

CUMULATIVE PROGRESS REPORT: 1985-PRESENT.

A. Overview.

As anticipated in my original application for an OIG, my laboratory's activities under the terms of this award have addressed several questions in two main arenas: retroviral replication and retroviral oncogenesis. In the sections that follow, I summarize our progress towards understanding the retroviral lifecycle (with emphasis upon viral entry, proviral integration, synthesis of reverse transcriptase by ribosomal frameshifting, and virus assembly) and some of the genes (especially *src*, *Wnt-1/int-1*, *myc*, and their relatives) implicated in retrovirus-induced neoplasia. (References to our published work on these topics are indicated by numbers in the text and listed at the end of the application.)

At the time of my original application, the remainder of my research program was in the field of hepadnaviruses and funded by separate grants. During the past few years, I have decreased and ultimately ceased our work on hepadnaviruses and have consequently not renewed the grants that supported it. Concomitantly, our group has devoted increased efforts to aspects of retroviral replication related to AIDS, and I have obtained NIH support (listed elsewhere) for those extensions of my program.

B. Highlights.

1. Retrovirus replication.

*** Discovery of the nucleoprotein complex that mediates retroviral integration (1) and the development of the first in vitro system that recapitulates the integration reaction (2).** These advances have allowed a precise definition of substrates and intermediates in the reaction (3) and stimulated the development of in vitro assays (e.g., for att site binding [2]) that use purified components. More recently, we have been able to simulate the in vivo reaction more closely in vitro by using minichromosomes as targets, and have developed a PCR-band assay that allows us to survey thousands of integration events in a single gel. (Fig. 1)

*** Discovery of the phenomenon of ribosomal frameshifting in the -1 direction to account for the synthesis of most retroviral pol gene products (4).** Our subsequent studies of this phenomenon in the context of RSV (5), MMTV (6), and HIV (7) produced the first insights into the viral signals that influence frameshifting (a "slippery" heptanucleotide sequence and RNA secondary structure downstream) and attracted attention to a translational mechanism for generating diversity that has now been encountered in *E.coli*, yeast, and coronaviruses.

*** Cloning of the chicken and quail genes encoding the putative receptor for RSV, subgroup A.** Although the sequencing of genes encoding this long-sought receptor is not yet complete, it will undoubtedly offer a new perspective on viral-host interactions, the mechanism of early events in the virus life cycle, and the nature of the presumptively polymorphic avian proteins that recognize the several subgroups of RSV.

*** Demonstration that an ectopically expressed transmembrane protein (human CD4) can be efficiently incorporated into RSV envelopes during virus assembly (8).** This finding casts a new light on the assembly process and provides new opportunities for targeting retrovirus vectors to host cells and for understanding the determinants of virus entry.

2. Retroviral oncogenesis.

*** Discovery that the mouse *int-1* (now *Wnt-1*) gene is expressed in a highly restricted manner (only in the embryonic nervous system and during spermatogenesis), implying roles in mammalian development (9,10).** These findings, in concert with Nusse's discovery that *Wnt-1* is the mouse homolog of the *Drosophila* segment polarity gene, *wingless*, inspired the gene targeting experiments of others, which demonstrated that *Wnt-1* is essential for formation of the cerebellum and midbrain.

*** Biochemical and genetic characterization of *Wnt-1* protein as a secretory glycoprotein (11,12) that can function in autocrine or paracrine modes, presumably by interaction with membrane receptor(s).** This work has placed the *Wnt-1* proto-oncogene and other members of the *Wnt* family among genes that affect differentiation and cell growth by cell-cell signalling and has focused attention upon the *Wnt* receptor(s).

*** Development of cell culture assays for biological effects of *Wnt-1* (13) and production of transgenic mice as a model for multistep mammary tumorigenesis involving the *Wnt-1* gene (14).** These advances secured the definition of *Wnt-1* as a proto-oncogene, displayed its unusual tissue-restricted potential for cell transformation, showed the dramatic effects of *Wnt-1* on mammary cell proliferation in male and female animals, and permitted demonstration of collaborative oncogenic effects with *int-2* and other proto-oncogenes.

*** Characterization of *src* mutants with host-dependent and dominant-negative phenotypes that define important functional domains outside the protein-tyrosine kinase domain of *src* proteins (15, 16, 17, 18).** In conjunction with work in several other labs, these mutants brought the Src-homology (SH) domains 2 and 3 to widespread attention, providing new opportunities for identifying the proteins that are crucial modifiers and targets for Src-like kinases.

*** Implication of *c-src* protein in the cell cycle through the demonstration that it is a target for phosphorylation by the mitotic kinase, p34, encoded by the mammalian homolog of *cdc2* (19).** This work, with concurrent studies from Shalloway's lab, identified p60^{C-*src*} as one of the first specific substrates for the mitotic kinase and suggested mechanisms to induce those properties of mitotic cells that mimic cell transformation.

*** Generation of a large collection of *c-myc* mutants that allowed definition of functional domains of *Myc* proteins (20).** These insertion and deletion mutants have been a valuable resource for the scientific community over the past five years, expediting the rapid advances in our understanding of *Myc* proteins as probable transcription factors.

C. Narrative.

1. Retrovirus replication.

a. Retroviral integration.

Retroviral integration remains the best understood example of recombination in higher eukaryotes, and the components that account for the efficiency and precision of the reaction are of great interest. About five years ago, we found that newly synthesized viral DNA in the cytoplasm and nuclei of MLV-infected cells was part of a high molecular weight nucleoprotein complex that sedimented at 160S in sucrose gradients (1). Predicting that the nucleoprotein complex might function as an integration machine, we found (in collaboration with Pat Brown in Mike Bishop's laboratory) that it was able to mediate correct integration of MLV DNA into a target of naked

lambda phage DNA in vitro (2). More detailed studies revealed that the complex contained viral capsid proteins, in addition to reverse transcriptase and integrase; that it could be considerably purified without loss of integration activity; that the linear viral DNA lacked two nucleotides from the 3' ends of each strand as a result of integrase activity in the cytoplasm; and, most importantly, that linear DNA (rather than circular DNA, as previously thought) was the immediate precursor to the provirus, with the recessed 3' ends of viral DNA joined to newly created, staggered 5' ends in the target DNA (1, 3).

Although the in vitro reaction was monitored initially with a genetic assay (in which a *supF* gene in the viral LTR conferred replication-competence on a target phage DNA with amber mutations), the products can also be followed with physical assays, and the target DNA molecules can be varied (linear or circular, large or small). In the past year, we have shown that chromatin can also serve as a target; it is used at least as efficiently as naked DNA and allows us to evaluate the presence and position of nucleosomes (or other chromatin associated proteins) for influences upon choice of integration sites. Using either cloning and sequencing or a novel PCR-based method (see Fig. 1) for assay of integration sites, we have studied integration of MLV DNA into a small yeast minichromosome with two mapped nucleosome-free domains. Unexpectedly, insertion occurs somewhat more frequently in regions with nucleosomes, and integration sites in nucleosomes occur with a periodicity of approximately ten bps, suggesting that one face of the DNA helix is favored for integration. (A similar periodicity is not observed with the same plasmid in naked form.) Similar studies of MLV integration in SV40 minichromosomes in vitro and in co-infected monkey cells also indicate that nucleosome-free regions are not favored. The use of SV40 DNA as a target in infected cells also permits controlled perturbation of the target, using temperature sensitive SV40 mutants; preliminary results suggest that the DNA replication activity of the minichromosomes does not affect integration frequency.

We have attempted to exploit the in vitro integration assay to understand the mechanism of restriction of MLV infection mediated by alleles at the mouse *Fv-1* locus. In our hands, a marked reduction in the level of full length viral DNA in the cytoplasm appears to be the major impairment during infection of restrictive hosts; nucleoprotein complexes from restrictive cells can mediate integration in vitro, and extracts from restrictive cells do not affect integration in trans. We have concluded that the Fv-1 protein affects a step in synthesis, transport, or catalysis of viral DNA, before the integration event, but we have been unable to identify an alteration in the viral nucleoprotein complex in restrictive cells.

To examine the integration step more precisely, we have used yeast expression vectors to produce large amounts of MLV and HIV integration proteins, and these have been purified to virtual homogeneity for structural studies (under the terms of another award) and for biochemical tests. Both proteins have been shown to bind to their cognate att sites in vitro, using gel retardation assays (21, 22). The binding sites at each end of MLV DNA are equivalent and relatively short (13 bp are sufficient); binding is sensitive to mutations that affect integration in vivo. The two HIV binding sites are not equivalent, a larger region is required for optimal binding, and mutations that affect binding have yet to be tested in vivo (i.e. the att sites have not been defined genetically). Recently, we have used our purified HIV integration protein for 3' end processing and DNA recombination reactions with putative att site oligonucleotides, as originally described by Craigie and colleagues. This has allowed us to characterize sequence determinants of the three defined steps (binding, 3'end processing, and recombination) and to demonstrate the importance of the conserved CA dinucleotide near the end of the viral att site in both the processing and recombination reactions.

b. Ribosomal frameshifting.

At the time of my original application for an OIG, we had just begun to take seriously the possibility that ribosomal frameshifting could explain the synthesis of *gag-pol* fusion proteins from

overlapping reading frames. By the time the award was implemented, we had shown that -1 frameshifting was indeed responsible for synthesis of the RSV *gag-pol* precursor protein, using a now standard assay in which *gag-pol* transcripts, synthesized in vitro by a bacteriophage RNA polymerase from cloned retroviral DNA, are translated in a rabbit reticulocyte lysate (4). We then used this assay to study HIV (7) and MMTV (6), as well as RSV (5), confirming the common use of -1 frameshifting to make reverse transcriptase; in the case of MMTV, two frameshifts occur, one to generate a *gag-pro* fusion protein and a second to generate *gag-pro-pol* protein. The shifts of frame occur efficiently (5 to 25% of ribosomes undergo a shift) at sites with the general sequence ..X XXY YYZ (X may equal Y). The importance of these sites was documented by the sequence of transframe proteins and by site-directed mutagenesis, and prompted the currently accepted "simultaneous slippage" model for -1 frameshifting (5) in many experimental contexts.

