Am J Hum Genet 26:13–34, 1974

Studies on Blood and Urine Glucose in Seminole Indians: Indications for Segregation of a Major Gene

R. C. ELSTON,[1,2] K. K. NAMBOODIRI,[1] H. V. NINO,[3] AND W. S. POLLITZER[2,4]

A high incidence of abnormal glucose tolerance and diabetes has been reported in recent years in many American Indian tribes: in Seneca Indians of New York [1], Cherokee Indians of North Carolina [2], Coushatta Indians of the Southwest [3], Pima Indians of Arizona [4], Cocopah Indians of Arizona [5], and Seminole Indians of Oklahoma [6]. The question of the nature of genetic factors in diabetes is an ancient one and various theories have been advanced over the past 50 years for its mode of inheritance, including autosomal recessive, codominant, dominant, sex-linked, and multifactorial (see table 1).

The aim of the present study is to elucidate the genetic transmission of diabetes from family data of the Seminole Indians on their reservations in Florida and Oklahoma. In particular, we examine the incidence of hyperglycemia and diabetes, as indicated by glucose levels in the serum and urine, and the association of these to some morphological variables and diet.

Historically, the Seminoles arose in Florida in the eighteenth and early nineteenth centuries, derived chiefly from the Creek Indians of neighboring Georgia and Alabama. One group, speaking Seminole, eventually settled in the Brighton area north of Lake Okeechobee; another, speaking the related Hitchiti dialect, settled in the Big Cypress and Dania regions near the Everglades. From the 1830s many of the Florida Seminoles were forced to migrate to Oklahoma, where they gradually came to occupy the area now known as Seminole County. Our investigation of serology and morphology confirms the historical trend that the Seminoles of today have undergone some admixture with both whites and Negroes [37, 38].

Received March 20, 1973; revised June 28, 1973.

This research was supported by U.S. Public Health Service research career development award 1-K3-GM-31, 732; research grant GM HD 16697 and grants GM 10601 and GM 15419 from the National Institute of General Medical Sciences; and by a postdoctoral fellowship from the Carolina Population Center.

[1] Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina 27514.

[2] Genetics Training Committee, University of North Carolina, Chapel Hill, North Carolina 27514.

[3] Nutritional Biochemical Laboratory, Center for Disease Control, 1600 Clifton Road NE, Atlanta, Georgia 30333.

[4] Department of Anatomy, University of North Carolina, Chapel Hill, North Carolina 27514.

TABLE 1

PROPOSED GENETIC MECHANISMS FOR DIABETES MELLITUS

Genetic Model and Source	Arguments Favoring Model	Arguments against Model
Autosomal dominant:		
Levit and Pessikova [7]	Dominant model with 10% penetrance fits data on 200 families studied; frequency of diabetes is almost same among parents and siblings of probands.	Very low penetrance could also be interpreted by other models.
Vallance-Owen [8]	Synalbumin, an albumin bound insulin inhibitor, is inherited as dominant trait and causes diabetes.	Not enough evidence to give a fair assessment of the role of inhibitors; synalbumin-positive persons with and without diabetes and diabetic persons with and without synalbumin antagonists are reported by Levin and Recant [9].
Roy et al. [10], O'Brien et al. [11]	Altered form of insulin (resistant to insulinase) is inherited as dominant and is found in both early and late onset cases.	No. families studied too few to permit generalization.
Pavel and Pieptea [12]	Many three-generational and a few four-generational families with high incidence of diabetes, no parental consanguinity, and diabetes often occurring only on one side of families favor dominant hypothesis.	Retrospective nature of study might have led to misreporting of data.
Autosomal recessive:		
Pincus and White [13], Post [14], Nilsson [15], Steinberg [16, 17], Barrai and Cann [18]	Occurrence of affected children to normal parents, steep increase in risk among offspring from families with none, one, or both parents affected, and increased consanguinity among parents of diabetics are all compatible with recessive inheritance. Nilsson [15] reports fit of his data to both recessive and dominant models with reduced penetrance; Barrai and Cann [18] report fit of Simpson's data [25] on juvenile diabetics to both recessive and polygenic models.	Juvenile diabetes cannot be easily interpreted by this model; incidence among parents of affected persons is as high as among siblings; only 50% of the offspring of conjugal diabetics are affected; incomplete penetrance of a recessive trait is difficult to prove or disprove.
Autosomal codominant:		
Cammidge [19], Harris [20], Stimmler and Elliot [21], Ehrlich and Martin [22] ...	Correlation between severe early onset cases and mild late onset cases caused by dosage differences of the same basic defect; insulin inhibitor shows codominant mode of inheritance.	Some predictions based on model not supported empirically: early onset cases should have on the average more parents affected than late onset cases; risk in siblings should be greater in early than in late onset cases; consanguinity among parents of early onset cases should be greater than among parents of late onset cases.

TABLE 1 (*Continued*)

Genetic Model and Source	Arguments Favoring Model	Arguments against Model
Sex-linked recessive or dominant: Penrose and Watson [23], Gronberg et al. [24]	Mother-son transmission of disease in some families and inequalities in sex ratio of diabetics thought to be evidence of sex-linked inheritance.	Elaborate family studies contradict sex-linked mode of inheritance; Penrose later retracted this view.
Multifactorial with equal and additive effects: Simpson [25–27], Neel et al. [28], Falconer [29], Thompson [30], Carter [31]	Compatible with family data; Neel et al. [28] found no bimodality in glucose tolerance test curves; positive correlation between risk in relatives and severity in index cases; increased risk with family incidence of affected; heritability of liability to diabetes based on incidence among first-degree relatives for all ages is about 36% according to Falconer [29] and about 50% according to Simpson [25–27].	Apparent unimodality does not disprove monogenic inheritance but only shows that more refined techniques in diagnosis of affected genotype and in analyses of parameters are needed to bring out gene effect presently confounded with other factors.
Lamy et al. [32]	Only a major gene with minor modifying genes can explain the inheritance in diabetes.	Practically impossible to distinguish this model from one-locus model with incomplete penetrance.
Gedda et al. [33]	Multiple alleles with recessive inheritance at a single locus can cause different forms of the disease; different combinations of these alleles at the same locus produce different degrees of severity in the disease.	Severe and mild cases occur side by side in the several families; chance occurrence of such double mutant alleles is very low; severity varies among monozygous twins who share the same gene constitution.
Others: Rimoin [34], Rimoin and Schimke [35]	Diabetes is a genetically heterogeneous group of disorders, which explains its clinical, physiological, and racial variability.	Most diabetics suffer from common and basic maladies; that different rare mutants pleiotropically produce the same common malady is a bit unrealistic.
Swift et al. [36]	Heterozygotes of many rare recessive fatal or near fatal syndromes may show higher than normal risk to diabetes; higher incidence of diabetes is expected in families with such rare syndromes as cystic fibrosis, Friedreich's ataxia, ataxia telangiectasia, Fanconi's anemia; positive correlation found between heterozygosity to Fanconi's anemia and diabetes in a retrospective study of families with Fanconi's anemia.	With recessive traits it is difficult to identify heterozygotes except for parents of affected persons.

SUBJECTS AND METHODS

The population survey on the Seminoles was conducted during 1964–1966. The Seminole Indian families were urged to participate in a health survey and those who responded are believed to be representative of the population in that area. In 1964, 263 persons from Brighton and Big Cypress reservations of Florida, and in 1966, 263 persons from the Seminole County, Oklahoma, were examined. Blood and urine specimens were collected from all persons 6 years of age or older approximately 1 hr after ingesting 75 g of glucose in the form of Glucola (generously supplied by Ames Co.). Blood was collected in EDTA and the plasma was shipped with fluoride to the Clinical Chemistry Laboratory at Chapel Hill for glucose determination by the method of Hoffman adapted for autoanalyzer. Urine was tested immediately with Clinistix, and positive specimens were checked with Clinitest (Ames Co.).

