Genetics: Published Articles Ahead of Print, published on December 15, 2008 as 10.1534/genetics.108.099218

Association mapping in outbred populations: power and efficiency when genotyping parents and phenotyping progeny

Stephen F. Chenoweth[*] and Peter M. Visscher[†]

* School of Integrative Biology, University of Queensland, St Lucia 4072, QLD, Australia.

† Queensland Statistical Genetics, Queensland Institute of Medical Research, Brisbane, Australia

ABSTRACT

We develop expressions for the power to detect associations between parental genotypes and offspring phenotypes for quantitative traits. Three different 'indirect' experimental designs are considered consisting of full-sib, half-sib and full-sib half-sib families. We compare the power of these designs to detect genotype-phenotype associations relative to the common, 'direct', approach of genotyping and phenotyping the same sample of individuals. When heritability is low, the indirect designs are powerful, to the extent that they can outperform the direct method. However, the extra power comes at a cost due to an increased phenotyping effort. We show how the extra costs associated with phenotyping a large number of individuals will influence experimental design decisions, developing expressions for optimal experimental designs given the cost of phenotyping relative to genotyping. Our results suggest that indirect association studies can be a powerful means of detecting allelic associations in outbred populations of species for which genotyping and phenotyping the same individuals is impractical and for life-history and behavioural traits that are heavily influenced by environmental variance and therefore best measured on groups of individuals. Indirect association studies are only likely to be favoured on purely economical grounds however, when phenotyping is substantially less expensive than genotyping. A web-based application implementing our expressions has been developed to aid in the design of indirect association studies.

Running head: Indirect association studies

Keywords: full-sib, half-sib, non-centrality parameter, statistical power, outbred populations, association, SNP

Correspondence: Stephen Chenoweth, School of Integrative Biology, University of Queensland, St Lucia 4072, QLD, Australia *email: s.chenoweth@uq.edu.au Tel: +61-7-3365-2188, Fax: +61-7-3365-1655*

INTRODUCTION

Predictions of the evolutionary dynamics of quantitative traits are highly sensitive to the genetic architecture assumed to underlie them (BARTON and TURELLI 1989). Accordingly, in recent times, there has been a large amount of empirical attention paid to the dissection of quantitative traits at the genomic level (BARTON and KEIGHTLEY 2002). Classical methods for trait dissection such as quantitative trait locus (QTL) analysis by linkage rely upon linkage-disequilibrium between marker loci and functional polymorphisms within pedigrees, and although useful for determining whether a particular genomic region may affect a trait, suffer from undesirable properties such as narrow sampling of naturally occurring allelic variation, and a limited scope for distinguishing physical linkage from pleiotropy (MACKAY 2001). As these weaknesses all concern critical aspects of genetic architecture (HANSEN 2006), they must be overcome if we are to better understand the genetic architecture of quantitative traits. A common alternative to analysis via linkage is to test for direct association between single nucleotide polymorphisms (SNPs) in candidate gene regions (or the entire genome) and quantitative traits. For example, when linkage disequlibrium is weak, an association study may be more useful for distinguishing linkage from pleiotropy (e.g. CARBONE *et al.* 2006).

Typically, association studies are conducted on a sample of individuals that are scored for a phenotype and also genotyped (LONG and LANGLEY 1999), we refer to this as the *direct* approach. Despite the significant sample sizes required, direct approaches have identified candidate polymorphisms contributing to quantitative trait variation in natural populations in a wide range of organisms including humans (LETTRE *et al.* 2008; WEEDON *et al.* 2008), and model organisms such as *Drosophila* (DWORKIN *et al.* 2003; DWORKIN *et al.* 2005; LONG *et al.* 1998; LONG *et al.* 2000; ROBIN *et al.* 2002), *Arabidopsis* (ARANZANA *et al.* 2005) and mice (LIU *et al.* 2006). Although replication of initial associations is usually required to eliminate false positives and explore

population specific effects (e.g. GRUBER *et al.* 2007; MACDONALD and LONG 2004). The direct approach is suitable when a candidate gene's contribution to naturally-occurring phenotypic variance is of interest (via the use of field-sampled individuals) and for traits that can be measured with minimal error, such as morphology. However, many fitness-related traits such as behavioural and life-history traits can be heavily influenced by environmental variance (HOULE 1992), making them difficult to measure on single field-caught individuals. Such traits are often impossible to measure under field conditions and are more reliably measured on groups of individuals.

For some organisms, performing association studies in a panel of inbred lines (ILs) or doubled haploid lines (DHLs) can be one solution to the requirement of phenotyping multiple individuals and or traits per genotype (e.g. DWORKIN *et al.* 2003). The IL approach does suffer some disadvantages in that the genetic structure of an IL panel does not accurately reflect the genetic structure of outbred populations, and the variances of quantitative traits are often altered as a consequence of inbreeding (HILL and CABALLERO 1992; VAN BUSKIRK and WILLI 2006). For many species, inbreeding is also impractical. Clearly, alternative approaches are required that can be implemented in outbred natural populations.

One potential solution is an *indirect* method whereby a set of parents are genotyped and multiple progeny are phenotyped, with an association tested between parental genotype and progeny mean phenotype. This approach has recently been used to detect genotype-phenotype associations in natural populations of *Drosophila* (KENNINGTON *et al.* 2007; RAKO *et al.* 2007; WEEKS *et al.* 2002). Indirect methods have also been suggested and applied before in the case of linkage analysis, i.e. the association between phenotypes and identity-by-descent (IBD) within pedigrees, in particular in livestock (VAN DER BEEK *et al.* 1995; WELLER *et al.* 1990) and experimental populations (HILL 1998). The reasons for indirect approaches in linkage studies are the same as for association studies. Firstly, it may be impossible to measure the phenotype or

genotype in certain individuals, for example for sex-limited traits or when it is necessary to sacrifice individuals to be genotyped before their phenotype can be measured. Secondly, due to the increased sample size per genotype, measuring phenotypes on relatives can increase the precision of estimating QTL effects, thereby increasing the power of detection. However, the extra precision may come at a cost due to incomplete linkage between marker and QTL (VAN DER BEEK *et al.* 1995). Moreover, linkage analyses require very large sample sizes, lack power to detect QTL that explain only a few percent of population variance, have poor map resolution (VISSCHER and GODDARD 2004).

While the practical utility of indirect approaches are obvious, within the context of association studies, what remains less clear is the cost one pays in terms of statistical power by indirectly associating parental genotypes with offspring phenotypes. Past investigations of the power of indirect methods have focussed only on the case of linkage rather than association (VAN DER BEEK *et al.* 1995; WELLER *et al.* 1990). Here, we derive analytical expressions for the power to detect associations between candidate polymorphisms and quantitative traits, in four breeding designs that involve genotyping different configurations of parents and phenotyping their offspring, comparing the performance of each with the direct approach of genotyping and phenotyping the same set of individuals. We then derive expressions for the optimal experimental designs given the relative costs of genotyping and phenotyping. Our results demonstrate that if phenotyping is inexpensive relative to genotyping, these designs can provide levels of power equal to, and in some cases better than, the direct approach indicating that indirect association may be a potentially powerful tool for detecting QTNs in outbred populations.

