Immunoglobulin Gene Transcripts Have Distinct $V_H^{DH}$ Recombination Characteristics in Human Epithelial Cancer Cells*

Received for publication, December 18, 2008, and in revised form, February 27, 2009. Published, JBC Papers in Press, March 16, 2009, DOI 10.1074/jbc.M809524200

Jie Zheng‡1, Jing Huang§1, Yuntao Miao³, Shiqing Liu³, Xin Sun³, Xiaohui Zhu³, Teng Ma³, Li Zhang³, Jiafu Ji³, Youhui Zhang§, C. Cameron Yin§, and Xiaoyan Qiu¹²

From the ¹Peking University Center for Human Disease Genomics, Beijing, 100191, China, the ²Department of Surgery, Peking University School of Oncology, Beijing, 100142, China, the ³Department of Immunology, Cancer Institute, Chinese Academy of Medical Science and Peking Union Medical College, Beijing, 100021, China, and the ⁴Department of Hematopathology, University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030

It was well accepted that only B-lymphocytes and plasma cells expressed immunoglobulin (Ig) gene. However, our group and others have confirmed that non-B-cells, such as epithelial cancer cells, can also express Ig. The aim of this work is to elucidate the role of non-B-cell-derived Ig by investigating the characteristics of the Ig heavy chain (IgH) gene repertoire in epithelial cancer cells. We cloned and sequenced 89 $V_H^{DJH}$ (V-D-J recombination of the IgH variable region) transcripts by microdissecting cells from eight different types of epithelial cancers and two cancer cell lines (HT-29 and HeLa S3). The cancer-derived Ig gene repertoire showed specific restricted patterns of $V_H^{DJH}$ recombination with seven sets of predominant $V_H^{DJH}$ sequences. Surprisingly, within a set of $V_H^{DJH}$ recombination, the variable (V) sequences derived from different cancer types had not only identical heavy chain variable (VH), diversity (D), and joining (JH) segments usage, but also identical junctions and mutation targets in the VH region. The $V_H^{DJH}$ (but not $V_H^{DJH}$) in the cancer-derived sequences had a high mutation rate; however, it was shown that the mechanism of hypermutation was different from antigen selection in B-cell-derived $V_H^{DJH}$ sequences. In contrast to $V_H^{DJH}$, the $V_H^{DJH}$ sequences did not appear to originate from classical class switching. These results suggest that cancer-derived Ig genes have a distinct repertoire that may have implications for their role in carcinogenesis.

Immunoglobulins (Ig) were discovered more than a century ago, yet the understanding of these proteins continues to evolve. Until 1950, most scientists believed that cells from various types of tissues could express Ig (1). However, it was shown that B-lymphocytes from bone marrow secreted Ig, although other hematopoietic cells did not (2), and that levels of serum Ig decreased with B-cell disfigurement (2, 3). These were thought to indicate that only B-lymphocytes could express Ig; non-immunocytes could not.

In 1976, Tonegawa discovered that Ig gene recombination was the mechanism behind antibody diversity in B-lymphocyte-derived plasma cells. Ig gene recombination, as theorized previously by Dreyer and Bennett, was confirmed to exist in mouse myeloma cells using a probe against the Ig mRNA kappa chain (4, 5).

Subsequently, Cleary et al. compared the restriction enzyme map of the Ig gene in B-lymphocytes with that of the genes in cell types such as germ-line using Southern blot analysis and found that B-cell and non-B-cell restriction maps differed. These results further strengthened the hypothesis that Ig gene recombination only occurred in B-lymphocytes. Consequently, Ig gene recombination became a criterion for identifying B-cells (6, 7). Some tumor cells expressing both epithelial cell markers and Ig gene recombination were thus believed to originate from B-cells (6, 8).

Immunoglobulin gene recombination has been detected in T-cell lymphomas and acute non-lymphocytic leukemias (9, 10). However, there is no substantial evidence that Ig gene recombination, transcription, and production could occur in non-immunocytes.

Patients with non-hematopoietic tumors, including carcinomas of the brain, breast, colon, and liver, may have elevated levels of serum IgG, IgA, and/or IgM (11–13). Additionally, many patients with malignant tumors of epithelial origin have been shown to have monoclonal or oligoclonal gamma globulinemia (14–16). These antibodies had been presumed to be produced by B-lymphocytes and plasma cells. However, recent studies from our group and others have demonstrated that both malignant and normal epithelial cells could express Ig.

In 1996, we first reported the detection of IgG-like molecules in breast and colon carcinoma cells and showed that these molecules were not present in their normal epithelial cell counterparts by immunohistochemical staining and Western blot analysis (17). In studies of human cancer cell lines, IgG-like proteins were detected in both the tumor cells and the culture supernatant (18). Kimoto (19) identified transcripts of the Ig constant region and the T-cell receptor (TCR) gene in five epithelial-derived cancer cell lines (SW1116, HEP2, MCF-7, MDA-MB-231, and HC48) using

* This work was supported by Fundamental Research Grants 30371609 and 30572094 from the Natural Sciences Foundation, China. The nucleotide sequence(s) reported in this paper has been submitted to the GenBank/EMBL Data Bank with accession number(s) AY270187-AY270190, AY247234, AY286495, and AY505568.

† These authors contributed equally to this work.

‡ To whom correspondence should be addressed: Center for Human Disease Genomics, Peking University, 38 Xue-yuan Road, 100191, Beijing, P. R. China. Tel.: 86-10-82802846-5038; Fax: 86-10-82801149; E-mail: qiuxy@bjmu.edu.cn or xqiu@mdanderson.org.
nested reverse transcription-PCR (RT-PCR). In 2003, we demonstrated that tumor cells isolated from epithelial cancers and cell lines could secrete IgG using Western blot analysis and N terminus sequencing, and we detected both cytoplasmic and secreted IgG in cells from carcinomas of the lung, breast, liver, and colon, as well as epithelial cell lines (20). IgG transcription was also detected by in situ hybridization, Northern blot analysis, and single cell RT-PCR (20).

