

Will Formal Genetics Become Dispensable?

Françoise Clerget-Darpoux^{a, b} Robert C. Elston^c^aINSERM U-781, Département de Génétique, Hôpital Necker-Enfants Malades, and ^bUniversité Paris Descartes, Faculté de Médecine, Paris, France; ^cCase Western Reserve University, Cleveland, Ohio, USA

Over the last half century or so, in the form of molecular biology, genetics has usurped what had until then been classified as, for example, biochemistry, physiology, embryology and immunology. Once it was recognized that DNA plays a crucial role in genetically caused diseases, anything that causes a disease by acting on DNA, even environmental triggers such as ionizing radiation, has been claimed by geneticists as a ‘genetic’ cause of disease. Prior to that, since the rediscovery of Mendel’s work up to the discovery of the structure of DNA, genetics largely involved what we shall here call ‘formal’ genetics, i.e. what Haldane [1] called ‘beanbag genetics’ or the mathematical formulation of genetic transmission through the study of families and, by extension, populations. There is little doubt that the study of families and evolutionary genetics still has a major part to play in the study of genetics. But for how long will this continue? Haldane [1], in his defense of ‘beanbag genetics’, wrote of the dichotomy between physiological and beanbag genetics. Whereas Haldane was mostly concerned with how mathematics could contribute to evolutionary biology, today we might consider more how mathematics and statistics contribute to molecular genetics to understand the phenotype of an individual. Here, we ask the question: what would be necessary for formal genetics to become dispensable?

Some perhaps believed that individualized ‘predictive medicine’ and individualized ‘precision medicine’ [2]

would become a reality and that it would be possible to dispense with any family information. This would assume that, given a person’s complete DNA sequence, we would know the person’s predisposition to any disease and/or the appropriate treatment. Whereas sequencing tumor DNA has had successes in providing precision medicine at the individual level in treating cancer [3], even 10 years after the completion of the Human Genome Project the clinical utility of genomic medicine to advance disease prevention and patient outcomes is still suboptimal [4]. This raises two questions. First, can we imagine a time will come when we shall be able to measure everything that a human inherits from the two parents at conception, and this would comprise all the genetic information that has been transmitted? Second, will it be possible with that knowledge to predict a person’s phenotype? In other words, could we consider that, once given a complete specification of what is transmitted, it will be possible to know the probability of any phenotype, whether we look at one or more traits, and/or at one or multiple points in time?

Regarding the first question, even if we agree that we shall one day be able to accurately determine the whole genomic sequence of a newborn (and we are not there yet [5]!), shall we ever be able to measure accurately *everything* that is transmitted to the newly formed zygote without any familial information? Once upon a time we had

over 100,000 genes in the genome, then slowly that figure decreased to somewhere between 20,000 and 22,000. Then we realized we have many stretches of DNA that are transcribed but not translated. For more than 5 years now, it has been recognized that these are really also genes, though not protein-coding genes. Gerstein et al. [6] reviewed the history of the word 'gene' and proposed a tentative update to the definition of a gene: 'A gene is a union of genomic sequences encoding a coherent set of potentially overlapping functional products'. If we include these RNA genes, by now we must know of at least 100,000 genes – and there could well be millions. About 95% of the genome is transcribed, but only about 1–2% is translated.

We have long known that what is transmitted from a parent to offspring comprises chromosomes and mitochondria, but that is much more than just DNA. DNA comprises only a small fraction of what is transmitted from one generation to the next: most of what is transmitted is protein. We know that methylation/acetylation of the histones that are transmitted from cell to cell during development – 'epigenetics marks' – play an important role in coding-gene expression, and hence in disease initiation and progression. But these epigenetic marks, which are transmitted from cell to cell as the zygote develops into an adult, have been found to disappear, and so have been assumed to play no part in the transmission of information from one generation to the next.

