NEW TOOLS FOR SOMATIC CELL GENETICS: APPLICATION TO HUMAN CHROMOSOME 4

By

Barbara L. Handelin, B.A.

A dissertation

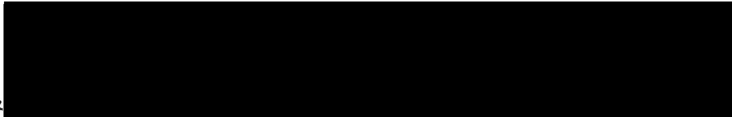
Presented to the Department of Medical Genetics
and the Oregon Health Sciences University
School of Medicine

in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

August 1987

APPROVED:



David E. Housman, Ph.D.
Professor in Charge of Thesis

John A. Resko, Ph.D.
Chairman, Graduate Council

ACKNOWLEDGEMENTS

Five years, two laboratories, two institutions, two U.S. coasts and two advisors have been traveled through in the making of this thesis. There are so many people to thank, so many jokes to remember to tell...

The transition from Portland to Boston, where the vast majority of the research presented here was accomplished, would never have been possible without the generous support and efforts of Dr. David Kabat, Dr. Robert Koler and my anonymous supporters on the OHSU Graduate Council who approved my transfer. To all of them, thank you.

To Dr. David Housman, I do not think I can ever extend enough appreciation and gratitude for giving me an opportunity to work in his laboratory despite the coast-to-coast predicament. Because he has the grace and wisdom to see each member of his lab as an individual with unique talents and limitations, it was possible for me to set appropriate goals and to begin to feel successful in realizing them. For opening all the doors of opportunity and for allowing me to remember the "subtle, sneaky, important reason why [I] was born a human being and not a chair" (Herb Gardner, A Thousand Clowns), I am more than thankful (> thankful).

The realization of this thesis was also made possible by tirelessly patient instructions from Tom Glaser and was spirited by the great humor provided in Room 536 by Jack Hensold, Nic Dracopoli and Tom Glaser (again?). To Nic, my pusher, special thanks for developing my addiction to "Considering Every Little Thing". I especially also need to thank Jack for all the music (even 'Wayne' Newton), for the Chloe stories, for the SDS, NaCl and for the concept of "Why do something once when you can do it three or four times?"

I am sure that Jack joins me in remembering and thanking our neighbor Lewis Chodosh for the various interior decoratings of our bay. "When clowns go bad" is Gary Larson's expression of Dr. Chodosh's efforts, which all go down in significant memory. To Lewis I owe the levitation of my spirits during very tough times, whether via a giant yellow chicken bearing balloons or simply via outstanding generosity of heart, listening ears or creative humor.

(The following should be heard in Garrison Keilor's voice as when he is reading the greetings on Prairie Home Companion:) To Flash Kent I owe thanks for allowing me to bestow an inspired name. Thanks to Edwin Geissler for dancing and singing with me in the lab, to Kathy Call for being a Woman in science and to all in the Housman lab for being a comfortable and supportive second home to me through times when I was emotionally without one.

Finally, special recognition goes to Mary Stenzel-Poore, who has been and will always be my professional and personal soulmate. And certainly, I could never have made it to the end financially or spiritually without the support of my parents nor without and somehow despite my husband, Tim Donlon, who always thought I "was going to be an Achievement".

And truly finally, an admission: as suspected by many, it is true that Erin did most of the Big Thoughts Work behind these experiments and provided most of the technical expertise as well. A truer companion and partner there never was.

TABLE OF CONTENTS

Item	Page Number
Acknowledgements	iii
Table of Contents	iv
Abstract	v
Introduction	1
General Materials and Methods	23
Chapter 1	
Introduction	26
Materials and Methods	31
Results	36
Figures and Tables	42
Discussion	53
Chapter 2	
Introduction	57
Materials and Methods	62
Results	66
Figures and Tables	72
Discussion	92
Chapter 3	
Introduction	96
Materials and Methods	98
Results	101
Figures and Tables	106
Discussion	120
Chapter 4	
Introduction	123
Materials and Methods	127
Results	130
Figures and Tables	135
Discussion	150
Chapter 5	
Introduction	155
Cell Surface Receptors...	158
Summary	163
References	166

ABSTRACT

We have sought efficient methods for introducing dominant selectable markers into targeted regions of human chromosomes. Introduced markers provide opportunities to selectively mobilize large fragments surrounding the marker for transfer into heterologous genetic/cellular environments. The isolation of such a limited chromosome region may be the first step toward cloning a gene of interest located in the region, or it may be an enriched source of random genomic clones from that region. For some very large genes, such transfer of chromosome fragments may be the only method of gene transfer which will permit expression of the gene. Finally, transferring a chromosome region into heterologous cells which contains a gene known only by its chromosome location and phenotype may permit in vitro studies on the biochemical and physiological effects of its gene product.

We chose the region of human chromosome 4 which contains the putative Huntington's disease (HD) gene as a model system for demonstrating the feasibility of introducing a dominant selectable marker into a particular chromosome region. Our goal was to introduce a marker near enough to the HD locus (at 4p14-16) that would enable the transfer of disease gene into heterologous cells.

In a stepwise process, we first derived a monochromosomal hybrid retaining human chromosome 4 and a dominant selectable marker carried on a retroviral vector (neo). We then transfered the selectable marker into heterologous cells and the resulting limited somatic cell hybrids were found to contain a small region (1-5,000 kb) of human DNA which had been co-transfered with the neo marker. By three methods we established that the neo marker was integrated in 4p14-16; therefore, these hybrids contain a terminal portion of chromosome 4, although they do not contain the

two markers most closely linked to HD. Such hybrids will be useful for isolating new markers near the HD locus and may even contain the HD gene itself.

We also explored the utility of a monochromosomal hybrid for conventional mapping of a unique gene and for mapping low and mid-order repetetive human sequences.

The construction and use of a new retroviral vector, SP-1, is also described. The vector carries a new dominant selectable marker and is easily cloned from integration sites in host DNA. SP-1 introduces recognition sequences for rarely-cutting restriction sites and we provide preliminary evidence that the degree of methylation of these sites in proviral SP-1 is reflective of the tendency of this retrovirus to integrate preferentially in hypomethylated host sites.

Finally, we describe initial experiments directed at applying this same methodology to mouse chromosome 5; specifically to the region containing the ecotropic murine leukemia virus receptor gene. To facilitate the somatic cell genetics, we develop two assays which easily distinguish receptor-bearing cells from non-receptor-bearing cells. These are described here.

INTRODUCTION

In a basic sense, all scientific disiplines endeavor to achieve the same fundamental goal; to break down a complex system into unit parts which can be studied in controlled experiments which test only a limited number of variables. The assumption is that, by isolating and manipulating the unit parts, the original system may be reconstructed with a clearer understanding of how it functions. Thus the biochemist purifies and crystalizes individual polypeptide chains to study how one interacts with another in a cascade of biochemical reactions; the physicist strives to identify fundamental particles as a basis for understanding the nature of mass and energy reactions; the organic chemist tracks individual electrons as they are passed between the atoms of simple organic molecules in the hope of being able to explain intramolecular energy bonds; and the molecular biologist purifies the unit of life - genes.

Since 1865 when Mendel demonstrated that some traits of plants and animals are transmitted from parent to offspring in a predictable and mathematically stable fashion, geneticists have striven to characterize and understand the organization of the transmitting genetic material. In this nearly century old discipline, which is currently in an exponential phase of growth, how can a few experiments such as those represented in this thesis find their significance or relevence? This thesis demonstrates the technical feasibility of new tools for somatic cell genetics; the relevence for such types of experiments (more technical than functional) is elevated by T.C. Hsu in his historical review of cytogenetics (Hsu, 1979). In his preface, Dr Hsu notes a change of attitude in the scientific community about the importance of technique. The historical attitude was

that "methodological and technological achievements were subordinate to fact-finding missions and conceptual advancements", whereas, in the face of a burgeoning arena of technique-based disciplines such as molecular biology, the modern community recognizes that a "new technique may beget new sets of pertinent facts, which in turn may beget new concepts and new theories."

