INVITED REVIEW

Quantitative genetics of disease traits

N.R. Wray & P.M. Visscher

Queensland Brain Institute, The University of Queensland, Brisbane, Qld, Australia

Summary

John James authored two key papers on the theory of risk to relatives for binary disease traits and the relationship between parameters on the observed binary scale and an unobserved scale of liability (James *Annals of Human Genetics*, 1971; **35**: 47; Reich, James and Morris *Annals of Human Genetics*, 1972; **36**: 163). These two papers are John James' most cited papers (198 and 328 citations, November 2014). They have been influential in human genetics and have recently gained renewed popularity because of their relevance to the estimation of quantitative genetics parameters for disease traits using SNP data. In this review, we summarize the two early papers and put them into context. We show recent extensions of the theory for ascertained case–control data and review recent applications in human genetics.

Historical context

Quantitative genetics as a scientific field of research was firmly established in the 1960s, with theory broadly consistent with empirical data from selection experiments in model organisms and genetic improvement programs in animal and plant breeding. In particular, a polygenic model underlying quantitative trait genetic variation was widely accepted. For binary (0–1) traits, a threshold (liability) model had been proposed (Wright 1934) and applied [see for example, the literature reviewed in the introductions of Dempster & Lerner (1950) and Fraser (1976)], but this model was not widely adopted by human geneticists as a model for disease.

For common diseases in humans, such as psychiatric disorders, heart disease and hypertension, the prevailing paradigm in the 1970s was Mendelian, that is the cause of disease in an affected individual is due to a single factor, usually a single mutation or sometimes an environmental insult (e.g. a head injury leading to a brain disease). In fact, for some researchers working in human genetics, this is still the paradigm today, despite strong empirical evidence against this model (as will be discussed later). There are exceptions to the prevalent paradigm of the time, in particular early landmark papers such as Gottesman & Shields 1967 and McGue *et al.* 1983, which suggested that the threshold model was a good model for schizophrenia and provided empirical evidence that the risk to relatives was consistent with such a model.

Liability threshold model

An unobserved liability threshold model to explain observations on discrete characters was first proposed by Sewall Wright in the context of the number of toes in guinea pigs (Wright 1934). The quantitative genetic theory that showed the correspondence between genetic parameters on a scale of liability and an observed binary scale was developed by Dempster and Lerner (1950), in particular in the appendix by Alan Robertson (in this day and age, he would have been a formal co-author). The theory used a linear approximation (from a Taylor series expansion) to go from an additive scale of liability to a discrete scale. Assuming that the liability threshold model is a reasonable model to explain observations on a binary scale, the Dempster & Lerner paper had important implications for response to selection in populations undergoing (mass) selection. In particular, if genetic variation on the scale of liability is fully additive, then genetic

variation on a binary scale can be highly epistatic. This implies that the estimation of additive genetic variance on the observed scale from the resemblance between relatives is biased upwards, the bias depending on whether close relatives (e.g. full-sibs) or more distant relatives (e.g. half-sibs) are used for estimation. In addition, predicted response to mass selection based upon narrow-sense heritability on the 0–1 scale (h_0^2) can be biased downwards or upwards, the bias being a function of heritability of liability (h_l^2), population prevalence and selection intensity.

The linear transformation of heritability from the observed 0–1 scale to that of liability is

$$h_l^2 \approx \frac{h_0^2 K(1-K)}{z^2} \qquad (1)$$

with K the population lifetime prevalence and z the height of the normal curve at the truncation point pertaining to K (Dempster & Lerner 1950). In livestock, this approximation is useful as extensive pedigrees with large half-sibship families allow reasonably unbiased estimation of narrow-sense heritability on the observed scale. In contrast, in human studies (of the pregenomics era), accurate estimates of narrow-sense heritability on the observed scale were not achievable as the sample sizes were limited and disease status could only be recorded on close family members, for example identical or non-identical twin pairs.

