

ARTICLE

# Multivariate QTL linkage analysis suggests a QTL for platelet count on chromosome 19q

David M Evans<sup>\*,1,2</sup>, Gu Zhu<sup>1</sup>, David L Duffy<sup>1</sup>, Grant W Montgomery<sup>1</sup>, Ian H Frazer<sup>3</sup> and Nicholas G Martin<sup>1</sup>

<sup>1</sup>Queensland Institute of Medical Research, Brisbane, Australia; <sup>2</sup>Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; <sup>3</sup>Centre for Immunology and Cancer Research, University of Queensland, Princess Alexandra Hospital, Brisbane, Australia

Platelet count is a highly heritable trait with genetic factors responsible for around 80% of the phenotypic variance. We measured platelet count longitudinally in 327 monozygotic and 418 dizygotic twin pairs at 12, 14 and 16 years of age. We also performed a genome-wide linkage scan of these twins and their families in an attempt to localize QTLs that influenced variation in platelet concentrations. Suggestive linkage was observed on chromosome 19q13.13–19q13.31 at 12 (LOD = 2.12,  $P = 0.0009$ ), 14 (LOD = 2.23,  $P = 0.0007$ ) and 16 (LOD = 1.01,  $P = 0.016$ ) years of age and multivariate analysis of counts at all three ages increased the LOD to 2.59 ( $P = 0.0003$ ). A possible candidate in this region is the gene for glycoprotein VI, a receptor involved in platelet aggregation. Smaller linkage peaks were also seen at 2p, 5p, 5q, 10p and 15q. There was little evidence for linkage to the chromosomal regions containing the genes for thrombopoietin (3q27) and the thrombopoietin receptor (1q34), suggesting that polymorphisms in these genes do not contribute substantially to variation in platelet count between healthy individuals.

*European Journal of Human Genetics* (2004) 12, 835–842. doi:10.1038/sj.ejhg.5201248

Published online 28 July 2004

**Keywords:** multivariate; QTL; linkage; variance components; platelets

## Introduction

Platelets are small, anucleate cell fragments that play an integral role in the blood clotting process. The number of platelets in a given volume of blood (ie the platelet count) is correlated with several clinically relevant phenotypes including blood clotting time,<sup>1</sup> and may be a risk factor in the development of thrombosis<sup>2,3</sup> and atherosclerosis.<sup>4,5</sup> A very low platelet count (thrombocytopenia) is a common side effect of cancer chemotherapy and is a limiting factor in the dose of cytotoxic agents that can be used. Current treatment typically involves the transfusion of platelet concentrates, but this carries significant risks including transmission of infectious agents as well as alloimmuniza-

tion and transfusion reactions. In addition, clinical trials of several new agents including recombinant thrombopoietin (a major regulator of platelet production) have yet to demonstrate a clear benefit in the treatment of thrombocytopenic cancer patients receiving dose-intensive chemotherapy for acute leukemia. There is clearly an urgent need to develop other agents which are effective in treating thrombocytopenic states.

We have shown that platelet count is highly heritable with genetic factors explaining 80% of the phenotypic variance in adolescent twins.<sup>6</sup> In this paper we extend these results by performing a genome-wide linkage scan of the dizygotic (DZ) twin pairs and their families from that study. We measured twins' platelet count longitudinally at 12, 14 and 16 years of age, and sought to increase the power to detect linkage by performing multivariate multi-point sib-pair linkage analyses across the genome.<sup>7–11</sup> Identification of QTLs for platelet count will improve our

\*Correspondence: D Evans, The Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford OX3 7BN, UK. Tel: +44 1865 287598; Fax: +44 1865 287697; E-mail: davide@well.ox.ac.uk  
Received 12 January 2004; revised 18 May 2004; accepted 25 May 2004

understanding of platelet biology and may suggest novel treatment strategies for thrombocytopenic conditions.

### Subjects and methods

Twins were recruited as part of an ongoing study concerned with the development of melanocytic naevi (moles), the clinical protocol of which has been described in detail elsewhere.<sup>12–14</sup> Twins were enlisted by contacting the principals of primary schools in the greater Brisbane area, by media appeals and by word of mouth. It is estimated that approximately 50% of the eligible birth cohort were recruited into the study and they were typical of the population with respect to platelet concentrations.<sup>6</sup> Informed consent was obtained from all participants and parents prior to testing.

Venous blood samples were collected using EDTA tubes and platelet count was subsequently measured using a Coulter Model STKS blood counter. Blood was collected from twins at 12, 14 and 16 years of age. Where possible, DNA was also obtained from parents and siblings for genotyping (platelet count was not measured in these relatives). No attempt was made to exclude subjects suffering from illness.

Platelet count was available from 745 twin pairs comprising 162 MZ female, 165 MZ male, 101 DZ female, 110 DZ male and 207 opposite sex (OS) twin pairs (consisting of 94 pairs where the female was born first and 113 pairs where the male was born first). Although twins were tested as close as possible to their 12th, 14th and 16th birthdays, not all twins were tested at all three measurement occasions (see Table 1 for a breakdown of these data).

### Genotypes

DNA was extracted from buffy coats using a modification of the 'salt method'.<sup>15</sup> For same sex twin pairs, zygosity was determined by typing nine independent DNA

microsatellite polymorphisms plus the X/Y amelogenin marker for sex-determination by polymerase chain reaction (ABI Profiler system). All twins were also typed for ABO, Rh and MNS blood groups.

