The Journal of Immunology

Analysis of Mutational Lineage Trees from Sites of Primary and Secondary Ig Gene Diversification in Rabbits and Chickens

Ramit Mehr, Hanna Edelman, Devinder Sehgal and Rose Mage

J Immunol 2004; 172:4790-4796; ;
doi: 10.4049/jimmunol.172.8.4790
http://www.jimmunol.org/content/172/8/4790

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.

Analysis of Mutational Lineage Trees from Sites of Primary and Secondary Ig Gene Diversification in Rabbits and Chickens[1]

Ramit Mehr,[2]* Hanna Edelman,* Devinder Sehgal,† and Rose Mage‡

Lineage trees of mutated rearranged Ig V region sequences in B lymphocyte clones often serve to qualitatively illustrate claims concerning the dynamics of affinity maturation. In this study, we use a novel method for analyzing lineage tree shapes, using terms from graph theory to quantify the differences between primary and secondary diversification in rabbits and chickens. In these species, Ig gene diversification starts with rearrangement of a single (in chicken) or a few (in rabbit) V_H genes. Somatic hypermutation and gene conversion contribute to primary diversification in appendix of young rabbits or in bursa of Fabricius of embryonic and young chickens and to secondary diversification during immune responses in germinal centers (GCs). We find that, at least in rabbits, primary diversification appears to occur at a constant rate in the appendix, and the type of Ag-specific selection seen in splenic GCs is absent. This supports the view that a primary repertoire is being generated within the expanding clonally related B cells in appendix of young rabbits and emphasizes the important role that gut-associated lymphoid tissues may play in early development of mammalian immune repertoires. Additionally, the data indicate a higher rate of hypermutation in rabbit and chicken GCs, such that the balance between hypermutation and selection tends more toward mutation and less toward selection in rabbit and chicken compared with murine GCs. ***The Journal of Immunology*****, 2004, 172: 4790–4796.**

In humans and mice, primary diversification of Ig genes occurs throughout life by gene rearrangement in progenitor B cells in the bone marrow (1–4), and secondary diversification occurs in the course of affinity maturation of the cells' Ag receptors, via somatic hypermutation of receptor genes and Ag-driven selection of the resulting mutants (5–10). However, this is not the rule in all species. In rabbits and chickens, for example, Ig gene diversification starts with rearrangement of a single (in chicken) or a few (in rabbit) V_H genes. Primary diversification occurs in the appendix and other gut-associated lymphoid tissues in young rabbits, or the bursa of Fabricius in young chickens (11, 12). Secondary diversification in germinal centers (GCs)[3] involves both somatic hypermutation and gene conversion events in rabbits (13–15) and in chickens (16, 17). Variations on these mechanisms, with each mechanism receiving a different weight in different life stages, may exist in other species.

The dynamics of Ig gene diversification in humans and mice have been studied in detail, in experiments, by mathematical models of cell population dynamics (18–26), or by bioinformatical methods for analyzing the hypermutation process (27–34). In contrast, little is known about B lymphocyte population dynamics in species such as rabbits or chickens. Although GC dynamics during the humoral immune response may be similar to that in mice, an assumption that is currently under study, the population dynamics of B cells undergoing primary diversification through gene conversion are probably very different from the clonal dynamics that characterize murine or human B cell development.

Lineage trees of mutant B lymphocytes often serve to qualitatively illustrate claims concerning the dynamics of affinity maturation (35–46). The generation of trees or dendrograms visualizing lineage relationships of B cell mutants in GCs has been used to confirm the role of the GC as the location of somatic hypermutation (35–37), to identify lineage relationships between cells from independent GCs (40) or different tissues (41), and to study affinity maturation in various disease situations (42–44) or following vaccination (45, 46). However, no quantitative methods for analyzing mutational lineage trees have existed before our work.

We have taken the use of mutational lineage trees one step further, by developing a quantitative algorithm, based on terms of mathematical graph theory, for quantifying the shape properties of mutational lineage trees (47). We have shown that quantitative information about the dynamics of hypermutation and Ag-driven clonal selection during the humoral immune response is indeed contained in mutational lineage tree shapes deduced from responding clones. Differences in the hypermutation and selection processes, between tissues, experimental conditions, or disease situations, can be studied using our method (48, 49). Tree shape analysis thus provides a quantitative means of elucidating the dynamics of diversification processes based on hypermutation and gene conversion.

In the present study, we have applied our tree shape analysis method to lineage trees from Ig primary and secondary diversification processes in rabbits and chickens. These trees are, in general, more complex than the classical lineage trees published. Many more base changes are involved, in contrast to the relatively small numbers (order of 10) of mutations per sequence in data

*Faculty of Life Sciences, Bar-Ilan University, Ramat-Gan, Israel; †Molecular Immunology Group, National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi, India; and ‡Molecular Immunogenetics Section, Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892

Received for publication July 22, 2003. Accepted for publication February 4, 2004.

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] The work was supported by Israel Science Foundation Grant 759/01-1 (to R.M.).

[2] Address correspondence and reprint requests to Dr. Ramit Mehr, Faculty of Life Sciences, Building 212, Bar-Ilan University, Ramat-Gan 52900, Israel. E-mail address: mehrra@mail.biu.ac.il

[3] Abbreviations used in this paper: GC, germinal center; AvgOD2, outgoing degree averaged over split nodes; BDL, bursal duct ligation; CDR3, complementarity-determining region 3; MaxOD, maximum outgoing degree; NP, 4-hydroxy-3-nitrophenylacetyl; RootD, root degree.