Crittenden (1961) and Falconer (1965) realized that the resemblance between relatives on discrete scales can be framed in the theory of response to truncation selection, and derived the estimation of heritability on the scale of liability directly from the lifetime prevalence (called 'incidence' by Falconer) in probands (ascertained individuals who have the trait/disease of interest) and their relatives. Falconer's method to estimate heritability is a linear regression of mean liability of relatives of probands on mean liability of probands, both as a deviation from the population mean. This is analogous to a ratio of response to selection and selection differential, and he showed that heritability of liability could be estimated from two measures: risk of disease in the population (i.e., K) and risk of disease in relatives of those affected (K_R),

$$h_l^2 = \frac{T - T_R}{a_R i} \qquad (2)$$

where a_R is the additive genetic relationship between the relatives, $T = \Phi^{-1}(1-K)$ and $T_R = \Phi^{-1}(1-K_R)$ are the thresholds of the normal distribution that truncate proportions K and K_R, respectively; i is the

mean liability of the diseased group in the population, calculated as $i = z/K$ where $z = \phi(T)$, as in (1). Falconer provided an approximation to the sampling variance of the estimate of heritability (which he attributed to Barnet 'Woggy' Woolf, an Edinburgh colleague). Falconer recognized that his derivation assumed that the variance in liability amongst relatives of probands was the same as the variance in liability amongst probands, arguing that the reduction in variance was negligible. Edwards (1969) and Smith (1970) showed that proper accounting for the reduction in variance increased estimates of heritability by approximately 10%. An important extension from Charlie Smith (who made major contributions to both human and livestock genetics) was the derivation of the expected disease concordance rate for monozygous (MZ) twin pairs under a liability threshold model and the important observation that the expected concordance rate can be low even when the heritability of liability is high. Conversely, a low MZ concordance rate for a disease with prevalence of, say, 1% or less does not imply that genetic factors are unimportant (Smith 1970). To this day, there is much confusion in human genetics about the relationship between heritability and disease concordance in relatives. Smith expanded on this study by showing how a proband concordance rate can be used to estimate heritability of liability from a design including both dizygotic (DZ) and MZ twin pairs (Smith 1974).

James 1971

In his landmark short paper from 1971, James addressed the question of the risk to relatives of probands. He showed that on the observed probability (or risk or disease) scale, the risk to a relative with relationship R is simply

$$K_R = K + \frac{cov_R}{K}, \qquad (3)$$

with cov_R the phenotypic covariance between the proband and the relative on the observed 0–1 scale. This equation is completely general and does not depend on assumptions about the sources of the phenotypic covariance. It can also be expressed as

$$\lambda_R - 1 = \frac{cov_R}{K^2} \qquad (4)$$

with λ_R the relative risk to relatives (K_R/K) (Risch 1990). Equations (3) and (4) are sometimes referred to as the 'James' Identity' (Lynch & Walsh 1998).

If we now assume a genetic model such that the only covariance between relatives is due to genetic

factors, then the phenotypic covariance on the 0–1 scale can be decomposed into genetic variance components (James 1971),

$$cov_R = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} r^k u^l V_{A(k)D(l)},$$

where $V_{A(k)D(l)}$ denotes the genetic variance components with k A and l D terms, given an additive genetic relationship coefficient of r and a dominance coefficient of relationship of u. So for R = MZ twin, $r = 1$ and $u = 1$ then,

$$cov_{MZ} = V_A + V_D + V_{AA} + V_{AD} + V_{DD} + V_{AAD} + V_{AAA}$$
$$+ \cdots = V_G \tag{5}$$

Likewise for R = full-sib (FS), where $r = \frac{1}{2}$ and $u = \frac{1}{4}$ then,

$$cov_{FS} = \frac{V_A}{2} + \frac{V_D}{4} + \frac{V_{AA}}{4} + \frac{V_{AD}}{8} + \frac{V_{DD}}{16} + \frac{V_{AAD}}{16} + \frac{V_{AAA}}{8}$$
$$+ \cdots \tag{6}$$

Despite being <3 pages in length, the importance of James (1971) is that this formulation provides a framework to test the observed frequencies (or relative risks) in relatives against single locus versus multiple locus models. A single locus model for human disease – the prevailing view at that time – does not give rise to epistatic variance and hence gives different predictions of risk across different classes of relatives than multilocus models (James 1971). Therefore, it provides a strategy, in principle (if data sets are sufficiently large and if there are multiple classes of relatives), to test the validity of a single locus model compared to a multilocus model from observable data (O'Rourke *et al.* 1982; Craddock *et al.* 1995). A particular multilocus model is the threshold model that Falconer (1965) developed, and James cites this paper. He also cites Dempster & Lerner (1950), as they showed that epistatic variance on the observed scale may be high when narrow-sense heritability of liability is high and the proportion of the population with the affected trait is close to zero, consistent with the epidemiology of many human diseases.

