Genetic and environmental causes of variation in basal levels of blood cells

David M Evans[1], Ian H Frazer[2] and Nicholas G Martin[1]

[1]Queensland Institute of Medical Research and Joint Genetics Program, University of Queensland, Brisbane
[2]Centre for Immunology and Cancer Research, University of Queensland, Princess Alexandra Hospital, Brisbane, Australia

The genetic and environmental determinants of variation in blood cell size and number were investigated in 392 pairs of 12-year-old twins. The following blood cell indices were measured: haemoglobin, red blood cell count, haematocrit, mean corpuscular volume, platelet number, total white cell count, level of neutrophils, monocytes, eosinophils, total lymphocytes, $CD3^+$ lymphocytes, $CD4^+$ lymphocytes, $CD8^+$ lymphocytes, $CD19^+$ lymphocytes, $CD56^+$ lymphocytes and $CD4^+/CD8^+$ ratio. Genetic factors contributed significantly to all blood cell measures accounting for between 61 and 96% of variance. Heritability estimates did not differ significantly between males and females, although the sample size of the present study was not large enough to exclude the possibility of sex-limited gene expression. Common environmental factors were important in determining red blood cell count and haematocrit, but were not important in determining basal levels of any white blood cell type.

Keywords: erythrocytes, immune system, twins, genetic analyses, heritability, lymphocytes, T-cells, $CD4^+/CD8^+$ ratio

Introduction

There is considerable variation in the number and size of blood cells between individuals. Intra-individually, however, many blood cell measurements remain stable longitudinally during periods of health. Red cell and platelet indices change little over the course of 12 months, whilst levels of the various leukocyte sub-types exhibit more variation.[1–4] The existence of steady state levels across time for many blood cell counts implies that these indices are subject to homeostatic regulation. An understanding of the factors underlying this homeostasis is desirable not only from a theoretical perspective, but also practically, in terms of identifying new therapeutic possibilities in pathological states characterised by abnormal blood cell counts (eg idiopathic aplastic anaemia, neutropenia etc).

The large inter-individual variation observed in many blood cell counts suggests that at least part of this homeostatic control may be genetic. Several small twin studies have found significant genetic influences in the maintenance of red cell indices including haemoglobin level,[5,6] red blood cell count,[5–7] haematocrit,[5] mean corpuscular volume,[5,7] and also platelet count.[6] However, Yokoyama and

Correspondence: Professor NG Martin, Queensland Institute of Medical Research, PO Royal Brisbane Hospital, Brisbane 4029, Australia. Tel: + 61 7 3362 0278; Fax: + 61 7 3362 0101; E-mail: nickM@qimr.edu.au

Akiyama[8] in a study involving 154 pairs of twins found significant heritability in these variables for males, but not females. Likewise, Dal Colletto et al[6] in a study of 105 same sex twin pairs found only a suggestion of additive genetic influences for haematocrit and mean corpuscular volume and no significant heritability for platelet count. The discrepancy in results between these studies is probably due to the small sample sizes employed. In order to obtain accurate estimates of the degree of genetic influence on a quantitative trait of high heritability, a sample size of around 200 pairs of twins is needed – the figure being much higher for traits of low heritability.[9]

Only a few studies have examined the possibility that levels of leukocytes are genetically regulated. Yokoyama and Akiyama[8] found high correlations for MZ (monozygotic) relative to DZ (dizygotic) twins for natural killer cells, helper T cells and $CD4^+$ to $CD8^+$ cell ratio. However, the small number of DZ twins (36 pairs of twins) used in this study may not have provided accurate estimates of DZ correlations. In contrast, two other studies failed to find significant genetic influences accounting for total white cell count, neutrophils, monocytes, eosinophils, lymphocytes or T and B cell counts.[6,10] Again, both these studies used small numbers of twins (26 and 105 pairs of twins, respectively).

The present study investigated the possibility that the number and size of erythrocytes, platelets and leukocytes are regulated by genetic mechanisms.

Twins were measured on a variety of indices including haemoglobin, red blood cell count, haematocrit, mean corpuscular volume, platelet count, white blood cell count, level of neutrophils, monocytes, eosinophils, basophils, total lymphocytes, CD3[+] lymphocytes, CD4[+] lymphocytes, CD8[+] lymphocytes, CD19[+] lymphocytes, CD56[+] lymphocytes and CD4[+]/CD8[+] ratio. Measurements were made in a sufficiently large number of twin pairs to provide accurate estimates of heritability. It was hypothesised that additive genetic influences would contribute significantly to the variation observed in all blood cell measures.

Materials and methods

Subjects and measures

Data were collected in the context of an ongoing study on the development of melanocytic naevi (moles), the clinical protocol of which has been described in detail elsewhere.[11–13] Twins were enlisted by contacting the principals of primary schools in the greater Brisbane area, media appeals and by word of mouth. Informed consent was obtained from all participants and parents prior to testing. Data were collected from May 1992 to February 1999. The data reported here concern the first 395 pairs of twins collected.

