JOURNAL OF VIROLOGY, Jan. 1988, p. 75–83
0022-538X/88/010075-09$02.00/0

Vol. 62, No. 1

The Avian Retrovirus *env* Gene Family: Molecular Analysis of Host Range and Antigenic Variants

CAROL A. BOVA, JOHN C. OLSEN, AND RONALD SWANSTROM*

Department of Biochemistry and Lineberger Cancer Research Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

Received 14 May 1987/Accepted 7 October 1987

The nucleotide sequence of the *env* gp85-coding domain from two avian sarcoma and leukosis retrovirus isolates was determined to identify host range and antigenic determinants. The predicted amino acid sequence of gp85 from a subgroup D virus isolate of the Schmidt-Ruppin strain of Rous sarcoma virus was compared with the previously reported sequences of subgroup A, B, C, and E avian sarcoma and leukosis retroviruses. Subgroup D viruses are closely related to the subgroup B viruses but have an extended host range that includes the ability to penetrate certain mammalian cells. There are 27 amino acid differences shared between the subgroup D sequence and three subgroup B sequences. At 16 of these sites, the subgroup D sequence is identical to the sequence of one or more of the other subgroup viruses (A, C, and E). The remaining 11 sites are specific to subgroup D and show some clustering in the two large variable regions that are thought to be major determinants of host range. Biological analysis of recombinant viruses containing a dominant selectable marker confirmed the role of the gp85-coding domain in determining the host range of the subgroup D virus in the infection of mammalian cells. We also compared the sequence of the gp85-coding domain from two subgroup A viruses, Rous-associated virus type 1 and a subgroup A virus of the Schmidt-Ruppin strain of Rous sarcoma virus. The comparison revealed 24 nonconservative amino acid changes, of which 6 result in changes in potential glycosylation sites. The positions of 10 amino acid differences are coincident with the positions of 10 differences found between two subgroup B virus *env* gene sequences. These 10 sites identify seven domains in the sequence which may constitute determinants of type-specific antigenicity. Using a molecular recombinant, we demonstrated that type-specific neutralization of two subgroup A viruses was associated with the gp85-coding domain of the virus.

The envelope glycoproteins on the surface of a retrovirus particle specify the major determinants of host cell recognition and penetration (59). The *env* gene of avian retroviruses encodes two proteins, gp85 and gp37, which are synthesized as a single precursor polypeptide, processed, and transported to the cell membrane, where they remain linked by disulfide bonds (9, 21, 27, 29, 35, 36, 39).

Individual isolates of the avian sarcoma and leukosis virus (ASLV) group interact with different host cell components, presumably specific host cell receptors, at an early stage in infection. The presence of polymorphism in these determinants within the host population gives rise to a host range phenomenon in which different virus isolates display a characteristic pattern of infectivity for cells derived from inbred lines of chickens (57). This pattern of host range is one criterion used to divide the various virus isolates into subgroups. The viral determinants that give rise to the subgroup phenotype lie within the *env* gene (11, 31, 58). Comparison of the nucleotide sequence of the gp85-coding domain from viruses of different subgroups has revealed specific regions of sequence variability (8, 17, 45). Recently, analysis of molecular recombinants has shown that the major determinants of host range lie within the gp85-coding domain (8, 16, 17).

Three dominant autosomal loci, *TV-A*, *TV-B*, and *TV-C*, have been identified in chickens as the determinants of host susceptibility to subgroup A, B, and C viruses, respectively (13–15, 41). Chicken cells resistant to subgroup B ASLV are equally resistant to subgroup E viruses and partially resistant to subgroup D viruses (57). In addition, subgroup B viruses interfere to various degrees with infection by subgroup D and E viruses (2, 19, 26). These observations suggest that subgroup B, D, and E viruses use the same receptor or related receptors when infecting chicken cells.

Subgroup D viruses are phenotypically distinct from subgroup B viruses in two ways. (i) Most inbred chicken lines show some level of susceptibility to subgroup D viruses even though certain lines are highly resistant to subgroup B viruses (19). (ii) Subgroup D viruses are able to penetrate a variety of mammalian cells (6, 19). One manifestation of this mammal tropism is the ability of subgroup D strains of Rous sarcoma virus (RSV) to produce tumors when injected into certain mammals (19, 52).

Another criterion used to define the subgroup phenotype is antigenicity. Immune chicken sera show specificity in reacting with viruses from the same subgroup (30). Frequently, these sera show even greater specificity in that they also recognize type-specific antigens present on the particular virus isolate used for immunization (30). Thus, antigenic variation is observed at two levels: between viruses that use different host cell receptors (subgroup specific) and within a group of viruses that use the same receptor (type specific).

In this report, we extend our molecular analysis of the ASLV *env* gene product, gp85, to include a subgroup D virus isolate. We compare the predicted amino acid sequence of Schmidt-Ruppin strain subgroup D (SR-D) RSV with the reported sequences of other subgroup viruses to identify determinants of mammal tropism for the subgroup D virus. We also present the sequence of the gp85-coding domain of the subgroup A virus Rous-associated virus type 1 (RAV-1) and discuss the role of sequence differences in the antigenic variability frequently seen in subgroup A and B viruses.

* Corresponding author.

Finally, we used recombinant viruses to investigate the role of the gp85-coding domain in defining host cell penetration by subgroup D and A viruses in a variety of avian and mammalian cell lines and in defining the site of action of subgroup- and type-specific neutralizing antibodies.

MATERIALS AND METHODS

Cells, viruses, and antisera. Fertile chicken eggs were obtained from the Regional Poultry Research Laboratory, East Lansing, Mich. Chicken embryo fibroblasts (CEF), prepared from 11-day-old embryos, and QT6 cells were maintained as described previously (8). Most mammalian cell lines were maintained in Dulbecco minimum essential medium supplemented with 10% fetal bovine serum. P19 cells were grown in Dulbecco minimum essential medium supplemented with 7.5% calf serum and 2.5% fetal bovine serum, and D98/HR1 cells were grown in RPMI medium containing 10% fetal bovine serum.

Viral genomes were obtained as molecular clones. pSRD-11, an infectious clone of SR-D RSV, was obtained from D. Shalloway (47). Plasmid p779NC327AC28F, which contains a complete copy of the SR-A RSV genome with a unique *Cla*I site in place of the v-*src* insert, was obtained from S. Hughes (28). pRAV10R, an infectious clone of the RAV-1 genome, was obtained from L. Sealy (46). Polyclonal chicken antisera against RAV-1 and a nontransforming derivative of SR-A RSV were a gift from R. Shuman. The antisera were raised in Spafas gs^- chf^- C/O chickens.

Plasmids and bacteria. The molecular recombinant of subgroup A and D viruses was constructed by a strategy similar to that described previously (8). Plasmid pAV-V1, which contains an intact copy of the SR-A RSV genome, was cleaved with restriction enzymes *Kpn*I and *Sal*I to excise a 1.1-kilobase (kb) fragment that spans most of the gp85-coding region. This fragment was replaced with an equivalent fragment from the pSRD-11 clone.