A few examples of technical advances which have provided the tools for whole new fields of research make the point most directly:

> Cytogenetics grew quickly after the accidental discovery by Hsu (1953) that pretreatment of mitotic cells with hypotonic buffer enhances spreading and morphology of metaphase chromosomes. For example, this simple improvement on technique combined with air-drying of the microscope slides (Rothfels and Siminovich, 1958) made it possible to establish for the first time the correct number of human chromosomes as 46 (Tjio and Levan, 1956; Ford and Hamerton, 1956).

> The development by Puck(1956) of a protocol for culturing true clones of eukaryotic cells was the foundation for the ascension of modern mammalian cell biology to a level of experimental sophistication which had previously only been achieved in microbiology.

> The simple observation by Harris and Watkins (1965) that inactivated Sendai virus promotes efficient fusing of co-cultivated cells triggered the rapid development of all of subsequent somatic cell genetics using interspecies cell hybrids.

> The desire to apply experimental designs which had been used in microbiology to eukaryotic systems led Graham and Van der Eb (1973)to develop a modification of the $CaPO_4$-mediated gene transfer technique for eukaryotic cells. This technique, plus the astute usage of species-specific repetitive DNA sequences (Gusella et al 1982), were the critical tools necessary for identifying and cloning dozens of eukaryotic genes which were previosly known only by their cellular function (Murray et al 1981).

> The Southern (1975)method for transferring DNA via capillary action from

electrophoresis gels onto stable membranes using paper hand towels as blotting wicks is perhaps one of the least elegant, yet fundamental techniques of molecular biology. This method, which incorporated a technique for molecular hybridization to these membranes, replaced liquid hybridization. and revealed for the first time the physical structure of eukaryotic genes.

> The power of dominant, selectable markers such as antibiotic resistance to select for rare transformation events in bacteria, led Mulligan and Berg (1980) to design plasmid vectors which conferred drug resistance to eukaryotic cells. The subsequent engineering of retrovirus vectors also carrying dominant selectable markers which can efficiently infect cells from a wide variety of species and tissue types, placed the cornerstone for the current fields of transgenic animal models and experimental gene therapy .

> Most recently, Cantor and Schwartz, developed a unique electrophoresis system which can separate DNA fragments in the 100 kb to 100 megabase size range (Schwartz and Cantor 1984). Applying this technology, yeast geneticists could visualize and measure for the first time the size of yeast chromosomes (Carle and Olson 1984), bacterial geneticists could physically map the entire E. Coli genome (Smith et al 1987), and human geneticists can make the first correlations between the physical and genetic maps of the human genome (Kenwich et al 1987; Burmeister and Lehrach, 1987; Lawrance et al 1987; Estivill et al 1987).

Most of these examples of technical advances have heavily influenced the field of somatic cell genetics where the goal has been to systematically categorize, map and isolate smaller and smaller pieces of mammalian chromosomes. A chronology of the hallmark achievements in this field reads like an experimental protocol which endeavors to stepwise purify any given region of the genome; the scheme starts with isolating one or a few intact chromosomes, followed by identifying a linear genetic map in order to locate a locus of interest and ends by isolating a limited region that contains that locus.

To be able to carry out such a scheme in a directed manner for any region of interest in the human genome and in a contemporary, rather than historical, time scale has been the goal of the entire field of human somatic cell genetics. In particular, it has been one of the focuses of work in David Housman's laboratory since 1970 and this thesis was built on that laboratory history.

Single cell cloning (Puck 1956) was the requisite first step in developing a system for isolating single chromosomes from one species in the genetic background of another. Somatic cell genetics was truly born, however, after the discovery by Okada (1958) that HVJ (heamagglutinating virus of Japan) caused two clones of infected tumor cells to spontaneously fuse in culture. Barskie et al also observed spontaneous cell fusion at a very low frequency (1960) and Ephrussi recognized the value of these heterokaryons and their resulting stable cell hybrids (Ephrussi and Weiss, 1965), but it was not until Harris and Watkins noted the highly efficient fusion of two cell clones in the presence of inactivated Sendai virus (1965) that this technique became widely applicable and somatic cell genetics got underway.

Capitalizing on the culture technique of Puck, several groups began to isolate mutants in cultured mouse and hamster cells. The first and still widely used mutants, developed by Szybalska and Szybalski (1962), were $HPRT^-$ mouse cells (deficient in hypoxanthine phosphoribosyl transferase activity required in the purine salvage pathway) . This marker can be selected both for and against; $HPRT^-$ mutants are resistant to 6-thioguanine in the culture media and therefore survive selection. On the other hand, $HPRT^+$ revertants are sensitive to 6-thioguanine and resistant to the HAT selective media containing hypoxanthine, aminopterin and thymidine (Littlefield, 1964). TK^- and $APRT^-$ mutants are also selectable markers in this pathway (Littlefield, 1964 and Chasin, 1974, respectively).

After the spontaneous transformation of a Chinese hamster ovary cell culture in 1958 and the isolation of the first auxotrophic mutant (CHO pro^-) in 1963, it was

soon demonstrated that this cell line could be easily mutagenized by a variety of agents providing many useful markers for somatic cell genetics. Dozens of other auxotrophic CHO mutants have since been derived (Puck and Kao 1982) and used for gene mapping as discussed below. The CHO cell line has, in general, proven to be an extremely valuable reagent. It has relatively few chromosomes which can be easily identified and are stably maintained. Its short generation time and high plating efficiency make growing and maintaining large cultures relatively easy. It hybridizes well with human cells and rapidly eliminates human chromosomes in the absence of selective pressure.

During fusion experiments between TK$^-$ B82 mouse cells and human auxotrophic cells, Weiss and Green (1967) noted the preferential loss of human chromosomes in the resulting cell hybrids. This selective loss of human chromosomes occured during an intial acute phase of culturing hybrids but this phase was usually followed by stabilization of a cell line containing one to 15 or 16 heterologous chromosomes. Hamster/human hybrids have since also been shown to segregate human chromosomes (Kao and Puck 1970). This still poorly understood phenomenon has provided the opportunity to create many interspecies hybrids, each containing a subset of the chromosome complement from the species which is segregating in the hybrid.

These three developments - the report of a method for propogating pure mutant CHO and mouse cell clones, and the discoveries of cell fusion and the preferential loss of human chromosomes from human/rodent interspecies cell hybrids - provided a methodology to create hybrids which selectively retained specific chromosomes. The focus from here forward will be on hybrids retaining specific human chromosomes. For example, CHO cells which were defective in proline uptake (pro$^-$) were fused with human fibroblasts and the fusion hybrids were selected in proline deficient medium. Human chromosome 10 was consistently retained in these hybrids which lead Jones et al (1975) to hypothesize the existence of a pro$^-$ complementing gene on chromosome 10. By similar protocols other human chromosomes have been selectively retained in fusion

hybrids such as chromosome 11 by a glycine$^-$CHO cell mutant (Jones et al 1972); chromosome 21 by an adenine C $^-$ hamster mutant (Moore, 1977); chromosome 12 (Law and Kao, 1978); chromosome 9 by a glycine-adenine-thymidine$^-$ CHO mutant (Jones et al, 1980); chromosome 14 by an adenine E$^-$ CHO cell mutant (Jones et al 1981)); and chromosome X by HPRT$^-$ mouse or hamster cells (Migeon and Miller,1981). The retention of single human chromosomes in these first hybrids also permitted the assignment of the human complementing gene to that particular chromosome and this was the beginning of human gene mapping by somatic cell genetics.