In 1942, Waddington [7] stressed the importance of studying the mechanics of development from zygote to the adult phenotype, suggesting the name 'epigenetics' for such studies. Some now define epigenetics as the transmission of parental somatic epigenetic patterns through the germ line to offspring, so that there is a heritable component to epigenetic changes that originally came from a past environment. A case for this transgenerational epigenetic inheritance is made in an article by Morgan and Whitelaw [8]. It has been suggested that the maternal granddaughters of women who were exposed to diethylstilbestrol during pregnancy have an increased risk of ovarian cancer [9]. The daughters, exposed to famine in utero, of Dutch women during the second world war have experienced neonatal adiposity and poor health in later life [10]. A paternal grandfather's food supply is associated with the mortality of his grandsons, while a paternal grandmother's food supply is associated with her granddaughters' mortality [11, 12]. So, given this mounting evidence, could it be that the non-DNA part of the chromosomes is transmitting information from one generation to the next? Can we imagine that this information

will one day be available on the zygotes, without any knowledge of their parents or other past ancestors?

Is it inconceivable that the day will come when all that information will be measurable? It is much more than the 3.2 billion bits of information obtained by sequencing the genome, but is it *immeasurably* more? Of course, we do not yet even know exactly where that information is. Is any of it somehow transmitted in the DNA, but not just in the DNA sequence? Multiple mechanisms for methylation erasure occur in mice several days after zygote formation [13, 14], so some of this information is transmitted in the chromatin during meiosis. But even if we were sure where all that information is, can we imagine being able to measure and use it, together with the DNA sequence? The speed at which technology is advancing, indeed *accelerating*, makes it difficult to be sure that the day will never come that this will happen. Who would have conceived, just a generation ago, that sequencing a person's genome could become a reality? Clearly, wherever it is, this is information we should strive to obtain.

Now the above discussion on formal genetics raises our second question: once we are given a complete specification of the genotype and all other information that is transmitted from one generation to the next, will it be possible to deduce from this a reliable probability for any phenotypic outcome that could occur, at any particular stage of life? More generally, knowing a person's genomic sequence is key in situations for which a specific genotype is necessary to develop disease. For some diseases, all the factors necessary for their expression are even known: since 1963, for instance, systematic screening at birth has allowed us to detect those who may develop phenylketonuria, and to prevent it through a phenylalanine-restricted diet. From the formal genetic point of view, this is an ideal situation. In other situations, the expression of a disease and the variability of its expression often remain very difficult to model mathematically, and thus difficult to predict [15] even when we know that a specific genotype is necessary for the disease to develop.

In Haldane's [1] time, the interest of formal genetics was limited to those diseases for which risk prediction could be appraised through a simple beanbag law and it was simple to model the genotype/phenotype relation. Some believed that increased knowledge about the human genome would ultimately allow establishing such a relation for any disease. The discovery of human leukocyte antigens (HLAs) and the demonstration of associations between these antigens and many diseases fueled this hope and raised much enthusiasm. Introducing the first international symposium on HLA and diseases

in 1976, Dausset declared: 'This scientific meeting is the first entirely devoted to this new field of modern biology, which evolves extremely rapidly and which opens new insights into the etiology and genetics of many diseases' [16]. Yet, 37 years later, it is fair to say that the HLA component of most HLA-associated diseases is still poorly understood.

The fast development of molecular biology and computer technology led to the Human Genome Project, which started in 1989 with the ambition of identifying the genetic etiology of diseases, and thence developing appropriate treatments. In 2000, the enthusiasm of geneticists was at its peak. Collins, then director of the National Human Genome Research Institute, claimed that genetic diagnosis would be accomplished within 10 years; and President Clinton declared it would 'revolutionize the diagnosis, prevention and treatment of most, if not all, human diseases' [17]. Should these claims be realized, then formal genetics would indeed become dispensable. It is obvious that research on the genome has revolutionized basic knowledge in genetics and biology, bringing many new and surprising results. However, the etiology of the majority of diseases remains poorly understood and individual risks are often estimated under simplistic and unrealistic assumptions.

Were the promises of the Human Genome Project made with an overly optimistic schedule in mind, or could it be that the desired aim is simply unrealistic? It is an inescapable fact that the majority of human diseases are due to multiple factors (both genetic and environmental) that are neither necessary nor sufficient for the disease to develop. For such 'multifactorial' diseases, it has sometimes been assumed that the additive polygenic model proposed by Fisher [18] underlies the disease genetic susceptibility. This is the basis of the allele-counting method [19, 20] that has been proposed for using the results of genome-wide studies, using the total number of catalogued susceptibility alleles as an overall measure of risk, and thus allowing people to be classified according to their personal risk compared to a reference population. It is under this assumption that many studies – and also direct-to-consumer genetic testing companies – provide estimates of individual risks, though other more realistic ways have been developed [21]. Alas, as we also discuss below, this model is not just inefficient, but can be rejected outright in some cases, and so using it would jeopardize risk estimation. Just as a lod score computed under a wrong model leads to a wrong interpretation [22], so are risk estimates prone to misinterpretation if a wrong model is adopted to obtain them. Furthermore, an individual

may be told s/he has a high-risk genome when in fact having zero probability of becoming affected because of never being exposed to the necessary unfavorable environmental factors.