The construction of pANV-A will be described elsewhere (J. C. Olsen, W. P. Osheroff, and R. Swanstrom, manuscript in preparation). This plasmid contains the neomycin resistance gene (neo^r) from pSV2neo (49) cloned into the *Cla*I site of p779NC327AC28F. The resulting plasmid contains the replicative genes of SR-A RSV, with the v-*src* gene replaced by the neo^r gene. pANV-D and pANV-A1 were constructed from pANV-A by removing the *Kpn*I-to-*Sal*I restriction fragment and replacing it with the equivalent fragment from pSRD-11 and pRAVIOR, respectively. A similar strategy was used to construct a neo^r plasmid containing a subgroup E recombinant virus genome, pANV-E (8). All plasmids were propagated in either the HB101 or DH-1 strain of *Escherichia coli*.

DNA transfection and host range analysis. Each bacterial plasmid containing an intact viral genome was transfected into C/O chicken cells by the polybrene-dimethyl sulfoxide protocol of Kawai and Nishizawa (33) to generate virus. The presence of virus in the medium was detected by measuring virus-associated reverse transcriptase activity (8, 54). Host range was tested by two assays. (i) Virus growth in different avian cells was monitored by the appearance of virus-associated reverse transcriptase in the culture medium. (ii) Cells were first plated at a density of 5×10^5 cells per 60-mm (diameter) plate. On the next day, the cells were pretreated with polybrene (55) at 20 μg of polybrene per ml in serum-free medium and then infected for 2 h with virus containing the neo^r gene. At 24 h later, the cells were placed in medium containing Geneticin G418 (GIBCO Laboratories, Grand Island, N.Y.). The concentration of G418 was chosen such that sensitive cells were dead 7 to 10 days after selection was applied. Mock-infected cells were maintained in parallel. The number of G418-resistant colonies was scored 14 days after infection.

Virus neutralization. QT6 cells were plated at a density of 2×10^5 to 5×10^5 cells per 60-mm plate. On the following day, dilutions of chicken antiserum were first incubated with virus for 30 to 40 min at 37°C in 1 ml of medium, followed by 2 h of incubation with the QT6 cells. For avian neo virus subgroup E (ANV-E) infections, QT6 cells were preincubated for 1 h with 10 μg of polybrene per ml. The cells were subjected to selection with 140 μg of G418 after infection, and G418-resistant colonies were scored approximately 2 weeks later.

Immunofluorescence. CEF C/O cells were infected with virus stocks of ANV-A, ANV-A1, and ANV-E. Virus production was monitored by analyzing reverse transcriptase activity in the medium of infected cells. Infected cells were grown on glass cover slips, washed with phosphate-buffered saline, and exposed to a 1:10 dilution of chicken antiserum for 30 min to 1 h in serum-free medium. After incubation with antibody, the cells were washed and fixed with 3.7% formaldehyde for 15 min. The cells were then exposed to the second antibody, a 1:20 dilution of fluorescein-conjugated rabbit anti-chicken immunoglobulin G (IgG; Cooper Biomedical, Inc., West Chester, Pa.) for 30 min to 1 h. Cover slips were mounted and viewed with a Leitz fluorescence microscope.

DNA sequence analysis. The nucleotide sequence of both SR-D RSV and RAV-1 was obtained by a combination of the chemical cleavage method (37) and the dideoxy chain termination method (1, 44, 61). The complete sequence was determined on both strands of DNA. Sequence analysis was performed with programs obtained through Bionet IntelliGenetics, Palo Alto, Calif. Conservative amino acid changes are defined as changes involving amino acids with similar side chains. The six other sequences that were used in various comparisons have been described previously as indicated: SR-A RSV and RAV-2 (8); RAV-0 (8, 17); Prague strain subgroup B (Pr-B) RSV (17); myeloblastosis-associated virus subgroup B (MAV-B) (32); and Pr-C RSV (45). For simplicity, we have adopted the nomenclature of Dorner and Coffin (16) for naming variable regions identified by comparing sequences from different subgroup *env* genes.

RESULTS

Nucleotide sequence of the gp85-coding domain from SR-D RSV and RAV-1. In Fig. 1, the nucleotide sequences of a portion of the *env* gene of two ASLV isolates, RAV-1 (subgroup A) and SR-D RSV (subgroup D), are compared with the sequences of subgroup B (RAV-2; 8), subgroup C (Pr-C RSV; 45), and subgroup E (RAV-0; 8, 17) isolates. In each case, the sequence begins at or adjacent to a conserved *Kpn*I site. Analysis of these sequence comparisons confirmed previous reports that sequence variability is clustered within four domains (vr1, hr1, hr2, and vr3), with a fifth domain (vr2) showing divergence in the subgroup B sequence only (16). Comparisons of examples of the most divergent sequences (subgroups A, C, and D) indicated that there is approximately 60% nucleotide sequence identity within the variable domains and 90% identity within the flanking conserved domains.

Analysis of the predicted amino acid sequence of the gp85-coding domain from SR-D RSV. The predicted amino acid