Several physical traits were also measured during the examination. Skin color was determined on all subjects by reflectometry (Photovolt model 410), using tristimulus and red filters separately. Height, weight, and head, nose, and face measurements were made on all subjects aged 16 or above. Blood types were determined at the National Institutes of Health and the University of North Carolina for the Florida and Oklahoma samples. Serum proteins and red cell enzymes were determined for the Florida group only at the Human Genetics Laboratory, Ann Arbor, Michigan. The distribution of the morphological traits and blood marker systems have already been reported by Pollitzer et al. [37, 38].

While detailed and precise data on diet were not collected, the high carbohydrate intake of the Florida Seminole today contrasts strikingly with the protein rich animal food, in addition to the fruits and vegetables, eaten by the Indians in 1880 [39]. Recent analysis of the diet of Oklahoma Seminoles [6] shows fat content and carbohydrates well above that of the neighboring whites.

Serum glucose is measured in milligrams per 100 ml of serum. The data for those aged 30 or above are examined for indications of tri- or bi-modality by fitting mixtures of normal distributions to them; this is accomplished by maximum-likelihood methods using a search algorithm that finds a local maximum on the likelihood surface [40]. In doing this, we assume that the two- or three-component distributions have the same variance but different means.

Demographic data such as age, marital status, sibship size, and details about spouse and children were collected, and subjects were queried about the presence of diabetes in themselves and their families. From the kinship data, pedigrees were constructed for family studies.

In Oklahoma, the family data are comprised of 30 pedigrees containing 77 males and 79 females; in Florida there are 27 pedigrees with 89 males and 97 females. Separating the Florida sample by reservation, the Brighton sample has 12 families of 74 persons and Big Cypress has 15 families with 112 persons. Most of the families are two-generational and only a few three-generational. The raw data used to determine the fit of a single autosomal locus model for serum glucose levels are deposited with a documentation service.* Each individual has a unique identification number, and the identification numbers of his two parents are given together with his record. This method of presenting the data is not only concise but also very convenient for computer handling. Com-

---

* See NAPS document no. 02261 for seven pages. Order from ASIS/NAPS, c/o Microfiche Publications, 305 East 46th Street, New York, New York 10017. Remit with order for each NAPS document number $1.50 for microfiche or $5.00 for photocopies *up to 30 pages* and 15¢ per page for each additional page over the first 30 pages. Make checks payable to Microfiche Publications.

puter programs to reconstruct pedigrees from data in this format have been written in *Fortran* by E. B. Kaplan (personal communication, 1973) and by Landré et al. [41].

The method of pedigree analysis used has been described by Elston and Stewart [42]. Here we assume a two-allele, one-locus autosomal model and a lognormal phenotypic distribution for the different genotypes. In Elston and Stewart's notation the phenotypic distribution is denoted $g_u(x)$, where in this paper $u = AA$, Aa, or aa. Thus the basic assumptions of the model underlying the analysis are the following:

1. For each person in the pedigree, his or her logarithmic serum glucose level, after allowing for a linear age correction, comes from one of three normal distributions with common variance. In the population of mating individuals, the proportions of such measures belonging to each of the three distributions are denoted ψ_{AA}, ψ_{Aa} *and* ψ_{aa}; these three proportions add up to unity. (In terms of a genetic hypothesis, these three distributions are the phenotypic distributions corresponding to the three genotypes AA, Aa, and aa.)

2. To determine in a probabilistic sense the distribution to which each offspring belongs, three transmission probabilities are defined: $\tau_{AA\,A}$, the probability that the genotype AA transmits an A allele to his offspring; $\tau_{Aa\,A}$, the probability that the genotype Aa transmits an A allele to his offspring; and $\tau_{aa\,A}$, the probability that the genotype aa transmits an A allele to his offspring. (The usual Mendelian hypothesis corresponds to the special case in which these probabilities are, respectively, $\tau_{AA\,A}=1$, $\tau_{Aa\,A}=1/2$, $\tau_{aa\,A}=0$.) There are also the three complementary probabilities that an a allele is transmitted: $\tau_{AA\,a} = 1 - \tau_{AA\,A}$, $\tau_{Aa\,a} = 1 - \tau_{Aa\,A}$, $\tau_{aa\,a} = 1 - \tau_{aa\,A}$.

Using both the formulation given by Elston and Stewart [42] for this model and the maximum-likelihood program developed by Kaplan and Elston [40], the likelihood of each set of data (maximizing over unknown parameters) is obtained under different hypotheses. Each particular hypothesis is equivalent to one or more restrictions on the underlying model, and so causes the likelihood to be smaller; the reduction in the log likelihood can be used to test for departure from the hypothesis. In this way, using the well known asymptotic properties of the likelihood ratio criterion, two different genetic hypotheses can be tested: (1) Hardy-Weinberg equilibrium is tested by comparing the maximum likelihood obtained when ψ_{AA} and ψ_{Aa} are allowed to be arbitrary (but $\psi_{aa} = 1 - \psi_{AA} - \psi_{Aa}$) with that obtained when the restriction $\psi_{Aa} = 2\sqrt{\psi_{AA}\,\psi_{aa}}$ is imposed on the estimates; and (2) the goodness-of-fit of the Mendelian hypothesis to the data is tested by comparing the maximum likelihood obtained when the transmission probabilities are arbitrary (but constrained between 0 and 1) with that obtained when they are fixed at the theoretical Mendelian values of 1, 1/2, and 0. In each case, twice the difference in log likelihoods is expected, asymptotically, to be distributed as a χ^2 under the appropriate hypothesis. It should be noted that the absolute value of the log likelihood has very little meaning. In the model we are assuming it is a density and not a probability; hence it can be positive or negative, depending very much on sample size and scale of measurement.

RESULTS

The summary statistics for serum glucose are given in table 2 for the two locations. The females in both places show a higher mean than the males, though the difference is not significant. The overall mean is lower in Florida (150.6) than in Oklahoma (182.6) and the difference is significant ($t = 3.44$, $P < .01$). The distributions are asymmetrical with significant positive skewness and leptokurtosis. An increase in serum glucose levels with increasing age is noticed in both the

TABLE 2

SERUM GLUCOSE DISTRIBUTION IN SEMINOLE INDIANS

GROUP AND AGE (YR)	MALES			FEMALES		
	N	Mean	SD	*N*	Mean	SD
Florida:						
0–9	25	127.52	39.08	19	122.78	29.00
0–19	27	122.55	29.56	33	130.91	28.81
20–29	10	125.70	31.17	21	147.00	55.14
30–39	10	141.20	46.07	12	147.58	40.39
40–49	19	161.10	58.84	22	191.13	94.17
50–59	7	163.42	34.77	19	202.89	87.45
60–69	12	199.75	87.38	7	172.00	24.82
70–79	3	150.00	45.03	2	244.50	7.77
80–89	2	109.00	26.87	3	128.00	39.85
90–99	1	186.00	...	...	...	...
Overall	116	143.29	52.58	138	156.87	65.30
Skewness		1.99**			2.56**	
Kurtosis		7.73**			12.81**	
Range		77–390			82–530	
Oklahoma:						
0–9	15	109.13	33.23	13	131.61	68.95
10–19	18	106.78	25.24	25	108.16	33.40
20–29	4	106.75	29.94	16	112.50	33.38
30–39	10	207.60	166.42	22	165.82	93.05
40–49	13	256.92	177.56	20	243.05	149.36
50–59	11	194.82	113.13	20	321.65	161.48
60–69	13	221.84	217.06	11	264.45	187.17
70–79	1	246.00	...	3	161.33	35.64
80–89	...	...	...	2	200.00	82.02
90–99	...	...	...	...	...	...
Overall	85	172.64	140.33	132	189.01	133.89
Skewness		2.75**			1.96**	
Kurtosis		11.64**			6.75**	
Range		33–880			54–726	

NOTE.—Skewness $= \sqrt{N}m_3/m_2^{3/2}$ and kurtosis $= Nm^4/m_2^2$, where m_i is the ith sample moment about the mean and N is the sample size.

