579

Transformation of QTL genotypic effects to allelic effects

Yoshitaka NAGAMINE*

National Institute of Livestock and Grassland Science, Ikenodai 2, Tsukuba, 305-0901, Japan

Abstract – The genotypic and allelic effect models are equivalent in terms of QTL detection in a simple additive model, but the QTL allelic model has the advantage of providing direct information for marker-assisted selection. However, the allelic matrix is four times as large as the genotypic IBD matrix, causing computational problems, especially in genome scans examining multiple positions. Transformation from genotypic to allelic effects, after estimating the genotypic effects with a smaller IBD matrix, can solve this problem. Although the validity of transformation from genotypic to allelic effects has been disputed, this work proves that transformation can successfully yield unique allelic effects when genotypic and allelic IBD matrixes exist.

QTL / transformation / marker assist selection / genotypic effect / allelic effect

1. INTRODUCTION

The variance component method with random QTL effects is preferred for estimation of QTL genotypic or allelic effects [6, 19], since this method does not require specification of the number of alleles, and the relationship matrixes allow separate estimations of the polygenic and QTL effects. After the first report of QTL analysis in commercial pig populations [2, 14] using the least squares method [7, 11], Nagamine *et al.* [15] used variance component analysis to estimate the heritabilities and genotypic effects of QTL in commercial breeds. These authors also showed the transformation from genotypic to allelic effects [13] and applied this transformation to large pig data sets [15]. Since the allelic or gametic IBD, identity-by-descent, matrix is four times as large as the genotypic IBD matrix, inversion of the allelic matrix demands computational resources, especially when multiple positions are examined. Thus, the estimation of genotypic effects with the smaller IBD matrix followed by

---

* Corresponding author: nagamine@affrc.go.jp

transformation from genotypic to allelic effects has the benefit of decreasing computational cost. However, Tuchscherer *et al.* [17] wrote that this transformation was impossible and did not provide unique results. This paper offers derivation details proving that this transformation is possible, and shows that the transformation gives unique solutions.

2. MODELS AND DERIVATION

The genotypic and allelic effect models are equivalent when it comes to QTL detection in a simple additive model [10, 18]. An animal model including polygenic, QTL, genotypic and allelic effects is presented below [6, 10]. For the QTL genotypic effect model:

$$\mathbf{y} = \mathbf{X}f + \mathbf{Z}u + \mathbf{Z}w + e.$$

For the QTL allelic effect model:

$$\mathbf{y} = \mathbf{X}f + \mathbf{Z}u + \mathbf{Z}\mathbf{T}v + e.$$

In these cases, $\mathbf{X}$ is a design matrix for fixed effects, $\mathbf{Z}$ is an incident matrix relating records in $\mathbf{y}$ to the animals, and vectors f, u, w and e are values for fixed, polygene, QTL genotypic and residual effects, respectively. The sizes of u and w are calculated as $n \times 1$ for n individuals. Vector v is the QTL allelic or gametic effect having a size of $2n \times 1$, and $\mathbf{T}$ is an incident matrix relating each animal to its two allelic effects [18]. If all animals have records, using the Kronecker product *, $\mathbf{T}$ is:

$$\mathbf{I} * [1\ 1] = \begin{pmatrix} 1\ 1 & 0\ 0 & 0\ 0 & \dots \\ 0\ 0 & 1\ 1 & 0\ 0 & \dots \\ 0\ 0 & 0\ 0 & 1\ 1 & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}.$$

The fixed and random values are estimated using mixed model equations. $\mathbf{G}$ represents the allelic IBD matrix, while $\mathbf{A}$ is the relationship matrix for the polygenic effect [9]. The genotypic IBD matrix, $\mathbf{Q}$, is (1/2) $\mathbf{TGT'}$ [18]. The j and k values are the ratios of error variance/polygenic variance and error variance/QTL genotypic variance, respectively.

