Title
Analysis of endogenous avian retrovirus DNA and RNA

Author
Baker, Barbara Joan

Publication Date
1981

Peer reviewed|Thesis/dissertation

Analysis of Endogenous Avian Retrovirus DNA and RNA: Viral and Cellular Determinants of Retrovirus Gene Expression

by

Barbara Joan Baker
B.S. University of California, San Diego 1974

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Microbiology and Immunology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco

[redacted]

Date University Librarian

SEP 23 1981

Degree Conferred: . . . . . .

ABSTRACT

I have characterized endogenous avian retrovirus DNAs and RNAs in uninfected white leghorn chicken embryos. The embryos I studied include representations of four phenotypic classification of chickens (gs^-chf^-, gs^+chf^+, gs^-chf^+ and v^+), and contained the endogenous proviruses ev1, 2, 3, 4, 5, 6, 8, 9 and 15. I confirmed the previously described phenotypes and restriction maps for ev1, 3, 4, 5, 6, 8 and 15, and I confirmed the phenotype for ev9 and derived a restriction map for this locus. Evaluation of the structure of the individual proviruses and characterization of the nuclear and cytoplasmic viral RNA isolated from embryos of defined genotype yielded the following major conclusions: (i) Viral RNA could not be detected for ev4, ev5, ev8 or ev15. (ii) The amount of stable RNA produced from ev1, ev3, ev6 or ev9 varied over a range of at least one to three hundred fold and was at least 10-1000 fold lower than the amounts found in cells productively infected with exogenous avian retroviruses. (iii) A deletion in the provirus at ev3 can account for three abnormalities: the production of a viral protein that results from fusion of internal regions of the gag and pol genes; the abnormal structure of a sub-genomic viral RNA; and failure of the abnormal RNA to appear in the cytoplasm. (iv) A deletion in the provirus at ev6 gives rise to at least two major anomalies: initiation

of transcription at an upstream cellular promoter, and the abnormal metabolism of an ev6 viral RNA. (v) Ev9 has an ostensibly normal structure; nevertheless, the metabolism of an RNA arising from this locus is aberrant. I conclude that the expression of endogenous retrovirus genes in chickens is subject to control by both viral and cellular determinants.

ACKNOWLEDGEMENT

I am indebted to Mike Bishop, under whose direction this work was performed. I am also grateful for his patience. I want to thank H. Varmus and L. Levintow as well as all colleagues in the laboratory for their advice and helpful suggestions.

Finally, I thank my parents for their support and understanding.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

RETROVIRUSES

Retroviruses comprise a family of viruses that are widely distributed among the vertebrate species. Several unique features of retroviruses have attracted the attention of investigators: retroviruses are important natural agents of oncogenesis in several species; they are transmitted both genetically and by horizontal infection; and they are dependent upon an RNA-directed DNA polymerase for the establishment of infection (Baltimore, 1970; Temin and Mizutani, 1970).

Four major characteristics define the retrovirus family; the structural properties of the enveloped virion, a diploid single stranded RNA genome (of messenger RNA polarity), the presence of the enzyme reverse transcriptase in virions, and the requirement for a DNA intermediate in viral replication (Bishop, 1978; Varmus, 1981).

Classification of Retroviruses

Retroviruses are classified by using a variety of criteria (Bishop, 1978; Fenner, 1976). i) Each strain of virus has been classified either as an endogenous virus or an exogenous virus. The endogenous viruses are present in the germ line DNA of a species or certain individuals

of a species and are transmitted vertically through the gametes (Weiss, 1975; Aaronson and Stephenson, 1976). Exogenous viruses persist by virtue of horizontal spread among susceptible members of a species. ii) Virus strains have been identified according to the species from which they were isolated (eg. avian leukosis-sarcoma virus (ALSV) from chickens or pheasants; murine leukemia virus (MLV) from mice). iii) The virion proteins possess various kinds of antigenic sites which are useful for classifying the viruses (eg. type-specific determinants, specific for each virus and located on virus encoded envelope glycoproteins; group-specific determinants shared by viruses of the same host species and located on the core proteins; interspecies determinants, shared with viruses of other species). iv) Viruses are classified on the basis of host range. These properties depend primarily on the virus-encoded envelope glycoprotein. Avian retroviruses have been divided into seven subgroups (A to G) according to host range specificities. Avian cells are susceptible to various assortments of virus subgroups depending upon the presence of specific cell encoded receptors (Weiss, 1975). The expression of receptors is controlled by several dominant cellular genes (Vogt, 1977). v) Retroviruses can be distinguished by the disease that they induce in susceptible animals (eg. leukemia, sarcomas, carcinomas) though many viruses produce no disease.

The Retrovirus Virion

In the following sections the avian leukosis-sarcoma group of retroviruses will be used as a prototype for description of and reference to certain virus molecules and virus functions. Available data indicate that similar properties apply to other retroviruses (Bishop, 1978).

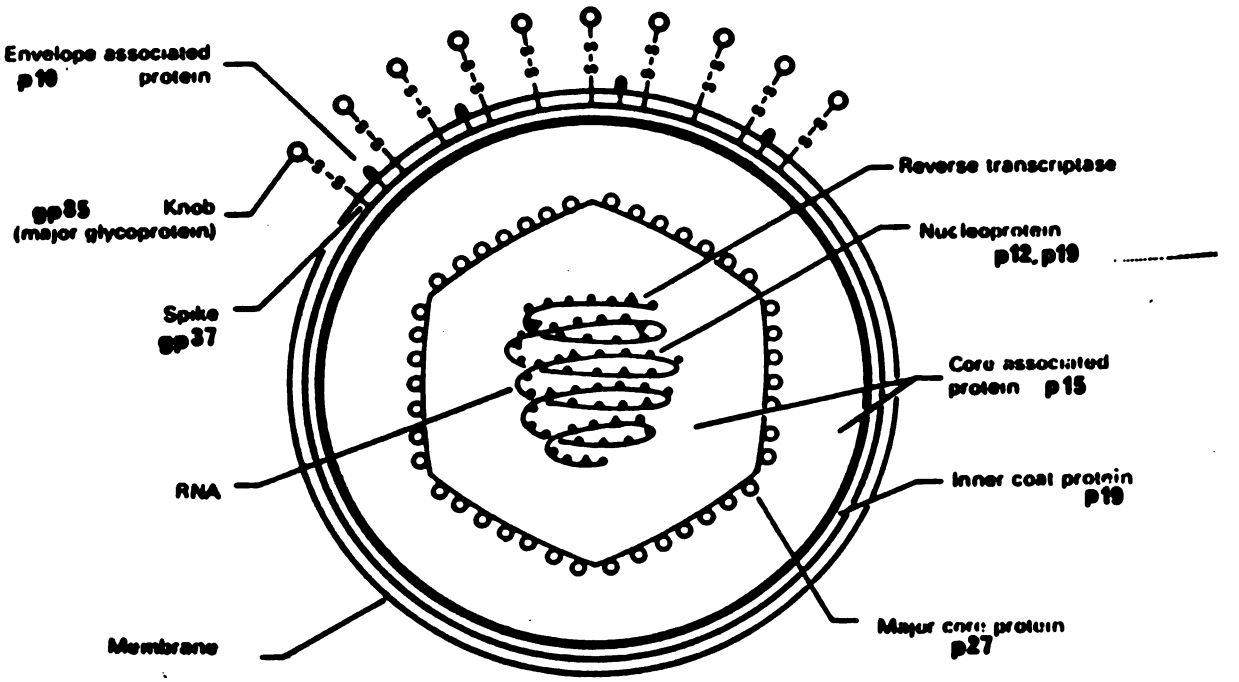
The retrovirus virion consists of an envelope, a core shell and a core with a ribonucleoprotein (RNP) (Fig. 1). The envelope of the virion is derived from the plasma membrane of the host cell during the process of virus budding (Vogt, 1977).

Information for the virion structural components is specified by three retrovirus genes: gag, encoding the major structural proteins of the core; pol encoding reverse transcriptase; and env, encoding the envelope glycoproteins (Bolognesi, 1974; Bishop, 1978).

The external surface of the virion envelope contains protruding spikes and knobs composed of two viral encoded glycoproteins, gp 85 and gp 37 which are linked by disulfide bonds. Determinants for type-specific antigenicity are located primarily within these proteins (Weiss, 1975). The glycoproteins are responsible for absorption to and penetration into susceptible host cells (Vogt, 1977). A putative gag protein, p10, has been localized to the virus envelope (Bolognesi, et al., 1973). The gag protein p19 is associated with the lipid envelope.

Figure 1. Structure of the retrovirus virion (from Bolognesi, et al., 1978).

Figure 1.


Envelope associated protein p10
gp85 (major glycoprotein)
Knob
Spike gp37
RNA
Membrane
Reverse transcriptase
Nucleoprotein p12, p19
Core associated protein p15
Inner coat protein p19
Major coat protein p27

Within the envelope is a geometrically symmetrical "core shell" composed of virus protein p27 which is a major locus for group specific antigenicity (Bolognesi 1974). Between the virion envelope and core shell resides virion gag protein p15. This protein is associated with protease activity and may function in processing of the gag precursor protein after it is removed from the same precursor molecule (see below).

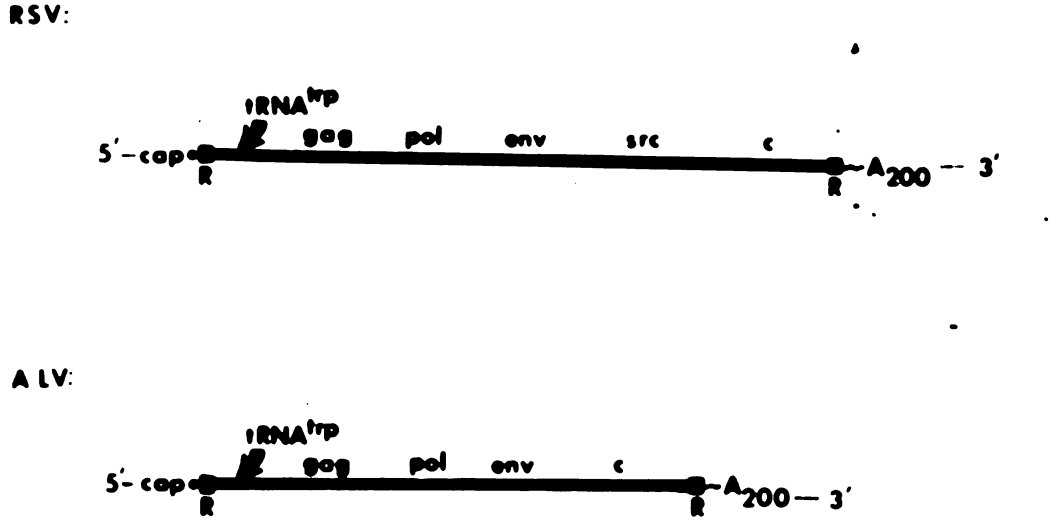
The core shell encloses a helical RNP composed of the genome (Bolognesi 1974, Bishop 1978 and ref. cited) low molecular weight RNA, reverse transcriptase and two virus encoded gag phosphoproteins pp19 and pp12. pp19 binds preferentially to duplex regions of RNA, whereas pp12 binds to single strand regions of RNA (Sen and Todaro, 1977; Smith and Baily, 1979).

The Genome of Retroviruses

The genome of a retrovirus is composed of 2 identical subunits of single stranded RNA, each ca. 9-10 kilobases (for Rous-sarcoma virus (RSV) or 8-9 kilobases (for avian leukosis virus (ALV) (Bishop 1978; see Fig. 2). The 3' termini of the haploid subunits are polyadenylated (ca. 200 residues) (Bender, et al., 1976) the 5' termini are capped by the structure 5'-m^7GpppGm (Furuichi, et al., 1975), and approximately 10 residues of adenosine, located at specific sites within the 3' half of the genome, are methylated (Furuichi, et al., 1975).

Figure 2. The genome of Rous Sarcoma virus (RSV) and avain leukosis virus (ALV).

Figure 2.


RSV:
tRNA^trp
gag
pol
env
src
c
5'-cap
R
R
A200 -- 3'
A LV:
tRNA^trp
gag
pol
env
c
5'-cap
R
R
A200 — 3'

A molecule of tRNA ^{Trp}, derived from the host cell, is bound to the genome of RSV at a site 101 nucleotides distant from the 5' terminus of the genome (Taylor, et al., 1975, Shine, et al., 1977). tRNA^{Trp} serves as a primer for initiation of DNA synthesis by reverse transcriptase in vitro (Dahlberg, et al., 1975).

The haploid subunit of RSV is terminally redundant for 16 to 21 nucleotides (Schwartz, et al., 1977). The redundant sequence, "R", is essential to the early stages of DNA synthesis (see below; Swanstrom, et al., 1981).

Four genes have been identified in the genome of RSV (Bishop 1978, Baltimore 1974): gag encoding structural proteins of the virion core; pol encoding reverse transcriptase; env encoding the glycoproteins of the viral envelope; and src which encodes a protein with kinase activity that is responsible for neoplastic transformation of the host cell (Brugge, et al., 1977; Collett and Erikson, 1979; Levinson, et al., 1978). The gag, pol and env gene products are responsible for viral replication. The ALV genome consists of these three genes (Bishop, 1978; see Fig. 2).

The Replicative Cycle

Viral DNA is transcribed in the cytoplasm from the retrovirus RNA genome by the enzyme reverse transcriptase during the early hours of infection (Verma 1976; Bishop 1978). The product, double stranded DNA, migrates to the

nucleus and is subsequently integrated into the chromosomal DNA (Varmus, et al., 1974; Shank, et al., 1977; Bishop, 1978) (Fig. 3).

Reverse transcriptase (Baltimore, 1970; Temin, 1970) is an RNA-dependent DNA polymerase encoded by the pol gene (Verma, et al., 1977; Oppermann, et al., 1977) and it is encapsidated in virus particles. Reverse transcriptase is an enzyme complex composed of two subunits α and β (Kacian, 1971; Faras, et al., 1972). The complex has RNA-directed DNA polymerase activity (Baltimore, 1970' Temin, 1970), DNA-directed DNA polymerase activity (Temin and Baltimore, 1972) and hybrid specific ribonuclease activity (RNase H) (Moelling, 1971).

Transcription (-strand DNA synthesis) of the RSV genome initiates on a $tRNA^{Trp}$ primer which is bound to viral RNA 101 nucleotides from the 5' end by a 17 nucleotide long complementarity (Cordell, et al., 1976; Peters, et al., 1979). Transcription proceeds to the 5'-end of the template. The RNA template at the 5' terminus of the viral genome is removed from the complementary DNA transcript (perhaps by the action of RNase H activity associated with reverse transcriptase (Friedrich, et al., 1979)). The newly synthesized minus strand DNA pairs with the direct terminally redundant sequence, R, located adjacent to the polyA tail at the 3' end of the same or another RNA template (Schwartz et al., 1977; Coffin, et al., 1977) and DNA transcription

Figure 3. An overview of the replication cycle of retroviruses.

The top line shows several structural features of a subunit of viral RNA: the capped nucleotide at the 5' terminus, the poly(A) tract at the 3' terminus, and the site of binding of the cellular tRNA used to prime DNA synthesis. In addition, the solid squares at the ends indicate the short sequence "R" directly repeated at the termini of viral RNA; the stippled box called U_5 denotes the sequence unique to the 5' terminus between the primer binding site and "R"; and the open box called U_3 denotes the sequence of several hundred bases unique to the 3' terminus and repeated at both ends of unintegrated linear DNA. The subsequent lines show the organization of these repeated sequences as LTR units in linear DNA, the two forms of monomeric circular DNA, and proviral DNA. A subunit of viral RNA, the putative primary transcript of the provirus, is shown in the last line. In order to clarify the structural details, the figure is not drawn to scale.

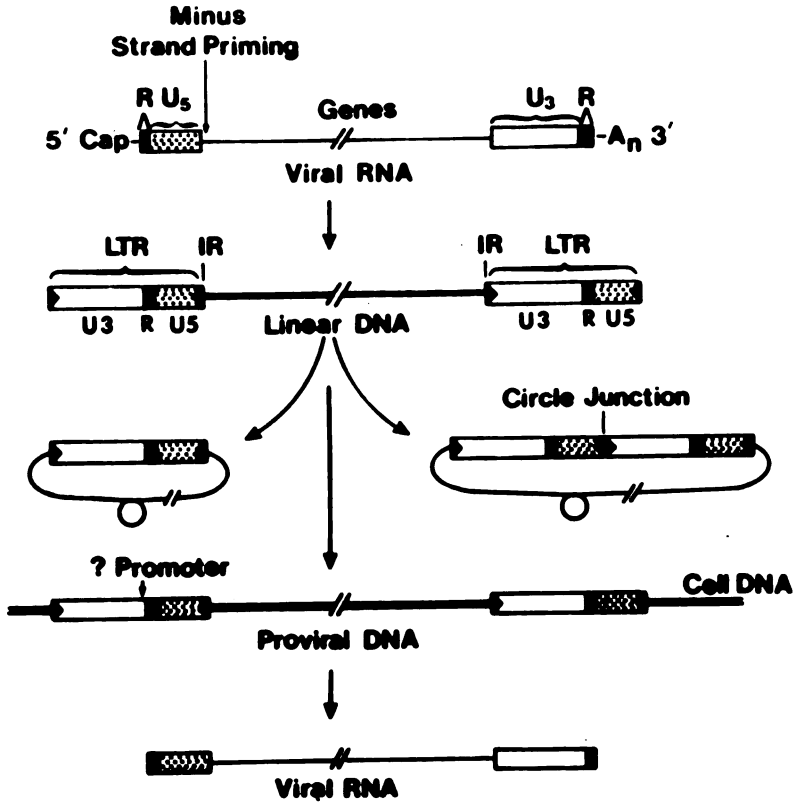


Figure 3. An overview of the replication cycle of retroviruses (Varmus, 1981).

—— viral RNA

▬▬ double stranded DNA

═══ cell DNA

continues towards the 5' terminus of the viral RNA (Swanstrom, et al., 1981). Synthesis of the second DNA strand (plus (+) strand) initiates prior to the completion of minus strand DNA synthesis (Varmus, et al., 1978). The primer(s) for plus strand synthesis and the mechanistic details of this process are still to be elucidated (Czernilofsky, et al., 1980; Swanstrom, et al., 1981; Kung, et al., 1981).

The end products of viral DNA synthesis in cells infected with retroviruses include both linear and circular duplex molecules (Varmus, et al., 1975; Guntaka, et al., 1975; see Fig. 3). The linear form of viral DNA is bounded by long terminal repeats (LTR's). Each LTR is composed of three domains: U_3, a unique sequence from the 3' end of the viral genome; R, a short redundant sequence present at each end of the viral RNA; and U_5, encoded at the 5' end of the viral RNA. The domains are arranged in the order: 5'-U_3-R-U_5-3' (for review, see Coffin, 1979). Some of the circular molecules contain two copies of an LTR, and others contain one copy (DeLorbe, et al., 1980; Varmus, 1981). The linear molecules migrate to the nucleus where their ends presumably join to form the circular species by an unknown mechanism (Shank, et al., 1978; Varmus, 1981 and ref. cited). Viral DNA is then inserted into chromosomal DNA to become continguous with the host genome (Varmus, et al., 1976; Varmus, 1981 and ref. cited). It is not clear whether a linear or circular molecule is the intermediate involved

in integration. The integrated provirus exists in a form colinear with the unintegrated linear viral DNA and is flanked by long terminal repeats (LTRs) (see Fig. 3; Varmus, 1981; Hughes, et al., 1978).

Long terminal repeats and cell-virus junctions isolated from integrated and unintegrated viral DNA of various retrovirus strains have been sequenced (Cold Spring Harbor Symp. Vol. 45, 1980). The data suggest that the LTR's play a role in regulation of virus gene expression (eg. RNA transcription initiation and termination). Structural similarities between retrovirus proviruses and bacterial, yeast and Drosophila moveable genetic elements were also revealed (a direct repeat of host DNA, inverted repeats of viral DNA located in direct repeats of viral DNA (Majors, et al., 1980; Temin 1980; Varmus, 1981)). These properties imply that retrovirus LTR's may function to promote integration and possibly transposition by mechanisms similar to transposible elements (Grindley, et al., 1979, Shapiro, et al., 1979). At present, there is no evidence to support the notion that retroviruses are "moveable elements" except by the process involving reverse transcriptase (Weinberg, 1980; Varmus, 1981).

Viral DNA integrates at numerous and perhaps random sites on the host genome (Taylor, 1979; Hughes, et al., 1980), and once integrated the provirus is apparently

stable. Relocation of a provirus from one position to another within the host genome has not been experimentally observed but has been suggested on the basis of close clustering of several endogenous avian proviruses in a region of chicken chromosome 1, (see Discussion). Excision of a provirus has been observed when part or all of the host chromosome on which it resides is deleted (Morris, et al., 1977).

Deletions removing LTR's and other regions of viral DNA have been observed among endogenous proviruses of chickens (Hughes, et al., 1981; Hayward, et al., 1980, and ALV proviruses in bursal tumors (Payne, et al., 1980; Neel, et al., 1980). The mechanisms responsible for these deletions are not known (see Discussion).

Viral Gene Expression

Viral RNA is synthesized in the nucleus of the cell. The template for synthesis is presumably the integrated provirus and the responsible enzyme is cellular RNA polymerase II (for review, see Bishop, 1978). The initial product of synthesis appears to be a genome length RNA molecule of 9-10 kilobases in length (ca. 38s) for RSV, or 8-9 kilobases in length (ca. 35s) for ALV (Hayward, 1977; Weiss, et al., 1977). Shortly after infection the cytoplasm of RSV infected cells contains the following viral RNAs: 9-10 kb (38s), 5.4 kb (28s), and 3.3 kb (21s).

The cytoplasmic viral TNAs of ALV infected cells are 8-9 kb (35s) and 3.3 (kb) in length (Weiss, et al., 1977; Hayward,1977; see Fig. 4). These virus-specific RNAs are found on polyribosomes (Lee, et al., 1979).

The following criteria have been used to assess the function of the intracellular RNAs. The genetic composition of the virus-specific RNAs was determined by hybridizing size fractionated infected cellular RNA with radioactively labeled cDNA probes specific for various regions of the RSV genome (Weiss, et al., 1977; Hayward, 1977). From these analyses it was determined that the 38s and 35s RNAs represented the entire viral genome of RSV and ALV respectively. The 28s RNA of RSV infected cells contained sequences from the 3' domain of the genome encoding env and src. The 21s species of RSV infected cells contained sequences for src alone, and the 21s species of ALV infected cells contained sequences for the envelope gene. All species hybridized with cDNA complementary to nucleotide sequences present only at the end of the viral genome (see Fig. 4; Weiss, et al., 1977; Hayward, 1977). These results suggested the following: the 38s and 35s RNAs are messenger RNAs for the gag and pol proteins; the RSV 28s and ALV 21s RNAs are mRNAs for the envelope glycoproteins and the RSV 21s is the mRNA for the src protein. It is also suggested that subgenomic RNAs were spliced.

Figure 4.

The diagram illustrates the genetic composition of the three classes of RSV mRNAs (38S, 28S, and 21S) and the two classes of ALV mRNAs (35S and 21S) found in the cytoplasm of RSV or ALV infected cells. The solid boxes labeled "L" are the sequences that are joined to internal coding sequences during RNA processing. The products of translation from the viral mRNAs are denoted by their molecular weight in thousands preceded by a p, gp for glycoprotein, pp for phosphoprotein, Pr for precursor. The details of protein processing events have been omitted.

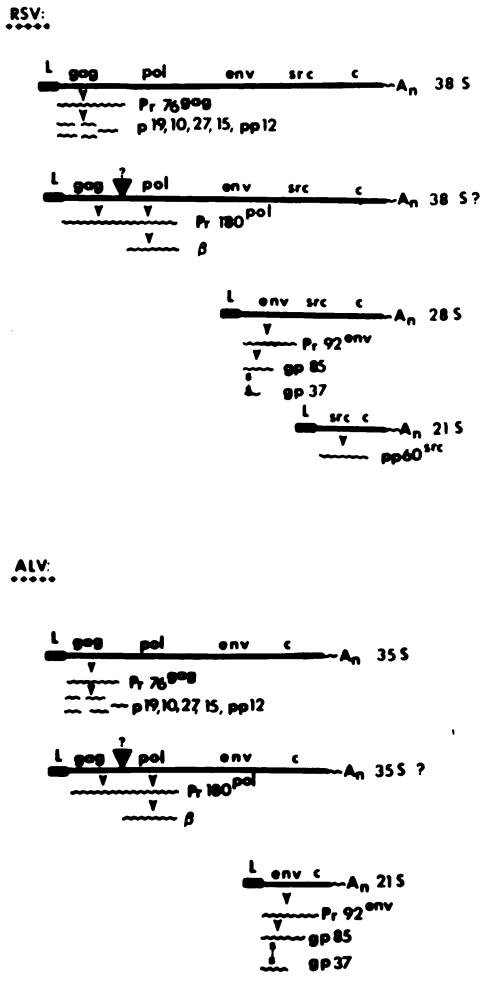


Figure 4. The mRNAs of Rous Sarcoma virus and avian leukosis virus and proposed strategy for the expression of RSV and ALV genes.

The functions of the major size classes of RSV and ALV viral RNA have been assessed principally by translation *in vitro*; the results conform to the predictions based upon the genetic content of each species. The RSV 38s and ALV 35s RNAs isolated from either virions or infected cells direct the synthesis of the *gag* and *pol* precursor proteins (Pr 76^{gag} and Pr 180^{pol}, respectively). A subgenomic RNA species, presumably RSV 28s, directs synthesis of the *env* precursor protein *in vitro* (Pawson, et al., 1977, 1980). The 21s RNA of RSV infected cells directs synthesis of $pp60^{src}$ in lysates of rabbit reticulocytes (Yamamoto, et al., 1980; Bishop, et al., 1980). Cells infected with a *src* deletion mutant of RSV (td RSV), synthesize 21s *env* RNA which is functionally equivalent to the 21s RNA of ALV infected cells. Microinjection of the 21s RNA of td RSV infected cells into chicken cells infected with an *env*-defective deletion mutant induces synthesis of complementing envelope glycoprotein (Stacey, et al., 1977).