METHODS

We assume random mating, no segregation distortion and a general model of gene action that may or may not include dominance. For the direct approach, we consider a number of unrelated individuals with a genotype and phenotype. For the indirect approach, we consider a balanced design of k dams per sire and m progeny per dam. We assume that the phenotypic standard deviation is unity and that a proportion, q^2, of the phenotypic variance is due to additive genetic effects at the QTL. This quantity is sometimes called the (additive) QTL heritability (e.g. ALMASY and BLANGERO 1998). For a bi-allelic QTL in Hardy-Weinberg equilibrium, allele frequency of p and additive effect of a and dominance deviation of d, $q^2 = 2p(1\text{-}p)[a + d(2p - 1)]^2$ (FALCONER and MACKAY 1996). In practice, the specific values of a, d and p do not matter but their combination does, because power of detection using an additive model for analysis depends on q^2 (LYNCH and WALSH 1998; NEIMANN-SORENSEN and ROBERTSON 1961) . We note that although dominance can be present in this model (i.e. $d \neq 0$), for reasons outlined below (see "Modelling dominance effects" section), we only model the power to detect the additive genetic effect at the QTL. For the direct approach, the model for a phenotype is,

$$y = \mu + g + e, \qquad (1)$$

with g the QTL and e the residual and $\text{var}(y) = \text{var}(g) + \text{var}(e) = q^2 + (1 - q^2)$. For the indirect approach, the model is,

$$y = \mu + s + m + g + e, \qquad (2)$$

with s and m the sire and dam effects and $\text{var}(y) = s^2 + m^2 + q^2 + (1 - s^2 - m^2 - q^2)$ (FALCONER and MACKAY 1996). The (intraclass) correlation between full-sibs is $t_{FS} = s^2 + m^2 + \frac{1}{2}q^2$ and that between half-sibs $t_{HS} = s^2 + \frac{1}{4}q^2$ (via extension from Eqs 18.33a,b LYNCH and WALSH 1998 p

573). If all resemblance between relatives is due to additive genetic factors, then $t_{FS} = \frac{1}{2}h^2$ and $t_{HS} = \frac{1}{4}h^2$, where h^2 is the narrow sense heritability that includes the effect of the QTL.

Power calculations

We test associations by regressing the individual (direct approach) or family mean phenotype (indirect approach) phenotype on the individual or expected genotype indicator, which is coded by an indicator variable (x) as the number of 'A' alleles for a bi-allelic locus with alleles 'A' and 'B'. For individual genotypes (e.g. sires), x can have the values of 0, 1 or 2. The variance of x in the population is $2p(1-p)$, with p the frequency of allele 'A'. For the expected mean genotype in the progeny, x can have values 0, ½, 1, 1.5 and 2 for full-sib families, reflecting the different types of parental matings possible. For example, E(x) in the progeny is 2 if both parents have genotype AA.

We take an analytical approach to the analysis of power, a key feature of which is that our derivations focus on the non-centrality parameter (NCP) of the test of association rather than statistical power $(1 - \beta)$ *per se*. The NCP of a particular design can be thought of as the amount of variation attributable to the model treatment effects (LYNCH and WALSH 1998), which in our case is the additive effect of the polymorphism. There are several advantages to the NCP-based approach. First, unlike statistical power, the NCP does not depend upon any arbitrary choice of Type 1 (α) error threshold, which is often contingent upon the type of study being conducted. Second, NCPs scale in a linear fashion with sample size, whereas power does not, making it far simpler to recalculate power for any variation in sample size without having to recalculate the NCP itself. The calculation of power itself remains straightforward under this approach. Once an NCP has been calculated, it can be used with the desired critical p value to calculate statistical power.

The test statistic for association is the square of a simple t-test (i.e., an F-test). We assume that the sample size is large enough ($N \geq 60$, SEVERO and ZELEN 1960) so that the test statistic is approximately distributed as a central χ^2 with 1 degree of freedom under the null hypothesis of no association. Under the alternative hypothesis, the test statistic is distributed as a non-central χ^2 (SEARLE 1971) with a non-centrality-parameter (NCP) of λ. If α and β are the type-I and type-II error rates, then the power ($= 1 - \beta$) to detect an association is,

$$1-\beta = \Pr(Z < z_{(\alpha/2)} - \sqrt{\lambda}) + \Pr(Z > z_{(1-\alpha/2)} - \sqrt{\lambda}) \approx \Pr(Z > z_{(1-\alpha/2)} - \sqrt{\lambda}) \qquad (3)$$

with Z a standard normal variate and $z_{(1-\alpha/2)}$ the threshold of a normal distribution corresponding to a type-I error rate of α ($\Pr(Z < z_{(\alpha)}) = \alpha$ and $\Pr(Z > z_{(1-\alpha)}) = \alpha$) (e.g. LYNCH and WALSH 1998, p. 870) The relationship between type-I error rate, power and the NCP is,

$$\lambda = (z_{(1-\alpha/2)} + z_{(1-\beta)})^2 \qquad (4)$$

General expression for the NCP: When linear regression is used to estimate the effect of the marker on the quantitative trait, $y = \mu + bx + e$, the general form of the NCP is, $\lambda = b^2 / \mathrm{var}(\hat{b})$ (LYNCH and WALSH 1998, p881). From standard regression theory, $\mathrm{var}(\hat{b}) = \mathrm{var}(e) / \Sigma(x - \bar{x})^2$ (KENDALL and STUART 1977). Treating the x as random, $\mathrm{var}(\hat{b}) = \mathrm{var}(e) / [(N-1)\mathrm{var}(x)] \approx \mathrm{var}(e) / [N\mathrm{var}(x)]$ when N is large (VISSCHER and DUFFY 2006; VISSCHER and HOPPER 2001). Hence, $\lambda = Nb^2\mathrm{var}(x) / \mathrm{var}(e) = Nb^2\mathrm{var}(x) / [\mathrm{var}(y) - b^2\mathrm{var}(x)]$, since $b^2\mathrm{var}(x)$ is the variance removed by the regression. In our regression on the genotype indicator variable, this quantity is approximately cq^2, with c the proportion of the QTL variance detected for a given design. A general expression of the NCP across our designs is then:

$$\lambda = Nb^2var(x) / var(e) = Nc q^2 / var(e) = Nc q^2 / [var(y) - c q^2] \quad (5)$$

For example, for the direct approach, var(y) = 1 and c = 1. For the indirect approach, both c and var(y) < 1 and their ratio determines the efficiency of the design. If cq^2 is small (i.e. assuming that $0 < q^2 < \sim 0.05$) relative to var(y) then $\lambda \approx Nc q^2 / var(y)$.

Direct method

Individual genotypes and individual phenotypes: The model for detecting an association is simply $y = \mu + x + e$, with x the indicator variable for the genotype. If there are 2N individuals phenotyped and genotyped, then

$$\lambda = 2Nq^2/(1-q^2) \approx 2Nq^2. \quad (6)$$

This is a standard expression for the NCP of association between a SNP and quantitative trait (LYNCH and WALSH 1998). The NCP per genotyped individual is approximately q^2. This is also the NCP per phenotyped individual.

Indirect designs

Full-sib families (FS), sire and dam genotyped and full-sib progeny phenotyped: The model for the family mean (Y) is $Y = \mu + b*E(x_o) + e_n$, with $E(x_o)$ the expected genotype indicator variable in the progeny. The variance of the family mean is $var(Y) = [(1-t_{FS})/m + t_{FS}]$, with t_{FS} the intraclass correlation of full-sibs (Appendix 1, A1.2). $E(x_o)$ is simply the average of the parents,

$$E(x_o) = \tfrac{1}{2}(x_{sire} + x_{dam}), \text{ and } var[E(x_o)] = \tfrac{1}{2}var(x) \quad (7)$$

The regression of Y on x_0 is b = a, and therefore c, the proportion of QTL variance detected is ½ and var(y) is $[(1-t_{FS})/m + t_{FS}]$. The NCP for the test of association is, per full-sib family,

$$\lambda_{FS} = \tfrac{1}{2}q^2 / [(1-t_{FS})/m + t_{FS} - \tfrac{1}{2}q^2] \approx \tfrac{1}{2}q^2 / [(1-t_{FS})/m + t_{FS}] \qquad (8)$$

The NCP per genotyped individual is approximately, $\tfrac{1}{4}q^2/[(1-t_{FS})/m + t_{FS}]$. The NCP per phenotyped individual is $\approx \tfrac{1}{2}q^2 / [1 + t_{FS}(m-1)]$. For the limiting case of m=1, the NCP per genotyped individual is ¼ that of the direct approach and the NCP per phenotyped individual is ½ that of the direct approach because the ratio of genotypes to phenotypes is 2.

