ONCOGENES AND HUMAN CANCER

(Oncogenen en Humane Kanker)

PROEFSCHRIFT

ter verkrijging van de graad van
doctor in de geneeskunde
aan de Erasmus Universiteit te Rotterdam
op gezag van de rector magnificus
Prof. Dr. M.W. van Hof
en volgens besluit van het college van dekanen.
De openbare verdediging zal plaatsvinden op
woensdag 30 mei 1984 te 14.00 uur.

door

Eleonora Classina Petronella Heisterkamp

geboren te Diemen

Promotor: Prof. Dr. D. Bootsma

Gedrukt bij Offsetdrukkerij Kanters B.V., Alblasserdam.

Dit proefschrift werd bewerkt in de Laboratory of Viral Carcinogenesis, FCRF, NIH, Frederick, MD., USA.

ONCOGENES AND HUMAN CANCER

(Oncogenen en Humane Kanker)

PROEFSCHRIFT

ter verkrijging van de graad van
doctor in de geneeskunde
aan de Erasmus Universiteit te Rotterdam
op gezag van de rector magnificus
Prof. Dr. M.W. van Hof
en volgens besluit van het college van dekanen.
De openbare verdediging zal plaatsvinden op
woensdag 30 mei 1984 te 15.15 uur.

door

Johannes Hendrikus Cornelis Groffen

geboren te Vlaardingen

Promotor: Prof. Dr. D. Bootsma

Dit proefschrift werd bewerkt in de Laboratory of Viral Carcinogenesis, FCRF, NIH, Frederick, MD., USA.

CONTENTS

Paper I: Isolation of v-fms and its Cellular Homolog
N. Heisterkamp, J. Groffen, J.R. Stephenson
Virology 126 (1983):248-258

Paper II: Chromosomal Localization of the Human c-fms Oncogene
J. Groffen, N. Heisterkamp, N. Spurr, S. Dana, J. J. Wasmuth, J.R. Stephenson.
Nucleic Acids Research 11 (1983): 6331-6341

Paper III: Isolation of Human Oncogene Sequences (v-fes homolog) from a Cosmid Library
J. Groffen, N. Heisterkamp, F. Grosveld, W. van de Ven. J.R. Stephenson.
Science 216 (1982): 1136-1138

Paper IV: Transforming Genes of Avian (v-fps) and Mammalian v-fes) Retroviruses Correspond to a Common Cellular Locus
J. Groffen, N. Heisterkamp, M. Shibuya, H. Hanafusa, J.R. Stephenson.
Virology 125 (1983):480-486

Paper V: Chromosomal Localization of Human Cellular Homologues of two Viral Oncogenes
N. Heisterkamp, J. Groffen, J.R. Stephenson, N.K. Spurr, P.N. Goodfellow, E. Solomon, B. Carritt, W.F. Bodmer.
Nature 299 (1982): 747-749

Paper VI: Genetic Analysis of the 15:17 Chromosome Translocation Associated with Acute Promyelocytic Leukemia
D. Sheer, L.R. Hiorns, K.F. Stanley, P.N. Goodfellow, D.M. Swallow, S. Povey, N. Heisterkamp, J. Groffen, J.R. Stephenson, E. Solomon.
Proc. Natl. Acad. Sci. USA 80 (1983): 5007-5011

Paper VII: The Human v-abl Cellular Homolog
N. Heisterkamp, J. Groffen, J.R. Stephenson.
J. Mol. Appl. Genetics 2 (1983): 57-68

Paper VIII: Homology between Phosphotyrosine Acceptor Site of Human c-abl and Viral Oncogene Products
J. Groffen, N. Heisterkamp, F.H. Reynolds, Jr., J.R. Stephenson.
Nature 304 (1983): 167-169

Paper IX: A Cellular Oncogene is Translocated to the Philadelphia Chromosome in Chronic Myelocytic Leukemia
A. de Klein, A. Geurts van Kessel, G. Grosveld, C.R. Bartram, A. Hagemeijer, D. Bootsma, N.K. Spurr, N. Heisterkamp, J. Groffen, J.R. Stephenson.
Nature 300 (1982): 765-767

Paper X: C-sis is Translocated from Chromosome 22 to Chromosome 9 in Chronic Myelocytic Leukemia
J. Groffen, N. Heisterkamp, J.R. Stephenson, A. Geurts van Kessel, A. de Klein, G. Grosveld, D. Bootsma.
J. Exper. Med 158 (1983): 9-15

Paper XI: Localization of the c-abl Oncogene Adjacent to a Translocation Breakpoint in Chronic Myelocytic Leukemia.
N. Heisterkamp, J.R. Stephenson, J. Groffen, P.F. Hansen, A. de Klein, C.R. Bartram, G. Grosveld.
Nature 306 (1983): 239-242.

Paper XII: Philadelphia Chromosomal Breakpoints are clustered within a Limited Region, bcr, on Chromosome 22.
J. Groffen, J.R. Stephenson, N. Heisterkamp, A. de Klein, C.R. Bartram, G. Grosveld.
Cell 36 (1984): 93-99

PREFACE - VOORWOORD

Dit proefschrift is tot stand gekomen op een voor Nederland ongebruikelijke wijze: het wetenschappelijke werk dat ervoor door ons is gedaan is uitgevoerd in Amerika, in het Laboratory of Viral Carcinogenesis van de National Institutes of Health te Frederick, M.D. Dat wij in de gelegenheid zijn gesteld om daar een aantal jaren onderzoek te doen is, als we teruggaan naar onze studententijd, indirect te danken aan Prof. Dr. M. Gruber: hij was ten zeerste bereid om mij, John (die niet eens een doctoraal onderzoek in zijn afdeling had gedaan) te helpen met mijn wens om "te leren kloneren". Het resultaat was, dat ik een periode in Londen kon doorbrengen in de Laboratory of Gene Structure and Expression te Mill Hill, waar ik een enorme hoeveelheid kennis en ervaring heb opgedaan. Een interesse in het kankeronderzoek, de mogelijkheid om de opgedane kloneringservaring toe te passen en het feit, dat wij onderzoeksplaatsen niet al te ver van elkaar zochten leidde ertoe dat wij uiteindelijk in de NIH zijn terechtgekomen. Hierbij zijn wij veel dank verschuldigd aan Dr. W.J.M. van de Ven, Elly van de Ven en Dr. H.P.J. Bloemers.

Het contact met de mensen in Mill Hill is nog steeds aanwezig; hoewel vele van de oude groep naar elders zijn verhuisd geldt het gezegde "uit het oog, uit het hart" bepaald niet. Dick Flavell en Frank Grosveld willen we in de eerste plaats bedanken voor alles wat ze in het verleden en heden voor ons hebben gedaan. Zonder hun bereidheid, om een student uit Nederland een kans te geven was dit alles niet mogelijk geweest.Ook een aantal andere mensen van de Mill Hill groep, ondanks het feit dat zij niet direct betrokken zijn geweest bij dit proefschrift, hebben toch in zekere zin een invloed gehad. Er moet dan speciaal worden gedacht aan Cora, Dimitris, Torben, Titia, Chris, Andy, Elizabeth, Stephanie, Rene, Henrik, Lex en Ernie. Ook Gerard Grosveld en Annelies de Klein hebben een grote rol gehad in het tot stand komen van dit proefschrift; wie had ooit

gedacht dat het oversturen van een "proobje" zulke spectaculaire resultaten zou opleveren! Wij willen jullie bedanken voor jullie enthousiasme en voor de hele plezierige samenwerking. Ook de andere mensen van de vakgroep Celbiologie en Genetica in Rotterdam zijn wij zeer erkentelijk.

Tot onze promotor, Prof. Dr. Dick Bootsma willen we een speciaal woord van dank richten: het is al een opgave, om promotor te zijn voor promovendi binnen je vakgroep, laat staan als de promovendi niet tot je vakgroep behoren en hun onderzoek verrichten op een heel ander continent. Ondanks deze belemmering hebben we het contact als bijzonder prettig en probleemloos ervaren; wij willen je nogmaals bedanken voor alles, wat je voor ons hebt gedaan en het feit, dat je bereid bent geweest als promotor op te treden.

We also wish to thank all our colleagues from the Laboratory of Viral Carcinogenesis; first of all, Dr. John Stephenson, who initially did not believe in the great virtues of molecular cloning. However, we presume he has changed his mind since then, if we interpret his threat that "one day he will run a DNA sequencing gel" correctly. We would like to thank Dr. Fred Reynolds, Dr. Fulvia Veronese and their technicians Chris, Dick and Bob, the latter who did some of the photography work for our manuscripts. A special word of thanks goes to Siobahn, who did all the tissue culture work and kept John's weird cell lines alive.

Last but not least we want to thank Gail Blennerhasset and Pamela Hansen, our technicians. Although DNA work is certainly not among the easiest, you have both done your very best and thanks to you we have been able to set up a well-running DNA laboratory. Dear Pam and Gail, we have very much appreciated your technical assistance, your patience and you involvement in the work.

Finally, in America, we thank all the secretaries who have typed and

retyped different parts of the manuscript even though they were busy: Mrs. Barbara Eader and Mrs. Beverly Bales in Frederick, Mrs. Ann Newson, Mrs. Peggy Brandeburg and Ms. Marianne Lewis in Mineola.

Tot slot willen wij nog een speciaal woord van dank aan onze ouders richten, hoewel zij zich misschien hieronder een beetje ongemakkelijk voelen. Dankzij de opvoeding die ik, Nora, heb gekregen heb ik ten volle gebruik kunnen maken van het beste wat het Nederlandse onderwijssysteem te bieden heeft; mijn ouders hebben de term "emancipatie" in mijn opvoeding in de practijk toegepast lang voordat de term mode werd. Ook ik, John, ben ervan overtuigd dat mijn ouders heel belangrijk zijn geweest voor mijn geestelijke vorming en moet daarbij speciaal denken aan de vele weekeinden die in de "natuur" werden doorgebracht.

Mineola, maart 1984 Nora Heisterkamp
John Groffen

I. INTRODUCTION

1.1 General Introduction

The first demonstrations that cancer could have an infectious nature was by Ellerman and Bang (1), who showed that leukemia in chickens was transmissible with cell-free extracts and by Rous (2), who found in a similar fashion that naturally occurring chicken sarcomas were transmissible. Although they were able to show that these cell-free extracts contained a transmissible agent, the idea that this induced cancer was received by the scientific world at that time with great skepticism. The interest in oncogenic viruses was strongly enhanced in the early 60's by the isolation of mammalian tumor viruses and the general acceptance that at least some of these viruses were tumorigenic. The discovery of the reverse transcriptase enzyme in RNA tumor viruses (3,4), gave a logical explanation for how these viruses became integrated in the chromosomes of eukaryotic cells.

Taxonomically, oncogenic viruses are members of diverse families. DNA viruses (herpes-, adeno- and papovaviruses) as well as many members of the retrovirus family (containing RNA such as the type C RNA viruses) are capable of inducing tumors. For the retroviruses two different routes to become transforming (oncogenic) have become clear. The majority of these viruses (the acute type C RNA transforming viruses) "acquire" certain genetic sequences (oncogenes) from their host, which are necessary to initiate and maintain the malignant transformation of the cell by the virus. Other retroviruses integrate their genome nearby oncogenic sequences in the chromosome of their host. Independent of the exact mechanism, these viruses share the capability of inducing tumorigenesis by triggering the transcription of certain sequences, and it is the proteins encoded by these sequences which are necessary to maintain the neoplastic phenotype of the infected cell.

The accumulating number of independent isolates of tumorigenic retroviruses induced in the mid-70's a worldwide search for these viruses in humans. Only very recently the isolation of a human tumor T-cell leukemia retrovirus (HTLV) was reported (5,6).

Another approach was initiated in the beginning of the 80's, with the finding that the acquired sequences of retroviruses are strongly conserved among species. In general, the cellular homologs of these sequences were easily detectable and could be studied in more detail by molecular cloning, using the oncogenic acquired sequences of retroviruses as probes. This approach seems to be very fruitful and will be discussed in more detail below.

Although the oncogenic potential of the acquired sequences in a number of these viruses in vertebrates is well estabished, the involvement of their human cellular homologs in human tumorigenesis has been and will be a rich source for discussion. However, at the moment they provide us the best available model for the induction of human cancer at the molecular level.

1.2 Scope of the Thesis

The detailed structure of type C RNA viruses became clear once they had been molecularly cloned: as discussed before, most acute type C RNA transforming viruses had acquired DNA sequences from their host (e.g. chicken, cat, mouse, rat) which were responsible for the tumorigenic properties of the virus. A new approach, in the research of human cancer as a molecular biological discipline, became evident; by using the acquired cellular sequences of, for example, a cat acute type C RNA transforming virus, as a molecular probe, one could investigate the possible existence of homologous sequences in man. If the evolutionary divergence between these feline sequences and the homologous sequences in human were sufficiently low, the human

cellular homologs could be molecularly cloned and their possible involvement in human cancer could be studied in a detailed way. We have chosen that approach in the research described in this thesis: initially, three distinct acute type C RNA transforming viruses (see Introduction 1.3 for a more detailed description of these viruses) were chosen as a source for molecular probes. Two of these, the Snyder-Theilen Feline Sarcoma virus (ST-FeSV) and the Sarma-McDonough Feline Sarcoma virus (SM-FeSV) were molecularly cloned from transformed cell lines containing a complete integrated provirus. The third, Abelson Murine Leukemia virus (Ab-MuLV) was obtained as a gift from P. Reddy. Each of these viruses contains different acquired cellular sequences, v-fes in ST-FeSV, v-fms in SM-FeSV and v-abl in Ab-MuLV. Molecular probes specific for the acquired cellular sequences of each hybridized strongly to sequences in human DNA, indicating that these sequences were strongly conserved during evolution. The human cellular homologs of v-fes, v-fms and v-abl were subsequently molecularly cloned (paper I, III and VII) and research on the potential involvement of these human oncogenes in cancer was initiated. As a preliminary characterization we made detailed restriction maps of these human oncogenes and the position of sequences homologous to the viral acquired cellular sequences were determined; extensive homology between the v-abl tyrosine phosphorylation acceptor region and a corresponding region in human c-abl was established more precisely by nucleic acid sequencing (paper VIII). Additionally, human c-fes was examined for homology with the acquired cellular sequences (v-fps) of Fujinami Sarcoma virus (FSV); earlier reports (7) had indicated, that v-fes (from cat) and v-fps (from chicken) were related. Our results indicated that one locus exists in man having homology both to v-fes and v-fps (paper IV).

One straightforward method of examining the transforming potential of the human cellular homologs is, to introduce these sequences

into rodent cells in culture and to investigate, if they can transform the cells to a malignant phenotype (see Introduction 1.4) However, the c-fes, c-fms or c-abl human oncogenes were incapable of transforming cells in culture. This is in concordance with the results of others obtained with other human oncogenes: generally, human oncogenes do not have transforming ability unless they are manipulated extensively in vitro before introduction into cells. We have chosen not to pursue the effort to attempt to modify the cloned oncogenes in vitro in such a way that they might become transforming: one major obstacle in such an effort would have been the large size of the genes encoding human c-abl and c-fms. Rather, we have examined the possible involvement of human c-fes, c-fms and c-abl in existing human malignancies. As a first step, the chromosomal localization of each of these human oncogenes was determined. Such information might be of value in view of the accumulating evidence, that specific chromosomal aberrations are associated with specific types of neoplasia (See Introduction 1.5). The localization of human c-fms to chromosome 5 (paper II) and human c-fes to chromosome 15 (paper V) has to date been of moderate value. Aberrations on chromosome 5 have not been intimately associated with one type of neoplasia. Although c-fes is translocated from chromosome 15 to chromosome 17 in patients with acute promyelocytic leukemia (APL) in which a t(15;17) is characteristic (paper VI), we have not found a chromosomal breakpoint in the proximity of c-fes.

Human c-abl, on chromosome 9 (paper V) has provided us with an excellent tool to study leukemogenesis at the molecular level: in the t(9;22) characteristic of chronic myelogenous leukemia (CML) human c-abl is translocated to chromosome 22, the Philadelphia (Ph') chromosome (paper IX). As another human oncogene, c-sis, is located on chromosome 22, we examined if c-sis was also involved in the Philadelphia translocation. Although we could demonstrate the translocation of c-sis to chromosome 9 (paper X) human c-sis was

subsequently localized far from the Philadelphia translocation breakpoint (8). Therefore, our attention concentrated on the human c-abl oncogene; in one CML patient, we found, after an elaborate search, a chromosomal breakpoint 5' to the human c-abl locus (paper XI). Furthermore, we were able to clone a region from chromosome 22 in which the breakpoints occur in all Ph'-positive CML DNAs examined to date (paper XII).

The three acute type C transforming viruses utilized in our research represent only part of the many type C RNA viruses isolated to date (see table I). We have not concentrated our research on the viral oncogenes but rather, have used them as a tool in characterizing human oncogenes. However, the molecular organization, protein products, relatedness and transforming mechanism of the viral oncogenes has been studied extensively by others. Although this does not directly bear on the question if human oncogenes are involved in human tumorigenesis, studies on viral oncogenes may ultimately explain how cellular sequences can become transforming. Therefore, we have included research regarding type C RNA viruses (Introduction 1.3). Cellular oncogenes in animal models have also yielded much information (Introduction 1.4).

Many human oncogenes have been molecularly cloned and characterized as we have done (Introduction 1.5). We have attempted to provide a limited amount of structural information for each. Of the human oncogenes, c-abl, c-myc and the ras gene family have been most strongly implicated in neoplastic diseases. The putative transforming mechanism of c-abl and c-myc seem to be similar. We have done most of the research on c-abl; the research on c-myc is very elaborate and we have attempted to summarize this information.

In the above described research projects, Drs. N. Heisterkamp has concentrated on the human c-fms oncogene (papers I and II) and

human c-abl (papers V, VII, IX and XI). Drs. J. Groffen has had a major contribution in research concerning human c-fes (papers III, IV and VI), sequencing of human c-abl (paper VIII) and sequences on chromosome 22 (papers X and XII). However, all projects were performed in close collaboration between both authors.

1.3 Type C RNA Viruses

Type C RNA viruses can be divided into two classes on the basis of tumorigenicity. The first class (the chronic type C RNA viruses) cannot transform cells in culture. in vivo, however, these viruses, such as Feline leukemia virus (FeLV) and Moloney murine leukemia virus (MoMuLV), can induce leukemias in infected animals after a relatively long period of latency ranging from 6 to 12 months. The viral genome consists of several genes, which together are necessary for the life cycle of the virus. After infection of the cell, the RNA genome is converted to a double stranded DNA copy by the viral-encoded enzyme reverse transcriptase.

The DNA copy of the viral genome is integrated into the genome of the infected cell; this provirus carries two identical LTR's (long terminal repeats), one at either end of the viral genome. The LTR contains promoter sequences, signals for transcription termination and polyadenylation signals (9). Transcription is initiated from the 5' LTR; RNA is transported to the cytoplasm and translated into the proteins utilized for the formation of a new virus particle. The order of the genes on the linear double stranded DNA copy of the viral RNA is:

5' LTR ---- gag pol env ---- LTR 3'

The gag gene encodes the structural proteins of the viral core. These proteins are synthesized as part of a polyprotein, which is

processed into the mature viral proteins. Pol encodes the viral reverse transcriptase and env the viral envelope glycoproteins.

The second class of type C RNA viruses (the acute type C RNA viruses) is called transforming because of their ability to transform cells in culture. In general, they cause sarcomas and acute leukemias in vivo after a relatively short latency period (a few days or weeks). What has been described for the "chronic" viruses also applies to these viruses; the latter, however, are generally replication deficient and require the presence of a helper virus for the production of infectious viral particles. Acute type C transforming RNA viruses are thought to have originated from chronic type C viruses through a recombination event between the latter viruses and host cellular sequences. As a consequence, part of the structural information of the viral genome has been deleted and replaced by "non-structural" or "acquired" sequences of host cellular origin. These acquired sequences have imparted a new ability on the virus, enabling it to transform fibroblasts in vitro and to have an acute oncogenic potential in vivo (10, 11).

At least 15 different type C transforming viruses have been isolated, from species as diverse as chicken and baboon. Each isolate contains acquired cellular sequences. To distinguish the cellular homologs from their viral counterparts a standard nomenclature (12) is accepted: viral oncogenes are denoted by "v" as in v-src, the cellular progenitors by "c" as in c-src.

The chicken has been a particularly rich source for the isolation of type C transforming viruses (Table I). Rous sarcoma virus (RSV), in contrast with the other virus isolates, is not replication deficient: chicken cellular sequences have been inserted 3' to the structural genes of avian leukosis virus, resulting in the formation of a replication competent transforming virus. The acquired cellular

TABLE I

Gene	Virus	Animal Origin	Protein	Protein Kinase
src	RSV	chicken	p60	PK
	B77	chicken	p60	PK
	rASV	chicken/quail	p60	PK
fps/fes	FSV	chicken	P140	PK
	PRC11	chicken	P105	PK
	URI	chicken	P250	PK
	16L	chicken	P142	PK
	ST-FeSV	cat	P85	PK
	GA-FeSV	cat	P115	PK
yes	Y73	chicken	P90	PK
	ESV	chicken	P80	PK
ros	UR2	chicken	P68	PK
myb	AMV	chicken	-	-
	E26	chicken	-	-
erb	AEV	chicken	P75+40	
myc	MC29	chicken	P110	-
	CM11	chicken	P90	-
	MH2	chicken	P100	-
	OK10	chicken	P200	-
rel	AEV	turkey	-	-
mos	Mo-MuSV	mouse	-	-
	GZ-MuSV	mouse	-	-
ras-Ki	Ki-MuSV	rat	p21	-
ras-Ha/*bas*	Ha-MuSV	rat	p21	-
	Ra-RaSV	rat	p29	-
	Balb-MuSV	mouse	p21	-
abl	Ab-MuLV	mouse	P120	PK
fms	SM-FeSV	cat	P170	-
sis	SiSV	monkey	-	-
raf	3611-MuSV	mouse	P90	-
fos	FBJ-MuSV	mouse	p55	-

sequences of RSV are called src. Two different strains of avian sarcoma virus, Y73 (13) and Esh (14) have independently recombined with the same cellular sequence, c-yes (15). Fujinami, PRCII and URl avian sarcoma viruses all contain a strongly related oncogene, v-fps (7, 16). A new avian sarcoma virus isolate, UR2 (17), contains a unique oncogene, v-ros (18). Avian erythroblastosis virus contains two "domains" of acquired cellular sequences, erbA and B. Each encodes a protein: erbA encodes a gag fusion protein (see later) of MW 75,000 (19). ErbB, which consists entirely of acquired cellular sequences, encodes a 40,000 MW protein (20) which is thought to be responsible for the transforming activity of the virus (21). The acquired cellular sequences of avian myeloblastosis virus (AMV) have been called both amv and myb (22,12). Finally, MC29 (myelocytomatosis virus) is one of a series of avian type C transforming viruses that have transduced a cellular oncogene called myc from chicken. One species of transforming type C virus has been isolated from turkey; the reticuloendotheliosis virus strain (REV-T) contains the rel oncogene (23).

In mammals, the isolation of type C transforming viruses has been reported from mouse, rat, cat and baboon. In mouse, murine leukemia virus has recombined with different cellular oncogenes, giving rise to Abelson murine leukemia virus (v-abl) (24,25), Moloney murine sarcoma virus (v-mos) (26), 3611-MuSV (v-raf) (27), and Finkel-Biskis-Jinkins murine osteosarcoma virus (v-fos) (28). Through passage of a murine leukemia virus on rat cells Harvey and Kirsten murine sarcoma virus were obtained (29, 30). Both contain an oncogene (v-Ha-ras (31) and v-Ki-ras) flanked at either end by rat "30S" DNA sequences, which are retrovirus-like but without transforming capacity. v-Ha-ras and v-Ki-ras are related but not identical: they contain homologous and unique sequences (32). From cat, two strains of feline sarcoma virus have been independently isolated which both contain portions of the same cellular oncogene

c-fes: Snyder-Theilen FeSV and Gardner FeSV (33, 34, paper IV). Another cat isolate yielded McDonough FeSV which carries a unique oncogene, v-fms (35, paper I). One type C transforming RNA virus has been isolated from primates, the simian sarcoma virus (v-sis oncogene) (36, 37).

From this summary it is clear that a large variety of type C transforming viruses has been isolated. In this context it should be mentioned that in most cases the viruses have been molecularly cloned, allowing a detailed comparison at the nucleic acid level with respect to the relatedness of various isolates; most references in this thesis refer to the molecular cloning of the virus and not the original isolation. From a comparison of the fes oncogene in ST-FeSV and GA-FeSV with the fps oncogene, in Fujinami sarcoma virus, it became clear that fps and fes represent the same oncogene from chicken and from cat (38, 39, paper IV). New viral isolates do not always yield new oncogenes: a type C RNA virus isolated from cat contains a sis-like oncogene (40). The acquired cellular sequences from Balb murine sarcoma virus (bas) seem to be analogous to ras from rats (41). Therefore, the amount of different oncogenes transduced by type C RNA viruses may be quite limited. This may reflect a situation wherein the number of oncogenes in the genome is low; on the other hand, it is possible that certain oncogenes cannot fulfill certain "requirements" necessary to become integrated in the retroviral genome. At the moment, it is not known if this type of requirement exists; however, with the development of retroviral vectors for eukaryotic cells, analogous to the E. coli cloning vectors, this type of information should become available. The isolation of new oncogenes with DNA-mediated transfection experiments (see Introduction 1.5, j) lends support to the hypothesis that there may be still a large repertoire of sequences with oncogenic potential in the genome not yet isolated.

The protein products of the type C transforming RNA viruses are very diverse in size. As mentioned, the cellular sequences have been integrated in phase in the structural genes of the virus. Where this recombination has taken place differs in various viral isolates. Fujinami sarcoma virus has the structure 5'-Δgag-fps-3' (38), indicating that part of the gag gene and the entire pol and env gene have been deleted and replaced by fps sequences. As a consequence, fps is expressed as a polyprotein, with amino acid-terminus encoded structural components covalently linked to the acquired cellular sequence encoded carboxy-terminus component: P140$^{gag-fps}$. Both ST-FeSV and Gardner FeSV have the structure 5'-Δ gag-fes- Δ env-3'. They express fes as part of a gag-fes polyprotein; the difference in molecular weight (P115$^{gag-fes}$ for Gardner FeSV, P85$^{gag-fes}$ for ST-FeSV) is caused by the different size of the gag portion and differences in the acquired sequence component (39). MC29 also encodes a gag-onc fusion protein, of MW 110,000 (P110$^{gag-myc}$); the genomic organization is similar to that of ST and Gardner FeSV: 5'-Δ gag-myc-Δ env-3' (42). Simian sarcoma virus has a variant structure, in which the sis sequences are situated in the env gene (36). As mentioned before, avian erythroblastosis virus encodes two proteins, a gag-erbA fusion protein and a protein completely encoded by the acquired sequence component (20,21). A number of other type C transforming viruses have oncogenes, which are not expressed as part of a polyprotein; Rous sarcoma virus (in which structural genes have not been replaced by an oncogene) encodes a pp60^{src} (43). Both Ha-MuSV and Ki-MuSV encode a p21 protein from the acquired cellular sequences (32).

The protein products of many of these viral oncogene sequences have been identified; in many cases, however, the putative enzymatic activity or function remains obscure. Even though for some an enzymatic activity has been found it is difficult to unravel the pathway by which the viral oncogene mediates its transforming effect.

However, some retroviruses have acquired sequences that encode proteins, which all possess a tyrosine-specific protein kinase activity. Protein kinases are thought to have a regulatory function in the cell through the activation and deactivation of cellular enzymes by phosphorylation. That $pp60^{src}$ (Rous sarcoma virus) possessed a protein kinase activity with specificity for tyrosine residues (44) came somewhat as a surprise; until then it was thought that phosphorylation involved only serine or threonine residues. In normal cells only one of the 3000 phosphorylated residues is a tyrosine; in RSV transformed cells this amount increases five-to tenfold. Soon after this enzymatic activity had been discovered it was found for a number of other viral oncogene products (Table I): v-fes/fps (45,46), v-abl (47,48), v-yes (49) and v-ros (50). This enzymatic activity seems required for the maintenance of the transformed state; temperature sensitive mutants of RSV (51) and FSV (52) are no longer able to transform cells and do not exhibit protein kinase activity at the non-permissive temperature. Transformation deficient mutants of AbMuLV lack protein kinase activity (53, 54). Although phosphotyrosine residues in the protein kinases themselves are thought to be of importance for their enzymatic activity (55), their exact functional significance remains to be determined.

Nucleic acid sequencing data have shown that the different oncogene encoded protein kinases may be related and have evolved from a common ancestor: v-fps and v-src are 40% homologous in the carboxy terminal 280 amino acids (38). Although v-yes and v-src did not exhibit significant homology in nucleic acid hybridization experiments (13), the amino acid sequence predicted on the base of nucleic acid sequencing data revealed that both oncogenes were strikingly similar (56). That v-abl and v-src are related stems from two lines of evidence. Firstly, in Drosophila melanogaster sequences have been cloned that exhibit nucleic acid homology to both v-abl and v-src (which might represent a common ancestor), and

ones that have exclusive homology to src (57, 58). Interestingly, the regions of v-abl and v-src that show cross homology with this putative progenitor sequence (called c-Dash) are those that are essential for the protein kinase activity in v-abl (25) and v-src (59). Secondly, nucleic acid sequence data have shown that v-src, v-fes, v-abl and human c-abl contain a common region that is highly conserved (paper VIII).

Even v-mos appears to have limited homology with v-src at the amino acid level (60). The phosphorylation of tyrosine residues does not seem to be an exclusive property of viral oncogene products. The cell surface receptor for epidermal growth factor (EGF) is phosphorylated in tyrosine after binding EGF, a small polypeptide with growth promoting activity (61, 62). Furthermore, the same reaction is found after the binding of insulin to the insulin receptor (63) and the binding of platelet derived growth factor (PDGF) to its receptor (64). Although it seems plausible that these receptors are themselves protein kinases specific for tyrosine residues, it remains to be established if this is the case (65) or that such protein kinases are intimately associated with them. Interestingly, cells transformed by either AbMuLV (v-abl) or ST-FeSV (v-fes) and some human tumor derived cell lines show an overall increase in phosphotyrosine level and a reduced binding of exogenously added EGF to its receptor. The cause of this reduced binding capability seems to be, that in cells transformed by these viruses, a new growth factor (TGF = transforming growth factor) is produced, that competes with EGF for the binding of its receptor (66). That a connection exists between some oncogene encoded products and growth factors became clear, when protein sequencing of a part of human PDGF revealed complete concordance in 104 amino acid residues with the deduced protein sequence of v-sis (67, 68).

The information regarding the other viral oncogene products is less elaborate. None encode proteins with tyrosine specific kinase activity. v-myc codes for a DNA-binding nuclear protein (69, 70). No protein product of v-rel has been reported to date. v-fms and v-raf both encode identifiable products of unknown function (71, 72). v-erbB seems to be a cell membrane glycoprotein (73). The p21 of Harvey and Kirsten MuSV (v-Ha-ras and v-Ki-ras) has a protein kinase activity; it specifically binds guanine containing nucleotides and can phosphorylate itself by transferring a phosphate of GTP to a threonine residue on its own molecule (74).

1.4 Cellular Progenitors of Viral Oncogenes

With the isolation of new oncogenic retroviruses, the possibility emerged that these viruses, during passage in animals, recombined with the genome of the host and that this recombination resulted in the formation of an acutely transforming virus. Molecular hybridization and the ability to prepare radioactive cDNAs were the tools used to show that the Rous sarcoma virus had acquired specific sequences (c-src) from the chicken genome (75). Subsequently it was possible to show that the presence of these sequences was not restricted to chicken; v-src homologous sequences were found in fish and human with a nucleotide divergence possibly as low as 15%, indicating that this gene is evolutionarily conserved (76). Very recently (57, 77) v-src related sequences were found in the insect Drosophila, which extends the conservation of this gene over an even longer evolutionary period.

With the introduction of molecular cloning it became possible to specifically clone the acquired sequences of tumorigenic retroviruses and to perform these studies in a much more straightforward and detailed way. These types of studies established that the majority of the retroviral oncogenes are

acquired from the cellular genome and that the extent of evolutionary conservation varies among the different oncogenes.

The cellular homologs of most of the known viral oncogenes exist as single loci within a given species. Although in some cases additional related sequences can be detected (78) these cellular sequences are only partly homologous to viral oncogenes, which makes it unlikely that viral oncogenes are derived from them. More likely these related sequences are derivatives of the original cellular gene and unable to encode a functional protein ("pseudo-genes"), "processed genes," or related loci. The latter possibility is illustrated by v-ras, of which two distinctive forms have been acquired, one by the Harvey and one by the Kirsten strain of sarcoma virus (79,80). Although the acquired sequences of these viruses are related they are derived from different loci. The acquired sequences of the transforming viruses do not exceed a stretch of more than 3 kb of DNA. However their cellular progenitors are in most cases found to be distributed over a much larger stretch of DNA, indicating the presence of intervening sequences. The finding that most of the cellular homologs contain one or more intervening sequences unambiguously demonstrates that their origin is not viral.

That the cellular oncogenes are strongly conserved in evolution implies that these sequences encode mRNAs and that their translational products are essential for the cell (i.e., "housekeeping" enzymes). First experimental evidence for their presence was obtained with the discovery of RNA homologous to v-src in uninfected fibroblasts of several avian species (81). Similar findings were subsequently made for other cellular oncogenes. Exceptions were, however, found; no c-fes transcript so far has been detected; moreover, a detailed search for c-mos mRNAs in many different tissues has not revealed any transcripts.

Detection of cellular oncogene proteins has not been easy. To date, three proteins have been identified. The most thoroughly studied is $pp60^{c\text{-}src}$, a 60,000 MW phosphoprotein; $p21^{c\text{-}ras}$, a 21,000 MW protein encoded by Harvey c-ras and $P150^{c\text{-}abl}$, a 150,000 MW protein most probably encoded by c-abl.

The translation products of v-src and c-src are remarkably identical; moreover, deletion mutants of v-src can be restored by recombination with c-src resulting in a functional oncogene (82).

The DNA-mediated transfection procedure, introduced by Graham and van der Eb in 1973 (83), permits transfer of genetic traits to cells in culture by exposing these cells to DNA containing the genes encoding the characteristics to be transferred. In this way, for instance, cells lacking the activity of thymidine kinase (TK) can be made TK^+ by transfecting DNA of Herpes simplex type II virus, which carries a gene for thymidine kinase. In some cases, the donor DNA is stably integrated into the genome of the recipient.

This procedure provides a new approach to investigate if DNA sequences have oncogenic activity; this DNA can be transferred to recipient non-transformed cells through DNA-mediated transfection procedures. The successful transfer of an oncogene would be apparent by the formation of foci of transformed recipient cells, which are usually NIH 3T3 cells.

The finding that the viral oncogenes were derivatives of cellular progenitors introduced the tentative possibility that the cellular oncogenes could be induced to cause malignant transformation in other ways than through integration in retroviruses. Data concerning this question accumulated rapidly after the molecular cloning of the first cellular oncogenes: introduction of viral transcriptional regulatory sequences (LTR) 5' of normal mouse c-mos

and rat c-ras by ligation, subsequently followed by transfection to NIH 3T3 cells resulted in the activation of the transforming potential of these cellular genes (84, 85). The number of foci was comparable to that obtained upon transfection of the homologous sarcoma virus DNA's. The cellular mos and ras genes do not exihibit transforming activity on their own. However, removal of cellular sequences 5' of the mouse c-mos gene results in a very low transforming activity (86). The hypothesis that this low transforming activity is a consequence of the integration of this gene nearby strong cellular promotors is very plausible. Further evidence that cellular oncogene expression could be responsible for at least some forms of tumorigenesis was obtained from the study of avian leukosis virus (ALV). This virus has no acquired sequences (oncogene) yet it is capable of inducing B-cell lymphoma in chickens. Its tumorigenicity was explained by the finding that in these ALV-induced lymphomas, ALV is integrated adjacent to the c-myc gene and that transcription, initiated from the viral promotor of ALV, causes enhanced expression of c-myc, leading to neoplastic transformation (87). These data showed that a carcinogenic agent (ALV) acts by altering the expression of a cellular oncogene, c-myc.

Recently the DNA transfection procedure was employed in a new way to detect DNA sequences possibly involved in tumorigenesis. Instead of transfection of a particular DNA segment to recipient contact inhibited eukaryotic cells, the cells were exposed to total cellular DNA. This type of experiment resulted in the identification of oncogenes homologous to the acquired sequences of retroviruses. However, in some experiments the results were not in agreement with the expectations. When DNA of an ALV induced chicken lymphoma was transfected to NIH 3T3 cells a number of foci were obtained. Upon cloning, DNA sequences responsible for focus formation were distinct from the expected c-myc gene. Nucleotide sequencing identified the transforming sequence to be coding for a

small protein characterized by partial homology with transferrin (88). In similar experiments, by the same group, the DNA of lymphoid neoplasms induced by Abelson murine leukaemia virus (AbMuLV) was used to transfect NIH 3T3 cells. Again it was shown that the transforming sequences were distinct from AbMuLV sequences (89). This apparent contradiction of results could be explained if some forms of tumorigenesis are multiple step events.

Strong evidence for the involvement of cellular oncogenes in tumorigenesis was recently obtained from mouse myelomas and plasmacytomas. In a mouse myeloma it was found that 5' of the c-mos gene normal cellular sequences were substituted by an insertion sequence (IS)-like element. Upon cloning of this rearranged gene it was subsequently shown that in transfection experiments it was capable of transforming mouse fibroblasts (90). Many mouse plasmacytomas were found to have rearranged immunoglobulin heavy chain constant regions. A certain DNA sequence (called LyR, NIARD or NIRD by different investigators) was found to be translocated into this region; (91, 92, 93) subsequently it could be shown that this sequence represented c-myc sequences (94, 95, 96). As a result of these findings an analogous translocation was found with human c-myc in Burkitt lymphoma. In the next chapter, the involvement of c-myc in tumorigenesis will be discussed in more detail.

1.5 Human sequences with potential transforming activity.

a. Overview

An increasing number of human sequences with a potential to transform cells in culture has been isolated. Using viral oncogene sequences as probes, the following human cellular homologs have been isolated: c-abl (paper VII), c-fes (97, 98, paper III), c-fms (paper I), c-fos (99), c-mos (100, 101), c-myb (102), c-myc (103, 104), c-Ha-ras (105), c-Ki-ras (105), and c-sis (106). The

isolation of the human cellular homologs of v-yes, v-erb and v-rel has not yet been reported. This may reflect difficulty using avian probes to isolate human sequences that are evolutionary distant. For instance, although v-fps (chicken) and v-fes (cat) both detect an identical locus in human (paper IV) the strength of hybridization is much less when v-fps sequences are used as a probe. Sequences homologous to v-rel seem not to be highly conserved in non-avian species (23, 107).

Other human sequences, with transforming activity, have been shown to exist by means of DNA-mediated transfection experiments; DNA from either human tumor tissue or cell lines established from human tumors transfected to mouse NIH 3T3 cells induces focus formation at a low frequency. These foci contain human sequences, detectable by an Alu-repeat probe, which is specific for human DNA. Some of these transforming sequences have been molecularly cloned.

What could be regarded as evidence that these human oncogenes represent causative agents in tumorigenesis? Because the protein products of many human oncogenes have not yet been identified, expression has mostly been studied on the RNA level, using either viral or human probes. An increased level of RNA production of a particular oncogene in transformed cells is thought to be an indication that the oncogene may have taken part in the neoplastic process. However, elevated levels of oncogene RNA may also be a phenomenon accompanying neoplastic transformation but not necessarily causing it: altered expression of many enzymes and proteins in tumors has been reported (for a review, see 108), reflecting changes in metabolism of the transformed cells. Hypomethylation of human growth hormone, α-globin and γ-globin sequences in cancerous tissues in comparison with healthy tissue of the same patient has also been reported (109); the tissues studied normally do not express these genes.

Increased mRNA production can be caused by an amplification of oncogene sequences, for example by duplication of the chromosome on which the oncogene is situated. Alternatively, only limited regions of a chromosome may be amplified. Alterations in the signals responsible for the regulation of transcription can also cause an increase in mRNA production. These alterations can be local, by the insertion or deletion of sequences near the oncogene or by mutations in the promoter sequences of an oncogene. Or as a part of a gross rearrangement, oncogene sequences can be translocated to other chromosomes and either activated through position effects (e.g., an oncogene that is normally not transcribed is translocated into a region of high transcriptional activity) or through replacement of its own promoter by a promoter at the breakpoint on the recipient chromosome. Novel species of mRNA could be synthesized in this case.

One firmly established fact is, that human neoplasms do very often contain chromosomal aberrations. Double minute chromosomes (DM), which contain as yet unidentified portions of chromosomes have been found in tumor cell lines; they could represent amplification of one limited large stretch of DNA. Translocations between chromosomes in human neoplasms, especially in leukemias and lymphomas, have been reported in an increasing number (110). The first discovered and most well known example is the Philadelphia translocation, t(9;22), associated with chronic myelogenous leukemia (CML) (111). In this translocation a piece of the long arm of chromosome 22 (q11-qter) is translocated to chromosome 9 (112); a very small piece of chromosome 9 is translocated to chromosome 22 (paper IX). The $22q^-$ chromosome (the Philadelphia chromosome) is present in the leukemic cells of about 96% of all CML patients; this generally appears to be the only chromosomal abnormality prior to the blast crisis during which other abnormalities become apparent (110).

In certain forms of acute promyelocytic leukemia (APL) a t(15;17) is characteristic (113, 114). Recently, a correlation has been found between Burkitt lymphoma and translocations involving chromosome 2, 14 or 22 with chromosome 8 (115). With the improvement of cytogenetic techniques, the number of specific chromosomal aberrations associated with specific neoplastic disease is increasing rapidly (116). For example, among the solid tumors small lung cell carcinoma (117) is correlated with a deletion in chromosome 3 and Wilms tumor with a deletion of chromosome 11 at band p13 (118). However, if a chromosomal translocation is the direct cause of a neoplastic disease or if the disease causes chromosomal aberrations is a question difficult to answer but important to the debate if human oncogenes have oncogenic potential.

Another mechanism, by which an oncogene might acquire tumorigenic properties is by a mutation, for example in the coding sequences; the oncogene sequences could then code for a protein with activity or characteristics different from the normal protein, which makes it turmorigenic. This will be discussed in the chapter concerning human c-Ki-ras and c-Ha-ras.

Models have been proposed which catagorize all the above described mechanisms by which an oncogene might become tumorigenic. One model states that the overproduction of an oncogene product may cause malignancy; the second model supposes that a mutation in the coding region of an oncogene leads to the production of a new oncogenic protein. A third alternative is that a normal oncogene product is produced in normal quantities, but at the wrong moment in a developmental process.

b. Human c-abl

Human c-abl has been molecularly cloned using a cosmid vector system because of the presence of an extensive region of DNA

homologous to v-abl (paper VII). Apart from the main human c-abl locus, other regions of more limited homology to v-abl were found, encompassed by two separate EcoRI fragments of 4.1 and 12.5 kb (paper VII). The 12.5 kb EcoRI fragment has been molecularly cloned in phage (unpublished results). Homology only to a central part of v-abl could be detected; in v-abl, this region contains the putative tyrosine phosphorylation acceptor site which is highly conserved between v-abl, v-src and v-fes. It is possible that sequences in the 12.5 kb EcoRI fragment code for a protein with tyrosine-specific protein kinase activity. That these sequences represent a "processed gene" is not likely: the region of homology to v-abl stretches out over 7.4 kb whereas v-abl itself is approximately 3.0 kb in length. Therefore these human sequences may contain introns.

The human v-abl homologue is located on chromosome 9q34 (paper V). In CML, it is translocated to chromosome 22 in the standard t(9;22) Philadelphia translocation (paper IX) and in complex translocations such as the t(1;9;22) and the t(9;11;22) (ref. 141). In one CML patient, a chromosomal breakpoint was found 5' to the human c-abl locus; a restriction enzyme fragment containing sequences from chromosome 9 and 22 was molecularly cloned from the DNA of this patient (paper XI). Using a probe isolated from the chromosome 22-specific sequences of this fragment, an extended region of chromosome 22 DNA was cloned. In the DNAs of all Ph'-positive CML patients examined a chromosomal breakpoint was found within a limited region of approximately 5 kb on chromosome 22 (paper XII). The human CML cell line K562 appears to be representative of CML in that a chromosomal breakpoint is present within that same region on chromosome 22 (unpublished observations); furthermore, the five fold amplification of the immunoglobulin λ light chain constant region on chromosome 22 (paper XI), of the breakpoint region remaining on chromosome 22 (unpublished observations) and of the human v-abl homologue (paper XI) suggests that at least part of the Philadelphia chromosome has selectively been amplified in this cell line.

Human c-abl is expressed in various human cell lines; at least four RNA species of different molecular weights (7.2, 6.4, 3.8 and 2.0 kb) were found using a v-abl probe, both in hematopoietic and in tumor cell lines; normal fibroblasts contained only two species of high molecular weight (119, 120). Others, however, have reported the presence of only two species of v-abl homologous RNA, of 5.4 and 5.2 kb molecular weight in all human cells examined (121). Two additional RNA species of 6.8 and 6.2 kb were found in a human cell line estabished from a patient with a pre B acute lymphoblastic leukemia (ALL). However, transfection of this DNA to NIH 3T3 cells yielded foci which did not contain detectable human v-abl homologous sequences. Production of a transforming growth factor was evident both in the original ALL cell line and in NIH 3T3 cells transfected by DNA of this cell line. It was suggested, that sequences responsible for the induction of foci in NIH 3T3 either coded for a (human) transforming growth factor or induced the production of a transforming growth factor in the NIH 3T3 cells (121).

c. Human c-fes

RNA transcripts of human c-fes have not been found in detectable quantities in 36 human cell lines examined (unpublished observations). Molecularly cloned human c-fes does not exhibit transforming activity, even if a viral LTR is ligated 5' to the putative coding region (unpublished observations). Human c-fes maps on chromosome 15 (paper V, 102) in the region q25-26 (122). In the t(15;17) characteristic of some forms of APL, c-fes was not found on the 15q+ chromosome in one patient (paper VI), suggesting that it had been translocated to chromosome 17. However no sequences up to 30 kb 5' and 12 kb 3' of human v-fes homologous sequences were found to be retained on the 15q+ chromosome from the same patient indicating the absence of a chromosomal breakpoint near human c-fes (unpublished observations). These observations are concordant with the localization of human c-fes to 15q25-26 (122) and the breakpoint

on chromosome 15 in APL to band q22 (paper VI).

d. Human c-fms

Human c-fms, as human c-abl, appears to be a very large gene; sequences homologous to v-fms stretch out over a region of at least 30 kb. Human c-fms exhibits no transforming ability when it is transfected to Rat-2 cells, even when all v-fms homologous sequences are joined together in one cosmid clone. However, as with human c-abl, the possibility must be considered that coding sequences at the 5' region, which are not represented in v-abl or v-fms, are missing from the clones. Human c-fms has been mapped on chromosome 5, band q34 (paper II). A relatively high level of c-fms transcripts was found in human placenta (123).

e. Human c-fos

Human c-fos has been cloned as a 9 kb EcoRI fragment (99). Sequencing of the human c-fos gene has shown, that it contains introns and can encode a protein of 380 aminoacids. Human c-fos transcripts of 2.2 kb have been found in human term fetal membranes (123). Human c-fos does not transform NIH 3T3 cells; however, if the carboxy-terminus encoding region is replaced by that of v-fos (human c-fos and v-fos differ in that region by a 104 bp sequence not present in v-fos) the hybrid gene is able to induce the formation of foci (124).

f. Human c-mos

Human c-mos was cloned as part of a 9 kb BamHI restriction fragment (100); the homology to v-mos proved to be confined to a very limited region (an 0.96 kb EcoRI/BglII restriction fragment) with no apparent intervening sequences. Human c-mos and v-mos are 77% homologous in nucleic acid sequence. Although the mouse c-mos

readily transforms cells in culture when an LTR is ligated 5' to the putative coding region, human c-mos has no detectable transforming activity in a similar transfection assay (100). RNA transcripts of human c-mos have not been detected in undifferentiated B-cell lymphomas (125); however, others have reported expression in Burkitt lymphoma cell lines (126). Human c-mos has been localized to chromosome 8 (96, 127) in band q22.

g. Human c-myc

Human c-myc has been isolated in two overlapping phage clones spanning a region of more than 20 kb (103) and as a 20 kb EcoRI restriction fragment from a partial EcoRI human library (104). Homology to v-myc was restricted to two stretches of DNA smaller than 1 kb, separated by an apparent intervening sequence of approximately 1.4 kb (104). The sequences mentioned above hybridized to probes specific for both the 5' and 3' end of v-myc. On a genomic blot, more sequences with exclusive homology to a probe specific for the 5' part of v-myc are seen. Three such regions, representing apparent non-contiguous DNA sequences have been cloned (103); heteroduplex mapping showed that only 0.2-0.4 kb of these clones contained homology to v-myc (the central region); no intervening sequences were apparent. Since no transcript other than a 2.7 kb RNA species, presumably encoded by the main c-myc locus (see below) has been found, the regions with less homology to v-myc could represent pseudogenes.

Sequencing of the main human c-myc locus has revealed an open reading frame, which, (with omission of the putative intron) could code for a mRNA of 2.3 kb (104). This size is reasonably concordant with the size of an RNA of 2.7 kb found in human tumor cell lines, in normal human fibroblasts and in haematopoietic cells (119, 120).

In the acute promyelocytic leukemia cell line HL-60 (this cell line does not contain the t(15;17) often found in APL (128)) the level of transcription of the 2.7 kb mRNA was found to be approximately tenfold enhanced (120). Moreover, c-myc was amplified 8-32-fold in this cell line (129, 130). These results imply that the increased level of c-myc expression in HL-60 is due to amplification of c-myc sequences. Because amplifications of the c-myc locus was also found in primary leukemic cells of the patient from which HL-60 was established before chemotherapeutic treatment, (129), this phenomenon does not seem to have occurred during passage of the cells in tissue culture. However, it is difficult to evaluate the possible role of c-myc in the causation of this leukemia because no non-leukemic cells of the same patient could be examined. Moreoever, no amplification of c-myc was found in cells from the peripheral blood of three different patients with APL. The correlation between c-myc and APL is therefore unclear.

The same applies to another study, in which an elevated (20 to 30-fold) level of c-myc expression was found in two cell lines established from a colon carcinoma with characteristics of a neuroendocrine tumor. By in situ hybridization the presence of v-myc homologous sequences could be demonstrated in the heterogenously staining region of a marker chromosome; the human c-myc locus was amplified in both cell lines 16 to 32-fold (131).

Human c-mos and c-myc have both been mapped to chromosome 8; c-mos is located at position 8q22 (127) and c-myc at 8q24 (96, 132). Chromosome 8 is frequently involved in translocations associated with neoplastic disease; in Burkitt lymphomas for instance, translocations of chromosome 8 with chromosome 2, 14 and 22 have been described (115). In all cases the breakpoint on chromosme 8 is 8q24; those on chromosome 2, 14 and 22 are 2p12, 14q32 and 22q11 (133, 134, 96).

Chromosomes 2, 14 and 22 each contain sequences coding for immunoglobulins; the immunoglobulin heavy chain locus maps to 14q32 (135), exactly in the region that contains the breakpoint in some forms of Burkitt lymphoma. Using somatic cell hybrids of the human Burkitt lymphoma cell line Daudi with murine cells, a chromosomal breakpoint in the immunoglobulin heavy chain variable region was demonstrated (136). Subsequently, human c-myc sequences were detected in somatic cell hybrids carrying the 14q+ chromosome of Daudi and of another Burkitt lymphoma cell line, P_3-HRI, indicating that c-myc had been translocated from chromosome 8 to 14 (132).

Rearrangement of the c-myc locus has been found in 5 out of 15 B-cell lymphoma cell lines tested, both of the Burkitt and non-Burkitt type. In three, c-myc and heavy chain immunoglobulin sequences were found on the same restriction enzyme fragment. The rearrangements had occurred 5' to c-myc, resulting in an apparent head-to-head joining of c-myc with immunoglobulin sequences (137). In another study (96), one allelic copy of c-myc was rearranged in 8 out of 12 Burkitt lymphoma cell lines studied, including the Raji cell line.

Before the connection between c-myc and Burkitt lymphoma had been made, various groups had found non-immunoglobulin DNA sequences near the heavy chain constant region in murine plasmacytomas and in some B-lymphomas which contain a 15:12 translocation (chromosome 12 contains the murine heavy chain locus). These sequences, named LyR (91), NIARD (92) or NIRD (93) were molecularly cloned and subsequently shown to be murine c-myc sequences originating from chromosome 15 (94-96). By testing the hybridization of probes from various regions of the murine c-myc to murine RNA, two additional coding regions were noted other than the two identified on basis of hybridization to v-myc: one stretch of DNA was separated by sequences which did not hybridize to RNA (the apparent first intron) from the two known exons; an additional putative coding region was

present 3' to the already identified exons. In murine plasmacytoma's, however, a probe from the 5' exon does not detect homologous RNA, indicating that the translocation has altered or removed the 5' exon and the 5' regulatory sequences. A probe from the first intron does detect multiple RNA species. This may indicate, that a cryptic promotor in the first intron has been activated (95). In two mouse myelomas, M167 and M603, breaks had occurred in the first c-myc intron but at different positions. In MPC-11 the translocation had occurred in the first c-myc exon, without any visible rearrangements in the immunoglobulin heavy chain constant region. In all cases, the 5' regulatory sequences and the first exon had been disrupted, which could possibly result in the activation of several putative transcription initiation sites within the first intron. The first (non v-myc homologous) exon contains stop codons in every reading frame and could represent a non-translated leader sequence; by removal of these sequences as a consequence of the translocation a shorter transcript could be synthesized which could have a lower probability of degradation and thus a prolonged half-life. The actual cause of the translocation remains unknown, as the recombination sites in J558, M167 and M603 do not share any obvious sequence-homology; furthermore, mouse c-myc does not contain the repetitive sequences found in the μ switch region (138).

In human c-myc, an exon 5' to the v-myc homologous exons was detected using a probe from the murine 5' exon. In contrast to the mouse plasmacytomas, no altered RNA product could be found or a significant increased level of transcription. The breakpoint seemed to be near immunoglobulin switch regions in two Burkitt lines (W_1, W_2). In the c-myc locus, translocation breakpoints were all in the 5' region, at different sites: for the Lou and W_1 lines, this was 1-2 kb 5' from the second exon; in Raji, the breakpoint seemed to be 3-9 kb 5' from the second exon. Of importance was that non-Burkitt cell lines of the lines W_1, W_2 and Lou were available and showed no

rearrangements in the human c-myc locus (95). The breakpoint in the Raji cell line was defined more accurately by the molecular cloning of a DNA fragment containing c-myc sequences and immunoglobulin constant region sequences. The presence of an exon 5' to the two defined on basis of hybridization to v-myc was confirmed. No increase in the level of c-myc mRNA could be found specifically in Burkitt lymphoma cell lines. Although two species of mRNA (2.4 and 2.2 kb) were detected, both seem to be transcribed starting at the most 5' exon and neither is expressed preferentially in Burkitt lymphoma (139). In contrast, others (126) have claimed a significant increase in the levels of c-myc RNA in Burkitt lymphoma cell lines compared to control lymphoblastoid cell lines.

Although no translocation of c-myc to chromosome 22 has been demonstrated yet in the t(8;22), immunoglobulin λ light chain sequences are certainly involved; as mentioned, the breakpoint on chromosome 22 is at q11 and the λ light chain constant locus has recently been mapped to band q11 using in situ hybridizing techniques (140). In one Burkitt lymphoma cell line, sequences hybridizing to the 8q+ and to the $22q^-$ chromosomes were detected using a λ constant region probe, indicating a breakpoint in the λ constant region. In another cell line, only the $8q^+$ and not the $22q^-$ chromosome retained λ-homologous sequences (140).

h. Human c-myb

v-myb has been used as a molecular probe to investigate expression of human v-myb homologous RNA. None of 18 human cell lines established from solid tumors exhibited expression of v-myb homologous RNA sequences in detectable amounts. Mature T- and B-cells also were negative in this respect. However, an RNA species of 4.5 kb was detected in more immature cells and lymphoid cells. As with human c-myc expression in HL-60, the expression of the 4.5 kb v-myb homologous RNA was abolished if the cell line was induced

to differentiate with Me_2SO or retinoic acid (22).