Unfortunately, there were not enough complementation schemes available to selectively retain each of the 24 distinct human chromosomes in seperate hybrids. Therefore a panel of hybrids containing overlapping sets of human chromosomes were developed by exploiting the observation that hybrid cells which retained one human chromosome due to selective pressure also usually retained others. The standard protocol for creating hybrids thus became one in which TK$^-$ or HPRT$^-$ rodent cells were fused with human cells and the resulting fusion hybrids were selected in HAT medium to select for the retention of the human X chromosome. The human fusion parent cells were selectively killed by oubain at 10^{-6}M (a dosage which spares rodent cells), and the HPRT$^-$ or TK$^-$ rodent parents by HAT selection. The resulting hybrids were karyotyped to identify which human chromosomes had been retained.

These hybrids have been the reagents for mapping hundreds of human genetic markers using the principle of co-segregation. A new marker is mapped by scoring each member of a panel of hybrids for the presence(+) or absence (-) of a particular marker and looking for the concordant presence or absence of one human chromosome across the panel. There are usually discordancies between the +/- pattern of the marker across the panel and the +/- pattern of every candidate chromosome. These discordancies arise for two reasons: (1) if two or more chromosomes are always present together in the same hybrids then it is not possible to discriminate between them, and

(2) if small rearrangements and deletions are missed in the karyotypical characterization of the hybrids, then markers which are in the deletion or rearrangement are scored as absent although the chromosome on which that marker actually resides is apparently present. It has, in addition, been difficult in the past to identify small human chromosome fragments in hybrids, especially if they become associated with a rodent chromosome; these go undetected , yet give (+) scores for any markers which may happen to reside on the fragment. Dissatisfaction with this ambiguity has created the need for a panel of 24, exhaustively characterized hybrid cells lines each containing only one intact human chromosome.

In an effort to create such mono-chromosomal human/rodent hybrids both for gene mapping purposes and as a first step in isolating genes which are known only by their phenotypes such as tumor suppression (Killary and Fournier, 1984; Ruddle, 1977), or HLA expression (Goodfellow et al 1982) the microcell mediated chromosome transfer method was developed. In this method, chromosome donor cells are treated with 0.01 - 0.2 ug/ml colcemid for 48 hours which prevents segregation of anaphase chromosomes but allows repeated constrictions of the nuclear membrane. This is followed by a brief exposure to cytochalasin-B, resulting in micronucleation. Micronuclei are small, membrane bound vesicles budding from the nuclear membrane, each containing one or a few chromosomes. These 'microcells' can be selectively collected by centrifugation and then allowed to fuse with an excess of recipient cells in a standard PEG fusion protocol to transfer the chromosome(s) contained in the microcell. With optimum parent cell lines and culture conditions, single donor chromosomes can be effectivley transferred to recipient cells via this technique. However, experience with a variety of donor and recipient cell lines has unfortunately established that most human cell lines are poor chromosome donors due to their resistance to micronucleation. Several human chromosomes, rather than one, are usually transferred in these fusion experiments. By performing a second round of transfers from the primary hybrid to a

new rodent recipient, the human chromosome content can often be reduced (Lugo et al 1987 and Chapter 1).

The derivation of a monochromosomal hybrid achieves the first step in the process of directedly isolating a single human chromosomal region. The next step in reducing the complexity or size of the target chromosome map region is to transfer large fragments of human chromosomes into rodent cells. Chromosomes may be broken with irradiation (X and gamma ray), alkylating agents or physical shearing. Goss and Harris demonstrated the feasability of this technique by selectively transferring large fragments generated by gamma irradiation which contained the human HPRT locus into recipient HPRT$^-$ rodent cells(Goss and Harris, 1975). They induced chromosome breaks in the donor cells with lethal doses of gamma rays and then fused these immediately to recipient cells for chromosome transfer. The resulting HAT resistant hybrids were used for fine structure gene mapping of the region surrounding the HPRT locus. They reasoned that if two markers were far apart on the chromosome they would frequently be seperated by a break and therefore not co-transferred, whereas if they were physically close they would rarely be separated by a break and thereby frequently co-transferred. The frequency of co-transfer of two markers could be translated into a physical map distance. Indeed, they derived from these initial experiments the gene order and relative distances between the selected HPRT locus and the unselected, but cotransferred, PGK, alpha-galactosidase, and G6PD gene loci on the human X chromosome. These loci were the first to be mapped relative to one another without genetic linkage information (Goss and Harris, 1977).

Hybrids which have deleted portions of human chromosomes can also be obtained by selecting for loss of a marker. A cell surface antigen (Sa_1), whose genetic determinant resides on human chromosome 11, has been used very successfully in this way. The J1 hybrid, which retains a single human chromosome 11 to complement the gly$^-$ mutation of its CHO parent cell, was treated with low doses of chromosome breaking

agents. Mutagenized cells were then incubated with Sa_1-specific antisera plus cytotoxic complement. Only cells which had deleted the Sa_1 locus on chromosome 11 were able to survive complement mediated cell killing; however,to avoid loss of the entire chromosome , survivors were cultured in gly⁻ media. A series of hybrids carrying a range of deletions of 11p surrounding the Sa_1 antigen-determining gene were derived in this way from J1 (Kao et al 1977; Jones and Kao, 1978). These hybrids, plus others generated with other surface antigen genes on 11p, have been used to derive a fine structure map of this region. In particular, a region containing the susceptibility to Wilm's tumor gene has been narrowly defined by such physical mapping (Glaser 1986).

Hybrids containing segments of human chromosomes have also been made by metaphase chromosome transfer (McBride and Ozer, 1973; Miller and Ruddle, 1978; Klobutcher and Ruddle 1979; Olsen et al 1981; Pritchard and Goodfellow, 1987, 1986; Lugo and Baker, 1983). This method is a derivative of the $CaPO_4$ mediated gene transfer method. Preparations of metaphase chromosomes replace purified DNA in the $CaPO_4$ precipitate which is overlayed on recipient cells. Crude preparations of highly condensed metaphase chromosomes are co-precipitated with Ca^{++} and PO_4 onto recipient cells. The method is otherwise exactly as for DNA mediated gene transfer. However, the frequency of chromosome transfer is orders of magnitude lower than gene transfer. Although the $CaPO_4$precipitate probably contains whole chromosomes, these are never transferred intact. because the cellular mechanisms for receiving and incorporating exogenous DNA result in their fragmentation. By capitalizing on these recombination processes, interspecies hybrids are created by the incorporation of large fragments of human chromosomes into the transferent genome.

Although metaphase chromosome transfer may produce more rearrangements than the gamma irradiation technique of Goss and Harris, fine structure mapping has been elegantly procured using these hybrids as well. For example,Weiss et al (1984) used surface antigens to select <u>for</u> a chromosome fragment of interest - the H-2

histocompatability gene region in the mouse. Mouse chromosomes previously marked with the dominant selectable marker, neo, were used as donors in metaphase chromosome transfers to recipient hamster cells. Transferents which had received the neo selectable marker were then incubated with H-2 specific antisera and secondary antibodies tagged with fluorescent reporter molecules. Transferents which had also received the H-2 marker bound H-2 specific antisera and were thereby selectively sorted by FACS analysis (fluoresence activated cell sorter) A series of these transferents were analyzed for the presence of neo and for each of the cloned H-2 family of genes. From the frequencies with which each H-2 gene was co-transferred into transferents with the introduced neo marker, it was possible to derive a map ordering these genes relative to the neo marker and to each other on mouse chromosome 17.

The Weiss experiments are another example of the elegant use of surface antigens in somatic cell genetics. Their genetic determinants can be selected for or against. Human surface antigens are often very species specific which make them useful for selecting interspecies hybrids against a background of rodent parental cells. If the antisera raised against the antigen has a sufficiently high avidity and expression of the antigen on the cell surface is sufficiently ubiquitous, the antigen can be used to select by FACS analysis cells which represent a very small percentage of a mass population. Many surface antigens have now been identified with monoclonal antibodies and their respective genes mapped to specific human chromsomes (Rettig et al 1984), making these a promising tool for future somatic cell genetics.