Twins were tested as close as possible to their 12th birthday. No attempt was made to exclude subjects suffering from illness. Venous blood was collected into a 5 ml EDTA tube. Total blood haemoglobin g/l (HB), red blood cell count $\times 10^{12}$/l (RBC), mean corpuscular volume fl (MCV), platelet count ($\times 10^9$/l) and white blood cell count ($\times 10^9$/l) were measured using a Coulter Model STKS blood counter. From the RBC and MCV estimates, the machine calculated the haematocrit (HT). A five-part differential, providing estimates for neutrophil, monocyte, eosinophil, basophil and lymphocyte concentrations was based on an automated differential count. The lymphocyte subset analysis was performed on whole blood using the AutoPrep (Coulter, Hileah, FL, USA), and direct fluorochrome conjugated monoclonal antibodies to CD3, CD4, CD8, CD19 and CD56 antigens (Coulter). Analysis was performed on an Epics 753 cytofluorograph (Coulter) using standardised control samples and machine settings. The absolute numbers of CD3[+], CD4[+], CD8[+], CD19[+] and CD56[+] cells were calculated as percentages of the total lymphocyte count. It was these absolute numbers that were used for analyses of the lymphocyte subsets.

Zygosity was determined by typing eight highly polymorphic DNA microsatellite markers and three blood groups (ABO, MNS, Rh). In 27 pairs of twins where DNA was not available, zygosity was judged by similarity of appearance.

Statistical analysis

Maximum likelihood (ML) analysis of individual observations was used to test hypotheses concerning means, variances and covariances.[12] Significant twin correlations establish the presence of familial aggregation for a trait, but do not distinguish between the possible mechanisms through which similarity arises. Phenotypic variation can be conceived as arising from four sources of variation: Additive genetic (A), dominant genetic (D), and common environmental (C) sources which serve to make siblings more alike, and random environmental sources (E) which make siblings more different. The task is to determine what combination of A, C, D and E provides the most parsimonious fit to the observed data. The task is somewhat complicated by the fact that C and D are negatively confounded so that they cannot be estimated concurrently in a study of MZ and DZ twins reared together. This does not imply that C and D cannot both contribute to the phenotypic variance of a trait, rather they cannot be estimated simultaneously with data from twins alone.

In the present study A, C, E and D were estimated through structural equation modelling (SEM) using the software package Mx.[14] For each variable, the full ACE model, or when the MZ correlation was greater than twice the DZ correlation, the ADE model, was fitted to the raw data. Subsequent, more simplified models were fitted to test whether A, C (or D) or both parameters could be dropped from the full model. The fit of each submodel was assessed by the difference in log likelihood between the sub and full models.[12] Twice the difference in log likelihood is distributed as $\Delta\chi^2$ with the degrees of freedom (df) for this test equal to the difference in df between the full and sub-model.

Results

Of the 395 pairs of twins recruited, blood specimens were unsuitable for analysis in three pairs. The data from one twin were excluded from further analyses because of aplastic anaemia. Thus, data were available for 72 MZ female, 66 MZ male, 62 DZ female, 72 DZ male and 120 opposite sex (OS) twin pairs (consisting of 57 pairs where the female was born first and 63 pairs where the male was born first). The means and standard deviations of all blood measurements for males and females are listed in Table 1. Basophils were excluded from analyses because the number of cells was too small.

Table 1 Means and standard deviations for male and female twins on blood cell measures

Variable	Units	Range	Females N	Mean	SD	Males N	Mean	SD
Haemoglobin	g/l	107–162	385	133.99	8.64	387	135.11	7.68
Red cell count	Cells/l $\times 10^{-12}$	3.56–6.30	385	4.67	0.32	387	4.75	0.33
Haematocrit	L/l	0.318–0.481	385	0.398	0.02	387	0.399	0.02
Mean corpuscular volume	fL	62–97	385	85.34	3.37	387	84.29	3.61
Platelet count	Cells/l $\times 10^{-9}$	145–509	384	285.92	56.82	387	283.56	57.49
White cell count	Cells/l $\times 10^{-9}$	3.5–14.3	385	7.30	1.58	387	7.04	1.80
Neutrophils	Cells/l $\times 10^{-9}$	1.34–10.61	368	3.71	1.18	353	3.40	1.25
Monocytes	Cells/l $\times 10^{-9}$	0.16–2.58	368	0.53	0.20	353	0.55	0.20
Eosinophils	Cells/l $\times 10^{-9}$	0.00–2.48	368	0.37	0.29	353	0.46	0.35
Lymphocytes	Cells/l $\times 10^{-9}$	0.90–5.88	385	2.62	0.62	392	2.56	0.64
$CD3^+$ T cells	Cells/l $\times 10^{-9}$	0.58–4.24	385	1.86	0.48	392	1.77	0.46
$CD4^+$ T cells	Cells/l $\times 10^{-9}$	0.43–2.61	385	1.13	0.32	392	1.03	0.28
$CD8^+$ T cells	Cells/l $\times 10^{-9}$	0.08–1.88	385	0.63	0.22	392	0.62	0.23
$CD19^+$ B cells	Cells/l $\times 10^{-9}$	0.17–1.35	385	0.50	0.17	392	0.51	0.18
$CD56^+$ NK cells	Cells/l $\times 10^{-9}$	0.01–0.88	354	0.20	0.13	337	0.20	0.13
$CD4^+/CD8^+$ ratio		0.27–6.00	385	1.93	0.63	392	1.78	0.59

