GENETICS AND RACE

PART I

HYMIE GORDON,* B.SC., M.D. (CAPE TOWN), M.R.C.P. (LOND. AND EDIN.), *Senior Lecturer, Comprehensive Medicine Group, Department of Medicine, University of Cape Town*

> 'There is one principal and as it were radical distinction between different minds, in respect of philosophy and the sciences; which is this: that some minds are stronger and apter to mark the differences of things, other to mark their resemblances. The steady and acute mind can fix its contemplations and dwell and fasten on the subtlest distinctions: the lofty and discursive mind recognises and puts together the finest and most general resemblances. Both kinds however easily err in excess, by catching the one at gradations the other at shadows.'
>
> Francis Bacon (1620): *Novum Organum* I, 55.

To those of us who practise medicine among the different peoples of South Africa, racial variation in the incidence and presentation of diseases is commonplace. Ischaemic heart disease and diabetes mellitus are common in Whites but are rare in Africans. Africans are more susceptible than Whites to carcinoma of the liver and oesophagus, whereas the latter are more likely to develop skin cancers and gliomata. Gallstones and kidney stones are rarely found in Africans; thyrotoxicosis and pernicious anaemia are similarly uncommon in this group. On the other hand, chronic tuberculous lymphadenopathy, keloids and fulminating amoebic dysentery occur frequently in Africans but are most unusual in Whites. Furthermore, it is recognized that Indians are exceptionally prone to develop diabetes mellitus, that carcinoma of the cervix uteri is particularly common in Cape Coloured women, and that chronic ulcerative colitis occurs with increased frequency in Jews. This list of variations is far from complete: many more could be added, either from clinical impression or from the results of formal epidemiological surveys.

Such experiences are not limited to South Africa. All over the world, racial variations of this sort have been recorded. In several instances, comparative racial studies have provided useful clues for unravelling complex problems of aetiology and pathogenesis. In South Africa, such studies have contributed a great deal to the understanding of ischaemic heart disease. Elsewhere, the important association of poor genital hygiene with cervical cancer first became apparent from inter-racial studies. In recent years, the increased movements of populations and the greater accessibility of previously remote peoples have brought to light several other racial variations, and more can be expected in the future.

These and other considerations have led to an increasing biological interest in the problem of 'race' and to a reappraisal of the meaning of this term. There is a great deal of confusion about this matter mainly because there is no general agreement about the definition of 'race' and because the term is used differently by different people.

Originally, people were classified into races on the basis of their language: hence the 'Latin race', the 'Aryan race'

*First Frank Forman Foundation Fellow (1958 - 1959).

and so on. The term 'Bantu race', which is so frequently used in South Africa, also has a linguistic origin and means no more than the 'Bantu-speaking people'. Sometimes the term 'race' has been applied to geographically related people: hence the 'Mediterranean race', 'Asiatic race', etc. Population movements have made these classifications almost useless.

Nowadays, the term 'race' is defined in 4 different ways.

1. **Sociological definition:** *A race is a group of people who, through many generations, have developed customs, traditions and language which are different from those of other groups of people.*

For biologists, this definition has only limited use. It is a feature of our time for groups of people to adopt the language and customs of other groups with whom they come in contact, but without much loss of their biological individuality. Thus, in America, Negroes have adopted the English language and the social customs of White Americans but, to a large extent, they have retained the physical attributes of their African ancestors.

2. **Anthropological definition:** *A race is a group of people whose physical characters distinguish it from other groups of people.*

This has been a serviceable definition for a long time, but it also has serious limitations. In particular, too many physical characters are subject to environmental change: stature, skin pigmentation, colour of hair are all susceptible to modification by natural or artificial external agents. In addition, it has proved difficult to express physical characters in a form that is suitable for the accurate quantitative assessment of racial affinities and mixtures.

Nevertheless, classical anthropology has made very substantial contributions to the study of races. The preliminary subdivision of mankind into a few 'primary races' is most conveniently based on anthropological criteria. In biological terms, man is a species, *Homo sapiens,* reproductively isolated from other species. Within this species, anthropologists can recognize a number of sub-species. It is to these sub-species of *Homo sapiens* that the term 'primary races' is applied. Most anthropologists recognize at least 4 primary races: Caucasoid, Negroid, Mongoloid and Australoid. However, there is no general agreement about what constitutes a 'primary race.' Some anthropologists recognize the American Indians as a separate 'primary race' (Amerindians); others regard the Bushmen and related peoples as constituting another 'primary race' (Capoids).