Initiation of Transcription of Viral RNA

The promoter sequence responsible for initiation of viral RNA transcription is most likely located within the U_3 region of the proviral left-ward LTR (Fig. 3). The results of several types of experiments support this idea. First, the U_3 regions of several strains of retroviruses have been sequenced. Features of the sequence of this region resemble

other putative promoter sites that have been identified for various encaryotic genes (e.g., a "Hogness" box situated 24 bases upstream from the mRNA capping site (for RSV), (Yamamoto, et al., 1980; Czernilofsky, et al., 1980; Swanstrom, et al., 1980)). Second there is evidence that RNA synthesis from eucaryotic genes is initiated with the nucleotide which is capped (Ziff and Evans, 1978). If the "upstream" U_3 sequence serves as a promoter, the primary transcript would not be longer than a subunit of virion RNA (assuming transcription termination occurs within the provirus, see below). Precursor viral RNA of greater than subunit length has not been identified in infected cells. Viral RNAs, slightly larger than virion subunits, have been observed in MLV-producing cells (Fan, 1977), but the significance of these species is uncertain, since it was not demonstrated that these RNAs functioned as precursors.

Since the provirus is terminally redundant, it is possible that the putative promoter sequence located in the left and right LTR function to initiate transcription of DNA positioned downstream. Use of the right-ward LTR would lead to RNA species linking the R-U_5 sequence from the right end of the provirus with cellular sequences from flanking DNA. RNA species of this type have been identified by Quintrell, et al. (1980) in cloned mammalian cells transformed by RSV, and by Payne, et al. (1981) and Neel, et al. (1981) in avian leukosis virus-induced tumor cells.

However, the exact site of initiation of these species has yet to be identified. There is evidence that the major species of viral RNA in infected cells are transcribed from a single promoter. The 5' ends of intracellular subgenomic RNAs are identical to the 5' ends of virion RNA, suggesting that these RNAs arise by splicing (see below) and that internal viral promoters are not used for their generation. UV mapping studies (Hackett and Sauerbier, 1975) indicate that all three major species of RSV cellular RNA have similar target sizes, again suggesting that a single promoter is responsible for viral RNA transcription.

Termination of Transcription

It is speculated that transcription terminates at the U_3R junction in the right-hand LTR. If transcription initiates as described above and terminates at the U_3R junction then retrovirus genomic RNA is generated in its entirety in the absence of processing events. Evidence to the contrary has been presented by Yamamoto, et al., (1980). The data suggested that transcription continued through the entire right-ward LTR and passed at least eighteen bases of flanking cellular DNA. The data were obtained by sequence analysis of a single cloned cDNA that was transcribed from a subgenomic viral mRNA. More information is required to assess the significance of this finding.

Processing of Viral RNA: Generation of Subgenomic mRNAs

Investigation of the structure of genomic and intracellular RNAs from virus and infected cells has indicated requirements for multiple processing events during the generation of mature species: (i) the genome and intracellular RNAs are capped and polyadenylated (see aboveP; (ii) the subgenomic mRNAs are generated by covalent joining of sequences non-contiguous in the provirus ("splicing").

All three species of RSV infected cell RNA (38S, 28S and 21S) and both species of ALV intracellular RNA (35S and 21S) annealed with cDNA specific for the 5' terminus of the viral RNA (Weiss, et al., 1977). It was suggested that sequences from the 5' end of the viral RNA were joined to sequences near env or src genes by a splicing mechanism. Direct evidence for this was provided by experiments that examined the 5' termini of ^{32}P-labeled 38S, 28S and 21S intracellular RSV RNAs (Mellon and Duesberg, 1977; Cordell, et al., 1978).

The S1 mapping technique has been recently used to determine the length of the sequence transposed to the 5' ends of RSV 28S and 21S RNAs. These RNAs were used to protect labelled restriction enzyme fragments of cloned RSV DNA from S1 nuclease digestion. From the size of the protected fragments, it was estimated that about 360 nucleotides from the 5' end of the genome were joined to internal coding sequences during RNA processing (P. Hackett, unpublished).

Synthesis of Viral Proteins

The gag gene product is a polyprotein Pr76 (76,000 daltons) synthesized from the RSV 38S or ALV 35S gag/pol mRNA. (The mRNAs are indistinguishable from genomic subunits); Pr76 is cleaved to yield the core proteins p19, p27, p12, p15 and p10. P10 is located in the envelope. The order of these proteins in Pr76 is NH_2-p19-p10-p27-p12-p15-COOH (Vogt, et al., 1975; Hunter, unpublished). The primary precursor to the avain retrovirus reverse transcriptase is a protein Pr180 (180,000 daltons) which is a common product of the gag and pol genes. Analysis of its structure demonstrated that Pr180 contained polypeptide regions corresponding to both $Pr76^{gag}$ and the β subunit of reverse transcriptase (Retenmier, et al., 1979). The mRNA for the precursor to the polymerase has not been distinguished structurally from gag mRNA. The mechanism by which the gag termination codon(s) are bypassed to generate Pr180 is not understood. It has been suggested that the mechanism involves elimination of gag termination codons from the 38S and 35S RNA by splicing. Production of functional reverse transcriptase involves cleavage and modification of the gag/pol polyprotein (Copeland, et al., 1980). The env gene is expressed via subgenomic 28S (RSV) and 21S (ALV) mRNAs. The primary translation product of env mRNA is presumably inserted in cell membranes during synthesis, and subsequently glycosylated and cleaved to generate the

the envelope glycoproteins gp85 and gp37 present on the virion envelope. The protein product of *src*, $pp60^{src}$, which is not required for replication and not readily found in viral progeny, has been shown to be phosphorylated and inserted into the cellular plasma membrane (for review, see Hunter, 1980).

Thus far, no structural differences have been observed among subunit sized RNA species which appear in virus particles and those which constitute the mRNAs for *gag* and *pol* polyproteins. Genomic RNA and mRNA have been distinguished functionally. If MLV RNA synthesis is inhibited by actinomycin D, viral proteins continue to be synthesized, while particles are assembled with cellular rather than viral RNA (Levin and Rosenak, 1976). A deletion mutant of RSV (missing ca. 150 nucleotides of non-coding sequences near the 5' terminus of the genome) produces functional mRNAs for *gag* and *pol* but fails to package genomic RNA into particles (Linial, et al., 1978).

Several questions remain unanswered about retrovirus gene expression: What constitutes the promoter for transcription of proviral DNA, and how is this process regulated? What is the primary transcriptional product and how does processing proceed? What is the precise nucleotide composition of viral mRNAs and where are the signals for translation located?

ENDOGENOUS RETROVIRUSES

Endogenous retroviruses exist as proviral elements within the germline DNA of many vertebrate species (Weiss, 1975; Aaronson and Stephenson, 1976).

Direct evidence of the genetic transmission of retroviruses in several species was obtained by molecular hybridization using uninfected cellular DNA and radioactively labeled virus-specific nucleic acids as probes (Lowy, et al., 1971; Gelb, et al., 1973; Rosenthal, et al., 1971; Neiman, 1972; Varmus, et al., 1972). The DNA of endogenous viruses is partially homologous to the genomes of exogenous viruses in the species of origin and is capable of directing synthesis of viral gene products (Weiss, 1975; Aaronson and Stephenson, 1976). Occasionally, retroviruses of certain species have infected the germline of secondary hosts. The endogenous viruses that have accomplished a cross species transfer can be traced to their proper species of origin using the criteria of nucleic acid homology (Todaro, 1975). The ubiquitous presence of endogenous viruses in vertebrates has raised as yet unanswered questions about their origin and function in the host genome.

Evolution of Retroviruses

The origin and function of endogenous viruses and the ancestral relationship between endogenous and exogenous retroviruses remain uncertain. Several hypothesis suggest

an explanation for the functional homology of viral proteins and genomic organization between groups of retroviruses in the apparent absence of nucleic acid homology (e.g., between avian and murine retroviruses) (Temin, 1975; Todaro, et al., 1975; Varmus, et al., 1981).

The "protovirus" hypothesis proposes that RNA-directed DNA synthesis, first discovered in retroviruses, exists in normal cells, and is a fundamental mechanism in cellular processes (e.g., differentiation). The enzyme allows for rapid mutation and rearrangement of genes and for transfer of genetic information between cells. Occasionally, these processes may generate retroviruses which escape from the cell to assume an autonomous existence (Temin, 1974a; Temin, 1974b). The presence of RNA-directed DNA polymerase in normal cells has not been conclusively demonstrated and evidence for the "protovirus" hypothesis remains circumstantial.

Temin has recently revised the protovirus hypothesis to include "cellular moveable elements" as fundamental components for retrovirus evolution (Temin, 1980). Sequence analysis of several different strains of retroviruses (including an avian endogenous retrovirus) has revealed structural similarities between retrovirus proviruses and bacterial (Calos and Miller, 1980), yeast (Cameron, 1979), and Drosophila (Finnegan, et al., 1978) moveable (transposeable)

elements. (For a recent survey of such elements, see Cold Spring Harbor Symp. Quant. Biol., Vol. 45, in press, 1980). Temin suggests that the similarities comprise a compelling argument for an evolutionary relationship between retroviruses and cellular moveable genetic elements. The significance of this theory is difficult to assess. There is no experimental evidence that retrovirus proviruses move within cellular DNA via a process resembling transposition, although retroviruses do transpose their genomes to other cells via a process that requires reverse transcriptase.

The "oncogene or virogene" hypothesis suggests that an ancient retrovirus, of unspecified origin, was present in early or pre-vertebrates (Todaro, et al., 1975). The retrovirus genes have evolved as the different species have evolved. If this hypothesis is correct then the relatedness of endogenous proviruses could be used as an assessment of the evolutionary relationships among many animals.

The lack of concordance of endogenous avian retrovirus DNA with established phylogenetic relationships of fowl argues against ancestral germline infection for establishment of the vast family of retroviruses (Frisby, et al., 1979; see below). It is possible that some kinds of proviral DNA other than endogenous avian retroviruses (e.g., endogenous

baboon virus), appeared in germlines by infection, or other mechanisms, sufficiently long ago to serve as useful markers for vertebrate evolution (Benveniste, et al., 1974).

A recent hypothesis has emerged from the results of biochemical studies of the cellular DNA of chickens and mice with respect to the amount, location, and structure of endogenous virus-related DNA (Hughes, et al., 1979; Cohen and Varmus, 1979; Steffen, et al., 1980). The hypothesis suggests that endogenous proviruses result from multiple, independent, but presumably infrequent infections of germ-line tissues after speciation. The similarities among retroviruses from many species would reflect their probable evolution from a single progenitor virus. The appearance of retroviruses as endogenous viruses would be a consequence of their capacity to establish their genomes in infected cells, including germ cells, as proviruses covalently integrated into host cell DNA (Varmus, et al., 1981 in press).

The evidence to support this hypothesis is derived primarily from data of three experimental systems: endogenous mammary tumor virus of mice, endogenous ecotropic AKV virus of AKR and asian mice, and endogenous avian leukosis virus of chickens (Cohen and Varmus, 1979; Steffen, et al., 1980; Hughes, et al., 1979). Endogenous proviruses related to mouse mammary tumor virus and Rous-associated

virus-0 (RAV-0; a member of the ALSV group) are found in varying numbers and are situated at different sites in the genomes of different inbred and outbred individuals of mice and chickens (Cohen and Varmus, 1979; Hughes, et al., 1979). The endogenous proviruses may also be found on multiple chromosomes (Chattopadhyay, et al., 1975; Tereba and Astrin, 1980; Tereba, unpublished; Smith, unpublished). These findings support the hypothesis that endogenous proviruses of mice and chickens have been acquired by relatively recent (post speciation) infection of germline cells. In chickens germline infection appears to have occurred sufficiently recently that proviruses are still segregating in contemporary domestic flocks. The observation that a provirus at one locus is common to most domestic chickens is discussed below. Frisby, et al., (1979) have reported that two varieties of jungle fowl, members of the genus Gallus (as are domestic chickens) are devoid of DNA related to RAV-0.

The above studies could lead to misunderstandings about the origins of endogenous retrovirus DNA if retroviral DNA exhibits a propensity to undergo deletion, rearrangement, or translocation. If proviruses or their flanking sequences were more likely to undergo major genetic changes than cellular genes then reinterpretation of the above studies would be necessary. (see Discussion)

The Function of Endogenous Retroviruses

To account for retention of functional proviral DNA in the vertebrate genome it has been proposed that the virus-specific DNA serves a purpose which confers some selective advantage to the host (Tolaro, 1974).

One idea is that the products of endogenous envelope genes serve as surface markers during cellular differentiation (Elder, et al., 1977), but other studies have failed to implicate the expression of endogenous viral genes in embryonic development (Strand, et al., 1977).

The ability of retroviruses to move from the germline of one species to that of another has prompted the idea that these viruses have been a force in the evolution of vertebrates, as horizontal transmission of bacterial genes by extrachromosomal elements may have played a major role in the evolution of procaryotes (Todaro, 1975).

The hypothesis that endogenous proviruses arose by infection and not from sequences common to primitive vertebrates does not favor any functional role for viral DNA in the growth and development of normal organisms. The presence of some strains of endogenous proviruses may even have pathological consequences for the host animal. In several strains of mice, mammary tumors develop late in life under the influence of genetically transmitted viruses (Benvelzen, 1974; Moore, 1976). An endogenous provirus, specifying the murine leukemia virus, AKV, is responsible for a 90%

incidence of leukemia in inbred mice of the AKR strain. Pathological consequences have never been associated with the presence of endogenous retroviruses in domestic chickens (Robinson, 1978; Crittenden, et al., 1979).

Expression of Endogenous Retroviruses

Expression of the genes of endogenous proviruses is controlled by three general classes of genetic determinants. (i) The provirus itself segregates as a genetic locus which determines the capacity of the cell to produce endogenous virus proteins or complete particles. The locus can be traced through the progeny of genetic crosses with molecular hybridization (Benveniste and Todaro, 1975; Cattopadhyay, et al., 1975; Astrin, 1980b). (ii) Cellular genetic determinants which are separate from the endogenous provirus can regulate the expression of viral genes (Weiss, 1975; Aaronson and Stephenson, 1976). In chickens, one of these determinants appears to be a cis-active regulator closely linked to the endogenous provirus at the ev2 locus (see below; Cooper and Temin, 1976). This suggests that the site at which an endogenous provirus is integrated in cellular DNA may have a considerable influence upon its expression. (iii) Many endogenous retroviruses cannot reinfect (or grow poorly upon) cells of the species from which they were derived (Todaro, 1975; Levy, 1978; Robinson, 1978). The restriction is accounted for by lack of appropriate cellular receptors and/or inefficient intracellular replication of the endogenous virus.

The expression of some endogenous viruses can be induced by a variety of agents, including halogenated pyrimidines (Lowy, et al., 1971; Robinson, et al., 1976), chemical carcinogens (Weiss, et al., 1971), ionizing radiation (Weiss, et al., 1971) and perhaps methylation inhibors (Jenkins and Coffin, unpublished). The specific mechanism of induction of endogenous retroviruses by these agenjs has not been described.

Endogenous viruses interact with exogenous viruses introduced by infection. Deletions in env can be complemented by phenotypic mixing with the product of an endogenous viral env gene (Weiss, 1969), and the genomes of endogenous and exogenous viruses can recombine (Weiss, Mason and Vogt, 1973).

Endogenous Avian Retroviruses of White Leghorn Chickens

Evidence that retroviruses were transmitted through the germline of chickens arose from genetic studies of uninfected white leghorn chickens. Some uninfected cells derived from white leghorn chick embryos were found to synthesize avian leukosis-sarcome virus-related (gag) proteins representing the group specific (gs) virion core proteins (Payne and Chubb, 1968), and/or a chicken helper factor (chf) which complemented env defective strains of Rous Sarcoma virus that lacked the virion envelope glycoproteins (Weiss, 1969; Hanafusa, et al., 1970). Furthermore, the release of an

infectious virus (RAV-0) from certain inbred lines suggested the existence of a complete viral genome in at least some chickens (Vogt and Friis, 1971). White leghorn chickens were phenotypically designated as gs^-chf^-, gs^+chf^+, gs^-chf^+, or V^+ depending upon expression of gag (gs), env (chf) or virus (V^+). It was soon established that DNA sequences related to the avian leukosis-sarcoma viruses existed in the genome of all normal chickens tested, whether or not these sequences were expressed (Rosenthal, et al., 1971; Baluda, 1972; Neiman, 1972; Varmus, et al., 1972).

It was also demonstrated that the overall concentration of virus-specific RNA in fibroblasts derived from chicken embryos of a determined phenotype correlated with the level of virus gag and env gene expression (Hayward, et al., 1973) (e.g., the concentration of virus-specific RNA in gs^-chf^- cells was extremely low ($<$1 copy/cell) and the concentration in gs^+chf^+ and gs^-chf^+ cells was 50-150 copies per cell (Hayward, et al., 1973; B. Baker, unpublished results). Furthermore, it was demonstrated that qualitative differences in the virus-specific RNAs existed between cells of different phenotype (Wang, et al., 1977). From these data, it was suggested that transcriptional controls were operative in the regulation of endogenous virus gene expression.

Biochemical analysis of the DNA isolated from white leghorn chickens has demonstrated that individual white leghorn chickens may harbor proviruses of endogenous retroviruses at any of at least 17 distinct genetic loci (denoted ev and assigned identifying numbers from 1 to 17) (Astrin, 1978; Astrin, et al., 1980b; Hughes, et al., 1980; Tereba and Astrin, 1980; Astrin and Tereba, personal communication) (Table 1). At each locus, the provirus has a constant and, in some instances, distinctive structure, but all of the proviruses are organized according to the structural principles first enunciated for proviruses of exogenous retroviruses (Hughes, et al., 1978; Sabran, et al., 1979). In particular, the order of the viral genes recapitulates that found in the RNA genomes of exogenous retroviruses (5'-gag-pol-env-3') and the proviruses are characteristically bracketed by terminal redundancies known as LTRs (long terminal repeats).

It has been demonstrated by breeding experiments, that several proviruses segregate with one of the four phenotypic designations for white leghorn chickens (see Table 1). Using this criteria, it was revealed that the proviruses at the ev1, ev4, ev5, or ev8 loci segregated in chickens of the $gs^{-}chf^{-}$ phenotype (i.e., no detectable virus gene expression), (Astrin, et al., 1980b; Tereba and Astrin, 1980). Ev15 contains only those sequences represented at the 3' and 5' termini of the virus genome and is therefore associated with a $gs^{-}chf^{-}$ phenotype (Hughes, et al., 1981; see below).

Table 1. Endogenous proviruses present in white leghorn chickens.

The endogenous provirus loci, abbreviated ev, of white leghorn chickens are listed in accordance with the nomenclature described by Astrin (Astrin, et al., 1980b; Hughes, et al., 1981). The phenotype associated with each locus is listed according to expression of gag, encoding the group specific (gs) structural proteins of the virion core; env, chicken helper factor (chf) encoding the glycoproteins the viral envelope; and expression of virus particles (V). The source of embryos is listed in column three. The RPRL chicken lines are inbred, whereas the Spafas, H&N and Kimber birds are non-inbred.

Table 1

Endogenous Retrovirus Loci in White Leghorn Chickens and Their Associated Phenotype

LOCUS	gs	chf	V	SOURCE
ev1	-	-	-	Ubiquitous
ev2	-	-	+	Line 7_2, Line 100 (RPRL)
ev3	+	+	-	Line 6_B (RPRL), H&N, Spafa, Kimber
ev4	-	-	-	H&N, Kimber, Spafa
ev5	-	-	-	H&N, Kimber, Spafa
ev6	-	+	-	Kimber, Spafa
ev7	-	-	+	Line 15_B (RPRL)
ev8	-	-	-	Kimber, Spafa
ev9	-	+	-	Kimber, Spafa
ev10	-	-	+	Line 15_{1_4} (RPRL)
ev11	-	-	+	Line 15_{1_4} (RPRL)
ev12	-	-	+	Line 15_1 (RPRL)
ev15	-	-	-	Kimber, H&N
ev16	-	-	-	Kimber

The ev3 provirus was the only provirus that segregated with the gs^+chf^+ phenotype (Astrin, et al., 1979; Astrin and Robinson, 1979). The ev6 or ev9 provirus were demonstrated to segregate with the gs^-chf^+ phenotype (Astrin, et al., 1980b). To date, only inbred white leghorn chickens have been designated V^+. The provirus at the ev2 locus segregates with a V^+ phenotype and is responsible for production of infectious RAV-0 virus. The proviruses at the ev7, ev10, ev11, and ev12 loci were found to segregate with the V^+ phenotype in various inbred chicken lines (Table 1).

In the present study, I have analyzed the structure of the viral DNA and RNA responsible for each of the white leghorn chicken phenotypes. The rigor of the analyses was enhanced by the use of molecularly cloned DNA to prepare radioactive DNA (cDNA) representing specific regions of the retrovirus genome. My data confirm and extend previous descriptions of the proviruses at eight ev loci and sustain the conclusion that the expression of endogenous avian retroviruses is subject to control by both viral and cellular determinants: the synthesis of viral RNA can be initiated at promoters contained within either viral DNA or adjacent cellular DNA; the position within the host genome of individual ev loci can strongly influence viral gene expression; and structural anomalies in proviral DNA can apparently give rise to aberrant metabolism of viral RNA subsequent to transcription.

MATERIALS AND METHODS

Cells and Viruses

Embryonated white leghorn chicken eggs were obtained from H & N Farms, Redmond, Washington and from Kimber Farm chickens maintained at the Worcester Foundation for Experimental Biology. Embryonated eggs from line 6_3 chickens and from a line 7_2 x line 100 mating were kindly provided by L.B. Crittenden, Regional Poultry Research Laboratories, East Lansing, Michigan. Japanese quail eggs were obtained from Life Sciences, Gainesville, Florida. Chicken embryo fibroblasts were propagated in Dulbecco's modified Eagle's medium (Gibco) supplemented with 10% tryptose phosphate broth, 5-10% fetal calf serum and 1% heat inactivated chicken serum as described by Robinson (1976).

Cells from an embryo derived from a line 7_2 x line 100 mating were tested for RAV-0 virus particle production by examining the culture medium for RNA-directed DNA polymerase activity. The subgroup E specificity of the particles produced was tested by infecting chicken (C/E) and quail (Q/B) fibroblasts with clarified culture medium (from line 7_2 x 100 cells that were positive for polymerase activity). After 3 passages of the infected (quail and chicken) cells, the media were tested for particles that contained reverse transcriptase. The medium from quail cells was positive,

whereas the medium from the chicken cells was negative for polymerase activity. We concluded that the line 7_2 x line 100 embryo was producing RAV-0 virus.

Preparation of High Molecular Weight DNA From Chicken Embryo Fibroblasts

Cells were washed 3 times with Tris-glucose (0.14 M NaCl, 5 mM KCl, 5 mM glucose, 25 mM Tris-HCl pH 7.4) and pelleted. The cells were resuspended in STE buffer (0.1 M NaCl, 20 mM Tris-HCl, 1 mM EDTA pH 7.4) at a final concentration of approximately 10^7/ml. The cells were lysed by addition of Proteinase K (200ug/ml) and 0.5% SDS and incubation for one hour at 37°C. The solution was extracted 3 times with phenol:chloroform (1:1). The aqueous phase was then extracted one time with chloroform, precipitated with 2.5 volumes of ethanol, and resuspended in 20 mM Tris-HCl pH 7.4 and 10 mM Na EDTA. Pancreatic RNase was added (final concentration, 20 ug/ml) and the solution was incubated at 37°C for 1 hour. The solution was extracted and precipitated as before. The DNA was resuspended in 10 mM Tris pH 7.4 and 0.1 mM EDTA.

Preparation of RNA From Chicken Embryo Tissue

I extracted RNA from chicken embryonic tissue after removing the head, limbs and viscera of 10-13 day old embryos. The tissue was placed in a dish and rinsed once with ice cold Tris-glucose and twice with ice cold STE,

and then homogenized in a Waring blender for 30 seconds in 50 mls STE buffer/embryo containing 250 ug/ml Proteinase K and 0.5% SDS. The homogenate was incubated for 30 minutes at 37°C and then extracted 3 times with an equal volume of phenol. Nucleic acid was precipitated by addition of NaOAc (final concentration 0.2 M) and 2.5 volumes of ethanol. The precipitate was collected by centrifugation and resuspended at 1 mg/ml of nucleic acid in DNase buffer (10 mM Tris pH 7.4, 10 mM $MgCl_2$). Iodo-acetic acid treated bovine pancreatic DNase I (Worthington) was added (final concentration 2 ug/ml), and incubated at room temperature for one hour. Sodium dodececylsulfate (final concentration 0.5%) and EDTA (final concentration 0.02 M) were added to stop DNase digestion. The solution was phenol extracted three times and ethanol precipitated as before. Polyadenylated RNA was selected by adsorption to columns of oligo (dT)-cellulose by centrifugation, resuspended at 1-3 mgs/ml in 20 mM Tris-HCl pH 7.4 and 10 mM EDTA and denatured (100°C 30 seconds). NaCl (final concentration 0.4 M) and SDS (final concentration 0.2%) were added and the RNA was applied to a 1-2 ml packed column of oligo (dT)-cellulose (T3 grade, Collaborative Research). The column was washed with 3 volumes of 0.4 M NaCl, 20 mM Tris-HCl pH 7.4, 10 mM EDTA, 0.2% SDS, and washed again with 3 volumes of 0.1 M NaCl, 20 mM Tris-HCl, 10 mM EDTA, 0.2% SDS. Poly-adenylated RNA

was eluted from the column in H_2O containing 0.2% SDS. The poly-adenylated fraction of RNA was recycled one time on oligo (dT) cellulose as described above. The final poly-A fraction was precipitated by the addition of NaOAc (final concentration 0.2 M) and 2 volumns of ethanol.