The QTL genotypic effect model can be written as:

$$\begin{pmatrix} \mathbf{X'X} & \mathbf{X'Z} & \mathbf{X'Z} \\ \mathbf{Z'X} & \mathbf{Z'Z} + \mathbf{A}^{-1}\mathrm{j} & \mathbf{Z'X} \\ \mathbf{Z'X} & \mathbf{Z'Z} & \mathbf{Z'Z} + \mathbf{Q}^{-1}\mathrm{k} \end{pmatrix} \begin{pmatrix} f \\ u \\ w \end{pmatrix} = \begin{pmatrix} \mathbf{X'y} \\ \mathbf{Z'y} \\ \mathbf{Z'y} \end{pmatrix}.$$

Using the third row of the left hand matrix, we see that:

$$\mathbf{Z'X}f + \mathbf{Z'Z}\,u + [\mathbf{Z'Z} + \mathbf{Q}^{-1}\mathrm{k}]w = \mathbf{Z'y}. \tag{1}$$

The QTL allelic effect model can be written as:

$$\begin{pmatrix} \mathbf{X'X} & \mathbf{X'Z} & \mathbf{X'ZT} \\ \mathbf{Z'X} & \mathbf{Z'Z} + \mathbf{A}^{-1}\mathrm{j} & \mathbf{Z'ZT} \\ \mathbf{T'Z'X} & \mathbf{T'Z'Z} & \mathbf{T'Z'ZT} + \mathbf{G}^{-1}(2\mathrm{k}) \end{pmatrix} \begin{pmatrix} f \\ u \\ v \end{pmatrix} = \begin{pmatrix} \mathbf{X'y} \\ \mathbf{Z'y} \\ \mathbf{T'Z'y} \end{pmatrix}.$$

If there is no interaction between alleles, *i.e.* $v_{11} + v_{12} = w_1$ for Animal 1, the QTL allelic variance is half of the QTL genotypic variance. Therefore, the variance ratio for the allelic IBD matrix is taken to be 2k. Here, $\mathbf{Q}^{-1}$ and $\mathbf{G}^{-1}$ are the inverses, not the generalized inverses, of $\mathbf{Q}$ and $\mathbf{G}$, respectively [16]. Using the third row of the left hand matrix, we see that:

$$\mathbf{T'Z'X}f + \mathbf{T'Z'Z}\,u + [\mathbf{T'Z'ZT} + \mathbf{G}^{-1}(2\mathrm{k})]v = \mathbf{T'Z'y}. \tag{2}$$

Then, multiply Eq. (1) by $\mathbf{T'}$:

$$\mathbf{T'Z'X}f + \mathbf{T'Z'Z}\,u + \mathbf{T'}[\mathbf{Z'Z} + \mathbf{Q}^{-1}\mathrm{k}]w = \mathbf{T'Z'y}.$$

This is equivalent to Eq. (2), therefore:

$$[\mathbf{T'Z'ZT} + \mathbf{G}^{-1}(2\mathrm{k})]v = \mathbf{T'}[\mathbf{Z'Z} + \mathbf{Q}^{-1}\mathrm{k}]w.$$

Here, $\mathbf{T'Z'ZT}\,v = \mathbf{T'Z'Z}\,w$, because in the absence of allelic interactions, $\mathbf{T}v = w$. Therefore:

$$\mathbf{G}^{-1}(2\mathrm{k})v = [\mathbf{T'Q}^{-1}\mathrm{k}]w, \ \text{and}$$
$$v = \mathbf{GT'Q}^{-1}(1/2)w.$$

The size of $\mathbf{G}$ is 2n × 2n; however $\mathbf{Q}$ (=(1/2) $\mathbf{TGT'}$) is an n × n matrix. The values of v can be calculated without the inverse of the $\mathbf{G}$ matrix. Essentially, it is not necessary to calculate $\mathbf{Q}^{-1}$ for this conversion, because the parameter is already specified to estimate w. When allelic effects are simple additive effects and $\mathbf{Q}^{-1}$ and $\mathbf{G}^{-1}$ exist, this transformation can give unique estimators of allelic effects, because genotypic and allelic effects are known to be unique solutions from mixed models [3, 9, 12], and the transformation does not affect the allelic values.

