The Geographic and Racial Distribution of ABO and Rh Blood Types and Tasters of PTC in Puerto Rico

FREDERICK P. THIEME

University of Michigan

A LARGE number of studies (Boyd, 1939, Wiener, 1943) of population blood types frequencies have given the student of human variability a valuable tool for estimating group differences. This report,[1] from a sample of over 2,500 Puerto Rican adults, adds data to this already vast body of knowledge of world blood type and PTC non-taster frequencies. It also supplements data for Puerto Rico reported by Torregrossa (1945 a, b) for ABO and Rh frequencies. In addition to the report of island-wide frequencies, the variability of groups making up the total Puerto Rican population are examined. This intergroup analysis is based on a population breakdown according to racial phenotype and region of residence.

It is well known that within any human population random mating does not occur. Elements of human culture such as class structure, economic grouping, and the whole of social organization act to support and maintain selective systems. In addition, assortative systems of mating based on certain phenotypical characteristics are known to operate in human groups (Harris, 1930). Although the society and its attitudes and values set the pattern for these phenotype selection systems, the result is to reduce randomness of mating in a phenotypically varied population.

The Puerto Rican people, though less varied in origin than the population of the United States, have a development springing from lines originating in Southern Europe and Africa. In addition, the American Indian aborigines must have contributed a not inconsiderable genetic element which, though presently impossible to single out, may help to account for the island group heterogeneity for some characteristics.

Although mixtures between the originally quite distinct racial groups have gone on for a considerable time in Puerto Rico and even though it is said by

Received February 6, 1952.

[1] This paper reports on material gathered by the University of Puerto Rico Biological Survey which was a project in physical anthropology sponsored by the Social Science Research Center of the University of Puerto Rico. The field party consisted of Miss Joan Finkle, Mr. Rupert Murrill, and the author. Dr. H. L. Shapiro of the American Museum of Natural History acted as project supervisor. In addition, the author has been aided in carrying out statistical analysis by a research grant and a Summer Faculty Research Fellowship from the Rackham School of Graduate Studies of the University of Michigan.

many Puerto Ricans that race prejudice does not exist there, a rigid class structure and a strong and active system of race prejudice (Gordon, 1949–1950) have prevented the achievement of panmixia. In addition, the historical background of the migrations and economic pursuits of the various groups have acted so that the same proportions of the original groups have never existed in each of the various island regions. Even though Puerto Rico is a small island—a rectangle approximately 30 by 100 miles—and with considerable rates of internal migrations (U. S. Census Bull. #4, 1940), the regional differences have been maintained. Yet Puerto Rico cannot be considered unique in these respects, but probably is representative of many areas where immigration and subsequent race mixture has taken place.

The picture of significant genetic heterogeneity, if seen in Puerto Rico, would seem to indicate that population genetic surveys in other areas of the world should take into account phenotypic features that affect selective mating patterns and be based more often on representative samples. Otherwise the genetic characterization of whole populations on the basis of selected or conveniently gotten samples may be more misleading than constructive. In general, genetic surveys of populations should assume the existence of regional and social breeding isolates and should consequently give serious attention to insure that geographic and social sub-groups of the larger population are proportionally represented. Of course, the essential problem here is one of defining the population. Usually, however, due to cultural and historic features, we must work with groupings not readily definable or isolated. We are of necessity forced to work with given populations which need defining and description in terms of our problem. It would seem that random mating in human populations should not be assumed *a priori*, but follow a critical examination of the factors that may affect the frequency of each genetic characteristic studied.

METHODS

The sample reported here is a large portion of the total sample examined in the University of Puerto Rico Biological Survey, which studied the adult population of Puerto Rico during the period June 1948 to February 1949. In addition to the blood typing, 26 selected anthropometric measures, a dental record, several physiological characteristics, and some classificatory background material were noted for each individual. No comprehensive report of this project is yet published. The total sample of 3,562 adult Puerto Ricans (1,518 males, non-prisoners; 395 male prisoners; 1,649 females) was chosen according to a sample structure calculated before the survey began. On the basis of sex, age groups, and general economic status, the required number of individuals for each of 77 different municipios was calculated. A municipio is the area including and contiguous to a town and usually named for it. The municipio includes both urban and rural residences, and is highly variable

in size (populations from 169,247 to 860 for 1940) and densities (32,000 to 60 per square mile in 1940) of their population. Race was not used because the criteria employed in census enumeration are not clear for this.

The census data for 1935 and 1940 were used to extrapolate the numbers of the various categories which would be expected in June 1948. These calculations were provided through the courtesy of the Bureau of Vital Statistics of the Department of Health of the Government of Puerto Rico. The total numbers of adults between the ages of 20 and 45 inclusive was calculated as 744,693. The numbers for each of several age groups, residence categories, and sex and income groupings were calculated according to the 1940 proportions as applied to the new 1948 figure for each of the 77 island municipality regions. While this method of extrapolation has obvious deficiencies, it was a method of approximating the 1948 population structure which was considered better than using the 1940 figures. Where other information, such as the rates of increase resulting from the war, was known for specific localities, it was used in the calculations. An examination of the 1950 Census returns will ultimately test the accuracy of the 1948 predictions.

The sample structure thus calculated gives a representative sample of the whole island with each of the categories considered being included in our sample in the approximate proportion that they are found in the total population. Although this size sample (3,562) represents a little less than one-half of one per cent (0.48%) of the adult population and also, by being proportionally representative, gives very small numbers for some categories, the fact that no area or known social group was neglected gives a basis for determining island frequencies of observed characters and inter-population variability.

The sampling in each municipio was done so that no known bias favored the inclusion of any one segment of the population, and there is good reason to believe that the sample obtained is a random one within the categories selected. Because our sample gave proportions for urban-rural residence and nutritional background which match very closely known figures for Puerto Rico and because samples were gathered in each of the regions, we feel justified in assuming it to be representative.

Not all persons in the total sample of 3,562 were typed, nor were all that were typed for ABO also examined for Rh. Various difficulties in maintaining the supply of serum as well as the problem of getting blood to the typing laboratory in useable conditions made the number in the typed sample smaller. The number reported here does, however, represent the total numbers typed and none typed were not included. Because those not typed were not randomly spread through the sample, or not random by municipio, the blood type sample is not as accurately representative of the total population as the calculated sample. However, when the municipios are combined into regional districts, as has been done, no one section of the island is insufficiently represented in the

final sample. As will be noted later, the frequencies in a corrected sample designed to give the exact proportions of the calculated sample are not significantly different from the actual results (see table 12).

