Heritability

WG Hill, University of Edinburgh, Edinburgh, UK

Introduction

Heritability is a commonly used and important term to describe properties of the inheritance of quantitative traits, such as stature in man or milk yield of cows. Heritability is also used similarly for complex traits, such as incidence of a disease, that have discrete phenotypes but do not have simple Mendelian inheritance. Informally, heritability (h^2) is the proportion of the variation in the trait due to genetic differences between individuals, but a more precise definition of heritability is important because the term is both widely used and widely misused. Correlations among relatives and response to directional selection are proportional to the heritability. The first use of 'heritability' is most often associated with J. L. Lush, who applied the theory of quantitative genetics of Wright and Fisher to animal breeding.

Definition

The observed performance or phenotypic value, P, of an individual for a quantitative trait expressed as a deviation from the population mean can be partitioned in an additive model into two components: genotypic value (G) and environmental deviation (E):

$$P = G + E$$

A genotype × environment interaction term, GE, can also be included, but cannot usually be distinguished from E as each environmental deviation is unique. Because individuals transmit only one gene at each locus to their offspring, in describing correlations among most relatives and predicting responses to selection, it is necessary to consider the average performance of individuals who receive one copy of a specified gene and the other at random from their other parent in a population, called the average effect of the gene. The breeding value, A, of an individual is the sum of average effects of its genes. More simply and practically, the breeding value of an individual is defined as twice the expected deviation, if randomly mated, of the mean of its progeny from the population mean; but these definitions are the same unless there is epistasis. The dominance deviation, D, defines differences between genotypic value and breeding value due to interactions between genes at individual loci, and the epistatic deviation, I, defines differences due to interactions between different loci. A Fuller model is

$$P = A + D + I + E$$

Variation among individuals in phenotypic value, V_P, can be partitioned into components:

$$V_P = V_G + V_E = V_A + V_D + V_I + V_E$$

assuming that correlations between or interactions of genotype and environment can be ignored or catered for in other ways.

There are two different definitions of heritability:

- Heritability in the broad sense: $H^2 = V_G / V_P$
- Heritability in the narrow sense: $h^2 = V_A / V_P$. This is the more useful and easily estimated quantity, and so is usually implied and termed simply 'heritability'.

Heritability appears as a squared term, because h was first defined by Wright as the path coefficient from genotype to phenotype. The correlation between breeding value (A) and phenotype (P), the accuracy of selection on phenotype, equals h; the regression of breeding value on phenotype equals h^2; and the variance in A that is not explained by P is $V(A|P) = (1 - h^2)V_A$.

Magnitude

The magnitude of the genetic variance depends on the frequencies and effects of genes at many loci, and the environmental variation depends on the environment in which individuals are kept; therefore, heritability differs among traits, species, populations, and over time. In practice, however, each trait takes typical

Table 1   Typical heritability values for different species and traits (various sources)

Species and traits	h^2 (%)
Drosophila	
Life history traits (longevity, fecundity, development time)	12
Behavioral traits (locomotion, mating activity, geo- and phototaxis)	18
Morphological traits (bristle number, wing, and thorax size)	32
Pig	
Reproductive rate (litter size)	10
Growth rate (daily gain, feed intake, and conversion efficiency)	30
Morphology (backfat, carcass lean %)	45
Human	
Intelligence quotient (meta-analysis)	34
	48 (broad sense)
Stature	65
Finger ridge count	>95

heritability values, often similar among very different species. For example, traits of conformation or morphology usually have heritability of 0.5 or more, those of growth ∼0.3, and those of reproduction 0.2 or less (examples in Table 1).