Preparation of RNA from Cultured Chicken Embryo Fibroblasts

Whole cell RNA was prepared as follows: Cells were washed 3 times with Tris-glucose and 2 times with STE. The cells were then removed from the petri dish and resuspended in STE containing 200 µg/ml Proteinase K and 0.5 SDS and incubated at 37°C for 30 min. The lysate was further processed as described for preparation of RNA from chicken embryo tissue.

RNA was extracted from nuclear and cytoplasmic fractions as follows: Cells were washed twice in Tris-glucose and twice in 0.1 x RSB (1 mM NaCl, 0.6 mM $MgCl_2$, 1 mM Tris-HCl pH 7.4), then scraped directly into a dounce, incubated for 5 min. and disrupted by 5-10 strokes with a dounce homogenized in 0.1 x RSB at 2 x 10^7 cells/ml. Cell breakage was monitered with a light microscope. The lysate was centrifuged at 10,000 rpm in a Sorval SS 5534 rotor for 1 minute to pellet nuclei. The supernatant (cytoplasmic fraction) was removed and digested with Proteinase K (200 µg/ml) in the presence of 0.5% SDS and 20 mM EDTA and then processed as whole cellular RNA, but without DNase I treatment.

The nuclear pellet was resuspended at 1/2 the original lysis buffer volume in 0.1 x RSB containing 0.2% NP40 and 0.1% DOC, and repelleted as before. The supernatent was removed and discarded, since it contained only 10% of the total cytoplasmic RNA and was not enriched or depleted for specific cytoplasmic viral RNAs. The nuclei were suspended in STE at 2 x 10^7 nucleic/ml. Proteinase K (final concentration 200 ug/ml), SDS (final concentration 0.5%) and EDTA (final concentration 10 mM) were added and the lysate was further processed as whole cellular RNA.

Molecular Hybrization in Solution

The method for establishing viral RNA concentration within cells as a function of Crt (concentration of RNA x time of hybridization with cDNA) has been previously described (Leong et al., 1972). In brief, increasing amounts of unlabeled RNA (50 ug-1 mg) were annealed with 1000 cpm of radioactively labeled cDNA at 68°C in a solution containing 0.6 M NaCl, 0.01 M Tris-HCl, pH 7.4, 0.01 M EDTA and 400 ug/ml Salmon sperm DNA in 50 to 100 ul reaction volumns. The mixtures were incubated in plastic tubes under mineral oil at 68°C for 4 to 5 days. RNA-DNA hybrids were measured as percentage of input cpm resistant to digestion with single-stranded specific nuclease S_1 from Aspergillus oryzae (Leong, et al., 1972). Background values of 2-4% were obtained when reaction mixtures containing labeled

cDNAs and no cellular RNA were incubated for 5 days at 68°C. The kinetics of hybridization were expressed as a function of Crt (Birnsteil, et al., 1972; Leong, et al., 1972) corrected to standard conditions of salt concentration (Britten and Smith, 1970) and the results were used to compute the concentration of viral RNA (Leong, et al., 1972).

Analysis of Cellular DNA with Restriction Endonucleases

DNA prepared as described above was cleaved with restriction endonucleases and fractionated by electrophoresis through 0.8% agarose gels in Tris-acetate buffer (pH 8.1) (Shank, et al., 1978). To visualize the DNA, the gels were stained with ethidium bromide and photographed under ultraviolet light. The DNA was then denatured, neutralized *in situ* and transferred onto nitrocellulose sheets in 6 x SSC as described by Southern (1975). After transfer, the nitrocellulose membranes were baked at 80°C under vacuum for 2 hours and then prehybridized in 50% formamide, 6 x SSC (1 x SSC is 0.15 M NaCl, 0.015 M sodium citrate), 0.01 M Hepes pH 7.0, 1 x Denhardts (0.02% polyvinylpyprolidone, 0.02% Ficoll, and 0.02% bovine serum albumin), 1 mg/ml yeast RNA, and 100 ug/ml Salmon sperm DNA for 12-18 hours at 41°C. The membranes were hybridized for 3 days with 10^5-10^6 cpm of ^{32}P-cDNA in the buffer described above. The hybridized filters were washed for 2 hours at 50°C in 0.1 x SSC and 0.1% SDS and subjected to autoradiography

under Kodak X-Omat x-ray film using lightning-plus screens at -70°C for 2 to 14 days. The nitrocellulose membranes were hybridized successively to multiple radioactive cDNAs. cDNA was removed from previously hybridized filters by soaking the paper one time in hybridization buffer for 1 hour at 68°C.

Analysis of Viral RNAs

Electrophoresis in denaturing agarose gels: RNAs were fractionated through 1.0%, 1.2% or 1.5% agarose gels containing 10 mM methyl mercury hydroxide, as described by Bailey and Davidson (1976). After examination of the gels by transillumination with ultraviolet light, the RNAs were transferred to diazobenzyloxymethyl (DBM) paper. The paper was prehybridized in the buffer as described for nitrocellulose membranes and also included 10 mg/ml of glycine. Filters were hybridized and subsequently treated as described above but the glycine was omitted from the hybridization buffer. Filters were hybridized successively to multiple radioactive cDNAs as described by Quintrell, et al. (1980). cDNA was removed from previously hybridized filters by soaking the paper three times in 99% formamide, 2.5 mM EDTA, 0.3% diethylpyrocarbonate for 15 min. at 68°C.

Velocity Sedimentation Analysis of RNA

RNA (50 ug-1 mg) was dissolved in TE (0.01 M Tris (pH 7.4), 0.01 M EDTA) containing 0.5% SDS, denatured by heating to 80-100°C for 2 min. quenched in ice water, then layered on a 15-30% sucrose gradient (in TE containing 0.01 M NaCl) and sedimented in the SW27 rotor for 27 hr. at 26,000 rpm. 35-40 fractions were collected from the bottom of the tube. Optical density (A_{260}) measurements were made during collection to locate the position of 18S and 28S ribosomal RNA. Virus-specific RNA was located in sucrose gradients by hybridizing RNA from equivalent volumes of each gradient fraction with the appropriate cDNAs (approximately 1000 cpm/fraction) for the same amount of time at 68°C under mineral oil in volumes of 0.01-0.1 ml of 0.6 M NaCl-0.002 M EDTA-0.05 M TRis (pH 7.4) containing calf thymus DNA (400 ug/ml). Hyridization was detected as resistance to hydrolysis by the single-strand-specific nuclease S1 (Leong, et al., 1972).

Preparation of Molecular Hybridization Probes

In order to facilitate characterization of virus-specific nucleic acids, a set of radioactive cDNAs specific for various portions of RSV and RAV-0 virus genomes have been developed. The cDNAs were prepared either from viral RNA or from molecularly cloned viral DNA by the enzyme reverse transcriptase. The following descriptions summarize the cDNAs used in the present study.

1) $cDNA_{b77}$, DNA complementary (cDNA) to the RNA genome of B77-RSV was synthesized using the endogenous DNA polymerase activity of detergent activated virus. The preparation and characterization of labeled cDNA for RSV has been described (Garapin, et al., 1973). Endogenous RNA-directed DNA polymerase activity was stimulated by addition of 0.025% Nonidet P-40 to purified B77 virus in the presence of 0.1 M Tris-HCl (pH 8.1), α-^{32}P-labeled dCTP (2×10^{-6}M) and unlabeled TTP, dGTP, and dATP (10^{-4}M). Reaction mixtures also contained 8 mM $MgCl_2$, 1% B-mercaptoethanol and 100 ug/ml actinomycin D. Reactions were incubated for 4 hours at 37°C and terminated by addition of 0.5% SDS. The DNA was purified by pronase treatment (500 ug/ml, 37°C, 45 min.), phenol extraction, precipitated with ethanol and digested with pancreatic RNase (100 ug/ml, 37°C, 1 hr. in 3 mM EDTA). The DNA was then fractionated on hydroxyapatite to remove double-stranded product and the single stranded DNA (cDNA) was treated with 0.2N NaOH overnight at 37°C to inactivate PNase (Garapin, et al., 1973). The solution was neutralized, precipitated with ethanol and resuspended in a small volumn of 0.02 M Tris-HCl, pH 7.4, 0.01 M EDTA for use in hybridization experiments. 2) $cDNA_3$, (specific activity $2\text{-}4 \times 10^9$cpm/ug) was prepared as described by Tal, et al., (1977b) and contained sequences complementary to

ca300 nucleotides adjacent to the poly A sequences at the 3' end of the RAV-0 genome. $cDNA_3$, was made from a reconstructed polymerase reaction containing denatured (100°C for 3 min.) 70S RAV-0 RNA and purified DNA polymerase, ^{32}P-dCTP (300 Ci/mmole; Amersham) and using oligo $(dT)_{12-18}$ (PL Biochemicals) as primer. Oligo (dT)-primed $cDNA_3$, was then purified by chromatography on oligo (dT)-cellulose after annealing to poly(A) as previously described (Tal, et al., 1977b; Shank, et al., 1978). 3) $cDNA_5$, was prepared by incubating 9 ug of 70S B77 or Pr-C viral RNA in a reaction mix similar to that used in the preparation of $cDNA_3$, except that oligo dT was omitted. The reaction product was separated from unincorporated nucleotides by gel filtration with Sepharose G-50 and loaded on an 8% acrylamide/urea gel (Maxam and Gilbert, 1977). A band representing $cDNA_{5(100)}$, was located by autoradiography and eluted from gel slices according to the procedure of Maxam and Gilbert (1977). 4) $cDNA_{env}$ was prepared as described by Tal., et al. (1977a) and represented sequences for most or all of the <u>env</u> deletion in the rdNY8 strain of RSV. Since this deletion probably removes most, if not all, of the <u>env</u> gene (Duesberg, et al., 1975; Tal, et al., 1977a), $cDNA_{env}$ is complementary to at least the bulk of <u>env</u>. $cDNA_{env}$ prepared in this manner was used in experiments for analysis of virus-specific RNA by hybridization in solution. $cDNA_{env}$

was also prepared from a molecularly cloned RSV DNA fragment (see below). The probe prepared in this fashion was used for hybridization to filters bound with RNA or DNA.

5) $cDNA_{gag-pol}$ is single-stranded DNA complementary to approximately 4700 nucleotides within the two genes located in the 5'-half of the genome. $cDNA_{gag-pol}$ was prepared by removing nucleotide sequences at the 3'-terminus of the RSV genome and sequences complementary to src and env by saturation hybridization as described previously (Stehelin, et al., 1976). The remaining DNA hybridized only with RNA from the 5'-half of the RSV genome and occupied approximately 47% of viral RNA when hybridized to saturation.

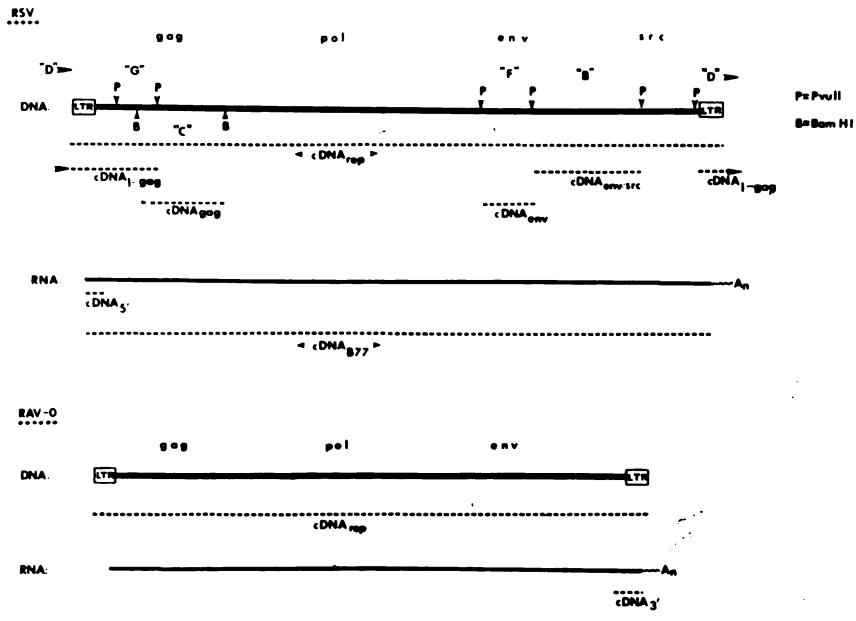
The following cDNAs were synthesized using molecularly cloned viral DNAs as templates. Circular RSV DNA from SR-A infected cells was isolated and cloned into λgt WES (DeLorbe, et al., 1980). The cloned insert representing the entirety of the RSV genome was then subcloned into the Sac I site of pBR322. RAV-0 DNA isolated from infected cells and representing the entire genome was cloned into λgt WES (unpublished data of S. Hughes). The RSV DNA, in pBR322, was digested with restriction endonuclease Bam Hl or Pvu II and the resulting fragments were subcloned into pBR322 (B. Baker and R. Parker, unpublished). The appropriate subcloned restriction fragments were purified from pBR322 after enzyme digestion either by velocity sedimentation in sucrose or by electrophoresis through agarose gels. DNA was recovered from

agarose gels by electroelution (Cordell, et al., 1979). The recovered viral DNA was denatured (100°C for 5 min. in 10 mM Tris-HCl pH 7.5, 0.1 mM EDTA) and a DNA copy was synthesized by purified avian myeloblastosis virus RNA-dependent DNA polymerase in the presence of α-^{32}P-dCTP (300 Ci/mmole), 50 mM Tris pH 8.1, 2 mM DTT, 5 mM $MgCl_2$ and 40 mM KCl, using as primer a 100-1000 fold mass excess of oligomers derived from calf thymus DNA. The specificity of the cDNA probes has been previously documented (Gonda, et al., 1981). Figure 5 illustrates the region of the viral genome that each cDNA represents. 6) $cDNA_{gag}$ was synthesized from the Bam Hl C fragment and represents 1.35 kpb located within the RSV gag gene. 7) $cDNA_{env}$ was synthesized from the Pvu II F fragment and represents 750 bp of the C-terminal domain of the env gene (DeLorbe, et al., 1980; Czernilofsky, et al., 1980). 8) $cDNA_{env-src}$ was synthesized from the Pvu II B fragment (1.6 kbp) and represents sequences of the C-terminus of env, the intercistronic region and ca. 1000 nucleotides of the src gene (Czernilofsky, et al., 1981). 9) $cDNA_{1-gag}$ was synthesized from the Pvu II D (1.0 kb) and c (600 bp) fragments which were cloned as one fragment (1.6 kb). The probe contains sequences for 2 copies of the RSV LTR structure, untranslated sequences at the 5' terminus of the viral genome (leader) and the amino-terminus of the gag gene.

Figure 5.

This figure illustrates a simplified genetic and physical map of RSV and the RAV-0 virus. The approximate positions of the structural genes for each virus, as well as the sequences represented in the various cDNAs used in this study are shown. $cDNA_{B77}$, synthesized from the B77 strain of RSV, represented the entire genome (Garapin, et al., 1973). $cDNA_{env}$, $cDNA_{gag/pol}$, and $cDNA_{5}$, were synthesized using viral RNA from either B77 or Pr-C virus. $cDNA_{3}$, was synthesized from RAV-0 viral RNA. Also shown are the restriction endonuclease cleavage sites defining the fragments of cloned RSV DNA used to prepare $cDNA_{gag}$, $cDNA_{env}$, $cDNA_{env-src}$, and $cDNA_{gag}$.

Figure 5. cDNAs used in the present study.



RESULTS

A. Analysis of Endogenous Avian Retrovirus DNA

1. Phenotype and genotype of white leghorn chickens used in the present study.

A number of endogenous virus loci (abbreviated *ev* and numbered 1-17) have been identified in white leghorn chickens, and each locus has been correlated with a phenotype for virus gene expression (Astrin, 1978; Astrin, et al., 1980b; Hayward, et al., 1980; Hughes, et al., 1981; Tereba and Astrin, 1980b; Astrin and Tereba, personal communication) (see Table 1). I have studied chicken embryos that represented four of these phenotypes: gs^- chf^-, gs^+ chf^+, gs^- chf^+, and V^+. Table 2 summarizes the characteristics of the embryos that I studied, and Figure 7 illustrates the structure of the pertinent endogenous proviruses based on previously reported data (Astrin, et al., 1980b; Hayward, et al., 1980; Hughes, et al., 1979; Hughes, et al., 1981; Shank, et al., 1981; McClements, et al., 1979; Skalka, et al., 1980) and results to be presented below.

I identified the endogenous proviruses of each of the embryos with the strategy first described by Astrin (1978). DNAs were cleaved with the *Sac* I restriction endonuclease, which generates distinctive DNA fragments from each of the *ev* loci; the cleaved DNA was then fractionated by agarose gel electrophoresis, transferred to nitrocellulose membranes, and hybridized to $cDNA_{rep}$ in order to detect virus-specific fragments.

Table 2

Genotype and Phenotype of Embryos
Used in the Present Study

Samples are grouped according to phenotype. The gs phenotype (gag gene expression) was determined directly for samples 1,2,10 by a complement fixation assay (Dougherty and DiStefano, 1966; Robinson, 1978) or by assaying ^{35}S-labeled cellular proteins by imunoprecipitation with rabbit antiserum to disrupted B77 virus (Oppermann, et al., 1977) (Samples 1,3,10). The chf phenotype (env gene expression) was determined directly for sample 10 as described by Robinson, et al. (1976). The gs and chf phenotypes of the other samples were assigned by knowledge of the phenotype of the parental birds. The V^+ phenotype (virus particle production) of sample 11 was determined directly (see Materials and Methods). The genotype (content of endogenous proviruses) of each embryo, with the exception of samples 6, 7 and 11, was determined by cleavage of cellular DNA with restriction enzymes Sac I, Eco RI and Bam Hl. The cleaved DNA was fractionated by agarose gel electrophoresis, transferred to nitrocellulose membranes, and bybridized to cDNAs for RSV and RAV-0 to detect virus-specific fragments.

Table 2

SAMPLE	PHENOTYPE	GENOTYPE	EMBRYO # OR LINE	NUMBER OF EMBRYOS ANALYZED	SOURCE
1	$gs^{-}chf^{-}$	ev1,5	#2453	1	H&N
2		ev1,5	#3578	1	H&N
3		ev1,3	K16	1	Kimber
4		ev1,3	#2497	1	H&N
5		ev1,3,15	#3429	1	H&N
6[a]	$gs^{+}chf^{+}$	N/D	line 6_3	6	RPRL
7[c]		N/D	#3406 #3431	2 1	H&N
8		ev1,6	y2, K18	1	Kimber
9[d]	$gs^{-}chf^{+}$	ev1,4,6,8	y2, K18	5	Kimber
10		ev1,9,15	#1865	1	Kimber
11	V^{+}	N/D	line 7_2 x 100	1	RPRL

(a) Six line 6_3 embryos were pooled prior to extraction of whole cellular RNA.

(b) Not determined.

(c) Three H&N $gs^{+}chf^{+}$ embryos were pooled prior to extraction of whole cellular RNA.

(d) Five Kimber embryos were individually analyzed for virus specific DNA and RNA.

Three virus-specific DNA fragments were common to all of the samples (see Fig. 6): 3.6 and 7.8 kb fragments, derived from c-src (the cellular homologue of the oncogene of Rous sarcoma virus described originally by Stehelin, et al. (1976) and demonstrated by Hughes, et al. (1979) to be contained on invariant restriction enzyme fragments among white leghorn chickens; and a 10.9 kb fragment, representing ev1 (Astrin, 1978). In addition, seven different endogenous proviruses were present in the eleven samples shown in Figure 6. The proviruses are identified by the nomenclature proposed by Astrin (1978); their distribution among the studied population is summarized in Table 2. Samples 5 and 10 also contained ev15 (formerly "element C", Hughes, et al., 1981) which can only be detected by hybridization with $cDNA_3$, and $cDNA_5$, and therefore does not appear in Figure 6. The proviruses at ev3, 8, and 9 each yielded an additional Sac I fragment that was not detected with $cDNA_{rep}$ but was detected with either $cDNA_3$, or $cDNA_5$, (see below).

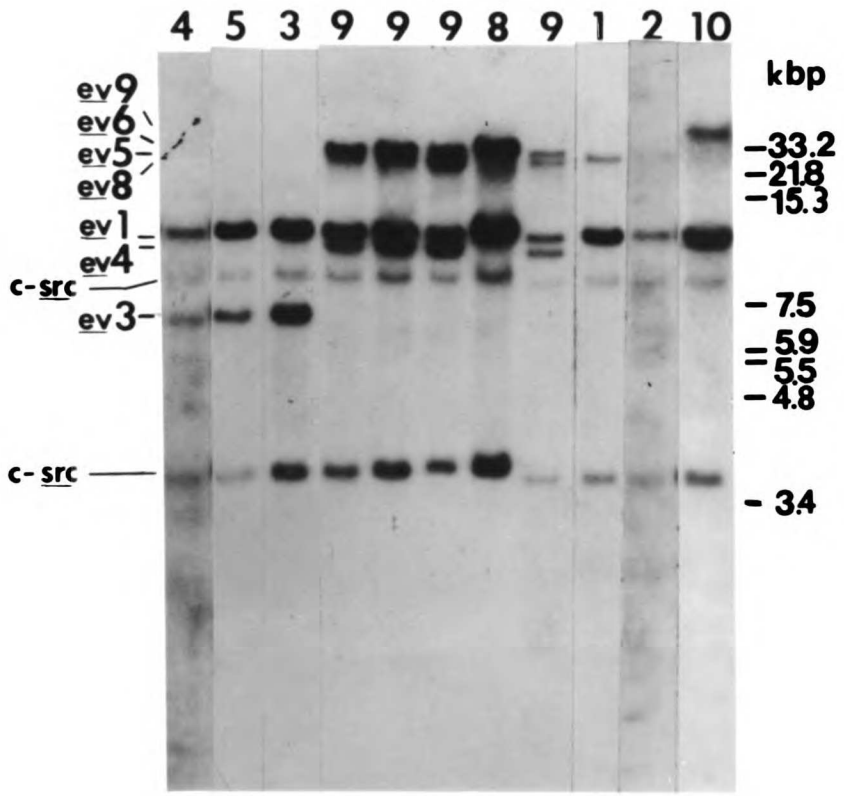
The identification of the Sac I DNA fragments was confirmed by further mapping with Eco RI and Bam H1 restriction endonucleases (see below). The results conformed to those reported previously (Astrin, et al., 1980b; Hughes, et al., 1981; Skalka, et al., 1980). The ev9 locus had not been mapped previously. My description of this locus is shown in Figure 7.

Figure 6

Analysis of the Endogenous Proviruses in White Leghorn Chickens with Restriction Enzyme Sac I

DNA (five to fifteen micrograms) from samples 1-5 and 8-10, described in Table 1, was digested to completion with Sac I. DNA fragments were separated on 0.8% agarose gels. The DNA was denatured, transferred to nitrocellulose, and hybridized with RSV-$cDNA_{rep}$. The numbers of the lanes correspond to the numbers assigned to the samples listed in Table 2. Only 4 of the 5 embryos that comprised sample 9 are shown. The approximate sizes of the major Sac I fragments for the different endogenous proviruses are as follows: ev1, 10.9 kb; ev3, 6.4 kb; ev4, 9.5 kb; ev5, 23 kb; ev6, 24 kb; ev8, 22 kb; and ev9, 34 kb. The sizes of the viral DNA fragments were calculated from the mobility of λphage DNA fragments produced by cleavage with Eco RI and Sal 1 enzymes (not illustrated).

Figure 6.


4 5 3 9 9 9 8 9 1 2 10
kbp
ev9
ev6
ev5
ev8
ev1
ev4
c-src
ev3
c-src
33.2
21.8
15.3
7.5
5.9
5.5
4.8
3.4

None of the samples that I analyzed carried a single provirus and it was therefore not possible to directly associate a provirus with a specific phenotype. By comparing genotype and phenotype of several samples, however, it was possible to identify which proviruses were responsible for the observed phenotype of each embryo. My results agree with the assignments made by others (Astrin, et al., 1980b; Hughes, et al., 1981). i) The presence of ev3 determines a gs^+chf^+ phenotype. ii) The presence of ev6 or ev9 determines a gs^-chf^+ phenotype. iii) Ev1, 4, 5, 8 and 15 are associated with a gs^-chf^- phenotype.

2. Structure and genetic composition of endogenous proviral DNAs.

Structural maps for endogenous avain retroviruses at eleven loci have been proposed (Hayward, et al., 1980; Skalka, et al., 1980; Hughes, et al., 1981). The maps were generated by comparing proviral DNA fragments produced by restriction enzyme digestion (Sac I, Eco RI, or Bam HI) of chicken embryo DNA that contained assortments of eleven endogenous proviruses (Hughes, et al., 1981). Fragments specific for each of the ev's were identified and cDNA probes from various portions of the RSV or RAV-0 genomes were used to order the specific viral DNA fragments for each provirus (Hughes, et al., 1981). Using a similar procedure, I was able to confirm the identity of the proposed structure of proviruses present in the embryos that I studied.