In addition, we found that ribosomal frameshifting may be augmented by secondary structural features of the RNA downstream from the frameshift site. This was first documented by site-directed and deletion mutagenesis of a predicted stem-loop in RSV RNA (5) and by finding that the frameshift sites in MMTV RNA were not sufficient for efficient frameshifting (6). Subsequently, we have shown that the MMTV *gag-pro* frameshift site is followed by an RNA pseudoknot, as earlier found by others for a coronavirus frameshift. On the other hand, a projected stem-loop in HIV RNA appears to have only weak effects on frameshift efficiency in vitro (23), although our more recent findings indicate that the HIV stem loop has a dramatic effect upon frameshifting in cells.

c. Retroviral receptors

About three years ago we initiated a project to clone chicken genes encoding the several receptors for RSV-related retroviruses. Since the receptors had been defined genetically and not biochemically, we have taken a genetic approach: mammalian cells (normally uninfected by most strains of RSV) are used as recipients of DNA from receptor-producing chick embryos, and recipient cells that take up DNA (identified by co-transfected markers) are tested for susceptibility to RSV vectors of an appropriate subgroup. In this way, we have obtained primary transformants in COS cells and secondary transformants in mouse 3T3 cells that are now permissive for infection by subgroup A strains of RSV. We have cloned DNA from the secondary transformants linked to markers present in the primary transfection, and have obtained a clone that appears by several criteria to be part of the gene for the subgroup A receptor. Using parts of this clone as probes, we have subsequently cloned homologous portions of quail and chicken DNA and have shown that a quail DNA clone can efficiently confer sensitivity to RSV-A upon mouse cells (Fig. 2). Sequencing of the avian clones reveals probable exons that are being used to isolate cDNA clones.

d. Virus assembly.

We have tested several chimeric envelope proteins and heterologous membrane-associated proteins for their ability to be incorporated into RSV particles. Initially, RSV *env*-encoded proteins with insertions of EGF were made: the only chimeric protein that was incorporated into RSV did not promote infection of mammalian cells expressing EGF receptors, perhaps because the EGF moiety was likely to be poorly exposed on the surface of the particle. In efforts to make chimeras in which the ectodomain was entirely encoded by a cellular gene, we also tested normal human CD4 and found that it (unlike CD4-RSV-*env* chimeras) was very efficiently assembled into RSV particles in quail fibroblasts chronically infected with RSV (8). Preliminary results indicate that several other foreign proteins (including human CD8, the mouse poly IgA receptor, and influenza hemagglutinin) can also be incorporated.

2. Retroviral oncogenesis.

a. The *src* gene and its relatives.

In a long-standing collaboration with Mike Bishop and his colleagues, we have identified and cloned several genes belonging to the *src* gene family, including *c-src* itself, *fgr*, and *hck* (24, 25, 26). Our studies of the viral and cellular *src* genes and their close relatives are now directed towards understanding the normal and neoplastic functions of the cytoplasmic protein-tyrosine kinases they encode, with emphasis upon genetic strategies for identifying regulatory properties and protein-protein interactions.

Mutational studies of *src* have refined the definition of aminoterminal sequences required for myristylation of pp60 (27) and identified multiple domains in the aminoterminal half of the protein that localize it variously to the plasma membrane, cytosolic vacuoles, and perinuclear membranes (28). Nucleotide sequencing of a fortuitously isolated host-range allele of *v-src*, one that permits transformation of chicken but not rodent cells (29), demonstrated the importance of a highly conserved amino acid sequence, FLVRES, upstream of the kinase domain of pp60, in a region likely to be involved in significant interactions between pp60 and cellular proteins (15). This sequence appears to be an important component of the portion of pp60 now known as *src* homology-2 (SH2), a region conserved among the protein-tyrosine kinases lacking transmembrane domains and related to regulatory domains of several other signal transduction proteins.

Using site-directed mutagenesis, we have produced a large collection of substitution and deletion mutants in the SH2 and the adjacent SH3 regions of *c-src* protein and characterized their biological and biochemical properties after expression in chicken and mouse cells (16, 17, 18). Several mutations in SH2 activate the oncogenic potential of pp60, some in SH2 or SH3 inhibit the transforming activity of pp60 activated by mutation of Tyr-527, and several mutations produce host-dependence for transformation. Two mutants (known as M6 and M9) with lesions in the FLVRES sequence are hyperactive in transformation of chicken cells, deficient in transformation of mouse cells, and capable of interfering with transformation of mouse cells by activated *c-src* (18). The M6 and M9 proteins also display differences in stability, protein-tyrosine kinase activity, and substrate preference when expressed in the two cell types, implying differential recognition of important regulators of pp60.

As an alternative means to identify interactions between pp60 and cellular proteins, we have sought host cell mutants resistant to transformation by *v-src*. A rat cell line expressing three copies of *v-src* was constructed, mutagenized, and screened for reversion to a non-transformed phenotype. One line was isolated and met the appropriate criteria (loss of transformed properties, continued expression of wild type *v-src*, and resistance to retransformation by additional *v-src* genes); somatic cell fusion experiments indicated that reversion is a dominant phenotype (M. Schofield, Ph.D. thesis).

To expedite biochemical and structural studies of *src* proteins, we have produced the products of chicken *c-src* and nearly a dozen mutant versions of *c-src* in insect cells, using baculovirus vectors (19, 30). With immunoaffinity chromatography, we obtain about 1 mg of protein (over 99% pure and enzymatically active) from a one liter culture. These proteins were first successfully used to look for an activity in mitotic cells responsible for the phosphorylation of threonine residues in pp60 reported by Shalloway and his colleagues; we found the mitotic kinase, the product of the mammalian homolog of *cdc2*, to be the relevant enzyme, thereby implicating *c-src* in the control of mitotic events (19).

During a sabbatical at the Whitehead Institute in 1988-89, I initiated attempts to make mutations of *c-src* in cultured cell lines by homologous recombination. Although this effort was made partially obsolete by Soriano's success in generating a *c-src* deficient mouse line (see below), it led

to the still unexplained observation that embryonic stem (ES) cells and embryonic carcinoma (EC) cells differ with respect to homologous recombination rates at different loci. We have found that targeting vectors for *c-src* and *HPRT* mediate efficient recombination at the endogenous loci in ES cells, but not in the EC line known as P19; however, an *N-myc* targeting vector undergoes homologous recombination with the *N-myc* locus at similar frequencies in the two lines.

b. The *Wnt-1* gene.

The gene originally known as *int-1*, recently renamed *Wnt-1* as the founding member of a large and highly conserved gene family (31), was discovered here about ten years ago as a common target for insertional activation in MMTV-induced mammary carcinomas (32). By cloning full length cDNA from mammary tumors, we predicted *Wnt-1* protein to be secreted, cysteine-rich, and modified by aminoterminal cleavage of a signal peptide and by four N-linked glycosylations (33). Direct analysis of *Wnt-1* proteins synthesized *in vitro* and in cells programmed to express the gene confirmed these findings, while also revealing processing and secretion to proceed inefficiently (11, 12). We have also recently found that the protein is associated with the chaperonin BiP (or HSP78), presumably in the endoplasmic reticulum.

The protein has proven difficult to express or purify from expression systems in bacteria, yeast, and insect cells, and we (and others) have failed to demonstrate biological activity in cell-free preparations of *Wnt-1* protein. Instead we have developed several biological assays for the *Wnt-1* gene. When introduced into a mouse mammary epithelial cell line (C57MG), the gene induces a dramatic morphological change and releases cells from growth restraints (13); in the rat pheochromocytoma line, PC12, the gene causes the cells to flatten and display greater adherence (Fig. 3). In addition to these direct, autocrine assays, we have developed a paracrine assay, taking advantage of the observation that many cell types express but fail to respond to the gene: mouse 3T3 cells producing *Wnt-1* protein induce characteristic changes in surrounding C57MG cells that are not expressing *Wnt-1*.

These assays have been used to examine site-directed mutants of *Wnt-1*. Glycosylation sites can be eliminated without loss of autocrine activity, although several mutants are deficient in paracrine activity. Changes that add or remove cysteine residues often inactivate the gene, findings consistent with the remarkable conservation of 21 cysteine residues throughout the *Wnt* gene family. (One of the glycosylation mutants and one of the cysteine mutants appear unexpectedly to display a temperature-sensitive transformation phenotype, which will be useful in attempts to study the *Wnt* receptor.) Deletion of the signal peptide prevents secretion, processing, and transformation in paracrine and autocrine assays, but addition of a signal for retention in the endoplasmic reticulum (SEKDEL) allows autocrine but limits paracrine transformation. This result suggests that productive interactions between *Wnt-1* protein and its receptor may occur in the endoplasmic reticulum of C57MG cells. A summary of the mutants and their protein products is presented in Figure 4.

As another means to validate the role of *Wnt-1* as a mammary oncogene, we have developed lines of transgenic mice carrying a *Wnt-1* allele with an MMTV LTR upstream from the gene in the opposite transcriptional orientation, mimicking the situation that occurs most frequently during viral oncogenesis (14). The mice express the transgene mainly in the mammary and salivary glands, manifest mammary hyperplasia in both males and females, and develop mammary (and occasionally salivary) adenocarcinomas. When crossed with transgenic mice carrying an MMTV-*int-2* transgene (*int-2* encodes a fibroblast growth factor and is also frequently activated by MMTV insertion mutation), bi-transgenic mice develop mammary cancer at a considerably accelerated rate in males and virgin females (Fig. 5) showing that two oncogenes encoding secretory proteins can act cooperatively during tumorigenesis (work done in collaboration with Philip Leder, Harvard Medical School).

To explore the role of the virus more fully, we generated molecular clones of MMTV DNA that proved to direct production of infectious, tumorigenic virus (34), solving a chronic difficulty in cloning MMTV DNA. In addition to making MMTV vectors that express heterologous genes in a glucocorticoid-responsive manner, we have used the cloned MMTV to infect *Wnt-1* transgenic mice, again accelerating tumorigenesis (Fig. 6). Tumors from the infected animals have new MMTV insertion mutations, often in known loci (*int-2*, *int-3*, and *hst*), but many tumors are clonal growths with proviruses in unknown sites now under investigation by a former post-doctoral fellow (Greg Shackelford, USC).