The regions used in this report are made up of ecologically similar municipios and these are listed in table 1. The location of the regions is indicated on the map in figures 1 and 2.

TABLE 1. MUNICIPIOS INCLUDED IN EACH OF THE 15 REGIONAL GROUPS

REGION NO.	MUNICIPIO
1	Aguadilla
	Camuy
	Hatillo
	Isabela
	Quebradillas
2	Arecibo
	Rio Piedras
	San Juan
3	Barceloneta
	Dorado
	Manati
	Toa Alta
	Toa Baja
	Vega Alta
	Vega Baja
4	Bayamon
	Catano
	Guaynabo
5	Carolina
	Loiza
	Rio Grande
	Trujillo Alto
6	Cieba
	Culebra
	Fajardo
	Humacao
6 (Con'd)	Luquillo
	Maunabo
	Naguabo
	Vieques
	Yabucoa
7	Arroyo
	Guayama
	Juana Diaz
	Patillas
	Penuelas
	Salinas
	Santa Isabel
8	Ponce
9	Guanica
	Guayanilla
	Lajas
	Sabana Grande
	Yauco
10	Coamo
	Villalba
11	Aguada
	Anasco
	Cabo Rojo
	Hormigueros
	Moca
	Rincon
	San German
12	Mayaguez
13	Adjuntas
	Ciales
	Jayuya
	Lares
	Las Marias
	Maricao
	San Sebastian
	Utuado
14	Aibonito
	Barranquitas
	Cayey
	Cidra
	Comerio
	Corozal
	Morovis
	Naranjito
	Orocovis
15	Aguas Buenas
	Caguas
	Gurabo
	Juncos
	Las Piedras
	San Lorenzo

The fifteen regions defined here were not chosen solely for this blood type frequency analysis but are the groupings used in the broader survey to denote similar environmental areas which also have a degree of historical similarity. Although a small island, Puerto Rico enjoys a wide range of geographical and climatic conditions. It was, in general, a total ecological viewpoint that decided the groupings. In addition, of course, the customary U. S. Census report units had to be used because their data made up the basis for the sample

structure. By combining various of the 15 regions listed here, the seven census regions used by the U. S. Census Bureau for reporting Puerto Rican demographic information can be constructed. It should be stated that the seven regions used in the U. S. Census reports were considered too inclusive and not representive of known ecological reality. The regional groups, then, do not represent a completely free approach to population grouping but were tied in with previous demographic report units. The dividing of the approximately 30 by 100 mile island into 15 regions gives reasonable certainty that population sub-group effects, if present, would be seen.

In addition to the geographic regional analysis, the population was sorted into phenotypically similar groups. The criteria used were those known (Gordon, 1950) to be held important by the Puerto Ricans themselves in assortative mating. These are skin color, hair form, nose shape and lip thickness. Inasmuch as the predominant racial groups constituting the present Puerto Rican population are White and Negro, these phenotypical characteristics also represent the most variable feature. Because the greatest preference, or social value, is given to skin color, with hair form, lip and nose characteristics being given less weight (Gordon 1949, 1950), these were evaluated accordingly in calculating the "ethnic factor" used in this report.

The system of calculating "ethnic factor" is as follows:

1. Skin color value was estimated by comparison with the von Luschan skin color scale. The range of recorded values was from 7 to 32, with 16 being the mean value for the total measured sample.
2. Hair form was judged visually and put in one category of a 5-point scale. The values for these positions were estimated and recorded as either 0, 4, 8, 12, or 16. Thus the values recorded into the "ethnic factor" total from hair form were approximately one half of those contributed by skin color.
3. Lip thickness and nose shape were judged on a 5-point scale and evaluated and recorded as either 0, 2, 4, 6, or 8. Thus these characteristics contributed approximately one half as much as hair form and one quarter as much as skin color in the total ethnic factor.

By adding up the values recorded for each of the four phenotypical characteristics, a total of from 7 to 64 was possible, the low number being the extreme White phenotype and the high number being the extreme Negro phenotype. This whole distribution range (7 to 64) was then evenly divided into 10 groups (0 to 9 inc.) with each individual being assigned to one group. These "ethnic groups" were then analyzed separately. It should be pointed out here that the sorting and typing procedure used here was not an attempt to place the individual racially but a mechanism to obtain a limited number of phenotypically similar groups. Inasmuch as the characters used are those most variable in a Negro-White mixed population, they do indicate the general racial affini-

ties of the groups. However, the individuals in the sample are not given a racial type on the basis of their ethnic factor. It is realized that the mode of inheritance or linkage between these characters is unknown, but it is assumed that within any one of these "ethnic groups" the genetic homogeneity is probably greater than in any combination of these groups, and certainly greater than in a combination of extreme groups. Also, because the characters used are the same ones given value by the Puerto Ricans in ascribing social status and influencing assortative mating, these "ethnic groups" have a valid use for this population. As will be seen in the genetic analysis which follows, these different phenotypical characteristics apparently have different values as group genetic "markers."

MATERIALS

The blood was drawn in a 10 cc. sample by veni-puncture. Five cc. were used in typing and five cc. used for other determinations not reported here. The samples were delivered to the Blood Bank Laboratory of the School of Tropical Medicine, where they were typed under the direction of Sr. N. Dobal[2] according to standard procedure recommended by Wiener (1943, 1946).

Difficulties were encountered with some of the sera used. Results from 3,079 MN determinations had to be discarded due to gross internal inconsistencies (test of binomality) in the frequency distribution of the types. Also the HR testing sera did not give consistent results so some 1,200 of these are also not reported.

In spite of these unfortunate difficulties, there is good reason to accept the reliability of the ABO and Rh frequencies here reported. In the first place, the results have internal consistency as seen in the $\frac{D}{\sigma}$ figures shown in tables 6 to 9; secondly, they are what would be expected from a Negro-White mixed population—that is, the Puerto Rican frequencies are intermediate to reported White and Negro figures.

For determination of the taster and non-taster reaction Phenyl-thiocarbamide crystals were placed by tooth pick applicator on the tongue of the 3,229 subjects tested. This method, while giving no indication of taste thresholds to known concentrations, does give all concentrations as the crystals dissolve and should discover all tasters. The infliction of this chemical on the palates of our trusting subjects gave the only threat to the otherwise excellent rapport

[2] Sr. Dobal was assigned to this study by the Director of the School of Tropical Medicine, Dr. Morales-Otero, and the considerable expense of laboratory analysis was carried by that institution. The sera were supplied by the Social Science Research Center of the Univeristy of Puerto Rico. The author wishes to express his gratitude and appreciation of the work done by Sr. Dobal and his assistants and to Dr. Morales-Otero for his unfailing assistance and encouragement in carrying this research to completion. Without his aid, this portion of our total project could not have been accomplished.

between investigators and Puerto Ricans during the study. It is unfortunate that so many were tasters.

