Am J Hum Genet 36:663–676, 1984

Heritability of Quantitative Variation at the Group-Specific Component (Gc) Locus

STEPHEN P. DAIGER,[1] MARVIN MILLER,[2] AND RANAJIT CHAKRABORTY[1]

SUMMARY

Human group-specific component (Gc) is the plasma transport protein for vitamin D; in addition, polymorphic electrophoretic variants of Gc are found in all human populations. Because of its physiologic importance and in view of the extensive genetic variation at the Gc locus, we have determined the heritability of quantitative variation in Gc by comparing a series of monozygotic (MZ) and dizygotic (DZ) twins of known Gc genotype. The series included 31 MZ twin pairs, 13 DZ twin pairs, and 45 unrelated controls. Since Gc concentration is increased by estrogens, pregnant women and women taking oral contraceptives were excluded. We found no age-related differences in Gc concentration or differences between males and females, but the concentrations of Gc in the three electrophoretically determined genotypes were significantly different from each other. Using classical methods of heritability analysis, the overall heritability of variation in Gc concentration is approximately 70%. Heritability in males is greater than in females, probably reflecting the additional environmental effect of estrogens in women. To determine if the differences in Gc concentration between the three genotypes explain the high heritability, a new variance decomposition procedure was developed following classical methods in quantitative genetics. Application of this method suggests that 19% of the total variation in Gc concentration, combining both sexes, is due to electrophoretic differences between individuals (30% in females and 20% in males). Thus, the genetic component of variation in Gc concentration can be decomposed into a major gene component—the result of electrophoretic variation at the structural locus—and a second, unexplained, polygenic component.

Received August 16, 1983; revised October 31, 1983.

[1] Center for Demographic and Population Genetics, Graduate School of Biomedical Sciences, The University of Texas Health Science Center at Houston, P.O. Box 20334, Houston, TX 77225.

[2] Department of Pediatrics, The University of Rochester School of Medicine, Rochester, NY 14642.

INTRODUCTION

Human group-specific component (Gc) protein is the plasma transport protein for vitamin D and vitamin D metabolites [1]. Vitamin D-binding proteins that are antigenically similar to human Gc are found in most mammals. Gc is a relatively abundant plasma protein (30–50 mg/100 ml or 6–10 μM approximately) whose concentration changes little throughout life except in pregnant women or in women taking oral contraceptives in whom Gc concentration may increase as much as twofold [2]. Apparently, Gc concentration is not affected by sunlight or by plasma vitamin D concentration, nor is it abnormal in acquired or inherited diseases of vitamin D metabolism or in bone diseases [3–4]. Although several hundred-thousand individuals have been sampled for Gc, its absence has not been reported. Clearly, Gc is of major physiologic significance although its exact biologic role has not been established.

In addition to its physiologic importance, Gc is also an excellent genetic marker, with extensive polymorphic variation. Two electrophoretic alleles, Gc^1 and Gc^2, are found in all human populations [5] and two isoelectric subtypes, Gc^{1f} and Gc^{1s}, are also widely distributed [6]. At least 40 uncommon and rare Gc types have been described [7].

We have determined the heritability of Gc concentration by comparing a series of MZ and DZ twins. We were interested in heritability for several reasons. First, a low heritability would suggest that common environmental factors, not yet detected, affect Gc concentration, whereas a high heritability would argue against such factors. Second, we wondered whether differences in the variance of Gc concentration between MZ and DZ twin pairs could be ascribed to variation in the polymorphic alleles at the structural locus. Finally, we hoped to compare heritability between age groups and between sexes to establish the effect of these variables on Gc concentration.

MATERIALS AND METHODS

Twin pairs of like sex from the Seattle metropolitan community were selected from participants in two unrelated twin studies ([8] and Carter-Saltzman, unpublished data, 1978). Small serum aliquots from these studies were provided for our investigation with the informed consent of donors or their parents. Zygosity was established by testing 19 polymorphic genetic markers. Loci tested were: ACP_1, GLO_1, NP, ADA, G6PD, PGM_1, PGM_2, AK_1, ENO_1, SGOT, GPT, UMPK, ESD, ABO, Rh, Kell, Duffy, MN, and Gc. (Note that with Gc subtypes alone the probability of discordance in dizygous twins is 34%.) Control blood samples were from unrelated studies conducted concurrently ([3] and S. P. Daiger, unpublished data, 1978). The estrogen status of adult female donors, that is, whether or not they were pregnant or taking oral contraceptives, was determined by interview by participating clinicians. All donors were Caucasian with an age range of 1 to 45 (see table 1).

Gc types were determined by electrophoresis-autoradiography [9] and Gc subtypes were determined by isoelectric focusing [10]. The concentration of Gc was established by radial immunodiffusion using M-Partigen Plates (CalBiochem-Behring, La Jolla, Calif.). Serum samples were diluted 50% with normal saline, 5 μl aliquots were loaded in duplicate, immunodiffusion was allowed to go to completion in 72 hrs at room temperature, and the immunoprecipitate disk was measured to establish relative Gc concentration. Standards of known Gc concentration were included in each plate.

