Genes and Immunity (2004) 5, 548–552

www.nature.com/gene

FULL PAPER

A major quantitative trait locus for CD4–CD8 ratio is located on chromosome 11

DM Evans[1,2], G Zhu[1], DL Duffy[1], IH Frazer[3], GW Montgomery[1] and NG Martin[1]
[1]Queensland Institute of Medical Research, Brisbane, Australia; [2]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; [3]Centre for Immunology and Cancer Research, University of Queensland, Princess Alexandra Hospital, Brisbane, Australia

CD4–CD8 ratio is an important diagnostic measure of immune system functioning. In particular, CD4–CD8 ratio predicts the time taken for progression of HIV infection to acquired immune deficiency syndrome (AIDS) and the long-term survival of AIDS patients. To map genes that regulate differences between healthy individuals in CD4–CD8 ratio, we typed 757 highly polymorphic microsatellite markers at an average spacing of ~5 cM across the genome in 405 pairs of dizygotic twins at ages 12, 14 and 16. We used multipoint variance components linkage analysis to test for linkage between marker loci and CD4–CD8 ratio at each age. We found suggestive evidence of linkage on chromosome 11p in 12-year-old twins (LOD = 2.55, P = 0.00031) and even stronger evidence of linkage in the same region at age 14 (LOD = 3.51, P = 0.00003). Possible candidate genes include CD5 and CD6, which encode cell membrane proteins involved in the positive selection of thymocytes. We also found suggestive evidence of linkage at other areas of the genome including regions on chromosomes 1, 3, 4, 5, 6, 12, 13, 15, 17 and 22.
Genes and Immunity (2004) **5,** 548–552. doi:10.1038/sj.gene.6364126
Published online 12 August 2004

Keywords: *CD4–CD8 ratio; linkage; genetic mapping; QTL*

Introduction

The ratio of CD4+ to CD8+ T lymphocytes (CD4–CD8 ratio) is an important diagnostic marker of immune system functioning. The majority of healthy individuals display a CD4–CD8 ratio in the range of 1.5–2.5 to 1,[1] whereas an inverted ratio is characteristic of intense chronic immune responses such as in graft *vs* host disease and also in several viral illnesses most notably HIV infection.[2] Importantly, CD4–CD8 ratio predicts the time taken for progression of HIV infection to acquired immune deficiency syndrome (AIDS) and the long-term survival of AIDS patients.[3–5]

Despite its clinical utility, there exists large interindividual variation in the CD4–CD8 ratio of normal individuals. In fact, approximately 5% of otherwise healthy individuals exhibit an inverted ratio. In a previous study of adolescent twins, we demonstrated that CD4–CD8 ratio was highly heritable with genetic factors responsible for 84% of variation in the phenotype.[6] Studies in the mouse[7] and humans[1] have also suggested that the trait is under the control of a major gene. We therefore consider CD4–CD8 ratio to be an excellent trait for genetic mapping.

Correspondence: Dr DM Evans, The Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK. E-mail: davide@well.ox.ac.uk
Received 11 May 2004; revised 22 June 2004; accepted 23 June 2004; published online 12 August 2004

In this paper, we extend the results of the Evans *et al*.[6] twin study by performing a genome-wide linkage scan of the dizygotic (DZ) twin pairs and their families. We measured twins' CD4–CD8 ratio longitudinally at 12, 14 and 16 years and performed genome-wide variance components QTL linkage analysis. We hope that the results from this study will constitute the first step in the eventual positional cloning of a gene involved in the regulation of CD4–CD8 ratio. Such knowledge will not only yield important insights into the homeostatic control of T-cell populations but may also prove important in developing strategies for immune cell reconstitution in HIV infection and other pathological states where the ratio is inverted.