The potential importance of the *Wnt-1* gene in development was initially suggested by our studies of the pattern of expression in the mouse: the gene is expressed principally in parts of the embryonic central nervous system in mid-gestation and in the round spermatid (a post-meiotic cell) during spermatogenesis (9, 10). (*Wnt-1* is now recognized to be the mouse homolog of the *Drosophila* segment polarity gene, *wingless*, and required for development of the cerebellum and midbrain in the mouse.) To pursue the developmental effects of *Wnt* genes in another instructive experimental setting, we have cloned (35) and sequenced most of the only detected *Wnt* gene and its cDNA in *Caenorhabditis elegans*; this gene is also expressed in embryos and has intron-exon boundaries that differ from those in other *Wnt* genes. (Fig. 7)

c. The *c-myc* gene.

In the first couple of years of the OIG, my laboratory was heavily involved in studies of *c-myc* and related genes. We helped to refine the structural analysis of chicken *c-myc* (36), described some complex rearrangements that had occurred during downstream insertional activation of *c-myc* in a bursal lymphoma (37), and characterized amplified *c-myc* loci in the COLO320 and SEWA cell lines (38, 39, 40) in collaboration with the Bishop lab. We learned the assay for co-transformation of rat embryo fibroblasts by *myc* and *ras* genes, as originally developed by Robert Weinberg's lab, and used it to show that elevated expression of a normal human or chicken *c-myc* gene is sufficient to collaborate with a mutant *ras* gene (i.e. mutations of the *c-myc* coding sequence are not required) (41) and that *N-myc* has oncogenic activity in the assay (42). The assay was central to our efforts to characterize a large collection of insertion and deletion mutants of human *c-myc* (20); the mutant proteins were also examined for nuclear localization, stability, ability to transform an established line of rat fibroblasts, and DNA binding properties. These experiments allowed us to propose the first rough maps of the functional domains of *c-myc* protein (20, 43, 44), and the mutants are now standard reagents for the community that studies the *myc* gene family.

d. Avian nephroblastoma and *c-Ha-ras*.

At the inception of this grant period, we were attempting to identify the proto-oncogenes activated by proviral insertion in MAV-2-induced avian nephroblastoma. We found a single tumor in which an apparent insertion of MAV-2 DNA initiated high level transcription of the chicken *Ha-ras* gene (45). In the absence of further definition of activated genes among our tumors, we have abandoned work on this problem.

3. Miscellany.

This award has partially supported work only tangentially related to its major themes by a few especially talented and independent post-doctoral fellows, two of whom have obtained additional funding as Scholars of the Markey Foundation. Titia DeLange staged a valiant but ultimately unsuccessful effort to clone cDNA from the Wilms' Tumor susceptibility locus by subtractive hybridization (discussed in Ref. 46) and then cloned and examined structural properties of human telomers (47). David Kaplan, working largely in collaboration with Deborah Morrison in Rusty Williams' laboratory, described interactions between the PDGF receptor and several elements in signal transduction pathways (the Raf kinase, phosphatidylinositol 3'kinase, and GTPase activator

protein; 48-51, 52) and, in collaborations involving Mike Bishop's and Luis Parada's laboratories, examined the tyrosine phosphorylations that follow treatment of PC12 cells with nerve growth factor (NGF), leading to the suggestion that the high affinity NGF receptor is the product of the *trk* proto-oncogene (53, 54). Peter Sorger has begun a genetic analysis of the role of the *CDC28* protein of *S. cerevisiae* in cell cycle control, with emphasis upon phosphorylation events, with additional advice from Andrew Murray.

D. Effects of the OIG.

It would be disingenuous for me to say that the OIG has had a dramatic effect upon the operation or goals of my laboratory or upon the tenor of my own life in science. Still, the award has simplified my relationships with granting agencies, since all my support (save my American Cancer Society Professorship and post-doctoral fellowships) now comes from NIH and mostly from this award. I have, no doubt, saved some time that would previously have been devoted to grant renewals and to more lengthy and more numerous progress reports, but I cannot say that these imposed major burdens (indeed the enforced exercises were often useful).

The stability of funding has, of course, been a highly favorable feature of the OIG, one that has permitted me to take a somewhat longer range view of our research activities than I might otherwise have done. This has probably increased my willingness to commit ourselves to some long-term, high-risk projects; among the successes are the *in vitro* integration reaction, some genetic approaches to *src*, the cloning of avian retroviral receptors, and development of *Wnt-1* transgenic mice, although there are also failures and projects whose outcome remains uncertain. I have also felt more comfortable about allowing a few highly talented post-doctoral fellows take on projects that may not have fit closely into work proposed or ongoing; in some of these situations I have affiliated myself with the work rather loosely and given the fellows the opportunity to publish independently or with collaborators from other labs.

The major problem I have encountered has been financial: with most of my research funding from a single source, budgeted over a seven year period, it is difficult to predict (and obtain financing for) purchase of major pieces of equipment and to adjust to unanticipated increases in expenditures. In my case, the most dramatic example of unanticipated increases has been in the domain of animal-based research, since the use of transgenic animals and our efforts to make gene-deficient animals by homologous recombination have escalated our vivarium costs and required employment of an animal care technician at a time when applications for supplementary appropriations were unlikely to be well received. In addition, the difficulty of buying equipment within an inflexible long-term budget has made much of our laboratory gear outmoded and subject to frequent breakdown. Perhaps the OIG program could be modified to allow reevaluation of the budgets every three years or so.

RESEARCH PROPOSAL

A. Perspective

The vast majority of the work proposed for the next seven years of support by this award represents a direct extension of studies already underway, so the ensuing sections generally coincide topically with the progress report. There is a modest change of emphasis, with more attention given to virus entry and assembly than to ribosomal frameshifting, and with the work on oncogenes generally restricted to the *src* and *Wnt* gene families. (Wherever references are required in this section, they are given by number to our papers and by letter to papers from other labs also listed at the end of the proposal.)

B. Retroviral replication.

1. Retroviral integration.

The nucleoprotein complex that mediates the integration reaction remains poorly characterized. We plan to purify substantial amounts of the MLV complex, as well as complexes formed after infection with other viruses (RSV, HIV) for determination of their protein constituents and for visualization in the electron microscope (in frozen hydrated form, in collaboration with Jim Hogle, Scripps Institute). To determine when the complex dissociates, we will ask whether integration intermediates can be precipitated with antisera known to recognize *gag*-derived or *pol*-derived components of the complex.

We will look more carefully at the apparent periodicity of retroviral integration into nucleosomal domains of minichromosomes, using the newly developed PCR assay (Fig. 1) to analyze mass populations of integration products (rather than cloning and sequencing of individual products as in the past) and using an MMTV LTR-based minichromosome on which a nucleosome has been stringently positioned (a). We will also be evaluating evidence that integration site preference is biased by sequence as well as nucleosome position, as suggested by recent experiments (e.g., Fig. 1 and ref. b). To this end, we will consider cloning integration sites preferred during MLV infection, in the manner described for RSV by Coffin's laboratory (b); the cloned DNA could then be studied as targets *in vitro*, in the form of naked DNA and reconstituted chromatin. In addition, we have recently found that we can monitor integration events mediated by either nucleoprotein complexes or purified HIV or MLV IN, using a wide variety of DNA's as targets. Such experiments should be informative about sequence preferences during integration.

The PCR-based assay of integration events also allows us to assess the effects of protein binding, DNA bending, and other perturbations of the target upon integration efficiency. (For example, the yeast $\alpha 2$ protein prevents integration of MLV DNA into the $\alpha 2$ binding site, simulating a DNase footprint.) We will use a variety of model systems (e.g., lambda phage *att* sites and integration proteins) to assess the influence of DNA structure and associations upon integration site use.

Our recent success in obtaining SV40-MLV cointegrates from coinfecting monkey cells allows us to study the effects of replication, transcription, and packaging of SV40 DNA upon its use as a target for MLV integration. For this purpose, we have obtained several conditional mutants of SV40 affecting T antigen and VP1 (c). In addition, we have isolated SV40 minichromosomes for use as targets for integration *in vitro*. Dramatic changes in MLV integration site preference in SV40 DNA occur during the SV40 life cycle, and we will attempt to identify the determinants of these changes.

Integration proteins are capable of specific binding to viral *att* sites, removing 2 nt from the 3' ends, cutting target DNA, and joining it to the viral substrate, both *in vivo* and *in vitro* (55, c').

To begin to assign domains for these various functions to these relatively small proteins (32-46 Kd), we are enlarging our repertoire of MLV and HIV integrase mutants by site-directed mutagenesis at conserved residues and producing chimeric (MLV-HIV) proteins; these will be expressed in yeast (or *E.coli*) and tested in reactions for each activity in vitro. (The MLV and HIV components in chimeras can be distinguished by differences in *att* site preferences and in the distances between staggered nicks in target DNA.) Ultimately, we hope to be able to correlate a genetic definition of the domains with a structural definition of the proteins obtained with Xray crystallography (a collaboration with Bob Stroud, UCSF, funded by another award).

We also plan to exploit the available assays for integrase function to define the nucleotide sequence requirements for *att* sites and chromosomal targets more rigorously. As a first step, we will continue to study the several MLV and HIV *att* site variants made for binding studies, and we are examining the utility of various oligonucleotides, varying in length and sequence, as target sites. (We have recently found that oligonucleotides other than cognate *att* sites as used in the assay described by Craigie [c'] can serve as targets for IN-mediated integration.) Ultimately, we plan to generate banks of oligonucleotides with randomized sequences at appropriate positions (d) to ask which sequences favor or interdict the steps in the integration reaction.)

We are continuing to study the mechanism of Fv-1 restriction of MLV infection. The major new goal is to attempt cloning of the Fv-1 locus, using a gene transfer assay in which resistance (the dominant trait) will be scored by infecting colonies of cotransfected cells with N or B tropic strains of MLV. We are also testing whether overexpression of part or all of *gag* p30 from restricted strains can abrogate resistance. This might allow improved mapping of the Fv-1 target and provide a means to isolate the Fv-1 gene product by affinity chromatography or the gene by protein affinity screening of cDNA expression libraries.

2. Retroviral receptors and virus entry.

As summarized in the progress report, we have recently obtained genomic clones that appear to encode the subgroup A receptor gene from the chicken and quail *rv-a* locus. We are now attempting to finish the sequencing of these clones, identify receptor mRNAs, clone and sequence the cDNAs, and test all relevant clones for biological activity. In addition, we will want to know whether exonic clones detect non-permissive alleles at the *rv-a* locus, the other chicken loci (*rv-b* and *rv-c*) believed to encode receptors for the RSV subgroups B, C, D, and E, and related genes in other birds and mammals. Attempts to understand the normal function of what we expect to be an extensive, polymorphic gene family will begin with analysis of predicted protein sequences, production of antibodies, and expression of receptor genes in various contexts. It would be premature to guess the nature of the receptors, especially in view of the heterogeneous nature of the retrovirus receptors whose genes have been cloned (receptors for HIV, ecotropic MLV, and gibbon ape leukemia virus [e, f, g]).

