Functional diversity of natural IgM

Caroline Brissac1,3, Alberto Nobrega2, Jorge Carneiro1,3 and John Stewart3,4

1Unité d’Immunobiologie, Institut Pasteur, 25 rue du Dr Roux, 75015 Paris, France
2Depto de Imunologia, Instituto de Microbiologia, CCS-Bloco I, Universidade Federal do Rio de Janeiro-UFRJ, Cidade Universitaria, Ilha do Fundação, CEP 21944, Rio de Janeiro, Brasil
3Estudos Avançados, Instituto Gulbenkian de Ciência, 2781 Oeiras, Portugal
4COSTECH, Département Technologie et Sciences de l’Homme, Université de Technologie de Compiègne, BP 60649, 60206 Compiègne Cedex, France

Keywords: antibody repertoire, diversity, natural antibody

Abstract
This paper proposes a method for the quantitative characterization of repertoire diversity of an unknown mixture of antibodies on the basis of its reactivity profile in the quantitative immunoblot (QIB). The QIB is calibrated by measuring the reactivity profiles of supernatants of known ‘diversity’ (i.e. known numbers of B cell clones). We define a quantitative ‘index of variability’ (IV) which decreases regularly as the diversity increases and the profiles tend towards a common ‘convergence profile’. The calibration procedure is consolidated by a mathematical model based on the Poisson distribution; this theoretical model accounts correctly for the observed convergence behavior. On the basis of this calibration curve, it is possible to estimate the diversity of an unknown antibody mixture from a measure of its IV. We conclude that the functional diversity of natural serum IgM in mice can be estimated at ~16,000 clones.

Introduction
Experiments in antigen-free animals have shown that in the complete absence of any stimulation by external antigens, the immune system nevertheless produces substantial titers of natural IgM antibodies in the serum (1,2). The functioning of the undisturbed immune system raises a number of basic questions, such as the origin, functional repertoire and physiological role of these natural antibodies, with important clinical implications concerning the development of autoimmune diseases as well as their prevention and treatment using, precisely, natural antibodies from healthy donors (i.v. Ig).

In view of the fact that these questions pertain to the most fundamental physiology of the immune system, it is somewhat surprising that they are currently largely unresolved issues in immunology. Several attempts have been made to build network models of natural antibody production (3–5). However, progress in this field has been hampered by the lack of experimental data with which to test these models of underlying mechanisms and in particular by the lack of an appropriate technique for characterizing the repertoire of serum antibodies unrelated to any particular antigen. Nobrega et al. (6,7) have recently developed a technique known as the quantitative immunoblot (QIB), which makes it possible to measure the reactions of a set of antibodies with a large panel of proteins. The aim of this paper is to establish a quantitative characterization of natural antibody repertoires based on data from this QIB, as a first step towards answering these unresolved questions.

The method we have adopted is to produce a large collection of supernatants from known numbers of unspecifically stimulated small spleen B cells from mice. The Ig in these supernatants thus represent a set of random samples of the unselected spleen repertoire. The reactivity profiles of these samples were then measured in the QIB. In this paper, we use the term ‘diversity’ to designate the number of different clones that contribute to a given sample; the term ‘variability’ designates the quantitative difference between profiles from samples of a given diversity. The results show that the variability among samples decreases as their diversity increases. Moreover, the profiles of samples of increasing diversity resemble more and more closely the profile obtained with a very high number of clones. In other words, it appears that as the diversity increases, the profiles of random samples converge towards a unique profile that we designate by the term ‘convergence profile’ (CP) and that this CP corresponds to the overall mean of all the reactivity profiles.

These results prompted us to establish a quantitative measure of the discrepancy between any given profile and the CP; we designate this measure, defined mathematically below,
by the term ‘index of variability’ (IV). The results show empirically that based on the sets of random samples, this index decreases regularly as the diversity increases. This regular relationship thus serves to ‘calibrate’ the index in terms of the diversity of the sample tested. On the basis of this calibration curve, it is then possible to estimate the diversity of an unknown antibody mixture from a measure of its IV. In particular, given certain assumptions, it becomes possible to estimate the diversity of a set of natural serum antibodies.

**Methods**

**Animals**

Specific pathogen-free (SPF) C57Bl/6j mice were obtained from IFFA/Credo (St Germain sur Arbristle, France) and bled by retro-orbital puncture.

