Haematology and Blood Transfusion Vol. 28
Modern Trends in Human Leukemia V
Edited by Neth, Gallo, Greaves, Moore, Winkler

Cellular *onc* Genes: Their Role as Progenitors of Viral *onc* Genes and Their Expression in Human Cells

F. Wong-Staal, S. Josephs, R. Dalla Favera, E. Westin, E. Gelmann, G. Franchini, and R. C. Gallo

Type C retroviruses are associated with naturally occurring leukemia-lymphomas in many animal species, including man (see Gallo et al., this volume), and they are also the first tangible tools for approaches to our understanding of the molecular mechanisms of cellular transformation (see Duesberg et al. and Vande Woude et al., this volume). While most retroviruses isolated in nature are slow acting in disease induction (chronic leukemia viruses), a subclass of viruses, including the sarcoma viruses and acute leukemia viruses, cause disease rapidly in vivo and transform appropriate target cells efficiently in vitro. These properties are conferred on the viruses by a viral transforming (*v-onc*) gene. There are now at least 17 different *v-onc* genes identified in retroviruses isolated from avian, rodent, feline, and primate species. All viral *onc* genes are derived from normal cellular genes (*c-onc* genes) of their host of origin. *C-onc* genes share several common features: (1) they are highly conserved among all vertebrates and some are conserved even in nonvertebrate species. For example, sequences related to a few *onc* genes have been identified in *Drosophila* [8], and an enzyme related to $pp60^{src}$ of Rous sarcoma virus has been detected in sponge (M. Schartl, personal communication). (2) With few exceptions, the homology between *c-onc* and *v-onc* genes is interrupted by nonhomologous stretches in *c-onc*, tentatively referred to as introns. (3) Most *c-onc* genes have been found to be expressed at least at some stages of normal cell growth, suggesting they are functional genes in normal cellular processes.

The interest of our laboratory in *c-onc* genes is chiefly in their role in growth, differentiation, and neoplastic transformation of human cells. We have molecularly cloned human DNA sequences homologous to a number of viral *onc* genes ([2, 3]; Franchini et al. [5] and unpublished), and in collaboration with others we have studied their patterns of expression [4, 10, 11]. Through the studies on the structures of the human cellular genes and their relationship to the corresponding viral genes we have gained some insight on the possible mechanism of generation of the acutely transforming viruses. In the first part of this paper we will selectively present a longitudinal study comparing the viral and cellular counterparts of one particular *onc* gene, namely *sis*, the *onc* gene of simian sarcoma virus, to illustrate this point. Studies on the expression of *onc* genes in human cells have also shed some light on the cellular role of these genes, and in the second part of this paper we will summarize studies on the expression of six *onc* genes in different human cells.

A. The Transforming Gene of Simian Sarcoma Virus (*sis*)

Simian sarcoma virus, SSV, is the only acutely transforming primate retrovirus characterized to date. It was isolated in association with a helper virus, SSAV, from the fibrosarcoma of a woolly monkey [12]. Closed circular SSV and SSAV viral DNA intermediates were cleaved with a one-cut enzyme and ligated to phage vector arms to