In the more distant future, we will attempt to define the site and nature of the interaction between RSV envelope proteins and cognate receptors, making use of cloned RSV *env* genes and receptor genes. Chimeric and mutant receptors will be produced, and interactions with RSV or *env* proteins will be examined functionally (by tests for permissivity to infection in cell culture) and biochemically (in collaboration with Judy White, Dept. of Pharmacology, UCSF). Another long-term goal is an explanation of how reverse transcriptase is activated during virus entry to produce a substrate for the integration machine. We will also use receptor clones to produce transgenic mice that can be infected with RSV; this will permit the development of useful experimental reagents--- mice and mouse cell lines that can be efficiently infected with high titre, replication-competent RSV vectors, which will not produce infectious virus due to intracellular blocks to later stages of the virus life cycle.

We have recently begun to apply our experience with avian viral receptors to a similar problem: the cloning of another potentially polymorphic family of transmembrane proteins that serve as receptors for feline leukemia viruses (FeLVs). Our primary strategy is to use receptorless cells (mouse 3T3 cells) as recipients for cat DNA, scoring positive cells with virus vectors bearing various FeLV glycoproteins and selectable genetic markers. In this venture we are collaborating with Jim Mullins (Stanford) who has shown that pathogenic potential of strains of FeLV can be ascribed to variations in *env* genes (h). We hope that the nature and distribution of the receptors for pathogenic variants will be informative about mechanisms of pathogenesis (e.g. anemia and immunodeficiency).

3. Viral assembly and cell targeting.

We are pursuing our recent discovery that human CD4 is efficiently incorporated into RSV particles in several ways. First, we are using a variety of other ectopic membrane-associated proteins to determine which can be assembled efficiently into RSV. We need to confirm preliminary evidence for incorporation of CD8, the poly IgA receptor, and influenza HA protein, and we are also testing the complement (and Epstein-Barr Virus) receptor CR2 and Wnt-1 protein. When the cDNA for the RSV-A receptor has been cloned, we will also ask whether that receptor can be incorporated. To explore the criteria for inclusion, we will test several variant versions of these proteins, including chimeras in which the transmembrane and cytoplasmic domains are replaced by RSV *env*-derived sequence (thus far, CD4-RSV chimeras perform much less well than wild type CD4), a collection of cytoplasmic tail mutants of the poly IgA receptor (provided by Keith Mostov, UCSF), and proteins (influenza HA and Wnt-1) linked to the membrane by phosphatidyl-inositol rather than transmembrane domains.

Our currently favored explanation for why ectopic proteins enter virus particles, whereas the vast majority of endogenous membrane-associated proteins do not, is that membrane proteins lacking a cytoplasmic partner proceed via a default pathway to regions of the membrane that ultimately form viral envelopes. We will test this idea by supplying suitable cytoplasmic partners (e.g. the Lck protein-tyrosine kinase for CD4 and CD8 [i]) or by attempting to anchor proteins in the cell by adding cytoplasmic tails that should recognize cytoplasmic proteins endogenous to avian fibroblasts. (For example, the *env-sea* fusion protein encoded by the S13 avian tumor virus is not present in virus particles [j], suggesting that addition of *sea*-encoded sequences to membrane proteins might impede assembly.)

Finally, we will ask whether any of the viruses carrying ectopic proteins in their envelopes can be directed to new cell types as a result of interactions with membrane proteins on the target cells. Although this is being tried initially using HIV gp120-producing human cells as targets for RSV carrying CD4, we will examine this question with viruses carrying any of the tested proteins. (Binding of virus particles, synthesis of viral DNA, and expression of a selectable marker in a virus vector are among the assays that will be used.) These experiments raise questions of obvious practical significance for virus vectoring, but also address interesting speculative issues about the range of membrane proteins that can be used as ligands and receptors to mediate virus entry. For example, can virus infection proceed normally if the usual receptor is in the virus particle and the viral envelope protein is on the surface of the target cell?

4. Ribosomal frameshifting.

Three components are believed to influence the -1 frameshifts that allow synthesis of *pol* gene products: (i) the nucleotide sequence at the frameshift site, (ii) secondary structure of downstream RNA, and (iii) host tRNA species. (i) We are examining a large number of potential frameshift sites (naturally occurring sites, synthetic sites that conform to the "rules", and mutant sites) to observe frameshift frequencies in various RNA contexts; we will then re-search the sequence data bank to attempt to identify cellular and viral genes that might exploit the frameshifting

opportunities. (An earlier more restrictive search, based upon a less complete understanding of the components involved, brought forth few candidates other than retroviruses and coronaviruses.) (ii) To document the role of RNA secondary structure, we have initiated a collaboration with Ignacio Tinoco's lab (UC Berkeley), using biochemical and biophysical (NMR) means (k) to study the pseudoknot that we have recently proposed to exist in MMTV RNA on the basis of site-directed mutants. We hope to use the structural information to design RNA structures that promote frameshifting, considering both components folding in *cis* and those that anneal in *trans*. We also plan to examine the physiological significance of the structure more closely: unpublished experiments that follow nascent polypeptides imply structure-dependent ribosomal pausing at the frameshift site, but we can now examine this point more precisely with the ribosome mapping procedure of Wolin and Walter (l). (iii) To identify the tRNAs that mediate frameshifts, we are collaborating with Dolph Hatfield (NIH), supplying him with a variety of RNAs with wildtype and mutant frameshift sites for assay in a tRNA-dependent system he has developed in *Xenopus* oocytes.

We also plan to judge the influence of frameshift efficiency upon virus production, using site and structure mutants of RSV and perhaps other viruses; preliminary results indicate that mutants with modest effects on frameshifting *in vitro* have dramatic effects on replication *in vivo*. The role of *trans*-acting factors in frameshifting will be evaluated by tests for RNA binding proteins specific for downstream structural elements (a search newly encouraged by evidence that HIV frameshifting is more dependent on RNA structure *in vivo* than *in vitro*) and by genetic modulation of frameshifting in budding yeast, using retroviral frameshift signals known to function in heterologous contexts (m).

C. Retroviral oncogenes.

1. The *src* gene family.

We are making a major effort to understand the role of the SH2 domain in mediating the regulation of cytoplasmic protein-tyrosine kinases and the interactions among proteins involved in signal transduction (n). (i) SH2 mutants of activated *c-src* that display a host-dependent capacity to transform cells will be exploited to look for cellular genes whose products have functionally significant interactions with pp60. By mutagenesis of mouse cells or by transfer of chicken genomic DNA (or cDNA), we will attempt to induce *src*-dependent transformation of mouse cells expressing SH2 mutants previously shown to be hyperactive in chicken cells but biologically inactive in mouse cells (16, 17). Chicken genes or mutant mouse genes that appear to suppress the host-range defect will be cloned by gene transfer techniques. (ii) We will pursue the dominant negative characteristics of SH2 mutants that interfere with transformation of mouse cells by activated *c-src* (18). Dominant-negative alleles, M6 and M9, with additional mutations (e.g., deletion of the kinase domain or loss of the myristylation signal) will be tested for inhibitory activity to identify determinants of the dominant negative phenotype and to isolate mutants with a more potent phenotype. Appropriate alleles will be tested for their ability to interfere with transformation by oncogenes in the *src* gene family (especially *hck* and *lck*), genes encoding transmembrane tyrosine kinases, and other classes of oncogenes. We will also ask whether such alleles can interfere with cell signalling events thought to be mediated by SH2-containing proteins, including *lck* in T cell activation (in collaboration with Dan Littman, UCSF) or *hck* in myeloid differentiation (with N.Quintrell in Mike Bishop's lab). (iii) We will attempt to identify cellular proteins responsible for the differences in enzymatic activity and substrate preferences of SH2-*src* mutants synthesized in chicken and mouse cells (17, 18). Mutant proteins immunoprecipitated from those cells or purified in large amounts from the baculovirus-based expression system (30) will be tested for modulation of kinase activity after incubation with extracts from avian and mammalian cells. We have synthesized wild type and mutant SH2-*src* domains in *E.coli* and will attempt to use these truncated proteins to inhibit modulating activities and to purify them by affinity chromatography. The wild type and mutant truncated proteins will also be used as binding reagents to examine the effects of functionally significant SH2 mutations upon the ability of SH2

domains to recognize other proteins, especially phosphotyrosine-containing proteins. (iv) One of our host-dependent SH2-*src* mutants (M9) induces tyrosine phosphorylation of many proteins in NIH3T3 cells, although it fails to transform these cells (18). We will look among the proteins known to be tyrosine phosphorylated in *src*-transformed cells for those that might be implicated in the failure of M9 to transform; GTPase activator protein appears to be one such protein and its role in *src* transformation will be further examined in this context.

Assumptions about the roles of *c-src* in normal cell growth and development have recently been reevaluated in the light of Soriano's revelation that *c-src* deficient mice survive through embryogenesis and early post-natal life, usually succumbing to malnutrition and osteopetrosis as a result of impaired osteoclast function (o). These findings suggest that in the vast majority of cells in vertebrate organisms *c-src* is not an essential gene, despite dramatic conservation of coding sequence and widespread expression of the gene. Our group and others are endeavoring to make mice deficient in other members of the *src* gene family to test the hypothesis that the functions of *src* can be complemented by other members of the family. Because of our familiarity with *hck* and *fgr*, two *src*-like genes expressed predominantly in the myeloid lineage (25,26, p), we have initiated efforts to interrupt these genes in ES cells (Bradley's AB1 line) and to produce chimeric, heterozygous, and ultimately homozygous mutant animals by blastocyst injection and breeding. As discussed in the progress report, we have learned to mutate *c-src* by homologous recombination in ES cells (obtaining recombinants at a frequency of about 1 per 100 cells incorporating DNA after electroporation), we have made similar targeting vectors with genomic clones of *hck* and *fgr*, and we have recently obtained many clones of ES cells with targeted mutations of *hck*. Stem cells identified (by PCR and Southern blotting) to have *hck* or *fgr* mutations resulting from homologous recombination will be injected into blastocysts (in collaboration with P. Soriano, Baylor); chimeric progeny will be tested for their ability to transmit the mutant gene; and mice homozygous for the deficiency will be bred and examined for phenotypic changes in several tissues and cell lineages at different developmental stages. Naturally, special emphasis will be given to the observation and manipulation of hematopoietic lineages *in vitro* and *in vivo*. The heterozygous animals will be crossed as appropriate with mice carrying null mutations in other *src* gene family members (and mice with mutations in other genes important in hematopoiesis, e.g. *abl*) as those animals become available. One obvious question to be addressed by such studies is whether certain differentiated properties or cell growth are affected in mice lacking multiple family members, suggesting that *src*-type genes can complement each other.