**Cell cultures**

Splenocytes were prepared by gentle disruption of the spleen into RPMI 1640 supplemented with 5% FCS, 2 mM L-glutamine, 10 mM HEPES, 1000 U/100 mg/ml penicillin/streptomycin (complete medium), at 4°C. Cells were washed in complete medium, counted and suspended at the appropriate concentrations. Purification of high-density B lymphocytes (>1.077 g/cm³) was done on Percoll gradients (8) and controlled by flow cytometry. B cell cultures were set up as described in (8) in flat-bottom 96-well plates (Corning, Brumath, France), in 0.2 ml of complete medium containing 3×10⁵ thymocytes from 5-week-old Lewis rats and 50 mg/ml of lipopolysaccharide (LPS; Difco, OSI, Maurepas, France). Culture supernatants were harvested at 9 days.

Variable numbers of B cells (80, 400, 2000 and 10,000) were put in culture and stimulated by LPS. It was estimated by limiting dilution (9) that one-third of the cells respond to stimulation by proliferating and producing IgM. We thus obtained supernatants resulting from the contributions of 27, 133, 666 and 3333 cells.

Since resting B cells show little or no proliferation (10,11), we assume that the clones in this compartment each comprise a single cell. Moreover, small peripheral B cells are restricted in numbers [10⁸ in SPF mice (12)] compared to the numbers of cells produced daily by the bone marrow (10⁷/day) from a very broad potential repertoire (well over 10⁹). It is therefore reasonable to assume that there are only very rare repeats in the variable regions expressed in the supernatants (certainly <10% and probably <1%), in other words, for all practical purposes it is safe to assume that each resting B cell is indeed a different clone.

By mixing supernatants of 3333 clones we obtained samples of higher diversity by a factor of 5, 25 and 100, in other words 16,665, 83,325 and 333,300 clones.

It is important for the theoretical basis of our approach to ensure that the contributions of the clones secreting into the supernatant are homogeneous. In previously published work, Nobrega et al. (9) have examined this point with care. Briefly, they have shown: (i) that the frequency of response to LPS is kept constant in the range of the numbers of cells put into culture; (ii) that the number of IgM-secreting cells (IgM-SC) increases linearly with the number of cells put into culture; and (iii) that the titer of IgM in the supernatants increases linearly with the number of cells in the culture wells. Under these conditions, it is reasonable to suppose that the contributions of the clones secreting into the supernatant are indeed homogeneous and that mixtures of supernatants do have a ‘diversity’ equal to the sum of individual diversities.

**QIB**

Nobrega et al. designed a technique, based on the Western Blot and called the QIB, that reveals global immunoreactivity of any antibody mixture against a large sample of proteins used as reactivity probes. The technique provides a reactivity profile for each tested sample. The profiles are cut into sections corresponding to the antigenic bands revealed by electrophoresis. By staining proteins with colloidal gold one can correct for the migration distortions occurring during electrophoresis which permits us to compare levels of reactivity of different antibody mixtures against a given section. A profile thus consists of a list of reactivity values, one for each section. A description of the overall technique has been published (6,7).

**Immunoreactivity testing with QIB**

For all experiments described here the blotted proteic extract was homologous liver extract. We tested the random samples as follows.

We made four gels on each of which we tested 18 supernatants of the same number of clones (27, 133, 666 and 3333 respectively) as well as a mixture of supernatants representing 3×10⁵ clones. In each of these four experiments the global IgM concentration was set so that the concentration per clone would be 10 ng/ml, except for the mixture of supernatants for which the concentration per clone was 0.1 ng/ml. Made in parallel, those four experiments were rescaled together, which allows us to compare reactivities of different samples against the antigenic sections common to the four membranes.

We also tested on a single membrane three groups of five supernatants from the same number of clones (3333, 16,665 and 83,325) clones, as well as the serum of the mouse whose cells were put in culture. In that experiment the total amount of IgM was the same for all samples and was set to 15 µg/ml, which means that the concentration per clone decreases when diversity increases.

**Measure of the distance to the CP**

We define the ‘distance’ between two profiles—in particular between a given profile and the CP—as a normalized SD. If the mean of the population—or the CP—is measured independently, this formula can be applied to a sample of only one point. Having measured the difference between a sample profile and the CP separately for each section, we then averaged this measurement over all sections. This overall measure was then normalized by the values of the CP, which lead us to express an average IV, IV(average), in non-dimensional units as a percentage. The detailed procedure can be found in Appendix I; the final result is:

$$IV = \frac{\sum (x_i - cpi)^2}{\sum cpi^2}$$

where $x_i$ is the observed optical density for section $i$ and $cpi$ is the corresponding value of the CP.
The ‘CP’ can be derived either from a single random sample of very high diversity or from the mean of all the profiles obtained on a single membrane, on condition that they have a sufficient diversity and are numerous enough to ensure that the mean is representative.

