A Meeting of the Nutrition Society, hosted by the Scottish Section, was held at King's College Conference Centre, University of Aberdeen on 26–27 March 2012

Conference on 'Future food and health'
Symposium II: Diet–gene interactions; implications for future diets and health

The genetic contribution to disease risk and variability in response to diet: where is the hidden heritability?

Anne Marie Minihane

Department of Nutrition, Norwich Medical School, University of East Anglia, Norwich NR4 7TJ, UK

Ten years ago, it was assumed that disease risk prediction and personalised nutrition based on genetic information would now be in widespread use. However, this has not (yet) transpired. The interaction of genetic make-up, diet and health is far more complex and subtle than originally thought. With a few notable exceptions, the impact of identified common genetic variants on phenotype is relatively small and variable in their penetrance. Furthermore, the known variants account for only a fraction of what we believe to be the total genetic contribution to disease risk and heterogeneity in response to environmental change. Here, the question 'how far have we progressed and are we likely to get there' (Rimbach and Minihane, 2009) is revisited with regard to the translation of genetic knowledge into public health benefit. It is concluded that progress to date has been modest. It is hoped that recent technological developments allowing the detection of rarer variants and future use of more hypothesis-driven targeted data analysis will reveal most of the currently 'hidden' significant genetic variability.

Heritability: Nutrigenetics: Nutrigenomics: GWAS: Sequencing: *APOE* genotype

The use of genetic information for the detection of disease risk and the provision of tailored therapeutic strategies or lifestyle advice holds enormous potential to result in population health benefits. It is hoped that a comprehensive understanding of nutrigenetics, which refers to the interactive impact of genetic variation (genotype) and diet composition in defining health and risk of disease (phenotype), may lead to the partial replacement of current generic 'one size fits all' dietary guidelines with more efficacious stratified dietary advice based in part on personal genetic information. However, to date, only a relatively small fraction of the estimated total genetic contribution to phenotype has been identified. This begs the critical question of whether our initial estimates of heredibility are inflated or whether our current methods for the detection of genetic variation or the interpretation of data have been insensitive or misleading in their approaches. Erroneously overestimated heritability may occur as a result of non-additive genetic effects, gene–environment interactions or shared environment by family members[1]; although the possibility cannot be discounted, there is little evidence to date that this is the case[2,3]. It appears more likely that a large proportion of genetic contribution to phenotype is as yet undiscovered.

The interaction between diet, genetics and health is complex and occurs at multiple levels (Fig. 1). Firstly, the influence of genotype on phenotype is not homogenous and can be more or less pronounced depending on diet composition and nutrition status of the individual. Genetic variation influences food preference, appetite and satiety and therefore overall diet composition. Once consumed, the amount of a particular dietary component absorbed and its subsequent metabolism, tissue uptake and elimination from the body are under genetic regulation. The influence of tissue concentrations of a dietary component on metabolism is again influenced by genotype, via for example, genetic variability influencing cell signalling mechanisms. Furthermore, adding to the complexity is the fact that the influence of genotype on homoeostasis and the metabolism and bioactivity of ditary components is not homogenous

---

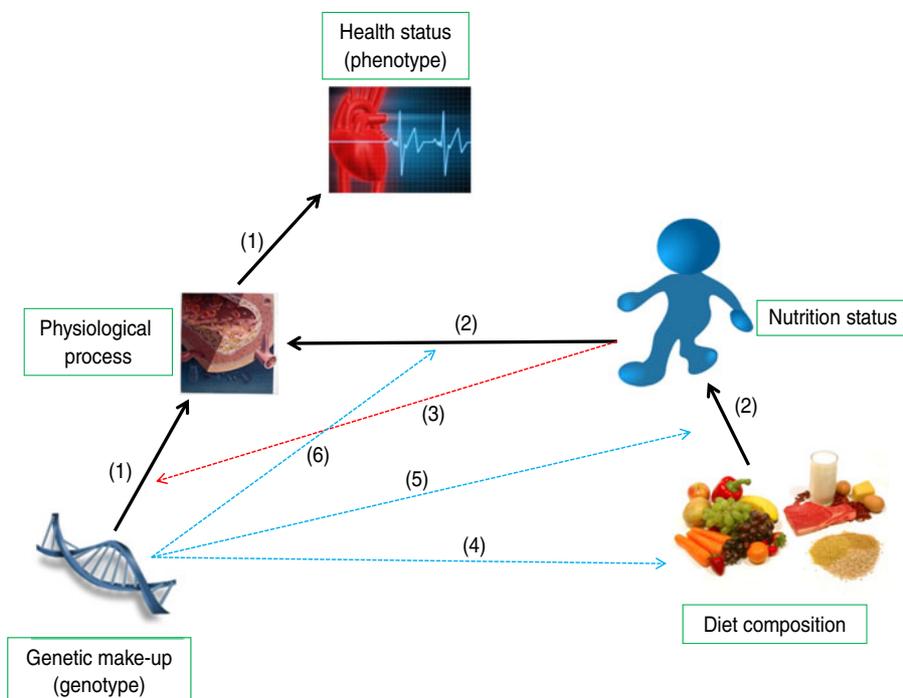**Corresponding author:** Professor A. M. Minihane, fax + 44 1603 593752, email a.minihane@uea.ac.uk