Using a human probe isolated from a 2.0 kb v-myb homologous EcoRI restriction fragment, human c-myb has been localized on chromosme 6 (102) in the region q22-24 (122).

i. Human c-sis

Human v-sis homologous sequences have been molecularly cloned as a 14 kb EcoRI restriction fragment (106) from a human partial EcoRI library. These sequences seem to be present in one single locus in man, contained on large HindIII and EcoRI restriction enzyme fragments. Sequences hybridizing to v-sis are distributed over 5 apparent exons.

Although in this study (106) no polymorphisms were found in all 12 human DNA samples examined, we found a polymorphism for HindIII to be rather common; this polymorphism was exhibited by the presence of an extra HindIII site near the second intervening sequence in human c-sis (unpublished observations).

Human c-sis has been mapped on chromosme 22 (101), distal to band q11, since it is translocated from chromosome 22 to chromosome 9 in CML patients carrying the Philadelphia chromosome (paper X). However, it is not localized in close proximity to the breakpoint on chromosome 22 (8) and is translocated to other chromosomes than chromosome 9 in complex Philadelphia translocations (141).

Expression of human c-sis is not common; in 8 out of 23 tumor cell lines tested, a 4.2 kb v-sis homologous transcript was found. The highest level of expression was in the glioblastoma cell line A 172 (119); expression was found in one of the haematopoietic cell lines tested, the T-cell lymphoma cell line HUT 102, which produces the human T-cell leukemia/lymphoma retrovirus HTLV (120).

j. Human c-Ha-*ras*, c-Ki-*ras* and other Transforming Human Sequences Detected by Transfection

Using probes specific for the acquired cellular sequences of Harvey and Kirsten MuSV, the human cellular homologs of v-Ha-*ras* and v-Ki-*ras* have been molecularly cloned from normal human DNA (105). Two non-overlapping stretches of DNA were isolated with homology to v-Ha-*ras*, designated c-Ha-*ras*-1 and c-Ha-*ras*-2. Two distinct v-Ki-*ras* homologous clones, containing c-Ki-*ras*-1 and c-Ki-*ras*-2 were also isolated. Human c-Ha-*ras*-1 has four apparent coding regions, separated by intervening sequences. In contrast, c-Ha-*ras*-2 lacked intervening sequences and exhibited a limited homology to v-Ha-*ras*. Human c-Ki-*ras*-1 has 0.9 kb of sequence homologous to the viral probe, organized in two exons separated by one intron; c-Ki-*ras*-2 has one region of 300 bp homologous to v-Ki-*ras* without apparent introns. As both c-Ha-*ras*-2 and c-Ki-*ras*-2 hybridize to a limited region, that is conserved between v-Ha-*ras* and v-Ki-*ras*, they may represent functional, related genes (105). Human c-Ha-*ras*-1 has been mapped on the short arm of chromosome 11 in the region p11-p15 (105, 142, 143); human c-Ki-*ras*-2 is localized on chromosome 12 (144). Human c-Ha-*ras*-2 and c-Ki-*ras*-1 are on chromosome 6 and X (145).

Human c-Ha-*ras*-1 has no transforming activity when it is transfected to NIH 3T3 cells without further manipulation. However, if a viral LTR is ligated to it, it can transform cells at high frequency and a p21 protein can be immunoprecipitated from these cells (146). Human v-*ras* homologous transcripts have been commonly found in all samples tested (120).

Using DNA-mediated transfection experiments, a transforming gene had been identified in the T24 and EJ bladder carcinoma cell lines (147, 148). Once these genes had been cloned (147-149), they

were examined for sequence homology to other human and viral oncogenes; soon it was reported that the T24 and EJ bladder carcinoma genes were highly homologous to c-Ha-ras-1 (150) and v-bas (151), an oncogene of murine origin analogous to v-Ha-ras from rats (41). On restriction enzyme level, no obvious differences were observed between the T24 oncogene and a molecularly cloned human c-Ha-ras-1 except for a 200 bp deletion in the latter (151). However, the T24 oncogene transformed NIH 3T3 cells efficiently, whereas the human c-Ha-ras-1 did not. Two proteins (a p21 and a p23) could be immunoprecipitated from cells transfected with the T24 oncogene (151).

Since no gross differences could be detected between the T24 oncogene and its normal cellular counterpart, the transforming element in the T24 oncogene was localized more precisely by exchanging restriction enzyme fragments between the transforming and nontransforming gene and assaying the hybrid genomes for focus-forming ability. In this way, the transforming activity was localized to a small region encompassing the first exon (encoding the first 37 amino acids) and sequences immediately 5' and 3' to it (152-154). This region was subsequently sequenced. Only one essential change could be detected: in the codon for the twelfth amino acid, a T was found in the bladder carcinoma gene and a G in the normal gene, resulting in a valine at position 12 in the former and a glycine in the latter (152-154). Interestingly, Kirsten, Harvey and Balb MuSV also differ in the amino acid at position 12; v-Ha-ras has an arginine, v-bas has a lysine and v-Ki-ras has a serine. Like the normal human c-Ha-ras-1, rat c-Ha-ras-1 has a glycine at position 12 (153). The change of one amino acid residue can be detected on protein level; in normal cells two proteins (collectively described as p21) are immunoprecipitated with v-Ha-ras-1 p21 specific antiserum. In cells transfected by DNA of the EJ bladder carcinoma cell line, four proteins can be detected, two of which are identical to those found in normal cells. The other two migrate slightly slower on gels and are encoded by the

transforming allele of the c-Ha-ras-1 in the EJ cell line (152). Other studies have confirmed that only the single basepair change in the transforming oncogene distinguishes it from its nontransforming counterpart (155).

An independent, apparently spontaneous, mutation was found in an originally non-transforming c-Ha-ras-1 allele; it had become transforming after transfection to NIH 3T3 cells. In this case, a point mutation had also occurred in the twelfth codon, but the normally occurring glycine was replaced by an aspartic acid (156). A transforming c-Ha-ras-1 gene has also been cloned from a lung carcinoma cell line, Hs242. In contrast to the T24 oncogene, the transforming sequences were not localized to the region encoding amino acid residue 12, but to a region more to the 3' in the gene; a single basepair change had occurred in codon 61, resulting in the substitution of a leucine for a glutamine (41).

Transfection experiments with DNA isolated from solid tumors has shown (157) that some of these also contain transforming oncogenes detectable through transfection; this gave answer to an important question, namely, if the transforming oncogenes in the cell lines could have been active in the original tumor and were not merely the result of an activation process during tissue culture propagation. However, only 5 of 28 DNA samples from solid tumors contained an oncogene, detectable with the transfection assay. DNA of second-cycle transformants were examined for similarity in restriction enzyme patterns after hybridization with total human DNA and for homology to a number of viral oncogenes. Many of these transforming DNAs contained similar sequences that had been transferred to the NIH 3T3 cells, sequences that hybridized to a v-Ki-ras probe (157).

The oncogenes transmissable from the SW 480 colon carcinoma cell line and the small lung carcinoma cell line LX-1 apparently also are homologous to v-Ki-ras (158, 159). Since two v-Ki-ras

homologous loci exist in man (105), c-Ki-*ras*-1 and c-Ki-*ras*-2, each was examined for identity with the colon and lung carcinoma genes; these genes hybridized to a probe specific for c-Ki-*ras*-2, (159). The SW 480 sequences necessary for transformation stretch out over an extended region of DNA; c-Ki-*ras*-2 is a very large gene (30-35 kb), comparable to human c-*abl* and human c-*fms* (160, 161). The normal c-Ki-*ras*-2 allele does not transform NIH 3T3 cells; the transforming gene has apparently undergone a point mutation (162).

The transforming gene from the neuroblastoma cell line SK-N-SH has been molecularly cloned (163); the transforming activity is situated on two EcoRI fragments of 9.2 and 7.0 kb. This neuroblastoma oncogene seems to be related to the human *ras* gene family (163, 164). A part of the gene has been cloned from secondary transfectants of the fibrosarcoma cell line, HT-1080; it is apparently identical to the SK-N-SH transforming gene and can be demonstrated in transfectants of DNA from the cell lines RD (rhabdomyosarcoma) and HL-60 (a promyelocytic line); as this N-*ras* gene is located on chromosome 1, it is clearly distinct from the other members of the *ras* gene family. RNA transcripts of 2.2 kb were detected in comparable amounts in HT-1080 cells, in EJ cells and in normal human fibroblasts (165).

Tumorigenesis is generally regarded to be a multi-stage process. That members of the *ras* gene family are able to transform NIH 3T3 cells with single hit kinetics is an argument against the possibility that these cells are representative of actual target cells *in vivo*. The *ras* genes are not capable of transforming normal fibroblasts in the transfection assay; however, upon cotransfection with a *myc* gene or the polyoma large-T antigen normal fibroblasts can be made tumorigenic (166). A similar phenomenon was reported by others; neither the T24 oncogene (Ha-*ras*) nor the polyoma middle-T antigen can transform primary cells alone; however, co-transfection with adenovirus early region 1A yielded transformed cells (167).

Additionally, the EJ oncogene (Ha-*ras*) can only transform normal fibroblasts after they have been immortalized by carcinogens (168). These results have led to the proposal that certain complementation groups exist among oncogenic sequences: large T-antigen, adenovirus early region 1A and *myc* would belong to one group, inducing immortality. The *ras* gene family and polyoma middle-T would be implicated in events leading to anchorage independent growth and morphological change (166).

2. CONCLUSION

The connection between human oncogenes and cancer has to date not been clarified for any human oncogene. Since all viral oncogenes cause transformation, the potential to transform may be present in these particular homologous sequences in human. However, it is clear that the viral oncogenes are in no case identical to their own direct cellular progenitors: viral oncogenes lack all intervening sequences and in some instances exons present in their cellular counterparts. Point mutations and deletions further distinguish them from the cellular oncogenes (for a review, see 169). However, similar critical processes which may have changed a virally transduced, non-transforming sequence into a transforming gene may take place in an individual. Therefore, it is possible that every human oncogene is ultimately found to be connected with some kind of malignancy. Only a limited number of human oncogenes have been implicated to date *in vivo* in neoplastic diseases.

Human c-*myc* appears to be closely associated with Burkitt lymphoma and with the type of APL represented by the cell line HL-60. In the first case, c-*myc* homologous RNA is not found in significantly higher amount, indicating that the transformation is not caused by elevated transcription. Moreover, no human c-*myc* RNAs of an abnormal size have been found in Burkitt lymphomas. Recently, it was shown that the c-*myc* allele translocated to chromosome 14 is

expressed, whereas the other allele is not; this led to the proposal that the translocated oncogene somehow escapes normal regulation as a consequence of the translocation (170). The breakpoints on both chromosome 8 and 14 seem to be variable: in some cell lines (such as Daudi) no breakpoint anywhere in the neighborhood of c-myc is detectable although it is translocated to chromosome 14. In the immunoglobulin locus, breaks in the variable region and in the switch region of the heavy chain have been reported. The mechanism by which human c-myc may cause transformation in Burkitt lymphoma may therefore be called unclear. However, the strong correlation between Burkitt lymphoma and the presence of a translocation of c-myc to chromosome 14 in the t(8;14) indicates that this oncogene must somehow be involved in this disease.

The oncogenes detected by transfection do not provide a consistent linkage between one type of malignancy and one type of oncogene; both c-Ha-ras and c-Ki-ras have been implicated in the transformation in bladder carcinomas. The N-ras gene has been activated in four different types of malignancies (see Table II). Some of these oncogenes are activated by single point mutations. How common such events are in vivo remains to be determined, as none of 10 primary bladder carcinomas, 9 colon carcinomas or 10 lung carcinomas had a point mutation in the codon for the twelfth amino acid (171), although a mutation in other codons may be possible.

The involvement of human c-fes in human tumorigenesis is at the moment unclear. Since the fes gene has been "acquired" independently by four type C transforming viruses in chicken and two in cat, the oncogenicity of these sequences seems unquestionable. Nonetheless, it is possible that in the evolutionary period between man and cat, safeguards have evolved to prevent the human homolog from being expressed at the wrong moment. The highly conserved nature of fes suggests that it has an important function in the cell, but, since it has not been found expressed in any cell line

examined to date, could represent a gene activated only in a particular tissue and/or during specific developmental stages. Although involvement of human c-fes could be possible in APL because it is translocated from chromosome 15, others (172) have indicated that the critical chromosome in this disease is the $15q^{+}$. This would argue that possibly another oncogene originating from chromosome 17 has a major part in causing the malignancy. Moreover, the localization of c-fes to 15q25-26 does not coincide with the breakpoint on chromosome 15 in APL, 15q22.

The connection between human c-abl and CML is much stronger; that the translocation between chromosome 9 and 22 resulting in the formation of the Ph' chromosome was reciprocal could only be detected using molecular hybridization techniques, because the piece of chromosome 9 translocated to chromosome 22 is too small to be visualized by cytogenetic means (paper IX). The relatively small size of this piece (in cytogenetic terms; it may be more than 1000 kb long) and the fact that it contains human c-abl automatically places c-abl in the proximity of a chromosomal breakpoint. Indeed,

Table II.

Cell Line	Type Line	Activated Oncogene
EJ/T24/J82	Bladder carcinoma	c-Ha-ras-1
Hs 242	Lung carcinoma	c-Ha-ras-1
Calu-1	Lung carcinoma	c-Ki-ras-2
SK-Lu-1	Lung carcinoma	c-Ki-ras-2
SK-Co-1	Colon carcinoma	c-Ki-ras-2
SW-480	Colon carcinoma	c-Ki-ras-2
Lx-1	Small cell lung carcinoma	c-Ki-ras-1
Sk-N-SH	Neuroblastoma	N-ras
HT-1080	Fibrosarcoma	N-ras
RD	Rhabdomyosarcoma	N-ras
HL-60	Promyelocytic leukemia	N-ras

we have found a chromosomal breakpoint in one CML patient located within 14 kb 5' from the most 5' v-abl homologous sequence (paper XI). That such a proximity is fortituous seems unlikely: recently, we have found a chromosomal breakpoint in a second CML DNA in the relative neighborhood of human c-abl (unpublished observations). If the break actually occurs in the human c-abl gene itself is at the moment unresolved; v-abl is approximately 3 kb long whereas v-abl homologous human RNA's of around 6 kb have been detected (unpublished observations), suggesting the presence of as of yet unidentified exons. Furthermore, studies with the mouse c-abl gene have indicated that it contains sequences to the 5', not present in v-abl (173). If the break is outside the human c-abl locus itself, a similar situation is found as in Burkitt lymphoma, in which a breakpoint is not found in c-myc coding sequences.

In contrast to the t(8;14) in Burkitt lymphoma, the "recipient" chromosome of the human oncogene in CML, chromosome 22, contains breakpoints within a narrowly defined region (4.2 - 5.8 kb). Preliminary results indicate that this region contains protein encoding sequences not related to the λ immunoglobulin region. The breakpoints seem to occur within these sequences, although we have not yet defined if they are in exons or introns. All Ph'-positive CML patients seem to share a common breakpoint region on chromosome 22; Ph'-negative CML seems to be a different type of leukemia in this respect.

We believe that at the moment we have a strong model system to examine the possible involvement of genetic sequences in the causation of human cancer: on chromosome 22 a region is present that consistently contains Ph' breakpoints and in addition c-abl is translocated into its relative vicinity, even in the DNAs of patients having complex translocations such as a t(9;11;22) and a t(1;9;22). The specific amplification in a human CML cell line of both the c-abl oncogene and chromosome 22 sequences in the breakpoint region (unpub-

lished observations) further implies that these genetic sequences, if any, may be responsible for the causation of CML.

In general, it seems reasonable to expect that in the near future discoveries on the molecular basis of human cancer will result in the development of new technologies for early diagnosis and ultimately in new approaches to cure this disease. A great amount of attention has been focused on oncogenes in particular and this thesis is a reflection of the expectation, that oncogenes do have a role in human cancer. Obviously, the correlation between oncogene expression and human tumorigenesis has yet to be established firmly. The practical value of the knowledge, that such a correlation exists, is large.

In diagnostics, tests could be developed based on abnormal oncogene expression either at the mRNA (detecting overproduction of, or abnormal size of mRNA with Northern blot) or at the protein level (detecting products with monoclonal antibodies). DNA probes may be used for the detection of gross DNA rearrangements in specific types of human cancer: for example, in Ph'-positive CML the presence of the breakpoint on chromosome 22 and thus the Philadelphia chromosome can be detected with DNA probes. Possibly these tests could be used for diagnosis of cancer at a very early stage. Knowledge of the type(s) of oncogene(s) activated in a certain neoplastic disease could lead to more accurate prognosis.

In patients with a certain type of cancer, identification of the oncogene responsible could result in a more specific treatment. If the biochemical pathway by which the oncogene causes cancer has become evident, very effective treatments may be possible. Even if such data were not available, monoclonal antibodies against specific oncogene products could be of potential value in controlling the disease and perhaps curing it.

In existing "cancer families", in which individual members have an increased probability of developing a certain type of cancer (such as hereditary retinoblastoma), identification of e.g. the retinoblastoma-causative oncogene(s) could be invaluable for genetic counselling, using DNA probes derived from these genes. If tumorigenesis requires the activation of more than one oncogene, some individuals could be members of a "high risk group" (e.g. highly susceptable to the development of lung cancer) because they were born with one activated oncogene. These individuals could be identified early and they could adapt their life-style (e.g. refrain from smoking) with the foreknowledge of an increased risk of developing that type of cancer. Even tumor-promoting substances in the environment could be identified based on their ability to mutate certain oncogenes (such as causing base-changes in the ras gene family). Taken together, the identification of oncogenes as the causative agents of tumorigenesis is not only valuable for improved cancer diagnosis, prognosis, treatment and cure, but also for a more profound understanding of the mechanisms in the cell which dictate growth and differentiation.

SUMMARY

The human oncogene field is a relatively new area investigating the molecular basis of human tumorigenesis. Since the discovery, that acute type C RNA transforming viruses have acquired sequences (oncogenes) of cellular origin, many human homologs of these viral oncogenes have been molecularily cloned. Although the viral oncogenes unequivocally have transforming activity, there is at the moment no definite proof available that human oncogenes cause human cancer. Because of their homology to the viral oncogenes, a connection between these human cellular sequences and tumorigenesis is not unlikely and evidence that this may be the case is accumulating; this thesis forms a contribution to that evidence.

A first step in the process of investigating the possible tumorigenic properties of human genetic sequences is their molecular cloning. In research presented in this thesis, we have molecularly cloned the human oncogenes c-abl, c-fes and c-fms from a human cosmid library using viral oncogene sequences as molecular probes. These oncogenes were further characterized by nucleic acid hybridization and restriction enzyme mapping. Nucleic acid sequencing of a homologous part of human c-abl and v-abl (from mouse) revealed a striking conservation on both nucleic acid and aminoacid level; a strong homology between the v-fes/v-fps, v-src oncogenes was also observed in the same region, as well as a more distant relatedness to the catalytic chain of the bovine - cAMP protein kinase. Furthermore, v-fps (from chicken) and v-fes (from cat) were shown to be represented by one single locus in man, the human c-fes oncogene.

Certain types of human cancer are strongly correlated with chromosomal aberrations. To determine, if the isolated human oncogenes could be connected with known chromosomal aberrations and to further characterize them, we determined the chromosomal localization of each; all three oncogenes were localized to different

chromosomes, c-abl to chromosome 9 band q34, c-fes to chromosome 15 and c-fms to chromosome 5 band q34. Although human c-fes was found to be translocated from chromosome 15 in the t(15;17) associated with some forms of acute promyelocytic leukemia, we could detect no chromosomal breakpoint in the relative vicinity of this oncogene.

Human c-abl however, was consistently translocated to chromosome 22 in the Philadelphia translocation occuring in the leukemic cells of a high percentage of all cases of chronic myelocytic leukemia (CML). The human c-sis oncogene, normally located on chromosome 22, was also involved in the translocation but was not situated in close proximity to the breakpoint on this chromosome. In one CML patient, a break on chromosome 9 had occurred only 15kb upstream of the most 5' located v-abl homologous sequences. The isolation of a chimeric restriction enzyme breakpoint fragment containing sequences originating both from chromosome 9 and 22 enabled us to moleculary clone the breakpoint region of chromosome 22. Not only the DNA of the patient examined originally, but all Ph'-positive CML DNAs were found to contain a chromosomal breakpoint within a very limited region on chromosome 22.

We have molecularly cloned human genetic sequences from both chromosomes involved in the Philadelphia translocation, immediately adjacent to the chromosomal breakpoints. It will therefore be possible to examine the significance of the Philadelphia translocation in the development of CML on a genetic level. Similar approaches have been used by others to study human oncogenes, resulting in the strong implication of especially human c-ras and c-myc in human tumorigenesis. There is no doubt that the study of human oncogenes will result in a better understanding of cancer.

SAMENVATTING

Het onderzoek aan menselijke oncogenen is een betrekkelijk nieuw gebied, dat zich bezighoudt met de moleculaire basis van menselijke tumorgenesis. Sinds de ontdekking, dat acuut type C RNA transformerende virussen sequenties van cellulaire oorsprong (oncogenen) hebben verkregen, zijn vele menselijke homologen van deze virale oncogenen gekloneerd. Alhoewel de virale oncogenen zonder twijfel transformerende activiteit bezitten is er momenteel geen definitief bewijs beschikbaar dat menselijke oncogenen kanker veroorzaken. Echter, vanwege hun homologie met de virale oncogenen is het niet onwaarschijnlijk dat er een verband bestaat tussen deze menselijke oncogenen en tumorgenesis. Het aantal aanwijzingen, dat een dusdanig verband aanwezig is neemt toe en dit proefschrift levert daarin een bijdrage.

Het kloneren van deze menselijke DNA sequenties is een eerste stap van het onderzoek naar hun kankerverwekkende eigenschappen. In het onderzoek beschreven in dit proefschrift hebben wij, gebruikmakend van de virale oncogen sequenties als moleculaire probes, de menselijke c-abl, c-fes en c-fms oncogenen gekloneerd uit een menselijke cosmide bank. Verdere karakterisatie van deze oncogenen vond plaats door middel van DNA hybridisaties en het in kaart brengen met restrictie enzymen. Na het bepalen van de nucleotide-volgorde van een homoloog gedeelte van het menselijke c-abl en het virale v-abl gen (dat van muis afkomstig is) bleek zowel de aminozuurvolgorde als de DNA sequentie sterk geconserveerd te zijn; ook de v-fes/v-fps en v-src oncogenen vertoonden in hetzelfde gebied een sterke homologie, evenals, in mindere mate, de katalytische keten van cAMP -afhankelijk protein kinase. Voorts kon er aangetoond worden, dat v-fps (uit de kip) en v-fes (uit de kat) voorkomen als een enkele genetische locus in de mens, het menselijke c-fes oncogen.

Er bestaat een sterke correlatie tussen bepaalde soorten menselijke kanker en de aanwezigheid van chromosomale afwijkingen. Wij hebben de chromosomale lokatie van elk der geisoleerde menselijke oncogenen bepaald, teneinde deze verder te karakteriseren en om te bepalen of er een verband kon worden gevonden met bekende chromosomale afwijkingen; elk oncogen kon worden geplaatst op een ander chromosoom, c-abl op chromosoom 9 band q34, c-fes op chromosoom 15 en c-fms op chromosoom 5 band q34. Hoewel wij konden aantonen, dat c-fes in de t(15;17) die voorkomt in sommige soorten van acute promyelocytic leukemie van chromosoom 15 naar chromosoom 17 wordt getransloceerd, konden wij geen chromosomale breukpunt vinden in de nabije omgeving van dit oncogen.

Het menselijke c-abl oncogen werd echter consequent getransloceerd naar chromosoom 22 in de Philadelphia translocatie; deze translocatie vindt plaats in de leukemische cellen van een groot percentage van alle gevallen van chronische myeloide leukemie (CML). Het menselijke oncogen c-sis, dat normaliter aanwezig is op chromosoom 22, was ook betrokken bij de translocatie maar bleek niet dichtbij het breukpunt op dit chromosoom te liggen. In een CML patient werd de chromosomale breuk op chromosoom 9 slechts 15 kb stroomopwaarts van de meest 5' gelegen v-abl homologe sequenties gevonden. Dankzij de isolatie van een restrictie enzym fragment dat sequenties bevatte van zowel chromosoom 9 als 22, konden wij het breukpunt gebied van chromosoom 22 kloneren. Niet alleen het DNA van de patient, dat wij aanvankelijk hadden onderzocht, maar alle Ph'-positieve CML DNAs bleken een chromosomaal breukpunt te hebben binnen een zeer beperkt gebied van de gekloneerde DNA sequenties op chromosoom 22.

Wij hebben menselijke sequenties gekloneerd die direct grenzen aan de chromosomale breukpunten van de beide chromosomen, die betrokken zijn bij de Philadelphia translocatie. Derhalve zal het mogelijk zijn om de betekenis van de Philadelphia translocatie in het

ontstaan van CML op een moleculair niveau te onderzoeken. Anderen hebben een soortgelijke benadering gebruikt om menselijke oncogenen te bestuderen, waardoor de betrokkenheid van vooral het menselijke c-ras en c-myc oncogen in menselijke tumorgenesis sterk wordt gesuggereerd. Zonder twijfel zal de studie aan menselijke oncogenen leiden tot het beter begrijpen van kanker.

REFERENCES

1. Ellermann, V., Bang, O. (1908). Zentralbl. Bakteriol. Parasitenkd. Infektionskr, Hyg., Abt. 1:46, 595-609.
2. Rous, P. (1910) J. Exp. Med. 12:696-705.
3. Temin, H. M., Mizutani, S. (1970). Nature (London) 226: 1211-1213.
4. Baltimore, D. (1970). Nature (London) 226:1209-1211.
5. Poiesz, B. J., Ruscetti, F. W., Gazdar, A. F., Bunn, P. A., Minna, J. D., Gallo, R. C. (1980). Proc. Natl. Acad. Sci. USA 77:7415-7419.
6. Yoshida, M., Miyoshi, I., Hinuma, Y. (1982). Proc. Natl. Acad. Sci. USA 79: 2031-2035.
7. Shibuya, M., Hanafusa, T., Hanafusa, H., Stephenson, J. R. (1980). Proc. Natl. Sci. USA 77: 6563-6540.
8. Bartram, C. R., de Klein, A., Hagemeijer, A., Grosveld, G., Heisterkamp, N., Groffen, J. (1983) Blood 63:223-225.
9. For an overview, see Temin H.M. (1981). Cell 27: 1-3.
10. Bishop, J. M. (1978). Retroviruses. Ann. Rev. Biochem. 47: 35-88.
11. Bishop, J. M. (1981). Cell 23: 5-6.
12. Coffin, J. M., Vogt, P. K. (1981). J. Virol. 40: 953-957.
13. Yoshida, M., Kawai, S., Toyoshima, K. (1980). Nature 287: 653-654.
14. Wallbank, A. M., Sperling, F. G., Hubben, K., Stubbs, E. L. (1966). Nature 209: 1265.
15. Ghysdael, J., Neil, J. C., Vogt, P. K. (198). Proc. Natl. Acad. Sci. USA. 768: 2611-2615.
16. Wang, L-H., Feldman, R., Shibuya, M., Hanafusa, H., Notter, M. F. D., Balduzzi, P. C. (1981). J. Virol. 40: 258-267.
17. Wang, L-H., Hanafusa, H., Notter, M. F. D., Balduzzi, P. C. (1982). J. Virol. 41: 833-841.

18. Sibuya, M., Hanafusa, H., Balduzzi, P. C. (1982). J. Virol. 42: 143-152.

19. Hayman, M. J., Royer-Pokora, B., Graf, T. (1979). Virology 92: 31-45.

20. Lai, M. M. C., Hu, S. S. F., Vogt, P. K. (1979). Virology 97: 366-377.

21. Frykberg, L., Palmieri, S., Beug, H., Graf, T., Hayman, M. J., Vennstrom, B. (1982). Cell 32: 227-238.

22. Westin, E. H., Gallo, R. C., Arya, S. K., Eva, A., Souza, L. M., Baluda, M. A., Aaronson, S. A., Wong-Staal, F. (1982). Proc. Natl. Acad. Sci. USA 79: 2194-2198.

23. Chen, I. S. Y., Mak, T. W., O'Rear, J. J., Temin, H. M. (1981). J. Virol. 40: 800-811.

24. Srinivasan, A., Premkumar Reddy, E., Aaronson, S. A. (1981). Proc. Natl. Acad. Sci. USA 78: 2077-2081.

25. Goff, S. P., Gilboa, E., Witte, O. N., Baltimore, D. (1980). Cell 22: 777-785.

26. Tronick, S. R., Robbins, K. C., Canaani, E., Devare, S. G., Andersen, P. R., Aaronson, S. A. (1979). Proc. Natl. Acad. Sci. USA 76: 6314-6318.

27. Rapp, U. R., Goldsborough, M. D., Mark, G. E., Bonner, T. I., Groffen, J., Reynolds, F. H. Jr., Stephenson, J. R. (1983). Proc. Natl. Acad. Sci. USA 80: 4218-4222.

28. Curran, T., Peters, G., van Beveren, C., Teich, N. M., Verma, I. M. (1982). J. Virol. 44: 674-682.

29. Shih, T. Y., Williams, D. R., Weeks, M. O., Maryak, J. M., Vass, W. C., Scolnick, E. M. (1978) J. Virol. 27: 45-55.

30. Chien, Y. H., Lai, M., Shih, T. Y., Verma, I. M., Scolnick, E. M., Roy-Burman, P., Davidson, N. (1979) J. Virol. 31: 752-560.

31. Ellis, R. W., deFeo, D. Maryak, J. M., Young, H. A., Shih, T. Y., Chang, E. H., Lowy, D. R., Scolnick, E. M. (1980) J.Virol. 36: 408-420.

32. Ellis, R. W., De Feo, D., Shih, T. Y., Gonda, M. A., Young, H.A., Tsuchida, N., Lowy, D. R., Scolnick, E. M. (1981). Nature 292: 506-511.

33. Sherr, C. J., Fedele, L. A., Oskarson, M., Mmaizel, J., Van de Woude, G. (1980). J. Virol. 34: 200-212.

34. Fedele, L. A., Even, J., Garon, C. F., Donner, L., Sherr, C.J. (1981) Proc. Natl. Acad. Sci. USA 78: 4036-4040.

35. Donner, L., Fedele, L. A., Garon, C. F., Anderson, S. J., Sherr, C. J. (1982). J. Virol. 41: 489-500.

36. Robbins, K. C., Devare, S. G., Aaronson, S. A. (1981). Proc. Natl. Acad. Sci. USA 78: 2918-2922.

37. Devare, S. G., Reddy, E. P., Robbins, K. C., Anderson, P. R., Tronick, S. R., Aaronson, S. A. (1982). Proc. Natl. Acad. Sci. USA 79: 3179-3182.

38. Shibuya, M., Hanafusa, H. (1982). Cell 30: 787-795.

39. Hampe, A., Laprevotte, I., Galibert, F., Fedele, L. A., Sherr, C. J., (1982). Cell 30: 775-785.

40. Besmer, P., Synder, H. W.. Jr., Murphy, J. E., Hardy, W. D., Jr., Parodi, A., J. Virol. 46: 606-613.

41. Yuasa, Y., Srinivastava, S. K., Dunn, C. Y., Rhim, J. S., Reddy, E. P., Aaronson, S. A. (1983). Nature 303: 775-779.

42. Alitalo, K., Bishop, J. M., Smith, D. H., Chen, E. Y., Colby, W. W., Levinson, A. D. (1983). Proc. Natl. Acad. Sci. USA 80: 100-104.

43. Brugge, J. S., Erikson, R. L. (1977). Nature 269: 346-347.

44. Collett, M. S., Purchio, A. F., Erikson, R. L. (1980) Nature 285: 167-169.

45. Reynolds, F. H., Jr., Van de Ven, W. J. M., Stephenson, J. R. (1980). J. Biol. 255: 11040-11047.

46. Feldman, R. A., Hanafusa, T., Hanafusa, H. (1980). Cell 22: 757-765.

47. Van de Ven, W. J. M., Reynolds, F. H., Jr., Stephenson, J. R. (1980). Virology 101: 185-197.

48. Witte, O. N., Dasgupta, A., Baltimore, D. (1980). Nature 283: 826-831.

49. Kawai, S., Yoshida, M., Segawa, K., Sugiyama, H., Ishizaki, R., Toyoshima, K. (1980). Proc. Natl. Acad. Sci. USA 77: 6199-6203.

50. Feldman, R. A., Wang, L-H., Hanafusa, H., Balduzzi, P. C. (1982). J. Virol. 42: 228-236.

51. Sefton, B. M., Hunter, T., Beeman, K., Eckhart, W. (1980). Cell. 20: 807-816.

52. Pawson, T., Guyden, J., Kung, T-H, Radke, K., Gilmore, T., Martin, G. S. (1980). Cell 22: 767-775.

53. Reynolds, F. H. Jr., Van de Ven, W. J. M., Stephenson, J. R. (19). J. Virol. 36: 374-386.

54. Witte, O. N., Goff, S., Rosenberg, N., Baltimore, D. (1980). Proc. Natl. Acad. Sci. USA 77: 4933-4997.

55. Snyder, M. A., Bishop, J. M., Colby, W. W., Levinson, A. D. (1983). Cell 32: 891-901.

56. Kitamura, N., Kitamura, A., Toyoshima, K., Hirayama, Y., Yoshida, M. (1982). Nature 297: 205-208.

57. Hoffman-Falk, H., Einat, P., Shilo, B-Z, Hoffman, F. M. (1983). Cell 32: 589-598.

58. Simon, M. A., Kornberg, T. B., Bishop, J. M. (1983). Nature 302: 837-839.

59. Levinson, A. D., Courtnenidge, S. A., Bishop, J. M. (1981). Proc. Natl. Acad. Sci. USA 78: 1624-1628.

60. Van Beveren, C., Galleshaw, J. A., Jonas, V., Berns, A. J. M., Doolittle, R. F., Donoghue, D. J., Verma, I. (1981). Nature 289: 258-262.

61. Cohen, S., Carpenter, G., King, L. (1980). J. Biol. Chem. 255: 4837-4842.

62. King, L. E., Carpenter, G., Cohen, S. (1980). Biochemistry 19: 1524-1528.

63. Kasuga, M., Fugita-Yamaguchi, Y., Blithe, D. L., Kahn, C.R. (1983). Proc. Natl. Acad. Sci. USA 80: 2137-2141.

64. Ek, B., Westermark, B., Wasteson, A., Heldin, C-H. (1982). Nature 295: 419-421.

65. Rosen, O. M., Herrera, R., Olowe, Y., Petruzzelli, L. M., Cobb, M. H. (1983). Proc. Natl. Acad. Sci. USA 80: 3237-3240.

66. De Larco, J. E., Todaro, G. J. (1980). J. Cell Physiol. 102: 267-277.

67. Doolittle, R. F., Hunkapiller, M. W., Hood, L. E., Devare, S. G., Robbins, K. C., Aaronson, S. A., Antoniades, H. N. (1983). Science 221: 275-277.

68. Waterfield, M. D., Scrace, G. T., Whittle, N., Stroobant, P., Johnsson, A., Wasteson, A., Westermark, B., Heldin, C-H., Huang, J. S., Deuel, T. F. (1983). Nature 304: 35-39.

69. Donner, P., Greiser-Wilke, I., Moelling, K. (1982). Nature 296: 262-266.

70. Abrams, H. D., Rohrschneider, L. R., Eisenman, R. N. (1982). Cell 29: 427-439.

71. Reynolds, F. H., Jr., Vande Ven, W. J. M., Blomberg, J., Stephenson, J. R. (1981). J. Virol. 41: 1084-1089.

72. Rapp, U. R., Reynolds, F. H., Jr., Stephenson, J. R. (1983). J. Virol. 45: 914-924.

73. Hayman, M. J., Ramsay, G. M., Savin, K., Kitchener, G., Graf, T., Beug, H. (1983). Cell 32: 579-588.

74. Shih, T. Y., Papageorge, A. G., Stokes, P. E., Weeks, M. O., Scolnick, E. M. (1980). Nature 287: 686-691.

75. Stehelin, D., Guntaka, R. V., Varmus, H. E., Bishop, J. M. (1976). J. Mole. Biol. 101: 349-365.

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

77. Simon, M. A., Kornberg, T. B., Bishop, J. M. (1983). Nature 302: 837-839.

78. Parker, R. C., Varmus, H. E., Bishop, J. M. (1981). Proc. Natl. Acad. Sci. USA 78: 6842-5846.

79. Gross (1970) "Oncogene Viruses" Pergamon, Oxford.

80. Young, H. A., Rasheed, S., Sowder, R., Benton, C. V. Henderson, L. E. (1981). J. Virol. 38: 286-293.

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

82. Hanafusa, H., Halpern, C. C., Buchhagen, D. L., Kawai, S. (1977). J. Exp. Med. 146: 1735:1737.

83. Graham, F. L., Van der Eb, A. J. (1973). Virology 54: 536-539.

84. Oskarsson, M., McClements, W. L., Blair, D. G., Maizel, J.V., Van de Woude, G. (1980). Science 207: 1222.

85. DeFeo, D., Gonda, M. A., Young, H. A., Chang, D. R., Lowy, E. M., Scolnick, E. M., Ellis, R. W. (1981). Proc. Natl. Acad. Sci. USA 78: 3328.

86. Blair, D. G., Oskarsson, M., Wood, T. G., McClements, W.L., Fischinger, P. J., Van de Woude, G. (1981). Science 212: 941.

87. Hayward, W. S., Neel, B. G., Astrin, S. M. (1981). Nature 290: 475-480.

88. Goubin, G., Goldman, D. S., Luce, J., Neiman, P. E., Cooper, G.M. (1983). Nature 302: 114-119.

89. Lane, M-A, Neary, D., Cooper, G. M. (1982). Nature 300: 659-661.

90. Rechavi, G., Givol, D., Canaani, E. (1982). Nature 300: 607-611.

91. Adams, J. M., Gerondakis, S., Webb, E., Mitchell, J., Bernard, O., Cory, S. (1982). Proc. Natl. Acad. Sci. USA 79: 6966-6970.

92. Harris, L. J., D'Eustachio, P., Ruddle, F. H., Marcu, K.B. (1982). Proc. Natl. Acad. Sci. USA. 79: 6622-6626.

93. Calame, K., Kim, S., Lalley, P., Hill, R., Davis, M., Hood, L. (1982). Proc. Natl. Acad. Sci. USA 79: 6994-6998.

94. Crews, S., Barth, R., Hood, L., Prehn, J., Calame, K. (1982). Science 218: 1319-1321.

95. Adams, J. M., Gerondakis, S., Webb, E., Corcoran, L. M., Cory, S. (1983). Proc. Natl. Acad. Sci. USA 50: 1982-1986.

96. Taub, R., Kirsch, J., Morton, C., Lenoir, G., Swan, D., Tronick, S., Aaronson, S., Leder, P. (1982). Proc. Natl. Acad. Sci. USA 79: 7837-7841.

97. Franchini, G., Gelmann, E. P., Dalla-Favera, R., Gallo, R. C., Wong-Staal, F. (1982). Mole Cell Biol. 2: 1014-1019.

98. Trus, M. D., Sodroski, J. G., Haseltine, W.A. (1982). J. Biol. Chem. 257: 2730-2733.

99. Curran, T., MacConnell, W. P., van Straaten, F., Verma, I. M. (1983). Mole. Cell Biol. (in press).

100. Watson, R., Oskarsson, M., Van de Woude, G. F. (1982). Proc. Natl. Acad. Sci. USA 79: 4078-4082.

101. Prakash, K., McBride, O. W., Swan, D. C., Devare, S. G., Tronick, S. R., Aaronson, S. A. (1982). Proc. Natl. Acad. Sci. USA 79: 5210-5214.

102. Dalla-Favera, R., Franchini, G., Martinotti, S., Wong-Staal, F., Gallo, R. C., Croce, C. M. (1982). Proc. Natl. Acad. Sci. USA 79: 4714-4717.

103. Dalla-Favera, R., Gelmann, E. P., Martinotti, S., Franchini, G., Papas, T. S., Gallo, R. C., Wong-Staal, F. (1982). Proc. Natl. Acad. Sci. USA 79: 6497-6501.

104. Colby, W. W., Chen, E. Y., Smith, D. H., Levinson, A. D. (1983). Nature 301: 722-725.

105. Chang, E. H., Gonda, M. A., Ellis, R. W., Scolnick, E. M., Lowy, D. R. (1982). Proc. Natl. Acad. Sci. USA 79: 4848-4852.

106. Dalla-Favera, R., Gelmann, E. P., Gallo, R. C., Wong-Staal, F. (1981). Nature 292: 31-35.

107. Wong, T. C., Lai, M. M. C. (1981). Virology 11: 289-293.

108. Honey, N. K., Shows, T. B. (1983). Cancer Genet. Cytogenet: 287-310.

109. Feinberg, A. P., Vogelstein, B. (1983). Nature 301: 89-92.

110. Sandberg, A. A. (1980). The chromosomes in human cancer and leukemia. Elsevier, North Holland, New York.

111. Rowley, J. D. (1973). Nature 243: 290-293.

112. Lawler, S. D. (1977). Clin. Haemat. 6: 55-75.

113. Rowley, J. D., Golomb, H. M., Dougherty, C. (1977). Lancet 1: 549-550.

114. The Second International Workshop on Chromosomes in Leukemia, Leuven, Belgium, October 2-6, 1979 (1980). Cancer Genet Cytogenet. 2: 89-113.

115. Manolov, G., Manolova, Y. (1972). Nature 237: 33-34.

116. Yunis, J. (1983). Science 221: 227-236.

117. Whang-Peng, J., Kao-Shan, C-S, Lee, E. C., Bunn, P. A., Carney, D. N., Gazdar, A. F., Minno, J. D. (1982). Science 215: 181-182.

118. Francke, U., Holmes, L. B., Atkins, L., Riccardi, V. W. (1979). Cytogenet Cell Genet. 24: 185.

119. Eva, A., Robbins, K. C., Andersen, P. R., Srinivasan, A., Tronick, S. R., Reddy, E. P., Ellmore, N. W., Galen, A. T., Lautenberger, J. A., Papas, T. S., Westin, E. H., Wong-Staal, F., Gallo, R. C., Aaronson, S. A. (1982). Nature 295: 116-119.

120. Westin, E. H., Wong-Staal, F., Gelman, E. P., Dalla-Favera, R., Papas, T. S., Lautenberger, J. A., Eva, A., Reddy, E. P., Tronick, S. R., Aaronson, S. A., Gallo, R. C. (1982). Proc. Natl. Acad. Sci. USA. 79: 2490-2494.

121. Ozanne, B., Wheeler, T., Zack, J., Smith, G., Dale, B. (1982). Nature 299: 744-747.

122. Harper, M. E., Franchini, G., Love, J., Simon, M. I., Gallo, R.C., Wong-Staal, F. (1983). Nature 304: 169-171.

123. Muller, R., Tremblay, J. M., Adamson, E. D., Verma, I. M. (1983). Nature 304: 454-456.

124. van Straaten, F., Muller, R., Curran, T., van Beveren, C., Verma, I. M. (1983). Proc. Natl. Acad. Sci. USA 80: 3183-3187.

125. Maguire, R. T., Robins, T. S., Thorgeirsson, S. S., Heilman, C.A. (1983). Proc. Natl. Acad. Sci. USA 80: 1947-1950.

126. Nishikura, K., ar-Rushdi, A., Erikson, J., Watt, R., Rovera, G., Croce, C. M. (1983). Proc. Natl. Acad. Sci. USA 80: 4822-4826.

127. Prakash, K., McBride, O. W., Swan, D. C., Devare, S. G., Tronick, S. R., Aaronson, S. A. (1982). Proc. Natl. Acad. Sci. USA 79: 5210-5214.

128. Gallagher, R. E. (1979). Blood 54: 713-733.

129. Dalla-Favera, R., Wong-Staal, F., Gallo, R. C. (1982) Nature 299: 61-63.

130. Collins, S., Groudine, M. (1982). Nature 298: 679-681.

131. Alitalo, K., Schwab, M., Lin, C. C., Varmus, H. E., Bishop, J.M. (1983). Proc. Natl. Acad. Sci. USA 80: 1707-1711.

132. Dalla-Favera, R., Bregni, M., Erikson, J., Patterson, D., Gallo, R. C., Croce, C. M. (1982). Proc. Natl. Acad. Sci USA: 7824-7827.

133. Kirsch, I. R., Morton, C. C., Nakahara, K., Leder, P. (1982). Science 216: 301-303.

134. Malcolm, S., Barton, P., Murphy, C., Ferguson-Smith, M.A., Bentley, D. L., Rabbits, T. H. (1982). Proc. Natl. Acad. Sci. USA 79: 4957-4961.

135. Kirsch, I. R., Morton, C. C., Nakahara, K., Leder, P. (1982). Science 216: 301-303.

136. Erikson, J., Finan, J., Nowell, P. C., Croce, C. M. (1982). Proc. Natl. Acad. Sci. USA 79: 5611-5615.

137. Dalla-Favera, R., Martinotti, S., Gallo, R. C., Erikson, J., Croce, C. M. (1983). Science: 963-967.

138. Stanton, L. W., Watt, R., Marcu, K. B. (1983) Nature 303: 401-406.

139. Hamyln, P. H., Rabbits, T. H. (1983). Nature 304: 135-139.

140. de La Chapelle, A., Lenoir, G., Boue, J., Boue, A., Gallano, P., Huerre, C., Szajner, M-F., Jeanpierre, M., Laloue, J-M., Kaplan, J-C. (1983). Nucleic Acid Res. 11: 1133-1141.

141. Bartram, C. R., de Klein, A., Hagemeijer, A., van Agthoven, T., van Kessel, A. G., Bootsma, D., Grosveld, G., Ferguson-Smith, M.A., Davies, T., Stone, M., Heisterkamp, N., Stephenson, J.R., Groffen, J. (1983). Nature 306: 277-280.

142. Rampino, M. R., Reynolds, R. C. (1983). Science 219: 498-501.

143. McBride, O. W., Swan, D. C., Santos, E., Barbacid, M., Tronick, S.R., Aaronson, S. A. (1982). Nature 300: 773-774.

144. Sakaguchi, A. Y., Naylor, S. L., Shows, T. B., Toole, J.J., McCoy, M., Weinberg, R. A. (1983) Science 219: 1081-1082.

145. O'Brien, S. J., Nash, W. G., Goodwin, J. L., Lowy, D.R., Chang, E.H. (1983). Nature 302: 839-842.

146. Chang, E. H., Furth, M. E., Scolnick, E. M., Lowy, D.R. (1982). Nature 297: 474-478.

147. Goldfarb, M., Shimizu, K., Perucho, M., Wigler, M. (1982). Nature 296: 404-409.

148. Pulcian, S., Santos, E., Lauver, A. V., Long, L.K., Robbins, K.C., Barbacid, M. (1982). Proc. Natl. Acad. Sci. USA : 2845-2849.

149. Shih, C., Weinberg, R. A. (1982). Cell 29: 161-169.

150. Parada, L. F., Tabin, C. J., Shih, C., Weinberg, R.A. (1982). Nature 297: 474-478.

151. Santos, E., Tronick, S. R., Aaronson, S. A., Pulciani, S., Barbacid, M. (1982). Nature 298: 343-347.

152. Tabin, C. J., Bradley, S. M., Bargmann, C. I., Weinberg, R.A., Papageorge, A. G., Scolnick, E.M., Dhar, R-D., Lowy, D.R., Chang, E.H. (1982). Nature 300: 143-149.

153. Reddy, E. P., Reynolds, R. K., Santos, E., Barbacid, M. (1982). Nature 300: 149-152.

154. Taparowsky, E., Suard, Y. Fasano, O., Shimizu, K., Goldfarb, M., Wigler, M. (1982). Nature 300: 762-765.

155. Capon, D. J., Chen, E. Y., Levinson, A. D., Seeburg, P. H., Goeddel, D. V. (1983). Nature 301: 33-37.

156. Santos, E., Reddy, E. P., Pulciani, S., Feldmann, R.J., Barbacid, M. (1983). Proc. Natl. Acad. Sci. USA 80: 4679-4683.

157. Pulciani, S., Santos, E., Lauver, A. V., Long, L.K., Aaronson, S.A., Barbacid, M. (1982). Nature 300: 539-542.

158. Der, C. J., Krontiris, T. G., Cooper, G. M. (1982). Proc. Natl. Acad. Sci. USA 79: 3637-3640.

159. McCoy, M. S., Toole, J. J., Cunningham, J. M., Chang, E.H., Lowy, D.R., Weinberg, R. A. (1983). Nature 302: 79-81.

160. Shimizu, K., Birnbaum, D., Ruley, M. A., Fasano, O., Suard, Y., Edlund, L., Taparowsky, E., Goldfarb, M., Wigler, M. (1983). Nature 304: 497-500.

161. McGrath, J. P., Capon, D. J., Smith, D. H., Chen, E.Y., Seeburg, P.H., Goeddel, D. V., Levinson, A. D. (1983). Nature 304: 501-506.

162. Capon, D. J., Seeburg, P. H., McGrath, J. P., Hayflick, J.S., Edman, U., Levinson, A.D., Goeddel, D. V. (1983). Nature 304: 507-512.

163. Shimizu, K., Goldfarb, M., Perucho, M., Wigler, M. (1983). Proc. Natl. Acad. Sci. USA 80: 383-387.

164. Shimizu, K., Goldfarb, M., Suard, Y., Perucho, M., Li, Y., Kamata, T., Feramisco, J., Stavnezer, E., Fogh, J., Wigler, M.H. (1983). Proc. Natl. Acad. Sci. USA. 80: 2112-2116.

165. Hall, A., Marshall, C. J., Spurr, N. K., Weiss, R. A. (1983). Nature 303: 396-400.

166. Land, H., Parada, L. F., Weinberg, R. A. (1983). Nature 304: 596-602.

167. Ruley, H. E. (1983). Nature 304: 602-606.

168. Newbold, R. F., Overall, R. W. (1983). Nature 304: 648-651.

169. Duesberg, P. H. (1983). Nature 301: 219-225.

170. Nishikura, K., ar-Rushdi, A., Erikson, J., Watt, R., Rovera, G., Croce, C. M. (1983). Proc. Natl. Acad. Sci. USA 80: 4822-4826.

171. Feinberg, A. P., Vogelstein, B., Droller, M. J., Baylin, S.B., Nelkin, B. D. (1983). Science 220: 1175-1177.

172. Rowley, J. D. (1982). Science 220: 749-751.

173. Wang, J. Y. J., Baltimore, D. (1983). Mole. Cell Biol. 3: 773-779.

CURRICULUM VITAE

Name: Eleonora C. P. Heisterkamp

Date and Place of Birth: March 22, 1954; Diemen, The Netherlands

Citizenship: Dutch

Marital Status: Single

Education:

1973	Start university studies, University of Groningen, The Netherlands.
1977	"Kandidaats B1" degree, University of Groningen.
1981	"Doctoraal" degree (cum laude), Specialization Molecular Genetics, University of Groningen, The Netherlands. Start of the research documented in this thesis.

Brief Chronology of Employment:

1977 - 1981	Full time "Doctoraal" research in the following laboratories: Laboratory of Molecular Genetics, University of Groningen, studied the restriction enzyme cleavage resistance of several Bacillus subtilis phages. Laboratory of Biochemistry, studied the binding interaction of E.coli RNA polymerase with rRNA cistron promotors. Laboratory of Developmental Biology, research on RNA populations during the differentiation of the fungus Schizophyllum commune.
May 1981- June 1981	Guest Worker, Carcinogenesis Mechanisms and Control Section, Laboratory of Viral Carcinogenesis, NCI, NIH, Frederick, MD.
July 1981- Aug 1981	Consultant, Litton Bionetics, Inc., Frederick, MD
Aug 1981- Nov 1983	Visiting Fellow, Carcinogenesis Mechanisms and Control Section, Laboratory of Viral Carcinogenesis, Chief Dr. J. R. Stephenson, NCI, NIH, Frederick, MD. During this period, collaborative projects with the department of Cellbiology and Genetics in Rotterdam (Head Prof. Dr. D. Bootsma) were initiated.

Dec 1983-Present	Senior Scientist, Laboratory of Molecular Genetics, Oncogene Science, Inc., Mineola, New York.

CURRICULUM VITAE

Name: Johannes H. C. Groffen

Date and Place of Birth: February 16, 1954; Vlaardingen, The Netherlands

Citizenship: Dutch

Marital Status: Single

Education:

1973	Start university studies, University of Groningen, The Netherlands.
1977	"Kandidaats B1" degree, University of Groningen.
1981	"Doctoraal" degree (cum laude), Specialization Molecular Genetics, University of Groningen, The Netherlands. Start of the research documented in this thesis.

Brief Chronology of Employment:

1977 - 1981	Full time "Doctoraal" research in the following laboratories: Laboratory of Molecular Genetics, University of Groningen, studied the restriction enzyme cleavage resistance of several Bacillus subtilis phages. Laboratory of Developmental Biology, University of Groningen, studied mitochondrial DNA and the genes coding for rRNA of the fungus Schizophyllum commune. Laboratory of Gene Structure and Expression, National Institute of Medical Research, Mill Hill, London, molecular cloning of human -globin in phage and cosmid vectors in order to study repetitive DNA sequences and methylation patterns.
May 1981- June 1981	Guest Worker, Carcinogenesis Mechanisms and Control Section, Laboratory of Viral Carcinogenesis, NCI, NIH, Frederick, MD.
July 1981- Aug 1981	Consultant, Litton Bionetics, Inc., Frederick, MD
Sept 1981- Nov 1983	Visiting Fellow, Carcinogenesis Mechanisms and Control Section, Laboratory of Viral Carcinogenesis, Chief Dr. J. R. Stephenson, NCI, NIH, Frederick, MD. During this period, collaborative projects with the department of Cellbiology and Genetics in Rotterdam (Head Prof. Dr. D. Bootsma) were initiated.

Dec 1983-Present	Senior Scientist, Laboratory of Molecular Genetics, Oncogene Science, Inc., Mineola, New York.

APPENDIX PAPER I

Isolation of v-*fms* and Its Human Cellular Homolog

NORA HEISTERKAMP, JOHN GROFFEN, AND JOHN R. STEPHENSON[1]

Laboratory of Viral Carcinogenesis, National Cancer Institute, Frederick, Maryland 21701

Received October 28, 1982; accepted December 20, 1982

The integrated form of McDonough FeSV proviral DNA, including cellular flanking sequences, was molecularly cloned from nonproductively transformed Fisher rat cells. Acquired cellular-derived (v-*fms*) sequences within the cloned proviral DNA were mapped from between 2.6 and 5.5 kb from the 5′ LTR. Upon transfection, the cloned proviral DNA was biologically active; it caused induction of the transformed phenotype and the resulting transformed cells expressed the major McDonough FeSV translational product, $P170^{gag\text{-}fms}$ at high level. Using a series of molecular probes representing subgenomic regions of the viral v-*fms* gene, a cosmid library of human lung carcinoma DNA was screened for v-*fms* homologous sequences. Three cosmid clones containing overlapping v-*fms* homologous cellular DNA inserts, representing a contiguous region of cellular DNA sequence of approximately 64 kb in length, were isolated. Within this region of human genomic DNA, v-*fms* homologous sequences are dispersed over a total region of around 32 kb. These represent the entire human cellular homolog of v-*fms*, are colinear with the viral v-*fms* transforming gene, and contain a minimum of four intervening sequences. At least 12 regions of highly repetitive DNA sequences have been mapped in close proximity to c-*fms* coding sequences.

INTRODUCTION

Approximately 16 independent type C retrovirus "oncogenes" have now been described (Coffin *et al.*, 1981; Varmus, 1982). These were initially identified as transformation-specific sequences within the genomes of various transforming retrovirus isolates of both avian and mammalian origin. The human cellular homologs of several such viral oncogenes have been molecularly cloned and studies initiated to determine their possible involvement in naturally occurring tumors of man (Klein, 1982). Based on comparison to transforming sequences identified by transfection techniques using *in vitro* cultured tumor cells as a DNA source, one oncogene, c-ras^{Ha}, has been implicated in the induction of human bladder carcinomas (Der *et al.*, 1982; Parada *et al.*, 1982; Santos *et al.*, 1982) while a second, c-ras^{Ki}, appears to be associated with carcinomas of the lung (Der *et al.*, 1982). Other cellular homologs of viral oncogenes, including c-*fes* (Dalla-Favera *et al.*, 1982; Heisterkamp *et al.*, 1982) and c-*abl* (Heisterkamp *et al.*, 1982; de Klein *et al.*, 1982), have been mapped on chromosomes involved in translocations frequently associated with human lymphoid neoplasms. These findings indicate that studies of cellular homologs of type C viral oncogenes may provide new insights into mechanisms involved in malignant transformation.