Results

The results from the genome-wide tests of linkage are displayed in Figure 1. The most prominent result was the large linkage peaks at ages 12 (LOD = 2.55, $P = 0.00031$) and 14 (LOD = 3.51, $P = 0.00003$) in the same region of chromosome 11p. Interestingly, there was little evidence for linkage in this region at age 16. Another large peak was present on chromosome 4 at age 14 (LOD = 2.21, $P = 0.00072$) and several smaller peaks (ie LOD scores >1.0) were also observed on chromosomes 1, 3, 4, 5, 6, 12, 13, 15, 17 and 22 (see Table 1). Interestingly, the peaks on chromosomes 1p (age 12: LOD = 1.50, $P = 0.0043$; age 14: LOD = 1.55, $P = 0.00379$) and 5p (age 12: LOD = 1.81, $P = 0.0020$; age 14: LOD = 1.08, $P = 0.013$) showed evidence of linkage at more than one age.

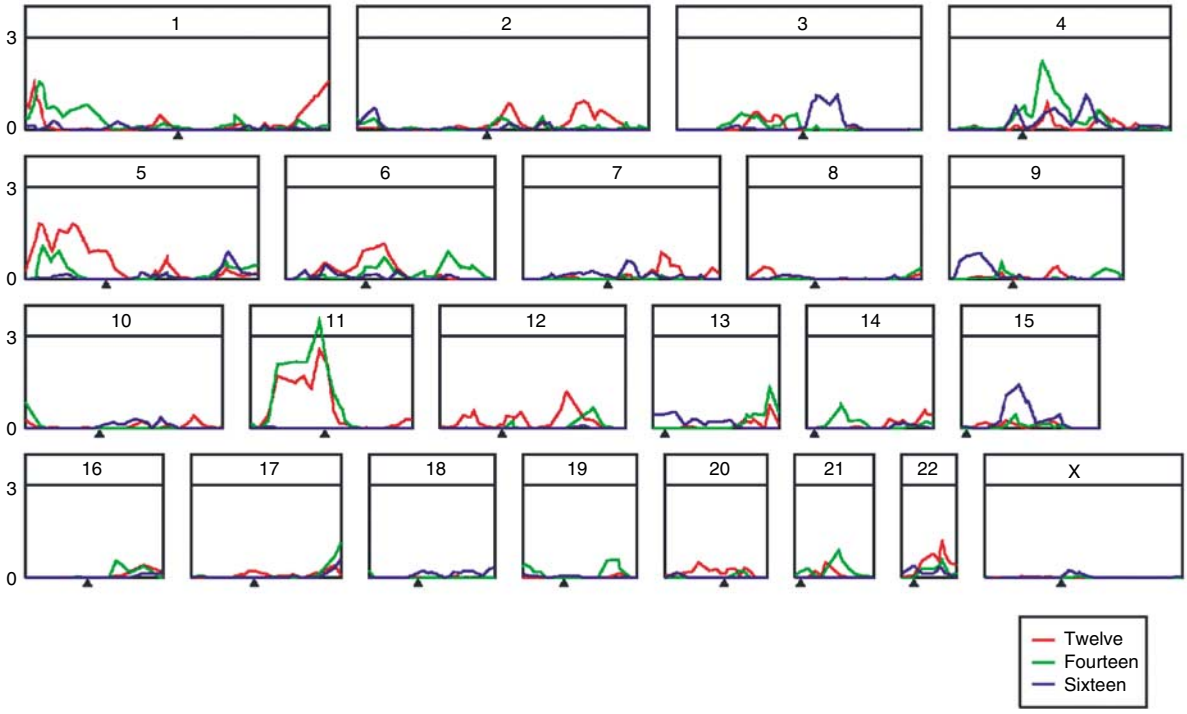

Figure 1 Genome scan results for CD4/CD8 ratio.