BLOOD TYPING RESULTS

The number of individuals for the ABO system are reported by 15 regions in Table 2 and by 10 ethnic groups in Table 3. The Rh frequencies for 15 regions are given in Table 4 and for 10 ethnic groups in Table 5. The ABO gene frequencies, calculated by the square root formulae of Wiener (1943, 1945), together with calculation of the sum difference from 100 per cent and the

TABLE 2. DISTRIBUTION OF AGGLUTINOGENS A_1—A_2—B IN 15 REGIONS OF PUERTO RICO

REGIONS	CALCULATED SAMPLE NO.	TOTAL TYPED	% OF SAMPLE TYPED*	BLOOD GROUPS											
				O		A_1		A_2		B		A_1B		A_2B	
				N	%	N	%	N	%	N	%	N	%	N	%
1	183	177	97	107	60.45	46	25.99	11	6.21	12	6.78	1	0.56	0	0.00
2	726	644	89	322	50.00	173	26.86	66	10.25	64	9.94	9	1.40	10	1.55
3	188	220	117	115	52.27	48	21.82	34	15.45	20	9.09	1	0.45	2	0.91
4	128	68	53	36	52.94	16	23.53	4	5.88	11	16.18	1	1.47	0	0.00
5	120	250	208	133	53.20	62	24.80	11	4.40	36	14.40	4	1.60	4	1.60
6	220	223	101	129	57.85	52	23.32	15	6.73	24	10.76	1	0.45	2	0.90
7	240	246	102	136	55.28	53	21.54	26	10.57	24	9.76	4	1.63	3	1.22
8	199	152	76	84	55.26	35	23.03	12	7.89	17	11.18	0	0.00	4	2.63
9	150	76	51	38	50.00	17	22.37	9	11.84	9	11.84	2	2.63	1	1.32
10	59	48	81	36	75.00	6	12.50	3	6.25	2	4.17	0	0.00	1	2.08
11	196	174	89	86	49.43	58	33.33	19	10.92	8	4.60	3	1.72	0	0.00
12	150	204	136	111	54.41	46	22.55	22	10.78	20	9.80	3	1.47	2	0.98
13	274	252	92	137	54.36	64	25.40	22	8.73	24	9.52	1	0.40	4	1.59
14	267	270	101	141	52.22	77	28.52	33	12.22	17	6.30	1	0.37	1	0.37
15	233	241	103	141	58.51	53	21.99	25	10.37	21	8.71	1	0.41	0	0.00
Totals....	3333	3245	97	1752	53.99	806	24.84	312	9.61	309	9.52	32	0.99	34	1.05

* This column gives the percentage of the total calculated representative sample of Puerto Rico which was actually typed for ABO. The total sample was calculated to include 3333 individuals (3562 actually obtained) and the percentages given here are the proportion of the calculated sample of each region which were actually typed. 100% would be perfect correspondence.

probable error of the difference are given in Table 6 and 7. The Rh gene frequencies by region and ethnic groups are given in Table 8 and 9. Total island frequencies are given in Table 10. These total island frequencies for Rh are not significantly different from Torregrossa's (1945 a) report; however her sample was small. Table 10 includes Torregrossa's reported results.

If it is recalled that the lower ethnic group numbers are the White end of the range and the higher numbers Negro, the changes in gene frequencies in both the ABO and the Rh show gradiants which would be expected. In particular, the gradiant for R^o and for I^{A_1} and I^B show expected slopes. The fact

that the type O frequencies in Puerto Rico are high for both Negro and White ethnic groups may be attributable to the genetic contribution of the aboriginal

TABLE 3. DISTRIBUTION OF AGGLUTINOGENS A_1–A_2–B IN 10 ETHNIC GROUPS OF PUERTO RICO

ETHNIC GROUP	TOTAL TYPED	BLOOD GROUPS											
		O		A_1		A_2		B		A_1B		A_2B	
		N	%	N	%	N	%	N	%	N	%	N	%
0	6*	5	—	1	—	—	—	—	—	—	—	—	—
1	101	50	49.50	27	26.73	11	10.89	10	9.90	1	0.99	2	1.98
2	824	450	54.34	214	25.85	88	10.63	63	7.61	9	1.09	4	0.48
3	1441	787	54.61	382	26.51	136	9.44	112	7.77	8	0.56	16	1.11
4	433	224	51.73	109	25.17	40	9.24	50	11.55	6	1.39	4	0.92
5	133	70	52.63	26	19.55	9	6.77	22	16.54	5	3.76	1	0.75
6	159	89	55.97	28	17.61	14	8.81	24	15.09	1	0.63	3	1.89
7	101	54	53.47	15	14.85	10	9.90	18	17.82	1	0.99	3	2.97
8	42	23	54.76	4	9.52	3	7.14	10	23.81	1	2.38	1	2.38
9	1*	—	—	—	—	1	—	—	—	—	—	—	—
Totals...	3245	1752	53.99	806	24.84	312	9.61	309	9.52	32	0.99	34	1.05

* Frequences not listed because of insufficient sample size.

TABLE 4. FREQUENCIES OF THE RH BLOOD TYPES FOR 15 REGIONS OF PUERTO RICO

REGION	TOTAL TYPED	rh		rh′		rh″		rh′rh″		Rh_0		Rh_1		Rh_2		Rh_1Rh_2	
		N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
1	150	15	10.00	0	0.00	1	0.67	0	0.00	20	13.33	70	46.67	24	16.00	20	13.33
2	326	24	7.36	2	0.61	0	0.00	0	0.00	41	12.58	163	50.00	59	18.10	37	11.35
3	212	26	12.26	3	1.42	0	0.00	0	0.00	41	19.34	99	46.70	21	9.91	22	10.38
4	51	5	9.80	1	1.96	2	3.92	0	0.00	6	11.76	25	49.02	6	11.76	6	11.76
5	244	25	10.25	7	2.87	1	0.41	0	0.00	48	19.67	114	46.72	28	11.48	21	8.61
6	177	16	9.04	0	0.00	0	0.00	0	0.00	31	17.51	90	50.85	29	16.38	11	6.21
7	212	17	8.02	7	3.30	1	0.47	1	0.47	37	17.45	113	53.30	18	8.49	18	8.49
8	126	11	8.73	1	0.79	2	1.59	0	0.00	13	10.32	58	46.03	18	14.29	23	18.25
9	61	5	8.20	0	0.00	1	1.64	0	0.00	8	13.11	25	40.98	12	19.67	10	16.39
10	31	6	19.35	1	3.23	0	0.00	0	0.00	4	12.90	13	41.94	4	12.90	3	9.68
11	154	13	8.44	1	0.65	0	0.00	0	0.00	16	10.39	70	45.45	28	18.18	26	16.88
12	191	10	5.24	1	0.52	1	0.52	0	0.00	16	8.38	109	57.07	33	17.28	21	10.99
13	253	21	8.30	3	1.19	0	0.00	2	0.79	14	5.53	124	49.01	38	15.02	51	20.16
14	209	20	9.57	2	0.96	3	1.44	0	0.00	18	8.61	98	46.89	34	16.27	34	16.27
15	131	8	6.11	2	1.53	0	0.00	0	0.00	21	16.03	65	49.62	22	16.79	13	9.92
Totals....	2528	222	8.78	31	1.23	12	0.47	3	0.12	334	13.21	1236	48.89	374	14.79	316	12.50