The McDonough strain of feline sarcoma virus (FeSV) was initially isolated from a naturally occurring fibrosarcoma of the domestic cat (McDonough *et al.*, 1971). By molecular hybridization it was shown to represent a recombinant between feline leukemia virus (FeLV) and cellular sequences of cat origin (Donner *et al.*, 1982). McDonough FeSV resembles the majority of other mammalian and avian transforming retroviruses in that it is replication-defective, requiring a helper virus for *in vitro* and *in vivo* propagation (Barbacid *et al.*, 1980; Van de Ven *et al.*, 1980). The major McDonough FeSV translational product has been identified as a 170,000 M_r polyprotein ($P170^{gag\text{-}fms}$ consist-

[1] To whom reprint requests should be addressed.

ing of amino terminal FeLV *gag* gene encoded structural components, p15, p12, and p30, covalently linked to an acquired sequence encoded nonstructural component (Barbacid *et al.*, 1980; Van de Ven *et al.*, 1980; Ruscetti *et al.*, 1980). By tryptic peptide analysis (Van de Ven *et al.*, 1980) and on the basis of immunologic criteria using both polyvalent (Barbacid *et al.*, 1980; Van de Ven *et al.*, 1980) and monoclonal antisera (Veronese *et al.*, 1982; Anderson *et al.*, 1982), the v-*fms* specific region of $P170^{gag\text{-}fms}$ has been shown to be distinct from that encoded by v-*fes*, an oncogene common to the Snyder-Theilen and Gardner isolates of FeSV (Barbacid *et al.*, 1980; Van de Ven *et al.*, 1980; Fedele *et al.*, 1981). Analogous results has been obtained upon direct comparison of v-*fms* and v-*fes* by molecular hybridization (Donner *et al.*, 1982). Functionally, $P170^{gag\text{-}fms}$ differs from the Gardner and Snyder-Theilen FeSV encoded transforming proteins in that it lacks tyrosine-specific protein kinase activity (Reynolds *et al.*, 1981).

We have recently applied a cosmid vector system to the molecular cloning of the human cellular homologs of two retroviral oncogenes, v-*fes* (Groffen *et al.*, 1982) and v-*abl* (Heisterkamp *et al.*, 1983). The use of a cosmid vector for this purpose was of particular value in the case of the human c-*abl* gene because of its overall size and relatively large intervening sequences. Using a similar approach in the present study we have molecularly cloned the McDonough FeSV proviral DNA from a transformed rat cell line and, by use of a v-*fms* specific probe, identified and isolated its cellular homolog from a cosmid library of human lung carcinoma DNA. Our results establish the c-*fms* gene to resemble c-*abl* in that it is distributed over a relatively large region (>30 kb) of human genomic DNA and contains multiple regions of noncoding (intervening) sequence.

MATERIALS AND METHODS

Cells and viruses. Cells were grown in Dulbecco's modification of Eagle's medium supplemented with 10% calf serum. Cell lines include a highly contact inhibited Fisher rat embryo line, FRE 3A, and several previously described subclones of FRE 3A nonproductively transformed by the McDonough strain of FeSV (Van de Ven *et al.*, 1980). A TK^- rat cell line, Rat-2 (Topp, 1981), was kindly provided by W. Topp.

Bacteriophages and plasmids. λHF60, a recombinant of Charon 4A containing an *Eco*RI restriction fragment consisting of an entire infectious FeLV provirus (Gardner-Arnstein B) with human host flanking sequences (Mullins *et al.*, 1981), was a generous gift of J. I. Mullins. A recombinant plasmid clone, designated pHEPSTH, containing the complete Snyder-Theilen FeSV proviral DNA, including FRE 3A rat cellular flanking sequences, within the *Hin*dIII site of plasmid pHEP, and other vectors used including λgtWESλB and pBR328, have previously been described (Groffen *et al.*, 1982).

Gel electrophoresis and hybridization. Restriction enzymes were purchased from either New England Biolabs or Bethesda Research Laboratories (BRL) and were used according to the suppliers' specifications. DNAs were digested with restriction enzymes, subjected to electrophoresis through 0.75% agarose gels and transferred to nitrocellulose (Schleicher and Schuell, PH79) as described by Southern (1975). Nick translation of probes and filter hybridization were as described (Bernards and Flavell, 1980; Flavell *et al.*, 1978). Specific activity of probes was $2\text{-}5 \times 10^8$ cpm/μg. After hybridization, filters were exposed to XAR-2 film (Kodak) for up to 5 days at −70° with Dupont Lightening Plus intensifying screens.

Preparation of DNA probes. DNA probes were prepared by digesting 150 μg of DNA with appropriate restriction enzymes, followed by electrophoresis through low melting point agarose (BRL). Desired bands were excised from gels and brought into solution by heating at 65° for 30 min. Agarose was removed by two extractions with phenol equilibrated with 0.3 *M* NaOAc, pH 5.0, and one extraction with phenol:chloroform:isoamylalcohol = 25:24:1. DNA was precipitated with ethanol and 0.2 *M* NaOAc, pH 4.8, in the

presence of 20 μg/ml Dextran T-500 as carrier.

Isolation of human v-fms homologous sequences. For isolation of v-*fms* homologous cellular sequences, a previously described cosmid library of human lung carcinoma DNA partially digested with *Mbo*I was used. This library is highly representative and contains relatively large (35–45 kb) cellular inserts. Construction of the cosmid library, as well as procedures for the screening, isolation, and growth of recombinants were as described (Grosveld *et al.*, 1981).

DNA-mediated transfection. Transfection of TK^- Rat-2 cells was based on the methods of Graham and Van der Eb (1973) as modified by Wigler *et al.* (1978).

RESULTS

U_3-LTR Probe for Detection of McDonough FeSV Proviral DNA

As an initial approach to the identification and molecular cloning of the McDonough FeSV proviral DNA from nonproductively transformed rat cells, a series of molecular probes containing FeLV-specific sequences were analyzed. These included *Sma*I and *Xho*I restriction fragments of around 3.7 and 5.0 kb, respectively, from the 5′ half of the FeLV genome (Mullins *et al.*, 1981) and a probe corresponding to the complete Snyder-Theilen FeSV proviral DNA containing both 5′ and 3′ FeLV-derived sequences (Groffen *et al.*, 1982). The DNA restriction fragments were purified by preparative gel electrophoresis and ^{32}P labeled by nicked translation as described under Materials and Methods. To test the potential use of these probes for detection of the integrated McDonough FeSV provirus, high-molecular-weight DNAs were prepared from each of several independently transformed and nontransformed control FRE 3A Fisher rat clones, digested with *Eco*RI and subjected to hybridization analysis by the method of Southern (1975). With all three ^{32}P-labeled probes, multiple bands of hybridization were observed in DNAs from both transformed and control cells (data not shown), thus precluding their use for molecular cloning the McDonough FeSV provirus.

As an alternative approach to identification of the integrated McDonough FeSV proviral DNA, we undertook construction of a probe specific for the U_3 region of the Snyder-Theilen FeSV LTR. We reasoned that the McDonough FeSV proviral DNA might hybridize strongly with this probe by virtue of the fact that it contains two copies of the LTR sequences. Moreover, an FeLV-derived probe of this nature was previously shown to be of value for the identification of newly integrated exogenous FeLV proviral DNAs in cat cells (Casey *et al.*, 1981). As shown in Fig. 1, plasmid pHEPSTH, which contains a single complete copy of Snyder-Theilen FeSV proviral DNA including both 5′ and 3′ cellular DNA flanking sequences, was digested with *Kpn*I and religated generating a new plasmid, pHEPLTR, from which all of the Snyder-Theilen FeSV proviral DNA sequences, with the exception of a single LTR, had been eliminated. This construction was possible due to the presence of corresponding *Kpn*I restriction sites within the 5′ and 3′ proviral LTRs, thus generating a single hybrid LTR in pHEPLTR. For preparation of a U_3-LTR specific probe, pHEPLTR was digested with *Bam*HI and *Kpn*I, and a 0.9-kb restriction fragment containing the U_3 region of the viral LTR and approximately 0.5 kb of mink cellular flanking sequence was purified by gel electrophoresis.

To test the potential use of the U_3-LTR probe for detection of the integrated McDonough FeSV provirus, high-molecular-weight DNAs were prepared from each of five independently transformed and one control clone of FRE 3A rat cells, digested with *Eco*RI and subjected to hybridization analysis. As shown in Fig. 2, single U_3-LTR cross-reactive restriction fragments were observed at different molecular weights in DNAs from transformed clones which were not present in control FRE 3A cellular DNAs. These results are consistent with those of a previous study in which the McDonough FeSV proviral DNA was shown to lack an internal *Eco*RI restriction site (Donner *et al.*, 1982). The differ-

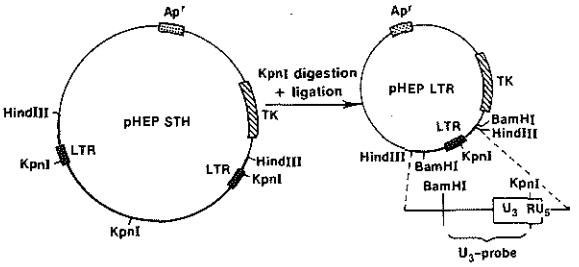


FIG. 1. Construction of U_3-LTR probe: Plasmid pHEPSTH containing a single complete copy of Snyder-Theilen FeSV proviral DNA including cellular flanking sequences terminating in *Hin*dIII restriction sites, was digested with *Kpn*I and religated, generating pHEPLTR. The position of the *Bam*I/*Kpn*I restriction fragment (U_3-probe) used for cloning of the McDonough FeSV proviral DNA is indicated.

ence in U_3-LTR homologous *Eco*RI restriction fragment sizes in the McDonough FeSV transformed clones thus reflects variations in the extents of cellular flanking sequences. In contrast, upon digestion with *Hin*dIII, an internal cutting enzyme, two U_3-LTR homologous restriction fragments were observed, a result consistent with the localization of LTRs at both the 5′ and 3′ termini of the integrated proviral DNAs. These findings establish the suitability of the U_3-LTR probe for molecular cloning of the integrated McDonough FeSV proviral DNA and suggest that it may have

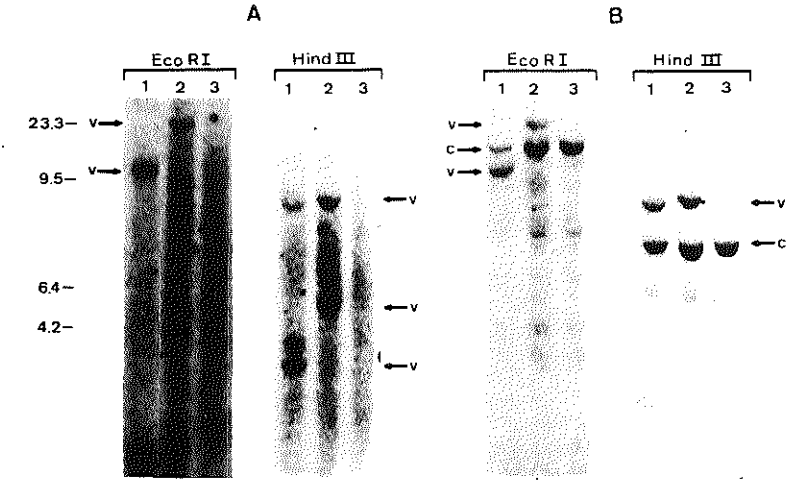


FIG. 2. Identification of v-*fms* homologous sequences in control and McDonough FeSV transformed FRE 3A rat cells. High-molecular-weight DNAs were prepared from SM-FeSV FRE 145-T (1), SM-FeSV FRE 60T (2) and control FRE 3A rat cells (3) digested with either *Eco*RI or *Hin*dIII as indicated and electrophoresed through an 0.75% agarose gel. Following transfer to nitrocellulose, filters were hybridized with ^{32}P-labeled probes including (A) U_3-LTR, and (B) probe D described in Fig. 3. ^{32}P-labeled *Hin*dIII digested λ DNA was used as a molecular weight standard. v, v-*fms*; c, c-*fms*.

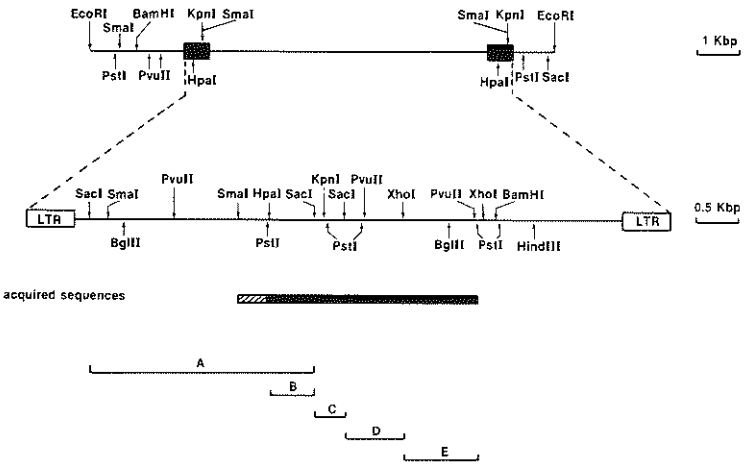


FIG. 3. Restriction endonuclease map of McDonough FeSV proviral DNA isolated from McD-FeSV FRE 145T cells and cloned in λgtWESλB. The position of acquired cellular (v-*fms*) sequences is indicated by the solid box (■) in the center of the figure. The hatched (▨) region at the 5′ terminus of these sequences indicates that the v-*fms* region terminates at a position intermediate between the corresponding *Sma*I and *Pst*I restriction sites. A series of five subgenomic probes isolated from the proviral clone are shown in the lower portion of the figure and include (A) *Sac*I/*Sac*I; (B) *Hpa*I/*Sac*I; (C) *Sac*I/*Sac*I; (D) *Sac*I/*Xho*I; and (E) *Xho*I/*Xho*I restriction fragments.

general application for the identification and molecular cloning of FeSV proviral DNAs.

Molecular Cloning of McDonough FeSV Proviral DNA

For further study, the U_3-LTR homologous *Eco*RI restriction fragment identified in one of the above clones, FRE 145-T, was molecularly cloned in λgtWESλB, subcloned in pBR328, and subjected to detailed restriction endonuclease analysis. On the basis of the data summarized in Fig. 3, a number of features of the molecularly cloned proviral DNA are apparent. First, the overall length of the cloned *Eco*RI restriction fragment is 11.2 kb, thus corresponding to that of the U_3-LTR cross-reactive *Eco*RI restriction fragment demonstrated in total FRE 145-T cellular DNA. Within the cloned DNA sequence, the proviral DNA, as defined by the positions of the two LTRs, is 8.0 kb in length and is flanked by cellular sequences of 2.2 kb and 1 kb at its 5′ and 3′ termini, respectively. Finally, the cloned sequence contains an internal *Hin*dIII restriction site as would be predicted based on hybridization analysis of FRE 145-T total cellular DNA (Fig. 2, lane A-1).

As a means of localizing the acquired sequence-specific (v-*fms*) region within the molecularly cloned McDonough FeSV proviral DNA, the viral genome was subjected to restriction enzyme digestion and various combinations of single and double digested DNA fragments analyzed for homology to a ^{32}P-labeled molecular probe corresponding to the complete FeLV provirus. These results thus localize the 5′ FeLV/v-*fms* junction site at a position intermediate between the *Sma*I and *Pst*I sites indicated by the cross-hatched box. By an analogous approach, the 3′ v-*fms*/FeLV junction site was localized between *Pst*I and *Xho*I sites also shown in Fig. 3. Because of the close proximity of these latter restriction sites to each other, the positioning of the 3′ junction site is much more

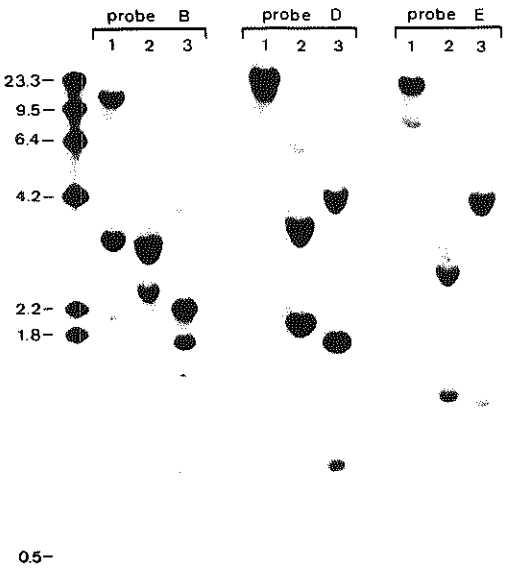


FIG. 4. Restriction endonuclease analysis of v-*fms* homologous sequences in human genomic DNA. Twenty micrograms of high-molecular-weight human DNA was digested with (1) *Eco*RI; (2) *Eco*RI and *Hin*dIII; or (3) *Eco*RI and *Bam*HI, electrophoresed through an 0.8% agarose gel, transferred to nitrocellulose, and analyzed by hybridization according to the method of Southern (1975). ^{32}P-labeled probes, prepared as described in Fig. 3, are defined on the top of each panel. ^{32}P-labeled *Hin*dIII digested λ DNA is included as a molecular weight standard.

precisely defined than that at the 5′ terminus of v-*fms*.

Having identified the acquired sequence (v-*fms*) region within the McDonough FeSV proviral DNA, we proceeded to prepare v-*fms* specific probes designated B, C, D, and E (Fig. 3). In addition, a somewhat larger probe, A, was generated encompassing both FeLV and v-*fms* specific sequences localized within the 5′ region of the McDonough FeSV proviral DNA. These probes were hybridized to the same *Eco*RI and *Hin*dIII digested cellular DNAs as initially analyzed using the U_3-LTR probe. Representative results with one such probe, D, are shown in Fig. 2B. In addition to the integrated McDonough FeSV provirus seen at the same position as with the U_3-LTR probe, its rat cellular homolog, c-*fms*, was detected in both transformed and control cells establishing the specificity of each of the v-*fms* probes for c-*fms* sequences. Although analogous results were obtained upon analysis of *Hin*dIII digested DNAs, only single McDonough FeSV proviral DNA restriction fragments hybridized with probe D, as would be expected on the basis of mapping data in Fig. 3.

To test the biological activity of the cloned proviral DNA sequence and thereby confirm its identity as McDonough FeSV, transfection experiments were performed. Upon transfection to either NIH/3T3 or Rat-2 cells, the cloned proviral DNA induced transformation at a specific activity of approximately 100 foci/μg (data not shown). For further analysis, transformed cells from one such focus were biologically cloned in microtiter plates, and tested by immunoprecipitation/SDS-PAGE analysis for expression of the major McDonough FeSV encoded polypro-

tein, P170$^{gag\text{-}fms}$. High levels of expression of P170$^{gag\text{-}fms}$ and its major cleavage product P120^{fms}, was seen in cells transformed by the cloned proviral DNA but not in control FRE 3A cells (Reynolds *et al.*, data not shown). On the basis of these results the identity of the cloned proviral DNA was confirmed as McDonough FeSV.

Human v-fms Cellular Homolog

With the availability of a series of v-*fms* specific probes, it was possible to test human cellular DNAs for v-*fms* homologous sequences. For this purpose, high-molecular-weight human DNA from A431 cells was digested with *Eco*RI and analyzed for sequence homology to probes B, D, and E. As shown in Fig. 4, v-*fms* homologous restriction fragments were detected with all three probes. For instance, with probe B, cross-hybridizing restriction fragments were observed at molecular weights of 3.0 and 13.0 kb (lane B-1). In contrast *Eco*RI restriction fragments of 13.0 and 16.0 kb hybridized with probe D (lane D-1), only the latter of which hybridized to probe E (lane E-1). These findings indicate that v-*fms* homologous sequences can be identified within total human DNA using the above probes and that such sequences encompass, at minimum, *Eco*RI restriction fragments of 3.0, 13.0, and 16.0 kb.

On the basis of the above considerations, a cosmid vector system appeared to provide the most practical approach to the isolation of human v-*fms* homologous sequences. We thus screened a previously described cosmid library of *Mbo*I partially digested human DNA for colonies hybridizing to probe D. Three positive cosmid clones identified by this means were colony purified, grown in culture, and their DNAs subjected to restriction endonuclease analysis. As summarized in Fig. 5, inserts within the three cosmid clones consisted of overlapping cellular sequences spanning a total region of 64 kb and exhibiting homology to probes A through E. Representative hybridization data for two such cosmids, cos 13 and cos 10-2, using probes A, D, and E, is presented in Fig. 6. Because of the overlapping nature of the

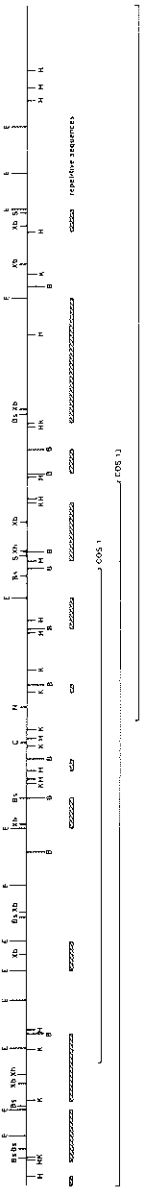


FIG. 5. Restriction enzyme map of the human c-*fms* gene. In the upper portion of the figure DNA restriction fragments homologous to v-*fms* sequences are represented as solid bars; v-*fms* subgenomic probes hybridizing to each region are indicated as A through E (See Fig. 3). Cross-hatched boxes below the c-*fms* restriction map show the positions of restriction fragments with homology to ^{32}P-labeled total human DNA. Restriction enzymes include *Bam*HI (B); *Bst*EII (Bs); *Cla*I (C); *Eco*RI (E); *Hin*dIII (H); *Kpn*I (K); *Not*I (N); *Sal*I (Sl); *Xba*I (Xb); *Xho*I (Xh). In the lower portion of the figure cellular v-*fms* homologous inserts within cos 1, cos 13, and cos 10-2 are shown.

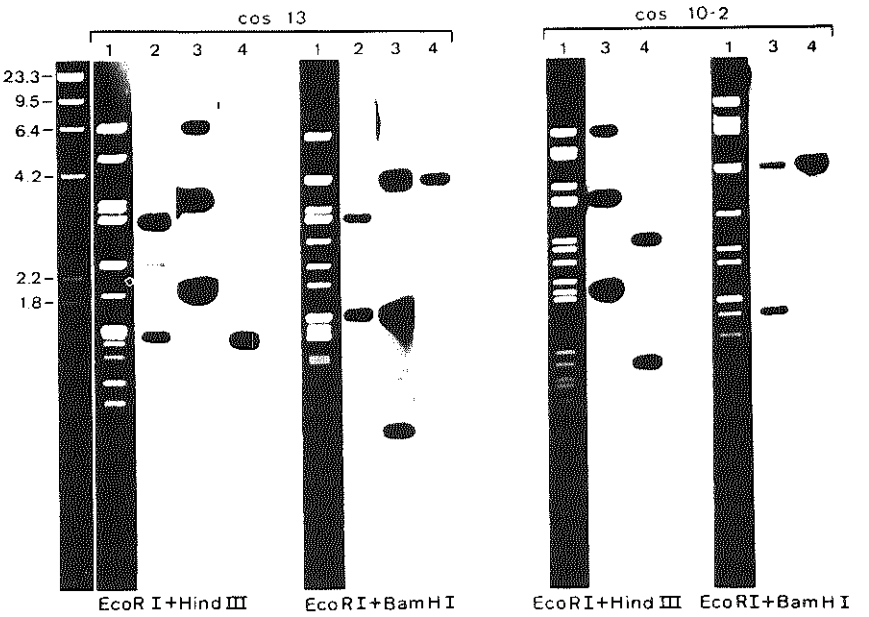


FIG. 6. Identification of v-*fms* homologous restriction fragments in human cellular DNA inserts of cos 13 and cos 10-2. Cosmid DNAs (0.4 μg) were digested with *Eco*RI in combination with either *Hin*dIII or *Bam*HI, as indicated at the bottom of each gel panel, electrophoresed through 0.75% gels, transferred to nitrocellulose, and hybridized to (2) probe A, (3) probe D, or (4) probe E. The first lane of each panel shows the ethidium bromide staining patterns of restricted cosmid DNAs. Molecular weight standards are as shown in Fig. 4.

cosmid inserts we were able to exclude the possibility of one or more internal fragments being missed from the cloned sequence. Moreover, based on extents of hybridization of restriction fragments within this cloned sequence to probes corresponding to subgenomic regions of v-*fms*, the orientation of the human c-*fms* gene could be established as shown.

To establish whether the complete human v-*fms* homolog was contained within the 64-kb cellular sequence represented by the above described cosmid cellular inserts, and that no rearrangements had taken place during the cosmid cloning, human cellular DNA was digested by *Eco*RI alone, as well as by *Eco*RI in combination with either *Hin*dIII or *Bam*HI, and analyzed by Southern blot hybridization using each of the above described probes (Fig. 5). On the basis of a comparison of v-*fms* homologous restriction fragments identified by this means (Fig. 4) to the composite human c-*fms* restriction map generated from analysis of the molecular cloned sequences, it is apparent that all of the major v-*fms* homologous fragments are represented in the cosmid clones. For instance, three of the four human v-*fms* homologous *Eco*RI fragments (2.5, 3.0, and 13.0 kb) identified in total human cellular DNA are represented in cos 13, while the fourth restriction fragment (16.0 kb) is present in cos 10-2. The smallest of these hybridizes only with probe A and thus is not seen in Fig. 4. Additionally, cos 10-2 contains 6.6 kb of the largest (13.0 kb) of the fragments in cos 13. These are identified by double digests using *Eco*RI in combination with either *Hin*dIII or *Bam*HI

and hybridization with probes corresponding to various regions of v-*fms* (Fig. 6). In addition, several minor restriction fragments were observed which hybridized to much lesser extents with one or more of the v-*fms* specific probes. These weakly hybridizing fragments could either reflect secondary loci within the human genome containing v-*fms* homologous sequences or may reflect low level cellular DNA contaminants of the v-*fms* probes.

We have previously demonstrated large numbers of highly repetitive sequences within the human c-*abl* cellular homolog, using as a molecular probe, ^{32}P-labeled total human cellular DNA (Heisterkamp *et al.*, 1983). These have been primarily localized within the noncoding (intron) regions of c-*abl*. To explore the possibility that c-*fms* may contain analogous repeats, various combinations of restriction enzyme digests of each of the cosmid clones were assayed for sequence homology to total human cellular DNA. As summarized in Fig. 5, c-*fms* was shown by this means to contain a minimum of 12 separate regions of repetitive sequence. Although the majority of these could be identified within non-v-*fms* homologous regions of the human c-*fms* locus, their precise localization will require further study.

DISCUSSION

In the present study we report the molecular cloning of a biologically active McDonough FeSV provirus from nonproductively transformed Fisher rat embryo cells. This was accomplished by use of a probe specific for the U_3 region of the Snyder–Theilen FeSV LTR. In addition, using a molecular probe prepared from the acquired sequence-specific region of the McDonough FeSV genome (v-*fms*), the human cellular homolog of v-*fms* has been isolated. It is of interest that the McDonough FeSV proviral DNA isolated in our study, although similar to, shows significant differences from a clone of McDonough FeSV described previously (Donner *et al.*, 1982). These differences are primarily manifested with respect to numbers and positions of particular restriction sites. For instance McDonough FeSV proviral DNA isolated in the present study lacks one of the *Hpa*I sites identified in the previously described clone and contains additional *Sac*I, *Pst*I, and *Bgl*II sites. Such variation may reflect evolutionary drift during the different passage histories of these virus isolates. With respect to their critical properties, however, the two clones appear highly related. Both are of the same overall size, the positioning of their acquired sequence (v-*fms*) components is similar, and upon transfection to appropriate recipient cell lines, both induce morphologic transformation at high efficiency. Moreover, both encode polyprotein gene products of the same size and immunologic properties. These include a major translational product of around 170,000 M_r and an acquired sequence specific cleavage product of 120,000 M_r.

Using a subgenomic probe corresponding to the acquired sequence-specific (v-*fms*) region of the molecularly cloned McDonough FeSV proviral DNA, we have extended the above studies to molecular cloning of the human v-*fms* cellular homolog. For this purpose a cosmid vector system was utilized because of the need to obtain cellular DNA inserts of relatively large (30–40 kb) size. The human c-*fms* gene isolated by this approach encompasses a total length of over 30 kb and contains numerous regions of noncoding (intervening) sequence. Colinearity with v-*fms* was established on the basis of relative extents of c-*fms* hybridization to subgenomic probes corresponding to different regions of the McDonough FeSV genome. It should be noted that the mapping procedure used provides an underestimate of the number of introns and overestimate of the size of v-*fms* homologous regions (exons) within the c-*fms* human locus. Moreover, when originally generated, the viral v-*fms* gene may have differed significantly from the cat c-*fms* and subsequently may have diverged further. Finally, differences in the cat v-*fms* and human c-*fms* gene may reflect differences in their species of origin. Thus conclusions regarding the positions and extents of human c-*fms* coding sequences must at present be considered

preliminary. It is, however, of interest that the c-*fms* gene contains extensive regions of highly repetitive DNA within what appears to be the noncoding (intron) regions of the genome. Similar sequences have been reported within the noncoding regions of a number of other eukaryotic genes including c-*abl* (Heisterkamp *et al.*, 1983). The nature of such sequences, their exact locations and significance will require further study.

With respect to overall size and extent of noncoding regions (introns), the human c-*fms* locus resembles c-*abl* which is also large (>25 kb) and consists of highly dispersed coding sequences (Heisterkamp *et al.*, 1983). Certain human cellular oncogenes such as c-*mos* lack detectable noncoding sequences (introns) (Prakash *et al.*, 1982; Watson *et al.*, 1982). Although the human homologs of many other viral oncogenes, including c-ras^{Ki}, c-ras^{Ha} (Chang *et al.*, 1982), and c-*fes* (Trus *et al.*, 1982; Groffen *et al.*, 1982; Franchini *et al.*, 1982), have been found to contain intervening sequences, the regions of homology have been localized within relatively small DNA fragments. Molecular cloning of each of the latter loci has been accomplished, using conventional λ phage vector systems. This approach, however, proved difficult for c-*abl* and c-*fms* because of their size, a problem which was to a large extent circumvented by use of the cosmid vector system. An additional advantage of the cosmid system for isolation of the cellular homologs of these viral oncogenes is that relatively extensive regions are obtained at their 5′ and 3′ termini, thus strengthening the possibility that the complete cellular homologs of the viral genes, including putative regulatory sequences, are represented. The availability of molecular clones of c-*fms* should now make possible studies to explore the possible involvement of this cellular oncogene in naturally occurring human tumors.

ACKNOWLEDGMENTS

The authors thank F. H. Reynolds, Jr. for analysis of $P170^{gag\text{-}fms}$ expression in v-*fms* transformed cells and for helpful discussions. The excellent technical assistance of G. T. Blennerhassett and P. Hansen is also gratefully acknowledged. This work was supported under Contract No. NOI-CO-76380 from the National Cancer Institute, Bethesda, Maryland.

REFERENCES

ANDERSON, S. J., FURTH, M., WOLFF, L., RUSCETTI, S. K., and SHERR, C. J. (1982). Monoclonal antibodies to the transformation-specific glycoprotein encoded by the feline retroviral oncogene v-*fms*. *J. Virol.* **44**, 696–702.

BARBACID, M., LAUVER, A. V., and DEVARE, S. G. (1980). Biochemical and immunological characterization of polyproteins coded for by the McDonough, Gardner-Arnstein, and Snyder-Theilen strains of feline sarcoma virus. *J. Virol.* **33**, 196–207.

BERNARDS, R., and FLAVELL, R. A. (1980). Physical mapping of the globin gene deletion in hereditary persistance of foetal haemoglobin (HPFH). *Nucleic Acids Res.* **8**, 1521–1534.

CASEY, J. W., ROACH, A., MULLINS, J. I., BURCK, K. B., NICOLSON, M. O., GARDNER, M. B., and DAVIDSON, N. (1981). The U3 portion of feline leukemia virus DNA identifies horizontally acquired proviruses in leukemic cats. *Proc. Nat. Acad. Sci. USA* **78**, 7778–7782.

COFFIN, J. M., VARMUS, H. E., BISHOP, J. M., ESSEX, M., HARDY, W. D., JR., MARTIN, G. S., ROSENBERG, N. E., SCOLNICK, E. M., WEINBERG, R. A., and VOGT, P. K. (1981). Proposal for naming host cell-derived inserts in retrovirus genomes. *J. Virol.* **40**, 953–957.

DALLA-FAVERA, R., FRANCHINI, G., MARTINOTTI, S., WONG-STAAL, F., GALLO, R. C., and CROCE, C. M. (1982). Chromosomal assignment of the human homologues of feline sarcoma virus and avian myeloblastosis virus *onc* genes. *Proc. Nat. Acad. Sci. USA* **79**, 4714–4717.

DE KLEIN, A., GEURTS VAN KESSEL, A., GROSVELD, G., BARTRAM, C. R., HAGEMEIJER, A., BOOTSMA, D., SPURR, N. K., HEISTERKAMP, N., GROFFEN, J., and STEPHENSON, J. R. (1982). A cellular oncogene (c-*abl*) is translocated to the Philadelphia chromosome in chronic myelocytic leukemia. *Nature (London)* **300**, 765–767.

DER, C. J., KRONTIRIS, T. G., and COOPER, G. M. (1982). Transforming genes of human bladder and lung carcinoma cell lines are homologous to the *ras* genes of Harvey and Kirsten sarcoma viruses. *Proc. Nat. Acad. Sci. USA* **79**, 3637–3640.

DONNER, L., FEDELE, L. A., GARON, C. F., ANDERSON, S. J., and SHERR, C. J. (1982). McDonough feline sarcoma virus: Characterization of the molecularly cloned provirus and its feline oncogene (v-*fms*). *J. Virol.* **41**, 489–500.

FEDELE, L. A., EVEN, J., GARON, C. F., DONNER, L., and SHERR, C. J. (1981). Recombinant bacteriophages containing the integrated, transforming provirus of Gardner-Arnstein feline sarcoma virus. *Proc. Nat. Acad. Sci. USA* **78**, 4036–4040.

FLAVELL, R. A., KOOTER, J. M., DE BOER, E., LITTLE, P. F. R., and WILLIAMSON, R. (1978). Analysis of the γ-β-globin gene loci in normal and Hb Lepore DNA: Direct determination of gene linkage and intergene distance. *Cell* **15**, 25–41.

FRANCHINI, G., GELMANN, E. P., DALLA-FAVERA, R., GALLO, R. C., and WONG-STAAL, F. (1982). Human gene (*c-fes*) related to the *onc* sequences of Snyder-Theilen feline sarcoma virus. *Mol. Cell. Biol.* **2**, 1014–1019.

FURTH, M. E., SCOLNICK, E. M., and LOWY, D. R. (1982). Tumorigenic transformation of mammalian cells induced by a normal human gene homologous to the oncogene of Harvey murine sarcoma virus. *Nature* **297**, 479–483.

GRAHAM, F. L., and VAN DER EB, A. J. (1973). Transformation of rat cells by DNA of human adenovirus 2. *Virology* **54**, 536–539.

GROFFEN, J., HEISTERKAMP, N., BLENNERHASSETT, G. T., and STEPHENSON, J. R. (1983). Regulation of viral and cellular oncogene expression by cytosine methylation. *Virology*, in press.

GROFFEN, J., HEISTERKAMP, N., GROSVELD, F., VAN DE VEN, W., and STEPHENSON, J. R. (1982). Isolation of human oncogene sequences (v-*fes* homolog) from a cosmid library. *Science* **216**, 1136–1138.

GROSVELD, F. G., DAHL, H-H. M., DE BOER, E., and FLAVELL, R. A. (1981). Isolation of β-globin-related genes from a human cosmid library. *Gene* **13**, 227–237.

HEISTERKAMP, N., GROFFEN, J., and STEPHENSON, J. R. (1983). The human v-*abl* cellular homologue. *J. Mol. Appl. Genet.*, in press.

HEISTERKAMP, N., GROFFEN, J., STEPHENSON, J. R., SPURR, N. K., GOODFELLOW, P. N., SOLOMON, E., CARRITT, B., and BODMER, W. F. (1982). Chromosomal localisation of human cellular homologues of two viral oncogenes (c-*fes* and c-*abl*). *Nature (London)* **299**, 747–749.

KLEIN, G. (1982). "Advances in Viral Oncology." Raven, New York.

MULLINS, J. I., CASEY, J. W., BURCK, K. B., DAVIDSON, N., and NICOLSON, M. O. (1981). The sequence arrangement and biological activity of cloned FeLV proviruses from a virus productive human cell line. *J. Virol.* **38**, 688–703.

MCDONOUGH, S. K., LARSEN, S., BRODEY, R. S., STOCK, N. D., and HARDY, W. D., JR. (1971). A transmissible feline fibrosarcoma of viral origin. *Cancer Res.* **31**, 953–956.

PARADA, L. F., TABIN, C. J., SHIH, C., and WEINBERG, R. A. (1982). Human EJ bladder carcinoma oncogene is homologue of Harvey sarcoma virus *ras* gene. *Nature (London)* **297**, 474–478.

PRAKASH, K., MCBRIDE, O. W., SWAN, D. C., DEVARE, S. G., TRONICK, S. R., and AARONSON, S. A. (1982). Molecular cloning and chromosomal mapping of a human locus related to the transforming gene of Moloney murine sarcoma virus. *Proc. Nat. Acad. Sci. USA* **79**, 5210–5214.

REYNOLDS, F. H., JR., VAN DE VEN, W. J. M., BLOMBERG, J., and STEPHENSON, J. R. (1981). Differences in mechanisms of transformation by independent feline sarcoma virus isolates. *J. Virol.* **41**, 1084–1089.

RUSCETTI, S. K., TUREK, L. P., and SHERR, C. J. (1980). Three independent isolates of feline sarcoma virus code for three distinct gag-x polyproteins. *J. Virol.* **35**, 259–264.

SANTOS, E., TRONICK, S. R., AARONSON, S. A., PULCIANI, S., and BARBACID, M. (1982). T24 human bladder carcinoma oncogene is an activated form of the normal human homologue of BALB- and Harvey-MSV transforming genes. *Nature (London)* **298**, 343–347.

SOUTHERN, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**, 503–517.

SWAN, D. C., MCBRIDE, O. W., ROBBINS, K. C., KEITHLEY, D. A., REDDY, E. P., and AARONSON, S. A. (1982). Chromosomal mapping of the simian sarcoma virus *onc* gene analogue in human cells. *Proc. Nat. Acad. Sci. USA* **79**, 4691–4695.

TOPP, W. C. (1981). Normal rat cell lines deficient in nuclear thymidine kinase. *Virology* **113**, 408–411.

TRUS, M. D., SODROSKI, J. G., and HASELTINE, W. A. (1982). Isolation and characterization of a human locus homologous to the transforming gene (v-*fes*) of feline sarcoma virus. *J. Biol. Chem.* **257**, 2730–2733.

VARMUS, H. E. (1982). Form and function of retroviral proviruses. *Science* **216**, 812–820.

VAN DE VEN, W. J. M., REYNOLDS, F. H., NALEWAIK, R., and STEPHENSON, J. R. (1980). Characterization of a 170,000-dalton polyprotein encoded by the McDonough strain of feline sarcoma virus. *J. Virol.* **35**, 165–175.

VERONESE, F., KELLOFF, G. J., REYNOLDS, F. H., JR., HILL, R. W., and STEPHENSON, J. R. (1982). Monoclonal antibodies specific to transforming polyproteins encoded by independent isolates of feline sarcoma virus. *J. Virol.* **43**, 896–904.

WATSON, R., OSKARSSON, M., and VANDE WOUDE, G. F. (1982). Human DNA sequence homologue to the transforming gene (mos) of Moloney murine sarcoma virus. *Proc. Nat. Acad. Sci. USA* **79**, 4078–4082.

WIGLER, M., PELLICER, A., SILVERSTEIN, S., and AXEL, R. (1978). Biochemical transfer of single-copy genes using total cellular DNA as donor. *Cell* **14**, 725–731.

APPENDIX PAPER II

Chromosomal localization of the human c-*fms* oncogene

John Groffen[1], Nora Heisterkamp[1], Nigel Spurr[2], Sharon Dana[3], John J.Wasmuth[3] and John R.Stephenson[1]

[1]Lab. Viral Carcinogenesis, National Cancer Institute-FCRF, Frederick, MD 21701, USA, [2]Imperial Cancer Research Fund, P.O. Box 123, Lincoln's Inn Fields, London WC2A 3PX, UK, and [3]Dept. Biological Chemistry, California College of Medicine, Univ. California, Irvine, CA 92717, USA

Received 27 June 1983; Revised and Accepted 10 August 1983

ABSTRACT

A molecular probe was prepared with specificity for the human cellular homologue of transforming sequences represented within the McDonough strain of feline sarcoma virus (v-fms). By analysis of a series of mouse-human somatic cell hybrids containing variable complements of human chromosomes it was possible to assign this human oncogene, designated c-fms, to chromsome 5. Regional localization of c-fms to band q34 on chromosome 5 was accomplished by analysis of Chinese hamster-human cell hybrids containing as their only human components, terminal and interstitial deleted forms of chromosome 5. The localization of c-fms to chromosome 5 (q34) is of interest in view of reports of a specific, apparently interstitial, deletion involving approximately two thirds of the q arm of chromosome 5 in acute myelogenous leukemia cells.

INTRODUCTION

A number of well defined genetic sequences, designated "oncogenes," have been identified within the human genome (1,2). Such sequences were demonstrated by use of molecular probes corresponding to transforming sequences of oncogenic retrovirus isolates of various avian and mammalian species. In general, cellular oncogenes have remained highly conserved throughout vertebrate evolution although their functional significance is not known. The viral counterparts of many of the oncogenes studied to date are known to encode tyrosine specific protein kinases while enzymatic functions have not been ascribed to transforming proteins encoded by the others (3). Support for the possible involvement of cellular oncogenes in human cancer has recently been derived by the demonstraton that under certain conditions, members of the ras family, can transform cells in culture (4-7). Moreover, two other human cellular oncogenes, c-abl and c-myc, are involved in translocations associated with chronic myelogenous leukemia (8) and Burkitt's lymphoma (9,10), respectively.

The human c-fms oncogene represents a cellular sequence initially identified on the basis of homology to the transforming gene (v-fms) of the

McDonough strain of feline sarcoma virus (11,12). We have recently applied a cosmid vector system for the molecular cloning of this gene (12). Although the c-fms gene product has not been identified, the corresponding viral gene, v-fms, has been shown to encode a protein with transforming function (13,14) but lacking protein kinase activity (15). In the present study we have developed a specific probe for the human c-fms gene. Using this reagent, studies were performed to determine the chromosomal localization of c-fms and its relation to chromosomal deletions and/or rearrangements associated with specific forms of human cancers.

MATERIALS AND METHODS

Cells and Viruses: Cells were grown in Dulbecco's modification of Eagles medium supplemented with 10% calf serum and included NIH/3T3 mouse cells (15), A673 human cells (15) and the human-mouse somatic cell hybrids described in Table 1. A series of human-Chinese hamster cell hybrids, including HHW105 containing chromosome 5 as its only human component and segregants of HHW105, designated HHW207, HHW209, HHW212, HHW213, HHW224, characterized by a variety of terminal and interstitial deletions of chromosome 5, have been described (16,17).

Preparation of a Human c-fms Specific Probe: The isolation of a cosmid clone with a cellular insert of approximately 38 kb containing v-fms homologous sequences (cos 13) from a library of human lung carcinoma DNA has been described (12). For use as a probe, an 0.9 kb Kpn-1 restriction fragment, was isolated from cos 13 according to previously described methods (12).

Molecular Hybridization: Restriction enzymes were purchased from New England Biolabs and Bethesda Research Laboratories and were used according to the suppliers' specifications. DNAs were digested with restriction enzymes, subjected to electrophoresis through 0.75% agarose gels and transferred to nitrocellulose essentially as described by Southern (18). Nick translation of probes and filter hybridization were as described (19). Specific activity of the probes was $2\text{-}5 \times 10^8$ cpm/ug. After hybridization, filters were washed under high stringency conditions (0.1 xSSC, 65°C) and exposed to XAR-2 film (Kodak) for up to 5 days at -70°C with Dupont Lightning Plus intensifying screens.

RESULTS

Identification of Human c-fms Specific Sequences: As a means of discriminating human c-fms sequences from related sequences in the mouse cellular genome, a series of probes were prepared from cos 13 (Fig. 1), a

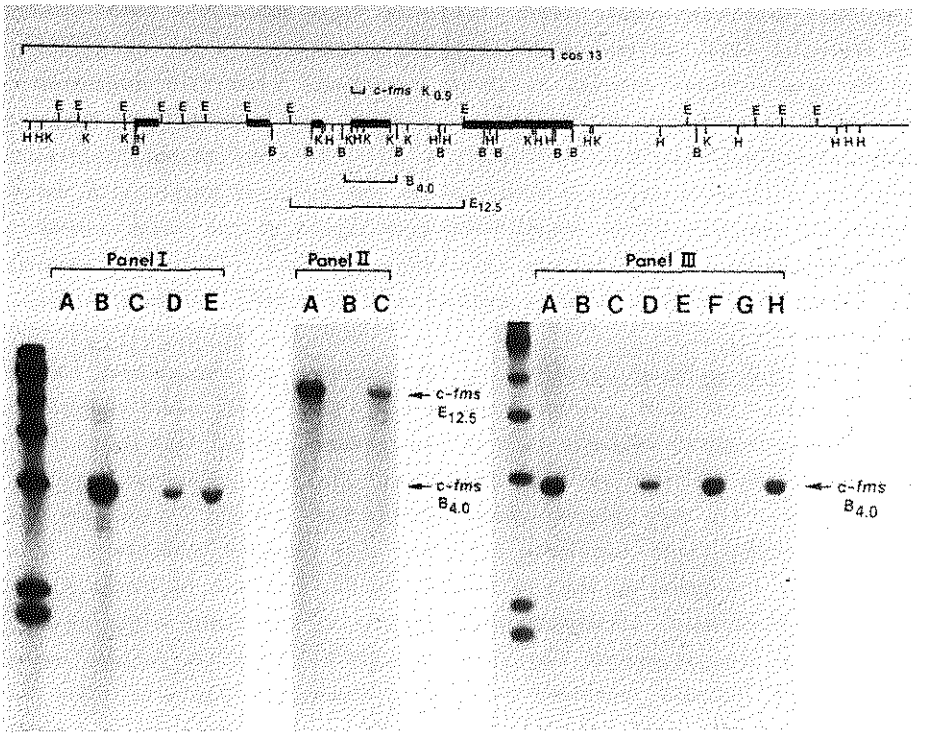


Figure 1 (Upper pannel) . Restriction endonuclease map of v-fms homologous human DNA sequences. The positions of v-fms homologous sequences, shown as solid boxes (▬) were determined as previously described (12). The position of sequences corresponding to the cellular insert in cos 13 and the 0.9 kb Kpn-1 restriction fragment used as a molecular probe for characterization of somatic cell hybrids are shown in upper portion of the figure. Single BamHl ($B_{4.0}$) and EcoRl ($E_{12.5}$) restriction fragments detected by the c-fms $K_{0.9}$ probe, are indicated below the restriction map. Restriction enzymes include: BamHl (B); HindIII (H); EcoRl (E);and Kpn-1 (K). Panels I & II: Analysis of mouse, human and representative mouse-human somatic cell hybrid cellular DNAs for human c-fms sequences. BamHl (I) and EcoRl (II) digested cellular DNAs (25 µg/lane) were electrophoresed on 0.7% gels, blotted to nitrocellulose and hybridized to the c-fms $K_{0.9}$ probe. Cell lines include: I: NIH/3T3 mouse (A); A673 human (B); and mouse-human somatic cell hybrids DTl.2R (C); Dur4R3 (D); and Dur4R4 (E); II: A673 human (A); NIH/3T3 mouse (B); and mouse-human hybrid, MOG-2E5 (C). The positions of single human DNA specific $K_{0.9}$ homologous BamHl ($B_{4.0}$) and EcoRl ($E_{12.5}$) restriction fragments are indicated to the right side of the figure. HindIII digested λDNA, included as a molecular weight marker is shown on the left. Panel III: Regional localization of c-fms on chromosome 5. BamHl digested DNAs of human A673 (A); Chinese hamster (B); and human-Chinese hamster somatic cell hybrids including HHW213 (C); HHW207 (D); HHW209 (E); HHW224 (F); HHW212 (G); HHW105 (H) were analyzed for human c-fms specific sequences as above.

previously described cosmid clone containing v-fms homologous human cellular DNA sequences (12). One such probe, c-fms $K_{0.9}$ was suitable for the purpose of the present study in that it lacked detectable repetitive sequences and hybridized specifically to single 4.0 kb BamHl and 12.5 kb EcoRl human DNA restriction fragments (Fig. 1, Panels I & II). The sizes of both restriction fragments correspond to those predicted on the basis of restriction endonuclease map analysis of the v-fms homologous region of the human genome. No detectable cross-homology between this probe and mouse cellular DNA sequences was observed when high stringency washing conditions were used as described in Methods.

c-fms Maps on Human Chromosome 5: To determine the chromosomal location of the human cellular homologue of v-fms, EcoRl and BamHl restriction endonuclease-digested cellular DNAs of human-mouse somatic cell hybrids were screened for human c-fms specific sequences. The above described c-fms $K_{0.9}$ probe was used for this purpose. These hybrids each contain a full complement of mouse chromosomes but have retained only limited numbers of human chromosomes. As shown in Fig. 1, Panels I & II, it was possible, by this means, to distinguish between hybrids containing the human c-fms gene and those lacking such sequences. As shown in Table 1, six of the fifteen somatic cell hybrids analyzed contained c-fms $K_{0.9}$ homologous sequences. Only four chromosomes (5,12,14 and 17) are common to all six c-fms positive hybrids. These findings localize c-fms to human chromosomes 5,12,14 or 17. To further localize c-fms we examined the chromosomal content of the series of nine hybrids which were non-crossreactive with the c-fms $K_{0.9}$ probe. As summarized in the lower portion of Table 1, only human chromosome 5 was missing from each of these nine hybrids. This excludes the possibility of c-fms mapping on chromosomes 12,14 or 17, each of which were present in at least three of the c-fms negative hybrids, and thus localizes c-fms to human chromosome 5.

Regional Localization of c-fms to Chromosome 5 (q34): The recent isolation of a series of segregants of Chinese hamster-human somatic cell hybrids expressing variable numbers of genes mapping on chromosome 5 (16,17) provided a means of regionally localizing c-fms. The majority of these segregants are characterized by terminal deletions involving well defined regions of the long arm of chromosome 5; these were isolated by positive selection for leuS, which maps on the p arm or very near the centromere on the q arm of the chromosome, and negative selection directed against markers on the distal half of the q arm (emtB and chr). DNAs prepared from Chinese hamster-human hybrid cells were analyzed for c-fms $K_{0.9}$ homologous sequences

TABLE 1: ANALYSIS OF HUMAN-MOUSE SOMATIC CELL HYBRIDS FOR HUMAN c-fms SPECIFIC SEQUENCES[1]

Hybrid	Ref	Human Chromosomes																						c-fms
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	
$CTP_{34}B_4$	21	+	+	+	−	+	+	−	−	+	−	−	+	−	+	−	+	+	+	−	−	−	−	+	+
DUR_4R_3	22	−	−	+	−	+	−	−	−	−	−	+	+	+	+	−	−	+	+	−	+	+	+	−	+
DUR_4R_4	22	−	−	+	−	+	−	−	−	−	−	+	+	+	+	−	−	+	−	−	−	+	+	−	+
$MOG^{-}2E5$	19	+	−	+	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	+	+	+	+	+	+
SIR-7 A2	19	+	+	−	−	+	−	+	−	−	+	+	+	−	+	+	+	+	+	+	+	+	−	+	+
MOG-2 G1	19	+	−	+	+	+	+	+	+	−	+	−	+	+	+	−	+	+	+	+	+	+	−	NT	+
DT1.2R	26	−	−	+	−	−	−	−	+	−	+	+	−	+	−	−	−	+	+	−	+	+	+	−	−
CTP41P1	21	−	−	+	−	−	+	+	−	−	−	−	−	−	+	−	+	+	−	−	+	−	−	+	−
W_3C1_6	25	−	−	−	−	−	−	−	−	−	+	+	+	−	+	+	−	+	−	−	−	+	−	+	−
$F4_4SC13_3C112$	10	+	−	−	−	−	−	−	−	+	−	−	−	−	+	−	−	−	−	−	−	−	−	+	−
F1R5R3	23	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	+	−	−	−	−	−	−
DUR 4.4	20	−	−	+	−	−	−	−	−	−	−	+	+	+	+	+	−	−	+	−	−	+	+	+	−
SIR-7 D1	19	+	+	−	+	−	−	+	−	−	−	−	+	+	+	+	−	+	+	+	+	+	−	+	−
SIR-7 G1	19	+	−	−	+	−	−	+	−	−	+	+	−	+	+	−	+	+	+	+	+	+	−	+	−
MOG 13/22	24	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	+	+	−

[1]The origins and details of the initial characterizations of the somatic cell hybrids are described in the references indicated in the second column. The human chromosomal contents were deduced from a combination of karyotypic, antigenic and enzymatic analyses. Karyotypic analysis was done by a combination of G11 staining and quinacrine banding. Analysis of hybrids for c-fms sequences was performed as described in the legend to Fig. 1. Concordance beween the presence or absence of a particular chromosome from a hybrid and c-fms sequences is indicated by the open boxes.

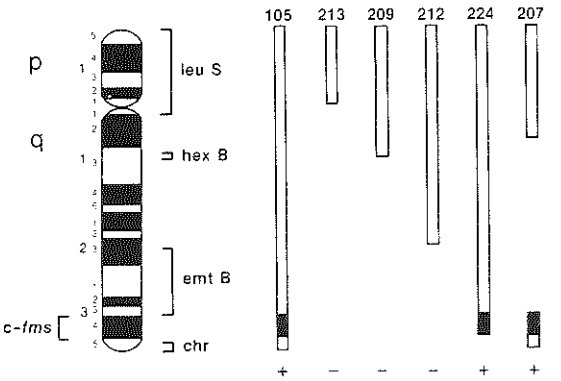


Figure 2. Schematic representation of human chromosome 5 indicating the localization of breakpoints of segregants utilized in the present study and previously reported map positions of leuS, hexB, emtB and chr (16,17). Localization of c-fms is based on the results of the present study. In the right panel of the figure the series of chromosome 5 deletion segregants are shown diagramatically and indicated as c-fms positive (+) or negative (-). Regions common to the three segregants containing human c-fms sequences are shown as solid bars.

following BamHl digestion (Fig. 1, Panel III). In contrast to human DNA, Chinese hamster cellular DNA lacked detectable homology with the c-fms $K_{0.9}$ probe. Hybrid HHW105, containing an intact chromosome 5 as its only human component, reacted strongly with the c-fms $K_{0.9}$ probe, thus confirming the localization of c-fms to human chromosome 5. In contrast, human c-fms sequences were absent from a segregant (HHW213) missing almost the complete q arm of chromosome 5 as well as from segregants HHW209 and HHW212 which have lost approximately 80% and 40% respectively, of the terminal portion of the q arm of chromosome 5. An additional segregant, HHW224, characterized by an apparent terminal deletion of chromosome 5 encompassing at most the q35 terminal band corresponding to the chr marker, was positive for c-fms, thus localizing c-fms between q23 and q34.

In a previous study (16), preliminary characterization of an additional segregant, HHW207, revealed a large interstitial deletion of chromosome 5 encompassing both the hexB and emtB markers. Upon hybridization analysis, this segregant was found to have retained c-fms (Fig. 1, Panel III). In order to more precisely localize c-fms, it was necessary to subject HHW207 to more detailed karyotypic analysis. Upon examination of chromosome preparations from HHW207 under conditions which gave much better banding resolution than was achieved previously we were able to unequivocally localize the interstitial

deletion as encompassing the region of chromosome 5 mapping from q12 to q33. As shown schematically in Fig. 2, these results establish the order of the individual markers on chromosome 5 as leuS, hexB, emtB, c-fms and chr and more specifically localize c-fms to band q34.