Table 1 Regions with LOD >1 for the analysis of CD4–CD8 ratio

Chromosome	*Age*	*Region*	*cM range*	*Peak LOD*
1	12	D1S214 – D1S1612	12.15 – 13.82	1.50
1	14	D1S1612 – D1S2667	13.82 – 19.88	1.55
1	12	D1S547 – D1S2836	259.16 – 271.84	1.60
3	16	D3S2460 – D3S1764	126.98 – 145.53	1.11
4	14	D4S2964 – D4S1647	87.95 – 103.77	2.21
4	16	D4S2394	128.08	1.12
5	12	D5S2849 – D5S2505	9.153 – 15.63	1.81
5	14	D5S2505	15.63	1.08
5	12	D5S817 – D5S1470	29.91 – 53.87	1.80
6	12	D6S1031 – D6S462	89.14 – 96.98	1.15
11	12	D11S1981 – D11S987	25.59 – 72.17	2.55
11	14	D11S1981 – D11S2371	25.59 – 79.66	3.51
12	12	PAH – D12S78	115.03 – 116.07	1.19
13	14	D13S1265	114.461	1.33
15	16	D15S978 – D15S643	47.63 – 59.51	1.42
17	14	D17S928	135.67	1.15
22	12	D22S683 – D22S283	41.73 – 42.07	1.20

Discussion

This is one of the first studies to have used a complete genome scan in order to map genes responsible for variation in CD4–CD8 ratio. Studies in mice[7] and humans[1] have indicated that CD4–CD8 ratio is under the control of a major gene. Our results suggest that such a gene may be located on chromosome 11p. In this regard, a major strength of our study has been that evidence for linkage was present in the same area of chromosome 11 at more than one age (ie at 12 and 14). While repeated measures do not formally constitute replication (since longitudinal data are not independent), they do indicate that the present results are robust with respect to measurement error and temporal changes in the phenotype.

Two promising candidates in the chromosome 11p region are the genes for CD5 and CD6– homologous genes, which have been implicated in the positive selection of thymocytes. Positive selection is the process by which double-positive thymocytes with T-cell receptors (TCRs) able to recognize MHC molecules on antigen-presenting cells are given survival signals and are able to differentiate further. These double-positive thymocytes subsequently differentiate to become single-positive thymocytes committed to either the CD4+ or

Table 2 Breakdown of participation data showing the number of complete twin pairs for whom platelet count and genotyping information were available

12	*14*	*16*	*DZF*	*DZM*	*DZFM*	*DZMF*	*Total*
√			12	15	11	10	48
	√		4	4	6	6	20
		√	14	10	20	17	61
√	√		28	35	21	36	120
√		√	5	1	3	6	15
	√	√	5	7	7	6	25
√	√	√	32	36	25	23	116
Total			100	108	93	104	405
Total at 12			77	87	60	75	299
Total at 14			69	82	59	71	281
Total at 16			56	54	55	52	217

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

CD8 + lineage. CD5 expression on the thymocyte surface increases throughout development and is upregulated at critical points through engagement of the pre-TCR and TCR.[8] The level of CD5 expression is proportional to the intensity of the pre-TCR/TCR signaling, with high avidity bindings associated with upregulated expression.[8] Engagement of thymocytes with anti-CD3 and anti-CD5 monoclonal antibodies reduces calcium influx into the cell,[9] whereas CD5 deficiency results in hyperreponsiveness to TCR stimulation.[10] These observations suggest that CD5 may play a role in attenuating the response to TCR signaling after initial stimulation, which may be important in ensuring thymocyte survival.[8] CD5 may also play a role in regulating lineage commitment, although its exact role remains unclear. For example, coligation of CD5 and CD3 *in vitro* enhances differentiation of double-positive thymocytes into the CD4 lineage and induction of the antiapoptotic protein Bcl-2.[9]