The genome scan consisted of 726 highly polymorphic autosomal microsatellite markers at an average spacing of ~5 cM in 539 families (2360 individuals). Markers on the X chromosome were also typed but linkage to these is not reported here. The microsatellites consisted of a combination of markers from the ABI-Prism and CIDR genotyping sets. Overlapping parts of the sample received either a 10 cM scan using the ABI-2 marker set (400 markers) at the Australian Genome Research Facility (Melbourne), a 10 cM scan using the Weber marker set at Center for Inherited Disease Research (Baltimore), or both. Only 30 markers were common to both marker sets and were used for quality control; the remaining markers intercalated to form a scan at approximately 5 cM spacing. The only families to receive one scan had both parents genotyped, and so had a high information content. The average heterozygosity of markers was 0.78, and the mean information content was 0.77. Full details of the genome scan are provided elsewhere.<sup>16</sup>

### Linkage analyses

Univariate multipoint variance components linkage analysis was used to test for linkage between marker loci and blood cell phenotypes.<sup>9,17–19</sup> In the variance components method, the expected phenotypic covariance matrix is partitioned into the following components:

$$\Sigma = \hat{\Pi}\sigma_q^2 + 2\Phi\sigma_a^2 + \mathbf{I}_n\sigma_e^2$$

where  $\sigma_q^2$  is the additive genetic variance due to the QTL,  $\sigma_a^2$  is the (residual) polygenic additive genetic variance, and  $\sigma_e^2$  is the unique environmental variance.<sup>17,20</sup>  $\hat{\Pi}$  is a matrix containing the elements  $\hat{\pi}_{ij}$ , which denote the estimated proportion of alleles shared IBD at the trait locus by

**Table 1** Breakdown of participation data showing the number of complete twin pairs for whom platelet count and genotyping information (in the case of DZ pairs) were available

12	14	16	MZF	MZM	DZF	DZM	DZFM	DZMF	Total
✓			19	31	12	13	12	10	97
	✓		10	5	3	5	3	4	30
		✓	49	56	15	11	21	22	174
✓	✓		30	26	28	38	22	40	184
✓		✓	4	3	4	2	2	5	20
	✓	✓	9	10	7	7	6	7	46
✓	✓	✓	41	34	32	34	28	25	194
Total			162	165	101	110	94	113	745
Total at 12			94	94	76	87	64	80	495
Total at 14			90	75	70	84	59	76	454
Total at 16			103	103	58	54	57	59	434

MZF = monozygotic female twins; MZM = monozygotic male; DZF = dizygotic female; DZM = dizygotic male; DZFM = opposite sex twins with the female born first; DZMF = opposite sex twins with the male born first.

individuals  $i$  and  $j$ .  $\hat{\mathbf{\Pi}}$  is an estimate of the true identity by descent sharing matrix  $\mathbf{\Pi}$  and is a function of the estimated IBD matrices as well as the distance between the marker and the QTL.<sup>17,21</sup>  $\Phi$  is the Kinship matrix containing elements  $\varphi_{ij}$  – the kinship coefficients between individuals  $i$  and  $j$  (ie equal to half the additive genetic sharing coefficients),<sup>22</sup> and  $\mathbf{I}_n$  is an identity matrix of order  $n$ .

Variance components were estimated by maximum-likelihood analysis of the raw data<sup>23</sup> as implemented in the software package MERLIN<sup>24</sup> along with fixed effects for sex and age. Since both circadian and seasonal effects have been reported for platelet count, linear, quadratic and sinusoidal fixed effects were included for the time of day and month from which blood was sampled.<sup>25</sup> Univariate linkage analyses were performed at each marker at each age. Only phenotypic data from DZ pairs were included in the analyses because MZ twins share all their genes identical by descent (IBD) across the genome, and are thus uninformative for linkage.

The null hypothesis that additive genetic variance in platelet count caused by a QTL linked to a given marker is zero (ie  $\sigma_q^2=0$ ) was tested by comparing the likelihood of a reduced model in which  $\sigma_q^2$  was constrained to zero with the likelihood of a model in which the genetic variance due to the QTL ( $\sigma_q^2$ ) was estimated. Twice the difference in natural log-likelihood between these models is distributed asymptotically as a 1/2:1/2 mixture of  $\chi_1^2$  and a point mass at zero,<sup>26</sup> while the difference between the two log<sub>10</sub> likelihoods produces a LOD score equivalent to the classical LOD score of parametric linkage analysis.<sup>27</sup>

Previous studies have suggested that variance components linkage analysis is sensitive to deviations from multivariate normality, particularly to high levels of kurtosis in the trait distribution.<sup>28</sup> The distribution of platelet counts for 12, 14 and 16 year olds exhibited a kurtosis of 0.102, 0.068 and  $-0.002$  respectively. Recent statistical genetics theory suggests that this level of kurtosis will have relatively minor effects on the distribution of LOD scores and that the standard nominal  $P$ -values for LOD scores are appropriate in these cases.<sup>29</sup>