Thus, several methods exist for selectively retaining portions of human chromosomes in interspecies hybrids. However, all of the techniques outlined above for transferring intact or large segments of chromosomes are limited by the availability of transferrable, selectable markers. Endogenous markers such as HPRT or ade C require the development of parallel mutant recipient cell lines and there are too few such markers which are amenable to selection to be able to transfer and study most regions of

the genome. Surface antigen markers are highly useful as outlined above, but again their genetic determinants are in limited regions of the human genome. Thus the development of vectors which could efficiently introduce dominant selectable markers into recipient cells (Mulligan and Berg 1980; Mann et al 1985) was quickly exploited for use in metaphase chromosome transfer (Weis et al 1984; Nelson et al 1984; Pritchard and Goodfellow, 1986) and for microcell mediated chromosome transfer (Saxon et al 1985; Athwal et al 1985; Lugo et al 1987 and Chapter I).

The first selectable markers were carried by plasmids which contained the selectable marker gene adjacent to SV40 eukaryotic promoter sequences (Mulligan and Berg,1980; Southern and Berg 1982). In the first uses of these markers for somatic cell genetics, these plasmids were inserted into human chromosomes by transfecting human cells which could subsequently be used as chromosome donors in the formation of hybrids. Although this protocol did produce the first hybrids containing introduced selectable markers, it was limited by the low transfection efficiency of most donor human cell lines. In addition, analysis of the integration site in hybrids was complicated by the frequent introduction of tandem, scrambled copies of the plasmid. Moreover, the recombination mechanism which permits integration of transfected DNA into host chromosomes promotes chromosome breakage and rearrangement (Capecchi et al 1980; Wigler et al, 1978; Calos et al 1983; Wake et al 1984; Miller and Temin 1983 and Nelson, D.L., PhD thesis). Remedies to these problems were found in the construction of murine retroviral vectors which could introduce selectable markers via the native retroviral infectious lifecycle.

Retroviral vectors circumvent the limitations of transfection as the method for introducing a marker by instead expliting the efficient infectious entry of retroviruses into their host cells. Infection is efficient in part because penetration of target cells is mediated by host cell surface receptors. Specificity for and binding to receptors is promoted by the presence of viral gp70 (envelope) protein on virion membranes.

gp70 also determines the tropism of these viruses as ecotropic (infect only mouse cells), xenotropic (infect only non-mouse cells), and amphotropic (infect cells from a wide variety of species including human). The only mammalian cells which cannot be infected by murine retorviruses are from hamster - a fact which has obvious relevence for somatic cell genetics as will be noted in Chapter 1.

Experiments which demonstrated this determinance of gp70 utilized virions which had core proteins synthesized by viruses of one tropism packaged in membranes containing gp70 of another tropism (Rein 1982). As described below, such 'pseudotyping' is regularly used in the manipulation of engineered retroviral vectors to control their host range . Pseudotyping experiments also revealed that after infection, retroviruses direct the synthesis of gp70 which is processed to the cell membrane where it occupied its own specific cell surface receptors. Superinfection by secondary viruses of the same tropism was thus blocked. If the host cell has two types of receptors, eg. ecotropic and amphotropic, superinfection by virions of the second tropism was still possible. However, since it was sometimes desirable to be able to introduce several retroviral vectors carrying dominant selectable markers into cells having only one type of receptor, this presented an experimental problem which was addressed in the development of later generations of retroviral vectors (see below). Clearly retroviral vectors have become one of the most useful tools of molecular and somatic cell genetics. The interaction of these virus particles with their host receptors is a critical limiting step in experimental manipulations of these gene transfer vehicles and so in Chapter 6 some initial experiments to study the host cell receptors and their genetic determinants are described.

Early generations of engineered retrovirus vectors (MSV-neo and MSV-gpt) were constructions derived from cloned proviral Moloney sarcoma virus (MSV) sequences. Virtually all of the provirus was left intact and the selectable markers neo or gpt (Mulligan and Berg, 1980; Southern and Berg, 1982) were simple additions to the

native provirus genome. The plasmid construction was transfected into virus producing cells where it integrated into host chromosomes. Expression of the integrated MSV(gpt) or (neo) genome was promoted by sequences present in the long terminal repeat (LTR) sequence. Spliced messenger RNAs which directed the production of all necessary viral proteins plus full length mRNA transcripts of the entire viral genome provided all the components which were necessary to produce infectious viral particles containing the engineered viral construct. Virus particles shed into the culture medium of transfected cells were harvested and used to infect other cells; transmission of the selectable marker by infection was demonstrated by the aquisition of resistance to the antibiotic analogue G418 in the case of MSV(neo) or by the ability to grow in HAT medium in the case of MSV(gpt). Titers of 10^6 virus particles per ml of culture media were routinely achieved by this protocol, a characteristic which makes these vectors attractive gene transfer vehicles.

However, as David Nelson (PhD thesis, 1984) soon discovered, the drawback to the MSV(neo) constructs for chromosome transfer experiments was that the infected human cells which were used as chromosome donors were also producing infectious virus. During chromosome transfer, then, recipient cells could acquire drug resistance either by receiving a chromosome containing the integrated marker (desirable) or by infection (undesirable). The production of infectious virus by human cells also presented an obvious laboratory safety hazard.

Second generation vectors developed by Mann et al (1985) circumvented this problem by deleting 351 base pairs from a cloned proviral Moloney murine leukemia virus (M-MuLV). Such vectors, when transfected into NIH 3T3 cells, were shown to be defective in virus production due to the lack of a sequence (ψ) required for packaging viral RNA into infectious particles. However, when this vector was co-transfected with another construct (pMSV-gpt) containing the ψ sequence but lacking all other essential viral genes, the ψ^- vector could provide all the gene products in trans which were

necessary for MSV-gpt virus production. Supernatants from co-transfected cells thus produced infectious particles containing MSV-gpt RNAs but not RNAs from the ψ^- provirus. Most importantly, when these supernatents were used to infect new cells, the defective MSV-gpt viruses could integrate into host chromosomes but could not direct the production of new virus. These results indicated that any retroviral construct containing only the ψ sequence and the LTR sequences required for full length mRNA transcription could be transmitted as a defective virus after being transfected into NIH cells containing the ψ^- vector (ψ-2 cell line). A second NIH-derived cell line carrying a ψ^- vector which encodes amphotropic gp70 protein was also developed. The use of both these virus 'packaging lines' to generate ecotropic and amphotropic virus vectors is described in Chapters 1 and 5 .

Finally, third generation vectors developed by Cepko et al (1984) addressed the problem of how to identify and analyze the integration site of a retroviral vector. By including the origin of replication sequences from SV40 virus (SV40-ori) and similar sequences from pBR322 plasmid in their SVX(neo) constructs they were able to specifically retrieve and clone the integrated provirus plus host flanking sequences. SVX infected cells were fused with CV-1 cells carrying a defective SV40 virus which constitutively expresses SV40 T-antigen but which lacks the SV40-ori (COS 7 cells). In fused cells T-antigen repeatedly induced replication from the SV40-ori in integrated copies of SVX, thereby producing an 'onion skin' of many newly replicated strands (Conrad et al 1982). By homologous recombination, usually between the two SVX LTRs, but sometimes between host flanking sequences, circular molecules were produced. These plasmids were harvested by the method of Hirt (1967) and used to transform E.Coli which were then selected by growth on kanamycin (analogue of neomycin). Although this method for rescuing proviruses is simple and efficient, it is difficult to obtain sufficient lengths of flanking sequence which reliably contain single copy sequences suitable for mapping. We illustrate this in Chapter 3, and in Chapter 4

describe the development of improved vectors for mapping integration sites.