TABLE 1

DATA ON SAMPLED INDIVIDUALS (Gc CONCENTRATION IN mg/100 ml)

A. Monozygotic twins

	Sex	Age	Gc type	[Gc] Twin 1	[Gc] Twin 2
1	F	7	1S-1S	32.9	31.7
2	F	8	1S-1S	29.6	30.7
3	F	8	2-1F	27.2	26.7
4	F	8	2-1F	27.8	28.4
5	F	9	1F-1S	25.6	30.9
6	F	9	1S-1S	32.0	33.7
7	F	9	1S-1S	32.3	32.8
8	F	9	2-1S	27.6	28.5
9	F	9	2-1S	31.3	30.7
10	F	10	1S-1S	32.3	36.0
11	F	11	1F-1S	35.8	35.2
12	F	13	2-1S	30.2	30.2
13	F	15	1F-1S	32.8	35.8
14	F	15	2-1F	25.8	23.8
15	F	17	2-2	27.1	27.1
16	F	19	1S-1S	32.6	32.0
17	F	20	1F-1S	25.8	24.5
18	F	20	1F-1S	32.0	34.0
19	F	20	2-1F	23.2	22.2
20	F	29	1S-1S	35.6	33.7
21	F	45	1S-1S	29.1	36.7
22	M	7	2-2	26.7	27.1
23	M	9	1F-1S	28.2	29.3
24	M	9	2-1S	28.4	29.6
25	M	10	2-1S	23.3	25.8
26	M	12	1S-1S	30.9	28.8
27	M	13	1S-1S	30.4	30.4
28	M	20	2-1S	28.2	28.5
29	M	22	1F-1S	28.6	29.4
30	M	23	1F-1S	24.2	21.9
31	M	36	1S-1S	34.0	34.4

B. Dizygotic twins

	Sex	Age	Twin 1 Gc type	Twin 1 [Gc]	Twin 2 Gc type	Twin 2 [Gc]
1	F	6	1F-1S	34.9	1F-1S	32.9
2	F	11	1F-1S	30.4	2-1S	28.2
3	F	17	1S-1S	33.0	2-1S	35.3
4	F	20	2-1S	25.5	1S-1S	30.9
5	F	21	1F-1S	30.9	1F-1S	29.0
6	F	24	2-2	27.2	2-2	29.7
7	F	28	1F-1F	29.3	1F-1F	32.9
8	F	42	2-2	28.1	2-1S	23.9
9	M	10	1F-1S	31.1	1F-1S	29.5
10	M	11	1S-1S	33.2	1S-1S	29.3
11	M	11	1S-1S	36.7	2-1S	32.5
12	M	15	2-2	24.9	2-2	24.3
13	M	19	2-2	18.4	2-1S	29.0

(*Table continues on p. 666.*)

TABLE 1 (*continued*)

C. Unrelated controls

	Sex	Age	Gc type	[Gc]
1	F	1	1-2	29.3
2	F	2	2-2	26.0
3	F	3	1-1	33.4
4	F	4	1-1	31.0
5	F	5	1-2	33.0
6	F	7	1-1	34.4
7	F	8	1-1	33.0
8	F	9	1-1	31.4
9	F	9	1-1	35.6
10	F	9	2.2	26.5
11	F	10	1-2	37.2
12	F	11	1-2	27.3
13	F	13	2-2	28.5
14	F	17	1-1	28.8
15	F	17	1-2	38.4
16	F	21	1-1	37.4
17	F	21	1-2	27.0
18	F	23	1-1	27.5
19	F	25	1-1	34.2
20	F	28	1-1	30.4
21	F	28	1-1	32.8
22	F	28	1-1	31.1
23	M	1	1-1	33.3
24	M	1	1-2	22.7
25	M	2	1-1	31.5
26	M	2	1-2	29.6
27	M	3	1-1	31.4
28	M	4	1-1	31.5
29	M	4	1-1	27.7
30	M	5	1-1	36.6
31	M	5	1-1	35.6
32	M	6	1-2	21.2
33	M	9	1-1	29.3
34	M	9	1-2	33.9
35	M	9	1-2	41.1
36	M	11	1-2	24.1
37	M	12	1-1	30.8
38	M	12	1-2	34.4
39	M	16	1-1	31.3
40	M	18	1-2	33.4
41	M	20	1-1	32.2
42	M	22	1-2	34.8
43	M	25	2-2	23.4
44	M	26	2-2	25.5
45	M	30	1-1	31.9

D. Pregnant women or women taking oral contraceptives

	Age	Gc type	[Gc]
1	17	1F-2	35.1
2	17	2-1F	31.0
3	20	1S-1S	54.8
4	21	1S-1S	48.9
5	21	2-1S	39.1
6	22	2-1S	55.6
7	23	1S-1S	45.2
8	26	1F-1S	39.6
9	26	1F-1S	45.2
10	28	1S-1S	49.4
11	28	1S-1S	52.1

To minimize measurement error or bias, concentrations were determined by one individual only, who was not aware of the donor's type. The mean coefficient of variation of samples tested in triplicate was 5.5%.