All variables were examined for the presence of genotype $\times$ environment (G $\times$ E) interactions. G $\times$ E interaction occurs when genotypes differ in their susceptibility to environmental agents. The presence of a G $\times$ E interaction confounds any attempt to decompose phenotypic variance into simple additive and environmental components. The presence of a G $\times$ E interaction was tested by plotting the absolute difference in raw scores for a variable against the sum of scores for each variable[15] for MZ twin pairs only (Table 2). Since MZ twins share all their genes in common, the difference between MZ twins can only be due to environmental causes. Similarly, the sum of MZ twin scores provides an estimate of genetic effects. A significant correlation between the two indicates the presence of G $\times$ E interaction.

Since G $\times$ E interactions are scale dependent, they can usually be removed by an appropriate transformation (Table 2). Total lymphocyte count, $CD4^+$, $CD8^+$, and $CD19^+$ lymphocyte counts were transformed using a square root transformation. Concentrations of neutrophils, monocytes, $CD56^+$ NK cells and $CD4^+/CD8^+$ ratio were transformed using log transformations (ie $\log_{10}(X)$). Eosinophil concentration was transformed using the calculation $\log_{10}(X + 0.05)$. WBC was transformed using the calculation $\log_{10}(X - 1.0)$. As this study was concerned with normal populations, scores greater than 3.5 standard deviations from the mean were excluded from further analyses. Note that all measures derived from the outlying variable were excluded from the analyses too (eg if a subject's lymphocyte count was greater than 3.5 standard deviations from the mean, the subject's scores on all lymphocyte subsets were also excluded from analyses). Four subjects' scores were removed from RBC (3.56, 3.56, 6.25, 6.30 cells/l $\times 10^{-12}$) and MCV (62, 64, 96, 97 fl) and therefore also from HT; one from platelet count (509 cells/l $\times 10^{-9}$); three from monocytes (0.16, 1.98, 2.58 cells/l $\times 10^{-9}$); four from total lymphocytes (5.25, 5.30, 5.43, 5.89 cells/l $\times 10^{-9}$) and thus all lymphocyte subsets; plus an extra two from $CD3^+$ lymphocytes (3.47, 3.48 cells/l $\times 10^{-9}$); four from $CD8^+$ lymphocytes (0.08, 1.58, 1.73, 1.88 cells/l $\times 10^{-9}$); one from $CD19^+$ B cells (1.35 cells/l $\times 10^{-9}$) and three from $CD56^+$ NK cells (0.01, 0.88 cells/l $\times 10^{-9}$). Subjects excluded from total lymphocytes, $CD4^+$ T cells and $CD8^+$ T cells were not included in the analysis of $CD4^+/CD8^+$ ratio.

Correlations between the transformed variables are displayed in Table 3. HB, RBC and HT correlated highly with one another, but did not correlate with

Table 2 Correlation between the absolute difference and sum of co-twin scores for female (r_F; 63–71 pairs) and male (r_M; 47–65 pairs) MZ twins before and after transformation

Variable	Before r_F	r_M	After r_F	r_M
Haemoglobin	−0.01	0.02	NA	NA
Red cell count	0.17	−0.03	NA	NA
Haematocrit	0.03	−0.07	NA	NA
Mean corpuscular volume	−0.06	0.04	NA	NA
Platelet count	0.16	0.15	NA	NA
White cell count[c]	0.26[a]	0.36[b]	0.07	0.20
Neutrophils[c]	0.36[b]	0.40[b]	0.07	0.10
Monocytes[c]	0.31[b]	0.37[b]	−0.10	0.24
Eosinophils[c]	0.82[b]	0.44[b]	0.20	−0.02
Lymphocytes[d]	0.05	0.25[a]	0.11	0.12
$CD3^+$ T cells	−0.06	0.18	NA	NA
$CD4^+$ T cells[d]	0.04	0.24	−0.04	0.10
$CD8^+$ T cells[d]	0.12	0.27[a]	−0.08	0.11
$CD19^+$ B cells[d]	0.07	0.21	−0.11	0.02
$CD56^+$ NK cells[c]	0.57[b]	0.43[b]	0.09	−0.20
$CD4^+/CD8^+$ ratio[c]	0.45[b]	0.38[b]	0.18	0.13

