METHOD OF MEASURING LINKAGE IN HUMAN GENETICS; WITH SPECIAL REFERENCE TO BLOOD GROUPS

ALEXANDER S. WIENER

Jewish Hospital of Brooklyn, New York

Received October 16, 1931

TABLE OF CONTENTS

INTRODUCTION

Whereas much has been learned regarding the genetics of many plants and animals in the past thirty years (the work of MORGAN and his co-workers on the *Drosophila melanogaster* is particularly notable), comparatively little progress has been made in man. The main reason for this is that the problems of human inheritance must be solved by somewhat different methods than those employed in plants and animals. Direct experimentation is obviously impossible, and data must often be accumulated by the less exact method of analyzing pedigrees and family histories. Studies in linkage, particularly, offer difficulties in human genetics. In plants or animals, when one wishes to determine whether or not any two mutations are linked in their heredity, all that is necessary is to cross two strains which possess the mutations in question, and examine the progeny for one or two generations. In man, however, data must be accumulated from crosses which are already made. Thus, if it were desired to study the linkage of white forelock and polydactylism, a search would have to be made for families presenting both of these anomalies. Since both of these anomalies are of infrequent occurrence, families possessing both traits at the same time must be exceedingly rare. Furthermore, because of the small size of human families, it is necessary to examine many families before any conclusions can be drawn, so that studies of this nature become practically impossible.

It is less difficult, of course, to determine whether or not sex-linkage is present, and at least fifteen human mutations have been found to be sex-

linked. The more important of the sex-linked characters are haemophilia, red-green color-blindness, optic atrophy (Leber's disease), and night blindness. Recently, an attempt was made by DAVENPORT (1930) to find families possessing two sex-linked traits at the same time, in order to determine the amount of crossing over between the sex-linked factors. Such families were found to be so rare, that no definite conclusions could be drawn. A remarkable family has been described by MADLENER (1928), however, in which a man who is both haemophilic and color-blind has a son who is haemophilic, two grandsons who are haemophilic and color-blind, and a great grandson who is haemophilic, but too young yet to determine whether or not color-blindness is present.

The discovery of the LANDSTEINER blood groups and the mechanism of their heredity opened the field for further studies in human linkage. Because of the simple heredity of the blood groups, and because of their universal distribution, it is possible now, by an examination of a sufficient number of families, to determine whether or not linkage exists between any hereditary trait and the blood groups. Although a number of such studies have already been made, no case of linkage has yet been found. Thus SNYDER (1929, 1931) found no evidence of linkage between polydactylism, telangiectasis, various eye anomalies and the blood groups; and LEVINE (1926) found no evidence of linkage between atopic hypersensitiveness and the blood groups. Because of the small number of such studies made thus far as compared with the number of chromosomes (twenty-four pairs), it is not surprising that not a single case of human linkage has been found, except for the cases of sex-linkage already mentioned. KUBANYI (1931) has determined the blood groups of members of haemophilic families, in order to determine the presence or absence of linkage. KUBANYI'S work, it would seem, is unnecessary, for haemophilia is a sex-linked trait, whereas the blood groups are not, so that linkage between these two characters is hardly possible.

More recently, two additional agglutinogens, *M* and *N*, have been described by LANDSTEINER and LEVINE (1928), and these agglutinogens also have a simple Mendelian heredity. From an inspection of their family material, LANDSTEINER and LEVINE concluded that the agglutinogens *M* and *N* are most likely inherited independently of the agglutinogens *A* and *B*. This discovery has therefore opened up further possibilities for linkage studies in human heredity.

Accordingly there are now three out of the total of twenty-four pairs of chromosomes marked by simple hereditary factors: one by sex, the second by the LANDSTEINER blood groups, and the third by the agglutinogens *M*

and *N* of LANDSTEINER and LEVINE. Whereas the color of eyes, hair and skin, finger-prints, etc. also have a universal distribution, their heredity is complex and not yet completely solved, so that they cannot be used as a basis for linkage studies. On the other hand, such characters as polydactylism, brachydactyly, zygodactyly, 'lobster claw,' etc., which have a simple heredity, are of too rare occurrence to be of much value. It seems obvious that conclusions regarding linkage can only be drawn from family material, and not from determinations of correlations between traits in a population. As is well known, there is a close association between hair-color and the color of the eyes. Thus, blue eyes are most commonly found associated with light brown or flaxen hair, and brown eyes with dark hair. Blue eyes combined with dark hair, and brown eyes combined with light hair are less commonly found. HAECKER (1925) suggests an explanation for this association on the ground that the genes for hair-color and eye-color are in the same chromosome. Then the "disharmonic phaenotypic combinations might be explained by crossing over." As DAVENPORT (1927) points out, however, HAECKER's interpretation of "crossing over" is not that of students of Drosophila. It seems more likely to the present author that the combinations of blue eyes with light hair and dark eyes with dark hair may have arisen from two different races living in different parts of the world. The less frequent combinations, light hair with brown eyes, and dark hair with blue eyes, would then represent hybrids resulting from crosses between individuals of these two races.