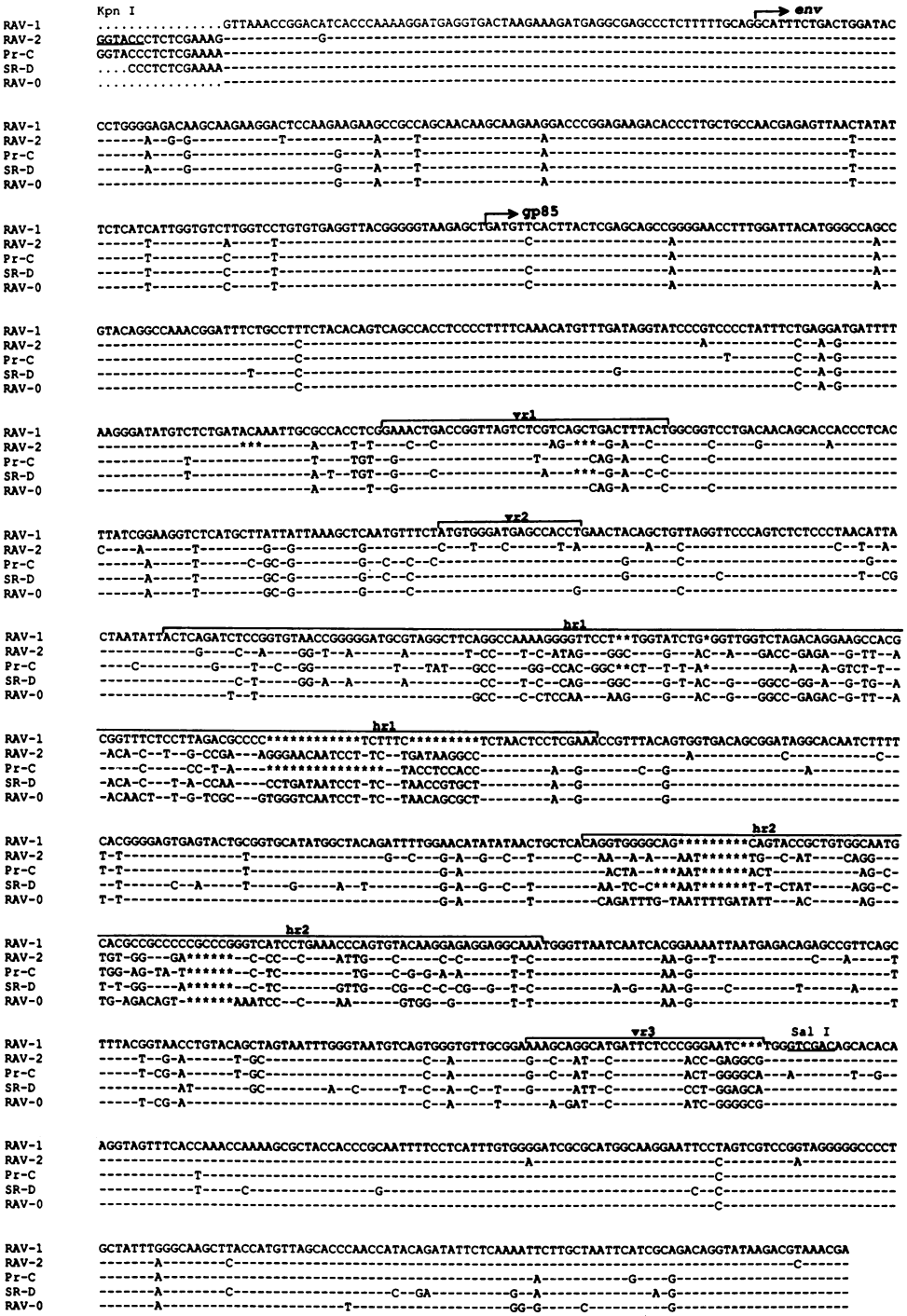

FIG. 1. Nucleotide sequence comparison of five avian sarcoma and leukosis viruses: RAV-1, subgroup A; RAV-2, subgroup B (8); Pr-C RSV, subgroup C (45); SR-D RSV, subgroup D; and RAV-0, subgroup E (8, 17). The entire sequence of RAV-1 is given, along with nucleotide differences of the other virus isolates. The sequence comparison begins at the unique *Kpn*I site corresponding to position 4995 of Pr-C RSV (45). Also shown are the beginning of the *env* gene and the first base encoding the mature gp85 protein (29). The *Kpn*I and *Sal*I restriction sites used to construct recombinant viruses (see the text) are underlined. The variable domains (vr1, vr2, hr1, hr2, vr3) have been described previously (8, 16, 17). Symbols: –, sequence identity with RAV-1; *, gap introduced during sequence alignment; . . ., sequence not completed to the *Kpn*I site.

sequence of the mature gp85-coding domain from SR-D RSV is shown in Fig. 2 in comparison with that of two subgroup B isolates (8, 32). Examination of positions of identity and conservative amino acid changes showed extensive (92%) homology over the entire length of the sequence. There are 15 positions at which B and D subgroup sequences are identical yet different from subgroup A, C, and E sequences (Fig. 2, squares). These positions occur within three of the

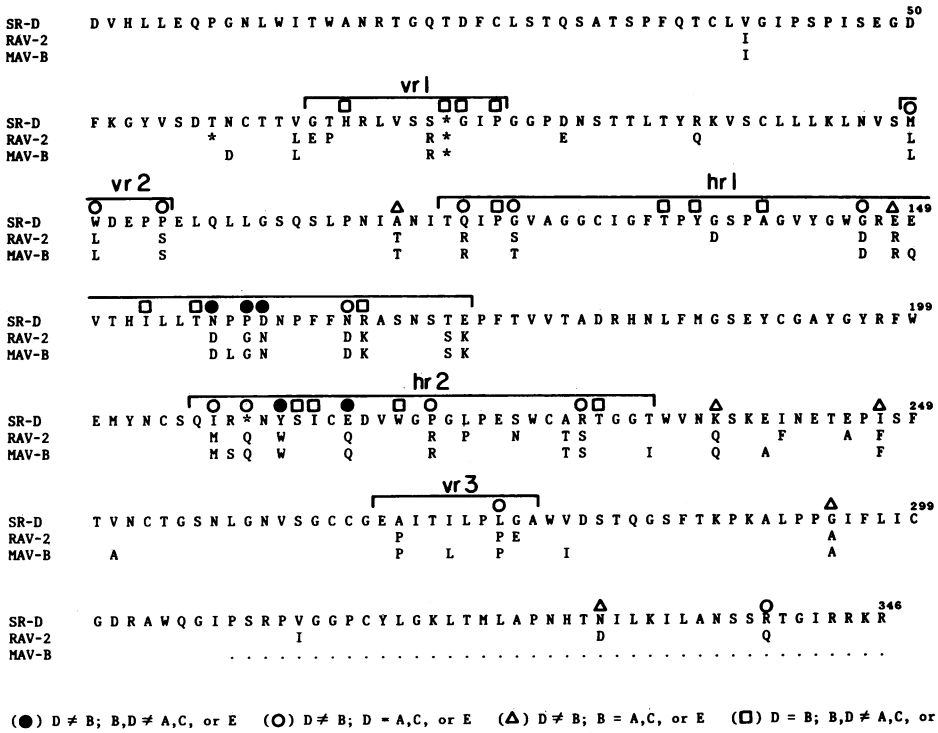

FIG. 2. Comparison of gp85-coding domains of subgroup D and B viruses. The entire predicted amino acid sequence is shown for the subgroup D virus, while only differences from the subgroup D sequence are shown for the subgroup B sequences. The RAV-2 and MAV-B sequences have been described previously (8, 32). The asterisk denotes a gap inserted during sequence alignment. The variable regions identified by comparison of subgroup A, B, C, and E virus *env* genes (vr1, vr2, vr3, hr1, and hr2; 8, 16, 17) are shown above the sequence. The symbols above the sequence represent results of sequence comparisons between SR-D RSV (Fig. 1), RAV-2 (8), MAV-B (32), Pr-B RSV (17), SR-A RSV (8), RAV-1 (Fig. 1), Pr-C RSV (45), and RAV-0 (8, 17). The symbols correspond to the sequence relationships shown at the bottom of the figure. There are five positions of difference between the subgroup D sequence and the two subgroup B sequences not noted with a symbol. The serine at position 71 is also found in the Pr-B RSV sequence, indicating that this site is neither D specific nor B specific. The threonine at position 171 represents a conservative amino acid change from the subgroup B sequence. There are three other sites not included (E172, A227, and A267), since at these sites both the subgroup D and B sequences resemble the sequences of other subgroup viruses.

five variable domains, including the two large variable domains (hr1 and hr2). Since subgroup B and D viruses probably use the same receptor on chicken cells, these B- and D-specific positions may define some of the amino acids that interact with the chicken cell receptor.

Of the 27 nonconservative amino acid differences between the subgroup D and B viruses, 13 are unique to the subgroup B sequence, with the amino acids at those positions in the D sequence identical to those in the subgroup A, C, or E sequence (Fig. 2, open circles). Twelve of the sites are within the variable regions. There are 11 amino acid positions unique to the subgroup D isolate (Fig. 2, closed circles and triangles). At six of these positions (Fig. 2, triangles) the subgroup B sequence is similar to one of the other subgroup sequences. In the remaining five positions (Fig. 2, closed circles), the subgroup B and D sequences differ not only from each other but also from the A, C, and E subgroup sequences (B and D unique). These positions are clustered into two small regions, one in each of the two large variable regions (hr1 and hr2).