Figure 7 illustrates the proposed structure of nine endogenous proviruses. The provirus at ev2, as described by Hughes, et al. (1981), is included. For comparison, the maps of unintegrated linear DNA of the SR-A RSV strain and RAV-0 are shown (DeLorbe, et al., 1980; Shank, et al., 1981). The conservation of Sac I, Eco R1 and Bam H1 restriction enzyme sites within the endogenous proviruses exemplifies the similarities between the exogenous and endogenous proviruses (Hughes, et al., 1981; Skalka, et al., 1980) (see Introduction). These sites were used for identification and mapping of the different ev's. The restriction analyses that I performed were not sufficient to establish the map for each provirus. The results confirm the previously proposed structures of eight proviruses. The sizes of Sac I, Eco R1 and Bam H1 restriction fragments for eight proviruses and c-src are summarized in Tables 3, 4 and 5. A preliminary Eco R1 digest of samples 1-9 was performed to ascertain the similarities between samples that shared ev1, ev3, ev4, ev5, ev6 and ev8 (Fig. 8). Hybridization with $cDNA_{rep}$ revealed that sample 1 (ev1, 5), sample 5 (ev1, 3, 15), sample 8 (ev1, 6), and sample 9 (ev1, 4, 6, 8) were appropriate representatives for further mapping studies. Sample 10 contained the only representation of ev9 and was included for further study. The c-src Eco R1 fragments (23 kb, 17 kb and 13.6 kb) were present in all samples and are visualized in Figure 8 (Hughes, et al., 1979).

Figure 7.

A comparison of the restriction maps of RSV, RAV-0 and 9 endogenous proviruses of white leghorn chickens. The Eco Rl, Sac I and Bam Hl sites in RSV, RAV-0 and 9 endogenous proviruses are shown schematically and indicated by arrows. The drawings are approximately to scale, except for the LTRs which are indicated with solid boxes. The sizes (in kilobase pairs) of the Sac I, Eco Rl, and Bam Hl restriction fragments are summarized in Tables 3, 4 and 5, respectively. Evl5 is detected only with $cDNA_3$, and $cDNA_5$, and may have a structure similar to a LTR. It is illustrated in this fashion for convenience.

Figure 7.

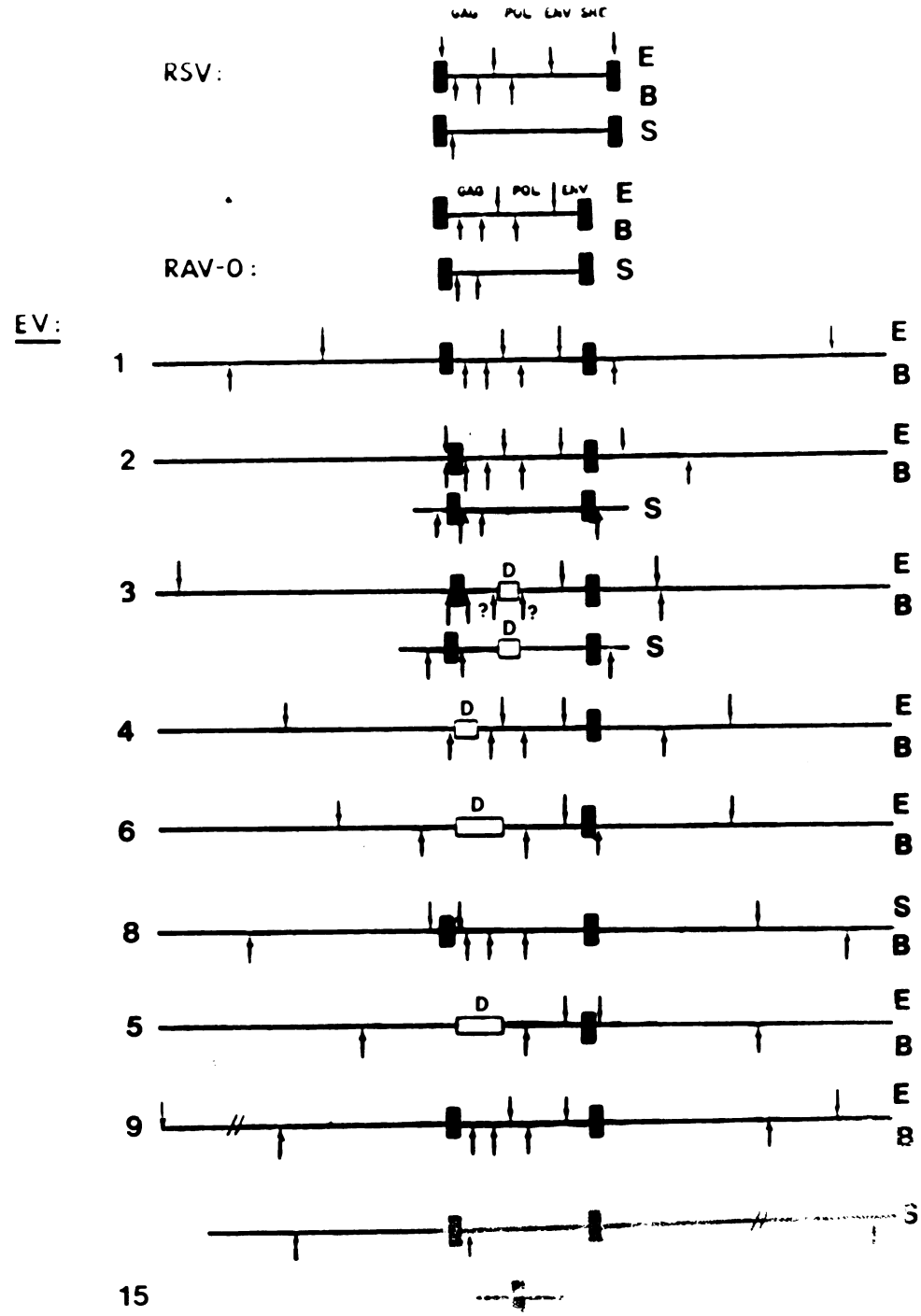
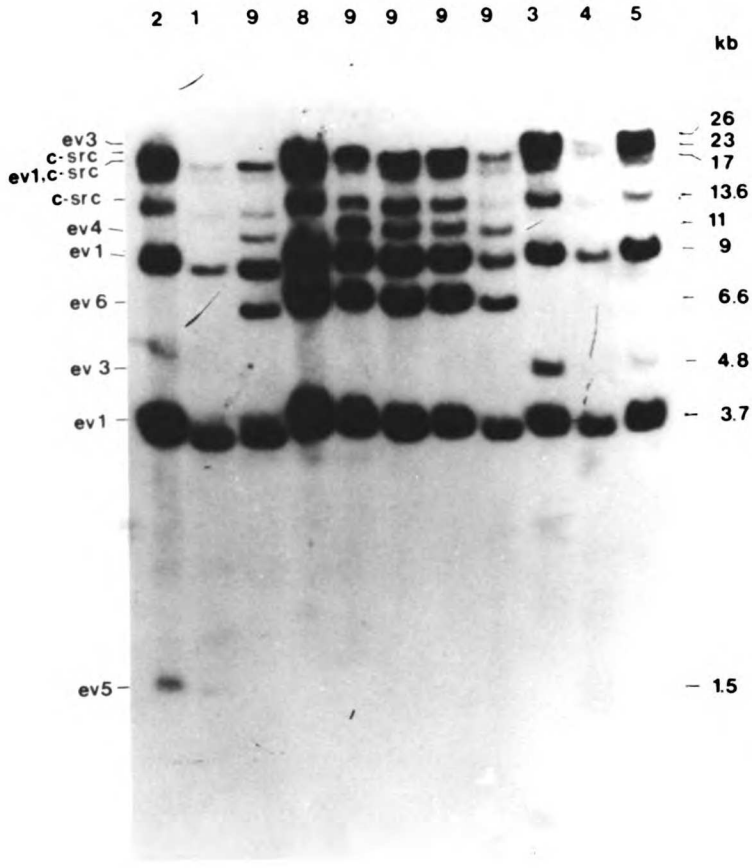

RSV:
E
B
S
GAG
POL
ENV
RAV-0:
EV:
1
2
3
4
6
8
5
9
15
D
?
E Eco RI
B Bam HI
S Sst I
D Deletion

Figure 8.

Analysis of DNA from samples 1-5 and 8-9 with Eco Rl. Five, ten, or 15 ug of DNA, described in Table 2 was digested to completion with Eco Rl and analyzed as in Figure 6 with $cDNA_{rep}$ for RSV. The fragments specific for each provirus are summarized in Table 4. The numbers of the lanes correspond to the numbers assigned to the samples listed in Table 2.

Figure 8.


2 1 9 8 9 9 9 9 3 4 5
kb
ev3
c-src
ev1,c-src
c-src
ev4
ev1
ev6
ev3
ev1
ev5
26
23
17
13.6
11
9
6.6
4.8
3.7
1.5

The 17 kb fragment comigrated with an evl fragment of similar size. The provirus at the evl locus was present in all embryos that I examined and was characterized by the presence of a 10.9 kb Sac I restriction fragment that hybridized to all cDNAs tested (Fig. 6; Fig. 9; panels A-E). Detailed restriction mapping of cloned DNA fragments that contained evl demonstrated that Sac I cleaved the proviral DNA once, close to the left-ward virus-cell junction (Skalka, et al., 1980). A "second" evl Sac I fragment has not been detected by myself or others (Astrin, 1978; Hughes, et al., 1981; Hayward, et al., 1980), presumably because the left-ward fragment comigrates with the 10.9 kb fragment, or is too small to permit detection.

Samples 1 and 2 yielded the least complex Eco Rl pattern and the three evl specific fragments (9 kb, 3.7 kb, and 17 kb) were visualized by hybridization with $cDNA_{rep}$ (Fig. 8). The 9 kb fragment was also detected with $cDNA_{3'}$, $cDNA_5$, and $cDNA_{gag}$ (Fig. 10; panel A). The 3.7 kb fragment was detected with $cDNA_{env}$, and the 17 kb fragment was detected with $cDNA_{3'}$, $cDNA_5$, and $cDNA_{env}$. It was concluded that the 9 kb and 17 kb DNAs were the left and right-ward junction fragments respectively and that the 3.7 kb fragment was the internal Eco Rl fragment of evl. Bam Hl generated 4 evl specific fragments of 14.5 kb, 1.35 kb, 1.8 kb and 4.8 kb which were detected with $cDNA_{rep}$ (Fig. 11; panels A-E).

Samples 1 and 2 carried ev5 which was characterized

Figure 9.

Analysis of DNA from samples 1, 5, 8, 9 and 10 (Table 2) by digestion with Sac I and hybridization with cDNAs representing various regions of the RSV genome. Five or ten micrograms of DNA was digested to completion with Sac I and was analyzed as in Figure 6. Each panel (A-E) represents a single lane on a filter that was successively hybridized with the designated cDNAs as described in Materials and Methods. Panel A, sample 2 (evl, 5); panel B, sample 5 (evl, 3, 15); panel C, sample 8 (evl, 6); panel D, sample 9 (evl, 4, 6, 8); panel E, sample 10 (evl, 9, 15). (See Table 3 for a summary of these data.)

Figure 9.

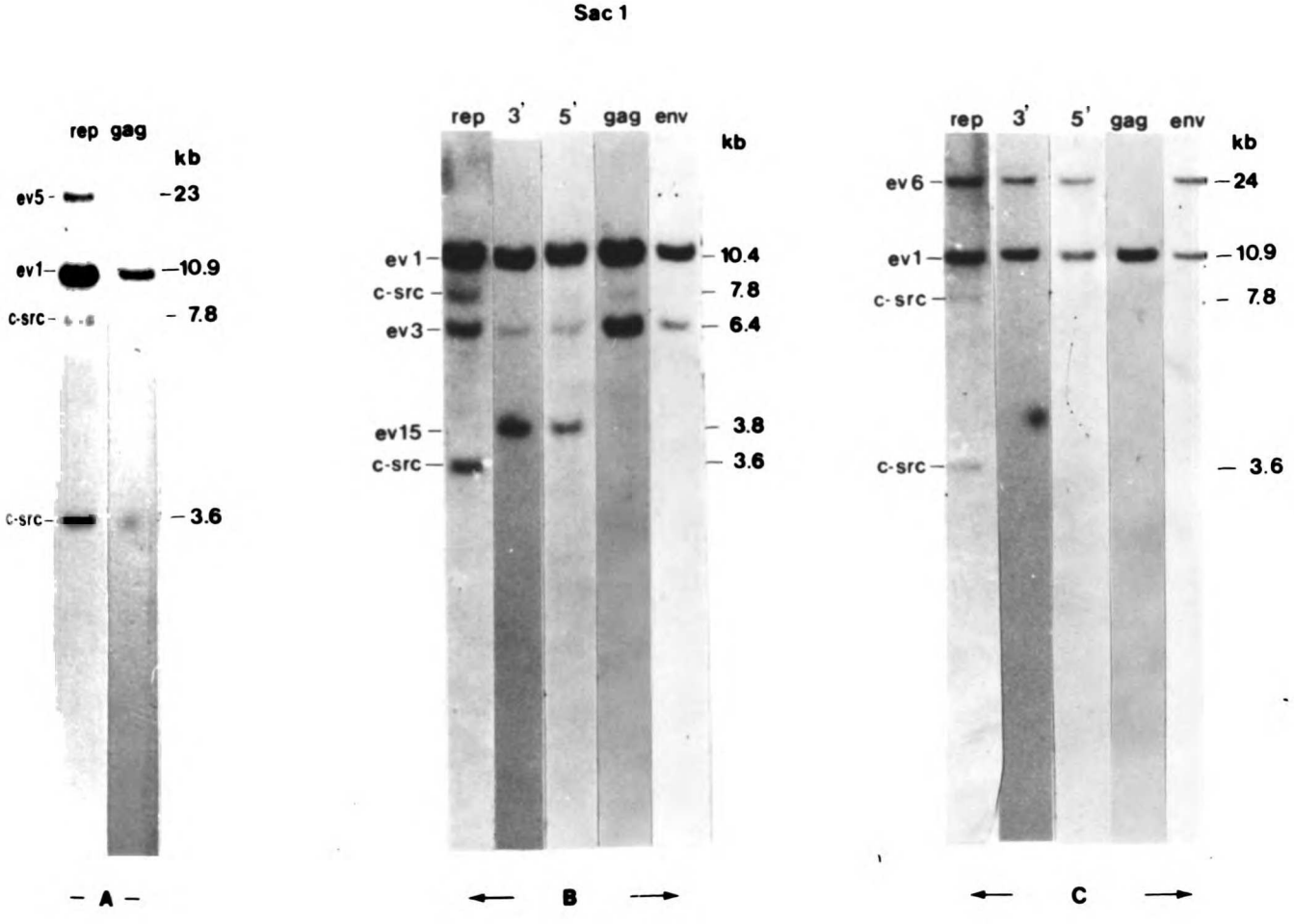

Sac 1
rep
gag
kb
ev5
ev1
c-src
23
10.9
7.8
3.6
A
rep
3'
5'
gag
env
kb
ev 1
c-src
ev3
ev15
c-src
10.4
7.8
6.4
3.8
3.6
B
rep
3'
5'
gag
env
kb
ev 6
ev1
c-src
c-src
24
10.9
7.8
3.6
C

Figure 9. (Cont'd.)

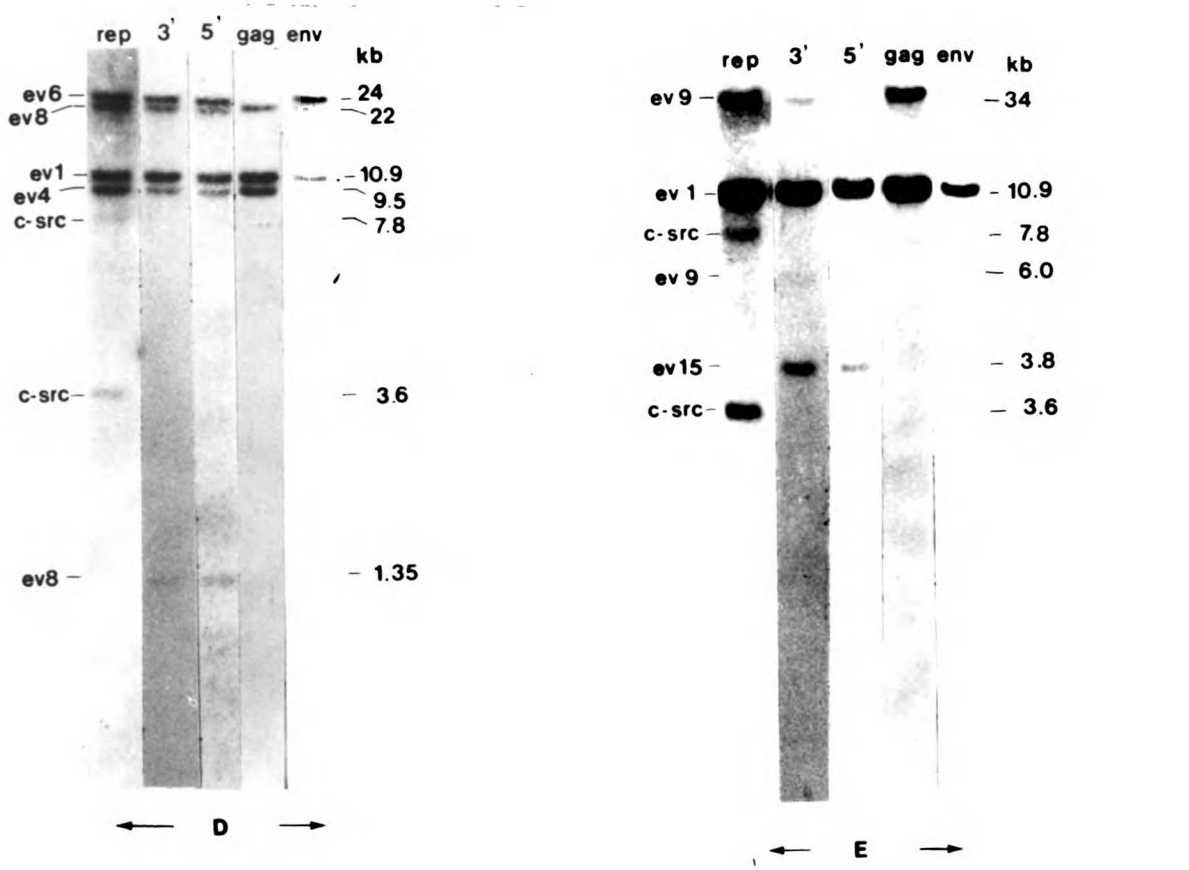


Table 3. The size of *Sac* I fragments of the endogenous proviruses and c-*src*.

The *Sac* I fragments attributed to each provirus are listed in kilobase pairs. The reactivity of a fragment with each cDNA is indicated with a + or - (n.d. = not determined).

Table 3

Locus	Sac I (kb)	cDNA rep	3'	5'	gag	env
ev1	10.9	+	+	+	+	+
ev3	6.4	+	+	+	+	+
	1.4	-	n.d.	n.d.	n.d.	n.d.
ev4	9.5	+	+	+	+	+
ev5	23.0	+	n.d.	n.d.	-	n.d.
ev6	24.0	+	+	+	-	+
ev8	22.0	+	+	+	+	+
	1.35	-	+	+	-	-
ev9	34.0	+	+	+	+	+
	6.0	-	+	+	-	-
ev15	3.8	-	+	+	-	-
c-src	7.8	+	-	-	-	-
	3.6	+	-	-	-	-

Figure 10.

Analysis of DNA from samples 1, 5, 8, 9, 10 (panels A-E respectively; Table 2), by digestion with Eco R1 and hybridization with various cDNAs. Ten micrograms of DNA was digested with Eco R1 and was analyzed as in Figure 6 and Figure 9. The left most lane in panel A represents results obtained with sample 1 in separate experiment. It is presented here in order to demonstrate the presence of the 1.5 kb ev5 Eco R1 fragment. The cDNA used for hybridization in the left-most lane was deficient for src sequences and therefore the c-src 23 kb and 13.6 kb fragments were not detected (see Table 4).

Figure 10.

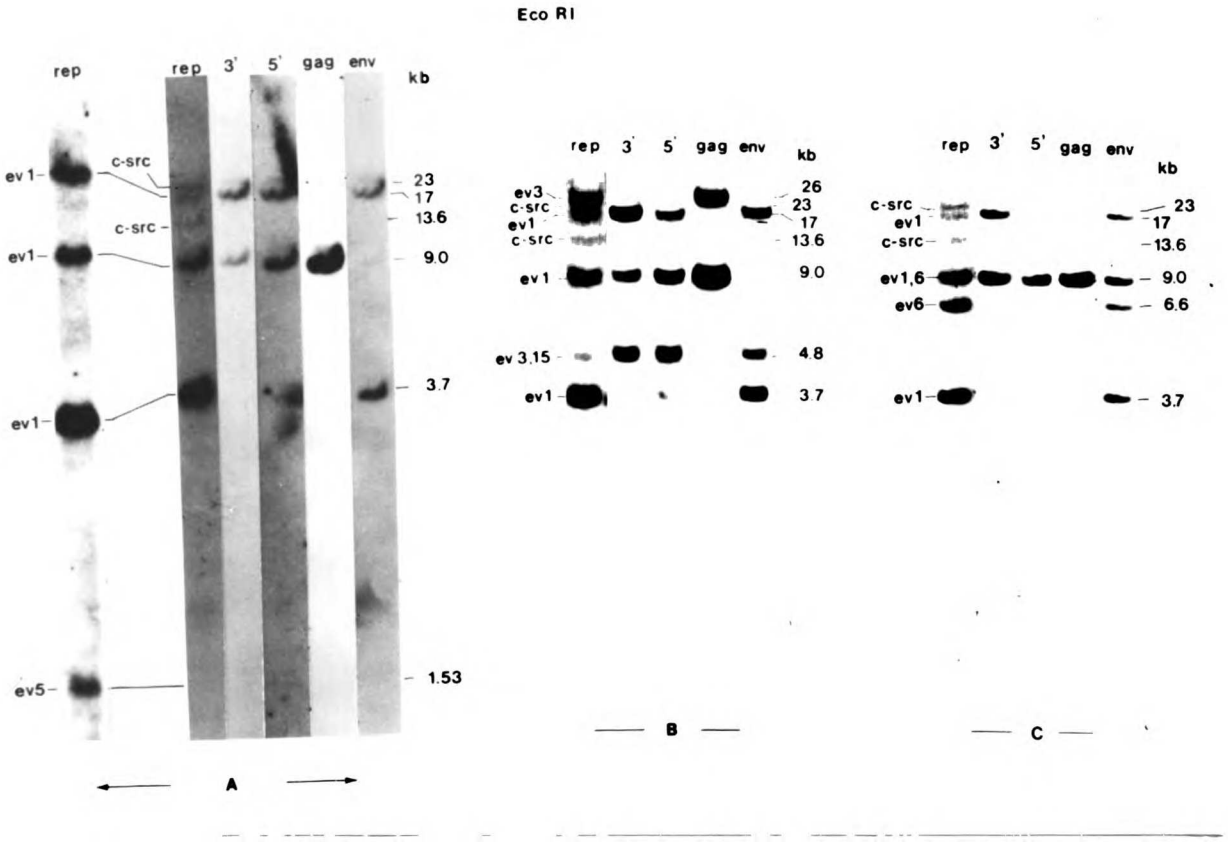
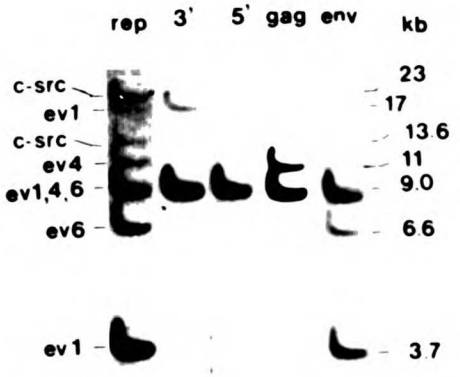
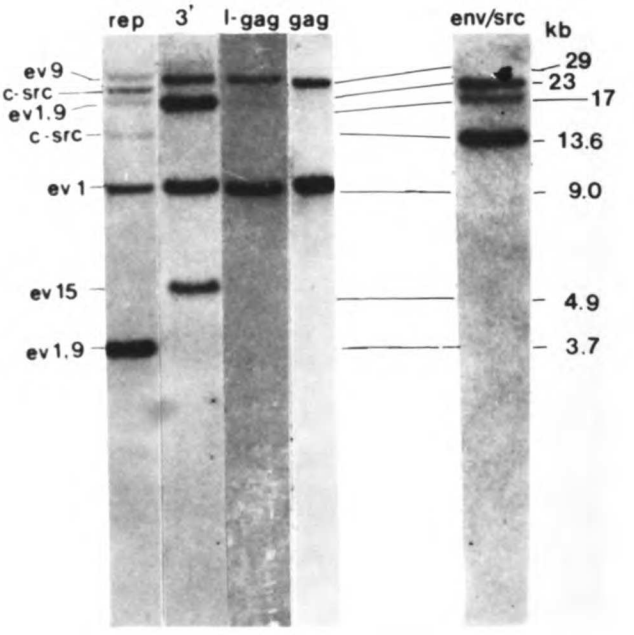

Eco RI
rep
rep 3' 5' gag env
kb
ev1
c-src
c-src
ev1
ev1
ev5
23
17
13.6
9.0
3.7
1.53
A
rep 3' 5' gag env
kb
ev3
c-src
ev1
c-src
ev1
ev 3,15
ev1
26
23
17
13.6
9.0
4.8
3.7
B
rep 3' 5' gag env
kb
c-src
ev1
c-src
ev1,6
ev6
ev1
23
17
13.6
9.0
6.6
3.7
C

Figure 10. (Cont'd.)



D



E

Table 4. The size of Eco R1 fragments of the endogenous proviruses and c-src.

The Eco R1 fragments attributed to each provirus are listed in kilobase pairs. The reactivity of a fragment with each cDNA is indicated with a + or - (n.d. = not determined).

