of cardiovascular diseases in adulthood. However, these associations could be due to confounding by genetic or socioeconomic factors.

Methods: The authors studied same-sex twins with known zygosity born in Sweden between 1926 and 1958. First, to obtain an overall effect of birth weight on risk of hypertension, they performed cohort analyses on all twins ($n = 16,265$). Second, to address genetic and shared environmental confounding, they performed a nested co-twin control analysis within 594 dizygotic and 250 monozygotic twin pairs discordant for hypertension. Birth characteristics were obtained from original birth records. Information from adulthood was collected from a postal questionnaire in 1973 (body mass index, height, smoking, and alcohol use) and from a telephone interview conducted from 1998 to 2002 (hypertension and socioeconomic status). Hypertension was defined as both high blood pressure and use of antihypertensive medication.

Results: In the cohort analysis, the adjusted odds ratio for hypertension for each 500-gram decrease in birth weight was 1.42 (95% CI 1.25–1.61). In the co-twin control analyses, similar odds ratios were observed within dizygotic 1.34 (95% CI 1.07–1.69) and monozygotic twins 1.74 (95% CI 1.13–2.70).

Conclusion(s): In the largest twin study on the fetal origins of hypertension, lower birth weight was associated with increased risk of hypertension independent of genetic factors, shared familial environment, and risk factors for hypertension in adulthood, including body mass index.

There has been much debate over the mechanisms behind the ‘fetal origins’ associations between birth weight and later disease. Proponents have been divided into genetic versus environment camps, the latter hoping that maternal nutrition provides both the explanation and solution. In one of the largest twins studies to date these authors used the ‘co-twin control’ model, which is an extension of the classical heritability analyses of twin studies, to attempt to disentangle the antenatal mechanisms that underlie associations with birth weight. Their somewhat surprising conclusion disappoints both the genetics and maternal environment camps; they confirmed the low birth weight-hypertension association seen in the general population, but concluded that this association was independent of both fetal genetic factors and also ‘shared environment’ (the latter includes maternal uterine factors that would be shared by both twins in each pair, e.g. mother’s metabolism and mother’s environment).

The answer may therefore lie in a ‘third’ category of non-genetic factors that are individual to each fetus (e.g. placental implantation and placental function). These findings would imply that interventions

at the level of the mother to increase fetal growth would be ineffective to reduce later disease risk, and certainly we need a much greater understanding of how nutritional supply is regulated by the placenta to the fetus. However, twins are in general more growth restricted in utero than singletons, and generalizability of findings in twins to the general population remains in question.

Concepts revised: genetic influences on birth weight

Type 2 diabetes TCF7L2 risk genotypes alter birth weight: a study of 24,053 individuals

Freathy RM, Weedon MN, Bennett A, Hypponen E, Relton CL, Knight B, Shields B, Parnell KS, Groves CJ, Ring SM, Pembrey ME, Ben-Shlomo Y, Strachan DP, Power C, Jarvelin MR, McCarthy MI, Davey Smith G, Hattersley AT, Frayling TM

Institute of Biomedical and Clinical Science, Peninsula Medical School, Exeter, UK

Am J Hum Genet 2007;80:1150–1161

Background: The role of genes in normal birth weight variation is poorly understood. Type 2 diabetes genes may influence birth weight through maternal genotype, by increasing maternal glycemia in pregnancy, or through fetal genotype, by altering fetal insulin secretion.

Methods: The authors genotyped the rs7903146 polymorphism in TCF7L2, the recently described type 2 diabetes gene, in 15,709 individuals whose birth weight was available from six studies and in 8,344 mothers from three studies.

Results: Each fetal copy of the predisposing allele was associated with an 18-gram (95% CI 7–29) increase in birth weight ($p = 0.001$) and each maternal copy with a 30-gram (95% CI 15–45) increase in offspring birth weight ($p = 2.8 \times 10^{-5}$). Stratification by fetal genotype suggested that the birth weight association was driven by maternal genotype, acting through increased maternal glycemia in pregnancy. In combination with the –30 G/A polymorphism in glucokinase (rs1799884), the 4% of offspring born to mothers carrying three or four risk alleles were 119 g (95% CI 62–172) heavier than were the 32% born to mothers with none (for overall trend, $p = 2 \times 10^{-7}$), comparable to the impact of maternal smoking during pregnancy.

Conclusion(s): The type 2 diabetes susceptibility allele at TCF7L2 was reproducibly associated with birth weight. Common gene variants can substantially influence normal birth weight variation.