Host range analysis of molecular recombinants between subgroup A and D viruses. The host range determinants of subgroup A, B, and E viruses lie within the 1.1-kb *Kpn*I-to-*Sal*I restriction enzyme fragment (Fig. 1) that spans most of the gp85-coding domain (8, 16). To extend these studies, we made a molecular recombinant between the SR-A and SR-D RSV genomes. The 1.1-kb *Kpn*I-to-*Sal*I restriction fragment that includes all of the variable regions was removed from SR-A RSV and replaced with the equivalent fragment from the subgroup D genome (Fig. 1). Virus was generated from the cloned genome after transfection of the cloned DNA onto susceptible cells. The parental and recombinant viruses were then tested for the ability to grow on different avian cell lines as measured by the appearance of reverse transcriptase activity in the medium of infected cells. The parental SR-A RSV and the subgroup D recombinant grew well on C/O CEF cells, which should be susceptible to all subgroup viruses. Both viruses grew on C/E cells, and neither virus grew on C/ABE cells. Only the SR-A RSV parent grew in the continuous quail cell line QT6, consistent with previous work which showed that QT6 cells are resistant to subgroup B and D viruses (40). We concluded that the recombinant with the subgroup D sequence cannot use either the subgroup A or the subgroup C receptor. These results indicate that for the subgroup D virus, as with subgroup A, B, and E viruses, the region of the genome encoding the major determinants of subgroup specificity can be localized to within the 1.1-kb *Kpn*I-to-*Sal*I restriction enzyme fragment.

To assess the infectivity of subgroup D viruses in mammalian cells, we constructed a virus that carries a dominant selectable marker, the *neo*^{r} gene (49). Eucaryotic cells are sensitive to the drug G418. Introduction of *neo*^{r} into these cells confers resistance to G418 (12, 49). We used *neo*^{r} as a marker to analyze the ability of a virus containing this gene to enter a cell and express the gene product. Plasmid pANV-A contains the replicative genes of SR-A RSV. In this

construction, the v-*src* gene of RSV has been replaced with *neo*^{r} under control of the simian virus 40 promoter, and the pBR322 origin of replication (Olsen et al., in preparation). pANV-D was derived from pANV-A by replacement of the 1.1-kb *Kpn*I-to-*Sal*I restriction fragment with the equivalent fragment from SR-D RSV. Transfection of the plasmids into susceptible chicken cells yielded infectious virus (ANV-A and ANV-D). These virus stocks were then used to infect various QT6 cells and various continuous mammalian cell lines.

Table 1 shows a summary of the number of colonies of G418-resistant cells obtained after exposure of the cells to virus-containing medium. No G418-resistant colonies were obtained in mock-infected cultures. As expected, the subgroup A virus readily infected the quail cell line QT6, while the subgroup D recombinant virus did not infect these cells. The opposite result was obtained with the NIH 3T3-derived clonal cell line 45b, which was readily infected by the subgroup D virus but not by the subgroup A virus. Several mouse cell lines, representing a variety of different tissue types, were analyzed. The level of infectivity of the recombinant subgroup D virus (as measured by the number of G418-resistant colonies per milliliter of virus-containing medium) varied greatly, depending on the different cell types, while the subgroup A virus was negative for infection. The primate cell line COS7, which expresses the simian virus 40 T antigen (24), was also readily infected by the recombinant subgroup D virus. Three human cell lines also showed some level of infection by ANV-D. The mammal-tropic determinants of subgroup D virus must, therefore, be largely encoded within the 1.1-kb fragment used to construct the recombinant subgroup D virus. Surprisingly, one of the human cell lines was also infected at a low level with the subgroup A virus. Thus, susceptibility to infection by these recombinant viruses varied greatly among different mammalian cell lines, and infectivity with a subgroup A virus could be detected, albeit at a low level, in one of the lines.

Type-specific antigenicity of two subgroup A viruses. RAV-1 and SR-A RSV are two subgroup A virus isolates that have been shown to produce type-specific neutralizing antibodies in chickens (30). We confirmed this observation by using virus generated from cloned DNA. ANV-A, the virus produced upon transfection of pANV-A into susceptible cells, contains the replicative genes, including the *env* gene, of SR-A RSV. We constructed a RAV-1 recombinant virus, ANV-A1, that contains the *Kpn*I-to-*Sal*I fragment of RAV-1 in an otherwise ANV-A genetic background. Table 2 provides a summary of neutralization experiments with eight different chicken antisera raised against either avian leukosis virus subgroup A (ALV-A; identical to ANV-A but lacking the *neo*^{r} gene) or RAV-1. Each of the four antisera raised against ALV-A (α-ALVA-1 through -4) effectively neutralized ANV-A at a dilution of 10^{-3}. Two anti-ALV-A antisera, α-ALVA-3 and -4, neutralized the RAV-1 recombinant ANV-A1, but only at the highest concentration tested in serum. The antisera raised against RAV-1 (α-RAV1-1 through -4) effectively neutralized the RAV-1 recombinant ANV-A1 at a dilution of 10^{-2} or 10^{-3}. Two of the four anti-RAV-1 antisera, α-RAV1-2 and -4, neutralized ANV-A at a serum dilution of 10^{-1}. None of the antisera neutralized the subgroup E recombinant virus ANV-E. From these results, we conclude that some, if not all, of the determinants that give rise to type-specific and subgroup-specific neutralizing antibody response in chickens reside within the sequences defined by the *Kpn*I-to-*Sal*I restriction fragment that encodes most of the gp85 *env* gene product.

We also investigated cell surface immunofluorescent staining of C/O CEF cells infected with ANV-A, ANV-A1, and ANV-E. Table 2 summarizes the results for each of the chicken antisera. For the anti-ALV-A antisera, good surface fluorescent staining (i.e., good overall staining of cells, with some brightly staining patches) was observed in cells infected with ANV-A. Two of the anti-ALV-A antisera gave good immunofluorescence for ANV-A1-infected cells, while none of the anti-ALV-A antisera stained ANV-E-infected cells. For the anti-RAV-1 antisera, good surface staining was found only for cells infected with ANV-A1 recombinant virus. Three of the anti-RAV-1 antisera stained ANV-A-infected cells weakly (poor overall staining; few patches of bright staining). ANV-E-infected cells did not show significant cell surface staining. Therefore, six of the eight antisera primarily stained cells infected with the homologous virus (i.e., α-ALVA and ANV-A; α-RAV1 and ANV-A1).

Comparison of the gp85-coding domain of RAV-1 and SR-A RSV. The predicted amino acid sequence of gp85 from RAV-1 was compared with the sequence of another subgroup A virus, SR-A RSV (Fig. 3). Of the 24 nonconservative amino acid differences between RAV-1 and SR-A RSV, 6 result in changes in potential glycosylation sites. There are 14 potential N-linked glycosylation sites that are highly conserved between subgroup B, C, D, and E viruses. Three of these sites are missing from the RAV-1 sequence, and a different three are missing from the SR-A RSV sequence. Among subgroup B, C, D, and E viruses, all 14 potential

TABLE 1. Host range of recombinant viruses

Cell line	Origin (tissue source)	Titer[a] of:		Amt (μg/ml) of G418 used for selection	Reference
		ANV-D	ANV-A		
QT6	Quail (fibrosarcoma)	0	3.1×10^3	250	40
NIH 3T3 45b	Mouse (fibroblast)	2.2×10^2	0	200	53
P19	Mouse (teratocarcinoma)	10	NT[b]	100	20
F9	Mouse (teratocarcinoma)	0	0	200	4
G2620	Mouse (glioma)	0	0	600	62
L-cell TK$^-$	Mouse (connective tissue)	0	0	400	34
Neuro-2a	Mouse (neuroblastoma)	82	0	150	Klebe and Ruddle[c]
COS7	Monkey (kidney)	2.1×10^3	0	350	24
D98/HR-1	Human (epithelial cell–B-lymphocyte fusion)	14	5	300	23
A549	Human (lung carcinoma)	9	0	400	22
U-251 MG	Human (glioma)	1.0×10^2	NT	350	5

[a] Virus titer, Number of G418-resistant colonies per milliliter of virus-containing medium.
[b] NT, Not tested.
[c] R. J. Klebe and F. H. Ruddle, J. Cell. Biol. **43**:69, 1969.