In principle, a dominant selectable marker can be introduced into any region of any genome of interest, thereby making that region selectively transferable. The co-transfer of the surrounding region may provide fine structure mapping information (Weis et al 1984) or provide the reagent for a library of molecular clones which, relative to a total human genomic library is many times enriched for markers from the specific transferred region (Chapter 3). Directed and selective transfer of a whole chromosomal region may also be useful for studying the expression of transferred genes which are too large to be transferred intact by standard $CaPO_4$ gene transfer techniques and/or which must remain under the control mechanisms determined by their native chromosome/chromatin structure (Gros et at, 1986).

Which regions of the human genome are of interest then? Certainly, just as the study of human genetics began with the hypothesis by Garrod (1902) that phenylketonuria was a metabolic disease which was inherited, rather than aquired, so still today human genetic disease remains a major focus for the discipline. As predicted by Botstein et al (1980) and reviewed by Wexler et al (1982), the ability to generate large numbers of single copy markers from randomly isolated human DNA sequences (Gusella et al 1980, Wyman and White 1980; Nakamura et al 1987) has made it possible to map a number of human single gene defects to a specific chromosome via family studies. Huntington's disease (Gusella et al, 1983); polycystic kidney disease (Reeders et al 1985); cystic fibrosis (Knowlton et al 1985, White et al l985,) Alzheimer's disease (St. George-Hyslop et al, 1987 ; Tanzi et al, 1987,); manic depressive disease (Egeland et al 1987) and neurofibromatosis (Barker et al, 1987; Seizinger et al 1987) have all been mapped by this method within the last four years. The rate limiting technical step toward mapping these loci has been the difficulty in identifying enough sufficiently polymorphic loci for use in linkage analysis. However, Nakamura et al have recently demonstrated that new, highly polymorphic loci can be

readily isolated from phage or cosmid libraries by screening with repeat sequences which are known to be the basis for complex polymorphisms at other loci (Nakamura 1987). Therefore, the localization of disease genes can be expected to accelerate in the near future although the most difficult first steps of identifying, characterizing and sampling multi-generational families which are transmitting the disease remain the most time consuming. Once a disease gene has been mapped to a specific chromosome, then all the techniques for focusing on a specific region outlined above become useful for identifying additional closely linked markers and for finding the disease gene itself. Of course, since few of the disease loci and their linked markers provide selectable markers, the ability to insert a dominant selectable marker in the general region or to fortuitously associate the region with a selectable marker by translocation (Wasmuth et al 1986) becomes extremely advantageous.

We have attempted to recapitulate the historical progress of somatic cell genetics in a contemporary time scale on a model system . For a model, we chose to contribute to the large collaborative effort directed at cloning the Huntington's disease (HD) gene by attempting to insert a dominant selectable marker in the region surrounding the HD gene; this would, in turn, enable us to selectively isolate this region in a somatic cell hybrid. We chose this model disease locus because mapping and cloning the HD gene presented challenges which could be particularly well addressed with somatic cell genetics.

Huntington's disease (HD) was so named for George Huntington who gave such a thorough and accurate description of this disease in 1872 that, although the condition had been previously described, this disease of the central nervous system has been associated with him since. The disease is characterized by the appearance of progressive choreic movements and dementia, usually late in life. The choreic movements in HD are characteristically "aimless, forceful and complicated" movements (Bell 1934), which include facial grimacing and twisting or jerking of the trunk and

limbs. These increase in severity and rigidity with age, often appearing truly grotesque in final stages. The neuropathological, pharmacological, biochemical and physiological bases of these symptoms remain largely obscure so that HD can still only best be described as a relentless, "genetically programmed cell death in the human central nervous system". (Martin 1982).

Although there is considerable variability in the age of onset of the disease, making determination of gene carriers difficult, an autosomal dominant mode of inheritance has been clearly defined for HD (Reed and Chandler, 1958; Elston et al 1980). One of the most tragic aspects of the disease is that because the age of onset is typically between 35 and 45 years of age, gene carriers are not detected until after childbearing, thus formally placing all of their offspring at 50% risk for HD. There is no sexual bias in the expression of the disease; however, there is a significant prevalence of juvenile age-of-onset cases among offspring of male gene carriers (Farrer et al 1984; Merritt et al,1969; Jones and PHillips 1970; Myers et al 1983). Indeed, the mean age of onset in all offspring of affected males is 3.5 years earlier than in offspring of affected females, suggesting that this phenomenon is not even confined to the juvenile form of HD. Hypotheses suggesting some maternal protective factor present either in the cytoplasm of the fertilized egg which has lasting effects on the developing fetus or in a maternally transmitted mitochondrial gene are the most likely explanations for this observation, but no experimental evidence exists to support either.

The search for the Huntington's disease gene has become the paradigm for identifying human disease genes of unknown biochemical or physiological function. In the late 1970's, after the development of strategies for selectively isolating random human clones from genomic DNA libraries by Gusella et al (1980), the idea arose that it should be possible to identify sufficient numbers of random genetic markers using this strategy to be able to establish genetic linkage between one such marker and a well defined genetic disease like HD. There were three requirements for the feasabiliy of

such a project: (1) the disease must be transmitted as a single gene defect with a clearly defined clinical phenotype; (2) a large pedigree including multiple generations of multiple affected individuals must be well characterized and sampled; and (3) a protocol for isolating random DNA markers which demonstrate sufficient polymorphism to be useful in linkage analysis must be available. Upon the rediscovery in 1982 of an enormous pedigree in Venezuela including 7,000 individuals among whom over 100 were living symptomatic individuals and almost 900 of whom were at 50% risk for HD, all three of these requirements were met and a presumed long search for a marker linked to HD was underway. However it was among the first 12 random markers which were tested in 1983 for linkage to HD that the G8 marker was found to be linked at a recombination distance of 10 cM with a cumulative LOD score of 8.53. The G8 marker could be assigned to human chromosome 4 by somatic cell hybrid mapping panel analysis and so, by inference, could the HD locus.

Rapid initial success has been followed by less serendipitous, yet steady, progress toward identifying the HD gene. A large collaborative project was conceived in 1983 and has been working since to share technologies and labor to progress from a single linked marker and a chromosome assignment to cloning the disease gene. The first goal was to identify bracketing markers which would define a limited region of chromosome 4 to be searched for the disease gene. The experiments described in this thesis are part of that collaborative agreement project and the results described below have been reported at yearly meetings of the participants. Most of the results are currently in press or in preparation for press and are reported here by agreement of the participants.

The most pertinent data for the development of this thesis was the chromosome localization by in situ hybridization of the G8 (D4S10) probe to 4p16 (Zabel et al 1986; Landegnet et al, 1986; Magenis et al 1986 and Wang et al 1986). These in situ results presented conflicting subband assignments to either 4p16.1 or 4p16.3, but

these have recently been resolved through the careful analysis of a series of somatic cell hybrids containing chromosomes 4 with terminal and interstitial deletions of 4p (McDonald, in press). These latter results place D4S10 in 4p16.3, the most terminal subband of chromosome 4.

Concomitantly, several hundred random clones have been isolated from a chromosome 4 flow sorted library (Gilliam et al 1987a) and more recently from a somatic cell hybrid carrying a 4p/5p translocation chromosome (Wasmuth et al1986). The goal has been to identify a marker which maps to the opposite side of the HD locus from D4S10 - a so called flanking marker. To this end, these clones have been mapped to regions of chromosme 4 and some, which are polymorphic, have been tested for linkage with HD. Only two clones from these sources have been shown to be located in 4p15-16 or to be linked to HD. The closer probe, called C4H, is a 3 kb fragment which has been limitedly informative due to infrequent polymorphisms detected by the probe, resulting in only 15% average heterozygosity in a sampled population. At 4 cM, C4H is closer to HD than D4S10 and, based on the observation of one crossover event between D4S10 and HD which did not also include recombination between HD and C4H, C4H may be a flanking marker. However, more rigorous evidence from a larger number of informative meioses will be required to make this definitive statement; such data are not yet available.