0022-1767/04/$02.00

from murine or human GCs. Blocks of base changes can often, but not always, be identified as resulting from gene conversion, with the donor V gene identifiable in most cases. For small blocks, e.g., 2–3 base changes, it is often hard to ascertain whether the cause is a gene conversion event or merely point mutations in neighboring positions. In the case of primary diversification, most events can be identified as gene conversions, but examples of stepwise single base changes leading to small blocks in D gene, and occasionally in V gene-encoded portions, are found. In both primary and secondary diversification, the reported clonal trees are often larger and more complex than those reported in studies of murine and human clones, even when we consider each gene conversion as a single event. Our quantitative treatment of Ig gene lineage trees from rabbits and chickens reveals the unique characteristics of the genetic and selection processes driving the formation of the Ig repertoire in these species, as well as some differences between these two species, as follows. First, primary diversification in the rabbit appendix appears to occur at a constant rate, and is not subject to the types of Ag-driven selection observed in GC responses after immunization of rabbits and other species. Second, in both rabbits and chickens, the balance between mutation and selection tends more toward diversification, which may reflect the importance of maintaining a diverse repertoire in the adult once appendix or bursal B cell receptor diversification has diminished or ceased. Third, gene conversion in rabbits may be more efficient than in chickens.

Materials and Methods

Data

The lineage trees analyzed in this study were published previously (11–17).

Tree shape analysis

The concepts of graphical analysis of lineage trees, and the details of our measurement and analysis methods, are described previously (47). However, we list in this study the basic details essential for understanding this work.

A lineage tree is defined, graphically, as a phylogenetic tree in which the nodes (points on the tree) correspond to B cell receptor gene sequences (Fig. 1). It is important to note that nodes do not correspond to individual cells, but to sequences; a node may represent many cells as long as their Ig gene sequences are identical. The root node of the tree is the original unmutated sequence of the founder B cell(s). A node in the tree can have descendants (daughters); a daughter node represents a sequence that differs in exactly one mutation from the mother node, that is, it is one mutation further away from the root than the mother node. A lineage tree describes the diversification process of a clone of B cells at a certain moment of observation: it consists only of the Ig gene sequences of cells that were sampled at that moment and their ancestor sequences back to the root. The ancestor sequences are not necessarily sampled at the time of observation. A tree may be represented graphically (Fig. 1*A*) or by a list, in which each line gives the number of a mother node, followed by the numbers of all its daughter nodes (Fig. 1*B*).

The shape of lineage trees is quantified by simply counting a number of properties of the graph describing the tree, such as the total number of nodes, the number of leaves, lengths of paths from root to leaf, etc. For quite a few of these properties, the maximum, minimum, and average values per tree may be measured. The complete list of parameters measured is given in Table I.

We developed a computer program that reads a tree as a text list, as described above, and measures the graphical parameters. We are willing to analyze data sent to us upon request (see http://repertoire.ls.biu.ac.il/TREES for details).

Statistical analysis

Significance tests were done using Student's *t* test.

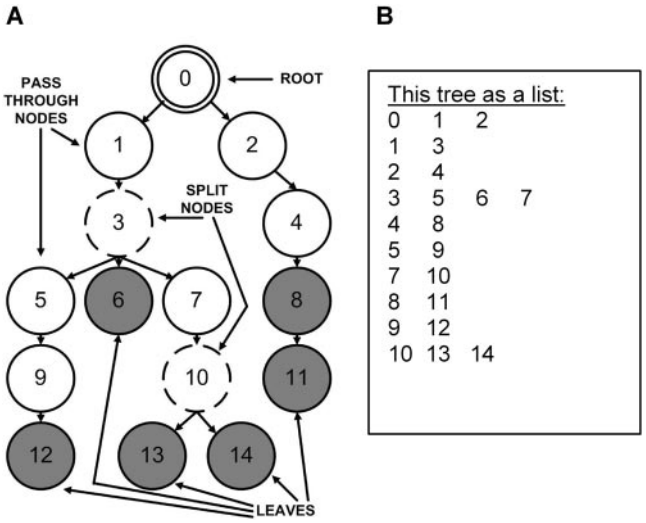

FIGURE 1. *A*, A sample lineage tree. Nodes in the tree can either be the root node (always the node numbered zero in our definition), leaves (sequences of cells that had no daughters at the time of observation, e.g., nodes 6, 11, 12, 13, and 14 here), or internal nodes, that is, nodes that are neither root nor leaves. Internal nodes can either be split nodes, those with more than one daughter (here, nodes 3 and 10); or pass-through nodes, those with exactly one daughter (nodes 1, 2, 4, 5, 7, 8, and 9 here). Thus, the total number of nodes, N, always equals the number of leaves, L, plus the number of internal nodes, IN, plus 1 (the root): $N = L + IN + 1$. Sequences of cells that were actually found in the experiment are represented by darker circles. These include, by definition, all the leaves, but may include some internal nodes as well (node 8 here). Putative intermediates, deduced from the existence of shared mutations, are indicated by dashed circles. Split nodes may be either putative intermediates (nodes 3 and 10 here) or sequences actually found in the experiment. *B*, The list representation of the tree in *A*.

Results

Can gene conversion events be regarded as one mutational event?

The first questions we asked were: How should we count gene conversion blocks in our analysis? Should we treat a gene conversion as one mutational event, or as if there were a number of separate mutations, treating each base pair substitution as a separate event? The two ways of treating gene conversion events will obviously result in different tree shapes; however, the question is whether the two methods give similar insights into the diversification process. In our preliminary analyses, we have found that, after scaling (dividing) by the total number of nodes or the number of leaves, trees generated by the two methods are similar, and give similar general insights. This is illustrated by the examples shown in Fig. 2: for the two trees analyzed, the scaled properties are the same whether we regard gene conversion blocks as one or several events; even though the trees do not look similar, most of their scaled properties are highly similar. Tree bushiness (defined by outgoing degree measures) is unaffected, as the number of splits per node is independent of the above choice. Therefore, we henceforth treat each gene conversion event as one mutational event.