It must be emphasized that such classifications owe more to convenience than to science. The differences between groups are by no means clear-cut: there is considerable overlap of physical characteristics and there are ever-increasing numbers of people who are descended from more than one primary stock.

3. **Genetical definition:** *A race is a group of people who differ from other groups of people in the frequency of their genes.*

This definition of race has evolved from the application of relatively modern methods of genetical investigation and analysis to human populations. It has permitted a new biological approach to the problem of race and is providing a useful tool for assessing the relative roles of environment and heredity in producing human variability. It is with these genetical aspects of race that the remainder of this essay will be concerned. Before proceeding, however, it is necessary to refer to a fourth definition of 'race'.

4. **Political definition:** *A race is a group of people who are different from US.*

This attitude to race can be rejected forthwith. The machinations of politicians have severely impeded the biological investigation of race. This is a great pity because these investigations, if properly pursued, must surely lead to the benefit of all peoples. Some biologists are so concerned by the odium with which the term 'race' is now associated in a political context, that they would like to see it excluded from the vocabulary of science. Alternatives such as 'ethnic group' or 'genogroup' have been suggested. But why should we allow ourselves to be inconvenienced in this way? The term 'race' has been used for hundreds of years: it is short, concise and useful. Only more confusion would result from any attempt to replace it by one of the more cumbersome alternatives. I have little doubt that in time, the term 'race' will lose its unsavoury associations and that it will continue to enjoy an honourable place in our language.

THE GENETIC DEFINITION OF RACE

There are certain aspects of the genetic definition of race which require further comment. It will have been noted that this definition, like the sociological and anthropological definitions, refers to *groups* of people—not to individuals. There are no scientific criteria, genetic or otherwise, by which the racial affinities of individuals can be determined with any certainty. It will also be noted that our definition deals with the difference in the *frequency* of genes among different groups of people. In these terms, 'race' becomes a statistical concept : the races differ from each other in the *frequency* of their genes and not in the *nature* of their genes. The genetic constitution (genome) of a human being is made up of more than a million genes. It is estimated that all men have at least 95% of their gene pool in common ; variability exists only in respect of about 5% of the human genetic endowment. In other words, although certain genetic differences can be detected between peoples, these are far less numerous and important than the many points of similarity which characterize the human species.

GENETIC MARKERS FOR POPULATION STUDIES

No genetic trait is absolutely race-specific, but the frequencies of certain genes are much greater in some groups of people than in others. Genetic traits which show such variability are selected as genetic markers for population studies. To qualify as a genetic marker, a trait must have the following properties:

(*i*) It must be the manifestation of a single gene or gene complex. Polygenic traits such as height, head-shape, etc., are not suitable for this purpose.

(*ii*) The expression of the gene must not be directly influenced by the environmental variables. For example, skin colour and the serum-cholesterol level are to some extent genetically determined, but they are subject to too much environmental modification to be useful as genetic markers.

(*iii*) The trait should be easy to ascertain. In population studies large numbers of people have to be examined, and for this reason biochemical or other traits, which can only be detected by complicated, time-consuming laboratory procedures, are not suitable for this type of investigation.

Blood groups, various serum factors, certain enzymes and a number of physical traits are all useful as genetic markers. Table I is by no means a complete list ; it merely tabulates those genetic markers which are used regularly in the genetics laboratory of the Comprehensive Medicine Group.

TABLE I. SOME GENETIC MARKERS FOR ROUTINE USE IN POPULATION STUDIES

Blood group systems: ABO, MNS, Rhesus, Kell, Duffy, P, Diego.
Haemoglobin types.
Serum-protein systems: Haptoglobins, transferrins.
Immunologically determined serum systems: Gm groups, Inv groups.
Enzymes: Glucose-6-phosphate dehydrogenase activity and G6PD types, serum-cholinesterase types.
Physical traits: Colour vision, PTC tasting ability, finger-print pattern index.

