FORMULAE AND TABLES FOR CALCULATING LINKAGE INTENSITIES[1*]

F. R. IMMER

Minnesota Agricultural Experiment Station, University Farm, St. Paul, Minnesota

Received April 29, 1929

TABLE OF CONTENTS

INTRODUCTION

At present many different methods are used by the various workers in the field of Genetics to calculate linkage intensities from F_2 data.

In some cases two or more methods applied to the same data give widely different results.

A general method which is satisfactory for all ordinary linkage problems would help greatly to reduce the confusion which exists now because of the great number of methods being used.

BATESON and PUNNETT (1911) were the first to give a method for measuring linkage intensity. Inspection was relied on to determine the closeness of fit between calculated and observed ratios. COLLINS (1912, 1924) made the coefficient of association (YULE, 1900) the basis of a method for calculating linkage intensities. BRIDGES (1914) used the same method as COLLINS (1912). EMERSON (1916) presented a formula which FISHER (1928) has referred to as the additive method and CASTLE (1916) used the same basis for his method. WOODWORTH (1923) and BRUNSON (1924) developed formulae for cases in which duplicate or complementary factors were involved which were based on EMERSON'S method. HOR (1924), BABCOCK and CLAUSEN (1927), ALBERTS (1926) and KAPPERT (1927) have presented still other formulae.

[1] Published with the approval of the Director as Journal Series Paper No. 861, Department of Agriculture, UNIVERSITY of MINNESOTA, University Farm, St. Paul, Minnesota.

* Part of the cost of the accompanying tables is paid by the GALTON AND MENDEL MEMORIAL FUND.

WELLENSIEK (1927) has given two methods for calculating the actual gametic F_2 series from a given zygotic series. The first method gives the actual ratio separately for each of the four types of gametes but does not show differences between the male and female gametic series. The second method makes it possible to calculate the actual gametic ratio of the four types of gametes for both male and female gametic series.

OWEN (1928) used the product moment correlation coefficient in developing formulae for calculating linkage intensities. His formulae have the added advantage of convenient algebraic manipulation. Further information regarding a linkage problem may be obtained with very little additional calculation.

HALDANE (1919) presented both a formula for calculating the crossover percentage from F_2 data and a formula for obtaining the probable error. Recently FISHER (1928) and FISHER and BALMUKAND (1928) have given a critical analysis of several of the above mentioned formulae. Methods are given by which the probable error formulae may be determined (FISHER 1928) and a number of ways for comparing the relative efficiency of different formulae. A further comparison of the relative efficiency of certain of the current methods of calculating linkage intensities is given in this paper. Tables are appended which greatly simplify the calculation by the product method of linkage intensities and their probable errors.

EFFICIENCY OF SEVERAL DIFFERENT FORMULAE

In order to be most useful a formula for calculating linkage intensities should be easy to use, it should be disturbed as little as possible by differential mortality of gametes or zygotes, and it should be statistically efficient. The term efficient is here used in the statistical sense as meaning that the formula has a probable error as small as possible. The comparative efficiency of other formulae may be determined by dividing the variance (squared standard deviation) of an efficient formula by the variance of the formula in question. The efficiency of several formulae will be discussed

In the discussion in this paper the symbol p will be used to designate the crossover percentage, expressed as a decimal fraction, where the cross was made in the repulsion phase. In repulsion crosses p will then vary from 0 to about .50. In the coupling phase p will represent the percentage of parental combinations, expressed as a decimal fraction. In that case p will vary from about .50 to 1.00 and $1-p$ will then be the crossover percentage. All formulae presented in this paper will use the symbol p in the manner defined above.

FISHER (1928) and FISHER and BALMUKAND (1928) state that the maximum likelihood method (for which they propose the symbol T_4, where $T_4=p^2$) will in all cases have a probable error, in the theory of large samples, as small as possible. This method consists in multiplying the logarithm of the number expected in each of the four F_2 phenotypes by the number observed, summing for the four classes and finding the value of p^2 which will make this sum a maximum. It is stated further that HALDANE (1919) could have arrived at his formula only by using the maximum likelihood solution. This method can then be used as a standard of comparison, from the probable error standpoint, for other methods.

FISHER (1928) and FISHER and BALMUKAND (1928) have shown that EMERSON'S method, called the additive method by these writers, is efficient *only* for close linkage in the coupling phase. The fraction of information utilized for various crossover percentages by the additive or EMERSON method may be obtained by dividing the variance of the maximum likelihood method by the variance of the additive method, which leads to the formula $2p^2(2+p^2)/(1+2p^2)(1+p^2)$, where p is used as previously defined. Substituting the values $p=.90$ (10 percent crossing over in the coupling phase) and $p=.10$ (10 percent crossing over in repulsion) in the above formula, we find that EMERSON'S formula is 96 percent efficient at 10 percent crossing over in the coupling phase and 4 percent efficient at 10 percent crossing over in the repulsion phase; that is, with 10 percent crossing over in repulsion it utilizes but 4 percent of the information which would be utilized by the maximum likelihood method or the product method. At 50 percent crossing over the additive or EMERSON method is 60 percent efficient. This emphasizes the errors which might be encountered by using EMERSON'S method as a general method, particularly in the repulsion phase.

FISHER (1928) and FISHER and BALMUKAND (1928) give a product method (for which the symbol T_3 is proposed, where $T_3=p^2$) which is equivalent to the coefficient of association method developed by COLLINS (1912, 1924) and BRIDGES (1914). It has the same probable error as the maximum likelihood or HALDANE'S method. Therefore it is of equal efficiency with the maximum likelihood method, which we have previously accepted as our standard for judging the efficiency of other formulae.