In 2004, it was reported that human cervical cancer cells could express Ig mRNA and protein (21). Recent studies have also confirmed the expression of Ig and activation-induced cytidine deaminase (AID) in six breast cancer cell lines (BT474, MDA-MB-231, MCF-7, SKBR3, T47D, and ZR75-1) (22). Furthermore, we recently reported that IgA and IgG were expressed in numerous oral epithelial tumor cells (23). Despite the detection of Ig in numerous cancer cell types, Ig specificity and variable region repertoire are poorly characterized.

B-cells are known to generate Ig diversity by several mechanisms. During the formation of Ig in B-cells from bone marrow, two recombinant events bring different VH, Ds1, and J1 exons together to form heavy chains. Additionally, short sequences are inserted between VH and DH and between DH and J1 to generate further diversity. Subsequent encounters with antigens in the germinal centers drive B-cell to undergo somatic hypermutation (SHM) and class switching, thus generating even greater diversity.

In the present study, we analyzed the V region transcripts in rearranged IgH in eight cancer cell samples from microdissected epithelial cancer tissue and in two cell lines (HT-29 and HeLaS3) using RT-PCR and sequence analysis. We found that cancer-derived Ig genes show classic VH-DH recombination. Reminiscent of classical recombination, additional short sequences were inserted between VH and DH segments and the DH and J1 segments. Rearranged Ig μ- and γ-chain gene sequences were expressed in cells, and SHM was observed in the VH segment of the γ-chain genes. However, the cancer-derived Ig gene repertoire also displayed several distinct characteristics.

**EXPERIMENTAL PROCEDURES**

**Sample Assays**—For laser capture microdissection (LCM) and RT-PCR analysis of the cancer-derived Ig gene, eight tumor samples from therapeutic excisions of breast invasive ductal carcinoma (n = 3), colon carcinoma (n = 2), squamous cell carcinoma of the lung (n = 1), from the tissue bank of Peking University School of Oncology), squamous cell carcinoma of the oral cavity (n = 1), and basal cell carcinoma of the oral cavity (n = 1) from the Department of Pathology at Peking University School of Stomatology were included with informed consent (20). Informed consent was obtained from the patients. Ethical approval of the study was granted by the Peking University Health Service Trust Research Ethics Committee. Carcinoma samples were embedded in Tissue-Tek OCT Compound (Sakura, IMEB International Medical Equipment, Inc., San Marcos, CA) and snap-frozen in liquid nitrogen immediately after surgery. Serial frozen sections (8 μm) were cut with a cryostat and mounted on slides treated with 0.1% DEPC for sterilization. Sections were air-dried, fixed in 70% ethanol, and evaluated with hematoxylin and eosin stain or immunohistochemical staining for LCM.

**Immunohistochemistry**—The slides were then incubated with 0.3% hydrogen peroxide for 5 min, washed with PBS, and blocked in PBS with 10% normal goat serum for 10 min. After removal of excess blocking buffer, indirect immunohistochemical staining was performed with monoclonal antibodies against human epithelial cell adhesion molecules (EpCAM) (1:100, Abcam, Cambridge, MA). Slides were incubated at 37 °C for 45 min, washed thoroughly, and then incubated with horse-radish peroxidase-labeled goat anti-mouse IgG (1:100, Dako, Carpinteria, CA) at 37 °C for 45 min. Slides were washed again and bound antibodies were detected using 3,3’-diaminobenzidine (DAB, Sigma Aldrich). The mouse IgG was used as an isotype control.

**HT-29 and HeLa S3 Cell Cultures**—HT-29 (from colon cancer) and HeLa S3 (from cervical cancer) cell lines were supplied by the Peking University Center for Human Disease Genomics. These two cancer cell lines were maintained in Dulbecco’s modified Eagle’s medium (Invitrogen, Carlsbad, CA) supplemented with 10% fetal bovine serum (HyClone/Thermo Fisher Scientific Inc., Waltham, MA) and L-glutamine (2 mM) at 37 °C in a humidified 5% CO2.

**Isolation and Preparation of Mononuclear Cells from Peripheral Blood**—Two samples of peripheral blood were obtained from two healthy donors. Mononuclear cells (MNC) were isolated from 5 ml of peripheral blood using two-step discontinuous Ficoll-Hypaque gradients (Second Chemistry Factory, Shanghai, China). The white gradient layer containing MNC were recovered and washed with 0.01 M PBS, and the isolated MNC used immediately for total RNA extraction.

**LCM and RT-PCR of Cancer Cells**—LCM was carried out as previously described (20). Briefly, to minimize contamination of infiltrating B-lymphocytes or plasma cells in cancer tissues, only large EpCAM+ cells in cancer cell nests without lymphocyte or plasma cell infiltration were dissected from fresh biopsy tissues of carcinomas of the breast, colon, oral cavity, and lung by LCM. Total RNA of microdissected cancer cells was extracted using RNeasy Micro kit (Qiagen) according to the manufacturer’s instructions. Reverse transcription was carried out with the Sensiscript RT kit (Qiagen) according to the manufacturer’s instructions. Touchdown PCR was then performed using 1 μl of each reverse transcription reaction with LA Taq polymerase (TaKaRa Bio USA, Madison, WI) as previously described (20). To amplify the human IgVH gene of the γ chain and the μ chain by nested PCR, the first round of PCR was carried out with upstream primers for VH1 (5’-GAGGTCGACGTTGAGGGAG-3’), VH2 (5’-CAGGGATGCCCCCGAAGAAGGAGGAGCAGTCTTGGG-3’), VH3 (5’-CAGGATACAGCTGAGCAGTCCAG-TGGAGCAGTCTGGG-3’), and VH4 (5’-CAGGTGCAGCTGAGCTGAGCAGTCCAG-TGGAGCAGTCTGGG-3’), coupled with CH1 region primer (Cγ1CH1, 5’-ACACCGTACGCGTCCGGG-3’; CμCH1, 5’-ACGCTGC-

---

3 The abbreviations used are: RT-PCR, reverse transcription-PCR; AID, activation-induced cytidine deaminase; PBS, phosphate-buffered saline; SHM, somatic hypermutation; LCM, laser capture microdissection; EpCAM, epithelial cell adhesion molecule; MNC, mononuclear cell; PBL, peripheral blood lymphocytes; CDR, complementary determining region; FWR, framework regions.
Ig Gene Recombination in Human Cancers

TCGATCCGACGGG-3'). Conditions for the second round of PCR were the same as the first round, except with the JH primer (5' -GTGACAGGGTNCCCTGCCCAGG-3') replacing the CH1 primer. To confirm that there was no B-lymphocyte contamination, we studied CD19 (a B-lymphocyte marker) transcription using the same cDNA used for the amplification of the IgH V gene, and the following primer set: CD19 up, 5'-AAGGGGCTAAATCTATTGCT-3' (sense), and CD19 down, 5'-CAGCTTCCCCATCTTGTTTCT-3' (antisense).

