Metabolomic and microbiome profiling are promising tools to identify biomarkers of food intake and health status. The individual’s genetic makeup plays a significant role on health, metabolism, gut microbes and diet and twin studies provide unique opportunities to untangle gene–environment effects on complex phenotypes. This brief review discusses the value of twin studies in nutrition research with a particular focus on metabolomics and the gut microbiome. Although, the twin model is a powerful tool to segregate the genetic component, to date, very few studies combine the twin design and metabolomics/microbiome in nutritional sciences. Moreover, since the individual’s diet has a strong influence on the microbiome composition and the gut microbiome is modifiable (60 % of microbiome diversity is due to the environment), future studies should target the microbiome via dietary interventions.

‘Twins history affords means of distinguishing between the effects of tendencies received at birth, and those that were imposed by the circumstances of their after lives; in other words between the effects of nature and of nurture’

Galton 1875

More than a century after Galton’s observation, twin studies remain an invaluable tool to determine the role of genes and environment on human development and behaviour. Classical twin studies compare the trait concordance between monozygotic (MZ) and dizygotic (DZ) twin pairs to estimate to what extent the observed phenotype difference is genetically determined(1). The assumption is that MZ twins share 100 % of their genes as they originate from a single zygote and are therefore genetic clones, whereas DZ twins originate from two separate eggs and thus share 50 % of their genes. If MZ twins are more alike than DZ twins with respect to the trait, then the trait is said to be influenced by genes, otherwise the trait is said to be determined by lifestyle and environment.

More formally, the twin phenotypic variance (i.e. individual differences of a trait) can be divided into three sources of variation: additive genetic variance (\(A\)), shared/common environmental variance (\(C\)) and non-shared/unique environmental variance (\(E\))(2). Additive genetic influences represent the sum of the effects of the individual alleles at all loci that influence the trait. The common environmental component estimates the contribution of family environment, which is assumed to be equal in both MZ and DZ twin pairs\(^{(3)}\), whereas the unique environmental component does not contribute to twin similarity, rather it estimates the effects that apply only to each individual including measurement error. Heritability (\(h^2\)) is defined as the proportion of the phenotypic variation attributable to genetic factors\(^{(4)}\), and is given by the equation, \(h^2 = \frac{A}{A + C + E}\).
number of assumptions underlie the twin study. These include: (i) MZ and DZ twin pairs share 100% of their common environment; (ii) twins are representative of the general population; (iii) gene–environment interactions are minimal for the trait under study; (iv) there is random assortment.

Twin studies have shown that the genetic make-up of individuals plays a significant role in a multitude of dietary phenotypes, including energy, macronutrient intakes, dietary patterns and specific food groups as extensively reviewed\(^\text{[6,7]}\). Heritability estimates of different energy and macronutrient intakes range between 8 and 70%, while both ‘healthy’ and ‘unhealthy’ diet patterns are relatively stable in adults and have heritability ranging from 33 to 50\%\(^\text{[7]}\). Specific foods intakes, including garlic (46%), fruit and vegetables (49%), and coffee (41%) are also highlyheritable in adults\(^\text{[8]}\). A recent study from our group identified four major food-liking patterns (fruit and vegetables, distinctive tastes, sweet and high carbohydrate, and meat) accounting for 26% of the total variance with heritability estimates ranging from 36 to 58% indicating genetic factors influence food liking–disliking\(^\text{[9]}\).

Besides estimating heritability of dietary assessment, using MZ twins discordant for dietary factor/nutrition status provides a naturally unique case-controlled experiment of assessing the links between diet and human biology. Due to their shared upbringing, including shared fetal exposure, matched genes and sex, MZ twins allow one to isolate the non-genetic contribution.

Diet is partially responsible for the dramatic rise in obesity\(^\text{[10]}\) and obesity-related diseases, such as type 2 diabetes\(^\text{[11]}\); however, studies do not consistently support associations between dietary intake and disease endpoints. Dietary intakes in epidemiological settings are generally measured via self-reported dietary assessment methods that are subject to recall bias and measurement error\(^\text{[12,13]}\). Nutritional epidemiological studies could be improved by the use of food intake biomarkers to better capture exposure.