Full and half-sib families (HS and FSHS1 designs) - sires genotyped, full and half-sib offspring phenotyped: If a sire is mated to k dams and each dam has m progeny (i.e., $n = km$) then the variance of the progeny average phenotype is (Appendix 1, A1.1):

$$\text{var}(Y) = (1-t_{FS})/(km) + (t_{FS}-t_{HS})/k + t_{HS} \qquad (9)$$

Only ¼ of the QTL variance is detected by using the expected genotype in the progeny, so c = ¼. The non-centrality parameter per sire family (and per genotyped individual) is,

$$\lambda_{sire} = \tfrac{1}{4}q^2 / [(1-t_{FS})/(km) + (t_{FS}-t_{HS})/k + t_{HS} - \tfrac{1}{4}q^2] \approx \tfrac{1}{4}q^2 / [(1-t_{FS})/(km) + (t_{FS}-t_{HS})/k + t_{HS}] \quad (10)$$

and the NCP of an experiment with N sires is $N\lambda_{sire}$. A special case of Equation 10 is the classical half-sib design (hereafter referred to as the half-sib design, HS) in which sires are mated to k dams with a single progeny per dam (m=1 and k=n). The NCP per genotyped sire for this design is,

$$\lambda_{HS} = \tfrac{1}{4}q^2/[(1-t_{HS})/n + t_{HS} - \tfrac{1}{4}q^2] \approx \tfrac{1}{4}q^2/[(1-t_{HS})/n + t_{HS}] \quad (11)$$

The expression for the NCP in the full-sib and half-sib case is very similar, the difference being the value of c (= ½ for the full-sib design and ¼ for the half-sib design) and the variance of the family mean.

Full and half-sib families (FSHS2 design) - sires and dams genotyped, full and half-sib offspring phenotyped: When dams are nested within sires and all parents are genotyped, then the contribution of dams and sires to the NCP can be treated separately, taking account of the data structure. For a sire family with k dams, the contribution from each of the dams is (from Equation 8 but with c = ¼),

$$\lambda_{dams} = \tfrac{1}{4}q^2 / [(1-t_{FS})/m + t_{FS} - \tfrac{1}{2}q^2] \approx \tfrac{1}{4}q^2 / [(1-t_{FS})/m + t_{FS}] \quad (12)$$

The contribution from the half-sibs is (Equation 10),

$$\lambda_{sire} = \tfrac{1}{4}q^2 / [(1-t_{FS})/(km) + (t_{FS}-t_{HS})/k + t_{HS} - \tfrac{1}{4}q^2] \approx \tfrac{1}{4}q^2 / [(1-t_{FS})/(km) + (t_{FS}-t_{HS})/k + t_{HS}] \quad (13)$$

And the total NCP per sire family is,

$$\lambda_{FSHS} = \lambda_{sire} + k\lambda_{dams} \approx \tfrac{1}{4}q^2 [\{(1-t_{FS})/(km) + (t_{FS}-t_{HS})/k + t_{HS}\}^{-1} + k\{(1-t_{FS})/m + t_{FS}\}^{-1}] \quad (14)$$

When k=1 the expression is equivalent to that for the full-sib design (Equation 8). When m=1 (and therefore k=n), i.e. a half-sib design but all dams genotyped, the NCP is $\tfrac{1}{4}q^2/[(1-t_{HS})/n + t_{HS}] + \tfrac{1}{4}q^2 n$, the two terms corresponding to the contribution of the half-sib family (Equation 11) and n parent-offspring pairs, respectively.

If all dams are genotyped then there are k+1 genotypes per sire family and km phenotypes per family. Hence, the NCP per genotype and per phenotype are $(\lambda_{sire} + k\lambda_{dams})/(k+1)$ and $[\lambda_{sire}/(km) + \lambda_{dams}/m]$, respectively.

Validation of equations

We performed a simulation study to verify our analytical results for the indirect design (Equations 8, 10, 11 and 14). Phenotypes for progeny were simulated as $y = \mu + s + m + g + e$, for a given QTL heritability and a model in which all family resemblance was due to additive genetic factors. Input parameters were the numbers of sires, dams and progeny, the proportion of variance explained by the QTL, the sire and dam intra-class correlations, the degree of dominance and the allele frequency at the QTL. Sire and dam genotypes were simulated by sampling alleles from the population using the binomial distribution (assuming HWE). Genotypes of progeny were simulated by sampling from the parental gametes, assuming Mendelian inheritance. Sire, dam and residual random effects were simulated from a normal distribution with the appropriate standard deviation. Phenotypic observations on progeny were calculated by summation of the individual-specific terms (i.e. sire, dam, QTL and residual). Data were analysed using linear regression of the progeny means on SNP genotype. 10,000 replicates were run for many combinations of parameters, and the average test statistic was recorded. The mean test statistic was found to be extremely close to our predictions (results not shown) and therefore simulations were not pursued further.

Modelling dominance effects

Our coding for the indicator variable x reflects the expected value of the transmitted allele from the parent and therefore models the additive effects of alleles. An alternative parameterisation

is to code the expected dosage of the *a* and *d* effects in the progeny, given the observed genotype in the parents and the allele frequency in the population. For example, for parents with genotypes BB, AB and AA, the expected mean value in the progeny due to the QTL are $\{-(1-p)a + pd\}$, $\{(p-\frac{1}{2})a + \frac{1}{2}d\}$ and $\{pa + (1-p)d\}$, respectively (FALCONER and MACKAY 1996). However, the coefficients for *a* and *d* are linear combinations of each other, so additive and dominance effects cannot be separated when genotypes are observed on a single parent and progeny are phenotyped, for example in the case of half sib designs. Therefore, only an allele substitution effect can be estimated, as coded, for example, by 0, ½ and 1 for these genotypes. This makes sense because twice the expected value of the mean progeny phenotype is, when deviated from the population mean, the additive breeding value of the parent.

This means that there are qualitative differences in the ability to model dominance effects between the different indirect designs. As it is only possible to fit a model parameter for dominance deviations when all parents are genotyped, we confine our results to the 1 d.f. model which only detects additive effects. A possible consequence of this is that when designing a study, one may make design decisions on the assumption that all genetic effects are additive. Although the loss in additive genetic variance as a consequence of dominance is simple to estimate for a given genetic variance due to a QTL, it is important to verify that the indirect designs perform in line with these theoretical expectations when dominance is non zero. We considered this issue using simulations and results are presented in Appendix 2.

Efficiency of designs

The direct and indirect designs can differ greatly in the number of genotyped and phenotyped individuals. To compare the relative efficiency of the different indirect designs, we

express the power (NCP) as a proportion of the cost of the experiment. If the cost per genotype is 1.0 unit and the relative cost of phenotyping to genotyping is C_p, then the total cost is,

$$C = N_g + C_p N_p, \quad (15)$$

with N_g and N_p the number of individuals genotyped and phenotyped, respectively. For a given design we can express the relative power (RNCP) as the NCP per 'QTL heritability' and per $. That is,

$$RNCP = NCP / (q^2 C) = NCP / (q^2 (N_g + C_p N_p)) \quad (16)$$

Taking the approximate expressions for the NCP (i.e. assuming that $0 < q^2 < \sim 0.05$), then the q^2 drops out of the equation. The resulting RNCP are then simple expressions of all these parameters.

For the direct approach,

$$RNCP = N q^2 / (q^2 (N + C_p N)) = 1 / (1 + C_p). \quad (17)$$

For the classic HS design,

$$RNCP = ¼ / [(1 + C_p n)((1-t_{HS})/n + t_{HS})] \quad (18)$$

For a given value of the heritability (and therefore t_{HS} if all family resemblance is due to additive genetic factors), the optimum value of n is:

$$n_{opt} = \sqrt{[(1 - t_{HS}) / (C_p t_{HS})]} \quad (19)$$

Similarly for the FS design,

$$\text{RNCP} = ½ / [(1 + C_p m)((1-t_{FS})/m + t_{FS}], \quad (20)$$

and the optimum value of m is

$$m_{opt} = \sqrt{[(1 - t_{FS}) / (C_p t_{FS})]} \quad (21)$$

For the nested design with only sires genotyped (FSHS1),

$$\text{RNCP} = ¼ / [(1 + C_p km)((1-t_{FS})/km + (t_{FS}-t_{HS})/k + t_{HS})] \quad (22)$$

If m is fixed then the optimal value for k is $\sqrt{[(1 - t_{HS}) / (mC_p t_{HS})]}$.