RT-PCR for Cell Lines and Peripheral Blood Lymphocytes—Total RNA were extracted from HT-29, HeLa S3, and peripheral blood MNC using TRIzol reagent (Invitrogen). Reverse transcription of total RNA from each of these samples was performed using a Superscript II RT kit (Invitrogen) according to the manufacturer's instructions. The human IgVH gene of the γ- or µ-chain and CD19 gene were amplified using the same PCR conditions and primers as those employed for the RT-PCR of cells obtained by LCM.

Sequencing and Analysis of Rearranged Genes—PCR products were cloned in a pGEM-T Easy Vector (Promega, Madison, WI) and sequenced with an ABI 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA). The V_{1H}D_{1H} sequences were compared with those found in the BLAST and Immunogenetics data bases (24) to identify the best matching germline sequences examined. The repertoire of the cancer-derived Ig V genes was compared with those found in the BLAST and ImmunoGenetics databases. The repertoire of the cancer-derived Ig V genes was compared with those found in the BLAST and ImmunoGenetics databases.

RESULTS

Cancer-derived IgVH Transcripts Were Amplified by RT-PCR—The Ig gene transcripts and repertoires were detected following LCM of cells from carcinomas from the colon, breast, oral, and lung (Fig. 1A) and HT-29 and HeLa S3 cell lines. Peripheral blood lymphocytes from two healthy donors served as positive controls. The rearranged V region genes of the γ and µ chains were successfully amplified and cloned from both the cancer cells and PBL (Fig. 1B). As expected, CD19 transcript was not detected in any of the cancer cell cDNA libraries, but was detected in PBL from the two donors (data not shown).

Analysis of IgVH Gene Mutations—The pattern of mutations of each sequence was compared with that of germline sequence to identify hybrid sequences derived from recombinant V_{1H} gene segments. To analyze whether the mechanism of SHM occurring in cancer-derived Ig variable region was similar to that caused by antigen selection in B-cell-derived Ig, the mutation frequency of both the RGYW and the WRCY motifs (the mutable position is G:C, which is underlined; r = A or G, Y = C or T, and W = T or A) used as a principal hotspot for AID-induced G:U lesions was calculated (29–31). In addition, we determined the replacement-to-silent mutation (R/S) ratio in the CDRII and FWRIII regions. A V_{1H} sequence was considered to be antigen-selected when the R/S ratio was higher than 2.9 in the CDRII and lower than 1.5 in the FWRIII region (32). The error-prone polymerases mainly induced A/T mutations that were identified as a principal site, and the dinucleotide target WA (AA or TA) mutation induced by the error-prone polymerases was involved in the mismatch repair of SHM (33–35). Therefore the WA/TW ratio was analyzed using JOINSOLVER (36).

Statistical Analysis—The distribution of V_{1H} and J_{1H} gene family usage and the calculation of mutations in WA versus TW were assessed using the Chi-squared test. Values were considered statistically significant when p < 0.05.
showed that JH4 was expressed most often (72.7%), although all six JH genes were detected (Fig. 2B).

Several Sets of Predominant Functional VHDJH Gene Recombinations Were Detected in Human Cancer Cells—A noteworthy finding in this study was that the cancer-derived Ig gene repertoire exhibited distinct characteristics. LCM-isolated cancer cells usually expressed one to three sets of dominant VHDJH recombination patterns, as well as some individual VHDJH recombination (Table 2). Notably, seven sets of predominant VHDJH recombination were detected in more than one cancer cell type (Table 3). Of the 26 VHDJH sequences studied, three sets of VHDJH recombinations were predominant: V1-3-15/D3-10/JH4, found in 5 of 26 (19.2%) cases of carcinomas from the breast and colon; V1-6-1/D6-13/JH4, found in 5 of 26 (19.2%) cases of carcinomas from the lung and oral cavity; and V1-4-30-2/D3-22/JH4, found in 4 of 26 (15.4%) cases of carcinomas from the breast and colon. Of the 41 VHDJH sequences studied, the following types of VHDJH recombinations were predominant: V1-5-51/D3-9/JH4 in 19 of 38 (50%) cases of carcinomas of the breast, colon, oral, and lung; and V1-3-30/D6-19/JH4 in 4 of 38 (10.5%) cases of carcinomas of the breast, colon, and lung. There was no correlation between the patterns of VHDJH recombinations and the histological origins of the cancers.

In contrast, all of the VHDJH sequences obtained, either from the HT-29 cell line (7 clones) or from the HeLa S3 cell line (6 clones), showed distinct diversity. None of the J chain clones from the cell lines was detected in any of the LCM-isolated cancer cells. All six VH1-4D1/JH1-4 sequences from HT-29 cells showed identical VH6-1/D6-13/JH4 recombination, which was also detected in LCM-isolated carcinomas of the lung and oral cavity. All three VH1-4D1/JH1-4 sequences from HeLa S3 cells showed identical VH3-30/D6-19/JH5 recombination, which was not identified in the LCM-isolated cancer cells. Ten VH/VH1-4H2 recombination sequences in two control PBL samples exhibited distinct diversity (Table 2).