DISCUSSION:

In the present study we have utilized a panel of mouse-human somatic cell hybrids to map the human oncogene, c-fms, on chromosome 5. Additionally, the availability of a set of hybrids with well defined deletions in chromosome 5 (16,17) has allowed regional localization of c-fms to band q34. The generation of sets of deletion hybrids, such as those used in the present study, is possible for any human chromosome for which appropriate positive and negative selection markers can be developed. By use of such hybrids in combination with Southern hybridization the localization of specific genetic sequences can be achieved with much less ambiguity than is possible by use of alternative techniques such as in situ hybridization.

The localization of c-fms on chromosome 5, in addition to the previous assignment of other human oncogenes to specific chromosomes, establishes the distribution of the known oncogenes among human chromosomes to be relatively random. In fact, the eight human oncogenes which have been mapped to date are distributed among seven different chromosomes. These include c-fes (20,21), c-myb (21), c-abl (20), c-sis (22,23) and c-ras-H (24) which have been mapped to chromosomes 15, 6, 9, 22 and 11, respectively, c-fms localized in the present study on chromosome 5, and two human oncogenes, c-myc (9,10) and c-mos (25,26) both of which map on chromosome 8.

Evidence for an association of human oncogenes with translocations and deletions specific to various human cancers is accumulating. For instance, c-abl has been localized within a small terminal region of chromosome 9 (20) which is translocated to chromosome 22 in chronic myelogenous leukemia (9,10) while c-sis maps within the reciprocally translocated region of chromosome 22 (27). Similarly, c-myc has been identified near the breakpoint of the segment of chromosome 8 which translocates to chromosomes 2, 14 or 22 in Burkitt's lymphoma (9,10) and c-fes has been mapped within the region of chromosome 15 which is translocated to chromosome 17 in acute promyelocytic leukemia (28). In other studies, c-ras-H has been mapped on chromosome 11, a human chromosome with a characteristic deletion in Wilms tumor (24). The mapping of c-fms on chromosome 5 is of interest in that deletions within the long arm of chromosome 5 are also frequently observed in patients with refractory anemia

(29) and acute myelogenous leukemia (30,31). Such deletions have been reported to appear to be interstitial, involving approximately two thirds of the long arm of chromosome 5, but not extending to band q34 (29,31).

As one possible model, to account for their association with chromosomal rearrangements occurring in specific human cancers, cellular oncogenes, such as c-fms could map adjacent to chromosomal breakpoints possibly resulting in the juxtapositioning of potentially oncogenic sequences with cellular promotors. Such a model has been proposed to account for the translocations associated with chronic myelogenous leukemia and Burkitt's lymphoma (8-10); in fact, in Burkitt's lymphoma the cellular sequences in the chromosome to which c-myc is translocated appear to be involved in immunoglobulin rearrangements (9,10). It will thus be of interest to accurately map the proximity of c-fms to chromosome 5q- deletions associated with acute myelogenous leukemia and to analyze tumors with the 5q- deletion for expression of c-fms gene products. Independent of its potential involvement in neoplasia, the localization of c-fms provides a unique molecular marker for chromosome 5 which should be useful for identifying syntenic linkage associations among diverse mammalian species.

ACKNOWLEDGEMENT:

The authors thank G.T. Blennerhassett and P. Hansen for their excellent technical assistance. This work was supported under Contract No. NOI-CO-76380 from the National Cancer Institute, Bethesda, Maryland, and by Public Health Service Grant No. GM-25339 from the National Institute of General Medical Sciences, Bethesda, Maryland.

REFERENCES:

1. Varmus, H.E. (1982) Science 216, 812-820
2. Bishop, J.M. (1982) Retrovirus and Cancer Genes. In Adv. Cancer Res. 37, 1-32
3. Klein, G. (1982) Advances in Viral Oncology. New York, Raven Press.
4. Parada, L.F., Tabin, C.J., Shih, C. and Weinberg, R.A. (1982) Nature 297,474-478
5. Santos, E., Tronick, S.R., Aaronson, S.A., Pulciani, S., and Barbacid, M. (1982) Nature 298, 343-347
6. Der, C.J., Krontiris, T.G., and Cooper, G.M. (1982) Proc. Natl. Acad. Sci. USA 79, 3637-3640
7. Chang, E.H., Furth, M.E., Scolnick, E.M. and Lowy, D.R. (1982) Nature 297, 479-483
8. de Klein, A. Geurts van Kessel, A., Grosveld, G., Bartram, C.R., Hagemeijer, A., Bootsma, D., Spurr, N., Heisterkamp, N., Groffen, J. and Stephenson, J.R. (1982) Nature 300, 765-767
9. Dalla-Favera, R., Bregni, M., Erikson, J., Patterson, D., Gallo, R.C., Croce, C.M. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 7824-7827
10. Taub, R., Kirsch, I., Morton, C., Lenoir, G., Swan, D., Tronick, S., Aaronson, S., and Leder, P. (1982) Proc. Natl. Acad. Sci. USA 79, 7837-7841
11. Donner, L., Fedele, L.A., Garon, C.F., Anderson, S.J., and Sherr, C.J. (1982) J. Virol. 41, 489-500

12. Heisterkamp, N., Groffen, J., and Stephenson, J.R. (1983) Virology 126, 248-258
13. Van de Ven, W.J.M., Reynolds, F.H., Nalewaik, R., and Stephenson, J.R. (1980) J. Virol. 35, 165-175
14. Barbacid, M., Lauver, A.V., and Devare, S.G. (1980) J. Virol. 33, 196-207
15. Reynolds, F.H. Jr., Van de Ven, W.J.M., Blomberg, J., and Stephenson, J.R. (1981) J. Vriol. 41, 1084-1089
16. Dana, S., and Wasmuth, J.J. (1982) Molec. & Cell Biol. 2, 1220-1228
17. Dana, S., and Wasmuth, J.J. (1982) Somatic Cell Gen. 8, 245-264
18. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517
19. Groffen, J., Heisterkamp, N., Grosveld, F., Van de Ven, W., and Stephenson, J.R. (1982) Science 216, 1136-1138
20. Heisterkamp, N., Groffen, J., Stephenson,J.R., Spurr, N.K., Goodfellow, P.N., Solomon, E., Carritt, B., and Bodmer, W.F. (1982) Nature 299, 747-749
21. Dalla-Favera, R., Franchini, G., Martionotti, S., Wong-Staal, F., Gallo, R.C., and Croce, C.M. (1982) Proc. Natl.Acad. Sci. USA 79, 4714-4717
22. Swan, D.C., McBride, O.W., Robbins, K.C., Keithley, D.A., Reddy, E.P., and Aaronson, S.A. (1982) Proc. Natl.Acad. Sci. USA 79, 4691-4695
23. Dalla-Favera,R., Gallo,R.C., Giallongo,A., and Croce, C.M. (1982) Science 218, 686-688
24. McBride, O.W., Swan, D.C.,Santos, E., Barbacid, M., Tronick, S.R., and Aaronson,S.A. (1982) Nature 300, 773-774
25. Prakash, K., McBride, O.W., Swan, D.C., Devare, S.G., Tronick, S.R., and Aaronson, S.A. (1982) Proc. Natl. Acad.Sci. USA 79, 5210
26. Neel, B.G., Jhanwar, S.C., Chaganti, R.S.K.,and Hayward, W.S. (1982) Proc. Natl. Acad. Sci. USA 79, 7842-7846
27. Groffen, J., Heisterkamp, N., Stephenson, J.R., Geurts van Kessel, A., de Klein, A., Grosveld, G., and Bootsma, D. (1983) J. Expt. Med. 158, 9-15.
28. Sheer, D., Hiorns, L.R., Goodfellow, P.N., Trowsdale, J., Solomon, E., Swallow, D.A., Povey, S., Heisterkamp, N., Groffen, J., and Stephenson, J.R. (1983) Proc. Natl. Acad. Sci. USA (in press)
29. Van den Berghe, H., Cassiman, J.J., David, G., Fryns, J.P., Michaux, J.L., and Sokal, G. (1974) Nature 251, 437-438
30. Sokal, G., Michaux, J.L., Van den Berghe, H., Cordier, A., Rodhain, J., Ferrant, A., Moriau, M., de Bruyere, M., Sonnet, J. (1975) Blood 46, 519-553
31. Van den Berghe, H., David, G., Michaux, J.L., Sokal, G., and Verwilghen, R., (1976) Blood 48, 624-626
32. Whitehead, A.S., Solomon, E., Chambers, S., Bodmer, W.F., Povey, S., and Fey, G. (1982) Proc. Natl. Acad. Sci. USA 79, 5021-5025
33. Solomon, E., Bobrow, M., Goodfellow, P.N., Bodmer, W.F., Swallow, D.M., Povey, S., and Noel, B. (1976) Somat. Cell Genet. 2, 125-140
34. Jones, E.A., Goodfellow, P.N., Kennett, R.H., and Bodmer, W.F. (1976) Somat. Cell Genet. 2, 483-496
35. Solomon, E., Bobrow, M., Goodfellow, P.N., Bodmer, W.F., Swallow, D.M., Povey, S., and Noel, B. (1976) Somat. Cell Genet. 2, 125-140
36. Hobart, M.J., Rabbits, T.H., Goodfellow, P.N., Solomon, E., Chambers, S., Spurr, N., and Povey, S. (1981) Ann. Hum. Genet. 45, 331-335
37. Povey, S., Jeremiah, S.J., Barker, R.F., Hopkinson, D.A., Robson, E., Cook, P..J.L., Solomon, E., Bobrow, M., Carrot, B., and Buckton, K.E., (1980) Ann. Hum. Genet. 43, 241-248
38. Nabholz, M.,Miggiano, V., and Bodmer, W.F. (1969) Nature 223, 358-363
39. Swallow, D.M., Solomon, E., and Pajunen, L. (1977) Cytogenet. Cell Genet. 18, 136-148

APPENDIX PAPER III

Isolation of Human Oncogene Sequences (v-*fes* Homolog) from a Cosmid Library

Abstract. *To define the human homolog (or homologs) of transforming sequences (v-*fes *gene) common to Gardner (GA) and Snyder Theilen (ST) isolates of feline sarcoma virus (FeSV), a representative library of human lung carcinoma DNA in a cosmid vector system was constructed. Three cosmid clones were isolated containing GA/ST FeSV v-*fes *homologous cellular sequences, within 32- to 42-kilobase cellular inserts representing 56 kilobases of contiguous human cellular DNA. Sequences both homologous to, and colinear with, GA or ST FeSV* $_{\gamma}$*-*fes *are distributed discontinuously over a region of up to 9.5 kilobases and contain a minimum of three regions of nonhomology representing probable introns. A thymidine kinase selection system was used to show that, upon transfection to RAT-2 cells, the human c-*fes *sequence lacked detectable transforming activity.*

That RNA transforming viruses contain acquired cellular genes accounting for their capacity to transform cells in culture and induce tumors of various histological classes in vivo is well established (*1*). Although such cellular derived "oncogenes" have been described only in animal model systems (*1*), their existence makes possible the isolation of related human genomic sequences. Extensive noncoding sequences within the cellular homologs of many such viral transforming genes (*2, 3*) represent a major difficulty for their cloning in conventional phage and plasmid systems. Because of this problem and the desirability of obtaining such genes with sufficiently extensive flanking sequences for studies of cellular regulatory controls influencing their expression, we used a cosmid cloning vector (*4–6*) in our study. This system involves the use of plasmids containing lambda cos sequences (cohesive ends) permitting insertion of large DNA fragments, in vitro packaging, and transduction to *Escherichia coli*.

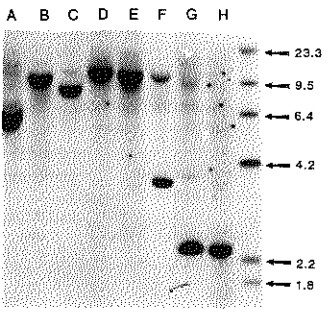


Fig. 1. Restriction patterns of cellular homologs of transformation-specfic DNA sequences common to GA and ST FeSV. (A) High molecular weight DNA's were prepared from kitten lung embryo cells; (B) CCL64 mink lung cells; (C) NIH/3T3 mouse embryo cells; (D and G) normal human lung tissue; and (E, F, and H) human lung carcinoma tissue. DNA samples (20 μg) were digested with Eco RI (A to E), Bam HI (F), or Kpn 1 (G and H), separated by electrophoresis on 0.75 percent agarose gels, transferred to nitrocellulose, and analyzed by hybridization according to previously described procedures (*19*). ^{32}P-Labeled DNA digested with Hind III is included as a molecular weight standard.

To define the cellular homolog of the GA/ST transforming (v-*fes*) gene, we used a molecular probe corresponding to a 0.5-kb Pst I restriction fragment within the ST FeSV (feline sarcoma virus) acquired sequence (v-*fes* S_L) (*2, 7*). By Southern blot analysis (*8*), single bands of hybridization were observed at molecular weights of between 7.0 and 12.5 kb in cat, mink, mouse, and human Eco RI restricted DNA's (Fig. 1, A to E). Restriction of DNA's from both normal human lung and human lung carcinoma with Bam HI or Kpn I resulted in generation of single bands hybridizing at 3.8 kb (Fig. 1F) and 2.4 kb (Fig. 1, G and H), respectively.

A cosmid library was constructed and screened for v-*fes* homologous sequences (Fig. 2). Three clones were initially selected on the basis of hybridization to v-*fes* S_L and were propagated for restriction endonuclease analysis. Two of the clones contained a 12.0-kb Eco RI restriction fragment with homology to both v-*fes* S_L and v-*fes* S_R while only the 9.8-kb 5′ region of this fragment was represented in the third clone. Further restriction enzyme analysis indicated that sequences within these clones were overlapping and represented a 58-kb contiguous region of the human genome. The orientation and positioning of the cellular inserts within these clones, both relative to each other and to the cosmid vector, are summarized in Fig. 3. For purposes of fine structure mapping, the above described 12.0-kb Eco RI v-*fes* homologous restriction fragment was subcloned in plasmid pBR328.

As a prerequisite to further analysis of the above described clones it was first necessary to isolate and prepare molecular probes corresponding to the complete

acquired cellular sequences of both GA and ST FeSV. The ST FeSV genome was cloned in phage Charon 9 as a 6.6-kb fragment of Hind III restricted DNA isolated from the nonproductively transformed mink cell line, ST-FeSV 64 C141. Similarly, GA FeSV was isolated as a 14-kb Eco RI restriction fragment from the mink-transformed clone, G-FeSV 64 F3, in phage λgtWESλB. ^{32}P-Labeled DNA's corresponding to the entire GA and ST FeSV genomes were prepared (Fig. 3).

The results of fine structure mapping of the 12.0-kb Eco RI restriction fragment containing GA and ST FeSV homologous sequences are summarized in Fig. 3. The indicated regions of hybridization are based on identification of restriction fragments hybridizing with each of the four indicated probes. These encompass a total of 4.5 kb of homology distributed discontinuously over a 9.5-kb region. No additional homology with either GA or ST FeSV was observed within the entire 58 kb of human DNA represented within the three cosmid clones. The similarity in arrangement of homologous sequences within the viral and cellular DNA's establishes colinearity between v-*fes* and its human homolog. These findings thus provide a maximum estimate of human DNA sequence homology to the GA and ST FeSV genomes and a minimal estimate of the number of nonhomologous regions representing probable introns.

From a comparison of the results summarized in Fig. 3, to the analysis of total human lung carcinoma DNA (Fig. 1), the major Eco RI, Kpn I, and Bam HI restriction fragments homologous to the v-

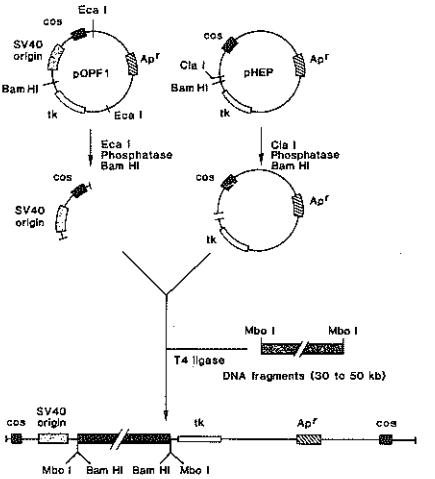


Fig. 2. Construction of a human lung carcinoma DNA library in a cosmid vector system. High molecular weight cellular DNA was purified from human lung carcinoma tissue and partially digested with Mbo I and the fragments were separated by sucrose gradient (5 to 20 percent) fractionation. The DNA's of plasmids pOPF1[19] and pHEP were digested with Eca I and Cla I, respectively, treated with calf intestine phosphatase, and redigested with Bam HI. A 1:1 molar ratio (1.0 μg each) of the restricted plasmid DNA's and restricted cellular DNA fragments (30 to 50 kb) were ligated, packaged in vitro, and transduced into *E. coli* ED8767 (*5, 6*). Ampicillin-selected bacteria containing recombinant DNA were grown on nitrocellulose (Millipore HAWP) at a density of approximately 15,000 colonies per filter. Approximately 250,000 individual colonies were obtained representing three to four times that required for representation of the entire human genome.

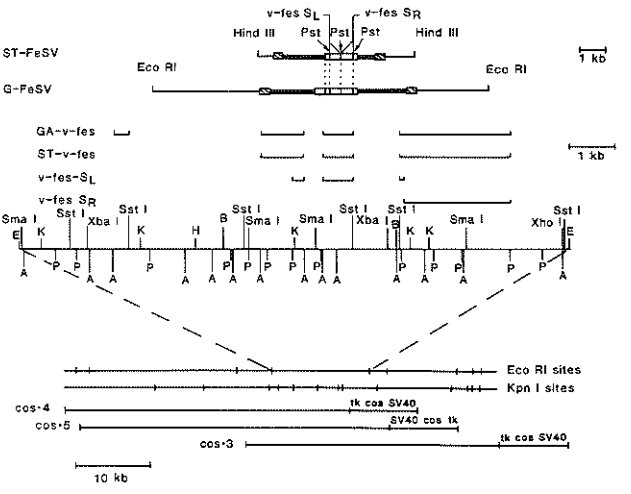


Fig. 3. Restriction map of human cellular sequences homologous to acquired cellular sequences represented within GA and ST FeSV. Open boxes (□) within the viral genomes represent the relative positions of their respective acquired cellular sequences, while feline leukemia virus cross-reactive sequences are shown as solid lines (▬). The positions of long terminal repeats (▨) and cellular flanking sequences (—) are also indicated. The 12.0-kb Eco RI DNA restriction fragment shown in the center of the figure contains the entire human homolog of the GA and ST FeSV acquired cellular sequences. Restriction fragments containing sequences homologous to the complete GA-FeSV genome (GA-v-*fes*), the ST-FeSV genome (ST-v-*fes*), v-*fes* S_L, and v-*fes* S_R are indicated. The lower portion of the figure shows the position of the 12.0-kb Eco RI fragment within a 58-kb contiguous sequence of human cellular DNA; the restriction map of this 58-kb sequence was deduced as a composite of individually mapped overlapping cellular sequences represented within the cosmids shown at the bottom. The relative positions of the thymidine kinase gene (*tk*), cos site, and SV40 restriction fragment within the 9-kb vector are shown for purposes of orientation. Restriction enzymes are abbreviated as follows: Ava I (*A*), Bam HI (*B*), Eco RI (*E*), Kpn I (*K*), Hind III (*H*), and Pst I (*P*).

fes S_L probe are seen to correspond. Similarly, correspondence was observed between v-*fes* S_R homologous restriction fragments as determined by genomic blots of total cellular DNA and the isolated human homolog of the GA/ST v-*fes* gene. The human v-*fes* homolog showed no detectable transforming activity when transfected to RAT-2 cells by means of a thymidine kinase selection system. Molecularly cloned GA/ST FeSV were included as positive controls.

The molecular cloning of human genomic sequences homologous to GA/ST FeSV v-*fes* is of particular interest because (i) the c-*fes* gene is highly conserved and (ii) it is subject to frequent recombination with type C retrovirus sequences resulting in the generation of transforming viruses (*7, 9*). Virus isolates of this class of not only mammalian (feline) but also avian (chicken) origin contain related cellular derived transforming sequences (*10*). A common feature of the major gene product of these recombinant transforming viruses is an associated tyrosine-specific protein kinase (*11–13*). In addition, the GA and ST FeSV gene products exhibit binding affinity for a 150,000 molecular weight cellular phosphoprotein (*12, 14*) and transformation by these viruses leads to abolition of epidermal growth factor binding (*15, 16*) and production of a low molecular weight transforming growth factor (*17*).

The finding of only a single genetic locus exhibiting significant homology with the GA and ST FeSV acquired sequences establishes that the highly related transforming sequences within the genomes of these independently isolated viruses were originally derived from the same cellular gene. Conversely, molecular probes specific for acquired cellular sequences represented within the Abelson MuLV genome, an independent RNA transforming virus with associated protein kinase activity (*11, 18*), recognized different human DNA restriction fragments and lacked detectable homology with sequences represented within any of the three cosmid isolates (data not shown). Thus there must exist at least two independent loci within the human genome homologous to viral genes with associated tyrosine-specific protein kinase activities.

The human carcinoma DNA cosmid library generated in our study provides a specific reagent for the molecular cloning of cellular homologs of viral transforming genes. In particular, the availability of this library should be of value for the isolation of those human transforming genes with extensive intervening sequences. An important feature of this system is the presence of a functionally active thymidine kinase gene that allows for selection of minority populations of eukaryotic cells containing such cosmids after transfection. In addition, SV40 DNA sequences situated in one of the cosmid arms have been shown to exert a positive influence on transcription in the β-globin system (*6*). If this sequence similarly influences expression of cellular homologs of viral transforming genes, its presence may be important for identification of the translational products of these sequences and a determination of their transforming potential.

JOHN GROFFEN
NORA HEISTERKAMP
Laboratory of Viral Carcinogenesis, National Cancer Institute, Frederick, Maryland 21701
FRANK GROSVELD
Laboratory of Gene Structure and Expression, Medical Research, Council, London NW7 1AA
WIM VAN DE VEN
Carcinogenesis and Intramural Program, Frederick Cancer Research Facility, Frederick, Maryland 21701
J. R. STEPHENSON
Laboratory of Viral Carcinogenesis, National Cancer Institute, Frederick, Maryland 21701

References and Notes

1. P. J. Fischinger, in *Molecular Biology of RNA Tumor Viruses*, J. R. Stephenson, Ed. (Academic Press, New York, 1980), p. 162.
2. G. Franchini, J. Even, C. J. Sherr, F. Wong-Staal, *Nature (London)* **290**, 154 (1981).
3. S. P. Goff, E. Gilboa, O. N. Witte, D. Baltimore, *Cell* **22**, 777 (1980); R. Dalla-Favera, E. P. Gelmann, R. C. Gallo, F. Wong-Staal, *Nature (London)* **292**, 31 (1981); R. W. Ellis, D. DeFeo, T. Y. Shih, M. A. Gonda, H. A. Young, N. Tsuchida, D. R. Lowy, E. M. Scolnick, *ibid.*, p. 506; D. Shalloway, A. D. Zelenetz, G. M. Cooper, *Cell* **24**, 531 (1981).
4. D. Horowicz and J. F. Burke, *Nucl. Acids Res.* **9**, 2989 (1981).
5. F. G. Grosveld, H. M. Dahl, E. de Boer, R. A. Flavell, *Gene* **13**, 227 (1981).
6. F. Grosveld, T. Lund, A. Mellor, H. Bud, H. Bullman, R. A. Flavell, *J. Mol. Appl. Genet.*, in press.
7. L. A. Fedele, J. Even, C. F. Garon, L. Donner, C. J. Sherr, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 4036 (1981).
8. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
9. M. A. Barbacid, A. V. Lauver, S. G. Devare, *J. Virol.* **33**, 196 (1980); W. J. M. Van de Ven, A. S. Khan, F. H. Reynolds, Jr., K. T. Mason, J. R. Stephenson, *ibid.* **33**, 1034 (1980).
10. M. Shibuya, T. Hanafusa, H. Hanafusa, J. R. Stephenson, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6536 (1980).
11. W. J. M. Van de Ven, F. H. Reynolds, Jr., J. R. Stephenson, *Virology* **101**, 185 (1980).
12. F. H. Reynolds, Jr., W. J. M. Van de Ven, J. R. Stephenson, *J. Biol. Chem.* **255**, 11040 (1980).
13. M. Barbacid, K. Beemon, S. G. Devare, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5158 (1980).
14. F. H. Reynolds, Jr., W. J. M. Van de Ven, J. R. Stephenson, *Nature (London)* **286**, 409 (1980).
15. F. H. Reynolds, Jr., W. J. M. Van de Ven, J. Blomberg, J. R. Stephenson, *J. Virol.* **37**, 643 (1981).
16. G. J. Todaro, J. E. DeLarco, S. Cohen, *Nature (London)* **264**, 26 (1976).
17. J. R. Stephenson and G. J. Todaro, in *Advances in Viral Oncology*, G. Klein, Ed. (Raven Press, New York, in press), vol. 1.
18. O. N. Witte, A. Dasgupta, D. Baltimore, *Nature (London)* **283**, 826 (1980); J. Blomberg *et al.*, *ibid.* **286**, 504 (1980).
19. R. Bernards and R. A. Flavell, *Nucl. Acids Res.* **8**, 1521 (1980).
20. We thank R. A. Flavell for advice and discussion, C. J. Sherr for providing plasmids containing the S_R v-*fes* and S_L v-*fes* subclones, and G. T. Blennerhassett and P. Hansen for technical assistance. Supported under Public Health Service Contract No. NOI-CO-75380 from the National Cancer Institute, by the British Medical Research Council (MRC), and by a grant (NUKC-BIOCH 80-1) awarded to J.G., N.H., and W.V.D.V. by the Royal Netherlands Cancer Foundation (KWF).

18 January 1982; revised 22 March 1982

APPENDIX PAPER IV

Transforming Genes of Avian (v-*fps*) and Mammalian (v-*fes*) Retroviruses Correspond to a Common Cellular Locus

JOHN GROFFEN,* NORA HEISTERKAMP,* MASABUMI SHIBUYA,† HIDESABURO HANAFUSA,† AND JOHN R. STEPHENSON*,[1]

*Laboratory of Viral Carcinogenesis, National Cancer Institute, Frederick, Maryland 21701, and †The Rockefeller University, 1230 York Avenue, New York, New York 10021

Received August 26, 1982; accepted November 29, 1982

The Gardner (GA) and Snyder-Theilen (ST) isolates of feline sarcoma virus (FeSV) represent genetic recombinants between feline leukemia virus (FeLV) and transformation-specific sequences (v-*fes* gene) of cat cellular origin. A related transforming gene (v-*fps*), common to the Fujinami, PRC II, and UR 1 strains of avian sarcoma virus has also been described. Translational products of each of these recombinant virus isolates are expressed in the form of polyproteins exhibiting protein kinase activities with specificity for tyrosine residues. In the present study, v-*fes* and v-*fps* homologous sequences of GA-FeSV, ST-FeSV, and Fujinami sarcoma virus (FSV) are defined and these independently derived transforming genes are shown to correspond to a common cellular genetic locus which has remained highly conserved throughout vertebrate evolution.

Several independent acute transforming retrovirus isolates have been shown to encode as their major translation products tyrosine-specific protein kinases which are closely associated with their transforming function (*12*). Such viruses can be separated into six classes on the basis of their cellular "oncogenic" sequences. Acquired cellular "oncogenes" and representative virus isolates of each group include: v-*src* (RSV) (*6, 20, 22*), v-*fes* (ST and GA FeSV) (*3, 12–14, 27–29, 33–35*), v-*fps* (FSV, PRCII, UR1) (*10, 18, 23–25, 30, 36*), v-*yes* (Y73, Esh) (*15, 21, 30, 39*), v-*ros* (UR2) (*11, 30*), and v-*abl* (A MuLV) (*5, 16, 26, 38*). Two of these genes (v-*fes* and v-*abl*) are of mammalian cellular origin while the remaining four (v-*src*, v-*fps*, v-*yes*, and v-*ros*) were derived from naturally occurring avian tumors. Earlier studies indicating low levels of sequence homology between the v-*fes* and v-*fps* genes (*31*), and both immunologic (*2*) and structural relatedness (*4*) between their translational products, raised the possibility that these either represent related genes or alternatively that they correspond to a common cellular locus which has remained highly conserved throughout vertebrate evolution. Precedence for the first model was derived from the relatedness of Harvey and Kirsten strains of rat sarcoma virus (*8*). The present study was undertaken to resolve these alternative possibilities.

To determine the extent of relatedness of FSV v-*fps* to the GA and ST FeSV v-*fes* genes, molecular probes corresponding to the complete GA-FeSV and ST FeSV genomes, designated GA-v-*fes* and ST-v-*fes*, respectively, and to 0.5-kb *Pst*I subgenomic fragments of ST-FeSV, designated v-*fes* S_L and v-*fes* S_R (Fig. 1), were utilized (*13, 17*). Each was tested for hybridization by Southern blot analysis to a series of restriction fragments of a previously described (*30*) molecular clone of FSV in λgt WES-λB. As shown in Fig. 1A, the extent of v-*fps* homology with the individual v-*fes* probes varies in a manner colinear with their arrangement in the ST- and GA-FeSV genomes. For instance, homology to v-*fes* S_R was restricted to the *Pst*I–*Sma*I fragment at the 3′ region of FSV v-*fps*, while v-*fes*-S_L lacked homology with the 3′-terminal *Bam*HI–*Sma*I fragment and v-*fes* S_L homology extended much fur-

[1] To whom reprint requests should be addressed.

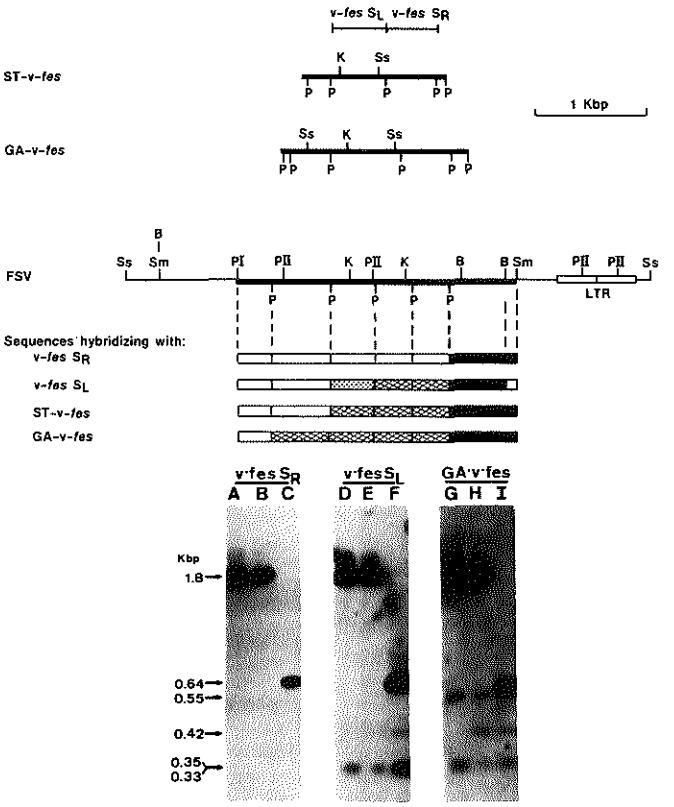


FIG. 1. Homology between FSV-v-*fps* and acquired cellular sequences (v-*fes*) of the Gardner and Snyder-Theilen strains of FeSV. (A) An *Sst*I restriction fragment, representing the entire molecularly cloned FSV genome, was purified from recombinant phage lambda-FSV-1 by gel electrophoresis (*17*). Restriction maps of DNA sequences mapping within the Snyder-Theilen (ST-v-*fes*) and Gardner (GA-v-*fes*) FeSV acquired cellular sequences, shown in the upper portion of the figure, have been previously described (*17*). Following digestion with the indicated restriction enzymes, hybridization was performed using previously described ^{32}P-labeled probes (*13, 17*) including GA-v-*fes*, ST-v-*fes*, v-*fes*-S_L, and v-*fes*-S_R. Extents of hybridization of FSV-v-*fps* with each of the four probes is shown as negative (□); weak (□); intermediate (▩); and strong (■). Restriction endonucleases include *Bam*HI (B); *Kpn*I (K); *Pst*I (P); *Pvu*I (PI); *Pvu*II (PII); *Sma*I (Sm); and *Sst*I (Ss). The 4.7-kb FSV *Sst*I restriction fragment shown in the center of the figure was digested with either *Pst*I and *Kpn*I (A, D, G) or *Pst*I alone (B, E, H), electrophoresed on a 1.0% agarose gel, transferred to nitrocellulose, and hybridized to the indicated probes. In addition, a 2.6-kb *Pvu*I/*Sma*I restriction fragment was isolated from the FSV clone by preparative gel electrophoresis, further digested with *Pst*I, and analyzed for hybridization to the same series of these probes.

ther into the 5′ region of FSV v-*fps*. The most extensive homology was observed with GA-v-*fes* which included all but the extreme 5′ region of v-*fps*. The strength of homology between v-*fes* and v-*fps* was most pronounced at their 3′ ends. Representative hybridization data are shown in the lower panel of Fig. 1B. In contrast to

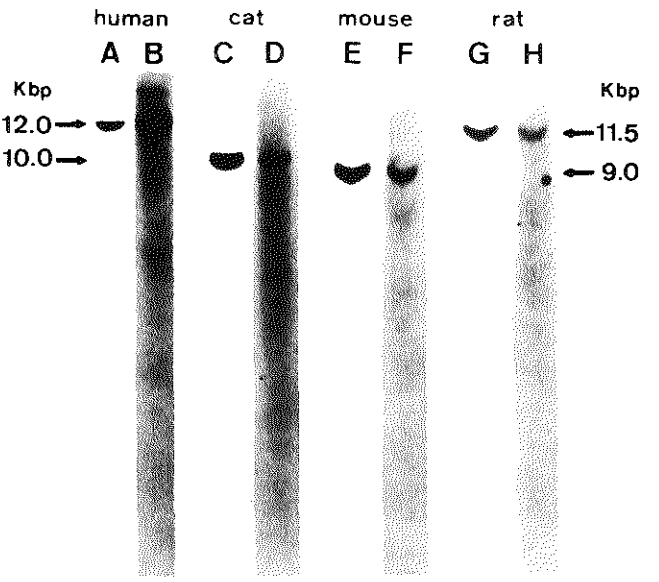


FIG. 2. Comparison of v-*fes* and v-*fps* homologous sequences in human (A, B), cat (C, D), mouse (E, F), and rat (G, H) cellular DNAs. DNAs (20 μg) from cell lines of each species were digested with *Eco*RI, separated by electrophoresis on 0.75% agarose gels, transferred to nitrocellulose, and hybridized to either v-*fes*-S_L (A, C, E, and G) or the 2.6-kb *Pvu*I-*Sma*I FSV-v-*fps* probe (B, D, F, and H).

the homology between the acquired cellular transforming genes of GA-FeSV and FSV, no detectable cross-hybridization was observed between the structural genes of these viruses.

The above results demonstrate extensive homology between v-*fes* and v-*fps* and establish colinearity in their structural organization. To resolve whether these viral transforming sequences correspond to the same, or alternatively, to related but genetically distinct cellular genes, high-molecular-weight DNAs of several mammalian species including human, cat, mouse, and rat were analyzed for sequence homology to v-*fes*- and v-*fps*-specific probes. In each case, single 9.0- to 12.0-kb *Eco*RI DNA restriction fragments hybridized to probes specific for both viral transforming genes (Fig. 2). Although the intensity of hybridization is much greater with the v-*fes* than with the v-*fps* specific probe, no additional v-*fps* homologous restriction fragments were observed. With both control and human lung carcinoma DNA, this fragment is 12.0 kb in length and corresponds to the previously described (*17*) 12.0-kb restriction fragment containing the complete v-*fes* human homolog. The lack of identifiable DNA restriction fragments exhibiting exclusive or preferential hybridization with either v-*fes*- or v-*fps*-specific probes strongly argues for the involvement of a common cellular genetic locus in derivation of viruses containing these genes.

To accurately define regions of sequence homology of the v-*fps* gene with the human cellular homologue of v-*fes*, a series of three cosmid clones containing the above-described 12.0-kb v-*fes* homologous *Eco*RI restriction fragment within a total of 56 kb of contiguous human DNA sequences, were examined (Fig. 3). Hybridization with FSV v-*fps* was restricted to a single 12.0-kb *Eco*RI restriction frag-

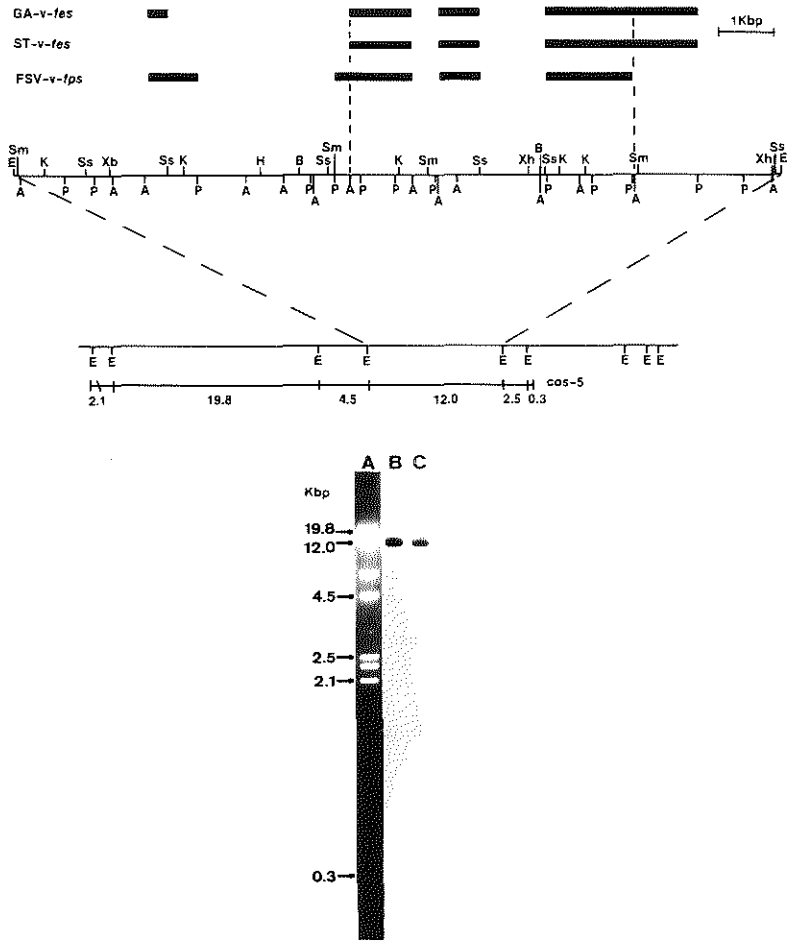


FIG. 3. Comparison of v-*fes* and v-*fps* homologous sequences within the c-*fes*/c-*fps* human genetic locus. The 12.0-kb *Eco*RI restriction fragment shown in the center of the figure contains all detectable v-*fes* and v-*fps* homologous human DNA sequences. Restriction fragments hybridizing to GA-v-*fes*, ST-v-*fes*, and FSV-v-*fps* are indicated as solid bars in the upper portion of the figure. The position of the 12.0-kb *Eco*RI fragment within a 58-kb contiguous sequence of human cellular DNA, represented within three independent overlapping cosmid clones, is indicated by the dotted diagonal lines. *Eco*RI restriction fragments within the inserted cellular sequence of one such cosmid clone (*cos*-5) are designated at the bottom of the figure and the specific hybridization of the 12.0-kb fragment with ^{32}P-labeled GA-v-*fes* (B) and the 2.6-kb *Pvu*I-*Sma*I FSV-v-*fps* (C) probes are shown on the right side. The diagonal line near the 5′ terminus of *cos*-5 represents the junction between the vector arm and inserted cellular sequences. Restriction enzymes are abbreviated as follows: *Ava*I (A); *Bam*HI (B); *Eco*RI (E); *Hind*III (H); *Kpn*I (K); *Pst*I (P); *Sma*I (Sm); *Sst*I (Ss); *Xba*I (Xb); and *Xho*I (Xh).

ment containing v-*fes* homologous sequences. Following subcloning, the 12.0-kb *Eco*RI fragment was digested with the indicated combinations of restriction endonucleases and analyzed for hybridization to v-*fes*- and v-*fps*-specific probes (Fig. 4). FSV v-*fps* homology to human DNA most closely resembled that of GA FeSV v-*fes*. However, the 3′-terminal v-*fes* homologous region of human DNA common to ST- and GA-v-*fes* did not hybridize to a detectable extent with the FSV v-*fps* probe. ST v-*fes* homologous sequences differ from those of both GA v-*fes* and FSV v-*fps* in that they are not represented within the 5′-terminal region hybridizing to the other two. Finally, two regions of human DNA homology unique to FSV-v-*fps* were identified in the 5′ half of the 12.0-kb human *Eco*RI DNA fragment. Thus, although all three viral "oncogenes" appear to be entirely represented by homologous sequences within a single c-*fes*/c-*fps* human genetic locus, the exact positions at which regions of homology map differ among the individual viral transforming genes. The lack of detectable hybridization of v-*fps* with the 3′ region of the c-*fes*/c-*fps* locus could reflect evolutionary divergence. Alternatively, there may exist within the v-*fes*/v-*fps* locus, sequences not necessary for transformation which were lost, either in the generation of FSV and the two FeSV isolates, or by subsequent deletions. Thus a comparison of the 5′-terminal position of ST-v-*fes*, and 3′ position of FSV-v-*fps* homologous sequences within the human c-*fes*/c-*fps* locus, indicated by the dotted lines, may provide a minimal estimate of those sequences required for transformation.

The larger size of the FSV v-*fps* gene (2.6 kb) (*30*) as compared to that of GA-v-*fes* (1.8 kb) (*9*) may be accounted for by regions of the human c-*fes*/c-*fps* locus which are uniquely homologous with FSV v-*fps*. These FSV v-*fps* unique sequences can be explained by one of several alternative possibilities. For instance such sequences may have been initially represented within GA-v-*fes* but deleted during passage of GA-FeSV subsequent to its initial derivation. Alternatively, the differences in sequences in these viruses may reflect differences in the positions of coding and noncoding regions of the c-*fes*/c-*fps* RNA transcripts of the divergent host species from which they were isolated. Finally, mechanisms of incorporation of cellular oncogenic sequences in retroviral ge-

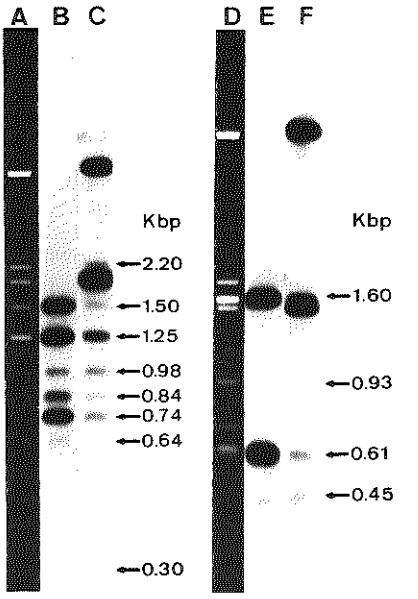


FIG. 4. Restriction endonuclease analysis of human cellular sequences hybridizing to ^{32}P-labeled GA-v-*fes* and FSV-v-*fps* probes. DNA from a variant of pBR328, containing the 12.0-kb human cellular *Eco*RI fragment from *cos*-5 (pBR328 *cos*-5), was digested either with *Ava*I (A–C); or *Pst*I (E–F), separated by electrophoresis on 0.75% agarose gels, and stained with ethidium bromide (A, D). Following transfer to nitrocellulose, hybridization was performed according to previously described (*17*) procedures. ^{32}P-labeled probes include the 2.6-kb *Pvu*I–*Sma*I FSV-v-*fps* fragment (B, E) and GA-v-*fes* (C, F). Restriction fragments which hybridized to either the FSV-v-*fps* or GA-v-*fes* probes and correspond to the 12.0-kb *Eco*RI insert of pBR328 *cos*-5, are indicated by arrows; additional hybridizing restriction fragments in lanes C and F correspond to pBR328 plasmid sequences. The 0.3-kb *Ava*I restriction fragment, although not visible in this figure, was found upon longer exposure to hybridize to both probes.

nomes may be imprecise and the resulting insert within the viral genomic RNA not necessarily correspond exactly to the cellular mRNA from which it was derived. Either the first or third possibilities appear most probable in that one of these must be envoked to explain the differences between the acquired cellular sequences of GA- and ST-v-*fes* (*9*, *17*). Resolution of these alternative models may provide insight into mechanisms by which cellular oncogenes become stabily associated with the genomes of retroviruses.

The present findings establish that the v-*fes* and v-*fps* sequences correspond to a common cellular genetic locus which has been highly conserved throughout vertebrate evolution. Thus, tyrosine-specific protein kinase activities associated with the translational product of such sequences, have probable functional significance to their host. Although the mammalian homologues of other avian retrovirus encoding tyrosine-specific protein kinases have not been characterized, the human homolog of the mouse v-*abl* gene has been shown to be independent of the c-*fes*/c-*fps* locus (*19*). By analysis of appropriate somatic cell hybrids, c-*fes*/c-*fps* has been localized on human chromosome 15 (*7*, *19*).

ACKNOWLEDGMENTS

We thank G. T. Blennerhassett and P. Hansen for excellent technical assistance. This work was supported under Public Health Service Contract NOI-CO-75380 and by Grant CA-14935 from the National Cancer Institute.

REFERENCES

1. BARBACID, M., BEEMON, K., and DEVARE, S. G., *Proc. Nat. Acad. Sci. USA* **77**, 5158-5162 (1980).
2. BARBACID, M., BREITMAN, M. L., LAUVER, A. V., LONG, L. K., and VOGT, P. K., *Virology* **110**, 411-419 (1981).
3. BARBACID, M. A., LAUVER, A. V., and DEVARE, S. G., *J. Virol.* **33**, 196-207 (1980).
4. BEEMON, K., *Cell* **24**, 145-153 (1981).
5. BLOMBERG, J., REYNOLDS, F. H. JR., VAN DE VEN, W. J. M., and STEPHENSON, J. R., *Nature (London)* **286**, 504-507 (1980).
6. COLLETT, M. S., and ERIKSON, R. L., *Proc. Nat. Acad. Sci. USA* **75**, 2021-2024 (1978).
7. DALLA-FAVERA, R., FRANCHINI, G., MARTINOTTI, S., WONG-STAAL, F., GALLO, R. C., and CROCE, C. M., *Proc. Nat. Acad. Sci. USA* **79**, 4714-4717 (1982).
8. ELLIS, R. W., DEFEO, D., SHIH, T. Y., GONDA, M. A., YOUNG, H. A., TSUCHIDA, N., LOWY, D. R., and SCOLNICK, E. M., *Nature (London)* **292**, 506-511 (1981).
9. FEDELE, L. A., EVEN, J., GARON, C. F., DONNER, L., and SHERR, C. J., *Proc. Nat. Acad. Sci. USA* **78**, 4036-4040 (1981).
10. FELDMAN, R. A., HANAFUSA, T., and HANAFUSA, H., *Cell* **22**, 757-765 (1980).
11. FELDMAN, R. A., WANG, L.-H., HANAFUSA, H., and BALDUZZI, P. C., *J. Virol.* **42**, 228-236 (1982).
12. FISCHINGER, P. J., *In* "Molecular Biology of RNA Tumor Viruses" (J. R. Stephenson, ed.), pp. 162-198. Academic Press, New York, 1980.
13. FRANCHINI, G., EVEN, J., SHERR, C. J., and WONG-STAAL, F., *Nature (London)* **290**, 154-157 (1981).
14. FRANKEL, A. E., GILBERT, J. H., PORZIG, K. J., SCOLNICK, E. M., and AARONSON, S. A., *J. Virol.* **30**, 821-827 (1979).
15. GHYSDAEL, J., NEIL, J. C., and VOGT, P. K., *Proc. Nat. Acad. Sci. USA* **768**, 2611-2615 (1981).
16. GOFF, S. P., GILBOA, E., WITTE, O. N., and BALTIMORE, D., *Cell* **22**, 777-785 (1980).
17. GROFFEN, J., HEISTERKAMP, N., GROSVELD, F., VAN DE VEN, W. J. M., and STEPHENSON, J. R., *Science* **216**, 1136-1138 (1982).
18. HANAFUSA, T., WANG, L.-H., ANDERSON, S. M., KARESS, R. E., HAYWARD, W. S., and HANAFUSA, H., *Proc. Nat. Acad. Sci. USA* **77**, 3009-3013 (1980).
19. HEISTERKAMP, N., GROFFEN, J., STEPHENSON, J. R., SPURR, N. K., GOODFELLOW, P. N., SOLOMON, E., CARRITT, B., and BODMER, W. F., *Nature (London)* **299**, 747-749 (1982).
20. HUNTER, T., and SEFTON, B. M., *Proc. Nat. Acad. Sci. USA* **77**, 1311-1315 (1980).
21. KAWAI, S., YOSHIDA, M., SEGAWA, K., SUGIYAMA, H., ISHIZAKI, R., and TOYOSHIMA, K., *Proc. Nat. Acad. Sci. USA* **77**, 6199-6203 (1980).
22. LEVINSON, A. D., OPPERMANN, H., LEVINTOW, L., VARMUS, H. E., and BISHOP, J. M., *Cell* **15**, 561-572 (1978).
23. NEIL, J. C., BREITMAN, M. L., and VOGT, P. K., *Virology* **108**, 98-110 (1981).
24. NEIL, J. C., GHYSDAEL, J., and VOGT, P. K., *Virology* **109**, 223-228 (1981).
25. PAWSON, T., GUYDEN, J., KUNG, T.-H., RADKE, K., GILMORE, T., and MARTIN, G. S., *Cell* **22**, 767-775 (1980).
26. REYNOLDS, F. H., JR., SACKS, T. L., DEOBAGKAR, D. N., and STEPHENSON, J. R., *Proc. Nat. Acad. Sci. USA* **75**, 3974-3978 (1978).
27. REYNOLDS, F. H., JR., VAN DE VEN, W. J. M., BLOM-

BERG, J., and STEPHENSON, J. R., *J. Virol.* **37**, 643–653 (1981).

28. REYNOLDS, F. H., JR., VAN DE VEN, W. J. M., and STEPHENSON, J. R., *J. Biol. Chem.* **255**, 11040–11047 (1980).

29. SHERR, C. J., SEN, A., TODARO, G. J., SLIKSI, A. H., and ESSEX, M., *Proc. Nat. Acad. Sci. USA* **75**, 1505–1509 (1978).

30. SHIBUYA, M., HANAFUSA, H., and BLADUZZI, P. C., *J. Virol.* **42**, 143–152 (1982).

31. SHIBUYA, M., HANAFUSA, T., HANAFUSA, H., and STEPHENSON, J. R., *Proc. Nat. Acad. Sci. USA* **77**, 6536–6540 (1980).

32. SHIBUYA, M., WANG, L.-W., and HANAFUSA, H., *J. Virol.*, in press.

33. STEPHENSON, J. R., KHAN, A. S., SLISKI, A. H., and ESSEX, M., *Proc. Nat. Acad. Sci. USA* **74**, 5608–5612 (1977).

34. VAN DE VEN, W. J. M., KHAN, A. S., REYNOLDS, F. H., JR., MASON, K. T., and STEPHENSON, J. R., *J. Virol.* **33**, 1034–1045 (1980).

35. VAN DE VEN, W. J. M., REYNOLDS, F. H., JR., and STEPHENSON, J. R., *Virology* **101**, 185–197 (1980).

36. WANG, L.-H., FELDMAN, R. A., SHIBUYA, M., HANAFUSA, H., NOTTER, M. F. D., and BALDUZZI, P. C., *J. Virol.* **40**, 258–267 (1981).

37. WITTE, O. N., DASGUPTA, A., and BALTIMORE, D., *Nature (London)* **283**, 826–831 (1980).

38. WITTE, O. N., ROSENBERG, N., PASKIND, M., SHIELDS, A., and BALTIMORE, D., *Proc. Nat. Acad. Sci. USA* **75**, 2488–2492 (1978).

39. YOSHIDA, M., KAWAI, S., and TOYOSHIMA, K., *Nature (London)* **287**, 653–654 (1980).

APPENDIX PAPER V

Chromosomal localization of human cellular homologues of two viral oncogenes

Nora Heisterkamp*, John Groffen*, John R. Stephenson*, N. K. Spurr†, P. N. Goodfellow†, E. Solomon†, B. Carritt‡ & W. F. Bodmer†

* Laboratory of Viral Carcinogenesis, National Cancer Institute–FCRF, Frederick, Maryland 21701, USA
† Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK
‡ MRC Human Biochemical Genetics Unit, Galton Laboratory, University College, London WC2, UK

Acute transforming RNA tumour viruses represent genetic recombinants between type C retroviral sequences and transformation-specific sequences of cellular origin. Of the known mammalian cell-derived transforming genes (oncogenes), two have been shown to encode proteins with either intrinsic or highly associated tyrosine-specific protein kinase activity. One such gene, *c-fes*, is of cat cellular origin, while the second, *c-abl*, was derived from mouse cellular sequences. We now show that the human equivalents of *c-fes* and *c-abl* are localized on human chromosomes 15 and 9, respectively. These findings exclude the possibility that these transformation-related genes are clustered at a single locus within the human genome. It is of interest that both of these chromosomes are involved in specific rearrangements found in certain forms of human cancer.

Two mammalian and four avian cell-derived viral oncogenes encoding proteins with tyrosine-specific protein kinase activity have been described[1]. One of these, *c-fes*, is represented within the genomes of two independent isolates of feline sarcoma virus (FeSV), Gardner and Snyder–Theilen, both of which represent genetic recombinants between feline leukaemia virus and cellular (*c-fes*) sequences of cat origin[2,3]. *c-fes* is closely related to *c-fps*, a transforming gene common to several independent avian transforming retrovirus isolates[4,5]. The second cell-derived viral oncogene of this class, *v-abl*, has so far been identified as a transforming component only of the Abelson strain of murine leukaemia virus (MuLV). This virus is a genetic recombinant between Moloney MuLV and mouse cellular (*c-abl*) sequences[6]. The primary translational products of these two cell-derived viral oncogenes are polyproteins containing amino-terminal structural proteins encoded by the *gag* gene, covalently linked to acquired cellular sequence-encoded components[7–10]. Tyrosine-specific protein kinase activities associated with such polyproteins[11–15] have been implicated in transformation through the analysis of transformation-defective viral mutants[16–18]. Recently, transforming sequences isolated from a human bladder carcinoma cell line have been shown to exhibit homology with acquired transforming (oncogenic) sequences of the Harvey strain of murine sarcoma virus[19,20]. Because of the potential involvement of the cellular homologues of viral oncogenes in human cancer, we decided to investigate the human genetics of these sequences and to test the possibility that chromosomes containing *c-fes* and *c-abl* might correspond to those with known translocations or other rearrangements associated with certain human cancers.

To determine the chromosomal location of the human cellular homologue of *v-fes*, we initially screened restriction endonuclease-digested cellular DNAs from a series of mouse–human somatic cell hybrids by molecular hybridization. These hybrids, which contained a full complement of mouse chromosomes, but retained limited numbers of different combinations of human chromosomes, were analysed for the presence of the human *v-fes*-cross-reactive *Eco*RI restriction fragment. Identification of the human chromosomes present in the hybrids was based on isoenzyme, antigenic and chromosome analysis (Table 1). For identification of *c-fes* sequences, we used a ^{32}P-labelled probe corresponding to a 0.5-kilobase (kb) *Pst*I restriction fragment localized within acquired sequences of the Snyder–Theilen FeSV genome, designated *v-fes* $P_{0.5}$. This probe has been described previously[2,21] and is shown diagrammatically in Fig. 1. A single 12.0-kb *Eco*RI restriction fragment of human DNA could be distinguished from the major 8.5-kb *v-fes*-cross-hybidizing *Eco*RI fragment in mouse cellular DNA which migrated at a position close to that of a much less cross-reactive restriction fragment of mouse DNA at ~11.5 kb (Fig. 1). In preliminary experiments using this probe, the human *c-fes* was identified only in those hybrid clones containing human chromosome 15 (Table 1).

