The Journal of Immunology

The Inference of Antigen Selection on Ig Genes

Izidore S. Lossos, Robert Tibshirani, Balasubramanian Narasimhan and Ronald Levy

J Immunol 2000; 165:5122-5126; ;
doi: 10.4049/jimmunol.165.9.5122
http://www.jimmunol.org/content/165/9/5122

The Journal of Immunology is published twice each month by
The American Association of Immunologists, Inc.,
1451 Rockville Pike, Suite 650, Rockville, MD 20852

Print ISSN: 0022-1767 Online ISSN: 1550-6606.

The Inference of Antigen Selection on Ig Genes[1]

Izidore S. Lossos,* Robert Tibshirani,† Balasubramanian Narasimhan,† and Ronald Levy[2]*

Analysis of somatic mutations in V regions of Ig genes is important for understanding various biological processes. It is customary to estimate Ag selection on Ig genes by assessment of replacement (R) as opposed to silent (S) mutations in the complementary-determining regions and S as opposed to R mutations in the framework regions. In the past such an evaluation was performed using a binomial distribution model equation, which is inappropriate for Ig genes in which mutations have four different distribution possibilities (R and S mutations in the complementary-determining region and/or framework regions of the gene). In the present work, we propose a multinomial distribution model for assessment of Ag selection. Side-by-side application of multinomial and binomial models on 86 previously established Ig sequences disclosed 8 discrepancies, leading to opposite statistical conclusions about Ag selection. We suggest the use of the multinomial model for all future analysis of Ag selection. ***The Journal of Immunology,*** **2000, 165: 5122–5126.**

Functional Ig genes are created by an ordered process of gene rearrangement. The large diversity of the primary Ab repertoire, which is independent of prior exposure to Ag, is achieved by combinatorial permutation of the Ig heavy and light chain V, D, and J segments and by addition or deletion of short coding sequences at the VD and DJ joints in heavy chains and VJ joints in the light chains. Following Ag encounter, the affinity of the Ab for the Ag increases during a process of affinity maturation (1). Affinity maturation results from a combination of somatic hypermutation of the rearranged V segments and Ag selection of mutants with improved binding properties. This process leads to preferential accumulation of replacement (R)[3] as opposed to silent (S) mutations in the complementary-determining regions (CDRs), which form the Ag binding sites. Concomitantly, S as opposed to R mutations, tend to cluster in the framework regions (FRs), which are required to maintain structural integrity. Initial estimates of Ag selection assumed a random pattern of R and S mutations and assumed they would localize to a region proportional to the relative size of the CDR and FR (2). Therefore, in the case of Ig heavy chain V region, mutations would localize three times more frequently to the FRs than to the CDRs and a CDR:FR ratio >0.3 would indicate Ag selection (3). Subsequently, Shlomchik et al. (4) proposed the use of a binomial distribution model for assessment of Ag selection. However, this method failed to consider the intrinsic properties of the CDR and the FR. The codon compositions of the CDR and the FR have mutational biases, since the CDRs generally consist of codons which are more susceptible to R mutations than those in the FRs. To account for the inherent susceptibility of the CDR and FR to R mutations, Chang and Casali (5) calculated the relative tendencies of the V regions of individual Ig germline genes to accumulate R mutations (Rf), and used these Rf values to estimate the expected frequency of R and S mutations, in particular CDR and FR, for a given total number of mutations. They used a binomial distribution model proposed by Shlomchik et al. (4) to determine the probability that a particular number of R mutations occurred by chance. This model, by definition, is applicable to variables that have only two distribution possibilities, whereas the total Ig mutations have four different distribution possibilities (R and S mutations in CDR and/or FR of the gene), thus requiring application of a multinomial distribution model (6). Moreover, previous methods failed to account for all the statistical possibilities to obtain a certain observed number of R mutations. Their equation consists of a single binomial probability whereas the correct version should consist of a sum of all the binomial probabilities, which include the observed value.