Indian element. In an attempt to follow this possibility, the blood type frequencies of those in the sample who had well defined shovel-shaped incisors, a characteristic of high frequency in American Indian groups, was analyzed. In

no way was this group, a sample of 424, significantly different from the total sample in either the ABO or Rh frequencies.

TABLE 5. FREQUENCIES OF THE RH BLOOD TYPES FOR 10 ETHNIC GROUPS OF PUERTO RICO

ETHNIC GROUP	TOTAL TYPED	rh		rh′		rh″		rh′rh″		Rh_0		Rh_1		Rh_2		Rh_1Rh_2	
		N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
0	6	0	0.00	0	0.00	0	0.00	0	0.00	1	16.67	3	50.00	1	16.67	1	16.67
1	84	4	4.76	0	0.00	0	0.00	1	1.19	4	4.76	52	61.90	10	11.90	13	15.48
2	635	58	9.13	2	0.31	5	0.79	1	0.16	64	10.08	315	49.61	102	16.06	88	13.86
3	1122	98	8.73	13	1.16	7	0.62	1	0.09	116	10.34	566	50.45	159	14.17	162	14.44
4	324	29	8.95	6	1.85	0	0.00	0	0.00	59	18.21	152	46.91	48	14.81	30	9.26
5	104	10	9.62	2	1.92	0	0.00	0	0.00	25	24.04	48	46.15	11	10.48	8	7.69
6	130	13	10.00	6	4.62	0	0.00	0	0.00	33	25.38	57	43.85	15	11.54	6	4.62
7	83	5	6.02	1	1.20	0	0.00	0	0.00	22	26.51	29	34.94	21	25.30	5	6.02
8	39	5	12.82	1	2.56	0	0.00	0	0.00	10	25.64	14	35.90	6	15.38	3	7.69
9	1													1	100.00		
Totals...	2528	222	8.78	21	1.23	12	0.47	3	0.12	334	13.21	1236	48.89	374	14.79	316	12.50

TABLE 6. ABO BLOOD GROUP GENE FREQUENCIES* FOR 15 REGIONS OF PUERTO RICO

REGION	GENE				TOTAL	DEVIATION FROM 100%	S. E.$_D$	$\frac{D}{\sigma}$
	I^O	I^{A1}	I^{A2}	I^B				
1	.777	.147	.039	.043	1.006	.006	.005	1.200
2	.707	.157	.069	.067	1.000	.000	.004	.000
3	.723	.123	.100	.060	1.006	.006	.006	1.000
4	.728	.140	.039	.103	1.010	.010	.014	.714
5	.729	.149	.030	.093	1.001	.001	.007	.143
6	.751	.134	.053	.077	1.015	.015	.007	2.143
7	.743	.124	.068	.063	0.998	.003	.006	.333
8	.743	.133	.052	.072	1.000	.000	.008	.000
9	.707	.132	.079	.079	.997	.003	.012	.250
10	.866	.067	.035	.024	.992	.008	.005	1.600
11	.703	.191	.074	.032	1.000	.000	.006	.000
12	.738	.130	.069	.063	1.000	.000	.006	.000
13	.737	.147	.057	.062	1.003	.003	.006	.500
14	.723	.161	.080	.042	1.006	.006	.005	1.200
15	.765	.123	.065	.055	1.008	.008	.005	1.600
Totals...	.735	.143	.062	.062	1.002	.002	.002	

* Gene frequencies calculated according to formulae given by Wiener (1943).

Regional differences are apparent in Tables 2 and 4, and give a picture of considerable variation. Region 10 stands out as a deviant one. To test the degree of homogeneity by regions, the Chi-Square test was applied. The formula used to calculate "t" was $\sqrt{2\chi^2} - \sqrt{2n - 1}$ (Fisher and Yates, Table IV, 1943). For the ABO distributions by region, the $\chi^2 = 69.99$ with d.f. 42, $t =$

2.72 so P is less than .01. Therefore, the distribution is non-homogeneous. For the Rh types, the same test was applied. The $\chi^2 = 135.32$ with d.f. = 70 gives $t = 4.66$, so again the distribution is definitely non-homogeneous.

TABLE 7. ABO BLOOD GROUP GENE FREQUENCIES† FOR 8 ETHNIC GROUPS OF PUERTO RICO

ETHNIC GROUP	GENE				TOTAL	DEVIATION FROM 100%	S. E.$_D$	$\frac{D}{\sigma}$
	I^O	I^{A1}	I^{A2}	I^B				
0*	—	—	—	—	—	—	—	—
1	.704	.156	.074	.067	1.001	.001	.013	.077
2	.377	.147	.068	.050	1.002	.002	.003	.667
3	.739	.151	.061	.050	1.001	.001	.002	.500
4	.719	.147	.062	.076	1.004	.004	.005	.800
5	.726	.118	.045	.106	.995	.005	.009	.556
6	.748	.103	.057	.095	1.003	.003	.008	.375
7	.731	.088	.065	.113	.997	.003	.011	.273
8	.740	.058	.047	.146	.991	.009	.016	.563
9*	—	—	—	—	—	—	—	—
Totals.....	.735	.143	.062	.062	1.002	.002	.002	

† Gene frequencies calculated according to formulae given by Wiener (1943).
* Frequencies not calculated because of small sample size.