Comparison of trees from L and H chains in the same clones

Another basic question was: How close are L and H chain trees from the same clone in their graphical properties? How do they relate to the properties of the full tree (which includes mutations in both chains)? We found that, in general, the trends in L chain trees are similar to those found in H chain trees, despite the fact that the trees differ in size and details. The examples in Fig. 2 demonstrate

Table I. *Measured parameters of mutational lineage trees*

Parameter type	Parameter Name	Value in Fig. 1
Indicators of tree size[a]	Total no. of nodes, N	15
	No. of leaves, L	5
	No. of internal nodes, IN	9
	No. of pass-through nodes, PTN	7
Root to leaf distance-path length (PL)[b]	Minimum, MinPL	3
	Maximum, MaxPL	5
	Average, AvgPL	4.4
Root to first split node distance[c]	Trunk length, T	0
Outgoing degree (number of daughters per node), OD[d]	Minimum, MinOD	1
	Maximum, MaxOD	3
	Average, AvgOD	14/15
	Average OD over split nodes only, AvgOD2	7/3
	Root OD	2
Leaf to nearest split node distance (DLSN)[e]	Minimum, MinDLSN	1
	Maximum, MaxDLSN	4
	Average, AvgDLSN	2
Leaf to first (closest to the root) split node (DLFSN) distance[f]	Minimum, MinDLFSN	3
	Maximum, MaxDLFSN	5
	Average, AvgDLFSN	4.4
Root to split node (DRSN) distance[g]	Minimum, MinDRSN	2
	Maximum, MaxDRSN	4
	Average, AvgDRSN	3
Distance from Root to the maximal split node[h]	DRMSN	2
Distance between two adjacent split nodes (DASN)[i]	Minimum, MinDASN	2
	Maximum, MaxDASN	2
	Average, AvgDASN	2

[a] The parameters N, IN, and PTN are very well correlated with each other. All size parameters are very sensitive to the number of cells that happened to be picked from the clone described by the tree. The number of leaves, L, is the most sensitive to the "pick size". To eliminate this problem in our analysis, we often scale (divide) the measured properties by either N or L (47).

[b] We measure minimum, maximum, and average PL (over all leaves in the tree). PL is a good measure of tree "length", reflecting the length of the diversification process from the original cell. For example, maximum or average path lengths tend to increase as the GC response progresses in mice (39,47).

[c] The length of the trunk $T > 0$ only if the Root's degree is 1, that is, the root itself is not a split node; otherwise $T = 0$, as in Fig 1. A long trunk in murine or human GC B cell clones often indicates that the clone did not originate from a naive B cell, but rather from an already "experienced" B cell, which as already accumulated several mutations in a previous response, to the same or to a cross-reactive Ag (48,49), but in rabbits and especially in chickens the clone might have originated from a cell with diversification that occurred during the primary preimmune stage in GALT or bursa.

[d] The maximum or average OD are the best indicators of tree "bushiness", and also do not need to be scaled by tree size parameters. For a given rate of mutations, the lower the OD, the stronger the selection operating on the diversifying clone (47–49).

[e] DLSN can be regarded as an inverse measure of tree "bushiness", and hence another indicator of selection, as, naturally, the bushier the tree, the smaller the distance between a given leaf to the last split node. It reflects the most recent events in the history of the branch.

[f] DLFSN variables are basically path lengths excluding the trunk, which we use instead of PL when we do not know for certain whether the trees originated from naive or experienced B cells (48–49).

[g] DRSN is measured on each path to a leaf, not considering the root itself, and included only split nodes that are closest to the root on each path. It is similar to *T*.

[h] DRMSN is defined as the distance from the root to the maximal (in terms of outgoing degree) split node.

[i] Like DLSN, DASN is an inverse measure of tree bushiness.

this point. Hence, in subsequent studies, we give equal weights to L and H chain gene trees, and regard the properties of two trees generated from the same clone (if available) as two data sets characterizing the clone.

Temporal development of rabbit GC trees: hypermutation vs selection

Having established the above, we conducted a comparative analysis of trees from rabbit appendix (primary diversification), at different time points during postnatal development, and from GCs (secondary diversification) at different time points during a specific immune response. The different trees showed rather large size variability, as seen in mice and humans. In immunized rabbits, trees from the same GC were usually similar in all properties, except general size (data not shown). Many unique sequences were found early in the response (days 7 and 10), and fewer unique sequences were found later. This is similar to findings in chickens (16, 17), and to a lesser extent in mice (39), in which it has been interpreted as reflecting selection via competition for Ag.

The rabbit splenic GC trees from days 7 and 10 of a specific immune response did not significantly differ in the various size (Fig. 3, *A–C*) parameters, or in bushiness in terms of outgoing degree measures (Fig. 3*D*), but on day 15 much larger trees were found. Trees from day 7 significantly differed in all size properties from day 15 trees ($p < 0.01$ for all properties except trunk length, for which $p < 0.07$; the single day 10 tree did not have a trunk). Day 15 trees were also slightly bushier than day 7 or 10 trees ($p < 0.05$ for outgoing degree averaged over split nodes (AvgOD2); $p < 0.07$ for maximum outgoing degree (MaxOD)).

The lengthening of trees with time is similar to what was found in murine immune responses (39, 47). However, rabbit GC trees were in general larger, not only in terms of path lengths, that is, number of mutations per leaf, which increased with time (data not shown), but also in the number of leaves. The latter grows with the

