

Dietary Patterns, Genes, and Health: Challenges and Obstacles to be Overcome

Alexis C. Frazier-Wood

Published online: 25 November 2014
© Springer Science+Business Media New York 2014

Abstract Several dietary approaches have been proposed to prevent the onset of chronic diseases. As yet, no single approach has emerged as having the most consistent health benefits. This arises, in part, due to the fact that diet influences health in the context of individual factors with genetic components. Therefore, the effects of diet on health may be dependent on an individual's genetic background. At this time we lack robust evidence for the effects of interactions between genes and dietary patterns on health. To understand why, I will briefly review the most methodologically strong attempts to identify gene–diet interactions, which will illuminate how the challenges facing all of genetic research apply to the search for gene–diet interactions. Then I will discuss some ways in which these challenges are being addressed that offer hope for the future in which the best diet for an individual is identified based on their genetic variation.

Keywords Genome-wide interaction study · GWIS methodology · Gene–diet interactions · Nutrigenomics · Obesity

Introduction

Human health arises from complex interactions between genetic predisposition and the environment in which genes manifest. From pre-conception [1] until death, human health is shaped by nutrition, leading dietary intake to be considered a critical environmental factor that may interact with genes in

determining human health [2–5]. Dietary patterns are one way of conceptualizing nutritional intake. The “dietary pattern” approach to public health is reflected in federal dietary guidelines [6, 7] and those of foundations such as the American Heart Association (AHA) [2], who focus their recommendations for intake on both the quality and variety of overall diet rather than solely on the contribution of individual nutrients [8]. These recommendations have led to an awareness of dietary patterns such as the Dietary Approaches to Stop Hypertension (DASH) and Mediterranean Diet patterns, which may convey more protection on health over recommendations which focus on a single macronutrient such carbohydrate or fat. It has long been recognized that individual differences in genetic variation will influence the association of these dietary recommendations with health, and this has yet to be reflected in dietary guidelines. Identifying the interplay between genes and dietary patterns holds promise for a new era of personalized medicine, whereby the recommended diet for best health is tailored towards how an individual's metabolism is genetically predisposed to respond to dietary intake.

Moving away from the current one-size-fits-all approach holds promise from increasing the efficacy of dietary approaches in the prevention of chronic disease. However, compared to genetic main effects, few attempts have been made to identify interactions between genes and dietary patterns on health. Those that have, have remained stuck at the macronutrient level — for example, looking at interactions between carbohydrate intake and genes on adiposity—with very little attention paid to the overall diet. Even research on specific dietary nutrients thus far has been labeled “fragmented, incomplete, and in many cases, controversial” [9]. Identifying how overall diet interacts with genetic variation to influence human health is an important direction for public health, and the lack of definitive findings needs to be placed in the context of unique challenges that confront all human genetic research. Only by identifying these challenges can researchers start to

This article is part of the Topical Collection on *Dietary Patterns and Behavior*

A. C. Frazier-Wood (✉)
USDA/ARS Children's Nutrition Research Center, Baylor College of
Medicine, 1100 Bates Street, Houston, TX 77030, USA
e-mail: lekki.wood@gmail.com

address and overcome them. In addition, as clinicians are aware of the difficulties faced in nutrigenomic research, they may be able to respond to those patients who ask why a personalized approach to diet is still lacking.

Lessons Learned from Decades of Genetic Research

The history of genetics has been a long and chequered one. While there are notable successes in identifying genes for health, such as the risk alleles housed within breast cancer type 1 and 2 susceptibility protein (*BRCA1* and *BRCA2*) genes for breast cancer [10] and the apolipoprotein E (*APOE*) gene for Alzheimer's disease [11], it has been repeatedly stated that the quest to identify the genetic variants that account for the estimated heritability of complex traits related to human health has fallen short of expectations [12–15]. Although the advent of sequencing analysis has added to the geneticist's toolbox, genetic association studies designed to identify specific variants associated with disease have typically been conducted by one of two methodologies: candidate gene or genome-wide association studies (GWAS; Table 1).