It has been shown by COLLINS (1924) and OWEN (1928) that the coefficient of association method (which is equivalent to the product method), of all the methods compared, seems to be affected the least by differential mortality of gametes or zygotes, although OWEN'S (1928) correlation coefficient method is but slightly inferior. The ease of calculating linkage intensities by the product method from the calculated tables presented in this paper will be discussed later. It seems, therefore, that the product method is the best general method available since it is the easiest to use when suitable tables are available, it is affected the least by differential mortality, and it has probable error, in the theory of large samples, as small as possible.

THE PRODUCT METHOD

Given the four observed F_2 phenotypic classes as AB, Ab, aB and ab and designating these by a, b, c, and d respectively, FISHER'S product method formula for two 3:1 ratios is $ad/bc = p^2(2+p^2)/(1-p^2)^2$. In calculating linkage intensities by this method, the observed frequencies for the four classes, a, b, c, and d are substituted in this formula and the value of p determined.

Since the product method seems to be the best general method available, because of its ease of calculation when tables are available, the magnitude of the probable error which is as small as possible, and the fact that it is disturbed the least by differential mortality of gametes or zygotes, it seems desirable to extend the method to the more complex ratios in which duplicate or complementary factors are involved. Such formulae are given in table 1. The formulae which have been available for these ratios were based either on EMERSON'S method, which in the case of two 3:1 ratios has been shown to be inefficient except for close linkage in the coupling phase, or on OWEN'S correlation coefficient method for which no method for developing probable error formulae has yet been presented. The determination of p for any of the factor relationships dealt with in table 1 then simply resolves itself into a solution of the proper quadratic equation. In these formulae a linkage is assumed between but one factor pair for each of the character pairs concerned.

Formulae for the probable errors of linkage intensities calculated by the product method may also be developed. FISHER (1928) and FISHER and BALMUKAND (1928) give the probable error of p calculated from two 3:1 ratios as $.6745\sqrt{(1-p^2)(2+p^2)/2(1+2p^2)N}$ where N is the total frequency. The formula for the probable error of p calculated from a backcross is usually given as $.6745\sqrt{p(1-p)/N}$. FISHER (1928) has given a general method by which the probable errors of other product method formulae may be determined. Such probable error formulae are given in table 1. The probable error concept applied to linkage problems should furnish the sound basis for judging the reliability of calculated linkage intensities which has been lacking so often in the past.

TABLE 1

Formulae for the calculation of crossover values and their probable errors based on the product method. In repulsion p is the crossover percentage and in coupling 1−p is the crossover percentage, expressed as decimal fractions.

PHENOTYPIC RATIOS	FORMULAE TO CALCULATE p	PROBABLE ERROR
3:1 and 1:1	$\frac{ad}{bc}=\frac{p+p^2}{2-3p+p^2}$	$.6745\sqrt{\frac{2p(2-p)(1-p^2)}{N(1+2p-2p^2)}}$
3:1 and 3:1	$\frac{ad}{bc}=\frac{2p^2+p^4}{1-2p^2+p^4}$	$.6745\sqrt{\frac{(1-p^2)(2+p^2)}{2N(1+2p^2)}}$
9:7 and 3:1	$\frac{ad}{bc}=\frac{2+7p^2+3p^4}{6-9p^2+3p^4}$	$.6745\sqrt{\frac{(1+3p^2)(1-p^2)(4-p^4)}{3Np^2(5+2p^2-4p^4)}}$
27:37 and 3:1	$\frac{ad}{bc}=\frac{14+25p^2+9p^4}{30-39p^2+9p^4}$	$.6745\sqrt{\frac{(1-p^2)(2+p^2)(10-3p^2)(7+9p^2)}{27Np^2(9+2p^2-4p^4)}}$
15:1 and 3:1	$\frac{ad}{bc}=\frac{11p^2+p^4}{4-5p^2+p^4}$	$.6745\sqrt{\frac{(1-p^2)(4-p^2)(11+p^2)}{N(11+2p^2-4p^4)}}$
63:1 and 3:1	$\frac{ad}{bc}=\frac{47p^2+p^4}{16-17p^2+p^4}$	$.6745\sqrt{\frac{(1-p^2)(16-p^2)(47+p^2)}{N(47+2p^2-4p^4)}}$
9:7 and 9:7	$\frac{ad}{bc}=\frac{20+28p^2+9p^4}{36-36p^2+9p^4}$	$.6745\sqrt{\frac{(10+9p^2)(4-p^4)}{18Np^2(3+2p^2)}}$
9:7 and 15:1	$\frac{ad}{bc}=\frac{11+34p^2+3p^4}{27-30p^2+3p^4}$	$.6745\sqrt{\frac{(1+3p^2)(9-p^2)(1-p^2)(11+p^2)}{6Np^2(13+p^2-2p^4)}}$
15:1 and 15:1	$\frac{ad}{bc}=\frac{56p^2+p^4}{16-8p^2+p^4}$	$.6745\sqrt{\frac{(4-p^2)(56+p^2)}{2N(7+2p^2)}}$

COMPARATIVE RELIABILITY OF LINKAGES CALCULATED FROM BACKCROSSES AND FROM F_2 DATA

Since the probable error of a linkage intensity calculated from a backcross will, in all cases, be less than from the same size of population in F_2, the backcross method may be used as a standard of comparison for F_2 data. In figure 1 is shown graphically the relative accuracy of backcross data compared with F_2 data. The comparison is made on the basis of the number of times as many individuals needed to establish a linkage intensity with the probable error, from F_2 data as from a backcross, the product method being used to determine the linkage from the F_2 data. It is seen that for the coupling phase F_2 data are but slightly inferior to backcross data, especially for close linkage. At 50 per cent, crossing over it would require 2.25 times as many individuals in F_2 to obtain the same reliability as from a backcross. This fact was first pointed out by HALDANE in 1919. F_2 repulsion data are much less reliable than F_2 coupling, especially for close linkages. As an illustration, a determination