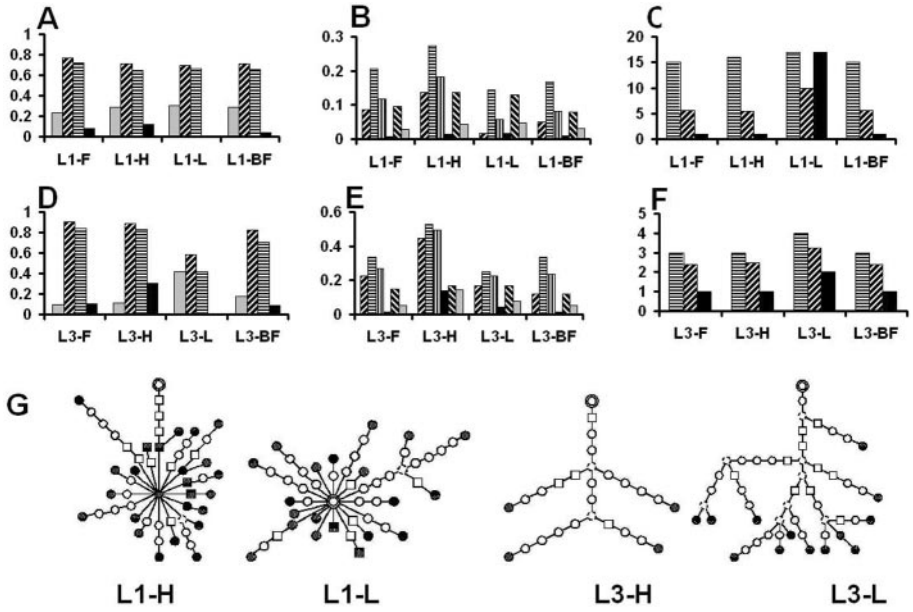

FIGURE 2. Properties of two trees derived from splenic DNP-specific GCs of immunized rabbits (14, 15). *A–C*, Clone L1; *D–F*, clone L3. *G*, From *left* to *right*: the trees L1-H, L1-L, L3-H, and L3-L, using the same conventions as in Fig. 1. Circles represent point mutations, and squares represent gene conversion events. Properties are shown for the H chain tree (L1-H and L3-H); the L chain tree (L1-L and L3-L); the full tree, which includes mutations in both the H and L chain genes (L1-F and L3-F), in which all the above were analyzed with gene conversion blocks taken as single events; and the full tree with each base change taken as a separate event, including base changes within gene conversion blocks (L1-BF and L3-BF). *A* and *D*, Scaled tree size parameters: gray, L/N (that is, number of leaves divided by number of nodes); ▨, IN/N; ▤, PTN/N; ■, T/N. *B* and *E*, Scaled lengths: ▨, MinPL/N; ▤, MaxPL/N; ▥, AvgPL/N; ■, MinDLSN/N; ▧, MaxDLSN/N; gray, AvgDLSN/N. *C* and *F*, Parameters defining tree bushiness: ▤, MaxOD; ▨, Avg2OD excluding pass-through nodes; ■, RootD.

number of cells sampled, but also reflects the intrinsic bushiness (number of branches in each branching point) of the tree. Thus, we have an overall growth in the number of nodes, which reflects both the number of leaves and the path length. In contrast to murine GC trees (39, 47), however, the number of leaves and the various measures of tree bushiness did not decrease with time (Fig. 3). This may reflect differences between the diversification processes in the two species: in murine GCs, selection dominates over hypermutation by day 15 of the response, and the trees become less bushy, while in rabbit GCs, selection either does not dominate at all over the higher rate hypermutation, or possibly does so only later in the response.

Indeed, rabbit GC trees were both longer (in terms of scaled path lengths) and slightly bushier (in terms of outgoing degree) than murine (47) or human (48) GC trees. In murine splenic GC trees, path lengths observed to date were at most 13 mutations long (35–40), while rabbit splenic GC trees have many paths of the order of 15–20 mutations accumulated in the same time period. Because we counted every gene conversion block as one mutation, the actual number of base changes was even higher. Scaling by the number of leaves, or of nodes, removes the size differences between different days, suggesting that the rate of hypermutation, which generates the tree, remains essentially the same throughout the response (Fig. 3). However, our analysis suggests that the rate of hypermutation, which is proportional to the path lengths, in rabbits (and chickens, as shown below) is larger than that in mice, even when we count every gene conversion block as one mutation.

Temporal development of chicken GC trees: similar to or faster than rabbit GC

Chickens, like rabbits, use gene conversion as a mechanism of primary and secondary diversification. Only a small number of lineage trees was published from chicken GCs, on the seventh and the eleventh day of the response (16, 17). Applying our analysis to these data, we see that tree growth is undetectable, and the bushiness (maximum and average OD) decreases only slightly between

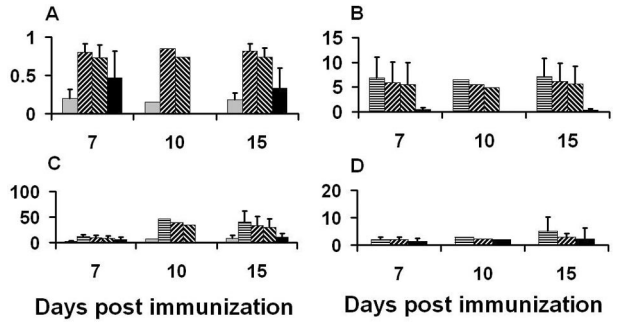

FIGURE 3. Data from all immunized rabbit splenic DNP-specific GC trees (13–15). *A–C*, Tree size parameters. *A*, Scaled by nodes; *B*, scaled by leaves; *C*, unscaled. Gray, L (leaves); ▤, N (nodes); ▨, IN (internal nodes); ▧, PTN (pass-through nodes); ■, T (trunk length). *D*, Tree bushiness parameters: ▤, MaxOD; ▨, AvgOD2; ■, RootD.

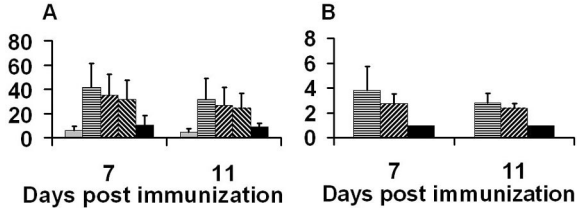

FIGURE 4. Trees from immunized chicken NP (16)- or FITC (17)-specific splenic GCs. Both graphs show unscaled properties. All symbols are as defined in the legend to Fig. 3. *A*, Tree size: L; N; IN; PTN (pass-through nodes); T. *B*, Tree bushiness: MaxOD; AvgOD2; RootD.