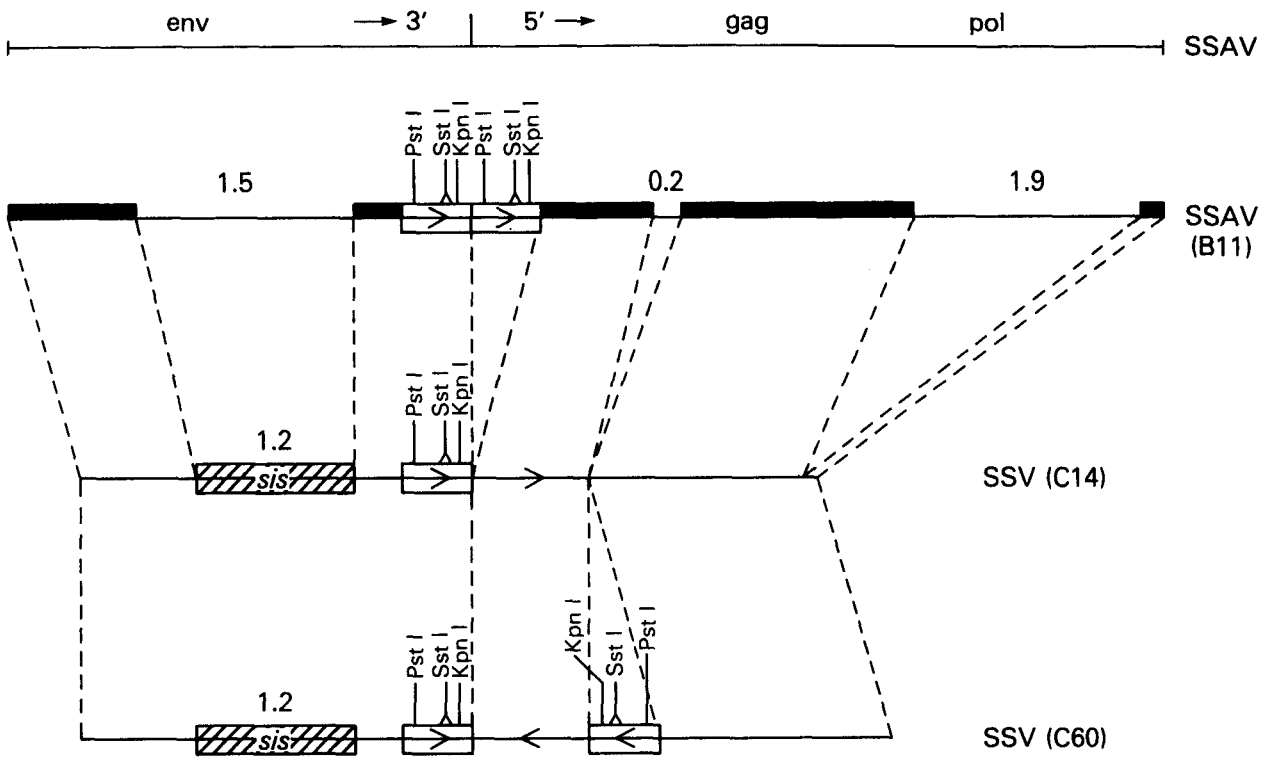
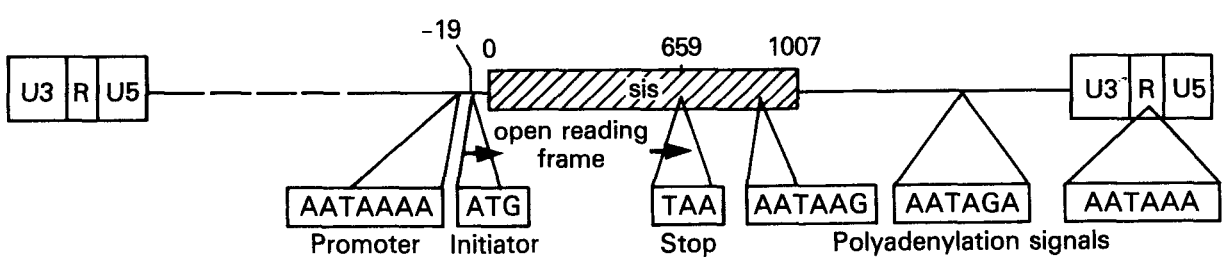