The other candidate gene CD6 is expressed primarily on the surface of cells of the T-lymphocyte lineage. CD6 has been well characterized as an accessory receptor able to provide costimulatory signals, which, along with signals through the TCR, lead to cell proliferation.[11] However, the observation that the ligand for CD6 (CD166) is present on the thymic epithelium has led to the suggestion that CD6 might play a role in the selection of thymocytes. Consistent with this hypothesis, expression of CD6 increases in human and mouse thymocytes that express heterogeneous TCR as double-positive thymocytes are selected into the single-positive stage.[12] Although its precise role is yet to be determined, it is interesting that there exists a close association between the level of CD4 and CD6 in CD4 single-positive thymocytes. It has therefore been suggested that interactions between CD6 and the thymic epithelium may contribute preferentially to positive selection of CD4 single-positive thymocytes.[12]

It was interesting that there was no evidence of linkage at age 16 in the chromosome 11p region, particularly since the correlation between CD4–CD8 ratio at different measurement occasions was high ($r_{12\&14} = 0.76$; $r_{14\&16} = 0.69$; $r_{12\&16} = 0.77$). While it is possible that different genes are responsible for variation in CD4/CD8 ratio at age 16, we consider this unlikely since preliminary analyses of this dataset have indicated that the majority of genetic variance from ages 12 and 14 was transmitted to age 16.[13] One explanation is simply that as there were considerably fewer DZ twins in the analyses at age 16, the power to detect linkage was not as great as at ages 12 or 14 (Table 2). Another possibility is the existence of a potent environmental agent at age 16, which might have masked evidence for linkage at this age, particularly if (by chance) the effect of the innovations was most pronounced in pairs who shared zero or both alleles identical by descent (IBD) at the relevant marker loci. For example, an environmental agent such as the Epstein–Barr virus could have produced such an effect. Finally, the positive results at 12 and 14 might reflect type I error.

Several other smaller linkage peaks (ie LOD >1.0) were identified in this study, including peaks on chromosomes 1, 3, 4, 5, 6, 12, 13, 15, 17 and 22. Curiously, save in the case of chromosomes 1p and 5, there was little evidence of linkage in these regions at other ages. One possibility is that different genes may regulate CD4–CD8 ratio at different ages. However, we consider this explanation unlikely given that the majority of the genetic variance was transmitted through the time series, and that the role of new genetic innovation was minimal.[13] 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 (and their absence) will simply be a result of random fluctuation and type I error.

A recent genome-wide study involving 15 large pedigrees identified two regions of suggestive linkage between CD4–CD8 ratio and chromosome 4.[14] However, the peaks in the Hall *et al* study were located between 2 and 5 cM and from 53 and 79 cM, whereas the two peaks in the present study were located between 89 and 104 cM and at 128 cM (see Table 1). Several simulation studies have demonstrated that the estimated location of a QTL can in fact be tens of centimorgans away from its actual location.[15] It is therefore unclear whether the peak in 53 – 79 cM region from the Hall *et al* study and our peak in the

89 – 104 cM region reflect the same underlying QTL. The key to assessing the significance of these and the other linkages from this study will be replication using independent data sets.

In conclusion, we have identified an area on chromosome 11p, which shows linkage to CD4–CD8 ratio at 12 and 14 years of age. We intend to fine map this region (ie increasing marker density) and perform association analyses of candidate genes in the area (eg CD5 and CD6) to better localize the genomic region involved. We hope that this study represents the first step in the eventual identification and subsequent cloning of a gene involved in regulating CD4–CD8 ratio.

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.[16–18] Twins were enlisted by contacting the principals of primary schools in the greater Brisbane area, 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 CD4–CD8 ratio.[6] The protocol was approved by the Queensland Institute of Medical Research Human Research Ethics Committee and informed consent was obtained from all participants and parents prior to testing.