We are using SV40-immortalized, *src*-deficient mouse fibroblasts obtained through a collaboration with P. Soriano (Baylor) to study the location and enzymatic activity of *c-src* protein during the cell cycle. As noted earlier, we and Shalloway's group have shown that during mitosis pp60^{*src*} is more active as a kinase and phosphorylated on serine and threonine residues near the amino terminus by the *cdc2* kinase (q,19). We have made mutations at the mitotic phosphorylation sites in human *c-src* protein to gauge the influence of these mutations on the kinase activity and the location of pp60 during interphase and mitosis in the "background free" cell context. Preliminary results with *c-src* endogenous to, or over-expressed in, normal rat cells using immunofluorescence methods indicate that pp60^{*c-src*} is located in several sites in addition to the plasma membrane. These include an association with microtubules (also see ref.r), the microtubule organizing center, the mitotic spindle, and perinuclear membranes that appear to be late endosomes, as they are enriched for the mannose-6-phosphate//IGF2 receptor. We are attempting to confirm and refine these results with biochemical fractionation of cell organelles and immunoelectron microscopy in various cell types. Potential substrates for the *src* kinase will also be sought in relevant fractions.

2. The *Wnt* gene family in development and oncogenesis.

The search for receptor(s) that mediate the developmental and oncogenic actions of *Wnt* family proteins is central to our current and future studies of *Wnt* genes. After considerable unsuccessful

efforts to purify biologically active *Wnt-1* protein to serve as a radioactive ligand in this search, we have initiated several other approaches to the receptor and its gene. In the long run, we aspire to characterize what is likely to be a family of receptor genes for the structure and function of the proteins, the pattern of expression, and the mode and pathway of signal transmission. A sampler of possible routes to these genes follows.

(i) In hopes that NIH3T3 cells, which do not respond to *Wnt-1*, will respond if provided with receptor, we will transfect cells programmed to express *Wnt-1* with cDNA expression libraries prepared from responsive cells (PC12 and C57MG cells) and with human genomic DNA. Although we will look initially for traditional properties of transformed cells (focus formation or growth in agar), we are also planning to identify genes that respond to a *Wnt-1* signal by augmented expression. This would allow us to use reporter genes linked to promoters from such genes to monitor for a *Wnt-1* response in cells that might not manifest an otherwise recognizable phenotype. We will use differential cDNA cloning and some educated guesses (e.g., genes known to respond to growth factors [s]) to find such genes. (We have recent evidence that this approach may work since we have found that one early response gene is induced after temperature shift-down of C57MG cells expressing a *ts* allele of *Wnt-1*.) We will be able to distinguish spontaneous transformants from *Wnt*-dependent transformants, using our newly isolated *ts* alleles of *Wnt-1* or *Wnt-1* retrovirus vectors that also carry a gene (HSV *tk*) we can select against. (ii) In hopes of producing a *Wnt* protein active in soluble form, we are making a series of deletion mutants and fusion proteins (e.g. *Wnt-1* - IL2 chimeras), and we will screen the mutants for biological activity in a new and simplified paracrine assay using transiently transfected quail cells and C57MG cells as responders. (This assay reduces the time required from 2 weeks or more to 4 days.) (iii) We are testing *Wnt* genes from other organisms (fly, worm) and other murine members of the *Wnt* gene family for their ability to transform C57MG cells. If any fail to transform and thus do not properly activate the *Wnt* receptor that must already be present in those cells, we will attempt to isolate the cognate receptor from the appropriate organism, using transformation of C57MG cells as an assay. (iv) We are attempting to make retroviruses that carry *Wnt-1* proteins on their surface, inspired by our success in producing RSV particles that contain CD4 (see above). Such viruses will be tested for biological activity in transformation assays and for use as radioactive ligands for *Wnt* receptors. (v) *Armadillo*, a gene that acts downstream of *wingless* in the fly, encodes the insect homolog of plakoglobin, a constituent of desmosomes and some cell-cell junctions (t). We have obtained antibodies (from Drs. P. Cowin, NYU; J. Stanley, NIH; and B. Gumbiner, UCSF) against several membrane glycoproteins associated with plakoglobin-containing structures (desmogleins, desmocollins, desmoplakin and cadherins) to attempt to inhibit transformation in paracrine assays, on the assumption that one of those glycoproteins might be a *Wnt* receptor. (vi) We have preliminary evidence (based upon *in vitro* kinase assays of C57MG membrane proteins precipitated from control and *Wnt-1*-expressing cells with anti-phosphotyrosine antibodies) to suggest that *Wnt* receptors might be protein-tyrosine kinases. We will explore this further in other cell types and through better characterization of the proteins phosphorylated in the *Wnt*-dependent assay. (vii) We are using a PCR-based approach to find mutants in mutagenized worm populations with deletions or base substitutions that eliminate restriction sites in the *C.elegans* *Wnt* gene (in collaboration with Cynthia Kenyon, UCSF). After characterizing any mutants obtained, we will look for additional mutants with similar phenotypes or for suppressor mutants, in hopes of finding abnormal receptor alleles.

A number of additional experiments will address the role of *Wnt* genes in development. (i) We are using a sensitive and quantitative PCR-based assay to measure the amount of *Wnt-1* RNA in a wide variety of mouse tissues, revealing previously undetected RNA in some organs (e.g. one transcript per 50 cells in the thymus) and stringently documenting the absence of RNA in others (e.g. less than one transcript per 100,000 cells in the mammary gland). This assay will soon be extended to other members of the *Wnt* family (31, u). (ii) We are trying to identify the region of the *Wnt-1* gene responsible for regulated expression during mouse development. Since over 5 kb of sequence upstream of the transcriptional start has proven insufficient to drive appropriate

expression of a beta-galactosidase reporter gene in transgenic animals, we are making constructs that contain introns, exons, and downstream sequence as well. If successful, the resulting transgenic mice will permit a more precise definition of the cells that express *Wnt-1* in the embryo and adult. In addition, we will be able to make transgenic mice that express other members of the *Wnt* family and *Wnt-1* genes with site-directed mutants active in autocrine but not paracrine assays, to ask, in collaboration with Mario Capecchi (Univ. of Utah), whether such transgenes can complement a *Wnt-1* deficiency (v). (iii) Preliminary evidence, obtained with the modified paracrine assay for *Wnt-1*, indicates that some of our site-directed mutants interfere with the paracrine effects of wild type *Wnt-1* (i.e., are dominant negative mutants). If so, these alleles will be tested for developmental effects in *Xenopus* (w), in collaboration with M. Kirschner (UCSF), or in mice. (iv) We will more fully characterize the pattern of expression of the *C.elegans Wnt* gene, using in situ hybridization and Wnt antibodies. Overexpression of the gene and expression in ectopic sites will be attempted using new methods for gene transfer (w'). These studies could be influenced by placement of *Wnt* on the "contig map" of the *C.elegans* genome (the gene is now mapped to an unattached DNA "island") or by success in our other attempts to make *Wnt* mutants in the worm.

We will also be extending our efforts to use MMTV and *Wnt-1* in models for understanding multi-step carcinogenesis. (i) We plan to cross *Wnt-1* transgenics and *Wnt-1/int-2* bi-transgenics with animals carrying other oncogenic transgenes: an MMTV-TGF-alpha transgene [x; from Bob Coffey, Vanderbilt] and an MMTV-*int-3* transgene [from Bob Callahan, NIH], or inactivated tumor suppressor genes (Rb and p53, from Tyler Jacks in Bob Weinberg's lab, MIT). Each of these has its attractions. TGF-alpha is another secretory protein, one that (like *int-2* protein) activates a transmembrane receptor with protein-tyrosine kinase activity and acts as a weak mammary oncogene; it is of interest to know whether three secretory proteins (*Wnt-1*, *Int-2*, and TGF α), at least two of which act through similar signalling pathways, can synergize during tumorigenesis. *Int-3* encodes a (truncated) transmembrane protein without a kinase domain; as a transgene it involutes mammary epithelium (in contrast to the proliferative effect of *Wnt-1*) but it also leads to mammary tumors; how will the *Wnt-1* and *int-3* transgenes interact in preneoplastic tissue? Mutant tumor suppressor genes are often found in human mammary carcinoma, but their role in tumorigenesis has been difficult to study, and our surveys with p53 antisera have failed to show p53 mutations in our transgenic tumors. A transgenic model for cooperation between dominant and recessive mutations is potentially powerful for simulating the interactions that appear to occur in the human disease. (ii) Viral determinants of MMTV-induced tumors are not well understood. We have made mutations in the open reading frame of the LTRs in our infectious clone of MMTV; provisional evidence indicates that the ORF mutants replicate well in culture and in the animal, but are deficient in tumor induction. We are repeating this experiment on a larger scale in BALB/c animals and in Czech 3 mice, devoid of endogenous MMTV proviruses (in collaboration with Bob Callahan, NIH). These experiments take on new significance in view of the recent discovery that MMTV ORFs encode superantigens that activate or delete T cells of defined V-beta subclasses (y). We will devise experiments to ask whether the nature of endogenous or exogenous MMTV ORFs influence mammary tumorigenesis after virus infection or in the transgenic model by deletion or activation of T cells that might be implicated in tumor immunity. (iii) MMTV-induced and transgenic mammary tumors are generally invasive and sometimes metastatic; we will ask whether such tumors are associated with activation of genes that have been implicated in tumor progression, including expression of matrix metalloproteases in stromal cells (such as stromolysin -3 [z]) and angiogenic factors.