In the present work, we propose a new method for estimation of Ag selection pressure on Ig genes that corrects the pitfalls mentioned above and apply this method to previously published Ig gene sequences.

Materials and Methods

Multinomial distribution model for estimation of excess or scarcity of R mutations in the Ig gene

The probability that an excess or scarcity of R mutations in V_H CDR or FR occurred by chance was calculated by a multinomial distribution model (6). The total number of mutations in each V_H gene is denoted by $n = r_1 + s_1 + r_2 + s_2$, in which r_1 and r_2 are R mutations in the FR and CDR, respectively, and correspondingly, s_1 and s_2 are S mutations in the FR and CDR. The theoretical probabilities for r_1, s_1, r_2, and s_2 mutations are denoted by p_1, q_1, p_2, and q_2, respectively. These probabilities were calculated using the following equations: $p_1 = Rf_{FR} \times Lr_{FR}$; $q_1 = (1- Rf_{FR}) \times Lr_{FR}$; $p_2 = Rf_{CDR} \times (1- Lr_{FR})$; and $q_2 = (1- Rf_{CDR}) \times (1- Lr_{FR})$ in which Lr_{FR} is a relative size of the FR, and Rf_{FR} and Rf_{CDR} are the inherent susceptibility to R mutations of the FRs and CDRs, respectively. Rf_{FR} and Rf_{CDR} were calculated for each of the identified human germline genes and were based on the chance of the occurrence in each codon of an amino acid replacement given any single nucleotide change not resulting in a termination codon.

The probability of observing r_1 or fewer R mutations in FRs is given by the multinomial tail probability:

$$P(R_1 \le r_1) = \sum_{k=0 \ldots r_1, k+S_1+R_2+S_2=n} \binom{n}{k, S_1, R_2, S_2} p_1^k q_1^{S_1} p_2^{R_2} q_2^{S_2}.$$

*Department of Medicine, Division of Oncology, and Department of Health Research and Policy, and †Department of Statistics, Stanford University Medical Center, Stanford, CA 94305

Received for publication December 16, 1999. Accepted for publication August 9, 2000.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] This study was supported by Grants CA33399 and CA34233 from the U.S. Public Health Service-National Institutes of Health. I.S.L. is a Harold Dobbs Oncology Fellow. R.L. is an American Cancer Society Clinical Research Professor.

[2] Address correspondence and reprint requests to Dr. Ronald Levy, Division of Oncology CCSR 1126, Stanford University School of Medicine, Stanford, CA 94305-5306. E-mail address: levy@leland.stanford.edu

[3] Abbreviations used in this paper: R, replacement; S, silent; CDR, complementary-determining region; FR, framework region.

0022-1767/00/$02.00

The sum is taken over values of k ranging from 0 to r_1 and all combinations of S_1, R_2, S_2 such that $k + S_1 + R_2 + S_2 = n$. To compute the P value of an observed number r_1, it is customary to split the probability at r_1: $P = P(R_1 < r_1) + 0.5 \times P(R_1 = r_1)$. It should be noted that $P(R_1 = r_1) = P(R_1 \leq r_1) - P(R_1 \leq r_1 - 1)$, and $P(R_1 < r_1) = P(R_1 \leq r_1) - P(R_1 = r_1)$.

The probability of observing r_2 or more R mutations in CDRs is similarly computed using the following equation:

$$P(R_2 \geq r_2) = \sum_{k=r_2 \ldots n, R_1+S_1+k+S_2=n} \binom{n}{R_1, S_1, k, S_2} p_1^{R1} q_1^{S1} p_2^{k} q_2^{S2}.$$

And the P value is computed using the formula: $P(R_2 > r_2) + 0.5 \times P(R_2 = r_2)$. For both FR and CDR, we used one-sided P values.