RESULTS

Data from all individuals tested are given in table 1. Although data from pregnant women and women taking oral contraceptives were excluded from the twin sample and from further analysis, they are included separately in this table to illustrate the effect of estrogens on Gc concentration. Gc subtyping was done on all individuals sampled, but the small numbers in some categories precludes useful analysis at the subtype level. Twin subtypes are given for their relevance to zygosity determination.

Table 2 summarizes the Gc concentrations found in various groups and subgroups. The "twins" in table 2 are actually the first member of each pair, listed alphabetically, chosen to eliminate genetic correlation in the comparison of concentration levels across subgroups. Ages and Gc concentrations of twins and controls are compared in figure 1.

Table 3 presents the analysis of variance results, examining the effects of sex, age (two groups: < 9 years, and ≥ 9 years), and genotype at the Gc structural locus (Gc 1-1, 1-2, and 2-2). The analysis was conducted following the procedure of an unbalanced three-way classification of data as detailed in Searle [11] and Neter and Wasserman [12].

Table 4 presents the results of conventional variance decomposition and heritability estimation along with computational methods and references. To determine whether or not the electrophoretic differences at the structural locus account for the total genetic variability of Gc concentration, we also computed the additive and dominance effects due to the Gc locus using the genotype-specific mean values. These estimates, together with the gene frequencies of the Gc-1 and Gc-2 alleles, were subsequently used to decompose the total genetic variability of Gc concentration into one component due to Gc electrophoretic differences and a second component due to an undetermined number of genetic loci, each with small effects, that is, polygenes. The APPENDIX describes the theoretical basis for this partitioning of heritability, and table 5 summarizes the numerical results of the analysis.

Comparison of Twins and Controls (Table 2 and Fig. 1)

The distributions of Gc concentrations in twins and in controls are comparable as are their age distributions. Within twins and within controls there are no significant differences in Gc concentration with sex or with age, nor are there age and sex differences in the combined data. The only significant difference, most apparent in the combined data but also clear in twins or in controls alone, is the effect of Gc genotype on concentration: mean Gc in 2-2 individuals is less than in 1-1 individuals and 1-2 individuals have intermediate values.

Analysis of variance (table 3) corroborates these findings in the combined sample of 89 individuals (the first twin from each of 31 MZ pairs and 13 DZ pairs plus the 45 controls). Neither the effect of age nor the effect of sex is significant at a 5% probability level, nor are age $\times$ sex, age $\times$ genotype, or sex

TABLE 2

SUMMARY OF MEAN Gc CONCENTRATION (mg/100 ml ± STANDARD ERROR) BY AGE, SEX, AND Gc-GENOTYPE IN TWINS, CONTROLS, AND TWINS PLUS CONTROLS

	Gc 1-1	Gc 1-2	Gc 2-2	Total
A. First twin per pair (random sample):				
Females (< 9 yrs)	32.5 ± 1.5 (3)*	27.5 ± 0.3 (2)	...	30.5 ± 1.5 (5)
Females (≥ 9 yrs)	31.3 ± 0.8 (15)	27.3 ± 1.2 (6)	27.5 ± 0.3 (3)	29.8 ± 0.7 (24)
Males (< 9 yrs)	...	...	26.7 (1)	26.7 (1)
Males (≥ 9 yrs)	30.8 ± 1.2 (9)	26.6 ± 1.7 (3)	21.7 ± 3.3 (2)	28.6 ± 1.3 (14)
Total	31.3 ± 0.6 (27)	27.1 ± 0.8 (11)	25.4 ± 1.5 (6)	29.4 ± 0.6 (44)
B. Unrelated controls:				
Females (< 9 yrs)	33.0 ± 0.7 (4)	31.2 ± 1.9 (2)	26.0 (1)	31.4 ± 1.1 (7)
Females (≥ 9 yrs)	32.1 ± 1.1 (9)	32.5 ± 3.1 (4)	27.5 ± 1.0 (2)	31.6 ± 1.1 (15)
Males (< 9 yrs)	32.5 ± 1.1 (7)	24.5 ± 2.6 (3)	...	30.1 ± 1.6 (10)
Males (≥ 9 yrs)	31.1 ± 0.5 (5)	33.6 ± 2.2 (6)	24.5 ± 1.1 (2)	31.2 ± 1.4 (13)
Total	32.2 ± 0.5 (25)	31.2 ± 1.5 (15)	26.0 ± 0.8 (5)	31.3 ± 0.7 (45)
C. Combined sample of first twin per pair plus controls:				
Females (< 9 yrs)	32.7 ± 0.7 (7)	29.3 ± 1.3 (4)	26.0 (1)	31.0 ± 0.9 (12)
Females (≥ 9 yrs)	31.6 ± 0.6 (24)	29.4 ± 1.6 (10)	27.5 ± 0.4 (5)	30.5 ± 0.6 (39)
All females	31.9 ± 0.5 (31)	29.4 ± 1.2 (14)	27.2 ± 0.4 (6)	30.6 ± 0.5 (51)
Males (< 9 yrs)	32.5 ± 1.1 (7)	24.5 ± 2.6 (3)	26.7 (1)	29.8 ± 1.5 (11)
Males (≥ 9 yrs)	30.9 ± 0.8 (14)	31.3 ± 1.9 (9)	23.1 ± 1.6 (4)	29.9 ± 0.9 (27)
All males	31.5 ± 0.7 (21)	29.6 ± 1.8 (12)	23.8 ± 1.4 (5)	29.9 ± 0.8 (38)
Total	31.7 ± 0.4 (52)	29.5 ± 1.0 (26)	25.7 ± 0.8 (11)	30.3 ± 0.4 (89)
D. Pregnant women and women taking oral contraceptives	47.9 ± 1.9 (7)	40.2 ± 5.4 (4)	...	45.1 ± 2.4 (11)

* Sample sizes are shown in parentheses. Mean coefficient of variation of Gc concentration measured in triplicate in a partial sample was 5.5 ± 2.2%.