Use of Vertebrate Animals

1. We are studying the genetic basis of cancer, using a variety of retroviruses and viral and cellular oncogenes to induce tumors in widely-studied laboratory animals (chickens, mice and rats). Our ultimate objective is to identify and characterize several genes that contribute to the development of cancers in animals, in hopes that this information will be useful in designing attacks upon human cancer and germane to an understanding of normal regulation of growth and development. Studies of the sort we describe here have been important in transforming a completely mysterious process (the origins of a cancer cell) to one we can now attribute with confidence to mutations in a few isolated genes in many circumstances. Continued study of such oncogenes biochemically in test tubes and functionally both in cultured cells and in experimental animals seems very likely to have a significant impact upon cancer in the next decade. Although we and others do as much as possible with cultured cells, there is no completely satisfactory correlation between the tumorigenic effects observed in the animal and changes noted in cells in a petri dish. Hence we must continue to rely upon assays in animals, particularly when certain cell types (mammary cells, blood cells, etc.) are involved. Finally, biochemical study of oncogenes and viruses depends heavily upon appropriate antibodies that are most reliably produced by immunization of rabbits - and sometimes by mice or chickens. This too involves time-honored and virtually painless methods.

Allocation of Animals and Rationales

For studies involving tests of the tumor-inducing capacity of viruses or cell lines in chickens, mice and rats, a minimum of 8-12 animals is required in each experiment to obtain a meaningful result, since at least two dilutions of the inducing material must generally be used. Naturally it is impossible to estimate how many experiments will be done in any given year; the number of animals requested is based upon past performance. For studies involving the induction of antibodies in rabbits, chickens, or mice, it is customary to immunize 3-4 animals for each antigenic preparation, since animals may differ appreciably in their immunological response. For studies involving the generation of transgenic mice, it is generally necessary to use 10-30 animals to obtain a single "founder" animal in which the foreign DNA has been inserted in the germ line. Since 5-10 "founders" are required to obtain meaningful results with each transgene, single experiments may use as many as 300 mice.

Anticipated Results

Our general goals are to understand the molecular basis of cancer and the manner in which retroviruses grow. The few types of experiments in which we must rely upon animals rather than cultured cells are intended to give us the following kinds of information:

- i). To assess the cancer-causing effects of an oncogene, it is necessary to examine the oncogene directly in an animal. Through such tests and correlations with biochemical properties of oncogenic proteins measured in the test tube, we expect to learn the attributes of such proteins that are responsible for their tumor-causing potential.
- ii). In studies of oncogenic proteins and of viral and cellular proteins that participate in the virus multiplication cycle, it is necessary to have antisera that can specifically recognize proteins of interest, both for purposes of detection and isolation of those proteins. We expect the immunization studies to provide us with the necessary sera.

The use of transgenic mice for the study of oncogenes has revolutionized cancer research because it is now possible to produce animals in which all of the target cells are afflicted by an initial step in carcinogenesis. This provides an ideal setting in which to seek the still largely mysterious secondary events in neoplasia. In the experiments we are performing, we expect to learn about such secondary events in breast cancer and, in addition, to learn about the role of proto-oncogenes (the normal forms of oncogenes) in development of an organism.

We will use both male and female mice from strains C57B6, BALB/cB45, SJL, and FVB. Ages will vary from 3 weeks to 12 months. We will maintain approximately 1,500 mice per year.

2. We have chosen mice because for our work, they are the cheapest mammalian species available and possess a fast reproductive cycle. The estimated size of our colony reflects our need for an ample statistical base for these studies.

3.-5. Veterinary care and euthanasia procedures comply with stringent guidelines promulgated by UCSF's Committee on Animal Research. These guidelines are based upon PHS regulations.

PCR-BASED ANALYSIS OF MLV INTEGRATION INTO
YEAST TA MINICHROMOSOMES AND NAKED DNA

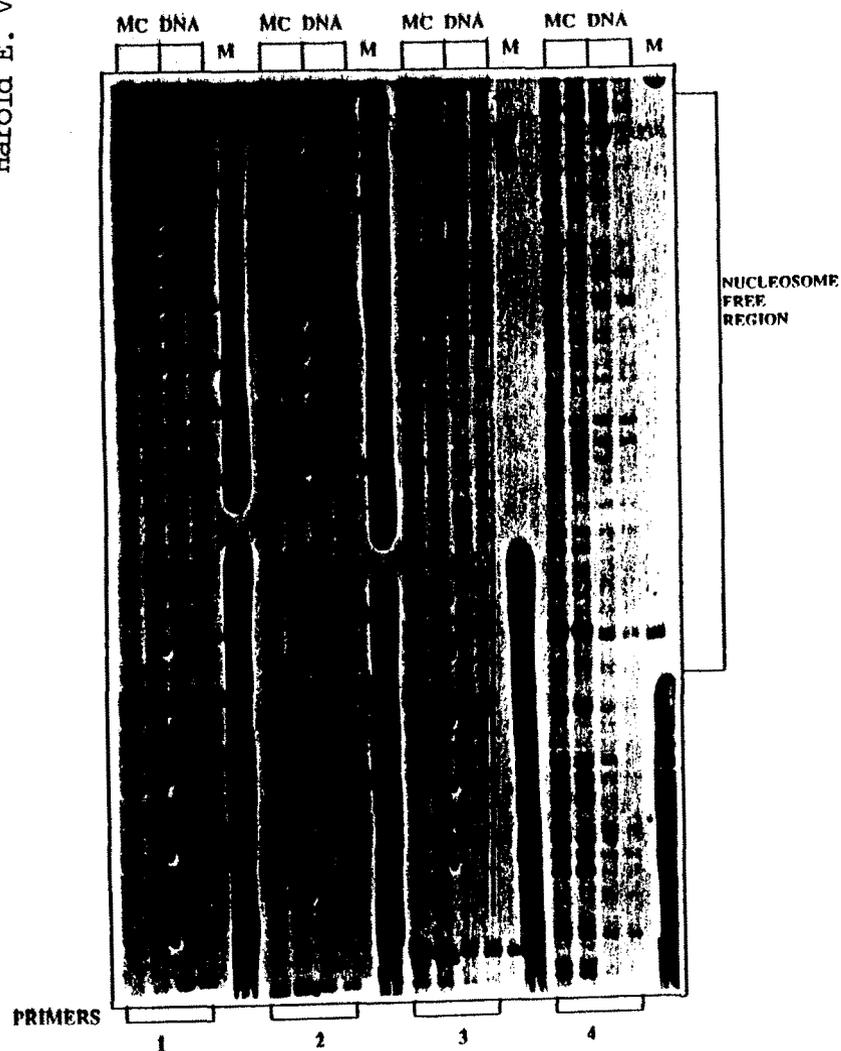


Figure 1. Polymerase chain reactions with ^{32}P -labeled oligonucleotide primers were used to characterize the products of integration of murine leukemia virus DNA by nucleoprotein complexes into naked Trp-Ars plasmid DNA or minichromosomes prepared from *S.cereveciae* (z). Each primer set included one primer for a sequence near the end of an MLV LTR and one for each of four positions on the 1.5 Kb Trp-Ars DNA. Each reaction was performed in duplicate on both naked DNA (DNA lanes) and minichromosomes (MC), and markers (M) of known size were electrophoresed in parallel. The sausage-like drawings indicate the positions of mapped nucleosomes; two nucleosome-free regions, representing the origin of replication and the transcriptional control region, are also shown.

At least four conclusions can be drawn from this analysis: (i) integration does not occur randomly in naked DNA, since some sites are clearly preferred; (ii) there is little or no difference between use of nucleosome-free regions of MC DNA and the homologous regions of naked DNA; (iii) different sites are favored for integration events in nucleosome-containing regions of MC DNA and many of these sites are used more efficiently than the favored sites in naked DNA; and (iv) many of the favored sites in nucleosomal domains are positioned 10-11 bp apart (as denoted by the black dots). This periodicity confirms earlier analyses by cloning and sequencing of integration products and predicts that integration events occur preferentially on one exposed face of the DNA helix wound around phased nucleosomes.

<u>Virus</u>	# HygromycinB-resistant colonies	
	<u>3T3 cells</u>	<u>3T3 transfectants</u>
AH (1ml)	0	confluent
AH (0.1ml)	0	357
AH (0.01ml)	0	80
BH (1ml)	0	30
None	0	0

Fig.2. A 5.5kb quail genomic DNA clone confers susceptibility to ALV infection upon mouse 3T3 cells. A 5.5kb quail genomic DNA clone derived from the ALV-receptor locus and a plasmid conferring resistance to G418 were cotransfected into mouse 3T3 fibroblasts and transfectants were selected using G418. 1.5×10^5 cells derived from this pool of transfectants were infected with either subgroup A-specific (AH) or subgroup B-specific (BH) ALV vectors containing a hygromycinB-resistance gene. An equivalent number of untransfected 3T3 cells were challenged with these viruses for control purposes. Infected cells were selected using hygromycinB. Preliminary analysis suggests that the AH and BH virus stocks used in these experiments were of equivalent titre. These findings indicate that introduction of the quail DNA clone into mouse cells confers a high degree of susceptibility to subgroup A RSV vectors: subgroup B vectors are at least 100-fold less able to infect the transfected cells, confirming that the cloned receptor locus is *sva*.