A second marker, D4S62, is an anonymous DNA locus detected by the monkey probe p8, a unique DNA sequence found in association with satellite DNA (Thayer, 1987). Polymorphisms detected by this probe and a third probe, RAF2 (Bonner et al 1985) which represents a processed psuedogene of the raf proto-oncogene, have provided sufficient linkage data for establishing their linkage relationships to HD and D4S10. The gene order relative to the centromere is RAF2-D4S62-D4S10-HD-telomere. (Gilliam et al 1987b).

Thus, four years have been spent to establish finally that the HD gene resides

near the telomere of chromosome 4. Very few DNA clones have been isolated from this region despite the retrieval of several hundred clones from several chromosome 4-specific libaries; in particular, no flanking marker has yet been identified. The regions surrounding D4S10 and C4H are now being avidly mapped by pulsed field gel electrophoresis in the hope of mapping the HD gene in the process. However, questions remain as to how to recognize the HD gene and how to study its function in vitro if and when it is cloned.

We reasoned that the introduction of a dominant selectable marker in the HD region of chromosome 4 would create two important opportunities for mapping and studying the HD gene. First, hybrids could be constructed containing only the introduced marker and small fragments surrounding it; libraries made from such hybrids would be highly enriched sources of clones linked to and including the HD gene. Secondly, by association with the selectable marker, the entire HD gene including any near or distant control elements could be selectively transferred into cultured neuronal cells to study its in vitro effects.

These particular challenges are not presented by other human genetic diseases which have been mapped by exploiting their cytogenetic or biochemical abnormalities. Duchenne Muscular Dystrophy, for example, is an X-linked disease affecting about 1 in 3000 males. In rare cases, daughters of carrier females may also inherit the disease. Indeed, it was by analyzing these rare patients that the DMD gene was first mapped to Xp21. Cytogenetic analysis revealed the presence of an X/autosome tranlocation in the peripheral lymphocytes of these patients which led Murray et al (1982) to hypothesize that if the DMD gene resided at the X/autosome breakpoint, then the translocation breakpoint would have disrupted this second copy of the DMD gene, thereby causing these females to be affected by the disease. The translocation thus mapped the DMD gene unequivocally to the Xp21 breakpoint and provided a focus for future attempts to clone the gene or closely linked markers.

For example, Worton and colleagues noted that the X/autosome breakpoints also interrupted the ribosomal gene cluster on chromosome 21 (Worton et , 1984; Ray et al 1985). Because this region had previously been cloned and fine structure mapped, they reasoned that it would be possible to molecularly "walk" across the breakpoint into the DMD gene starting from a cloned ribosomal gene. By isolating a contiguous set of cosmid clones they were, in fact, successful at identifying a cosmid containing the translocation junction between chromosomes 21 and the X. The fragment from the X chromosome (XJ) was subsequently shown to be part of a DMD candidate gene.

A concurrent approach taken by Kunkel et al (1985) also exploited a cytogenetic defect. The X chromosome of a rare patient affected with several diseases - DMD, chronic granulomatous disease, retinitis pigmentosa and McLeod syndrome - contained an interstitial deletion in Xp21. Kunkel et al developed a novel method for dramatically and selectively enriching a phage library for sequences missing in the deletion X chromosome and were successful in isolating a fragment called pERT87 which detected deletions in about 5% of classical DMD patients. This clone and the Worton clone, XJ, also show tight linkage to DMD in family studies (Monaco et al 1985). Further detailed mapping and sequencing around these initial clones has defined a putative DMD gene which covers more than 1000 kb of genomic DNA and codes for a 16 kb message (Monaco et al1986 and Doenig et al 1987).

DMD is a dramatic example of how chromosomal deletions and translocations can be used to directly target the search for a disease gene. Similarly, the loci determining inherited susceptibility to the childhood cancers, retinoblastoma (Cavanee 1984) and Wilm's tumor (Fearon et al 1984; Glaser et al 1986), were also mapped by their association with cytologically detectable deletions. Creative use of these physical genetic markers can result in direct cloning of their associated disease genes; a candidate for the retinoblastoma susceptibility gene has already been identified on the basis of its chromosomal location and homozygous deletion in tumor DNA (Friend et al

1986 and Lee et al 1987).

The biochemical bases for some inherited diseases such as sickle cell anemia, the alpha and beta-thalassemias, Lesch-Nyhan syndrome, analbuminemia (persistance of alpha-fetoprotein), familial hypercholesterolemia, the collagen disorders, phenylketonuria (PKU), ornithine transcarbamylase deficiency and other metabolic disorders have been known for decades. Detailed information about the biochemical pathways and specific protein products involved in the etiology of these diseases now provides the tools for designing schemes to directly isolate the relevant genes or cDNA representatives. For example, Brown and Goldstein's classic study of the feedback inhibition pathway for cholesterol synthesis was the basis for identifying and cloning the gene causing familial hypercholesterolemia (FH). By studying the overproduction of cholesterol in cells taken from patients with FH, they were able to determine the normal negative feedback control mechanism for cholesterol synthesis. (Goldstein and Brown, 1973) Their observation that these FH mutant cells lacked cell surface receptors which bind feedback control molecules (LDL) established the LDL receptor gene as a clear candidate for being the FH disease gene. The gene coding for LDL receptors has since been cloned and polymorphisms flanking the gene have demonstrated complete co-transmission with FH (reviewed in Goldstein et al, 1985).

In the absence of either a cytogenetic or biochemical marker around which to design a scheme for mapping and cloning a disease gene, systematic and efficient methods for narrowing the chromosomal region which must be searched become highly desirable. The experiments described in this thesis demonstrate the challenge, yet feasibiliy, of selectively isolating a chromosome region which contains a human disease gene, given only one linked genetic marker as a starting point for that goal.

GENERAL MATERIALS AND METHODS

Preparation of plasmid, phage and genomic DNAs

Mini-preps of plasmid DNAs were prepared from 1.5 ml of an overnight culture by resuspending pelleted bacteria in 0.5 ml resuspension buffer (50 mM TRIS, pH 7.5; 10 mM EDTA; 15% sucrose; and 1 mg/ml lysozyme added just before use) and pre-incubating at room temperature for 5 min. 5 μl of diethylpyrocarbonate (DEPC) were added, followed by vortexing and then 10 μl of 10% SDS was added, again followed by vortexing. Finally, 50 μl of 5M potassium acetate (pH 5) was added, and the mixed solutions were incubated on ice for 30 minutes. Plasmid DNA was recovered by centrifugating samples in an Eppendorf centrifuge for 25 minutes at 4° C to pellet debris and precipitating DNA from the supernatents with 0.7 ml of absolute ethanol at room temperature. DNA was pelleted by centrifugation in Eppendorf centrifuge for 10 minutes at room temperature and the pellets were resuspended in 20-35 μl TE (10mM TRIS, pH 7.5; 1mM EDTA). Before use in restriction digests, samples were treated with RNAse at 20 μg/ml final concentration.

Large quantities of plasmid DNA were isolated from 1 liter cultures (in Luria Broth) by resuspending in 9 ml total volume of 25% sucrose, 50mM TRIS (pH 7.5) plus 1 ml of 10 mg/ml lysozyme (in 0.25 M TRIS, pH 7.5). The resuspended bacteria were left on ice for 5 minutes. 1.8 ml 0.5 M EDTA was added and incubation on ice was continued for 5 more minutes. 14.5 ml Triton X100 solution (0.2% Triton X100; 100 mM EDTA; 50 mM TRIS, pH 7.5) was added and lysis left to occur on ice for 15 minutes. Cell debris was removed by centrifugation in Oakridge tubes in a Sorvall SS34 rotor for 40 minutes at 18,000 rpm at 4° C. Plasmid DNA was purified directly by making the supernatent up to 32 ml with TE and adding 30.24 g CsCl plus 3 ml of 10 mg/ml

ethidium bromide. Plasmid DNA bands were pulled after centrifugation in a Beckman VTi 50 rotor for >20 hours at 45,000 rpm.

