From the Serological Laboratory of the Office of the Chief Medical Examiner of the City of New York, and the Division of Immunohematology of the Jewish Hospital of Brooklyn, Brooklyn, New York

SEROLOGY, GENETICS AND NOMENCLATURE OF THE M-N-S TYPES

by

Alexander S. Wiener, M. D.

As has been shown in a previous paper, [1] the serology and genetics of the M-N-S types are analogous to that of the Rh-Hr types. To date, however, an entirely satisfactory nomenclature for the M-N-S genotypes has not been devised. Is is customary to refer to the 4 known alleles by the symbols MS, Ms, NS, and Ns, respectively. While these symbols are adequate if one accepts the theory of linked gene couplets, they are not if one believes, like the present author, that heredity of these types is by a series of 4 allelic genes. As been pointed out previously, [2] the 4 alleles determine corresponding complex agglutinogens, each characterized by a minimum of two blood factors as follows, **MS, Ms, NS,** and **Ns.** Therefore, it is important not to use the terms blood factors and agglutinogens interchangeably when discussing the M-N-S system, as well as the Rh-Hr types and A-B-O groups. One could designate the 4 M-N-S alleles as L^{MS}, L^{Ms}, L^{NS}, and L^{Ns}, respectively, but these symbols are cumbersome. In the present paper simplified, rational designations for the 4 allelic genes are suggested, as well as simple formulae for estimating the gene frequencies from the distribution of the phenotypes.

When tests for the three M-N types are carried out, the genes are designated by the symbols M and N, or L^M and L^N, respectively, the symbol L being selected in honor of Landsteiner. This suggests the use of the same base symbol when naming the 4 allelic genes of the M-N-S types. It is therefore proposed to give the 4 allelic M-N-S genes the simple designations, L^S, L, l^S and l, respectively. These symbols are easy to learn and to work with, since all one has to remember is that large L stands for the presence of blood factor **M** in the agglutinogen and small l for the presence of blood factor **N** in the agglutinogen. Similarly, the superscript large S in the gene symbol stands for the presence of blood factor **S** in the corresponding agglutinogen, while the omission of any superscript indicates the presence of blood factor **s** in the agglutinogen. The four allelic genes and their corresponding agglutinogens may then be represented as shown in table 1.

When preparing a table of the M-N-S phenotypes and genotypes, it is necessary to take into account the various possible levels of the serological tests. If tests are carried out with serums anti-**M** and anti-**N** alone, only three phenotypes are differentiated; if tests are carried out also with anti-**S** serum, then the number of phenotypes is increased to 6; while if tests are carried out also with anti-**s** serum, the number of phenotypes is increased to 9. In table 2 the possible phenotypes and corresponding genotypes are listed at these three levels. The simplicity and clarity of the designations of the genotypes using

Table 1 - The Four Allelic M-N-S Genes

Designations of genes under multiple allele theory	Designations of " chromosomes " under theory; of linked couplets	Corresponding agglutinogens	
		Designations	Blood factors present
L^S	MS	M.S	**M** and **S**
L	Ms	M.s	**M** and **s**
l^S	NS	N.S	**N** and **S**
l	Ns	N.s	**N** and **s**

Table 2 - Nomenclature of the M-N-S types

3 M-N Types				6 M-N-S Types			9 M-N-S Types		
Pheno-types	Reactions with		Geno-types	Pheno-types	Reactions with anti-**S**	Geno-types	Pheno-types	Reactions with anti-**s**	Geno-types
	Anti-**M**	Anti-**N**							
M	+	—	MM or $L^M L^M$	M.S	+	$L^S L^S$ or $L^S L$	M.SS	—	$L^S L^S$
				M.ss	—	LL	M.Ss	+	$L^S L$
							M.ss	+	LL
N	—	+	NN or $L^N L^N$	N.S	+	$l^S l^S$ or $l^S l$	N.SS	—	$l^S l^S$
				N.ss	—	ll	N.Ss	+	$l^S l$
							N.ss	+	ll
MN	+	+	MN or $L^M L^N$	MN.S	+	$L^S l^S$, $L^S l$ or $L l^S$	MN.SS	—	$L^S l^S$
				MN.ss	—	Ll	MN.Ss	+	$L^S l$ or $L l^S$
							MN.ss	+	Ll

the new symbols for the 4 allelic genes is apparent on inspection. It will be seen that the designations for the M-N-S phenotypes in common use are retained without change. This is as it should be, since the names of the phenotypes should be quite independent of the genetic theory. For example, there is nothing in the names of the blood groups O, A, B, and AB to indicate whether they are inherited by triple allelic genes, or by linked gene couplets as was maintained at one time by a large body of workers.