To obtain a molecular probe having increased specificity for the human *c-fes* gene, two 0.5-kb *Kpn*I human restriction fragments (*c-fes* $K_{0.5}$) were isolated from a previously described cosmid clone containing the complete 12.0-kb *c-fes* *Eco*RI restriction fragment. Use of these restriction fragments as molecular probes considerably improved detection of the human *c-fes* 12.0-kb *Eco*RI restriction fragment in DNA both from human cells and from mouse–human somatic cell hybrid clones (Fig. 1, lanes *c–h*). As shown for representative hybrids in Fig. 1 and summarized in Table 1, the 12.0-kb *Eco*RI *c-fes*-homologous restriction fragment identified in this way was again observed only in DNA isolated from hybrid clones containing human chromosome 15.

Particularly compelling was the positive reaction with Hor19D2 hybrid DNA with the 0.5-kb *Kpn*I human *c-fes* probe (Fig. 1, lane *e*). This hybrid contains human chromosomes 11, 15 and X as its only human contribution. Segregants of the Hor19D2 containing only the X chromosome (Hor19X) (lane *h*) or only chromosome 11 (Hor1 9D2RI; lane *f*) failed to react with the probe. The segregant Horl-I which retained human chromosome 15 and a human–mouse translocation which includes part of the long arm of chromosome 11, also reacted strongly with the *c-fes* $K_{0.5}$ (lane *g*) and *v-fes* $P_{0.5}$ (Table 1) probes. The results of further restriction endonuclease analysis of DNA from segregant Horl-I are shown in the right panel of the upper gel in Fig. 1. Two major *Bam*HI restriction fragments of the human *c-fes* gene which hybridize with the *c-fes* $K_{0.5}$ probe (lane *b*) are both identified in segregant Horl-I. Similarly, two *Sst* fragments of 3.6 and 7.5 kb were identified in human DNA (lane *b*) and in the Horl-I segregant (lane *c*), but not in DNA of mouse (lane *a*) origin.

An analogous approach was pursued to establish the chromosomal localization of the human cellular homologue of *v-abl*. In this case *Eco*RI-restricted DNAs from a similar series of mouse–human somatic cell hybrids were analysed. The molecular probes used included a *Bam*HI ($B_{3.4}$) restriction fragment containing all but the 3'-terminal portion of the Abelson MuLV genome and a 1.7-kb *Pst*I ($P_{1.7}$) restriction fragment localized within the Abelson MuLV-acquired cellular (*v-abl*) sequences (Fig. 1). A total of seven *Eco*RI restriction fragments of human DNA, ranging in size from 2.5 to 12.0 kb, and several of mouse DNA, were identified, each of which hybridized to the $P_{3.5}$ probe. The five lower molecular weight human fragments were completely separated from mouse *c-abl* *Eco*RI restriction fragments, thus providing a means of identifying hybrid clones containing the human *c-abl* gene. The results of a typical analysis are shown in the lower left panel of Fig. 1 (lanes *a–c*). In other studies the complete human *v-abl* cellular homologue has been cloned from a cosmid library and at least six of the seven *Eco*RI restriction fragments map to a single position. As an independent means of identifying hybrids containing the human *c-abl* gene, a 0.6-kb *Bam*HI probe was prepared from the second lowest molecular weight (2.9 kb) *Eco*RI human restriction fragment and used as a probe. Representative data obtained with this probe are shown in the right panel of Fig. 1.

As summarized in Table 1, complete concordance was

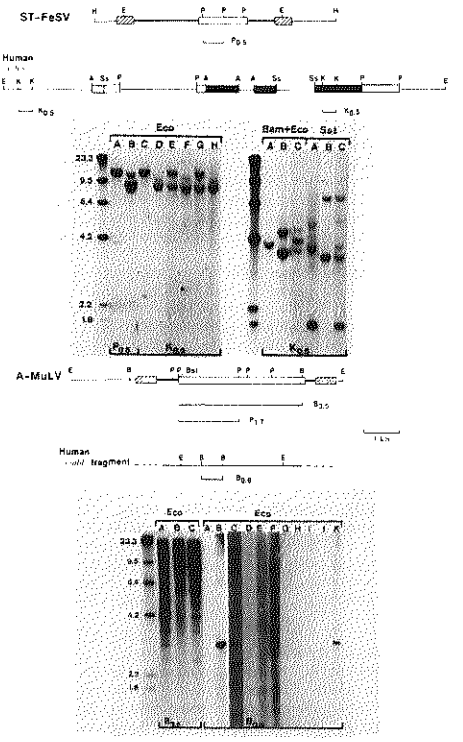


Fig. 1 Identification of human chromosomes containing *v-fes*- and *v-abl*-homologous sequences. Molecular clones of the ST-FeSV and human *c-fes* sequences, shown diagrammatically at the top of the figure, are based on previously published data[3]. The open box (□) within the viral genome represents the position of the acquired cellular sequences, while the feline leukaemia virus cross-reactive sequences are shown as solid lines (—). The positions of the long terminal repeats (▨), cellular flanking sequences (⊔) and the 0.5-kb *Pst*I restriction fragment used as a probe ($P_{0.5}$) are also shown. The organization of the human *c-fes* has been published previously[5]. For reference, the positions of previously identified restriction fragments of the human *c-fes* gene with homology to ST-FeSV, GA-FeSV and the Fujinami strain of avian sarcoma virus, as well as two 0.5-kb *Kpn*I restriction fragments which were co-purified and used as a probe ($K_{0.5}$), are shown. Restriction patterns of cellular sequences hybridizing to the *v-fes* and *c-fes* probes are shown in the uppermost gel. Cell lines from which DNAs were prepared are described in detail in Table 1 legend and include in the left panel: *b*, *d*, NIH/3T3(−) mouse; *a*, *c*, A673 human(+)[40] and mouse–human hybrids; *e*, Horl9D2(+); *f*, Horl9D2R1(−); *g*, Horl1(+); *h*, Horl9x(−); in right panel: *a*, NIH/3T3(−) mouse; *b*, A673 human(+); *c*, mouse–human hybrid Horl1(+). (+ Indicates a positive assignment; − indicates absence of the particular chromosome.) Molecular probes and restriction enzymes are as indicated. Molecular clones of the Abelson MuLV (A-MuLV) genome and a region of human *v-abl*-homologous sequences are shown diagrammatically in the centre of the figure. Restriction fragments used as hybridization probes, including 1.7-kb *Pst*I (*v-abl* $P_{1.7}$) and 3.5-kb *Bam*HI (*v-abl* $B_{3.5}$) restriction fragments as well as a 0.6 kb *Bam*HI (*c-abl* $B_{0.6}$) restriction fragment mapping within the 2.9-kb *Eco*RI human *c-abl* restriction fragment, are indicated. Restriction patterns of cellular DNA sequences hybridizing to the *v-abl* $B_{3.5}$ and *c-abl* $B_{0.6}$ probes are shown in the lower portion of the figure. Cell lines shown in the left panel include: *a*, A673 human(+); *b*, NIH/3T3 mouse(−); *c*, mouse–human hybrid F4sc13C112(+). Those in the right panel include: *a*, NIH/3T3 mouse(−); *b*, A673 human(+); mouse–human hybrids: *c*, F4sc13C112(+); *d*, FIR5R3(−); *e*, F4sc13C19(−); *f*, F4sc13C112(+); *g*, hamster: hamster–human hybrids: *h*, D1S20(−); *i*, C10b2BU(−); *j*, C4al(−); *k*, C10b(+). The derivations of these cell lines and their complements of human chromosomes are described in detail in Table 1 legend. Following hybridization with the $P_{0.5}$, $B_{3.5}$ and $P_{1.7}$ probes, filters were washed with 2.5×SSC at 65 °C; using $K_{0.5}$ and $B_{0.6}$ probes, washes were performed using 0.3×SSC at 65 °C according to the method of Bernards and Flavell[41]. Restriction enzymes are abbreviated as follows: E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; K, *Kpn*I; A, *Ava*I; Ss, *Sst*I. *Hin*dIII-digested λ DNA is included as a molecular weight standard.

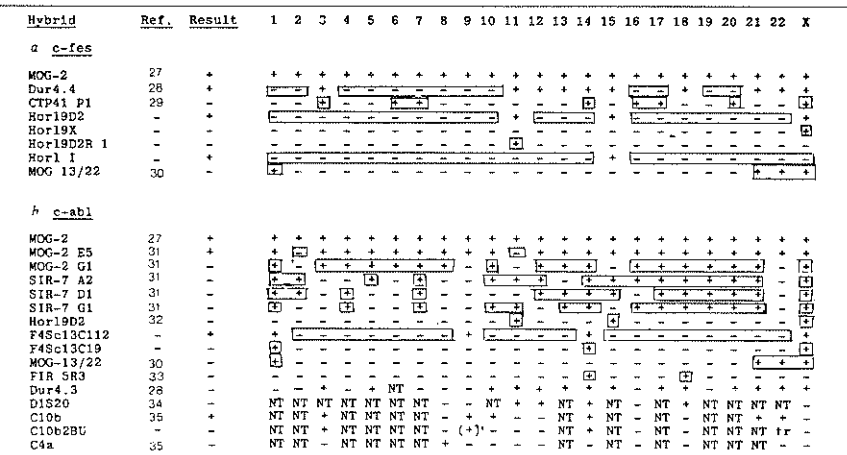
observed between the presence of the five *v-abl* cross-reactive *Eco*RI restriction fragments in the 2.5–5.0-kb size range and human chromosome 9, while each of the other chromosomes could be excluded by one or more examples of discordance. For example, human *c-abl*-specific restriction fragments were present in hybrid F4Sc13 Cl12 which contained only four human chromosomes: 1, 9, 14 and X. Of these, all but chromosome 9 were present in at least one hybrid which lacked the human *c-abl* gene, thus by elimination identifying chromosome 9 as the location of human *c-abl*. Finally, a series of four previously described hamster–human somatic cell hybrids were analysed for human *c-abl* sequences. As summarized in Table 1 and shown in the lower portion of Fig. 1 (lanes *h*–*k*), only one clone, C10b, contained human *c-abl*-specific sequences. Interestingly, this hybrid was positive for each of three isoenzyme markers for human chromosome 9, cytoplasmic soluble *cis*-aconitase (ACONs), AK3 and AK1. A subclone of C10b, designated C10b2BU, lacked detectable human *c-abl* sequences and by isoenzyme analysis was positive for only ACONs and AK3, both of which are localized on the short arm of chromosome 9. The third marker, AK1, which was not detected in this hybrid, is in contrast, localized to the long arm of chromosome 9. The remaining two hybrids were nonreactive in enzyme assays for all three markers for chromosome 9 and lacked detectable sequences hybridizing to the human *c-abl* $B_{0.6}$ probe.

These findings establish the independent chromosomal localization of human cellular homologues of two viral transforming genes, both of which encode tyrosine-specific protein kinases. The possibility that these functionally similar genes are all organized in a cluster such as that observed for the γ-globin, immunoglobulin or histocompatibility gene families is thus ruled out. Although the exact number of human genes encoding proteins with tyrosine-specific protein kinase activity remains to be determined, transforming proteins encoded by each of at least four independent avian type C viral oncogenes with such activity have been identified[1]. One of these, *v-fps*, is structurally related to and maps at the same human cellular locus as the *c-fes* gene analysed here[4,5]. Further study will be required to determine whether human cellular homologues of any of the three other avian viral oncogenes map to the same positions as either *c-fes* or *c-abl*.

In other studies, the mouse *c-abl* gene has been assigned to mouse chromosome 2 (S. P. Goff, P. D'Eustachio, F. Ruddle and D. Baltimore, personal communication). It is interesting that the human homologues of several linked markers assigned to mouse chromosome 2 are found on the long arm of human chromosome 9, indicating that these are, at least in part, homologous linkage groups[22]. The present results obtained from analysis of the hamster–human hybrids suggest that *c-abl* could be on the long arm of human chromosome 9.

Our findings are of interest in that both *c-fes* and *c-abl* map on chromosomes which are frequently involved in specific translocations associated with human neoplasms[23,24]. For example, a high percentage of human chronic myelogenous leukaemias (CML) involve translocations of the distal portion of the long arm of chromosome 22 to chromosome 9. The resulting abnormal form of chromosome 22 has been called the Philadelphia chromosome. The long arm of human chromosome 9 which is specifically involved in this rearrangement is the region of

Table 1 Hybridization of DNA from human–mouse and human–hamster somatic cell hybrids with *c-abl*- and *c-fes*-specific probes



The origin and details of the initial characterization of the somatic cell hybrids listed in column 1 are described in the references indicated in the second column. Some of these hybrids have been subcloned after the initial characterization. (C10b2BU is bromodeoxyuridine-selected subclone isolated from C10b. F4Sc13 C19 and F4Sc13 C112 were subclones, the original cells were a gift from Dr H. Koprowski. These hybrids were isolated from a fusion of human fibroblasts and RAG mouse cells. Hor19D2 is a clone or Horl9 (ref. 32) with an identical karyotype. Hor19D2R1, Horll and Hor19X are subclones of Hor19D2.) Four of the hybrids described, including D1S20, C10b2BU, C4a and C10b, represent human–hamster somatic cell hybrids; all remaining are human–mouse hybrids. The chromosomal content was deduced from a combination of karyotypic, antigenic and enzymatic analyses. karyotypic analysis was done by a combination of G11 staining and quinacrine banding[36,37]. Techniques and markers for antigenic analyses are reviewed in ref. 38. Assays for human enzymes were performed using standard methods[39]. Analysis of hybrids for *c-fes* and *c-abl* sequences was performed as described in Fig. 1 legend using *v-fes* $B_{0.6}$ and *c-fes* $K_{0.5}$ probes in *a* and *v-abl* $B_{3.5}$ and *c-abl* $B_{0.6}$ probes in *b*. Disconcordance between the presence of any one chromosome and hybridization with probes for either locus is shown by open boxes. Karyotypic analysis of F4Sc13C112 indicates that only about 4% of the cells contain chromosome 1. NT, not tested.

human chromosome 9 that appears to correspond to mouse chromosome 2. Similarly, a translocation involving the long arm of chromosome 15 has been observed in cells of patients with acute promyelocytic leukaemia (APL)[23,24]. Abelson MuLV was initially derived from a spontaneously occurring lymphosarcoma of the mouse and induces lymphoid tumours *in vivo* inoculation[25]. Although *c-fes* was derived from a fibrosarcoma of the cat[25], the possibility that its disease spectrum in natural conditions and in the human population may be very different from that as a recombinant virus must be considered. It will thus be of interest to examine levels of expression of *c-fes* and *c-abl* in tumour cells with translocations involving chromosomes 15 and 9, respectively, and to establish by *in situ* annealing techniques whether these genes are, in fact, near the specific translocations points found in chromosome 9 for CML and in chromosome 15 for APL.

Since this work was completed, Dalla-Favera *et al.*[26] have assigned *c-fes* to chromosome 15.

We thank Gail T. Blennerhassett and Pam Hansen for technical assistance, Karina Stanley for karyotype analysis, Sue Povey, Dallas Swallow and Mohamed Parker at the Galton Laboratories for enzyme analyses and Alan Tunnacliffe for deriving the clone HORLI. This work was supported under PHS contract NOI-CO-75380 from the NCI.

Received 2 July; accepted 25 August 1982.

1. Coffin, J. M. *et al. J. Virol.* **40**, 953–957 (1981).
2. Fedele, L. A., Even, J., Garon, C. F., Donner, L. & Sherr, C. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4036–4040 (1981).
3. Groffen, J., Heisterkamp, N., Grosveld, F., Van de Ven, W. J. M. & Stephenson, J. R. *Science* **216**, 1136–1138 (1982).
4. Shibuya, M., Hanafusa, T., Hanafusa, J. & Stephenson, J. R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6536–6540 (1980).
5. Groffen, J., Heisterkamp, N., Shibuya, M., Hanafusa, H. & Stephenson, J. R. *Virology* (submitted).
6. Goff, S. P., Gilboa, E., Witte, O. N. & Baltimore, D. *Cell* **22**, 777–785 (1980).
7. Stephenson, J. R., Kham, A. S., Sliski, A. H. & Essex, M. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5608–5612 (1977).
8. Sherr, C. J., Todaro, G. J., Sliski, A. H. & Essex, M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1505–1509 (1978).
9. Witte, O. N., Rosenberg, N., Paskind, M., Shields, A. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2488–2492 (1978).
10. Reynolds, F. H. Jr, Sacks, T. L., Deobagkar, D. N. & Stephenson, J. R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3974–3978 (1978).
11. Van de Ven, W. J. M., Reynolds, F. H. Jr & Stephenson, J. R. *Virology* **101**, 185–197 (1980).
12. Witte, O. N., Dasgupta, A. & Baltimore, D. *Nature* **283**, 826–831 (1980).
13. Reynolds, F. H. Jr, Van de Ven, W. J. M. & Stephenson, J. R. *J. biol. Chem.* **255**, 11040–11047 (1980).
14. Barbacid, M., Beemon, K. & Devare, S. G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5158–5162 (1980).
15. Sefton, B. M., Hunter, T. & Raschke, W. C. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1552–1556 (1981).
16. Reynolds, F. H. Jr., Van de Ven, W. J. M. & Stephenson, J. R. *J. Virol.* **36**, 374–386 (1980).
17. Witte, O. N., Goff, S., Rosenburg, N. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4993–4997 (1980).
18. Reynolds, F. H. Jr, Van de Ven, W. J. M., Blomberg, J. & Stephenson, J. R. *J. Virol.* **37**, 643–653 (1981).
19. Der, C. J., Krontiris, T. C. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3637–3640 (1982).
20. Parada, L. F., Tabin, C., Shih, C. & Weinberg, R. A. *Nature* **297**, 474–478 (1982).
21. Franchini, G., Even, J., Sherr, C. J. & Wong-Stall, F. *Nature* **290**, 154–157 (1981).
22. O'Brien, S. & Nash, W. *Science* **216**, 257–265 (1982).
23. Rowley, J. D. *A. Rev. Genet.* **14**, 17–39 (1980).
24. Sandberg, A. A. *The Chromosomes in Human Cancer and Leukaemia* (Elsevier, New York, 1980).
25. Fischinger, P. J. in *Molecular Biology of RNA Tumor Viruses* (ed. Stephenson, J. R.) 162–198 (Academic, New York, 1980).
26. Dalla-Favera, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 4714–4717 (1982).
27. Solomon, E., Swallow, D., Burgess, S. & Evans, L. *Ann. hum. Genet.* **42**, 273–281 (1979).
28. Solomon, E. *et al. Somatic Cell Genet.* **2**, 125–140 (1976).
29. Jones, E. A., Goodfellow, P. N., Kennett, R. H. & Bodmer, W. F. *Somatic Cell Genet.* **2**, 483–496 (1976).
30. Povey, S. *et al. Ann. hum. Genet.* **43**, 241–248 (1980).
31. Whitehead, A. S. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 5021–5025 (1982).
32. Goodfellow, P. N. *et al. Nature* **254**, 267–269 (1975).
33. Hobart, M. J. *et al. Ann. hum. Genet.* **45**, 331–335 (1981).
34. Carritt, B. & Povey, S. *Cytogenet. Cell Genet.* **23**, 171–181 (1979).
35. Carritt, B. in *Tissue Culture in Medical Research* Vol. 2 (eds Richards, R. J. & Rajan, K. T.) 169–175 (Pergamon, Oxford, 1980).
36. Bobrow, M. & Cross, I. *Nature* **251**, 74–79 (1974).
37. Casperson, T., Lomakka, G. & Zech, L. *Hereditas* **67**, 89–102 (1971).
38. Goodfellow, P. N. & Solomon, E. in *Monoclonal Antibodies in Clinical Medicine* (eds McMichael, A. & Folve, J.) 365–393 (Academic, London, 1982).
39. Harris, H. & Hopkinson, D. A. *Handbook of Enzyme Electrophoresis in Human Genetics* (North-Holland, Amsterdam, 1976).
40. Giard, D. J. *et al. J. natn. Cancer Inst.* **51**, 1417–1423 (1973).
41. Bernards, R. & Flavell, R. A. *Nucleic Acids Res.* **8**, 1521–1533 (1980).

APPENDIX PAPER VI

Genetic analysis of the 15;17 chromosome translocation associated with acute promyelocytic leukemia

(somatic cell hybrids/c-*fes*/localization of breakpoints/gene mapping)

DENISE SHEER*, LYNNE R. HIORNS*, KARINA F. STANLEY*, PETER N. GOODFELLOW*, DALLAS M. SWALLOW†, SUSAN POVEY†, NORA HEISTERKAMP‡, JOHN GROFFEN‡, JOHN R. STEPHENSON‡, AND ELLEN SOLOMON*

*Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, United Kingdom; †Medical Research Council Human Biochemical Genetics Unit, The Galton Laboratory, University College, London WC2, United Kingdom; and ‡Laboratory of Viral Carcinogenesis, National Cancer Institute, Frederick, Maryland 21701

Communicated by Walter F. Bodmer, April 20, 1983

ABSTRACT **Somatic cell hybrids have been constructed between a thymidine kinase-deficient mouse cell line and blood leukocytes from a patient with acute promyelocytic leukemia showing the 15q+;17q− chromosome translocation frequently associated with this disease. One hybrid contains the 15q+ translocation chromosome and very little other human material. We have shown that the c-*fes* oncogene, which has been mapped to chromosome 15, is not present in this hybrid and, therefore, probably is translocated to the 17q− chromosome. Analysis of the genetic markers present in this hybrid has enabled a more precise localization of the translocation breakpoints on chromosomes 15 and 17. Our experiments also have enabled an ordering and more precise mapping of several genetic markers on chromosomes 15 and 17.**

Several oncogenes, each of which was initially detected by use of transforming retrovirus-derived molecular probes, have been identified within the human genome (1). Many of these human sequences have been molecularly cloned, and studies have been initiated to determine their possible relationship to oncogenicity in man. At least two sequences, c-*ras*^{Ki} and c-*ras*^{Ha}, transform NIH 3T3 mouse cells in culture upon direct isolation from specific tumor cell lines (2–4). In other studies, human c-*ras*^{Ha} sequences have been isolated that lack such activity but do transform cells when linked to retrovirus promoter sequences (5). Similar results have been obtained with c-*mos*, the mouse cellular homologue of the Moloney murine sarcoma virus oncogene (6), whereas activation of the mouse c-*mos* gene by a transposition event resulting in its positioning adjacent to an insertion sequence (IS)-like element has been described more recently (7).

By somatic cell hybridization, many of the known oncogenes have been assigned to specific human chromosomes (8–15). Although the distribution of such sequences appears somewhat random, many of the chromosomes to which these genes have been localized are involved in translocations associated with specific human neoplasms. For instance, c-*abl* has been localized in a small terminal fragment of the long arm, q, of chromosome 9, which is translocated to chromosome 22 in chronic myelogenous leukemia (CML) (16), whereas c-*sis* maps within the region of chromosome 22 (q11 to qter), which translocates to chromosome 9 in CML (17). In other studies, c-*myc* has been identified within a fragment of human chromosome 8, which is translocated to chromosomes 2, 14, or 22 in Burkitt lymphoma (8, 9). Of potential significance, each of the chromosomes to which c-*abl* or c-*myc* are translocated in CML and Burkitt lymphoma, respectively, contain immunoglobulin sequences that map at the position of the translocation breakpoints (9, 18, 19).

The human cellular homologues of the feline sarcoma virus v-*fes* gene and of v-*fps*, a transforming gene shared by several independent isolates of avian sarcoma viruses, correspond to a common cellular locus, c-*fes* (20, 21), mapping on human chromosome 15 (13, 22). Functionally, v-*fes* and v-*fps* are also related in that both encode transforming proteins with tyrosine-specific protein kinase activity (23–27). The present study was undertaken to ascertain whether c-*fes* is involved in the 15q+;17q− chromosome translocation frequently associated with acute promyelocytic leukemia (APL) (28–31). This translocation occurs as a reciprocal exchange of the distal segments of the q arms of chromosomes 15 and 17. We also have been able to map more precisely the translocation breakpoints relative to known genetic markers on these chromosomes.

MATERIALS AND METHODS

Cells. Frozen blood leukocytes from a 26-yr-old male patient (J.D.) with the typical "M3" form of APL (32) were cultured for 48 hr in RPMI 1640 medium containing 20% fetal calf serum before karyotyping. The growth medium was supplemented with filtered supernatant from a 7-day culture of phytohemagglutinin-stimulated normal blood lymphocytes to give a final concentration of 10%. A subclone of the 3T3 mouse cell line (33), which is deficient for thymidine kinase (TK), was used for preparation of somatic cell hybrids.

Somatic Cell Hybridization. Frozen blood leukocytes from patient J.D. were fused with 3T3 (TK$^-$) cells by using polyethylene glycol (34). TK$^+$ hybrids were selected in HAT medium [RPMI 1640 medium containing 20% fetal calf serum and HAT (100 μM hypoxanthine/10 μM methotrexate/16 μM thymidine)] (35, 36). The unfused 3T3 cells die in HAT medium as they lack TK activity, whereas normal leukocytes are unable to grow in culture without mitogenic stimulation. Back-selection of TK$^-$ variants was carried out by growing hybrids in medium with 60 μg of BrdUrd per ml (37, 38).

Genetic Markers. Genetic markers previously mapped to chromosome 15 were analyzed by published immunologic and electrophoretic procedures (39–42). These were genes for β_2-microglobulin (B2M) (39), hexosaminidase α (HEXA) (41), mannosephosphate isomerase (MPI) (42), pyruvate kinase-3 (PKM2) (42), and a cell surface antigen of apparent M_r 95,000 encoded by

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: APL, acute promyelocytic leukemia; CML, chronic myelogenous leukemia; TK, thymidine kinase; HAT, hypoxanthine/methotrexate/thymidine; B2M, β_2 microglobulin; HEXA, hexosaminidase α; MPI, mannosephosphate isomerase; PKM2, pyruvate kinase-3; GALK, galactokinase; GAA, acid α-glucosidase; kb, kilobase(s).

the gene *MIC7* (40). As an independent means of identifying *B2M*, DNA Southern blot hybridization analysis was performed with a molecularly cloned cDNA probe provided by B. Wallace (43). Human c-*fes* sequences were identified by DNA Southern blot hybridization with a 0.5-kilobase (kb) *Kpn* c-*fes* probe (designated c-*fes* $K_{0.5}$) as described (44). Of the genetic markers mapped to chromosome 17, galactokinase (GALK) and acid α-glucosidase (GAA) were assayed by starch gel electrophoresis (45); the presence of the gene *MIC6* coding for a cell surface antigen of apparent M_r 125,000 (46) was measured by radioimmunoassay. The presence of the collagen α1(I) gene *COLIA1* was measured by DNA Southern blot hybridization with a human 15-kb genomic DNA probe designated pgα1H-1. This probe was isolated and provided by D. Rowe (University of Connecticut).

Chromosome Analyses. Chromosomes were prepared by standard procedures. Leukocytes from the patient J.D. were karyotyped by using G-banding (47) and Q-banding (48). Human chromosomes were identified in the hybrids by using G-11 (49), Q-banding, and G-banding techniques.

RESULTS

Karyotypic Analysis of Cultured Leukocytes from the APL Patient. A karyotype of 46,XY,t(15q+;17q−) (t designates translocation) was found in 18 of 20 cells examined, with no evidence for additional chromosome rearrangements. A representative partial karyotype is shown in Fig. 1 *Left*. By G-banding and Q-banding techniques, the translocation is characteristic of that previously described in APL (28–31).

Breakpoint in Chromosome 17 Relative to the *TK* and *GALK* Genes. To isolate somatic cell hybrids suitable for further studies of the 15q+;17q− translocation, peripheral blood leukocytes from the patient described above were fused with TK-deficient 3T3 mouse cells (33). Of a total of 12 hybrid clones examined, 7 were found to contain chromosome 17 in the absence of either 15q+ or 17q−. Four hybrids contained chromosome 17, both 15q+ and 17q− chromosomes in a small proportion of the cells, and 16–30 other human chromosomes. Finally, a single hybrid clone, PJT_2/A_1 (Fig. 1 *Right*), contained the 15q+ translocation chromosome in 37 of 40 cells examined, chromosome 21 in 9 of 40 cells, and a marker chromosome that did not appear to be related to either chromosomes 15 or 17 in 5 of 40 cells. The presence of TK in the hybrid PJT_2/A_1 localizes the 15q+;17q− translocation breakpoint proximal to or within band 17q210–220. This localization is supported by our finding that this hybrid also expresses GALK, which maps to 17q210–220 (50).

Genetic Markers Translocated from Chromosome 17 to 15q+. Fig. 2 shows the regional localization of previously described genetic markers on chromosome 17 (46, 50, 53–56). As predicted on the basis of its mapped position (54), *GAA* is expressed in PJT_2/A_1 hybrid cells. Additionally, the PJT_2/A_1 hybrid contains the *MIC6* gene (Fig. 3), thus establishing its localization within the segment of chromosome 17 that is translocated to the 15q+ chromosome in APL. Another locus, *COLIA1*, coding for the α(I) chain of type I collagen, has been mapped directly on chromosome 17 by molecular hybridization analysis (56). By using a 15-kb genomic probe for this gene, this assignment was confirmed with hybrids containing chromosome 17 as their only human material (unpublished data). We used this probe to show that the gene is present in PJT_2/A_1 hybrid cells (Table 1). This latter observation is consistent with results mapping a type I collagen gene to 17q21–q22 (55).

Regional Localization of c-*fes* and Other Loci on Chromosome 15. The availability of PJT_2/A_1 hybrid cells also enabled a more precise localization of several genetic markers previously mapped on chromosome 15, including c-*fes* (13, 22) and *B2M* (51). As shown in Fig. 4 *Left*, the human cellular homologue of v-*fes* (lane B), seen as *Sst* I restriction fragments of around 4.0 and 7.0 kb, was readily distinguishable from the mouse c-*fes* gene (lane A) by using a 0.5-kb c-*fes*-specific molecular probe (designated c-*fes* $K_{0.5}$). Human c-*fes* sequences were present in hybrid HORL-I (22), which contains chromosome 15 and a fragment of chromosome 11 as its only karyotypically detectable human material (lane C), and in hybrid DT2.12 (41), which contains a translocation between the long arms of chromosomes 15 and X (lane E). The absence of detectable human c-*fes*-specific sequences in hybrid DT2.1R3, a

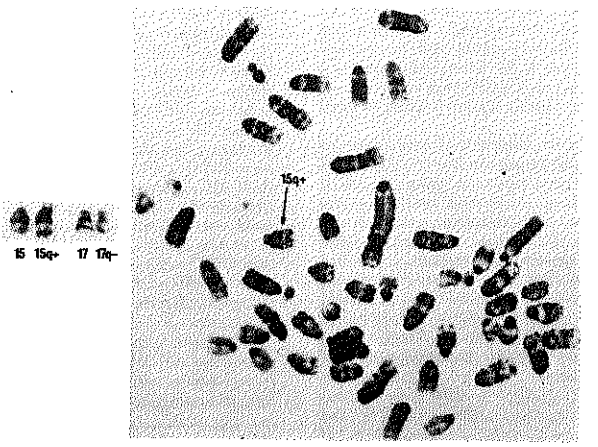


FIG. 1. Karyotypic analysis of the 15q+;17q− chromosome translocation in APL. (*Left*) Partial karyotype of a representative cell from APL patient J.D. (*Right*) G-banded metaphase spread of hybrid PJT_2/A_1 showing 15q+ translocation chromosome. Chromosome 21 and the unidentified marker chromosome are not present in this cell.

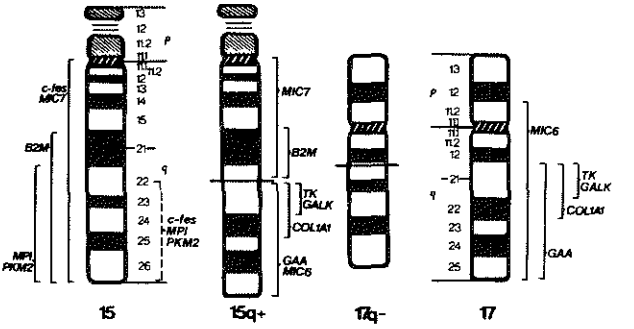


FIG. 2. Schematic representation of t(15q+;17q−). Translocation breakpoints are placed in bands 15q22 and 17q21 on the basis of both genetic and cytogenetic data (see text). Localizations of genes tested in the hybrid PJT_2/A_1 including *B2M* (51); *MIC7* (40), *MPI* (52), *PKM2* (52), c-*fes* (13, 22), *MIC6* (46), *GALK* (50), *TK* (53), *GAA* (54), and *COL1A1* (55, 56) are shown on the normal chromosomes 15 and 17 (solid lines). Those genes absent from hybrid PJT_2/A_1 are shown to the right of the normal chromosome 15 (broken line). Those genes present in hybrid PJT_2/A_1 are shown on the 15q+ translocation chromosome.

derivative of DT2.1R (41) lacking the 15/X translocation chromosome, confirms the earlier assignment of c-*fes* to chromosome 15 (lane F) (13, 22). Finally, the absence of human c-*fes* sequences in hybrid PJT_2/A_1 (lane D) establishes that c-*fes* is not present within the 15q+ translocation chromosome and strongly suggests that these sequences map within the region of chromosome 15 that is translocated to chromosome 17.

The results in Fig. 4 *Right* show that, in contrast to c-*fes*, human *B2M* sequences are present not only in hybrids containing chromosome 15 but also in hybrid PJT_2/A_1 (lane D), localizing *B2M* on chromosome 15q+ and, thus, proximal to the translocation breakpoint of chromosome 15. These are seen as a series of three *Sst* I restriction fragments of 3.0, 7.0, and 12.0 kb. Two additional markers, *MPI* and *PKM2*, previously mapped within the long arm of chromosome 15 (52), were not detected in hybrid PJT_2/A_1, while a third chromosome 15 marker, *MIC7*, was present (Table 1). These findings allow a more precise regional localization of each of these markers on chromosome 15 (Fig. 2). Analysis of *HEXA*, previously localized to 15q22–q25.1 (61), gave ambiguous results.

PJT_2/A_1 Revertant Lacking the 15q+ Chromosome. To confirm the above assignments of various genetic markers to the 15q+ translocation chromosome, PJT_2/A_1 hybrid cells were back-selected by growth in medium containing 60 μg of BrdUrd per ml. Cells containing the human *TK* gene incorporate BrdUrd and, thus, should be killed under these conditions (37, 38). One back-selected hybrid isolated by this approach, designated PJT_2/A_1R, was shown by karyotype analysis to have lost the human 15q+ translocation chromosome and the unidentified marker chromosome. This hybrid also was found to have lost *B2M*, *MIC7*, *GALK*, *GAA*, and *MIC6*. These results confirm our assignment of each of these markers as well as the human *TK* and *GALK* genes to the 15q+ translocation chromosome.

DISCUSSION

Although there is evidence for a specific association of the 15q+;17q− chromosome translocation with APL (28–31), the breakpoints in this translocation are difficult to localize and have been placed at different positions by various groups of investigators—e.g., 15q22 and 17q21 (29), 15q25–26 and 17q22 (28), and 15q22–23 and 17q12 (30). The approach used in the present study allowed more precise determination of these breakpoints in one patient. For instance, the presence of *TK* and *GALK* in hybrid PJT_2/A_1 establishes the breakpoint on chromosome 17 at a position proximal to that previously shown for these loci

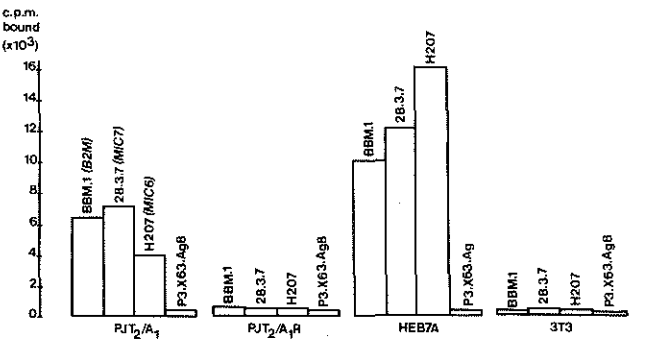


FIG. 3. Radioimmunoassay to detect the presence of *B2M*, *MIC6*, and *MIC7* genes by using monoclonal antibodies BBM1 (ref. 39), H207 (ref. 46), and 28.3.7 (ref. 40), respectively. Control monoclonal antibody is P3.X63.Ag8 (ref. 57). Binding of monoclonal antibody to cells visualized using ^{125}I-rabbit anti-mouse Ig. Human control cells are HEB7A (ref. 58). Mouse control cells are 3T3 (ref. 33).

Table 1. Identification of chromosome 15 and 17 genetic markers on the 15q+ translocation chromosome*

Hybrids	Human chromosomes	Chromosome 15 markers						Chromosome 17 markers				
		MIC7	*HEXA*	*MPI*	*PK*	*c-fes*	*B2M*	*MIC6*	*TK*	*GALK*	*COL1A1*	*GAA*
Horl-I	15, 11q	+	+	+	+	+	+	–	–	–	–	–
DT2.12	t(15/X), 2, 10, 11, 12, 13, 17, 21, 22	+	NT	+	NT	+	+	+	NT	+	NT	NT
DT2.1R3	2, 10, 11, 12, 18	–	NT	–	NT	–	–	–	NT	–	NT	NT
PJT_2/A_1	15q+, 21, marker	+	±	–	–	–	+	+	+	+	+	+
PJT_2/A_1R	21	–	–	NT	NT	NT	–	–	–	–	NT	–

NT, not tested.

*Details concerning the derivation and chromosomal characterization of Horl-I, DT2.12, and DT2.1R3 somatic cell hybrids has been reported (22, 41). PJT_2/A_1 and PJT_2/A_1R are described in the present study.

(q210–220) (50, 53). This information combined with a karyotypic analysis of the translocation chromosomes suggests that the breakpoint on chromosome 15 is within 15q22. These breakpoints are consistent with those proposed most recently by Rowley, who used extended chromosome preparations from several patients—i.e., 15q22.2 and 17q12–21 (J.D. Rowley, personal communication). On the basis of these considerations, in the schematic representation of the t(15q+;17q−) in Fig. 2, the breakpoints are shown in bands 15q22 and 17q21.

Our results also allow a more precise ordering of the genes on human chromosome 15. For instance, *MIC7* and *B2M* have been localized proximal to the translocation breakpoint on chromosome 15, whereas *MPI*, *PKM2*, and *c-fes* are distal to this position. Furthermore, we have localized *MIC6* to the long arm of chromosome 17, probably distal to band q210.

The absence of *c-fes* from the human 15q+ translocation chromosome strongly argues for its localization within the fragment of chromosome 15 that translocates to chromosome 17 in APL. The possibility that *c-fes* sequences are simply deleted or are localized within a small undetected fragment of chromosome 15 that is translocated to a third human chromosome, although unlikely, cannot be rigorously excluded. If translocation of *c-fes* from chromosome 15 to 17 is relevant to APL, chromosome 17q− would be predicted to be the critical chromosome. Some evidence suggesting that chromosome 15q+ is the critical chromosome has been obtained from two cases of APL with complex translocations involving a third chromosome (62, 63). In both these cases, the constant recombinant is the 15q+ translocation chromosome derived from the proximal segment of 15q and the distal segment of 17q. Further studies of this nature are required.

The demonstration of *B2M* sequences within the 15q+ translocation chromosome, in combination with earlier data positioning *B2M* distal to 15q21 (51), maps these sequences at or near the breakpoint in chromosome 15. In other studies, immunoglobulin sequences have been mapped at the breakpoints of chromosome 22 in the Ph^1 translocation associated with CML and in the regions of chromosomes 14 and 22 that are translocated to chromosome 8 in Burkitt lymphoma (9, 19). Similarly, the breakpoint in mouse chromosome 12 leading to translocations associated with lymphoid tumors has been shown to map within immunoglobulin heavy chain sequences (64–66). Although *B2M* is of functional significance primarily with respect to expression of HLA antigens and shows no evidence for structural rearrangements in development, it does exhibit some structural resemblance to immunoglobulins (67); therefore, its mapping near the breakpoint in chromosome 15 may be relevant to the 15;17 translocation associated with APL.

The possible involvement of *c-fes* in the 15q+;17q− translocation in APL is of interest in view of accumulating evidence for the mapping of human cellular oncogenes within chromosomal fragments subject to highly specific translocations associated with particular classes of tumors (8–10, 16). Moreover, the availability of somatic cell hybrids containing the 15q+ and

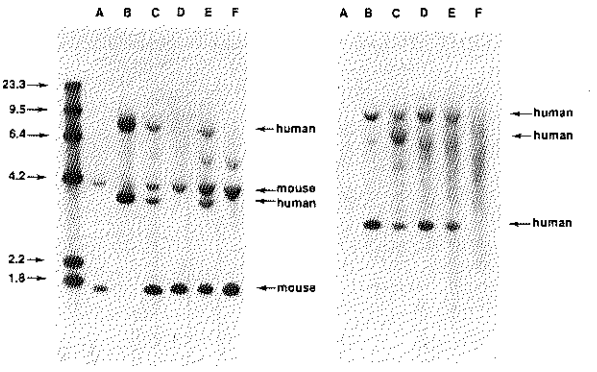


FIG. 4. Localization of *c-fes* and *B2M* relative to the breakpoint on human chromosome 15. *Sst* I-digested cellular DNAs (25 μg per lane) were electrophoresed on 0.7% agarose gels, blotted to nitrocellulose, and hybridized to either *c-fes* $K_{0.5}$ (*Left*) or *B2M* (*Right*). Cell lines are: mouse, 1R (ref. 59; lanes A); human, LNSV (ref. 60; lanes B); HORL-I (ref. 22; lanes C); PJT_2/A_1 (lanes D); DT2.12 (ref. 41; lanes E), and DT2.1R3 (ref. 41; lanes F). The positions of mouse and human specific restriction fragments are as indicated. *Hin*dIII-digested phage λ DNA fragments are the size markers in kb (*Left*).

17q− translocation chromosomes should facilitate studies of the breakpoints in this translocation and a determination of the proximity of c-*fes* to the translocation breakpoint.

We thank Mohamed Parkar, Barbara Pym, G. T. Blennerhassett, and P. Hanson for their excellent technical assistance; Dr. Janet Rowley for stimulating discussions; and Christine Furse for typing. We are also grateful to the Imperial Cancer Research Fund Medical Oncology Unit, St. Bartholomew's Hospital, London, for generously providing us with cells from APL patients. This work was partially supported under Contract NO1-CO-76380 from the National Cancer Institute.

1. Varmus, H. E. (1982) *Science* **216**, 812–820.
2. Santos, E., Tronick, S. R., Aaronson, S. A., Pulciani, S. & Barbacid, M. (1982) *Nature (London)* **298**, 343–347.
3. Der, C. J., Krontiris, T. G. & Cooper, G. M. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 3637–3640.
4. Parada, L. F., Tabin, C. J., Shih, C. & Weinberg, R. A. (1982) *Nature (London)* **297**, 474–478.
5. Chang, E. H., Furth, M. E., Scolnick, E. M. & Lowy, D. R. (1982) *Nature (London)* **297**, 479–483.
6. Blair, D. G., Oskarsson, M., Wood, T. G., McClements, W. L., Fischinger, P. J. & Vande Woude, G. F. (1981) *Science* **212**, 941–943.
7. Rechavi, G., Givol, D. & Canaani, E. (1982) *Nature (London)* **300**, 607–611.
8. Dalla-Favera, R., Bregni, M., Erikson, J., Patterson, D., Gallo, R. C. & Croce, C. M. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 7824–7827.
9. Taub, R., Kirsch, I., Morton, C., Lenoir, G., Swan, D., Tronick, S., Aaronson, S. & Leder, P. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 7837–7841.
10. Neel, B. G., Jhanwar, S. C., Chaganti, R. S. K. & Hayward, W. S. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 7842–7846.
11. Swan, D. C., McBride, O. W., Robbins, K. C., Keithley, D. A., Reddy, E. P. & Aaronson, S. A. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4691–4695.
12. Dalla-Favera, R., Gallo, R. C., Giallongo, A. & Croce, C. M. (1982) *Science* **218**, 686–688.
13. Dalla-Favera, R., Franchini, G., Martinotti, S., Wong-Staal, F., Gallo, R. C. & Croce, C. M. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4714–4717.
14. Prakash, K., McBride, O. W., Swan, D. C., Devare, S. G., Tronick, S. R. & Aaronson, S. A. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 5210–5214.
15. McBride, O. W., Swan, D. C., Santos, E., Barbacid, M., Tronick, S. R. & Aaronson, S. A. (1982) *Nature (London)* **300**, 773–774.
16. de Klein, A., van Kessel, A. G., Grosveld, G., Bartram, C. R., Hagemeijer, A., Bootsma, D., Spurr, N. K., Heisterkamp, N., Groffen, J. & Stephenson, J. R. (1982) *Nature (London)* **300**, 765–767.
17. Groffen, J., Heisterkamp, N., Stephenson, J. R., van Kessel, A. G., de Klein, A., Grosveld, G. & Bootsma, D. (1983) *J. Exp. Med.*, in press.
18. Erikson, J., Martinis, J. & Croce, C. M. (1981) *Nature (London)* **294**, 173–175.
19. Kirsch, I. R., Morton, C. C., Nakahara, K. & Leder, P. (1982) *Science* **216**, 301–303.
20. Shibuya, M., Hanafusa, H., Hanafusa, T. & Stephenson, J. R. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 6536–6540.
21. Groffen, J., Heisterkamp, N., Shibuya, M., Hanafusa, H. & Stephenson, J. R. (1983) *Virology* **125**, 480–486.
22. Heisterkamp, N., Groffen, J., Stephenson, J. R., Spurr, N. K., Goodfellow, P. N., Solomon, E., Carritt, B. & Bodmer, W. F. (1982) *Nature (London)* **299**, 747–749.
23. Van de Ven, W. J. M., Reynolds, F. H., Jr., & Stephenson, J. R. (1980) *Virology* **101**, 185–197.
24. Feldman, R. A., Hanafusa, T. & Hanafusa, H. (1980) *Cell* **22**, 757–765.
25. Pawson, T., Guyden, J., Kung, T.-H., Radke, K., Gilmore, T. & Martin, G. S. (1980) *Cell* **22**, 767–775.
26. Reynolds, F. H., Jr., Van de Ven, W. J. M. & Stephenson, J. R. (1980) *J. Biol. Chem.* **255**, 11040–11047.
27. Barbacid, M., Beemon, K. & Devare, S. G. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5158–5162.
28. Second International Workshop on Chromosomes in Leukemia, 1979: Chromosomes in Acute Promyelocytic Leukemia (1980) *Cancer Genet. Cytogenet.* **2**, 103–107.
29. Rowley, J. D., Colomb, H. M., Vardiman, J., Fukuhara, S., Dougherty, C. & Potter, D. (1977) *Int. J. Cancer* **20**, 869–872.
30. Kondo, K. & Sasaki, M. (1982) *Cancer Genet. Cytogenet.* **6**, 39–46.
31. Van Den Berghe, H., Louwagie, A., Broeckaert-Van Orshoven, A., David, G., Verwilghen, R., Michaux, J. L. & Sokal, G. (1979) *Cancer* **43**, 558–562.
32. Bennet, J. M., Catovsky, D., Daniel, M. T., Flandrin, G., Galton, D. A. G., Gralnick, H. R. & Sultan, C. (1976) *Br. J. Haematol.* **33**, 451–458.
33. Todaro, G. J. & Green, H. (1963) *J. Cell Biol.* **17**, 299–316.
34. Galfre, G., Howe, S. C., Milstein, C., Butcher, G. W. & Howard, J. C. (1977) *Nature (London)* **266**, 550–552.
35. Szybalski, W., Szybalski, E. & Rayni, G. (1962) *Natl. Cancer Inst. Monogr.* **7**, 75–89.
36. Littlefield, J. (1964) *Science* **145**, 709–710.
37. Kit, S., Dubbs, D. R., Piekarski, J. R. & Hsu, T. C. (1963) *Exp. Cell Res.* **31**, 297–312.
38. Weiss, M. C. & Green, H. (1967) *Proc. Natl. Acad. Sci. USA* **58**, 1104–1111.
39. Brodsky, F. M., Bodmer, W. F. & Parham, P. (1979) *Eur. J. Immunol.* **9**, 536–545.
40. Blaineau, C., Avner, P., Tunnacliffe, A. & Goodfellow, P. N. (1983) *EMBO J.*, in press.
41. Swallow, D. M., Solomon, E. & Pajunen, L. (1977) *Cytogenet. Cell Genet.* **18**, 136–146.
42. van Heyningen, V., Bobrow, M., Bodmer, W. F., Gardiner, S. E., Povey, S. & Hopkinson, D. A. (1975) *Ann. Hum. Genet.* **38**, 295–302.
43. Suggs, S. V., Wallace, R. B., Hirose, T., Kawashima, E. H. & Ikatura, K. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 6613–6617.
44. Groffen, J., Heisterkamp, N., Grosveld, F., Van de Ven, W. & Stephenson, J. R. (1982) *Science* **216**, 1136–1138.
45. Solomon, E., Swallow, D., Burgess, S. & Evans, L. (1979) *Ann. Hum. Genet.* **42**, 273–281.
46. Bai, Y., Sheer, D., Hiorns, L., Knowles, R. W. & Tunnacliffe, A. (1982) *Ann. Hum. Genet.* **46**, 337–347.
47. Seabright, M. (1972) *Chromosoma (Berlin)* **36**, 204–210.
48. Caspersson, T., Lomakka, G. & Zech, L. (1971) *Hereditas* **67**, 89–102.
49. Bobrow, M. & Cross, J. (1974) *Nature (London)* **251**, 77–79.
50. Elsevier, S. M., Kucherlapati, R. S., Nichols, E. A., Creagan, R. P., Giles, R. P., Ruddle, F. H., Willecke, K. & McDougall, J. K. (1974) *Nature (London)* **251**, 633–635.
51. Pajunen, L., Solomon, E., Burgess, S., Bobrow, M., Povey, S. & Swallow, D. (1978) *Cytogenet. Cell Genet.* **22**, 511–512.
52. Chern, C. J., Kennett, R., Engel, E., Mellman, W. J. & Croce, C. M. (1977) *Somatic Cell Genet.* **3**, 553–560.
53. Human Gene Mapping 5 (1979) *Cytogenet. Cell Genet.* **25**, 59–73.
54. Nickel, B. E., Chudley, A. E., Pabello, P. D. & McAlpine, P. J. (1982) *Cytogenet. Cell Genet.* **32**, 303.
55. Church, R. L., Sundar Raj, N. & McDougall, J. K. (1980) *Cytogenet. Cell Genet.* **27**, 24–30.
56. Huerre, C., Junien, C., Weil, D., Chu, M.-L., Morabito, M., van Cong, N., Myers, J. C., Foubert, C., Gross, M.-S., Prockop, D. J., Boue, A., Kaplan, J. C., de la Chapelle, A. & Ramirez, F. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 6627–6631.
57. Kohler, G. & Milstein, C. (1975) *Nature (London)* **256**, 495–497.
58. Wallace, D. C., Bunn, C. L. & Eisenstadt, J. M. (1975) *J. Cell Biol.* **67**, 174–188.
59. Nabholz, M., Miggiano, V. & Bodmer, W. F. (1969) *Nature (London)* **223**, 358–363.
60. Croce, C. M., Knowles, B. B. & Koprowski, H. (1973) *Exp. Cell Res.* **82**, 457–461.
61. Magenis, R. E., Vidgoff, J., Chamberlin, J. & Brown, M. G. (1979) *Cytogenet. Cell Genet.* **25**, 181.
62. Bernstein, R., Mendelow, B., Pinto, M. R., Morcom, G. & Bezwoda, W. (1980) *Br. J. Haematol.* **46**, 311–314.
63. Rowley, J. D. (1982) *Science* **216**, 749–751.
64. Crews, S., Barth, R., Hood, L., Prehn, J. & Calame, K. (1982) *Science* **218**, 1319–1321.
65. Shen-Ong, G. L. C., Keath, E. J., Piccoli, S. P. & Cole, M. D. (1982) *Cell* **31**, 443–452.
66. Calame, K., Kim, S., Lalley, P., Hill, R., Davis, M. & Hood, L. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 6994–6998.
67. Ploegh, H. L., Orr, H. T. & Strominger, J. L. (1981) *Cell* **24**, 287–299.

APPENDIX PAPER VII

The Human v-*abl* Cellular Homologue

Nora Heisterkamp, John Groffen, and John R. Stephenson

Laboratory of Viral Carcinogenesis, National Cancer Institute, Frederick, Maryland, U.S.A.

Summary: A contiguous region of cellular DNA sequence, 64 kb in length and representing overlapping cellular inserts from three independent cosmid clones, has been isolated from a representative library of human lung carcinoma DNA partially digested with *Mbo*I. Within this region of the cellular genome, v-*abl* homologous sequences are dispersed over a total region of around 32 kb. These sequences represent the entire v-*abl* human cellular homologue, are colinear with the viral v-*abl* transforming gene, and contain a minimum of seven intervening sequences. At least eight regions of highly repetitive DNA sequences have been shown to map in close proximity to c-*abl* coding sequences. In addition to the major c-*abl* human locus, three regions of human DNA sequence, corresponding to only portions of the v-*abl* gene, have been identified. Two of these have been molecularly cloned and shown to be distinct from the primary human c-*abl* locus. Upon transfection to rat embryo fibroblasts in culture, none of the cosmid DNAs containing v-*abl* homologous sequences exhibited transforming activity. These findings identify and map a single genetic locus of human DNA, c-*abl*, representing the complete v-*abl* homologue, and demonstrate the existence of additional human DNA sequences corresponding to more limited, subgenomic regions of v-*abl*. **Key Words:** v-*abl*-Human c-*abl*—Oncogenes—Malignant transformation.

RNA tumor viruses have been shown to represent recombinants between nontransforming type C viruses and cellular genetic sequences (oncogenes) (1,2). One approach to the identification of cellular genes with potential involvement in naturally occurring tumors has involved the molecular cloning of cellular homologues of such "acquired" viral transforming genes (2). Alternatively, efforts have been made to isolate cellular oncogenic sequences directly from total transformed cell DNA by a combination of transfection and molecular cloning procedures (3–5). In one case involving transformation-specific sequences associated with human bladder carcinoma cells, these independent approaches have led to the identification of the same cellular sequences (3–5). Thus, these methods provide the possibility of identifying human genes that may be involved in expression of the transformed phenotype and thereby represent powerful approaches for determining the molecular basis of malignancy.

The oncogenic retrovirus considered in the present study, Abelson murine leukemia virus

Received August 18, 1982; revised November 15, 1982; accepted December 7, 1982.

Address correspondence and reprint requests to John R. Stephenson, Laboratory of Viral Carcinogenesis, National Cancer Institute, Frederick, Maryland 21701, U.S.A.

(Abelson MuLV), is a recombinant between mouse cellular sequences (v-*abl*) and Moloney MuLV (6,7). Abelson MuLV is of particular interest in that, in contrast to other mammalian transforming viruses, it transforms lymphoid cells *in vitro* (8) and induces a rapid B-cell lymphoma *in vivo* (9). The major Abelson MuLV translational product has been identified as a 120,000-M_r polyprotein ($P120^{gag\text{-}abl}$) containing Moloney MuLV *gag* gene-encoded structural proteins (p15 and p12) and a carboxy terminal transformation-specific component (10,11). $P120^{gag\text{-}abl}$ is one of several viral encoded transforming proteins with tyrosine-specific protein kinase activity (12–15). By use of a series of transformation-defective mutants, the $P120^{gag\text{-}abl}$ enzymatic activity has been shown to be required for Abelson MuLV-induced transformation (14,16) and for the production of a 7600-M_r transforming growth factor (TGF) closely resembling TGFs produced by a number of human tumors (17,18).

As an initial step in efforts to determine the potential involvement of human v-*abl* homologous sequences in naturally occurring tumors, we undertook their identification and molecular cloning. Because of our preliminary results indicating the presence, within human genomic DNA, of numerous v-*abl* homologous restriction fragments, and evidence for extensive intervening sequences within the cellular homologues of certain other viral transforming genes (2), a cosmid vector system was used (19,20). By this means, the isolation of relatively large (35–45 kb) v-*abl* homologous DNA fragments is possible, thus facilitating efforts to determine the structural organization of such sequences within the human genome. Using this approach, we have isolated a single human locus (c-*abl*) that contains sequences corresponding to the complete v-*abl* gene distributed over a region of around 32 kb, including multiple noncoding (intervening) sequences. Other human sequences corresponding to less extensive subgenomic regions of v-*abl* are also identified, and a possible model that may account for their origin is considered.