For the diseases that most afflict us, we presently do not know: (a) how many genetic risk factors are involved; (b) how important their individual effects are and, especially, (c) how these factors interact among themselves and/or with the environment. Thus, assuming – as the additive polygenic model does – that a multifactorial disease is triggered by the small and independent effects of many factors is an overly simplistic assumption. The study of the HLA component of HLA-associated diseases has clearly shown the need to consider the possibility of statistical interaction between genes. A good example is given by celiac disease, for which the 0501 allele of the DQA1 gene confers a risk only if the 0201 allele of the DQB1 gene is also present, either in *cis* or *trans* position, the two alleles encoding for the two components of a heterodimer denoted DQ2. In terms of modeling, this example is interesting for several reasons. The transmission of the risk heterodimer does not follow classical monogenic rules, since it can be inherited either from a single parent (alleles in *cis* position) or from both parents (alleles in *trans* position). Incidentally, it is interesting to note that the frequency of the first situation relative to the second varies geographically [23]. The heterodimer DQ2 is present in most celiac disease patients (about 90% of European patients), but is also quite frequent in the general unaffected population (about 25%). To describe the effect of this heterodimer, we need a four-allele model or, in terms of two single nucleotide polymorphisms (SNPs), a model that assumes the presence of essential statistical interaction [23, 24]. Such a model is poorly reflected by the action of any single SNP in terms of relative risks. Using only the information on the most significantly associated SNP will lead to attributing a very different risk to an individual who carries both the DQA 501 and DQB 0201 alleles, either in *cis* or *trans* position, whereas the risk is actually the same. Similarly, the effect of PTPN22 in rheumatoid arthritis is poorly described either by a single SNP or by several independently acting SNPs. Assuming a model with independent effects of variants in PTPN22 leads to misclassifying individuals in terms of risks [25].

One promise of the HapMap Project was to 'facilitate [the] development of diagnostic tools' [26] and, in this respect, many consider that the very large number of SNPs shown to be associated with multifactorial diseases is a major outcome of the project. We have already pointed out how huge the gap is between our observing differ-

ent genotypic distributions in case-control samples and our being able to interpret what it means [27, 28]. Once the involvement of a gene in a disease has been detected through an association with a SNP, estimating its effect in terms of genotypic relative risks is currently very imprecise if all we have is population information. The genotypic relative risks based on the associated SNP may poorly reflect the differential risks of the genotypes; and more importantly, as shown in the two previous examples, it may lead to misclassifying individuals in terms of their risks. The more complex the true genetic model, the more information we need. It is important to note that neither association nor linkage is redundant: rather, they can provide complementary information for modeling a genetic effect [29]. In particular, examining the identity by descent (IBD) sharing of two affected sibs conditional on the genotype of one of them may offer good power to discriminate between different models that fit the overall population observations equally well. Even when the global IBD distribution is close to the one expected under no linkage, large differences may be observed among these conditional IBD distributions: this information proved very useful for improving genotypic relative risk estimation for HLA and the PTPN22 gene in the case of rheumatoid arthritis [25, 30], and for the IL2RA gene in multiple sclerosis [31]. Such information can also be very useful to confirm or reject the role of a rare variant [32, 33]. As described above, statistical interaction exists between the HLA DQA and DQB genes for celiac disease. The DQ2 heterodimer, present at the cell surface, corresponds to a complementation of a DQ α chain and a DQ β chain and is functional in peptide binding. More generally, for human traits, interactive effects of genetic variation at more than one genetic locus are more likely to be the rule than the exception. The interaction may or may not be well represented by an additive model on a given scale of measurement [34]. However, in order to show that two genes do not act additively, one first needs to have the best possible model for their marginal effects, which in turn requires us to collect samples that are best enriched for the information we need.