[a]$P < 0.05$; [b]$P < 0.01$; [c]variable transformed using a logarithmic function; [d]variable transformed using a square root function; NA = Not appropriate

Table 3   Correlations between blood cell measures in 12-year old female (N = 339–385) and male (N = 303–387) twins after pairwise deletion of missing data

		Hb	Rbc	Ht	Mcv	Plt	Wbc	Neu	Mon	Eos	Lym	CD3	CD4	CD8	B	Nk	4/8
									FEMALES								
Haemoglobin			0.80	0.95	0.12	0.11	0.05	0.07	0.11	−0.01	−0.01	−0.05	−0.08	−0.01	−0.06	0.15	−0.06
Red cell count		0.79		0.82	−0.42	0.12	0.04	0.02	0.11	−0.04	0.02	−0.02	−0.04	0.00	−0.04	0.17	−0.05
Haematocrit		0.94	0.81		0.16	0.12	0.09	0.09	0.10	0.01	0.01	−0.03	−0.06	0.00	−0.05	0.14	−0.05
Mean corpuscular volume		−0.01	−0.55	0.03		−0.02	0.07	0.11	−0.03	0.01	−0.02	−0.03	−0.02	−0.01	0.00	−0.07	0.00
Platelets (Plt)		−0.02	0.02	−0.02	0.06		0.36	0.27	0.20	0.13	0.29	0.24	0.25	0.14	0.22	0.16	0.07
White cell count	M	0.07	0.06	0.13	0.06	0.29		0.87	0.35	0.36	0.57	0.53	0.51	0.41	0.43	0.21	0.00
Neutrophils (Neu)	A	0.03	0.02	0.08	0.07	0.21	0.87		0.23	0.16	0.20	0.18	0.21	0.13	0.20	0.05	0.05
Monocytes (Mon)	L	0.06	0.08	0.09	−0.03	0.15	0.49	0.38		0.12	0.21	0.19	0.13	0.19	0.08	0.15	−0.09
Eosinophils (Eos)	E	0.08	0.04	0.10	0.06	0.14	0.44	0.16	0.16		0.16	0.18	0.17	0.13	0.13	−0.04	0.02
Lymphocytes (Lym)	S	0.08	0.06	0.13	0.07	0.24	0.61	0.25	0.20	0.29		0.94	0.86	0.73	0.66	0.37	−0.03
CD3+ T cells (CD3)		0.01	0.00	0.07	0.07	0.18	0.50	0.18	0.13	0.24	0.92		0.89	0.78	0.49	0.21	−0.06
CD4+ T cells (CD4)		−0.01	−0.08	0.03	0.16	0.16	0.41	0.12	0.10	0.25	0.80	0.84		0.51	0.51	0.17	0.33
CD8+ T cells (CD8)		0.03	0.04	0.07	0.00	0.12	0.36	0.10	0.11	0.14	0.69	0.78	0.49		0.33	0.21	−0.63
CD19+ B cells (B)		0.06	−0.05	0.08	0.19	0.18	0.48	0.25	0.17	0.34	0.63	0.45	0.49	0.26		0.02	0.08
CD56+ NK cells (NK)		0.25	0.28	0.29	−0.07	0.03	0.26	0.11	0.23	0.12	0.34	0.18	0.12	0.20	−0.03		−0.09
CD4+/CD8+ ratio (4/8)		−0.06	−0.11	−0.05	0.12	−0.01	−0.07	−0.01	−0.05	0.06	−0.08	−0.15	0.29	−0.68	0.11	−0.12	

the majority of white blood cell subsets. MCV correlated highly with RBC, but not with HB, HT or the white blood cell groups. Most white blood cell subsets correlated highly with one other, but did not correlate with the haematological variables. Platelets correlated moderately with most white cell subsets, but not with the red cell indices.

Hypotheses concerning means

It is necessary to demonstrate that means and variances are equal across different zygosity groups to establish regularity and randomness of sampling. Once this is demonstrated, means and variances can be pooled across the different zygosity groups to give a better estimate of the underlying population mean and variance.