Fig. 1. Genetic structures of SSAV and two molecular clones of SSV

generate clones of complete, permuted SSV, and SSAV genomes, which were compared by restriction enzyme and heteroduplex mapping [6]. The SSAV DNA genome is a 9.0-kb molecule with two long terminal repeat units (LTR). Two SSV clones from viral DNA intermediates were extensively analyzed (Fig. 1). When compared to SSAV the two clones share three regions of deletion and one substitution: a 0.2-kb deletion near the beginning of the *gag* gene, a 1.9-kb deletion probably comprising most of the *pol* gene, and a 1.5-kb deletion in the *env* gene where a substitution of 1.0-kb of SSV-specific (*v-sis*) sequences is found (Fig. 1). Comparison of *sis* to other viral *onc* genes showed no detectable homology [13]. Hybridization of *v-sis* sequences to DNA from different primate and nonprimate species showed that *v-sis* has highest homology to woolly monkey DNA [13]. Since SSV was originally isolated from a pet woolly monkey which cohabited with a pet gibbon ape, and since SSAV is highly homologous to isolates of gibbon ape leukemia virus (GaLV), we concluded that SSV arose from a recombination between a woolly monkey cellular gene and a retrovirus of the GaLV group transmitted from the gibbon to the woolly monkey.

We have determined the nucleotide sequence of the entire *v-sis* region and adjacent SSAV-derived sequences (Josephs et



sis protein: 226 residues
27,000 daltons
hydrophilic

Fig. 2. Fine structure of the SSV genome

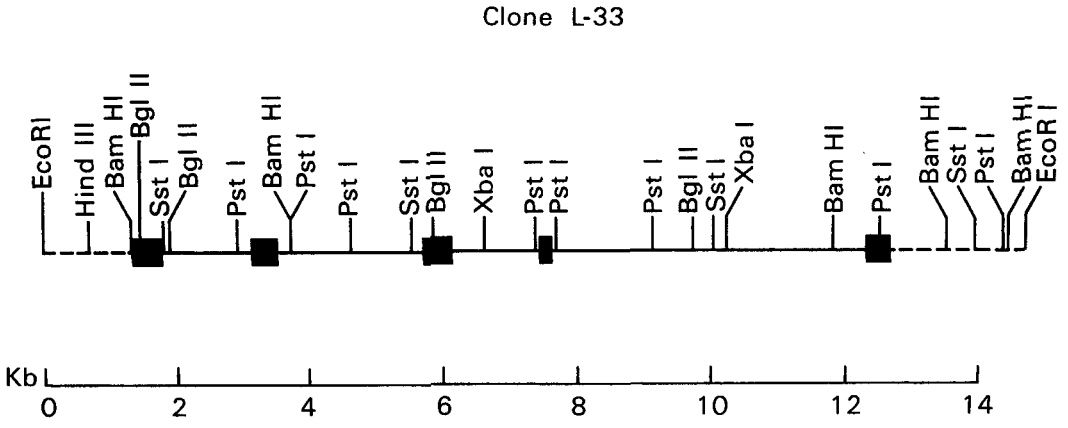


Fig. 3. Organization of human *c-sis* locus. Regions of homology with *v-sis* are represented as *black boxes*

al., submitted). The results revealed a single large open-reading frame which initiates 19 bases within the helper-derived sequences and terminates about two-thirds into *v-sis* (Fig. 2). The translated protein product is highly hydrophilic and has a molecular weight of 27,000 daltons. The first six amino acid residues of the *v-sis* protein are coded by the SSAV-derived sequences. The recombination of *v-sis* and SSAV occurs within a triplet at the 5′ end. Potential promoter sequences AATAAAA are found upstream from the ATG initiator codon and in the U3 region of the LTR. It is not clear which is the functional promoter for the *sis* protein.