This brief review discusses the value of twin studies in nutrition research with a particular focus on metabolomics and the gut microbiome.

Metabolomics

Metabolomic profiling has the potential to identify biomarkers of food intake\(^\text{[14]}\). Metabolomics is a new high-throughput technology that measures endogenous metabolites in cells, tissues and other biosamples such as blood, urine and saliva\(^\text{[15]}\). These metabolites provide a direct signature of the biochemical activities of the individual at a particular time, reflecting the metabolic effects of nutritional intake, physical activities and environmental exposures as well as the biological pathways associated with diet-related diseases including obesity, CVD and type 2 diabetes\(^\text{[16–19]}\). Metabolites have been proven to be of use in nutritional research and novel dietary biomarkers within the metabolome have been identified\(^\text{[20–23]}\). However, as the individual genetic makeup influences the levels of metabolites, explaining up to 81% of the variation in their blood levels\(^\text{[24–27]}\), the twin model can be employed to estimate the nutritional impact on metabolites, segregating the genetic component.

To date there are relatively few studies that combine a twin study design and metabolomics in nutritional sciences. Studies from TwinsUK have successfully identified potential biomarkers of alcohol intake\(^\text{[28]}\) including lysophosphatidylcholines, diethyl ether lipids, diacylphosphatidylcholines and sphingolipids, as well as biomarkers of self-reported dietary patterns\(^\text{[22]}\), food preference patterns\(^\text{[9]}\) and self-reported food intakes\(^\text{[29,30]}\).

Taking advantage of the twin nature of the TwinsUK data, we first looked for association in the larger twin population excluding MZ twins discordant for a nutritional intake. Then, for each significant metabolite–dietary variable we run the same analysis on the discordant MZ twin pairs to replicate results in this set.

Using a targeted metabolomics approach (Bioocrates platform; 163 metabolites), we identified forty-two dietary pattern–metabolite significant associations in 1003 female twins and successfully replicated eleven associations in a subcohort of MZ twin discordant for dietary intake (between twenty-eight and forty discordant twins independent from the first analysis)\(^\text{[22]}\). The strongest associations were observed for fruit and vegetables intake with the glycerophospholipid phosphatidylcholine diaclyl C\(_{38}:6\) and with the sphingolipid sphingomyeline C\(_{26}:1\). Significant associations were also found with coffee and garlic intake and for hypoenergetic dieting.

In a more recent study from our group, we looked for association between 601 blood metabolites (456 measured with the untargeted Metabolon platform and 145 measured with Biocrates) and seventy-one reported food intakes from FFQ in 3500 female twins to look for novel biomarkers of nutrition intake. We identified 180 associations with thirty-nine food group after meta-analysing the discovery and replication cohort (MZ discordant), overall consisting of 106 different metabolites, mainly lipids, including seventy-three novel associations\(^\text{[30]}\). In particular, we identified the amino-acid trans-4-hydroxyproline, a component of collagen, as a novel biomarker for red meat intake; ergothioneine as a potential marker of mushroom intake and two metabolites derived from the gut bacteria transformation of phenolic compounds as biomarkers of fruit intake. All the findings are compiled into the open access DietMetab database (www.twinsuk.ac.uk/dietmetab-data).

Using a similar design and replicating in two independent European populations, we have also identified four novel biomarkers of milk intake; trimethyl-N-aminovalerate, uridine, hydroxylysingomyelin C\(_{14}:1\) and diacylphosphatidylcholine C\(_{28}:1\)\(^\text{[28]}\).