** Significant at .01 level.

Florida and Oklahoma groups and the regression of serum glucose on age was found to be significant. A sharp rise in mean serum glucose level is noticed for those aged 30 or above in Oklahoma, while the rise appears to occur at a slightly later age in Florida.

The results of the urine test are given in table 3. Similar to the serum glucose results, the Florida sample shows a lower percentage of Clinistix positives (13.5% for males and 9.9% for females) as compared to the Oklahoma sample (36.5% for males and 33.6% for females). It is interesting to note that males have a higher percentage of glycosuria than females. This is in contrast to the elevated

TABLE 3

URINE GLUCOSE TEST RESULTS IN FLORIDA AND OKLAHOMA SEMINOLE INDIANS

TEST RESULT	MALES		FEMALES	
	N	%	*N*	%
		Florida		
Clinistix:				
Negative	97	86.5	120	90.1
Positive	15	13.5	13	9.9
Total	112	100.0	133	100.0
Clinitest:*				
Negative	2	13.3	3	23.1
Trace	3	20.0	2	15.4
1+	4	26.7	3	23.1
2+	3	20.0	0	...
3+	0	...	2	15.4
4+	3	20.0	3	23.1
Total	15	100.0	13	100.1
		Oklahoma		
Clinistix:				
Negative	54	63.5	87	66.4
Positive	31	36.5	44	33.6
Total	85	100.0	131	100.0
Clinitest:*				
Negative	6	19.4	11	25.6
Trace	3	9.7	3	7.0
1+	5	16.0	3	7.0
2+	4	12.9	7	16.3
3+	3	9.7	4	9.2
4+	10	32.3	15	34.9
Total	31	100.0	43	100.0

* Performed on those positive for Clinistix.

serum glucose means of females in both places. This suggests that women probably have a higher renal threshold than men, as has also been found by others [1–6, 16].

Clinitests were performed on those who gave positive results for Clinistix in both places. The results given in table 3 show that about 75%–85% of Clinistix positives were also positive with Clinitest. Even though Clinitest was intended as a refinement over Clinistix, Clinistix is specific for glucose, while Clinitest picks up a variety of substances other than glucose in urine and is therefore not always consistent with Clinistix results.

The incidence of hyperglycemia (arbitrarily defined here as a serum glucose level of 160 mg/100 ml or more) and its association with urine test results, age,

and weight are given in table 4. According to the above criteria, about 29% of males and 39% of females in Florida and 29% of males and 44% of females in Oklahoma are hyperglycemic. Among males who are hyperglycemic, 41% in Florida and 84% in Oklahoma are also glycosuric based on Clinistix results. Among hyperglycemic females, 23% in Florida and 62% in Oklahoma have

TABLE 4

INCIDENCE OF HYPERGLYCEMIA AND ITS ASSOCIATION WITH GLYCOSURIA, AGE, AND WEIGHT

	FLORIDA				OKLAHOMA			
	Males		Females		Males		Females	
	≤160*	>160	≤160	>160	≤160	>160	≤160	>160
N	82	34	83	55	60	25	73	59
%	71.4	28.6	60.9	39.1	70.6	29.4	55.7	44.5
Glycosuria:								
No. Clinistix tested	80	32	81	52	60	25	73	58
No. positive	2	13	1	12	10	21	8	36
% positive of total tested	2.6	40.6	1.2	23.1	16.6	84.0	10.9	62.1
Mean age (SD)	26.3 (20.7)	45.5 (22.1)	26.6 (18.3)	42.3 (19.8)	29.9 (21.1)	46.8 (18.1)	27.5 (18.1)	47.4 (17.2)
Mean weight (SD)	127.1 (52.9)	153.6 (48.9)	120.1 (40.1)	155.7 (37.4)	131.5 (53.1)	171.4 (46.4)	126.1 (42.1)	160.8 (48.5)

* Serum glucose (mg/100 ml).

glycosuria. The mean age for the hyperglycemic group is similar (42–47 years) in both sexes and both locations. The hyperglycemic group has a strikingly higher mean weight than the "normal" group.

The correlations between serum glucose levels and some morphological traits and age are given in table 5 for the sexes separately. The correlation of serum glucose with weight and age is strongly positive and significant ($P \leqslant .01$) for both sexes in Florida and for females in Oklahoma. The only other significant correlation was a negative trend between skin color (red filter) and serum glucose in Florida males. The trend is negative for all the groups with respect to these two variables, that is, decreasing serum glucose levels with lighter skin color.

Initially, search for a major gene was performed by fitting a mixture of two or three normal distributions to the population data. In order to minimize age effects, only serum glucose levels on those aged 30 and above were used for this analysis, since little change occurs after that age.

In view of the high skewness of the distribution, a transformation was made of the scale of measurement. The choice of the scale of measurement was made by comparing the cumulative plot of the original data (i.e., the rank of each measure-

TABLE 5

CORRELATIONS OF SERUM GLUCOSE WITH PHYSICAL TRAITS AND AGE IN FLORIDA AND OKLAHOMA SEMINOLE INDIANS

TRAIT	FLORIDA		OKLAHOMA	
	Males ($N = 116$)	Females ($N = 137$)	Males ($N = 56$)	Females ($N = 110$)
Height	.137	.148	—.103	—.172
Weight	.265**	.353**	.077	.257*
Skin color:				
Red filter	—.252*	—.025	—.028	—.116
Tristimulus filter	—.088	—.003	—.092	—.201
Head index	.164	.134	.177	.002
Face index	—.082	.067	—.085	—.164
Nose index	—.001	—.120	—.078	.017
Age	.373**	.384**	.163	.426**

* Significant at .05 level.
** Significant at .01 level.

ment, appropriately scaled on the Y axis, plotted against the measurement itself on the X axis) with best fitting cumulative mixtures of two normal distributions taking (1) the serum glucose on the original scale, (2) the square root of the serum glucose value, and (3) the logarithm of the serum glucose value. The plots are shown in figure 1 for the Oklahoma female population. It is clear from the plots that the logarithmic transformation of the data gives the best fit to the observed values. This is found to be true in both sexes and in both populations, and has been found elsewhere by other investigators [43, 44].