Each of the five zygosity groups has two sample means (for twin 1 and twin 2), two variances and a covariance, giving a total of 10 means, 10 variances and 5 covariances. Thus, initially a fully saturated model with 25 parameters (H0) was estimated. Subsequent more simplified models were tested by calculating a difference in log likelihood χ^2 test against the preceding model (Table 4a). Because of the large number of statistical tests performed, each

Table 4a   Contrasts used to test hypotheses about means (m) and variances (v) in the analysis of individual observations for twin pairs

Zygosity group	$H0_{m,v}$		$H1_{m,v}$		$H2_{m,v}$		$H3_{m,v}$		$H4_{m,v}$	
	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2
1. MZFF	1	2	1	1	1	1	1	1	1	1
2. MZMM	3	4	2	2	2	2	2	2	1	1
3. DZFF	5	6	3	3	1	1	1	1	1	1
4. DZMM	7	8	4	4	2	2	2	2	1	1
5. DZOS	9	10	5	6	3	4	1	2	1	1

Table 4b   Contrasts used to test hypotheses about covariances (c) in the analysis of individual observations for twin pairs

Zygosity group	$H0_c$	$H1_c$	$H2_c$	$H3_c$	$H4_c$
1. MZFF	1	1	1	1	0
2. MZMM	2	1	1	1	0
3. DZFF	3	2	2	1	0
4. DZMM	4	2	2	1	0
5. DZO	5	3	2	1	0

test was evaluated with respect to an alpha level of 0.01.

The first sub-model (H1m) equated means for twin 1 and twin 2 within same sex pairs. As twin 1 was always the first twin born in each pair, H1m tested for the presence of birth order effects. H2m equated means across same sex MZ twins and equated means across same sex DZ twins. Thus, H2m tested for homogeneity of means between MZ and DZ twins within like-sex twin pairs. H3m equated means across all female twins and equated means across all male twins, thus testing homogeneity of means between MZ and DZ twins within all female twins and within all male twins. Finally, H4m equated all means, testing for differences between male and female means.

Males and females differed significantly in their means for RBC, MCV, neutrophil count, eosinophil count, level of CD4+ lymphocytes and CD4+/CD8+ ratio (Table 5). Thus, different means for males and females were retained in future analyses of these variables. All other comparisons were not significant, so the means of the remaining variables were pooled across the different sex and zygosity groups.

Hypotheses concerning variances

The same set of contrasts that were used to test differences between means were used to test for

Table 5 Difference in log likelihood $\Delta\chi^2$ for tests concerning means (M), variances (V) and covariances (C) of blood cell measurements in 392 pairs of 12-year-old twins

Model	d.f.	HB	RBC	HT	MCV	PLT	WBC	NEUT	LYM	MON	EOS	CD3	CD4	CD8	B	NK	4/8
$H1_M$	4	2.21	1.64	1.62	1.79	7.32	1.19	1.90	2.39	6.72	4.48	1.63	0.74	4.64	3.06	7.34	7.29
$H2_M$	2	2.62	0.25	1.91	3.50	0.32	1.43	2.00	3.20	0.87	1.79	5.89	5.43	4.03	1.28	2.31	0.96
$H3_M$	2	0.72	0.99	1.32	3.08	1.38	0.47	0.03	0.77	1.05	3.40	0.33	0.06	1.30	0.46	3.27	2.22
$H4_M$	1	3.26	9.79[b]	0.41	15.02[b]	0.85	5.50	12.90[b]	1.61	1.21	12.78[b]	4.77	20.84[b]	0.00	0.01	0.00	16.26[b]
$H1_V$	4	3.05	1.74	3.94	2.17	4.11	6.91	11.45	2.25	6.60	4.29	3.74	1.73	7.03	3.14	0.20	4.55
$H2_V$	2	1.78	0.50	7.46	6.67	8.43	5.68	6.55	1.96	2.25	1.34	1.07	4.16	2.96	0.29	1.70	4.24
$H3_V$	2	1.06	0.04	7.69	4.66	1.94	1.80	0.53	1.31	3.89	1.72	1.96	8.97	1.06	2.52	5.29	2.64
$H4_V$	1	3.11	0.15	2.97	0.36	0.02	4.82	4.52	0.19	3.88	2.58	0.04	1.54	1.23	1.06	1.47	0.43
$H1_C$	2	0.14	0.71	1.54	3.17	2.77	0.95	4.19	0.11	4.24	2.94	0.42	0.20	1.03	1.37	0.21	0.44
$H2_C$	1	3.99	4.48	1.95	0.34	0.12	3.78	1.05	2.78	0.31	0.05	5.24	0.78	2.40	0.23	0.41	0.66
$H3_C$	1	45.95[b]	53.42[b]	24.93[b]	180.08[b]	32.76[b]	19.16[b]	15.99[b]	19.08[b]	8.03[a]	21.02[b]	27.22[b]	27.27[b]	35.83[b]	30.97[b]	10.06[b]	41.08[b]
$H4_C$	1	169.26[b]	220.56[b]	184.58[b]	204.57[b]	151.98[b]	136.26[b]	93.75[b]	121.50[b]	93.45[b]	90.28[b]	120.27[b]	122.98[b]	142.24[b]	121.01[b]	101.71[b]	178.84[b]

[a]$P<0.01$; [b]$P<0.001$

differences between variables (Table 4a). As none of these comparisons were significant (Table 5) variance estimates were pooled across all sex and zygosity groups.