RESULTS

Statistical Power

Common genotyping effort: It can be seen from the per genotype and per phenotype approximations in Table 1 that the power of the indirect approaches relative to the direct approach are simple functions of two factors: first, the degree to which phenotypes within families are correlated (i.e. the intraclass correlations) which in the absence of non-genetic and non-additive genetic causes of family resemblance, is the narrow-sense heritability of the trait, and second, the numbers of progeny phenotyped per family when using an indirect design. There is no difference in the relative power of direct and indirect approaches as the QTN effect size varies. Thus, we explore performance of the indirect approaches due to variation in the narrow-sense heritability and progeny phenotyping effort, for an arbitrary QTN effect size whilst holding genotyping effort constant.

The power for three indirect designs (FS, HS and FSHS1) and the corresponding direct design are plotted for different progeny numbers and narrow-sense heritabilities in Figure 1. In these examples, genotyping was fixed at 100 individuals and the phenotyping effort was free to vary for the indirect designs depending on the specific values of *m*, *n*, and *k* used. For all three designs there are combinations of progeny number and heritability for which the power exceeds that of the corresponding direct design, with the indirect approach performing better as heritability declines (Figure 1. A-C). We investigated the FSHS2 design for a wide range of parameters but found that genotyping both sires and dams is not efficient because the additional information on association from the progeny of the dams is not compensated for by the increase in the number of genotypes. Thus we do not consider this design further.

Due to its effect on the phenotypic correlation among full and half sibs, heritability has a strong impact on the power of the indirect designs. While there is little difference in performance

of the different indirect designs when heritability is low, all methods suffer reduced power when heritability is high. This is because the variance in progeny mean is relatively large, a factor that is particularly important for the performance of the full-sib design. Equating (6) and (8) for a particular value of t_{FS} gives the number of progeny that need to be phenotyped to give equal power of the direct and indirect design. That number is, $m = 4(1-t_{FS})/(1-4t_{FS})$. Thus, for small values of t_{FS}, 4 progeny per full sib family need to be phenotyped to have equal power. Relative to the direct approach in which 2N individuals are phenotyped this is an increase of a factor of 2. For large values of t_{FS} ($t_{FS} > 0.25$, e.g. $h^2 > ½$) there is no number of phenotyped progeny that can compensate for the loss of information. Thus, for highly heritable traits there will always be a higher genotyping effort for the full-sib design compared with the direct method. This limitation is not as severe however, for the half-sib case due to a lower expected correlation between half sibs over a much wider range of heritabilities.

The FSHS1 design requires a choice of both the number of dams, *k*, and the number of offspring per dam, *m*. For a given value of *km* it appears that there are marginal gains in power by increasing the increasing the number of dams rather than increasing the number of offspring per dam. For example, Figure 1C illustrates this point with $km = 10$. With sires mated to five dams and two full-sib progeny family phenotyped per dam, power was always higher compared with the alternative situation in which a sire is mated to two dams and five progeny are phenotyped.

Common phenotyping effort: The indirect designs can result in a large variation in the number of phenotyped progeny, thus we also analysed the performance of the designs on a per phenotype basis. Under no circumstances is power better than the direct approach for a common phenotyping effort (Figure 2). The per phenotype NCP of each indirect design tends to plateau well below the corresponding NCP for the direct design. On a per phenotype basis, the full-sib design always performs better than the half sib designs regardless of whether a nested design is used.

Relative efficiency and economy

As cost is often the major limiting factor in an association study and the indirect designs can lead to very different genotyping and phenotyping efforts compared with the direct method, we also modelled how the cost of phenotyping relative to genotyping affects the efficiency of different designs. Figure 3 illustrates the relative noncentrality parameters (RNCPs) for the two simplest cases, the full-sib and classic half-sib designs with each optimized for *m* and *n* respectively using (21) and (19). The full-sib design provides a much larger contribution to power per unit expenditure when $C_p = 0.1$, which represents a tenfold lower cost of phenotyping relative to genotyping. For all other cases considered, equal costs and when genotyping is ten times cheaper than phenotyping, the direct design is favoured on economical grounds.

DISCUSSION

We have provided analytical solutions for the power to detect genotype-quantitative trait associations using an indirect approach, in which parents are genotyped and different configurations of offspring are phenotyped in an outbred population. Inspired by the possibility that indirect approaches may represent a flexible and cost effective strategy for traits which are difficult to measure on single individuals and the opportunity to score a large number of phenotypes per genotype, our goal was to determine whether an indirect approach could provide power equal to or better than the direct approach of genotyping and phenotyping the same individuals. Upon finding that there are regions of parameter space for which the indirect approaches perform well, we explored how the cost of phenotyping relative to genotyping affects the efficiency of the designs.

Power considerations

It is immediately apparent from the per genotype non-centrality parameter expressions and their approximations (Table 1) that the QTN heritability, q^2, does not affect the relative performance of indirect methods as compared with direct methods. Thus our discussion concerning the performance of different indirect designs has generality across the entire range of effect sizes. However, the power of the indirect approaches is heavily influenced by the narrow-sense heritability of the trait, with power decreasing in all designs as heritability increases. When heritability is high, the contribution of environmental variance to trait values is low. Thus the benefit achieved via the indirect design, due to a decrease in the contribution of environmental variance to family means, is also diminished. By contrast, heritability has the opposite effect on the power of indirect linkage analysis, rather than association as we have considered here, in full and half-sib families (VAN DER BEEK *et al.* 1995). The likely source of this discrepancy is that, for linkage analysis, the power depends on the proportion of within-family variance explained by the QTL, so for a fixed QTL effect size, a larger heritability implies a larger proportion of within-family variance explained and therefore more power (VISSCHER and HOPPER 2001).

Although any "best" design is likely to depend on factors such as the reproductive biology of the organism studied, heritability and budget, four guiding principles emerge from our analyses. First, half-sib designs are more powerful on a per genotype basis than full-sib designs. Second, when using the FSHS1 design, it is always better to increase the number of dams rather than the number of offspring per dam. Because there are more independent half-sib families in this design, the power contribution to the test statistic comes from half rather than full sibs (see Equation 14). Third, on a per phenotype basis no indirect approach can match the power of the direct approach. Finally, there may be little to be gained from genotyping both dams and sires in a full-sib half sib design. The poor performance of the FSHS2 design, is partially a consequence of the fact that,

within a group of half sibs, the genotyping effort allocated the dams delivers little in terms of power gains compared with the alternative of genotyping more independent sires from the population.

Ane key difference between our approach and other considerations of the power of association studies is that we have assumed that the causal SNP is genotyped. For example, LONG and LANGLEY (1999) investigated the power of direct association studies to detect QTLs in outbred populations as a function of sample size, the QTL heritability and recombination rate between the causal and genotyped variants that are in linkage disequilibrium with the causal variant. Our power calculations can be easily adjusted for imperfect linkage disequilibrium by multiplying the non-centrality parameter of the test statistic by r^2, the squared correlation between alleles at the causal and genotyped variant (HILL and ROBERTSON 1968). In other words, if there is not perfect linkage disequilibrium between the causal and genotyped variant then the experimental sample size needs to increase by a factor of $1/r^2$ to achieve the same power.

Considerations due to dominance

In our analyses we have fitted an additive model and therefore our power calculations are relevant for the proportion of additive genetic variance due to the QTN. The predictable degree to which parental genotype reflects progeny mean phenotype is altered by dominance. Dominance has been shown to influence the power of tests of association and linkage, and in general terms can either increase or decrease power depending upon the magnitude of dominance variance relative to the cost of fitting an extra model parameter to account for it (SHAM *et al.* 2000). Although it has been argued on theoretical grounds and from empirical observations that dominance effects may be generally relatively weak (HILL *et al.* 2008), in the indirect case, as we have considered here, we model only the ability to detect additive effects, so it is likely that power will be compromised by dominance, an issue we consider below.