Reich, James & Morris 1972

Several papers of the 1970s era point, often obliquely, to the controversy of the liability threshold model, see for example the discussion of Curnow & Smith's (1975) 'Multifactorial models for familial diseases in man' after presentation at the Royal Statistical Society. In trying to set the record straight,

Fraser (1976) in 'The multifactorial/threshold concept – uses/misuses' openly acknowledges the controversy, in which some considered the model as the 'first (and perhaps only) rational explanation for the familial patterns shown by various relatively common human malformations and other disorders', whilst others [citing (Melnick & Shields 1976)] had referred to it as 'tautological, based on grandiose assumptions, having no experimental support, and providing no testable hypotheses.' It was into this melee that Reich *et al.* (1972) made an important contribution.

In the study by Reich *et al.* (1972), the liability threshold model developed by Falconer (1965) is expanded to a two- or more- threshold model. Under this extension, it is recognized that many diseases (their examples were diabetes, schizophrenia, malignant hypertension, cleft palate and centrencephalic epilepsy) can be viewed as semi-continuous traits, so that affected individuals manifest variable severity or subforms. The subdivision of a disease category together with the hypothesis that the severity of disease is inherited allows differentiation of alternative genetic models in a way not possible with a simple all-or-none classification. The multi-threshold model has particular relevance if the less severe disease phenotype is considerably more common than the severe disease phenotype. Hence, the study quantifies under what circumstances it may be possible to differentiate between single gene and liability threshold models and proposes goodness-of-fit statistics to evaluate the model that best fits the observed data. They show that MZ twins are particularly useful to differentiate between models because they share all non-additive effects. James co-authored a second paper with Reich (Reich *et al.* 1979), which is a more general formulation of the 1972 paper, that allows different ascertainment for, and different environmental variables contributing to, disease subforms.

Reich *et al.*'s (1972) paper is also important because it provided a neat formulation based on normal distribution theory for the reduction in variance in relatives of affected probands, to generate a more accurate estimate of the heritability of liability (Falconer 1967; Edwards 1969; Smith 1970), such that equation (2) becomes

$$h_l^2 = \frac{T - T_R \sqrt{1 - (1 - T/i)(T^2 - T_R^2)}}{a_R(i + (i - T)T_R^2)}$$

This more complex formulation still requires only two measures – the risks of disease in the population and

in the relatives of those affected, i.e., K and K_R, and has been widely used.

Personal context

The two papers that John James contributed to human genetics resulted from his first visit to the Institute of Animal Genetics at the University of Edinburgh, a year-long sabbatical in 1970. Although we studied in Edinburgh in the late 1980s and regularly chatted with Douglas Falconer and Charlie Smith, the importance of the liability threshold model theory seemed to pass us by. It was one of the last chapters in Falconer's text book (Falconer & Mackay 1996) and hence one of the last lectures of the year in the MSc course of Quantitative Genetics (so maybe PMV was not paying attention, although he does remember a very useful tutorial by Chris Haley in which the lecture was explained!). It is only after moving into human disease genetics that we have fully appreciated the contributions from Edinburgh to this field in the 1950s–1970s. Why was this theory deemed so unimportant in 1980s Edinburgh? Firstly, even by the time of Falconer's (1965) publication, the liability threshold model was rather standard fare, at least in the non-human genetics circles [Bill Hill, personal communication – but this sentiment can be detected in the introduction of Falconer (1965)]. In fact, in a Perspectives article published by Falconer in 1993 entitled 'Quantitative Genetics in Edinburgh 1947–1980', he fails to mention the liability threshold model (although he does say that he introduced the concept of realized heritability as a way of describing selection response, and perhaps he considered his work on the threshold model as a small aside to this work). Secondly, at that time, livestock genetics was the focus of the Edinburgh School, and in livestock the binary traits of economic interest compared to human disease traits have higher prevalences and little, if any, ascertainment biases. Moreover, the large sibships possible in cattle and poultry allowed approximate estimation of narrow-sense heritability on the observed scale for direct application of equation (1) [see Gianola (1982) for applications of the liability threshold model in livestock]. In contrast, the human genetics of the 1980s was limited; it was the pregenomics era and the large national Scandinavian data resources, that more recently have heavily utilized liability threshold theory (Lichtenstein *et al.* 2009), were in their infancy. It is only in the last decades that the empirical data have really met their match with the theory laid down in the 1950s–1970s. In moving our research focus from livestock to humans as the genomic era evolved, we have revisited the liability threshold model papers and have been humbled by the nuance of understanding conveyed in detail of those texts.