TABLE 2. Virus neutralization and cell surface immunofluorescent staining

Serum	ANV-A		ANV-A1		ANV-E	
	Neutralization[a]	Fluorescence[b]	Neutralization	Fluorescence	Neutralization	Fluorescence
α-ALVA-1	3	++	–	–	–	–
α-ALVA-2	3	++	–	–	–	–
α-ALVA-3	3	++	1	++	–	–
α-ALVA-4	3	++	1	++	–	–
α-RAV1-1	–	+	2	++	–	–
α-RAV1-2	1	+	3	++	–	–
α-RAV1-3	–	–	2	++	–	–
α-RAV1-4	1	+	2	++	–	–

[a] Negative logarithm of the greatest antiserum dilution that resulted in more than a 70% reduction of virus titer as measured by the number of G418-resistant colonies of QT6 cells. Each antiserum was tested in duplicate or triplicate for both ANV-A and ANV-A1 virus stocks. ANV-E was tested with antiserum dilutions of 10^{-2} and 10^{-1}. The symbol – indicates that virus neutralization was not found at the lowest antiserum dilution tested (10^{-1}).

[b] Immunofluorescent cell surface staining of C/O CEF infected with ANV-A, ANV-A1, and ANV-E. Intensity of staining was judged as follows: ++, easily detected; +, weak; –, background levels. Primary antibody was used at a 1:10 dilution, and fluorescein-conjugated second antibody was used at a 1:20 dilution.

glycosylation sites are completely conserved, with the exception of the loss of 1 site in the subgroup D sequence and of 1 site in the MAV-B sequence. The only other change in glycosylation is the addition of one site in hr2 in the subgroup B and D sequences. Overall, there appears to be far greater potential for heterogeneity in glycosylation between these two different subgroup A viruses than between either different subgroup B viruses or viruses from different subgroups.

The remaining 18 nonconservative amino changes between RAV-1 and SR-A RSV are scattered over the gp85 sequence. By comparison, there are 16 nonconservative

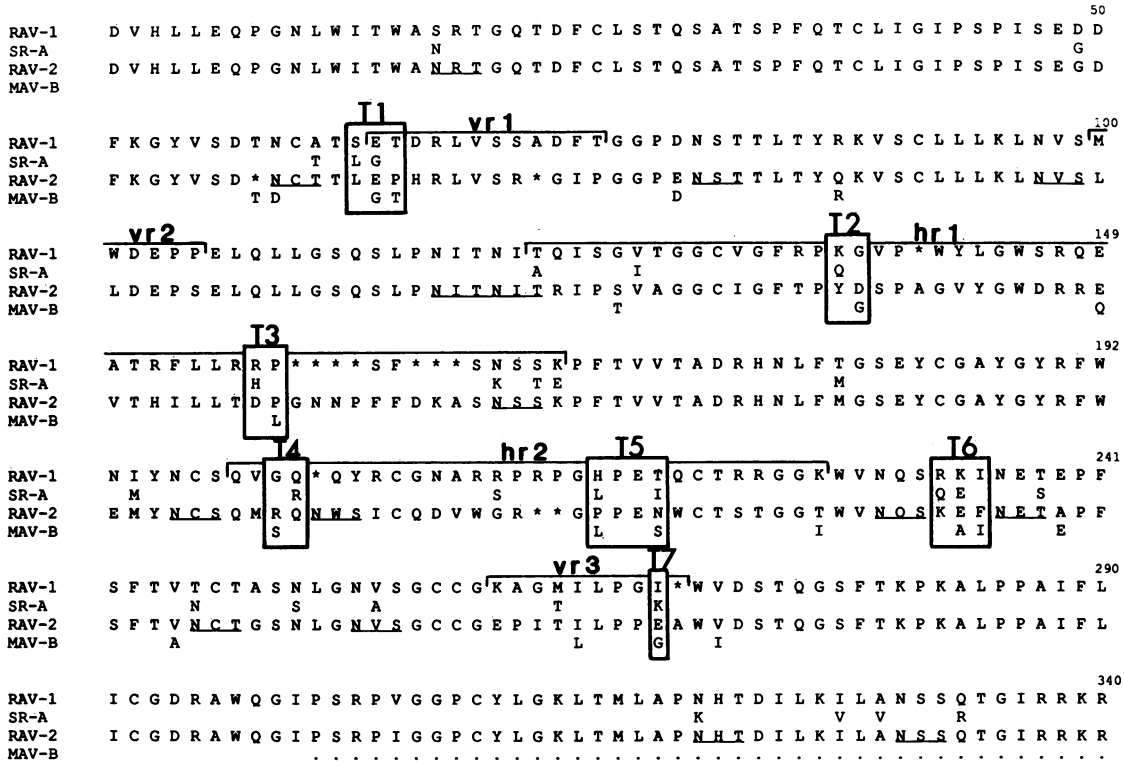

FIG. 3. Comparison of the predicted gp85 sequences from viruses within the same subgroup. The predicted amino acid sequence of the gp85-coding domain of RAV-1 is shown, along with only the differences in the SR-A RSV sequence (8) as compared with the RAV-1 sequence; both are subgroup A viruses. Also shown is the predicted amino acid sequence of RAV-2 (8), as well as the differences in the MAV-B sequence (32) as compared with that of RAV-2; both are subgroup B viruses. Fifteen asparagine-linked potential glycosylation sites are underlined in the RAV-2 sequence. The differences in potential glycosylation sites are as follows: RAV-1, loss of potential glycosylation at sites 1, 2, and 12; SR-A, loss of potential glycosylation at sites 6, 7, and 14; MAV-B, loss of a potential glycosylation site at 2. Potential glycosylation site 9 is specific to subgroup B and D viruses. The boxed regions (T1 through T7) indicate nonconservative amino acid sequence changes between viruses belonging to one subgroup (RAV-1 and SR-A RSV) that cluster at equivalent or nearly equivalent positions of nonconservative amino acid changes between viruses belonging to a second subgroup (RAV-2 and MAV-B). The variable domains described in Fig. 1 are also shown. The MAV-B sequence ends, denoted by . . ., at the equivalent amino acid at position 301 of RAV-1. The amino acid sequence numbering follows only the RAV-1 (and SR-A) sequence.

amino acid differences found between RAV-2 and MAV-B, two subgroup B viruses (Fig. 3). This comparison includes only the first 300 amino acids of the MAV-B gp85-coding domain (32). There appears to be some clustering of sites of nonconservative amino acid changes when different subgroup viruses are compared. When the positions of nonconservative amino acid changes between the two subgroup A viruses are compared with the positions of nonconservative amino acid changes between the two subgroup B viruses, 10 of the differences occur at exactly the same site or in adjacent amino acids (Fig. 3). Thus, many of the differences that distinguish two virus isolates from the same subgroup occur at common sites in two other viruses from a different subgroup. These differences identify seven regions that may be type-specific sequence determinants. These regions are labeled T1 through T7 in Fig. 3.