TABLE 8. THE RH GENE FREQUENCIES* FOR 15 REGIONS OF PUERTO RICO

REGIONS	Rh GENE						TOTAL	DEVIATION FROM 100%	S. E.$_D$	$\frac{D}{\sigma}$
	r	r'	r''	R^0	R^1	R^2				
1	.3162	.0000	.0104	.1668	.3537	.1391	.9862	.0138	.0179	.771
2	.2713	.0110	.0000	.1752	.3824	.1073	1.0102	.0102	.0143	.713
3	.3501	.0198	.0000	.2120	.3110	.0822	.9751	.0249	.0102	2.441
4	.3130	.0299	.0574	.1513	.3575	.0885	.9976	.0024	.0325	.074
5	.3202	.0420	.0063	.2268	.3027	.0933	.9913	.0087	.0109	.817
6	.3007	.0000	.0000	.2146	.3645	.1399	1.0197	.0197	.0162	1.173
7	.2832	.0533	.0082	.2215	.3479	.0739	.9880	.0120	.0119	1.008
8	.2955	.0130	.0258	.1410	.3621	.1287	.9661	.0339	.0209	1.622
9	.2561	.0000	.0303	.1874	.3457	.1664	.9859	.0141	.0330	.427
10	.4399	.0353	.0000	.1280	.2767	.1040	.9839	.0161	.0291	.553
11	.2905	.0110	.0000	.1434	.3609	.1745	.9803	.0197	.0202	.975
12	.2289	.0111	.0112	.1402	.4637	.1803	1.0353	.0353	.0237	1.532
13	.2881	.0200	.0000	.0838	.4083	.1652	.9654	.0346	.0171	1.023
14	.3094	.0151	.0236	.1170	.3711	.1491	.9853	.0147	.0177	.831
15	.2470	.0292	.0000	.2235	.3564	.1534	1.0095	.0095	.0208	.457
Totals.....	.2956	.0202	.0079	.1729	.3605	.1336	.9907	.0093	.0147	

* Calculated according to formula given in Wiener, Zepeda, Sonn, and Polivka (1945).

In this test, Rh′, Rh″, and Rh′Rh″ were combined so that no numbers in the cells were zero. When this combining was not done, $t = 5.05$ resulted, compared to $t = 4.66$ in the combination.

At this point the obvious question is whether the lack of homogeneity is the result of there being different proportions of White and Negro in the various regions, inasmuch as the ABO and Rh frequencies do vary between these groups. Consequently, only persons of ethnic group 2 were put in one chi-square regional distribution and only those of ethnic group 3 in another. All 15 regions

TABLE 9. THE RH GENE FREQUENCIES* FOR 8 ETHNIC GROUPS OF PUERTO RICO

ETHNIC GROUP	Rh GENE						TOTAL	DEVIATION FROM 100%	S. E.$_D$	$\frac{D}{\sigma}$
	r	r'	r''	R^0	R^1	R^2				
1	.2182	.0000	.0000	.0903	.5366	.1543	.9994	.0006	.0355	.017
2	.3022	.0050	.0128	.1361	.3881	.1494	.9936	.0064	.0099	.646
3	.2955	.0190	.0103	.1412	.3850	.1349	.9859	.0141	.0072	1.958
4	.2992	.0295	.0000	.2200	.3206	.1266	.9979	.0021	.0110	.191
5	.3102	.0295	.0000	.2700	.2943	.0849	.9889	.0111	.0146	.760
6	.3162	.0662	.0000	.2786	.2547	.0902	1.0059	.0059	.0134	.440
7	.2454	.0233	.0000	.3250	.2350	.1901	1.0188	.0188	.0222	.847
8	.3581	.0341	.0000	.2621	.2227	.1136	.9906	.0094	.0238	.395
Totals.....	.2956	.0202	.0079	.1729	.3605	.1336	.9907	.0093	.0147	

* Calculated according to formula given in Wiener, Zepeda, Sonn, and Polivka (1945).

TABLE 10. PUERTO RICAN TOTAL ISLAND PERCENTAGE FREQUENCIES REPORTED BY TORREGROSSA (1945 A, B) COMPARED WITH THIS REPORT

Rh Blood Types

	rh	rh'	rh''	rh'rh''	Rh_0	Rh_1	Rh_2	Rh_1Rh_2	TOTAL TYPED
Torregrousa.....	10.1	1.7	0.5	0.00	15.1	39.1	19.6	14.0	179
Thieme.....	8.78	1.23	0.47	.012	13.21	48.89	14.79	12.50	2528

ABO Blood Types

	O	A	B	AB	TOTAL TYPED
Torregrossa.....	48.72	38.69	9.56	3.03	429
Thieme.....	53.99	34.45	9.52	2.04	3245

were used in all of these chi-square tests. The ABO and Rh distributions were as follows:

ABO

Ethnic Group 2 $\chi^2 = 55.00$ d.f. $= 42$ $t = 1.38$ $P = > 0.10$, Homogeneous
Ethnic Group 3 $\chi^2 = 47.84$ d.f. $= 42$ $t = 0.67$ $P = 0.50$, Homogeneous

Rh

Ethnic Group 2 $\chi^2 = 81.03$ d.f. $= 70$ $t = 0.97$ $P = > 0.10$, Homogeneous
Ethnic Group 3 $\chi^2 = 103.16$ d.f. $= 70$ $t = 2.57$ $P = 0.01$, Non-homogeneous

This gives us a test for the assumption that the ethnic groupings represent somewhat homogeneous genetic groups. In the above it will be seen that three

of the four groupings are homogeneous by the chi-square test and establishes that the assumption is correct to a degree. As a further test for this point, the test was applied to samples grouped by region for similar recorded categories of skin color, hair form, lip thickness and nose shape. The Rh distributions were used as they give a more rigorous test for homogeneity as can be seen above. The Chi Square test gave the following:

Rh					
Skin Color	$\chi^2 = 88.77$	d.f. = 70	$t = 1.53$	$P = .126$,	Homogeneous
Hair Form	$\chi^2 = 105.53$	d.f. = 70	$t = 2.72$	$P = .0066$,	Non-homogeneous
Nose Shape	$\chi^2 = 109.44$	d.f. = 70	$t = 3.01$	$P = .0026$,	Non-homogeneous
Lip Thickness	$\chi^2 = 116.81$	d.f. = 70	$t = 3.49$	$P = .0005$,	Non-homogeneous

These results would seem to verify the statements made that skin color is the most important characteristic in determining mating choice in Puerto Rico, and that hair form is next in importance. This has acted to create group genetic similarity for other characteristics, in this case Rh frequency distributions. These results also tend to verify that skin color should be given the greatest importance in assessing ethnic factor and assortative mating, with hair form next, as was done from the outset in this study. In view of these results, doubt is cast upon the efficiency of the multiple character selection used in manufacturing an ethnic factor. It would seem that of the factors used, skin color alone would give a better sort into genetically homogeneous groups, as indicated by Rh blood type, than using all four combined. It is realized, of course, that this is for only one genetic constellation and would not necessarily hold for other genes.