B. The Human Cellular Homologue of *sis*

Labeled *v-sis* sequences detects a single locus in human DNA [14], and an RNA transcript of 4.2-kb in some human tumor tissues [4, 10]. A clone of the human *c-sis* gene was isolated from a recombinant phage library [2]. The DNA insert of this clone (L33) contains all the *v-sis* specific sequences. Two techniques were used to locate the regions of homology: restriction endonuclease mapping and heteroduplex formation between L33 and an SSV clone. Both analyses revealed that the 1.0-kb of *v-sis* homologous sequences in L33 span a

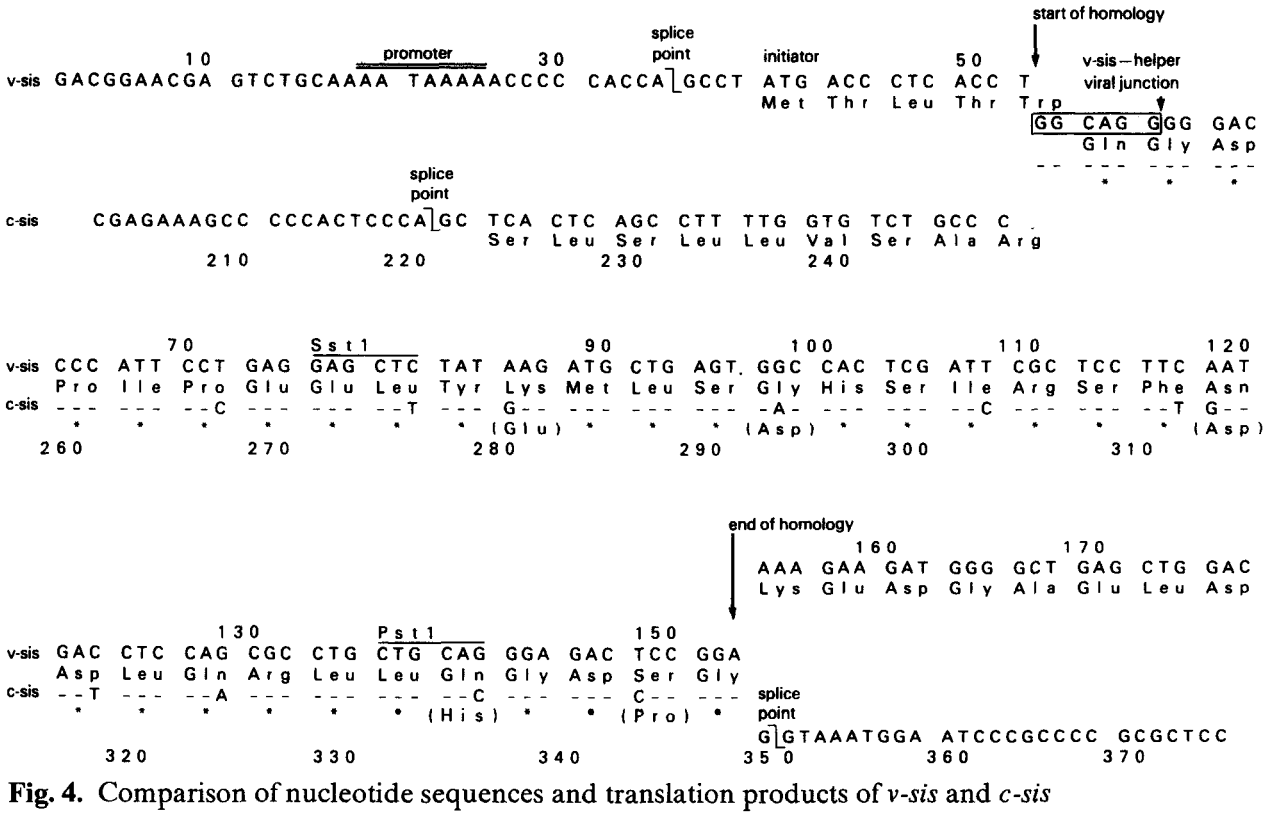


Fig. 4. Comparison of nucleotide sequences and translation products of *v-sis* and *c-sis*