### Multivariate analyses

Several groups have demonstrated that multivariate methods are a powerful method of detecting QTLs that influence a set of phenotypes pleiotropically.<sup>7,8,11,30</sup> Given that the Pearson correlations between platelet count at the various time points are quite high ( $r_{12\&14}=0.74$ ;  $r_{14\&16}=0.71$ ;  $r_{12\&16}=0.65$ ), our longitudinal data for platelet count are well suited to multivariate QTL linkage analysis.<sup>31</sup> In this study, multivariate QTL linkage analyses were performed using structural equation modeling as implemented in Mx,<sup>9,10,32</sup> using data from both MZ and DZ pairs. The QTL was modeled as a single latent factor that pleiotropically affected the phenotypes (Figure 1). In this model, the correlation between the QTL effects was set

to one for MZ twins (since MZ twins are genetically identical), and to  $\hat{\mathbf{\Pi}}$ , the estimated proportion of alleles shared IBD at the marker locus, for DZ twins. Probability of sharing zero ( $p_0$ ), one ( $p_1$ ) or two ( $p_2$ ) marker alleles IBD for each DZ twin pair was calculated in a multipoint fashion using a modification of the Lander-Green algorithm as implemented in MERLIN.<sup>24</sup> These probabilities were then used to obtain  $\hat{\mathbf{\Pi}} = p_2 + 1/2p_1$  for each DZ pair at each marker location. Residual sources of variation were modeled using saturated Cholesky structures (Figure 1). Since phenotypic information was present from MZ twins, it was possible to partition the residual familial resemblance into separate variance components due to additive genetic (A) and common environmental (C) sources of variation.

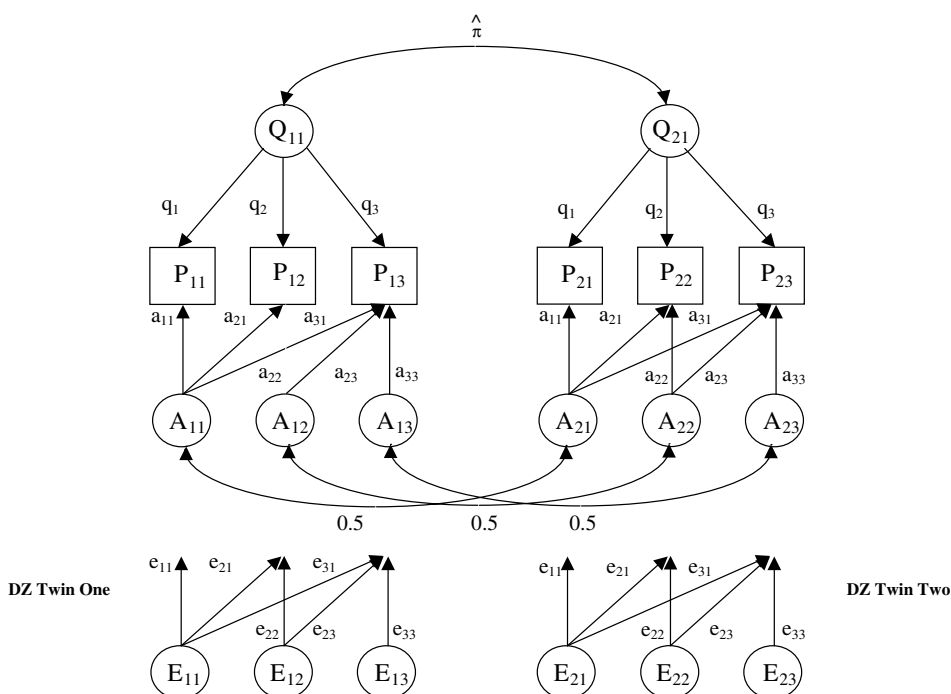
Trivariate linkage analyses of platelet count at age 12, 14 and 16 were conducted at each marker across the genome. For these longitudinal analyses we included the same fixed effects as in the univariate models above as well as age deviations at 14 and 16 years of age in the analysis.<sup>20</sup>

Two multivariate tests for linkage were performed. In the first ('Multivariate Test I'), the factor loadings of the QTL on platelet count at 12, 14 and 16 were unconstrained. As the true values of some of these parameters under the null hypothesis of no linkage are located on the boundary of the parameter space defined by the alternative hypothesis, the likelihood ratio test statistic is distributed as a complicated mixture of  $\chi^2$  distributions.<sup>26</sup> We therefore generated the empirical distribution of the test statistic under the null hypothesis of no QTL effects on any trait, while assuming the same background genetic and environmental correlations as in the actual data (10 000 simulations). We then used this empirical distribution ( $\mu=2.83$ ,  $\sigma^2=6.50$ ) to evaluate the significance of Multivariate Test I.