The validity of this measure of the distance to the CP requires that: (i) a profile be compared only to the CP obtained on the same membrane or on different membranes made and rescaled together (they have to show reactivity against the same antigenic sections); and (ii) to avoid introducing a non-relevant source of variability, all the samples on a given membrane have to be tested at the same global IgM concentration (otherwise the distance to the CP would be impared by the influence of concentration).

This comparison to the CP is internal to a membrane (or a set of membranes run and rescaled together), but the overall IV is a robust measure which gives comparable results for data derived from independent membranes. This feature consolidates the interest of normalizing SD by the CP. The blotted extract has to be the same for all membranes brought into a single set of data because the diversity of epitopes gathered in antigenic sections has an impact on the dispersion of profiles around their mean.

The mathematical model

In order to consolidate our interpretation of variation around the CP in terms of repertoire diversity, we have built a mathematical model of the experimental behavior of convergence towards the mean profile. This model was based on the assumption that each clonal recognition event (of an antigenic section by an antibody) is rare, but the total number of Ig clones is very large. With the sole additional assumption that the recognition events are mutually independent, the total number of recognition events then necessarily follows a Poisson distribution. It is a mathematical property of this distribution that the variance is equal to the mean; hence, the normalized SD is equal to the square root of the mean and the normalized distribution that the variance is equal to the mean; hence, the property accounts for the decrease in variability as the diversity increases. In Appendix II, we show that this property holds when the IV is calculated on the basis of a limited number of sections recognized.

Moreover, as the number of clones increases the profiles of samples of a given diversity look more and more alike, and, as diversity increases, the profiles tend to resemble more and more the profile obtained with a highly diverse repertoire. As the number of clones increases, the profiles of samples of a given diversity look more and more alike, and, as diversity increases, the profiles tend to resemble more and more the profile obtained with a highly diverse repertoire.

The mathematical model of the experimental behavior of convergence towards the mean profile was based on the assumption that each recognition event is rare, but the total number of Ig clones is very large. With the sole additional assumption that the recognition events are mutually independent, the total number of recognition events then necessarily follows a Poisson distribution. It is a mathematical property of this distribution that the variance is equal to the mean; hence, the normalized SD is equal to the square root of the mean and the normalized distribution that the variance is equal to the mean; hence, the property accounts for the decrease in variability as the diversity increases. In Appendix II, we show that this property holds when the IV is calculated on the basis of a limited number of sections recognized. Moreover, as the number of clones increases the profiles of samples of a given diversity look more and more alike, and, as diversity increases, the profiles tend to resemble more and more the profile obtained with a highly diverse repertoire. It appeared also that the profile towards which profiles converge with increasing diversity is identical to their average.

The ’CP’ can be derived either from a single random sample of very high diversity or from the mean of all the profiles obtained on a single membrane, on condition that they have a sufficient diversity and are numerous enough to ensure that the mean is representative.

The validity of this measure of the distance to the CP requires that: (i) a profile be compared only to the CP obtained on the same membrane or on different membranes made and rescaled together (they have to show reactivity against the same antigenic sections); and (ii) to avoid introducing a non-relevant source of variability, all the samples on a given membrane have to be tested at the same global IgM concentration (otherwise the distance to the CP would be impaired by the influence of concentration).

This comparison to the CP is internal to a membrane (or a set of membranes run and rescaled together), but the overall IV is a robust measure which gives comparable results for data derived from independent membranes. This feature consolidates the interest of normalizing SD by the CP. The blotted extract has to be the same for all membranes brought into a single set of data because the diversity of epitopes gathered in antigenic sections has an impact on the dispersion of profiles around their mean.

Curve fitting

If equation (2) is rewritten in log–log form, we obtain:

\[ \log (IV - k_0) = -0.5 \log (N) + \log (k_1) \]  

The ‘calibration curve’ then becomes a straight line, theoretically with a slope of −0.5; in addition, it turns out that the variance of the ordinate, log(IV − k_0), is roughly constant for all values of N. This form of the curve was therefore used for fitting, which was performed with the software IGOR PRO (WaveMetrics, Lake Oswego, OR), using a least-squares procedure for optimization.