Mechanism of SHM of Functional V Genes in Cancer Cells Was Different from Antigen Selection in B-cell-derived VH DJH Sequences—Somatic hypermutation of the VH region is known to be an important event in B-cell-derived Ig following antigen stimulation. Mutational analysis demonstrated that all of the VH mu gene sequences showed fewer than 5% mutations, and 73.9% of VH mu gene sequences had fewer than 2% mutations, which should be considered as “unmutated” by definition (38). In contrast, most functional VH gamma sequences were highly mutated. Approximately 90.6% of VH gamma sequences had greater than 5% mutation, which should be considered as “mutated” by definition (38) (Table 2). Unexpectedly, within a set of VHDJH recombinant, the V regions from different cancer cell types showed identical
VH/DJH recombination and junction. The recombination was either identical or differed only by several mutation targets over the VH region. For example, within the VH5-51/D3-9/JH4 recombination set, 19 of the VH5-51 sequences showed almost identical mutation targets, which were identical to our previously published VH5-51/D3-9/JH4 recombination sequence described for lung cancer cells (Fig. 3) (20). Despite the high mutation rate, the rate of homology within the set of VHDJH recombinants ranged from 92.7 to 100%.

In B-cell-derived Ig, the mutations may either be silent or missense that changes the affinity of the Ig for the antigen; the latter may occasionally give rise to cells expressing higher affinity antibodies, usually with mutations clustered in the CDRs and as a result of antigen selection. Moreover, mutations introduced by AID activity under antigen selection typically target the known SHM hotspot, WRCY, and its complement, RGYW. Furthermore, the mutation frequency is expected to be higher in the CDRs than in the framework regions (FWRs). However, no AID transcript was detected by RT-PCR in the LCM-isolated cancer cells. In contrast, we detected an AID transcript in HT-29 cell line at a low level (GenbankTM Accession Number, AY748364), as well as in Raji cell line (B-lymphocyte-derived, Burkitt’s lymphoma) which was used as a positive control (data not shown). In addition, in only 17 of 43 (39.5%) VH/H9253 sequences were there more frequent mutations in the RGYW hot spot. In 3 of 43 VH/H9253 sequences, the mutation frequency was higher in the CDRs than in the FWRs. In only 7 of 43 (16.3%) VH/H9253 sequences were the R/S ratios in the CDRII and FWRIII/H11022 2.9 and /H11021 1.5, respectively. In the VH5-51/D3-9/JH4 recombination set, a higher mutation frequency, /H11011 40%, occurred in the RGYW motif (if the ratio was under 25.6%, it was not considered classical SHM) (Fig. 3). In contrast, the mutations in cancer-derived Ig mainly occurred in the FWRII and not in the CDRs. This finding suggests that a large number of the mutations introduced into the cancer-derived VH/H9253 sequences were not typical of B-cell-derived VH mutations caused by antigen selection.

To define the role of the error-prone polymerase in cancer-derived VH/H9253 hypermutation, we analyzed the mutation frequency in the WA and TW motifs, and found that mutation of the WA motif (36/43 VH/H9253 sequences) was significantly more frequent than that of the TW motif (11/43 VH/H9253 sequences, p < 0.001). These data reveal a strand bias and suggest that error-prone polymerases are involved in the mutation of cancer-derived VH/H9253 sequences. The mutational frequency in the WRCY/RYGW

| TABLE 1 |
| Rate of functional rearrangement in 89 V/H,DJH recombination from different cancer types |
| Sample origin | Ig type | No. of sample | No. of clone | No. of functional VHDJH recombination | No. of non-functional VHDJH recombination |
|----------------|---------|--------------|-------------|-------------------------------------|----------------------------------------|
| PBL chain 2   | γ chain | 2            | 6           | 5                                   | 1                                      |
| PBL chain 2   | μ chain | 2            | 4           | 4                                   | 0                                      |
| Total         |         | 4            | 10          | 9                                   | 1 (10%) |
| Breast cancer | γ chain | 3            | 12          | 10                                  | 2                                      |
| Breast cancer | μ chain | 3            | 10          | 10                                  | 0                                      |
| Lung cancer   | γ chain | 1            | 6           | 6                                   | 0                                      |
| Lung cancer   | μ chain | 1            | 2           | 2                                   | 0                                      |
| Colon cancer  | γ chain | 2            | 11          | 10                                  | 1                                      |
| Colon cancer  | μ chain | 1            | 5           | 4                                   | 1                                      |
| Oral carcinoma| γ chain | 2            | 12          | 10                                  | 2                                      |
| Oral carcinoma| μ chain | 1            | 9           | 7                                   | 2                                      |
| HT29 cell     | γ chain | 1            | 7           | 7                                   | 0                                      |
| Hela S3 cell  | γ chain | 1            | 6           | 5                                   | 1                                      |
| Hela S3 cell  | μ chain | 1            | 6           | 4                                   | 2                                      |
| Total (cancer)|         | 18           | 89          | 78                                  | 11 (12%) |

*a The percentage of all sequences corresponding to each group of genes is shown in parentheses.

FIGURE 2. VH (A) and JH (B) gene family usage profiles of cancer-derived Ig compared with the normal PBL-derived Ig. We analyzed the usage of 78 functional VH genes from the cancer-derived Ig and collected five groups of usage profiles from healthy donors PBL-derived Ig from previous studies (22, 23, 25, 26, 33). The VH and JH gene family usage data from the combined patient group differed significantly from that of the healthy donors PBL-derived Ig (p < 0.05 for both VH and JH).

13614 JOURNAL OF BIOLOGICAL CHEMISTRY VOLUME 284 • NUMBER 20 • MAY 15, 2009
region, the WA/TW motifs, and calculations of the R/S mutation ratio in CDRII and FWRIII are summarized in Table 4.

Cancer-derived $V_{H}D_{J_{H}}$ Sequences Did Not Appear to Originate from the Classical Class Switching—In general, the B-cell-derived $V_{H}D_{J_{H}}$ recombination pattern of the $\gamma$ chain is similar to that of the $\mu$ chain (the precursor of $\gamma$ chain), because class switching changed only the constant region sequence from the $\mu$ chain to the $\gamma$ chain. However, there was no identical pattern between $V_{H}D_{J_{H}}$ and $V_{H}D_{J_{H}}$ in any of the cases studied, although each sample expressed a restricted $V_{H}D_{J_{H}}$ recombination pattern (Table 2). This suggests that cancer-derived $V_{H}D_{J_{H}}$ sequences do not originate from the classical class switching.