There is an approximate method for computing P values for this problem. We calculate the expected number of R mutations in the FR: $E = p_1 \times n$. Then we compute the standardized deviation:

$$\frac{r_1 - E}{\sqrt{E}}.$$

Under the usual Poisson model, this quantity should have approximately a standard normal distribution and can be compared with a normal table. However, we found that this approximation can be quite poor, and hence do not recommend it.

Assessment of the equation

To assess the applicability of this equation and to compare the results obtained by this method to those previously reported by application of the Chang and Casali equation (5), we evaluated Ag selection pressure on 7 autoantibodies evaluated by Chang and Casali (5), 24 autoantibodies and lymphoma-derived V_H genes randomly selected from previously published articles (7–12), and 55 V_H gene sequences derived from diffuse large B cell lymphoma cases established in our laboratory (13). For this comparison, we used the Rf_{FR} and Rf_{CDR} values implied in these articles. Recalculation of these values resulted in slightly different Rf values for some of the V_H genes. The discrepancies most probably result from the use of slightly different germline sequences before the final sequence of the V_H gene locus was established and from the use of sequences in which there are polymorphic variations. For the future calculation of the Rf values, we suggest using our JAVA applet, available at http://www-stat.stanford.edu/imuunoglobin, which calculates the Rf values for imported germline sequences.

Results

A total of 86 V_H gene sequences were analyzed using the multinomial distribution model equation for the presence of Ag selection as demonstrated by the conservation of the FR sequence and/or excess of R mutations in CDR (Table I). These results were compared with the results obtained by the binomial distribution equation suggested by Chang and Casali (5). A total of eight discrepancies leading to opposite statistical conclusions were observed. These included six V_H gene sequences in which an excess of R mutations in the CDR (five sequences) or a scarcity of R mutations in the FR (one sequence) were suggested by the binomial distribution equation, but rejected by the multinomial distribution model equation. In an another two V_H gene sequences, evidence for scarcity of R mutations in the FR was obtained using the multinomial but not by the binomial distribution model equation. In the majority of the remaining V_H gene sequences, the P values obtained using the two equations differed in magnitude but did not lead to discrepant statistical conclusions. The similarity cannot be explained mathematically, but is quite fortuitous. There is no guarantee in general that the binomial formula will give a good approximation. The multinomial distribution model equation suggested an excess and/or a scarcity of R mutations in the CDR and FR, respectively, in 13 of the 14 V_H gene sequences derived from high-affinity autoantibodies. By contrast, the binomial distribution suggested Ag selection in only 11 of these sequences. One of the autoantibody sequences did not fulfill the statistical criteria for Ag selection by either of the analytical models.

Discussion

Analysis of somatic mutations in V regions of Ig genes is important for studying the evolution of the Ab response, for assessment of the molecular features of autoantibodies, and for the investigation of lymphoma pathogenesis. Analysis of mutations in V_H genes can provide insights regarding the role of Ag before or during lymphoma clonal outgrowth. In the absence of Ag-positive or -negative selective pressure on Ig V regions, a random mutational process would result in an even distribution of R and S mutations throughout the coding sequence. However, Ag-selected Abs demonstrated a higher frequency of R mutations in CDRs than in FRs, whereas preservation of a functional Ig molecule is associated with a higher frequency of S mutations and scarcity of R mutations in FRs. In the past, such an evaluation was performed using the binomial distribution model equation, as suggested by Shlomchik et al. (4) and further modified by Chang and Casali (5). The necessity for such an evaluation and its wide usage are demonstrated by the fact that the Chang and Casali article was cited 130 times (The Web of Science[SM] on the Internet). However, their formula is incorrect and application of the binomial distribution model to variables (mutations) that have more than two distribution possibilities is incorrect. It should be viewed as the application of an improper statistical method for the data analysis, similar, for example, to the use of parametric statistical methods for the analysis of the nonparametric variables. Moreover, Chang and Casali (5) considered in their equation a single binomial probability, while correct statistical analysis should consist of a sum of all the observed probabilities, as is proposed in the new method presented herein. Consequently, incorrect biological conclusions may have been reached, as indeed had happened in eight tested V_H gene sequences (Table I). Fortuitously, the magnitude of the observed difference between the two statistical methods in the present study was relatively small. However, we would argue that a proper statistical method should require the application of a multinomial distribution equation for all future estimates of Ag selection.