region of 12 kb and are interrupted by at least four nonhomologous regions (Fig. 3). The 5′ region of L33 delimited by the *Eco*RI and *Pst*I sites was subcloned, and a portion of its nucleotide sequence, comprising sequences upstream from and including the entire first homologous region as well as part of the adjacent intron, was determined [3]. When aligned with the *v-sis* sequences, the first *c-sis* homologous region is found to be 1000 bases long also starting within a triplet, and contains six silent base changes and five base changes resulting in altered amino acid residues from *v-sis* (Fig. 4). These changes probably represent divergence of the human and woolly monkey genes rather than that of viral and cellular genes. There is no initiator ATG or promoter sequences within the region of homology or within the 246 bases upstream. However, there is a possible splice site 29 bases upstream of the start of the region of homology. These results taken together suggest that recombination of SSAV and *c-sis* occurred within an internal exon of the functional *c-sis* gene and that *c-sis* consists of additional exon(s) at the 5′ region. These results may also explain why the RNA transcript detected in human cells is much larger than the 1.0-kb *v-sis* gene. An additional interesting finding is the detection of a six-basepair sequence (boxed in Fig. 4) that is found in both *c-sis* and SSAV right at the junction of recombination. This sequence homology may be essential for the recombination event as described in the following model.

C. A Model for Generation of Transforming Retroviruses

Based on the nucleotide sequence data, we present a model for generation of SSV from recombination of SSAV and *c-sis* sequences (Fig. 5). This model may also apply to formation of other transforming retroviruses. First a nondefective (helper) virus infects a cell and integrates at a site on the chromosome. This initial event may be random. If by chance a short sequence in this provirus is homologous to a cell sequence downstream, recombination between these sequences may delete out the