In the early years of genetic research, a candidate gene approach was taken to identify putative loci associated with health. In this approach, a select number of variants within a pathway implicated in the development of a trait of interest would be analyzed for association with the trait. The lack of replication among candidate gene studies across a range of phenotypes led to the conclusion that candidate approaches on their own were unsuitable for identifying the complement of genetic variants associated with human disease. It was acknowledged that the pathways between candidate genes and disease-related traits were more numerous, and more complex, than previously thought, presenting challenges to generating a priori hypotheses about the putative gene regions.

The advent of DNA chip technology held promise for resolving this problem. Relative to candidate gene methods, chips allowed for quick and cheap genotyping of 500,000–3,000,000 common single-nucleotide polymorphisms (SNPs)

spaced somewhat evenly across the whole genome. This technological advance allowed researchers to survey loci across the genome for associations with health outcomes in a hypothesis-free manner, holding the promise of identifying new pathways between genes and disease. The initiation and success of the International HapMap Project further bolstered the science community's belief that GWAS would enable the identification of the SNPs accounting for the genetic variation underlying human health [16]. The International HapMap Project outlined the correlational structure between SNPs across the genome, which enabled researchers to impute data from the SNPs genotyped on a single chip to provide information on approximately 80 % of the human genome [16]. With the potential for hypothesis-free survey of almost all of the genome, there was enormous optimism at the start of the "GWAS era" that science would fairly rapidly start to account for the majority of the genetic variance associated with traits of interest.

To date, science has failed to do so. Even for some of the most commonly studied traits, such as body mass index (BMI), more than 95 % of the genetic variance remains unaccounted for [15]. In addition, the majority of initial genome-wide significant findings failed to replicate. In recognition of this, and of the limitation that many putative loci may lie just under the threshold for genome-wide significance, many have called for a two-stage approach to gene-hunting: a discovery phase in which a genome-wide scan is conducted, and then a replication phase in which all hits from the first phase reaching a given threshold (often $P < 1.0 \times 10^{-5}$; but can be between $P < 10^{-4}$ – 10^{-6}) are examined for associations in one or more independent samples [17]. It is hoped that this approach will prevent the false negatives which the stringent Bonferroni correction typically applied to GWAS studies has given rise to, but equally, prevent the "winner's curse" of initial associations that never replicate [18]. The failure of GWAS to account for the heritable variance has been termed the "missing heritability" [14]. The reasons for missing heritability are numerous and highly debated, but include the use of genetically complex phenotypes [19], the notion that rare

Table 1 A comparison of candidate gene and genome-wide approaches to genetic association studies

Candidate Gene	vs.	GWAS
A very small fraction of the genome is typically analyzed		Covers up to 80 % of the genome
Required a strong a priori hypothesis		Is a hypothesis-free or hypothesis-generating approach
A more statistically powerful approach		Compared to candidate genes, the needed correction for multiple testing leads to low power
Suitable for identifying variants of large and small effect		Typically more suitable for identifying variants of larger effect ^a
Can include rarer variants and nonSNP variants such as variable number of tandem repeats (VNTRs)		Traditionally includes common variation (SNP occurring in <1 % of the population, although some GWAS chips contain rarer variants and/or are customizable)

^a This is dependent on sample size

variants may play a more important role than previously realized and smaller effect sizes than initially anticipated [14, 15]. These well-recognized challenges pose grave problems for identifying how genetic background moderates the association between dietary patterns and health.

Placing Gene–Diet Interaction Research in the Light of These Lessons

In identifying loci that interact with a given dietary pattern, researchers are confronted with a difficult choice: does one take a candidate locus approach, or survey the whole genome in a hypothesis-free manner with a GWAS?

Candidate gene approaches have been the mainstay of gene–environment interaction studies, which generally choose SNPs and diets associated with a disease—or trait—in previous research in order to examine whether SNPs and dietary patterns interact. To my knowledge, although many studies have examined the effects of interactions between genetic variants and individual nutrients on health, only two single-cohort studies have examined gene–diet interactions with dietary patterns. The first study reported that weight loss on a Mediterranean style diet was moderated by variation at rs1801260 in the Circadian Locomotor Output Cycles Kaput (*CLOCK*) gene [20]. The second study reported that weight loss on the Mediterranean diet over a three-year period was significantly lower in A allele carriers of rs9939609 in the fat mass and obesity-associated (*FTO*) gene compared to those homozygous for the T allele [21]. Promising though these results are, they are notable for a lack of replication and validation.

