Major quantitative trait locus for eosinophil count is located on chromosome 2q

David M. Evans, PhD,[a,b] Gu Zhu, MD,[a] David L. Duffy, MBBS, PhD,[a] Grant W. Montgomery, PhD,[a] Ian H. Frazer, MD,[c] and Nicholas G. Martin, PhD[a]

Brisbane, Australia, and Oxford, United Kingdom

Background: Eosinophils are granulocytic white blood cells implicated in asthma and atopic disease. The degree of eosinophilia in the blood of patients with asthma correlates with the severity of asthmatic symptoms. Quantitative trait loci (QTL) linkage analysis of eosinophil count may be a more powerful strategy of mapping genes involved in asthma than linkage analysis using affected relative pairs.

Objective: To identify QTLs responsible for variation in eosinophil count in adolescent twins.

Methods: We measured eosinophil count longitudinally in 738 pairs of twins at 12, 14, and 16 years of age. We typed 757 highly polymorphic microsatellite markers at an average spacing of ~5 centimorgans across the genome. We then used multipoint variance components linkage analysis to test for linkage between marker loci and eosinophil concentrations at each age across the genome.

Results: We found highly significant linkage on chromosome 2q33 in 12-year-old twins (logarithm of the odds = 4.6; *P* = .000002) and suggestive evidence of linkage in the same region in 14-year-olds (logarithm of the odds = 1.0; *P* = .016). We also found suggestive evidence of linkage at other areas of the genome, including regions on chromosomes 2, 3, 4, 8, 9, 11, 12, 17, 20, and 22.

Conclusion: A QTL for eosinophil count is present on chromosome 2q33. This QTL might represent a gene involved in asthma pathophysiology. (J Allergy Clin Immunol 2004;114:826-30.)

Key words: *Eosinophils, asthma, genetics, linkage, QTL, twins*

Abbreviations used
cM: Centimorgan
IBD: Identical by descent
LOD: Logarithm of the odds
QTL: Quantitative trait locus

Nonparametric linkage analysis of affected relative pairs is a popular strategy used in the mapping of disease-causing genes. The rationale behind the approach is that if a trait and marker are linked, affected relatives should share more marker alleles identical by descent (IBD) than expected by chance.[1] However, large numbers of affected relatives need to be genotyped to detect loci of even moderate effect with appreciable power.[1] This has led to the suggestion that a more promising strategy might be to analyze continuously distributed traits associated with the disease in question (endophenotypes), because quantitative measures contain more information than dichotomous traits and hence may provide more power to detect genes of small effect.

In the case of asthma, 1 such quantitative phenotype is the concentration of eosinophils in a given volume of blood (eosinophil count). Eosinophils are the primary mediators of airway inflammation and bronchial damage in asthma. Circulating eosinophilia is commonly observed in asthma, with the concentration of eosinophils in the airways and the blood correlating with the severity of asthmatic symptoms.[2,3] In addition, segregation analyses have suggested the presence of a major gene influencing eosinophil count.[4] We therefore believe that linkage analysis of eosinophil count may be a powerful strategy in the identification of genes involved in asthma pathophysiology.

Compared with the large number of studies concerning the genetics of asthma, surprisingly few studies have investigated the genetics of eosinophil count. A few studies have found evidence of linkage in the 5q31-33 region, which contains the genes for several growth factors involved in eosinophil production.[5-7] These results are consistent with the results from linkage studies involving other asthma-related phenotypes including serum IgE,[8] bronchial hyperresponsiveness,[9] and severity of parasitic infection.[10,11] Other genome-wide linkage studies involving eosinophil count have implicated regions on chromosomes 6,[12,13] 7,[12] and 12.[14]

In a recent study of Australian adolescent twins, we demonstrated that approximately 60% of the phenotypic

From [a]the Queensland Institute of Medical Research and Joint Genetics Program, University of Queensland, Brisbane; [b]the Wellcome Trust Centre for Human Genetics, Oxford; and [c]the Centre for Immunology and Cancer Research, University of Queensland, Princess Alexandra Hospital, Brisbane.

Funding source: 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 National Health and Medical Research Council'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.