DISCUSSION

B-lymphocytes have been considered the primary source of serum Ig. However, we have found that cells from epithelial cancer and hyperplasia could also express Ig (17–18, 21). In this study, we confirmed that functional Ig gene recombination and transcription occurred in a variety of cancer types. Our analysis of 78 functional cancer-derived $V_{H}D_{J_{H}}$ sequences showed that cancer-derived Ig shared some features with B-cell-derived Ig, such as $V_{H}D_{J_{H}}$ recombination, insertion of the N region into junctions, and $J_{H}$ gene recombination at TG nucleotide sequences. In addition, there were either low levels or a complete absence of mutations in the $V_{H}D_{J_{H}}$ sequences and $J_{H}$ sequences.

On the other hand, cancer-derived $V_{H}D_{J_{H}}$ recombinations also exhibited distinct features that differed from B-cell-derived Ig recombination. For example, cancer-derived Ig $V_{H}D_{J_{H}}$ and $J_{H}$ usage showed distinct preferences, such as that $V_{H}5-51$ and $V_{H}6-1$ frequencies were higher and that $J_{H}4$ was expressed significantly higher than expected compared with normal B-cell-derived $V_{H}D_{J_{H}}$ recombinations; however, we found that $V_{H}5-51$ and $V_{H}6-1$ recombination frequency was significantly higher than expected compared with normal B-cell-derived $V_{H}D_{J_{H}}$ (22, 26–28).

It is known that each B-cell expresses a unique $V_{H}D_{J_{H}}$ recombination, including a random N region sequence. Therefore, the likelihood of identical junction sequences from two independent B-cell clones occurring in an individual should be lower than 1 in 4 million. Unexpectedly, several restricted

### Table 2

Assignment of the likely matching germline variable region genes to the $V_{H}D_{J_{H}}$ recombinants from different cancer types and analysis of the $V$ gene somatic mutation rate

| Cases | No. of clones | $V_{H}D_{J_{H}}$ (no. of clone) | Mutation rate | No. of clones | $V_{H}D_{J_{H}}$ (no. of clone) | Mutation rate |
|-------|--------------|-------------------------------|---------------|--------------|-------------------------------|---------------|
| Breast cancer (case 1) | 5 | VH7-51/D3-9/JH4(3) | 15.6% (3) | 3 | VH4-30-2/D3-10/JH4(1) | 3.3% (1) |
| Breast cancer (case 2) | 5 | VH5-51/D3-9/JH4(4) | 15.6% (2); 17.5% (1) | 3 | VH4-30-2/D3-12/JH4(1) | 4.4% (1) |
| Breast cancer (case 3) | 2 | VH5-51/D3-9/JH4(1) | 15.6% (1) | 4 | VH4-30-2/D3-22/JH4(1) | 3.3% (1) |
| Lung cancer | 6 | VH5-51/D3-9/JH4(2) | 15.6% (2) | 2 | VH6-1/D6-13/JH4(1) | 3.3% (1) |
| Colon cancer (case 1) | 4 | VH5-51/D3-9/JH4(1) | 15.6% (4) | 5 | VH3-15/D3-10/JH4(1) | 2.5% (2) |
| Colon cancer (case 2) | 7 | VH5-51/D3-9/JH4(1) | 15.6% (3); 17.5% (1) | nd |
| Oral cancer (case 1) | 3 | VH5-51/D3-9/JH4(2) | 15.6% (1), stop (1) | 9 | VH6-1/D6-13/JH4(4) | 0 (4) |
| Oral cancer (case 2) | 9 | VH3-33/D6-19/JH5(5) | 11.5% (5) | nd |
| HT-29 | 7 | VH3-33/D6-19/JH5(5) | 16.3% (1); 7.3% (3) | 6 | VH6-1/D6-23/JH4(6) | 3.3% (4); 1% (1); 7.6% (1) |
| HelaS3 | 6 | VH3-33/D6-19/JH5(5) | 8.1% (1) | 3 | VH3-30-3/D6-19/JH5(3) | 1.5% (3) |
| PBL (case 1) | 3 | VH1-8/D1-14/JH4(1) | 6.2% (1) | 4 | VH3-21/D1-26/JH4(1) | 0 (1) |
| PBL (case 2) | 3 | VH1-8/D1-14/JH4(1) | 6.2% (1) | nd |
| | | VH1-8/D1-14/JH4(1) | 10.4% (1) | nd |
| | | VH1-8/D1-14/JH4(1) | 7.5% (1) | nd |
| | | VH1-8/D1-14/JH4(1) | 9.1% (1) | nd |
| | | VH1-8/D1-14/JH4(1) | 4.8% (1) | nd |
| | | VH1-8/D1-14/JH4(1) | 11.5% (1) | nd |
| | | VH1-8/D1-14/JH4(1) | 12.5% (1) | nd |
| | | VH1-8/D1-14/JH4(1) | 20% (1) | nd |
| | | VH1-8/D1-14/JH4(1) | 17.9% (1) | nd |
| | | VH1-8/D1-14/JH4(1) | 3.1% (1) | nd |

* Stop, non-functional $V_{H}D_{J_{H}}$ sequences.