Fig. 1. (colour online) The complexity of genotype–diet–phenotype interactions: (1) physiological status and phenotype are influenced by genotype; (2) diet composition influences tissue concentration and form of individual dietary components which in turn influences physiological status; (3) the penetrance of an individual gene variant is influenced by nutritional status; (4) although as yet relatively under investigated there is evidence that the food consumed is influenced by genotype, with genetic variation affecting food preferences, appetite and satiety[4,5]; (5) once ingested, the digestion of food, the absorption efficiency of nutrients and non-nutrients, their post-absorptive metabolism and tissue uptake, utilisation and storage and elimination from the body are under genetic control; (6) the influence of a particular tissue status of a dietary component on phenotype is influenced by genotype via an array of mechanisms including genetic variability in cell signalling pathways, transcription factor activity, biotransformation enzymes, etc.

and is influenced by a whole range of variables such as sex, ethnicity, drug use and other lifestyle variables (see section on Genome-wide association studies).

The current paper will provide an update of an earlier review entitled 'Nutrigenetics and personalised nutrition: how far have we progressed and are we likely to get there' published in 2009[6]. In the interim, a large body of data from genome-wide association studies (GWAS) has been published. Their contribution to our understanding of disease pathology and its genetic component will be considered, along with the strengths and limitations of GWAS approaches. The likely advances associated with emerging sequencing technologies will be briefly discussed.

Although there are a small number of notable exceptions, it is becoming apparent that the individual impact of the most common type of genetic variation, namely SNP, is small, with effect sizes in the 0–10 % range. Therefore, it is unlikely that personalisation of risk of disease or dietary advice will be based on a small number of common variants with individual large effects. An alternative hypothesis as will be discussed is that a large proportion of variability is accounted for by rarer variants with large biological impacts, information not currently captured by GWAS.

Although the situation is rapidly changing, to date, due to lack of collection of dietary data, GWAS have contributed little to our nutrigenetic understanding which is largely derived from candidate-gene studies. Candidate-gene approaches have been used to identify and quantify the impact of variants such as *APOE epsilon* and the *Methylenetetrahydrofolate reductase (MTHFR) C677T* genotypes on disease risk and response to diet. An update on the recent literature for these gene loci will be provided.

Although the authors recognise that in addition to variation to the DNA code itself, the epigenetic status of genes influence phenotype, diet × epigenetic × phenotype interactions are outside the scope of the current paper, but have been reviewed extensively elsewhere[7–9].

Genome-wide association studies: how have they informed and misinformed us

In contrast to candidate-gene studies which focus on variants in genes with a known metabolic role, GWAS are not hypothesis driven. The advantage (but also challenge) of the hypothesis free approach used in GWAS is that it has the power to identify novel biological pathways associated with a phenotype or dietary component of interest.

An example of this paradigm is the large research interest in uncovering the biological action of the fat mass and obesity associated (FTO) protein following the identification of the association of a SNP in the FTO gene with BMI in a 2007 GWAS study[10], with the authors Frayling *et al.* concluding that 'FTO is a gene of unknown function in an unknown pathway'[11].

In GWAS, genetic variability is quantified in a group of cases *v.* matched controls, in order to establish disease associated variants. The standard SNP arrays typically have $300,000–2 \times 10^6$ tagging SNP which are correlated with (derived from HapMap[12]) and provide information on 80–90% of common variation (frequency >5%), but far less for the low frequency variants (0·01–5%) and virtually none for the rare variants[13].

The first GWAS output was published in 2005 and identified a polymorphism in complement H to be linked with age-related macular disease[14]. This led the way to an explosion of activity, and to date, GWAS has identified over 1600 variants associated with 250 traits[15] and has had some success in identifying a large component of the genetic basis of particular phenotypes. For example, as reviewed by Manolio *et al.* for age-related macular degeneration five loci have been identified which collectively explain 50% of total heritability[2]. In 2010, a meta-analysis of forty-six cohort studies confirmed ninety-five loci predictors of plasma lipids (total cholesterol, HDL-cholesterol, LDL-cholesterol and TAG) which explained 25–30% of the genetic component of these traits[16]. However, for the majority of polygenic traits, the identified variants only account for a much lower proportion of the total estimated heritability which varies from 20 to 80% depending on the phenotype of interest[2,17]. For BMI and obesity, thirty-two individual loci have been identified and confirmed, but the effect size of each individual variant is small. The largest effect is evident for FTO, with each risk allele increasing BMI by on average 0·39 kg/m^2 and obesity risk by 1·20[18]. However, collectively the thirty-two loci explain only 1·45% of the phenotypic variation in BMI, equivalent to 2–4% of heritability[19].