Discrete Traits

Although primarily used for traits with continuous expression such as stature, heritability can also be applied to traits with two or a few discrete phenotypes that do not have simple Mendelian expression, such as survival to weaning and incidence of twinning of a congenital defect such as club foot. One way is to simply regard the traits as having two values, say 1 (affected) and 0 (unaffected), and ignore any nonlinearity or heterogeneity of variance. Within the quantitative genetics framework, the discrete (all-or-none) trait can be considered as the expression of some underlying conceptual continuous variable or liability, with a threshold value above which affected individuals lie. Heritability on the all-or-none scale is lower than on the liability scale, increasingly so as the incidence of the trait departs from 0.5.

Estimation

Heritability is estimated from the similarity of relatives. The two major problems are to avoid confounding the genetic similarity of relatives by nongenetic causes such as shared environment; and to obtain sufficient data to provide accurate estimates.

Heritability in the broad sense can be estimated from the correlation of phenotypes of individuals that have the same genotype (i.e., clones or identical twins), but different environmental effects. In plants, this can be feasible, but human identical twins share prenatal and also, unless adopted, postnatal environment.

Relatives resemble each other because they have genes in common, and the closer their relationship the more likely they are to share genes and the more highly correlated are their phenotypes for quantitative traits. Similarly, the higher the heritability, the more highly correlated are the phenotypes of relatives. The (narrow-sense) heritability is therefore estimated from the resemblance between relatives, scaled to take account of the relationship. Although estimates from distant relatives may be less confounded by common environment, they have a higher sampling error. Traditionally, relationship has been established by pedigree, but it is now becoming feasible to determine actual relationship consequent on Mendelian segregation from dense single nucleotide polymorphism (SNP) marker information.

Parent and Offspring

The covariance of parent and offspring, which have precisely one (autosomal) gene in common at each locus, equals $V_A/2$ (here and subsequently, ignoring epistasis). Hence, for randomly sampled individuals, the correlation between phenotype of offspring and phenotype of individual parent is $h^2/2$. Similarly the regression of phenotype of offspring on phenotype of one parent is $h^2/2$ and on parental mean is h^2. (The word regression was coined by Galton to describe the observation that extreme parents tended to have less extreme offspring.) Hence, if a set of data on parent and offspring are collected and the regression ($\pm$SE) of progeny on one parent's phenotype is 0.2 (± 0.1), the estimate of heritability is 0.4 (± 0.2). Maternal effects can bias estimates from offspring–parent regression or correlation, for example, in body weight due to birth weight. If there is nonrandom mating among parents, the regression or correlation of offspring on individual parent is biased (upward with positive assortative mating), but the regression on midparent is not.

Full and Half Sibs

Full sibs share 0, 1, or 2 parental genes at each locus, with their respective probabilities 1/4, 1/2, and 1/4, and half sibs 0 or 1, each with probability 1/2. Therefore the genetic covariance of full sibs equals $V_A/2 + V_D/4$ and of half sibs $V_A/4$. In mammals, for example, full sibs also share a common environment, which contributes an environmental correlation between sibs, often denoted c^2. Sib data are classically subjected to analysis of variance, and heritability is estimated from the intraclass correlation, which, if there is no environmental confounding, estimates $h^2/2$ for full sibs and $h^2/4$ for half sibs. If each male has several mates, both full- and half-sib correlations can be

estimated, but the half-sib estimate is usually taken because it is less likely to be confounded by common environment and dominance, although it has a higher sampling error. Positive assortative mating can increase the correlation among sibs.

Twins

There are considerable problems in eliminating common environment effects for heritability estimation in man. The use of twins provides a route, specifically by comparing the correlations of identical (monozygous, MZ) and nonidentical (dizygous, DZ) twins. If the MZ correlation is assumed to equal $h^2 + c^2$ and the DZ correlation $h^2/2 + c^2$, an estimate of heritability is $2[\text{corr(MZ)} - \text{corr(DZ)}]$. This is biased upward by all nonadditive genetic effects (dominance and epistasis), and by any extra environmental similarity of MZ over DZ.