In the present work, we propose a new statistical method for estimation of Ag selection. It corrects the pitfalls present in the previous method while still taking into account the inherent susceptibility of the codons of the CDR and the FR to R mutations. One consideration not addressed here is the known propensity for certain positions to mutate—hot spots (14). Our equation assumes that mutations in V_H genes occur randomly, thus disregarding the possible contribution from intrinsic biases in the hypermutation mechanisms due to the presence of mutational hot spots. Since the hot spots are located in CDRs but not in FRs, the assumption that mutations in FRs occur randomly is absolutely correct. Regarding the CDRs, to consider mutational hot spots, one would need to know all the hot spots in each V_H gene sequence and their relative propensity to undergo mutations in comparison to each remaining non-hot spot codon in the sequence. Consideration of these hot spots may require a custom equation for each V_H gene sequence, thus precluding its wide applicability. Until the data required to construct the custom equation for each germline V_H gene sequence exists, our model can provide good approximation of the Ag selection on Ig genes.

In conclusion, we suggest the use of the multinomial model for all future analysis of Ag selection. The investigators should compare the tested Ig gene sequence to the most similar germline sequence, with particular attention to the presence of known polymorphic variants. The JAVA applet for computing the multinomial P values and Rf values of CDRs and FRs is available at http://www-stat.stanford.edu/immunoglobin. Usage of this applet will

Table I. *Comparison of multinomial and binomial distribution models for estimation of Ag selection on human Ig genes*

Ref.	Sample	Most Similar Germline V_H Gene	FR/CDR	Observed Mutations R	Observed Mutations S	P_M[a]	P_B[b,c]
13	2	1-69	FR	10	10	0.002	0.002
			CDR	8	3	0.121	0.082
	3	1-69	FR	16	9	0.003	0.003
			CDR	13	5	0.021	0.016
	4	1-03	FR	19	5	0.429	0.138
			CDR	7	2	0.276	0.144
	5	1-03	FR	24	19	0.008	0.006
			CDR	10	3	0.442	0.138
	6	1-f	FR	27	12	**0.076**[d]	**0.039**
			CDR	13	3	0.131	0.069
	7	1-f	FR	7	12	0.006	0.007
			CDR	2	1	0.842	0.145
	8	2-05	FR	6	5	0.009	0.010
			CDR	7	2	0.042	0.041
	9	2-05	FR	1	2	0.063	0.097
			CDR	2	0	0.141	0.187
	10	2-05	FR	15	4	0.220	0.106
			CDR	8	3	0.136	0.091
	11	2-05	FR	1	1	0.237	0.315
			CDR	1	0	0.276	0.370
	12	2-70	FR	8	9	0.030	0.028
			CDR	4	1	0.496	0.216
	13	2-70	FR	6	7	0.005	0.006
			CDR	5	3	0.265	0.168
	14	3-23	FR	24	23	<0.001	<0.001
			CDR	14	10	0.302	0.104
	15	3-23	FR	3	3	0.081	0.092
			CDR	3	0	0.126	0.141
	16	3-23	FR	17	9	0.107	0.060
			CDR	7	3	0.364	0.158
	17	3-23	FR	4	6	0.030	0.034
			CDR	3	0	0.290	0.223
	18	3-23	FR	2	6	0.001	0.001
			CDR	2	4	0.591	0.278
	19	3-23	FR	9	6	0.001	0.001
			CDR	11	4	0.006	0.006
	20	3-33	FR	3	3	0.085	0.095
			CDR	2	1	0.364	0.292
	21	3-33	FR	5	0	0.759	0.237
			CDR	1	1	0.567	0.382
	8	3-33	FR	11	8	0.020	0.016
			CDR	9	1	0.048	0.039
	22	3-33	FR	18	13	0.002	0.002
			CDR	16	2	0.008	0.006
	23	3-33	FR	3	6	0.015	0.019
			CDR	3	0	0.271	0.217
	24	3-48	FR	9	12	<0.001	<0.001
			CDR	9	5	0.140	0.088
	25	3-48	FR	6	10	<0.001	<0.001
			CDR	19	3	<0.001	<0.001
	26	3-48	FR	8	12	0.001	0.001
			CDR	4	4	0.693	0.177
	27	3-48	FR	6	7	<0.001	<0.001
			CDR	10	6	0.021	0.019
	28	3-49	FR	12	3	0.517	0.174
			CDR	4	2	0.439	0.216
	29	3-49	FR	6	5	0.158	0.123
			CDR	2	1	0.617	0.272
	30	3-30	FR	11	7	<0.001	<0.001
			CDR	15	4	0.001	0.001
	31	3-30	FR	21	12	0.015	0.010
			CDR	14	3	0.038	0.026
	32	3-11	FR	14	6	0.124	0.074
			CDR	7	3	0.229	0.133
	33	3-11	FR	4	8	<0.001	<0.001
			CDR	8	3	0.029	0.029
	34	3-07	FR	2	6	0.006	0.008
			CDR	2	1	0.486	0.299