In the studies in human heredity that have thus far been made, the presence or absence of linkage was determined from an inspection of the family material. The need for an exact mathematical method of measuring linkage, applicable in human genetics, is the cause of this paper, as by a mere inspection of the families a loose linkage with much crossing over would be overlooked. BERNSTEIN (1931) has already worked out a method which he applied to the question of linkage between the agglutinogens *M* and *N* and the agglutinogens *A* and *B*, and arrived at the conclusion that *M* and *N* are most likely inherited independently of *A* and *B*. We have also derived a method that can be used wherever the question of linkage arises in human genetics. By way of illustration we shall apply the method to the question of linkage of the agglutinogens *M* and *N* to *A* and *B*.

ARE THE AGGLUTINOGENS *M* AND *N* LINKED TO *A* AND *B*?

The only type of families in which the presence of linkage can make any difference in the types of the offspring is that in which one of the parents is doubly heterozygous. That is, one of the parents must belong to one of

the following types: $AB++$, A_h++, or B_h++. The letters O, A, B, AB represent the blood groups; the first $+$ or $-$ sign the reaction for M, and the second $+$ or $-$ sign the reaction for N. We use the subscript "h" to represent the heterozygous genotype of groups A and B, and the subscript "p" to represent the homozygous or pure genotype. If there are group O children from group A parents, such parents are necessarily heterozygous. In table 1 the offspring of all matings involving these three types of parents are given, and the offspring are subdivided into three classes: "linked," "crossover," and "indeterminate" types. Thus, in the cross $AB++ \times O++$, if the genotype of the $AB++$ parent is $(AM)(BN)$, the linked gametes from that parent are (AM) and (BN), and the crossover gametes are (AN) and (BM). The $O++$ parent produces (RM) and (RN) gametes in equal numbers. The linked zygotes, therefore, are $(AM)(RM)$, $(AM)(RN)$, $(BN)(RM)$, and $(BN)(RN)$, and the crossover zygotes are $(AN)(RM)$, $(AN)(RN)$, $(BM)(RM)$, and $(BM)(RN)$. Hence, the linked types are $A+-$ and $B-+$, and the crossover types are $A-+$ and $B+-$; whereas the types $A++$ and $B++$ might be either linked or crossovers and are therefore placed in the column headed "indeterminate." The remainder of the table was worked out in a similar manner.

If we could determine whether an individual of type $AB++$ belongs to genotype $(AM)(BN)$ or $(AN)(BM)$, and similarly for the types A_h++ and B_h++, it would be comparatively simple to determine the crossover value. This could often be done by determining the types of the grandparents (see table 2). Thus, an individual of type $AB++$ with parents $A+-$ and $B++$, can only belong to genotype $(AM)(BN)$; such an individual with parents $A-+$ and $B++$ must belong to genotype $(AN)(BM)$; but if the parents are $A++$ and $B++$, the genotype is indeterminate. Unfortunately, however, it is very difficult, for obvious reasons, to obtain three generations in studies on human heredity. In the studies on the heredity of the agglutinogens A, B, M, and N, that have appeared thus far, the grandparents were not examined. It is therefore necessary to devise a method of measuring linkage based on data from only two generations and involving only comparatively small families.

In table 3 is presented an analysis of all the data that have thus far been accumulated on the heredity of the agglutinogens A, B, M, and N, that can be used for the purpose of measuring linkage. The second series of LANDSTEINER and LEVINE has not been published yet but was made available to us through the courtesy of the authors. The table only includes matings involving the types $AB++$, A_h++, and B_h++. With the aid of table 1 the offspring of these crosses were subdivided into three