* nd, not done.
**TABLE 3**

| No. of sets | Clone | V gene name | V region | N1 | P | D region | J region | J gene name | D gene name |
|-------------|-------|-------------|----------|----|---|----------|----------|-------------|-------------|
| 1           | BG2 (4) | IGHV5-51*01 | tgtgcgaga tgggatg . . . . . . . . . . . ttatgattggattttata.. cagc . . . . . . . tgactactgg | IGHJ4*02 | IGHD3-9*01 |
| 1           | CG1 (4) | IGHV5-51*01 | tgtgcgaga tgggatg . . . . . . . . . . . ttatgattggattttata.. cagc . . . . . . . tgactactgg | IGHJ4*02 | IGHD3-9*01 |
| 1           | LG1 (1) | IGHV5-51*01 | tgtgcgaga tggaatg . . . . . . . . . . . ttatgattggattttata.. cagc . . . . . . . tgactactgg | IGHJ4*02 | IGHD3-9*01 |
| 1           | BM1 (1) | IGHV4-30-2*01 | tgtgccgg . . . cc . . . . . . . . . . . . ..gaagtggttattact c .cccctttgactactgg | IGHJ4*02 | IGHD3-22*01 |
| 1           | BM21 (1) | IGHV3-15*01 | tgtaccacaaa cctga ac gtattactatggttcggggaccga . . . accccc . . . . . . . gactactgg | IGHJ4*02 | IGHD3-10*01 |
| 1           | CM1 (4) | IGHV3-15*01 | tgtaccacaaa cctga ac gtattactatggttcggggaccga . . . accccc . . . . . . . gactactgg | IGHJ4*02 | IGHD3-10*01 |
| 1           | LG1 (2) | IGHV3-30*01 | tgtgccagaga . . . ..aacaatggctg. . . . cctc . . . . ctttgacaattgg | IGHJ4*03 | IGHD6-19*01 |
| 1           | LG1 (2) | IGHV3-30*01 | tgtgccagaga . . . ..aacaatggctg. . . . cctc . . . . ctttgacaattgg | IGHJ4*03 | IGHD6-19*01 |
| 1           | LG1 (2) | IGHV3-30*1 | tgtgccagaga . . . ..aacaatggctg. . . . cctc . . . . ctttgacaattgg | IGHJ4*02 | IGHD6-19*01 |
| 2           | OM1 (2) | IGHV6-1*01 | tgtgcaagaga at . . . . .tagcagcagct. . . . . . . . . .tttgactactgg | IGHJ4*02 | IGHD6-13*01 |
| 2           | OM1 (2) | IGHV6-1*01 | tgtgcaagaga at . . . . .tagcagcagct. . . . . . . . . .tttgactactgg | IGHJ4*02 | IGHD6-13*01 |
| 2           | OM1 (2) | IGHV6-1*01 | tgtgcaagaga at . . . . .tagcagcagct. . . . . . . . . .tttgactactgg | IGHJ4*02 | IGHD6-13*01 |
| 3           | OM1 (3) | IGHV6-1*01 | tgtgcaagag. ggcgtacgtgg. . . . .. ccagc . . . . .... ..ggacgtctgg | IGHJ6*02 | IGHD1-7*01 |
| 3           | OM1 (3) | IGHV6-1*01 | tgtgcaagag. ggcgtacgtgg. . . . .. ccagc . . . . .... ..ggacgtctgg | IGHJ6*02 | IGHD1-7*01 |
| 4           | CM1 (4) | IGHV3-15*01 | tgtaccacaaa cctga ac gtattactatggttcggggaccga . . . accccc . . . . . . . gactactgg | IGHJ4*02 | IGHD3-10*01 |
| 4           | BM2 (1) | IGHV3-33*01 | tgtgcgagaga . . . . .tggggtggctacgattac aaggaggtg . . . . ctggttcgacccctgg | IGHJ5*02 | IGHD5-12*01 |

V<sub>H</sub>DI<sub>H</sub> sequences were identified in each cancer sample. More interestingly, different cancer samples showed identical V<sub>H</sub>DI<sub>H</sub> recombination patterns, with identical junctions and V<sub>H</sub> region mutation targets. Several dominant V<sub>H</sub>DI<sub>H</sub> recombination sets were frequently expressed in different cancer types. These results suggest that there is an unknown mechanism allowing epithelial cancer cells to express several repeated sets of dominant V<sub>H</sub>DI<sub>H</sub> sequences. We eliminated the possibility of cross-contamination among cancer samples in a number of ways. First, using RT-PCR, the control tube (containing no cDNA) did not show a positive band when the V<sub>H</sub>DI<sub>H</sub> sequences were amplified. Second, there were no samples with identical Ig heavy chain gene repertoires among the eight cancer samples, suggesting that there was no cross contamination among different samples. In contrast, both HT-29 and HeLa S3 cells showed monoclonal characteristics in the V<sub>H</sub>D<sub>J</sub>H<sub>J</sub> derived from these two cancer cell lines showed distinct diversity, which was similar to that of B-cells, but different from primary cancer cells. Identical V<sub>H</sub>DI<sub>H</sub> sequences were not detected between the two cancer cell lines, which implied that the genetic characteristics of the cancer cell lines had changed under long-term culture *in vitro*.

The mechanism of SHM is another distinct feature in cancer-derived V<sub>H</sub>D<sub>J</sub>H<sub>J</sub> sequences as opposed to B-cell-derived V<sub>H</sub>D<sub>J</sub>H<sub>J</sub> sequences. In B-cell-derived Ig V regions, mutations induced by antigen selection occurred more frequently in the CDRs than in the FWRs. Moreover, there is a higher R/S mutation ratio in the CDRs. However, only a few cancer-derived V<sub>H</sub>D<sub>J</sub>H<sub>J</sub> sequences matched the pattern of antigen selection. In addition, AID is thought to be necessary for SHM of the Ig gene, since it can deaminate C to U on both DNA strands, resulting in the symmetrical mutation of C on both strands. The AID enzyme site preference results in hypermutation of the RGYW hotspot motif. In this study, only the V<sub>H</sub>5-51/D<sub>3</sub>3-9/JH4 pattern matched the AID-induced RGYW hypermutation pattern. Moreover, mutations in V<sub>H</sub>5-51 frequently occurred in the FWRIII, but not in the CDRs. Additionally, no AID transcript was detected by RT-PCR in the LCM-isolated cancer cells. In contrast, we detected AID transcripts in the HT-29 cell line at a low level, as well as in Raji cell line. It is possible that AID expression is unnecessary for V<sub>Y</sub> SHM in cancer cells and that other mechanisms may be involved in SHM of the cancer-derived Ig V region. Babbage et al. recently demonstrated constitutive expression of AID in six breast cancer cell lines and Matsumoto et al. demonstrated expression of AID in stomach cancer cells (22, 39). The mutational bias of A versus T in Ig genes results from DNA pol η activity (40), which functions as a secondary mutator (41, 42). Thus, the excess of mutations in the WA motif suggests that DNA pol η may be involved in mutating A:T base pairs in the Ig V gene in cancer cells.