Even though we tried to fit a mixture of three normal distributions to the data, in every case we found that two of the distributions turned out to be practically identical and so in effect we were able to fit only a mixture of two different distributions to each sex in the two groups. The estimated parameters of these mixtures are given in table 6. It is clear from the estimates of the means that for the Florida females and for males and females of Oklahoma, approximately the *same* two normal distributions fit the data. Moreover, the proportion of admixture (comparable to genotype frequencies) is approximately the same for the two sexes in Oklahoma. The Florida males, however, are discrepant in that the distribution with the higher mean has a mean value much lower than the other three cases, and the proportion of persons who belong to this distribution is much larger than the proportion of females that fall in the corresponding distribution. In view of the small sample size, this discrepancy could be due to chance. If we fit two normal distributions to the data on the Florida males with the constraint that the distribution with the higher mean has a fixed mean value of 5.95 (385 mg/100/ml on original scale), the same as that of Oklahoma males, the resulting estimates of admixture closely correspond to the proportions of admixture in the Florida females. In view of the fact that the higher distribution is only 3% of a total sample of 54, it is hardly surprising if the values are poorly estimated. So it is

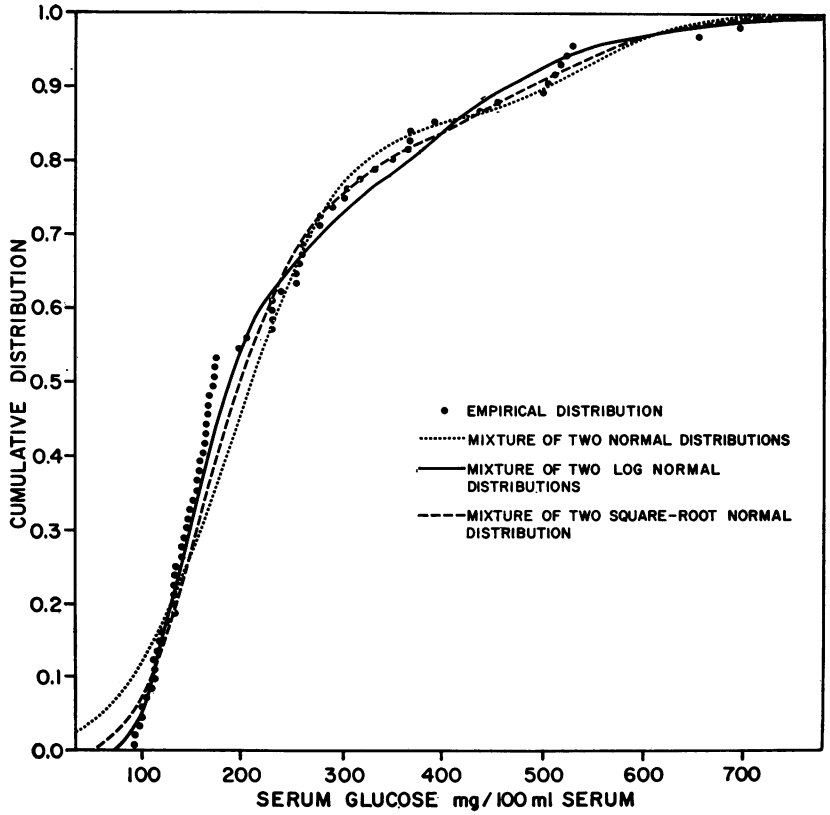

FIG. 1.—Serum glucose distribution of Oklahoma Seminole females aged 30 and above

TABLE 6

ESTIMATES OF PARAMETERS FROM FLORIDA AND OKLAHOMA SEMINOLE INDIAN SAMPLES FOR A MIXTURE OF TWO NORMAL DISTRIBUTIONS TO SERUM GLUCOSE VALUES

	MALES			FEMALES		
	Mean	Variance	Proportion Admixture	Mean	Variance	Proportion Admixture
Florida:						
Distribution 1	4.91	0.05	0.78	5.08	0.08	0.94
Distribution 2	5.49*	0.05	0.22	5.98	0.08	0.06
Oklahoma:						
Distribution 1	4.84	0.15	0.71	5.01	0.09	0.68
Distribution 2	5.95	0.15	0.29	5.96	0.09	0.32

NOTE.—Serum glucose values (natural logarithmic scale) approximately 1 hr after challenge. Sample size: Florida, 54 males and 67 females; Oklahoma, 49 males and 78 females.

* When the mean for distribution 2 in Florida males was fixed at 5.95, the same as the corresponding value in Oklahoma males, the proportions of admixture turned out to be .97 and .03, respectively, for distributions 1 and 2. These values closely correspond to the admixture proportions in the Florida females.

reasonable to conclude that the data fit approximately the *same* two distributions in all the four groups. These results are highly suggestive in that this is exactly what would be expected if in each case the two distributions were made up of individuals who differed with respect to segregation at one autosomal locus with dominance. The means for the two distributions in such a case would be approximately the same in the two populations; and within each population the proportions of admixture would be approximately the same, although between populations the proportions would differ. These results suggest the possibility of a major gene causing high serum glucose that occurs at a much higher frequency in Oklahoma than in Florida.

In view of the fact that the pedigrees consist of members of both sexes and different ages, the family data in the two locations were examined carefully to decide not only which scale of measurement of serum glucose is best suited for pedigree analyses but also whether the sexes could be pooled for the analyses and which regression model of serum glucose on age would be best suited.

Various regression models were fitted to the serum glucose values (both on the original scale of measurement and after logarithmic transformation) in order to allow for the age effect. The percentage of the total variation explained by the regression models is given in table 7 for the Oklahoma pedigree sample. It is

TABLE 7

PERCENTAGE OF TOTAL VARIATION OF SERUM GLUCOSE ACCOUNTED FOR BY REGRESSION MODELS IN OKLAHOMA

Terms in Regression Model and Serum Glucose Scale	Males	Females
Age:		
Original	12.6	19.9
Logarithmic	17.6	27.9
$(Age)^2$:		
Original	11.1	17.8
Logarithmic	14.7	22.9
Age + $(Age)^2$:		
Original	14.1	21.6
Logarithmic	18.9	29.5

clear that the logarithmic transformation explains more of total variation than the untransformed values; similar results were found for the Florida sample also. In view of this and the earlier finding that a mixture of two lognormal distributions best fit the population data, the logarithmic transformation was chosen for further analyses as the scale of measurement for serum glucose. It is also seen in table 7 that the regression on age accounts for more of the total variation in serum glucose than regression on $(age)^2$ in both sexes. The quadratic model with both terms does not substantially increase the amount of variation explained;

therefore it was decided to use the linear model of regression of logarithmic transformation of serum glucose on age in pedigree analysis. A similar trend is found in the Florida pedigree sample also.

The difference in the regression of serum glucose on age between sexes in Oklahoma is not significant ($t = 1.08$, df $= 214$). The difference in the residual mean square of the two sexes is also not significant ($F = 1.18$, df $= 82,130$). Similar results were obtained for the two locations and so the same procedures were followed for pedigree analyses in both places; in particular, a common linear regression on age is included in the model for simultaneous estimation along with the other unknown parameters.

Using the test based on the likelihood-ratio criterion described above, there is no significant departure from Hardy-Weinberg equilibrium when the transmission probabilities are assumed to be fixed at the Mendelian values; the χ^2 values with 1 df are 0.8 in Oklahoma and 0.96 in Florida. Hence all subsequent analyses assume Hardy-Weinberg equilibrium as part of the model.