TABLE 1

CROSS	LINKED	CROSSOVERS	INDETERMINATE
$AB++\times O++$	$A+-\ B-+$	$A-+\ B+-$	$A++\ B++$
$AB++\times O+-$	$A+-\ B++$	$A++\ B+-$	..
$AB++\times O-+$	$A++\ B-+$	$A-+\ B++$	..
$AB++\times A_p++$	$AB-+\ A+-$	$AB+-\ A-+$	$AB++\ A++$
$AB++\times A_p+-$	$AB++\ A+-$	$AB+-\ A++$	..
$AB++\times A_h+-$	$AB++\ A+-\ B++$	$AB+-\ A++\ B+-$	..
$AB++\times A_p-+$	$AB++\ A-+$	$AB-+\ A++$	..
$AB++\times A_h-+$	$AB++\ A-+\ B++$	$AB-+\ A++\ B-+$	..
$AB++\times B_p++$	$AB-+\ B+-$	$AB+-\ B-+$	$AB++\ B++$
$AB++\times B_p+-$	$AB++\ B+-$	$AB+-\ B++$	..
$AB++\times B_h+-$	$AB++\ A++\ B+-$	$AB+-\ A+-\ B++$	..
$AB++\times B_p-+$	$AB++\ B-+$	$AB-+\ B++$	..
$AB++\times B_h-+$	$AB++\ A++\ B-+$	$AB-+\ A-+\ B++$	..
$AB++\times AB+-$	$A+-\ B++$	$A++\ B+-$	$AB++\ AB+-$
$AB++\times AB-+$	$A-+\ B++$	$A++\ B-+$	$AB++\ AB-+$
$A_h++\times O++$	$A+-\ O-+$	$A-+\ O+-$	$A++\ O++$
$A_h++\times O+-$	$A+-\ O++$	$A++\ O+-$	..
$A_h++\times O-+$	$A-+\ O++$	$A++\ O-+$	..
$A_h++\times A_h+-$	$O++$	$O+-$	$A++\ A+-$
$A_h++\times A_h-+$	$O++$	$O-+$	$A++\ A-+$
$A_h++\times B_p++$	$AB+-\ B-+$	$AB-+\ B+-$	$AB++\ B++$
$A_h++\times B_p+-$	$AB+-\ B++$	$AB++\ B+-$	..
$A_h++\times B_h+-$	$AB+-\ A+-\ B++\ O++$	$AB++\ A++\ B+-\ O+-$	..
$A_h++\times B_p-+$	$AB-+\ B++$	$AB++\ B-+$	..
$A_h++\times B_h-+$	$AB-+\ A-+\ B++\ O++$	$AB++\ A++\ B-+\ O-+$	..
$A_h++\times AB+-$	$AB+-\ B++$	$AB++\ B+-$	$A++\ A+-$
$A_h++\times AB-+$	$AB-+\ B++$	$AB++\ B-+$	$A++\ A-+$
$B_h++\times O++$	$B+-\ O-+$	$B-+\ O+-$	$B++\ O++$
$B_h++\times O+-$	$B+-\ O++$	$B++\ O+-$	..
$B_h++\times O-+$	$B-+\ O++$	$B++\ O-+$	..
$B_h++\times A_p++$	$AB+-\ A-+$	$AB-+\ A+-$	$AB++\ A++$
$B_h++\times A_p+-$	$AB+-\ A++$	$AB++\ A+-$	..
$B_h++\times A_h+-$	$AB+-\ A++\ B+-\ O++$	$AB++\ A+-\ B++\ O+-$	..
$B_h++\times A_p-+$	$AB-+\ A++$	$AB++\ A-+$	..
$B_h++\times A_h-+$	$AB-+\ A++\ B-+\ O++$	$AB++\ A-+\ B++\ O-+$	..
$B_h++\times B_h+-$	$O++$	$O+-$	$B++\ B+-$
$B_h++\times B_h-+$	$O++$	$O-+$	$B++\ B-+$
$B_h++\times AB+-$	$AB+-\ A++$	$AB++\ A+-$	$B++\ B+-$
$B_h++\times AB-+$	$AB-+\ A++$	$AB++\ A-+$	$B++\ B-+$

NOTE: In this table the parents of types $AB++$, A_h++, and B_h++ are assumed to be of linkage types (AM) (BN), (AM) (RN), and (BM) (RN), respectively. For the linkage types (AN) (BM), (AN) (RM), and (BN) (RM) the same results hold except that the columns headed "linked" and "crossovers" are interchanged.

TABLE 2

PHENOTYPE	GENOTYPE	WILL BE PRODUCED BY CROSSES			
$AB++$	$(AM)(BN)$	$A+- \times \begin{cases} B++ \\ B-+ \\ AB++ \\ AB-+ \end{cases}$	$A++ \times \begin{cases} B-+ \\ AB-+ \end{cases}$	$AB+- \times \begin{cases} B-+ \\ B++ \end{cases}$	$AB++ \times B-+$
	$(AN)(BM)$	$A-+ \times \begin{cases} B++ \\ B+- \\ AB++ \\ AB+- \end{cases}$	$A++ \times \begin{cases} B+- \\ AB+- \end{cases}$	$AB-+ \times \begin{cases} B+- \\ B++ \end{cases}$	$AB++ \times B+-$
	Indeterminate	$\left.\begin{matrix} A++ \\ AB++ \end{matrix}\right\} \times \begin{cases} B++ \\ AB++ \end{cases}$	$AB++ \times \begin{cases} AB-+ \\ AB+- \end{cases}$	$AB-+ \times AB+-$	
$A++$	$(AM)(RN)$	$\left.\begin{matrix} A+- \\ AB+- \end{matrix}\right\} \times \begin{cases} O-+ \\ O++ \\ B-+ \\ B++ \end{cases}$	$\left.\begin{matrix} A++ \\ AB++ \end{matrix}\right\} \times \begin{cases} O-+ \\ B-+ \end{cases}$		
	$(AN)(RM)$	$\left.\begin{matrix} A-+ \\ AB-+ \end{matrix}\right\} \times \begin{cases} O+- \\ O++ \\ B+- \\ B++ \end{cases}$	$\left.\begin{matrix} A++ \\ AB++ \end{matrix}\right\} \times \begin{cases} O+- \\ B+- \end{cases}$		
	Indeterminate	$\left.\begin{matrix} A++ \\ A+- \end{matrix}\right\} \times \begin{cases} A++ \\ A-+ \\ AB++ \\ AB-+ \end{cases}$	$\left.\begin{matrix} A++ \\ AB++ \end{matrix}\right\} \times \begin{cases} O++ \\ B++ \end{cases}$		