[a] P_M value, probability calculated by multinomial distribution model that excess (for CDR) or scarcity (for FR) of mutations occurred by chance.
[b] P_B value, probability calculated by binomial distribution model that excess (for CDR) or scarcity (for FR) of mutations occurred by chance.
[c] *P* values found in the cited articles or calculated by us for those articles in which it was not provided.
[d] Values in boldface are *P* values leading to opposite statistical conclusions when binomial instead of multinomial distribution model was applied.

Table I. *Continued.*

Ref.	Sample	Most Similar Germline V_H Gene	FR/CDR	Observed Mutations R	Observed Mutations S	P_M[a]	P_B[b,c]
	35	3-07	FR	21	13	0.101	0.052
			CDR	6	4	0.795	0.117
	13	3-07	FR	4	1	0.089	0.091
			CDR	5	1	0.023	0.030
	36	3-66	FR	37	21	0.030	0.012
			CDR	18	3	**0.100**	**0.048**
	37	3-73	FR	2	2	0.023	0.033
			CDR	5	0	0.007	0.011
	38	4-34	FR	1	2	0.234	0.313
			CDR	0	0	0.722	0.557
	39	4-34	FR	3	16	<0.001	<0.001
			CDR	1	5	**0.979**	**0.041**
	40	4-34	FR	8	6	0.064	0.053
			CDR	5	1	0.201	0.145
	41	4-34	FR	4	2	0.343	0.250
			CDR	2	0	0.292	0.273
	42	4-61	FR	4	4	0.005	0.006
			CDR	5	4	0.148	0.122
	43	4-30-4	FR	10	4	0.121	0.081
			CDR	7	2	0.087	0.071
	44	4-59	FR	14	10	0.025	0.020
			CDR	9	2	0.114	0.076
	45	4-59	FR	44	21	0.711	0.078
			CDR	6	2	**0.960**	**0.011**
	46	4-39	FR	4	6	<0.001	0.001
			CDR	8	2	0.016	0.017
	47	4-39	FR	8	13	<0.001	<0.001
			CDR	7	4	0.338	0.155
	48	4-39	FR	27	18	<0.001	<0.001
			CDR	19	11	**0.090**	**0.043**
	49	4-39	FR	22	27	<0.001	<0.001
			CDR	15	8	0.400	0.107
	50	4-39	FR	14	14	0.016	0.012
			CDR	6	3	0.671	0.159
	51	4-30-1	FR	5	3	0.252	0.185
			CDR	2	1	0.490	0.298
	12	5-51	FR	12	4	0.245	0.128
			CDR	6	1	0.147	0.109
	52	5-51	FR	13	2	0.694	0.162
			CDR	4	1	0.369	0.209
	53	7-04-1	FR	13	19	<0.001	<0.001
			CDR	6	3	0.700	0.153
5	mAb57	1-69	FR	5	2	0.018	0.019
			CDR	8	1	0.002	0.002
	mAb412.67	3-21	FR	2	2	0.011	0.018
			CDR	6	0	0.002	0.003
	mAb412.66	3-23	FR	2	2	0.021	0.024
			CDR	5	0	0.006	0.009
	mAb61	4-39	FR	0	1	0.002	<0.001
			CDR	5	1	0.002	<0.001
	mAb112	1-69	FR	5	4	0.018	0.004
			CDR	6	1	0.030	0.004
	mAb13	3-74	FR	0	1	0.007	<0.001
			CDR	4	0	0.002	0.004
	mAb426.12.3	4.11	FR	2	7	0.003	0.006
			CDR	3	0	0.255	0.118
7	UPN2	4-61	FR	3	3	0.006	0.009
			CDR	7	1	0.005	0.004
	UPN3	3-33	FR	7	1	0.651	0.223
			CDR	3	0	0.222	0.177
	UPN4	3-07	FR	3	1	0.223	0.218
			CDR	3	0	0.076	0.082
	UPN6	3-74	FR	9	4	0.099	0.073
			CDR	6	2	0.117	0.076
	UPN7	1-69	FR	12	2	0.287	0.140
			CDR	7	2	0.064	0.054
	UPN8	3-07	FR	2	0	0.236	0.255
			CDR	2	1	0.141	0.169
8	1	5-51	FR	6	2	0.070	0.066
			CDR	6	1	0.021	0.025