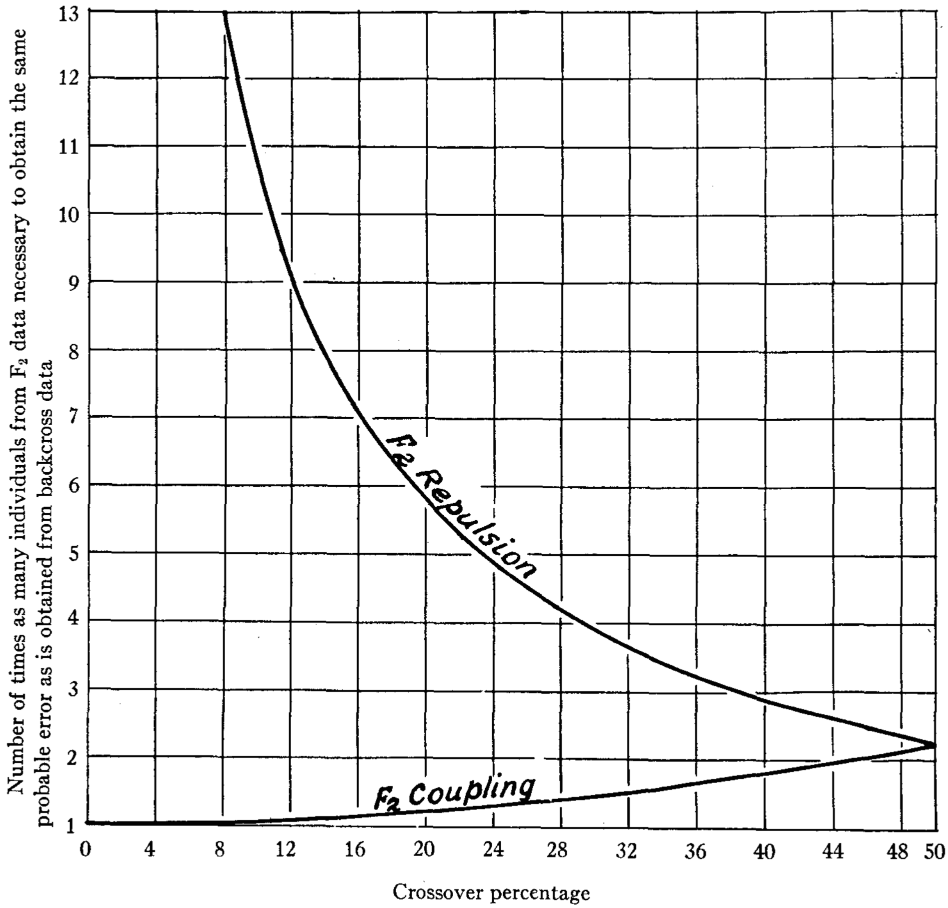

FIGURE 1.—Graph showing the relative reliability of F_2 and backcross data for determining linkage intensities.

of 10 percent crossing over based on 1,000 individuals from a backcross would be as reliable as if the same were obtained from 1,130 individuals from F_2 data from a cross in the coupling phase or 10,830 individuals from F_2 repulsion. This emphasizes again the relative reliability of F_2 data for the determination of close linkage in the coupling phase as contrasted with the repulsion phase. The repulsion phase is only as reliable as the coupling phase for 50 percent crossing over. For close linkage the repulsion phase is very inferior to coupling.

It can be concluded then that in linkage studies where backcrossing is difficult or time consuming that linkages may be determined accurately from F_2 data, the accuracy depending on whether the cross was made in the coupling or repulsion phase and whether a close or a loose linkage is found. If the factors entered in the coupling phase in a given cross the number of times as many individuals in F_2 necessary to obtain the same reliability as from a backcross would vary from 1 to 2.25. If the cross were made in the repulsion phase, the number of times as many individuals in F_2 necessary to obtain the same reliability as from a backcross would vary from 2.25 to ∞ as the linkage varied from 50 to 0 percent.

RELIABILITY OF LINKAGE DETERMINATIONS WHEN COMPLEMENTARY OR DUPLICATE FACTORS ARE INVOLVED

Having established the relative reliability of F_2 data as compared with backcross data, it would seem logical to determine next the relative reliability of linkage determinations when three or four factor pairs are concerned as compared with only two factor pairs.

In figure 2 is shown graphically the relative reliability of linkage intensity calculations from 9:7 and 3:1, 27:37 and 3:1, 15:1 and 3:1, or 63:1 and 3:1 ratios, using the probable error of the same crossover percentage for two 3:1 ratios as the standard of comparison. The efficiency of these more complex ratios is given in terms of the number of times as many individuals necessary to establish a linkage with the same probable error as would be obtained from two 3:1 ratios. It will be seen that linkage intensities can be determined fairly accurately when 9:7 and 3:1 or 27:37 and 3:1 ratios are involved if the cross is made in the coupling phase. If the cross is one of repulsion the probable errors are much larger than the probable errors for F_2 repulsion in the case of two 3:1 ratios, which are themselves very inferior to F_2 coupling as shown in figure 1. It is clear, as shown by the graph, that linkage intensities calculated from 9:7 and 3:1 or 27:37 and 3:1 ratios and based on F_2 repulsion data are very unreliable for close linkages unless very large numbers are obtained. This emphasizes the decided superiority of the coupling over the repulsion phase for three or four factor problems when complementary factors are involved.