As a demonstration of the well-known fact that the ABO and Rh frequencies vary by race, or by ethnic group as used here, chi-square tests for this were made. For the distribution of ABO types by the ethnic groups, the $\chi^2 = 54.30$, d.f. = 15, $t = 5.03$, $P = <0.0001$. For the Rh types, the results were also non-homogeneous, as $\chi^2 = 101.57$, d.f. = 25, $t = 7.25$ and $P = <0.0001$. In this test, ethnic groups 0 and 1, 5 and 6, and 7, 8 and 9 were combined so no cells would contain zero. And in the Rh test, Rh′, Rh″ and Rh′Rh″ were combined. This merely confirms the fact that different racial groups, in this case Negro and White, are characterized by different frequency distributions of the ABO and Rh blood types. It should be stated that, in all probability, each of the ethnic groups used here represents a mixed Negro-White group with the amount of mixture being less for the extreme group 0, 1, 2, and 7, 8, 9 than for the intermediate ones. It is difficult for any family long resident in Puerto Rico to point with certainty to a background lacking mixture. The rates of mixture were undoubtedly high during the early periods of colonization and slavery when Spanish women were scarce and Negro and Indian women were more plentiful. As in the United States, the Negro group in Puerto Rico has probably experienced considerable admixture since its arrival.

As a visual demonstration of the frequency distribution by regions, the island pattern for *r* and for non-taster is shown in figures 1 and 2. Approximately the same amount of variability by region would be found for genes I^O, I^A, I^B, R^2 and R^O. In addition, ethnic group comparisons would show even more dramatic differences. The results of tests of significance for inter-group regional and ethnic differences are not reported in detail because of the large number involved. Nevertheless, they do exist, even at the $P = .01$ level of significance. For example, when comparing the type O frequency of region 10 to region 2 and to region 9, the difference is significant in both cases. Also, significant differences are found for various blood types when comparing the total island

TABLE 11. FREQUENCY OF RH TYPES FOR PUERTO RICAN INDIVIDUALS BORN AND WITH PRESENT RESIDENCE IN 15 DIFFERENT GEOGRAPHIC REGIONS

REGIONS	rh		rh′		rh″		rh′rh″		Rh_0		Rh_1		Rh_2		Rh_1Rh_2	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
1	12	10.81	0	0.00	1	0.90	0	0.00	17	15.32	50	45.05	17	15.32	14	12.61
2	14	7.25	1	0.52	0	0.00	0	0.00	22	11.40	103	53.36	34	17.62	19	9.84
3	19	13.97	2	1.47	0	0.00	0	0.00	28	20.59	61	44.85	14	10.29	12	8.82
4	1	5.00	1	5.00	2	10.00	0	0.00	1	5.00	9	45.00	2	10.00	4	20.00
5	20	10.00	5	2.49	1	0.50	0	0.00	38	18.90	94	46.77	23	11.44	20	10.00
6	9	10.47	0	0.00	0	0.00	0	0.00	15	17.44	46	53.49	12	13.93	4	4.65
7	14	8.92	6	3.82	0	0.00	1	0.64	28	17.20	82	52.23	13	8.28	14	8.92
8	7	10.45	0	0.00	2	2.98	0	0.00	3	4.48	29	43.28	15	22.39	11	16.42
9	4	7.41	0	0.00	0	0.00	0	0.00	7	12.96	25	46.30	9	16.67	9	16.67
10	6	23.08	1	3.85	0	0.00	0	0.00	2	7.69	13	50.00	3	11.54	1	3.85
11	10	9.52	1	0.95	0	0.00	0	0.00	13	12.38	44	41.90	18	17.14	19	18.10
12	8	6.90	0	0.00	1	0.86	0	0.00	10	8.62	65	56.03	18	15.52	14	12.07
13	19	10.00	3	1.58	0	0.00	1	0.53	12	6.32	95	50.00	31	16.32	29	15.26
14	12	8.33	1	0.69	0	0.00	0	0.00	12	8.33	71	49.31	26	18.06	22	15.28
15	7	7.61	2	2.17	0	0.00	0	0.00	11	11.95	48	52.17	13	14.13	11	11.97
Totals.	162	9.54	23	1.35	7	0.41	2	0.12	218	12.84	835	49.18	248	14.61	203	11.96

sample frequency to various regional or ethnic sub-samples. The non-homogeneity indicated by the chi square results reported above confirms this general picture of island group variability.

In table 11, Rh frequencies for a sample selected according to different residence history are given. This group is composed of males and females who were born and remained in various geographic regions. They are the non-migratory segment of the total sample and represent the residual core of genetic material tending to perpetuate regional differences as they may occur. By visually comparing the frequencies shown here (table 11) with those given in table 4, some differences can be seen. Table 4 gives the total sample of all individuals classified by region of present residence regardless of island birth place. As a detailed Rh type by region comparison would involve the showing

of 120 "t" values, this is not included here. Yet it is apparent that significant differences between some regions in these two samples exist. Comparing the total island frequencies in each table, no significant difference occurs.

From this comparison, it would seem that individual migratory history may be an important sample consideration and that the inclusion of migrants may tend to reduce differences that actually exist between regions. This would be true for populations with a recent history of considerable internal migration. Of course, population with a long history of random internal migration would probably not be so affected. The chi square test applied to the distribution given in table 8 gives the following results compared to table 3:

Table 8 $\chi^2 = 109.54$ d.f. $= 70$ $t = 4.66$ $P = 0.00006$
Table 3 $\chi^2 = 135.32$ d.f. $= 70$ $t = 3.18$ $P = 0.0014$

It might be expected that the non-migrants would show greater heterogeneity by virtue of their representing existing isolates historically originated and unaltered by the high rates of recent internal migration. However, this is not the case as indicated by this test. Random migration is not characteristic of Puerto Rico, where localized and limited economic opportunity are motivating individual movements within prescribed spheres. How this may affect frequencies of particular genes is not answered here, except to say it apparently has no effect on Rh frequencies.

A check was performed to see whether the actual sample, by virtue of not being an exact duplicate of the calculated representative sample (see table 2), could have introduced a significant deviation which would affect the representativeness of the total island frequencies shown in table 7. To do this each region's proportions of ABO and Rh were applied to extrapolate its actual sample up or down in the calculated sample size. The proportions of types for each region remained unchanged but as between regions they were changed. These extrapolated numbers for each region were then used to calculate the total island frequencies based on a sample of 3,333, the total calculated sample. Table 12 gives the results of this comparison. None of the differences shown is significant so the actual sample obtained is used as the basis for comparisons in this report.