Genomic DNA was isolated from cultured fibroblasts by removing the culture media from confluent 15 cm tissue culture plates and replacing rinsed cells with 10 ml lysis buffer (10mM TRIS, pH8.0; 10mM EDTA, 10mM NaCl, 0.5% SDS, 50 μg/ml proteinase K added just before use) and incubated at 37° C overnight. After scraping the lysed cells into a 50 ml conical tube (Falcon polypropylene), the proteins were removed by extraction one time with phenol, followed by phenol/chloroform and then chloroform. These extractions were performed by slow rotation for several hours each. High molecular weight DNA was retrieved by precipitation with 0.6 volumes isopropanol and spooling out strands.

Phage DNA was prepared by inoculating 1 liter of NZYM media (10 gr NZ amines; 5 gr Difco yeast extract; 5 gr NaCl; 2.5 gr $MgSO_4$ per liter) with 0.3 ml phage plate lysate pre-absorbed to 0.3 ml overnight culture of LE392 grown in 0.2% maltose. The cultures were incubated with vigorous shaking at 37° C for 8-10 hours until complete lysis was observed. 3 ml of chloroform was then added and the culture incubated with shaking a further 30 minutes. Bacterial cell debris was removed by centrifugation in a Sorval GSA rotor at 9k rpm for 20 minutes and the supernatents were filtered through cheesecloth into 60 g NaCl and 70 g polyethylene glycol (MW 8000) per liter. The PEG solutions were stirred slowly at 4° C overnight and the phage/PEG complexes collected by centrifugation in a Sorval GSA rotor for 45 minutes at 4° C at 9000 rpm. The pellets were resuspended in 10 ml SM (see Maniatis Molecular Cloning Manual) plus 7.5 g CsCl and phage particles were banded in quickseal tubes by centrifugation in Beckman Type 65 rotor at 45k rpm, 20° C for > 20 hours. DNA was purified from particles after dialyzing against SM by extracting 2X with phenol and 2X with chloroform followed by

precipitation with ethanol. High molecular weight DNA was hooked out of ethanol, rinsed in 70% ethanol and dissolved in TE.

Agarose gel electrophoresis, Southern blotting and hybridizations

10 μg of genomic DNA was digested with restriction enzymes as per manufacturers directions, precipitated with ethanol and resuspended in 10-30 ul TE. Samples were size separated by electrophoresis through 1% agarose (SeaKem) in 1X TAE (as per Maniatis MCM) for 12-18 hours at 20 volts. Southern blotting was performed as described (Southern 1975) after transfer to Zetabind nylon membranes. Hybridizations were performed in 50% formamide, 1M NaCl, 1X Denhardt's (see Maniatis MCM), 0.05% SDS, 0.5M TRIS (pH 7.5), and 4% dextran sulphate at 42° C for 36 hours. Stringent washes were performed at 58-60° C in 1X SSC (20X SSC, see MCM). Autoradiography was done using Kodak XAR X-ray film with Lightening Plus intensifying screens at $^{-}$70° C.

Radiolabeled DNA for use as probes were prepared by first isolating specific sequences from plasmid or phage vectors via restriction digestion followed by gel electrophoresis and purifying the relevant fragments out of agarose by the NaI glass bead procedure (Vogelstein and Gillespie 1979). ^{32}P-dCTP (New England Nuclear #013 H)-labeled probe was then prepared by the random primer method using the large Klenow fragment of polymerase I (Feinberg and Vogelstein 1983). Unincorporated nucleotides were removed by spin column dialysis over a 1 ml column of Sepharose-CL/6B equilibrated in 1X STE (10X STE = 10 mM TRIS, pH 8.0; 100mM NaCl; 1 mM EDTA).

INTRODUCTION
CHAPTER 1

Weis et al (1984) demonstrated that it was possible to directly isolate a sub-chromosomal region containing a dominant selectable marker if the region already contained another selectable marker such as the murine surface antigen H-2. A pooled population of neo-infected mouse cells were used as donors in metaphase chromosome transfers to hamster recipient cells. The transferents were selected in G-418 and drug resistant transferents were then sorted for H-2 antigen expression by FACS analysis. Those rare transferent cells which were both G418 resistant and antigen positive were further analyzed to demonstrate that both markers had been co-transferred on one contiguous fragment - i.e. that the neo marker had integrated near to the H-2 region.

This direct protocol which selects for the co-transfer of two syntenic markers (one endogenous and one introduced) is limited by the lack of markers like H-2 in all target chromosomal regions. For example, there is no such known marker in the region of the Huntington's Disease gene on human chromosome 4 . Therefore, we chose to attempt to insert a dominant selectable marker in the HD region by a several step process; first we would attempt to insert a dominant selectable marker into any region of chromosome 4 and use the marker to select a microcell hybrid containing only that one human chromosome. This chapter describes the cell hybridization experiments which created this and other hybrids as well as the experiments which were done to characterize each of these cell lines.

Dominant selectable markers can be introduced into cultured cells by the mechanical methods of $CaPO_4$-mediated gene transfer (Graham and van der Eb 1973), microinjection (Mueller et al 1978) or by the biological method of retrovirus infection (Mann et al 1983). Selectable markers have been successfully introduced into human

chromosomes by transfecting human cells with plasmids containing gpt or neo markers and these marked chromosomes have then been transferred via microcell mediated chromosome transfer to create monochromosomal hybrids (Saxon 1985; Athwal 1985a and b, Schultz 1987 and Tunnacliffe 1983). However, insertion of a marker by retroviral infection is the method of choice for three reasons: 1) many human diploid cell lines are inefficiently transformed by any of the mechanical methods of gene transfer; 2) transformants often contain more than one copy of the plasmid; and 3) integration of plasmids may induce chromosome breaks. Therefore, the population of marked human chromosomes which were the beginning reagent for this project were established by infecting diploid human fibroblasts with amphotropic SV(X)-neo retroviruses.

Microcell mediated chromosome transfer of these marked human chromosomes into mouse NIH 3T6 cells produced a large population of hybrid G418^R clones, of which only 1 in 25-50 would be expected to contain a chromosome 4. Searching for those hybrids of interest presents a challenge. If there were a surface antigen anywhere on the chromosme 4 and if it were expressed in the hybrid cell but not in the recipient parent cell, then the presence of the antigen could be used to differentiate hybrids containing that chromosome from those that did not. Cell surface antigens coded for by genes on chromosomes 11 (Jones and Kao 1978) and Y (Pritchard and Goodfellow1986) have been elegantly used to isolate hybrids containing these chromosomes. There is one surface antigen which has been mapped to chromosome 4 but this marker is expressed only in lymphocytes (Peters et al 1984) and is therefore unsuitable for selecting fibroblast hybrids.

Hybrids containing specific chromosomes could also be selected by complementation of an auxotrophic mutation in the rodent fusion partner with a human gene on the chromosome of interest. However, as discussed previously, many human chromosomes cannot be selected this way in the absence of a sufficient number and

variety of mutant rodent partners. The ade C$^-$ CHO mutant can be complemented by a gene on human chromosome 4, but initial experiments using this cell line as a fusion partner to select for chromosome 4 revealed an unexpected and problematic characteristic: these CHO cells spontaneously break exogenous chromosomes and retain only very small, non-contiguous and unrelated fragments. (McDonald, M. unpublished)

Since it was not possible to utilize either of the more elegant methods of complementation or surface antigen selection, the methods which remain for characterizing the chromosome content of interspecies hybrids are isozyme analysis, DNA marker analysis and karyotyping. Isozymes are variant forms of enzymes; the variants can be detected by increased or decreased activity, by thermostability, by antigenicity or most commonly by electrophoretic mobility and isoelectric point. Isoforms which have been separated in starch or acrylamide gels by electrophoresis or isoelectric point can be stained in situ for activity. Fortunately, many enzymes have different isoforms in rodent vs human cells so that the human form can be discriminated in a hybrid cell. Multiple isozymes have been mapped to each of the human chromosomes so that there is a convenient marker for any chromosome of interest. Many cell lysate samples can be loaded on one electrophoresis gel so that a large number of hybrid cells can be screened for each isozyme quite easily. Although the presence of one marker from a particular chromosome does not verify the presence of the entire chromosome, this is a satisfactory initial screen for chromosomes of interest.