Linkage intensities calculated from F_2 data when duplicate factors determine one of the character pairs are also less reliable when the repulsion phase is used than with the coupling phase although the difference between the coupling and repulsion phases is less here than when two 3:1 ratios are involved. Larger populations are needed to obtain the same reliability when duplicate factors are involved as when only two 3:1 ratios are involved.

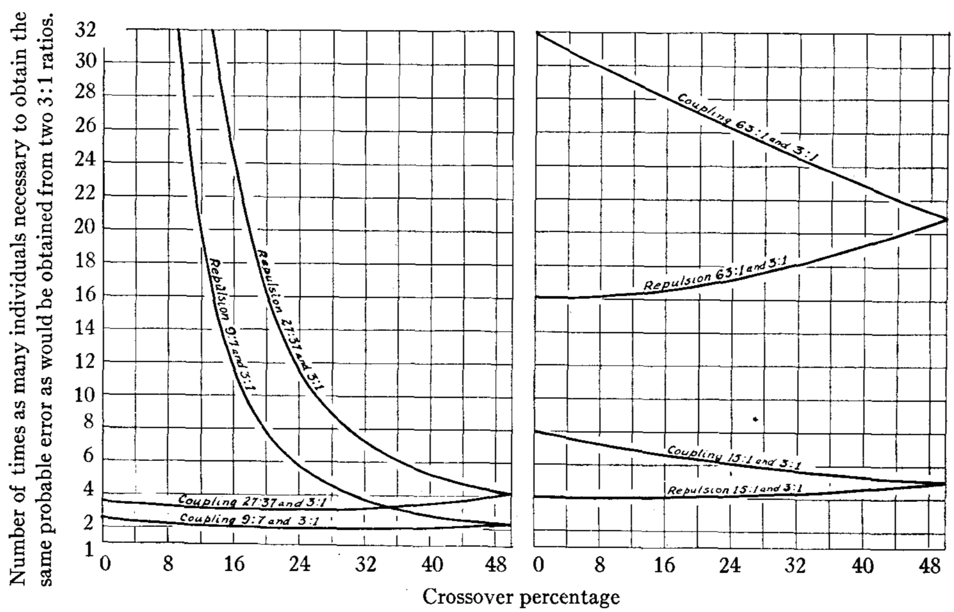

FIGURE 2.—Graph showing relative reliability of linkage intensities calculated from ratios in which more than 2 factors are involved as compared with determinations from two 3:1 ratios.

TABLES FOR CALCULATING LINKAGE INTENSITIES

In calculating linkage intensities from the formulae given in table 1, it is necessary to reduce the proper equation to the form of a quadratic, solve for p^2 and extract the square root. While this is not difficult, it may be rather laborious under some conditions and will always be time consuming. Fortunately the product method lends itself readily to the calculation of tables from which linkage intensities and their probable errors may be determined with the minimum of effort. FISHER and BALMUKAND (1928) gave a small table of this kind. More extended tables have been calculated and are presented here. The method of using these tables will be illustrated.

TABLE 2

Constants to facilitate the calculation of linkage intensities, by the product method, when each of two character pairs is determined by a single factor difference. Constants are given also to be used in obtaining probable errors for 3:1 and 3:1 ratios.

CROSSOVER VALUE	RATIO OF PRODUCTS		FACTOR TO BE DIVIDED BY $\sqrt{N}$ TO OBTAIN PROBABLE ERROR		
	ad/bc REPULSION	*bc/ad* COUPLING	F_2 REPULSION	F_2 COUPLING	BACKCROSS
.005	.00005000	.00003361	.6745	.04771	.04757
.010	.00020005	.0001356	.6744	.06751	.06711
.015	.0004503	.0003076	.6743	.08271	.08199
.020	.0008008	.0005516	.6742	.09555	.09443
.025	.001252	.0008692	.6740	.1069	.1053
.030	.001804	.001262	.6737	.1171	.1151
.035	.002458	.001733	.6735	.1266	.1240
.040	.003213	.002283	.6731	.1354	.1322
.045	.004070	.002914	.6728	.1436	.1398
.050	.005031	.003629	.6724	.1515	.1470
.055	.006096	.004429	.6719	.1590	.1538
.060	.007265	.005318	.6715	.1661	.1602
.065	.008540	.006296	.6709	.1730	.1663
.070	.009921	.007366	.6704	.1796	.1721
.075	.01141	.008531	.6698	.1860	.1777
.080	.01301	.009793	.6691	.1922	.1830
.085	.01471	.01116	.6684	.1982	.1881
.090	.01653	.01262	.6677	.2040	.1930
.095	.01846	.01419	.6670	.2097	.1978
.100	.02051	.01586	.6662	.2153	.2024
.105	.02267	.01765	.6653	.2207	.2068
.110	.02495	.01954	.6644	.2260	.2110
.115	.02734	.02156	.6635	.2312	.2152
.120	.02986	.02369	.6625	.2363	.2192
.125	.03250	.02594	.6616	.2413	.2231
.130	.03527	.02832	.6605	.2463	.2268
.135	.03816	.03083	.6594	.2511	.2305
.140	.04118	.03347	.6583	.2558	.2340
.145	.04434	.03624	.6572	.2605	.2375
.150	.04763	.03915	.6560	.2651	.2408
.155	.05105	.04220	.6548	.2697	.2441
.160	.05462	.04540	.6535	.2741	.2473
.165	.05832	.04875	.6522	.2785	.2503
.170	.06218	.05225	.6509	.2829	.2534
.175	.06618	.05591	.6495	.2872	.2563
.180	.07033	.05973	.6482	.2914	.2591
.185	.07464	.06371	.6467	.2956	.2619
.190	.07911	.06787	.6453	.2998	.2646
.195	.08374	.07220	.6438	.3039	.2672
.200	.08854	.07671	.6422	.3079	.2698
.205	.09351	.08140	.6407	.3119	.2723
.210	.09865	.08628	.6391	.3159	.2747