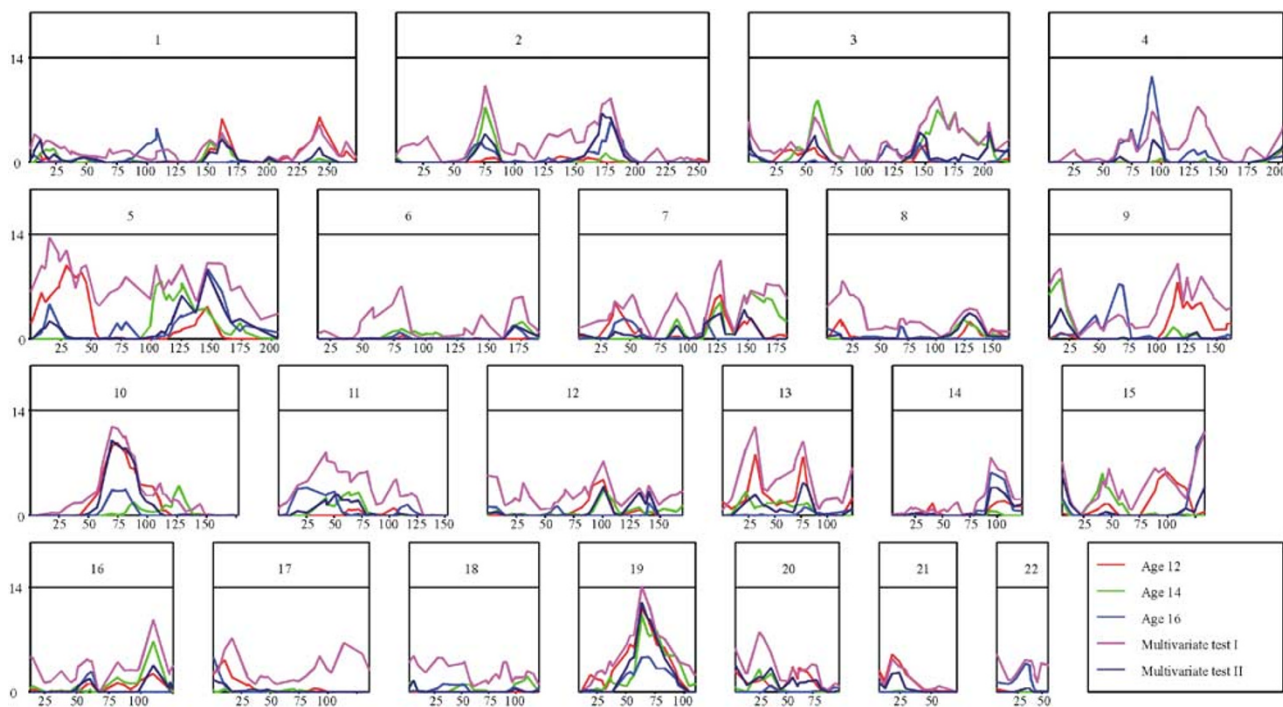
In order to perform the second test (ie Multivariate Test II) we first tested whether we could equate the factor loadings of the three measurement occasions on the QTL (ie we calculated the difference in minus two log-likelihood between the full model and the model with the equated loadings and evaluated it against a  $\chi^2$  distribution with two degrees of freedom). This was equivalent to testing whether the QTL was responsible for the same amount of phenotypic variation at each age. If this were the case, then the test for linkage was whether the (equated) loadings could then be set to zero. Since only one QTL variance component was estimated, the test statistic was distributed as in the univariate case (ie a 50:50 mixture of a point mass at zero and  $\chi_1^2$ ). Note that this test is equivalent to taking the mean of the phenotypes across the three ages and performing a univariate test of linkage on this statistic.<sup>11</sup>

### Results

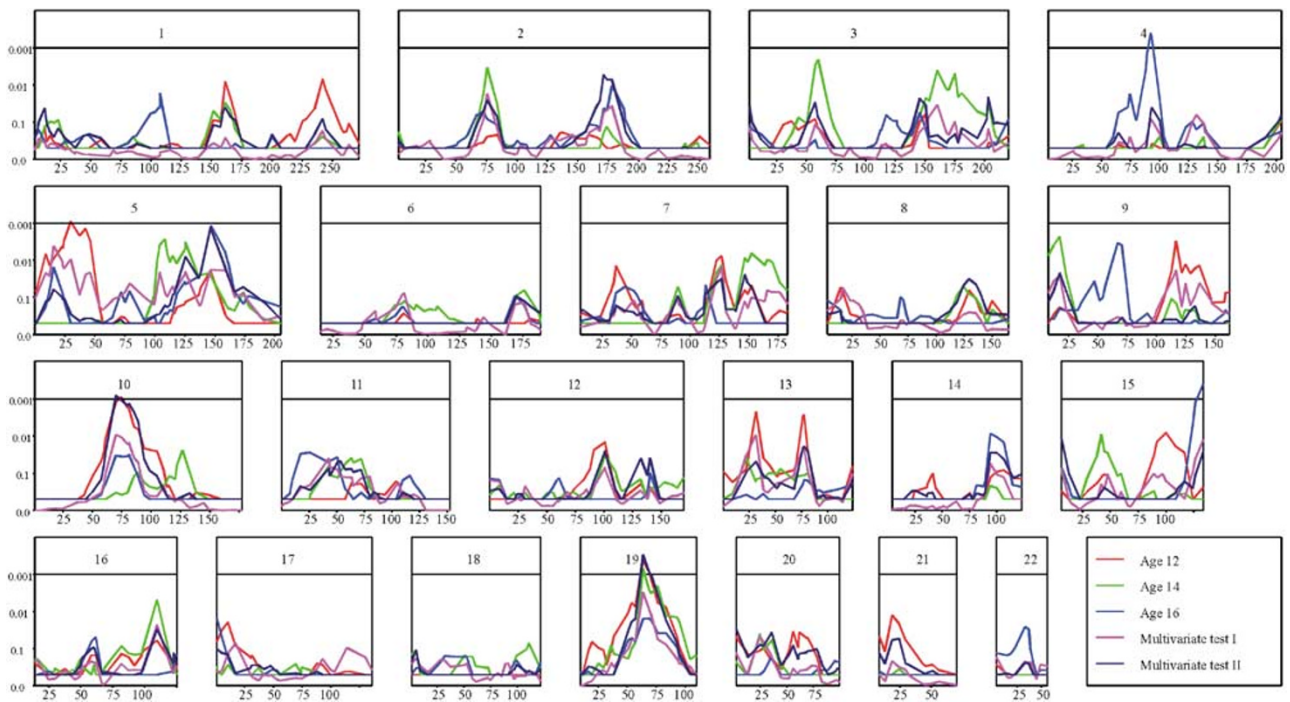
The results from the genome-wide univariate and multivariate tests of linkage are displayed in Figures 2 and 3. Figure 2 shows the difference in  $\chi^2$  at each marker along the