Table 4

Locus	Eco R1 (kb)	cDNA rep	3'	5'	gag	env	env/-src	1/gag
ev1	9.0	+	+	+	+	-	-	+
	3.7	+	-	-	-	+	-	-
	17.0	+	+	+	-	+	+	-
ev3	26.0	+	+	+	+	+	n.d.	n.d.
	4.8	+	+	+	-	+	n.d.	n.d.
ev4	11.0	+	-	-	+	-	n.d.	n.d.
	3.7	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	9.2	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
ev5	1.5	+	-	-	-	-	n.d.	n.d.
ev6	6.6	+	-	-	-	+	n.d.	n.d.
	9.0	+	+	+	-	+	n.d.	n.d.
ev8	?	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
ev9	29.0	+	+	+	+	n.d.	-	+
	3.7	+	-	n.d.	-	n.d.	-	-
	17.0	+	+	n.d.	-	n.d.	+	-
ev15	4.9	-	+	+	-	n.d.	-	-
c-src	23.0	+	-	-	-	-	+	-
	17.0	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	13.6	+	-	-	-	-	+	-

Figure 11.

Analysis of DNA from sample 1, 4, 9 and 10 (panels A, B, D and E respectively; Table 2) by digestion with Bam Hl and hybridization with various cDNAs. The results from sample 9 (panels D_1 and D_2) were obtained by analysis of two separate nitrocellulose membranes. The results of hybridization of sample 10 with $cDNA_{rep}$ (panel E) was obtained by analysis of a different DNA-bound nitrocellulose membrane. These data are summarized in Table 5.

Figure 11.

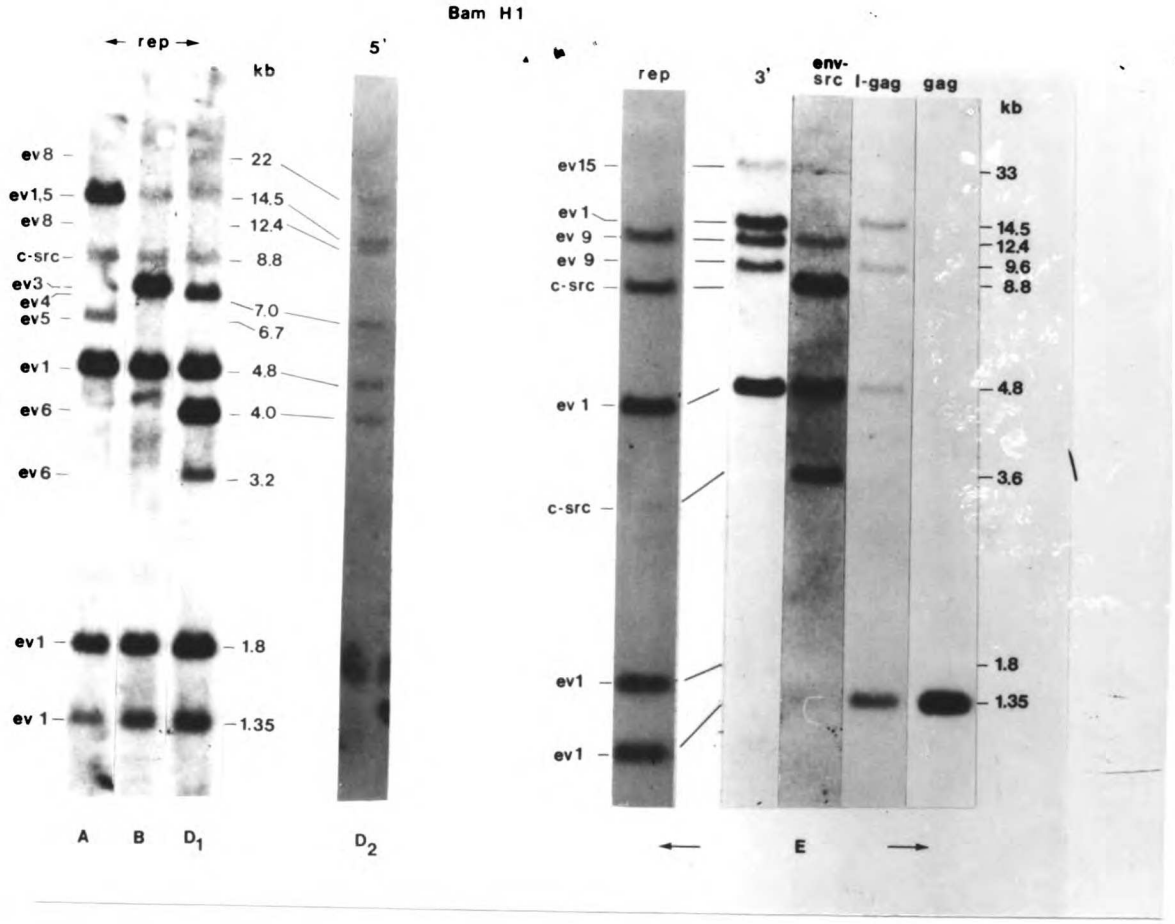

Bam H1
rep
5'
kb
ev8
ev1,5
ev8
c-src
ev3
ev4
ev5
ev1
ev6
ev6
ev1
ev 1
22
14.5
12.4
8.8
7.0
6.7
4.8
4.0
3.2
1.8
1.35
A
B
D1
D2
rep
3'
env-
src
l-gag
gag
kb
ev15
ev 1
ev 9
ev 9
c-src
ev 1
c-src
ev1
ev1
33
14.5
12.4
9.6
8.8
4.8
3.6
1.8
1.35
E

Table 5. The size of Bam Hl fragments of the endogenous proviruses and c-src.

The Bam Hl fragments attributed to each provirus are listed in kilobase pairs. The reactivity of a fragment with each cDNA is indicated with a + or - (n.d. = not determined).

Table 5

Locus	Bam HI (kb)	cDNA rep	3'	5'	gag	env	env/-src	1/gag
ev1	14.5	-/+	+	n.d.	-	n.d.	-	-
	1.35	+	-	n.d.	+	n.d.	-	+
	1.8	+	-	n.d.	-	n.d.	-	-
	4.8	+	+	n.d.	-	n.d.	+	+
ev3	1.0	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	7.0	+	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
ev4	1.0	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	1.8	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	7.0	+	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
ev5	6.7	+	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	14.5	+	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
ev6	3.2	+	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	4.0	+	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
ev8	12.4			+				
	1.35			-				
	1.8	n.d.	n.d.	-	n.d.	n.d.	n.d.	n.d.
	22.0			+				
ev9	9.6	-	+	n.d.	-	n.d.	-	+
	1.35	+	-	n.d.	+	-	-	+
	1.8	+	-	n.d.	-	-	-	-
	12.4	+	+	n.d.	+	n.d.	+	-
ev15	33.0	-	+	n.d.	-	n.d.	-	-
c-src	8.8	+	-	n.d.	-	n.d.	+	-
	3.6	-/+	-	n.d.	-	n.d.	+	-

by the presence of a Sac I 23 kb fragment that hybridized with $cDNA_{rep}$ but not $cDNA_{gag}$ (Fig. 6; Fig. 9; panel A) and substantiated previous observations that ev5 was deleted for gag sequences. The ev5 provirus present in samples 1 and 2 was further established by a 1.5 kb Eco Rl fragment (Fig. 8; Fig. 10; panel A) (Hughes, et al., 1981). Bam Hl generated two ev5 specific fragments, 14.5 kb and 6.7 kb (Fig. 11, panel A). Hughes, et al. (1981) have mapped the ev5 specific fragments and the provirus is illustrated in Figure 7.

Samples 3, 4 and 5 carried ev1 and ev3. Sample 5 was used for analysis of ev3. (Sample 5 also carried ev15; see below.) Ev3 was characterized by a 6.4 kb Sac I fragment that hybridized to all cDNA probes (Fig. 6; Fig. 9; panel B). A 1.4 kb Sac I fragment generated from ev3 (described by Hughes, et al., 1981) was not detected in the analysis presented here, but was detected in another sample with $cDNA_3$, and $cDNA_{1-gag}$ (data not shown). Eco Rl generated 2 ev3 fragments, 26 kb and 4.8 kb (Fig. 8; Fig. 10, panel B). The 26 kb fragment was detected with all cDNAs tested. The 4.8 kb fragment was detected with $cDNA_{rep}$, $cDNA_3$,, $cDNA_5$, and $cDNA_{env}$. It was concluded from these analyses and those of others (Hughes, et al., 1981) that the right-ward Eco Rl site is conserved in ev3 and that the sequences including the left-ward Eco Rl restriction

enzyme site were deleted from the ev3 provirus (Fig. 7). Ev3 is also characterized by a Bam Hl fragment of 7 kb (Fig. 11, panel B) (Hughes, et al., 1981).

Samples 8 and 9 contained ev1 and ev6. The five embryos included in sample 9 also carried ev4 and ev8. Ev6 was characterized by a single Sac 1 DNA fragment of 24 kb (Fig. 6; Fig.9; panels C and D). The ev6 fragment hybridized to all cDNAs except $cDNA_{gag}$ (Fig. 9, panels C and D). Eco Rl cleaved within the ev6 provirus and two specific fragments (9 kb and 6.6 kb) were detected with $cDNA_{rep}$ and $cDNA_{env}$ (Fig. 9, panels C and D). It was presumed that the 9 kb fragment also hybridized to $cDNA_3$, and $cDNA_5$,. This suggested that the ev6 provirus was deleted for the left-ward internal Eco Rl site and one of the LTR's. Digestion with Bam Hl yielded two ev6 fragments that were detected with $cDNA_{rep}$ (4.0 kb and 3.2 kb; Fig. 11, panel D) (Hughes, et al., 1981), only one of which (the 4.0 kb fragment) annealed with $cDNA_5$, (Fig. 11, panel D). These observations suggested that the deletion in ev6 extended from the left-ward LTR through the Eco Rl site near the gag/pol gene boundary (Fig. 7; Hughes, et al., 1981). The five embryos of sample 9 contained ev1, 4, 6 and 8 (Fig. 6). Since ev4 and ev8 were not distributed individually in other samples, it was not possible to distinguish Eco Rl or Bam Hl fragments specific for each

provirus. The assignments presented here are based on those made by Hughes, et al. (1981).

The ev4 provirus is contained on a 9.5 kb Sac I fragment and is detected by hybridization to all cDNA probes (Fig. 6; Fig. 9, panel D). Eco Rl generated two ev4 fragments (11 kb and 9.0 kb; Fig. 10, panel D). The 11 kb fragment was detected by $cDNA_{rep}$ and $cDNA_{gag}$. The 9.0 fragment hybridized with $cDNA_{rep}$, $cDNA_{3'}$, $cDNA_{5'}$ and $cDNA_{env}$. Ev4 is also distinguished by a 7.0 kb Bam Hl fragment (Fig. 11, panel D). The structure of the ev4 provirus was deduced by Hughes, et al. (1981), and is illustrated in Figure 7.

Two ev8 specific fragments are generated by Sac I (22 kb and 1.35 kb; Fig. 6; Fig. 9, panel D). The 1.35 fragment annealed with $cDNA_{3'}$ and $cDNA_{5'}$ (Fig. 9, panel D). the 22 kb fragment annealed with all cDNAs tested. Eco Rl fragments for ev8 have not been observed by myself or others (Hughes, et al., 1981). It is presumed that they comigrated with Eco Rl fragments of ev1, ev6 or ev4. Analysis of ev8 by Bam Hl digestion and hybridization with specific probes led Hughes, et al. (1981) to conclude that the provirus did not contain any major deletions.

The embryo of sample 10 carried ev1, ev9 and ev15. The ev9 provirus was known to be contained on a large major Sac I fragment (Astrin, et al., 1980b), but the location

of restriction sites with the provirus has not been previously described. Ev9 was cleaved by Sac I once and two fragments (34 kb and 6.0 kb) were generated. The 34 kb fragment reacted with all cDNAs tested. The 6.0 kb DNA was detected with $cDNA_3$, and $cDNA_5$, (Fig. 9, panel E) and $cDNA_{1\text{-}gag}$ (data not shown). Eco Rl generated 2 ev9 specific fragments (ca. 29 kb and 17 kb). The 29 kb fragment hybridized to $cDNA_{rep}$, $cDNA_{3'}$, $cDNA_{gag}$ and $cDNA_{1\text{-}gag}$. The 17 kb fragment comigrated with the right-ward evl Eco Rl fragment and reacted $cDNA_{rep}$, $cDNA_{3'}$, and $cDNA_{env/src}$ (Fig. 10, panel E) and $cDNA_{env}$ (data not shown). Bam Hl generated 2 ev9 virus-cell junction fragments, ca. 9.6 kb and 12.4 kb. The 12.4 kb fragment reacted with $cDNA_{rep}$, $cDNA_{3'}$, and $cDNA_{env/src}$ (Fig. 11, panel E). The 9.6 kb fragment hybridized with $cDNA_3$, and $cDNA_{1\text{-}gag}$ (Fig. 11, panel E). It was concluded that the 12.4 fragment was the right-ward virus-cell junction fragment and that the 9.6 kb fragment was the left-ward junction fragment. Internal Bam Hl fragments of ev9 comigrated with those of evl. I concluded that the provirus contained no major deletions.

Ev 15 was carried in the DNA isolated from samples 5 and 10. It is present on a 3.8 kb Sac I fragment (Fig. 9, panels B and E), a 4.9 kb Eco Rl fragment (Fig. 10,

panels B and E), and a 33 kb Bam Hl fragment (Fig. 11, panel E). The ev15 locus is detected only with $cDNA_3$, and $cDNA_5$, and may structurally resemble an LTR.

B. Characterization of Virus-Specific RNA in Uninfected White Leghorn Chicken Cells

1. Identification of virus-specific RNA in normal chicken embryos.

The overall concentration of virus specific RNA was determined for certain embryos of defined phenotype by the kinetics of hybridization in solution with $cDNA_{env}$ and $cDNA_{gag\text{-}pol}$. I analyzed polyadenylated RNA from either whole cells or nuclear and cytoplasmic fractions, using embryos of defined phenotype and genotype (see Table 2).

RNAs were fractionated by electrophoresis in agarose gels under denaturing conditions, transferred to and immobilized on chemically substituted paper, and hybridized with radioactive DNAs (cDNAs) specific for various portions of avian retrovirus genomes. For some analyses RNAs were denatured and fractionated by sedimentation in sucrose. It was possible to identify and assign virus-specific RNAs to ev1, ev3, ev6, and ev9, and to proviruses presumably derived from transcripts of ev2.

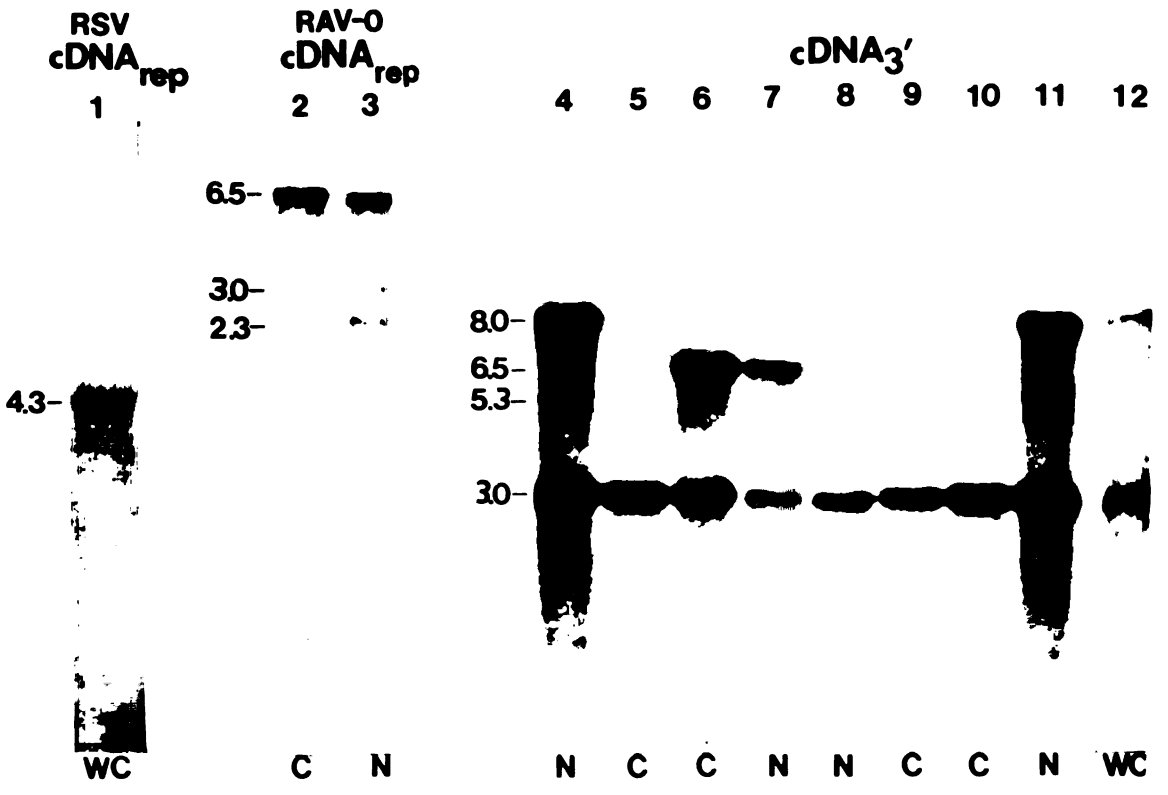
As expected, all of the samples contained a 4.3 kb RNA that is produced from the cellular homologue (c-src) of the RSV oncogene (src) and that can therefore be detected with either $cDNA_{rep}$ for RSV or a cDNA for src (Wang, et al., 1977; Spector, et al., 1978; Bishop, et al., 1980; Hayward, et al., 1980). An example is shown for sample 1 (Fig. 12, lane 1). In order to screen for RNAs produced from the

Figure 12. Virus-specific RNAs in embryos with the gs^-chf^-, gs^+chf^+, gs^-chf^+ ov v^+ phenotype.

Polyadenylated RNAs were prepared from whole cells (WC) or nuclear (N) and cytoplasmic (C) fractions of cultured embryo fibroblasts as described in Materials and Methods. The RNA was fractionated by electrophoresis in agarose gels containing methyl mercury hydroxide, transferred and immobilized on chemically substituted paper, and hybridized with RSV $cDNA_{rep}$, RAV-0 $cDNA_{rep}$ or RAV-0 $cDNA_{3'}$. The probes are indicated above each panel. The size of each viral RNA, expressed as kilobases (kb), was determined by the relative electrophoretic mobilities of chicken ribosomal RNAs and is indicated along the sides of the panels. Lane 1, sample 1 (ev1, 5) whole cell RNA; lane 2, sample 3 (ev1, 3) cytoplasmic RNA; lane 3, sample 3 nuclear RNA; lanes 4 and 10, sample 19 (ev1, 9, 15) nuclear RNA; lanes 5 and 11, sample 10 nuclear RNA; lane 6, sample 4 (ev1, 3) cytoplasmic RNA; lane 7, sample 4 nuclear RNA; lane 8, sample 8 (ev1, 6) nuclear RNA; lane 9, sample 8 cytoplasmic RNA; lane 12, sample 11 (ev1, 2) whole cell RNA.

Each lane contains 10 ug of RNA except lanes 1 (50 ug); lane 7 (5 ug) and lane 12 (0.3 ug).

Figure 12.


RSV cDNA rep
RAV-0 cDNA rep
cDNA 3'
1
2
3
4
5
6
7
8
9
10
11
12
6.5-
3.0-
2.3-
4.3-
8.0-
6.5-
5.3-
3.0-
WC
C
N
N
C
C
N
N
C
C
N
WC

proviruses of ev loci, either $cDNA_{rep}$ for RAV-0 or $cDNA_3$, were used; neither of these cDNAs will detect the transcript from c-src. When possible I chose to illustrate the data obtained with $cDNA_3$, because they provide an immediate impression of the relative amounts of different virus-specific RNA species, and because the quality of these data were generally better than that obtained with $cDNA_{rep}$ for RAV-0. It was found that $cDNA_3$, could detect all but one of the RNAs produced from the ev loci in the study population (see below).

Virus-specific RNAs in embryos with the gs^-chf^- phenothype.

Samples 1 and 2 displayed the gs^-chf^- phenotype and harbored ev1 and ev5 (Table 2). No virus-specific RNAs (other than c-src RNA) were detected in the whole cellular RNA of sample 1 (Fig. 12, lane 1) or the nuclear and cytoplasmic fractions of sample 2 (data not shown). The concentration of virus-specific RNA in gs^-chf^- cells is very low ($<$ 1 copy per cell), (Fig. 13). I have observed 35S and 21S RNAs (identified by rate-zonal centrifugation) at extremely low concentrations in whole cellular RNA isolated from gs^-chf^- embryos (Fig. 14). In addition, it will be shown below that an 8.0 kb RNA (ca. 35S) can be attributed to transcription from ev1. Other workers have also described 35S and 21S RNAs that are transcribed in small amounts from ev1 (Hayward, et al., 1980; Wang, et al., 1977). I conclude that ev1 is transcribed at barely

Figure 13.

Concentration of viral RNA in embryos of the $gs^{-}chf^{-}$ and $gs^{+}chf^{+}$ phenotype. Whole cell RNA was extracted from embryos of the $gs^{+}chf^{+}$ or $gs^{-}chf^{-}$ phenotype. ^{3}H-labeled $cDNA_{gag/pol}$ or ^{3}H-labeled $cDNA_{env}$, (specific activity, 2×10^{7} cpm/µg, 1000 cpm/reaction) was incubated with increasing quantities of RNAs in a 40 ul volumn at 68°C for 96 hours ($gs^{+}chf^{+}$ cell RNA) or 120 hours ($gs^{-}chf^{-}$ cell RNA). Hybridization was measured as described in Materials and Methods. The arrows indicate the half-Crt values. The curves on the left represent the results obtained with cellular RNA from $gs^{+}chf^{+}$ embryos and the curves on the right represent the results obtained with cellular RNA from $gs^{-}chf^{-}$ embryos.

Figure 13.

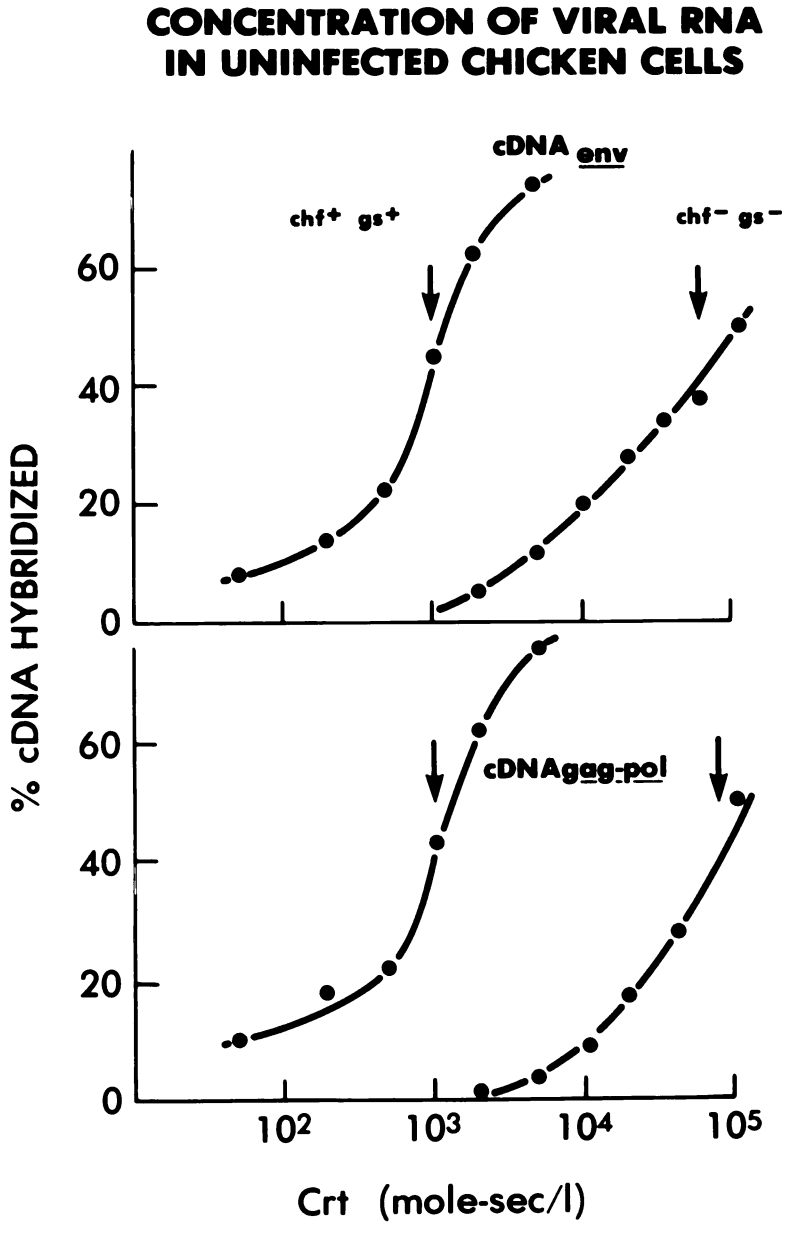

CONCENTRATION OF VIRAL RNA
IN UNINFECTED CHICKEN CELLS
cDNA env
chf+ gs+
chf− gs−
% cDNA HYBRIDIZED
60
40
20
0
cDNAgag-pol
10^2
10^3
10^4
10^5
Crt (mole-sec/l)

Figure 14. Rate-zonal sedimentation of RNA from gs^-chf^- chicken embryos.

One milligram of polyadenylated RNA, isolated from a gs^-chf^- chicken embryo, was sedimented through a 35 ml gradient of 15-30% sucrose in 0.01 M Tris-HCl (pH 7.4), 0.01 M NaCl, 0.01 M EDTA at 23,000 rpm for 41 hours. Fractions were collected, and one half of each fraction was hybridized to $^3H\text{-}cDNA_{env}$ (1000 cpm/fraction). The other half was hybridized to a mixture of $^{32}P\text{-}cDNA_{B77}$ (specific activity 10^8 cpm/ug) and $^3H\text{-}cDNA_{gag/pol}$ (1000 cpm of each cDNA/fraction). All samples were hybridized for 96 hours.