TABLE 2 (continued)

PHENOTYPE	GENOTYPE	WILL BE PRODUCED BY CROSSES
	(*BM*) (*RN*)	$\left.\begin{matrix}B+-\\AB+-\end{matrix}\right\}\times\left\{\begin{matrix}O-+\\O++\\A-+\\A++\end{matrix}\right.$ $\left.\begin{matrix}B++\\AB++\end{matrix}\right\}\times\left\{\begin{matrix}O-+\\A-+\end{matrix}\right.$
B++	(*BN*) (*RM*)	$\left.\begin{matrix}B-+\\AB-+\end{matrix}\right\}\times\left\{\begin{matrix}O+-\\O++\\A+-\\A++\end{matrix}\right.$ $\left.\begin{matrix}B++\\AB++\end{matrix}\right\}\times\left\{\begin{matrix}O+-\\A+-\end{matrix}\right.$
	Indeterminate	$\left.\begin{matrix}B++\\B+-\end{matrix}\right\}\times\left\{\begin{matrix}B++\\B-+\\AB++\\AB-+\end{matrix}\right.$ $\left.\begin{matrix}B++\\AB++\end{matrix}\right\}\times\left\{\begin{matrix}O++\\A++\end{matrix}\right.$

TABLE 3

	s=2			s=3			s=4			s=5			s=6			s=7			s=8			s=9			s=10		
	FAMILY NUMBER	U	V	FAMILY NUMBER	U	V	FAMILY NUMBER	U	V	FAMILY NUMBER	U	V	FAMILY NUMBER	U	V	FAMILY NUMBER	U	V	FAMILY NUMBER	U	V	FAMILY NUMBER	U	V	FAMILY NUMBER	U	V
	45	2	0	3	2	1	20	3	1	16	3	2	91	4	2	119	4	3	..	..	..	118	6	3	..	..	..
	124	2	0	96	3	0	27	2	2	122	3	2	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
	148	2	0	103	2	1	31	3	1	164	3	2	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
	..	..	..	154	2	1	35	2	2	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
LANDSTEINER and	..	..	..	156	2	1	72	3	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
LEVINE (First Series)	..	..	..	..	..	..	104	3	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
	..	..	..	..	..	..	107	3	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
	..	..	..	..	..	..	110	3	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
	..	..	..	..	..	..	114	3	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
	..	..	..	..	..	..	146	2	2	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
	..	..	..	..	..	..	160	3	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
Totals		6	0		11	4		29	15		9	6		4	2		4	3					6	3			

TABLE 3 (continued)

	s=2			s=3			s=4			s=5			s=6			s=7			s=8			s=9			s=10		
	FAMILY NUMBER	U	V	FAMILY NUMBER	U	V	FAMILY NUMBER	U	V	FAMILY NUMBER	U	V	FAMILY NUMBER	U	V	FAMILY NUMBER	U	V	FAMILY NUMBER	U	V	FAMILY NUMBER	U	V	FAMILY NUMBER	U	V
	6	1	1	10	2	1	21	4	0	8	3	2	4	5	1	44	5	2	51	6	2	..	..	..	108	6	4
	76	2	0	14	2	1	23	3	1	13	4	1	11	3	3	60	5	2	103	4	4	..	..	..	..	..	..
	..	..	..	25	3	0	31	3	1	17	3	2	79	4	2	81	4	3	..	..	..	..	..	..	..	..	..
	..	.	..	48	2	1	39	2	2	20	4	1	94	5	1	109	5	2	..	..	..	..	..	..	..	..	..
WIENER and	..	..	..	74	2	1	52	2	2	53	5	0	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
VAISBERG	..	..	..	90	2	1	82	2	2	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
	..	..	..	96	3	0	89	3	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
	..	..	..	110	2	1	114	3	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
	..	..	..	113	2	1	124	3	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
	..	..	..	119	2	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
Totals		3	1		22	8		25	11		19	6		17	7		19	9		10	6					6	4
	3	1	1	11	2	1	1	3	1	10	3	2	38	4	2	22	4	3	14	4	4	..	..	..	..	..	..
	42	1	1	31	2	1	27	2	2	15	3	2	..	..	..	67	5	2	..	..	..	..	..	..	..	..	..
	52	1	1	78	3	0	49	3	1	68	3	2	..	..	..	75	4	3	..	..	..	..	..	..	..	..	..
LANDSTEINER and	74	2	0	84	2	1	55	2	2	77	3	2	..	..	..	100	5	2	..	..	..	..	..	..	..	..	..
LEVINE (Second	81	1	1	111	3	0	57	4	0	89	3	2	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
Series)	97	1	1	..	..	..	65	2	2	91	3	2	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
	101	2	0	..	..	..	93	2	2	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
	108	2	0	..	..	..	94	3	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
	110	2	0	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
Totals		13	5		12	3		21	11		18	12		4	2		18	10		4	4						
Grand Totals		22	6		45	15		75	37		46	24		25	11		37	19		14	10		6	3		6	4
N=	14			20			28			14			6			8			3			1			1		
Q=	21.43 percent			25.00 percent			33.04 percent			34.29 percent			30.56 percent			33.93 percent			41.66 percent			33.33 percent			40.00 percent		