Results

 Supernatants and QIB

The qualitative results of testing 18 supernatants of 27, 133, 666 and 3333 clones plus one supernatant from 3 × 10^5 clones on the QIB have been published by Nobrega (9). For low numbers of clones not all sections are recognized and each supernatant recognizes different sets of sections. The total number of sections recognized increases with the diversity of the samples tested and, in the same way, the frequency of recognition of a given section increases with diversity.

Moreover, as the number of clones increases the profiles of samples of a given diversity look more and more alike, and, as diversity increases, the profiles tend to resemble more and more the profile obtained with a highly diverse repertoire. It appeared also that the profile towards which profiles converge with increasing diversity is identical to their average (9) and this is the profile that we have denoted by the term CP. Similar results were obtained by testing supernatants from 3333, 16,665 and 83,325 clones, as shown in Fig. 1.

The IV as a function of diversity

In Fig. 2, we show the IV (calculated from our experimental data) plotted as a function of diversity on a log–log scale as explained above. Qualitatively, the IV decreases linearly with increasing diversity, as predicted by the theory. The quantitative fit is also satisfactory: the free estimate for the slope is −0.470 ± 0.045, which is very close to the theoretical value of −0.500. The estimated value of k_1, 20.1 ± 1.3, corresponds to an average frequency of section recognition, Φ, of 1 in 400. The useful range of the IV runs from 2.2 (m = 0.2 recognition events/section) down to 0.1 (m = 100); with Φ = 2.5 × 10^{-3}, this corresponds to a range of N from 10^2 to 10^5. The estimated value of the residual error k_0 is 0.07, i.e. quite small, and practically negligible except for the smallest values of the IV at the highest values of N.

In Fig. 2, it is evident that at any given value of N, the values of the IV are rather widely scattered about the model curve. Theoretical calculations (not given here) show that this is precisely to be expected, on statistical grounds, since the IV has the form of a SD estimated on the basis of a limited ‘sample size’ (i.e. the 44 sections of the QIB). Overall, therefore, it is fair to conclude that the fit between the experimental and theoretical curves is satisfactory; this substantially consolidates our calibration procedure and allows us to use the IV as a measure of the functional diversity of a set of natural antibodies.
Fig. 1. Supernatants of random mixtures of clones of increasing diversity (five samples of each number of clone: 3333, 16,665 and 83,325) showing that the profiles (in dots) converge towards the profile of a very high number of clones (3\times10^5 clones, solid line). The plot represents the reactivity value for each profile against each protein section as separated by electrophoresis.

Fig. 2. Plot of IV (k₀ subtracted) as a function of diversity, on a log-log scale. The figure shows 18 random supernatants of 666 clones, 23 random supernatants of 3333 clones, five random supernatants of 16,665 clones and five random supernatants of 83,325 clones. Each point represents a single sample. The model curve (a straight line) is drawn with the theoretical slope of −0.5.

Table 1. Estimation of serum IgM functional diversity on various protein extracts

<table>
<thead>
<tr>
<th>Extract</th>
<th>IV</th>
<th>Functional diversity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liver</td>
<td>0.23</td>
<td>16,000 ± 8000</td>
</tr>
<tr>
<td>Brain</td>
<td>0.60</td>
<td>16,000 ± 8000</td>
</tr>
<tr>
<td>Xanthomonas campestris</td>
<td>0.57</td>
<td>14,000 ± 7000</td>
</tr>
<tr>
<td>Bacillus nitritolens</td>
<td>0.46</td>
<td>28,000 ± 14,000</td>
</tr>
</tbody>
</table>

Estimation of serum functional diversity

Having established this calibration procedure, it is possible to obtain an estimate of the functional diversity of natural serum antibodies by comparison of the distance between the serum profile and the CP with the fitted calibration curve. The IV for the single IgM serum profile is 0.23; as shown in Fig. 2, this corresponds to a functional diversity of 16\times10^3 clones. The error in this estimate has two sources: (i) error in the ‘calibration curve’ and (ii) error in the measure of the IV for the serum sample. Since only a single serum sample was included in the experiment analyzed here, it is not possible to estimate error (ii) empirically. As a conservative upper estimate of error (ii), we may take the standard deviation of the IV for random samples of given diversity, since sera from different mice of the same age, sex and strain will vary less than this if there is any homeostatic repertoire selection at all. [On this basis, we may note that the error due to (i), which is of the order of the standard error of the IV, i.e. ±5%, becomes negligible]. The SD of the IV is ±35% and correspondingly the maximal error on the estimated functional diversity is ±50%.