**Figure 1** Path diagram for multivariate variance components QTL linkage analysis. The QTL is parameterized as a single latent variable that pleiotropically affects the phenotypes. Residual sources of variation are parameterized using Cholesky structures. Only additive genetic and unique environmental residual sources of variation are shown. When MZ twins are present, similar structures for the common environment/dominance may also be included in the model.



**Figure 2** Difference in minus two log-likelihood  $\chi^2$  for platelet count. Distance in cM is shown on the x-axis.



**Figure 3** Significance levels for univariate and multivariate tests of genetic linkage for platelet count. Distance in cM is shown on the x-axis.

genome, while Figure 3 displays the corresponding significance levels for these tests of linkage. The most prominent result was the linkage peak at ages 12 (LOD = 2.12,  $p = 0.000896$ ), 14 (LOD = 2.23,  $p = 0.00068$ ) and 16 (LOD = 1.01,  $p = 0.015563$ ) in the same region of chromosome 19q. Other regions of suggestive linkage (ie LOD scores > 2.0) were also observed on chromosome 4 at age 16 (LOD = 2.49,  $p = 0.000357$ ), on different regions of chromosome 5 at ages 12 (LOD = 2.14,  $p = 0.000852$ ) and 16 (LOD = 2.02,  $p = 0.001151$ ), on chromosome 10 at age 12 (LOD = 2.1,  $p = 0.00942$ ), and on chromosome 15 at age 16 (LOD = 2.44,  $p = 0.000404$ ). With the exception of one of the peaks on chromosome 5, which showed some evidence of linkage at the other two ages (12: LOD = 0.94,  $p = 0.018789$ ; 14: LOD = 0.89,  $p = 0.021518$ ), there was little evidence for linkage in similar regions at other ages.

Comparison of the pink and dark blue lines with the others in Figure 2 show that both multivariate tests were associated with an increase in  $\chi^2$  across most regions of the genome. However, in the case of the first multivariate test, this increase in  $\chi^2$  was also accompanied by an increase in the associated degrees of freedom. This resulted in a decreased level of significance across many areas of the genome, including the region on chromosome 19q (Figure 3). In contrast, the second multivariate test increased the power to detect linkage in several regions including chromosome 19q at the D19S220 (LOD = 2.59,  $p = 0.000277$ ) and D19S420 markers (LOD = 2.11,

$p = 0.00092$ ), as well as regions on chromosomes 2q, 10 and 15 (see Table 2 for a summary of regions of the genome with a LOD score in excess of one for the second multivariate analysis).

Figure 4 displays the standardized path coefficients under the full multivariate model, under the model where the QTL factor loadings were equated, and under the model of no QTL effects. Results are shown for the D19S220 marker (ie the marker where evidence for linkage was maximal). Note in particular how the factor loadings for the QTL under the full multivariate model were similar at all ages, suggesting that the QTL was responsible for similar amounts of variation in the phenotype at all ages. Consistent with this interpretation, the difference in  $\chi^2$  between the full multivariate model and the model where the QTL factor loadings were equated was not significant ( $\chi^2 = 2.19$ ,  $p = 0.335$ ).

### Discussion

To our knowledge, this is the first study to have used a complete genome scan to attempt to map genes responsible for variation in platelet count. A major strength of this study has been the availability of multiple measurements on the same individuals across time. Not only has this permitted multivariate QTL linkage analysis,<sup>9–11</sup> but has also provided supporting evidence for linkage when a signal has been present in the same region at multiple ages.

While these common regions of linkage do not formally constitute replication (since the longitudinal data are not independent), they do indicate the extent to which the present results are robust with respect to measurement error and temporal changes in the phenotype.