classes: linked, crossovers, and indeterminates. The indeterminates were omitted from consideration, and the number of children remaining taken equal to s, and the column the family was listed in determined. Since the crossovers and the linked types could not be differentiated, the consistent system was adopted of placing the larger number in the column headed U, and the smaller number in the column headed V. Thus, family 96 of the study of WIENER and VAISBERG is a cross of an $A+-$ father by an $A++$ mother, producing three $O+-$, four $A++$, and one $A+-$ children. According to table 1, therefore, there are no linked type children, the three $O+-$ children are crossovers, and the $A+-$ and four $A++$ children are indeterminate. This family is therefore listed under the column headed s = 3, with U = 3, V = O. The other families listed in table 3 were analyzed in a similar manner.

The columns headed U and V were then totalled, and Q, the percentage value of V, was determined. It is possible to determine the most probable value of Q for various sized families (different values of s) and for any crossover value. The results of such calculations are tabulated in table 4.

TABLE 4

Value of Q for various sized families and various crossover values.

NUMBER OF CHILDREN	CROSSOVER VALUE									
	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
s=2–3	0.0475	0.0900	0.1275	0.1600	0.1875	0.2100	0.2275	0.2400	0.2475	0.2500
s=4–5	0.0498	0.0981	0.1237	0.1856	0.2226	0.2541	0.2792	0.2976	0.3088	0.3125
s=6–7	0.0499	0.0996	0.1479	0.1938	0.2358	0.2726	0.3028	0.3252	0.3391	0.3437
s=8–9	0.0500	0.0999	0.1492	0.1971	0.2420	0.2823	0.3162	0.3418	0.3578	0.3633
s=10–11	0.0500	0.1000	0.1496	0.1985	0.2453	0.2881	0.3247	0.3530	0.3708	0.3769
s=12–13	0.0500	0.1000	0.1498	0.1992	0.2471	0.2917	0.3305	0.3610	0.3805	0.3872
s=14–15	0.0500	0.1000	0.1499	0.1996	0.2482	0.2940	0.3347	0.3671	0.3880	0.3951
s=16–17	0.0500	0.1000	0.1499	0.1998	0.2488	0.2957	0.3378	0.3718	0.3940	0.4016

NOTE: The value of Q in general, is the same for two values of s such as 2n and 2n+1.

In order to evaluate the significance of the difference between the value of Q found and its theoretically to be expected value, it is necessary to know the probable error. In table 5 the standard error of Q is given. To derive the probable error from the value given in table 5, it is necessary to multiply by the factor $0.6745/\sqrt{N}$, where N represents the number of families analyzed. The derivation of formulae for Q and its standard error are given at the end of the paper.

In table 6, the actual value of Q is compared with the values theoreti-

TABLE 5

Standard error of Q.

NUMBER OF CHILDREN	CROSSOVER VALUE									
	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
s=2	0.1466	0.1921	0.2179	0.2333	0.2420	0.2468	0.2490	0.2498	0.2500	0.2500
s=3	0.1165	0.1480	0.1639	0.1665	0.1654	0.1609	0.1552	0.1497	0.1456	0.1443
s=4	0.1211	0.1448	0.1658	0.1765	0.1802	0.1788	0.1750	0.1731	0.1668	0.1653
s=5	0.0964	0.1284	0.1452	0.1545	0.1548	0.1467	0.1388	0.1305	0.1239	0.1218
s=6	0.0871	0.1210	0.1408	0.1514	0.1548	0.1527	0.1467	0.1411	0.1337	0.1313
s=7	0.0803	0.1118	0.1294	0.1417	0.1397	0.1353	0.1301	0.1172	0.1094	0.1069
s=8	0.0769	0.1068	0.1241	0.1383	0.1392	0.1392	0.1313	0.1225	0.1151	0.1118
s=9	0.0705	0.0995	0.1156	0.1262	0.1295	0.1271	0.1188	0.1080	0.0989	0.0955
s=10	0.0689	0.0937	0.1119	0.1225	0.1273	0.1268	0.1206	0.1112	0.1092	0.0993