Venous blood samples were collected using EDTA tubes. A lymphocyte subset analysis was performed on whole blood using the AutoPrep (Coulter, Hileah, Fl, USA), and direct fluorochrome conjugated monoclonal antibodies to CD3, CD4, and CD8 antigens (Coulter). Analysis was performed on an Epics 753 cytofluorograph (Coulter) using standardized control samples and machine settings. The absolute numbers of CD3+ T lymphocytes (CD3), CD4+ T lymphocytes (CD4) and CD8+ T lymphocytes (CD8) were calculated as percentages of the total lymphocyte count. It was the absolute numbers that were subsequently used for calculation of the CD4/CD8 ratio. Blood was collected from twins longitudinally at 12, 14 and 16 years of age. Where possible, DNA was also obtained from parents and siblings for genotyping (CD4–CD8 ratio was not measured). No attempt was made to exclude subjects suffering from illness.

CD4–CD8 ratio was available from 405 complete twin pairs comprising 100 DZ female, 108 DZ male and 197 opposite sex (OS) twin pairs (consisting of 104 pairs where the female was born first and 93 pairs where the male was born first). Although twins were tested as close as possible to their 12, 14 and 16 birthdays, not all twins were tested at all three measurement occasions (see Table 2 for a breakdown of these data).

Genotypes

DNA was extracted from buffy coats using a modification of the 'salt method'.[19] 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 Plus 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 (55% of families had both parents typed, 14% had neither parent typed and 31% had one parent typed). The average heterozygosity of markers was 0.78, and the mean information content was 0.77. Although genome scan data were available from parents, twins and siblings, phenotype data (ie blood cell counts) were only available from twins. Full details of the genome scan are provided elsewhere.[20]

Linkage analyses

Univariate multipoint variance components linkage analysis was used to test for linkage between marker loci and blood cell phenotypes.[21–24] Variance components were estimated by maximum-likelihood analysis of the raw data[25] as implemented in the software package MERLIN[26] along with fixed effects for sex and age. Since both circadian and seasonal effects have been reported for lymphocytes, linear, quadratic and sinusoidal fixed effects were included for the time of day and month from which blood was sampled.[27] 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 IBD across the genome, and are thus uninformative for linkage. Note also that although CD4–CD8 ratio was only measured in twins, parental and sibling genotypes still helped determine IBD sharing between the DZ twin pairs.

The null hypothesis that additive genetic variance in CD4–CD8 ratio 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 σ_q^2 was constrained to zero with the likelihood of a model in which the genetic variance due to the QTL (σ_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 χ_1^2 and a point mass at zero,[28] while the difference between the two $\log_{10}$ likelihoods produces a LOD score equivalent to the classical LOD score of parametric linkage analysis.[29]

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.[30] We therefore $\log_{10}$ transformed the CD4–CD8 ratio data. The distribution of the transformed data for 12-, 14- and 16-year-olds exhibited a kurtosis of −0.130, 0.098 and 0.073, respectively. Recent statistical genetics theory[31] 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.

Acknowledgements

We thank Ann Eldridge, Marlene Grace and Anjali Henders for assistance and the twins, their siblings and their parents for their cooperation. 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.