Results of testing the goodness-of-fit of the Mendelian hypothesis, assuming Hardy-Weinberg equilibrium and dominance [i.e., $g_{AA}(x) = g_{Aa}(x)$] for low serum glucose levels, are given in table 8 for the two locations. In the Oklahoma

TABLE 8

MAXIMUM-LIKELIHOOD ESTIMATES OF PARAMETERS OF PEDIGREE ANALYSES FOR SERUM GLUCOSE: DOMINANCE ASSUMED

PARAMETER	OKLAHOMA		FLORIDA	
	Unrestricted	τ Fixed	Unrestricted	τ Fixed
$\tau_{AA\,A}$	1.00*	1.00	1.00*	1.00
$\tau_{Aa\,A}$	0.75	0.50	1.00*	0.50
$\tau_{aa\,A}$	0.12	0.00	0.11	0.00
ψ_{AA}	0.24	0.34	0.21	0.20
ψ_{Aa}	0.50	0.49	0.50	0.50
ψ_{aa}	0.26	0.17	0.29	0.30
Mean of AA and Aa	4.80	4.75	4.95	4.90
Mean of aa	5.88	5.76	5.54	5.34
Common variance	0.134	0.133	0.070	0.059
Regression on age	0.010	0.011	0.004	0.006
Log likelihood	28.22	27.09	97.27	93.58
Difference in log likelihood	1.13		3.69	

NOTE.—Hardy-Weinberg equilibrium assumed. $\tau_{u(k)}$ is the probability that the genotype u transmits allele k to the offspring. ψ_u is the probability of genotype u in the population. Thus $\Sigma\psi_u = 1$ and the assumption of Hardy-Weinberg restriction is equivalent to the restriction $\psi_{Aa} = 2\sqrt{\psi_{AA}\psi_{aa}}$.

* Estimate converged to a bound.

sample the unrestricted $\tau_{AA\,A}$, the probability that the AA genotype transmits an A allele to offspring, converged to the upper limit of 1. Therefore, with respect to this estimate, the likelihood is not at a local maximum. For this reason the approximate distribution of the test statistic, comparing the log likelihoods obtained when the τ are unrestricted and when they are fixed, lies somewhere be-

tween a χ^2 with 2 df and 3 df. Even with 2 df, the χ^2 (2.26, twice the difference in log likelihood) is not significant, indicating no significant departure from the Mendelian hypothesis. The data points and the genetic hypothesis are graphically depicted in figure 2 for the Oklahoma sample.

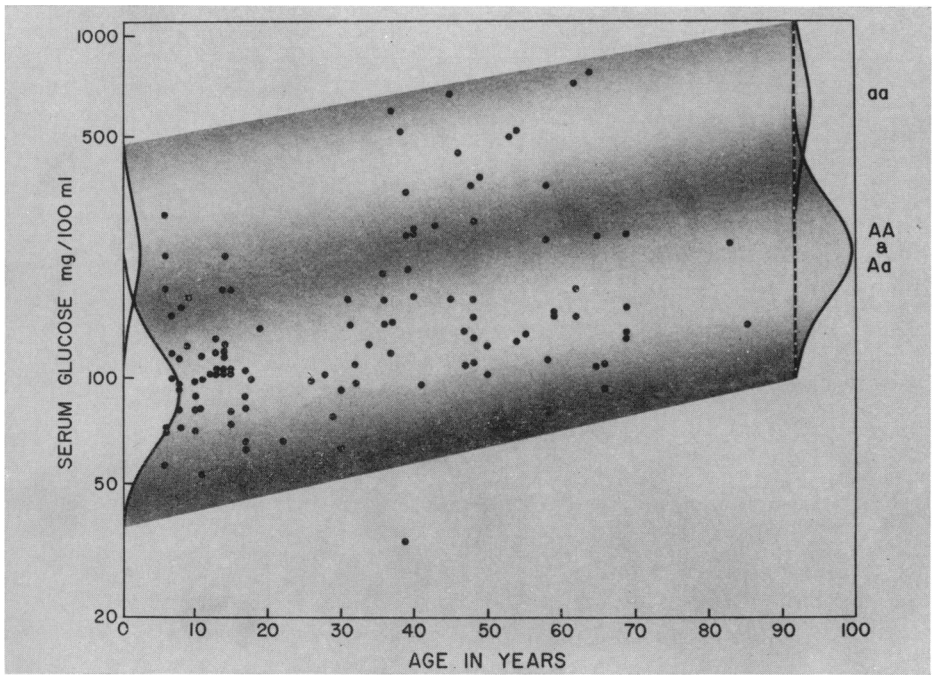

FIG. 2.—Bivariate distribution of age and serum glucose in Oklahoma Seminole pedigree sample. The genetic model of one major locus with dominance is shown superimposed on the data points.

In the Florida pedigree sample, among the three unrestricted transmission probabilities (τ), only one is at a local maximum; the other two ($\tau_{AA\,A}$ and $\tau_{Aa\,A}$) converged to the upper bound, thus reducing the degrees of freedom for the χ^2 to somewhere between 1 and 3. The statistic is again twice the difference in log likelihoods (7.38), and the significance level is about 6% for a χ^2 with 3 df and 1% for a χ^2 with 1 df. Thus in Florida this simple hypothesis does not fit the data adequately.

All analyses were repeated in the two samples without assuming dominance: the heterozygote is assumed to come from a distribution with a distinct mean different from the two homozygote mean values [i.e., $g_{AA}(x) \neq g_{Aa}(x) \neq g_{aa}(x)$]. The χ^2 value testing whether the three-distribution model is better than the two-distribution model can be obtained from tables 8 and 9: using the log likelihoods obtained when the τ are fixed, the result is not significant in Oklahoma [$\chi^2_{(1)} = 2(27.58—27.09) = 0.98$] but highly significant in Florida [$\chi^2_{(1)} = 2(97.56—$

TABLE 9

MAXIMUM-LIKELIHOOD ESTIMATES OF PARAMETERS OF PEDIGREE ANALYSES FOR SERUM GLUCOSE: NO DOMINANCE

PARAMETER	OKLAHOMA		FLORIDA	
	Unrestricted	τ Fixed	Unrestricted	τ Fixed
$\tau_{AA\,A}$	1.00*	1.00	0.54	1.00
$\tau_{Aa\,A}$	0.53	0.50	0.29	0.50
$\tau_{aa\,A}$	0.40	0.00	0.00*	0.00
ψ_{AA}	0.32	0.28	0.43	0.96
ψ_{Aa}	0.49	0.50	0.45	0.04
ψ_{aa}	0.19	0.22	0.12	0.00
Mean of *AA*	4.71	4.70	4.80	4.89
Mean of *Aa*	5.33	4.85	5.26	5.11
Mean of *aa*	6.02	5.89	5.77	5.55
Common variance	0.089	0.131	0.201	0.069
Regression on age	0.011	0.013	0.003	0.056
Log likelihood	30.44	27.58	102.23	97.56
Difference in log likelihood	2.86		4.65	

NOTE.—Hardy-Weinberg equilibrium assumed. See table 8 for definition of terms.
* Converged to a bound.

93.58) = 7.96]. This indicates that both two- and three-distribution Mendelian models fit the data equally well in Oklahoma but not in Florida.