The question then becomes: how far will we ever be able to model and understand the complexity of multifactorial diseases? As underscored by the above examples, familial information can be very useful – even crucial at present – to go from population information to modeling the effects of alleles. It has also become clear over the years that formal genetics cannot simply ignore information provided by other scientific fields (epidemiology, pathophysiology, biology, immunology, etc.). There must al-

ways be strong links between formal geneticists and pathophysiologists. Indeed, modeling a multifactorial disease should aim to reconstruct the pathways of the disease process. Twenty years ago, the candidate gene approach was in favor among geneticists, before being replaced by a systematic (so-called ‘unbiased’) genome-wide approach. But when genes are detected through a genome-wide approach without family data, their involvement in a disease can only be justified by the fact that they are good biological candidates. All the multiple sclerosis susceptibility genes detected through genome-wide association studies are good candidates, most being involved in the immune and inflammatory response [35]. The fruitful exchange of information between formal geneticists and pathophysiologists is also well illustrated in Alzheimer’s disease, for which the genes involved in its monogenic forms, as well as the differential risks for APOE genotypes in its multifactorial forms, can be reconciled to explain how amyloid plaque is formed [36, 37]. To detect, with enough power, a statistical interaction between susceptibility genes with very weak marginal effects, one has no alternative but to study genes for which biological interaction exists [38]. More generally, to cope with the complexity of most human disease processes, we require ‘good’ candidate models [39]. Contrary to what Haldane [1] advocated, progress in the mathematical modeling of human genetics requires a strong interaction between the formal geneticist and the biologist [40].

Besides, the role of environmental factors in diseases has already created a strong interface between geneticists and epidemiologists. Unfortunately, in our view, the recent merging of these two fields – genetics and epidemiology – was long overdue and, when it did occur, did not occur in a sensible manner. Instead of enriching the information in the samples studied by the use of family structures, most studies during the last decades have used huge samples of unrelated persons, analyzing them as simple case-control studies. In fact, when related persons are detected in these samples, instead of making use of the extra genetic information they contain, it is considered standard practice to remove them! Furthermore, disease etiology very likely involves not only complex interactions between genetic variants, but also interactions between the genome and the environment. For example, it is clear that predicting the risk of leprosy according to a person’s genetic profile does not make sense if this individual is never exposed to the bacillus. Our knowledge of the environmental factors involved is very poor for many diseases and, contrary to what can be done in animal species, cannot be controlled other than by statistical model-

ing. Each individual has his/her own genotype and environment, different from one individual to another. Inferring an individual's risk from data on large population samples with heterogeneous environmental exposures could be very misleading if the personal environment is ignored. Furthermore, an individual's environment changes over time, so that any risk assessment is a moving target and must be continually updated, both as secular environmental changes occur and as the individual ages. At present, the information that can be obtained on the phenotype and genotype correlations for different kinds of relatives is, in many cases, important in modeling the joint action of the genotype and environment to produce the phenotype.

Environmental exposures during life are generally not predictable at birth and are thus difficult to model. Yet, they can be essential for the eventual phenotypic expression. Knowledge of family members' phenotypes, apart from their genotypes, may be important predictors of an individual's phenotype because they could contain environmental information. This means that family studies are needed to specify individual environments. Putting a hierarchy on genetic and environmental factors in disease causation is often extremely complex. Does, for instance, a unipolar depressive mother predispose a child to this illness for genetic reasons, for environmental reasons, or for both? Similarly, multiple sclerosis could be initiated

by various different environmental factors for which the autoimmune response is itself very heterogeneous. Can one possibly come up with a meaningful and reliable genetic risk without correctly modeling the gene-environment interaction? Assigning a disease risk of 50% at birth solely from genomic information makes no sense if it only takes a single additional piece of information to bring this risk to zero! Examples of such breakthroughs have happened in the past, for example by identifying the agents responsible for infectious diseases. More recent examples are the discovery of the role of asbestos in mesothelioma, of *Helicobacter pylori* in cancer of the stomach and herpes in uterine cancer. Current studies on the microbiome could also provide key information in relation to certain pathological processes. All this new information should be taken into account if we are to make progress in mathematical disease modeling.