METHODS

Gel Electrophoresis and Hybridization

Restriction enzymes were purchased from either New England Biolabs or Bethesda Research Laboratories (BRL) and were used according to the suppliers' specifications. DNAs were digested with restriction enzymes, subjected to electrophoresis through 0.75% agarose gels, and transferred to nitrocellulose (Schleicher and Schuell, ph79) essentially as described by Southern (21). Nick translation of probes and filter hybridization were as described (22,23). Specific activity of the probes was $2-5 \times 10^8$ cpm/μg. After hybridization, filters were exposed to XAR-2 film (Kodak) for up to 5 days at −70°C with Dupont Lightning Plus intensifying screens.

Preparation of DNA Probes

DNA probes were prepared by digestion of 150 μg of DNA with appropriate restriction enzymes, followed by electrophoresis through low-melting-point agarose (BRL). Desired bands were excised from gels and brought into solution by heating at 65°C for 30 min. Agarose was removed by two extractions with phenol equilibrated with 0.3 *M* NaOAc, pH 5.0, and one extraction with phenol/chloroform/isoamyl alcohol (25:24:1). DNA was precipitated with ethanol and 0.2 *M* NaOAc, pH 4.8, in the presence of 20 μg/ml Dextran T-500 as carrier.

Isolation of Human v-*abl* Homologous Sequences

For isolation of v-*abl* homologous cellular sequences, a previously described cosmid library of human lung carcinoma DNA partially digested with *Mbo*I was used. This library is highly representative and contains relatively large (35–45 kb) cellular inserts. Construction of the cosmid library, as well as procedures for the screening, isolation, and growth of recombinants, was as described (19).

Subcloning

Plasmid vector pHEP (20) was digested with *Bam*HI and treated with alkaline phosphatase (Boehringer) prior to ligation with the insert to prevent reannealing of the vector arms. The phosphatased DNA was extracted once with phenol/chloroform/isoamyl alcohol (25:24:1) after addition of EDTA to a final concentration of 10 m*M*. The DNA was precipitated with ethanol in the presence of 20 μg/ml Dextran T-500 (Pharmacia) as carrier and $^1/_{10}$ volume of 2 *M* NaOAc, pH 4.8. The phosphatase treated vector was ligated with *Bam*HI-digested λ-AM-1 (a clone of λgtWES λB containing a 7.8-kb cellular insert corresponding to the complete Abelson MuLV proviral DNA, generously provided by P. Reddy, NCI; 24) in a molar ratio of, respectively, 5:1. Transformation of *E. coli* HB101 to ampicillin resistance was as described previously (25).

DNA-Mediated Transfection

Transfection of TK$^-$ Rat-2 cells (26) was based on the methods of Graham and Van der Eb (27) as modified by Wigler et al. (28) with the exception that carrier DNA was omitted.

RESULTS

Isolation of v-*abl* Homologous Sequences from a Cosmid Library of Human Cellular DNA

For use as a molecular probe, a 3.5-kb *Bam*HI restriction fragment encompassing all but a short region of the Abelson MuLV proviral DNA was excised from λ-Am-1 (24) and subcloned into pHEP. Transfection of the newly generated plasmid, pHEP-Ab, to Rat-2 cells resulted in comparable numbers of transformed foci as observed with λ-AM-1 DNA, thus establishing the *Bam*HI restriction fragment to include the complete v-*abl* transforming gene. A DNA probe prepared from pHEP-Ab by digestion with *Bst*EII and *Bam*HI contained, in addition to v-*abl*, approximately 0.6 kb of sequence derived from the Moloney MuLV *gag* gene (Fig. 1). For identification of v-*abl* homologous human sequences, high molecular weight mouse and human cellular DNAs were digested with *Eco*RI, electrophoresed, transferred to nitrocellulose filters, and hybridized to the v-*abl* *Bam*HI/*Bst*EII probe. As shown in Fig. 1, lane C, v-*abl* homologous human DNA restriction fragments of 11, 7.2, 5.0, 4.0, 3.4, 2.9, and 2.5 kb were clearly

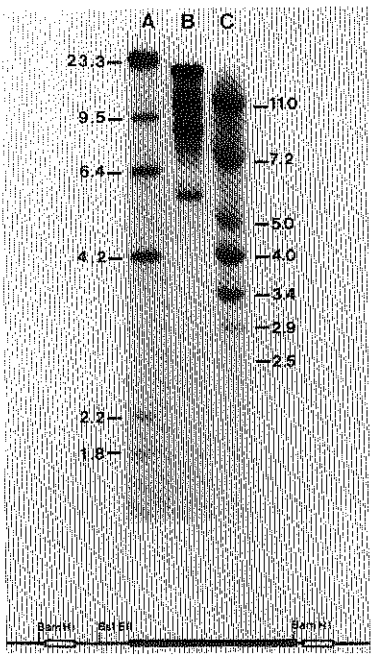
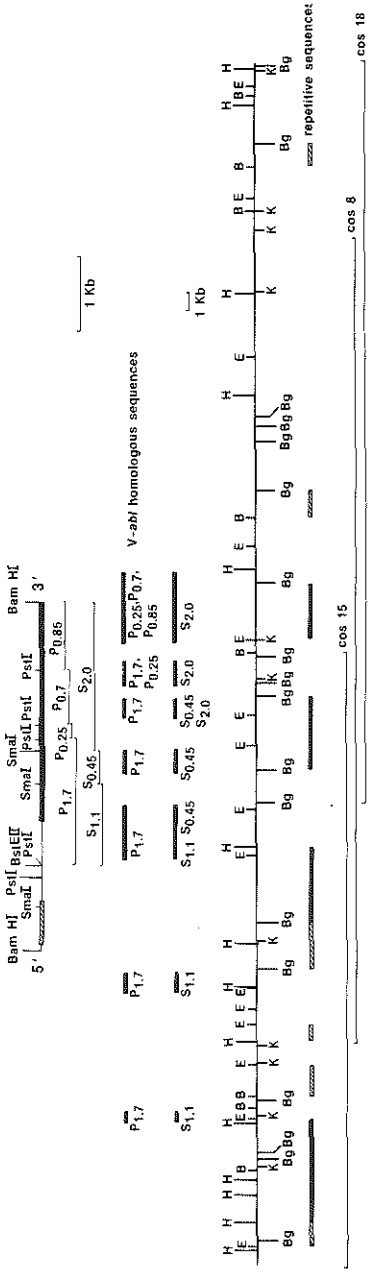


FIG. 1. Identification of v-*abl* homologous sequences in human and mouse cellular DNAs. High molecular weight DNAs (20 μg) were prepared from mouse dense N (lane B) and human A431 cells (Lane C), digested with *Eco*RI, and electrophoresed through a 0.75% agarose gel. After transfer of DNA to nitrocellulose, filters were hybridized with a ^{32}P-nick-translated v-*abl* *Bst*EII/*Bam*HI probe, indicated in the lower part of the figure. The boxed areas indicate the viral long terminal repeats; the solid bar represents the acquired cellular v-*abl* sequences of Abelson-MuLV. ^{32}P-Labeled *Hin*dIII-digested λDNA is included as a molecular weight standard (lane A). All molecular weights are in kb.



resolved. As discussed below, the 4.0-kb restriction fragment actually corresponds to two v-*abl* homologous sequences of similar size (4.0 and 4.1 kb), which are not resolved in this figure. The demonstration of numerous human DNA *Eco*RI restriction fragments hybridizing to v-*abl*, which itself encompasses a region of only 3.0 kb, argues that the human v-*abl* homologue must contain extensive intervening sequences. It is also apparent from the data presented in Fig. 1 that although strength of homology of human and mouse cellular DNAs with v-*abl* is comparable, the restriction patterns are distinct.

To facilitate further characterization of v-*abl* homologous human DNA sequences, we utilized a previously described (20) cosmid library of human lung carcinoma DNA. Upon screening with the v-*abl* *Bam*HI/*Bst*EII probe, ten positive cosmid clones were identified. None of these exhibited detectable hybridization with a subclone of pBR322 containing the complete Moloney MuLV proviral DNA (data not shown). For further characterization of these clones, a series of probes with specificity for subgenomic regions of v-*abl* was prepared. As shown in Fig. 2, digestion of plasmid pHEP-Ab with *Pst*I and *Bam*HI yielded a 5′ probe of 1.7 kb (v-*abl* $P_{1.7}$), which, in addition to v-*abl* sequences, contained around 0.6 kb of Moloney MuLV sequences. Two probes corresponding to the central portion of v-*abl* ($P_{0.25}$ and $P_{0.7}$) and a fourth probe specific for the 3′ region (v-*abl* $P_{0.85}$) were also prepared. Of the nine v-*abl* homologous cosmid clones isolated,

FIG. 2. Restriction enzyme map of the human c-*abl* region. In the upper portion of the figure the *Bam*HI subclone of Abelson MuLV is indicated. The hatched box represents the long terminal repeat; the solid bar the acquired cellular sequences. Directly beneath the Abelson-MuLV genome, subgenomic restriction enzyme fragments, used as probes in this study, are indicated. Human DNA restriction fragments homologous to v-*abl* sequences are represented as solid boxes; below each region probes that exhibit homology are shown. Boxed areas directly below the restriction enzyme map show repetitive sequences found in the human c-*abl* region. Hatched boxes represent regions with intermediate homology; solid boxes represent regions with strong homology. Beneath the map, cellular DNA inserts within cos 8, 15, and 18 are shown. Bg, *Bgl*II; B, *Bam*HI; H, *Hin*dIII; E, *Eco*RI; K, *Kpn*I.

TABLE 1. *Homology of human DNA* EcoRI *restriction fragments with* v-abl *subgenomic probes and distribution among individual cosmid isolates*

*Eco*RI restriction fragments	Extent of hybridization to the following subgenomic probes							Cos 8,15,18	Cos 4,5,7,16	Cos 2,3,10
	$P_{1.7}$	$P_{0.25}$	$P_{0.85}$	$P_{0.7}$	$S_{1.1}$	$S_{0.45}$	$S_{2.0}$			
12.0	–	–	–	W	–	–	W	–	+	–
11.0	M	–	–	–	–	M	–	–	–	–
7.2	S	–	–	–	S	–	–	+	–	–
5.0	–	W	W	S	–	–	S	+	–	–
4.0	S	W	–	–	–	M	M	+	–	–
3.4	S	–	–	–	–	M	–	+	–	–
2.9	S	–	–	–	S	–	–	+	–	–
2.5	M	–	–	–	W	W	–	+	–	–
4.1	M	–	–	–	–	S	–	–	–	+

High molecular weight DNAs (20 μg) were prepared from human A431 cells, digested with *Eco*RI, electrophoresed through a 0.75% agarose gel, transferred to nitrocellulose, and hybridized to the indicated v-*abl* subgenomic probes. Extents of hybridization are designated as strong (S), medium (M), weak (W), and negative (–). The panels on the right portion of the table indicate the three separate series of cosmid clones isolated in this study and the v-*abl* homologous *Eco*RI restriction fragments represented in each series collectively.

three, cos 8, 15, and 18, hybridized to the original v-*abl* *Bam*HI/*Bst*EII probe to a much greater extent than the others; in addition, each hybridized to at least two of the more specific v-*abl* subgenomic probes (Fig. 3). Cos 15 contained *Eco*RI fragments of 7.2, 3.4, 2.9, and 2.5 kb hybridizing to the v-*abl* $P_{1.7}$ probe. Cos 18 contained *Eco*RI fragments of 4.0 and 3.1 kb homologous to $P_{1.7}$, one of which, the 4.0-kb fragment, also hybridized to the v-*abl* $P_{0.25}$ probe, and a fragment of 5.0-kb homologous to v-*abl* $P_{0.25}$, $P_{0.7}$, and $P_{0.85}$, whereas the third clone, cos 8, contained v-*abl* homologous *Eco*RI fragments of 7.2, 5.0, 4.0, 3.4, and 2.5 kb. Finally, as shown in Fig. 3, these three clones exhibited considerable sequence cross-homology with each other.

Restriction maps of v-*abl* homologous sequences within cos 8, 15, and 18 were generated using various combinations of *Eco*RI, *Kpn*I, *Bam*HI, *Hin*dIII, and *Bgl*II. As summarized in Fig. 2, the three cosmid clones contain overlapping cellular sequences corresponding to a single contiguous region of human cellular DNA of around 64 kb. Based on hybridization with the four individual subgenomic v-*abl* probes, a minimum of seven distinct regions of v-*abl* homology (exons), interspersed by six nonhomologous regions representing probable intervening sequences (introns), was identified. In addition, colinearity of sequences within the human v-*abl* homologue and the viral v-*abl* gene itself is established.

Cos 8, 15, and 18 Contain the Complete v-*abl* Homologue

To establish whether the complete human v-*abl* homologue is represented within the cos 8, 15, and 18, and whether rearrangements might have occurred during cloning, the composite restriction map of the cloned sequences was compared with that of v-*abl* homologous sequences in human cellular DNA. The latter were identified on the basis of hybridization to individual v-*abl* subgenomic probes. As shown in Fig. 4, digestion of human DNA with *Eco*RI and subsequent hybridization to the v-*abl* $P_{1.7}$ probe yielded fragments of 11.0, 7.2, 4.0, 3.4, 2.9, and 2.5 kb, whereas the v-*abl* $P_{0.7}$ probe hybridized to a single fragment of 5.0 kb. Upon digestion of human cellular DNA, together with a small amount of DNA from cos 15 and 18, and hybridization to $P_{1.7}$, all v-*abl* homologous fragments in cos 15 and 18 comigrated with those in human genomic DNA (Fig. 5). The extra fragment of 3.1 kb represents a truncated 3.4-kb fragment located at the very 5′ end

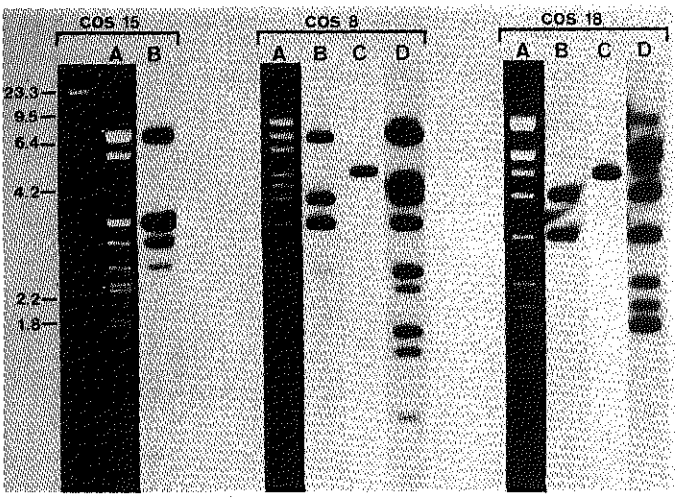


FIG. 3. Homology of cosmid clones 8, 15, and 18 with v-*abl* $P_{1.7}$, v-*abl* $P_{0.7}$, and clone 15. Cosmid DNAs (0.4 μg) were digested with *Eco*RI, electrophoresed through 0.75% agarose gels, and transferred to nitrocellulose. Lane A, ethidium-bromide stained gel; lane B, hybridization with the v-*abl* $P_{1.7}$ probe; lane C, hybridization with v-*abl* $P_{0.7}$ probe; lane D, hybridization with cos 15. A *Hin*dIII digest of λDNA is included as a molecular weight standard in the extreme left lane.

of clone 18. *Kpn*I digestion of cellular DNA generated restriction fragments of 18.6, 14.2, 5.6, and 2.8 kb hybridizing to v-*abl* $P_{1.7}$, whereas *Bam*HI digestion yielded 24.0- and 3.4-kb fragments hybridizing with v-*abl* $P_{1.7}$ and a 7.3-kb v-*abl* $P_{0.7}$ homologous fragment (Fig. 4). These results are in concordance with the mapping and hybridization data on the cloned DNA shown in Fig. 3, arguing against the possibility that major rearrangements of the c-*abl* locus might have taken place during cloning.

It should be noted that the 4.1- and 11-kb *Eco*RI fragments identified on digestion of total human DNA and hybridization with v-*abl* $P_{1.7}$ are missing from the cloned human c-*abl* sequence represented in cos 8, 15, and 18 (Table 1). Since the v-*abl* $P_{1.7}$ probe corresponds to

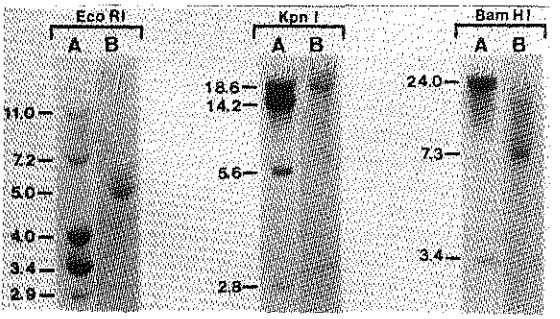


FIG. 4. Restriction endonuclease analysis of v-*abl* homologous sequences in human genomic DNA. Twenty micrograms of high molecular weight human A431 DNA was digested with enzymes as indicated, electrophoresed through 0.75% gels, and hybridized with either v-*abl* $P_{1.7}$ (lane A) or v-*abl* $P_{0.7}$ (lane B).

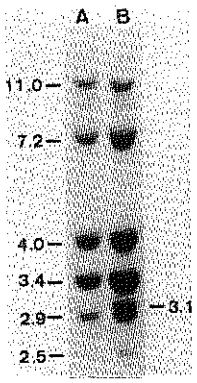


FIG. 5. Comparison of v-*abl* homologous *Eco*RI restriction fragments in total A431 human DNA with those of the molecular cloned human c-*abl* gene. Twenty micrograms of high molecular weight human DNA (lane A) or 0.3 ng of a mixture of cosmid 15 and 18 DNAs plus 20 μg of high molecular weight human DNA (lane B) was digested with *Eco*RI, electrophoresed through a 0.75% gel, transferred to nitrocellulose, and hybridized to the v-*abl* $P_{1.7}$ probe.

the 5′ region of v-*abl*, the possibility was considered that these missing fragments might represent DNA sequences at the extreme 5′ end of the human c-*abl* region, not present in these clones. Alternatively, detection of these sequences in total human DNA after hybridization could reflect the fact that v-*abl* $P_{1.7}$ probe contains approximately 0.6 kb of Moloney MuLV *gag* sequences. To test these possibilities, probes were prepared from plasmid pHEP-Ab by digestion with *Sma*I, *Bst*EII, and *Bam*HI (Fig. 2). In this way, the v-*abl* $P_{1.7}$ region was subdivided into different domains. Hybridization of each of these probes to *Eco*RI-digested high molecular weight human DNA (Fig. 6, panel A) to cos 15 (Fig. 6, panel B), and cos 18 (data not shown) confirmed that all *Eco*RI fragments except the 4.1- and 11-kb fragments were present in the clones. In the total cellular DNA, however, the 4.1- and 11-kb fragments lacked detectable homology to a probe specific for the 5′ part of v-*abl* and the 0.6 kb of Moloney MuLV sequences (v-*abl* $S_{1.1}$), but did hybridize to the v-*abl* $S_{0.45}$ probe. The possibility that the 4.1- and 11-kb fragments are Moloney MuLV specific is thus excluded. Furthermore, these latter fragments cannot represent 5′ c-*abl* sequences located upstream from the major c-*abl* locus, because they do not hybridize to the v-*abl* $S_{1.1}$ extreme 5′ probe. We therefore conclude that cosmid clones 8, 15, and 18 encompass a DNA sequence representing the complete v-*abl* human

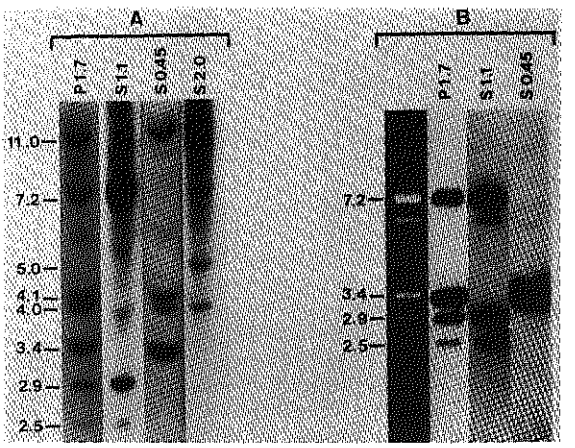


FIG. 6. Comparison of restriction endonuclease patterns of human genomic DNA and molecularly cloned v-*abl* homologous sequences in cosmid 15. Twenty micrograms of high molecular weight A431 human DNA **(A)** or 0.4 μg of DNA from cosmid clone 15 **(B)** was digested with *Eco*RI, electrophoresed through a 0.75% agarose gel, and blotted onto nitrocellulose. Probes used are indicated above each lane. The first lane in (B) shows the ethidium bromide-stained gel of cosmid 15 DNA.

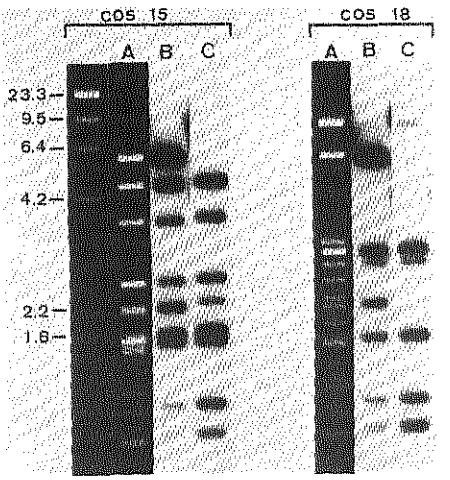


FIG. 7. Repetitive DNA sequences in the human c-*abl* region. Cosmid 15 and 18 DNAs (0.4 μg each) were digested with *Bgl*II and *Eco*RI, electrophoresed through a 0.75% agarose gel, and transferred to nitrocellulose. Lane A, ethidium bromide-stained gel; lane B, hybridization to ^{32}P-labeled DNA of a previously described cosmid clone containing a 35-kb cellular insert representing the complete human c-*fes* gene; lane C, hybridization to a ^{32}P-labeled probe of total human DNA. High molecular weight human DNA was digested with *Eco*RI and size fractionated on a 10–40% sucrose gradient in 20 m*M* Tris-HCl, pH 8.0, 5 m*M* EDTA, 1 *M* NaCl. DNA (300 ng) in the size range 4–5 kb was nick translated. ^{32}P-Labeled *Hind*III-digested λDNA is included as a molecular weight standard in extreme left lane.

homologue and that the 4.1- and 11-kb v-*abl* homologous *Eco*RI restriction fragments correspond to separate genetic loci.

Human c-*abl* Locus Contains Repetitive Sequences

There is accumulating evidence for the presence of repetitive DNA sequences mapping in proximity to cellular coding sequences (exons) (29,30). To test the 64-kb c-*abl* homologous region for such sequences, restriction fragments were analyzed for hybridization to a probe of total human cellular DNA. As shown in Fig. 7, lane C, several *Eco*RI/*Bgl*II restriction fragments in cos 15 and 18 were found to hybridize to this probe. Moreover, as shown in lane B, each of the repetitive sequences within the c-*abl* locus also hybridized to a cosmid clone (cos 5), containing the entire human homologue of v-*fes*, a second cellular oncogene whose viral counterpart encodes a tyrosine-specific protein kinase. The extra positive restriction fragments seen in lane B reflect hybridization of vector sequences common to the c-*abl* and c-*fes* cosmid clones. Upon further restriction mapping analysis, the positions of these repetitive sequences within the region of c-*abl* have been determined (Fig. 2). Repeats mapped at essentially the same positions, regardless of whether the total DNA or c-*fes* probe was used for their detection.

Identification of v-*abl* Related Sequences ($S_{0.45}$ Homologous) in Cosmid Clones 2, 3, and 10

To determine the relationship of the 4.1- and 11-kb v-*abl* homologous *Eco*RI restriction fragments to the major c-*abl* locus, additional cosmid clones were analyzed for v-*abl* $S_{0.45}$ homologous sequences. Although none of the six clones screened contained the 11-kb restriction fragment, three clones were found to contain a 4.1-kb restriction fragment with homology to the v-*abl* $S_{0.45}$ probe (Fig. 8). The lack of detectable overlapping restriction fragments between the cellular inserts in these clones and the major human c-*abl* locus indicates the existence, within human genomic DNA, of at least two independent regions with homology to one of the v-*abl* internal restriction fragments, v-*abl* $S_{0.45}$. The relationship of these sequences both to each other and to the major c-*abl* locus remains to be established.

Cos 4, 5, 7, and 16 Contain v-*abl* $P_{0.7}$ Homologous Sequences

In contrast to the above-described cosmid clones, the four remaining clones isolated in the initial screening (cos 4, 5, 6, and 16) hybridized exclusively to the v-*abl* $P_{0.7}$ probe (Table 1). As these clones exhibit extensive

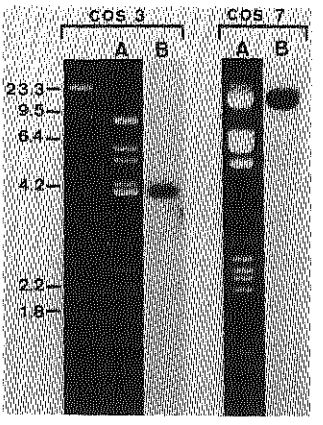


FIG. 8. Identification of v-*abl* homologous sequences in cos 3 and cos 7. Cosmid DNAs (0.4 μg) were digested with *Eco*RI, electrophoresed through a 0.75% agarose gel, and transferred to nitrocellulose. Lane A, ethidium bromide-stained gels; lane B (left), hybridization to v-*abl* $S_{0.45}$; lane B (right), hybridization to v-*abl* $P_{0.7}$. ^{32}P-Labeled *Hin*dIII-digested λDNA is included as a molecular weight standard in extreme left lane.

homology to one another, they appear to represent a single contiguous DNA region. All contain, in part or in whole, a 12.0-kb *Eco*RI restriction fragment with homology to v-*abl* $P_{0.7}$. The fact that this particular fragment was not detected in *Eco*RI-digested total cellular DNA using the same probe argues that the extent of v-*abl* homology within the sequence represented in these four cosmid clones is much less than that described above for cos 5, 15, and 18. To exclude the possibility that the 12.0-kb *Eco*RI fragment might represent an artifact generated during cloning, a probe of 0.7 kb representing an internal *Ava*I restriction fragment of the 12.0-kb cloned v-*abl* homologous sequence was prepared. Using this latter probe, a 12.0-kb *Eco*RI fragment could be demonstrated in total human cellular DNA (data not shown), establishing that the v-*abl* $P_{0.7}$ homologous sequences in cosmid clones 4, 5, 7, and 16 are representative of nonrearranged cellular sequences.

Although cos 4, 5, 7, and 16 showed some cross-homology to clones 8, 15, and 18, even when posthybridization washings were performed under stringent conditions, no overlapping restriction fragments were detected. Since the human c-*abl* locus contains numerous repetitive sequences, this hybridization is probably due to repeats common to these clones. Thus the v-*abl*-related sequences represented in cos 4, 5, 7, and 16 appear to represent a different region of human genome from the primary c-*abl* locus. Moreover, if the sequences in cos 4, 5, 7, and 16 are contiguous with sequences at the 3′ end of clone 18, they must be separated by at least 26 kb.

Transfection Analysis of Molecularly Cloned Human c-*abl* Sequences

Since the cosmid vector used for the cloning of v-*abl* homologous sequences in humans contains the TK gene, transfection of these clones to a recipient TK^- Rat-2 cell line was possible, using a HAT selection system to ensure that the transfection was successful. Upon such analysis, using molecularly cloned v-*abl* (pHEP-Ab) as a positive control, none of the three cosmids containing inserts representing portions of the major c-*abl* locus exhibited detectable transforming activity. Similarly, cosmid clones containing either v-*abl* $S_{0.45}$ (2, 3, and 10) or $P_{0.7}$ (4, 5, 7, and 16) homologous sequences were analyzed and found to lack transforming activity.

DISCUSSION

A cosmid library constructed of human lung carcinoma DNA has been applied to the molecular cloning and restriction endonuclease analysis of human cellular sequences (c-*abl*) homologous to the Abelson MuLV transforming gene v-*abl*. A contiguous human DNA sequence, 64 kb in length, represented by overlapping cellular inserts of three independent cosmid clones was identified and shown to contain v-*abl* homologous sequences distributed discontinuously over a region of 32 kb, as compared with

the v-*abl* gene, which is approximately 3 kb in length. This region of human genomic DNA exhibits homology with probes representing the complete acquired sequence region (v-*abl* gene) of the Abelson MuLV genome. Similarly, mouse cellular c-*abl* homologous sequences are spread over a region of 11–20 kb (7). Based on hybridization with probes corresponding to subgenomic regions of v-*abl*, the arrangement of human c-*abl* sequences is shown to be colinear with that of the viral v-*abl* gene.

On the basis of extent of hybridization with ^{32}P-labeled probes corresponding to the complete v-*abl* gene, six regions of nonhomology (intervening sequences) are identified within the human c-*abl* gene. This, however, is based on hybridization to restriction fragments and thus represents a probable overestimation of the extent of homology and an underestimation of the number of intervening sequences. Moreover, the degree to which v-*abl* resembled the mouse c-*abl* gene at the time of the initial recombination event involved in its derivation, the extent of modification of the v-*abl* gene subsequent to recombination with Moloney MuLV, and the degree to which the mouse and human c-*abl* genes have diverged relative to each other are unknown. For instance, in addition to the Abelson MuLV isolate utilized in the present study, a second variant has been described that contains around 0.8 kb of additional acquired cellular sequence and encodes a somewhat larger polyprotein, P160 (31). Presumably, the smaller form of v-*abl* represents a deletion of the larger variant. On the basis of these considerations, it is clear that coding and intervening sequences of the human c-*abl* gene cannot be accurately defined strictly on the basis of homology to v-*abl*. A more precise determination of the exact positions of the human c-*abl* coding sequence will require identification and characterization of its mRNA transcript(s). It is apparent, however, based on a comparison to v-*abl*, that >90% of the human c-*abl* gene represents probable noncoding sequence.

The present findings indicate differences in the extent of hybridization between different regions of the v-*abl* and human c-*abl* genes. Although probes corresponding to two of the v-*abl* *Pst*I fragments, v-*abl* $P_{1.7}$ and v-*abl* $P_{0.7}$, exhibit strong homology with c-*abl*, the two others, v-*abl* $P_{0.25}$ and v-*abl* $P_{0.85}$, hybridize to much lesser extents. As one possible explanation of these observations, regions of v-*abl* corresponding to the latter two probes may not be required for Abelson MuLV-induced transformation and consequently may be subject to much greater rates of evolutionary divergence. The possibility that at least one of these probes, v-*abl* $P_{0.85}$, may correspond to a region nonessential for expression of transformation is indicated by the isolation of viral mutants with deletions in the 3′ region of v-*abl* that retain transforming activity (31).

The demonstration of extensive regions of DNA repetitive sequences within the human c-*abl* locus is consistent with recent reports by others indicating the frequent existence of DNA repeats in close proximity to cellular coding sequences (29,30). It is not known, however, whether the repetitive sequences observed in the present study correspond to multiple copies of a single sequence or alternatively represent diverse members of a broader family of repetitive sequences. If the latter possibility should prove to be the case, the finding that each of the restriction fragments containing repetitive sequences, as measured by a total cellular DNA probe, also contains one or more repeats in common with the 35-kb human c-*fes*-containing cellular sequence may be of significance.

In addition to cosmid clones representing the major human c-*abl* genetic locus, evidence for as many as three other regions of v-*abl* homology is presented. These include a 12.0-kb *Eco*RI restriction fragment that hybridizes exclusively to a probe corresponding to a single 0.7-kb *Pst*I restriction fragment mapping within the 3′ half of v-*abl* ($P_{0.7}$) and fragments of 4.1 and 11.0 kb with homology to v-*abl* $S_{0.45}$. Although the relationship of these secondary regions of homology to the primary c-*abl* gene is

unclear, a number of possibilities can be considered. For instance, the secondary regions may correspond to highly divergent independent loci that arose because of transposition of portions of, or even the complete, c-*abl* locus, and as such may represent pseudogenes. Alternatively, the homology of these secondary sequences to v-*abl* may reflect relatedness of distinct but functionally analogous cellular genes. As an example of the latter possibility, one or more of the secondary v-*abl* homologous sequences could correspond to other cellular oncogenes with tyrosine-specific protein kinase activity. Finally, it is not resolved as to whether these secondary v-*abl* homologous sequences map in close proximity to each other or are randomly distributed throughout the human genome.

None of the above-described c-*abl* containing cosmid clones exhibit transforming activity on transfection to Rat-2 cells. This does not rule out the possibility that the human c-*abl* gene has such activity, since two of the clones, 8 and 18, are missing 5′ regions that are likely to contain promoter sequences. Clone 15 is truncated and lacks sequences homologous to the 3′ terminal of the v-*abl* gene. Each of the remaining cosmid clones contains sequences corresponding to only limited regions of v-*abl*, and thus their lack of transforming activity is not unexpected. Although on molecular cloning from normal tissues the Harvey murine sarcoma virus c-*ras* gene lacks transforming activity (32), an active form of the gene is found in certain bladder carcinomas cell lines (3–5); even by appropriate *in vitro* modification, the inactive molecularly cloned c-*ras* gene can be converted to an active form (32). Similarly, the human c-*abl* gene may have oncogenic potential that becomes realized only subsequent to specific modification.

Acknowledgments: The authors thank F. Grosveld for helpful discussions and G. T. Blennerhassett and P. Hanson for their excellent technical assistance. This work was supported under Contract No. NO1-CO-76380 from the National Cancer Institute, Bethesda, Maryland.

REFERENCES

1. Fischinger PJ. Type C RNA transforming viruses. In: Stephenson JR, ed. *Molecular biology of RNA tumor viruses*. New York: Academic Press, 1980:162–94.
2. Klein G. *Advances in viral oncology*. New York: Raven Press, 1982.
3. Parada LF, Tabin CJ, Shih C, Weinberg RA. Human EJ bladder carcinoma oncogene is homologue of Harvey sarcoma virus *ras* gene. *Nature* 1982;297: 474–8.
4. Der CJ, Krontiris TG, Cooper GM. Transforming genes of human bladder and lung carcinoma cell lines are homologous to the *ras* genes of Harvey and Kirsten sarcoma viruses. *Proc Natl Acad Sci USA* 1982;79:3637–40.
5. Santos E, Tronick SR, Aaronson SA, Pulciani S, Barbacid M. T24 human bladder carcinoma oncogene is an activated form of the normal human homologue of BALB- and Harvey-MSV transforming genes. *Nature* 1982;298:343–7.
6. Scolnick EM, Howk RS, Anisowicz A, Peebles P, Scher CD, Parks WP. Separation of sarcoma virus-specific and leukemia virus-specific genetic sequences of Moloney sarcoma virus. *Proc Natl Acad Sci USA* 1975;72:4650–4.
7. Goff SP, Gilboa E, Witte ON, Baltimore D. Structure of the Abelson murine leukemia virus genome and the homologous cellular gene: Studies with cloned viral DNA. *Cell* 1980;22:777–85.
8. Rosenberg N, Baltimore D, Scher C. *In vitro* transformation of lymphoid cells by Abelson murine leukemia virus. *Proc Natl Acad Sci USA* 1975;72: 1932–6.
9. Abelson HT, Rabstein LS. Lymphosarcoma virus-induced thymic independent disease in mice. *Cancer Res* 1970;30:2213–22.
10. Reynolds FH Jr, Sacks TL, Deobagkar DN, Stephenson JR. Cells nonproductively transformed by Abelson murine leukemia virus express a high molecular weight polyprotein containing structural and nonstructural components. *Proc Natl Acad Sci USA* 1980;75:3974–8.
11. Witte ON, Rosenberg N, Paskind M, Shields A, Baltimore D. Identification of an Abelson murine leukemia virus-encoded protein present in transformed fibroblasts and lymphoid cells. *Proc Natl Acad Sci USA* 1978;75:2488–92.
12. Van de Ven WJM, Reynolds FH Jr, Stephenson JR. The nonstructural components of polyproteins encoded by replication-defective mammalian transforming retroviruses are phosphorylated and have associated protein kinase activity. *Virology* 1980; 101:185–97.
13. Witte ON, Dasgupta A, Baltimore D. Abelson murine leukemia virus protein is phosphorylated *in vitro* to form phosphotyrosine. *Nature* 1980;283:826–31.
14. Reynolds FH Jr, Van de Ven WJ, Stephenson JR. Abelson murine leukemia virus transformation-defective mutants with impaired P120-associated

protein kinase activity. *J Virol* 1980;36:374–86.
15. Sefton BM, Hunter T, Raschke WC. Evidence that the Abelson virus protein functions *in vivo* as a protein kinase that phosphorylates tyrosine. *Proc Natl Acad Sci USA* 1981;78:1552–6.
16. Witte ON, Goff S, Rosenberg N, Baltimore D. A transformation defective mutant of Abelson murine leukemia virus lacks protein kinase activity. *Proc Natl Acad Sci USA* 1980;77:4993–7.
17. Blomberg J, Reynolds FH Jr, Van de Ven WJM, Stephenson JR. Abelson murine leukemia virus transformation involves loss of epidermal growth factor-binding sites. *Nature* 1980;286:504–7.
18. Twardzik DR, Todaro GJ, Marquardt H, Reynolds FH Jr, Stephenson JR. Abelson MuLV induced transformation involves production of a polypeptide growth factor. *Science* 1982;216:894–7.
19. Grosveld FG, Dahl H-HM, de Boer E, Flavell RA. Isolation of β-globin-related genes from a human cosmid library. *Gene* 1981;13:227–37.
20. Groffen J, Heisterkamp N, Grosveld F, Van de Ven WJM, Stephenson JR. Isolation of human oncogene sequences (v-*fes* homolog) from a cosmid library. *Science* 1982;216:1136–8.
21. Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 1975;98:503–17.
22. Bernards R, Flavell RA. Physical mapping of the globin gene deletion in hereditary persistence of foetal haemoglobin (HPFH). *Nucleic Acids Res* 1980;8:1521–34.
23. Flavell RA, Kooter JM, de Boer E, Little PFR, Williamson R. Analysis of the β-δ-globin gene loci in normal and Hb Lepore DNA: Direct determination of gene linkage and intergene distance. *Cell* 1978;15:25–41.
24. Srinivasan A, Reddy EP, Aaronson SA. Abelson murine leukemia virus: Molecular cloning of infectious integrated proviral DNA. *Biochemistry* 1981;78:2077–81.
25. Morrison DA. Transformation and preservation of competent bacterial cells by freezing. In: Wu R, ed. *Methods in enzymology, vol 68*. New York: Academic Press, 1980:326–31.
26. Topp WC. Normal rat cell lines deficient in nuclear thymidine kinase. *Virology* 1981;113:408–11.
27. Graham FL, Van der Eb AJ. Transformation of rat cells by DNA of human adenovirus 2. *Virology* 1973;54:536–9.
28. Wigler M, Pellicer A, Silverstein S, Axel R. Biochemical transfer of single-copy genes using total cellular DNA as donor. *Cell* 1978;14:725–31.
29. Shen C-KJ, Maniatis T. The organization of repetitive sequences in a cluster of rabbit β-like globin genes. *Cell* 1980;19:379–91.
30. Jelinek WR, Toomey TP, Leinwand L, et al. Ubiquitous, interspersed repeated sequences in mammalian genomes. *Proc Natl Acad Sci USA* 1981;77:1398–402.
31. Ziegler SF, Whitlock CA, Goff SP, Gifford A, Witte ON. Lethal effect of the Abelson murine leukemia virus transforming gene product. *Cell* 1981;27:477–86.
32. Chang EH, Furth ME, Scolnick EM, Lowy DR. Tumorigenic transformation of mammalian cells induced by a normal human gene homologous to the oncogene of Harvey murine sarcoma virus. *Nature* 1982;297:479–83.

APPENDIX PAPER VIII

Homology between phosphotyrosine acceptor site of human c-*abl* and viral oncogene products

John Groffen, Nora Heisterkamp, Fred H. Reynolds Jr & John R. Stephenson

Laboratory of Viral Carcinogenesis, National Cancer Institute-FCRF, Frederick, Maryland 21701, USA

The human homologues of several independent viral oncogenes, each of which encodes tyrosine-specific protein kinases, have been identified[1,2]. Of these, three (v-*src*, v-*yes* and v-*fes*/*fps*) are known to exhibit considerable sequence homology, particularly in the regions of their phosphorylation acceptor sites[3-5]. In the present study, sequences encoding the tyrosine phosphorylation acceptor sites of the Abelson murine leukaemia virus oncogene, v-*abl*, and its human cellular homologue, c-*abl*, have been identified and their nucleic acid sequences determined. Our results establish extensive homology between this region of c-*abl* and acceptor domains of the v-*src*[3,6], v-*yes*[6] and v-*fes*/*fps*[4,5] family of viral oncogenes, as well as more distant relatedness to the catalytic chain of the mammalian cyclic AMP-dependent protein kinase[7]. These findings suggest that, of the homologues of retroviral oncogenes with tyrosine protein kinase activity examined to date, all were probably derived from a common progenitor and may represent members of a diverse family of cellular protein kinases.

We have described previously the application of a cosmid vector system to the molecular cloning of the human cellular homologue of acquired cellular (v-*abl*) sequences of the Abelson strain of murine leukaemia virus[8]. v-*abl* homologous sequences were found to encompass six separate coding regions (exons) distributed over a total of 32 kilobases (kb)[9] localized on human chromosome 9(q34) (refs 9, 10). Figure 1 shows the positioning of such sequences, as delineated by restriction endonuclease and hybridization analysis. Hybridization of individual exons to a series of *Sma*I restriction fragments corresponding to specific regions of v-*abl* is also indicated. By transfection analysis, v-*abl*-specific sequences required for transforming activity have been localized to the 5′ half of the acquired region within the viral genome[11]. Moreover, a variant of Abelson-MuLV characterized by an in-phase interstitial deletion mapping within the 5′ region of the v-*abl* acquired sequence has been shown to lack transforming activity[12]. Since this latter deletion approximates that of a 0.45-kb *Sma*I restriction fragment of v-*abl*, shown in Fig. 1, we reasoned that human c-*abl* sequences homologous to the $S_{0.45}$ probe might correspond to its phosphotyrosine acceptor site. As shown in Fig. 1, most of the $S_{0.45}$ homologous sequences could be localized to a single 3.4-kb *Eco*RI restriction fragment encompassing portions of three exons and mapping within the middle of the human v-*abl* homologous region. Subcloning in pBR328 and fine mapping revealed that this fragment contained a 0.14-kb *Bgl*II/*Ava*I restriction fragment which hybridized strongly to the $S_{0.45}$ probe.

In view of its homology to the transformation-specific region of v-*abl* and consequently the possibility that it might correspond to the active region of the c-*abl* protein kinase, the 0.14-kb *Bgl*II/*Ava*I restriction fragment was subjected to nucleic acid sequence analysis. Figure 2 shows that of the three possible reading frames, the first contained six termination codons, the second contained a single TGA stop codon near the middle of the sequence, not followed by any obvious splice acceptor site, while the third lacked any termination codons and thus appeared to represent the authentic coding sequence for the c-*abl* gene product. To extend this analysis, we subsequently sequenced the coding region of an adjacent downstream *Ava*I/*Ava*I c-*abl* restriction fragment, and by an analogous approach, the corresponding region of v-*abl*. The predicted amino acid sequences corresponding to each of these regions of nucleic acid sequence are shown in Fig. 2. The predicted amino acid sequences of the internal 61 residues were identical based on either the v-*abl* or c-*abl* nucleic acid sequence. Sixteen additional amino-terminal and 18 carboxy-terminal residues correspond to regions for which only v-*abl* sequence was determined.

We investigated whether the deduced c-*abl*/v-*abl* amino acid sequence corresponds to phosphotyrosine acceptor site regions common to the v-*src*-, v-*yes*- and v-*fes*/*fps*-encoded transforming proteins by comparing relevant sequences. As shown in Fig. 3, the c-*abl*/v-*abl* sequence contains 38 out of 96 residues in common with all four of the viral acceptor site regions examined. Additional relatedness with individual viral sequences appears rather random; some c-*abl*/v-*abl* residues are common to either v-*yes* or v-*src* while others are shared only with the v-*fes* and v-*fps* gene products. The c-*abl*/v-*abl*-encoded proteins also contain a tyrosine residue at the position corresponding to the previously identified tyrosine acceptor sites of each of the other oncogene-encoded products examined. In addition, there is considerable homology (28 of 96 residues) between c-*abl* and the catalytic subunit of the mammalian cyclic AMP-dependent protein kinase (BOV-PK), although the tyrosine acceptor residue of c-*abl*/v-*abl* is replaced by a tryptophan residue in BOV-PK. Interestingly, this substitution is directly adjacent to a threonine residue known to represent one of two phosphorylation sites previously identified in BOV-PK[7].

To establish whether the putative c-*abl* tyrosine phosphorylation site is recognized by functionally active viral protein kinases, a synthetic peptide with the sequence NH_2-Ser-Arg-

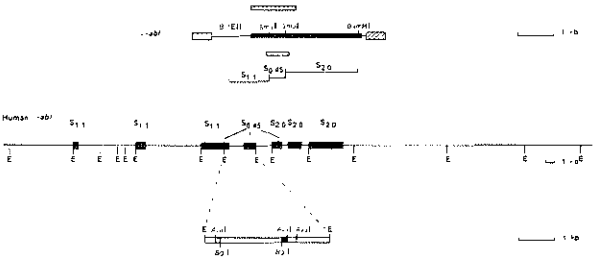


Fig. 1 Schematic representation of Abelson-MuLV proviral DNA and the human v-*abl* homologous (c-*abl*) sequence. Positions of long terminal repeats (▨), acquired cellular sequences (■), region of the acquired sequences necessary to retain transforming activity (▩), a deletion within a transformation-defective variant of Abelson MuLV (▩), and a series of three *Sma*I restriction fragments (⊔) used as probes for analysis of c-*abl* sequences, are shown in the upper portion of the figure. v-*abl* homologous sequences, shown as solid boxes in the human c-*abl* genetic locus, were delineated by restriction endonuclease digestion and molecular hybridization analysis as described previously[8]. Specific v-*abl* probes, to which the individual c-*abl* exons hybridize, are also shown. A restriction map of the phosphotyrosine acceptor site region of c-*abl* is shown in more detail directly below the c-*abl* map; extents of hybridization to the $S_{0.45}$ probe are indicated as negative (⊐), weak (▨), medium (▩), or strong (■).

Leu-Met-Thr-Gly-Asp-Thr-Tyr-Thr-Ala-His-Ala-Gly-Ala-Lys-COOH was chemically synthesized and purified by HPLC. On addition of Abelson-MuLV P120 to immunoprecipitates, the c-*abl* acceptor peptide was phosphorylated 10-fold more efficiently than in the same reaction conditions using [Val^5]angiotensin II as substrate[13]. Following re-purification of the labelled peptide by HPLC and subsequent two-dimensional phosphoamino acid analysis, ^{32}P incorporation by the synthetic c-*abl* acceptor site peptide was found to be exclusively in tyrosine.

Table 1 summarizes the relatedness of the amino acid sequences within the acceptor regions of representative tyrosine-specific protein kinases and the corresponding nucleic acid sequences from which these were deduced. Although c-*abl* is somewhat more related to v-*src*, especially with respect to nucleic acid sequence, than to the other viral oncogene sequences examined, it is less related to v-*src* than is v-*yes*. Interestingly, of the five oncogene acceptor sequences analysed, c-*abl* is most closely related to BOV-PK. It is also apparent that, in general, the oncogene acceptor site regions are much less conserved with respect to nucleic acid sequence than in the amino acids they encode, indicating that much of the nucleic acid sequence divergence reflects third position or other silent nucleotide changes not resulting in an altered gene product.

The present findings establish the evolutionary relatedness of c-*abl*/v-*abl* to four retroviral oncogenes, each of which contains transformation-specific sequences with tyrosine-specific protein kinase activity. In fact, despite its human cellular origin, the phosphotyrosine acceptor site region of c-*abl* is as closely related to analogous regions of the v-*src*, v-*yes* and v-*fes*/v-*fps* viral oncogenes as these are to each other. Although an additional retroviral oncogene, v-*ros*, with similar enzymatic activity has been described, its relationship to those genes examined in the present study is unknown[14]. Human c-*abl*, c-*src* and c-*fes*/c-*fps* are sufficiently diverged to lack detectable cross-hybridizing sequences and have been mapped to different human chromosomes[9,10,15,16]. The mammalian v-*fes* and avian v-*fps* genes, although independently identified, correspond to a common mammalian cellular genetic locus, c-*fes*/c-*fps*[17,18]. The human homologue of the fourth viral oncogene, v-*yes*, remains poorly defined.

Evidence that the c-*abl* phosphotyrosine acceptor site identified in the present study is functionally active derives from its *in vitro* phosphorylation by the v-*abl* protein kinase. Moreover, the identity of this site to that of v-*abl*, and its close

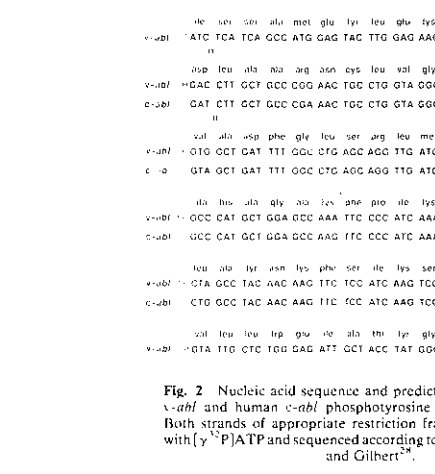


Fig. 2 Nucleic acid sequence and predicted amino acids of the v-*abl* and human c-*abl* phosphotyrosine acceptor site regions. Both strands of appropriate restriction fragments were labelled with [γ-^{32}P]ATP and sequenced according to the method of Maxam and Gilbert[28].

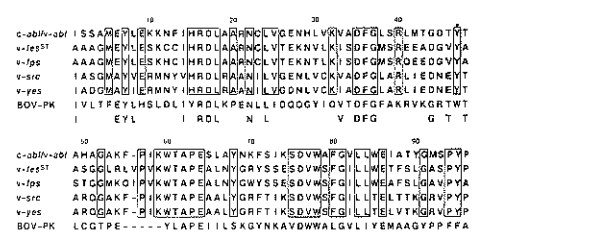


Fig. 3 Comparison of the deduced amino acid sequence of the v-*abl* and human c-*abl* phosphotyrosine acceptor sites to analogous regions of the v-*fes*ST- (ref. 4), v-*fps*-[5], v-*src*-[6] and v-*yes*-[6] encoded transforming proteins and to the catalytic subunit of the mammalian cyclic AMP-dependent protein kinase[7]. Amino acids common to each of the oncogene-encoded acceptor domains are shown in open boxes, while those residues shared by the c-*abl*-encoded sequence and BOV-PK are indicated in the lower portion of the figure. The position of the putative phosphotyrosine acceptor residue (*) is also shown. A, alanine; G, glycine; T, threonine; C, cysteine; P, proline; M, methionine; V, valine; I, isoleucine; L, leucine; Y, tyrosine; F, phenylalanine; W, tryptophan; D, aspartic acid; E, glutamic acid; R, arginine; K, lysine; H, histidine; S, serine; N, asparagine; Q, glutamine.

resemblance to the well characterized v-*src*, v-*yes* and v-*fes*/v-*fps* tyrosine phosphorylation acceptor sites strongly favours the possibility that this site has functional activity. The major *in vitro* and *in vivo* phosphorylated tyrosine residue of v-*abl* has been localized seven residues distal to a trypsin cleavage site[19,20], consistent with the positioning of the v-*abl* and human c-*abl* tyrosines seven amino acids distal to arginine residues. An analogous approach was first used to identify the acceptor sites common to the v-*src*, v-*yes* and v-*fes*/v-*fps* gene products[21–24]. Although the requirement for a functionally active tyrosine-specific protein kinase for v-*abl*-associated transformation has been demonstrated using transformation-defective mutants[25,26], the involvement of the major tyrosine acceptor in this process remains to be determined.

The conservation of phosphotyrosine acceptor sequences throughout vertebrate evolution, and the finding that most of the observed divergence between different genes encoding such sequences involves silent base substitutions, argues that each confers on its host some essential biological function. Presumably, divergence of these sequences occurred early in vertebrate evolution before separation of avian and mammalian classes. The possible involvement of tyrosine-specific cellular protein kinases in human cancer remains to be determined. However, the high frequency with which recombinant transforming

Table 1 Segment comparison scores of human c-*abl* and representative viral oncogene phosphotyrosine acceptor sites and BOV-PK

Sequence	SD units				
	c-*abl*	v-*fes*	v-*fps*	v-*src*	v-*yes*
c-*abl*	—				
v-*fes*	17.6(6.7)	—			
v-*fps*	15.6(7.8)	19.5(16.9)	—		
v-*src*	18.9(11.4)	11.9(3.2)	13.6(3.0)	—	
v-*yes*	17.1(5.2)	15.7(0.2)	9.8(1.6)	42.5(15.8)	—
BOV-PK	7.6(−)	5.4(−)	4.8(−)	5.7(−)	4.5(−)

Extents of relatedness of amino acid sequences designated 17–64 in Fig. 3 and corresponding nucleic acid sequences from which they were derived. Results are expressed in SD units based on 20-residue segment lengths according to the Relate program as described by Barker and Dayhoff[27]. A score of >5 SD indicates evolutionary relatedness of two proteins and scores between 3 and 5 SD support relationship if there are other indications such as a similarity of function[27].

viruses containing such sequences have been isolated and the recent localization of c-*abl* within a small fragment of human chromosome 9 which is translocated to chromosome 22 in chronic myelogenous leukaemia[9,10], would seem to favour such a model.

We thank G. T. Blennerhassett and P. Hansen for technical assistance and K. L. McNitt (Computer and Statistical Services, IMS) for help with calculation of the sequence homology. This work was supported under NCI contract NOI-CO 76380.

Received 2 March; accepted 16 May 1983.

1. Varmus, H. E. *Science* **216,** 812–820 (1982).
2. Bishop, J. M. *Adv. Cancer Res.* **37,** 1–32 (1982).
3. Czernilofsky, A. P. *et al. Nature* **287,** 198–203 (1980).
4. Hampe, A., Laprevotte, I., Galibert, F., Fedele, L. A. & Sherr, C. J. *Cell* **30,** 775–785 (1982).
5. Shibuya, M. & Hanafusa, H. *Cell* **30,** 787–795 (1982).
6. Kitamura, N., Kitamura, A., Toyoshima, K., Hirayama, Y. & Yoshida, M. *Nature* **297,** 205–208 (1982).
7. Shoji, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **78,** 848–851 (1981).
8. Heisterkamp, N., Groffen, J. & Stephenson, J. R. *J. molec. appl. Genet.* **2,** 57–68 (1983).
9. Heisterkamp, N. *et al. Nature* **299,** 747–749 (1982).
10. de Klein, A. *et al. Nature* **300,** 765–767 (1982).
11. Srinivasan, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **79,** 5508–5512 (1982).
12. Goff, S. P., Gilboa, E., Witte, O. N. & Baltimore, D. *Cell* **22,** 777–785 (1980).
13. Wong, T. W. & Goldberg, A. R. *J. biol. Chem.* **258,** 1022–1025 (1983).
14. Feldman, R. A., Wang, L.-H., Hanafusa, H. & Balduzzi, P. C. *J. Virol.* **42,** 228–236 (1982).
15. Dalla-Favera, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **79,** 4714–4717 (1982).
16. Sakaguchi, A. Y., Naylor, S. L. & Shows, T. B. *Prog. Nucleic Acid Res. molec. Biol.* (in the press).
17. Shibuya, M., Hanafusa, T., Hanafusa, H. & Stephenson, J. R. *Proc. natn. Acad. Sci. U.S.A.* **77,** 6536–6540 (1980).
18. Groffen, J., Heisterkamp, N., Shibuya, M., Hanafusa, H. & Stephenson, J. R. *Virology* **125,** 480–486 (1983).
19. Reynolds, F. H. Jr, Oroszlan, S. & Stephenson, J. R. *J. Virol.* **44,** 1097–1101 (1982).
20. Patschinsky, T., Hunter, T., Esch, F. S., Cooper, J. A. & Sefton, B. M. *Proc. natn. Acad. Sci. U.S.A.* **75,** 973–977 (1982).
21. Smart, J. E. *et al. Proc. natn. Acad. Sci. U.S.A.* **78,** 6013–6017 (1981).
22. Blomberg, J., Van de Ven, W. J. M., Reynolds, F. H. Jr, Nalewaik, R. P. & Stephenson, J. R. *J. Virol.* **38,** 886–894 (1981).
23. Neil, J. E., Ghysdael, J., Vogt, P. K. & Smart, J. E. *Nature* **291,** 675–677 (1981).
24. Reynolds, F. H. Jr, Oroszlan, S., Blomberg, J. & Stephenson, J. R. *Virology* **122,** 134–146 (1982).
25. Reynolds, F. H. Jr, Van de Ven, W. J. M. & Stephenson, J. R. *J. Virol.* **36,** 374–386 (1980).
26. Witte, O. N., Goff, S., Rosenberg, N. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **77,** 4993–4997 (1980).
27. Barker, W. C. & Dayhoff, M. O. *Proc. natn. Acad. Sci. U.S.A.* **79,** 2836–2839 (1982).
28. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65,** 499–560 (1980).