Barker's 'fetal origins' hypothesis lead to much interest in the determinants of birth weight. Initial attention was aimed at identifying fetal genes that could influence both fetal growth and also later diabetes risk as, for example, described by this group's fetal insulin hypothesis [16]. They tested this hypothesis by examining the birth weight effect of risk alleles in TCF7L2, the recently described robust type 2 diabetes susceptibility gene. The pertinent finding here is that maternal genotype, but not fetal genotype, was associated with higher birth weight, probably by increasing the mother's glucose levels during pregnancy. The same findings have been seen with glucokinase alleles. Fetal genotype was also apparently associated with higher birth weight, but stratified analyses showed that this association was explained by its correlation with maternal genotype. While fetal genotype might have been hypothesized to lower birth weight by reducing fetal insulin secretion, there was no independent effect of fetal genes.

The take home messages here are that: yes, birth weight is genetically regulated and in combination these factors may have a substantial effect. However, maternal genotype appears to be more important than fetal genotype. Therefore, when exploring the potential effects of fetal genotype on birth weight we always have to consider whether the fetal genotype is simply a marker of maternal genotype effects on the mother's metabolism.

Medical sequencing at the extremes of human body mass

Ahituv N, Kavasar N, Schackwitz W, Ustaszewska A, Martin J, Hebert S, Doelle H, Ersoy B, Kryukov G, Schmidt S, Yosef N, Ruppin E, Sharan R, Vaisse C, Sunyaev S, Dent R, Cohen J, McPherson R, Pennacchio LA
Genomics Division, Lawrence Berkeley National Laboratory, Calif., USA
Am J Hum Genet 2007;80:779–791

Background: Body weight is a quantitative trait with significant heritability in humans. In this study, we sought to explore the relationship between sequence variation in multiple candidate genes and the extremes of human body mass.

Methods: The authors sequenced the coding exons and splice junctions of 58 candidate genes (total 96 Mb) in 379 extremely obese (mean BMI 49) adults attending an obesity clinic and in 378 lean (mean BMI 19.4) and apparently healthy white men and women. Candidate genes included 21 genes associated with monogenic forms of obesity in humans or mice, as well as 37 genes that function in body weight-related pathways.

Results: Cumulatively, the study identified 1,074 genetic variants, averaging one variant per 125 bp of the reference human genome sequence. The monogenic obesity-associated gene group was enriched for rare non-synonymous variants unique to the obese population compared with the lean population. In addition, computational analysis predicted a greater fraction of deleterious variants within the obese cohort.

Conclusion(s): Together, these data suggest that multiple rare alleles contribute to obesity in the population and provide a medical sequencing-based approach to detect them.

Genome-wide association studies rely on chips that genotype hundreds of thousands of SNPs in each individual. However, current technologies preferentially select only common SNPs, with minor allele frequency (MAF) of $>5\%$. There is growing recognition that less frequently occurring SNPs may be causal for common diseases. While, single gene defects are very rare (MAF $<0.1\%$) and explain only a limited number of cases of complex disease such as obesity, 'intermediate rare' SNPs (MAF around $0.1\text{--}5\%$) are more likely to be biologically functional than more common SNPs (MAF $>5\%$) [17] and could therefore contribute substantially to genetic inheritance of common traits and complex diseases. Of the 1,074 genetic variants identified in the current study, 252 were common (MAF $>1\%$), whereas the remaining 822 were rare variants, including 272 non-synonymous (amino acid changing) variants. Among the 21 genes associated with monogenic forms of obesity unique non-synonymous variants were significantly more common in the obese group (46 variants in 41 individuals) than in the lean group (26 variants in 27 individuals; $p < 0.05$). The genes that accounted for the highest differences were MC4R (8 variants in obese vs. 2 in lean), SIM1 (6 in obese vs. 0 in lean), and UCP3 (5 in obese vs. 2 in lean). In familial segregation analysis, the authors found only variable penetrance of these non-synonymous variants (i.e. not all carriers were obese).

The authors therefore conclude that the majority of the genetic etiology of obesity is likely to be influenced by a combination of multiple susceptibility alleles, which even when rare are not likely to be independently causative. Identification of these variants will depend on further technological advances and future widespread availability of high-throughput genome sequencing.