3. NUMERICAL EXAMPLE

This example utilizes the allelic relationships among four animals previously described by Fernando and Grossman [3]. Animals 1 and 2 are the parents of Animal 3. In order to create a highly inbred offspring, Animal 3 was mated with Animal 1 to produce Animal 4. The allelic relationship matrix is $\mathbf{G}$ and the genotypic relationship matrix is $\mathbf{Q}$, which was easily transformed from $\mathbf{G}$.

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0.9 & 0 & 0.5 & 0.81 \\ 0 & 1 & 0 & 0 & 0.1 & 0 & 0.5 & 0.09 \\ 0 & 0 & 1 & 0 & 0 & 0.1 & 0 & 0.01 \\ 0 & 0 & 0 & 1 & 0 & 0.9 & 0 & 0.09 \\ 0.9 & 0.1 & 0 & 0 & 1 & 0 & 0.5 & 0.9 \\ 0 & 0 & 0.1 & 0.9 & 0 & 1 & 0 & 0.1 \\ 0.5 & 0.5 & 0 & 0 & 0.5 & 0 & 1 & 0.45 \\ 0.81 & 0.09 & 0.01 & 0.09 & 0.9 & 0.1 & 0.45 & 1 \end{pmatrix}$$

$$(1/2)\,\mathbf{TGT'} = \mathbf{Q} = \begin{pmatrix} 1 & 0 & 0.5 & 0.95 \\ 0 & 1 & 0.5 & 0.05 \\ 0.5 & 0.5 & 1 & 0.75 \\ 0.95 & 0.05 & 0.75 & 1.45 \end{pmatrix}.$$

The QTL allelic effects were assumed to be completely additive. For simplicity, the example only utilizes the QTL effect, and ratio k is set to one. Vector $\mathbf{y}$ is for animal observations, but in this case Animals 1 and 2 have no records. Matrix $\mathbf{Z}$ is a design matrix, with all elements set to zero except for the two diagonal elements, (3, 3) and (4, 4), which have values of one. This yields:

$$\mathbf{y'} = \begin{pmatrix} 0 & 0 & 100 & 120 \end{pmatrix}.$$

If the model for QTL genotypic effects is

$$(\mathbf{Z'Z} + \mathbf{Q}^{-1}k)(\mathbf{w}) = \mathbf{Z'y},$$

then the estimated genotypic effects are $\mathbf{w'} = (54.006, 19.770, 64.265, 81.959)$.

Similarly if the model for QTL allelic effects is

$$(\mathbf{T'Z'ZT} + \mathbf{G}^{-1}2k)(\mathbf{v}) = \mathbf{T'Z'y},$$

then the estimated allelic effects are $\mathbf{v'} = (40.997, 13.009/1.977, 17.793/44.496, 19.769/36.513, 45.447)$. Thus, we can confirm that the sum of the allelic effects is equal to the genotypic effect, *e.g.* $54.006 = 40.997 + 13.009$ for Animal 1.

For the transformation from genotypic to allelic effects,

$$\mathbf{Q}^{-1} = \begin{pmatrix} 2.6666 & 0.1666 & -0.1666 & -1.6666 \\ 0.1666 & 1.5952 & -1.2381 & 0.4762 \\ -0.1666 & -1.2381 & 2.5952 & -1.1904 \\ -1.6666 & 0.4762 & -1.1904 & 2.3809 \end{pmatrix}.$$

The transformation gives v as the following:

$$v = \mathbf{GT'Q}^{-1}(1/2)w = \mathbf{GT'Q}^{-1}(1/2) \begin{pmatrix} 54.006 \\ 19.770 \\ 64.265 \\ 81.959 \end{pmatrix} = \begin{pmatrix} 40.997 \\ 13.009 \\ 1.977 \\ 17.793 \\ 44.496 \\ 19.769 \\ 36.513 \\ 45.447 \end{pmatrix}.$$

We can confirm that the transformation form, $v = \mathbf{GT'Q}^{-1}(1/2)w$, gives the same v as the estimators from the allelic model.