There are likely to be a number of reasons why GWAS has resulted in only modest capture of heritability[20]. Firstly, although the gene variant may have been identified by GWAS, it may not have emerged as significant or its effect size may be underestimated for a number of reasons which include:

1. Use of stringent *P* value (typically $P < 1 \times 10^{-8}$) to compensate for multiple testing and eliminate false positive results. This may result in failure to detect many true signals and we may be 'correcting away the hidden heritability'[21]. Alternative approaches to the use of strict *P* values, such as multi-stage confirmation of significant SNP in subsequent datasets[22,23], have led to the identification of further variants of interest;
2. Imprecise phenotyping[13]. For disease outcomes, patients in the 'case group' often present with a range of related conditions with variable genetic aetiology. For example, in the cardiovascular field, myocardial infarction, ischaemic heart disease and coronary stenosis are often pooled, although they have both

common and separate aetiological components. For many outcomes, such as blood pressure, the precision of the measurement is problematic, while for others there is a large intra-individual variability such as pro-inflammatory cytokines and C-reactive protein, which means the trait is imprecisely captured.
3. Control group of questionable quality[24]. Often an individual in the control group does not have a clinical diagnosis of the primary outcome but is a registered patient for an alternative condition whose risk may also be impacted by the identified gene variants, therefore underestimating the effect size of the variant.
4. True causal variants incompletely surveyed are not in full linkage disequilibrium with *tagging* SNP.
5. A number of causal variants may exist in one locus, with only one *tagging* SNP chosen, which may result in an underestimation of the total heritability accounted for by that particular region.

Secondly, it is plausible and increasingly demonstrated that rarer single nucleotide changes, or structural variants such as copy number variations[25,26], which are far more common than originally thought, could make a significant contribution to hidden variability (Fig. 2)[26–28]. Next generation sequencing is becoming increasingly feasible and affordable and is in more widespread use as a research tool[2,29,30]. This technology provides a complete map of an individual's genome, overcoming limitations associated with SNP *tagging* in GWAS and allowing the detection of less common variants. The 1000 Genomes Project, which will include far more than 1000 aims to capture all variants with <1% frequency and >0·1% in protein coding regions (exome)[30]. It is hoped that this technology will detect much of the hidden heritability.

Thirdly, the broad sense heritability model posits that once 'unveiled' neither common variants with modest impact nor rare alleles with high penetrance are likely to explain away missing heritability. It theorises that known genetic variation in the form of interactions, between allele pairs (dominance), between alleles in different genes (epistasis) and between genotype and environment (including diet composition) or physiological variables, explain a large proportion of inheritance.

Impact of physiological variables on genotype–phenotype associations

Currently, in genetic studies, populations are considered as single entities. It is becoming increasingly apparent that population genetic associations often under- or over-estimate the effect in subgroups. At this stage, it is too early to say what the contribution of differential penetrance in population subgroups to missing heritability is likely to be, but is likely to be significant.

For biological processes with a known influence of sex, such as adiposity and plasma lipids, it is plausible to assume that the impact of genetic variation on these phenotypes may vary between sexes, with numerous demonstrations of this now available. In the Framingham Heart Offspring cohort, no variant showed genome wide significance for measures of obesity. However, sex
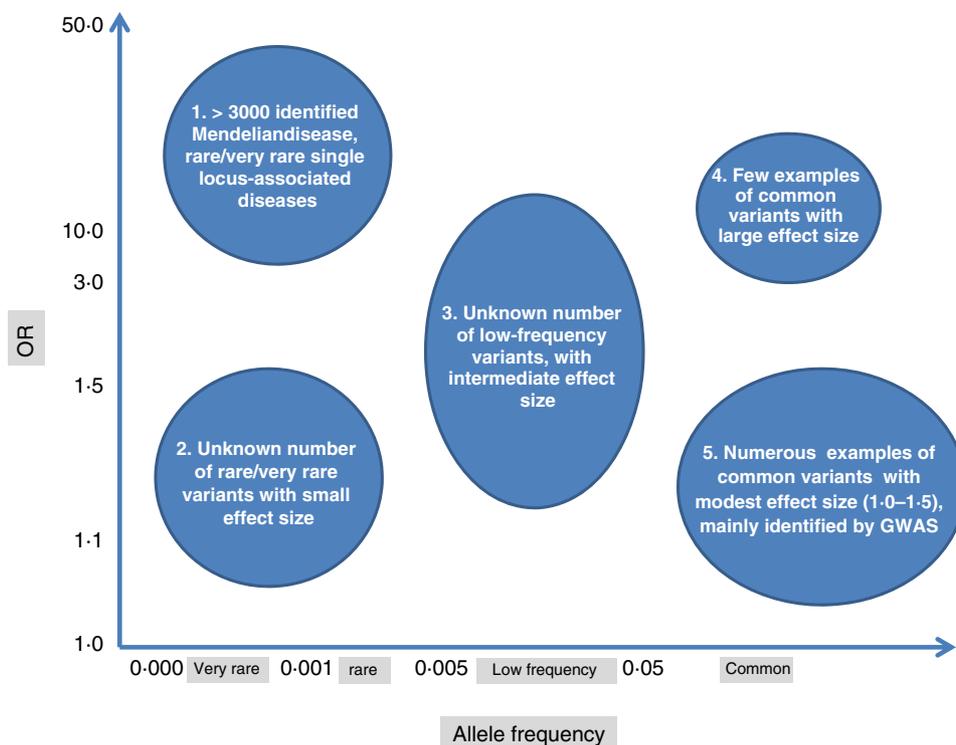
Fig. 2. (colour online) Identification of genetic variants of various frequencies and effect sizes. (Adapted from[2,24].) (1) Examples of Mendelian diseases include Huntington's disease, sickle cell anaemia and cystic fibrosis. Lifestyle, including diet composition often has a minimal effect on disease severity; (2) these variants are difficult to identify and given their rarity and small effect size their identification is not a priority; (3) currently a few of these disease-associated variants have been identified possibly due to their lack of representation on currently used genome-wide association studies (GWAS) arrays. Increased use of sequencing technologies and redesign of traditional arrays is predicted to substantially increase their detection rates; (4) one example of such a genotype is the association between the *APOE4* allele and risk of Alzheimer's disease; (5) to date >95% of the identified common variants associated with disease have modest effects sizes 1·0–1·5, and explain only a small proportion of the total heritability of the phenotype.