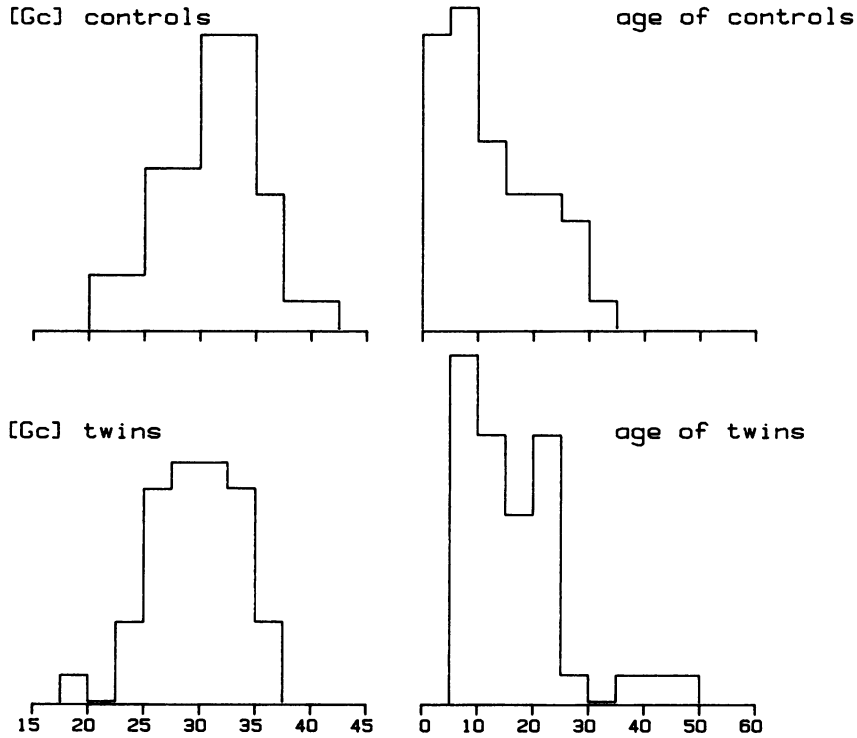

FIG. 1.—Comparison of Gc concentration (mg/100 ml) and age distributions between twins and controls.

× genotype interactions. Only the effect of the Gc genotypes is statistically significant (F-ratio with 2 and 73 df, respectively, is 9.88, $P < .001$). Further, since the mean Gc concentration for heterozygotes is roughly intermediate to the mean of the two homozygotes, the effect on concentration of the two alleles at the Gc locus may be largely additive in nature.

The gene frequencies in males, in females, and in the combined sample (shown in table 5) are very similar to the ones reported for other Caucasian populations (e.g., [5]), suggesting that the participants in this study were an unbiased sample from the population.

As noted above (see also table 2), pregnant women and women taking oral contraceptives have increased levels of Gc. The data are not sufficient in this case, however, to establish whether the various Gc types are affected differentially by estrogens.

Conventional Heritability Analysis (Table 4)

There is considerable controversy over the best estimate of heritability or whether any estimate accurately reflects the relative roles of genotype and environment. To deal with these issues, at least in part, we have applied several different methods to compute the degree of genetic determination of Gc concen-

TABLE 3

ANALYSIS OF VARIANCE EXAMINING THE EFFECT OF SEX, AGE, AND GENOTYPE ON Gc CONCENTRATION IN 89 RANDOM INDIVIDUALS

Source of variation	df	Sum of squares	Mean sum of squares	F-ratio
Sex*	1	7.81	7.81	0.58
Age*	1	0.01	0.01	0.001
Genotype*	2	350.15	175.08	12.92†
Two-factor interactions	5	100.02	20.00	1.52
Residual	79	1038.15	13.14	...

NOTE: Individuals include 45 unrelated controls and the first twin member from each of 31 MZ and 13 DZ twin pairs. Analysis of variance was conducted according to the computational procedures for three-way classified unbalanced data as detailed in Searle [11] and Neter and Wasserman [12].

* Since none of the two-factor interactions (sex × age, sex × genotype, or age × genotype) is significant, the F-ratios for each effect were computed by re-estimating the mean sum of squares by pooling the last two sources of variation.

† $P < .001$; none of the other F-ratios is significant.

tration using intrapair vs. interpair comparisons of MZ and DZ twin pairs. The computational basis for these calculations is given in table 4. Since Gc levels do not vary with age, we have not included this variable in our analysis but we *have* partitioned the data according to sex, although our analysis in table 3 suggests no significant sex-effect on Gc concentrations. Note that the number of male DZ pairs is small (only five) making estimates for this category somewhat unreliable.