4. CONCLUSION

Genotypic QTL effects can be estimated with computer software packages such as SOLAR [1], and the two-step method [4] using LOKI [8] and ASReml [5]. In practice, after estimating genotypic effects by the two-step method, Nagamine *et al.* [15] applied the transformation from the genotypic effect to the allelic effect. However, only the final derivation of transformation from QTL genotypic effects to allelic effect was written in the previous papers [13, 15]. Here, I have detailed the derivation with a numerical example, showing that the transformation is possible and that it successfully gives a unique estimator of allelic effects in the additive effect model when $\mathbf{Q}^{-1}$ and $\mathbf{G}^{-1}$ exist.

REFERENCES

[1] Almasy L., Blangero J., Multipoint quantitative-traits linkage analysis in general pedigrees, Am. J. Hum. Genet. 62 (1998) 1198–1211.
[2] Evans G.J., Giuffra E., Sanchez A., Kerje S., Davalos G., Vidal O., Illan S., Noguera J.L., Varona L., Velander I., Southwood O.I., de Koning D.J., Haley C.S., Plastow G.S., Anderson L., Identification of quantitative trait loci for production traits in commercial pig populations, Genetics 164 (2003) 621–627.

[3] Fernando R.L., Grossman M., Marker assisted selection using best linear unbiased prediction, Genet. Sel. Evol. 21 (1989) 467–477.

[4] George A.W., Visscher P.M., Haley C.S., Mapping quantitative trait loci in complex pedigrees, Genetics 156 (2000) 2081–2092.

[5] Gilmour A.R., Cullis B.R., Welham S.J., Thompson R., ASREML Program user manual, Orange Agricultural Institute, NSW, 1999.

[6] Grignola F.E., Hoeschele I., Tier B., Mapping quantitative trait loci in outcross populations via residual maximum likelihood. I. Methodology, Genet. Sel. Evol. 28 (1996) 479–490.

[7] Haley C.S., Knott S.A., Elsen J.M., Mapping quantitative trait loci in crosses between outbred lines using least squares, Genetics 136 (1994) 1195–1207.

[8] Heath S.C., Markov chain Monte Carlo segregation and linkage analysis for oligogenic models, Am. J. Hum. Genet. 61 (1997) 748–760.

[9] Henderson C.R., Application of linear models in animal breeding, University of Guelph, Ontario, 1984.

[10] Hoeschele I., Mapping quantitative trait loci in outbred pedigrees, in: Balding D.J., Bishop M., Cannings C. (Eds.), Handbook of statistical genetics, John Wiley & Sons, Chichester, 2001, pp. 599–644.

[11] Knott S.A., Elsen J.M., Haley C.S., Methods for multiple-marker mapping of quantitative trait loci in half-sib populations, Theor. Appl. Genet. 93 (1996) 71–80.

[12] Mrode R.A., Linear Models for the prediction of animal breeding values, CAB International, Oxon, 2000.

[13] Nagamine Y., Knott S.A., Visscher P.M., Haley C.S., Simple deterministic identity-by-descent coefficients and estimation of QTL allelic effects in full and half sibs, Genet. Res. 80 (2002) 237–243.

[14] Nagamine Y., Haley C.S., Sewalem A., Visscher P.M., Quantitative trait loci variation for growth and obesity between and within lines of pigs (*Sus scrofa*), Genetics 164 (2003) 629–635.

[15] Nagamine Y., Visscher P.M., Haley C.S., QTL detection and allelic effects for growth and fat traits in outbred pig populations, Genet. Sel. Evol. 36 (2004) 83–96.

[16] Searle S.R., Matrix algebra useful for statistics, John Wiley & Sons, New York, 1982.

[17] Tuchscherer A., Mayer M., Reinsch N., Identification of gametes and treatment of linear dependencies in the gametic QTL-relationship matrix and its inverse, Genet. Sel. Evol. 36 (2004) 621–642.

[18] van Arendonk J.A.M., Tier B., Kinghorn B.P., Use of multiple genetic markers in prediction of breeding values, Genetics 137 (1994) 319–329.

[19] Xu S., Atchley W.R., A random model approach to interval mapping of quantitative trait loci, Genetics 141 (1995) 1189–1197.