Hypotheses concerning covariances

Whilst it is important to test that means and variances are equal across different zygosity groups to establish regularity and randomness of sampling, it is the pattern of covariance across different zygosity groups that is the focus of twin studies. The correlation for all five zygosity groups is displayed in Table 6. If a trait is genetically influenced, the correlation between MZ twins should be significantly greater than the correlation between DZ twins. However, the size and aetiology of genetic variation between males and females are not necessarily the same. This is known as sex limitation. 'Scalar' sex limitation occurs when the same set of genes operates in males and females, but the magnitude of the genetic effect differs between the sexes. Non-scalar sex limitation occurs when different genes in males and females affect expression of the phenotype.[16] The contrasts used to test for the presence of these effects are listed in Table 4b.

All sub-models were tested with respect to a saturated full model where all covariances were unequal (H0c). H1c equated the covariances of MZ twins, and equated the covariances of same-sex DZ twins. Thus, the comparison between H1c and H0c tested for the presence of scalar sex limitation. H2c equated the covariances of MZ twins and the covariances of all DZ twins. The comparison between H2c and H1c therefore tested for the presence of non scalar sex limitation. H3c equated all covariances. H3c tested whether variance on a trait was influenced by genetic factors. Finally, H4c set all covariances to zero and thus tested for the presence of familial aggregation.

Results of these tests are shown in Table 5. All variables displayed highly significant differences in χ^2 values between H3c and H2c indicating that genetic effects explained significant portions of variance on all measures. Similarly, there was a significant difference between H4c and H3c across all variables, indicating that familial aggregation was present for all variables tested.

Structural equation modelling of genetic hypotheses

The presence of sex limitation was tested by fitting a sex limitation model to each of the variables. The full sex limitation model estimated seven parameters – A, C and E for females, A, C and E for males – and

Table 6 Twin correlations[a] for five sex/zygosity groups (from H0$_c$) and for total MZ and total DZ pairs (from H2$_c$)

ZYG	Pairs	HB	RBC[b]	HT	MCV[b]	PLT	WBC	NEU[b]	MON	EOS	LYM	CD3	CD4[b]	CD8	B	NK	4/8[b]
MZF	72	0.84	0.86	0.72	0.97	0.86	0.68	0.67	0.66	0.63	0.73	0.74	0.73	0.83	0.77	0.70	0.85
MZM	66	0.83	0.87	0.76	0.97	0.77	0.75	0.75	0.60	0.76	0.70	0.74	0.75	0.78	0.74	0.65	0.85
DZF	62	0.58	0.60	0.63	0.38	0.43	0.55	0.46	0.33	0.38	0.50	0.49	0.46	0.56	0.40	0.35	0.42
DZM	72	0.60	0.66	0.65	0.60	0.46	0.49	0.38	0.59	0.36	0.50	0.57	0.46	0.47	0.45	0.36	0.50
DZO	120	0.43	0.47	0.60	0.44	0.48	0.32	0.31	0.39	0.34	0.33	0.31	0.37	0.36	0.38	0.48	0.54
MZ	138	0.83	0.87	0.82	0.96	0.80	0.70	0.66	0.64	0.69	0.70	0.74	0.73	0.80	0.75	0.67	0.84
DZ	254	0.51	0.56	0.51	0.47	0.47	0.44	0.38	0.42	0.35[b]	0.43	0.41[b]	0.40	0.45	0.40	0.43	0.51

[a]based on models H4m and H4v; [b]based on models H3m and H4v

an estimate of the additive genetic correlation between opposite sex twin pairs. Constraining the additive genetic correlation between opposite sex twin pairs to 0.5 did not result in a significantly worse fit compared with the full model for any variable, although the difference in log likelihood χ^2 values for haemoglobin and red blood cell count approached significance (HB: $\Delta^2_{\chi 1} = 2.96$, P = 0.09; RBC: $\Delta^2_{\chi 1} = 2.14$, P = 0.14). Constraining A, C and E equal across sexes also did not produce a significantly worse fit compared with the full sex limitation model for any variable. There was, however, a significant difference between an AE submodel with different estimates of A and E for males and females, and an AE model with equal estimates for males and females for the neutrophil variable ($\Delta^2_{\chi 2} = 6.31$, P = 0.04).

Thus for all variables, A, C and E were constrained equal across sexes and sub-models were fitted to test whether A, C or both parameters could be dropped from the full model. The results of the model fitting and standardised variance estimates for the significant parameters are presented in Table 7. An ACE model provided the best fit for RBC and HT. An AE model provided the best fit for all the other continuous measures. Estimates of additive genetic influences ranged from 0.61 to 0.96. It must be noted that the results for CD4$^+$/CD8$^+$ ratio (unlike the other variables) depended upon the exclusion of outliers. When outliers were included in the analysis, common environmental effects reached significance ($\Delta^2_{\chi 1} = 4.40$, P = 0.04; A = 0.64, C = 0.20, E = 0.17).