So, finally, our initial question becomes: how far shall we be able to go on the way to defining and measuring the genetic predisposition of an individual, and how long will it take? The answer depends on how optimistic one is, and this editorial reflects a discussion between two co-authors – one a pessimist and the other an optimist – on this issue. At any rate, we both believe in the importance of convincing the scientific community that the overwhelming success of biotechnology should not lead to neglecting other sources of information.

References

- Haldane JB: A defense of beanbag genetics. *Perspect Biol Med* 1964;7:343–359.
- Mirnezami R, Nicholson J, Darzi A: Preparing for precision medicine. *N Engl J Med* 2012; 366:489–491.
- Ginsburg GS: Realizing the opportunities of genomics in health care. *JAMA* 2013;309: 1463–1464.
- Tuckson RV, Newcomer L, De Sa JM: Accessing genomic medicine: affordability, diffusion, and disparities. *JAMA* 2013;309:1469–1470.
- O'Rawe J, Jiang T, Sun G, Wu Y, Wang W, Hu J, Bodily P, Tian L, Hakonarson H, Johnson WE, Wei Z, Wang K, Lyon GJ: Low concordance of multiple variant-calling pipelines: practical implications for exome and genome sequencing. *Genome Med* 2013;5:28.
- Gerstein MB, Bruce C, Rozowsky JS, Zheng D, Du J, Korbelt JO, Emanuelsson O, Zhang ZD, Weissman S, Snyder M: What is a gene, post-ENCODE? History and updated definition. *Genome Res* 2007;17:669–681.
- Waddington CH: The epigenotype. 1942. *Int J Epidemiol* 2012;41:10–13.
- Morgan DK, Whitelaw E: The case for trans-generational epigenetic inheritance in humans. *Mamm Genome* 2008;19:394–397.
- Rogers K: Epigenetics: a turning point in our understanding of heredity. *Sci Am* 2012.
- Painter RC, Osmond C, Gluckman P, Hanson M, Phillips DI, Roseboom TJ: Transgenerational effects of prenatal exposure to the Dutch famine on neonatal adiposity and health in later life. *Bjog* 2008;115:1243–1249.
- Pembrey ME, Bygren LO, Kaati G, Edvinsson S, Northstone K, Sjöström M, Golding J: Sex-specific, male-line transgenerational responses in humans. *Eur J Hum Genet* 2006;14:159–166.
- Kaati G, Bygren LO, Pembrey M, Sjöström M: Transgenerational response to nutrition, early life circumstances and longevity. *Eur J Hum Genet* 2007;15:784–790.
- Guibert S, Forne T, Weber M: Global profiling of DNA methylation erasure in mouse primordial germ cells. *Genome Res* 2012;22: 633–641.
- Hackett JA, Sengupta R, Zyllicz JJ, Murakami K, Lee C, Down TA, Surani MA: Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science* 2013;339:448–452.
- Genin E, Feingold J, Clerget-Darpoux F: Identifying modifier genes of monogenic disease: strategies and difficulties. *Hum Genet* 2008;124:357–368.
- Dausset J, Svejgaard A: HLA and Disease. Baltimore, Williams and Wilkins, 1977.
- Butler D: Human genome at ten: science after the sequence. *Nature* 2010;465:1000–1001.
- Fisher R: The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh* 1918;52:399–433.
- Meigs JB, Shrader P, Sullivan LM, McAteer J, Fox CS, Dupuis J, Manning AK, Florez JC, Wilson PW, D'Agostino RB Sr, Cupples LA: Genotype score in addition to common risk factors for prediction of type 2 diabetes. *N Engl J Med* 2008;359:2208–2219.