Some of the strongest gene–environment research has come out of consortia such as the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium [22]. Consortia such as CHARGE employ data meta-analyzed across a number of cohorts, which may therefore be considered more robust and more powerful than single-cohort studies. A very focused candidate gene study used data from 15 cohorts to examine whether the associations of polymorphisms in the lipoprotein receptor-related protein 1 (*LRP1*) gene with BMI, waist circumference, and hip circumference were modified by fatty acid intake. The study reported that saturated fatty acid intake was associated with all three anthropometric traits, and interacted with rs2306692 in *LRP1* such that the association between saturated fatty acids and each anthropometric data was stronger for each of the anthropometric traits for each copy of the T allele [23•]. This interaction was present in whites but not in African-Americans [23•].

More often, however, the candidate gene approach to identifying gene–diet interactions has not yielded significant results. In data meta-analyzed across 14 CHARGE cohorts of

European ancestry, total zinc intake was associated with fasting glucose ($p=0.0003$) but did not significantly interact with glucose-raising alleles after a correction for multiple testing [24]. Similarly, a meta-analysis of the same cohorts reported that whole-grain food intake was strongly associated with fasting glucose and insulin ($P=0.0003$ - <0.00001), but no significant interactions were found with SNPs known to influence fasting glucose and insulin levels [25]. In possibly the only study to examine interactions between genes and a dietary pattern rather than a single macro- or micronutrient, an analysis across 15 cohorts from CHARGE investigated whether a “healthy diet” altered the associations between glucose- and insulin-associated SNPs and fasting levels. “Healthy diet” was a single score defined a priori based on self-reported intake of nine food groups. Five food groups were designated “favorable,” and a high healthy diet score indicted higher intake. These favorable foods were whole grains, fish, fruits, vegetables, and nuts/seeds. Four groups were defined as “unfavorable,” and high intake contributed to a lower overall healthy diet score. Red and processed meats, sweets, sugared beverages, and fried potatoes all contributed negatively to the healthy diet score [26••]. As expected based on previous research, the healthy diet score was inversely associated with fasting glucose and insulin; for each unit increase, there was a 0.004 mmol/L decrease in glucose and a 0.0058 mmol/L decrease in insulin. However, the score did not interact with known glucose- and insulin-related SNPs on the glycaemia measures [26••].

In many studies, SNP–diet interactions reached suggestive levels of significance, which did not survive a multiple testing correction, and thus it is not clear whether power is an issue [25]. Nonetheless, it is likely that the challenges faced by main-effect association studies are also faced by gene–diet interaction studies. As over 95 % of the mechanistic pathways from genes to chronic diseases such as obesity, cardiovascular disease, and cancer remain opaque, it is challenging to select which putative loci are involved in these pathways and thus should be analyzed in the context of gene–diet interactions [15, 27, 28]. In addition, selecting variants for interaction analysis is challenging in light of data simulations showing that the most significant associations from GWAS are not likely to be those variants which interact with the environment to influence phenotypes of interest [29–31]. Rather, since GWAS detects variants of relatively large effect, and the presence of an interaction with diet is likely to weaken the detected main effect, it is likely that the “sub-threshold” variants are those which hold promise for gene–diet interactions [29–31]. Focusing on genetic loci which are not the top GWAS hits may be a more fruitful approach. As yet, however, there is no empirical or theoretical threshold for selecting these variants.

Genome-wide interaction study (GWIS) approaches also pose unique difficulties: with a given sample and effect size, a

study has less power to detect gene–environment interactions than genetic main effects [32]. Full GWIS approaches therefore require the accrual of large sample sizes. While many successful consortia have been established, these bring the additional challenge of harmonizing diet data across cohorts. Dietary data are typically measured in one of a number of ways, including through food frequency questionnaires (FFQs), dietary diary analysis, and interview-led diet history interviews. The concordance between these measures has yet to be established. In addition, if a dietary pattern is defined in part by a macronutrient variable, it may be that the constituent components of that macronutrient differ across cohorts defined by, for example, ethnicity or geographic location. This heterogeneity makes it difficult to detect interactions in a combined sample.