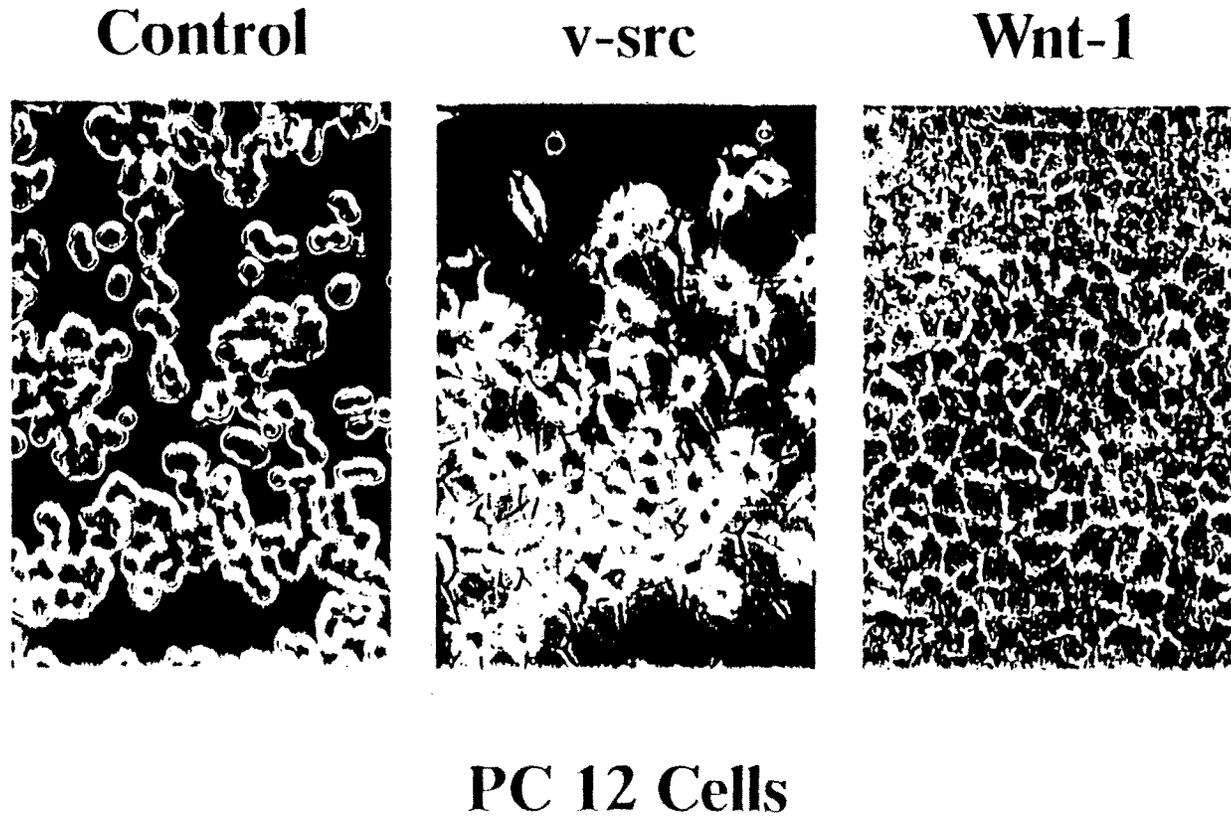


Figure 3. The rat pheochromocytoma line, PC12, was infected with MLV vectors carrying the v-src or Wnt-1 genes, plus a bacterial neomycin phosphotransferase gene as a selectable marker. Colonies of cells pictured here show the induction of neurite outgrowths previously observed after expression of v-src and the flattened, adherent behavior of cells expressing Wnt-1.

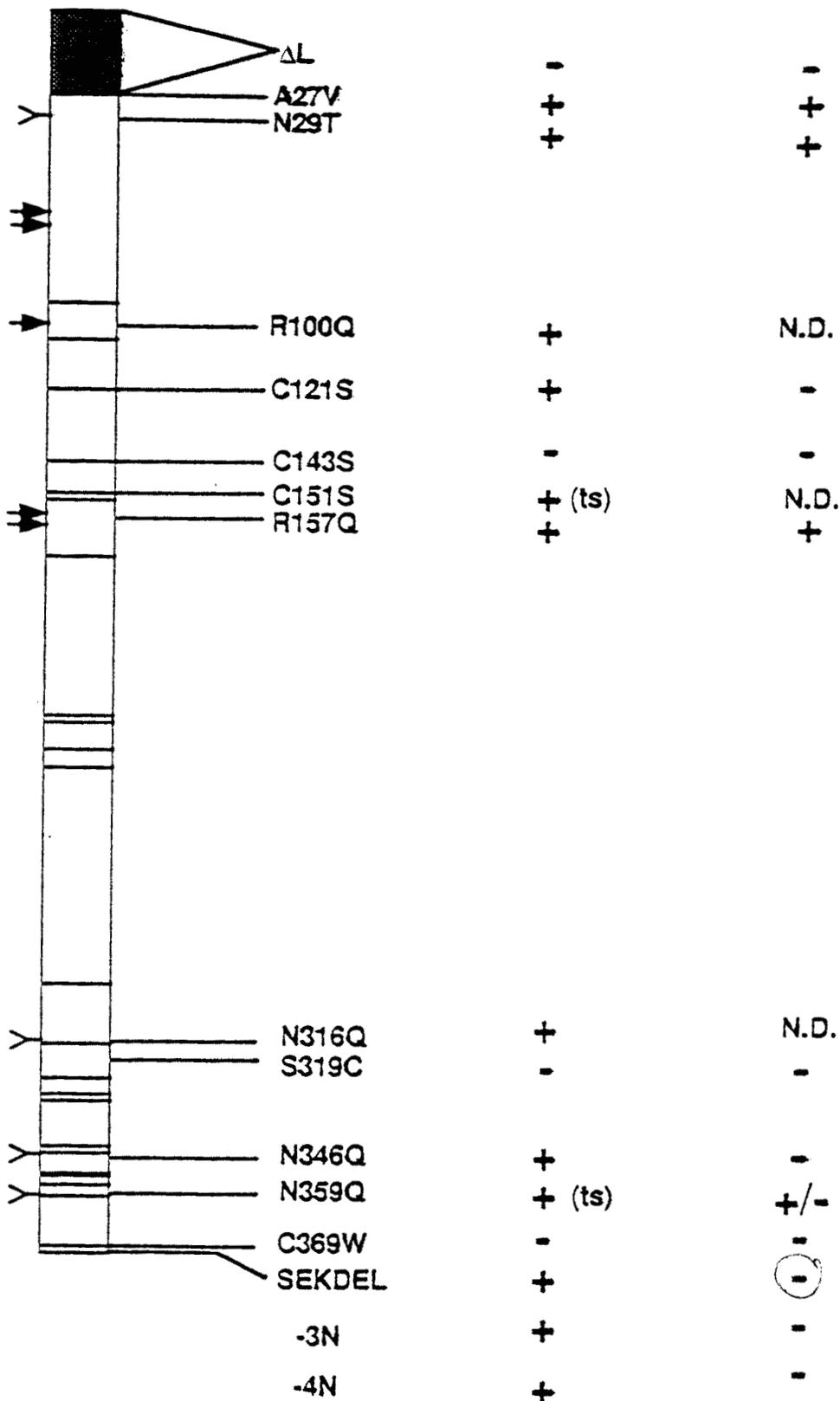


Fig.4A Mutants of mouse Wnt-1 were generated by site-directed mutagenesis and tested biologically for autocrine and paracrine transformation of the mammary epithelial cell line, C57MG, after introduction into C57MG or NIH3T3 cells in MLV vectors. The diagram at the left shows the predicted primary product of translation of Wnt-1 mRNA, with the shaded signal peptide at the aminoterminalus (top). N-linked glycosylation sites are denoted by forked lines; the arrows indicate dibasic peptides that are possible sites for serine proteases (probably not used); and horizontal lines show the positions of cysteine residues. ts, inactive at 40°C but active at 34-37°C; N.D., not done.

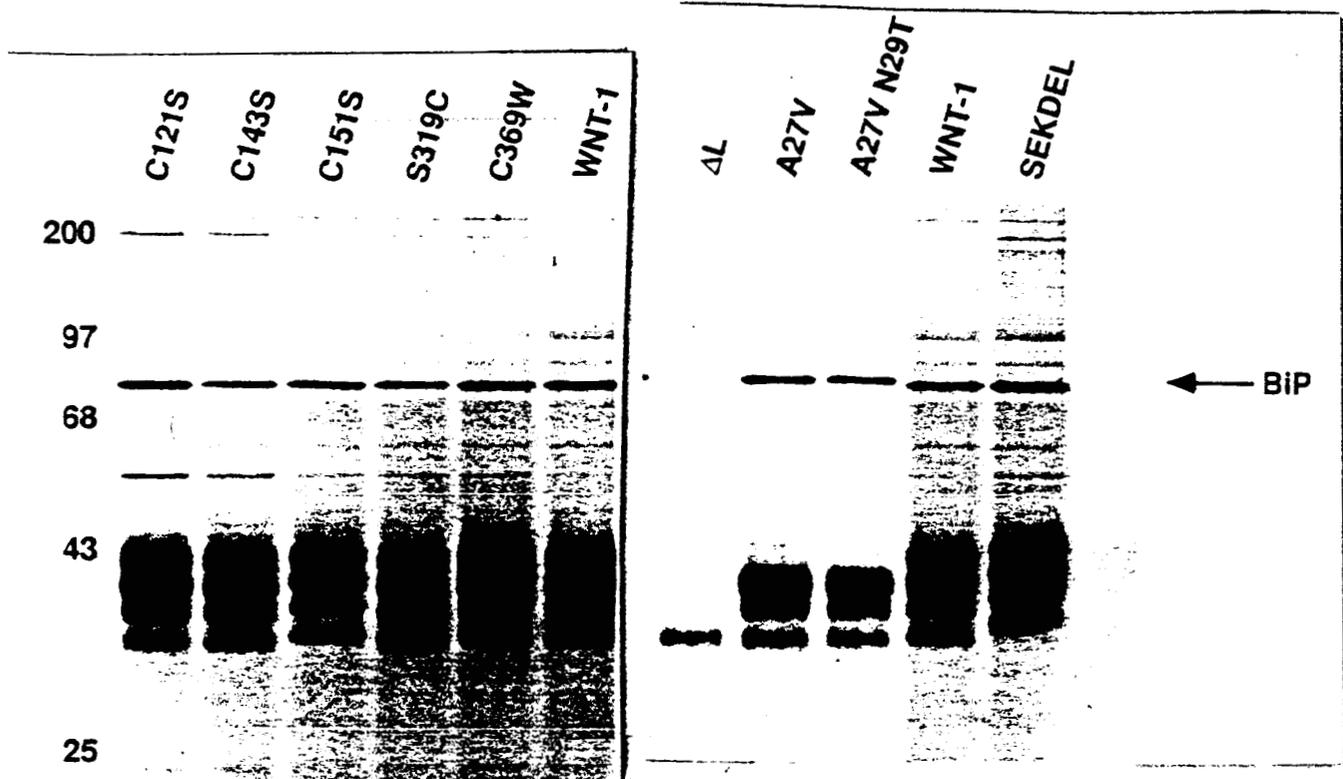
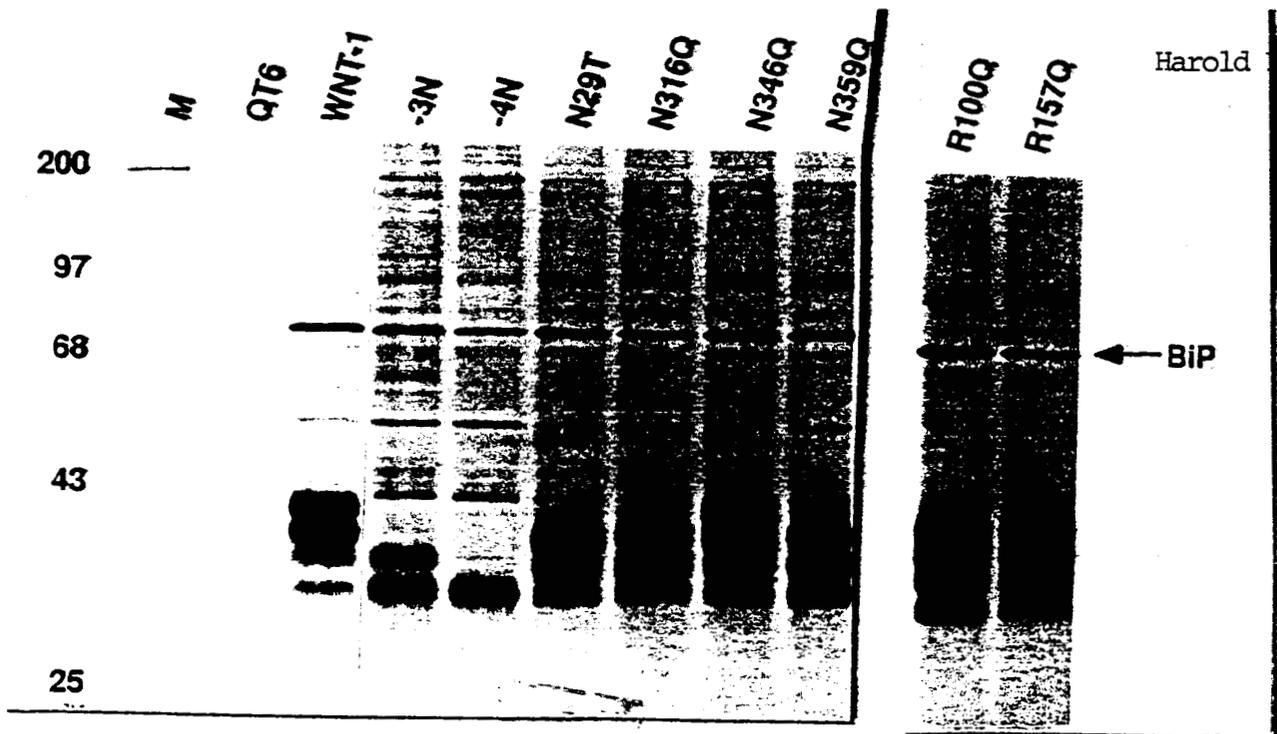