Summarizing, the interpair variance, a rough measure of overall population variance, is almost equal for DZ and MZ pairs in the combined data, whereas

TABLE 4

ESTIMATES FROM TWIN DATA OF VARIANCE, CORRELATION COEFFICIENT, AND HERITABILITY OF Gc CONCENTRATION, IGNORING GENETIC VARIABILITY AT THE Gc LOCUS

A. Variance and intraclass correlation coefficients

	Sample size		Intrapair variance		Interpair variance		Intraclass correlation	
	n_{MZ}	n_{DZ}	V_{MZ}	V_{DZ}	V_{MZ}'	V_{DZ}'	r_{MZ}	r_{DZ}
Females	21	8	3.1	5.2	26.1	15.9	0.79	0.50
Males	10	5	1.0	14.8	19.2	42.8	0.90	0.49
Combined	31	13	2.4	8.9	24.8	24.3	0.82	0.46

B. Heritability

	$(V_{DZ} - V_{MZ})/V_{DZ}$	$2(r_{MZ} - r_{DZ})$	$(r_{MZ} - r_{DZ})/(1 - r_{DZ})$
Females	0.40	0.56	0.57
Males	0.93	0.83	0.81
Combined	0.73 ± 0.20	0.72	0.67

NOTE: Computational methods—(1) intrapair variance = $\Sigma(a_i - b_i)^2/2n$, (2) interpair variance = $[\Sigma(a_i + b_i)^2/2 - (\Sigma a_i + \Sigma b_i)^2/2n]/(n - 1)$, (3) intraclass correlation = $[(2.) - (1.)]/[(2.) + (1.)]$ (all of the above according to Snedecor and Cochran [13]), (4) variance of $(V_{DZ} - V_{MZ})/V_{DZ}$ is approximately $2(V_{MZ}/V_{DZ})^2 n_{DZ}^2(n_{MZ} - 1)(n_{MZ} + n_{DZ} - 4)/[n_{MZ}^2(n_{DZ} - 3)^2(n_{DZ} - 5)]$, according to Cavalli-Sforza and Bodmer [18].

TABLE 5

CONTRIBUTION OF ELECTROPHORETIC DIFFERENCES AT THE Gc LOCUS TO THE VARIABILITY IN Gc CONCENTRATION (mg/100 ml)

A. Summary of values from tables 2 and 4 (twins plus controls)

	Females	Males	Both sexes
Mean [Gc] in 1-1 genotype	31.9 (31)*	31.5 (21)	31.7 (52)
1-2 genotype	29.4 (14)	29.6 (12)	29.5 (26)
2-2 genotype	27.2 (6)	23.8 (5)	25.7 (11)
Intrapair MZ variance (V_{MZ})	3.1 (21)	1.0 (10)	2.4 (31)
Intrapair DZ variance (V_{MZ})	5.2 (8)	14.8 (5)	8.9 (13)

B. Estimates of:

	Females	Males	Both sexes
Gene frequency of Gc-1	.75	.71	.73
Gene frequency of Gc-2	.25	.29	.27
Additive variance at the Gc locus, σ_{ma}^2	2.22 (30.3%)*	3.70 (17.2%)	2.79 (18.2%)
Dominance variance at the Gc locus, σ_{md}^2	0.01 (0.1%)	0.66 (3.1%)	0.09 (0.6%)
Polygenic variance at the Gc locus, σ_{pg}^2	1.99 (27.1%)	16.19 (75.2%)	10.05 (65.4%)
Environmental variance, σ_e^2	3.12 (42.5%)	0.98 (4.5%)	2.43 (15.8%)
Total phenotypic variance, σ_t^2	7.34 (100%)	21.53 (100%)	15.36 (100%)
Genetic variability of [Gc] due to the Gc locus, $\sigma_{ma}^2 + \sigma_{md}^2$	2.33 (30.4%)	4.36 (20.3%)	2.88 (18.8%)
Total genetic component of [Gc], $\sigma_{ma}^2 + \sigma_{md}^2 + \sigma_{pg}^2$	4.22 (57.5%)	20.55 (95.5%)	12.93 (84.2%)
Contribution of allelic effects at the Gc locus to total phenotypic variance, $(\sigma_{ma}^2 + \sigma_{md}^2)/\sigma_t^2$	30%	20%	19%

* Nos. in parentheses are sample sizes.
Nos. in parentheses are percent of total phenotypic variance.

the intrapair variance for DZ pairs is substantially higher than for MZ pairs. This is expected if Gc concentration is under genetic control since the MZ intrapair variance is mainly due to environmental effects whereas the DZ intrapair variance includes both genetic and environmental contributions. The intraclass correlation coefficient, a more robust measure of genetic effects, is high in MZ twins and is nearly twice as great in MZ twins than in DZ twins. The observed intraclass correlation coefficients are consistent with a trait under significant genetic control.