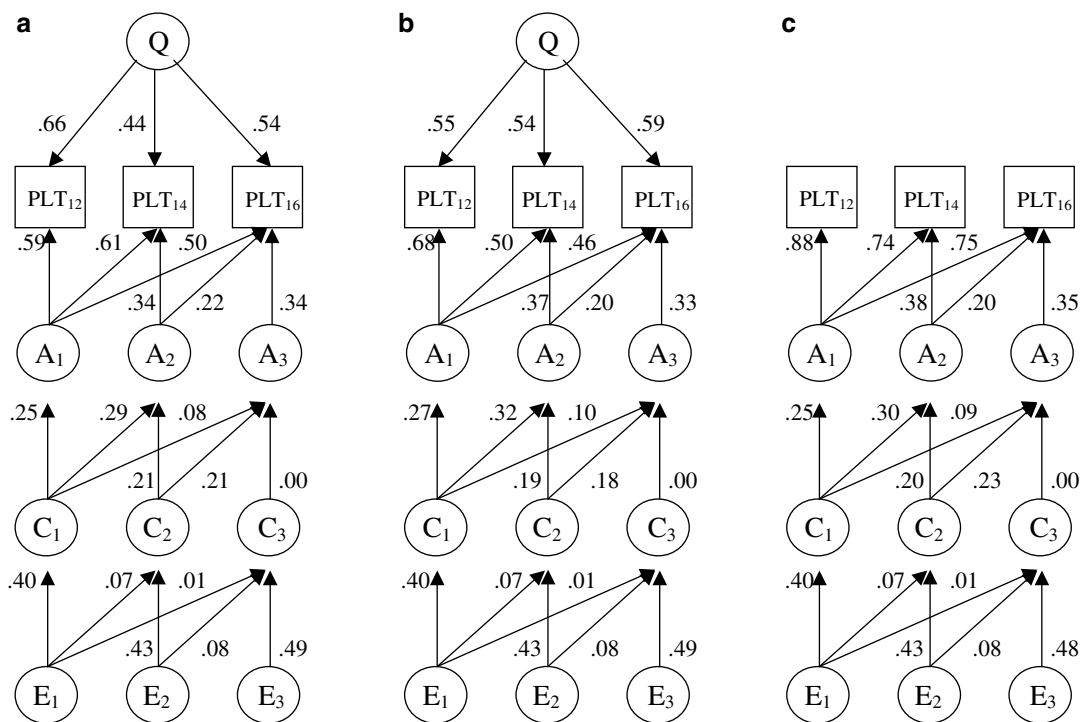
The most striking results from this study were the linkage peaks on the same region of chromosome 19q which were present at all three ages. The evidence for linkage in this region was greatest at ages 12 and 14, possibly reflecting the smaller number of DZ pairs from whom platelet count was available at age 16 (see Table 1). In addition, multivariate analyses (ie Multivariate Test II) increased the evidence for linkage at some of the markers in this region increasing support for the existence of a QTL. The most

promising candidate in this region is the gene for glycoprotein VI. Glycoprotein VI is a membrane-bound receptor that plays an important role in platelet aggregation in response to collagen.<sup>33</sup> It has been demonstrated that polymorphism in this gene is associated with differences in collagen-induced platelet aggregation.<sup>34</sup> Very recently, Nurden *et al* (2004) reported a patient with symptoms of gray platelet syndrome (ie a low platelet count and absence of  $\alpha$ -granules) accompanied by a severe deficiency in glycoprotein VI.<sup>35</sup> These and similar observations have led to speculation that polymorphism in the gene for glycoprotein VI may be a risk factor in the development of arterial thrombosis.<sup>36</sup> Our results suggest that glycoprotein VI or a gene close by may be responsible for nonpathological variation in platelet count.

Several other regions of 'suggestive' linkage (defined as  $\text{LOD} > 2.0$ ) were identified in this study, including a peak on chromosome 4 at age 16, two peaks on chromosome 5 at ages 12 and 16, respectively, a peak on chromosome 10 at age 12, and a peak on chromosome 15 at age 16. Curiously, in most cases there was little evidence of linkage in these regions at other ages (the exception was one of the peaks on chromosome 5). One possibility is that different genes may regulate platelet count at different ages. However, we consider this explanation unlikely given that the majority of the genetic variance was transmitted

**Table 2** Regions with  $\text{LOD} > 1$  for the second multivariate analysis of platelet count

Chromosome	Region	cM range	Peak LOD
2	D2S2330–D2S335	171.1–177.98	1.42
5	D5S471–D5S1505	124.98–125.35	1.26
5	D5S1480–D5S410	146.55–158.44	1.97
10	D10S196–D10S537	70.07–89.16	2.18
15	D15S128	5.82	1.15
19	D19S245–D19S902	56.88–76.15	2.59



**Figure 4** Standardized parameter estimates for the multivariate analysis of platelet count at the D19S220 marker. Results are shown for the full multivariate model: (a)  $-2\text{LL} = 2963.589$   $\text{df} = 2761$ ; the model where the QTL factor loadings are equated: (b)  $-2\text{LL} = 2965.777$   $\text{df} = 2763$ ; and the model where the QTL effect has been dropped (c)  $-2\text{LL} = 2977.704$   $\text{df} = 2764$ .

through the time series, and that the role of new genetic innovation was minimal (fitting a time series model to the data did not change interpretation of the results — data not shown).<sup>31</sup> More likely is that some of these peaks represent stochastic variation associated with the sampling of a complex phenotype. That is, while some of these regions will harbor QTLs, other peaks will simply be a result of random fluctuation and type I error. While multivariate analysis can go some way to addressing these questions (see below), the key to assessing the significance of these linkages will be replication using independent data sets.

Since this is the first study of its kind, there are no previous linkage studies with which to compare the location of the peaks. However, it is interesting to note that there was little evidence for linkage in the chromosomal regions containing the genes for thrombopoietin (3q27) and its receptor (1q34). While mutations in the genes for thrombopoietin and the thrombopoietin receptor are known to produce clinical pathology (eg Congenital Amegakaryocytic Thrombocytopenia, Essential Thrombocythemia), the present results suggest that other regions of the genome may be more important in determining variation in platelet count amongst healthy adolescent twins.<sup>37,38</sup>