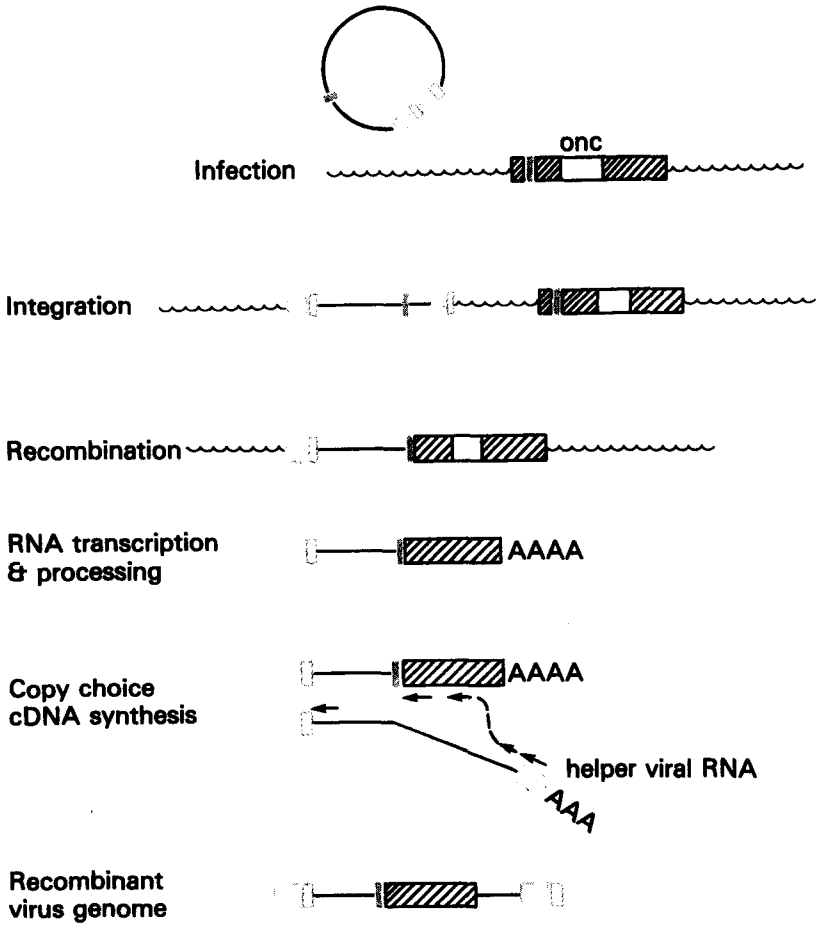


Fig. 5. Model for generation of acutely transforming retroviruses. (a) Infection: a circular helper virus genome shown here with two LTRs in the vicinity of a cellular *onc* gene. For simplicity, the cellular gene is shown to have two exons (*hatched area*) and one intron. The *small solid bar* represents homologous sequences in the helper virus and *c-onc*. *Wavy lines* are cellular sequences flanking *c-onc*. (b) The linearized helper provirus integrates upstream from *c-onc*. (c) Recombination between the homologous sequences deleting out the intermediary stretch (including the transcriptional stop signals of the virus genome). (d) Cotranscription of viral and cellular sequences and removal of the intron of *c-onc*. (e) Copy choice cDNA synthesis. (f) Generation of recombinant virus genome containing viral sequences flanking a cellular insert

intermediary sequences. The site of recombination may occur within a gene, an exon, or even a codon, as exemplified by *sis*. Transcription of the newly juxtaposed sequences results in a chimeric RNA molecule with viral and cellular sequences which is then processed to remove the introns of that part of the cellular gene. Simultaneous presence of complete viral and chimeric RNA probably held together by a dimer linkage at the 5′ end makes possible copy choice DNA synthesis, resulting in a recombinant DNA with viral sequences at both ends and a cellular sequence inserted. This DNA can then integrate into host chromosomes, and RNA transcripts from it can be efficiently packaged into virions. Some aspects of the latter steps of this model have been proposed by others [1, 7].

D. Expression of *onc* Gene Homologues in Human Cells

The identification of human *onc* gene homologues obviously raised the question of whether these are functional genes and whether they play a role in normal or neoplastic cell growth. We have examined a wide variety of human cells for the expression of *onc* gene homologues of Abelson murine leukemia virus (*abl*), avian myelocytomatosis virus (*myc*), avian myeloblastosis virus (*myb*), Harvey murine sarcoma virus (Ha-*ras*), simian sarcoma virus (*sis*), and feline sarcoma virus (*fes*). Molecularly cloned probes containing the *v-onc* sequences are labeled and hybridized to poly (A) containing RNA by the gel blotting technique described by Thomas [9].

I. Hematopoietic Cells

Our source of human hematopoietic cells includes fresh uncultered cells from normal individuals and leukemic patients as well as various cell lines of defined marker characteristics representing cells of myeloid, lymphoid, and erythroid lineages arrested at different stages of cell differentiation. The results [10, 11] are summarized in Table 1. Several points can be generalized from these studies: (1) There is no obvious difference between fresh and cultured cells of the same lineage, e.g., fresh AML cells and T-ALL cells behave as myeloblast cell lines (KG-1) and T-lymphoblast lines (CCRF-CEM, Molt-4) respectively. This argues for the validity of studying cell lines in these experiments. (2) The size(s) of mRNA for a given *onc* gene is the same in all human cells, and very similar (though not necessarily identical) to that of mRNA in other vertebrate cells. (3) The patterns of expression of different *onc* genes vary. Thus, each *c-onc* gene should be considered separate from the others. Specifically, the *abl* and Ha-*ras* genes are detectably expressed (one to five copies per cell) in all hematopoietic cells examined as multiple mRNA species. These genes are probably important for some basic cellular functions. The *myc* and *myb* genes code for single size transcripts of 2.7-kb and 4.5-kb respectively. However, the expression of *myb* is more restricted than *myc*. The *myc* gene is transcribed in all hematopoietic cells examined, including normal peripheral blood lymphocytes prior to or after stimulation with PHA. The only exception is terminally differentiated HL60 cells where *myc* transcription is turned off. The *myb* gene is expressed in the early precursor cells of lymphoid, myeloid, and erythroid lineages, but there is little or no expression relatively early in B-lymphoid cell differentiation, and late in T-cell or myeloid cell differentiation. Like *myc, myb* is transcribed in undifferentiated HL60 cells but not in HL60 cells induced to differentiate with either DMSO or retinoic acid. The *sis* and *fes* genes are not commonly transcribed in hematopoietic cells. There are two instances where enhanced transcription is observed: *myc* transcription in the promyelocytic cell line HL60 and *myb* transcription in leukemic T-lymphoblasts. However, it is premature to conclude that these enhanced expressions are disease related. Normal cells of equivalent lineages and stages of differentiation and/or more samples of similar disease types will be needed to clarify the correlation further.