The ease with which the new gene symbols can be manipulated is best demonstrated by illustrating their use in a sample mating. Suppose that in tests carried out with the three antisera, anti-**M**, anti-**N**, and anti-**S**, the types of the parents proved to be MN.S and N.ss, respectively. As shown in figure 1, three genetic possibilities then exist, and in each case only two different M-N-S types can occur among the children. If any child
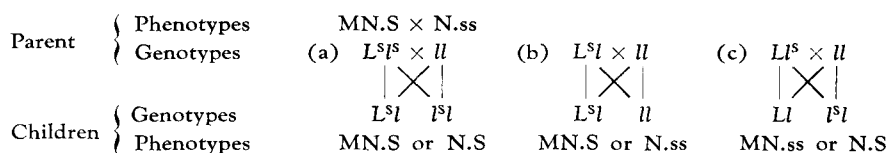
Parent	Phenotypes	MN.S × N.ss		
	Genotypes	(a) $L^s l^s \times ll$	(b) $L^s l \times ll$	(c) $Ll^s \times ll$
		$\times$	$\times$	$\times$
Children	Genotypes	$L^s l$ $l^s l$	$L^s l$ ll	Ll $l^s l$
	Phenotypes	MN.S or N.S	MN.S or N.ss	MN.ss or N.S

Figure 1

is encountered with blood of type N and lacking blood factor **S,** then the genotypes of the parents must be $L^s l$ and ll (mating b, figure 1), respectively, so that all the type N children in such a family would lack the S factor while all the MN children would have the factor. The other possibilities can be recognized in the same way from the types in the children.

Gene frequencies

A method of calculating the gene frequencies for the M-N-S types, using the method of maximum likelihood has recently been proposed by Boyd. [3-4] While this method gives values with minimum standard errors, it is cumbersome and too difficult to use, except for skilled mathematicians. Therefore, the method is impracticable, especially considering that the available reports on the distribution of the M-N-S types are all based on relatively small series, and since in addition a certain number of technical errors in determining the types are inevitable, such exact and difficult mathematical techniques seem to be out of place. At any rate, calculating gene frequencies to the seventh significant figure, based on tests on a few hundred or thousand individuals, as has been done by certain other workers in the field, is incongruous to say the least.

The method to be proposed here develops a principle presented in an earlier paper. [5] The principle is best explained by citing the example given in that paper. In a population, the M-N types have the distribution, M = 0.2681, N = 0.2754, and MN = 0.4565. count, the frequencies of the two genes are obtained as follows:

$$m = M + \frac{MN}{2} = 0.4964$$

and $n = 0.5036$

The square root formulae, on the other hand, give the following estimates:

$$m = \sqrt{M} = 0.5178$$
$$n = \sqrt{N} = 0.5247$$

The principle of adjusting the square root formulae is to make the sum equal to unity with a pair of simple simultaneous equations.

$$n - m = \sqrt{N} - \sqrt{M} = .0069$$
$$n + m = 1$$

From these simple equations are obtained the adjusted estimates,

$$m = 0.4966 \text{ and } n = 0.5034$$

316

As can be seen, these estimates agree very closely with the values obtained by direct count.

Suppose we now consider the case where tests are carried out with all four antisera, anti-**M**, anti-**N**, anti-**S**, and anti-**s**, so as to differentiate all nine M-N-S types. The frequencies of « major genes » M, N, S, and s, are first determined by direct count as follows:

$$m = M + \frac{MN}{2}; n = N + \frac{MN}{2}; S = SS + \frac{Ss}{2}; \text{ and } s = ss + \frac{Ss}{2}$$

The next step is to make a first estimate of the frequencies of the genes L and l by means of the square root formulae, as follows.

$$L = \sqrt{M.ss}$$
$$l = \sqrt{N.ss}$$

These estimates are then adjusted so as to make the sum $L + l$ equal to the value of s obtained by direct count.