In most positions along the genome the multivariate analyses (ie Multivariate test I and Multivariate test II) did not increase the power to detect linkage relative to the univariate results. This was perhaps surprising given that several theoretical papers have documented a clear advantage of multivariate approaches in the power to detect QTLs.<sup>7,8,11</sup> There are several possible reasons for this discrepancy. First, we note that since there are three phenotypes (ie platelet count at 12, 14 and 16), strictly speaking each univariate test should be evaluated with a higher threshold of significance than if a single trait were analyzed. Second, any increase in power from a multivariate analysis will tend to be small when the traits are highly correlated.<sup>39</sup> Third, univariate analyses are more susceptible to random variation. In other words, it is not uncommon to see more similar trait values in individuals who share two alleles IBD at the marker locus than individuals who share zero marker alleles IBD as a result of random variation in the bivariate distribution of phenotypes. In contrast, multivariate analyses are less susceptible to this randomness because additional information is incorporated into the test through the trait covariances. In other words, to increase the evidence of linkage, any random variation would have to induce an appropriate pattern of IBD sharing not only in the cross-sibling covariance, but also in the cross-trait cross-sibling covariances as well. It is for this reason that the multivariate result should be a more precise reflection of the true effect due to the QTL.<sup>10</sup>

However, a complicating factor in the interpretation of the multivariate analyses is the large amount of missing

data in the study. Multivariate approaches provide increased power to detect linkage because they capture additional information on the QTL from the cross-time covariances as well as the cross-sibling cross-time covariances. Since a substantial proportion of sib-pairs were measured on only one occasion (see Table 1), information from the cross-time covariances was not always available. Since the degrees of freedom associated with the first multivariate test increased with the number of traits in the analysis, this may have resulted in a decrease in the power to detect linkage. In this regard, we note that the significance of the first multivariate test was evaluated against an empirical distribution of  $\chi^2$  scores rather than against an asymptotic distribution. This was because the asymptotic distribution of the likelihood ratio statistic for multivariate tests of linkage is complicated and has yet to be well characterized.<sup>10,30</sup> We suggest that there is an urgent need to characterize the asymptotic distribution associated with these multivariate tests of linkage so that the full power associated with these approaches can be realized.

In contrast, the test statistic for Multivariate test II is distributed exactly as in the univariate case (ie a 50:50 mixture of a point mass at zero and  $\chi^2_1$ ). This test is a natural and powerful method of increasing the power to detect QTLs, which is particularly appropriate when analyzing repeated measures data (so long as the QTL affects the measurements pleiotropically). The test provided increased evidence of linkage at several regions across the genome including areas on chromosomes 2, 10 and 15, but most importantly at two markers on chromosome 19q13.13–19q13.31. We also note that another group has recently demonstrated suggestive evidence of linkage to platelet count in a homologous region of the mouse genome.<sup>40</sup>

In conclusion, we have identified an area on chromosome 19q that shows suggestive linkage to platelet count at 12, 14 and 16 years of age. Multivariate analyses provided a small increase in the power to detect linkage at two of the markers in this region. We intend to fine map this location (ie increasing marker density) and perform association analyses involving the gene for glycoprotein VI to better localize the area involved. In addition, the univariate and multivariate analyses revealed several other regions of interest on chromosomes 2, 4, 5, 8, 10 and 15. We hope that this study represents the first step in the eventual identification of genes that will not only increase our understanding of platelet biology but also suggest novel therapeutic possibilities in the treatment of thrombocytopenic states.

#### Acknowledgements

Collection of phenotypes and DNA samples was supported by grants from the Queensland Cancer Fund, the Australian National Health and Medical Research Council (950998, 981339 and 241944) and the US National Cancer Institute (CA88363) to Dr Nick Hayward.

The genome scans were supported by the Australian NHMRC's Program in Medical Genomics and funding from the Center for Inherited Disease Research (Director, Dr Jerry Roberts) at Johns Hopkins University to Dr Jeff Trent. We thank Ann Eldridge, Marlene Grace and Anjali Henders for assistance and the twins, their siblings, and their parents for their cooperation.