The third distinct feature that differentiates cancer-derived from B-cell-derived V<sub>H</sub>D<sub>J</sub>H<sub>J</sub> sequences is that the cancer-derived V<sub>H</sub>D<sub>J</sub>H<sub>J</sub> sequences do not seem to originate from classical class switching. Class switch recombination and SHM are two important events in Ig production. In the classical Ig class switching, the IgM-producing cells are precursors of the IgG producers, and the V<sub>Y</sub> gene assembly mode should be similar to that of V<sub>H</sub>D<sub>J</sub>H<sub>J</sub>. Although
and VH/H9253 D/H9253 JH/H9253 could be synchronously detected in the same cancer cells in this study, we did not detect any identical patterns of VH/H9253 D/H9253 JH/H9253 and VH/H9262 D/H9262 JH/H9262 in the cases studied. This unexpected result suggested that the IgGs were completely different from the IgMs in the same cancer cells, and that IgG production in these cells did not follow the classical class switching mechanism. These results raised the possibility that IgGs might be coded by an allele on another chromosome. Our finding (by Southern blot analysis using a JH DNA segment probe) that two Ig alleles in HT-29 had been rearranged (data not shown) supported the presence of such a mechanism. Alternatively, there may be another class switching mechanism. Jhagvaral et al. (36) recently described that naive B-lymphocytes could develop into IgG-secreting cells through successive cell divisions.

In this study, a pivotal precaution was to avoid B-cell contamination. B-lymphocytes and plasma cells are capable of infiltrating cancer tissues and are mainly located in the stroma. By using the LCM method, we specifically captured the EpCAM+ cancer cells in the cancer nest regions and avoided capturing tumor-infiltrating B-lymphocytes and plasma cells. We did not detect B-cell contamination in any of the cDNA libraries from the LCM-isolated cancer cells. Importantly, none of the cancer-derived VH/D/J sequences were homologous to the VH/D/J recombination sequences of the two control PBL samples or published recombination sequences present in B-lymphocytes and B-cell CLL/SLL. Additionally, there was no sequence homology with published sequences for tumor-infiltrating B-lymphocytes in breast cancer tissues (25). This lack of recombinant sequence homology with B-lymphocytes indicates that the detected Ig sequences are specific to epithelial cancer cells.

To be reactive to multiple antigens, B-cells generate Ig diversity through several mechanisms. However, the biological significance of non-B-cell-derived Ig is not yet clear. The cancer cell-derived Ig repertoire displayed distinct features suggesting that non-B-cell-derived Ig may have important undiscovered activities because of their diverse origins. We previously noted that Ig expression and activity in cancer cell lines could be blocked by specific antisense DNAs and antibodies, causing the cancer cells to undergo apoptosis (20). These data suggest that non-B-cell-derived Ig is involved in the growth and survival of cells.

In summary, many nonhematopoietic tumor cells express Ig. The cancer-derived Ig gene repertoire displays several distinct characteristics, suggesting that there is an idiosyncratic mechanism for cancer-derived Ig gene expression. The gene expression patterns of Ig in different cancer cells may prove useful in...
Ig Gene Recombination in Human Cancers

| Clone | Mutation frequency in CDRII | Mutation frequency in WA/TW | R/S ratio in CDRII and FWRIII |
|-------|-----------------------------|-----------------------------|-------------------------------|
|       | HG1-1 | 12.5 (±25) | 0 | 0/1 (>-2.9) | 9/4 |
|       | HG1-2 | 12.5 (±25) | 0 | 0/1 (>-2.9) | 9/4 |
|       | HG1-3 | 12.5 (±25) | 0 | 0/1 (>-2.9) | 9/4 |
|       | HG1-4 | 12.5 (±25) | 0 | 0/1 (>-2.9) | 9/4 |
|       | HG1-5 | 12.5 (±25) | 0 | 0/1 (>-2.9) | 9/4 |
|       | HG1-6 | 12.5 (±25) | 0 | 0/1 (>-2.9) | 9/4 |
|       | HG1-7 | 12.5 (±25) | 0 | 0/1 (>-2.9) | 9/4 |

understanding the structure and function of nonhematopoietic-derived Ig. In addition, these findings may have important implications for the diagnosis, targeted therapy, as well as monitoring of residual disease of cancers.

Acknowledgments—We thank Dr. Dalong Ma, Dr. Wenling Han, and Shuang Shi from the Peking University Center for Human Disease Genomics for help with experiments; we thank Dr. Lieping Chen from the Dept. of Dermatology/Oncology, Johns Hopkins University School of Medicine, Dr. Yongguang Yang from Surgery, Harvard Medical School, Dr. Jiang Gu, and Dr. Michael A. McNutt from the Dept. of Pathology, Peking University, Dr. Yu Zhang and Dr. Yanhui Yin from the Dept. of Immunology for contributions to manuscript revision. We also thank I. B. Rogozin from the National Institutes of Health/NLM/NCBI for contributions to SHM analysis.