DISCUSSION

We examined at the molecular level two natural variants of the ASLV retrovirus family. The variants are a virus that has a tropism for mammalian cells (SR-D RSV) and a virus that expresses type-specific antigenicity (RAV-1). Using nucleotide sequence comparison, we searched for sequence differences that may determine the variant phenotype of each virus.

Molecular basis of mammal tropism. Although subgroup D and B viruses can use the same receptor to infect chicken cells, the pattern of infectivity of subgroup D viruses is distinct from that of subgroup B viruses in two ways. (i) Subgroup D viruses are able to infect mammalian cells (19). (ii) They show a low level of infectivity on chicken cells that are resistant to subgroup B viruses (19). It has been shown previously that the *env* gene is a major determinant in mammal tropism (6, 19, 25). We further refined the genetic map for the mammal-tropic determinants by constructing a molecular recombinant between subgroup A and D viruses. The recombinant virus had a mammal-tropic host range, which allows us to place the major mammal-tropic determinants within the left half of the *env* gene, i.e., within the gp85-coding domain.

There is obvious homology between the subgroup D virus sequence and two subgroup B sequences within the gp85-coding domain (Fig. 2). There are only 27 positions at which the sequence of the subgroup D virus is distinct from that of the subgroup B viruses within all of gp85, and in 16 of these sites the sequence of the subgroup D virus is the same as that of one of the other subgroup viruses. The subgroup D virus has only 11 positions at which the amino acid is different (Fig. 2, closed circles and triangles) when the subgroup D sequence is compared with those of the other known subgroups. Thus, the determinants that confer mammal tropism may represent small changes dispersed along the protein chain, although the determinants could be juxtaposed in the native protein.

Small sequence changes can alter the receptor specificity of viral attachment proteins. In the case of influenza virus, a single amino acid change results in altered specificity in the recognition of sialic acid as a component of the receptor (42). In addition, variants of influenza virus can recognize multiple receptors, each with a different affinity (G. N. Rogers and J. C. Paulson, Fed. Proc. **41**:5880, 1982). Small changes along the peptide chain of gp85 may, in a similar manner, alter or expand the specificity of receptor interaction. By analogy, the clustering of five subgroup D-specific amino acids in hr1 and hr2 (at sites where the subgroup B sequence is also unique) may suggest two determinants involved directly or indirectly in receptor interactions (Fig. 2, closed circles).

The ability to infect mammalian cells is probably of little biological significance, since such infections are nonproductive (52). However, an alternative pathway of infection may provide certain viruses with a mechanism for infecting chicken cells that are already infected, thus abrogating the interference of superinfection that occurs when the secreted viral glycoprotein masks the homologous receptor on the surface of an infected cell (51). This interpretation is supported by the work of Bauer and Graf (2), in which they showed that subgroup B viruses only partially interfere with a superinfecting subgroup D virus, consistent with the possibility that subgroup D viruses have an alternative pathway into cells in addition to the receptor used by subgroup B viruses. Another explanation is that subgroup D viruses are antigenic variants of subgroup B viruses that fortuitously have an expanded host range. The five D- and B-unique sites (Fig. 2, closed circles) are either in or adjacent to two potential determinants of type-specific antigenicity, T3 and T4 (see below).

Type-specific variation in antigenicity. ASLV of a given subgroup are able to induce a neutralizing antibody response in chickens that primarily recognizes viruses from that same subgroup, thus defining one of the criteria for the subgroup classification of these viruses (57). However, the antigenicity of viruses within a given subgroup is not homogeneous, as displayed by virus neutralization and cell surface immunofluorescence with polyclonal antisera induced in chickens (30; Table 2). Usually the neutralization titer is greatest against the virus used to produce the antibody response; i.e., it is type specific. Subgroup- and type-specific antigenicity associated with virus neutralization has been identified both with the gp85-gp37 complex (18, 56) and individually with gp85 and gp37 (7). Our results with molecular recombinants between two subgroup A viruses and between a subgroup A virus and a subgroup E virus (Table 2) suggest that neutralizing antibodies in the polyclonal chicken antisera tested react only with gp85.

To investigate the molecular basis of type-specific antigenicity, we compared the positions of differences in the predicted amino acid sequence of the gp85-coding region of two pairs of viruses representing two different subgroups (Fig. 3). Comparison of two subgroup A viruses (RAV-1 and SR-A RSV) revealed heterogeneity in six positions of potential glycosylation. In addition to these differences, there are 18 other scattered nonconservative amino acid changes. Comparison of two subgroup B viruses (RAV-2 and MAV-B) showed only 1 change in potential glycosylation sites and 15 other scattered nonconservative amino acid changes. Surprisingly, 10 of the nonconservative amino acid changes between the two subgroup A sequences occurred at or adjacent to the positions of 10 of the nonconservative changes between the two subgroup B sequences (Fig. 3). The concurrence of the sites of change in the two comparisons identifies seven small domains (T1 through T7, Fig. 2) that correlate with type specificity in these viruses. Two of these domains are adjacent to positions of glycosylation site heterogeneity.

A number of distinct antigenic differences associated with virus neutralization must lie within the gp85 protein-coding region of the two subgroup A viruses, as demonstrated by the ANV-A1 recombinant virus (Table 2). It is likely that these antigenic differences are associated with the differ-

ences seen in the glycosylation pattern, the differences in the primary amino acid sequence within gp85, or both types of differences (Fig. 3). The subgroup B viruses RAV-2 and avian myeloblastosis virus type 2 have also been shown to express antigenic variability (30). However, we do not know whether the subgroup B helper virus of avian myeloblastosis virus that was sequenced by Kan et al. (32) and used in our analysis corresponds to the avian myeloblastosis virus type 2, subgroup B, tested previously for antigenic variability by Ishizaki and Vogt (30).

Structural similarities between the influenza virus hemagglutinin protein and the *env* gene product of avian retroviruses have been noted previously (29). The similarity between these two proteins is even more striking in light of the potential role of sequence changes in altering the phenotype of the protein. (i) Changes in the glycosylation pattern can contribute to changes in the antigenicity of the hemagglutinin protein (48). The variability in six potential glycosylation sites between RAV-1 and SR-A RSV may also contribute to the type-specific variation in antigenicity observed in these viruses. (ii) Naturally occurring antigenic variants of influenza virus arise from mutations in four specific domains in the hemagglutinin protein (60). We identified seven domains in which sequence variation coincided between two pairs of viruses in different subgroups. Some or all of these domains may be analogous to the four domains of the hemagglutinin protein that lie on the outer surface of the protein molecule and represent the primary sites of interaction with antibodies. It is perhaps noteworthy that six of the seven type-specific variable regions lie within the larger variable domains, hr1 and hr2, that correlate with subgroup specificity, suggesting that the domains that confer subgroup-specific antigenicity and receptor interaction also play a role in defining type-specific antigenicity.