Received for publication February 25, 2004; revised May 21, 2004; accepted for publication May 24, 2004.

Available online August 9, 2004.

Reprint requests: David M. Evans, PhD, Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford, OX3 7BN, United Kingdom. E-mail: davide@well.ox.ac.uk.

0091-6749/$30.00

© 2004 American Academy of Allergy, Asthma and Immunology

doi:10.1016/j.jaci.2004.05.060

TABLE I. Breakdown of participation data showing the number of complete twin pairs for whom eosinophil count and genotyping information (in the case of dizygotic pairs) were available

12 y	14 y	16 y	MZF	MZM	DZF	DZM	DZFM	DZMF	Total
√			19	30	13	12	11	9	94
	√		9	4	3	5	5	4	30
		√	50	57	16	14	21	22	180
√	√		31	25	28	37	22	40	183
√		√	3	2	4	2	2	7	20
	√	√	11	18	8	7	10	7	61
√	√	√	37	27	29	30	24	23	170
Total			160	163	101	107	95	112	738
Total at 12 y			90	84	74	81	59	79	467
Total at 14 y			88	74	68	79	61	74	444
Total at 16 y			101	104	57	53	57	59	431

MZF, Monozygotic female twins; *MZM*, monozygotic male twins; *DZF*, dizygotic female twins; *DZM*, dizygotic male twins; *DZFM*, opposite sex twins with the female subject born first; *DZMF*, opposite sex twins with the male subject born first.

variance in eosinophil count was a result of genetic factors.[15] In the current article, we extend these results by performing a genome-wide linkage scan of the dizygotic twin pairs and their families from that study. We hope that the identification of quantitative trait loci (QTLs) underlying variation in eosinophil count will yield important insights into the pathophysiology of asthma and other atopic conditions.

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, by media appeals, and by word of mouth. It is estimated that approximately 50% of the eligible birth cohort was recruited into the study and was typical of the population with respect to eosinophil concentrations.[15] Informed consent was obtained from all participants and parents before testing.

Venous blood samples were collected by using EDTA tubes, and eosinophil count was measured by using a Coulter model STKS blood counter (Beckman Coulter, Fullerton, Calif). 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 (eosinophil count was not measured). No attempt was made to exclude subjects with illness.

Phenotypic and genotypic data (in the case of dizygotic twins) were available from 738 twin pairs composed of 160 monozygotic female, 163 monozygotic male, 101 dizygotic female, 107 dizygotic male, and 207 opposite sex twin pairs (95 pairs in which the female subject was born first and 112 pairs in which the male subject 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 3 measurement occasions (see Table I for a breakdown of these data). To normalize the distribution, eosinophil concentration was transformed by using the calculation $\log_{10}(X + 0.05)$. Descriptive statistics and heritabilities have been reported elsewhere.[15]

Genotypes

DNA was extracted from buffy coats by using a modification of the salt method.[19] For same-sex twin pairs, zygosity was determined by typing 9 independent DNA microsatellite polymorphisms plus the X/Y amelogenin marker for sex determination by PCR (ABI Profiler system; Applied Biosystems, Foster City, Calif). All twins were also typed for ABO, Rh, and MNS blood groups.

The genome scan consisted of 726 highly polymorphic autosomal and 31 X-linked microsatellite markers at an average spacing of ~5 centimorgans (cM) in 539 families (2360 individuals). The microsatellites consisted of a combination of markers from the ABI-Prism and Center for Inherited Disease Research genotyping sets. Overlapping parts of the sample received a 10-cM scan by using the ABI-2 marker set (400 markers) at the Australian Genome Research Facility (Melbourne), a 10-cM scan by using the Weber marker set at the Center for Inherited Disease Research (Baltimore, Md), 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 1 scan had both parents genotyped and so had high information content. 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, eosinophil counts) were available only from twins. Full details of the genome scan have been published 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] In the variance components method, the expected phenotypic covariance matrix is partitioned into the following components:

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

where σ_q^2 is the additive genetic variance caused by the QTL, σ_a^2 is the (residual) polygenic additive genetic variance, and σ_e^2 is the unique environmental variance.[21,25] $\hat{\mathbf{\Pi}}$ is a matrix containing the elements that denote the estimated proportion of alleles shared IBD at the trait locus by 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.[21,26] Φ is the kinship matrix containing elements ϕ_{ij}—the kinship coefficients between individuals i and j[27]—and $\mathbf{I_n}$ is an identity matrix of order n. The variance components were estimated by maximum likelihood analysis of the raw data[28] as implemented in the software package MERLIN[29] (http://csg.sph.umich.edu/pn/index.php?furl=/abecasis/Merlin/index.html) along with fixed effects for sex and age, as well as time of day and month when blood was sampled (quadratic and sinusoidal terms were also estimated for these

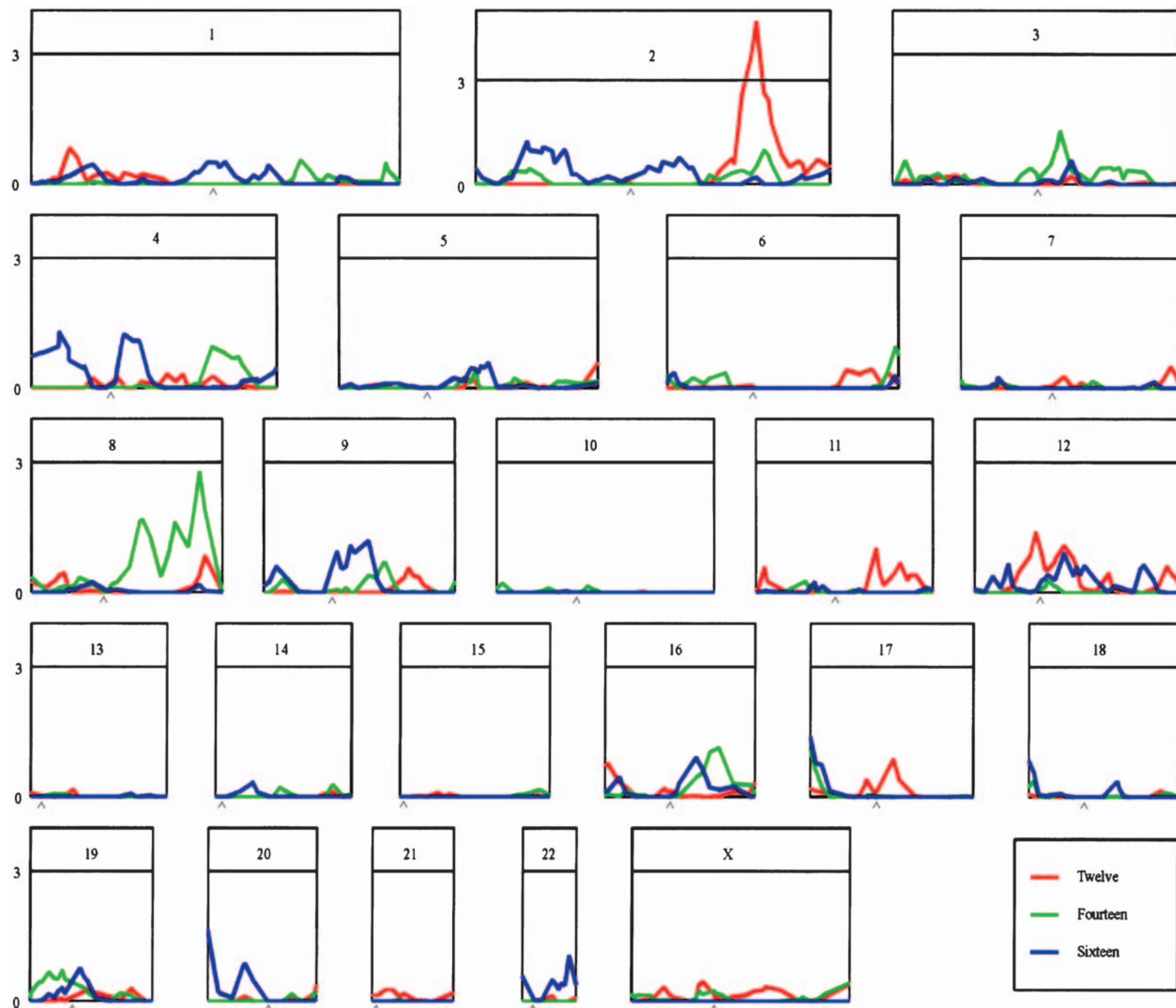

FIG 1. Genome scan results for eosinophil count.

last 2 covariates). Univariate analyses were performed at each marker, for each variable, at each age.