The GWIS approach still holds promise for identifying gene–diet interactions. In a genome-wide analysis of 9,287 colorectal cancer cases and 9,117 unaffected control controls across ten independent studies, a significant interaction was found between a variant near the trans-acting T-cell-specific transcription factor (*GATA3*) gene and processed meat intake on colorectal cancer status [33•]. The association between processed meat intake and colorectal cancer was only present among those who were T allele carriers [33•]. However, while compelling, this result has yet to replicate. In addition, this study focused on a single food group rather than an overall pattern of intake, making it difficult to derive clinical advice from the results.

Developing Solutions to Methodological Challenges in Gene–Diet Interaction Analysis

Numerous challenges remain in the quest to determine the dietary pattern that will convey the most protection against chronic disease, given an individual’s genetic background. Even with a candidate gene approach which is powered to detect smaller effect sizes than a GWIS approach, the very small effect sizes found for suggestive interactions (those that are significant but do not survive a correction from multiple testing) suggest that power still remains an issue [24, 25]. One solution to dealing with small effect sizes for interactions with individual SNPs has been to use a genetic risk score (GRS). A genetic risk score is a sum of alleles taken across several genes known to convey deleterious effects on the same phenotype. By summing the effects of several SNPs, GRS typically show larger effect sizes than a lone SNP. Recently, studies adopting this approach have met with some success, as witnessed in the following two illustrative examples.

A GRS constructed of 63 obesity-associated alleles was found to interact with a diet high in total—and especially saturated—fat intake on BMI in two independent populations [34]. When saturated fat intake was low, every additional risk allele in the GRS was associated with an increase in BMI of 0.11–0.20; when

saturated fat intake was high, each risk allele conveyed an additional increase of 0.19–0.41 BMI points. In a similar study, a higher GRS score (more risk alleles across 32 BMI-associated loci) increased the association between BMI and sugar-sweetened beverage intake; per increment of 10 risk alleles, the increase in BMI points was 1.00 for an intake of less than one serving per month, 1.12 for one to four servings per month, 1.38 for two to six servings per week, and 1.78 for one or more servings per day ($P < 0.001$ for interaction) [35]. GRSs can also include SNPs identified from mechanistic candidate gene studies. This may convey an advantage since, as previously discussed, the top GWAS hits may not contain SNPs that associate with the phenotype of interest through gene–environment interactions [29–31]. However, a limitation of the GRS approach is that it fails to consider potential interactions between the alleles comprising the GRS (gene–gene interactions), and further, offers little mechanistic insight into how genes modify the effect of dietary pattern on health, since GRSs often subsume many mechanistic pathways.

Further Challenges in Identifying the Best Dietary Pattern for Good Health for Each Genotypic Background

There still remain several issues that researchers need to resolve if we are to identify the dietary pattern that is most protective against chronic diseases for each individual, given their genetic background. Only a few articles could be identified in which the interaction between *dietary patterns* and genetic variation on health were analyzed [26••]. This may reflect difficulties with how “dietary pattern” should be defined and what it should be compared to. Dietary intake, in general, reflects a continuum of adherence to a given dietary pattern rather than a yes/no approach. Defining how closely someone adheres to, for example, the DASH diet requires the quantification of guidelines, which may be tricky to achieve across diverse cohorts. In addition, the implication of the anticipated small effect sizes goes beyond the need for large sample sizes, and reflects the larger question of clinical utility: what is the role of expensive genetic tests if they may only inform on a difference of 0.004 mmol/L in fasting glucose [26••], when dietary changes alone may reduce the incidence of type 2 diabetes by up to 31 % [36]?