Discussion

There was enormous inter-individual variation in the blood cell measures of 12-year-old twins, particularly in white cell counts and lymphocyte subsets. The present study has demonstrated that part of this large variation is due to genetic factors. Additive genetic influences were important in determining basal levels of all blood cell measures accounting for between 61 and 96% of the total variance. Common environmental factors were also important in determining red blood cell count and haematocrit, accounting for 25 and 20% of the total variation in these variables, respectively.

The finding that genetic influences were important in determining basal levels of haematological variables is similar to the findings of some[5,7] but not all studies.[6,8] Discrepancy between our results and the results of the latter two studies was most likely due to the latter studies' smaller sample sizes and associated lack of power. As hypothesised, genetic factors explained significant proportions of the variance for all white blood cells. A few studies have previously found evidence for genetic control of the levels of one or two white blood cell types,[5,8] but the generalisability of these studies has been limited because of the small sample sizes involved and the fact that only a few white cell subsets were measured. Thus, the present study expands on previous work by demonstrating that the basal levels of many white blood cell subsets are regulated by genetic mechanisms. The task of the future will be to identify the genes responsible for this control.

Table 7 Results of fitting genetic and environmental models to raw data (95% confidence intervals are given in parentheses)

Variable	Proportion of total variance			−2LL fit to ACE model	df	$\Delta\chi^2$ (compared with full model)		
	A	C	E			AE $\Delta\chi^2_1$	CE $\Delta\chi^2_1$	E $\Delta\chi^2_2$
Haemoglobin	0.84 (0.79–0.87)	–	0.16 (0.13–0.21)	5207.422	769	3.832	45.946[b]	235.203[b]
Red cell count	0.61 (0.46–0.79)	0.25 (0.07–0.40)	0.14 (0.10–0.18)	122.478	764	7.366[b]	53.417[b]	273.977[b]
Packed cell volume	0.62 (0.44–0.82)	0.20 (0.01–0.36)	0.18 (0.10–0.18)	−3822.957	765	4.211[a]	39.922[b]	224.503[b]
Mean corpuscular volume	0.96 (0.94–0.97)	–	0.04 (0.03–0.06)	3607.326	763	0.001	189.15[b]	393.718[b]
Platelets	0.80 (0.75–0.85)	–	0.20 (0.15–0.25)	8221.981	767	1.751	37.720[b]	189.700[b]
White cell count	0.71 (0.64–0.77)	–	0.29 (0.23–0.36)	−1269.294	768	2.934	17.292[b]	136.260[b]
Neutrophils	0.67 (0.58–0.74)	–	0.33 (0.26–0.42)	−892.200	717	0.638	15.435[b]	109.190[b]
Monocytes	0.66 (0.56–0.73)	–	0.34 (0.27–0.44)	−921.227	715	3.179	8.034[b]	101.486[b]
Eosinophils	0.69 (0.60–0.76)	–	0.31 (0.24–0.40)	60.159	717	0.042	21.021[b]	111.302[b]
Lymphocytes	0.71 (0.63–0.77)	–	0.29 (0.23–0.37)	−558.704	769	1.975	17.548[b]	139.051[b]
CD3$^+$ T cells	0.75 (0.68–0.80)	–	0.25 (0.20–0.32)	798.444	767	0.704	27.220[b]	120.267[b]
CD4$^+$ T cells	0.74 (0.67–0.80)	–	0.26 (0.20–0.33)	−1011.019	769	0.572	27.266[b]	150.243[b]
CD8$^+$ T cells	0.80 (0.74–0.84)	–	0.20 (0.16–0.26)	−1122.732	766	0.932	38.861[b]	181.10[b]
CD19$^+$ B cells	0.75 (0.69–0.81)	–	0.25 (0.19–0.32)	−1260.073	769	0.243	31.862[b]	121.012[b]
CD56$^+$ NK cells	0.68 (0.59–0.75)	–	0.32 (0.25–0.41)	−13.035	682	3.077	10.061[b]	111.766[b]
CD4$^+$ CD8$^+$ ratio	0.84 (0.79–0.87)	–	0.16 (0.13–0.21)	−1141.699	765	3.288	46.196[b]	225.034[b]

[a]P<0.05; [b]P<0.01

In the present study, there was some evidence of sex limitation. Analysis of the neutrophil variable suggested that the magnitude of genetic effects may differ between the sexes. Similarly, the possibility that different genes underlie the regulation of some blood cell indices in 12-year-old males and females should not be ruled out either. Although a simple model constraining the magnitude of A, C and E equal between the sexes could not be formally rejected for any variable, the sample size of the present study was not large enough to properly address the issue of sex limitation.[9] Future studies should bear this issue in mind.