APPENDIX PAPER IX

A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukaemia

Annelies de Klein*, Ad Geurts van Kessel*, Gerard Grosveld*, Claus R. Bartram*, Anne Hagemeijer*, Dirk Bootsma*, Nigel K. Spurr†, Nora Heisterkamp‡, John Groffen‡ & John R. Stephenson‡

* Department of Cell Biology and Genetics, Erasmus University, PO Box 1738 - 3000 DR Rotterdam, The Netherlands
† Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK
‡ Laboratory of Viral Carcinogenesis, National Cancer Institute-FCRF, Frederick, Maryland 21701, USA

The transforming genes of oncogenic retroviruses are homologous to a group of evolutionary conserved cellular *onc* genes[1]. The human cellular homologue (c-*abl*) of the transforming sequence of Abelson murine leukaemia virus (A-MuLV) was recently shown[2] to be located on chromosome 9. The long arm of this chromosome is involved in a specific translocation with chromosome 22, the Philadelphia translocation (Ph[1]), t(9; 22) (q34, q11), which occurs in patients with chronic myelocytic leukaemia (CML)[3-5]. Here we investigate whether the c-*abl* gene is included in this translocation. Using c-*abl* and v-*abl* hybridization probes on blots of somatic cell hybrids, positive hybridization is found when the 22q⁻ (the Philadelphia chromosome), and not the 9q⁺ derivative of the translocation, is present in the cell hybrids. From this we conclude that in CML, c-*abl* sequences are translocated from chromosome 9 to chromosome 22q⁻. This finding is a direct demonstration of a reciprocal exchange between the two chromosomes[6] and suggests a role for the c-*abl* gene in the generation of CML.

The human c-*abl* sequences represent a cellular homologue of the transforming component of A-MuLV. This retrovirus is a recombinant between Moloney MuLV and mouse cellular c-*abl* sequences[7] and induces lymphoid tumours on *in vivo* inoculation of the mouse[8,9]. The major A-MuLV translational product has been identified as a poly-protein, $P120^{gag\text{-}abl}$, consisting of amino-terminal structural proteins encoded by the M-MuLV *gag* gene, linked to an acquired cellular sequence encoded carboxy-terminal component[10,11]. This protein is one of several virus-encoded transforming proteins with tyrosine-specific protein kinase activity[12-15]. Similar oncogenic sequences of Harvey and Kirsten sarcoma virus are homologous to transforming sequences (c-Ha-*ras*, c-Ka-*ras*) isolated from human bladder and lung carcinoma cell lines[16-18]. Both these sequences induce transformation of mouse NIH 3T3 cells after transfection, establishing that the human genes have potential transforming activity. Recently, the human c-*abl* gene has been cloned in cosmids[19]. Using v-*abl* DNA as a probe, several clones containing overlapping sequences representing the entire c-*abl* gene were isolated from a human lung carcinoma cosmid library. The restriction enzyme map of the human v-*abl* cellular homologue, presented in Fig. 1, identifies areas of the gene which hybridize to v-*abl* sequences. The gene is distributed over a region of 40 kilobases (kb) of human DNA and contains multiple intervening sequences. On transfection of Rat-2 cells with the c-*abl* cosmids, no transforming activity was detected, not unexpectedly, as none of the cosmid clones tested contained the entire c-*abl* gene[19].

By Southern blot analysis of a series of somatic cell hybrids, the human c-*abl* gene has been localized on chromosome 9[2]. This finding is of interest because of the involvement of the long arm of chromosome 22 (band 22q11) in a specific translocation with the long arm of chromosome 9 (band 9q34), the Philadelphia translocation (Ph[1]), occurring in human CML[3,4]. The abnormal chromosomes are designated $9q^+$ and $22q^-$; of these, the $22q^-$ chromosome is observed in 92% of CML cases[5]. We investigated the chromosomal location of the human c-*abl* gene in cases of CML, where the Philadelphia translocation is present. Southern blot analyses with c-*abl* and v-*abl* probes were performed on *Eco*RI-digested DNAs from somatic cell hybrids segregating the $9q^+$ and $22q^-$ chromosomes.

The cell hybrids used here contain a full complement of mouse or Chinese hamster chromosomes and a limited number of human chromosomes. The hybrid cell lines have been obtained by fusion of cells from mouse (Pg 19 and WEHI-3B) or Chinese hamster (E36 and a3) origin with leukocytes from different CML patients and from a normal donor[6,20]. The human chromosome content of these cells is summarized in Table 1 and is based on chromosome analysis. In addition, the hybrid cells were tested for the expression of human adenylate kinase-1 (AK1) enzyme activity, a marker localized proximal in band

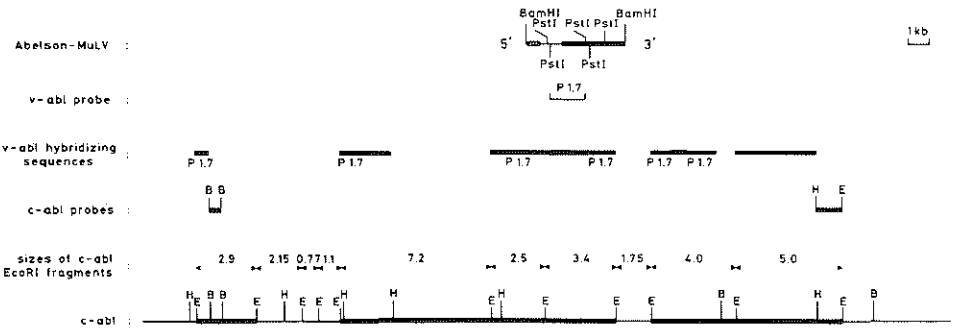


Fig. 1 Restriction enzyme map of the human c-*abl* region[19]. The upper line of the figure shows the *Bam*HI subclone of A-MuLV; the hatched box presents the long terminal repeat, the solid bar the acquired cellular sequences. Directly beneath the A-MuLV genome, a subgenomic *Pst* 1.7-kb fragment, used as a probe in this study, is shown. Human c-*abl* DNA restriction fragments homologous to v-*abl* sequences are indicated as black boxes and those that show homology to the 1.7-kb *Pst* v-*abl* fragments are designated by P 1.7. The third line shows the human c-*abl* 0.6-kb *Bam*HI and 2.2-kb *Hin*dIII–*Eco*RI probes, which hybridize to 5' and 3' c-*abl* *Eco*RI fragments, respectively. The sizes of all *Eco*RI c-*abl* fragments are indicated on the fourth line. The bottom line represents the restriction enzyme map of the human c-*abl* gene. Restriction enzymes include *Bam*HI (B), *Hin*dIII (H) and *Eco*RI (E). A more detailed characterization of the human c-*abl* locus will be published elsewhere[19].

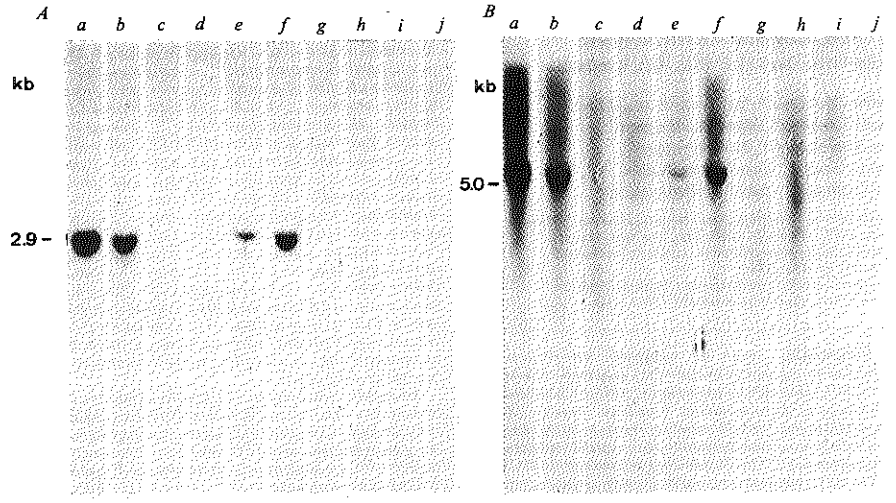


Fig. 2 Localization of human c-*abl* sequences on the Philadelphia chromosome, using hybrid cell lines and human c-*abl* probes. *A*, detection of the human 5′ end 2.9-kb *Eco*RI c-*abl* fragment in DNA from *a*, human placenta; *b*, 10CB-23B (chromosome 9); *c*, PgMe-25NU (chromosome 22q); *d*, 14CB-21A (chromosome 9q$^+$); *e*, 1CN-17aNU (chromosome 22q$^-$); *f*, WESP-2A (chromosome 22q$^-$); *g*, mouse Pg19; *h*, Chinese hamster E36; *i*, mouse WEHI-3B; *j*, Chinese hamster a3. *B*, detection of the 3′-end 5.0-kb *Eco*RI c-*abl* fragment in DNAs as indicated in *A* (*a–j*). The derivations of all these cell lines and their complements of human chromosomes are summarized in Table 1.
Methods: All cell lines used in this experiment were grown in large batches (10^7–10^8 cells) and DNA was prepared as described by Jeffreys and Flavell[28]. *Eco*RI-restricted DNAs (10 μg per lane) from human placenta, hybrid cell lines, mouse and Chinese hamster fusion partners were electrophoresed on 0.7% agarose gels. *Hin*dIII and *Hin*dIII–*Eco*RI-digested λ DNAs were included as molecular weight markers (not shown). After blotting to nitrocellulose, the filters were hybridized to the 0.6-kb *Bam*HI c-*abl* (*A*) or 1.1-kb *Hin*dIII–*Eco*RI c-*abl* (*B*) restriction fragments described in Fig. 1. Hybridization and washing procedures (to 0.1 × SSC at 65 °C) were carried out according to the method of Bernards and Flavell[23].

9q34 (ref. 21). This latter test was necessary to exclude the possibility of hidden (broken or rearranged) chromosome 9 fragments in the 22 and 22q$^-$ cell lines.

Detection of the human c-*abl* restriction fragments in hybrid cell DNAs is often inconclusive using v-*abl* probes, because the human sequences are present in submolar amounts (20–50%) and also because many of the human c-*abl* restriction fragments electrophorese in close proximity with strongly hybridizing mouse or Chinese hamster fragments. To obtain molecular probes with specificity for human c-*abl* sequences, two restriction fragments were isolated from subclones of c-*qbl*-containing cosmids, with homology to the presumptive 5′ and 3′ proximal *Eco*RI fragments of c-*abl*. These are 2.9 and 5.0 kb, respectively, in size (Fig. 1). After hybridization and washing to high stringency (0.1 × SSC)[12], the 5′-terminal 0.6-kb *Bam*HI probe and the 3′-terminal 1.1-kb *Hin*dIII–*Eco*RI probe cross-hybridize to a very low extent with mouse or hamster c-*abl* sequences. Figure 2*A* shows an example of a hybridization experiment with the 0.6-kb *Bam*HI probe. This Southern blot illustrates hybridization of *Eco*RI-restricted DNAs of hybrid cell lines containing chromosomes 22, 9, 9q$^+$ or 22q$^-$. As controls, hybridization of the probe with human placenta DNA and DNA from the mouse and Chinese hamster fusion partners is shown. It is clear that the 2.9-kb *Eco*RI fragment, detected in human placenta DNA, is also present in the lanes containing DNA from the hybrid cell lines 10 CB-23B (chromosome 9), 1CB–17a NU and WESP-2A (both containing chromosome 22q$^-$). The band is not detected in lanes containing DNA from

Table 1 Human chromosome content of human–mouse and human–Chinese hamster somatic cell hybrids

Hybrid	Human chromosomes																							Human isoenzyme AK1	Ref.
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	Y	9q$^+$	22q$^-$	
PgMe-25NU	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–	–	–	–	–	26
10CB-23B	–	–	–	–	+	–	–	–	+	–	+	–	–	–	–	–	–	–	+	–	–	–	–	–	–	–	+	6
14CB-21A	–	–	–	+	–	–	+	+	–	–	–	–	–	+	–	–	–	–	–	+	–	–	–	–	+	–	+	20
1CB-17a NU	–	–	–	–	–	–	–	+	–	–	–	+	–	+	+	–	–	–	+	–	+	–	–	–	–	+	–	6
WESP-2A	–	–	–	–	–	–	+	+	–	–	–	–	–	+	–	–	–	–	–	–	–	–	+	–	–	+	–	×

The origin and details of the initial characterization of the somatic cell hybrids are described in the references listed in the last column. PgMe-25NU and WESP-2A are hybrids obtained from fusions with mouse Pg19 and WEHI-3B cells, respectively. Chinese hamster cell line E36 was used to produce hybrid clones 10CB-23B and 14CB-21A, while Chinese hamster cell line a3 was used to obtain 1CB-17aNU. Chromosome analysis was done using reverse (R) banding with acridine orange, after heat denaturation. At least 16 metaphases were analysed per cell line. The presence of human AK1 activity was assayed by cellulose acetate (Cellogel) electrophoresis[27]. This test is inconclusive for the WESP-2A cell line (×), because the expression of AK1 was found to be repressed in hybrids derived from fusion with WEHI-3B cells (A.H.M., G.v.K., unpublished results). Chromosome and isoenzyme analyses[22] were performed on the same batches of hybrid cells that were used for the isolation of DNA.

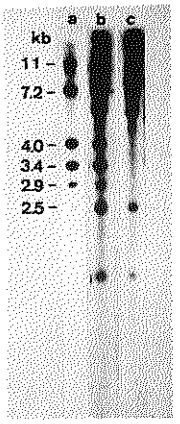


Fig. 3 Localization of human c-*abl* sequences on the Philadelphia chromosome, using a 22q$^-$ somatic cell hybrid and a v-*abl* probe. *Eco*RI-digested DNAs (10 μg) from human placenta (*a*), hybrid WESP-2A (*b*) and mouse WEHI-3B cells (*c*) were hybridized with the 1.7-kb *Pst* v-*abl* fragment (Fig. 1), as described in Fig. 2 legend. After hybridization, the filters were washed to 1 × SSC at 65 °C. Molecular weights of human c-*abl* fragments were deduced from co-electrophoresed *Hin*dIII and *Hin*dIII–*Eco*RI-digested λ DNA markers.

PgMe-25Nu (chromosome 22), 14CB-21A (chromosome 9q$^+$), Pg19 and WEHI-3B (mouse controls) or E36 and a3 (Chinese hamster controls). Analogous results are obtained when the same *Eco*RI-digested DNAs are hybridized to the 3′-terminal 1.1-kb *Hin*dIII–*Eco*RI probe (Fig. 2*B*). The 5.0-kb *Eco*RI fragment is detected only in DNA from human placenta and from hybrid cell lines containing chromosome 22q$^-$ or 9.

The above results show that both the 5′ and 3′ ends of the c-*abl* gene are translocated to chromosome 22q$^-$. Because all other c-*abl* *Eco*RI fragments, which hybridize to v-*abl* sequences, are flanked by the 2.9-kb and 5.0-kb *Eco*RI fragments, it seems highly probable that these fragments are also included in the translocation to the Philadelphia chromosome. To test this possibility directly, hybridization was performed using a 1.7-kb *Pst* v-*abl* probe (Fig. 1). Because of the problems with v-*abl* probes indicated above, only WESP-2A, the hybrid containing the most 22q$^-$ sequences (50% of the molar amount), was tested. As shown in Fig. 3, the viral probe detects human *Eco*RI c-*abl* fragments of 11, 7.2, 4.0, 3.4, 2.9 and 2.5 kb (weakly). Of these fragments, the 11-kb band has been shown to map outside the main human c-*abl* locus[19] and will not be considered here. The human 2.9-, 3.4-, and 4.0-kb c-*abl* fragments are readily detected in the WESP-2A DNA. In contrast, the 7.2-kb *Eco*RI fragment can only be seen in a short exposure of this filter (not shown), due to spill-over of radiation from strongly hybridizing mouse c-*abl* fragments in this area. The 2.5-kb *Eco*RI human c-*abl* fragment co-migrates with a mouse fragment of similar size and thus cannot be identified in this analysis.

The hybrid cell lines containing the 9q$^+$ and 22q$^-$ chromosomes examined in the present study, were obtained from fusion experiments with CML cells from three different individuals. Therefore, we conclude that in the Philadelphia translocation a fragment of chromosome 9 is translocated to chromosome 22q$^-$ and that this fragment includes the human c-*abl* sequences. This finding establishes that the translocation is reciprocal, a general assumption which is now demonstrated unequivocally. Moreover, the data map the human c-*abl* sequences distal to AK1 (not translocated to 22q$^-$, 6, 20) on chromosome 9. The most interesting aspect is that it raises the possibility of involvement of the human c-*abl* gene in the generation of CML.

In principle, the chromosomal translocation associated with CML could lead to elevated levels of c-*abl* expression which, by analogy to the c-Ha-*ras* gene in bladder carcinoma, would induce malignant transformation[24]. Elevated levels of c-*abl* expression could be the result of coupling of the gene to an enhancer sequence present on chromosome 22 or, alternatively, the gene could be linked to a strong promoter of another gene. To test these possibilities, we have initiated studies to clone the c-*abl* gene from the 22q$^-$ chromosome using WESP-2A DNA and a cosmid vector system. Finally, it is of interest that in some CML patients variant Ph^1 translocations are observed, in which the participation of chromosome 9 cannot be detected by classical cytogenetic analysis[3,5]. In another group of CML patients the Ph^1 translocation appears to be completely absent[25]. We are now investigating whether the c-*abl* gene is translocated to chromosome 22 in these cases also.

These studies were initiated as part of a collaborative effort with W. F. Bodmer; his helpful discussions throughout the work are greatly appreciated. We also thank F. Grosveld for important contributions to this study, R. A. Flavell for useful suggestions, Ton van Agthoven, Gail T. Blennerhassett and Pam Hansen for technical assistance and Ad Konings and Rita Boucke for help with the preparation of the manuscript. The work was supported by the Netherlands Cancer Society (Koningin Wilhelmina Fonds) and under NCI PHS contract NOI–CO-75380. C.R.B. is a recipient of a fellowship from the Deutsche Forschungsgemeinschaft.

Received 12 November; accepted 25 November 1982.

1. Coffin, J. M. *et al. J. Virol.* **40**, 953–957 (1981).
2. Heisterkamp, N. *et al. Nature* **299**, 747–750 (1982).
3. Rowley, J. D. *Nature* **243**, 290–293 (1973).
4. Lawler, S. D. *Clin. Haemat.* **6**, 55–75 (1977).
5. Rowley, J. D. *Clin. Haemat.* **9**, 55–86 (1980).
6. Geurts van Kessell, A. H. M. *et al. Cytogenet. Cell Genet.* **30**, 83–91 (1981).
7. Goff, S. P., Gilboa, E., Witte, O. M. & Baltimore, D. *Cell* **22**, 777–785 (1980).
8. Potter, M., Sklar, M. D. & Rowe, W. P. *Science* **182**, 592–594 (1973).
9. Premkumar, E., Potter, M., Singer, P. A. & Sklar, M. D. *Cell* **6**, 149–159 (1975).
10. Witte, O. N., Rosenberg, N., Paskind, M., Shields, A. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2488–2492 (1978).
11. Reynolds, F. H. Jr, Sacks, T. L., Neobogkar, D. N. & Stephenson, J. R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3974–3978 (1978).
12. Van de Ven, W. J. M., Reynolds, F. H. Jr & Stephenson, J. R. *Virology* **101**, 185–197 (1980).
13. Witte, O. N., Dasgupta, A. & Baltimore, D. *Nature* **283**, 826–831 (1980).
14. Blomberg, J., Reynolds, F. H. Jr, Van de Ven, W. J. M. & Stephenson, J. R. *Nature* **286**, 504–507 (1980).
15. Sefton, B. M., Hunter, T. & Raschke, W. C. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1552–1556 (1981).
16. Der, C. J., Krontiris, T. C. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3637–3640 (1982).
17. Parada, L. F., Tabin, C., Shih, C. & Weinberg, R. A. *Nature* **297**, 474–478 (1982).
18. Santos, E., Tronick, S. R., Aaronson, S. A., Pulciani, S. & Barbacid, M. *Nature* **298**, 343–347 (1982).
19. Heisterkamp, N., Groffen, J. & Stephenson, J. R. *J. molec. appl. Genet.* (in the press).
20. Geurts van Kessel, A. H. M., van Agthoven, A. J. & Hagemeijer, A. *Cancer Genet. Cytogenet.* **6**, 55–58 (1981).
21. Ferguson-Smith, M. A. & Aitken, D. A. *Cytogenet. Cell Genet.* **22**, 49–451 (1978).
22. Povey, S., Boyd, Y., Duncan, M. E., Jeremiah, S. J. & Carritt, B. *Cytogenet. Cell Genet.* **22**, 461–464 (1978).
23. Bernards, R. & Flavell, R. A. *Nucleic Acids Res.* **8**, 1421–1533 (1980).
24. Chang, E. H., Furth, M. E., Scolnick, E. M. & Lowy, D. R. *Nature* **297**, 479–483 (1982).
25. Sandberg, A. A. *Cancer Genet. Cytogenet.* **1**, 217–228 (1980).
26. Geurts van Kessel, A. H. M., den Boer, W. C., van Agthoven, A. J. & Hagemeijer, A. *Somatic Cell Genet.* **7**, 645–656 (1981).
27. Meera Khan, P. *Archs. Biochem. Biophys.* **145**, 470–483 (1971).
28. Jeffreys, A. J. & Flavell, R. A. *Cell* **12**, 429–439 (1977).

APPENDIX PAPER

c-*sis* IS TRANSLOCATED FROM CHROMOSOME 22 TO CHROMOSOME 9 IN CHRONIC MYELOCYTIC LEUKEMIA*

By JOHN GROFFEN, NORA HEISTERKAMP, JOHN R. STEPHENSON, AD GEURTS VAN KESSEL, ANNELIES DE KLEIN, GERARD GROSVELD, AND DIRK BOOTSMA

From the Laboratory of Viral Carcinogenesis, National Cancer Institute, Frederick, Maryland 21701; and the Department of Cell Biology and Genetics, Erasmus University, Rotterdam, The Netherlands

Several relatively specific chromosomal translocations are known to be associated with particular human cancers (1–3). One of these, the Philadelphia translocation, t(9;22) (q34;q11) is observed in over 90% of chronic myelocytic leukemias (CML)[1] (3). Translocation of the q11 to qter segment of chromosome 22 to chromosome 9 results in a deleted form of chromosome 22, referred to as the Philadelphia (Ph') chromosome (4). Recently, we have localized a human oncogene, c-*abl*, on chromosome 9 (q34 to qter) (5) and demonstrated its translocation to chromosome 22q− (the Ph' chromosome) in CML (6). Because of the small size of the segment of chromosome 9 that translocates to chromosome 22 (6) and the localization of immunoglobulin λ light chain sequences on chromosome 22 (7), c-*abl* appears to map in close proximity to λ sequences in the Ph' chromosome.

Another acute transforming retrovirus, the Simian sarcoma virus (SSV), is a genetic recombinant between a nontransforming retrovirus and cellular sequences of woolly monkey origin (8, 9). The SSV transforming gene, v-*sis*, and its human cellular homologue, c-*sis*, have been molecularly cloned (8–10), and c-*sis* has been localized on the q arm of chromosome 22 (11, 12). In the present study, we report the localization of c-*sis* on the q11 to qter segment of chromosome 22 and its translocation from chromosome 22 to chromosome 9 in CML.

Materials and Methods

Cells. Cell lines, propagated in Dulbecco's modified Eagle's medium supplemented with 10% calf serum, included NIH/3T3 mouse cells and a human cell line, A673 (5). Somatic cell hybrids containing full complements of either mouse or Chinese hamster chromosomes and a limited number of human chromosomes were derived by fusion of either mouse or Chinese hamster cells with leukocytes from different CML patients or from normal donors (Table I); details concerning their origin and initial characterization have been previously reported (13, 14).

Preparation of a Human c-sis Probe (c-sis $B_{1.2}$). A cosmid clone with a cellular insert of ~30 kb containing v-*sis*-homologous sequences was isolated from a library of human lung

* Supported by contract NOI-CO-75380 from the National Cancer Institute and by the Netherlands Cancer Society (Koningin Wilhelmina Funds).

[1] *Abbreviations used in this paper:* CML, chronic myelocytic leukemia; Ph', Philadelphia chromosome; SSV, Simian sarcoma virus.

carcinoma DNA (15) using, as a probe, a 1.2 kb PstI v-*sis* restriction fragment in pBR322 (8), generously provided by K. Robbins and S. A. Aaronson. Isolation of a 1.7 kb BamHI v-*sis*-homologous restriction fragment from the cosmid clone was performed according to previously described methods (15).

Molecular Hybridization. Restriction enzymes were purchased from New England Biolabs, Beverly, MA and Bethesda Research Laboratories, Rockville, MD and were used according to the suppliers' specifications. DNA were digested with restriction enzymes, subjected to electrophoresis through 0.75% agarose gels, and transferred to nitrocellulose essentially as described by Southern (16). Nick translation of probes and filter hybridization were as described (15). Specific activity of the probes was $2–5 \times 10^8$ cpm/μg. After hybridization, filters were washed under high stringency conditions (10% standard saline citrate, 65°C) and exposed to XAR-2 film (Eastman Kodak Co., Rochester, NY) for up to 5 d at −70°C with Dupont Lightning Plus intensifying screens (Dupont Instruments, Wilmington, DE).

Results

To prepare a probe suitable for identification of somatic cell hybrids containing human c-*sis* sequences, a previously described (15) cosmid library of human lung carcinoma DNA was screened for clones containing sequences homologous to the 1.2 kb PstI v-*sis* probe. As shown in Fig. 1, a single cosmid clone was obtained containing a 30 kb cellular insert with v-*sis*-homologous cellular sequences. By restriction endonuclease analysis this clone was shown to correspond to a previously described v-*sis*-homologous human sequence, designated c-*sis* (12). For generation of a c-*sis*-specific probe, a single 1.7 kb BamHI fragment, possessing strong homology to v-*sis* (c-*sis* $B_{1.7}$), was isolated from the cosmid cellular DNA insert.

Human and mouse control cellular DNA were digested with Sst-I and analyzed for homology to the above described c-*sis* $B_{1.7}$ probe. As shown in Fig. 1, a single mouse cellular restriction fragment of around 10.0 kb (lane B) is detected, while the only human c-*sis* $B_{1.7}$ cross-reactive Sst-I restriction fragment is 3.6 kb in length (lane C). The size of the latter restriction fragment corresponds to that predicted on the basis of the human c-*sis* restriction map shown in Fig. 1. Cellular DNA from a mouse × human somatic cell hybrid, PgMe-25NU, previously shown to have chromosome 22 as its only human component (14), contains the 3.6 kb human c-*sis* $B_{1.7}$ cross-reactive Sst-I restriction fragment (lane A), thus confirming the mapping of c-*sis* on chromosome 22. Localization of c-*sis* to the region of chromosome 22 (q11 to qter) which is translocated to chromosome 9 in CML, is established by the absence of c-*sis*-homologous sequences from hybrid WESP-2A (lane D), which contains chromosome 22q− (the Ph′ chromosome) but lacks detectable amounts of chromosomes 9, 22, or 9q+ (5).

To independently show the localization of c-*sis* on chromosome 22 (q11 to qter) and demonstrate its translocation to chromosome 9 in CML, a series of Chinese hamster × human somatic cell hybrids were analyzed for human c-*sis* sequences. As shown in Fig. 1, lanes E and F, the only c-*sis* $B_{1.7}$ cross-hybridizing Sst-I restriction fragment in Chinese hamster cellular DNA is around 2.7 kb in size and thus clearly resolved from the 3.6 kb human Sst-I fragment. Hybrid 1CB-17aNU, which contains chromosome 22q−, lacks detectable human c-*sis* sequences (lane F), while a second hybrid, 14CB-21A, containing chromosome 9q+ but not chromosome 9 or 22 (13), is positive for human c-*sis* sequences (lane

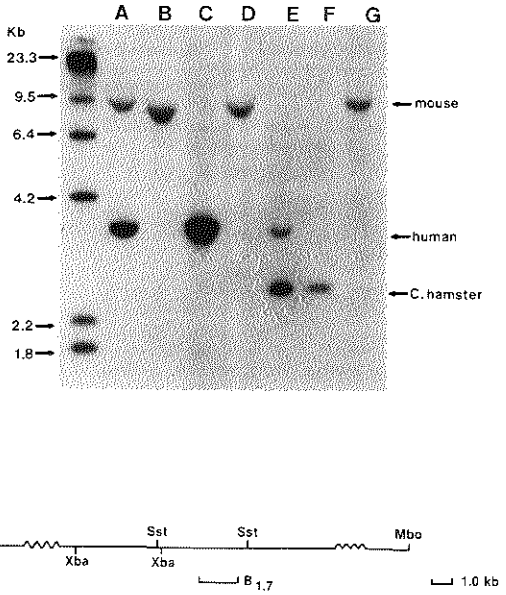


FIGURE 1. Localization of human c-*sis* on chromosome 22 (q11 to qter) and its translocation to chromosome 9 in CML. SstI-digested cellular DNA (10 μg/lane) were electrophoresed on 0.7% agarose gels, blotted to nitrocellulose, and hybridized to the c-*sis* $B_{1.7}$ probe shown in the lower portion of the figure. Cell lines analyzed are described in Table I and include PgMe25Nu (*A*), NIH/3T3 (*B*), A673 (*C*), WESP 2A (*D*), 14CB21A (*E*), 1CB17ANu (*F*). The positions of single mouse, human, and Chinese hamster c-*sis* $B_{1.7}$ homologous SstI restriction fragments are shown. Hind III-digested DNA, included as a molecular weight marker, is shown on the left side of the figure. In the lower portion of the figure, the restriction map of the 30 kb cellular insert from the v-*sis* cross-reactive cosmid clone of MboI-digested human lung carcinoma DNA is shown. MboI (Mbo) sites indicate the ends of the insert; the positions of XbaI (Xba) and SstI (Sst) restriction sites are shown for purposes of orientation of this clone with the more detailed previously published restriction maps of c-*sis* (10). The position of a single 3.6 kb human SstI restriction fragment overlapping with the 1.7 kb BamHI restriction fragment ($B_{1.7}$), used as a c-*sis*-specific probe for analysis of somatic cell hybrids, is also shown.

E). Finally, hybrid 10CB-23B, which contains chromosome 9 in the absence of detectable 22, 9q+, or 21q−, lacks human c-*sis* (Table I). As internal controls, each of the above hybrids were also analyzed for c-*abl*, a marker for the portion of chromosome 9 translocated to chromosome 22 in CML (6), and for AK1, which maps near the breakpoint but within the nontranslocated portion of chromosome 9 (Table I).

Discussion

Several of the human cellular homologues of viral oncogenes studied to date including c-*abl* (5, 6), c-*sis* (11, 12), c-*fes* (5, 17), and c-*mos* (18, 19) have been localized on human chromosomes frequently involved in translocations associated with specific human cancers. One of these, c-*myc*, is translocated from chromo-

TABLE I

Translocation of c-sis from Chromosome 22 to Chromosome 9 in Chronic Myelocytic Leukemia

Cell line	AK1	Human chromosomes				Oncogenes	
		9	22	9q+	22q−	*c-abl*	*c-sis*
Mouse NIH/3T3	NT	−	−	−	−	−	−
Human A673	NT	+	+	+	+	+	+
Mouse × human hybrid							
PgMe-25NU	−	−	+	−	−	−	+
WESP-2A	−	−	−	−	+	+	−
Chinese hamster × human hybrid							
10CB-23B	+	+	−	−	−	+	−
14CB-21A	+	−	−	+	−	−	+
1CB-17a NU	−	−	−	−	+	+	−

PgMe-25NU cells contain chromosome 22 as their only human component, while each of the other five hybrid clones contain a few human chromosomes in addition to those relevant to the t(9;22)(q34;q11) translocation (13, 14). Identification of hybrid clones containing *c-abl* sequences (6) and analysis of human adenylate kinase (AK1) enzymatic activity (6) have been previously reported. Cells were analyzed for human *c-sis*-specific sequences as described in Fig. 1.

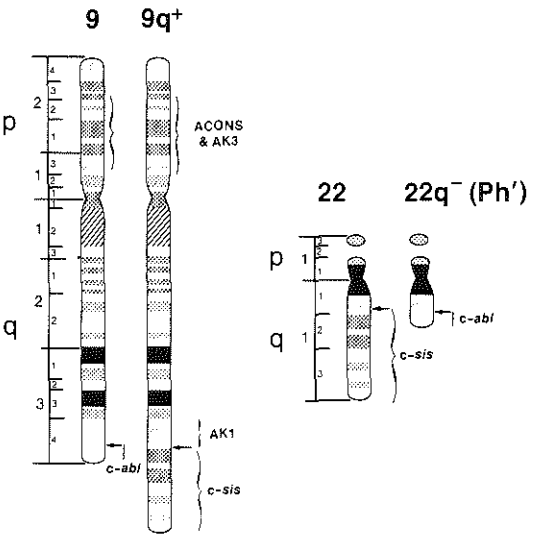


FIGURE 2. Diagrammatic representation of the involvement of c-*abl* and c-*sis* in the Ph' translocation. Chromosome banding patterns are as previously shown by Yunis (3); map positions of ACONS, AK3, and AK1 are as previously reported (24, 25). Localization of c-*abl* within the terminal portion of chromosome 9 (q34), which is translocated to chromosome 22 in CML, is as described by de Klein et al. (6) while localization of c-*sis* in the region of chromosome 22 (q11 to qter) translocated to chromosome 9 is based on the results of the present study.

some 8 to chromosomes 14, 2, or 22, each of which contain immunoglobulin sequences, in Burkitt's lymphoma (1, 19–21). Similarly, c-*abl* maps on the region of chromosome 9 which translocates to chromosome 22 in CML (5, 6). Other translocations, such as the t(15;17) reciprocal translocation associated with acute promyelocytic leukemia (3), involve regions to which human cellular oncogenes (in this case c-*fes*) have been mapped (22), but appear to be independent of immunoglobulin sequences. The present demonstration that c-*sis* is translocated from chromosome 22 to chromosome 9 in CML raises the possibility that c-*sis* rather than c-*abl* may be involved in CML. Resolution of these alternatives will require a determination of the proximity of these genes to the breakpoints in chromosomes 22 and 9, respectively, and analysis of the expression of their transcriptional and translational products in CML cells.

In addition to possible implications regarding the cause of CML and the significance of the associated t(9;22) (q34;q11) translocation, the localization of c-*sis* within the translocated region of chromosome 22 (Fig. 2) provides a unique molecular marker for studies of the more complex translocations associated with minority populations of CML patients. These can involve translocation of the q11 to qter region of chromosome 22 to chromosomes other than chromosome 9, or can in some instances involve more complex rearrangements including three or occasionally even four or five chromosomes (23). Analysis of these translocations using c-*abl* and c-*sis* probes should allow a determination of the critical translocation event resulting in the generation of CML.

Summary

By analysis of a series of somatic cell hybrids derived by fusion of either mouse or Chinese hamster cells with leukocytes from different chronic myelocytic leukemia (CML) patients or from normal donors, we have localized the human oncogene, c-*sis*, on the q11 to qter segment of chromosome 22 and demonstrated its translocation from chromosome 22 to chromosome 9 (q34) in CML.

We thank P. Hansen and G. T. Blennerhassett for excellent technical assistance.

Received for publication 10 January 1983 and in revised form 23 March 1983.

References

1. Rowley, J. D. 1982. Identification of the constant chromosome regions involved in human hematologic malignant disease. *Science (Wash. DC).* 216:749.
2. Yunis, J. J., C. Bloomfield, and K. Ensrud. 1981. All patients with acute nonlymphocytic leukemia may have a chromosomal defect. *N. Engl. J. Med.* 305:135.
3. Yunis, J. J. 1982. Most cancers may have a chromosomal defect. *In* Gene Amplification. R. T. Schimke, editor. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. 297pp.
4. Nowell, P. C., D. A. Hungerford. 1960. A minute chromosome in human chronic granulocytic leukemia. *Science (Wash. DC).* 132:1497.
5. Heisterkamp, N., J. Groffen, J. R. Stephenson, N. K. Spurr, P. N. Goodfellow, E. Solomon, B. Carritt, and W. F. Bodmer. 1982. Chromosomal localization of human cellular homologues of two viral oncogenes. *Nature (Lond.),* 299:747.
6. de Klein, A., A. Geurts van Kessel, G. Grosveld, C. R. Bartram, A. Hagemeijer, D.

Bootsma, N. K. Spurr, N. Heisterkamp, J. Groffen, and J. R. Stephenson. 1982. A cellular oncogene (c-*abl*) is translocated to the Philadelphia chromosome in chronic myelocytic leukemia. *Nature (Lond.).* 300:765.

7. Erikson, J., J. Martinis, and C. M. Croce. 1981. Assignment of the genes for human immunoglobulin chains to chromosome 22. *Nature (Lond.).* 294:173.
8. Robbins, K. C., S. G. Devare, and S. A. Aaronson. 1981. Molecular cloning of integrated simian sarcoma virus: genome organization of infectious DNA clones. *Proc. Natl. Acad. Sci. USA.* 78:2918.
9. Gelmann, E. P., F. Wong-Staal, R. A. Kramer, and R. C. Gallo. 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.
10. Dalla-Favera, R., E. P. Gelmann, R. C. Gallo, and F. Wong-Staal. 1981. A human *onc* gene homologous to the transforming gene (v-*sis*) of simian sarcoma virus. *Nature (Lond.).* 292:31.
11. Swan, D. C., O. W. McBride, K. C. Robbins, D. A. Keithley, E. P. Reddy, and S. A. Aaronson. 1982. Chromosomal mapping of the simian sarcoma virus *onc* gene analogue in human cells. *Proc. Natl. Acad. Sci. USA.* 79:4691.
12. Dalla-Favera, R., R. C. Gallo, A. Giallongo, and C. M. Croce. 1982. Chromosomal localization of the human homolog (c-*sis*) of the Simian sarcoma virus *onc* gene. *Science (Wash. DC).* 218:686.
13. Geurts van Kessel, A. H. M., H. den Brinke, W. A. M. Boere, W. C. den Boer, P. G. de Groot, A. Hagemeijer, P. Meera Khan, and P. L. Pearson. 1981. Characterization of the Philadelphia chromosome by gene mapping. *Cytogenet. Cell Genet.* 30:83.
14. Geurts van Kessel, A. H. M., W. C. den Boer, A. J. van Agthoven, and A. Hagemeijer. 1981. Decreased tumorigenicity of rodent cells after fusion with leukocytes from normal and leukemic donors. *Somatic Cell Genet.* 7:645.
15. Groffen, J., N. Heisterkamp, F. Grosveld, W. Van de Ven, and J. R. Stephenson. 1982. Isolation of human oncogene sequences (v-*fes* homolog) from a cosmid library. *Science (Wash. DC)* 216:1136.
16. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503.
17. Dalla-Favera, R., G. Franchini, S. Martinotti, F. Wong-Staal, R. C. Gallo, and C. M. Croce. 1982. Chromosomal assignment of the human homologues of feline sarcoma virus and avian myeloblastosis virus *onc* genes. *Proc. Natl. Acad. Sci. USA.* 79:4714.
18. Prakash, K., O. W. McBride, D. C. Swan, S. G. Devare, S. R. Tronick, and S. A. Aaronson. 1982. Molecular cloning and chromosomal mapping of a human locus related to the transforming gene of Moloney murine sarcoma virus. *Proc. Natl. Acad. Sci. USA.* 79:5210.
19. Neel, B. G., S. C. Jhanwar, R. S. K. Chaganti, and W. S. Hayward. 1982. Two human c-*onc* genes are located on the long arm of chromosome 8. *Proc. Natl. Acad. Sci. USA.* 79:7842.
20. Dalla-Favera, R., M. Bregnni, J. Erikson, D. Patterson, R. C. Gallo, and C. M. Croce. 1982. Human c-*myc onc* gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. *Proc. Natl. Acad. Sci. USA.* 79:7824.
21. Taub, R., I. Kirsch, C. Morton, G. Lenoir, D. Swan, S. Tronick, S. Aaronson, and P. Leder. 1982. Translocation of the c-*myc* gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. *Proc. Natl. Acad. Sci. USA.* 79:7837.
22. Sheer, D., L. R. Hiorns, K. F. Stanley, P. N. Goodfellow, D. M. Swallow, S. Povey, N. Heisterkamp, J. Groffen, J. R. Stephenson, and E. Solomon. 1983. Genetic analysis of the 15;17 chromosome translocation associated with acute promyelocytic leukemia.

Proc. Natl. Acad. Sci. USA. In press.
23. Koeffler, H. P., and D. W. Golde. 1981. Chronic myelogenous leukemia—new concepts. *New Engl. J. Med.* 304:1201.
24. Carritt, B., and S. Povey. 1979. Regional assignments of the loci AK_3, ACONS and ASS on human chromosome 9. *Cytogenet. Cell Genet.* 23:171.
25. Shows, T. B., A. Y. Sakaguchi, and S. L. Naylor. 1982. Mapping the human genome, cloned genes, DNA polymorphisms, and inherited disease. *Adv. Hum. Genet.* 12:341.

APPENDIX PAPER XI

Localization of the c-*abl* oncogene adjacent to a translocation break point in chronic myelocytic leukaemia

Nora Heisterkamp*, John R. Stephenson*, John Groffen*, Pamela F. Hansen†, Annelies de Klein†, Claus R. Bartram‡ & Gerard Grosveld‡

* Laboratory of Viral Carcinogenesis, National Cancer Institute-FCRF, and † Carcinogenesis Intramural Program, Frederick Cancer Research Facility, Frederick, Maryland 21701, USA
† Department of Cell Biology and Genetics, Erasmus University, PO Box 1738, 3000 DR Rotterdam, The Netherlands

The human c-abl *oncogene maps within the region (q34-qter) of chromosome 9 which is translocated to chromosome 22, the Philadelphia (Ph') chromosome, in chronic myelocytic leukaemia (CML). The position of the Ph' chromosomal break point is shown to be variable and, in one CML patient, has been localized immediately 5' of, or within, the* c-abl *oncogene. A DNA restriction fragment corresponding to this site has been molecularly cloned and shown to represent a chimaeric fragment of DNA from chromosomes 9 and 22.*

HUMAN cellular sequences designated oncogenes have been identified using retroviral transformation-specific sequences as molecular probes. Evidence for the potential involvement of such oncogenes in human cancer is accumulating rapidly. For instance, human tumour-derived DNA sequences, isolated on the basis of their ability to transform mouse cells phenotypically in tissue culture, have been shown to correspond to particular classes of cellular oncogenes[1–5]. More recently, amplification of specific cellular oncogenes in tumour cells has been demonstrated[6–8] and other human oncogenes have been shown to be involved in highly specific chromosomal translocations characteristic of particular classes of human neoplasia[9–13].

The Abelson strain of murine leukaemia virus (MuLV) represents a recombinant between Moloney MuLV and c-*abl*, a cellular oncogene of mouse origin[14]. In contrast to other oncogenic retroviruses, Abelson MuLV transforms lymphoid cells *in vitro*[15] and induces a rapid B-cell lymphoma *in vivo*[16]. We have previously reported the molecular cloning of the v-*abl* homologue (c-*abl*) from a human cosmid library[17]. Human v-*abl* homologous sequences were shown to be dispersed over a total region of around 32 kilobases (kb) and to contain a minimum of six introns. Extensive homology in nucleic acid and amino acid sequence was demonstrated between the tyrosine phosphorylation region of c-*abl* and corresponding regions of the v-*src*, v-*yes* and v-*fes*/*fps* family of viral oncogenes, as well as a more distant relatedness to the catalytic chain of the mammalian cyclic AMP-dependent protein kinase[18]. These findings suggest that each of these oncogenes was probably derived from a common progenitor and may represent members of a diverse, but highly related, family of cellular genes encoding protein kinases.

The human c-*abl* oncogene was initially mapped to the long (q) arm of chromosome 9 (ref. 19). This observation was of interest in view of the frequent involvement of chromosome 9 in the translocation with chromosome 22, t(9; 22), characteristic of chronic myelocytic leukaemia (CML)[20]. Analysis of a series of rodent–human somatic cell hybrids containing either the 9q+ or the 22q− (Philadelphia, Ph') chromosome demonstrated the translocation of human c-*abl* from chromosome 9 to the Ph' chromosome[9]. We show here that in one Ph'-positive CML patient, the chromosomal break point maps within 14 kb 5' of the human c-*abl* oncogene. Such a break point could not be identified at this position in the DNAs of two other patients, arguing that the chromosomal break in Ph'-positive CML can occur at variable sites on chromosome 9. These findings strongly implicate the human c-*abl* oncogene in Ph'-positive CML.

Search for the t(9; 22) junction site

To investigate the possibility that the chromosome 9 break point in the t(9; 22) may be localized either within or in close proximity to the human c-*abl* oncogene, high molecular weight DNA was isolated from biopsy samples of the spleens of three CML patients, each of which contained high percentages of leukaemic cells. In addition, DNA was isolated from an erythroleukaemic cell line, K562, established from a CML patient in blast crisis[21]. Each DNA was digested with *Bam*HI and subjected to Southern blot analysis. As shown in Fig. 1*a*, the restriction enzyme patterns of K562 DNA and DNA of patient 0319129 were indistinguishable from those of control DNA, when hybridized to a previously described[17] v-*abl* probe, $P_{1.7}$. Similarly, no differences in restriction enzyme patterns could be detected in the other two CML DNAs, or in any of the DNAs using other restriction enzymes or probes isolated from different regions of the human c-*abl* locus (not shown).

Chromosomal walking

As no rearrangements could be detected in the v-*abl* homologous cellular DNA region encompassed by our cosmid clones, a probe was prepared corresponding to an 0.4-kb *Hin*dIII–*Eco*RI fragment (0.4 HE) mapping within the 5' domain of the human c-*abl* locus (Fig. 2) and was used to screen a previously described human cosmid library[22]. Although recombinants representative of approximately 10 times the complete human genome were analysed, no positive clones were obtained. Analysis of total human DNA revealed that a *Bam*HI fragment of ~14.5 kb hybridized to the 0.4 HE probe. To characterize this restriction fragment further, a library of size-fractionated *Bam*HI-digested total human DNA was constructed in Charon 30. Screening of the library with the 0.4 HE probe revealed seven positive recombinant phage clones. All positive plaques were extremely small and turbid. The insert of one of these positives was subcloned into pBR328 (p5'-c-*abl*) and subjected to detailed restriction enzyme analysis as shown in Fig. 2*b*. No sequences in the 14.5-kb fragment hybridized to any of a series of previously described[17] probes corresponding to the complete v-*abl* genome.

As indicated in Fig. 2*b*, an 0.2-kb probe was prepared from the 5' terminus of p5'-c-*abl* by digestion with *Bam*HI and *Eco*RI. This probe hybridized to a *Hin*dIII fragment, of ~9.5 kb, in all DNAs examined, suggesting that sequences immediately 5' to the newly cloned region are non-rearranged (Fig. 1*c*, lanes 3–7). As expected, a fragment of 14.5 kb was

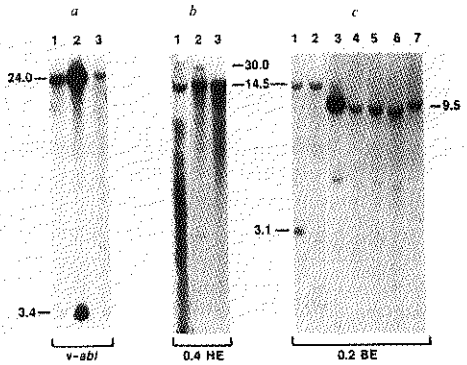


Fig. 1 Identification of DNA sequences mapping within and 5′ of the human c-*abl* locus. *a*, *Bam*HI digest of 10 μg of DNA from human cell line A673 (lane 1), K562 (lane 2) and patient 0319129 (lane 3). *b*, *Bam*HI digest of 10 μg DNA from CML 02120185 (lane 1), CML 0319129 (lane 2) and CML 0311068 (lane 3). *c*, *Bam*HI digest of DNA from patient 0319129 (lane 1), 0311068 (lane 2), *Hin*dIII digest of DNA from cell line K562 (lane 3), CML 02120185 (lane 4), CML 0319129 (lane 5), CML 0311068 (lane 6) and cell line A204 (lane 7). Molecular probes including a previously described v-*abl* probe (v-*abl* $P_{1,7}$)[17], 0.4 HE and 0.2 BE (see Fig. 2) are indicated at the bottom of the figure. Frozen spleen tissue was obtained through the NIH Resources and Logistics Branch, DCCP. Patient 0319129 was a 26-yr-old male, patient 0218105 was a 57-yr-old male and patient 0311068 was a 67-yr-old female. Cell line K562, pedigree 722, (GM 5372) was obtained from the Human Genetic Mutant Cell Repository, Camden, New Jersey. DNA was prepared from tissue and cell lines according to Jeffreys and Flavell[30]. DNA was digested with restriction enzymes, electrophoresed on 0.75% agarose gels and blotted according to the method of Southern[31]. Isolation of probes, hybridization and washings, to 0.3×SSC at 65 °C, were as described elsewhere[17,32]. Molecular weights of the hybridizing fragments (kb) are as indicated.

detected in *Bam*HI-digested normal human DNA, while in DNA of CML patient 0319129, an additional *Bam*HI restriction fragment of ~3.1 kb (Fig. 1*c*, lane 1) was identified. Similarly, the 0.4-kb HE probe isolated from the 5′ region of c-*abl* hybridized with a 14.5-kb *Bam*HI fragment in the other two CML DNAs (Fig. 1*b*). In DNA from CML patient 0319129, however, an extra *Bam*HI fragment of ~30 kb was visible (lane 2). The most likely explanation for these findings is that one allelic copy of DNA sequences 5′ to c-*abl* in CML patient 0319129 DNA is normal while the second is rearranged; these appear to be present in about a 1 : 1 molar ratio.

To define more accurately the region where DNA of patient 0319129 differs from the normal region, probes including an 0.52-kb *Eco*RI fragment (0.52 E) and an 0.76-kb *Eco*RI–*Hin*dIII fragment (0.76 HE) (see Fig. 2*b*) were prepared from p5′-c-*abl*, corresponding to the approximate region of rearranged sequences in the DNA of patient 0319129. As shown in Fig. 3, extra DNA fragments were found with these probes on restriction enzyme digestion of DNA 0319129 with *Bgl*II, *Xba*I, *Sst*I and *Hin*dIII. The 9.5-kb 5′ *Hin*dIII fragment, which hybridized to the 0.52-kb *Eco*RI probe, seemed to be present in a normal quantative level, as did the 2.0-kb fragment hybridizing to the 0.76-kb HE probe. A rearrangement thus appeared to have taken place within the 1.25-kb *Hin*dIII fragment since, in addition to the normal 1.25-kb *Hin*dIII fragment, a fragment of 2.2-kb can be identified.

Molecular cloning

As human c-*abl* is translocated to the Philadelphia chromosome (22q⁻) in CML, we reasoned that the transposed sequences in DNA of CML patient 0319129 should include most of the 14.5-kb *Bam*HI fragment; the 5′ sequences of this fragment would presumably remain on chromosome 9, linked to sequences from chromosome 22, thus accounting for the 3.1-kb *Bam*HI fragment (Fig. 1*c*, lane 1). The 9.5-kb *Bgl*II fragment (Fig. 3*b*) would thus represent the chromosome 22/9 chimaeric fragment localized at the Ph′ translocation break point; the 6.0-kb *Bgl*II fragment (Fig. 3*b*) would contain sequences from chromosomes 9 and 22, on the 9q+ chromosome. Alternatively, the differences between the DNA of patient 0319129 and control DNA could be the result of DNA rearrangements, deletions or insertions having no connection with the t(9; 22) Ph′ translocation.

To distinguish between these possibilities, DNA from CML patient 0319129 was digested with *Bgl*II, size fractionated in

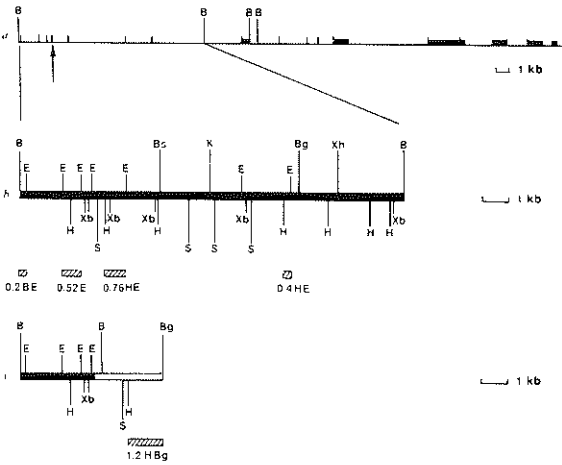


Fig. 2 Restriction enzyme analysis of p5′-c-*abl* and a part of pBglII 9q+. The position of p5′-c-*abl* in relation to the v-*abl* hybridizing region in the human c-*abl* locus is indicated in *a*. *Eco*RI sites are marked by small vertical lines; v-*abl* homologous sequences are indicated by solid bars. The vertical arrow indicates the break point in the DNA of CML 0319129. A restriction map of p5′-c-*abl* is shown in *b*; probes used in this study, indicated as hatched boxes, are shown beneath the map. The pBglII 9q+ fragment in c is a subclone of a 6.0-kb *Bgl*II fragment isolated from CML patient 0319129; in *b* and c the *Bgl*II site within 200 bp of the 5′ *Bam*HI site and a *Hpa*I site in between these two restriction sites are not shown. The solid bar indicates sequences from chromosome 9 while the open bar represents sequences from chromosome 22. Cloning and subcloning of the p5′-c-*abl* and *Bgl*II-9q+ sequences was according to previously described methods[33]. Restriction enzyme maps are based on hybridization data and on results of double digestions of individual fragments isolated from low melting point agarose gels. Restriction enzymes include: *Bam*HI (B), *Bgl*II (Bg), *Bst*EII (Bs), *Hin*dIII (H), *Sst*I (S), *Xba*I (Xb), *Xho*I (Xh), *Eco*RI (E) and *Kpn*I (K).