The heritability estimates support the conclusion drawn from the intraclass correlation coefficients: they are similar to each other and are consistently high, 67%–73% in the combined sample. One feature is particularly striking: in all cases heritability in females is substantially less than in males. This difference may be accounted for by the additional environmental effect of estrogens, although neither pregnancy nor estrogen therapy was reported by these women.

We conclude that the variation in Gc concentration is essentially the result of genetic factors but that the genetic factors are more predominant in males than in females.

Decomposition of Total Heritability (Table 5)

Conventional heritability estimates do not take into account the possibility that the intrapair variance in DZ twins may be due in part to genotypic differences between DZ twin members at the Gc structural locus. Therefore, we estimated the additive and dominance effects of the Gc electrophoretic alleles from the pooled samples of males and females using the computational procedure described in the APPENDIX. In conjunction with the gene frequencies, then, these values provide estimates of additive and dominance variance due to polymorphic differences at the Gc locus (referred to as σ_{ma}^2 and σ_{md}^2, respectively). Taken together, these two statistics establish the contribution of the genotypic differences at the Gc structural locus to the variation in Gc concentration. The estimated values for these variances are shown in table 5 for females, for males, and for the combined sample.

Following the theory outlined in the APPENDIX, we can estimate four components of the total phenotypic variance in Gc concentration: variance due to additive effects of the Gc alleles (σ_{ma}^2), variance due to dominant effects of the Gc allele (σ_{md}^2), variance due to additional polygene effects (σ_{pg}^2), and residual (environmental) variance (σ_e^2). The absolute values of these components and their contribution as a percent of the total phenotypic variance are shown in table 5. In summary, this table indicates that while the overall genetic component of phenotypic variation in Gc concentration ($\sigma_{ma}^2 + \sigma_{md}^2 + \sigma_{pg}^2$) is 84% in the combined sample, 19% of the phenotypic variation is due to polymorphic allelic differences at the Gc structural locus. As in the conventional heritability estimates, table 5 also suggests that the total variation in Gc concentration due to nongenetic factors is much greater in females (43%) than in males (5%). The contribution of the allelic effects at the Gc locus to the total phenotypic variation [$(\sigma_{ma}^2 + \sigma_{md}^2)/\sigma_t^2$] is larger in females (30%) than in males (20%), but this difference may not be significant considering the small sample size of the DZ twin pairs.

DISCUSSION

We have attempted to answer three questions in this study. First: Is Gc concentration affected by age, by sex (exclusive of estrogens), or by Gc genotype? Second: To what extent is the variance in Gc concentration genetically determined? Third: Can the heritable component of the variance in Gc concentration be explained by the polymorphic, electrophoretic alleles at the structural locus?

The average Gc concentration in our combined twins and controls was 30.3 ± 0.4 mg/100 ml. This is comparable to values given by others using similar immunological methods: 32.8 ± 1.2* [14], 28.8 ± 1.0 [15] and 34.0 ± 0.5 [2], although less than other reported values: 44.0 ± 1.3 [16] and 52.5 ± 2.4

* Mean ± standard error; published standard deviations have been converted to standard errors.

[4]. We failed to detect any effect of age on Gc concentration over an age range of 1 to 45 years, and males and females did not differ significantly after pregnant women and women taking oral contraceptives were excluded. These findings confirm those of other research groups [2, 14].

However, we did detect an effect of Gc type on Gc concentration: Gc 2-2 individuals (25.7 ± 0.8, no. = 11) had significantly less Gc than did 1-1 individuals (31.7 ± 0.4, no. = 52) and 1-2 individuals had intermediate levels (29.5 ± 1.0). Most published studies of Gc concentration have not addressed this problem or have had sample sizes too small for statistical tests.

Is there actually less Gc in 2-2 and 1-2 individuals or is this difference simply the result of reduced antigen-antibody affinity? There are carbohydrate differences between the two Gc's that might affect either plasma half-life or immunologic activity or both [17]. At present, we cannot distinguish between these two possibilities, but as far as heritability is concerned, this uncertainty is immaterial since either phenomenon would affect Gc variance equally. The question is, however, of biologic significance.

Ignoring genotypic variability at the structural locus, we found a high heritability of Gc concentration, roughly 70%, independent of which particular method was used to calculate this quantity. As expected of a trait under genetic control, the intrapair correlation coefficient was nearly twice as great in MZ twins (r_{MZ} = .82) as in DZ twins (r_{DZ} = .46). This is true in both males and females and in the combined data. However, the heritability in females was significantly less than in males, also independent of method of calculation. We interpret this to mean that estrogen effects are still present in tese women even though pregnant women and women taking oral contraceptives were excluded presumably. This does not rule out the possibility of other environmental influences in women, but the most parsimonious explanation for these findings, given the profound effect of estrogens on Gc concentration, is of hormonal effects in women but not in men.

The conventional heritability measures employed in this study make several common assumptions (see Vogel and Motulsky [18] for a review). First, it is assumed that the twin pairs are a random sample of the population and that mating is random. It is also assumed that MZ twins do not seek out a more similar environment than do DZ twins.

The similarity of the twins to the controls, the similarity of the Gc "gene frequencies" to those expected in Caucasians, and the fact that the MZ intrapair correlation coefficient is nearly twice that of the DZ pairs corroborate these assumptions in this study.