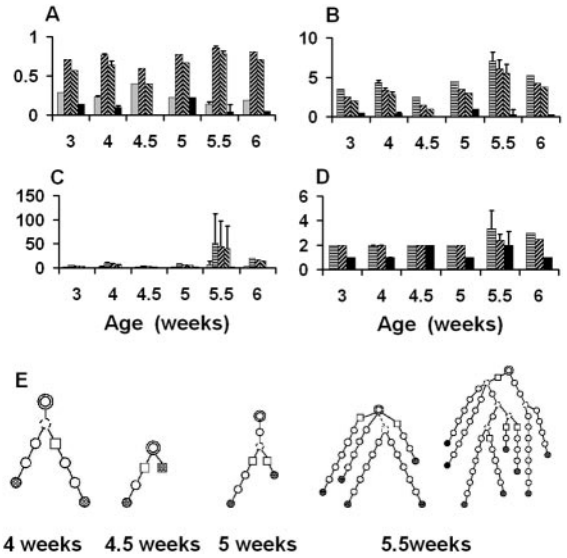

FIGURE 5. Rabbit appendix trees (from Fig. 4 in Ref. 11). Parameters and all symbols are as in Fig. 3. *A–C*, Tree size parameters. *A*, Scaled by nodes; *B*, scaled by leaves; *C*, unscaled. Gray, L; ▤, N; ▨, IN; ▧, PTN (pass-through nodes); ■, T. *D*, Tree bushiness parameters: ▤, MaxOD; ▨, AvgOD2; ■, RootD.

the two time points (Fig. 4). This is despite the fact that many mutations have accumulated per path, as in rabbit splenic GCs, so that the trees are as large on day 7 in chicken GC as on day 15 in rabbit GC ($p > 0.05$ for all unscaled size comparisons) and significantly larger ($p < 0.05$) from trees in day 7 of rabbit GC. There were no significant differences in bushiness (MaxOD, AvgOD2, and root degree (RootD)).

Primary diversification in rabbit appendix

The results of comparing the trees from rabbit primary (11) to secondary diversification revealed the greatest insights. During rabbit postnatal B cell development in the appendix, the earliest trees were very small and few, but became increasingly larger with time (Fig. 5). However, the average bushiness (outgoing degree) did not significantly change with time. This is in striking contrast to the behavior of trees from murine immune responses to specific Ags, in which path lengths, even scaled ones, grow in size, yet the outgoing degree decreases, as the response progresses (47). This is thought to be due to the intense selection, which trims branches that did not lead to successful cells, so that the number of branches decreases while their lengths increase. This selection seems to be weaker in rabbit splenic GCs, as shown above, and completely missing in appendix clonal diversification. Note that although the trees from rabbit age 5.5 wk look quite large and bushy, they

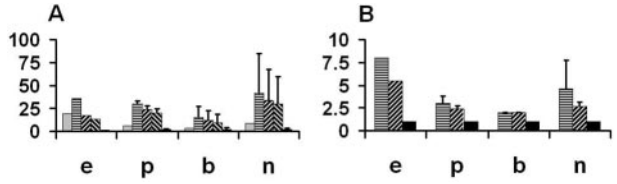

FIGURE 6. Trees from chicken bursa (12). Tree sources are: e, embryonic; p, postnatal; b, postnatal with bursal duct ligation; n, postnatal with bursal duct ligation and addition of the Ag NP. Both graphs show unscaled properties. All symbols are as defined in the legend to Fig. 3. *A*, Tree size: L; N; IN; PTN (pass-through nodes); T. *B*, Tree bushiness: MaxOD; AvgOD2; RootD.

significantly differ from rabbit GC trees in their lack of trunks and in their shorter paths ($p < 0.05$ for all path length comparisons between 5.5-wk appendix trees and day 15 GC trees).

Taken together, these results illustrate a diversification process that seems to go on and on at an apparently constant rate. In contrast to rabbit splenic GC trees, in the appendix even the scaled tree size properties show growth with time (compare Fig. 3, *A* and *B*, with Fig. 5, *A* and *B*). This can be interpreted to show that, in primary diversification, the balance between diversification and selection tends toward diversification even more strongly and completely than in rabbit splenic GCs. In contrast, the comparison lends support to the notion that in rabbit splenic GCs, some selection must be occurring, even though it is not reflected too strongly by tree bushiness parameter differences.

Primary diversification in chicken bursa

To compare primary diversification in rabbits with that in a different species that uses gene conversion, we have analyzed the trees from chicken bursa (12). The data are again limited, as only one embryonic tree is compared in this study with four posthatching trees. What we can say is that posthatching trees are both larger and less bushy than the embryonic tree (Fig. 6). Path lengths also increase (data not shown). Larger tree size is observed in scaled properties as well as in the unscaled ones, as in rabbit appendix (data not shown).

It is hard to say, based on these data, whether primary diversification in chickens is similar in all aspects to that in rabbits. The involvement of Ag in the posthatching diversification process is implied by the fact that trees are both smaller and less bushy in chickens in which bursal duct ligation (BDL) was performed than in non-BDL chickens, or BDL chickens in which the Ag 4-hydroxy-3-nitrophenylacetyl (NP) was added to the bursa, bypassing the ligation (BDL-NP). However, differences between trees from normal posthatching chickens and BDL chickens are not as large as those between BDL and BDL-NP chickens, suggesting that the BDL-NP situation is rather artificial.

Scaling of path lengths by nodes results in an opposite trend, that is, shorter paths in NP-BDL trees; we attribute this to a less stringent selection. Note that when measured on trees, the stringency of selection is only relative to the force of proliferation. Hence, it may be that the same selection forces act in both cases; however, in the case of NP-BDL trees, these selection forces are less able to curb the diversification.