Microbiome

The study of the gut microbiome is an exciting area of medicine because of its immediate potential for therapeutic interventions\(^\text{[31]}\). The term microbiome describes the DNA material of microbial communities within an
animal and recent advances in technology are revealing its complexity (32). Human subjects have about 100 trillion gut microbes that outnumber their cells ten to one (33). Human subjects and their microbiome have co-evolved and live intimately in a symbiotic relationship (34) with the microbes producing a wide range of enzymes, chemicals, hormones and vitamins that can potentially interact with the host (35). The composition of the microbiome varies by anatomical site, with the primary determinant of community composition being anatomical location (36). Diversity measures the number of microbes present at a particular site (37). The α diversity refers to the number of different species at a particular site (36), while β diversity considers how many taxa are shared between populations, acting as a similarity score (37). A low diversity of gut microbes (sometimes called dysbiosis) has been implicated in many human diseases including CVD, obesity, inflammatory bowel disease, colitis and type 2 diabetes (38–40). Only recently, thanks to genetic sequencing have we been able to study microbes properly and realise that the vast majority are not harmful and many are beneficial (31). Research has shown that the largest influence on the gut microbiome comes from diet (41) and hence we could potentially alter many diseases with food and monitor the effects via microbiomes. For instance, dietary fibre is fermented in the intestine by the colonic microbiota resulting in increased production of SCFA (42). SCFA are involved in the energy metabolism as they regulate the balance between fatty acid synthesis, fatty acid oxidation, and lipolysis in the body. A reduction in the concentrations of free fatty acids in plasma results in a decrease in body weight (43).

Studies so far on food and the microbiome have mainly focused on meat-based/plant-based diet and highlighted differences in the microbiota of plant- and animal-based diets (44–46). In particular, they have demonstrated that human gut responds rapidly to major dietary changes (47). A meat-based diet, rich in animal protein, fats, artificial additives, and lacking in fibre, has been shown to drive chronic conditions such as obesity, the metabolic syndrome and atherosclerosis by encouraging gut dysbiosis. Dietary interventional studies, however, have shown that the effects of meat-based diets on the gut microbiome could be rapidly reversed by adopting a plant-based, minimally processed diet (48–50). Long-term dietary habits, however, tend to determine the composition of the gut microbiota with studies showing that long-term dietary habits associated with the individual microbiome signature (47). Also, changes in diet are dependent on the individual gut microbiota composition suggesting that not everyone reacts the same to the same dietary change (51).

Moreover, the way in which individuals respond to dietary interventions may be complicated by host genetics and so the twin model could be a valuable tool to study the association of the gut microbiome with nutritional phenotypes. Novel twin data shows that although many microbes are driven by environment, a substantial proportion of microbiota with disease associations have a heritable component (52–54). The most heritable taxon, the Christensenella family was shown to be associated with leanness in human subjects and prevents weight gain under high-fat diet in mice (55).

Using discordant twins, we have also successfully identified a striking negative association between frailty due to ageing and gut microbiota diversity (56) as well as a significant alteration in the gut microbiota in proton pump inhibitors users (57). In general, therefore, a more diverse gut microbiome signals a healthier state and this is modified by age, diet and use of medication.

Conclusions

In conclusion, twin research is an invaluable tool to untangle gene–environment effects on complex phenotypes including metabolomics, the gut microbiome and complex phenotypes such as obesity and diet related diseases. Indeed, twins have provided valuable evidence that diet is influenced by genetics suggesting that future dietary counselling should target this domain.

Metabolomics and microbiomics have the potential to identify biomarkers of food intake and health status. Combining these omics with twin data allows the separation of features influenced by genetics from those influenced by the environment. This supports the efficient longitudinal monitoring of individuals at high genetic risk as they progress from health to disease making this model an invaluable means of testing personalised medicine strategies.

Finally, because the gut microbiome is modifiable (60% of microbiome diversity is due to the environment) and since an individual’s diet has a strong influence on the microbiome composition, future studies should target the microbiome via dietary interventions.

Financial Support

TwinsUK was funded by the Wellcome Trust; European Community’s Seventh Framework Programme (FP7/2007-2013). The study also receives support from the National Institute for Health Research (NIHR) Clinical Research Facility at Guy’s & St Thomas’ NHS Foundation Trust and NIHR Biomedical Research Centre based at Guy’s and St. Thomas’ NHS Foundation Trust and King’s College London. C. M. is funded by the MRC AimHY (MR/M016560/1) grant.

Conflicts of Interest

None.

Authorship

C. M. was solely responsible for all aspects of preparation of the present paper.

References