The estimation of IgM functional diversity has been repeated on membranes prepared from other sources of antigens, both homologous and bacterial; the results are shown in Table 1. These results are consistent, the differences being well within the range of the estimated errors, and the overall average of 16\times10^3 clones confirms the result obtained with the liver extract.

Discussion

The calculations presented in this paper lead us to conclude that the difference between the profile obtained from serum antibodies and the CP corresponds quantitatively to the difference that would be expected if the serum profile was produced by 1.6\times10^4 clones. This constitutes an estimate of the actual number of clones that contributes to the serum profile if, but only if, certain assumptions are valid. We have designated our estimate by the term ‘functional diversity’ precisely in order to signify that it does not necessarily correspond to the actual number of clones. Actually, the ‘functional repertoire’ is arguably the most relevant feature from the immunological point of view, as it is this repertoire which determines perception of the antigenic universe by the immune system.

The assumptions relating functional diversity to actual diversity are, notably, the following: (i) each clone contributes equally to the overall profile and (ii) the clones that constitute the serum profile are a representative sample of the relatively
unselected small B cells in the spleen. These assumptions are not necessarily valid. To take an extreme example, nothing in these calculations permits us to exclude the hypothesis that the serum profile results from a single multireactive clone. More realistically, there is evidence (13) that the actual serum profile results from an actively homeostatic selection process. It is therefore interesting to compare our provisional estimate of functional diversity with estimates of actual diversity obtained by other means.

Our estimate of 1.6×10^4 clones might seem rather small if the function of natural antibodies is to provide a first line of defense against pathogens. However, Cohn and Langman (14) have argued on theoretical grounds that the humoral immune system must be modular in construct and proposed the term 'protecton' to designate this minimum repertoire size; their estimate of 5×10^4 clones is not much larger than our estimate of 1.6×10^4 clones. Moreover, Hodgkin (15) has recently argued, again on theoretical grounds, that the valence of the antigens will profoundly affect the minimum number of receptors needed to anticipate the entire universe of possible epitopes. Hodgkin gives as a quantitative estimate that for monovalent binding, 10^{11} clonal receptors are needed to cover all epitopes at an affinity of 10^{-16} M, whereas only 10^5 clones are required for divalent and 10^2 clones for trivalent antigens. He concludes that the natural antibody repertoire is functionally relevant for antigens that are at least divalent and probably trivalent. Our experimental estimate of 1.6×10^4 clones is therefore not incompatible with a functionally complete repertoire and provides a useful complement to these theoretical considerations.

On purely experimental grounds there are, in the literature, two types of data from which we can estimate the number of cells contributing to serum IgM: direct counts of IgM-SC in the organs of a mouse and IgM titers detected in the serum.

Heijden et al. (16) estimated that IgM-SC are distributed in the following organs: 6.5×10^5 in spleen, 1.8×10^5 in bone marrow, 1.3×10^4 in Peyer's patches, 1.3×10^4 in small intestine and 4.7×10^5 in the mesenteric lymph nodes. The proportion of cells other than spleen and bone marrow being negligible, we shall not discuss whether peritoneal cavity cells contribute to the serum IgM or not, and we finally get an estimation of cells other than spleen and bone marrow being negligible, to the serum IgM titers detected in animals could not account for the titers of Ig found in the blood. He thus questioned the half-life of Ig. But now, having estimations of IgM half-life, we would rather question the Ig production rate and consider that a cell produces in vivo ~20–25,000 Ig molecules/s. With this value the number of PFC and the required cells to maintain measured levels of IgM are in agreement. We conclude that the number of IgM-SC may fairly be estimated as 3–4×10^5.

This raises the question of the number of IgM-SC per clone. The straight comparison of the two numbers—a functional repertoire of 1.6×10^4 and number of IgM-SC of 3–4×10^5—corresponds to 25–40 IgM-SC per clone. Since these clones are stimulated to secretion, it is likely that some proliferation also occurs. On the other hand, it has been pointed out (20) that natural IgM-SC in normal SPF mice show very limited cell proliferation. Five rounds of division would give 32 cells per clone; consequently the figure of 30 IgM-SC per clone seems quite reasonable. This point is clearly worthy of further investigation; but our provisional estimate of the functional diversity of natural IgM as 1.6×10^4 clones is quite plausible. This would suggest that although the clones contributing to natural serum IgM are probably selected, with respect to their functional diversity they nevertheless constitute a fairly representative sample of the relatively unselected small B cells in the spleen.