TABLE 2 (*continued*)

CROSSOVER VALUE	RATIO OF PRODUCTS		FACTOR TO BE DIVIDED BY $\sqrt{N}$ TO OBTAIN PROBABLE ERROR		
	ad/bc REPULSION	*bc/ad* COUPLING	F_2 REPULSION	F_2 COUPLING	BACKCROSS
.215	.1040	.09136	.6375	.3198	.2771
.220	.1095	.09663	.6358	.3237	.2794
.225	.1152	.1021	.6341	.3276	.2817
.230	.1211	.1078	.6324	.3314	.2839
.235	.1272	.1137	.6307	.3352	.2860
.240	.1334	.1198	.6289	.3390	.2881
.245	.1400	.1262	.6271	.3427	.2901
.250	.1467	.1328	.6253	.3464	.2921
.255	.1536	.1396	.6234	.3501	.2940
.260	.1608	.1467	.6215	.3537	.2959
.265	.1682	.1540	.6196	.3573	.2977
.270	.1758	.1616	.6177	.3609	.2995
.275	.1837	.1695	.6157	.3645	.3012
.280	.1919	.1777	.6137	.3680	.3029
.285	.2003	.1861	.6117	.3716	.3045
.290	.2089	.1948	.6097	.3750	.3061
.295	.2179	.2038	.6076	.3785	.3076
.300	.2271	.2132	.6055	.3820	.3091
.305	.2367	.2228	.6034	.3854	.3105
.310	.2465	.2328	.6012	.3888	.3119
.315	.2567	.2432	.5991	.3922	.3133
.320	.2672	.2538	.5969	.3955	.3146
.325	.2780	.2649	.5947	.3989	.3159
.330	.2892	.2763	.5925	.4022	.3172
.335	.3008	.2881	.5902	.4055	.3184
.340	.3127	.3003	.5879	.4088	.3195
.345	.3250	.3128	.5857	.4121	.3206
.350	.3377	.3259	.5833	.4153	.3217
.355	.3508	.3393	.5810	.4185	.3228
.360	.3643	.3532	.5787	.4218	.3238
.365	.3783	.3675	.5763	.4250	.3247
.370	.3927	.3823	.5739	.4281	.3256
.375	.4076	.3977	.5715	.4313	.3265
.380	.4230	.4135	.5691	.4345	.3274
.385	.4389	.4298	.5666	.4376	.3282
.390	.4553	.4467	.5641	.4407	.3290
.395	.4723	.4641	.5617	.4438	.3297
.400	.4898	.4821	.5592	.4469	.3304
.405	.5079	.5007	.5566	.4500	.3311
.410	.5266	.5199	.5541	.4531	.3317
.415	.5460	.5398	.5516	.4561	.3323
.420	.5660	.5603	.5490	.4592	.3329
.425	.5867	.5815	.5464	.4622	.3334
.430	.6081	.6034	.5438	.4652	.3339
.435	.6302	.6260	.5412	.4682	.3344

TABLE 2 (*continued*)

CROSSOVER VALUE	RATIO OF PRODUCTS		FACTOR TO BE DIVIDED BY $\sqrt{N}$ TO OBTAIN PROBABLE ERROR		
	ad/bc REPULSION	*bc/ad* COUPLING	F_2 REPULSION	F_2 COUPLING	BACKCROSS
.440	.6531	.6494	.5386	.4712	.3348
.445	.6768	.6735	.5359	.4741	.3352
.450	.7013	.6985	.5333	.4771	.3356
.455	.7266	.7243	.5306	.4800	.3359
.460	.7529	.7510	.5279	.4829	.3362
.465	.7801	.7786	.5252	.4859	.3364
.470	.8082	.8071	.5225	.4888	.3366
.475	.8374	.8366	.5197	.4916	.3368
.480	.8676	.8671	.5170	.4945	.3370
.485	.8990	.8986	.5142	.4974	.3371
.490	.9314	.9313	.5115	.5002	.3372
.495	.9651	.9651	.5087	.5031	.3372
.500	1.0000	1.0000	.5059	.5059	.3372
.505	1.0362	1.0362	.5031	.5087	.3372
.510	1.0738	1.0736	.5002	.5115	.3372
.515	1.1128	1.1124	.4974	.5142	.3371
.520	1.1533	1.1526	.4945	.5170	.3370
.525	1.1953	1.1942	.4916	.5197	.3368
.530	1.2390	1.2373	.4888	.5225	.3366
.535	1.2844	1.2819	.4859	.5252	.3364
.540	1.3316	1.3282	.4829	.5279	.3362
.545	1.3806	1.3762	.4800	.5306	.3359
.550	1.4317	1.4260	.4771	.5333	.3356
.555	1.4847	1.4776	.4741	.5359	.3352
.560	1.5400	1.5312	.4712	.5386	.3348
.565	1.5975	1.5868	.4682	.5412	.3344
.570	1.6574	1.6446	.4652	.5438	.3339
.575	1.7198	1.7045	.4622	.5464	.3334
.580	1.7848	1.7668	.4592	.5490	.3329
.585	1.8526	1.8316	.4561	.5516	.3323
.590	1.9234	1.8989	.4531	.5541	.3317
.595	1.9972	1.9689	.4500	.5566	.3311
.600	2.0742	2.0417	.4469	.5592	.3304

By calculating tables giving the values of the ratio of products ad/bc or bc/ad for different values of p, the labor of determining linkage intensities can be materially reduced. Such aids for the rapid determination of linkages from 3:1 and 3:1, 9:7 and 3:1, 27:37 and 3:1 and 15:1 and 3:1 ratios are given in tables 2, 3, 4, and 5. In these tables the ratio of products for repulsion is given as ad/bc and for coupling as bc/ad. This is done because it is more convenient to keep the ratios less than one for all crossover percentages less than 50.