This is done with the aid of the following simple pair of equations:

$$L - l = \sqrt{M.ss} - \sqrt{N.ss}$$
$$L + l = s$$

From these adjusted estimates of the frequencies of the genes L and l the frequencies of the other two genes are obtained by simple subtraction.

$$L^S = m - L$$
$$l^S = n - l,$$

where m and n are the values for the frequencies of genes M and N obtained by direct count.

Where the tests are limited to the three antiserum, anti-**M**, anti-**N**, and anti-**S**, so that only six phenotypes are differentiated, the principle of estimating gene frequencies is the same, except that here the frequencies of the « major genes » S and s cannot be determined by direct count. Instead, therefore, the frequency of s must be estimated by the sguare root formula below.

$$s = \sqrt{ss} = \sqrt{M.ss + N.ss + MN.ss}$$

The ease with which the formulae for the M-N-S gene frequencies can be applied is best demonstrated by an actual example. Among 230 Bengalis in Dacca Boyd found the following distribution of the M-N-S types:

Type	Number	Percent
M.ss	35	15.217
M.S	44	19.130
MN.ss	47	20.435
MN.S	62	26.957
N.ss	21	9.130
N.S	21	9.130
Totals	230	100.

317

By direct count we find:

$m = 0.58043$ and $n = 0.41957$

Moreover, $s = \sqrt{ss} = \sqrt{0.44783} = 0.66920$

The first estimates of L and l are as follows:

$$L = \sqrt{M.ss} = \sqrt{0.15217} = 0.39009$$

$$l = \sqrt{N.ss} = \sqrt{0.09130} = 0.30216$$

The improved estimates are then simply derived as follows:

$L + l = 0.66920$

$L - l = 0.08793$

Therefore, $L = 0.37857$ and $l = 0.29063$

and $L^S = m - L = 0.58043 - 0.37857 = 0.20187$

$l^S = n - l = 0.12894$

Since these gene frequencies are based on tests on only 230 individuals, the last two significant figures can now be dropped from the estimates of the gene frequencies.

In table 3 the estimates of the frequencies using the simplified formulae are compared with those obtained by Boyd using the maximum likelihood method, as well as Mourant's method. As will be seen the differences in the estimates affect the third significant figure only whereas the size of the series and the standard deviation indicate that the estimates can at most be correct only to the second significant figure. Therefore, the additional effort required by the maximum likelihood method is not justified, since it does not increase the accuracy of the estimates.

Table 3 - M-N-S Gene frequencies Among 230 Bengalis in Dacca
East Pakistan (data of Boyd)

Genes	Calculated gene frequencies (percent)				Standard deviation
	Author's method	Maximum likelihood method [1]	Mourant's method [1]		
			Original	Alternative	
L^S	20.2	20.0	20.3	20.2	1.8
L	37.9	38.0	37.8	37.8	2.2
l^S	12.9	13.0	12.8	12.9	1.4
l	29.1	28.9	29.1	29.1	2.0

[1] Calculated by Boyd.

Summary

The symbols for the 4 allelic M-N-S genes in common use, namely, *MS*, *Ms*, *NS*, and *Ns*, are satisfactory only if one subscribes to the theory of linked gene couplets, but not for those who, like the present author, believe that the theory of multiple alleles is valid. Therefore, it is proposed to designate the 4 allelic genes by the symbols L^S, L l^S and l, respectively. These symbols are easy to learn and to use, if one bears in mind that L stands for the presence of blood factor **M** in the corresponding agglutinogen, and l for the presence of blood factor **N**. Similarly the letter S in the superscript stands for the presence of blood factor **S** in the corresponding agglutinogen, while the absence of any superscript denotes the presence of blood factor **s**. It is unnecessary to change the names of the phenotypes. This is fortunate, since duplication of nomenclature is avoided, such as exists unnecessarily in the case of the Rh-Hr types. The purpose of a phenotype name is to identify and not to provide a full description, or to indicate the genetic mechanism. For this reason, the phenotype designations of the M-N-S types should be the same, regardless of the genetic theory to which one subscribes.