We first considered the question of whether the decrease in power as a consequence of dominance is any worse than when using the equivalent direct method (i.e. a 1 d.f. model). From theory, power loss due to dominance can be substantial for both designs when dominance is strong ($a = d$) and the common allele is dominant (Appendix 1). Simulations suggest however that the direct and indirect approaches do not differ systematically their sensitivity to this effect (Appendix 2). However, one has the option of fitting a model with a specific coding variable for *d* (i.e., a 2 d.f. model) when using the direct approach, an option only available for indirect designs in which both parents are genotyped. For the indirect FS design, dominance effects can be separated from additive effects because the progeny genotype can be predicted, in some cases without error. For example, AA x AA matings always give AA progeny and AA x BB matings always give AB progeny. Progeny from AB x AB matings have the same expectation for *a* as progeny from AA x BB matings but differ in their expectation for *d* (½ and 1, respectively). In practice however the power to detect dominance will be lower than the direct approach for two reasons. First, the variance of the indicator variable for *d* is lower for the indirect case in a manner that depends upon allele frequency. For the direct approach this is $\mathrm{Var}(x_{\mathrm{d\ direct}}) = H(1-H)$ whereas for the indirect full sib case $\mathrm{Var}(x_{\mathrm{d\ indirect}}) = H(1-H) - \frac{1}{2}H(1 - \frac{1}{2}H)$ where H, the heterozygosity (= $2p(1-p)$). Second, when allele frequencies are non symmetrical, the coding variables for *a* and *d* become correlated with each other. This has the effect of significantly reducing the amount of dominance that is recoverable after fitting an additive term. Thus, although it is theoretically possible to recover some information on dominance using the full-sib design it is likely to be small unless moderate frequency SNPs are being tested. In summary, it appears that the main impact on power between the direct and indirect methods due to dominance will be a function of the limited ability to fit a 2 d.f. model in some cases rather than any systematic difference between the approaches in their ability to detect additive effects.

Relative efficiency and economy

As power relative to the direct approach is a function of the phenotyping effort, the indirect designs can require considerable phenotyping effort to gain equal or better power than the direct approach. Thus, by developing expressions for the relative noncentrality parameter per dollar, and optimal sizes for the three simplest indirect designs, we were able to investigate how variability in the relative costs of genotyping and phenotyping may influence design choice. Only when phenotyping is cheaper than genotyping (i.e. $C_p < 1$) is it possible for the indirect approach to be more economical. In fact, often phenotying may have to be considerably cheaper than genotyping. For the limit of $C_p \rightarrow 0$, i.e. very cheap phenotyping relative to genotyping, the comparison between designs is for the same number of genotypes (see Table 1). For the case considered in Figure 3, where heritability is 0.2, RNCP(HS) $\rightarrow 1/h^2 = 5$ and RNCP(FS) $\rightarrow 1/2h^2 = 2.5$, both for a large (strictly infinite) number of progeny.

Conclusion

We have developed exact expressions and approximations for the statistical power of tests of association, in which parental genotypes are associated with progeny mean phenotypes and compared their performance with a direct approach in which associations are tested between genotypes and phenotypes from the same individuals. For situations in which both direct and indirect approaches are feasible, our results suggest that the indirect approaches are more powerful when traits have low heritability but are only more economical to implement when genotyping costs far outweigh phenotyping costs. For studies in which the specific trait or study organism precludes a direct association study, the indirect approach nonetheless remains a viable option. We have implemented power calculations for both the direct method and the indirect designs considered here

on a web-based application, PIAS (Power calculator for Indirect Association Studies) that can be accessed by the wider community (http://www.chenowethlab.org/pias/index.html).

ACKNOWLEDGEMENTS

We thank Bill Hill, Ary Hoffmann and Emma Hine for comments on the manuscript. S.F.C. is supported by the Australian Research Council, P. M. V. is supported by the Australian National Health and Medical Research Council.

TABLES

Table 1: Approximate NCP/q^2 per genotype and per phenotype for different indirect breeding designs.

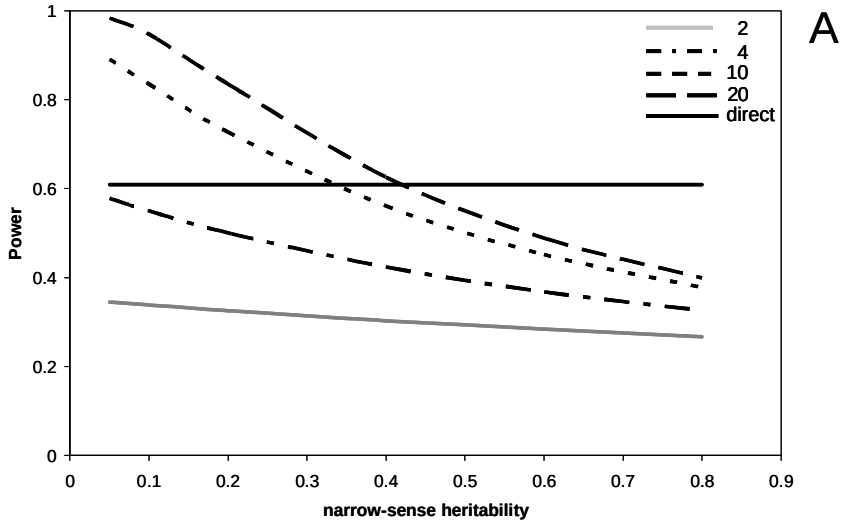
Design	NCP per genotype	NCP per phenotype
Direct	1	1
Half-sib families (HS)	$¼ / [(1-t_{HS})/n + t_{HS}]$	$¼ / [1 + (n-1)t_{HS}]$
Full-sib families (FS)	$¼ / [(1-t_{FS})/m + t_{FS}]$	$½ / [1 + (m-1)t_{FS}]$
Dams nested within sires; dams not genotyped (FSHS1)	$¼ / [(1-t_{FS})/(km) + (t_{FS}-t_{HS})/k + t_{HS}]$	$¼ / [1 + (m-1)t_{FS} + mt_{HS}(k-1)]$
Dams nested within sires; dams genotyped (FSHS2)	$¼ [1/(k+1)] [k\{(1-t_{FS})/m + t_{FS}\}^{-1} + \{(1-t_{FS})/(km) + (t_{FS}-t_{HS})/k + t_{HS}\}^{-1}]$	$¼ [1/km] [k\{(1-t_{FS})/m + t_{FS}\}^{-1} + \{(1-t_{FS})/(km) + (t_{FS}-t_{HS})/k + t_{HS}\}^{-1}]$