TABLE 6

CROSSOVER VALUE	s=2			s=3			s=4		
	Q	DEV.	P.E.	Q	DEV.	P.E.	Q	DEV.	P.E.
0.50	25.00	−3.57	4.51	25.00	0.00	2.18	31.25	+1.79	2.12
0.45	24.75	−3.32	4.46	24.75	+0.25	2.20	30.88	+2.16	2.13
0.40	24.00	−2.57	4.33	24.00	+1.00	2.26	29.76	+3.28	2.21
0.35	22.75	−1.32	4.10	22.75	+2.25	2.34	27.92	+5.12	2.23
0.30	21.00	+0.43	3.79	21.00	+4.00	2.43	25.41	+7.63	2.28
Actual	21.43			25.00			33.04		

CROSSOVER VALUE	s=5			s=6			s=7		
	Q	DEV.	P.E.	Q	DEV.	P.E.	Q	DEV.	P.E.
0.50	31.25	+3.04	2.18	34.37	−3.81	3.53	34.37	−0.46	2.55
0.45	30.88	+3.41	2.23	33.91	−3.35	3.68	33.91	+0.02	2.61
0.40	29.76	+4.53	2.35	32.52	−1.96	3.80	32.52	+1.41	2.80
0.35	27.92	+6.37	2.50	30.28	+0.28	4.04	30.28	+3.65	3.10
0.30	25.41	+8.88	2.64	27.26	+3.30	4.20	27.26	+6.67	3.22
Actual	34.29			30.56			33.93		

CROSSOVER VALUE	s=8			s=9			s=10		
	Q	DEV.	P.E.	Q	DEV.	P.E.	Q	DEV.	P.E.
0.50	36.33	+5.33	4.35	36.33	−3.00	6.44	37.69	+2.31	6.70
0.45	35.78	+5.88	4.48	35.78	−2.45	6.67	37.08	+2.92	7.36
0.40	34.18	+7.48	4.77	34.18	−0.85	7.28	35.30	+4.70	7.50
0.35	31.62	+10.04	5.11	31.62	+1.71	8.01	32.49	+7.51	8.13
0.30	28.23	+13.43	5.42	28.23	+5.10	8.57	28.81	+11.19	8.55
Actual	41.66			33.33			40.00		

TABLE 7

	ACTUAL		CROSSOVER FREQUENCY									
			0.50		0.45		0.40		0.35		0.30	
	TOTAL NUMBER OF CHILDREN	V	V	$(P.E._v)^2$	V	$(P.E._v)^2$	V	$(P.E._v)^2$	V	$(P.E._v)^2$	V	$(P.E._v)^2$
s=2	28	6	7.00	1.59	6.93	1.56	6.72	1.46	6.27	1.32	5.88	1.12
s=3	60	15	15.00	1.72	14.85	1.74	14.40	1.79	13.56	1.96	12.60	2.13
s=4	112	37	35.00	5.62	34.59	5.71	33.33	6.15	31.27	6.25	28.46	6.50
s=5	70	24	21.88	2.34	21.62	2.43	20.83	2.69	19.54	3.06	17.79	3.34
s=6	36	11	12.37	1.61	12.21	1.74	11.71	1.88	10.90	2.10	9.81	2.28
s=7	56	19	19.25	2.04	18.89	2.13	18.20	2.47	16.96	3.03	15.27	3.24
s=8	24	10	8.72	1.08	8.59	1.16	8.20	1.30	7.59	1.51	6.78	1.69
s=9	9	3	3.27	0.34	3.22	0.36	3.08	0.42	2.85	0.51	2.53	0.59
s=10	10	4	3.77	0.45	3.71	0.54	3.53	0.56	3.25	0.66	2.88	0.74
Totals	405	129	126.26	16.79	124.71	17.37	120.00	18.72	112.28	20.40	102.00	21.63
Deviation			2.74		4.29		9.00		16.72		27.00	
Probable error				4.09		4.17		4.33		4.51		4.65
Dev./P.E.			0.67		1.03		2.08		3.68		5.81	

cally to be expected for various crossover frequencies. For an assumed crossover value of 50 percent (assumption of independence), the deviation was found to be less than the probable error for $s = 2, 3, 4, 7, 9, 10$, and less than twice the probable error for $s = 5, 6, 8$. The results, therefore, do not contradict the assumption that M and N are inherited independently of A and B. For an assumed crossover frequency of 30 percent, on the other hand, the deviation was found to be more than three times the probable error for $s = 3, 4$; and more than twice the probable error for $s = 7, 8$. A linkage intensity as great as 70 percent can therefore certainly be excluded. Furthermore, for an assumed crossover frequency of 35 percent, the deviation is almost two and a half times the probable error for $s = 3, 4$, so that the existence of a linkage intensity even as small as 65 percent is rendered unlikely.