Table I. *Continued.*

Ref.	Sample	Most Similar Germline V_H Gene	FR/CDR	Observed Mutations R	Observed Mutations S	P_M[a]	P_B[b,c]
	2	3-07	FR	10	8	0.007	<0.01
			CDR	7	4	0.223	0.130
	3	3-23	FR	18	19	0.001	<0.001
			CDR	8	7	0.715	0.126
	4	3-07	FR	2	8	<0.001	<0.001
			CDR	4	1	0.216	0.170
	5	3-30	FR	1	8	<0.001	<0.001
			CDR	5	3	0.119	0.105
	6	4-61	FR	12	15	0.007	<0.01
			CDR	2	6	**0.984**	**0.021**
	7	1-02	FR	20	9	0.304	0.12
			CDR	7	1	0.438	0.164
11		3-07	FR	2	2	0.003	0.004
			CDR	8	0	<0.001	<0.001
10	JK428H	4-34	FR	4	7	**0.031**	**0.05**
			CDR	1	1	0.809	0.2
	JK410H	4-34	FR	5	4	**0.035**	**0.05**
			CDR	6	0	0.022	0.04
	KC17H	4-34	FR	8	5	0.140	0.1
			CDR	3	2	0.519	0.2
9	71	5-51	FR	6	5	0.043	0.043
			CDR	5	0	0.088	0.085
	56	4-59	FR	5	5	0.001	0.002
			CDR	9	2	0.004	0.005
	54	3-23	FR	1	1	0.029	0.047
			CDR	4	0	0.006	0.01
12	1	3-07	FR	5	10	<0.001	<0.001
			CDR	9	1	0.021	0.01
	2	4-34	FR	14	10	0.006	0.004
			CDR	12	2	0.019	0.009
	6	5-51	FR	18	4	0.520	0.14
			CDR	7	1	0.198	0.11
	11	4-34	FR	5	1	0.057	0.05
			CDR	8	0	0.001	<0.001

allow uniform analysis of Ig sequences and prevent possible errors that may occur while calculating the Rf values.