TABLE 3

Constants to facilitate the calculation of linkage intensities and their probable errors when one character pair is determined by two complementary factor differences and the other character pair by a single factor difference — 9:7 and 3:1 ratios.

CROSSOVER VALUE	RATIO OF PRODUCTS		FACTOR TO BE DIVIDED BY $\sqrt{N}$ TO OBTAIN PROBABLE ERROR	
	ad/bc REPULSION	*bc/ad* COUPLING	F_2 REPULSION	F_2 COUPLING
.01	.3335	.005185	34.83	.1088
.02	.3340	.01075	17.42	.1520
.03	.3348	.01670	11.62	.1843
.04	.3360	.02306	8.719	.2107
.05	.3375	.02983	6.980	.2336
.06	.3394	.03703	5.822	.2538
.07	.3416	.04467	4.995	.2722
.08	.3441	.05277	4.376	.2891
.09	.3470	.06133	3.895	.3048
.10	.3503	.07038	3.510	.3195
.11	.3539	.07992	3.196	.3335
.12	.3579	.08997	2.935	.3468
.13	.3623	.1005	2.714	.3596
.14	.3671	.1117	2.525	.3719
.15	.3723	.1233	2.362	.3838
.16	.3779	.1356	2.219	.3953
.17	.3840	.1484	2.093	.4066
.18	.3904	.1618	1.982	.4175
.19	.3974	.1759	1.882	.4283
.20	.4048	.1905	1.792	.4390
.21	.4127	.2059	1.711	.4494
.22	.4211	.2219	1.638	.4597
.23	.4300	.2385	1.571	.4700
.24	.4394	.2559	1.510	.4801
.25	.4495	.2740	1.454	.4902
.26	.4601	.2928	1.402	.5003
.27	.4713	.3123	1.354	.5103
.28	.4832	.3327	1.309	.5203
.29	.4958	.3538	1.268	.5303
.30	.5090	.3756	1.229	.5404
.31	.5231	.3983	1.193	.5505
.32	.5378	.4219	1.159	.5606
.33	.5534	.4462	1.127	.5709
.34	.5699	.4715	1.097	.5812
.35	.5873	.4975	1.069	.5916
.36	.6056	.5245	1.042	.6022
.37	.6249	.5524	1.016	.6129
.38	.6453	.5812	.9922	.6237
.39	.6668	.6108	.9692	.6348
.40	.6894	.6415	.9474	.6459
.41	.7134	.6730	.9265	.6573
.42	.7387	.7055	.9066	.6689

TABLE 3 (*continued*)

CROSSOVER VALUE	RATIO OF PRODUCTS		FACTOR TO BE DIVIDED BY $\sqrt{N}$ TO OBTAIN PROBABLE ERROR	
	ad/bc REPULSION	*bc/ad* COUPLING	F_2 REPULSION	F_2 COUPLING
.43	.7653	.7390	.8875	.6808
.44	.7935	.7734	.8692	.6930
.45	.8233	.8087	.8516	.7054
.46	.8548	.8451	.8347	.7181
.47	.8880	.8824	.8184	.7312
.48	.9232	.9206	.8026	.7446
.49	.9605	.9598	.7874	.7585
.50	1.0000	1.0000	.7727	.7727
.51	1.0418	1.0412	.7585	.7874
.52	1.0862	1.0831	.7446	.8026
.53	1.1333	1.1261	.7312	.8184
.54	1.1833	1.1699	.7181	.8347
.55	1.2365	1.2147	.7054	.8516
.56	1.2930	1.2602	.6930	.8692
.57	1.3532	1.3066	.6808	.8875
.58	1.4174	1.3538	.6689	.9066
.59	1.4859	1.4018	.6573	.9265
.60	1.5589	1.4505	.6459	.9474

The values of the ratio of products ad/bc and bc/ad corresponding to crossover values of .01, .02, .03, .04, ... , .50 for two 3:1 ratios (table 2) are taken from the paper by FISHER and BALMUKAND (1928) by permission of the JOURNAL OF GENETICS.

The determination of linkage intensities by the product method then simply resolves itself into calculating ad/bc or bc/ad from the four observed F_2 phenotypic classes and finding the crossover percentage by interpolation in the tables appended. (See tables 2, 3, 4 and 5). Since most determinations of linkage intensities are made from two 3:1 ratios, the interval between the values of p in table 2 has been made .005. In tables 3, 4 and 5 the interval is .01. In the linkage calculations from 9:7 and 3:1, 27:37 and 3:1 or 15:1 and 3:1 ratios, where complementary or duplicate factors are concerned in the production of one of the character pairs, a linkage is assumed between but one of the factor pairs responsible for the 9:7, 27:37 or 15:1 ratio and the single factor pair responsible for the 3:1 ratio.

The probable errors for the four types of ratios dealt with in tables 2, 3, 4 and 5 are obtained by dividing the probable error factor corresponding to the calculated crossover value by the square root of the number of individuals ($\sqrt{N}$).

TABLE 4

Constants to facilitate the calculation of linkage intensities and their probable errors when one character pair is determined by three complementary factor differences and the other character pair by a single factor difference — 27:37 and 3:1 ratios.