REFERENCES
1. Jerne, N. K. (1955) Proc. Natl. Acad. Sci. U. S. A. 41, 849–857
2. Glick, B., Chang, T. S., and Jaap, G. (1996) Poultry Sci. 35, 224–225
3. Mitchell, G. F., and Miller, J. F. (1968) J. Exp. Med. 128, 821–837
4. Tonegawa, S. (1976) Proc. Natl. Acad. Sci. U. S. A. 73, 203–207
5. Hozumi, N., and Tonegawa, S. (1976) Proc. Natl. Acad. Sci. U. S. A. 73, 3628–3632
6. Neri, A., Jakobiec, F. A., Pellicci, P. G., Dalla-Favera, R., and Knowles, D. M., 2nd (1987) Blood 70, 1519–1529
7. Cleary, M. L., Chao, J., Warnke, R., and Sklar, J. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 593–597
8. Fraternali-Orcioni, G., Falini, B., Quaini, F., Campo, E., Piccioli, M., Gambri, B., Pasquinelli, G., Poggi, S., Asceni, S., Sabattini, E., and Pileri, S. A. (1999) Ann. I. Surg. Pathol. 23, 717–721
9. Bartram, C. R., Raghavachar, A., and Heimpel, H. (1986) Blut 52, 203–210
10. Zuniga, M. C., D’Eustachio, and Puddie, R. H. (1985) Proc. J. Med. Sci. 22, 57–62
11. O’Brien, P. M., Tsirimonaki, E., Coomber, D. W., Millan, D. W., Davis, U. S. A. (1998) Int. J. Biochem. Cell Biol. 143, 83–93
12. Manjula, S., Aroor, A. R., Raja, A., Rao, S. N., and Rao, A. (1992) Acta Neurochir. (Wien) 115, 103–105
13. Taylor, D. D., and Gercel-Taylor, C. (1998) Oncol. Rep. 5, 1519–1524
14. Tenenbaum, N., Meignant, S., and Vincent, J. P. (1992) Ann. Med. Interne (Paris) 3015–3019
15. Hori, H., Kihara, Y., Nagashio, Y., Hirohata, Y., Abe, S., Murata, I., and Otsuki, M. (2000) Nippon Shokakibyo Gakkai Zasshi 97, 1373–1377
16. Gregersten, H., Mellekendea, I., Isen, J. S., Dahlber, J. F., Thomassen, L., and Sorensvn, H. T. (2001) Haematologica 86, 1172–1179
17. Qiu, X., and Yang, G. (1996) Chinese J. Immunol. 5, 296
18. Wang, D., Qiu, X., Zhu, X., Lu, P., and Gao, X. (2000) J. Med. Biomed. 32, 310
19. Kimoto, Y. (1998) Genes Chromosomes Cancer 22, 83–86
20. Qiu, X., Zhu, X., Zhang, L., Mao, Y., Zhang, J., Hao, P., Li, G., Lu, P., Li, Z., Sun, X., Wu, L., Zheng, J., Deng, Y., Hou, C., Tang, P., Zhang, S., and Zhang, Y. (2003) Cancer Res. 63, 6488–6495
21. Li, M., Feng, D. Y., Ren, W., Zheng, L., Zhang, H., Tang, M., and Cao, Y. (2004) Int. J. Biochem. Cell Biol. 36, 2250–2257
22. Babbage, G., Osttensmeier, C. H., Blydes, J., Stevenson, F. K., and Sahota, S. S. (2006) Cancer Res. 66, 3996–4000
23. Zhu, X., Li, C., Sun, X., Mao, Y., Li, G., Liu, X., Zhang, Y., and Qiu, X. (2008) Appl. Immunohistochem. Mol. Morphol. 16, 232–238
24. Lefranc, M. P., Giudicelli, V., Ginestoux, C., Bodmer, J., Muller, W., Bontrop, R., Lemaire, M., Malik, A., Barbier, V., and Chaume, D. (1999) Nucl. Acids Res. 27, 209–212
25. Nzula, S., Going, J., and Stott, D. I. (2003) Cancer Res. 63, 3275–3280
26. Messner, B. T., Albesiano, E., Messner, D., and Chiorazzi, N. (2004) Blood 103, 3490–3495
27. Hansen, M. H., Nielsen, H., and Ditzel, H. J. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 12659–12664
28. O’Brien, P. M., Tsirimonaki, E., Coomber, D. W., Millan, D. W., Davis, J. A., and Campo, M. S. (2001) Cancer Immunol. Immunother. 50, 523–532
29. Rogozin, I. B., and Kolchanov, N. A. (1992) Biochem. Biophys. Acta 1171, 11–18
30. Yoshikawa, K., Okazaki, I. M., Eto, T., Kinoshita, K., Muramatsu, M., Nakagawa, H., and Honjo, T. (2002) Science 296, 2033–2036
31. Martin, A., Bardwell, P. D., Woo, C. J., Fan, M., Shulman, M. J., and Scharff, M. D. (2002) Nature 415, 802–806
32. Tamaru, I., Hummel, M., Marafioti, T., Kalvelage, B., Leoni, L., Minacci, C., Tosi, P., Wright, D., and Stein, H. (1995) Ann. I. Pathol. 147, 1398–1407
33. Dorner, T., Brezinschek, H. P., Brezinschek, R. I., Foster, S. J., Domiati-Saad, R., and Lipsky, P. E. (1997) J. Immunol. 158, 2779–2789
34. Spencer, J., Dunn, M., and Dunn-Walters, D. K. (1999) J. Immunol. 162, 6596–6601
35. Oprea, M., Cowell, L. G., and Kepler, T. B. (2001) J. Immunol. 166,
892–899
36. Souto-Carneiro, M. M., Longo, N. S., Russ, D. E., Sun, H. W., and Lipsky, P. E. (2004) *J. Immunol.* **172**, 6790–6802
37. Brezinschek, H. P., Brezinschek, R. I., and Lipsky, P. E. (1995) *J. Immunol.* **155**, 190–202
38. Yamada, M., Wasserman, R., Reichard, B. A., Shane, S., Caton, A. J., and Rovera, G. (1991) *J. Exp. Med.* **173**, 395–407
39. Matsumoto, Y., Marusawa, H., Kinoshita, K., Endo, Y., Kou, T., Morisawa, T., Azuma, T., Okazaki, I. M., Honjo, T., and Chiba, T. (2007) *Nat. Med.* **13**, 470–476
40. Mayorov, V. I., Rogozin, I. B., Adkison, L. R., and Gearhart, P. J. (2005) *J. Immunol.* **174**, 7781–7786
41. Neuberger, M. S., Di Noia, J. M., Beale, R. C., Williams, G. T., Yang, Z., and Rada, C. (2005) *Nat. Rev. Immunol.* **5**, 171–178
42. Delbos, F., De Smet, A., Faili, A., Aoufouchi, S., Weill, J. C., and Reynaud, C. A. (2005) *J. Exp. Med.* **201**, 1191–1196