Discussion

This work presents the application of a quantitative algorithm for analyzing Ig gene lineage tree shapes, which we have previously applied to the relatively simple murine and human systems, to data on Ig gene diversification in rabbits and chickens. The process of gene conversion, used by these species for both primary and secondary diversification, is shown by our analysis to have several interesting characteristics: primary diversification, at least in rabbits, seems to go on at a constant rate, without ligand-driven selection; the balance between diversification and selection seems to be more in favor of diversification and less in favor of selection, than it is in murine or human GCs; and gene conversion in rabbits may be more efficient than in chickens.

The idea that the diversification process in the appendix appears to occur at a constant rate is based on our observation that both tree size (measured by numbers of nodes) and bushiness (measured by outgoing degrees) grow with time. The type of Ag-specific selection seen in splenic GCs seems to be absent in the appendix, and this supports the view that the primary repertoire is generated within the expanding clonally related B cells in young rabbits, and

emphasizes the important role that gut-associated lymphoid tissues play in early development of mammalian immune repertoires.

The greater bushiness of rabbit trees may reflect a greater efficiency of the rabbit's gene conversion process in generating new sequences that are still functional. Compared with the chicken data, in rabbits we have not seen many out-of-frame sequences resulting from gene conversions. The present analysis of tree shape further supports this notion, because in chickens we do find indications of selection, which may merely be favoring cells that maintained productive sequences after gene conversion events.

Rabbit and chicken splenic GC trees are both longer and bushier than mouse GC trees. A higher rate of hypermutation in rabbit and chicken GCs, compared with murine GCs, may be the reason for the longer paths (higher numbers of accumulated mutations per cell), and possibly for the higher bushiness. Even if the same selective forces apply in both systems, the balance between hypermutation and selection tends more toward mutation and less toward selection in rabbit and chicken GCs compared with murine GCs. This is consistent with our past suggestion that the GC in rabbit spleen might be a site for continued generation of B cell receptor diversity in adult life (13); some of the alterations of sequence may also lead to new members of the B cell repertoire in adult rabbits comparable to those produced in gut-associated lymphoid tissues of young rabbits (14). That the balance tends more toward hypermutation than toward selection does not, however, necessarily mean that selection is impaired, or that the danger of producing autoreactive clones is not avoided. The strength of selection may well be the same in all model systems, while either the rate or duration of hypermutation may be increased.

The high variability in all measured tree properties is a characteristic that we have encountered in all of our studies to date (47–49). In the rabbit GC data, for example, a large clone often (but not always) dominated a given GC, with the other clones in the same GC yielding fewer cells and smaller trees (13–15). A large clone may dominate a GC because it was probably good to begin with, i.e., the founder cell may have already had H and L chain complementarity-determining region 3 (CDR3) conformations with good affinity for the Ag used. This was suggested by the minimal number of changes in HCDR3 and LCDR3 in some cases. Alternatively, one key mutation, which increases the affinity by up to an order of magnitude, may result in clonal dominance of a GC (50). Thus, the initial fit with the Ag, and hence the potential for improvement by mutation (if any), may determine in a large part the dynamics of the response of each individual clone. The variability in potential, hence in dynamics, and hence in tree shape, is a reflection of the sensitivity of the GC reaction to the particular Ig gene sequences of the founder cells.

The variability and sensitivity issues have to date been largely ignored, in both experimental and theoretical studies of GC dynamics. Studies of only a few clones were taken to represent the dynamics of a typical GC reaction (35–40), and most models relied on these data (22–26), with only a few addressing the sensitivity of GC reaction dynamics to individual sequence characteristics (51, 52). The focus of the studies to date has been on the affinity-improving mutations, largely ignoring the fact that most mutations are actually destructive, or at best neutral (21). The proper way to address the variability and sensitivity issues, and the question of whether the sample size affects the results, is via computer simulations in which all the parameters are known, and all mutations in all cells are counted; we intend to perform such studies in the near future.

An additional factor, compounding the sensitivity of response dynamics to sequence potential and hence increasing tree shape variability, is the competition for Ag between clones in each GC. Part of tree pruning, the elimination of less successful branches, is probably due to competition for Ag rather than functionality of the sequence. This would be true especially in the late stages of the response, when there is not much Ag around. At least this is our interpretation of the decrease in bushiness of murine GC trees with time. The differences between tree sizes in the same GC in the rabbit data seem to indicate that a similar competition may be going on here.

All standard algorithms of tree generation from sequence data assume that each shared change indicated cells with a shared precursor. In drawing our trees from rabbit data, especially from the appendix sequences, we could not always accept this assumption, as we had indications of repeated independent occurrences of point mutations and even gene conversions. We could sometimes infer that specific point mutations occurred independently at a given position (e.g., in HCDR3) because the cells carrying them were on different branches. Conversely, the best overall fit of the sequence data sometimes led to the conclusion that gene conversions recurred on different branches. As long as the trees are produced by a consistent algorithm, we consider the comparison between shape properties of different trees to be legitimate.

In conclusion, use of the algorithm described in this study has revealed new insights about Ig gene diversification in rabbits and chickens that were not obvious from less rigorous examination of the available data on clonal trees. These insights include: first, primary diversification in the rabbit appendix appears to occur at a constant rate, and is not subject to the types of Ag-driven selection observed in GC responses after immunization of rabbits and other species. Second, in both rabbits and chickens, the balance between mutation and selection tends more toward diversification, which may reflect the importance of maintaining a diverse repertoire in the adult once appendix or bursal B cell receptor diversification has diminished or ceased. Finally, gene conversion in rabbits may be more efficient than in chickens. This may be explained by the introduction of blocks of changes in CDRs without introduction of deleterious replacement changes in the framework regions, and few, if any, insertions and deletions that lead to frameshifts during the conversion process in rabbits.

Acknowledgments

We are grateful to Drs. D. Dunn-Walters, A. George, M. Mage, M. Meyer-Schellersheim, B. Newman, W. Paul, K. Rao, and S. Rath for critical review of the manuscript.