The practicability of the new gene symbols for the M-N-S types is demonstrated by applying them to a sample mating. A simple method of estimating the frequencies of the 4 allelic M-N-S genes is also described. These formulae are applied to Boyd's data on the distribution of the M-N-S types among 230 Bengalis. The results obtained do not differ significantly from those obtained by the more complex maximum likelihood method. Since the maximum likelihood method does not improve the accuracy of the estimates, the additional effort and skill which is required are wasted. The simple formulae proposed here yield results which fulfill all the requirements of anthropologists and geneticists.

References

1. WIENER, A. S. (1952). Heredity of the M-N-S Types. Theoretico-Statistical Considerations. *Amer. Jour. Human. Genet.*, 4, 37.
2. WIENER, A. S., and WEXLER, I. B. (1952). The Mosaic Structure of Red Blood Cell Agglutinogens. *Bacteriol. Rev.*, 16, 69.
3. BOYD, W. C. (1953). Estimation of Gene Frequencies from M-N-S Data. *Sicence*, 118, 756.
4. BOYD, W. C. (1954). Maximum Likelihood Method for Estimation of Gene Frequencies from MNS Data. *Amer. Jour. Human Genet.*, 6, I.
5. WIENER, A. S., LEDERER, M., and POLAYES, S. H. (1930). Stud.es in Isohemagglutination. IV. On the Chances of Proving Non-paternity; with Special Reference to Blood Groups. *Jour. Immunol.*, 19, 259.

RIASSUNTO

I simboli generalmente adoperati per i quattro geni allelomorfi M-N-S, i. e., MS, Ms, NS e Ns, sono accettabili solo a condizione di condividere la teoria dell'appaiamento dei geni; ma ciò non è per quanti, come l'A., pensano che si debba accettare la teoria degli alleli multipli. Perciò si propone di designare i quattro geni allelomorfi con i simboli L^s, L, l^s e l. Questi simboli si imparano e si adoperano facilmente, quando si ritenga che L indica la presenza del fattore M nell'agglutinogeno corrispondente e che l indica la presenza del fattore N. In modo analogo la lettera S adoperata come esponente indica la presenza del fattore S nell'agglutinogeno corrispondente, mentre l'assenza di qualsiasi esponente significa la presenza del fattore sanguigno s. È inutile modificare i nomi dei fenotipi. Si evita, in tal modo, felicemente, il raddoppiamento della nomenclatura che esiste inutilmente nel caso dei tipi Ah-Hr. Il compito di un nome relativo al fenotipo è d'identificazione e non di fornire una descrizione completa o di indicare il meccanismo genetico. Di conseguenza, le designazioni dei fenotipi dei tipi M-N-S dovranno essere gli stessi, qualunque sia la teoria genetica che si accetta.

La praticità dei nuovi simboli dei geni relativi ai tipi M-N-S viene dimostrata con l'applicazione ad un incrocio esemplificativo. Viene inoltre descritto un metodo semplice per valutare le frequenze dei quattro geni allelomorfi M-N-S. Queste formule vengono applicate ai dati di Boyd relativi alla distribuzione dei tipi M-N-S in un gruppo di 230 bengalesi. I risultati ottenuti non differiscono in modo significativo da quelli ottenuti mediante il metodo più complesso della massima probabilità. Poichè questo metodo non migliora la precisione delle valutazioni, lo sforzo e l'esperienza richiesti vanno sprecati. Le semplici formule che qui vengono proposte, conducono a dei risultati di tale natura da soddisfare le esigenze degli antropologi e dei genetisti.