The null hypothesis that additive genetic variance caused by a QTL linked to a marker for a given phenotype was 0 (ie, $\sigma_q^2 = 0$) was tested by comparing the likelihood of a restricted model in which σ_q^2 was constrained to 0 with the likelihood of a model in which the genetic variance caused by the QTL (σ_q^2) was estimated. Twice the difference in natural log-likelihoods between these models is distributed asymptotically as a ½:½ mixture of χ_1^2 and a point mass at 0.[30] The difference between the 2 $\log_{10}$ likelihoods produces a LOD score equivalent to the classic LOD score of parametric linkage analysis.[26]

Of course, only phenotypic data from dizygotic twins were included in the analysis, because monozygotic twins share all their genes IBD across the genome and are thus uninformative for linkage. Note that although parents and siblings were not phenotyped, their genotypes still contributed to IBD estimation in the dizygotic twin pairs.

RESULTS

The results of the univariate variance components tests for linkage are displayed in Fig 1. Each plot displays LOD scores for all 3 ages so that it is possible to compare the consistency of results across measurement occasions. The strongest evidence of linkage occurred at age 12 years on chromosome 2q (LOD = 4.70; P = .000002). There was also a smaller peak in the same location at age 14 years (LOD = 1.0; P = .016), but not at age 16 years, at which there were far fewer twins tested and hence less power. There were also several regions of suggestive linkage (ie, defined here as LOD scores >1.0), summarized in Table II. These included a region of chromosome 17p that was present at ages 14 and 16 years, as well as a largish peak on chromosome 8q (LOD = 2.78; P = .00017). There was little evidence of linkage in the chromosome 5q region.

DISCUSSION

The most promising result from this study is the large linkage peak on chromosome 2q33 for eosinophil count in 12-year-old twins. Several studies have reported

TABLE II. Regions with LOD >1 for the analysis of eosinophil count

Chromosome	Age, y	Region	cM range	Peak LOD
2	12	D2S117-D2S434	194.63-216.01	4.7
2	14	D2S2944	210.578	1.00
2	16	D2S1360	39.96	1.23
2	16	D2S405	51.48	1.07
2	16	D2S2259	67.06	1.00
3	14	D3S1267	130.014	1.24
4	16	D4S403	26.81	1.30
4	16	D4S2367-D4S1534	79.335-93.19	1.24
8	14	D8S1119-GAAT1A4	94.822-104.28	1.65
8	14	D8S514-D8S272	124.62-149.758	2.78
9	16	D9S922	77.77	1.08
9	16	D9S283	92.422	1.18
11	12	D11S898	103.59	1.01
12	12	D12S1042	50.66	1.39
12	12	D12S183	74.03	1.07
17	14	D17S849-D17S1308	0.63-1.386	1.1
17	16	D17S849-D17S1308	0.63-1.386	1.38
20	16	D20S103-D20S117	2.52-2.9	1.65
22	16	D2D22S423	49.14	1.05

suggestive linkages to asthma[31] and asthma-related phenotypes in the 2q33 region, including circulating IgE,[13] bronchial hyperresponsiveness,[14] and reaction to a skin prick test.[31] Several candidate genes are present within this region, including genes involved in T-cell signaling (*CD28* and *CTLA*-4) and apoptosis (*caspase 8*).