TASTER (P. T. C.) RESULTS

Non-taster percentage frequencies are given in table 13 by regions and in table 14 by ethnic groups. These are the same groupings used for blood type analysis.

The sub-group heterogeneity is not significant. The chi-square test gives the following results for males only:

Ethnic groupings: $\chi^2 = 9.63$ d.f. $= 7$ $P = 0.20$
Regional Groupings: $\chi^2 = 19.71$ d.f. $= 14$ $P = 0.10$

The females more frequently taste than do the males, yet the difference is not significant at the 5% level (t = 1.54). However, this difference is not constant between males and females by ethnic group. The difference is less

TABLE 12. COMPARISON OF TOTAL ACTUAL SAMPLE WITH TOTAL CALCULATED SAMPLE FOR FREQUENCIES OF ABO AND RH TYPES.

ABO

	O	A_1	A_2	B	A_1B	A_2B
Calculated	53.88	24.88	9.70	9.49	1.01	1.04
Actual	53.99	24.84	9.61	9.52	.99	1.05
Difference	+.11	−.04	−.09	+.03	−.02	+.01

Rh

	rh	rh′	rh″	rh′rh″	Rh_0	Rh_1	Rh_2	Rh_1Rh_2
Calculated	8.53	1.16	0.46	0.11	12.82	48.75	15.46	12.71
Actual	8.78	1.23	0.47	0.12	13.21	48.89	14.79	12.50
Difference	+.25	+.07	+.01	+.01	+.39	+.14	−.67	−.21

TABLE 13. FREQUENCY FOR MALES AND FEMALES OF NON-TASTERS OF P.T.C. FOR 15 REGIONS OF PUERTO RICO

REGIONS	MALES			FEMALES		
	Total Tested	Non-Tasters	% Frequency	Total Tested	Non-Tasters	% Frequency
1	80	10	12.50	86	6	6.98
2	299	59	19.73	265	29	10.94
3	151	19	12.58	100	11	11.00
4	35	2	5.71	27	1	3.70
5	126	18	14.29	121	6	4.96
6	120	12	10.00	98	7	7.14
7	98	13	13.26	151	16	10.60
8	90	13	14.44	115	21	18.26
9	62	9	14.52	11	2	18.18
10	32	5	15.63	38	4	10.53
11	97	8	8.25	65	8	12.31
12	99	12	12.12	86	10	11.62
13	139	15	10.79	141	15	10.63
14	140	19	13.57	123	13	10.57
15	125	11	8.80	109	9	8.26
Totals	1693	225	13.29	1536	158	10.29

for the Whites than for the Negro. This suggests that a possible difference in penetrance between the sexes, which might explain this difference, may be unequal in different races; and here, specifically, between persons of African

and European origin. In addition, the relatively low non-taster frequency found in Puerto Rico (see tables in Boyd, 1950, and Gates, 1946) suggests again the presence of American Indian in the make-up of the present population. However, Lee (1934) gives a frequency of 9.2% for non-tasters in the American Negro, which is under the figures reported here. Consequently, any interpretation pointing to the effects of the American Indian strains in Puerto Rico is doubtful from this evidence alone. Yet, the historical records give clear evidence that this is a plausible explanation in the case of Puerto Rico because significant numbers of American Indians were present there in colonial times.

TABLE 14. FREQUENCY FOR MALES AND FEMALES OF NON-TASTERS OF P.T.C. FOR 10 ETHNIC GROUPS OF PUERTO RICO

REGIONS	MALES			FEMALES		
	Total Tested	Non-Tasters	% Frequency	Total Tested	Non-Tasters	% Frequency
0	2	0	0.00	—	—	—
1	40	5	12.50	44	4	9.09
2	377	43	11.40	441	53	12.02
3	818	113	13.81	671	68	10.13
4	216	29	13.43	208	21	10.10
5	58	9	15.58	68	5	7.35
6	83	6	7.23	61	4	6.56
7	63	12	19.04	35	3	8.57
8	36	8	22.22	7	0	0.00
9	—	—	—	1	0	0.00
Totals.....	1693	225	13.29	1536	158	10.29

DISCUSSION

The frequency of the ABO and Rh blood types in Puerto Rico, as well as non-tasters of P. T. C., is as would be expected in a mixed Negro-White-American Indian population. However, variability within this population is significant. It seems that the general implications of this are quite clear. Specifically, sampling for genetic traits in a population must be done in a fashion which will measure the internal variability and sample all groupings to give an overall representative result. It is also indicated from these findings that at least regional and phenotypic sub-groups must be representatively sampled before any remarks about the gene frequencies of a population are valid.

In this study, the genetic variability of the Puerto Rican population is shown to be large (see figures 1 and 2). Consequently, if the sample had been gotten from one region, rather than from all, the results may have been significantly different. For example, if region 10 had been the sole scene of sampling, the obtained type O frequency would have been 21.01% more than that

actually obtained. To put it another way, the frequencies of type O (see table 2 and 4) may have been from 49% to 75%, for type A_1 from 12% to 33%, for rh from 5% to 19% and for Rh_1 from 41% to 57%, if the sample had been

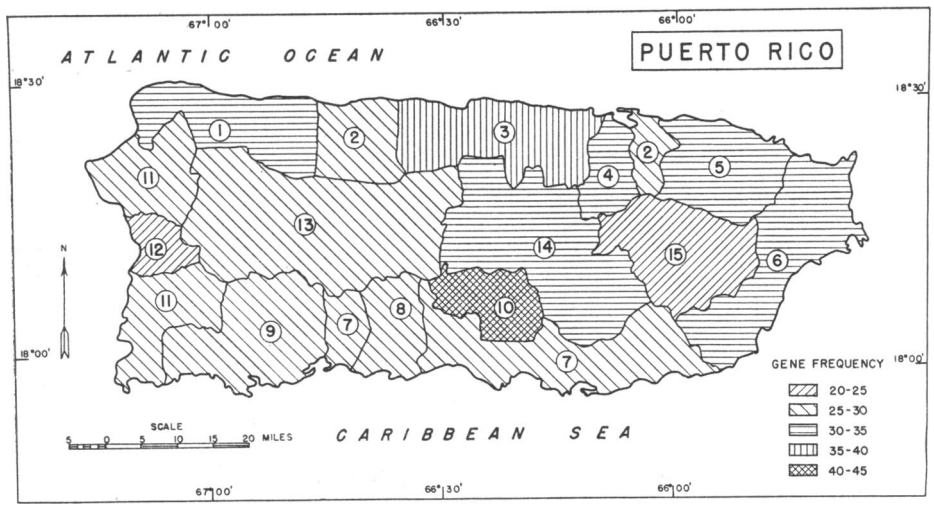

FIG. 1. The gene frequency distribution for *r* in the population of Puerto Rico for 15 geographic regions.