Fig.4B. Proteins synthesized by the Wnt-1 mutants listed in Figure 4A were detected by transient transfection of quail tumor cells (QT6) with plasmid vectors containing the indicated alleles driven by a cytomegalovirus immediate early promoter. Proteins were labeled for three hours with ^{35}S methionine and cysteine before immunoprecipitation with a Wnt-1-specific monoclonal antibody and subjected to polyacrylamide gel electrophoresis. The unglycosylated and three most abundant glycosylated forms of Wnt-1 protein are observed in extracts from cells expressing the wild type Wnt-1 gene; the analysis of glycosylation site mutants shows that N346 is the site least efficiently used. (The faint band representing Wnt-1 protein glycosylated at all four sites is not visible in these analyses.) Also indicated is the 78Kd co-precipitated protein that we have identified as the BiP chaperonin.

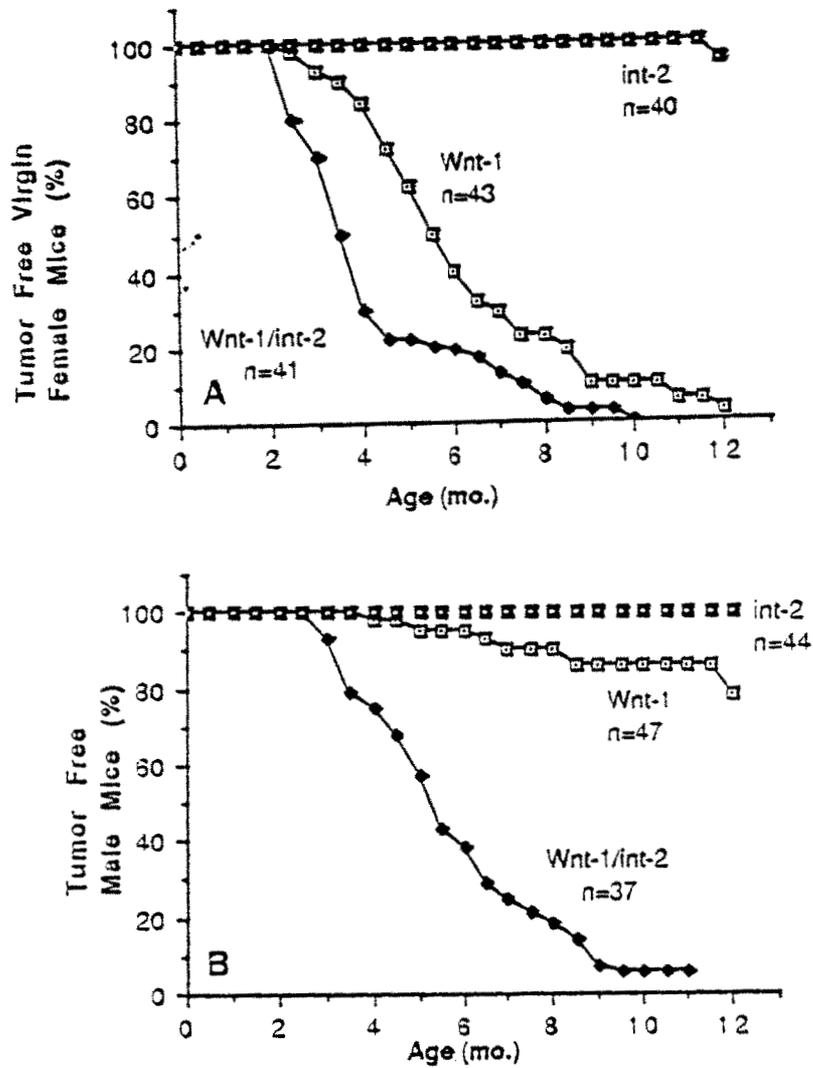


Figure 5. Oncogenic cooperation between Wnt-1 and int-2 transgenes. Sibling transgenic mice bearing an MMTV-Wnt-1 transgene, an MMTV-int-2 transgene, or both transgenes were monitored for one year for the appearance of palpable mammary tumors, which were then confirmed histologically as adenocarcinomas. As indicated in the graphs, the presence of the int-2 transgene shortens the latency for tumorigenesis in Wnt-1 transgenic animals by about two months in virgin females; an even more dramatic effect is observed in males.

MMTV Infection Accelerates Mammary Tumorigenesis In Female Wnt-1 Transgenic Mice

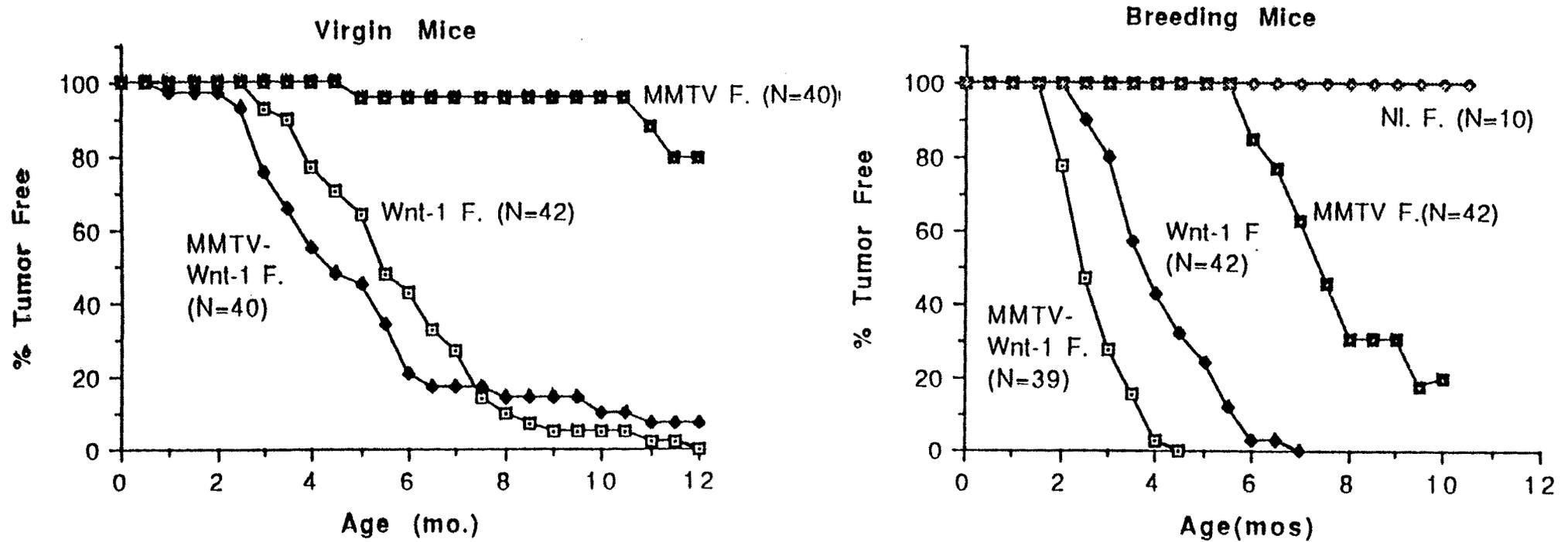


Figure 6. Wnt-1 transgenic female mice were infected with MMTV as neonates, by intraperitoneal injection with a rat cell line expressing an infectious, tumorigenic clone of MMTV DNA. The infected animals and uninfected control animals were observed for the appearance of mammary carcinomas for one year. The graphs show that virus infection accelerates the tumorigenesis by one to two months; in addition, the number of tumors per animal was generally greater (up to ten per animal) in the infected cohort. Tumors from the infected animals contained new MMTV proviruses, often in known proto-oncogenic loci as discussed in the text.