II. Solid Tumors

A parallel study of human solid tumors and normal fibroblast cell lines has been carried

Table 1. Expression of *onc* genes in human hematopoietic cells [10, 11]

Cell type	Source	Stage of differentiation	mRNA species detected with					
			v-abl (kb) 7.2, 6.4 3.8, 2.0	*v-myc* (kb) 2.7	*v-myb* (kb) 4.5	*v*-H-*ras* (kb) 6.5 5.8, 1.5	*v-sis* (kb) 4.3	*v-fes* ?
Myeloid	KG-1, fresh AML	Myeloblast	+ +	+ +	+ +	+	–	–
	HL60	Promyelocyte	+ +	+ + + +	+ +	+	–	–
	HL60 + DMSO, RA	Granulocyte	+ +	±	–	+	–	–
Erythroid	K562	(Immature erythroid precursor)	+ +	+ +	+ +	+	–	–
Lymphoid								
T cells:	CEM, Molt-4 fresh ALL	Immature T cell	+ +	+ +	+ + +	+	–	–
	HUT78	Mature T cell	+ +	+ +	–	+	–	–
	HUT102	Mature T cell	+ +	+ +	–	+	+	–
B cells:	Raji, Daudi	Burkitt lymphoma	+ +	+ +	–	+	–	–
	NC37, CRB	EBV transformed	+ +	+ +	–	+	–	–
Normal peripheral lymphocytes			NT	+ +	–	NT	NT	NT
Normal peripheral lymphocytes + PHA			NT	+ +	–	NT	NT	NT

NT, not tested

	myb	*myc*	*sis*[a]	*abl*	Ha-*ras*
Sarcomas					
Rhabdomyo-	–	+	+	+	+
Osteogenic	–	+	+	+	+
Fibro-	–	+	+	+	+
Synovial	–	+	+	+	+
Carcinomas					
Skin	–	+	–	+	+
Lung	–	+	–	+	+
GI	–	+	–	+	+
Renal	–	+	–	+	+
Bladder	–	+	–	+	+
Ovarian	–	+	–	+	+
Others					
Melanoma	–	+	–	+	+
Glioblastoma	–	+	+	+	+
Teratoma	–	+	–	+	+
Normal fibroblast	–	+	–	+	+

[a] *sis* is expressed in a high percentage, but not all, of the tumors indicated

Table 2. Expression of *onc* gene homologues in human solid tumor cell lines [4]

out in collaboration with others and in particular Aaronson and colleagues [4]. The results as summarized in Table 2 reinforced the universality of expression of *abl* and Ha-*ras* genes. The *myb* gene is not expressed in these cells and may be specifically involved in hematopoietic cell differentiation. Of interest is the finding that *c-sis* is frequently expressed at moderate to high levels in sarcomas and glioblastomas but not in any melanomas, carcinomas, or normal fibroblast cell lines. Therefore, expression of this gene shows the greatest correlation with specific types of neoplasias.

E. Summary and Conclusion

Viral transforming (*v-onc*) genes are derived from cellular (*c-onc*) genes that are highly conserved among vertebrates. Comparative studies of *v-onc* and *c-onc* genes have shed some light on the mechanism leading to formation of the transforming viruses. A specific example of the *sis* gene is presented here for illustration. Studies on the expression of six *c-onc* genes in human cells revealed at least three categories of *onc* genes: (a) those that are universally expressed and probably are important in basic cellular functions, (b) those that are not detectably expressed in the cells examined and may have very transient expression in development, and (c) those that are only expressed in specific cell types and may be important in tissue differentiation. Our studies do not show conclusively a role of these *onc* genes in human neoplasias.