The maximum-likelihood estimates obtained for the pedigrees when dominance is not assumed are given in table 9. In Oklahoma, when the τ are fixed, the estimate of the mean of the heterozygote distribution *AA* is much closer to the mean of the homozygote *AA* (.15 units away) than to the mean of the other homozygote *aa* (1.04 units away, nearly seven times the distance to *Aa* mean). This suggests there is a dominance component in the total variation. Since there is no significant difference between two and three distributions, the results in Oklahoma suggest that if a major locus is involved in serum glucose levels, the high serum glucose value is probably inherited as a recessive. This also corroborates the collapse of the mixture of three distributions to two, to fit the population data. Comparing the two log likelihoods in table 9 for the Oklahoma sample, the χ^2 with 2 or 3 df is not significant (χ^2=5.72, significance level 12% with 3 df and 6% with 2 df), suggesting adequate agreement with the hypothesis of segregation at an autosomal locus. The proportion* of total variation of serum glucose

* Percentage of total variation due to the three genotypes at one major autosomal locus is computed using estimates in place of the parameters as $100\ \sigma_g^2/(\sigma_g^2 + \sigma_e^2)$, where σ_e^2 is the common environmental variance and

$$\sigma_g^2 = \sum_{u=1}^{3} \psi_u(\mu_u - \bar{\mu})^2;$$

values (corrected for age effect) accounted for by the hypothesis of one autosomal locus with dominance is 54.5%. With three distributions, the one-locus hypothesis accounts for 85.2% of the total variation in serum glucose.

In the Florida sample, when dominance is not assumed, one autosomal locus clearly does not fit the data. Among the estimates when the τ are fixed, the homozygote *aa* frequency nearly converged to its lower limit of zero. The χ^2 obtained by comparing the two log likelihoods is highly significant ($\chi^2 = 9.30$, $P < .01$ with 2 or 3 df), indicating poor fit of the Mendelian hypothesis to the data.

DISCUSSION

The present study indicates a high incidence of abnormal glucose tolerance in the Seminole Indians, especially those of Oklahoma. Although the blood glucose level varied with the time lapsed after the challenge and the nature of the previous meal, in population screening studies a sharp distinction in the blood glucose levels between the normal and abnormals is obtained within 1 hr after the challenge [1, 3, 28, 45, 46]. While in the United States clinically diagnosed diabetics in both sexes and all ages account for only about 2% of the general population [46], among Oklahoma Seminoles they account for about 13%. Only 3% of the Florida Seminoles are reported diabetics. High incidence of hyperglycemia has been reported in recent years in many North American Indian tribes (table 10). The incidence figures have to be treated with some caution, since the studies vary in definition of hyperglycemia, fasting or nonfasting state when glucose load is given, amount of glucose ingested, time lapsed between ingestion of glucose and blood drawing, and use of whole blood or plasma. Even with all these limitations, the very high incidence of this defect in North American tribes as compared to the general U.S. population is clearly evident. The Oklahoma Seminoles compare well with the Pima Indians and Coushatta Indians while the Florida come closer to the Cocopah Indians than to the general U.S. population (table 10).

It is puzzling how such a serious disease is so frequent in some populations. One possible explanation proposed by Neel [47], assumes a selective advantage some time in the past. According to him, the diabetic genotype may have been beneficial in the hunting and gathering periods of human existence, without leading to diabetes. Diabetics actually overproduce effective circulating insulin at certain stages. This greater availability of circulating insulin may have facilitated effective carbohydrate storage under conditions when plentifulness of food varied widely. With acculturation, this genotype may have resulted in overproduction of insulin without prompt utilization, resulting in a subsequent development of insulin antagonists, consequent free insulin shortage, and diabetes.

ψ_u is the probability of genotype u in the population, μ_u is the mean log serum glucose for genotype u, and

$$\bar{\mu} = \sum_{i=1}^{3} \psi_u \mu_u.$$

TABLE 10

INCIDENCE OF ABNORMAL GLUCOSE TOLERANCE IN SOME NORTH AMERICAN INDIAN POPULATIONS

Population and Reference	*N*	Age	Fasting (F) or Nonfasting (NF)	Glucose Load (g)	Time Elapsed after Ingestion (hr)	Hyperglycemic Status (mg/100 ml)*	% Hyperglycemics†
Seneca Indians, New York [1]	209	≥25	NF	75	1	≥200 (P)	33.5
Cherokee Indians, North Carolina [2]	448	≥34	NF	1‡	2	≥150 (WB)	29.0
Coushatta Indians, Southwest [3]	171	All ages	F	100	1	≥160 (WB)	26.1
Pima Indians, Arizona [4]	1,881	≥5	NF	75	2	≥160 (P)	23.5
Cocopah Indians, Arizona [5]	182	≥5	NF	75	2	≥160 (P)	17.0
Seminole Indians, Oklahoma [6]	301	≥14	NF	...	2	>130 (WB)	17.0
Seminole Indians, Florida (present study)	254	≥6	F	75	1–1½	≥200 (P)	16.1
Seminole Indians, Oklahoma (present study)	217	≥6	F	75	1–1½	≥200 (P)	25.8
General U.S. population [46]	6,672	≥18	...	50	1	≥180 (WB)	8.4§

* P = plasma glucose; WB = whole blood glucose.

† Includes clinically known diabetes.

‡ Per kg body weight.

§ No known diabetes included in the sample; incidence of diabetes for all ages and both sexes in general U.S. population is near 2%.

Steinberg [16] reported that the average number of offspring was least when neither parent was diabetic and most when both parents were diabetic, strongly suggesting differential fertility among the diabetics.

In the present study, Oklahoma Indians show a much higher incidence of diabetes and hyperglycemia than their Florida counterparts. The diet has a high carbohydrate and fat content in both places. The admixture with the white and Negro populations is nearly the same in both places, as indicated by morphology and serology [37, 38]. Most of the serological traits are in Hardy-Weinberg equilibrium (with the exception of Rh and haptoglobin loci), suggesting no significant departures in factors such as mating pattern or selection, in the two places. A very reasonable explanation for the high incidence of elevated serum glucose levels in Oklahoma is that there is a major gene (probably from founder effect) segregating there that is either rare or nonexistent in Florida. Whether this gene is an important cause of diabetes in other populations is a moot point.

Females show a higher serum glucose mean than males. The difference is not significant but the trend is consistent with many recent studies [1–6]. However, there is a higher percentage of men who have glycosuria, suggesting a higher renal threshold for women than for men. Steinberg [16] reported similar findings in a large sample who participated in a U.S. health survey.

Although the importance of familial transmission in diabetes is fairly certain, the mode of inheritance of the disease is still unclear as is seen from the various proposed modes of transmission given in table 1. The most important impediment to genetic analysis is a lack of knowledge concerning the basic defect in diabetes and lack of a reliable marker for prediabetes. There is no simple genetic hypothesis which will explain all the available data. Nilsson [15] comments on the difficulties in distinguishing dominant and recessive inheritance when the gene frequency is high. From a study of a Swedish sample, he suggests that diabetes could be due to an autosomal recessive gene with a frequency of 30% and a lifetime penetrance of 70% in males and 90% in females, or an autosomal dominant gene with a frequency of 5% and a lifetime penetrance of 25% in males and 30% in females. In the Oklahoma sample, where the data fit the hypothesis of segregation at a major locus, analyses with two distributions showed that elevated serum glucose levels may be due to a recessive gene with a frequency of 41% in the population (table 8, τ fixed). However, virtually the same log likelihood is obtained if we assume the elevated levels are due to a dominant gene with a frequency of 9% in the population. In either case 17% of the population is estimated to belong to the elevated distribution, and in either case the transmission probabilities are Mendelian. The only reason for preferring the recessive hypothesis is that when a separate distribution is assumed for the heterozygote, its mean is much closer to the mean of the lower distribution (table 9, τ fixed).