References

1. Janeway, C. A., P. Travers, M. Walport, and J. D. Capra. 1999. *Immunobiology: The Immune System in Health and Disease*. Elsevier/Garland, New York, pp. 196–198.
2. Oltz, E. M. 2001. Regulation of antigen receptor gene assembly in lymphocytes. *Immunol. Res. 23:121*.
3. Hesslein, D. G., and D. G. Schatz. 2001. Factors and forces controlling V(D)J recombination. *Adv. Immunol. 78:169*.
4. Gellert, M. 2002. V(D)J recombination: RAG proteins, repair factors, and regulation. *Annu. Rev. Biochem. 71:101*.
5. Kelsoe, G. 1996. Life and death in germinal centers. *Immunity 4:107*.
6. Wabl, M., M. Cascalho, and C. Steinberg. 1999. Hypermutation in antibody affinity maturation. *Curr. Opin. Immunol. 11:186*.
7. Neuberger, M. S., M. R. Ehrenstein, C. Rada, J. Sale, F. D. Batista, G. Williams, and C. Milstein. 2000. Memory in the B cell compartment: antibody affinity maturation. *Philos. Trans. R. Soc. Lond. B Biol. Sci. 1395:357*.
8. Diaz, M., and P. Casali. 2002. Somatic immunoglobulin hypermutation. *Curr. Opin. Immunol. 14:235*.
9. Winter, D. B., and P. J. Gearhart. 1998. Dual enigma of somatic mutation of immunoglobulin variable genes: targeting and mechanism. *Immunol. Rev. 162:89*.
10. Neuberger, M. S., M. R. Ehrenstein, N. Klix, C. J. Jolly, J. Yelamos, C. Rada, and C. Milstein. 1998. Monitoring and interpreting the intrinsic features of somatic hypermutation. *Immunol. Rev. 162:107*.
11. Seghal, D., H. Obiakor, and R. G. Mage. 2002. Distinct clonal diversification patterns in young appendix compared to antigen-specific splenic clones. *J. Immunol. 168:5424*.