RÉSUMÉ

Les symboles généralement usités pour les quatre gènes allélomorphes M-N-S, i. e., MS, Ms, NS et Ns, ne sauraient satisfaire qu'à la condition d'adhérer à la théorie de la liaison des gènes par paires; mais il n'en est pas ainsi pour ceux qui, comme l'auteur, pensent que c'est la théorie des allèles multiples qui est fondée. C'est pourquoi l'on propose de désigner les quatre gènes allélomorphes par les symboles L^s, L, l^s, et l. Ces symboles sont faciles à apprendre et à employer, si l'on veut retenir que L est indicatif de la présence du facteur M dans l'agglutinogène correspondant et que l indique la présence du facteur N. D'une manière analogue, la lettre S comme exposant, indique la présence du facteur S dans l'agglutinogène correspondant, alors que l'absence de tout exposant signifie la présence du facteur sanguin s. Il est inutile de modifier les noms des phénotypes. On évite ainsi, heureusement, la duplication de nomenclature telle qu'il en existe inutilement dans le cas des types Rh-Hr. Le rôle d'un nom de phénotype est d'identifier, et non pas de fournir une description complète ou d'indiquer le mécanisme génétique. Par conséquent, les désignations des phénotypes des types M-N-S devront être les mêmes, quelque soit la théorie génétique qu'on accepte.

On démontre la praticabilité des nouveaux symboles de gènes relatifs aux types M-N-S en les appliquant à un croisement explicatif. On décrit en outre une méthode simple pour évaluer les fréquences des quatre gènes allélomorphes M-N-S. On applique ces formules aux données de Boyd relatives à la distribution des types M-N-S parmi 230 Bengalis. Les résultats obtenus ne diffèrent pas d'une manière significative de ceux obtenus au moyen de la méthode plus complexe de la probabilité maxima. Etant donné que cette méthode n'améliore pas la précision des évaluations, l'effort et l'expérience supplémentaires nécessaires sont gaspillés. Les formules simples qu'on propose ici, donnent des résultats de nature à satisfaire les exigences des anthropologistes et génétistes.

ZUSAMMENFASSUNG

Die allgemein gebräuchlichen Symbole für die 4 M-N-S-Allelogene, nämlich *MS, Ms, NS* und *Ns* befriedigen nur dann, wenn man der Theorie der gepaarten Gene anhängt, nicht aber wenn man, wie der Autor, an die Gültigkeit der Theorie der multiplen Allelen glaubt. Es wird daher vorgeschlagen, die 4 Allelogene mit den Symbolen L^s, L, l^s und l zu bezeichnen. Diese Symbole sind leicht zu erlernen und zu behalten, wenn man sich merkt, dass L die Anwesenheit des Blutfaktors M im entsprechenden Agglutinogen und l die Anwesenheit des Blutfaktors N bedeutet. In ähnlicher Weise bedeutet das hochgestellte S die Anwesenheit des Blutfaktors S im entsprechenden Agglutinogen, während das Fehlen eines Index die Anwesenheit des Blutfaktors s bedeutet. Es fügt sich gut, dass es nicht notwendig ist, die Bezeichnungen der Phänotypen zu ändern, wodurch eine Duplikation der Nomenklatur vermieden wird, wie sie überflüssigerweise im Falle der Rh-Hr-Typen vorkommt. Die Bezeichnungen der Phänotypen sollen der Identifizierung dienen, und nicht eine volle Beschreibung oder einen Hinweis auf den genetischen Mechanismus geben. Aus diesem Grunde sollten die Phänotypenbezeichnungen der M-N-S-Typen die gleichen sein, ohne Rücksicht auf die Genentheorie, der man anhängt.

Die Brauchbarkeit der neuen Genensymbole für die M-N-S-Typen wird an Hand eines Paarungsbeispieles erläutert. Eine einfache Methode zur Abschätzung der Häufigkeiten der 4 M-N-S-Allelotypen wird ebenfalls beschrieben. Diese Formeln werden auf Boyd's Arbeit über die Verteilung der M-N-S-Typen unter 230 Bengalen angewandt. Die Ergebnisse weisen keine wesentliche Verschiedenheit gegenüber den mittels der komplizierteren Maximalwahrscheinlichkeitsmethode erhaltenen auf. Da die Maximalwahrscheinlichkeitsmethode die Genauigkeit der Schätzung nicht erhöht, wären die zusätzlich erforderlichen Bemühungen und Kenntnisse vergeudet. Die einfache, hier vorgeschlagene, Methode ergibt Resultate, die allen Forderungen der Anthropologen und der Genetiker gerecht werden.