dimorphism was evident with four polymorphisms in the lysophospholipase-like protein 1 (LYPLAL1) locus which encodes for a lipase/esterase in adipose tissue, having divergent effects in men and women[31]. In a meta-analysis of the available GWAS, and using the waist:hip ratio as a measure of body fat topography, along with the LYPLAL1 signal, thirteen new loci for the waist:hip ratio were evident. Seven of these displayed dimorphism with a stronger effect in women[32]. Using a candidate-gene approach we have reported a number of significant associations between common SNP and the postprandial lipaemic response, a CVD determinant of ever increasing prevalence[33]. For the leptin receptor (Gln223Arg, rs1137101) and *APOA5* (−1131T>C, rs662799) variants, the effect of genotype was only evident in men[34,35]. For example, for leptin receptor, a 20% lower postprandial TAG response was evident in ArgArg *v.* GlnGln homozygotes with men and women combined, with a 35% difference in the men only group and no effect of genotype in women (Table 1)[35].

There is also evidence of racial/ethnic differences in the physiological impact of particular variants. Results from the Population Architecture Using Genomics and Epidemiology Consortium[36] and a meta-analysis of forty-six individual GWAS[16], which aimed to identify genome-wide signals for BMI and plasma lipids, respectively, showed a considerable overlap between associations in those of European and Asian ancestry, with more modest replication in more traditional populations such as African Americans and American Indians. Linkage disequilibrium patterns suggest that *tagging* SNP used in GWAS for Europeans may not adequately capture the genetic variation in other ethnic groups. Apparent ethnic differences in genotype × phenotype associations from GWAS may be in part attributable to differences in the habitual diets between populations.

Genome-wide association studies in the study of nutrigenetic interactions

Although GWAS methodology and modelling do not lend themselves well to the direct study of genotype × diet × disease associations, in part due to the fact that the sample size needed would be enormous[37], an increasing number of GWAS have nutrient status as their primary endpoint[38,39]. Using a genome-wide approach, Wang *et al.* identified three loci near the genes for cholesterol

Table 1. Effect of the leptin receptor rs113701 genotype on post-prandial lipaemia in UK adults (adapted from[35])
(Mean values with their standard errors)

Group	GlnGln	GlnArg	ArgArg
All			
Mean	1226·7	1063·0	904·2
SEM	71·6*	42·8	56·4
Males			
Mean	1484·0	1249·5	964·0
SEM	66·8	47·3	97·1
Females			
Mean	887·2	827·6	888·9
SEM	65·8	54·8	89·6

*Values are group mean postprandial TAG area under the curves (mmol/l × 480 min) following consumption of test meals containing 49 g (0 min) and 29 g (330 min) total fat.

synthesis, hydroxylation and vitamin D transport as significant predictors of vitamin D status, which could potentially be used to set vitamin D intake recommendations[39]. The output from GWAS has also informed the choice of variants for a more focused study of genotype × diet interaction in human epidemiology and intervention studies and in targeted replacement animal models. The identification of the *FTO* genotype by GWAS has led to a flurry of activity examining its impact on food intake, satiety and appetite and its interaction with macronutrient composition in determining BMI and risk of obesity[10,40,41]. In the *RISCK* randomised control trial, the impact of forty GWAS identified lipid associated SNP on the response of plasma lipids to a low-saturated fat diet were determined[42]. Relatively recent availability of GWAS data for cohorts for which participants have detailed dietary data, such as the Nurse's Health Study, Framingham Heart Studies and EPIC is likely to make a significant contribution to our nutrigenetic understanding in the near future.

To date, the majority of nutrigenetic information is derived from candidate-gene studies. Two of the most widely researched variants using this approach are *APOE epsilon* and *MTHFR* SNP, with approximately 6000 and 3000 associated published articles, respectively. However, despite extensive research focus there remains considerable uncertainty regarding the relative impact of these common genotypes on health and response to dietary change, which demonstrates the complexity of the interactions.

The renowned *APOE epsilon* genotype

As its name suggests apoE is an important modulator of many stages of lipoprotein metabolism and is the main lipid transporter in the central nervous system. Since its original identification, its pleiotropic nature has been realised with apoE now known to regulate immunity and inflammation, oxidative status and β-amyloid metabolism in the central nervous system. Two non-synonymous SNP in the *APOE* gene, result in three specific apoE protein isoforms namely apoE2, apoE3 and apoE4. The *APOE* genotype was originally described as a genetic contributor to CVD, with *APOE4* carriers at increased risk. Over time,

and with ever larger meta-analyses it has become apparent that at a population level the impact on CVD risk is marginal[43,44], and often does not emerge as a significant signal in GWAS (Table 2), although in individual population subgroups such as smokers, *the APOE* genotype remains a highly significant risk factor. In both the Northwick Park and Framingham Offspring cohorts, risk of CVD was about 2-fold higher in smokers who were wild-type *E3/E3 v. E4* carriers (Table 3)[46,47].