It is possible to bring our conclusions within narrower limits by combining the results for the various sized families. This is done in table 7. The desired result was accomplished by converting the percentage expectancies (Q) into the corresponding absolute values (V) for each sized family, and then obtaining the total absolute expectancies for V for the various crossover values. These results were compared with the total absolute value of V actually found. The absolute value of the probable error was also calculated. It was found that the deviation was more than three times the probable error for an assumed crossover value of 35 percent, so that a linkage intensity of 65 percent can now certainly be excluded. The actual value of V found agrees best with the expectancy for a crossover value of 50 percent (equivalent to the assumption of independence). Considering the fact that there are twenty-four pairs of chromosomes, so that *a priori* the likelihood of linkage is very small, our results indicate that the agglutinogens M and N are most probably inherited independently of A and B.

DERIVATION OF FORMULA FOR Q

If p represents the true linkage intensity and q the crossover frequency (so that $p+q=1$) and s equals the number of "determinate" children, then the probability that μ of these children be linked types and ν be crossovers is given by:

$$w_{\mu,\nu} = \frac{s!}{\mu!\nu!}p^{\mu}q^{\nu}.$$

Then,

$$Q = \Sigma\frac{\mu}{s}\frac{s!}{\mu!\nu!}(p^{\mu}q^{\nu} + p^{\nu}q^{\mu}), \text{ where } \mu \text{ is less than } \nu.$$

This may be reduced to:

$$Q = pq + (pq)^2 + 2(pq)^3 + 5(pq)^4 + 14(pq)^5 + 42(pq)^6 + 132(pq)^7 + 429(pq)^8 + \cdots . \quad (1)$$

To determine the value of Q from this formula, s/2 terms are taken if s is even, and (s−1)/2 terms are taken if s is odd. Thus, if s = 7, and p = 0.50 (so that pq = 0.25), $Q = (0.25) + (0.25)^2 + 2(0.25)^3$ or Q = 0.34375.

The remaining values of Q given in table 4 were calculated in a similar manner from formula (1).

DERIVATION OF FORMULA FOR STANDARD ERROR OF Q

The standard error of Q may now be derived as follows:

Let

$$\frac{s!}{\mu!\nu!}p^\mu q^\nu = A_\mu$$

Then,

$$Q = \Sigma\frac{\mu}{s}(A_\mu + A_{(s-\mu)}).$$

Let A_μ' be the actual frequency of families with μ children of linked types and $(s-\mu)$ children of crossover types found in a study of N families of s children each.

Then,

$$A_\mu' = A_\mu + x_\mu$$

$$\sqrt{N}\,\sigma_{A_\mu} = \sqrt{A_\mu(1 - A_\mu)} = \sqrt{\Sigma x_\mu^2}.$$

Similarly, if Q′ equal the actual value found and Q the theoretically to be expected value,

$$Q' - Q = \Sigma\frac{\mu}{s}(x_\mu + x_{(s-\mu)}) = x_Q$$

$$N \cdot \sigma_Q^2 = \Sigma xQ^2 = \Sigma\frac{1}{s^2}[\Sigma\mu(x_\mu + x_{(s-\mu)})]^2$$

Since,

$$\Sigma x_\mu x_{(s-\mu)} = 0$$

$$N\sigma_Q^2 = \Sigma\frac{1}{s^2}[\Sigma\mu^2(x_\mu^2 + x_{(s-\mu)}^2) + \Sigma 2\mu\nu(x_\mu + x_{(s-\mu)})(x_\nu + x_{(s-\nu)})]$$

$$\Sigma x_\mu^2 = N\sigma_{A_\mu}^2 = A_\mu(1 - A_\mu)$$

$$\Sigma x_\mu x_\nu = N r_{\mu\nu}\sigma_\mu\sigma_\nu .$$

Since $A_1, A_2, \cdots, A_{s-1}$ form a Bernoulli system,

$$r_{\mu\nu} = -\sqrt{\frac{A_\mu A_\nu}{(1-A_\mu)(1-A_\nu)}}$$

so that

$$\Sigma x_\mu x_\nu = -A_\mu A_\nu$$

$$\begin{aligned} N\sigma_Q^2 &= \frac{1}{s^2}\{\Sigma\mu^2[A_\mu(1-A_\mu) + A_{(s-\mu)}(1-A_{(s-\mu)})] \\ &\qquad - 2\Sigma\mu\nu(A_\mu + A_{(s-\mu)})(A_\nu + A_{(s-\nu)})\} \\ &= \frac{1}{s^2}\{\Sigma\mu^2(A_\mu + A_{(s-\mu)}) - [\Sigma\mu^2(A_\mu^2 + A_{(s-\mu)}^2) \\ &\qquad + \Sigma 2\mu\nu(A_\mu + A_{(s-\mu)})(A_\nu + A_{(s-\nu)})]\} \\ &= \frac{1}{s^2}\{\Sigma\mu^2(A_\mu + A_{(s-\mu)}) - [\Sigma\mu(A_\mu + A_{(s-\mu)})]^2\} \end{aligned}$$

$$\sqrt{N}\sigma_Q = \frac{1}{s}\sqrt{\Sigma\mu^2(A_\mu + A_{(s-\mu)}) - [\Sigma\mu(A_\mu + A_{(s-\mu)})]^2}.$$