Conclusions

While the clinical utility of understanding how genes moderate the association of dietary pattern with health has not yet been fulfilled, the potential benefits of using genetic background to design a personalized dietary approach for the prevention of chronic disease remain. The popularity of home-based nutrigenomic tests, while scientifically

questionable, reflects public desire for personalized nutrition advice [37]. There are many methodological challenges to overcome, shared with all of human genetic research, before we can identify the putative variants that interact with dietary patterns. These challenges center on the lack of power for GWIS approaches, but the theoretical and empirical difficulties in identifying loci a priori for a candidate gene approach. Dietary nutrients do not act in isolation, and as guidelines for preventing chronic disease focus on overall holistic diet, understanding which dietary pattern an individual should adhere to may present the most promise for improving public health. To address this question, we are going to need to combine data from studies employing a host of statistical approaches. Studies employing GRSs have perhaps met with the most success in identifying gene–diet interactions, but to gain mechanistic insights, these will have to complement candidate gene approaches. In addition, as we do not fully understand the pathways between genes and chronic disease, and have identified SNPs which only account for a small proportion of the underlying genetic variance, GWIS approaches will be vital for identifying new related loci. Finally, the need for replication must be addressed. Until then, dietary advice for optimum health will remain at the population—rather than individual—level, and many benefits of dietary change on health may be lost.

Acknowledgments Dr. Alexis C. Wood is funded by the American Heart Association (grant no: 14BGIA18740011) and the USDA/ARS Children’s Nutrition Research Center at Baylor College of Medicine. This work is a publication of the USDA/ARS Children’s Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine, Houston, Texas, and funded in part with federal funds from the USDA/ARS under Cooperative Agreement No. 58-6250-0-008. The contents of this publication do not necessarily reflect the views or policies of the USDA, nor does mention of trade names, commercial products, or organizations imply endorsement from the U.S. government.

Compliance with Ethics Guidelines

Conflict of Interest Alexis C. Frazier-Wood declares that she has no conflict of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:

- Of importance
- Of major importance

1. Cooney CA. Germ cells carry the epigenetic benefits of grandmother’s diet. *Proc Natl Acad Sci U S A*. 2006;103:17071–2.

2. Lichtenstein AH et al. Summary of American Heart Association Diet and Lifestyle Recommendations revision 2006. *Arterioscler Thromb Vasc Biol*. 2006;26:2186–91.
3. Lichtenstein AH et al. Diet and lifestyle recommendations revision 2006: a scientific statement from the American Heart Association Nutrition Committee. *Circulation*. 2006;114:82–96.
4. Kushi LH et al. American Cancer Society Guidelines on Nutrition and Physical Activity for Cancer Prevention: reducing the risk of cancer with healthy food choices and physical activity. *CA Cancer J Clin*. 2006;56:254–81.
5. McGuire S. U.S. Department of Agriculture and U.S. Department of Health and Human Services, Dietary Guidelines for Americans, 2010. 7th Edition, Washington, DC: U.S. Government Printing Office, January 2011. *Adv Nutr*. 2011;2:293–4.
6. Peterkin BB. Dietary Guidelines for Americans, 1990 edition. *J Am Diet Assoc*. 1990;90:1725–7.
7. Flock MR, Kris-Etherton PM. Dietary Guidelines for Americans 2010: implications for cardiovascular disease. *Curr Atheroscler Rep*. 2011;13:499–507.
8. Marshall S, Burrows T, Collins CE. Systematic review of diet quality indices and their associations with health-related outcomes in children and adolescents. *J Hum Nutr Diet*. 2014. doi:10.1111/jhn.12208.
9. Ordovas JM. Introduction. *Nutr Metab Cardiovasc Dis*. 2007;17:71–3.
10. Chen S, Parmigiani G. Meta-analysis of BRCA1 and BRCA2 penetrance. *J Clin Oncol*. 2007;25:1329–33.
11. Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet*. 2007;39:17–23.
12. Maher B. Personal genomes: the case of the missing heritability. *Nature*. 2008;456:18–21.
13. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. *Am J Hum Genet*. 2012;90:7–24.
14. Manolio TA et al. Finding the missing heritability of complex diseases. *Nature*. 2009;461:747–53.
15. Hebebrand J, Volckmar A-L, Knoll N, Hinney A. Chipping away the “missing heritability”: GIANT steps forward in the molecular elucidation of obesity - but still lots to go. *Obes Facts*. 2010;3:294–303.
16. Consortium, T. I. H. The International HapMap project. *Nature*. 2005;63 Suppl 1:29–34.
17. Asking for more. *Nat Genet*. 2012;44:733.
18. Kraft P. Curses—winner’s and otherwise—in genetic epidemiology. *Epidemiology*. 2008;19:649–51. discussion 657–8.
19. Gottesman II, Shields J. Genetic theorizing and schizophrenia. *Br J Psychiatry*. 1973;122:15–30.
20. Garaulet M et al. CLOCK gene is implicated in weight reduction in obese patients participating in a dietary programme based on the Mediterranean diet. *Int J Obes (Lond)*. 2010;34(516–23).
21. Razquin C et al. A 3-year intervention with a Mediterranean diet modified the association between the rs9939609 gene variant in FTO and body weight changes. *Int J Obes (Lond)*. 2010;34:266–72.
22. Psaty BM et al. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ Cardiovasc Genet*. 2009;2:73–80.
23. Smith CE, et al. Lipoprotein receptor-related protein 1 variants and dietary fatty acids: meta-analysis of European origin and African American studies. *Int J Obes (Lond)*. 2013;37(9):1211–20. *A positive gene–diet interaction in from a 15 cohort meta-analysis*.
24. Kanoni S et al. Total zinc intake may modify the glucose-raising effect of a zinc transporter (SLC30A8) variant: a 14-cohort meta-analysis. *Diabetes*. 2011;60:2407–16.
25. Nettleton JA, McKeown NM, Kanoni S, Lemaitre RN, Hivert MF, Ngwa J et al. Interactions of dietary whole-grain intake with fasting glucose- and insulin-related genetic loci in individuals of European descent: a meta-analysis of 14 cohort studies. *Diabetes Care*. 2010;33:2684–91.