A major strength of the current study is the availability of multiple measurements on the same individuals across time. Although these results do not formally constitute replication (because the data are not independent), they do indicate the extent to which the results are robust with respect to measurement error and temporal changes in the phenotypes. In this vein, it is interesting to note that although there was evidence for linkage at 12 and 14 years on chromosome 2q33, there was little evidence of linkage at age 16 years. This is interesting given that the phenotypic correlation across measurement occasions is quite high ($r_{12 \text{ and } 14} = .63$; $r_{14 \text{ and } 16} = .56$; $r_{12 \text{ and } 16} = .56$), and previous longitudinal genetic analyses have suggested that the majority of genetic variance from ages 12 and 14 years is transmitted to age 16 years.[32] One explanation is that there were far fewer dizygotic twins available at age 16 years than at ages 12 or 14 years (Table I), and hence the power to detect linkage at this age was not as great. Another possibility is the existence of a potent environmental agent at age 16 years that influences eosinophil count. Such an agent might have masked evidence for linkage at this age particularly if (by chance) it affected pairs who shared 0 or both alleles IBD at the relevant marker loci.

Several other regions of suggestive linkage were identified in this study, including regions on chromosomes 2, 3, 4, 8, 9, 11, 12, 17, 20, and 22. Although some of these regions will harbor QTLs, other peaks will be the result of random fluctuation and type I error. In this respect, it is important to compare these results with those from previous genome scans. Several authors have reported linkage to eosinophil count in other regions of the genome, most notably in the 5q31 to 5q33 region[5-7] and regions on chromosomes 6,[13,33] 7,[33] and 12.[12-14] Unfortunately, there was little evidence for linkage to any of these regions in the current study. This lack of replication of course does not disprove these earlier results—there are several reasons why this might be the case (eg, racial differences)—but is most probably a result of lack of power. Even though the sample size used in this study was larger than in the majority of genome scans to date, it is still well short of the large sample sizes required to detect loci of even moderate effect with appreciable power.[34,35]

In conclusion, we found strong evidence for the existence of a QTL influencing eosinophil count on chromosome 2q33 in 12-year-old twins. We also found evidence of linkage in the same region at age 14 years. This area of chromosome 2 has previously been implicated in studies of asthma and asthma-related phenotypes. We therefore intend to fine-map this region (ie, increasing marker density) and perform association analyses of candidate genes in the area (eg, *CD28*, *CTLA-4*, *caspase 8*) to localize better the genomic region involved. We hope that this study represents the first step in the eventual identification of genes that will not only increase our understanding of eosinophil biology but also provide insight into the pathogenesis of asthma and other atopic conditions.

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

REFERENCES

1. Risch N. Linkage strategies for genetically complex traits, II: the power of affected relative pairs. Am J Hum Genet 1990;46:229-41.
2. Durham SR, Kay AB. Eosinophils, bronchial hyperreactivity and late-phase asthmatic reactions. Clin Allergy 1985;15:411-8.
3. Laitinen LA, Laitinen A, Heino M, Haahtela T. Eosinophilic airway inflammation during exacerbation of asthma and its treatment with inhaled corticosteroid. Am Rev Respir Dis 1991;143:423-7.
4. Palmer LJ, Cookson WO, James AL, Musk AW, Burton PR. Gibbs sampling-based segregation analysis of asthma-associated quantitative traits in a population-based sample of nuclear families. Genet Epidemiol 2001;20:356-72.
5. Martinez FD, Solomon S, Holberg CJ, Graves PE, Baldini M, Erickson RP. Linkage of circulating eosinophils to markers on chromosome 5q. Am J Respir Crit Care Med 1998;158:1739-44.
6. Palmer LJ, Daniels SE, Rye PJ, Gibson NA, Tay GK, Cookson WO, et al. Linkage of chromosome 5q and 11q gene markers to asthma-associated quantitative traits in Australian children. Am J Respir Crit Care Med 1998;158:1825-30.
7. Rioux JD, Stone VA, Daly MJ, Cargill M, Green T, Nguyen H, et al. Familial eosinophilia maps to the cytokine gene cluster on human chromosomal region 5q31-q33. Am J Hum Genet 1998;63:1086-94.
8. Meyers DA, Postma DS, Panhuysen CI, Xu J, Amelung PJ, Levitt RC, et al. Evidence for a locus regulating total serum IgE levels mapping to chromosome 5. Genomics 1994;23:464-70.
9. Postma DS, Bleecker ER, Amelung PJ, Holroyd KJ, Xu J, Panhuysen CI, et al. Genetic susceptibility to asthma—bronchial hyperresponsiveness coinherited with a major gene for atopy. N Engl J Med 1995;333:894-900.
10. Marquet S, Abel L, Hillaire D, Dessein H, Kalil J, Feingold J, et al. Genetic localization of a locus controlling the intensity of infection by