We have focused much of our disease-related research to psychiatric disorders, and it is, therefore, of particular interest in us to realize that the driving force in John's important contribution to human genetics, Theodore (Ted) Reich, was a prominent psychiatrist from the Washington University in St Louis, who had chosen to spend two of his formative years (1969 and 1970) in Edinburgh on a US-funded fellowship. We are told (personally, PMV met Ted Reich in the late 1990s) that he chose Edinburgh for his fellowship because of the growing support for a polygenic genetic architecture for psychiatric disorders such as schizophrenia. The third author of the paper was then PhD student Chris Morris, with whom Ted Reich shared an office during this time, and who undertook the calculations for the paper (Chris Morris, personal communication). We are told (Bill Hill, personal communication) that Ted stimulated many discussions about human disease genetics, not least in the famous Alan Robertson coffee sessions (as described in Falconer 1993). It is only in the last 5 years that the empirical data have provided unequivocal support for this hypothesis, and we are proud to have contributed to analyses that demonstrate this (Purcell *et al.* 2009; Lee *et al.* 2012). The steady stream of citations to James (1971) and Reich *et al.* (1972) demonstrate that these papers continue to be of relevance in the genomics era.

James 1971 and Reich, James & Morris 1972 in the genomics era

Genome-wide association studies (GWAS) and exome or whole-genome sequencing studies are used in human (medical) genetics to discover genes and gene variants that are associated with risk to disease. But data from genome-wide coverage by genetic markers can also be used to estimate genetic parameters. For example, one parameter of interest is the proportion of variation that is captured with SNP arrays (sometimes called the 'SNP-heritability'), first developed and applied in livestock (Fernando 1998; Habier *et al.* 2007; VanRaden 2008; Goddard 2009), then applied to quantitative traits in humans (Yang *et al.* 2010; Lee *et al.* 2011). By contrasting genetic variation estimated from pedigree data with that associated with a SNP array, the proportion of 'missing' heritability can be quantified. If most genetic variation is captured by a SNP array, then larger experimental sample size will

allow the discovery of additional variants, genes and biological pathways that contribute to genetic variation. If, in contrast, only a small proportion of overall genetic variation is captured due to LD with SNPs on the array, consistent with rare variants contributing to disease risk, then a different experimental design (based upon sequencing) may be required to discover those variants.

Analyses to estimate SNP-heritability are usually performed on the observed 0–1 scale using regression or REML and then transformed to a scale of liability (Lee *et al.* 2011). Most disease studies in humans are based on case–control studies in which cases are heavily over-sampled relative to the prevalence of disease in the population. This creates a non-normal distribution of liability in the sample and the ascertainment needs to be taken into account when transforming the parameters on the observed 0–1 scale to that of liability. Lee *et al.* (2011) derived a generalization of the Dempster & Lerner (1950) transformation as

$$h_l^2 \approx \frac{h_0^2 K^2 (1-K)^2}{z^2 P(1-P)} \qquad (7)$$

with P being the proportion of cases in the case–control sample (usually $P \sim \frac{1}{2}$). For a random sample $P = K$ and the expression is identical to (1). The constant factor in equation (7) is 1.0 when $K \sim 0.085$, so at this prevalence the heritability on the two scales are the same. Note that expression (7) is an approximation and can result in bias when ascertainment is extreme and heritability on the liability scale is high. When $P = \frac{1}{2}$ and K small so that $(1 - K) \sim 1$ and expressing 'heritability in the case–control sample' as a function of liability in the population,

$$h_0^2 \approx \frac{(1/4) h_l^2 z^2}{K^2} = (1/4) h_l^2 i^2.$$

For small values of K, this predicts the heritability on the 0–1 scale in a case–control population >1. The reason for this apparent inflation of heritability on the 0–1 scale is that the ascertainment has created much more genetic variation in the case–control sample than exists in the population, so that the ratio V_g/V_p on the 0–1 scale is larger than 1. However, in practice, the proportion of variation in liability captured by SNP arrays is usually $<\frac{1}{2}$ and estimating variance components on the 0–1 scale by REML without constraining heritability to be between 0 and 1 can give estimates with little bias (Lee *et al.* 2011; Speed *et al.* 2012; Zaitlen & Kraft 2012; Gusev *et al.* 2013). Hence, for most practical applications, the procedure of a linear analysis on the 0–1 scale followed by a

transformation to liability that accounts for over-sampling of cases is a reasonable and useful exercise.