Since,

$$\Sigma\frac{\mu}{s}(A_\mu + A_{(s-\mu)}) = Q$$

and since it can be shown that

$$\Sigma\frac{\mu^2}{s^2}(A_\mu + A_{(s-\mu)}) = Q - \frac{s-1}{s}pq,$$

then,

$$\sqrt{N}\sigma_Q = \sqrt{Q - \frac{s-1}{s}pq - Q^2}.$$

Thus, if $s=5$ and $p=q=\frac{1}{2}$

$$Q = 0.3125 \text{ (from table 4)}$$

$$Q^2 = 0.09766$$

$$\frac{s-1}{s}pq = 0.20$$

$$\begin{aligned} \sqrt{N}\sigma_Q &= \sqrt{0.01484} \\ &= 0.1218. \end{aligned}$$

In making our calculations we have not taken into account the effect of selection of families sometimes necessary because a doubly heterozygous parent can often only be recognized from the phenotypes of the children. Thus three kinds of cases may be distinguished:

(1) When studying linkage of the LANDSTEINER blood groups and the agglutinogens M and N, if one parent is of type $AB++$, we know immediately that this parent is doubly heterozygous, as the blood group AB and type $M+N+$ are both heterozygous.

(2) If the parent in question is of type $A++$, however, we cannot tell if he is doubly heterozygous unless at least one child is of group O.

(3) When studying linkage of blood groups and eye color, for example, if the parent in question is of group A and has brown eyes, we can only be certain he is doubly heterozygous if at least one child is of group O and at least one child has blue eyes.

In the second and third cases, therefore, there will be selection of families. The values of Q and its standard error given in tables 4 and 5 apply only to the former two matings. For the third kind of mating a correction is necessary. Thus for the first series of LANDSTEINER and LEVINE a correction is necessary for families 16, 20, 27, 31, 35, 45, 72, 93 and 96, which were examined for agglutinogen M but not for agglutinogen N. This correction is small and does not materially affect the final results, however.

SUMMARY

A method of measuring linkage that may be applied in human genetics is presented. The method of application is illustrated by an analysis of data on the heredity of the agglutinogens A, B, M, and N in order to determine whether or not linkage exists between the agglutinogens M and N and the agglutinogens A and B.

Considering the fact that there are twenty-four pairs of chromosomes so that *a priori* the probability of linkage between any two factors is small, and also considering the fact that the results of the analysis fit best with the assumption of independence of M and N from A and B, it becomes most probable that M and N are inherited independently of A and B. Our findings therefore confirm those of LANDSTEINER and LEVINE, and of BERNSTEIN.

The method used depends upon the elimination of the necessity for the determination of the genotype of the heterozygous plant. This is accomplished by dividing the children into two main classes, linked and crossover types. Since the two classes cannot be differentiated, the system was

adopted of calling the larger class U and the smaller class V. The total value of V for all the families examined can then be compared with the expectancies for various assumed linkage intensities. BERNSTEIN attained a similar result by taking the product of the two classes, thus eliminating the necessity for determining the genotype of the heterozygous parent.

LITERATURE CITED

BERNSTEIN, F., 1931 Zur Grundlegung der Chromosomentheorie der Vererbung beim Menschen mit besonderer Berücksichtung der Blutgruppen. Z. indukt. Abstamm.-u. Vererb-Lehre. **57**: 113–138.

DAVENPORT, C. B., 1927 Heredity of human eye color. Bibl. genet. **3**: 443–463.

1930 Sex linkage in man. Genetics **15**: 401–444.

HAECKER, V. 1925 Pluripotenerscheinungen, Synthetische Beitrage zur Vererbungs- und Abstammunglehre. Fischer: Jena.

KUBANYI, A., 1931 Weitere Untersuchungen über die Beziehung zwischen Hemophilieverebung und Blutgruppencharakter, Klin. Wchnschr. **10**: 597–601.

LANDSTEINER, K., and LEVINE, P., 1928 On individual differences in human blood. J. Exp. Med. **47**:757–775.

1928 On the inheritance of agglutinogens of human blood demonstrable by immune agglutinins. J. Exp. Med. **48**: 731–748.

LEVINE, P., 1926 Atopic hypersensitiveness. J. Immunol. **11**: 283–297.

MADLENER, M., 1928 Eine Bluter-Familie, Arch. Rass.-u. Ges. Biol. **20**(4): 390–394.

MORGAN, T. H., 1926 The theory of the gene. 358 pp. Yale Univ. Press: New Haven, Conn.

SNYDER, L., 1929 Blood grouping and its relation to clinical and legal medicine. Williams and Wilkins: Baltimore, Md.

1931 Linkage in man. Eugenical News **16**: 177–179.

WIENER, A. S. and VAISBERG, M., 1931 Heredity of the agglutinogens *M* and *N* of LANDSTEINER and LEVINE. J. Immunol. **20**: 371–388.