References

1. Rajewsky, K. 1996. Clonal selection and learning in the antibody system. *Nature 381:751.*
2. Jukes, T. H., and J. L. King. 1979. Evolutionary nucleotide replacements in DNA. *Nature 281:605.*
3. Davi, F., K. Maloum, A. Michel, O. Pritsch, C. Magnac, E. Macintyre, F. Salomon-Nguyen, J. L. Binet, G. Dighiero, and H. Merle-Beral. 1996. High frequency of somatic mutations in the VH genes expressed in prolymphocytic leukemia. *Blood 88:3953.*
4. Shlomchik, M. J., A. H. Aucolin, D. S. Pisetsky, and M. G. Weigert. 1987. Structure and function of anti-DNA antibodies derived from a single autoimmune mouse. *Proc. Natl. Acad. Sci. USA 84:9150.*
5. Chang, B., and P. Casali. 1994. The CDR1 sequences of a major proportion of human germline Ig VH genes are inherently susceptible to amino acid replacement. *Immunol. Today 15:367.*
6. Hogg, R., and A. Craig. 1978. *Introduction to Mathematical Statistics.* Macmillan, New York.
7. Thiede, C., B. Alpen, A. Morgner, M. Schmidt, M. Ritter, G. Ehninger, M. Stolte, E. Bayerdorffer, and A. Neubauer. 1998. Ongoing somatic mutations and clonal expansions after cure of *Helicobacter pylori* infection in gastric mucosa-associated lymphoid tissue B-cell lymphoma [Published erratum appears in *1999 J. Clin. Oncol. 17:1092*]. *J. Clin. Oncol. 16:3822.*
8. Aarts, W. M., R. Willemze, R. J. Bende, C. J. Meijer, S. T. Pals, and C. J. van Noesel. 1998. VH gene analysis of primary cutaneous B-cell lymphomas: evidence for ongoing somatic hypermutation and isotype switching. *Blood 92:3857.*
9. Krenn, V., A. Konig, F. Hensel, C. Berek, M. M. Souto Carneiro, W. Haedicke, Y. Wang, H. Vollmers, and H. K. Muller-Hermelink. 1999. Molecular analysis of rheumatoid factor (RF)-negative B cell hybridomas from rheumatoid synovial tissue: evidence for an antigen-induced stimulation with selection of high mutated IgVH and low mutated IgVL/λ genes. *Clin. Exp. Immunol. 115:168.*
10. Mockridge, C. I., C. J. Chapman, M. B. Spellerberg, B. Sheth, T. P. Fleming, D. A. Isenberg, and F. K. Stevenson. 1998. Sequence analysis of V(4–34)-encoded antibodies from single B cells of two patients with systemic lupus erythematosus (SLE). *Clin. Exp. Immunol. 114:129.*
11. Ravirajan, C. T., M. A. Rahman, L. Papadaki, M. H. Griffiths, J. Kalsi, A. C. Martin, M. R. Ehrenstein, D. S. Latchman, and D. A. Isenberg. 1998. Genetic, structural and functional properties of an IgG DNA-binding monoclonal antibody from a lupus patient with nephritis. *Eur. J. Immunol. 28:339.*
12. Taniguchi, M., K. Oka, A. Hiasa, M. Yamaguchi, T. Ohno, K. Kita, and H. Shiku. 1998. De novo $CD5^+$ diffuse large B-cell lymphomas express VH genes with somatic mutation. *Blood 91:1145.*
13. Lossos, I., C. Okada, R. Tibshirani, R. Warnke, J. Vose, T. Greiner, and R. Levy. 2000. Molecular analysis of immunoglobulin genes in diffuse large B cell lymphomas. *Blood 95:1797.*
14. Betz, A. G., M. S. Neuberger, and C. Milstein. 1993. Discriminating intrinsic and antigen-selected mutational hotspots in immunoglobulin V genes. *Immunol. Today 14:405.*