Conclusion

The two papers from John James in human genetics have been important contributions to that field because they laid theoretical foundations about polygenic models underlying liability to disease and their relationship with observations on risk to relatives. They are vital links between quantitative genetics and epidemiology and have been cited extensively in methodological exploration of the unknown underlying genetic architecture of human disease (O'Rourke *et al.* 1982; Risch 1987; Rich 1990; Risch 1990; Holmans 1993; Craddock *et al.* 1995; Pritchard 2001; Slatkin 2008; Wray *et al.* 2010). It is only in the last 5–10 years that the tools have become available to dissect genetic variation for common diseases into contributions from specific variants. To date, genomic data sets that record disease status as semi-continuous have not been generated, but the very large population-based samples currently in the pipeline may enable testing of the multiple threshold model presented in the study by Reich *et al.* (1972). However, it is clear that the empirical data from the last 10 years are consistent with highly polygenic model for most common diseases. With the *caveat* that all models are wrong but some may be useful, the liability threshold model continues to be very useful and John James has made a lasting contribution to the theory underlying this model.

Acknowledgements

References

Craddock N., Khodel V., Van Eerdewegh P., Reich T. (1995) Mathematical limits of multilocus models: the genetic transmission of bipolar disorder. *Am. J. Hum. Genet.*, **57**, 690–702.

Crittenden L. (1961) An interpretation of familial aggregation based on multiple genetic and environmental factors. *Ann. N. Y. Acad. Sci.*, **91**, 769–780.

Curnow R.N., Smith C. (1975) Multifactorial models for familial diseases in man. *J. R. Stat. Soc. Ser. A. Stat. Soc.*, **138**, 131–169.

Dempster E.R., Lerner I.M. (1950) Heritability of threshold characters. *Genetics*, **35**, 212–236.

Edwards J.H. (1969) Familial predisposition in man. *Br. Med. Bull.*, **25**, 58–64.

Falconer D. (1965) The inheritance of liability to certain diseases, estimates from the incidence among relatives. *Ann. Hum. Genet.*, **29**, 51–76.

Falconer D.S. (1967) Inheritance of liability to diseases with variable age of onset with particular reference to diabetes mellitus. *Ann. Hum. Genet.*, **31**, 1–20.

Falconer D. (1993) Quantitative genetics in Edinburgh – 1947–1980. *Genetics*, **133**, 137–142.

Falconer D., Mackay T.F. (1996) Introduction to Quantitative Genetics. Pearson Education Ltd, Harlow, UK.

Fernando R.L. (1998) Genetic evaluation and selection using genotypic, phenotypic and pedigree information. *Proc. 6th World Congr. Genet. Appl. Livest. Prod.*, **26**, 329–336.

Fraser F.C. (1976) Multifactorial-threshold concept – uses and misuses. *Teratology*, **14**, 267–280.

Gianola D. (1982) Theory and analysis of threshold characters. *J. Anim. Sci.*, **54**, 1079–1096.

Goddard M.E. (2009) Genomic selection: prediction of accuracy and maximisation of long term response. *Genetica*, **136**, 245–257.

Gottesman I.I., Shields J. (1967) A polygenic theory of schizophrenia. *Proc. Natl Acad. Sci. USA*, **58**, 199–205.

Gusev A., Bhatia G., Zaitlen N., Vilhjalmsson B.J., Diogo D., Stahl E.A., Gregersen P.K., Worthington J., Klareskog L., Raychaudhuri S., Plenge R.M., Pasaniuc B., Price A.L. (2013) Quantifying missing heritability at known GWAS loci. *PLoS Genet.*, **9**, e1003993.

Habier D., Fernando R.L., Dekkers J.C. (2007) The impact of genetic relationship information on genome-assisted breeding values. *Genetics*, **177**, 2389–2397.