More recently there has been much interest in the *APOE* genotype as a longevity gene. Although not fully consistent, the *APOE4* allele has emerged as being associated with a shorter life-span[48,49].

Perhaps the most consistent and consequential common genotype-disease association described to date is the impact of the *APOE* genotype on risk of age-related cognitive decline and Alzheimer's disease. As summarised in Table 2, the *APOE3/E4* and *APOE4/E4* individuals are at approximately 3–4- and 12–16-fold increased risk of Alzheimer's disease and have a much earlier age of onset[45]. The clinical significance of the genotype is demonstrated by the fact that almost 50 and 10 % of Alzheimer's disease patients are *APOE3/E4* and *APOE4/E4*, whereas these genotype subgroups represent approximately 20 and 2 %, respectively, of the general population[50,51]. Interestingly, when James D. Watson was presented with his genetic information, his being the first genome sequenced by next-generation sequence technologies, he elected not to know his *APOE* genotype[52].

The aetiological basis of this association is likely to be multi-factorial with *APOE4* carriers having altered central nervous system lipid metabolism, vascular dysfunction, increased neuroinflammation and oxidative stress, β-amyloid deposition, synaptic dysfunction and impaired neurogenesis[53]. This is of wide interest in the quest for further establishment of the Alzheimer's disease pathological process and its treatment and prevention.

APOE genotype and its response to dietary change

Given the association between the *APOE4* allele and cognitive decline and CVD risk in particular individuals, there is wide interest in the identification of dietary strategies to reduce disease risk in this large genotype subgroup. Research into the impact of the *APOE* genotype in response to diet has almost exclusively focused on plasma lipid response to altered dietary fat composition.

Overall, the evidence is suggestive that *APOE4* carriers are most responsive to the plasma cholesterol modulating impact of total fat, cholesterol, saturated fat intake and long chain *n*-3 PUFA, EPA and DHA intake[50]. However, with a few exceptions, the study of the impact of genotype has been conducted using retrospective genotyping, where lack of power has often led to inconclusive findings. In our original study, using retrospective genotype analysis, we observed a LDL-cholesterol raising effect of high-dose fish oil supplementation (3 g EPA + DHA per d) in *APOE4* carriers which may in part negate the cardioprotective benefits[54]. In a subsequent adequately powered recruitment on the basis of genotype approach, we confirmed these earlier findings and demonstrated that

Table 2. Meta-analysis of the impact of the *APOE* genotype on CVD and Alzheimer's disease risk

Reference	E2 carriers*			E4 carriers*
Song et al.[44]				
n 48 studies, CHD risk	0·93 (0·83, 1·05)†			1·30 (1·18, 1·43)
Bennet et al.[43]				
n 17 studies <500 cases, CAD risk	1·00 (0·91, 1·11)			1·66 (1·50, 1·84)
n 104 studies ≥ 500 cases, CAD risk	0·80 (0·70, 0·90)			1·06 (0·99, 1·13)

	E2/E2	E2/E3	E3/E4	E4/E4
Bertram et al. (Caucasians)[45]				
n 20 studies, clinical/autopsy, AD risk		0·6 (0·3, 1·2)	4·3 (3·3, 3·5)	15·6 (10·9, 22·5)
n 8 studies, population based, AD risk		0·3 (0·2, 0·6)	2·8 (2·3, 3·5)	11·8 (7·0, 19·8)

CAD, coronary artery disease; AD, Alzheimer's disease.
*E2 carriers include E2/E2 and E2/E3; E4 carriers include E3/E4 and E4/E4.
†OR (95% CI) with the wild-type E3/E3 genotype as the reference.

Table 3. Impact of *APOE* genotype on CVD risk in smokers

Reference	E3/E3	E2 carriers*	E4 carriers*
Humphries et al.[46]†			
Northwick Park Study			
Ex-smokers	1·74 (0·93, 2·37)	0·47 (0·11, 1·94)	0·74 (0·35, 1·55)
Smokers	1·47 (0·87, 2·51)	0·85 (0·30, 2·43)	2·79 (1·59, 4·91)
Talmud et al.[47]‡			
Framingham Offspring Study			
Ten cigarettes per d	1·16 (1·08, 1·30)	1·61 (1·09, 2·36)	1·60 (1·17, 2·20)
Forty cigarettes per d	1·98 (1·36, 2·88)	4·97 (2·75, 8·90)	4·53 (2·72, 7·55)

*E2 carriers include E2/E2 and E2/E3; E4 carriers include E3/E4 and E4/E4.
†Hazard ratio (95% CI) with all genotypes combined, never smokers as reference.
‡Hazard ratio (95% CI) with E3/E3 never smokers as reference.

it is likely to be the DHA rather than EPA in fish which raises cholesterol[55]. In more recent publications, also using prospective recruitment, we have reported no significant *APOE* genotype × DHA × LDL-cholesterol interaction following lower intake of fish oils (<2 g EPA + DHA per d)[9] or against a background of high saturated fat intake[56].