Common environmental factors were important in influencing red blood cell count and haematocrit. Possible environmental factors responsible include dietary iron intake and exercise, as these variables are known to affect red cell indices and are likely to be shared among 12-year-old co-twins. The present result is similar to the results of Whitfield and Martin's[5] study of young adults, save that they found significant shared environmental effects influencing mean corpuscular volume but not red blood cell count. The discrepancy in results between studies is not surprising since 12-year-old co-twins are unlikely to be exposed to many of the environmental agents that affect mean corpuscular volume (eg smoking, alcohol intake). Similarly, the dietary iron intake of 12-year-old co-twins is likely to be similar, since young twins usually live in the same house and presumably share the same meals (ie in this case dietary iron is a shared environmental factor influencing red blood cell count). Adult co-twins, however, are likely to live apart and thus would have different dietary iron intakes (ie dietary iron in this case is a unique environmental factor).

A surprising finding of the present study was that common environmental effects did not contribute significantly to variation in the basal levels of white blood cells. Assuming that 12-year-old co-twins living in the same house are exposed to similar amounts and types of antigen, the present result implies that prior antigen exposure is not important in determining basal levels of white blood cells. The same argument also suggests that immunisation history is not an important factor either, since this is also likely to be shared by co-twins reared together. The possibility that common environmental factors influence the $CD4^+/CD8^+$ ratio deserves further investigation, however, as their statistical significance depended upon the inclusion of a small number of outlying cases.

In conclusion, the present study demonstrates that the basal levels of many blood cell measurements in 12-year-old twins are due to the additive action of several genes of small effect. Common environmental effects are important in regulating red blood cell count and haematocrit, but do not appear to be important in determining the basal level of white cell types at this age. We are now performing a genome scan of the DZ twins and siblings from this study with the aim of identifying quantitative trait loci underlying homeostatic control of these blood cell types.

References

1 Statland LP, Winkel P, Harris SC, Burdsall ML, Saunders AM. Evaluation of biologic sources of variation of leukocyte counts and other hematologic quantities using very precise automated analyzers. Am J Clin Pathol 1977; 69: 48–54.

2 Costongs GMPJ, Janson PCW, Bas BM, Hermans J, Brombacher PJ, van Wersch JKJ. Short-term and long-term intra-individual variations and critical differences of hematological laboratory parameters. J Clin Chem Clin Biochem 1985; 23: 69–76.

3 Fraser CG, Wilkinson SP, Neville RG, Knox JDE, King JF, MacWalter RS. Biological variation of common hematologic laboratory quantities in the elderly. Am J Clin Pathol 1989; 92: 465–470.

4 Dot D, Miro J, Fuentes-Arderiu X. Biological variation of the leukocyte differential count quantities. Scand J Clin Lab Invest 1992; 52: 607–611.

5 Whitfield JB, Martin NG. Genetic and environmental influences on the size and number of cells in the blood. Genet Epidemiol 1985; 2: 133–144.

6 Dal Colletto GM, Fulker DW, Barretto OC, Kolya M. Genetic and environmental effects on blood cells. Acta Genet Med Gemellol (Roma) 1993; 42: 245–252.

7 Lindemann JP, Kang KW, Christian JC. Genetic variance of erythrocyte parameters in adult male twins. Clin Genet 1977; 12: 73–76.

8 Yokoyama Y, Akiyama T. Intrapair differences of the blood cell components and lymphocyte subsets in monozygotic and dizygotic twins. Acta Genet Med Gemellol (Roma) 1995; 44: 203–214.

9 Martin NG, Eaves LJ, Kearsey MJ, Davies P. The power of the classical twin study. Heredity 1978; 40: 97–116.

10 Anton-Guirgis H, Culver BD, Kurosaki T, Elston R. A study of multiple biological markers in twins. Acta Genet Med Gemellol (Roma) 1985; 34: 153–165.

11 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.

12 McGregor B, Pfitzner J, Zhu G, Grace M, Eldridge A, Pearson J, Mayne C, Aitken JF, Green AC, Martin NG. Genetic and environmental contributions to size, color, shape, and other characteristics of melanocytic naevi in a sample of adolescent twins. Genet Epidemiol 1999; 61(1): 40–53.

13 Zhu G, Duffy DL, Eldridge A, Grace M, Mayne C, O'Gorman L, Aitken JF, Neale MC, Hayward NK, Green AC, Martin NG. A major QTL 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(2): 483–492.

14 Neale MC. Mx: Statistical Modelling, 4th edn. Department of Psychiatry, Medical College of Virginia: Richmond, VA, 1997.

15 Jinks JL, Fulker DW. Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of human behaviour. Psychol Bull 1970; 73: 311–349.

16 Eaves LJ. Inferring the causes of human variation. JR Stat Soc 1977; 140(3): 324–355.