Holmans P. (1993) Asymptotic properties of affected-sib-pair linkage analysis. *Am. J. Hum. Genet.*, **52**, 362–374.

James J.W. (1971) Frequency in relatives for an all-or-none trait. *Ann. Hum. Genet.*, **35**, 47–49.

Lee S.H., Wray N.R., Goddard M.E., Visscher P.M. (2011) Estimating missing heritability for disease from genome-wide association studies. *Am. J. Hum. Genet.*, **88**, 294–305.

Lee S.H., DeCandia T.R., Ripke S., Yang J., Sullivan P.F., Goddard M.E., Keller M.C., Visscher P.M., Wray N.R. (2012) Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nat. Genet.*, **44**, 247–250.

Lichtenstein P., Yip B.H., Bjork C., Pawitan Y., Cannon T.D., Sullivan P.F., Hultman C.M. (2009) Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet*, **373**, 234–239.

Lynch M., Walsh B. (1998) Genetics and Analysis of Quantitative Traits. Sinauer Associates, Sunderland, Massachusetts.

McGue M., Gottesman I.I., Rao D.C. (1983) The transmission of schizophrenia under a multifactorial threshold-model. *Am. J. Hum. Genet.*, **35**, 1161–1178.

Melnick M., Shields E.D. (1976) Allelic restriction – biologic alternative to multifactorial threshold inheritance. *Lancet*, **1**, 176–179.

O'Rourke D.H., Gottesman I.I., Suarez B.K., Rice J., Reich T. (1982) Refutation of the general single-locus model for the etiology of schizophrenia. *Am. J. Hum. Genet.*, **34**, 630–649.

Pritchard J.K. (2001) Are rare variants responsible for susceptibility to complex diseases? *Am. J. Hum. Genet.*, **69**, 124–137.

Purcell S.M., Wray N.R., Stone J.L., Visscher P.M., O'Donovan M.C., Sullivan P.F., Sklar P. (2009) Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, **460**, 748–752.

Reich T., James J.W., Morris C.A. (1972) Use of multiple thresholds in determining mode of transmission of semi-continuous traits. *Ann. Hum. Genet.*, **36**, 163–184.

Reich T., Rice J., Cloninger C.R., Wette R., James J. (1979) Use of multiple thresholds and segregation analysis in analyzing the phenotypic heterogeneity of multifactorial traits. *Ann. Hum. Genet.*, **42**, 371–390.

Rich S.S. (1990) Mapping genes in diabetes. Genetic epidemiological perspective. *Diabetes*, **39**, 1315–1319.

Risch N. (1987) Assessing the role of HLA-linked and unlinked determinants of disease. *Am. J. Hum. Genet.*, **40**, 1–14.

Risch N. (1990) Linkage strategies for genetically complex traits. 1. Multilocus models. *Am. J. Hum. Genet.*, **46**, 222–228.

Slatkin M. (2008) Exchangeable models of complex inherited diseases. *Genetics*, **179**, 2253–2261.

Smith C. (1970) Heritability of liability and concordance in monozygous twins. *Ann. Hum. Genet.*, **34**, 85–91.

Smith C. (1974) Concordance in twins – methods and interpretation. *Am. J. Hum. Genet.*, **26**, 454–466.

Speed D., Hemani G., Johnson M.R., Balding D.J. (2012) Improved heritability estimation from genome-wide SNPs. *Am. J. Hum. Genet.*, **91**, 1011–1021.

VanRaden P.M. (2008) Efficient methods to compute genomic predictions. *J. Dairy Sci.*, **91**, 4414–4423.

Wray N.R., Yang J., Goddard M.E., Visscher P.M. (2010) The genetic interpretation of area under the ROC curve in genomic profiling. *PLoS Genet.*, **6**, 1–9.

Wright S. (1934) An analysis of variability in number of digits in an inbred strain of guinea pigs. *Genetics*, **19**, 0506–0536.

Yang J., Benyamin B., McEvoy B.P., Gordon S., Henders A.K., Nyholt D.R., Madden P.A., Heath A.C., Martin N.G., Montgomery G.W., Goddard M.E., Visscher P.M. (2010) Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.*, **42**, 565–569.

Zaitlen N., Kraft P. (2012) Heritability in the genome-wide association era. *Hum. Genet.*, **131**, 1655–1664.