Second, it is assumed that the covariance of genotype with environment is negligible and that there is no interaction between environment and genotype. In the case of Gc, a simple Mendelian trait, these are reasonable assumptions. Third, it is assumed that there is no *dominant* polygenic effect on the trait. The variance decomposition method discussed in the following section confirms this assumption. Finally, it is assumed that the effect of measurement error on the variance is small. The mean coefficient of variation for our assay was 5.5%, not negligible certainly, but small relative to the interpair and intrapair variances.

Even under these assumptions, the conventional methods for determining heritability underestimate the genetic contribution to the total phenotypic variability, especially in our combined sample, because the residual variance due to nongenetic causes is added to the denominator twice in some procedures (see Cavalli-Sforza and Bodmer [18]). This is confirmed by our revised estimate (84%) obtained using the variance decomposition method.

This new method addresses the third question: Is the genetic variability in Gc concentration due mainly to the polymorphism found at the Gc structural locus? Our answer is "no," since only 21% of the genetic component of the variation in Gc concentration ($\sigma_{ma}^2 + \sigma_{md}^2 + \sigma_{pg}^2$ in table 5) is due to differences in the Gc electrophoretic types ($\sigma_{ma}^2 + \sigma_{md}^2$).

The biological implications of our findings are the following. First, quantitative variation in Gc is not strongly influenced by environment, thus arguing against a significant effect of sunlight, vitamin D status, and so forth. Nor is Gc concentration affected by age or by sex except that in females hormonal effects are observed. Second, the electrophoretic types differ in mean concentration but whether this represents an immunologic difference or an actual concentration difference has not been established. Finally, in spite of these differences, electrophoretic differences at the Gc structural locus account for only a small fraction of the total genetic variance in Gc, that is, about one-fifth. Therefore, quantitative variation in Gc concentration can be decomposed into an additive component due to allelic differences at the Gc locus, a polygenic component, and environmental effects, with no evidence for a dominant genetic component. The polygenic component of the variation in Gc concentration may relate to protein synthesis, or to clearance or to other genetically mediated influences.

APPENDIX
A VARIANCE DECOMPOSITION METHOD TO ESTIMATE THE EFFECT OF POLYMORPHIC QUALITATIVE VARIATION ON THE HERITABILITY OF A QUANTITATIVE TRAIT

RANAJIT CHAKRABORTY

Let X be a quantitative trait that is under genetic control. As an example, consider a protein for which electrophoretic variation exists in the population. It is postulated that this electrophoretically detected polymorphic locus is also a major gene responsible for the character X. For simplicity, let us assume that there are three genotypes, G_1, G_2, and G_3, corresponding to the protein locus (G_2 is the heterozygote, and G_1 and G_3 are the two alternate homozygotes). Further, let us assume that X has a polygenic component as well, so that the variable X can be represented by

$$X = M + P + E , \tag{A1}$$

where M is the contribution of the major locus, P is the polygenic contribution, and E is a residual component (environmental and/or nongenetic factors). These components are also assumed to be independently distributed. Denoting the main effects of genotypes G_1, G_2, and G_3 by a, d, and $-a$, respectively, following classical quantitative genetic theory (e.g., Falconer [20]), we have

$$\bar{M} = a(p - q) + 2dpq, \bar{P} \text{ and } E = b,$$
$$V(M) = \sigma_{ma}^2 + \sigma_{md}^2, V(P) = \sigma_{pg}^2, \text{ and } V(E) = \sigma_e^2, \tag{A2}$$

where σ_{ma}^2 = additive genetic variance = $2pq[a + d(q - p)]^2$, at the major locus, σ_{md}^2 = variance of dominance deviation = $(2pqd)^2$, at the major locus, and p, q are the allele frequencies at the major locus ($p + q = 1$).

With data on genotypes and quantitative measurements for a sample of monozygotic (MZ) and dizygotic (DZ) twin pairs and with an additional sample of random individuals, we estimate the variance components σ_{ma}^2, σ_{md}^2, σ_{pg}^2, and σ_e^2. As in classical twin analysis, we assume that E, the residual component, includes no common familial environmental factors that are responsible for determination of X.

In a random sample (e.g., taking one twin member per pair in combination with other control individuals), if $\bar{X}_1$, $\bar{X}_2$, and $\bar{X}_3$ denote the mean values on the X measurements for the three genotypes G_1, G_2, and G_3, respectively, the estimates of a and d for equation (A2) are given by

$$\hat{a} = (\bar{X}_1 - \bar{X}_3)/2$$
$$\hat{d} = \bar{X}_2 - (\bar{X}_1 + \bar{X}_3)/2 . \tag{A3}$$

The estimates for σ_{ma}^2 and σ_{md}^2 can then be obtained by substituting these expressions for a and d in equation (A2) together with the estimated p and q values that can be obtained by the gene-counting method from the proportion of individuals with genotypes G_1, G_2, and G_3. This will complete the process of estimating the genetic variance contributed by the major locus.