## References

- Rodgers RP, Levin J: A critical reappraisal of the bleeding time. *Semin Thromb Hemost* 1990; **16**: 1–20.
- Randi ML, Fabris F, Cella G, Rossi C, Girolami A: Cerebral vascular accidents in young patients with essential thrombocythemia: relation with other known cardiovascular risk factors. *Angiology* 1998; **49**: 477–481.
- Daniel S, O'Brien JR, John JA: Platelets in the prediction of thrombotic risk. *Atherosclerosis* 1982; **45**: 91–99.
- Kario K, Matsuo T, Nakao K: Cigarette smoking increases the mean platelet volume in elderly patients with risk factors for atherosclerosis. *Clin Lab Haematol* 1992; **14**: 281–287.
- Nieto FJ, Szklo M, Folsom AR, Rock R, Mercuri M: Leukocyte count correlates in middle-aged adults: the Atherosclerosis Risk in Communities (ARIC) Study. *Am J Epidemiol* 1992; **136**: 525–537.
- Evans DM, Frazer IH, Martin NG: Genetic and environmental causes of variation in basal levels of blood cells. *Twin Res* 1999; **2**: 250–257.
- Boomsma DI: Using multivariate genetic modeling to detect pleiotropic quantitative trait loci. *Behav Genet* 1996; **26**: 161–166.
- Boomsma DI, Dolan CV: A comparison of power to detect a QTL in sib-pair data using multivariate phenotypes, mean phenotypes, and factor scores. *Behav Genet* 1998; **28**: 329–340.
- Eaves LJ, Neale MC, Maes H: Multivariate multipoint linkage analysis of quantitative trait loci. *Behav Genet* 1996; **26**: 519–525.
- Marlow AJ, Fisher SE, Francks C *et al*: Use of multivariate linkage analysis for dissection of a complex cognitive trait. *Am J Hum Genet* 2003; **72**: 561–570.
- Martin N, Boomsma D, Machin G: A twin-pronged attack on complex traits. *Nat Genet* 1997; **17**: 387–392.
- Aitken JF, Pfitzner J, Battistutta D, O'Rourke PK, Green AC, Martin NG: Reliability of computer image analysis of pigmented skin lesions of Australian adolescents. *Cancer* 1996; **78**: 252–257.
- McGregor B, Pfitzner J, Zhu G *et al*: Genetic and environmental contributions to size, color, shape, and other characteristics of melanocytic naevi in a sample of adolescent twins. *Genet Epidemiol* 1999; **16**: 40–53.
- Zhu G, Duffy DL, Eldridge A *et al*: A major quantitative-trait locus for mole density is linked to the familial melanoma gene CDKN2A: a maximum-likelihood combined linkage and association analysis in twins and their sibs. *Am J Hum Genet* 1999; **65**: 483–492.
- Miller SA, Dykes DD, Polesky HF: A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988; **16**: 1215.
- Zhu G, Evans DM, Duffy DL *et al*: A genome scan for eye colour in 539 twin families. *Twin Res* 2004; **7**: 197–210.
- Amos CI: Robust variance-components approach for assessing genetic linkage in pedigrees. *Am J Hum Genet* 1994; **54**: 535–543.
- Almasy L, Blangero J: Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet* 1998; **62**: 1198–1211.
- Fulker DW, Cherny SS: An improved multipoint sib-pair analysis of quantitative traits. *Behav Genet* 1996; **26**: 527–532.
- Hopper JL, Matthews JD: Extensions to multivariate normal models for pedigree analysis. *Ann Hum Genet* 1982; **46**: 373–383.
- Williams JT, Begleiter H, Porjesz B *et al*: Joint multipoint linkage analysis of multivariate qualitative and quantitative traits. II. Alcoholism and event-related potentials. *Am J Hum Genet* 1999; **65**: 1148–1160.
- Jacquard A: Genetic information given by a relative. *Biometrics* 1972; **28**: 1101–1114.
- Lange K, Westlake J, Spence MA: Extensions to pedigree analysis. III. Variance components by the scoring method. *Ann Hum Genet* 1976; **39**: 485–491.
- Abecasis GR, Cherny SS, Cookson WO, Cardon LR: Merlin – rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002; **30**: 97–101.
- Dot D, Miro J, Fuentes-Arderiu X: Within-subject biological variation of hematological quantities and analytical goals. *Arch Pathol Lab Med* 1992; **116**: 825–826.
- Self SG, Liang K: Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *J Am Stat Assoc* 1987; **82**: 605–610.
- Williams JT, Blangero J: Power of variance component linkage analysis to detect quantitative trait loci. *Ann Hum Genet* 1999; **63**: 545–563.
- Allison DB, Neale MC, Zannolli R, Schork NJ, Amos CI, Blangero J: Testing the robustness of the likelihood-ratio test in a variance-component quantitative-trait loci-mapping procedure. *Am J Hum Genet* 1999; **65**: 531–544.
- Blangero J, Williams JT, Almasy L: Robust LOD scores for variance component-based linkage analysis. *Genet Epidemiol* 2000; **19** (Suppl 1): S8–S14.
- Amos C, de Andrade M, Zhu D: Comparison of multivariate tests for genetic linkage. *Hum Hered* 2001; **51**: 133–144.
- Evans DM: The Genetics of Blood Cell Concentrations, Unpublished PhD thesis, University of Queensland, Brisbane, 2003.
- Neale MC: Statistical modelling. Richmond, VA: Department of Psychiatry, Medical College of Virginia, 1997.
- Kato K, Kanaji T, Russell S *et al*: The contribution of glycoprotein VI to stable platelet adhesion and thrombus formation illustrated by targeted gene deletion. *Blood* 2003; **102**: 1701–1707.
- Joutsu-Korhonen L, Smethurst PA, Rankin A *et al*: The low-frequency allele of the platelet collagen signaling receptor glycoprotein VI is associated with reduced functional responses and expression. *Blood* 2003; **101**: 4372–4379.
- Nurden P, Jandrot-Perrus M, Combrie R *et al*: Severe deficiency of glycoprotein VI in a patient with gray platelet syndrome. *Blood* 2004; **2**: 813–819.
- Kunicki TJ: The influence of platelet collagen receptor polymorphisms in hemostasis and thrombotic disease. *Arterioscler Thromb Vasc Biol* 2002; **22**: 14–20.
- Ballmaier M, Germeshausen M, Schulze H *et al*: c-mpl mutations are the cause of congenital amegakaryocytic thrombocytopenia. *Blood* 2001; **97**: 139–146.
- Wiestner A, Schlemper RJ, van der Maas AP, Skoda RC: An activating splice donor mutation in the thrombopoietin gene causes hereditary thrombocythaemia. *Nat Genet* 1998; **18**: 49–52.
- Evans DM: The power of multivariate quantitative-trait loci linkage analysis is influenced by the correlation between variables. *Am J Hum Genet* 2002; **70**: 1599–1602.
- Buckley MF, Cheung C, Martin I *et al*: Quantitative trait loci for steady state platelet count in mice. *Am J Hum Genet* 2003; **73**: A1918:1496.