12. Arakawa, H., K. Kuma, M. Yasuda, S. Ekino, A. Shimizu, and H. Yamagishi. 2002. Effect of environmental antigens on the Ig diversification and the selection of productive V-J joints in the bursa. *J. Immunol. 169:818*.
13. Seghal, D., E. Schiaffella, A. O. Anderson, and R. G. Mage. 1998. Analysis of single B cells by PCR reveals rearranged V_H with germline sequences in spleen of immunized rabbits: implications for B cell repertoire maintenance and renewal. *J. Immunol. 161:5347*.
14. Schiaffella, E., D. Seghal, A. O. Anderson, and R. G. Mage. 1999. Gene conversion and hypermutation during diversification of V_H sequences in developing splenic germinal centers of immunized rabbits. *J. Immunol. 162:3984*.
15. Seghal, D., E. Schiaffella, A. O. Anderson, and R. G. Mage. 2000. Generation of heterogeneous rabbit anti-DNP antibodies by gene conversion and hypermutation of rearranged V_L and V_H genes during clonal expansion of B cells in splenic germinal centers. *Eur. J. Immunol. 30:3634*.
16. Arakawa, H., S. Furusawa, S. Ekino, and H. Yamagishi. 1996. Immunoglobulin gene hyperconversion ongoing in chicken splenic germinal centers. *EMBO J. 15:2540*.
17. Arakawa, H., K. Kuma, M. Yasuda, S. Furusawa, S. Ekino, and H. Yamagishi. 1998. Oligoclonal development of B cells bearing discrete Ig chains in chicken single germinal centers. *J. Immunol. 160:4232*.
18. Mehr, R., M. Shannon, and S. Litwin. 1999. Models for antigen receptor gene rearrangement. I. Biased receptor editing in B cells: implications for allelic exclusion. *J. Immunol. 163:1793*.
19. Piper, H., S. Litwin, and R. Mehr. 1999. Models for antigen receptor gene rearrangement. II. T cell receptor editing: allelic exclusion or inclusion? *J. Immunol. 163:1799*.
20. Kalmanovich, G., and R. Mehr. 2003. Models for antigen receptor gene rearrangement. II. Heavy and light chain allelic exclusion. *J. Immunol. 170:182*.
21. Shannon, M., and R. Mehr. 1999. Reconciling repertoire shift with affinity maturation: the role of deleterious mutations. *J. Immunol. 162:3950*.
22. Sulzer, B., L. van Hemmen, A. U. Neumann, and U. Behn. 1993. Memory in idiotypic networks due to competition between proliferation and differentiation. *Bull. Math. Biol. 55:1133*.
23. Kepler, T. B., and A. S. Perelson. 1993. Cyclic re-entry of germinal center B cells and the efficiency of affinity maturation. *Immunol. Today 14:412*.
24. Oprea, M., and A. S. Perelson. 1997. Somatic mutation leads to efficient affinity maturation when centrocytes recycle back to centroblasts. *J. Immunol. 158:5155*.
25. Shlomchik, M., P. Watts, M. Weigert, and S. Litwin. 1998. Clone: a Monte-Carlo computer simulation of B cell clonal expansion, somatic mutation, and antigen-driven selection. *Curr. Top. Microbiol. Immunol. 229:173*.
26. Kesmir, C., and R. J. De Boer. 1999. A mathematical model on germinal center kinetics and termination. *J. Immunol. 163:2463*.
27. Dunn-Walters, D. K., A. Dogan, L. Boursier, and J. Spencer. 1998. Base-specific sequences that bias somatic hypermutation deduced by analysis of out-of-frame human IgV_H genes. *J. Immunol. 160:2360*.
28. Dorner, T., S. J. Foster, H. P. Brezinschek, and P. E. Lipsky. 1998. Analysis of the targeting of the hypermutational machinery and the impact of subsequent selection on the distribution of nucleotide changes in human V_HDJ_H rearrangements. *Immunol. Rev. 162:161*.
29. Spencer, J., M. Dunn, and D. K. Dunn-Walters. 1999. Characteristics of sequences around individual nucleotide substitutions in IgV_H genes suggest different GC and AT mutators. *J. Immunol. 162:6596*.
30. Oprea, M., and T. B. Kepler. 1999. Genetic plasticity of V genes under somatic hypermutation: statistical analyses using a new resampling-based methodology. *Genet. Res. 9:1294*.
31. Kim, N., G. Bozek, J. C. Lo, and U. Storb. 1999. Different mismatch repair deficiencies all have the same effects on somatic hypermutation: intact primary mechanism accompanied by secondary modifications. *J. Exp. Med. 190:21*.
32. Foster, S. J., T. Dorner, and P. E. Lipsky. 1999. Somatic hypermutation of $V\kappa J\kappa$ rearrangements: targetting of RGYW motifs on both DNA strands and preferential selection of mutated codons within RGYW motifs. *Eur. J. Immunol. 29:4011*.
33. Michael, N., T. E. Martin, D. Nicolae, N. Kim, K. Padjen, P. Zhan, H. Nguyen, C. Pinkert, and U. Storb. 2002. Effects of sequence and structure on the hypermutability of immunoglobulin genes. *Immunity 16:123*.
34. Monson, N. L., T. Dorner, and P. E. Lipsky. 2000. Targeting and selection of mutations in human $V\lambda$-rearrangements. *Eur. J. Immunol. 30:1597*.
35. Kocks, C., and K. Rajewsky. 1988. Stepwise intraclonal maturation of antibody affinity through somatic hypermutation. *Proc. Natl. Acad. Sci. USA 85:8206*.
36. Manser, T. 1989. Evolution of antibody structure during the immune response. *J. Exp. Med. 170:1211*.
37. Jacob, J., G. Kelsoe, K. Rajewsky, and U. Weiss. 1991. Intraclonal generation of antibody mutants in germinal centres. *Nature 354:389*.
38. Jacob, J., and G. Kelsoe. 1992. In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl) acetyl. II. A common clonal origin for periarteriolar lymphoid sheath-associated foci and germinal centers. *J. Exp. Med. 176:679*.
39. Jacob, J., J. Przylepa, C. Miller, and G. Kelsoe. 1993. In situ studies of the primary immune response to (4-hydroxy-3- nitrophenyl)acetyl. III. The kinetics of V region mutation and selection in germinal center B cells. *J. Exp. Med. 178:1293*.
40. Vora, K. A., K. Tumas-Brundage, and T. Manser. 1999. Contrasting the in situ behavior of a memory B cell clone during primary and secondary immune responses. *J. Immunol. 163:4315*.
41. Dunn-Walters, D. K., L. Boursier, and J. Spencer. 1997. Hypermutation, diversity and dissemination of human intestinal lamina propria plasma cells. *Eur. J. Immunol. 27:2959*.
42. Sims, G. P., H. Shiono, N. Willcox, and D. Stott. 2001. Somatic hypermutation and selection of B cells in thymic germinal centers responding to acetylcholine receptor in myasthenia gravis. *J. Immunol. 167:1935*.
43. Brauninger, A., T. Spieker, K. Willenbrock, P. Gaulard, H. H. Wacker, K. Rajewsky, M. L. Hansmann, and R. Kuppers. 2001. Survival and clonal expansion of mutating "forbidden" (immunoglobulin receptor-deficient) Epstein-Barr virus-infected B cells in angioimmunoblastic T cell lymphoma. *J. Exp. Med. 194:927*.
44. 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. 16:3822*.] *J. Clin. Oncol. 17:1092*.
45. Hougs, L., L. Juul, H. J. Ditzel, C. Heilmann, A. Svejgaard, and T Barington. 1999. The first dose of a *Haemophilus influenzae* type b conjugate vaccine reactivates memory B cells: evidence for extensive clonal selection, intraclonal affinity maturation, and multiple isotype switches to IgA2. [Published erratum appears in *2001 J. Immunol. 162:224*.] *J. Immunol. 166:2147*.
46. Barington, T., L. Hougs, L. Juul, H. O. Madsen, L. P. Ryder, C. Heilmann, and A. Svejgaard. 1996. The progeny of a single virgin B cell predominates the human recall B cell response to the capsular polysaccharide of *Haemophilus influenzae* type b. *J. Immunol. 157:4016*.
47. Dunn-Walters, D. K., A. Belelovsky, H. Edelman, M. Banerjee, and R. Mehr. 2003. The dynamics of germinal centre selection as measured by graph-theoretical analysis of mutational lineage trees. *Clin. Dev. Immunol. 9:233*.
48. Banerjee, M., R. Mehr, A. Belelovsky, J. Spencer, and D. K. Dunn-Walters. 2002. Age- and tissue-specific differences in human germinal centre B cell selection revealed by analysis of IgV_H gene hypermutation and lineage trees. *Eur. J. Immunol. 32:1947*.
49. Mehr, R., M. Banerjee, and D. K. Dunn-Walters. 2003. Age effects on antibody affinity maturation. *Biochem. Soc. Trans. 31:447*.
50. Allen, D., T. Simon, F. Sablitzky, K. Rajewsky, and A. Cumano. 1988. Antibody engineering for the analysis of affinity maturation of an anti-hapten response. *EMBO J. 7:1995*.
51. Radmacher, M. D., G. Kelsoe, and T. B. Kepler. 1998. Predicted and inferred waiting times for key mutations in the germinal centre reaction: evidence for stochasticity in selection. *Immunol. Cell Biol. 76:373*.
52. Radmacher, M. D., and T. B. Kepler. 2001. Waiting times to appearance and dominance of advantageous mutants: estimation based on the likelihood. *Math. Biosci. 170:59*.