(A) Hybridization to $cDNA_{env}$ (o).

(B) Hyrbridization to $cDNA_{B77}$ (o), and $cDNA_{gag/pol}$ (o).

Figure 14.

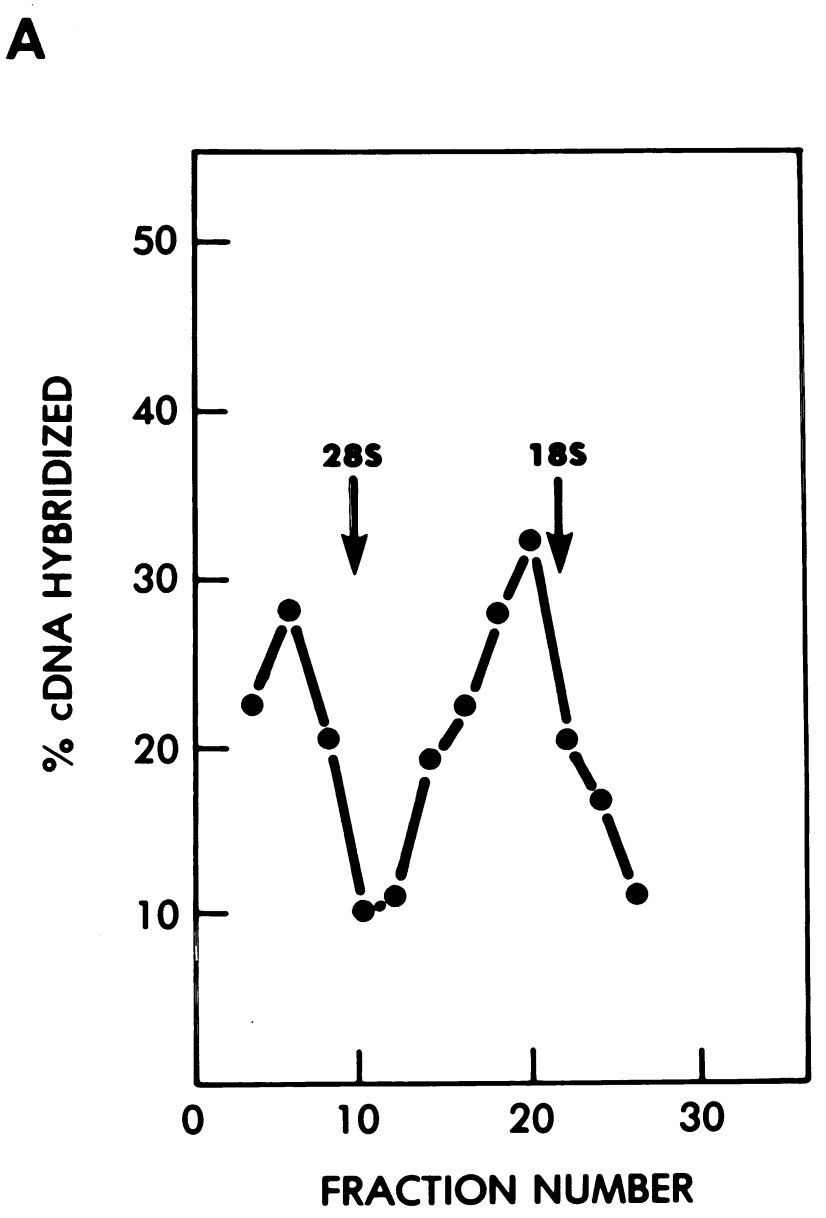
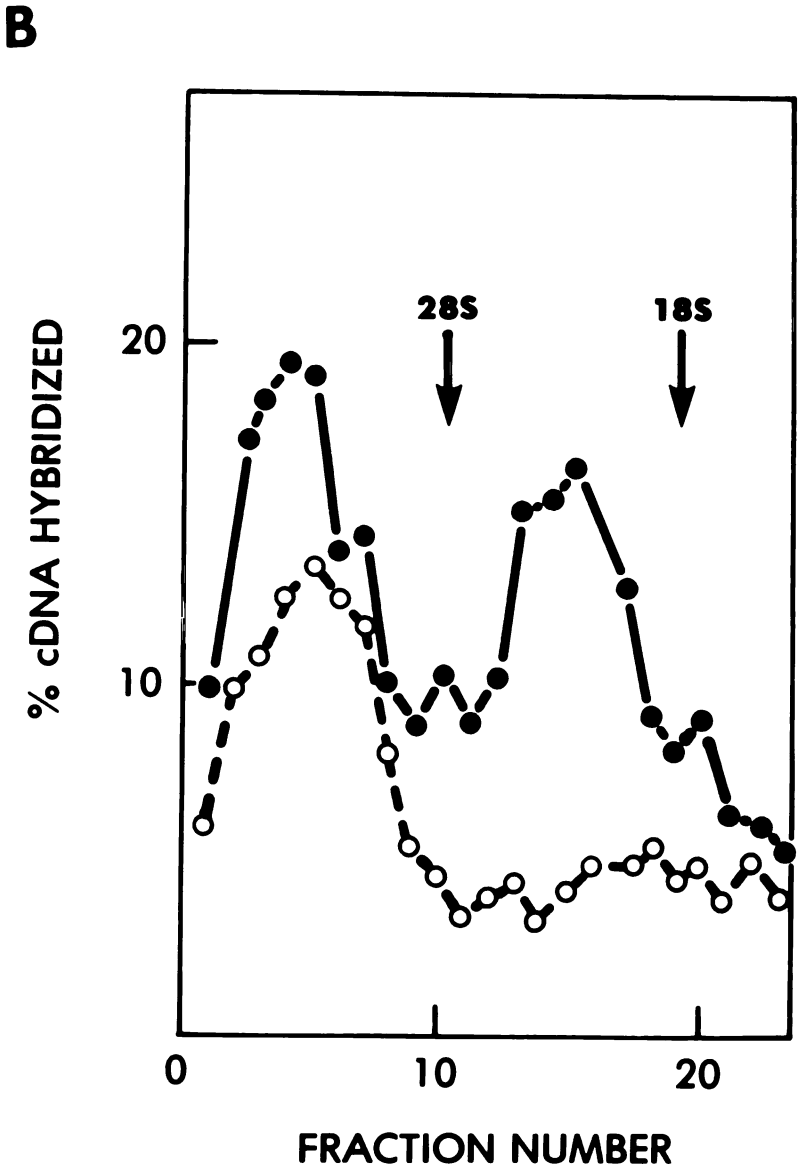

A
50
40
30
20
10
% cDNA HYBRIDIZED
28S
18S
0
10
20
30
FRACTION NUMBER

Figure 14. (Cont'd.)



detectable levels, but I have been unable to identify any RNAs that might arise from ev5.

Virus-specific RNAs in Embryos with the gs^+chf^+ Phenotype

The embryos of samples 3 and 4 had the gs^+chf^+ phenotype and contained ev1 and ev3 (Table 1). The embryos contained 50-100 copies of viral RNA per cell, as judged from measurements of the kinetics of hybridization of gs^+chf^+ whole cell RNA in solution (Fig. 13) (Hayward and Hanafusa, 1973; Bishop, et al., 1973). Two virus-specific RNAs (ca. 6.5 kb and 3.0 kb) were detected in samples 3 and 4 by the use of $cDNA_3$, (Fig. 12, lanes 6 and 7; data for sample 3 not shown), whereas three viral RNAs (6.5 kb, 3.0 kb, and 2.3 kb) were found with $cDNA_{rep}$ for RAV-0 (Fig. 12, lane 3; data for sample 4 not shown). The 6.5 kb and 3.0 kb RNAs were found at similar concentrations in the nuclear and cytoplasmic fractions (Fig. 12, lanes 2,3,6,7), whereas the 2.3 kb species was detected only in nuclear RNA (Fig. 12, lanes 2 and 3). These RNAs are most likely transcribed from ev3, but I cannot rigorously exclude the possibility that the presence of ev3 has induced transcription of the ev1 provirus.

Sample 5 also represented the gs^+chf^+ phenotype, but carried ev1, ev3 and ev15. The pattern of virus-specific RNA in this sample was indistinguishable from the patterns in samples 3 and 4 (see below, Fig. 17C). Thus, the presence of ev15 did not alter the pattern of transcription that we attribute to ev3, nor could we identify transcripts

unique to ev15. I again conclude that ev3 is transcribed into three virus-specific RNAs; either the transcription products of ev15 comigrate with ev3 RNAs, or their concentration is too low to permit detection.

The genotypes of samples 6 and 7 were not determined. Both samples had a gs^+chf^+ phenotype and were therefore assumed to contain ev3 (Astrin and Robinson, 1979). Because of the ubiquity of ev1, I assume that it was also present in the embryos of samples 6 and 7. Moreover, sample 6 was comprised of embryos from inbred line 6_3, which is known to carry only ev1 and ev3 (Astrin, et al., 1979a). The pattern of virus-specific RNA attributed to ev3 of samples 6 and 7 was indistinguishable from that of samples 3, 4 and 5 (see Fig. 12, lane 14 for sample 6; and Fig. 17A for sample 7). In addition to the ev3 RNAs an 8.0 kb RNA, attributed to ev1, was detected in sample 7 (Fig. 17A) and sample 4 (Fig. 17B).

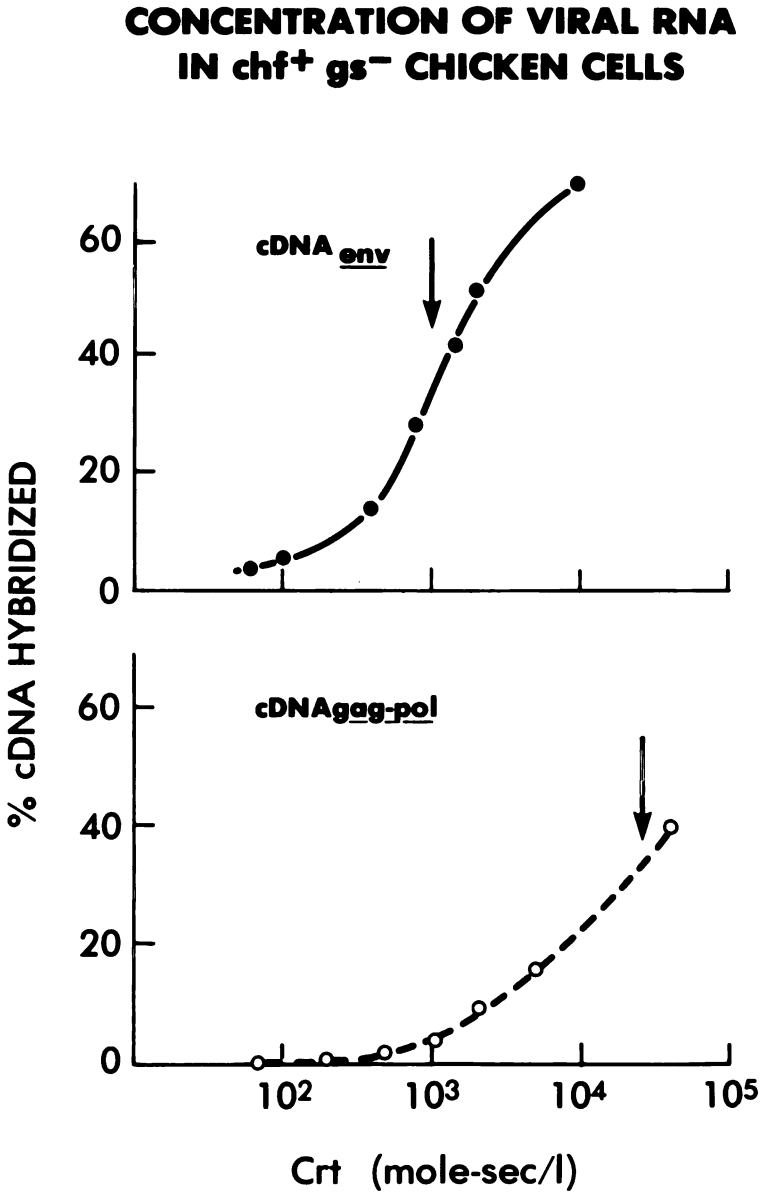
Virus-specific RNAs in embryos with the gs^-chf^+ phenotype.

I measured the amount of virus-specific cytoplasmic RNA in gs^-chf^+ embryos by the kinetics of hybridization with $cDNA_{env}$ and $cDNA_{gag\text{-}pol}$ (Fig. 15). The RNA was isolated from a pool of embryos (gs^-chf^+) of undetermined genotype. The results indicated that cytoplasmic virus-specific RNA contained env related sequences but was deficient in sequences represented by $cDNA_{gag\text{-}pol}$.

Figure 15.

Concentration of viral RNA in embryos of the gs^- chf^+ phenotype. Whole cell RNA was extracted from 3 embryos of the gs^- chf^+ phenotype. $^3H\text{-}cDNA_{env}$ or $^3H\text{-}cDNA_{gag/pol}$ was incubated with increasing quantities of RNA in a 40 ul volumn at 68°C for 96 hours. Hybridization was measured as described in Materials and Methods. Arrows indicate the half-Crt values.

Figure 15.


CONCENTRATION OF VIRAL RNA
IN chf+ gs− CHICKEN CELLS
% cDNA HYBRIDIZED
cDNA env
cDNAgag-pol
60
40
20
0
10²
10³
10⁴
10⁵
Crt (mole-sec/l)

Sample 8 consisted of a single embryo that had the $gs^{-}chf^{+}$ phenotype and carried ev1 and ev6 (Table 2). Two virus-specific RNAs, ca. 5.3 kb and 2.7 kb, were detected in sample 8 (Fig. 12, lane 8). The 5.3 kb RNA is difficult to visualize in the photographic reproduction (but see below, Fig. 18A and B). (The analysis illustrated in Fig. 12 failed to resolve the 3.0 and 2.7 kb RNA; but see below, Fig. 18C.) I measured the amounts of these RNAs by scanning autoradiograms with a densitometer and comparing the results to standards containing known amounts of viral RNA. The 5.3 kb RNA was determined to be present at 1-2 copies per cell, the 2.7 kb RNA at 50 copies per cell (data not illustrated). The 5.3 kb RNA appeared to be localized in the nuclear fraction (not evident in Fig. 12, see Fig. 18A). The 2.7 kb RNA was detected at similar concentrations in the nuclear and cytoplasmic fractions (Fig. 12, lanes 8 and 9). I attribute these two RNAs to the transcription of ev6, although I cannot exclude the unlikely possibility that transcription of ev1 has been induced by ev6.

Sample 9 consisted of five $gs^{-}chf^{+}$ embryos, each of which contained ev1, ev4, ev6 and ev8. Each embryo was analyzed separately for virus-specific RNA. The viral RNAs in all of the embryos were identical with regard to sizes, concentrations, and sub-cellular localizations (see below, Fig. 18A and B). The results were indistinguishable from

those obtained with sample 8. Ev4 and ev8 are either inefficiently transcribed (1 copy/cell), or their transcription products are obscured by comigration with the RNAs transcribed from ev6. Ev4 and ev8 do not appear to influence the transcription of ev6.

The $gs^{-}chf^{+}$ embryo of sample 10 carried ev1, ev9 and ev15 (Table 2). Sample 10 provided the only representation of ev9. I therefore extracted RNA from the cells of this embryo on several separate occasions. Two virus-specific RNAs, ca. 8.0 kb (20-40 copies per cell) and 3.0 kb (50-100 copies per cell) were detected in the cytoplasmic and nuclear fractions (Fig. 12, lanes 4,5,10 and 11). (The amounts of RNA were measured by autoradiography, as described above.) The 3.0 kb RNA was present at similar concentrations in the nucleus and the cytoplasm, whereas the 8.0 kb RNA was approximately 10-25 fold more concentrated in the nuclear fraction. The 8.0 kb and 3.0 kb RNAs were identified as transcription products of ev9. Other possible explanations for the origins of these RNAs are less likely: it is unlikely that they are due to the induction of transcription from ev1; ev15 consists only of sequences homologous to ends of viral RNA and hence could not encode the RNA described here; and ev15 does not appear to affect the transcription of other loci (see above, Fig. 17C). Hayward, et al. (1980) have described two viral RNAs, 35S and 21S, in cells that harbor ev1 and ev9.

I presume that these RNAs correspond to the 8.0 kb and 3.0 kb RNAs described here.

Virus-specific RNAs in cells producing RAV-0.

Sample 11 was an embryo derived from a line 7_2 x line 100 mating. The genotype of this embryo was not determined, but I did demonstrate that cells from this embryo produced virus particles presumed to be RAV-0 (Materials and Methods). It has been shown that line 7_2 and line 100 harbor ev1 and ev2, and that ev2 is responsible for production of RAV-0 (Astrin, et al., 1980a; Hughes, et al., 1979; Hughes, et al., 1981). Ev2 itself is inefficiently expressed (Vogt and Friis, 1971; Crittenden, et al., 1974; Smith, et al., 1974; Robinson, et al., 1976; Cooper and Temin, 1976; Hayward, et al., 1980). Sample 11 presumably carried ev2, but a genetic determinant in line 100 facilitated spread of RAV-0 by horizontal infection (Crittenden, et al., 1974; Robinson, et al., 1976; Crittenden, et al., 1979) and therefore effected the amplification of RNA transcribed originally from ev2. Therefore, it is presumed that sample 11 contained ev1, ev2, and proviruses that were integrated following transcription from the RNA of ev2. I assume that the newly acquired proviruses reflect the structure and composition of ev2.

Approximately 250-500 copies of viral RNA per cell were detected in cultured fibroblasts of sample 11 (data not shown). Hybridization with $cDNA_3$, detected two species of

viral RNA (8.0 kb and 3.0 kb) (Fig. 12, lane 12). Similar results were obtained by hybridization with $cDNA_{rep}$ for RAV-0 (data not shown). I found that these RNAs (35S and 21S) had genetic compositions and intra-cellular distributions similar to those of the viral RNAs in cells productively infected by ALV or in RAV-0 producing line 7_2 x 15 cells (Hayward, 1977; Weiss, et al., 1977) (Fig. 16). I conclude that the transcription products of newly acquired proviruses derived from ev2 were detected, although I cannot exclude the unlikely possibility that transcription of the ev1 provirus or other unidentified proviruses had been induced.

2. The genetic composition of virus-specific RNAs.

The genetic composition of virus-specific RNAs was characterized by hybridization with specific cDNAs described in Materials and Methods. The same RNA-bearing filters were used in two to five successive hybridizations with different cDNAs.

Figure 17A illustrates the results for the transcripts of ev3, using sample 7. RSV $cDNA_{rep}$ detected the 6.5 kb, 3.0 kb and 2.3 kb RNAs previously attributed to ev3. Two additional RNAs were found; an 8.0 kb species that reacted with $cDNA_{rep}$, $cDNA_5$, and $cDNA_{3'}$, and presumably arose from ev1 (see above); and a 4.3 kb RNA that reacted only with

Figure 16. Rate zonal sedimentation of cytoplasmic and nuclear RNA isolated from RAV-0 infected cells.

Ten micrograms of heat denatured cytoplasmic RNA (A), or nuclear RNA (B), from RAV-0 infected cells was sedimented through a 35 ml gradient of 15-30% sucrose. Fractions were collected and the RNA from each fraction was hybridized with $^{32}P\text{-}cDNA_{B77}$ (1000 cpm/fraction) at 68°C for 72 hours. Sedimentation coefficients were determined for viral RNAs (35S and 21S) based on the location of 28S and 18S ribosomal RNAs detected by recording optical density at 260 nm.

Figure 16.

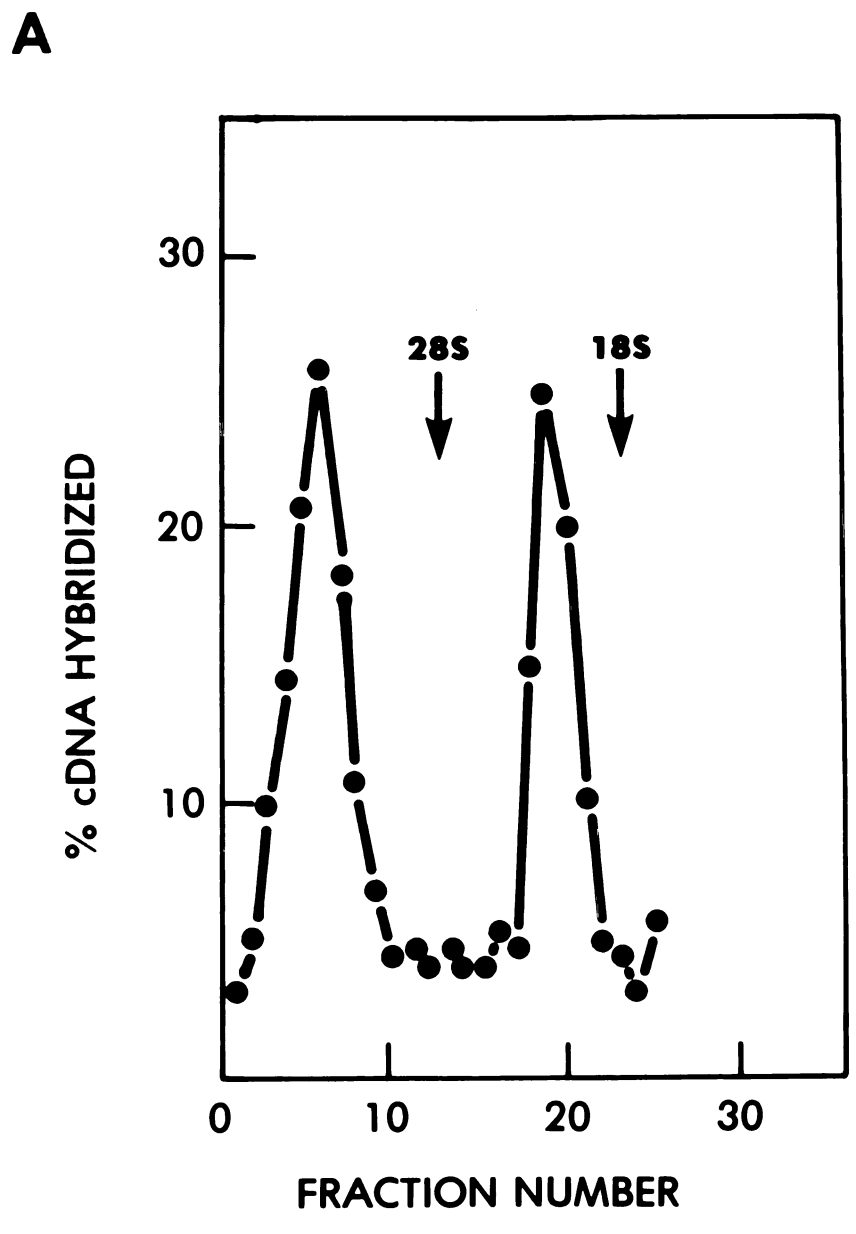


Figure 16. (Cont'd.)