References

1 Amadori A, Zamarchi R, De Silvestro G *et al*. Genetic control of the CD4/CD8 T-cell ratio in humans [see comments]. *Nat Med* 1995; **1**: 1279–1283.
2 Pedersen C, Dickmeiss E, Gaub J *et al*. T-cell subset alterations and lymphocyte responsiveness to mitogens and antigen during severe primary infection with HIV: a case series of seven consecutive HIV seroconverters. *Aids* 1990; **4**: 523–526.
3 Baillargeon J, Grady J, Borucki MJ. Immunological predictors of HIV-related survival. *Int J STD AIDS* 1999; **10**: 467–470.
4 Drabick JJ, Williams WJ, Tang DB, Sun W, Chung RC. CD4 lymphocyte decline and survival in human immunodeficiency virus infection. The Military Medical Consortium for Applied Retroviral Research. *AIDS Res Hum Retroviruses* 1992; **8**: 2039–2047.
5 Taylor JM, Fahey JL, Detels R, Giorgi JV. CD4 percentage, CD4 number, and CD4:CD8 ratio in HIV infection: which to choose and how to use. *J Acquir Immune Defic Syndr* 1989; **2**: 114–124.
6 Evans DM, Frazer IH, Martin NG. Genetic and environmental causes of variation in basal levels of blood cells. *Twin Res* 1999; **2**: 250–257.
7 Kraal G, Weissman IL, Butcher EC. Genetic control of T-cell subset representation in inbred mice. *Immunogenetics* 1983; **18**: 585–592.
8 Azzam HS, Grinberg A, Lui K, Shen H, Shores EW, Love PE. CD5 expression is developmentally regulated by T cell receptor (TCR) signals and TCR avidity. *J Exp Med* 1998; **188**: 2301–2311.
9 Zhou XY, Yashiro-Ohtani Y, Toyo-Oka K *et al*. CD5 costimulation up-regulates the signaling to extracellular signal-regulated kinase activation in CD4+CD8+ thymocytes and supports their differentiation to the CD4 lineage. *J Immunol* 2000; **164**: 1260–1268.
10 Tarakhovsky A, Kanner SB, Hombach J *et al*. A role for CD5 in TCR-mediated signal transduction and thymocyte selection. *Science* 1995; **269**: 535–537.
11 Osorio LM, Rottenberg M, Jondal M, Chow SC. Simultaneous cross-linking of CD6 and CD28 induces cell proliferation in resting T cells. *Immunology* 1998; **93**: 358–365.
12 Singer NG, Fox DA, Haqqi TM *et al*. CD6: expression during development, apoptosis and selection of human and mouse thymocytes. *Int Immunol* 2002; **14**: 585–597.
13 Evans DM. The genetics of blood cell concentrations. Unpublished PhD thesis, University of Queensland, Brisbane, Australia, 2003.
14 Hall MA, Norman PJ, Thiel B *et al*. Quantitative trait loci on chromosomes 1, 2, 3, 4, 8, 9, 11, 12, and 18 control variation in levels of T and B lymphocyte subpopulations. *Am J Hum Genet* 2002; **70**: 1172–1182.
15 Roberts SB, MacLean CJ, Neale MC, Eaves LJ, Kendler KS. Replication of linkage studies of complex traits: an examination of variation in location estimates. *Am J Hum Genet* 1999; **65**: 876–884.
16 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.
17 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.
18 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.
19 Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988; **16**: 1215.
20 Zhu G, Evans DM, Duffy DL *et al*. A genome scan for eye colour in 539 twin families. *Twin Res* 2004; **7**: 197–210.
21 Amos CI. Robust variance-components approach for assessing genetic linkage in pedigrees. *Am J Hum Genet* 1994; **54**: 535–543.
22 Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet* 1998; **62**: 1198–1211.
23 Eaves LJ, Neale MC, Maes H. Multivariate multipoint linkage analysis of quantitative trait loci. *Behav Genet* 1996; **26**: 519–525.
24 Fulker DW, Cherny SS. An improved multipoint sib-pair analysis of quantitative traits. *Behav Genet* 1996; **26**: 527–532.
25 Lange K, Westlake J, Spence MA. Extensions to pedigree analysis. III. Variance components by the scoring method. *Ann Hum Genet* 1976; **39**: 485–491.
26 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.
27 Dot D, Miro J, Fuentes-Arderiu X. Biological variation of the leukocyte differential count quantities. *Scand J Clin Lab Invest* 1992; **52**: 607–611.
28 Self SG, Liang K. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *J Amer Stat Assoc* 1987; **82**: 605–610.
29 Williams JT, Blangero J. Power of variance component linkage analysis to detect quantitative trait loci. *Ann Hum Genet* 1999; **63**: 545–563.
30 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.
31 Blangero J, Williams JT, Almasy L. Robust LOD scores for variance component-based linkage analysis. *Genet Epidemiol* 2000; **19** (Suppl 1): S8–S14.