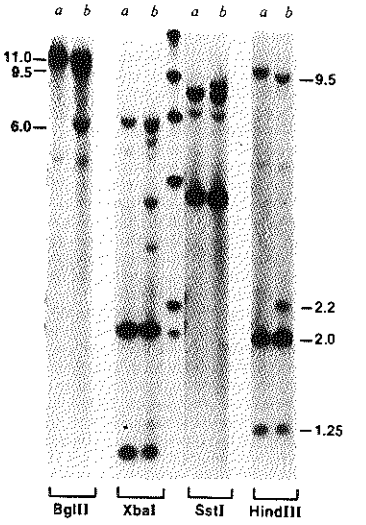


Fig. 3 Restriction enzyme analysis of the DNA of CML patient 0319129. Restriction enzymes used and the molecular weights of the fragments in the *Bgl*II and *Hin*dIII digestions are as indicated. ^{32}P-labelled *Hin*dIII digestion fragments of λ DNA in the centre of the figure serve as a molecular weight marker for the *Xba*I and *Sst*I digestions. Southern blots were hybridized with a mixture of the 0.52 E and 0.76 HE probes (see Fig. 2*b*), as described in Fig. 1 legend. DNA of the cell line A673 is shown in lanes *a*; DNA of CML 0319129 is in lanes *b*.

the 6–7-kb range, and a library constructed in *Bam*HI-digested Charon 30 phage arms. Three positive clones were identified on screening with the 0.2-kb BE probe. Because each of the positive clones contained an additional inserted *Bgl*II fragment, some *Bgl*II sites had remained intact, allowing the subcloning of most of the 6.0-kb *Bgl*II fragment by *Bgl*II–*Hpa*I digestion into plasmid ORF[23]. Colonies containing this recombinant plasmid were visible only on a two- to threefold longer incubation at 37 °C than neighbouring non-positive colonies. This particular DNA region may thus contain sequences that inhibit bacterial growth, possibly explaining our failure to detect these sequences in cosmid libraries. As shown in the restriction enzyme map of this fragment (Fig. 2*c*), the *Eco*RI site 3′ of the 0.52-kb *Eco*RI fragment is still present. In contrast, an *Sst*I site immediately 3′ of it and all other restriction enzyme sites to the 3′ are either missing or different from those found in the same region in p5′-c-*abl*. These findings localize a putative break point in p5′-c-*abl* between the *Eco*RI and *Sst*I sites. Sequences 5′ of this point in the 6.0-kb *Bgl*II fragment hybridize to p5′-c-*abl*; sequences 3′ of it do not (results not shown).

To determine the origin of the 6.0-kb *Bgl*II fragment, a 1.2-kb *Hin*dIII–*Bgl*II probe (1.2 HBg) was prepared (Fig. 2*c*). As shown in Fig. 4, using stringent washing conditions, this probe hybridized to a 5.0-kb *Bgl*II fragment in DNA of the human cell line A204 (lane 2), but not to sequences in mouse DNA (lane 1) or in rodent–human somatic cell hybrids containing human chromosome 9 (lane 4)[9]. In an independent hybrid (PgMe-25Nu), however, containing chromosome 22 as its only human component[9], a 5.0-kb *Bgl*II fragment was detected (lane 3). Moreover, using the same probe both 5.0-kb and 6.0-kb *Bgl*II fragments were identified in the DNA of patient 0319129 (lane 6). The 1.2 HBg probe did not hybridize to sequences in the human–mouse somatic cell hybrid WESP-2A (Fig. 4, lane 5), which contains the Philadelphia chromosome in the absence of 9.9q+ or 22. Thus, the 6.0-kb *Bgl*II fragment cloned from patient 0319129 DNA must contain, in addition to sequences from chromosome 9, sequences originating from the translocated region of chromosome 22.

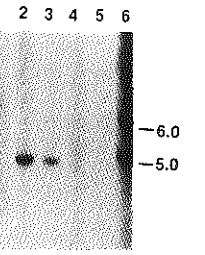


Fig. 4 Chromosomal localization of the 3′ domain of pBglII-9q+. High molecular weight DNAs were digested with *Bgl*II, electrophoresed on an 0.75% agarose gel, blotted and hybridized to the 1.2 HBg probe (see Fig. 2*c*) and included mouse (lane 1), A204 (lane 2), PgMe-25Nu (chromosome 22, lane 3), 10CB-23B (chromosome 9, lane 4), WESP-2A (chromosome 22q−, lane 5) and CML patient 0319129 (lane 6).

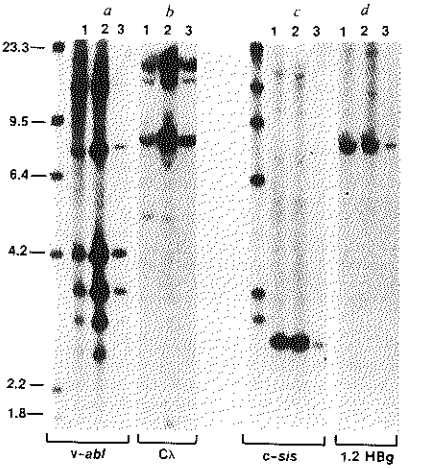


Fig. 5 Amplification of c-*abl* in the erythroleukaemic cell line K562. Control human DNA (10 μg, lanes 1 in each panel) was run next to 10 μg (lanes 2) and 2.5 μg (lanes 3) K562 DNA. ^{32}P-labelled *Hin*dIII-digested λ DNA is included as a molecular weight standard for *a* and *b* (0.6% agarose gels) and for *c* and *d* (0.75% agarose gels). *a*, DNAs were digested with *Eco*RI and hybridized to a previously described v-*abl* probe, v-*abl* $P_{1\ 7}$ (ref. 17). *b*, *Eco*RI-digested DNAs were hybridized to a *Bgl*II–*Hin*dIII probe, isolated from a human C_λ clone Hu λ 5 (ref. 25). *c*, *Bam*HI-digested DNA hybridized to a previously described[26] 1.7-kb *Bam*HI probe prepared from a human c-*sis* cosmid clone. *d*, *Bgl*II-digested DNAs hybridized to the 1.2 HBg probe (Fig. 2*c*).

Amplification of c-*abl*

Although the K562 cell line lacks a cytogenetically identifiable Ph′ chromosome (ref. 24 and A. Hagemeijer, personal communication), it was reported to contain a small ring chromosome r(22), which is probably derived from a Ph′ chromosome[24]. Our results suggest that, at a minimum, K562 contains a chromosome(s) resembling the Ph′ chromosome marker. As shown in Fig. 5*d*, using the 1.2 HBg probe, a non-amplified 5.0-kb *Bgl*II fragment was detected in K562 DNA. In contrast, human c-*abl* sequences are amplified at least fourfold in K562 DNA, as are

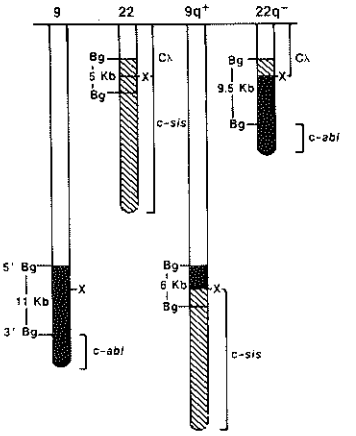


Fig. 6 Schematic representation of the t(9; 22) in CML patient 0319129. The sizes of the *Bgl*II fragments detected with probes from chromosome 9 and 22 are indicated. Regions in black originate from chromosome 9; hatched regions are from chromosome 22. An *X* indicates the t(9; 22) chromosomal break point. The human c-*sis* oncogene and the C_λ region on chromosome 22 are shown for purposes of orientation; human c-*sis* is localized on chromosome 22 (ref. 34) and is translocated to chromosome 9 in CML[26]. C_λ is localized in band q 11 of chromosome 22 (ref. 10), and is not involved in the Ph' translocation (unpublished results).

sequences up to 17 kb upstream of the most 5' v-*abl* homologous *Eco*RI region (Fig. 1*c*). A 11.0-kb *Eco*RI fragment (Fig. 5*a*) that hybridizes to the v-*abl* P_{L7} probe but has no linkage with the main human c-*abl* locus[17] is non-amplified and thus serves as an internal control. The immunoglobulin light-chain constant region (C_λ)[25] is also amplified at least fourfold (Fig. 5*b*). This latter observation is consistent with the localization of human c-*abl* and C_λ sequences on the same amplification unit, presumably a part of the Ph' chromosome. In concordance with this conclusion, C_λ remains on the Ph' chromosome in the t9:22 (unpublished observations). In contrast, another human oncogene, c-*sis*, which is located on chromosome 22 and transferred to chromosome 9 in the Ph' translocation[26], is non-amplified (Fig. 5*c*). Finally, the fact that the 5.0-kb *Bgl*II fragment normally localized on chromosome 22, is non-amplified in K562, supports the assumption that it is translocated to chromosome 9 in this cell line.

Conclusions

The present findings establish that in one of three CML patient DNAs examined, the Ph' chromosomal break point is localized within 14 kb of the most 5' v-*abl* homologous region. Because the position of the most 5' exon in human c-*abl* is not known, the possibility that the t(9; 22) break point may even map within a coding region of the human c-*abl* locus cannot be excluded. We have not localized the break points in the DNAs of the two other CML patients or in DNA of the cell line WESP-2A, containing the 22q− chromosome. Our findings, however, establish that the site of the Ph' translocation break point is variable and suggest that break points in the latter DNAs may be located 5' of the ~80 kb of cloned DNA encompassing the human c-*abl* locus. Similarly, the t(8; 14) translocation in Burkitt lymphoma, involving the human c-*myc* oncogene, does not appear to involve a single site[10–12,27]. The orientation of human c-*abl* on chromosome 9, with its 5' region towards the centromere of the chromosome (Fig. 6), is established by the fact that the break point has been found 5' to the human c-*abl* locus.

K562 cells lack a cytogenetically identifiable Ph' chromosome despite their derivation from a CML patient in blast cell crisis[21]. However, genes on either side of the break point in the Ph' chromosome, c-*abl* and C_λ, are not only present in this cell line, but are amplified. In contrast, no amplification was found in the case of c-*sis*, normally on chromosome 22 but translocated to chromosome 9 in Ph'-positive CML, or a DNA region translocated to chromosome 9 corresponding to a restriction fragment containing the t(9; 22) break point isolated from a CML patient. The amplification seems to be specific; neither human c-*fes*, on chromosome 15 (ref. 19), nor c-*fms*, on chromosome 5 (ref. 28), is amplified (data not shown). The finding that c-*abl* is amplified in K562 can be explained in two ways: it can reflect the fact that these cells were grown in tissue culture, possibly resulting in specific chromosomal aberrations and/or amplifications; alternatively, the c-*abl* may have been amplified during the blast crisis of the patient, the overproduction of the c-*abl* gene product being the cause of the disease. The relative amplification of a region of DNA, which includes both c-*abl* and C_λ sequences, suggests that the two are nearly contiguous in this cell line.

We have previously demonstrated the translocation of a small piece of chromosome 9 to chromosome 22 in CML patients carrying the Ph' chromosome[9]. The size of this fragment was calculated to be less than 5,000 kb, suggesting a potential role of the c-*abl* oncogene in CML. The isolation, in the present study, of a region of DNA from a CML patient that contains a chromosomal break point which is localized within 14 kb of the c-*abl* locus, the specific amplification of c-*abl* sequence in the K562 cell line, and the finding that c-*abl* is translocated to chromosome 22 in complex Ph' translocations[29] further implicate the c-*abl* oncogene in chronic myelocytic leukaemia.

We thank A. Geurts van Kessel for providing somatic cell lines, A. Hagemeijer for valuable discussion, P. Leder for the generous gift of the recombinant HuA5, Gail Blennerhassett for technical assistance and Beverly Bales for help with the preparation of the manuscript. The work was supported by the Netherlands Cancer Society (Koningin Wilhelmina Fonds) and under NCI PHS contract NOI-CP-75380. C.R.B. is a recipient of a fellowship from the Deutsche Forschungsgemeinschaft.

Received 8 July; accepted 2 September 1983.

1. Pulciani, S. *et al. Nature* **300,** 539–542 (1982).
2. Der, C. J., Krontiris, T. G. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **79,** 3637–3640 (1982).
3. Tabin, C. J. *et al. Nature* **300,** 143–149 (1982).
4. Taparowsky, E. *et al. Nature* **300,** 762–765 (1982).
5. Shimizu, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **80,** 2112–2116 (1983).
6. Collins, S. & Groudine, M. *Nature* **298,** 679–681 (1982).
7. Dalla-Favera, R., Wong-Staal, F. & Gallo, R. *Nature* **299,** 61–63 (1982).
8. Alitalo, K., Schwab, M., Lin, C., Varmus, H. E. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **80,** 1707–1711 (1983).
9. de Klein, A. *et al. Nature* **300,** 765–767 (1982).
10. Taub, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **79,** 7837–7841 (1982).
11. Adams, J. M., Gerondakis, S., Webb, E., Corcoran, L. M. & Cary, S. *Proc. natn. Acad. Sci. U.S.A.* **80,** 1982–1986 (1983).
12. Dalla-Favera, R. *et al. Science* **219,** 963–967 (1983).
13. Sheer, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **80,** 5007–5011 (1983).
14. Goff, S. P., Gilboa, E., Witte, O. N. & Baltimore, D. *Cell* **22,** 777–785 (1980).
15. Rosenberg, N., Baltimore, D. & Scher, C. *Proc. natn. Acad. Sci. U.S.A.* **72,** 1932–1936 (1975).
16. Abelson, H. T. & Rabstein, L. S. *Cancer Res.* **30,** 2213–2222 (1970).
17. Heisterkamp, N., Groffen, J. & Stephenson, J. R. *J. molec. appl. Genet.* **2,** 57–69 (1983).
18. Groffen, J., Heisterkamp, N., Reynolds, F. Jr. & Stephenson, J. R. *Nature* **304,** 167–169 (1983).
19. Heisterkamp, N. *et al. Nature* **299,** 747–750 (1982).
20. Rowley, J. D. *Nature* **243,** 290–293 (1973).
21. Lozzio, C. B. & Lozzio, B. B. *Blood* **45,** 321–334 (1975).
22. Groffen, J., Heisterkamp, N., Grosveld, G., van de Ven, W. & Stephenson, J. R. *Science* **216,** 1136–1138 (1982).
23. Weinstock, G. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **80,** 4432–4436 (1983).
24. Klein, E. *et al. Int. J. Cancer* **18,** 421–431 (1976).
25. Hieter, P. A., Hollis, G. F., Korsmeyer, S. J., Waldmann, T. A. & Leder, P. *Nature* **294,** 536–540 (1981).
26. Groffen, J. *et al. J. exp. Med.* **158,** 9–15 (1983).
27. Stanton, L. W., Watt, R. & Marcu, K. B. *Nature* **303,** 401–406 (1983).
28. Groffen, J. *et al. Nucleic Acids Res.* (in the press).
29. Bartram, C. R. *et al. Nature* **306,** 277–280 (1983).
30. Jeffreys, A. J. & Flavell, R. A. *Cell* **12,** 429–439 (1977).
31. Southern, E. M. *J. molec. Biol.* **98,** 503–517 (1975).
32. Bernards, R. & Flavell, R. A. *Nucleic Acids Res.* **8,** 1521–1533 (1980).
33. Groffen, J., Heisterkamp, N., Blennerhassett, G. & Stephenson, J. R. *Virology* **126,** 213–227 (1983).
34. Swan, D. C. *et al. Proc. natn. Acad. Sci. U.S.A.* **79,** 4691–4696 (1982).

APPENDIX PAPER XII

Philadelphia Chromosomal Breakpoints Are Clustered within a Limited Region, bcr, on Chromosome 22

John Groffen,*‡ John R. Stephenson,*‡ Nora Heisterkamp,*‡ Annelies de Klein,† Claus R. Bartram,† and Gerard Grosveld†
*Laboratory of Viral Carcinogenesis
National Cancer Institute-FCRF
Frederick, Maryland 21701
†Department of Cell Biology and Genetics
Erasmus University, P.O. Box 1738
3000 DR Rotterdam, The Netherlands

Summary

We have identified and molecularly cloned 46 kb of human DNA from chromosome 22 using a probe specific for the Philadelphia (Ph′) translocation breakpoint domain of one chronic myelocytic leukemia (CML) patient. The DNAs of 19 CML patients were examined for rearrangements on chromosome 22 with probes isolated from this cloned region. In 17 patients, chromosomal breakpoints were found within a limited region of up to 5.8 kb, for which we propose the term "breakpoint cluster region" (bcr). The two patients having no rearrangements within bcr lacked the Ph′ chromosome. The highly specific presence of a chromosomal breakpoint within bcr in Ph′-positive CML patients strongly suggests the involvement of bcr in this type of leukemia.

Introduction

Chronic myelocytic leukemia (CML) is characterized by the presence of the Philadelphia (Ph′) chromosome in the leukemic cells of 96% of all CML patients. The Ph′ chromosome is the result of a translocation between chromosome 22 and chromosome 9 (Rowley, 1973, 1982; Sandberg, 1980); its presence has important prognostic and diagnostic value. Previously we described the localization of the human c-*abl* oncogene (Heisterkamp et al., 1983a), to chromosome 9 (Heisterkamp et al., 1982) and demonstrated its translocation to the Philadelphia ($22q^-$) chromosome in CML (de Klein et al., 1982). This demonstrated unequivocally that the t(9;22) is reciprocal. As the breakpoint on chromosome 9 is at the most telomeric band on this chromosome, q34, human c-*abl* must be translocated on a relatively small fragment of less than 5000 kb to chromosome 22, suggesting a potential role for the c-*abl* oncogene in CML. This hypothesis was strengthened by the isolation of a chimeric DNA fragment from one CML patient containing sequences from chromosome 9 and 22 and located 14 kb immediately 5′ of human v-*abl* homologous sequences (Heisterkamp et al., 1983b). In the present study, we have used the chromosome 22-specific sequences of the chimeric DNA fragment to isolate a second chimeric chromosome 9/22 ($9q^+$) fragment from a different CML patient. The chromosome 9-specific sequences in this fragment must be localized at a minimal distance of 40 kb from the human v-*abl* homologous sequences. Using the same probe, we have isolated an extended region on chromosome 22 from non-CML human DNA. In contrast to the situation on chromosome 9, the breakpoints on chromosome 22 in the DNAs of these two CML patients had occurred in the same region, although not at an identical site. Subsequently, we investigated the genomic organization of this area in a number of other Ph′-positive CML patients: all exhibited abnormal restriction enzyme patterns, indicating that in Ph^+-positive CML a breakpoint occurs in a single well defined region of chromosome 22.

‡ Present address: Oncogene Science, Inc., 222 Station Plaza N., Mineola, NY 11501.

Results

Isolation of a $9q^+$ Chimeric Fragment

Previously we have isolated a chimeric DNA fragment containing sequences originating from chromosomes 9 and 22 (Figure 1B) from a CML patient, 0319129. On chromosome 9, the breakpoint was located immediately to the 5′ of human v-*abl* homologous sequences (Figure 1A) and may even be within the human c-*abl* oncogene. However, the DNAs of two other CML patients did not contain rearrangements in this region; furthermore, we have molecularly cloned an additional 11 kb of DNA to the 5′ and have found no rearrangements in this area for these two DNAs (results not shown). Therefore, we decided to investigate whether we could localize the Ph′ translocation breakpoint to a specific site on chromosome 22. Using a 1.2 kb Hind III–Bgl II fragment (1.2 HBg, see Figure 1B) containing chromosome 22 sequences from the breakpoint region of CML patient 0319129 as a probe, we examined the DNA of the leukemic cells of a second patient (02120185). As shown in Figure 2A, this probe detects a normal 5.0 kb Bgl II fragment in control DNA (lane 1), in DNA of patient 0319129 (lane 2) and in DNA of patient 02120185 (lane 3). As expected, it also detects the breakpoint fragment of DNA 0319129 (Figure 1B and Figure 2, lane 2). In DNA of patient 02120185, an extra Bgl II fragment of 6.6 kb is visible (Figure 2, lane 3). Similarly, this probe hybridizes to a normal 3.3 kb Bam HI fragment in all three DNAs (Figure 2B), but detects an additional abnormal 11.3 kb Bam HI fragment in DNA 02120185 (lane 3). As we could detect additional restriction enzyme fragments with other restriction enzymes in DNA 02120185 (data not shown), we decided to examine whether these abnormal fragments were the result of the presence of a chromosomal breakpoint. Using the 1.2 kb HBg fragment as a probe, we molecularly cloned the 11.3 kb Bam HI fragment in charon 30. In this fragment only 1.2 kb of DNA, homologous to the probe, was defined as originating from chromosome 22; to determine all chromosome 22-specific sequences in the 11.3 kb Bam HI fragment, it was necessary to isolate the homologous region on chromosome 22 from non-CML DNA. For this

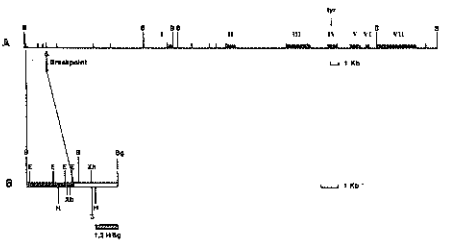


Figure 1. Position of the 9q+ Chromosomal Breakpoint in Spleen DNA of Patient 0319129 in Relation to Human *v-abl* Homologous Sequences

(A) The human *v-abl* homologous regions are designated I through VII and are indicated by solid bars; Eco RI sites are marked by small vertical lines. The vertical arrow points to the breakpoint in the DNA of CML patient 0319129. (B) The molecularly cloned region of DNA of patient 0319129 that contains a breakpoint. The solid bar indicates sequences from chromosome 9, while the open bar represents sequences from chromosome 22. The 1.2 kb HBg probe is indicated as a hatched box. Restriction enzymes include: Bam HI (B), Bgl II (Bg), Hind III (H), Sst I (S), Xba I (Xb), Xho I (Xh) and Eco RI (E).

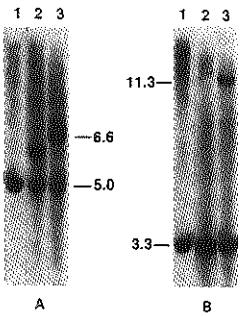


Figure 2. Restriction Enzyme Analysis of CML DNAs

Ten microgram of high molecular weight DNA was digested with Bgl II (A) or Bam HI (B), electrophoresed on 0.75% agarose gels, and blotted. DNAs analyzed were from human cell line GM3314 (lane 1), CML patient 0319129 (lane 2), and CML patient 02120185 (lane 3). Frozen spleen tissues, including those used as a source for the isolation of DNAs shown in Figure 5, were obtained through the Biological Carcinogenesis Branch, DCCP. Hybridization was with the 1.2 kb HBg probe (see Figure 1); filters were washed to 0.3 × SSC at 65°C. Molecular weights of fragments hybridizing to the probe are indicated in kilobases.

purpose, a previously described (Groffen et al., 1982) human lung carcinoma cosmid library was screened with the 1.2 kb HBg probe. Three cosmid clones were isolated, which contained overlapping portions of the same region.

Molecular Cloning of Ph′ Breakpoint Region of Normal Chromosome 22

As shown in Figure 3A, a region of approximately 46 kb was molecularly cloned; the 1.2 kb HBg probe hybridizes to a Bgl II fragment of 5.0 kb, located centrally in the cloned region. No homology is apparent between the restriction map of this region and that of human *c-sis* (Dalla-Favera et al., 1981), an oncogene situated on chromosome 22 but translocated to chromosome 9 in the Ph′ translocation (Groffen et al., 1983a). This confirms earlier reports that indicated that *c-sis* is not located in the immediate proximity of the Ph′ breakpoint (Bartram et al., 1983a). In variant Burkitt lymphoma, a t(8;22) has been described in which the immunoglobulin light chain was found to be involved (de la Chapelle et al., 1983). The light-chain constant region (Cλ) and the Ph′ chromosomal breakpoint have been localized to chromosome 22 band q11 (Rowley, 1973; McBride et al., 1982; Yunis, 1983); this suggests that *c-abl* could be translocated into Cλ in patients with CML. However, a probe isolated from the λ constant region showed no cross-homology with the above described chromosome 22 sequences. Additionally, no hybridization to a murine λ-variable region probe (Miller et al., 1981) was observed (results not shown).

To facilitate comparison of the 11.3 kb Bam HI fragment with homologous sequences on chromosome 22, the 5.0 kb Bgl II fragment was subcloned into pSV2-neo (Figure 3B). In concordance with our previous results (Heisterkamp et al., 1983b), in the 6.0 kb Bgl II breakpoint fragment from CML patient 0319129, restriction enzyme sites 3′ to the most 3′ Eco RI site originate from chromosome 22 (Figure 3C). In the 11.3 kb Bam HI fragment (Figure 3D) approximately 2.5 kb of DNA, including the 3′ Bam HI site and extending to the 3′ Xho I site, originates from chromosome 22.

The 11.3 kb Bam HI Fragment Also Contains a Breakpoint

To establish conclusively that the 11.3 kb Bam HI fragment represents a chimeric fragment of chromosomes 22 and 9, we isolated a 1.3 kb Eco RI fragment from the chromosome 22 nonhomologous region. Using this fragment as a probe, homologous sequences were detected in Bgl II-digested mouse DNA (Figure 4, lane 1) and Chinese hamster DNA (not shown). These bands, however, were clearly resolved from the Bgl II fragment visible in human DNA (lane 2). No human sequences homologous to the probe were detected in rodent–human somatic cell hybrids PgMe-25Nu, having human chromosome 22 as its only human component (lane 4) or in WESP-2A, (lane 3) containing a Ph′ chromosome but not chromosome 9 or 9q+ (de Klein et al., 1982). In the rodent–human somatic cell hybrids 10CB-23B (lane 5), containing human chromosome 9 and in 14CB-21A, containing a 9q+ chromosome (not shown), a Bgl II fragment of human origin homologous to the probe clearly was present. The only human DNA sequences these two hybrids have in common are those originating from chromosome 9. Therefore, the 11.3 kb Bam HI fragment possesses a breakpoint and represents a chimeric fragment containing chromosome 9- and 22-specific sequences isolated from a second CML patient.

Clustering of Ph′ Breakpoints on Chromosome 22 in CML Patients

Since in each of the above two CML DNAs the breakpoint in the t(9;22) on chromosome 22 was localized within a

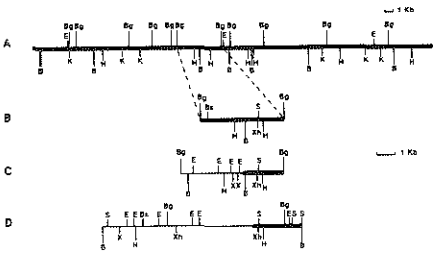


Figure 3. Comparative Restriction Enzyme Analysis of the Breakpoint Region on Chromosome 22 with Two 9q+ Breakpoint Regions

A restriction enzyme map of the cloned region in which chromosomal breaks occur on chromosome 22 is shown in (A); a subclone containing the 5.0 kb Bgl II fragment in (B) is compared with the 6.0 kb Bgl II and 11.3 kb Bam HI restriction enzyme fragments of the 9q+ chromosomes in (C) and (D); heavy lines indicate sequences from chromosome 22, whereas light lines indicate sequences originating from chromosome 9. B = Bam HI; Bg = Bgl II; Bs = BstE II; E = Eco RI; H = Hind III; K = Kpn I; S = Sst I; X = Xba I; Xh = Xho I.

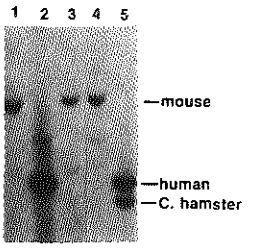


Figure 4. Origin of the 5′ Sequences of the 11.3 kb Bam HI Fragment

High molecular weight DNAs, including mouse (lane 1), human cell line A204 (lane 2), WESP-2A (chromosome 22q−, lane 3), PgMe-25Nu (chromosome 22, lane 4), and 10CB-23B (chromosome 9, lane 5) were digested with Bgl II, electrophoresed on an 0.75% agarose gel, blotted, and hybridized to a 1.3 kb Eco RI probe isolated from the 11.3 kb Bam HI fragment.

common region, we decided to investigate this area in other CML DNAs. As the 1.2 HBg probe had detected abnormal ($9q^+$) Bgl II restriction fragments in DNAs 0319129 and 02120185, we subjected 17 additional independent CML DNAs to similar analysis; six of these were from spleen tissue, nine were from patient blood, and two were from bone marrow. As shown in Figure 5, lanes 1–13, CML DNAs from spleen, blood, and bone marrow all contained additional Bgl II fragments hybridizing to the 1.2 HBg probe; the DNAs of the patients shown in lanes 14–17 did not exhibit abnormal Bgl II fragments. Two of these (H81-258, lane 14 and C080, lane 15) showed deviant restriction enzyme patterns with other restriction enzymes (this will be discussed below).

To confirm that the 1.2 HBg probe detected chromosomal rearrangements and not merely DNA polymorphisms for the restriction enzyme Bgl II, all DNAs were subjected to digestion with at least one, but in most cases two or more different restriction enzymes. After hybridization with the 1.2 HBg probe, abnormal restriction enzyme fragments were detected in all Ph′-positive CML DNAs (also see below). Therefore, a polymorphism for Bgl II seems an unlikely explanation for the abnormal fragments shown in Figure 5; moreover, in the DNAs of most patients, abnormal fragments of different molecular weights are detected with the 1.2 HBg probe.

No extra Bgl II fragments were found in DNA isolated from cultured fibroblasts of patient H80-251 (lane 18) although an extra Bgl II fragment is clearly visible in DNA isolated from the leukemic cells of this patient (lane 7). Moreover, DNA isolated from the fibroblast cell line, AG 1732, established from a Ph′-positive CML patient, also lacked abnormalities (lane 19) in this region. Finally, in DNA isolated from leukemic cells of a Ph′-negative CML patient (lane 17) and of a two-year-old child with juvenile Ph′-negative CML, no visible rearrangements were found (lane 16), confirming our results of previous experiments (Bartram et al., 1983b) in which no translocations concerning c-*abl* to chromosome 22 or c-*sis* to chromosome 9 were found in Ph′-negative CML.

Sublocalization of Ph′ Breakpoints on Chromosome 22

As is apparent from the detailed restriction enzyme analysis of the breakpoint fragments of the DNAs of patients 0319129 and 02120185, the exact breakpoints are not localized at identical sites on chromosome 22. To sublocalize the breakpoints in the DNAs of the other CML patients more precisely, we arbitrarily divided the 5.0 kb Bgl II fragment into segments bordered by restriction enzyme sites for Bgl II, Bam HI, and Hind III (see Figure 6, bottom). Region 0 thus extends from the 5′ Bgl II site to the first 5′ Hind III site, region 1, a 0.6 kb Hind III–Bam HI fragment, is bordered by the same Hind III site at the 5′ and a Bam HI site 3′ to it. Region 2 is delineated by this Bam HI site at the 5′ and a Hind III site 3′ to it; region 3 is the 1.2 kb Hind III–Bgl II fragment (1.2 HBg) used as a probe in the experiments described above. Region 4 is outside the 5.0 kb Bgl II fragment and is bordered at the 5′ by the Bgl II site and at the 3′ by a Bam HI site.

As is evident from the restriction enzyme map of the 6.0 kb breakpoint fragment of CML DNA 0319129 (Figure 3C), the Bam HI site from chromosome 22 in region 1 is found on the $9q^+$ chromosome, whereas the Hind III site 5′ to it is missing; therefore, a break must have occurred between these two restriction enzyme sites in region 1. In accordance with this, only a (normal) 3.3 kb Bam HI fragment is detected with the 1.2 HBg probe, which originates from a region 3′ to the Bam HI site (Figure 2B, lane 2). In DNA 02120185, however, this Bam HI site is missing on the $9q^+$ chromosome (Figure 3D) and, therefore, the 1.2 HBg probe detects, in addition to the normal 3.3 kb Bam HI fragment, the 11.3 chimeric Bam HI fragment. As the 3′ Hind III site at region 2 is present in this $9q^+$ fragment, the breakpoint in this DNA is in region 2.

In DNA 0311068, as in DNA 0319129, only a normal 3.3 kb Bam HI fragment is detected with the 1.2 HBg probe,

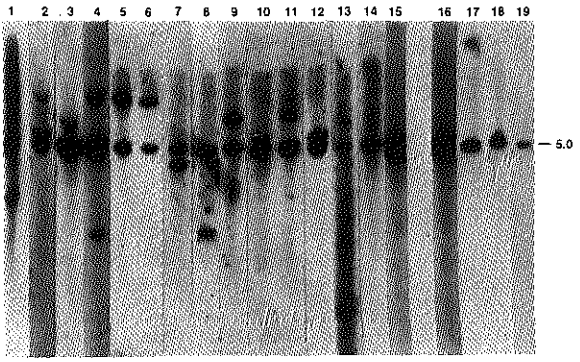


Figure 5. Analysis of DNAs from CML Patients

DNAs analyzed include isolates from CML patient 0311068 (lane 1), 7701C (lane 2), C999 (lane 3), C481 (lane 4), B79-100 (lane 5), B79-216 (lane 6), H80-251 (lane 7), CML 0 (lane 8), H81-164 (lane 9), H81-122 (lane 10), H81-118 (lane 11), H79-179 (lane 12), H77-94 (lane 13), H81-258 (lane 14), C080 (lane 15), C011 (lane 16), and H79-147 (lane 17). Also shown are DNAs isolated from fibroblasts of patient H80-251 (lane 18) and from the fibroblast cell line AG 1732 (lane 19).

DNAs in lanes 1–4 and 15–16 were isolated from the frozen spleens of CML patients; those from CML patients 0311068 and 7701C contained a very high percentage of leukemic cells; the percentage of leukemic cells in the other spleen tissues is not known. There are no data concerning the presence of a Ph′ chromosome in cells of these spleens. DNAs in lanes 7–14 and 17 were isolated from blood and those in lanes 5–6 from bone marrow. Cells from which DNA was isolated were examined for the presence of the Ph′ chromosomal marker: bone marrow and blood cells of all patients except patient H79-147 were Ph′-positive. The human fibroblast cell line AG 1732, obtained from the Human Genetic Mutant Cell Repository (Camden, New Jersey) was established from a CML patient carrying the Ph′ chromosome in her leukemic cells. DNA (10 μg) was digested with Bgl II, electrophoresed on an agarose gel, blotted, and hybridized to the 1.2 kb HBg probe.

indicating that no break has occurred within this fragment (Figure 6, lane 1). When the 0.6 kb Hind III–Bam HI fragment encompassing region 1 is used as a molecular probe, a normal 5.0 kb and two abnormal Bgl II fragments of 4.0 and 3.2 kb are visible. The 3.2 kb Bgl II fragment represents a $9q^+$ chimeric fragment containing the 3′ Bgl II site from the 5.0 kb Bgl II fragment on chromosome 22, as it is also detected with the 1.2 HBg probe (Figure 5, lane 1). The 4.0 kb Bgl II fragment is a $22q^-$ chimeric fragment with the 5′ Bgl II site originating from chromosome 22; it is not detected by the 1.2 HBg probe. The breakpoint in DNA 0311068 must be located in region 1.

The 1.2 HBg probe detects, in addition to a normal 3.3 kb Bam HI fragment, an abnormal 6.2 kb Bam HI fragment in CML DNA 7701C (Figure 6, lane 3). The Bam HI site bordering region 2 at the 5′ must, therefore, be absent from the $9q^+$ chromosome. The Hind III site at the 3′ of region 2 has been retained, as only one normal 4.5 kb Hind III fragment is visible after hybridization to 1.2 HBg (Figure 6, lane 4). This CML DNA must contain a breakpoint in region 2.

Patient C481 and H77-92 apparently have a breakpoint in region 3, encompassing the 1.2 HBg probe. For example, in DNA of patient C481 the 1.2 HBg probe hybridizes to three restriction enzyme fragments in every restriction enzyme digest: abnormal Bgl II fragments of 6.0 and 2.8 kb and a normal one of 5.0 kb (Figure 5, lane 4), abnormal Hind III fragments of 7.0 and 3.5 kb in addition to a normal 4.5 kb fragment (Figure 6, lane 7), abnormal Bam HI fragments of 7.5 and 5.2 kb and a normal 3.3 kb fragment (Figure 6, lane 8). Therefore, in this CML DNA, the 1.2 HBg probe detects both the $22q^-$ and $9q^+$ breakpoint fragments.

The situation in the DNA of patient C080 is less clear. As only one normal 5.0 kb Bgl II fragment is visible after hybridization to 1.2 HBg (Figure 5, lane 15), a chromosomal breakpoint most likely has occurred outside the Bgl II fragment. As the 1.2 HBg probe detects an abnormal 5.0 kb Hind III fragment in addition to the normal 4.5 kb Hind III fragment (Figure 6, lane 6), a chromosomal breakpoint may be situated immediately 3′ of the 5.0 kb Bgl II fragment. This is supported by the hybridization of the same probe to an abnormal 13 kb Bam HI fragment (and the normal 3.3 kb fragment, Figure 6, lane 5). The breakpoint has therefore been tentatively placed in region 4; in DNA of this patient, the 1.2 HBg probe would detect only $22q^-$ restriction enzyme fragments.

Using different restriction enzymes and probes from the 5.0 kb Bgl II fragment, we have analyzed the location of the breakpoint in all CML DNAs shown in Figure 5, lane 1–15. These results are summarized in Table 1.

Discussion

In the present studies we have identified and cloned a breakpoint cluster region (bcr) on chromosome 22, involved in the chromosomal translocation, t(9;22), of Ph′-positive CML. In total we have studied 19 CML patients; ten of these were shown to be Ph′-positive by cytogenetical analysis. All of the patients of this latter group pos-

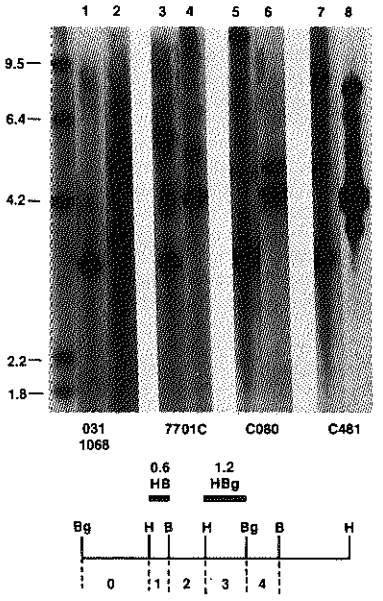


Figure 6. Analysis of Ph′ Translocation Breakpoints on Chromosome 22

Top: DNA of patients was digested with Bam HI (lanes 1, 3, 5, 7), Bgl II (lane 2), or Hind III (lanes 4, 6, 8). Hybridization was with the 1.2 kb HBg probe; lane 2 is hybridized with a 0.6 kb HB probe indicated in the bottom of the figure. Molecular weights of marker fragments in kilobases are indicated in the left of the figure. Bottom: bcr, the region within the cloned chromosome 22 sequences containing Ph′ translocation breakpoints identified in this study, is shown schematically. The numbers refer to the subregions mentioned in the text; the probes used are indicated above the figure.

sessed a chromosomal break within bcr. Of the remaining nine patients, one was cytogenetically characterized as Ph′-negative and a second patient has Ph′-negative juvenile CML; as expected, they did not exhibit a breakpoint in this region. Seven of seven nonkaryotyped patients were Ph′-positive because a chromosomal break could be identified within bcr. The involvement of bcr in the Ph′ translocation is highly specific for CML, as analogous rearrangements were not found in DNAs isolated from other neoplastic tissues or cell lines, including DNAs from four acute myeloid leukemia patients, one acute myelomonocytic leukemia patient, glioblastoma, melanoma, multiple myeloma, and teratocarcinoma cell lines (data not shown). Since abnormalities were not seen in a fibroblast cell line established from a Ph′-positive CML patient, in cultured fibroblasts of a Ph′-positive CML patient, or in leukemic cells of two Ph′-negative CML patients, we believe these rearrangements to be highly specific for the leukemic cells in Ph′-positive CML patients. In two patients these rearrangements were rigorously analyzed and shown to represent chromosomal breakpoints. Probes, isolated

Table 1. Breakpoint Location within bcr of Ph′-positive CML Patients

CML Patient	Breakpoint Location	CML Patient	Breakpoint Location
0311068	1	H81-122	2, 3
7701C	2	H81-118	2, 3
C999	1	H79-179	1
C481	3	H77-94	3
B79-100	2	H81-258	1, 2
B79-216	0, 1, 2	C080	4
H80-251	0, 1	0319129	1
CML-0	2	02120185	2
H81-164	2, 3		

The different breakpoint subregions (0–4) are as indicated in Figure 6, bottom. For some patients, the exact breakpoint subregion has not yet been determined; more than one subregion is indicated for these patients in the table.

from the bcr, in particular the 1.2 kb HBg probe, are highly specific tools for the identification of the Ph′ translocation in leukemic DNA and as such, may be of diagnostic value, in particular when no metaphase chromosome spread can be obtained.

We have molecularly cloned a region of chromosome 22 from non-CML DNA that contains the Ph′ breakpoints in CML DNA. The orientation of the chromosome 22-specific sequences in the two $9q^+$ breakpoint fragments determines the orientation of the breakpoint cluster region on the chromosome: the most 5′ end will remain on chromosome 22 (Figure 7) and, depending on the exact position of a breakpoint, a smaller or larger region of bcr will be translocated to chromosome 9. Although the breakpoint on the Ph′ chromosome is in band q11 (Rowley, 1973; Yunis, 1983) and the λ light-chain constant region has been localized to the same band (McBride et al., 1982; our unpublished results) no cross-homology was observed between the chromosome 22 cosmid clones and Cλ. As of yet, these clones contain unidentified sequences. Experiments to determine if this region of chromosome 22 contains protein coding regions and/or enhancer sequences are in progress.

Previously we have reported the isolation of a Ph′ breakpoint fragment containing chromosome 9 sequences originating approximately 14 kb 5′ of human v-*abl* homologous sequences (Heisterkamp et al., 1983b). However, we were unable to detect chromosomal breakpoints in other CML DNAs up to 11 kb 5′ of this breakpoint. The $9q^+$ breakpoint fragment from a second CML patient, isolated in the present study, contains 9 kb of DNA originating from chromosome 9. Preliminary results suggest that these sequences must be separated by, at minimum, 27 kb of DNA 5′ of the previously reported breakpoint. Therefore, the breakpoint on chromosome 9 appears to be variable in different Ph′-positive CML patients and may be found within a relatively large but limited region on chromosome 9; the region of chromosome 9

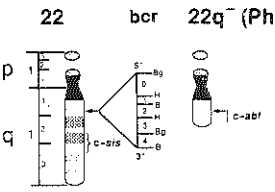


Figure 7. Diagrammatic Representation of the Ph' Translocation

The horizontal arrows indicate the chromosomal breakpoint in chromosome 22. Mapping of c-*sis* to the region of chromosome 22 (q12.3 to q13.1) translocated to chromosome 9 is as previously described (Groffen et al., 1983a; Bartram et al., 1983a). Localization of c-*abl* to the terminal portion of chromosome 9 (q34), which is translocated to chromosome 22 in the Ph' translocation, is as described in the text. The maximum size of bcr, from the 5' Bgl II to the 3' Bam HI site, is 5.8 kb.

containing human c-*abl* (q34-qter) that is translocated to chromosome 22 is too small to be visualized by cytogenetic analysis and does not exceed 5000 kb (Heisterkamp et al., 1983b; de Klein et al., 1982).

At present, we do not know whether the two breakpoints we have identified actually occur within human c-*abl* coding sequences: the most 5' exon of human c-*abl* has not yet been determined. However, the possibility must be considered that the human c-*abl* oncogene extends over a much larger region than that characterized by homology to the viral oncogene v-*abl*; whereas the v-*abl* oncogene is 3.5 kb in length (Goff et al., 1980). Homologous RNA species ranging in size from 5–7 kb (Ozanne et al., 1982; Westin et al., 1982), have been reported in humans. In contrast to the situation on chromosome 9, the breakpoints on chromosome 22 seem to be clustered in a very limited region. It is evident, however, even at the restriction enzyme level, that the breakpoints have not occurred in one specific site but rather are distributed over a region of up to 5.8 kb.

Analogous to the t(9;22) in Ph'-positive CML, a t(8;14) is found in many patients with Burkitt lymphoma: in the latter case, the human oncogene c-*myc*, located on chromosome 8, (Dalla-Favera et al., 1982; Taub et al., 1982) is often translocated into the immunoglobulin heavy-chain locus on chromosome 14 (Taub et al., 1982; Dalla-Favera et al., 1983; Adams et al., 1983; Erikson et al., 1983; Hamlyn and Rabbits, 1983). The breakpoints on chromosome 8 may be distributed over a relatively large region 5' of human c-*myc* (Dalla-Favera et al., 1983; Adams et al., 1983; Erikson et al., 1983), a situation analogous to that of human c-*abl* on chromosome 9. On chromosome 14, breakpoints in the variable (Erikson et al., 1982) and in the constant region of the heavy-chain locus have been reported in Burkitt lymphoma, indicating that neither the breakpoints on chromosome 8 nor those 14 can be found within a breakpoint cluster region as discussed in the present study. The bcr on chromosome 22 seems as of yet to be unique in human. However, the existence of bcrs on other human chromosomes is not unlikely taking into consideration the increasing number of reports of other highly specific translocations in neoplastic diseases (for a review, see Yunis, 1983).

The specificity of the presence of a chromosomal breakpoint on chromosome 22 within bcr in Ph'-positive CML indicates that this region may be involved in CML. Additionally, a human oncogene, c-*abl*, is consistently translocated to chromosome 22, even in patients with complex Ph' translocations (Bartram et al., 1983b) and is amplified in a CML cell line, K562, (Heisterkamp et al., 1983b). Although the breakpoints on chromosome 9 are distributed over a relatively large region of DNA 5' to human v-*abl* homologous sequences, the specific translocation of this oncogene in the t(9;22) must be of functional significance. Therefore, we believe that both human c-*abl* and bcr may be associated with Ph'-positive CML.

Experimental Procedures

Somatic Cell Hybrids

PgME-25 Nu is a human–mouse somatic cell hybrid obtained from fusion with mouse Pg19 cells; it contains as its only human component chromosome 22. Chinese hamster cell line E36 was used to obtain hybrids 10CB-23B and 14CB-21A. The hybrid 10CB-23B contains human chromosomes 5, 9, 11 and 19, whereas 14CB-21A has retained chromosomes 4, 7, 8, 14, 20, and 9q⁺ (Geurts van Kessel et al., 1981b; Geurts van Kessel et al., 1981c; de Klein et al., 1983). WESP 2A was obtained by fusion of mouse WEHI-3B cells with leukocytes of a Ph'-positive CML patient and contains human chromosomes 7, 8, and 14 in addition to the Ph' chromosome (de Klein et al., 1983).

Southern Blot Analysis

High molecular weight DNAs were isolated as described (Jeffreys and Flavell, 1977), digested with restriction enzymes, and electrophoresed on agarose gels. Blotting was according to Southern (1975) on nitrocellulose (Schleicher and Schuell, pH 7.9). Nick translation of probes and filter hybridizations were as described (Flavell et al., 1978; Bernards and Flavell, 1980). Specific activity of the probes were $2–5 \times 10^8$ cpm/μg. Filters were exposed to XAR-2 film (Kodak) at −70°C with Dupont Lightning Plus intensifying screens.

Isolation of Probes

DNA probes were prepared by digestion with appropriate restriction enzymes, followed by electrophoresis through low-melting-point agarose gels. Desired bands were cut from the gel and brought into solution by heating at 65°C for 30 min. Agarose was removed by two extractions with phenol equilibrated with 0.3 M NaOAc (pH 5.0), and one extraction with phenol/chloroform/isoamylalcohol (25:24:1). DNA was precipitated with ethanol and 0.2 M NaOAc (pH 4.8) in the presence of 20 μg/ml Dextran T-500 as carrier. Restriction enzymes and low-melting-point agarose were purchased from Bethesda Research Laboratories and were used according to the supplier's specifications.

Molecular Cloning

Subcloning of the 5.0 kb Bgl II fragment and cloning of the 11.3 kb Bam HI fragment was according to published procedures (Groffen et al., 1983b). A previously described (Groffen et al., 1982) human lung carcinoma cosmid library was screened with the 1.2 HBg probe according to the method of Grosveld et al. (1981). Three positive cosmid clones were isolated and mapped independently by digestion of individual restriction enzyme fragments isolated from low-melting-point agarose gels.

Acknowledgments

We thank Prof. Dick Bootsma for helpful comments on the manuscript; Anne Hagemeijer and Ton van Agthoven for supplying patient material and for cytogenetical analysis; P. Leder and U. Storb for the generous gifts of HuA5 and pVA1; Pamela Hansen and Gail Blennerhassett for technical

assistance; Beverly Bales for help with the preparation of the manuscript. The work was supported by the Netherlands Cancer Society (Koningin Wilhelmina Fonds) and under NCI PHS contract NOI-CP-75380. C. R. B. is a recipient of a fellowship from the Deutsche Forschungsgemeinschaft.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received November 14, 1983

References

Adams, J. M., Gerondakis, S., Webb, E., Corcoran, L. M., and Cory, S. (1983). Cellular *myc* oncogene is altered by chromosome translocation to an immunoglobulin locus in murine plasmacytomas and is rearranged similarly in human Burkitt lymphomas. Proc. Nat. Acad. Sci. USA *80*, 1982–1986.

Bartram, C. R., de Klein, A., Hagemeijer, A., Grosveld, G., Heisterkamp, N., and Groffen, J. (1983a). Localization of the human c-*sis* oncogene in Ph'-positive and Ph'-negative chronic myelocytic leukemia by *in situ* hybridization. Blood, in press.

Bartram, C. R., de Klein, A., Hagemeijer, A., van Agthoven, T., Geurts van Kessel, A., Bootsma, D., Grosveld, G., Ferguson-Smith, M. A., Davies, T., Stone, M., Heisterkamp, N., Stephenson, J. R., and Groffen, J. (1983b). Translocation of the human c-*abl* oncogene occurs in variant Ph'-positive but not Ph'-negative chronic myelocytic leukaemia. Nature, *306*, 277–280.

Bernards, R., and Flavell, R. A. (1980). Physical mapping of the globin gene deletion in hereditary persistence of foetal haemoglobin (HPFH). Nucl. Acids Res. *8*, 1521–1534.

Dalla-Favera, R., Gelmann, E. P., Gallo, R. C., and Wong-Staal, F. (1981). A human onc gene homologous to the transforming gene (v-*sis*) of simian sarcoma virus. Nature *292*, 31–35.

Dalla-Favera, R., Bregni, M., Erikson, J., Patterson, D., Gallo, R. C., and Croce, C. M. (1982). Human c-*myc* onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. Proc. Nat. Acad. Sci. USA *79*, 7824–7827.

Dalla-Favera, R., Martinotti, S., Gallo, R. C., Erikson, J., and Croce, C. M. (1983). Translocation and rearrangements of the c-*myc* oncogene locus in human undifferentiated B-cell lymphomas. Science *219*, 963–967.

de Klein, A., Geurts van Kessel, A., Grosveld, G., Bartram, C. R., Hagemeijer, A., Bootsma, D., Spurr, N. K., Heisterkamp, N., Groffen, J., and Stephenson, J. R. (1982). A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukaemia. Nature *300*, 765–767.

de la Chapelle, A., Lenoir G, Boué, J., Boué, A., Gallano, P., Huerre, C., Szajner, M.-F., Jeanpierre, M., Laloue, J. M., and Kaplan, J.-C. (1983). Lambda Ig constant region genes are translocated to chromosome 8 in Burkitt's lymphoma with t(8;22). Nucl. Acids Res. *11*, 1133–1142.

Erikson, J., Finann, J., Nowell, P. C., and Croce, C. M. (1982). Translocation of immunoglobulin V_H genes in Burkitt lymphoma. Proc. Nat. Acad. Sci. USA *79*, 5611–5615.

Erikson, J., ar-Rushdi, A., Drwinga, H. L., Nowell, P. C., and Croce, C. M. (1983). Transcriptional activation of the translocated c-*myc* oncogene in Burkitt lymphoma. Proc. Nat. Acad. Sci. USA *80*, 820–824.

Flavell, R. A., Kooter, J. M., de Boer, E., Little, P. F. R., and Williamson, R. (1978). Analysis of the β-δ-globin gene loci in normal and Hb Lepore DNA: direct determination of gene linkage and intergene distance. Cell *15*, 25–41.

Geurts van Kessel, A. H. M., den Boer, W. C., van Agthoven, A. J., and Hagemeijer, A. (1981a). Decreased tumorigenicity of rodent cells after fusion with leukocytes from normal and leukemic donors. Somatic Cell Genet. *7*, 645–656.

Geurts van Kessel, A. H. M., ten Brinke, H., Boere, W. A. M., den Boer, W. C., de Groot, P. G., Hagemeijer, A., Meera Kahn, P., and Pearson, P. L. (1981b). Characterization of the Philadelphia chromosome by gene mapping. Cytogenet. Cell Genet. *30*, 83–91.

Geurts van Kessel, A. H. M., van Agthoven, A. J., and Hagemeijer, A. (1981c). Clonal origin of the Philadelphia translocation in chronic myeloid leukemia demonstrated in somatic cell hybrids using an adenylate kinase-1 polymorphism. Cancer Genet. Cytogenet. *6*, 55–58.

Goff, S. P., Gilboa, E., Witte, O. N., and Baltimore, D. (1980). Structure of the Abelson murine leukemia virus genome and the homologous cellular gene: studies with cloned viral DNA. Cell *22*, 777–785.

Groffen, J., Heisterkamp, N., Grosveld, F., Van de Ven, W. J. M., and Stephenson, J. R. (1982). Isolation of human oncogene sequences (v-*fes* homolog) from a cosmid library. Science *216*, 1136–1138.

Groffen, J., Heisterkamp, N., Stephenson, J. R., Geurts van Kessel, A., de Klein, A., Grosveld, G., and Bootsma, D. (1983a). c-*sis* is translocated from chromosome 22 to chromosome 9 in chronic myelocytic leukemia. J. Exp. Med. *158*, 9–15.

Groffen, J., Heisterkamp, N., Blennerhassett, G., and Stephenson, J. R. (1983b). Regulation of viral and cellular oncogene expression by cytosine methylation. Virology *126*, 213–227.

Grosveld, F. G., Dahl, H.-H. M., de Boer, E., and Flavell, R. A. (1981). Isolation of β-globin-related genes from a human cosmid library. Gene *13*, 227–237.

Hamlyn, P. H., and Rabbits, T. H. (1983). Translocation joins c-*myc* and immunoglobulin 1 genes in a Burkitt lymphoma revealing a third exon in the c-*myc* oncogene. Nature *304*, 135–139.

Heisterkamp, N., Groffen, J., Stephenson, J. R., Spurr, N. K., Goodfellow, P. N., Solomon, E., Carrit, B., and Bodmer, W. F. (1982). Chromosomal localization of human cellular homologues of two viral oncogenes. Nature *299*, 747–749.

Heisterkamp, N., Groffen, J., and Stephenson, J. R. (1983a). The human v-*abl* cellular homologue. J. Mol. App. Genet. *2*, 57–68.

Heisterkamp, N., Stephenson, J. R., Groffen, J., Hansen, P. F., de Klein, A., Bartram, C. R., and Grosveld, G. (1983b). Localization of the c-*abl* oncogene adjacent to a translocation breakpoint in chronic myelocytic leukemia. Nature, *306*, 239–242.

Jeffreys, A. J., and Flavell, R. A. (1977). A physical map of the DNA regions flanking the rabbit β-globin gene. Cell *12*, 429–439.

McBride, O. W., Hieter, P. A., Hollis, G. F., Swan, D., Otey, M. C., and Leder, P. (1982). Chromosomal location of human kappa and lambda immunoglobulin light chain constant region genes. J. Exp. Med. *155*, 1480–1490.

Miller, J., Bothwell, A., and Storb, U. (1981). Physical linkage of the constant region genes for immunoglobulins λ1 and λIII. Proc. Nat. Acad. Sci. USA *78*, 3829–3833.

Ozanne, B., Wheeler, T., Zack, J., Smith, G., and Dale, B. (1982). Transforming gene of a human leukaemia cell is unrelated to the expressed tumour virus related gene of the cell. Nature *299*, 744–747.

Rowley, J. D. (1973). A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. Nature *243*, 290–293.

Rowley, J. D. (1982). Identification of the constant chromosome regions involved in human hematologic malignant diseases. Science *216*, 749–751.

Sandberg, A. A. (1980). The Chromosomes in Human Cancer and Leukemia. (New York: Elsevier).

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. *98*, 503–517.

Taub, R., Kirsch, I., Morton, C., Lenoir, G., Swan, D., Tronick, S., Aaronson, S., and Leder, P. (1982). Translocation of the c-*myc* gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. Proc. Nat. Acad. Sci. USA *79*, 7837–7841.

Westin, E. H., Wong-Staal, F., Gelmann, E. P., Dalla-Favera, R., Papas, T. S., Lautenberger, J. A., Eva, A., Reddy, E. P., Tronick, S. R., Aaronson, S. A., and Gallo, R. C. (1982). Expression of cellular homologues of retroviral *onc* genes in human hematopoietic cells. Proc. Nat. Acad. Sci. USA *79*, 2490–2494.

Yunis, J. J. (1983). The chromosomal basis of human neoplasia. Science *221*, 227–236.