Blood Groups

During World War I (1914 - 18), Professor and Mrs. L. Hirszfeld, working in Salonika, had the opportunity to determine the ABO blood groups of several thousand soldiers from different countries. They soon observed that groups of soldiers from different parts of the world had different ABO blood group distributions. This was the first indication of the potential usefulness of genetic traits for population studies. Since then, the results of literally millions of blood-group tests have been published. Most of these refer to the ABO system, but more or less extensive data are becoming available for all the common blood groups. In Table II, representative data are

TABLE II. REPRESENTATIVE DATA FOR THE COMMON BLOOD GROUP GENES IN THE FOUR PRIMARY RACES

Blood group system	*Gene*	*Gene frequency (percent)*			
		Caucasoid	*Negroid*	*Mongoloid*	*Australoid*
ABO	B	6	12	25	2
MNS	M	50	60	65	25
	S	35	20	5	0
	Henshaw	0	5	0	0
Rhesus	cde (r)	40	10	0	0
	cDe (Ro)	2	72	4	8
	CDe (R_1)	40	5	75	55
Kell	K	4	0	0	0
Duffy	Fy^a	40	7	85	100
Diego	Di^a	0	0	10	0
P	P_1	55	65	25	?

shown to illustrate the frequency distribution of the common blood group genes among the four primary races.

It will be seen that in some cases there are very striking differences in gene frequency. Thus, the Rhesus 'variant' cDe (Ro) has a very high frequency in Negroids, and the CDe (R_1) variant occurs most commonly in Mongoloids. Australoids generally have comparatively little of the M gene and almost no S gene. The Diego gene seems to be almost an exclusive 'Mongoloid trait' and is found mainly in the Far East and in South America.

In Table III a number of population groups have been classified arbitrarily on geographical lines. Using the distribution of the Rhesus positive trait as a marker, the genetic affinities of the different groups can clearly be seen. It is believed that the Rhesus negative trait originated among

TABLE III. THE PERCENTAGE DISTRIBUTION OF RHESUS POSITIVE INDIVIDUALS IN VARIOUS POPULATION GROUPS

Europe			
English	84	French	86
Scots	83	Italians	86
Irish	84	Russians (Moscow)	85
Welsh	85	Basques	65
Africa			
Negroes (Katanga)	97	Hottentots	100
Bantu (S. Africa)	95	Bushmen	100
North America			
US Whites	85	US Negroes (Chicago)	85
US Negroes (whole country)	90	Eskimo	99
Asia		*Australasia*	
Indians	94	Maori	100
Japanese	99	Australian aborigines	100
Chinese	100		

the Basques (at the border of Spain and France), 35% of whom are Rh-. In the course of centuries, as a result of the migrations of individuals and groups of individuals, the Rh- trait spread throughout Europe; hitherto its spread to Africa and Asia has been much less.

Haemoglobin Types

Some dozens of abnormal haemoglobins have been described, and the number is constantly being increased. The first to be discovered—HbS and HbC—have been studied in most detail. In the homozygous state they cause severe haemolytic anaemia (sickle-cell anaemia or HbC disease). The heterozygotes are clinically quite normal, but the presence of the abnormal gene can be detected by a fairly simple laboratory procedure and it is useful as a genetic marker. The thalassaemias are also a group of inherited haemolytic anaemias; in these the basic defect is not an abnormal haemoglobin but a depression of normal haemoglobin synthesis. Recognition of the unaffected heterozygotes is a little more difficult but it is extensively used as a genetic marker in some countries.

HbS is found mainly in Central and West Africa, in the Mediterranean countries and in parts of Arabia and India. HbC occurs mainly in West Africa around Ghana. The thalassaemia trait is found mainly in the Mediterranean area and South-East Asia. Abnormal haemoglobins are not common in Southern Africa and their ascertainment is of little use in population studies in this part of the world.

Serum-protein Systems

The electrophoretic separation of the serum-protein components on a medium such as starch gel shows several globulins whose structures are genetically determined. Two of the most extensively studied are the haptoglobins and the transferrins.

Haptoglobins are alpha-2 globulins which bind haemoglobin. Electrophoresis shows 3 common inherited types: haptoglobin 1:1, which shows up as a single, relatively fast-moving band and is the phenotypic manifestation of the gene Hp^1 in its homozygous state; haptoglobin 2:2 appears as a series of slower moving bands and represents the homozygous state of the gene Hp^2; haptoglobin 2:1 is a mixture of the fast and slow bands and represents the heterozygous genotype Hp^1Hp^2. The Hp^1 gene occurs with a frequency of about 40% and 65% in Caucasoids and Negroids respectively; it is much less common in Mongoloids and Australoids among whom the frequency is about 15%. In certain parts of Africa there is a high frequency of individuals with absent haptoglobins. This has been regarded as a manifestation of a genetically determined trait, HpO. However, studies in our laboratory have cast some doubt on this. We have found that absence of haptoglobins is often encountered as an acquired feature of liver disease. Liver disease is common in those areas where a high frequency of HpO has been reported, and it seems likely that in these areas HpO is an acquired rather than an inherited trait. These observations illustrate the importance of ensuring the absence of environmental influences before accepting a trait as a genetic marker.