Interestingly, there is also inconsistent evidence to suggest that the purported cognitive benefits of increased EPA + DHA status may be *APOE* genotype dependent, with no benefit in the *APOE4* carriers[57]. This genotype × diet interaction requires substantiation but may underlie a differential long-chain *n*-3 PUFA uptake and partitioning in the central nervous system.

Owing to the population prevalence of the homozygous *APOE4/E4* genotype (2%), studies to date have largely compared response in the *APOE4* carriers (largely *APOE3/E4* individuals) *v.* non-carriers. Although there is limited supporting evidence, it is likely that the *APOE4/E4* individuals are most responsive to dietary fat manipulation. Quantification of the response in this genotype group is important given its impact on risk of cognitive decline.

The often cited methylenetetrahydrofolate reductase C677T variant

MTHFR is an important enzyme in folate/homocysteine metabolism. It provides a clear demonstration of how

genetic information could be used to provide targeted dietary advice in an at-risk population subgroup. A homozygous mutant genotype (TT, rs 1801133), which has a frequency of approximately 10% worldwide[58], is associated with reduced enzyme activity[59]. Its subsequent impact on homocysteine concentrations, blood pressure and risk of diseases such as cancer and CVD, is variable and has been shown to be dependent on factors such as sex and ethnicity[60,61]. The penetrance of the genotype is also dependent on vitamin B status (folate, riboflavin, vitamin B_6 and vitamin B_{12})[60,62]. In two complementary intervention trials, Scott and colleagues elegantly demonstrated that riboflavin (a co-factor of MTHFR) lowered blood pressure in patients with the MTHFR *TT* genotype which is independent of background use of prescribed drugs[63,64]. Overall, there is considerable evidence to indicate that adequate vitamin B status is likely to abrogate the negative physiological impact of this genotype on disease risk.

Conclusion: the road ahead

The first draft of the majority (about 90%) of the sequence of the human genome was published in a *Nature* article entitled 'Initial sequencing and analysis of the human genome' a little over a decade ago (February 2001)[65] with the complete sequence (about 99·7%) available in 2004[66]. At the time of availability such information was considered by many to be the panacea and one of the greatest ever

medical achievements. Ten years on, many consider progress based on the human genome to be limited, but as reviewed by Eric Lander (who was lead author of the original 2001 publication), this may be a rather harsh assessment[17]. Of the 3000 Mendelian (monogenic) disorders whose genetic basis is known, the locus of the vast majority have been identified since 2001. Furthermore, GWAS and HapMap approaches have led to the identification of 1600 variants associated with 250 traits, which has established numerous novel pathological pathways and contributed to the development of novel therapies. Although sufficient to indicate the potential of the technology, there has been limited success in the use of genetic information for disease prediction and personalisation of therapeutics or preventative advice, with much of the estimated heritable component of disease risk and response to diet unaccounted for. Based on the available information, it appears that rather than being overestimated, the heritability is dark matter (i.e. it is real but we cannot see it yet), attributable to as yet undetected rare variants, or the underestimation of the impact of known variants. The wider use of sequencing will provide information on variants with a frequency of <5%. More detailed and precise characterisation of study participants in genetic studies and more sophisticated modelling will undoubtedly lead to the detection of variants of particular importance in population subgroups. Most of the benefits of genetics in public health remain to be realised and we undoubtedly have a long way to go. In the words of Churchill, it feels like the 'end of the beginning' rather than the 'beginning of the end'.