26. •• Nettleton JA et al. Meta-analysis investigating associations between healthy diet and fasting glucose and insulin levels and modification by loci associated with glucose homeostasis in data from 15 cohorts. *Am J Epidemiol.* 2013;177:103–15. *The largest gene–diet interaction study which includes analysis of a dietary pattern rather than single nutrient.*
27. Teslovich TM et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature.* 2010;466:707–13.
28. Speliotes EK et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet.* 2010;42:937–48.
29. Kraft P, Yen Y-C, Stram DO, Morrison J, Gauderman WJ. Exploiting gene-environment interaction to detect genetic associations. *Hum Hered.* 2007;63:111–9.
30. Murcray CE, Lewinger JP, Gauderman WJ. Gene-environment interaction in genome-wide association studies. *Am J Epidemiol.* 2009;169:219–26.
31. Gauderman WJ, Zhang P, Morrison JL, Lewinger JP. Finding novel genes by testing $G \times E$ interactions in a genome-wide association study. *Genet Epidemiol.* 2013;37:603–13.
32. Khoury MJ, Wacholder S. Invited commentary: from genome-wide association studies to gene-environment-wide interaction studies—challenges and opportunities. *Am J Epidemiol.* 2009;169:227–30. discussion 234–5.
33. • Figueiredo JC et al. Genome-wide diet-gene interaction analyses for risk of colorectal cancer. *PLoS Genet.* 2014;10:e1004228. *The largest GWIS with reporting a significant gene–diet interaction.*
34. Casas-Agustench P et al. Saturated fat intake modulates the association between an obesity genetic risk score and body mass index in two us populations. *J Acad Nutr Diet.* 2014. doi:[10.1016/j.jand.2014.03.014](https://doi.org/10.1016/j.jand.2014.03.014).
35. Qi Q et al. Sugar-sweetened beverages and genetic risk of obesity. *N Engl J Med.* 2012;367:1387–96.
36. Pan X-R et al. Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance: the Da Qing IGT and diabetes study. *Diabetes Care.* 1997;20:537–44.
37. Robitaille J. Nutrigenomics and personalized diet: what are the anticipated impacts for research on chronic diseases and public health? *Curr Pharmacogenomics Person Med.* 2009;7: 106–14.