Acknowledgments

We thank Anna Mazzuca for expert secretarial assistance.

References

1. Coffin JM (1979) Structure, replication and recombination of retrovirus genomes: some unifying hypotheses. J Gen Virol 42:1–26
2. Dalla Favera R, Gelmann EP, Gallo RC, Wong-Staal F (1981) A human *onc* gene homologous to the transforming gene (*v-sis*) of simian sarcoma virus. Nature 292:31–35
3. Dalla Favera R, Gelmann EP, Martinotti S, Franchini G, Papas T, Gallo RC, Wong-Staal F (1982) Cloning and characterization

of different human sequences related to the *onc* gene (v-*myc*) of avian myelocytomatosis virus (MC29). Proc Natl Acad Sci USA 79:6497–6501
4. Eva A, Robbins KC, Andersen PR, Srinvasan A, Tronick SR, Reddy EP, Ellmore NW, Galen AT, Lautenberger JA, Papas TS, Westin EH, Wong-Staal F, Gallo RC, Aaronson SA (1981) Cellular genes analogous to retroviral *onc* genes are transcribed in human tumor cells. Nature 95:116–119
5. Franchini G, Gelmann EP, Dalla-Favera R, Gallo RC, Wong-Staal F (1982) A human gene (*c-fes*) related to the *onc* sequences of Snyder-Theilen feline sarcoma virus. Mol Cell Biol 2:1014–1019
6. Gelmann EP, Wong-Staal F, Kramer R, Gallo RC (1981) Molecular cloning and comparative analyses of the genomes of simian sarcoma virus and its associated helper virus. Proc Natl Acad Sci USA 78:3373–3377
7. Goldfarb MP, Weinberg RA (1981) Generation of novel, biologically active Harvey sarcoma viruses via apparent illegitimate recombination. J Virol 38:136–150
8. Shilo BZ, Weinberg RA (1981) DNA sequences homologous to vertebrate oncogenes are conserved in *Drosophila melanogaster*. Proc Natl Acad Sci USA 78:6789–6792
9. Thomas PS (1980) Hybridization of denatured RNA and small DNA fragments transformed to nitrocellulose. Proc Natl Acad Sci USA 77:5201–5205
10. Westin EH, Wong-Staal F, Gelmann EP, Dalla-Favera R, Papas TS, Lautenberger JA, Eva A, Reddy P, Tronick SR, Aaronson SA, Gallo RC (1982a) Expression of cellular homologs of retroviral *onc* genes in human hematopoietic cells. Proc Natl Acad Sci USA 79:2490–2494
11. Westin EH, Gallo RC, Arya SK, Eva A, Souza LM, Baluda MA, Aaronson SA, Wong-Staal F (1982b) Differential expression of the *amv* gene in human hematopoietic cells. Proc Natl Acad Sci USA 79:2194–2198
12. Wolfe L, Deinhardt F, Theilen G, Kawakami T, Bustad L (1971) Induction of tumors in marmoset monkeys by simian sarcoma virus type 1 (*Lagotherix*): A preliminary report. J Natl Cancer Inst 47:1115–1120
13. Wong-Staal F, Dalla-Favera R, Gelmann E, Manzari V, Szala S, Josephs S, Gallo RC (1981a) The transforming gene of simian sarcoma virus (*sis*): A new *onc* gene of primate origin. Nature 294:273–275
14. Wong-Staal F, Dalla-Favera R, Franchini G, Gelmann EP, Gallo RC (1981b) Three distinct genes in human DNA related to the transforming genes of mammalian sarcoma retroviruses. Science 213:226–228