Acknowledgements

References

1. Zuk O, Hechter E, Sunyaev SR et al. (2012) The mystery of missing heritability: genetic interactions create phantom heritability. *Proc Natl Acad Sci USA* **109**, 1193–1198.
2. Manolio TA, Collins FS, Cox NJ et al. (2009) Finding the missing heritability of complex diseases. *Nature* **461**, 747–753.
3. Visscher PM, Medland SE, Ferreira MA et al. (2006) Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *PLoS Genet* **2**, e41.
4. Keller KL, Liang LC, Sakimura J et al. (2012) Common variants in the CD36 gene are associated with oral fat perception, fat preferences, and obesity in African Americans. *Obesity (Silver Spring)* **20**, 1066–1073.
5. Yeung EH, Zhang C, Chen J et al. (2011) Polymorphisms in the neuropeptide Y gene and the risk of obesity: findings from two prospective cohorts. *J Clin Endocrinol Metab* **96**, E2055–E2062.
6. Rimbach G & Minihane AM (2009) Nutrigenetics and personalised nutrition: how far have we progressed and are we likely to get there? *Proc Nutr Soc* **68**, 162–172.
7. Burdge GC, Hoile SP & Lillycrop KA (2012) Epigenetics: are there implications for personalised nutrition? *Curr Opin Clin Nutr Metab Care* **15**, 442–447.
8. Jimenez-Chillaron JC, Diaz R, Martinez D et al. (2012) The role of nutrition on epigenetic modifications and their implications on health. *Biochimie* **94**, 2242–2263.
9. Caslake MJ, Miles EA, Kofler BM et al. (2008) Effect of sex and genotype on cardiovascular biomarker response to fish oils: the FINGEN study. *Am J Clin Nutr* **88**, 618–629.
10. Tung YC & Yeo GS (2011) From GWAS to biology: lessons from FTO. *Ann NY Acad Sci* **1220**, 162–171.
11. Frayling TM, Timpson NJ, Weedon MN et al. (2007) A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* **316**, 889–894.
12. Altshuler DM, Gibbs RA, Peltonen L et al. (2010) Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52–58.
13. Marian AJ & Belmont J (2011) Strategic approaches to unraveling genetic causes of cardiovascular diseases. *Circ Res* **108**, 1252–1269.
14. Klein RJ, Zeiss C, Chew EY et al. (2005) Complement factor H polymorphism in age-related macular degeneration. *Science* **308**, 385–389.
15. Hindorff LA, MacArthur J, Wise A et al. (2011) A Catalog of Published Genome-Wide Association Studies. http://www.genome.gov/gwastudies/
16. Teslovich TM, Musunuru K, Smith AV et al. (2010) Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707–713.
17. Lander ES (2011) Initial impact of the sequencing of the human genome. *Nature* **470**, 187–197.
18. Loos RJ (2012) Genetic determinants of common obesity and their value in prediction. *Best Pract Res Clin Endocrinol Metab* **26**, 211–226.
19. Speliotes EK, Willer CJ, Berndt SI et al. (2010) Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* **42**, 937–948.
20. Gibson G (2011) Rare and common variants: twenty arguments. *Nat Rev Genet* **13**, 135–145.
21. Williams SM & Haines JL (2011) Correcting away the hidden heritability. *Ann Hum Genet* **75**, 348–350.
22. Easton DF, Pooley KA, Dunning AM et al. (2007) Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* **447**, 1087–1093.
23. Consortium IMSG (2010) Comprehensive follow-up of the first genome-wide association study of multiple sclerosis identifies KIF21B and TMEM39A as susceptibility loci. *Hum Mol Genet* **19**, 953–962.
24. McCarthy MI, Abecasis GR, Cardon LR et al. (2008) Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* **9**, 356–369.
25. Redon R, Ishikawa S, Fitch KR et al. (2006) Global variation in copy number in the human genome. *Nature* **444**, 444–454.
26. Almal SH & Padh H (2012) Implications of gene copy-number variation in health and diseases. *J Hum Genet* **57**, 6–13.
27. Fanciulli M, Petretto E & Aitman TJ (2010) Gene copy number variation and common human disease. *Clin Genet* **77**, 201–213.
28. Jakobsson M, Scholz SW, Scheet P et al. (2008) Genotype, haplotype and copy-number variation in worldwide human populations. *Nature* **451**, 998–1003.
29. Bras J, Guerreiro R & Hardy J (2012) Use of next-generation sequencing and other whole-genome strategies to dissect neurological disease. *Nat Rev Neurosci* **13**, 453–464.
30. Clarke L, Zheng-Bradley X, Smith R et al. (2012) The 1000 Genomes Project: data management and community access. *Nat Methods* **9**, 459–462.