Combination of Information (Animal Model)

To make best use of information on all relatives, whether close or distant in the same or different generation, particularly from field data, sophisticated models and computer-intensive statistical methods using (restricted) maximum likelihood or Bayes' theorem (via Gibbs sampling) are adopted using an 'animal model'. This incorporates the relationship matrix to define the correlations among observations on all individuals in the data, and accounts for identifiable environmental differences such as location or year of birth. These methods are replacing simple regression or correlation analyses in most applications because they are efficient and enable successively more complicated models to be fitted and assessed using likelihood ratio tests.

Within Families

A novel method of estimation was proposed by Visscher which utilizes dense markers to estimate the actual proportion of genome shared by full sibs. As this varies around 0.5 by Mendelian sampling, with a standard deviation of about 0.04 for human, heritability can be estimated within families free of environmental confounding.

Selection Response

As the regression of offspring on midparent phenotype equals heritability and is approximately linear under polygenic inheritance (exactly linear under multivariate normality), the regression of the offspring of a group of individuals on the mean of their parents' phenotype equals h^2. Hence, if a group of individuals are selected which differ in phenotype by the selection differential S, their offspring will be expected to deviate in performance from those of unselected parents by an amount h^2S. This is the selection response, $R = h^2S$, from the 'breeder's equation'. If, therefore, environmental change over generations can be eliminated or corrected for using an unselected control population, heritability can be estimated by Falconer's 'realized heritability', $h^2 = R/S$.

Uses and Possible Abuses

Heritability tells us no more than the additive genetic and phenotypic variances do separately, but it is a useful summary and descriptive parameter. Just as the correlation among relatives can be used to estimate heritability, the heritability can be used to predict the correlation of relatives and, perhaps most importantly, the selection response from selection differential. A comparison between the heritability predicted from collateral relatives such as half sibs and the realized heritability or selection response provides a check on quantitative genetics theory.

Heritability applies within populations, and therefore cannot be used to predict genetic differences between races or other populations from phenotypic differences, whether or not they share the same environment.

The prediction formula $R = h^2S$ usually holds only if selection is practiced on the trait on which response is measured. If selection is practiced on some other trait or some combination of traits, the regression of response on selection differential depends on the genetic and phenotypic correlations among the traits. This is a serious problem in inferences about selection in nature, where the actual selection applied is not known.

Heritability is a summary parameter over loci and does not tell us about either the numbers of genes that affect a quantitative trait or the magnitude of their effects. It can change as a population evolves, for example, by selection and inbreeding, although heritability typically changes slowly with polygenic inheritance.

See also: Artificial Selection.

Further Reading

Barton NH, Briggs DEG, Eisen JA, Goldstein DB, and Patel NH (2007) *Evolution.* Cold Spring Harbor, NY: Cold Spring Harbor.

Bell AE (1977) Heritability in retrospect. *Journal of Heredity* 68: 297–300.

Devlin B, Daniels M, and Roeder K (1997) The heritability of IQ. *Nature* 388: 468–471.

Falconer DS and Mackay TFC (1996) *Introduction to Quantitative Genetics*, 4th edn. Harlow: Longman.

Hartl DL and Clark AG (2007) *Principles of Population Genetics*, 4th edn. Sunderland, MA: Sinauer Associate, Inc.

Kearsey M and Pooni HS (1996) *The Genetical Analysis of Quantitative Traits.* London: Chapman & Hall.

Lynch M and Walsh B (1998) *Genetics and Analysis of Quantitative Traits.* Sunderland, MA: Sinauer Associates, Inc.

Roff DA (1997) *Evolutionary Quantitative Genetics.* New York: Chapman & Hall.

Visscher PM, Hill WG, and Wray NR (2008) Heritability in the genomics area – Concepts and misconceptions. *Nature Review Genetics* 9: 255–266.

Walsh B and Lynch M (2009) *Genetics and Analysis of Quantitative Traits, Vol. 2: Evolution and Selection of Quantitative Traits.* http://nitro.biosci.arizona.edu/zbook/NewVolume_2/newvol2.html (accessed March 2010).