Mechanisms of asthma and allergic inflammation

Schistosoma mansoni on chromosome 5q31-q33. Nat Genet 1996;14:181-4.
11. Marquet S, Abel L, Hillaire D, Dessein A. Full results of the genome-wide scan which localises a locus controlling the intensity of infection by Schistosoma mansoni on chromosome 5q31-q33. Eur J Hum Genet 1999;7:88-97.
12. Daniels SE, Bhattacharrya S, James A, Leaves NI, Young A, Hill MR, et al. A genome-wide search for quantitative trait loci underlying asthma. Nature 1996;383:247-50.
13. Wjst M, Fischer G, Immervoll T, Jung M, Saar K, Rueschendorf F, et al. A genome-wide search for linkage to asthma. German Asthma Genetics Group. Genomics 1999;58:1-8.
14. Dizier MH, Besse-Schmittler C, Guilloud-Bataille M, Annesi-Maesano I, Boussaha M, Bousquet J, et al. Genome screen for asthma and related phenotypes in the French EGEA study. Am J Respir Crit Care Med 2000;162:1812-8.
15. Evans DM, Frazer IH, Martin NG. Genetic and environmental causes of variation in basal levels of blood cells. Twin Res 1999;2:250-7.
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;78:252-7.
17. McGregor B, Pfitzner J, Zhu G, Grace M, Eldridge A, Pearson J, 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, Grace M, Mayne C, O'Gorman L, 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;65:483-92.
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, Montgomery GW, Medland SE, Gillespie NA, et al. A genome scan for eye colour in 502 twin families: most variation is due to a QTL on chromosome 15q. Twin Res 2004;7:197-210.
21. Amos CI. Robust variance-components approach for assessing genetic linkage in pedigrees. Am J Hum Genet 1994;54:535-43.
22. Eaves LJ, Neale MC, Maes H. Multivariate multipoint linkage analysis of quantitative trait loci. Behav Genet 1996;26:519-25.
23. Fulker DW, Cherny SS. An improved multipoint sib-pair analysis of quantitative traits. Behav Genet 1996;26:527-32.
24. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet 1998;62:1198-211.
25. Hopper JL, Matthews JD. Extensions to multivariate normal models for pedigree analysis. Ann Hum Genet 1982;46:373-83.
26. Williams JT, Blangero J. Power of variance component linkage analysis to detect quantitative trait loci. Ann Hum Genet 1999;63:545-63.
27. Jacquard A. Genetic information given by a relative. Biometrics 1972;28:1101-14.
28. Lange K, Westlake J, Spence MA. Extensions to pedigree analysis, III: variance components by the scoring method. Ann Hum Genet 1976;39:485-91.
29. 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.
30. Self SG, Liang K. Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. J Am Stat Assoc 1987;82:605-10.
31. Ober C, Tsalenko A, Parry R, Cox NJ. A second-generation genomewide screen for asthma-susceptibility alleles in a founder population. Am J Hum Genet 2000;67:1154-62.
32. Evans DM. The genetics of blood cell concentrations, Unpublished PhD thesis, University of Queensland, Brisbane, Australia. 2003.
33. Daniels SE, Bhattacharrya S, James A, Leaves NI, Young A, Hill MR et al. A genome-wide search for quantitative trait loci underlying asthma. Nature 1996;383:247-50.
34. Blackwelder WC, Elston RC. A comparison of sib-pair linkage tests for disease susceptibility loci. Genet Epidemiol 1985;2:85-97.
35. Martin N, Boomsma D, Machin G. A twin-pronged attack on complex traits. Nat Genet 1997;17:387-92.