Identical data have before now been interpreted to support different modes of inheritance [18, 25]. Simpson [25] concluded from incidence among relatives of juvenile diabetic probands that the data are compatible with multifactorial inheritance. Barrai and Cann [18], using Morton's [48] segregation analyses on the

same data, found that the segregation frequency among backcross matings is twice that of intercross matings and penetrance is as low as 25% in both mating groups. They claim these findings as support for the hypothesis that diabetes is transmitted recessively at a single locus. Edwards [49] has discussed the difficulties in differentiating between recessive inheritance with low penetrance and multifactorial inheritance. According to him the incidence of a polygenic disorder among sibs of probands approximates the square root of the population incidence of the disorder. Using Edward's formula, Barrai and Cann [18] find that the observed and expected incidences of diabetes among sibs of probands do not differ significantly, indicating that the data fit the expectation under a multifactorial model also.

The fact that there is no clear cut pattern of Mendelian segregation has driven many researchers to conclude that this disease has a polygenic mode of inheritance. It has long been recognized that "diabetes" may include multiple distinct entities implying genetic heterogeneity (see [35, 36]). It is also plausible that while most of these entities are governed by a recessive mode of inheritance, some may be governed by more than one locus. There is also obviously a strong interplay of environmental factors (especially diet) as well as age influencing the manifestation of the disease. However, as Carter [31] points out, single gene inheritance is an attractive idea, especially when the abnormality is suspected to be in a single plasma protein. There are a number of ways in which a gene mutation could effect insulin synthesis, secretion, transport, or action so as to produce carbohydrate intolerance.

In the present study we have tried to isolate a major genetic component and follow its segregation pattern in families. As discussed by Elston and Stewart [42], the method of pedigree analysis used is a natural extension of classical segregation analysis, the underlying model being less restrictive. There is no need to separate pedigrees into two-generational families and assume these families are independently sampled, or to assume that $\tau_{AA\ A} = 1$ and $\tau_{aa\ A} = 0$ as is usually done. Indeed, it is fair to say that classical segregation analysis corresponds to testing, in two-generational families, the hypothesis $\tau_{Aa\ A} = 1/2$ assuming $\tau_{AA\ A}$ and $\tau_{aa\ A}$ are known to be 1 and 0, respectively. Had this been done in the present analysis, the results would automatically have favored the monogenic hypothesis even more. Steinberg et al. [43] have suggested a different method of analysis specifically for the case where many individuals cannot be individually classified as "normal" or "affected" because of overlapping curves. Apart from requiring the two restrictions just mentioned, their method also depends on establishing a cutoff point above and below which the prior probabilities (ignoring all pedigree information) that an individual has a particular genotype are assumed constant for all individuals. The method we have used takes into account the prior phenotypic distributions in their entirety: an individual with a very high value automatically has a larger prior probability of belonging to the *"aa"* distribution than an individual with a moderately high value, even though both most probably belong to the *"aa"* distribution.

Bimodality has been demonstrated by Steinberg et al. [43, 50] for the log glucose tolerance levels of the Pima Indians. Among our two samples, the Oklahoma data similarly fit a mixture of two lognormal distributions, suggesting segregation at a major locus. There is a good fit of the family data to segregation at an autosomal locus, whether high serum glucose is assumed to be a recessive or a dominant trait. We must conclude that with our mostly two-generational data it is practically impossible, when heritability may be as low as about 50%, to distinguish between dominant and recessive inheritance patterns. Furthermore we must admit that it is never possible to *prove* a null genetic hypothesis by the kinds of statistical analyses we have performed: this is as true for qualitative traits as it is for quantitative traits. In the final analysis, biochemical and/or linkage evidence is always necessary. On the other hand, the fact that the Florida data depart significantly from the same model indicates the power of the method of pedigree analysis used: a simple single-locus autosomal model was not found to account for a sizeable proportion of the variation in serum glucose in a population where, on the basis of fitting mixtures of distributions, it could hardly exist. It is therefore unlikely that the apparent fit of the Oklahoma data to a monogenic model, accounting for 50% or more of the variation, is merely a statistical artefact. This does not, of course, imply that the remaining variation in serum glucose is necessarily environmental; a part of it could well be polygenic. Elston and Stewart [42] have indicated a method of analysis that would allow for the simultaneous estimation of a polygenic part, but it is not yet clear whether such an analysis is feasible with present-day computing facilities. In any case, it is difficult to see how the distributional properties of serum glucose in the Oklahoma Indians, as in the Pima Indians, can be accounted for merely on the basis of a large number of small additive factors. The skewness of the empirical serum glucose distribution (illustrated in fig. 1) surely indicates that at least one factor is considerably larger than all the others, and pedigree analysis indicates that segregation at an autosomal locus can account for it.

SUMMARY

A survey of the serum glucose levels of 254 Seminole Indians from Florida and 217 Seminole Indians from Oklahoma has shown that, like many North American Indian tribes, Seminoles have a high incidence of hyperglycemia: 16% in Florida and 26% in Oklahoma. While mean serum glucose levels are higher among females in both places, there is a higher percentage of glycosuria in males, confirming an earlier suggestion of a lower renal threshold for males.

Initial search for a major gene for serum glucose levels is performed by fitting a mixture of two lognormal distributions to the data on those aged 30 and above. The means for the two distributions are found to be approximately the same in the two places, and within each location the proportions of admixture are found to be similar in the two sexes. This is compatible with the hypothesis that there is a major gene governing serum glucose levels occurring in different frequencies in the two populations.

Pedigree analyses are performed on 30 pedigrees from Oklahoma and 27 pedigrees from Florida by the method of Elston and Stewart [42]. There is a good fit to segregation at an autosomal locus in Oklahoma with a recessive allele (frequency 41%) governing high serum glucose levels. There is an equally good fit for a dominant allele (frequency 9%) governing elevated serum glucose levels. The recessive hypothesis is preferred for the reason that when a separate distribution is assumed for the heterozygote, its mean is much closer to the mean of the lower distribution. In Florida, however, there is no adequate fit of the Mendelian hypothesis to the family data. Probable reasons for this discrepancy are discussed.

ACKNOWLEDGMENTS

The authors are deeply grateful to the large number of people who made this investigation possible: Lois Tillman in both Florida and Oklahoma; R. M. Menegaz-Bock, Margot Hall, Fred Hammer, Lollie Hindel, and Cassi Parrish in Florida; William H. Brown in Oklahoma; the Bureau of Indian Affairs and its agents, especially Paul Nelson and Louis McGahey; the Indian Health Service and its agents, especially Dr. W. F. Hill, Dr. Benjamin McBrayer, Calvin Beames, Charles Bailey, and Margaret Kilius; Dr. K. W. Navin and staff of the Seminole County Health Department; the former Chief Billy Osceola of the Florida Seminoles and Chairman John Brown of the Oklahoma Seminoles; Mary Jumper Johns, Judy Bill Gopher, Elsie McGirt, Mabel Lena Ward, and Benny Harjo; and the Seminole Indian people. We are grateful to Ames Company for the generous gift of Glucola.