CROSSOVER VALUE	RATIO OF PRODUCTS		FACTOR TO BE DIVIDED BY $\sqrt{N}$ TO OBTAIN PROBABLE ERROR	
	ad/bc REPULSION	*bc/ad* COUPLING	F_2 REPULSION	F_2 COUPLING
.01	.4668	.008939	51.20	.1265
.02	.4672	.01826	25.60	.1780
.03	.4680	.02797	17.07	.2170
.04	.4690	.03808	12.80	.2495
.05	.4703	.04860	10.24	.2779
.06	.4719	.05952	8.537	.3033
.07	.4738	.07086	7.318	.3266
.08	.4760	.08263	6.405	.3482
.09	.4785	.09482	5.694	.3685
.10	.4813	.1075	5.126	.3876
.11	.4844	.1205	4.661	.4059
.12	.4878	.1341	4.274	.4233
.13	.4916	.1480	3.946	.4402
.14	.4957	.1625	3.666	.4564
.15	.5001	.1774	3.422	.4722
.16	.5049	.1928	3.209	.4876
.17	.5100	.2086	3.022	.5027
.18	.5155	.2249	2.855	.5175
.19	.5214	.2418	2.706	.5320
.20	.5277	.2591	2.571	.5464
.21	.5343	.2768	2.450	.5606
.22	.5414	.2951	2.339	.5747
.23	.5489	.3139	2.238	.5887
.24	.5568	.3332	2.146	.6026
.25	.5652	.3530	2.061	.6165
.26	.5740	.3732	1.982	.6304
.27	.5834	.3940	1.909	.6443
.28	.5932	.4153	1.842	.6583
.29	.6036	.4371	1.779	.6723
.30	.6145	.4594	1.720	.6864
.31	.6260	.4821	1.665	.7007
.32	.6381	.5054	1.613	.7151
.33	.6508	.5292	1.564	.7296
.34	.6642	.5534	1.519	.7443
.35	.6782	.5781	1.475	.7592
.36	.6930	.6033	1.435	.7744
.37	.7085	.6290	1.396	.7898
.38	.7248	.6551	1.359	.8054
.39	.7419	.6817	1.324	.8214
.40	.7599	.7087	1.291	.8380

TABLE 4 (*continued*)

CROSSOVER VALUE	RATIO OF PRODUCTS		FACTOR TO BE DIVIDED BY $\sqrt{N}$ TO OBTAIN PROBABLE ERROR	
	ad/bc REPULSION	*bc/ad* COUPLING	F_2 REPULSION	F_2 COUPLING
.41	.7788	.7361	1.259	.8544
.42	.7987	.7640	1.228	.8714
.43	.8196	.7922	1.199	.8889
.44	.8416	.8209	1.172	.9068
.45	.8647	.8499	1.145	.9252
.46	.8891	.8793	1.119	.9441
.47	.9147	.9090	1.094	.9636
.48	.9416	.9390	1.071	.9836
.49	.9701	.9694	1.048	1.004
.50	1.0000	1.0000	1.026	1.026
.51	1.0316	1.0309	1.004	1.048
.52	1.0649	1.0620	.9836	1.071
.53	1.1001	1.0933	.9636	1.094
.54	1.1373	1.1248	.9441	1.119
.55	1.1766	1.1565	.9252	1.145
.56	1.2182	1.1882	.9068	1.172
.57	1.2623	1.2201	.8889	1.199
.58	1.3089	1.2520	.8714	1.228
.59	1.3585	1.2840	.8544	1.259
.60	1.4111	1.3160	.8380	1.291

As an example of the use of the product method, assume a cross AABB × aabb would give this ratio in F_2:

AB(a)	Ab(b)	aB(c)	ab(d)	Total(N)
1757	118	119	506	2500

Since the cross is one of coupling, bc/ad = 118×119/1757×506 = 0.01579. By interpolation in table 2 we find the crossover value to be 0.0998 which would mean 9.98 percent crossing over. From the same table we find by interpolation that the proper factor for the probable error is 0.2151 which, divided by $\sqrt{2500}$ gives a probable error of the above linkage intensity of 0.0043 or 0.43 percent.

If two or three factor pairs had been involved in determining one of the character pairs from a cross made in the coupling phase, we would have calculated bc/ad as before and looked up the crossover percentage from table 3, 4 or 5, depending on whether two or three complementary or duplicate factor pairs were concerned.

TABLE 5

Constants to facilitate the calculation of linkage intensities and their probable errors when one character pair is determined by two duplicate factor differences and the other character pair by a single factor difference — 15:1 and 3:1 ratios.

CROSSOVER VALUE	RATIO OF PRODUCTS		FACTOR TO BE DIVIDED BY $\sqrt{N}$ TO OBTAIN PROBABLE ERROR	
	ad/bc REPULSION	*bc/ad* COUPLING	F_2 REPULSION	F_2 COUPLING
.01	.0002750	.005118	1.349	.1895
.02	.001101	.01048	1.349	.2664
.03	.002478	.01609	1.348	.3243
.04	.004409	.02197	1.348	.3723
.05	.006898	.02811	1.347	.4139
.06	.009948	.03455	1.346	.4509
.07	.01356	.04127	1.345	.4846
.08	.01775	.04831	1.343	.5154
.09	.02252	.05567	1.342	.5440
.10	.02787	.06336	1.340	.5708
.11	.03382	.07140	1.338	.5959
.12	.04038	.07981	1.336	.6197
.13	.04755	.08860	1.334	.6423
.14	.05535	.09778	1.331	.6638
.15	.06379	.1074	1.329	.6843
.16	.07288	.1174	1.326	.7041
.17	.08265	.1279	1.323	.7230
.18	.09311	.1389	1.320	.7412
.19	.1043	.1504	1.317	.7588
.20	.1162	.1624	1.313	.7758
.21	.1288	.1749	1.309	.7923
.22	.1422	.1881	1.306	.8083
.23	.1564	.2018	1.302	.8238
.24	.1714	.2162	1.297	.8389
.25	.1873	.2312	1.293	.8535
.26	.2041	.2470	1.289	.8678
.27	.2217	.2635	1.284	.8817
.28	.2403	.2808	1.279	.8953
.29	.2599	.2989	1.274	.9085
.30	.2805	.3180	1.269	.9215
.31	.3022	.3379	1.264	.9342
.32	.3250	.3588	1.258	.9465
.33	.3489	.3808	1.252	.9586
.34	.3740	.4039	1.247	.9705
.35	.4004	.4281	1.241	.9821
.36	.4282	.4536	1.235	.9935
.37	.4573	.4804	1.228	1.005
.38	.4878	.5086	1.222	1.016
.39	.5199	.5383	1.215	1.026
.40	.5536	.5696	1.208	1.037