31. Benjamin AM, Suchindran S, Pearce K *et al.* (2011) Gene by sex interaction for measures of obesity in the Framingham Heart study. *J Obes* **2011**, 329038.
32. Heid IM, Jackson AU, Randall JC *et al.* (2010) Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat Genet* **42**, 949–960.
33. Jackson KG, Poppitt SD & Minihane AM (2012) Post-prandial lipemia and cardiovascular disease risk: inter-relationships between dietary, physiological and genetic determinants. *Atherosclerosis* **220**, 22–33.
34. Olano-Martin E, Abraham EC, Gill-Garrison R *et al.* (2008) Influence of apoA-V gene variants on postprandial triglyceride metabolism: impact of gender. *J Lipid Res* **49**, 945–953.
35. Jackson KG, Delgado-Lista J, Gill R *et al.* (2012) The leptin receptor Gln223Arg polymorphism (rs1137101) mediates the postprandial lipaemic response, but only in males. *Atherosclerosis* **225**, 135–141.
36. Fesinmeyer MD, North KE, Ritchie MD *et al.* (2012) Genetic risk factors for BMI and obesity in an ethnically diverse population: results from the population architecture using genomics and epidemiology (PAGE) study. *Obesity (Silver Spring)* Epublication ahead of print version.
37. Thomas D (2010) Gene–environment-wide association studies: emerging approaches. *Nat. Rev. Genet.* **11**, 259–272.
38. Lemaitre RN, Tanaka T, Tang W *et al.* (2011) Genetic loci associated with plasma phospholipid *n*-3 fatty acids: a meta-analysis of genome-wide association studies from the CHARGE Consortium. *PLoS Genet* **7**, e1002193.
39. Wang TJ, Zhang F, Richards JB *et al.* (2010) Common genetic determinants of vitamin D insufficiency: a genome-wide association study. *Lancet* **376**, 180–188.
40. McCaffery JM, Papandonatos GD, Peter I *et al.* (2012) Obesity susceptibility loci and dietary intake in the Look AHEAD Trial. *Am J Clin Nutr* **95**, 1477–1486.
41. Razquin C, Marti A & Martinez JA (2011) Evidences on three relevant obesogenes: MC4R, FTO and PPARgamma. Approaches for personalized nutrition. *Mol Nutr Food Res* **55**, 136–149.
42. Walker CG, Loos RJ, Olson AD *et al.* (2011) Genetic predisposition influences plasma lipids of participants on habitual diet, but not the response to reductions in dietary intake of saturated fatty acids. *Atherosclerosis* **215**, 421–427.
43. Bennet AM, Di Angelantonio E, Ye Z *et al.* (2007) Association of apolipoprotein E genotypes with lipid levels and coronary risk. *JAMA: J Am Med Assoc* **298**, 1300–1311.
44. Song Y, Stampfer MJ & Liu S (2004) Meta-analysis: apolipoprotein E genotypes and risk for coronary heart disease. *Ann Intern Med* **141**, 137–147.
45. Bertram L, McQueen MB, Mullin K *et al.* (2007) Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet* **39**, 17–23.
46. Humphries SE, Talmud PJ, Hawe E *et al.* (2001) Apolipoprotein E4 and coronary heart disease in middle-aged men who smoke: a prospective study. *Lancet* **358**, 115–119.
47. Talmud PJ, Stephens JW, Hawe E *et al.* (2005) The significant increase in cardiovascular disease risk in APOEepsilon4 carriers is evident only in men who smoke: potential relationship between reduced antioxidant status and ApoE4. *Ann Hum Genet* **69**, 613–622.
48. McKay GJ, Silvestri G, Chakravarthy U *et al.* (2011) Variations in apolipoprotein E frequency with age in a pooled analysis of a large group of older people. *Am J Epidemiol* **173**, 1357–1364.
49. Murabito JM, Yuan R & Lunetta KL (2012) The search for longevity and healthy aging genes: insights from epidemiological studies and samples of long-lived individuals. *J Gerontol A Biol Sci Med Sci* **67**, 470–479.
50. Minihane AM, Jofre-Monseny L, Olano-Martin E *et al.* (2007) ApoE genotype, cardiovascular risk and responsiveness to dietary fat manipulation. *Proc Nutr Soc* **66**, 183–197.
51. Ward A, Crean S, Mercaldi CJ *et al.* (2012) Prevalence of apolipoprotein E4 genotype and homozygotes (APOE e4/4) among patients diagnosed with Alzheimer's disease: a systematic review and meta-analysis. *Neuroepidemiology* **38**, 1–17.
52. Wheeler DA, Srinivasan M, Egholm M *et al.* (2008) The complete genome of an individual by massively parallel DNA sequencing. *Nature* **452**, 872–876.
53. Hauser PS, Narayanaswami V & Ryan RO (2011) Apolipoprotein E: from lipid transport to neurobiology. *Prog Lipid Res* **50**, 62–74.
54. Minihane AM, Khan S, Leigh-Firbank EC *et al.* (2000) ApoE polymorphism and fish oil supplementation in subjects with an atherogenic lipoprotein phenotype. *Arterioscler Thromb Vasc Biol* **20**, 1990–1997.
55. Olano-Martin E, Anil E, Caslake MJ *et al.* (2010) Contribution of apolipoprotein E genotype and docosahexaenoic acid to the LDL-cholesterol response to fish oil. *Atherosclerosis* **209**, 104–110.
56. Carvalho-Wells AL, Jackson KG, Lockyer S *et al.* (2012) APOE genotype influences the triglyceride and C-reactive protein response to altered dietary fat intake in UK adults. *Am J Clin Nutr* (In the Press).
57. Barberger-Gateau P, Samieri C, Feart C *et al.* (2011) Dietary omega 3 polyunsaturated fatty acids and Alzheimer's disease: interaction with apolipoprotein E genotype. *Curr Alzheimer Res* **8**, 479–491.
58. Wilcken B, Bamforth F, Li Z *et al.* (2003) Geographical and ethnic variation of the 677C>T allele of 5,10 methylenetetrahydrofolate reductase (MTHFR): findings from over 7000 newborns from 16 areas world wide. *J Med Genet* **40**, 619–625.
59. Frosst P, Blom HJ, Milos R *et al.* (1995) A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. *Nat Genet* **10**, 111–113.
60. Holmes MV, Newcombe P, Hubacek JA *et al.* (2011) Effect modification by population dietary folate on the association between MTHFR genotype, homocysteine, and stroke risk: a meta-analysis of genetic studies and randomised trials. *Lancet* **378**, 584–594.
61. Xuan C, Bai XY, Gao G *et al.* (2011) Association between polymorphism of methylenetetrahydrofolate reductase (MTHFR) C677T and risk of myocardial infarction: a meta-analysis for 8,140 cases and 10,522 controls. *Arch Med Res* **42**, 677–685.
62. McNulty H, Pentieva K, Hoey L *et al.* (2008) Homocysteine, B-vitamins and CVD. *Proc Nutr Soc* **67**, 232–237.
63. Horigan G, McNulty H, Ward M *et al.* (2010) Riboflavin lowers blood pressure in cardiovascular disease patients homozygous for the 677C–>T polymorphism in MTHFR. *J Hypertens* **28**, 478–486.
64. Wilson CP, Ward M, McNulty H *et al.* (2012) Riboflavin offers a targeted strategy for managing hypertension in patients with the MTHFR 677TT genotype: a 4-y follow-up. *Am J Clin Nutr* **95**, 766–772.
65. Lander ES, Linton LM, Birren B *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921.
66. International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. *Nature* **431**, 931–945.