REFERENCES

1. Doeblin TD, Evans K, Ingall GB, Dowling K, Chilcote ME, Elsea W, Bannerman RM: Diabetes and hyperglycemia in Seneca Indians. *Hum Hered* 19:613–627, 1969
2. Stein JH, West KM, Robey JM, Tirador DF, McDonald GW: The high prevalence of abnormal glucose tolerance in the Cherokee Indians of North Carolina. *Arch Intern Med* 116:842–845, 1965
3. Johnson JE, McNutt CW: Diabetes mellitus in an American Indian population isolate. *Tex Rep Biol Med* 22:110–125, 1964
4. Miller M, Burch T, Bennett PH, Steinberg AG: Prevalence of diabetes mellitus in the American Indians: results of glucose tolerance tests in the Pima Indians of Arizona. *Diabetes* 14:439–440, 1965
5. Henry RE, Burch TA, Bennett PH, Miller M: Diabetes in the Cocopah Indians. *Diabetes* 18:33–37, 1969
6. Mayberry RH, Lindeman RD: A survey of chronic disease and diet in Seminole Indians in Oklahoma. *Am J Clin Nutr* 13:127–134, 1963
7. Levit SG, Pessikova LN: The genetics of diabetes mellitus. *Proc Maxim Gorky Med Biol Inst* (Moscow) 3:132–147, 1934
8. Vallance-Owen J: Synalbumin in insulin antagonism. *Diabetes* 13:241–246, 1964
9. Levin ME, Recant L: Hormonal and biochemical studies in a diabetic family. *Ann Intern Med* 66:69–76, 1967
10. Roy CC, Elliot RB, Shapcott DJ, O'Brien D: Resistance of insulin to insulinase. A genetic discriminant in diabetes mellitus. *Lancet* 2:1433–1435, 1966
11. O'Brien D, Shapcott DJ, Roy CC: Further studies on an abnormal insulin of diabetes mellitus. *Diabetes* 16:572–575, 1967
12. Pavel I, Pieptea R: Etude sur la diabète héréditaire au cours de 3 et 4 générations successives. *Diabetologia* 2:281–286, 1966

13. PINCUS G, WHITE P: On the inheritance of diabetes mellitus. I. An analysis of 675 family histories. *Am J Med Sci* 186:1–14, 1933
14. POST RH: An approach to the question, does all diabetes depend upon a single genetic locus. *Diabetes* 11:56–65, 1962
15. NILSSON SE: On the heredity of diabetes mellitus and its interrelationship with some other diseases. *Acta Genet Statist Med* (Basel) 14:97–124, 1964
16. STEINBERG AG: Genetics and diabetes. *Excerpta Medica Int Congr Ser* 83:601–611, 1965
17. STEINBERG AG: Heredity in diabetes mellitus. *Postgrad Med* 43:145–147, 1968
18. BARRAI I, CANN HM: Segregation analysis of juvenile diabetes mellitus. *J Med Genet* 2:8–11, 1965
19. CAMMIDGE PJ: Heredity as a factor in the aetiology of diabetes mellitus. *Lancet* 1:393–395, 1934
20. HARRIS H: The familial distribution of diabetes mellitus: a study of the relatives of 1241 diabetic propositi. *Ann Eugen* 15:95–119, 1950
21. STIMMLER L, ELLIOT RB: Inheritance of a diabetic serum factor inhibiting normal utilization of insulin. *Lancet* 1:956–958, 1964
22. EHRLICH RM, MARTIN JM: Presence of synalbumin antagonist in siblings of diabetic children. *Diabetes* 15:400–405, 1966
23. PENROSE LS, WATSON EM: A sex linked tendency in familial diabetes. *Proc Am Diabetes Assoc* 5:165–177, 1945
24. GRONBERG A, LARSSON T, JUNG J: Diabetes in Sweden. A clinicostatistical, epidemiological and genetic study of hospital patients and death certificates. *Acta Med Scand* 477, suppl.:126–127, 1967
25. SIMPSON NE: The genetics of diabetes: a study of 233 families of juvenile diabetics. *Ann Hum Genet* 26:1–21, 1962
26. SIMPSON NE: Multifactorial inheritance: a possible hypothesis for diabetes. *Diabetes* 13:462–471, 1964
27. SIMPSON NE: Heritabilities of liability to diabetes when sex and age at onset are considered. *Ann Hum Genet* 32:283–303, 1969
28. NEEL JV, FAJANS SS, CONN JW, DAVIDSON RT: Diabetes mellitus, in *Genetics and the Epidemiology of Chronic Diseases*, edited by NEEL JV, SHAW MW, SCHULL WJ, Washington, D.C., Department of Health, Education and Welfare, Public Health Service, 1965, pp 105–132
29. FALCONER DS: The inheritance of liability to diseases with variable age of onset with particular reference to diabetes mellitus. *Ann Hum Genet* 31:1–20, 1967
30. THOMPSON GS: Genetic factors in diabetes mellitus studied by oral glucose tolerance test. *J Med Genet* 2:221–226, 1965
31. CARTER CO: Genetics of common disorders. *Br Med Bull* 25:52–57, 1969
32. LAMY M, FREZAL J, REY J: Hérédité du diabète sucre. *Jour Diabète* 2:5–15, 1961
33. GEDDA L, CASA D, BRENCI A: Chronan and the problems of anticepalin (or two family cases of diabetes). *Acta Genet Med Gemellol* (Roma) 16:217–228, 1967
34. RIMOIN DL: Genetics of diabetes mellitus. *Diabetes* 16:346–351, 1967
35. RIMOIN DL, SCHIMKE RN: *Genetic Disorders of the Endocrine Glands*. Saint Louis, Mosby, 1971, pp 150–216
36. SWIFT M, SHOLMAN L, GILMOUR D: Diabetes mellitus and the gene for Fanconi's anemia. *Science* 178:308–310, 1972
37. POLLITZER WS, RUCKNAGEL D, TASHIAN R, SHREFFLER DC, LEYSHON WC, NAMBOODIRI KK, ELSTON RC: The Seminole Indians of Florida: morphology and serology. *Am J Phys Anthropol* 32:65–82, 1970
38. POLLITZER WS, NAMBOODIRI KK, ELSTON RC, BROWN WH, LEYSHON WC: The Seminole Indians of Oklahoma: morphology and serology. *Am J Phys Anthropol* 33:15–30, 1970

39. MacCauley C: The Seminole Indians of Florida. Annual Report, Bureau of American Ethnology, Washington, D.C., 1884
40. Kaplan EB, Elston RC: A subroutine for maximum-likelihood estimation (MAXLIK). *Inst Statist Mimeo Ser* no. 823, Department of Biostatistics, Chapel Hill, N.C., 1972
41. Landré MF, Valat MT, Jutier P: Reconnaissance automatique des liens de parenté. Tracé automatique d'arbre généalogique (abstr.). *Excerpta Medica Int Cong Ser* 233:107, 1971
42. Elston RC, Stewart J: A general model for the genetic analysis of pedigree data. *Hum Hered* 21:523–542, 1971
43. Steinberg AG, Rushforth NB, Bennett PH, Burch TA, Miller M: On the genetics of diabetes. *Nobel Symp* 13:237–264, 1970
44. Rushforth NB, Bennett PH, Steinberg AG, Burch TA, Miller M: Diabetes in the Pima Indians. Evidence of bimodality in glucose tolerance distributions. *Diabetes* 20:756–765, 1971
45. Lilly Research Laboratories: *Diabetes Mellitus,* 7th ed. Indianapolis, E. Lilly Co., 1970
46. Gordon T: Glucose tolerance of adults, U.S. 1960–1962, *Public Health Service Publication* 1000, Ser 11, no. 2. Washington, D.C., Department of Health, Education and Welfare, 1964
47. Neel JV: Diabetes mellitus: a "thrifty" genotype rendered detrimental by "progress"? *Am J Hum Genet* 14:353–362, 1962
48. Morton NE: Segregation analysis in human genetics. *Science* 127:79–80, 1958
49. Edwards JH: The simulation of mendelism. *Acta Genet Statist Med* (Basel) 10:63–70, 1960
50. Steinberg AG, Rushforth NB, Bennett PH, Burch TA, Miller M: Preliminary report on the genetics of diabetes among the Pima Indians. *Adv Metab Disord* 1, suppl.:11–21, 1970