- 20 Zheng SL, Sun J, Wiklund F, Smith S, Stattin P, Li G, Adami HO, Hsu FC, Zhu Y, Balter K, Kader AK, Turner AR, Liu W, Bleecker ER, Meyers DA, Duggan D, Carpten JD, Chang BL, Isaacs WB, Xu J, Gronberg H: Cumulative association of five genetic variants with prostate cancer. *N Engl J Med* 2008;358:910–919.
- 21 Lu Q, Cui Y, Ye C, Wei C, Elston RC: Bagging optimal ROC curve method for predictive genetic tests, with an application for rheumatoid arthritis. *J Biopharm Stat* 2010;20:401–414.
- 22 Clerget-Darpoux F, Bonaiti-Pellie C, Hochez J: Effects of misspecifying genetic parameters in lod score analysis. *Biometrics* 1986;42:393–399.
- 23 Margaritte-Jeannin P, Babron MC, Bourgey M, Louka AS, Clot F, Percopo S, Coto I, Hugot JP, Ascher H, Sollid LM, Greco L, Clerget-Darpoux F: HLA-DQ relative risks for coeliac disease in European populations: a study of the European Genetics Cluster on Coeliac Disease. *Tissue Antigens* 2004;63:562–567.
- 24 Wang X, Elston RC, Zhu X: The meaning of interaction. *Hum Hered* 2010;70:269–277.
- 25 Bourgey M, Perdry H, Clerget-Darpoux F: Modeling the effect of PTPN22 in rheumatoid arthritis. *BMC Proc* 2007;1(suppl 1):S37.
- 26 International HapMap Consortium: The International HapMap Project. *Nature* 2003;426:789–796.
- 27 Bourgain C, Genin E, Cox N, Clerget-Darpoux F: Are genome-wide association studies all that we need to dissect the genetic component of complex human diseases? *Eur J Hum Genet* 2007;15:260–263.
- 28 Clerget-Darpoux F, Elston RC: Are linkage analysis and the collection of family data dead? Prospects for family studies in the age of genome-wide association. *Hum Hered* 2007;64:91–96.
- 29 Clerget-Darpoux F: Use of association and linkage information for the study of multifactorial diseases; in Pawlowitzki IH, Edwards JH, Thompson EA (eds): *Genetic Mapping of Disease Genes*. Academic Press, 1997, pp 179–187.
- 30 du Montcel ST, Michou L, Petit-Teixeira E, Osorio J, Lemaire I, Lasbleiz S, Pierlot C, Quillet P, Bardin T, Prum B, Cornelis F, Clerget-Darpoux F: New classification of HLA-DRB1 alleles supports the shared epitope hypothesis of rheumatoid arthritis susceptibility. *Arthritis Rheum* 2005;52:1063–1068.
- 31 Babron MC, Perdry H, Handel AE, Ramagopal SV, Damotte V, Fontaine B, Muller-Myhsok B, Ebers GC, Clerget-Darpoux F: Determination of the real effect of genes identified in GWAS: the example of IL2RA in multiple sclerosis. *Eur J Hum Genet* 2012;20:321–325.
- 32 Perdry H, Muller-Myhsok B, Clerget-Darpoux F: Using affected sib-pairs to uncover rare disease variants. *Hum Hered* 2012;74:129–141.
- 33 Feng T, Elston RC, Zhu X: Detecting rare and common variants for complex traits: sib-pair and odds ratio weighted sum statistics (SPWSS, ORWSS). *Genet Epidemiol* 2011;35:398–409.
- 34 Satagopan JM, Elston RC: Evaluation of removable statistical interaction for binary traits. *Stat Med* 2013;32:1164–1190.
- 35 Sawcer S, Hellenthal G, Pirinen M, Spencer CC, Patsopoulos NA, et al: Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* 2011;476:214–219.
- 36 Lannfelt L, Nordstedt C: Genetics of Alzheimer's disease – routes to the pathophysiology. *J Neural Transm Suppl* 2000;59:155–161.
- 37 Dong S, Duan Y, Hu Y, Zheng Z: Advances in the pathogenesis of Alzheimer's disease: a re-evaluation of amyloid cascade hypothesis. *Transl Neurodegener* 2012;1:18.
- 38 Kotti S, Bickeboller H, Clerget-Darpoux F: Strategy for detecting susceptibility genes with weak or no marginal effect. *Hum Hered* 2007;63:85–92.
- 39 Cowper-Salari R, Cole MD, Karagas MR, Lupien M, Moore JH: Layers of epistasis: genome-wide regulatory networks and network approaches to genome-wide association studies. *Wiley Interdiscip Rev Syst Biol Med* 2011;3:513–526.
- 40 Tyler AL, Asselbergs FW, Williams SM, Moore JH: Shadows of complexity: what biological networks reveal about epistasis and pleiotropy. *Bioessays* 2009;31:220–227.

Erratum

In the article 'Functional Analysis of Seven Genes Linked to Body Mass Index and Adiposity by Genome-Wide Association Studies: A Review' by Speakman (*Hum Hered* 2013;75:57–79) reference [123] to figure 10 was omitted.