TABLE 5 (*continued*)

CROSSOVER VALUE	RATIO OF PRODUCTS		FACTOR TO BE DIVIDED BY $\sqrt{N}$ TO OBTAIN PROBABLE ERROR	
	ad/bc REPULSION	*bc/ad* COUPLING	F_2 REPULSION	F_2 COUPLING
.41	.5889	.6027	1.201	1.047
.42	.6261	.6375	1.194	1.057
.43	.6651	.6743	1.187	1.067
.44	.7060	.7132	1.179	1.077
.45	.7491	.7543	1.172	1.086
.46	.7943	.7979	1.164	1.096
.47	.8419	.8440	1.156	1.105
.48	.8919	.8929	1.148	1.114
.49	.9446	.9448	1.140	1.123
.50	1.0000	1.0000	1.131	1.131
.51	1.0584	1.0587	1.123	1.140
.52	1.1200	1.1212	1.114	1.148
.53	1.1849	1.1878	1.105	1.156
.54	1.2534	1.2590	1.096	1.164
.55	1.3257	1.3350	1.086	1.172
.56	1.4022	1.4164	1.077	1.179
.57	1.4830	1.5037	1.067	1.187
.58	1.5686	1.5973	1.057	1.194
.59	1.6593	1.6980	1.047	1.201
.60	1.7555	1.8065	1.037	1.208

SUMMARY

1. The development of probable error formulae for various methods of calculating linkage intensities from F_2 data has provided a means of determining the relative efficiency of these methods.

2. The product method (equivalent to the coefficient of association method) seems to be the best general method available. The ratio of products ad/bc or bc/ad is calculated from the observed data and the linkage intensity obtained by interpolation in the tables appended.

3. F_2 coupling data are but slightly less reliable than backcross data while F_2 repulsion data are less reliable, the relative reliability depending on the closeness of the linkage. In linkage studies where backcrossing is difficult or time consuming, linkage intensities may be accurately determined from F_2 data by growing slightly larger populations. This is particularly true of the coupling phase.

4. Linkage intensities calculated from F_2 coupling data are more reliable as measured by the probable error, than from F_2 repulsion data, the relative reliability depending on the closeness of the linkage. For close linkage, coupling data are much more reliable than repulsion data.

5. Linkages between 2 factor pairs may be determined quite accurately from 9:7 and 3:1 or 27:37 and 3:1 ratios if the cross is made in the coupling phase. In the repulsion phase much larger numbers are needed to obtain the same reliability. When duplicate factors are involved (15:1 and 3:1 or 63:1 and 3:1 ratios) much larger populations are needed to obtain the same reliability as would be secured from two 3:1 ratios.

LITERATURE CITED

ALBERTS, HUGO W., 1926 A method for calculating linkage values. Genetics **11**: 235–248.

BABCOCK, E. B., and CLAUSEN, R. E., 1927 Genetics in relation to agriculture. New York: McGraw-Hill Book Co.

BATESON, W., and PUNNETT, R. C., 1911 On gametic series involving reduplication of certain terms. J. Genet. **1**: 293–302

BRIDGES, C. V., 1914 The chromosome hypothesis of linkage applied to cases in sweet peas and primula. Amer. Nat. **48**: 524–534.

BRUNSON, A. N., 1924 The inheritance of a lethal pale green seedling character in maize. Mem. Cornell Agric. Expt. Sta. 72.

CASTLE, W. E., 1916 Tables of linkage intensities. Amer. Nat. **50**: 575–576.

COLLINS, G. N., 1912 Gametic coupling as a cause of correlation. Amer. Nat. **46**: 569–590.

1924 Measurement of linkage values. J. Agric. Res. **27**: 881–891.

EMERSON, R. A., 1916 The calculation of linkage intensities. Amer. Nat. **50**: 411–420.

FISHER, R. A., 1928 Statistical methods for research workers. Second edition. Edinburgh: Oliver and Boyd.

FISHER, R. A., and BALMUKAND, BHAI, 1928 The estimation of linkage from the offspring of selfed heterozygotes. J. Genet. **20**: 79–92.

HALDANE, J. B. S., 1919 The probable errors of calculated linkage values and the most accurate method of determining gametic from certain zygotic series. J. Genet. **8**: 291–297

HOR, K. S., 1924 Interrelation of genetic factors in barley. Genetics **9**: 151–180.

KAPPERT, H., 1927 Uber die Auswertung dihybrider Spaltungsreihen bei Koppelungsstudien. Z. indukt. Abstamm.-u. Vererblehre. **44**: 303–314.

OWEN, F. V., 1928 Calculation of linkage intensities by product moment correlation. Genetics **13**: 80–110.

WELLENSIEK, S. J., 1927 Methods for calculating the actual gametic F_2 series from a given zygotic series. Genetica **9**: 329–340.

WOODWORTH, C. M., 1923 Calculation of linkage intensities where duplicate factors are concerned. Genetics **8**: 106–115.

YULE, G. U., 1900 On the association of attributes in statistics. Phil. Trans. Roy. Soc. A. **194**: 257–319.