Transferrins are iron-binding beta-globulins. Most people are homozygous for the common transferrin TfC which shows up as a single band on electrophoresis. However, in the serum of some individuals, double trans-

ferrin bands are found. The additional band migrates more slowly than the usual one and represents the gene Tf^{D_1}. This gene is particularly common in Australoids (frequently about 37%), but it is rarely found in the other races. A faster moving band representing the gene Tf^{B_2} has been found with a frequency of 1% in some European communities.

Immunologically Determined Serum Systems

The serum from some patients with rheumatoid arthritis contains an antibody which causes the agglutination of specially sensitized human red blood cells. Most, but not all, human beings have in their sera genetically determined gamma globulins which can inhibit this agglutination reaction. These individuals are either homozygous or heterozygous for the gene Gm^a; those whose sera fail to inhibit the agglutination reaction are homozygous for the contrasting allele Gm. In Caucasoids the frequency of the gene Gm^a is about 60%; in Negroids and Mongoloids it is almost 100%. Using different anti-sera and different combinations of sensitized red blood cells, several more genetically controlled gamma globulins have been detected. Some of these are controlled by other alleles at the Gm locus (Gm^a, Gm^b, Gm^x, etc.) while others represent a different gene complex at another locus (Inv^1, Inv^a).

Enzymes

Hundreds of enzymes are known, all of which are genetically determined. In the majority of cases elaborate laboratory procedures are necessary to determine the presence or absence of the enzyme. However, there are some whose presence can be demonstrated by relatively simple techniques. One of these is the enzyme *glucose-6-phosphate dehydrogenase* (G6PD), absence of which is associated with the development of haemolytic anaemia on exposure to fava beans, primaquin, sulphonamides and many other substances. G6PD deficiency is particularly common in Eastern Mediterranean communities, in parts of Asia and in the Negroids of West, Central and East Africa. There is evidence that different genetic loci are concerned in the Eastern Mediterranean and in the African forms of G6PD deficiency. G6PD deficiency is rare in Southern Africa and the ascertainment of this trait has not proved useful in our population studies in Cape Town.

Individuals who are homozygous for the gene which controls an atypical variant of the enzyme *serum cholinesterase* are likely to develop prolonged apnoea if suxamethonium is administered to them. The presence of the atypical cholinesterase can be determined by a fairly simple test, and this may prove to be a useful marker in population studies.

Physical Traits

Genes for the different forms of *colour blindness* are carried on the X chromosome. Colour blindness is therefore very uncommon in women, but it occurs in about 5 - 10% of men. The frequency varies from race to race, but the range of variability is rather small to be of much use in population surveys.

Phenylthiocarbamide (PTC) is an interesting substance for geneticists. In predominantly Caucasoid communities about 70% of people find it to have an intensely bitter taste, but about 30% cannot taste it at all. The ability to taste PTC is controlled by a single dominant gene and it has proved to be a useful marker in genetic surveys. The 30% of non-tasters among Caucasoids can be compared with only 5 - 10% of non-tasters among Negroids and Mongoloids and 50% non-tasters in Australoids.

Finger-print patterns are also genetically determined, but the genetics is more complex than in the traits which have already been described. Nevertheless, they are useful in genetic studies. There are 3 basic finger-print patterns: whorls, loops and arches. Two tri-radii are associated with a whorl, one with a loop and none with an arch. The simplest method of analysis is to count the number of tri-radii on all 10 digits; the total number of tri-radii is the 'pattern index.' If all 10 finger-print patterns are whorls, the index is 20; if they are all arches, the index is 0. The average pattern index of Negroids is about 10 - 11; in Caucasoids it is 12 - 13; in Mongoloids it is about 15. The variability is not great, but within each group the range of the pattern index is said to be small. Its usefulness in population genetics has yet to be established.

(To be continued)