Antigenic variability may be important in determining the disease course of retrovirus infection. This appears to be especially true for the chronic disease state induced by infection with lentiviruses. Both visna virus and equine infectious anemia virus show sequence and antigenic variations during infection of a single host (10, 43). Among different isolates of human immunodeficiency virus, sequence heterogeneity is greatest within the *env* gene (3, 50). This variability in *env* is not random but appears to be clustered in regions that may be determinants of antigenicity (38). In all of these cases, the antigenic variability is probably a result of negative selection by the host immune system. Thus, the changes seen in the *env* gene sequence of these highly pathogenic viruses are likely to be results of a common mechanism that is also responsible for the type specificity of avian retroviruses.

ACKNOWLEDGMENTS

The stenographic assistance of Sylvia Gault and Brenda Siler is gratefully acknowledged, as are the thoughtful comments of Stephanie Curtis. We thank Ruth Shuman for providing the chicken antisera and Wendy Osheroff for constructing pANV-E. We also thank Peter Vogt for helpful discussion and encouragement.

This investigation was supported by Public Health Service grant RO1 CA33147 from the National Cancer Institute. R.S. is the recipient of an American Cancer Society junior faculty award. J.O. is a fellow of the Leukemia Society of America. Computer resources were provided by Bionet, whose funding is from the Division of Research Resources, National Institutes of Health (Public Health Service grant U41 RR-01685).

LITERATURE CITED

1. **Bankier, A. T., and B. G. Barrell.** 1983. Shotgun DNA sequencing, p. 1–34. *In* R. A. Flavell (ed.), Techniques in nucleic acid biochemistry, vol. 85. Elsevier Scientific Publishing, Inc., New York.
2. **Bauer, H., and T. Graf.** 1969. Evidence for the possible existence of two envelope antigenic determinants and corresponding cell receptors for avian tumor viruses. Virology **37**:157–161.
3. **Benn, S., R. Rutledge, T. Folks, J. Gold, L. Baker, J. McCormick, P. Feorino, P. Piot, T. Quinn, and M. Martin.** 1985. Genomic heterogeneity of AIDS retroviral isolates from North America and Zaire. Science **230**:949–951.
4. **Bernstine, E. G., M. L. Hooper, S. Grandchamp, and B. Ephrussi.** 1973. Alkaline phosphatase activity in mouse teratoma. Proc. Natl. Acad. Sci. USA **70**:3899–3903.
5. **Bigner, D. D., S. H. Bigner, J. Pontén, B. Westermark, M. S. Mahaley, Jr., E. Ruoslahti, H. Herschman, L. F. Eng, and C. J. Wikstrand.** 1981. Heterogeneity of genotypic and phenotypic characteristics of fifteen permanent cell lines derived from human gliomas. J. Neuropathol. Exp. Neurol. **40**:201–229.
6. **Boettiger, D., D. N. Love, and R. A. Weiss.** 1975. Virus envelope markers in mammalian tropism of avian RNA tumor viruses. J. Virol. **15**:108–114.
7. **Bolognesi, D. P., H. Bauer, H. Gelderblom, and G. Huper.** 1972. Polypeptides of avian RNA tumor viruses. IV. Components of the viral envelope. Virology **47**:551–566.
8. **Bova, C. A., J. P. Manfredi, and R. Swanstrom.** 1986. *env* genes of avian retroviruses: nucleotide sequence and molecular recombinants define host range determinants. Virology **152**: 343–354.
9. **Buchhagen, D. L., and H. Hanafusa.** 1978. Intracellular precursors to the major glycoprotein of avian oncoviruses in chicken embryo fibroblasts. J. Virol. **25**:845–851.
10. **Clements, J. E., F. S. Pedersen, O. Narayan, and W. A. Haseltine.** 1980. Genomic changes associated with antigenic variation of visna virus during persistent infection. Proc. Natl. Acad. Sci. USA **77**:4454–4458.
11. **Coffin, J. M., M. Champion, and F. Chabot.** 1978. Nucleotide sequence relationships between the genomes of an endogenous and an exogenous avian tumor virus. J. Virol. **28**:972–991.
12. **Colbere-Garapin, F., F. Horodniceanu, P. Kourilsky, and A. C. Garapin.** 1981. A new dominant hybrid selective marker for higher eukaryotic cells. J. Mol. Biol. **150**:1–14.
13. **Crittenden, L. B., W. Okazaki, and R. Reamer.** 1963. Genetic resistance to Rous sarcoma virus in embryo cell cultures and embryos. Virology **20**:541–544.
14. **Crittenden, L. B., W. Okazaki, and R. Reamer.** 1964. Genetic control of responses to Rous sarcoma and strain RPL12 viruses in the cells, embryos, and chickens of two inbred lines. Natl. Cancer Inst. Monogr. **17**:161–177.
15. **Crittenden, L. B., H. A. Stone, R. H. Reamer, and W. Okazaki.** 1967. Two loci controlling genetic cellular resistance to avian leukosis-sarcoma viruses. J. Virol. **1**:898–904.
16. **Dorner, A. J., and J. M. Coffin.** 1986. Determinants for receptor interaction and cell killing on the avian retrovirus glycoprotein gp85. Cell **45**:365–374.
17. **Dorner, A. J., J. P. Stoyle, and J. M. Coffin.** 1985. Molecular basis of host range variation in avian retroviruses. J. Virol. **53**:32–39.
18. **Duesberg, P. H., G. S. Martin, and P. K. Vogt.** 1970. Glycoprotein components of avian and murine RNA tumor viruses. Virology **41**:631–646.
19. **Duff, R. G., and P. K. Vogt.** 1969. Characteristics of two new avian tumor virus subgroups. Virology **39**:18–30.
20. **Edwards, M. K. S., and M. W. McBurney.** 1983. The concentration of retionic acid determines the differentiated cell types formed by a teratoma carcinoma cell line. Dev. Biol. **98**: 187–191.
21. **England, J. M., D. P. Bolognesi, B. Dietzschold, and M. S. Halpern.** 1977. Evidence that a precursor glycoprotein is cleaved to yield the major glycoprotein of avian tumor virus. J. Virol. **21**:810–814.