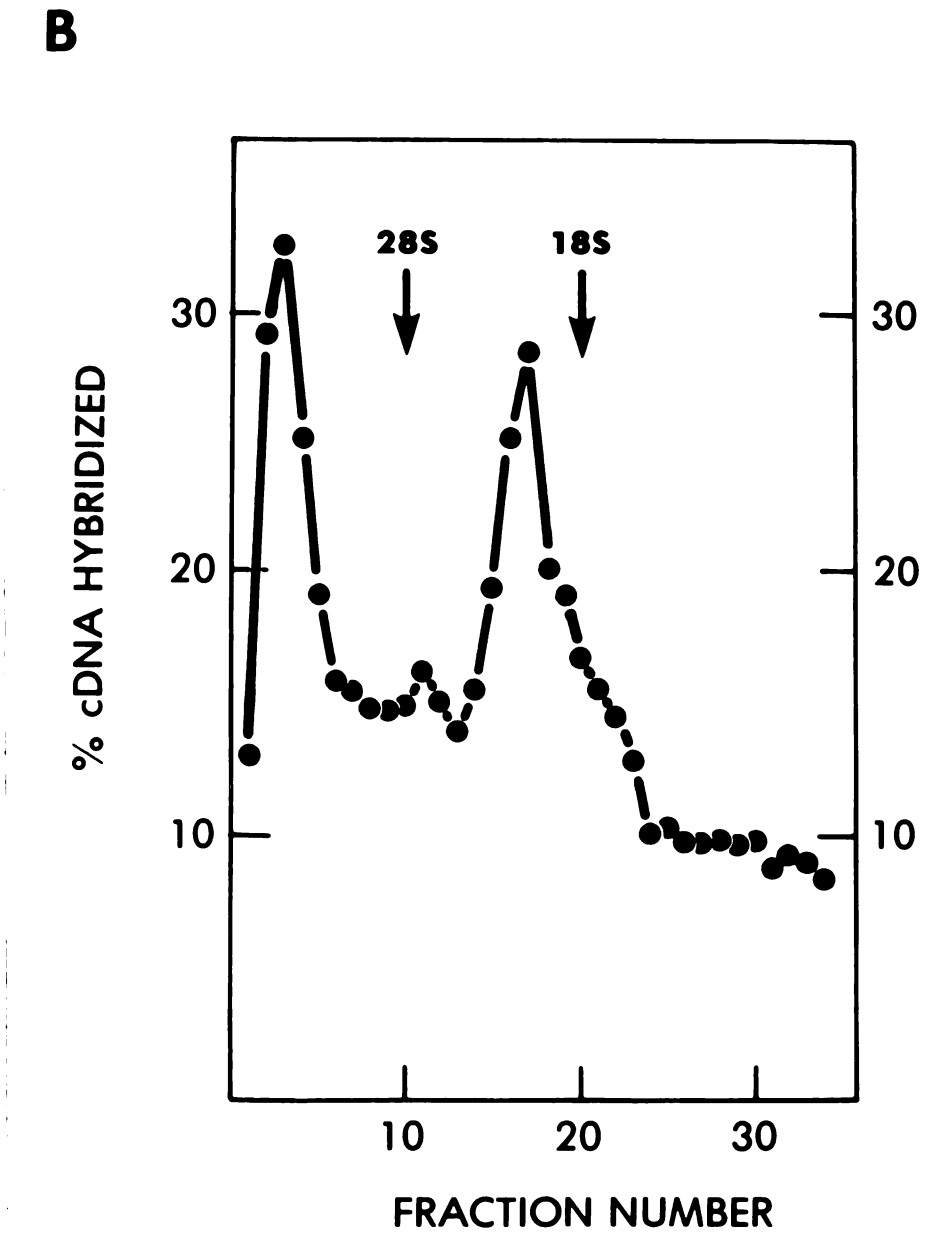
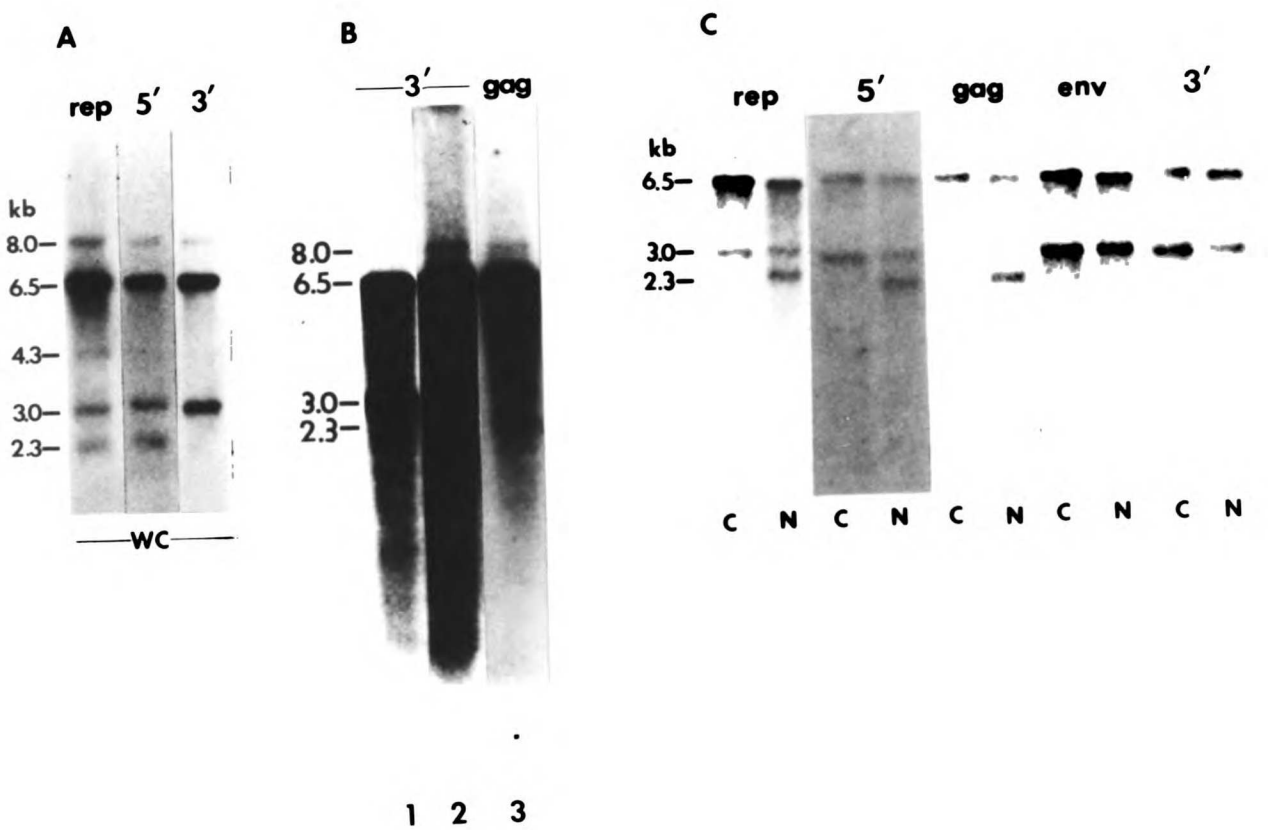


Figure 17. Genetic composition of virus-specific RNAs attributed to ev3.

Polyadenylated RNAs were prepared from whole cells derived from embryonic tissue (WC) or cytoplasmic (C) and nuclear (N) fractions of cultured embryo fibroblasts as described in Materials and Methods. RNA was analyzed as in Figure 12. The RNA-bound filter was sequentially annealed to the cDNAs listed above each lane or set of lanes. Panel A shows the results of hybridizations of 10 ug of RNA from sample 7 (ev1, ev3, Table 2). Panel B shows the results of hybridization of 17 ug of cytoplasmic RNA isolated from sample 4 (ev1, ev3). Lane 1, hybridization with $cDNA_3$,; lane 2 is a longer exposure of lane 1; lane 3, hybridization with $cDNA_{gag}$. Panel C shows the results of hybridization of 10 ug of cytoplasmic and nuclear RNA isolated from sample 5 (ev1, ev3, ev15, Table 2).

Figure 17.


A
rep
5′
3′
kb
8.0–
6.5–
4.3–
3.0–
2.3–
WC
B
3′
gag
8.0–
6.5–
3.0–
2.3–
1
2
3
C
rep
5′
gag
env
3′
kb
6.5–
3.0–
2.3–
C
N
C
N
C
N
C
N
C
N

$cDNA_{rep}$, a reaction with I ascribed to c-src (see above, Fig. 12, lane 1).

Figure 17B illustrates the results for the transcripts of ev1 and ev3 using sample 4, which carried only ev1 and ev3. $cDNA_3$, detected the 6.5 kb and the 3.0 kb RNAs attributed to ev3 (Fig. 17B, lane 1). Upon longer exposures of the filter, an 8.0 kb RNA attributed to ev1 was detected (Fig. 17B, lane 2). The 6.5 kb and 8.0 kb RNAs were also detected with $cDNA_{gag}$ (Fig. 17B, lane 3). A small amount of the 2.3 kb ev3 RNA was also detected.

The genetic composition of the ev3 RNAs was further analyzed with sample 5 (Fig. 17C). The 6.5 kb RNA reacted with all of the cDNAs that were used. I presume that this RNA is a complete transcript of ev3 and that the RNA serves as messenger for synthesis of the 120,000 dalton virus-specific protein (P120) detected exclusively in gs^+chf^+ cells (Eisenman, et al., 1978; Eisenman, et al., 1980). The 3.0 kb RNA was detected by all cDNAs except $cDNA_{gag}$. Its size and genetic composition is similar to the env mRNAs in chicken cells infected with avian leukosis virus (Hayward, 1977; Weiss, et al., 1977). I conclude that the 3.0 kb RNA is the messenger that is responsible for the chf^+ phenotype associated with ev3. The 2.3 kb RNA was detected with $cDNA_{rep}$ $cDNA_5$, and $cDNA_{gag}$, but not by $cDNA_{env}$ or $cDNA_3$, (Fig. 17C). The size and genetic composition of the 2.3 kb RNA suggest

that it is composed entirely of sequences from the leftward domain of the ev3 provirus. The nuclear localization of the 2.3 kb RNA suggests that it does not contribute to the ev3 gs^+ phenotype.

Figure 18A illustrates the composition of RNAs transcribed from ev6 in sample 9. The 5.3 kb and 2.7 kb RNAs reacted with $cDNA_{rep}$, $cDNA_{env}$, and $cDNA_3$, but not with $cDNA_5$, or $cDNA_{gag}$. The size of the 5.3 kb RNA, and its reactions with cDNAs as just described, suggest that the RNA is a complete transcript of ev6. The failure of the RNA to react with either $cDNA_5$, or $cDNA_{gag}$ is attributable to the large deletion in the provirus of ev6 (see Fig. 7). I did not test the reactivity of the 5.3 kb RNA with a cDNA for the pol gene, but it has been shown by others that this RNA does react with such a cDNA (Hayward, et al., 1980). The 2.7 kb RNA contains env gene sequences and untranslated sequences from the 3' viral domain. It is presumably an env mRNA, and its expression most likely accounts for the chf^+ phenotype associated with ev6. Figure 18B shows the reactivity of $cDNA_{env}$ with RNA isolated from two additional embryos of sample 9.

The env mRNA (2.7 kb) of ev6 appeared to be slightly smaller than the env mRNAs of ev3 and ev9 (3.0 kb) (see Fig. 18C). In order to verify this apparent difference, equal amounts of cytoplasmic RNA isolated from sample 4

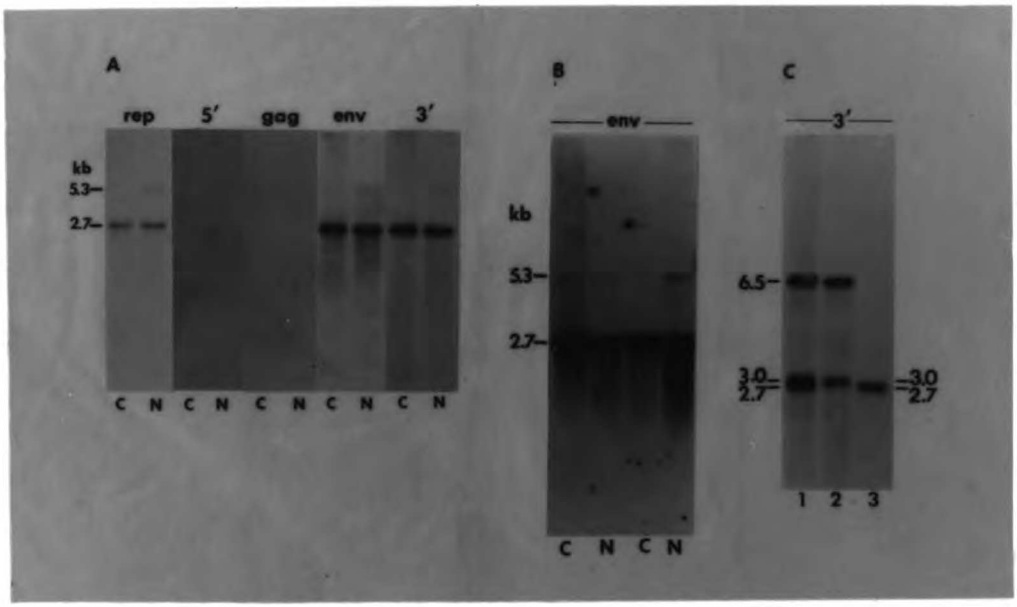
Figure 18. Genetic Composition of virus-specific RNA attributed to ev6.

(A) Ten micrograms of polyadenylated RNA isolated from cytoplasmic (C) and nuclear (N) fractions of sample 9 (ev1, ev4, ev6, ev8; Table 2) were analyzed by successive hybridizations with the cDNAs indicated above each set of lanes. Panel C shows hybridization of $cDNA_{3'}$. Lane 1, a mixture of 10 ug of cytoplasmic RNA 9f sample 8 (ev1, ev6 and 10 ug of sample 4 (ev1, ev3); lane 2, sample 3 alone; lane 3; sample 8 alone.

(B) Polyadenylated, cytoplasmic and nuclear RNA isolated from 2 other embryos of sample 9 (ev1, ev4, ev6, ev8), were analyzed with $cDNA_{env}$. Lane 1, 15 ug cytoplasmic RNA; lane 2, 7 ug of nuclear RNA; lane 3, 10 ug cytoplasmic RNA; lane 4, 10 ug nuclear RNA.

(C) Polyadenylated, cytoplasmic RNA isolated from sample 4 (ev1, ev3) and sample 8 (ev1, ev6). Lane 1, a mixture of 10 ug of RNA from sample 8 and 12 ug of sample 4; lane 2, sample 4 alone; lane 3, sample 8 alone.

Figure 18.


A
rep
5′
gag
env
3′
kb
5.3
2.7
C N C N C N C N C N
B
env
kb
5.3
2.7
C N C N
C
3′
6.5
3.0
2.7
3.0
2.7
1 2 3

(ev1, ev3) and sample 8 (ev1, ev6) were combined and electrophoresed in a lane (Fig. 18C, lane 1) adjacent to lanes carrying the two individual samples (Fig. 18C, lanes 2 and 3). The two env mRNAs were not completely resolved in the mixed sample, but the intensity and width of the 2.7-3.0 kb band was increased and it therefore appears that the two env mRNAs are indeed different in length. I attribute the smaller size of the ev6 env mRNA to the anomalous structure of the ev6 provirus (see Fig. 7 and Discussion).

The genetic composition of the ev9 RNAs is illustrated in Figure 19. The 8.0 kb RNA was detected by all cDNA probes and is probably a complete transcript of the ev9 provirus. This RNA is similar in size and genetic composition to the gag/pol mRNAs found in ALV-infected cells (Weiss, et al., 1977). Its apparent low cytoplasmic concentration may account for the gs^- phenotype associated with ev9. The 3.0 kb RNA reacted with all cDNAs except $cDNA_{gag}$. The size and genetic composition of this RNA are similar to env mRNAs detected in ALV-infected cells. I conclude that translation of the 3.0 kb RNA gives rise to the chf^+ phenotype associated with ev9.

Figure 20 summarizes the properties of viral DNA and RNA ov ev1, ev3, ev6, ev9 and proviruses derived from ev2.

Figure 19

Genetic Composition of Virus-Specific RNA Attributed to Ev9

Ten micrograms of polyadenylated cytoplasmic (c) and nuclear (N) RNA isolated from sample 10 (ev1, ev9, ev15; Table 1) were analyzed for virus-specific RNA as in Fig. 13. The filter, bound with RNA, was used in successive hybridizations with the cDNAs indicated above each set of lanes.

Figure 19.

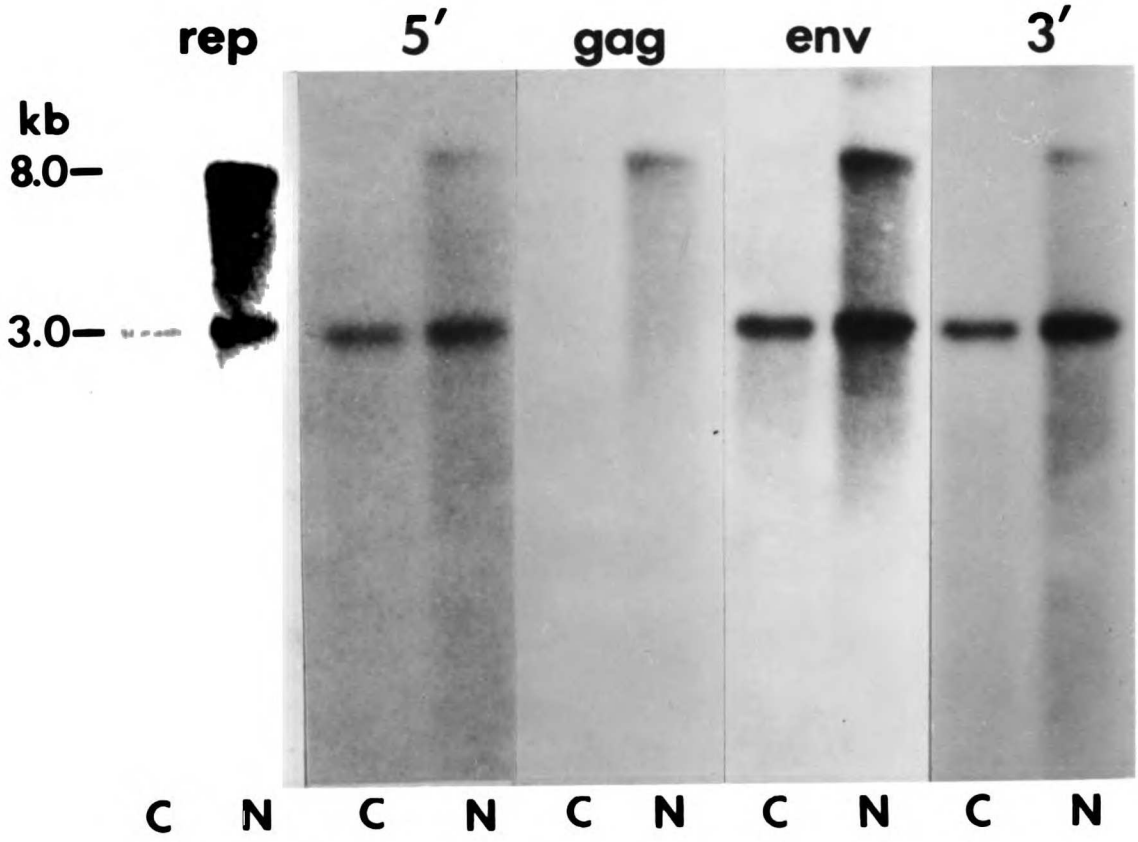
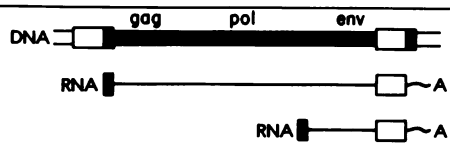
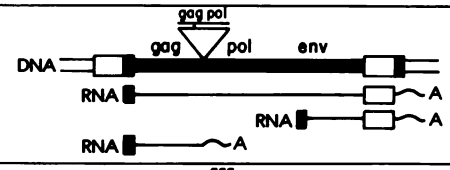
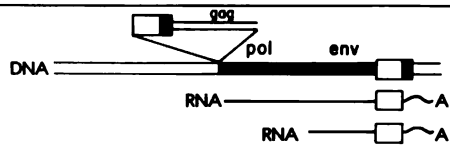
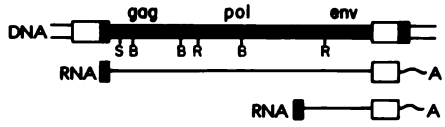


Figure 20. Summary of endogenous provirus DNAs and characteristics of virus-specific RNA.

Five endogenous proviruses are listed and identified by ev locus. The proviral DNA is represented by solid heavy lines. The left-ward and right-ward long terminal repeats are represented by adjacent open and solid boxes (not to scale). Open boxes denote sequences that are derived from the 3' terminus of viral genome (U_3), the closed boxes represent sequences derived from the 5' terminus of the viral genome (U_5). Cellular DNA is represented by an open double line extending to the left and to the right of the proviral DNA.

The structure and genetic content of each provirus and the location of major deletions are based on previous reports (Hughes, et al., 1979; Hughes, et al., 1981; Skalka, et al., 1980; Hayward, et al., 1980) and data presented here. The viral RNAs are represented by thin lines drawn below each provirus DNA. The solid and open boxes represent sequences described above. The genetic content, size and subcellular localization of virus-specific RNA is based on data presented in Figs. 12 to 19. All viral RNAs are polyadenylated (A). The copy numbers of viral RNA per cell were determined by measuring the kinetics of molecular hybridization in solution, or by scanning autoradiograms with a densitometer and comparing the result to standards containing known amounts of viral RNA.

Figure 20.

LOCUS	PHENOTYPE	GENETIC CONTENT OF VIRAL DNA AND RNA	SIZE OF VIRAL RNA (KB)	SUBCELLULAR LOCALIZATION OF VIRAL RNA		TOTAL COPIES OF VIRAL RNA/CELL
				Detected in Whole Cell and Cytoplasm		
				Cyto-plasm	Nucleus	
ev 1	gs− chf−		8.0 3.0			<1
Proviruses derived from ev 2	V+		8.0 3.0	+ +	+ +	250-500
ev 3	gs+chf+		6.5 3.0 2.3	+ + −	+ + +	50-100
ev 6	gs−chf+		5.3 2.7	− +	+ +	50
ev 9	gs−chf+		8.0 3.0	− +	+ +	50-100

DISCUSSION

Structure of Endogenous Proviral DNA

The endogenous proviral DNAs are located at different sites in host DNA and in most cases, their structural properties are similar to exogenous proviruses acquired by horizontal infection (Hughes, et al., 1979; Hayward, et al., 1980; Hughes, et al., 1981). Thus, their gene order and content, restriction maps and long terminal repeats (LTRs) resemble those of the unintegrated and integrated DNAs of RSV and RAV-0 (see Fig. 7; Shank, et al., 1978; Hsu, et al., 1978; Sabran, et al., 1979).

The structures of the ev1, ev2, ev8 and ev9, as determined by restriction enzyme analysis, indicated that these proviruses contained no major deletions (Hughes, et al., 1981; Hayward, et al., 1980). The structure of ev1 has been more thoroughly investigated by others (Skalka, et al., 1980; Hishinuma, et al., 1981). The ev1 Eco R1 fragments have been molecularly cloned and a detailed restriction map of the provirus at the ev1 locus has been described (Skalka, et al., 1980). The map of ev1 was found to be very similar to restriction maps for the RAV-0 and RSV virus strains. The nucleotide sequences at the junctions between host cellular and integrated ev1 proviral DNA were determined (Hishinuma, et al., 1981) and showed that, like proviral DNA of exogenous

origin, that the left and right long terminal repeat sequences (273 base pairs) contained sequences similar to proposed transcription control signals (Hishinuma, et al., 1981).

It was also found that a 6 bp repeat of host DNA immediately flanked each provirus terminus. The 6 bp sequence occurred only once in the DNA of a chicken that lacked ev1. It has been proposed that the host repeat sequences flanking integrated retrovirus DNAs are produced during integration by a mechanism similar to that of procaryotic and eucaryotic transposable genetic elements (Bukhari, et al., 1977; Shapiro, 1979; Calos and Miller, 1980).

At least 6 of the 12 proviruses that have been mapped lack sequences found in RAV-0 DNA (Fig. 7). The proviruses at ev4, ev5, and ev6 are deleted for the left long terminal repeat and part or all of the gag-pol regions of viral DNA. Ev3 appears to have retained the left terminal repeat but lacks sequences from the gag-pol gene boundary. The viral DNA at the ev15 and ev16 loci appear to consist only of sequences from the RAV-0-related terminal repeat (Fig. 7; Hughes, et al., 1981; Hayward, et al., 1980). It is not known if the deleted proviruses were flawed at the time of their insertion into the germline or if deletions have occurred after original insertion (Hughes, et al., 1981).

Possible Origins of Endogenous Avian Retroviruses

The demonstration that endogenous avian proviruses differ in number and location in the germlines of different individuals of white leghorn chickens suggest that these proviruses have been recently introduced into germlines of chickens, long after speciation. A similar conclusion has been reached by analysis of endogenous mouse mammary tumor virus DNA and endogenous AKV RNA (Cohen, et al., 1979; Steffen, et al., 1980) in various strains of mice.

Analysis of the structure of endogenous proviruses has revealed close similarities to proviruses acquired by experimental infection (e.g., gene order and content, presence of long terminal repeats, similarity of restriction endonuclease maps, and the structure of flanking cellular DNA) and further supports the hypothesis that they were introduced into germlines by infection.

The structural similarities between retrovirus proviruses and procaryotic and eucaryotic transposable elements has caused some investigators to speculate that proviruses may move within the genome of a host species by translocation of viral DNA sequences to other sites in the host genome by a mechanism that would not require reverse transcriptase. The available data indicate that such events would be rare (Varmus, 1981).

It has been speculated that some of the loci have arisen by such an event (Temin, 1980).

Transposable elements of bacteria can be precisely or imprecisely deleted generating flanking cellular DNA deletions. Deletions of endogenous avian retroviruses leftward LTR's are observed (e.g., for ev4, 5, 6). However, the mechanisms responsible for deletions of these proviruses are not known. Excision of most of a provirus could theoretically occur via homologous recombination or unequal crossing over between LTR's. Homologous recombination may explain the presence of endogenous elements that anneal only with $cDNA_3$, and $cDNA_5$, (e.g., ev15 and ev16).

Accounting for Patterns of Retrovirus Gene Expression in Uninfected Chick Embryos

Each endogenous retrovirus locus in chick embryos is reproducibly associated with a particular pattern of viral gene expression. In accord with previous reports (Astrin and Robinson, 1979; Astrin, et al., 1980b; Hughes, et al., 1981), my findings attribute the chf^-gs^- phenotype to ev1, ev4, ev5, ev8 and ev15; the gs^-chf^+ phenotype to ev6 and ev9; and the gs^+chf^+ phenotype to ev3. Ev2 apparently encodes the provirus of an infectious virus (RAV-0) that is produced in either very small quantities in line 7_2 (Astrin, et al., 1980a; Robinson, et al., 1976; Crittenden, et al., 1977) or larger amounts in line 7_2 x 100 (whose genetically determined susceptibility permits spread of virus (Crittenden, et al., 1974; Smith, et al., 1974). Moreover, my analysis of viral DNA and RNA in representative embryos indicates

that the patterns of viral gene expression are explicable by either structural features of the endogenous proviruses, or anomalous metabolism of viral RNAs (see Fig. 1).

i.) The gs^-chf^- phentotype of ev1, ev4, ev5, and ev8 probably arises from failure to produce sufficient amounts of viral RNA, although it remains formally possible that the RNA is synthesized and then rapidly degraded. The mechanisms that might account for this apparent deficiency in transcription are discussed below.

ii.) The gs^-chf^- phentotype of ev15 is due to the absence of viral structural genes from the provirus. The DNA of ev15 may consist of a single LTR, hypothesized to be the residue from excision of a more complete provirus by homologous recombination (Hughes, et al., 1981).

iii.) The gs^+chf^+ phenotype of ev3 reflects the production of viral mRNAs for both the gag and env genes. A deletion of ca. 1.5 kb overlaps the gag and pol genes of ev3, resulting in the synthesis of a 120,000 dalton protein that fuses the remainder of the gag and pol proteins and presumably accounts for the gs^+ phenotype (Eiseman, et al., 1978).

iv. A deletion in ev6 has removed the leftward LTR, all of the gag gene, and a portion of pol. As a consequence, the locus has a gs^- phenotype. The chf^+ phenotype of ev6 is attributable to the 2.7 kb env mRNA, which must be generated through the agency of cellular signals for

transcription and, perhaps, splicing (see below).

v.) The gs^-chf^+ phenotype of ev9 is apparently due to abnormal metabolism of viral RNA. The provirus has a normal size and genetic composition, and transcription produces nuclear RNAs with sizes suitable for messengers of gag (8.0 kb) and env (3.0 kb). Only the smaller RNA of these RNAs is abundant in the cytoplasm, however, and the cells consequently express env, but not gag.

vi.) Cells of line 7_2 x 100 produce substantial amounts of infectious virus (RAV-0) and contain 3.0 and 8.0 kb RNAs. These RNAs are presumably generated by transcription of proviruses that resulted from spread of the virus derived from ev2. I therefore presume that the viral RNAs in line 7_2 x 100 reflect the structure of the ev2 provirus.

vii.) I was unable to assess transcription from ev4 and ev8 in the absence of transcription from any other viral locus. Both ev4 and ev8 have been associated with the gs^-chf^- phenotype (Astrin, et al., 1980b). Accordingly, I found no evidence that either locus gives rise to appreciable amounts of stable RNA. My findings also indicate that the presence of ev4 and ev8 has no apparent effect on viral gene expression attributable to ev6. It remains formally possible that, in company with ev6, either ev4 or ev8 can produce RNAs that are indistinguishable from the transcripts of ev6.