If (x_{11}, x_{21}), (x_{12}, x_{22}), . . . represent the X-measurements for n_{MZ} monozygotic twin pairs, the intrapair variance of MZ twins, $V_{MZ} = \Sigma(x_{1i} - x_{2i})^2/2n_{MZ}$, has expectation σ_e^2. Therefore V_{MZ}, the intrapair variance for MZ twins, provides an estimate for the component σ_e^2. Under these assumptions, all of the estimators are unbiased.

From n_{DZ} dizygotic twin pairs, we can similarly compute the intrapair variance for DZ twins. Denoting this by V_{DZ}, simple algebra yields the following expectation for V_{DZ}:

$$E(V_{DZ}) = \frac{1}{2}\sigma_{ma}^2 + \frac{3}{4}\sigma_{md}^2 + \frac{1}{2}\sigma_{pg}^2 + \sigma_e^2 . \tag{A4}$$

Replacing σ_{ma}^2, σ_{md}^2, and σ_e^2 by their respective estimators (as obtained above), equation (A4) provides the estimator for σ_{pg}^2, that is, the polygenic component of variance of X.

Admittedly, the above estimation procedure assumes that the environmental similarities between members of a twin pair are no larger than that between two random individuals. This assumption is not universally accepted, and it remains a common problem in all twin-related study designs. Using data on twins reared together and apart, such an assumption may be partially relaxed and the above logic for estimating the variance components can be extended to disentangle the confounding effects of common environmental variance between twin members. This variance decomposition method should also be applicable, in principle, to data on other pairs of relatives (e.g., sib-pairs, cousin-pairs, etc.), demonstrating the effect of a recognizable major locus on a quantitative trait with mixed mode of inheritance.

REFERENCES

1. DAIGER SP, SCHANFIELD MS, CAVALLI-SFORZA LL: Group-specific component (Gc) proteins bind vitamin D and vitamin D metabolites. *Proc Natl Acad Sci USA* 72:2076–2080, 1975
2. BOUILLON R, VAN BAELEN H, DEMOOR P: The measurement of the vitamin D-binding protein in human serum. *J Clin Endocrinol Metab* 45:225–231, 1977
3. DAIGER SP, MILLER M, ROMEO G, PARSONS M, CAVALLI-SFORZA LL: Vitamin D-binding protein in Williams syndrome. *N Engl J Med* 298:687–688, 1978
4. HADDAD JG, WALGATE J: Radioimmunoassay of the binding protein for vitamin D and its metabolites in human serum. *J Clin Invest* 58:1217–1222, 1976
5. MOURANT AE, KOPEĆ AC, DOMANIEWSKA-SOBCZAK K: *The Distribution of the Human Blood Groups*, 2nd ed. London, Oxford Univ. Press, 1976, pp 686–695
6. CONSTANS J, VIAU M: Group-specific component: evidence for two subtypes of the Gc^1 gene. *Science* 198:1070–1071, 1977
7. CONSTANS J, CLEVE H: Group-specific component. *Hum Genet* 48:143–149, 1979
8. MILLER ME, OPHEIM K, RAISYS V, MOTULSKY AJ: Theophylline metabolism: variation and pharmacogenetics. *Clin Pharmacol Ther*. In press, 1984
9. DAIGER SP, CAVALLI-SFORZA, LL: Detection of genetic variation with radioactive ligands. II. Genetic variants of vitamin D-labeled group-specific component (Gc) proteins. *Am J Hum Genet* 29:593–604, 1977
10. CLEVE H, PATUTSCHNICK W, NEVO S, WENDT GG: Genetic studies on the Gc subtypes. *Hum Genet* 44:117–122, 1978
11. SEARLE SR: *Linear Models*. New York, John Wiley, 1971, pp 477–479
12. NETER J, WASSERMAN W: *Applied Linear Statistical Models*. Homewood, Ill., Irwin, 1974, pp 632–633
13. SNEDECOR GW, COCHRAN WG: *Statistical Methods*. Ames, Iowa Univ. Press, 1967, pp 294–296
14. COPPENHAVER D, KUEPPERS F, SCHIDLOW D, ET AL.: Serum concentrations of vitamin D binding protein (group-specific component) in cystic fibrosis. *Hum Genet* 57:399–403, 1981
15. BARRAGRY JM, FRANCE MW, CARTER ND, ET AL.: Vitamin-D metabolism in nephrotic syndrome. *Lancet* Sept 24:629–632, 1977
16. MIHAS AA, HIRSCHOWITZ BI: Group-specific component levels (Gc) in chronic liver disease. *J Med* 9:109–115, 1978
17. SVASTI J, BOWMAN BH: Human group-specific component. Changes in electrophoretic mobility resulting from vitamin D binding and from neuraminidase digestion. *J Biol Chem* 253:4188–4194, 1978
18. VOGEL F, MOTULSKY AG: *Human Genetics. Problems and Approaches*. Berlin, Heidelberg, New York, Springer-Verlag, 1979, pp 586–590
19. CAVALLI-SFORZA LL, BODMER WF: *The Genetics of Human Populations*. San Francisco, W. H. Freeman, 1971, pp 565–602
20. FALCONER DS: *Introduction to Quantitative Genetics*. New York, Ronald Press, 1960, pp 114–115, 134–136