References:

1. Ong K: Population Genetics and Pharmacogenetics; in Carel J-C, Hochberg Z (eds): Yearbook of Pediatric Endocrinology 2007. Basel, Karger, 2007, pp 147–159.
2. Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, et al: Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 2007;447:799–816.
3. Shah NA, Antoine HJ, Pall M, Taylor KD, Azziz R, Goodarzi MO: Association of androgen receptor CAG repeat polymorphism and polycystic ovary syndrome. *J Clin Endocrinol Metab* 2008;93:1939–1945.
4. Lappalainen S, Utriainen P, Kuulasmaa T, Voutilainen R, Jaaskelainen J: Androgen receptor gene CAG repeat polymorphism and X-chromosome inactivation in children with premature adrenarche. *J Clin Endocrinol Metab* 2008;93:1304–1309.
5. Tobias JH, Steer CD, Vilarino-Guell C, Brown MA: Effect of an estrogen receptor-alpha intron 4 polymorphism on fat mass in 11-year-old children. *J Clin Endocrinol Metab* 2007;92:2286–2289.
6. Yaffe K, Lindquist K, Sen S, Cauley J, Ferrell R, Penninx B, et al: Estrogen receptor genotype and risk of cognitive impairment in elders: findings from the Health ABC study. *Neurobiol Aging* 2007 Sep 20. [Epub ahead of print].

7. Grarup N, Rose CS, Andersson EA, Andersen G, Nielsen AL, Albrechtsen A, et al: Studies of association of variants near the HHEX, CDKN2A/B, and IGF2BP2 genes with type 2 diabetes and impaired insulin release in 10,705 Danish subjects: validation and extension of genome-wide association studies. *Diabetes* 2007;56:3105–3111.
8. Lopez-Bermejo A, Petry CJ, Diaz M, Sebastiani G, de Zegher F, Dunger DB, et al: The association between the FTO gene and fat mass in humans develops by the postnatal age of two weeks. *J Clin Endocrinol Metab* 2008;93:1501–1505.
9. Lasky-Su J, Lyon HN, Emilsson V, Heid IM, Molony C, Raby BA, et al: On the replication of genetic associations: timing can be everything! *Am J Hum Genet* 2008;82:849–858.
10. Schulz LO, Bennett PH, Ravussin E, Kidd JR, Kidd KK, Esparza J, et al: Effects of traditional and western environments on prevalence of type 2 diabetes in Pima Indians in Mexico and the US. *Diabetes Care* 2006;29:1866–1871.
11. Raz B, Janner M, Petkovic V, Lochmatter D, Eble A, Dattani MT, et al: Influence of growth hormone (GH) receptor deletion of exon 3 and full-length isoforms on GH response and final height in patients with severe GH deficiency. *J Clin Endocrinol Metab* 2008;93:974–980.
12. Schreiner F, Stutte S, Bartmann P, Gohlke B, Woelfle J: Association of the growth hormone receptor d3-variant and catch-up growth of preterm infants with birth weight of less than 1500 grams. *J Clin Endocrinol Metab* 2007;92:4489–4493.
13. van der Klaauw AA, van der Straaten T, Baak-Pablo R, Biermasz NR, Guchelaar HJ, Pereira AM, et al: Influence of the d3-growth hormone receptor isoform on short-term and long-term treatment response to growth hormone replacement in growth hormone deficient adults. *J Clin Endocrinol Metab* 2008 Apr 8. [Epub ahead of print].
14. Carrascosa A, Audi L, Esteban C, Fernandez-Cancio M, Andaluz P, Gussinye M, et al: Growth hormone (GH) dose, but not exon 3-deleted/full-length GH receptor polymorphism genotypes, influences growth response to two-year GH Therapy in Short Small-for-Gestational-Age Children. *J Clin Endocrinol Metab* 2008;93:147–153.
15. Carrascosa A, Audi L, Fernandez-Cancio M, Esteban C, Andaluz P, Vilaro E, et al: The exon 3-deleted/full-length growth hormone receptor polymorphism did not influence growth response to growth hormone therapy over two years in prepubertal short children born at term with adequate weight and length for gestational age. *J Clin Endocrinol Metab* 2008;93:764–770.
16. Hattersley AT, Tooke JE: The fetal insulin hypothesis: an alternative explanation of the association of low birthweight with diabetes and vascular disease. *Lancet* 1999;353:1789–1792.
17. Gorlov IP, Gorlova OY, Sunyaev SR, Spitz MR, Amos CI: Shifting paradigm of association studies: value of rare single-nucleotide polymorphisms. *Am J Hum Genet* 2008;82:100–112.