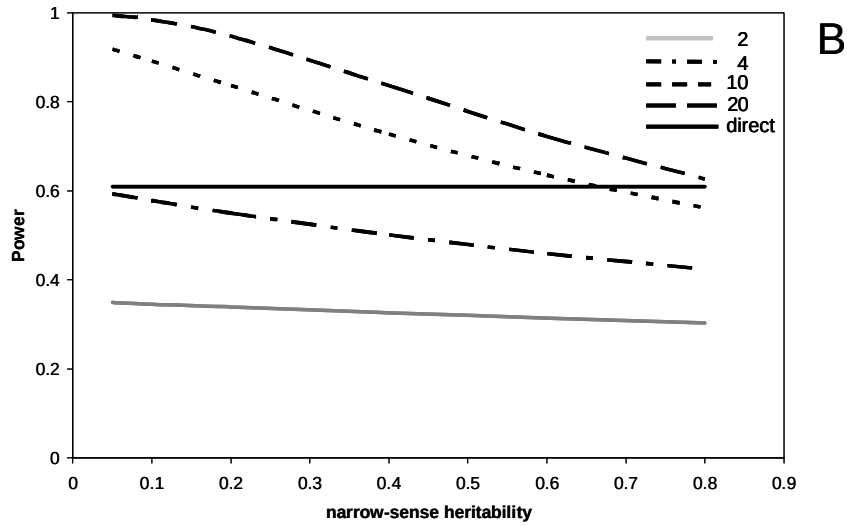
FIGURES

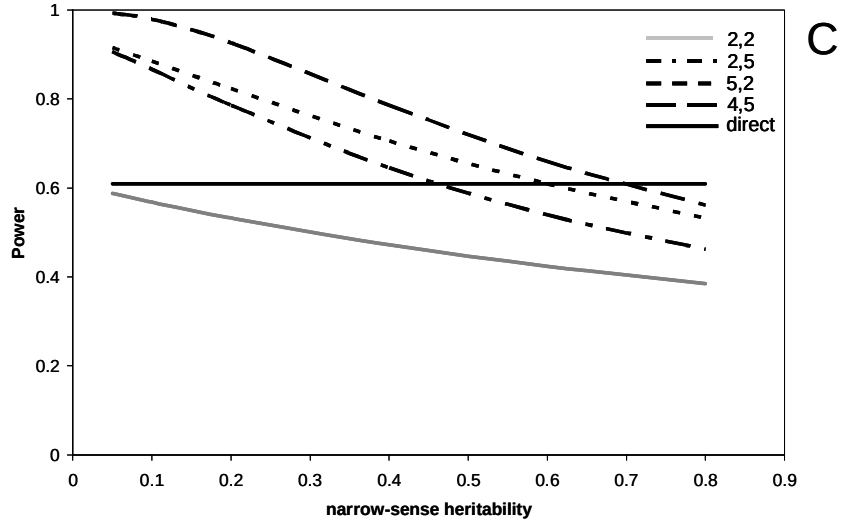
Figure 1. Effect of narrow-sense heritability and progeny number on the power (1 - β) of the full-sib (FS) **A**, classic half-sib (HS) **B**, and nested half sib design in which only sires are genotyped (FSHS1), **C**. In all cases, a QTN is assumed to explain 5% of the phenotypic variance (i.e. q^2 = 0.05), the genotyping effort is set at 100 individuals and $\alpha = 0.05$. In all plots the solid horizontal line is the power for the direct design with identical genotyping effort and QTN effect size. For the full and half sib designs (A and B), *m* and *n* were fixed at the following values: 2, 4, 10, 20. For the FSHS1 design different combinations of *k* and *m* were used (2,2; 2,5; 5,2 and 4,5).

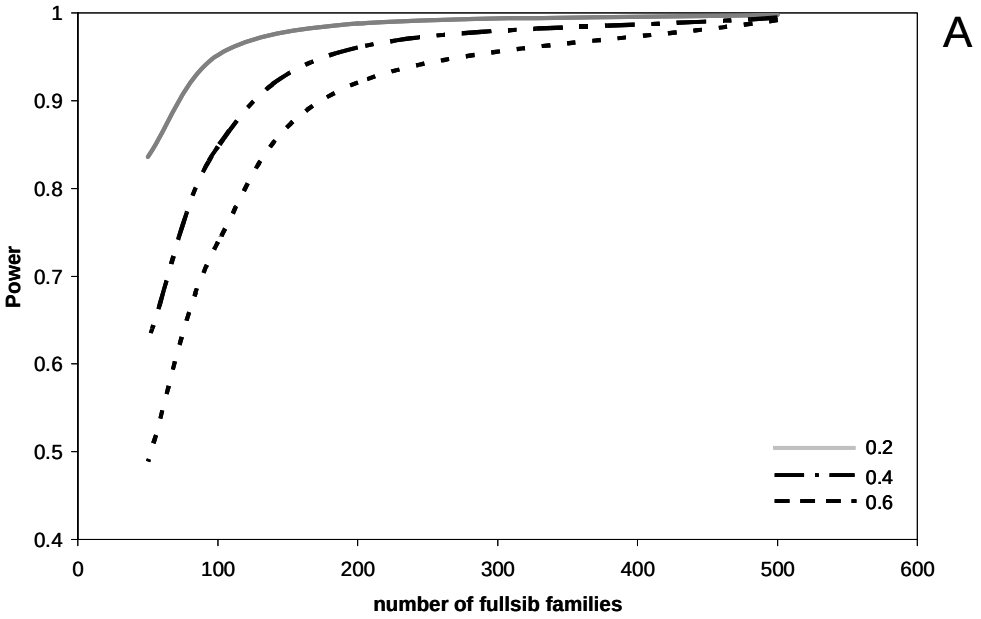
Figure 2. Effect of the number of full- and half-sib families on the power (1 - β) of the full-sib (FS), **A** and classic half-sib (HS), **B** designs for a common phenotyping effort of 1000 individuals where $\alpha = 0.05$. As family number varies, the numbers of *m* and *n* have been scaled to reflect the common phenotyping effort, thus an analysis involving 100 families implies a value of *m* of 10 and so on. The QTN explains 5% of the phenotypic variance (i.e. $q^2 = 0.05$). Corresponding power for the direct approach is 1.0 in both cases and profiles for three narrow–sense heritability values are shown.

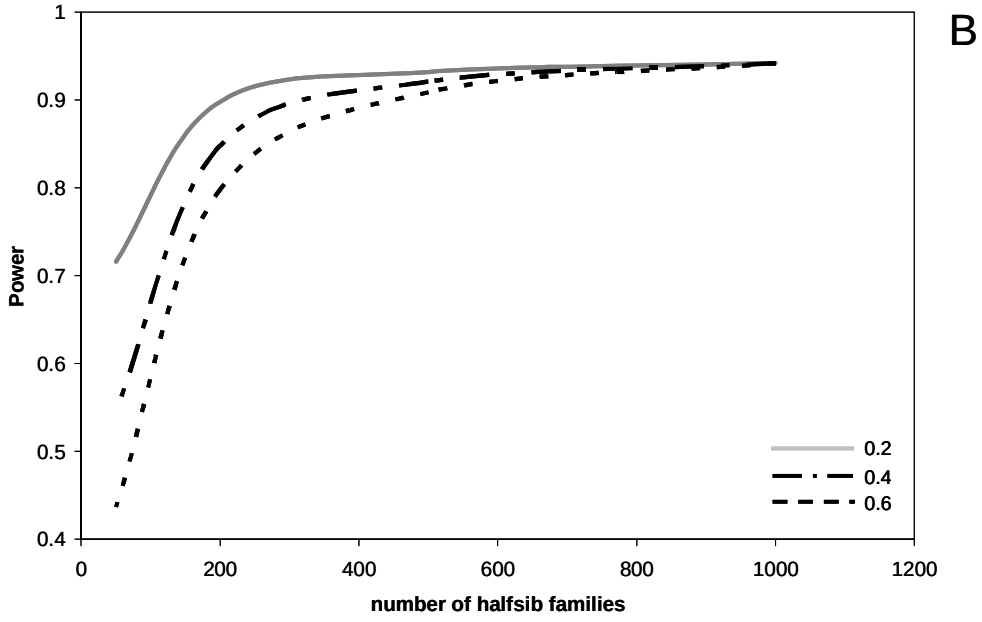
Figure 3. Relative efficiency, expressed as the relative non-centrality parameter (RNCP), of the direct (Equation 17), full-sib (FS: Equation 20) and classic half-sib designs (HS: Equation 18) when the cost of phenotyping varies relative to the cost of genotyping. For each indirect design, we first estimated the optimal values of *m* and *n* for the relative cost of phenotyping versus genotyping (C_P) using Equations 21 and 19 respectively. For this example, the narrow-sense heritability was set at 0.2 and the QTN explained 5% of the phenotypic variance.