22. **Giard, D. J., S. A. Aaronson, G. J. Todaro, P. Arnstein, J. H. Kersey, H. Dosik, and W. P. Parks.** 1973. In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. J. Natl. Cancer Inst. **51**:1417–1423.
23. **Glaser, R., and F. Rapp.** 1972. Rescue of Epstein-Barr virus from somatic cell hybrids of Burkitt lymphoblastoid cells. J. Virol. **10**:288–296.
24. **Gluzman, Y.** 1981. SV40-transformed simian cells support the replication of early SV40 mutants. Cell **23**:175–182.
25. **Hanafusa, H., and T. Hanafusa.** 1966. Determining factor in the capacity of Rous sarcoma virus to induce tumors in mammals. Proc. Natl. Acad. Sci. USA **55**:532–538.
26. **Hanafusa, T., H. Hanafusa, and T. Miyamoto.** 1970. Recovery of a new virus from apparently normal chick cells by infection with avian tumor viruses. Proc. Natl. Acad. Sci. USA **67**:1797–1803.
27. **Hayman, M. J.** 1978. Synthesis and processing of avian sarcoma virus glycoproteins. Virology **85**:475–486.
28. **Hughes, S., and E. Kosik.** 1984. Mutagenesis of the region between *env* and *src* of the SR-A strain of Rous sarcoma virus for the purpose of constructing helper independent vectors. Virology **136**:89–99.
29. **Hunter, E., E. Hill, M. Hardwick, A. Bhown, D. E. Schwartz, and R. Tizard.** 1983. Complete sequence of the Rous sarcoma virus *env* gene: identification of structural and functional regions of its product. J. Virol. **46**:920–936.
30. **Ishizaki, R., and P. K. Vogt.** 1966. Immunological relationships among envelope antigens of avian tumor viruses. Virology **30**:375–387.
31. **Joho, R. H., M. A. Billeter, and C. Weissman.** 1975. Mapping of biological functions on RNA of avian tumor viruses: location of regions required for transformation and determining host range. Proc. Natl. Acad. Sci. USA **72**:4772–4776.
32. **Kan, N. C., M. A. Baluda, and T. Papas.** 1985. Sites of recombination between the transforming gene of avian myeloblastosis virus and its helper virus. Virology **145**:323–329.
33. **Kawai, S., and M. Nishizawa.** 1984. New procedure for DNA transfection with polycation and dimethyl sulfoxide. Mol. Cell. Biol. **4**:1172–1174.
34. **Kit, S., D. R. Dubbs, L. J. Piekarski, and T. C. Hsu.** 1963. Deletion of thymidine kinase activity from L cells resistant to bromodeoxyuridine. Exp. Cell. Res. **31**:297–312.
35. **Klemenz, R., and H. Diggelmann.** 1978. The generation of the two envelope glycoproteins of Rous sarcoma virus from a common precursor polypeptide. Virology **85**:63–74.
36. **Leamnson, R. N., and M. S. Halpern.** 1976. Subunit structure of the glycoprotein complex of avian tumor virus. J. Virol. **18**:956–968.
37. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
38. **Modrow, S., B. H. Hahn, G. M. Shaw, R. C. Gallo, F. Wong-Staal, and H. Wolf.** 1987. Computer-assisted analysis of envelope protein sequences of seven human immunodeficiency virus isolates: prediction of antigenic epitopes in conserved and variable regions. J. Virol. **61**:570–578.
39. **Moelling, K., and M. Hayami.** 1977. Analysis of precursors to the envelope glycoproteins of avian RNA tumor viruses in chicken and quail cells. J. Virol. **22**:598–607.
40. **Moscovici, C., M. G. Moscovici, H. Jimenez, M. M. C. Lai, M. J. Hayman, and P. K. Vogt.** 1977. Continuous tissue culture cell lines derived from chemically induced tumors of Japanese quail. Cell **11**:95–103.
41. **Payne, L. N., and P. M. Biggs.** 1964. Differences between highly inbred lines of chickens in the response to Rous sarcoma virus of the chorioallantoic membrane and of embryonic cells in tissue culture. Virology **24**:610–616.
42. **Rogers, C. N., J. Paulson, R. S. Daniels, A. Douglas, I. A. Wilson, D. C. Wiley, and J. J. Skehel.** 1983. Single amino acid substitutions in influenza haemagglutinin change receptor binding specificity. Nature (London) **304**:76–78.
43. **Salinovich, O., S. L. Payne, R. C. Montelaro, K. A. Hussain, C. J. Issel, and K. L. Schnorr.** 1986. Rapid emergence of novel antigenic and genetic variants of equine infectious anemia virus during persistent infection. J. Virol. **57**:71–80.
44. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1974. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5468.
45. **Schwartz, D. E., R. Tizard, and W. Gilbert.** 1983. Nucleotide sequence of Rous sarcoma virus. Cell **32**:853–869.
46. **Sealy, L., M. L. Privalsky, G. Moscovici, C. Moscovici, and J. M. Bishop.** 1983. Site-specific mutagenesis of avian erythroblastosis virus: erb-B is required for oncogenicity. Virology **130**:155–178.
47. **Shalloway, D., A. D. Zelenetz, and G. M. Cooper.** 1981. Molecular cloning and characterization of the chicken gene homologous to the transforming gene of Rous sarcoma virus. Cell **24**:531–541.
48. **Skehel, J. J., D. J. Stevens, R. S. Daniels, A. R. Douglas, M. Knossow, I. A. Wilson, and D. C. Wiley.** 1984. A carbohydrate side chain on hemagglutinins of Hong Kong influenza viruses inhibits recognition by monoclonal antibody. Proc. Natl. Acad. Sci. USA **81**:1779–1783.
49. **Southern, P. J., and P. Berg.** 1982. Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. J. Mol. Appl. Genet. **1**:327–341.
50. **Starcich, B. R., B. H. Hahn, G. M. Shaw, P. D. McNeely, S. Modrow, H. Wolf, E. S. Parks, W. P. Parks, S. F. Joseph, R. C. Gallo, and F. Wong-Staal.** 1986. Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS. Cell **45**:637–648.
51. **Steck, F. T., and H. Rubin.** 1966. The mechanism of interference between an avian leukosis virus and Rous sarcoma virus. II. Early steps of infection by RSV of cells under conditions of interference. Virology **29**:642–653.
52. **Svoboda, J.** 1962. Further findings on the induction of tumors by Rous sarcoma in rats and on the Rous virus producing capacity of one of the induced tumors (XC) in chicks. Folia Biol. **8**:215–220.
53. **Tabin, C. J., J. W. Hoffmann, S. P. Goff, and R. A. Weinberg.** 1982. Adaptation of a retrovirus as a eucaryotic vector transmitting the herpes simplex virus thymidine kinase gene. Mol. Cell. Biol. **2**:426–436.
54. **Tereba, A., and K. G. Murti.** 1977. A very sensitive biochemical assay for detecting and quantitating avian oncornaviruses. Virology **80**:166–176.
55. **Toyoshima, K., and P. K. Vogt.** 1969. Enhancement of inhibition of avian sarcoma viruses by polycations and polyanions. Virology **38**:414–426.
56. **Tozawa, H., H. Bauer, T. Graf, and H. Gelderblom.** 1970. Strain-specific antigen of the avian leukosis sarcoma virus group. 1. Isolation and immunological characterization. Virology **40**:530–539.
57. **Vogt, P. K.** 1977. Genetics of RNA tumor viruses, p. 341–455. *In* H. Fraenkel-Conrat and R. R. Wagner (ed.), Comprehensive virology, vol. 9. Plenum Publishing Corp., New York.
58. **Wang, L. H., P. Duesberg, S. Kawai, and H. Hanafusa.** 1976. The location of envelope-specific and sarcoma-specific oligonucleotides on the RNA of Schmidt-Ruppin Rous sarcoma virus. Proc. Natl. Acad. Sci. USA **73**:447–451.
59. **Weiss, R. A.** 1982. Experimental biology and assay, p. 210–260. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
60. **Wiley, D. C., I. A. Wilson, and J. J. Skehel.** 1981. Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature (London) **289**:373–378.
61. **Zagursky, R. J., K. Baumeister, N. Lomax, and M. L. Berman.** 1985. Rapid and easy sequencing of large linear double-stranded DNA and supercoiled plasmid DNA. Gene Anal. Tech. **2**:89–94.
62. **Zimmerman, H. M.** 1985. The nature of gliomas as revealed by animal experimentation. Am. J. Pathol. **31**:1–29.