Viral and Cellular Signals for RNA Synthesis and Metabolism

It appears likely that the promoter for retrovirus RNA synthesis resides in the U_3 domain of the leftward LTR (Taylor, 1979; Yamamoto, et al., 1980; Czernilofsky, et al., 1980b). This domain has been removed from the proviruses of ev4 and ev5 by large deletions (Hughes, et al., 1979; Hayward, et al., 1980). Accordingly, neither provirus is transcribed into detectable amounts of RNA. A similar deletion has affected the provirus of ev6, but in this instance, viral RNA is synthesized and apparently processed in amounts sufficient to permit detectable viral gene expression. I attribute the transcription of ev6 to a properly situated cellular promoter. Hayward and his colleagues have reached a similar conclusion (Hayward, et al., 1980).

The deletion in ev6 includes a splice donor site that normally participates in the genesis of sub-genomic retrovirus mRNAs (Weiss, et al., 1977; Mellon and Duesberg, 1977; Cordell, et al., 1978). I can envision several possible mechanisms by which this might occur: splicing from a donor site located in cellular DNA to the left of the ev6 provirus; splicing from a normally cryptic site within viral RNA; processing of the RNA without splicing; and initiation of transcription within the anomalous provirus. None of these possibilities can be excluded by available information. In any event, it is clear that the env mRNA of ev6 can be

translated in the absence ofthe normal leader sequence, which has been deleted from the provirus.

Regulation of Transcription from Endogenous Proviruses

The proviruses of ev1 and ev2 are ostensibly intact, yet they give rise to only small amounts of viral RNA. What factors limit the transcription of these loci? Cooper and colleagues have suggested that cellular DNA adjacent to ev2 exerts an inhibitory influence on transcription of the viral locus (Cooper and Temin, 1976; Cooper and Silverman, 1978). Thus, expression of ev2 is facilitated if a cis-regulatory element is eliminated: by physical removal of cellular DNA linked to the provirus, in the experiments of Cooper and colleagues; or by the insertion of proviruses representing ev2 at new sites, as when virus produced from ev2 spreads through a culture of cells from 7_2 x 100 embryos (see above, Jenkins and Cooper, 1980; Humphries, et al., 1979). In unpublished work, Conklin and Coffin have obtained evidence that the expression of ev1 is also constrained by reversible factors - in this instance, methylation of DNA in and/or adjacent to the provirus (personal communication). Even when freed of external regulatory influences, however, the endogenous avian retroviruses replicate less efficiently than do the homologous exogenous retroviruses. Available evidence suggests that this discrepancy may be due to the relative inefficiency of the

promoters found in the endogenous viral loci (Tsichlis and Coffin, 1980; Robinson, 1976; Lineal and Neiman, 1976).

Differential Metabolism of Endogenous Viral RNAs

Three of the endogenous viral loci that I have studied give rise to RNAs whose metabolism does not conform to the usual pattern for retroviruses.

i.) Two RNAs are produced from ev6: 5.3 kb, the length of the partially deleted provirus; and 2.7 kb, the subgenomic mRNA for env. The 2.7 kb RNA is more abundant in the cytoplasm than in the nucleus, in accord with the usual pattern (Deng, et al., 1977; unpublished results of authors); by contrast, the 5.3 kb RNA appears only in the nucleus and is consequently not represented by a protein product. I cannot explain the anomalous distribution of the 5.3 kb RNA: it may be entirely processed to the 2.7 kb form; it may resist transport from the nucleus; or it may be unstable upon reaching the cytoplasm. Each of these possibilities could be laid to the fact that important regulatory sequences may be affected by the deletion in ev6.

ii.) The provirus of ev9 has no gross defect and gives rise to viral RNAs of the expected sizes: 8.0 kb, representing the entire provirus, and the 3.0 env mRNA. In contrast to the usual pattern, however, the 8.0 kb RNA appears in only small amounts in the cytoplasm. Again, I have no explanation for this anomaly, but it is formally equivalent

to the defect associated with ev6 and it has the same unusual consequence: env is expressed, but gag is not.

iii.) Ev3 produces three RNAs, one of which (2.3 kb) is unusual in two regards: it fails to appear in the cytoplasm, despite its subgenomic size; and its polyadenylated 3' terminus lies in the midst of the ev3 locus rather than at the right-hand end, as is usually the case with subgenomic RNAs of retroviruses (Weiss, et al., 1977; Hayward, et al., 1977; Quintrell, et al., 1980). I cannot explain the atypical subcellular localization of this RNA, other than to suggest that the aberrant structure of the RNA might impede exodus from the nucleus even though the RNA is polyadenylated. I am similarly stymied by the unusual 3' terminus of the 2.3 kb RNA. I presume, but cannot prove, that the deletion in ev3 must be responsible, by causing either abnormal termination of transcription or abnormal processing of precursor RNA. It does appear that a detailed study of transcription from ev3 might unveil some of the elements that direct the metabolism of retrovirus RNA.

Do the Ev1 and Ev9 Loci Contain Structural Defects?

Both ev1 and ev9 give rise to RNAs large enough to constitute a full genome of endogenous viruses, but I have reason to believe that neither RNA is entirely normal.
i.) RNA is produced from ev2 in amounts no larger than those produced from ev1 (Hayward, et al., 1980); yet the

expression of ev2 generates at least small amounts of infectious virus, whereas the expression of ev1 produces no virus. I therefore presume that the structure of ev1 might be defective. Coffin has recently substantiated this presumption by showing that the pol gene of ev1 contains a defect of uncertain nature (personal communication). ii.) My deduction that ev9 is also defective is based on more circumstancial evidence: the failure of the ostensibly full-length transcript (8.0 kb) to appear in the cytoplasm (see above). More detailed studies of the structure of ev9 should provide a test of deduction and might also reveal some of the structural requirements for the intracellular transport of retrovirus RNA.

BIBLIOGRAPHY

Alwine, J.C., Kemp, J., and Stark, G.R. (1977). Proc. Nat. Acad. Sci. USA 74, 5350-5354.

Astrin, S. (1978). Proc. Nat. Acad. Sci. USA 75, 5941-5945.

Astrin, S.M., Crittenden, L.B., and Buss, E.G. (1979). Virology 99, 1-9.

Astrin, S.M., Crittenden, L.B., and Buss, E.G. (1980a). J. Virol. 33, 250-255.

Astrin, S.M., Robinson, H.L., Crittenden, L.B., Buss, E.G., Wyban, J., and Hayward, W.S. (1980b). Cold Spring Harbor Symp. Quant. Biol. 44, 1105-1109.

Astrin, S.M. and Robinson, H.L. (1979). J. Virol. 31, 420-425.

Bailey, J.M. and Davidson, N. (1976). Anal. Biochem. 70, 75-85.

Baltimore, D. (1970). Nature 226, 1209-1214.

Baltimore, D. (1974). Cold Spring Harbor Symp. Quant. Biol. 39, 1187-1200.

Baluda, M. (1972). Proc. Nat. Acad. Sci. USA 69, 576-580.

Bender, W. and Davidson, N. (1976). Cell 7, 595-607.

Benvelzen, P. (1974). Biochim. Biophys. Acta 335, 236-251.

Benveniste, R.E. and Todaro, G.J. (1974). Proc. Nat. Acad. Sci. USA 7, 4513-4517.

Benveniste, R.E. and Todaro, G.J. (1975). Proc. Nat. Acad. Sci. USA 72, 4090-4094.

Birnstiel, M.L., Sells, B.H., and Purdom, I.F. (1972). J. Molec. Biol. 63, 21-39.

Bishop, J.M. (1978). Ann. Rev. Biochem. 47, 35-88.

Bishop, J.M., Courtneidge, S.A., Levinson, A.D., Opperman, H., Quintrell, N., Sheiness, D.K., Weiss, S.R., and

Varmus, H.E. (1980). Cold Spring Harbor Symp. Quant. Biol. 44, 919-930.

Bishop, J.M., Jackson, N., Quintrell, N., and Varmus, H.E. (1973). In "Proceedings of the Fourth Lepetit Colloquim," (L. Silverstri, ed.), 41-50, North-Holland, Amsterdam.

Bolognesi, D.P. (1974). Adv. Virus Res. 19, 315-359.

Bolognesi, D.P., Luftig, R., and Shaper, J.H. (1973). Virology 56, 549-564.

Britten, R.J. and Smith, K. (1970). Carnegie Inst. Yearbook 68, 378-386.

Brugge, J. and Erikson, R.L. (1977). Nature 269, 346-347.

Bubbers, J.E. and Lilly, F. (1977). Nature 266, 458-459.

Bukhari, A.T., Shapiro, J.A., and Adhya, S.L. (eds.) (1977). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Maine.

Calos, M.P., Johnsrud, L., and Miller, J.H. (1978). Cell 13, 411-418.

Cameron, J.R., Loh, E.Y., and Davis, R.W. (1979). Cell 16, 739-751.

Coffin, J.M. and Haseltine, W.A. (1977). Proc. Nat. Acad. Sci. USA 74, 1908-1912.

Coffin, J.M., Champion, M., and Chabot, R. (1978). J. Virol. 28, 972-991.

Coffin, J.M. (1979). J. Gen. Virol. 42, 1-26.

Cold Spring Harbor Symp. Quant. Biol. 45 (1980).

Cohen, J.C. and Varmus, H.E. (1979). Nature 278, 418-423.

Collett, M.S., Erikson, E., and Erikson, R.L. (1979). J. Virol. 29, 770-781.

Cooper, G. and Temin, H.M. (1976). J. Virol. 17, 422-430.

Cooper, G.M. and Silverman, L. (1978). Cell 15, 573-577.

Copeland, T.D., Grandgenett, D.P., and Oroszlan, S. (1980). J. Virol., in press.

Cordell, B., Stavnezer, E., Friedrich, R., Bishop, J.M., and Goodman, H.M. (1976). J. Virol. 19, 548-558.

Cordell, B., Weiss, S.R., Varmus, H.E., and Bishop, J.M. (1978). Cell 15, 79-92.

Crittenden, L.B., Smith, E.J., Weiss, R.A., and Sarma, P.S. (1974). Virology 57, 128-138.

Crittenden, L.B., Motta, J.V., and Smith, E.J. (1977). Virology 76, 90-97.

Crittenden, L.B., Smith, E.J., Gulvas, F.A., and Robinson, H.L. (1979). Virology 95, 434-444.

Czernilofsky, A.P., Levinson, A.D., Varmus, H.E., Bishop, J.M., Tischer, E., and Goodman, H.M. (1980a). Nature 287, 198-203.

Czernilofsky, A.P., DeLorbe, W., Swanstrom, R., Varmus, H.E., Bishop, J.M., Tischer, E., and Goodman, H.M. (1980b). Nucl. Acid Res. 8, 2967-2984.

Dahlberg, J.E., Sawyer, R.C., Taylor, J.M., Faras, A.J., Levinson, W.E., Goodman, H.M., and Bishop, J.M. (1974). J. Virol. 13, 1126-1133.

DeLorbe, W.J., Luciw, P.A., Varmus, H.E., Bishop, J.M., and Goodman, H.M. (1980). J. Virol. 36, 50-61.

Deng, C.T., Stehelin, D., Bishop, J.M., and Varmus, H.E. (1977). Virology 76, 313-330.

Doughtery, R.M. and DiStefano, H.S. (1966). Virology 29, 586-595.

Eisenman, R., Shaikh, R., and Mason, W.S. (1978). Cell 14, 89-104.

Eisenman, R.N., Lineal, M., Groudine, M., Shaikh, R., Brown, S., and Neiman, P.E. (1980). Cold Spring Harbor Symp. Quant. Biol. 44, 1235-1247.

Elder, J.H., Jensen, F.C., Bryant, M.L., and Lerner, R.A. (1977). Nature 267, 23-28.

Fan, H. (1977). Cell 11, 297-305.

Faras, A.J., Taylor, J.M., McDonnell, J.P., Levinson, W.E., and Bishop, J.M. (1972). Biochemistry 11, 2234.

Fenner, F. (1976). Intervirology 7(1-2), 61-64.

Finnegan, D.J., Rubin, G.M., Young, M.W., and Hogness, D.S. (1978). Cold Spring Harbor Symp. Quant. Biol. 42, 1053-1063.

Friedrich, R. and Moelling, K. (1979). J. Virol. 31, 630.

Friedrich, R., Kung, H.J., Baker, B., Varmus, H.E., Goodman, H.M., and Bishop, J.M. (1977). Virology 79, 198-215.

Frisby, D.P., Weiss, R.A., Roussel, M., and Stehelin, D. (1979). Cell 17, 623-634.

Furuichi, Y., Shatkin, A.J., Stavnezer, E., and Bishop, J.M. (1975). Nature 257, 618-620.

Garapin, A.C., McDonnell, J.P., Levinson, W.E., Quintrell, N., Fanshier, L., and Bishop, J.M. (1970). J. Virol. 6, 589-597.

Gelb, L.D., Milstien, J.B., Martin, M.A., and Aaronson, S.A. (1973). Nat. New Biol. 244, 76-79.

Gonda, T.J., Sheiness, D., Fanshier, L., and Bishop, J.M. (1981). Cell, in press.

Grindley, N.D.F. and Sherratt, D.J. (1979). Cold Spring Harbor Symp. Quant. Biol. 43, 1257-1268.

Hackett, P. and Sauerbier, W. (1975). J. Molec. Biol. 91, 235-256.

Hanafusa, H., Miyamoto, T., and Hanafusa, T. (1970). Proc. Natl. Acad. Sci. USA 66, 314-321.

Hayward, W.S. (1977). J. Virol. 24, 47-52.

Hayward, W. and Hanafusa, H. (1973). J. Virol. 11, 157.

Hayward, W.S., Braverman, S.B., and Astrin, S. (1980). Cold Spring Harbor Symp. Quant. Biol. 44, 1111-1222.

Hishinuma, F., DeBona, P.J., Astrin, S., and Skalka, A.M. (1981). Cell, in press.

Hughes, S.H., Vogt, P.K., Stubblefield, E., Robinson, H., Bishop, J.M., and Varmus, H.E. (1980). Cold Spring Harbor Symp. Quant. Biol. 44, 1077-1088.

Hughes, S.H., Shank, P.R., Spector, D.H., Kung, H.J., Bishop, J.M., Varmus, H.E., Vogt, P.K., and Breitman, M.L. (1978). Cell 15, 1397-1410.

Hughes, S.H., Payvar, F., Spector, D., Payne, G., Schimke, R.T., Robinson, H.C., Bishop, J.M., and Varmus, H.E. (1979). Cell 18, 347-359.

Hughes, S.H., Bishop, J.M., and Varmus, H.E. (1981). Virology, in press.

Humphries, E.H., Glover, C., Weiss, R.A., and Arrand, J.R. (1979). Cell 18, 803-815.

Hunter, T. (1980). Cell 12, 647-649.

Hsu, T.W., Sabran, J.L., Mark, G.E., Guntaka, T.V., and Taylor, J.M. (1978). J. Virol. 28, 810-817.

Jenkins, N. and Cooper, G. (1980). J. Virol. 36, 684-691.

Kacian, D.L., Watson, K.F., Burny, A., and Spiegelman, S. (1971). Biochem. Biophys. Acta 246, 365-383.

Kung, H.J., Fung, Y.K., Majors, J.E., Bishop, J.M., and Varmus, H.E. (1980). J. Virol., in press.

Leong, J.A., Garapin, A.C., Jackson, N., Fanshier, L., Levinson, W., and Bishop, J.M. (1972). J. Virol. 9, 891-902.

Levin, J.G. and Rosenak, J. (1976). Proc. Nat. Acad. Sci. USA 73, 1154-1158.

Levinson, A.D., Oppermann, H., Levintow, L., Varmus, H.E., and Bishop, J.M. (1978). Cell 15, 561-572.

Levy, J. (1978). Curr. Top. Microbiol. Immunol., in press.

Linial, M. and Neiman, P.E. (1976). Virology 73, 508-520.

Linial, M., Madeiros, E., and Hayward, W.S. (1978). Cell 15, 1371-1382.

Lowy, D.R., Rowe, W.P., Teich, N., and Hartley, J.W. (1971). Science 174, 155-156.

Majors, J.E. and Varmus, H.E. (1980). Submitted to Nature.

Maxam, A.M. and Gilbert, W. (1977). Proc. Nat. Acad. Sci. USA 74, 560-564.

McClements, W., Hanafusa, H., Tilgham, S., and Skalka, A. (1979). Proc. Nat. Acad. Sci. USA 76, 2165-2169.

Mellon, P. and Duesberg, P.H. (1977). Nature 270, 631-634.

Moelling, K., Bolognesi, D.P., Bauer, H., Busen, W., Plassman, H.W., and Hausen, P. (1971). Nature New Biol. 234, 240-243.

Moore, D. (1976). In "Cancer: A Comprehensive Treatise," (F.F. Becker, ed.), 131-167.

Morris, V.M., Medeiros, E., Ringold, G.M., Bishop, J.M., and Varmus, H.E. (1977). J. Molec. Biol. 114, 73-92.

Neel, B., Astrin, S., and Hayward, W. (1980). Cell, in press.

Neiman, P.E. (1972). Science 178, 750-753.

Neiman, P.E., Das, S., and McMillin-Helsel, C. (1977). Cell 11, 321-329.

Opperman, H., Bishop, J.M., Varmus, H.E., and Levintow, L. (1977). Cell 12, 993-1006.

Pawson, T., Harvey, R., and Smith, A.E. (1977). Nature 268, 416-420.

Payne, G.S., Courtneidge, S.A., Crittenden, L.B., Fadly, A.M., Bishop, J.M., and Varmus, H.E. (1981). Cell, in press.

Payne, L.N. and Chubb, R. (1968). J. Gen. Virol. 3, 379-391.

Peters, G. and Dahlberg, J.F. (1979). J. Virol. 31, 395-407.

Quintrell, N., Hughes, S.H., Varmus, H.E., and Bishop, J.M. (1980). J. Molec. Biol. 143, 363-373.

Rettenmier, C.W., Karess, R.E., Anderson, S.M., and Hanafusa, H. (1979). J. Virol. 32, 102-113.

Rettenmier, C.W., Anderson, S.M., Reimen, M.W., and Hanafusa, H. (1979). J. Virol. 32, 749-761.

Robinson, H.L., Swanson, C.A., Hruska, J.F., and Crittenden, L.B. (1976). Virology 69, 63-74.

Robinson, H.L. (1978). Curr. Topics Microbiol. Immunol. 83, 1-36.

Rosenthal, P.N., Robinson, H.L., Robinson, W.S., Hanafusa, T., and Hanafusa, H. (1971). Proc. Nat. Acad. Sci. USA 68, 2336-2340.

Sabran, J.L., Hsu, T.W., Yeater, C., Kaji, A., Mason, W.S., and Taylor, J.M. (1979). J. Virol. 29, 170-185.

Schwartz, D.E., Zamecnik, P.C., and Weith, H.L. (1977). Proc. Nat. Acad. Sci. USA 74, 994-998.

Sen., A. and Todaro, G.J. (1977). Cell 10, 91-99.

Shapiro, J.A. (1979). Proc. Nat. Acad. Sci. USA 76, 1933-1937.

Shank, P.R. and Varmus, H.E. (1978). J. Virol. 25, 104-114.

Shank, P.R., Hughes, S.H., Kung, H.J., Majors, J.E., Quintrell, N., Guntaka, R.V., Bishop, J.M., and Varmus, H.E. (1978). Cell 15, 1383-1395.

Shank, P.R., Hughes, S.H., and Varmus, H.E. (1981). Virology 107, in press.

Shine, J., Czernilofsky, A.P., Friedrich, R., Goodman, H., and Bishop, J.M. (1977). Proc. Nat. Acad. Sci. USA 74, 1473-1477.

Skalka, A., DeBona, P., Hishinuma, F., and McClements, W. (1980). Cold Spring Harbor Symp. Quant. Biol. 44, 1097-1104.

Smith, B.J. and Bailey, J. (1979). Nucleic Acids Research 7, 2055-2072.

Smith, E.J., Crittenden, L.B., and Brinsfield, T.H., Jr. (1974). Virology 61, 594-596.

Southern, E.M. (1975). J. Molec. Biol. 98, 503-517.

Spector, D.H., Baker, B., Varmus, H.E., and Bishop, J.M. (1978). Cell 13, 381-386.

Stacey, B.W., Allfrey, V.G., and Hanafusa, H. (1977). Proc. Nat. Acad. Sci. USA 74, 1614-1618.

Steffen, D.L., Bird, S., and Weinberg, R.A. (1980). J. Virol. 35, 824-835.

Stehelin, D., Varmus, H.E., Bishop, J.M., and Vogt, P.K. (1976). Nature 260, 170-173.

Strand, M., August, J.T., and Jaenisch, R. (1977). Virology 76, 886-890.

Swanstrom, R., et al. (1981). J. Biol. Chem., in press.

Swanstrom, R., DeLorbe, W., Bishop, J.M., and Varmus, H.E. (1980). Proc. Nat. Acad. Sci. USA, in press.

Tal, J. Fujita, D.J., Kawal, S., Varmus, H.E., and Bishop, J.M. (1977). J. Virol. 21, 497-505.

Tal, J., Kung, H.J., Varmus, H.E., and Bishop, J.M. (1977b). Virology 79, 183-197.

Taylor, J.M. and Illmensee, R. (1975). J. Virol. 16, 553-558.

Taylor, J.M, Hsu, T.W., and Lai, M.M.L. (1978). J. Virol. 26, 479-484.

Taylor, J.M. (1979). Curr. Topics Microbiol. Immunol. 87, 23-41.

Temin, H.M. (1970). In The Biology of Large RNA Viruses (R.D. Barry and B.W. Mahy, eds.), 233-249.

Temin, H.M. and Baltimore, D. (1972). Advances in Virus Research (K. Smith, M. Lauffer and F. Bang, eds.), 17, 129-186.

Temin, H. (1974). Ann. Rev. Genetics 8, 155-178.

Temin, H. (1980). Cell 21, 599-600.

Tereba, A. and Astrin, S. (1980). J. Virol. 35, 888-894.

Todaro, G.J., Benveniste, R.E., Callahan, R., Leiber, M.M., and Sherr, C.J. (1975). Cold Spring Harbor Symp. Quant. Biol. 39, 1159-1168.

Todaro, G.J. (1975). Am. J. Pathol. 81, 590-605.

Tsichis, P. and Coffin, J. (1980). J. Virol. 33, 238-249.

Varmus, H.E., et al. (1972). Proc. Natl. Acad. Sci. USA 69, 20-24.

Varmus, H.E., Guntaka, R.V., Fan, W., Heasley, S., and

Bishop, J.M. (1974). Proc. Natl. Acad. Sci. USA 71, 874-878.

Varmus, H.E., Heasley, S. Kung, H.J., Opperman, H., Smith, V.C., Bishop, J.M., and Shank, P.R. (1978). J. Molec. Biol. 120, 55-82.

Varmus, H.E. (1981). In "RNA-Tumor Viruses (Part I)," Cold Spring Harbor (J. Tooze, ed.), Chapter 5, in press.

Verma, I.M., Varmus, H.E., and Hunter, E. (1976). Virology 74, 16-29.

Verma, I.M. (1977). Biochim. Biophys. Acta 473, 1-34.

Vogt, P.K. and Friis, R.R. (1971). Virology 43, 223-234.

Vogt, V.M., Eisenman, R., and Diggelmann, H. (1975). J. Molec. Biol. 96, 471-493.

Vogt, P.K. (1977). Comprehensive Virology 9, 341-456.

Wang, S., Hayward, W.S., and Hanafusa, H. (1977). J. Virol. 24, 64-73.

Weinberg, R.A. (1980). Cell, 22, 643-644.

Weiss, R.A. (1969). J. Gen. Virol. 5, 511-528.

Weiss, R.A., Friis, R.R., Katz, E, and Vott, P.K. (1971). Virology 46, 920-938.

Weiss, R.A., Mason, W.S., and Vogt, P.K. (1973). Virology 52, 535-552.

Weiss, R.A. (1975). Perspectives in Virology 9, 165-205.

Weiss, S.R., Varmus, H.E., and Bishop, J.M. (1977). Cell 12, 983-992.

Yamamoto, T., Jay, G., and Pastan, I. (1980). Proc. Nat. Acad. Sci. USA 77, 176-180.

Ziff, E.B. and Evans, R.M. (1978). Cell 15, 1463-1476.