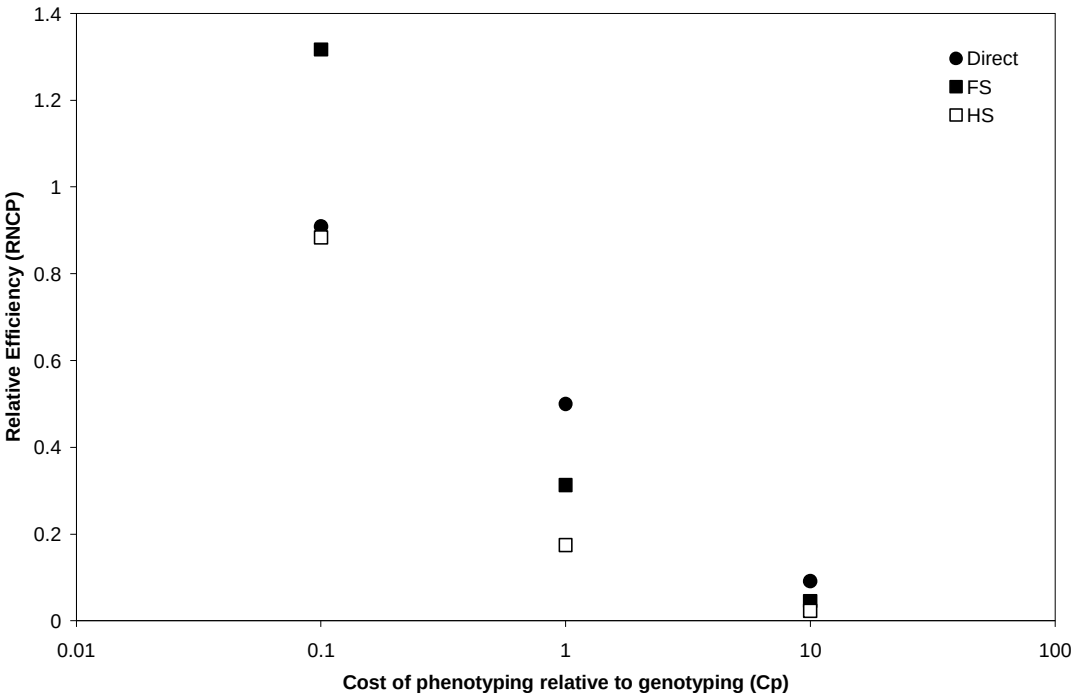

A
2
4
10
20
direct
Power
narrow-sense heritability
0
0.2
0.4
0.6
0.8
1
0.1
0.3
0.5
0.7
0.9


B
2
4
10
20
direct
Power
narrow-sense heritability
0
0.2
0.4
0.6
0.8
1
0.1
0.3
0.5
0.7
0.9


C
2,2
2,5
5,2
4,5
direct
Power
narrow-sense heritability
0
0.2
0.4
0.6
0.8
1
0.1
0.3
0.5
0.7
0.9


A
1
0.9
0.8
0.7
0.6
0.5
0.4
Power
0
100
200
300
400
500
600
number of fullsib families
0.2
0.4
0.6


B
1
0.9
0.8
0.7
0.6
0.5
0.4
Power
0
200
400
600
800
1000
1200
number of halfsib families
0.2
0.4
0.6


1.4
1.2
1
0.8
0.6
0.4
0.2
0
Relative Efficiency (RNCP)
0.01
0.1
1
10
100
Cost of phenotyping relative to genotyping (Cp)
Direct
FS
HS

LITERATURE CITED

ALMASY, L., and J. BLANGERO, 1998 Multipoint quantitative-trait linkage analysis in general pedigrees. American Journal of Human Genetics **62:** 1198-1211.

ARANZANA, M. J., S. KIM, K. Y. ZHAO, E. BAKKER, M. HORTON *et al.*, 2005 Genome-wide association mapping in *Arabidopsis* identifies previously known flowering time and pathogen resistance genes. PloS Genetics **1:** 531-539.

BARTON, N. H., and P. D. KEIGHTLEY, 2002 Understanding quantitative genetic variation. Nature Reviews Genetics **3:** 11-21.

BARTON, N. H., and M. TURELLI, 1989 Evolutionary quantitative genetics - how little do we know. Annual Review of Genetics **23:** 337-370.

CARBONE, M. A., K. W. JORDAN, R. F. LYMAN, S. T. HARBISON, J. LEIPS *et al.*, 2006 Phenotypic variation and natural selection at Catsup, a pleiotropic quantitative trait gene in Drosphila. Current Biology **16:** 912-919.

DWORKIN, I., A. PALSSON, K. BIRDSALL and G. GIBSON, 2003 Evidence that Egfr contributes to cryptic genetic variation for photoreceptor determination in natural populations of *Drosophila melanogaster*. Current Biology **13:** 1888-1893.

DWORKIN, I., A. PALSSON and G. GIBSON, 2005 Replication of an Egfr-wing shape association in a wild-caught cohort of *Drosophila melanogaster*. Genetics **169:** 2115-2125.

FALCONER, D. S., and T. F. C. MACKAY, 1996 *Introduction to Quantitative Genetics*. Longman, Essex, UK.

GRUBER, J. D., A. GENISSEL, S. J. MACDONALD and A. D. LONG, 2007 How repeatable are associations between polymorphisms in achaete-scute and bristle number variation in *Drosophila*? Genetics **175:** 1987-1997.

HANSEN, T. F., 2006 The evolution of genetic architecture. Annual Review of Ecology Evolution and Systematics **37:** 123-157.

HILL, W. G., 1998 Selection with recurrent backcrossing to develop congenic lines for quantitative trait loci analysis. Genetics **148:** 1341-1352.

HILL, W. G., and A. CABALLERO, 1992 Artificial Selection Experiments. Annual Review of Ecology and Systematics **23:** 287-310.

HILL, W. G., M. E. GODDARD and P. M. VISSCHER, 2008 Data and theory point to mainly additive genetic variance for complex traits. PloS Genetics **4**.

HILL, W. G., and A. ROBERTSON, 1968 Linkage disequilbrium in finite populations. Theoretical and Applied Genetics **38:** 183-201.

HOULE, D., 1992 Comparing evolvability and variability of quantitative traits. Genetics **130:** 195-204.

KENDALL, M. G., and A. STUART, 1977 *The advanced theory of statistics*. Macmillan, New York.

KENNINGTON, W. J., A. A. HOFFMANN and L. PARTRIDGE, 2007 Mapping regions within cosmopolitan inversion In(3R)Payne associated with natural variation in body size in *Drosophila melanogaster*. Genetics **177:** 549-556.

LETTRE, G., A. U. JACKSON, C. GIEGER, F. R. SCHUMACHER, S. I. BERNDT *et al.*, 2008 Identification of ten loci associated with height highlights new biological pathways in human growth. Nature Genetics **40:** 584-591.

LIU, P. Y., Y. WANG, H. VIKIS, A. MACIAG, D. L. WANG *et al.*, 2006 Candidate lung tumor susceptibility genes identified through whole-genome association analyses in inbred mice. Nature Genetics **38:** 888-895.

LONG, A. D., and C. H. LANGLEY, 1999 The power of association studies to detect the contribution of candidate genetic loci to variation in complex traits. Genome Research **9:** 720-731.

LONG, A. D., R. F. LYMAN, C. H. LANGLEY and T. F. C. MACKAY, 1998 Two sites in the Delta gene region contribute to naturally occurring variation in bristle number in *Drosophila melanogaster*. Genetics **149:** 999-1017.

LONG, A. D., R. F. LYMAN, A. H. MORGAN, C. H. LANGLEY and T. F. C. MACKAY, 2000 Both naturally occurring insertions of transposable elements and intermediate frequency polymorphisms at the achaete-scute complex are associated with variation in bristle number in *Drosophila melanogaster*. Genetics **154:** 1255-1269.

LYNCH, M., and B. WALSH, 1998 *Genetics and Analysis of Quantitative Traits*. Sinauer, Sunderland, Ma.

MACDONALD, S. J., and A. D. LONG, 2004 A potential regulatory polymorphism upstream of hairy is not associated with bristle number variation in wild-caught *Drosophila*. Genetics **167:** 2127-2131.

MACKAY, T. F. C., 2001 The genetic architecture of quantitative traits. Annual Review of Genetics **35:** 303-339.

NEIMANN-SORENSEN, A., and A. ROBERTSON, 1961 The association between blood groups and several production characters in three Danish cattle breeds. Acta Agriculturae Scandinavica **11:** 163-196.

RAKO, L., M. J. BLACKET, S. W. MCKECHNIE and A. A. HOFFMANN, 2007 Candidate genes and thermal phenotypes: identifying ecologically important genetic variation for thermotolerance in the Australian *Drosophila melanogaster* cline. Molecular Ecology **16:** 2948-2957.

ROBIN, C., R. F. LYMAN, A. D. LONG, C. H. LANGLEY and T. F. C. MACKAY, 2002 hairy: A quantitative trait locus for *Drosophila* sensory bristle number. Genetics **162:** 155-164.

SEARLE, S. R., 1971 *Linear models*. Wiley, New York.

SEVERO, N. C., and M. ZELEN, 1960 Normal approximation to the Chi-square and non-central F-probability functions. Biometrika **47:** 411-416.