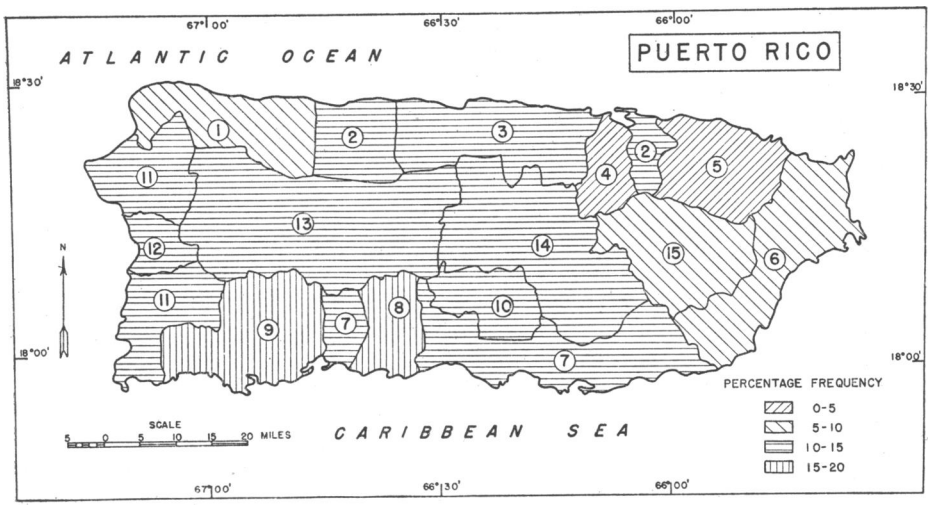

FIG. 2. The percentage frequency of male non-tasters in the population of Puerto Rico by 15 geographic regions.

obtained only from certain regions. From this, it seems evident that a sample derived from one region probably would have given quite different results. Although some of the regions did turn out to be representative of the frequencies for the whole island, this fact was not known, nor could it be known, until

the total representative sample had been gathered. Now we know which regions are typical in terms of certain genetic frequencies, but this, of course, does not mean that they will be typical for other traits. There seems to be no escaping the task of obtaining a representative sample if our problem is to determine the population genetic frequency.

In a review of the literature in blood type frequencies, no reports based on pre-determined demographically representative samples has been found. Although it is readily acknowledged that some reported samples may truly be representative, it cannot be verified from procedures reported that this is so. It would seem that reliable demographic information is available for many populations of the world—at least as these populations are defined by national boundaries—and consequently, the efforts in the future should direct attention both to the definition of the population and to its representative genetic description. It is also important to emphasize the need for valid sampling in terms of choice of subjects. As Glass remarks in commenting on a paper by the author (Thieme, 1950), the difficulty of obtaining valid samples in blood typing laboratories, hospitals or clinics is very great. Any selection such as must obviously underlie such samples does not give us the picture of the total population which is ultimately desirable. In the last analysis, we can seldom hope that findings from studies carried out for one problem will give answers to different problems. Clinics, hospitals and insurance company examinations are not designed to gather representative population data. They may approach this goal but the degree of attainment should be measured, not assumed. To the same extent, the results of blood group surveys on people about which demographic knowledge is lacking must be viewed with caution. Continental maps with isogene lines showing large trends for changes in gene frequencies by areas are occasionally based on small samples and frequently on unrepresentative samples. The standard error of estimates for small samples is large and frequently so large as to overlap the range of variation seen in any one of the isogenic maps for a continent. It is not necessary to belabor the need for large numbers in sampling, or the desirability of representativeness. Most investigators are aware of this. However, it is apparent that this desirable goal for all studies is far from being achieved.

REFERENCES

BOYD, W. C. 1939. Blood Groups. *Tabulae biol. Berl.* 17: 113–240.

BOYD, W. C. 1950. *Genetics and the Races of Man.* Boston: D. C. Heath & Co., pp. xiii, 453.

FISHER, R. A. & F. YATES. 1943. *Statistical tables for Biological, Agricultural and Medical Research.* Edinburgh: Oliver and Boyd.

GATES, R. R. 1946. *Human Genetics.* New York: The Macmillan Co. 2 vols. pp. xi, 1518.

GORDON, M. W. 1949. Race patterns and prejudice in Puerto Rico. *Amer. Sociol. Rev.* 14: 294–301.

GORDON, M. W. 1950. Cultural aspects of Puerto Rico's race problem. *Amer. Sociol. Rev.* 15: 382–392.

HARRIS, J. A. 1930. The measurement of man in the mass. In: Harris, J. A., Jackson, C. M., Paterson, D. G., and Scammon, R. E., *The Measurement of Man.* Minneapolis: Univ. of Minnesota Press.

LEE, B. F. 1934. A genetic analysis of taste deficiency in the American Negro. *Ohio J. Sc.*, 34: 337–342.

THIEME, F. P. 1950. Problems and methods of population surveys. In: *Origin and Evolution of Man.* Cold Spring Harbor Symposia on Quantitative Biology. Vol. XV.

TORREGROSSA, M. V. 1945a. Incidence of the Rh agglutinogen among Puerto Ricans. *Puerto Rico J. Pub. Health*, 21: 166–168.

TORREGROSSA, M. V. 1945b. Incidence of the eight Rh types among 179 Porto Ricans. *Proc. Soc. Exp. Biol. N. Y.* 60: 215–218.

U. S. Bureau of the Census. 1946. Sixteenth Census of the United States: 1940. *Puerto Rico Population, Bulletin No. 4,* Migration between Municipalities. Washington, U. S. Government Printing Office.

WIENER, A. S. 1943. *Blood Groups and Transfusions.* ed. 3, Springfield: Charles C. Thomas. pp. 438.

WIENER, A. S., ZEPEDA, J. P., SONN, E. B. & POLIVKA, H. R. 1945. Individual blood differences in Mexican Indians, with special reference to the Rh blood types and Hr factor. *J. Exp. M.* 81: 559–571.

WIENER, A. S. 1946. Technique of Rh testing. *Amer. J. of Med. Technology.* 12: 95–99.