In conclusion, this article presents a new technique which makes it possible, for the first time, to estimate the number of clones which contribute to an unknown mixture of antibodies. The application of this method to natural IgM antibodies in the mouse gives an estimate of the ‘functional diversity’ at 1.6×10^4 clones. This estimate is consistent with indirect estimates obtained on theoretical or other experimental grounds. This article therefore fills a methodological gap and makes it possible to address some fundamental questions in immunology that have so far remained unanswered.

Appendix I: Measure of the distance to the CP

Here we define mathematically the measure of the ‘distance’ between a given profile and the CP.

Let \(|x_i|\) be the list of values of reactivity against the \(N_s\) sections that constitute a profile \(X\), and \([cp_i]\) the values for the CP. Considering a single profile and a single section of the protein profile of IV \(i = 1, 2 \ldots N_s\), we define the ‘distance’ as:

\[
s_i^2 = (x_i - cp_i)^2
\]

In order to obtain a composite measure, we sum over the whole list of sections:

\[
s_2 = \left( \frac{\sum (x_i - cp_i)^2}{\sum cp_i^2} \right)
\]

We will show in Appendix II that this measure has a simple theoretical value.

This formula gives a higher weight to the sections that show high reactivity and for which the estimated percentage of variance to CP is therefore more reliable.

The division by the sum of \(cp_i\) has the convenient effect of compensating for the influence of the amount of proteins in...
the antigenic section, since we may suppose that the CP (the mean) is proportional to the amount of antigen. Expressing the distance as a percentage of the CP also allows us to compare distances estimated on different membranes and different CPs.

The normalized SD is then the square root of the normalized variance:

$$ IV = \sqrt[5]{\frac{\sum_i (x_i - \bar{cp}_i)^2}{\sum_i cp_i^2}} $$

which is the result given for IV as equation (1) in the text.

### Appendix II: The mathematical model

We postulate that the measured optical density, $x_i$, results from the summation of individual recognition events: $x_i = n_i \delta_j$ where $n_i$ is the number of clones interacting with the antigenic section number $i$ and $\delta_j$ is the unit signal given by a single clonal recognition event.

The measure $s^2 = (x_i - \bar{cp}_i)^2$ has the form of a variance and the theoretically expected value of this variance is given by

$$ \sigma^2 = \text{var}(x) = \text{var}(n_i \delta_i) = \delta^2 \text{var}(n_i) = \delta^2 n_i $$

since, by the properties of the Poisson distribution, $\text{var}(n_i) = n_i$.

The number of interacting clones, $n_i$, can in turn be decomposed as follows: $n_i = N f_i$ where $N$ is the total number of clones and $f_i$ is the frequency of recognition of the section. It follows that $\sigma^2 = \delta^2 N f_i$ and the expected value of the sum is

$$ \sum_i (x_i - \bar{cp}_i)^2 = \delta^2 N \sum_i f_i $$

if we assume that $\delta_i$ is the same for all sections and is equal to $\delta$.

The expected value of $cp_i$ is $\delta N f_i$ and hence

$$ \sum_i cp_i^2 = \delta^2 N^2 \sum_i f_i^2 $$

Thus $\sigma^2$, the expected value of the measure

$$ \sigma^2 = \frac{\sum_i (x_i - \bar{cp})^2}{\sum_i cp_i^2} $$

is given by:

$$ \sigma^2 = \frac{\delta^2 N \sum_i f_i}{\delta^2 N^2 \sum_i f_i^2} = \frac{1}{N} \frac{\sum_i f_i}{\sum_i f_i^2} $$

Note that $\Sigma f_i / \Sigma a_i$ is a weighted mean of $f_i$ and therefore that the term $\sum f_i / \Sigma f_i$ is simply the average value of $f_i$, the frequency with which section $i$ is recognized by a random IgM clone, weighted so as to privilege the higher values of $f_i$ which do contribute more to the IV as calculated. In what follows we shall designate this average frequency/section by the symbol $\Phi$. It follows that $\sigma$, the theoretically expected value of the IV, is given by:

$$ \sigma = 1/\sqrt(\Phi N) $$

The term $N \Phi$ is of course equal to $m$, the expected mean number of recognition events for a typical section in the QIB.
Functional diversity of natural IgM