SHAM, P. C., S. S. CHERNY, S. PURCELL and J. K. HEWITT, 2000 Power of linkage versus association analysis of quantitative traits, by use of variance-components models, for sibship data (vol 66, pg 1616, 2000). American Journal of Human Genetics **66:** 2020-2020.

VAN BUSKIRK, J., and Y. WILLI, 2006 The change in quantitative genetic variation with inbreeding. Evolution **60:** 2428-2434.

VAN DER BEEK, S., J. A. M. VAN ARENDONK and A. F. GROEN, 1995 Power of 2-generation and 3-generation QTL mapping experiments in an outbred population containing full-sib or half-sib families. Theoretical and Applied Genetics **91:** 1115-1124.

VISSCHER, P. M., and D. L. DUFFY, 2006 The value of relatives with phenotypes but missing genotypes in association studies for quantitative traits. Genetic Epidemiology **30:** 30-36.

VISSCHER, P. M., and M. E. GODDARD, 2004 Prediction of the confidence interval of quantitative trait loci location. Behavior Genetics **34:** 477-482.

VISSCHER, P. M., and J. L. HOPPER, 2001 Power of regression and maximum likelihood methods to map QTL from sib-pair and DZ twin data. Annals of Human Genetics **65:** 583-601.

WEEDON, M. N., H. LANGO, C. M. LINDGREN, C. WALLACE, D. M. EVANS *et al.*, 2008 Genome-wide association analysis identifies 20 loci that influence adult height. Nature Genetics **40:** 575-583.

WEEKS, A. R., S. W. MCKECHNIE and A. A. HOFFMANN, 2002 Dissecting adaptive clinal variation: markers, inversions and size/stress associations in *Drosophila melanogaster* from a central field population. Ecology Letters **5:** 756-763.

WELLER, J. I., Y. KASHI and M. SOLLER, 1990 Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy-cattle. Journal of Dairy Science **73:** 2525-2537.

Appendix 1

DERIVATION OF VARIANCE AMONG FAMILY MEANS

From Falconer and Mackay (1996 p167)

The variance of sire means is given by:

$$\text{Var}(Y) = MS_{SIRE} / km = \sigma^2_w /km + \sigma^2_D /k + \sigma^2_S$$

But assuming $V_P = 1$:

$$\sigma^2_S = \text{cov}_{(HS)} = t_{HS}$$

$$\sigma^2_D = \text{cov}_{(FS)} - \text{cov}_{(HS)} = t_{FS} - t_{HS}$$

$$\sigma^2_w = 1 - \text{cov}_{(FS)} = 1 - t_{FS}$$

Hence,

$$\text{var}(Y) = (1-t_{FS})/(km) + (t_{FS}-t_{HS})/k + t_{HS} \qquad \text{(A1.1)}$$

For the full sib case, k = 1 and there is no contribution from half-sibs thus, $t_{HS} = 0$.

Hence,

$$\text{Var}(Y) = (1-t_{FS})/m + t_{FS} \qquad \text{(A1.2)}$$

Appendix 2

ASSOCIATION WITH DOMINANCE

For a standard quantitative genetic model, with mean values of $-a$, d and a for genotypes AA, AB and BB, using $k = d/a$ as the degree of dominance, p the allele frequency of allele B and H the heterozygosity ($2p(1-p)$):

$$\text{var(A)} = Ha^2[1 + k(1-2p)]^2 \qquad \text{(A2.1)}$$

$$\text{var(D)} = (Hka)^2 \qquad \text{(A2.2)}$$

(FALCONER and MACKAY 1996; LYNCH and WALSH 1998). Let the phenotypic variance be one and the total proportion of variance due to the QTL be Q^2 = var(A) + var(D). As before, the proportion of phenotypic variance due to additive variance at the QTL is q^2 = var(A). For a direct design, the NCP for an additive (1 d.f.) model is approximately $Nq^2/(1-q^2)$ and the NCP for a 2 d.f. model is approximately $NQ^2/(1-Q^2)$ (e.g. SHAM *et al.* 2000). Some examples, keeping the total proportion of variance due to the QTL constant ($Q^2 = 0.05$) are given in Table A1.

Only if there is strong dominance and the common allele is dominant (e.g., p=0.1, k=-1) is there substantial loss in power by fitting an additive model. If Q^2 is small then the ratio of the NCP for fitting a or fitting $a + d$ is $\sim q^2/Q^2$. This ratio can be expressed as,

$$(1 + k(1-2p))^2 / [(1 + k(1-2p))^2 + Hk^2], \qquad \text{(A2.3)}$$

and depends on k and p.

In the indirect model we detect ½q^2 or ¼q^2 when we fit an additive model, thus apart from this reduction in the amount of q^2 detected, there is no theoretical reason for expecting the performance of a 1 d.f. test to perform differently than the direct approach. We tested this prediction using simulations.

We simulated a quantitative trait according to a model in which the trait was influenced by both additive and dominance effects. Phenotypes for progeny were simulated as $y = \mu + s + m + g + e$, for a given QTL heritability and a model in which all family resemblance was due both additive and dominance effects (given by a value of k). Data were analysed using linear regression of the progeny means on SNP genotype. 10,000 replicates were run for each combination of parameters, and the average test statistic was recorded. We considered both the FS and classic HS designs with parameter values similar to Figs 1A and B. That is, a sample of 100 genotypes, total QTL variance of 0.05, trait heritability of 0.4. We considered three allele frequencies (p = 0.1; 0.5 and 0.7) for the cases of no (k =0), partial (k = 0.5) and complete (k=1) dominance. We then compared predicted (theory) test statistics with simulated values for each set of parameter values. For both designs, simulated results were very similar to theory and in no case did theoretical test statistics significantly exceed simulated ones (Tables A2 and A3).

Table A1. Examples of power loss due to dominance when fitting an additive (1 d.f.) model using the direct approach. Note that the power of the 2 d.f. model is unaffected by dominance when using this approach.

p	k	q^2	NCP (N=100)	
			1 d.f. fit (a)	2 d.f. fit ($a + d$)
0.1	0.0	0.050	5.26	5.26
	0.5	0.049	5.13	5.26
	1.0	0.047	4.97	5.26
0.5	0.0	0.050	5.26	5.26
	0.5	0.044	4.65	5.26
	1.0	0.033	3.45	5.26
0.7	0.0	0.050	5.26	5.26
	0.5	0.042	4.39	5.26
	1.0	0.023	2.36	5.26

Table A2. Theory and simulation results for power loss due to dominance in an indirect FS design. A 1 d.f. model is fitted for all tests of association. Simulation details given in the text.

				Predicted: mean test statistic	Simulation: mean test statistic
p	k	a	d	F-ratio	F-ratio
0.1	0.0	0.527	0.000	5.90	5.98
	0.5	0.372	0.186	5.79	5.93
	1.0	0.285	0.285	5.63	5.85
0.5	0.0	0.316	0.000	5.90	6.01
	0.5	0.298	0.149	5.33	5.50
	1.0	0.258	0.258	4.21	4.34
0.7	0.0	0.345	0.000	5.90	6.09
	0.5	0.400	0.200	5.18	5.15
	1.0	0.391	0.391	3.20	3.20

S.E.s from runs of 10,000 replicates ranged from 0.4-0.5

Table A3. Theory and simulation results for power loss due to dominance in an indirect HS design.

A 1 d.f. model is fitted for all tests of association. Simulation details given in the text.

				Predicted: mean test statistic	Simulation: mean test statistic
p	*k*	*a*	*d*	*F* -ratio	*F* –ratio
0.1	0.0	0.527	0.000	8.04	8.18
	0.5	0.372	0.186	7.87	8.05
	1.0	0.285	0.285	7.65	7.78
0.5	0.0	0.316	0.000	8.04	8.13
	0.5	0.298	0.149	7.21	7.22
	1.0	0.258	0.258	5.59	5.62
0.7	0.0	0.345	0.000	8.04	8.10
	0.5	0.400	0.200	6.99	7.09
	1.0	0.391	0.391	4.13	4.16

S.E.s from runs of 10,000 replicates ranged from 0.4-0.6