The number of isozyme markers which had been developed and mapped to human chromosomes from the 1950's to 1977 was quickly surpassed in the explosion of technology which permitted the rapid isolation, characterization and mapping of cloned genes and random DNA fragments. The human genome is being saturated with DNA markers distributed every few centimorgans so that a battery of markers are now available for detecting any chromosome. If the human marker cross-hybridizes with rodent sequences in a DNA sample from a hybrid dell, these are usually on different size

restriction fragments and can therefore be distinguished. Indeed, when the human fragment is being scored as absent, the presence of the endogenous rodent fragment acts as an internal positive control for hybridization. We have used a variety of DNA markers to verify the chromosome content of the hybrids described in this chapter.

Perhaps more useful than DNA markers which detect single copy sequences are those which detect a related family of sequences. For example, the arginosuccinate synthetase gene is part of a family of genes, both functional and non-functional (pseudo) copies. Its cDNA hybridizes to 18 Eco RI fragments in human genomic DNA which correspond to a minimum of 11 argininosuccinate synthetase-related sequences in the human genome. Fifteen Eco RI fragments have been assigned to single human chromosomes and three fragments to one of two human chromosomes (Beaudet 1982). Thus, the presence or absence of particular AS-hybridizing fragments in genomic DNA from somatic cell hybrids indicates the presence or absence of 10 distinct human chromosomes. Thus with one hybridization experiment 10 chromosomes can be either included or excluded from the list of possible chromosomes in a new hybrid. We used the pAS probe for this purpose to verify the chromosome content of the microcell hybrids generated in this project. When the multiple fragments detected by the cDNA for actin (Soriano et al 1982) or other representatives of gene families are similarly individually mapped it will be possible to characterize hybrids with just a few such probes.

An equally important method for characterizing hybrids is by visualizing the chromosomes directly. Human chromosomes can be quickly distinguished from rodent chromosomes by G-11 staining which stains human chromosomes blue and rodent pink (Friend et al 1976). However, this technique does not identify which human chromosomes in particular are present. Metaphase or interphase chromosomes can be individually identified after staining by quinacrine mustard (Caspersson et al 1970), Giemsa (Sumner et al 1971, Patil et al 1971, Drets and Shaw1971) and modifications

using trypsin (Seabright 1971; Wang and Fedoroff 1972), or chromomycin counterstained with distamycin (Donlon et al1983). Careful karyotyping defines quite accurately which intact human chromosomes are present and also detects gross rearrangements. The R-band staining method of Donlon (1983) was used to characterize the hybrids described in this chapter.

Combined isozyme and karyological analyses provide a relatively complete description of the human genetic content in interspecies hybrids. However, the process of cell hybridization itself and the subsequent selection for growth in culture probably promote chromosome breakage and recombination resulting in human chromosome fragments which may be unidentifiable by isozyme or cytogenetic analysis. Small fragments which have become associated through translocation with a host rodent chromosome are particularly difficult to detect even with the species discrimination visualized with G-11 staining. However, in situ hybridization with either radiolabeled or non-isotopically tagged DNA fragments containing human reiterated sequences such as the Alu family is a very sensitive method for specifically visualizing human chromosome fragments. In this chapter we show an example of a $G418^R$ microcell hybrid which contains several intact human chromosomes, none of which carry the neo marker. By in situ hybridization we were able to visualize an additional fragment of an unknown human chromosome which had become translocated onto a mouse chromosome. This fragment presumably carries the neo marker. We have also used this technique to identify a small fragment of chromosome 11 in a CHO hybrid which has been used to do fine structure mapping of this region. (Gerhard et al 1987)

MATERIALS AND METHODS
CHAPTER I

Cell lines and Culture Conditions

Primary diploid human fibroblasts were obtained from a sample of foreskin and propagated in F/DV medium (a 1:1 mixture of DME and Ham's F12) supplemented with 15% fetal calf serum and 2mM glutamine. 3T6 cells are an established line of moue fibroblasts; these were maintained in F/DV medium plus 10% fetal calf serum.

Preparation of virus and infection of cells

Culture medium containing virus was harvested from ψ-AM producer cells filtered through a 0.45 um filter assembly (Gelman #4184) and incubated with human fibroblasts for two hours at 37° C in the presence of 8 μg/ml polybrene. Infected cell clones were selected by growth in 2 mg/ml G418.

Microcell-mediated chromosome transfer

The infected human fibroblasts were plated onto plastic "bullets" cut from tissue culture dishes and incubated in the presence of 10 μg/ml colcemid for 48 hours. After micronuclei had formed in the cells, the bullets were placed back-to-back in centrifuge tubes containing 5 μg/ml cytochalasin B in serum-free growth medium. They were centrifuged at 18,000 x g for 35 minutes at 28-32° C to enucleate the cells. The resulting preparation of microcells was suspended in a solution of 100 μg/ml phytohemagglutinin P and applied to monolayers of recipient cells growing in 25 cm^2

tissue culture flasks. Following 15 minutes incubation at 37° C to permit agglutination, the cells and microcells were fused by 60 seconds exposure to 44% (w/w) polyethylene glycol (molecular weight 1300-1600). After an overnight incubation in non-selective medium, the recipient cells were plated into medium containing 500 µg/ml G418. Colonies appeared fourteen days later. The cultures included G418-resistant donor fibroblasts that survived the enucleation procedure intact; these could be distinguished by their appearance from the microcells hybrids, which resembled 3T6 cells. Individual hybrid colonies were picked using glass cloning rings and passaged through medium containing 500 µg/ml G418 plus 3 x 10^{-6}M ouabain to eliminate any contaminating human donor cells. Subsequently the hybrids were maintained in 250 µg/ml G418.

Isozyme Analysis

Microcell hybrids were screened electrophoretically for the presence of human isozymes of phosphoglucomutase-2 (PGM-2), esterase-10 (ES-10), nucleoside phosphorylase (NP), mannose phosphate isomerase (MPI), glucose phosphate isomerase (GPI) and adenosine deaminase (ADA). Cell extracts were subjected to electrophoresis on starch gels or cellulose acetate strips, followed by specific staining as described (Ruddle and Creagan, 1975). Filter hybridization (see below) was used to detect human alleles for argininosuccinate synthetase (pAS) (Su et al 1981), adenosine deaminase (ADA) (Orkin et al, 1983), nucleoside phosphorylase (pNP) (Goddard et al, 1983), insulin (Owerbach et al, 1980) and for the arbitrary DNA markers D4S10 (Gusella et al, 1983), 3.6/1.2 (Gilliam et al, 1984), pAW101 (Wyman and White, 1980), p267 (Watkins et al, 1984) and D20S2 (Bruns et al, 1982).

Cytogenetic analyses

Microcell hybrid cultures were semi-synchronized by maintaining cultures at confluence for 24 hours prior to re-plating for chromosome harvest. 36 hours before harvest, cultures in 100 mm dishes were re-plated at 1:4 dilution. 24 hours later half of the culture media was replaced with fresh media containing 10% fetal calf serum. Mitotic nuclei were harvested by incubating semi-synchronized cultures in 20 ng/ml colcemid for 45 minutes at 37° C, followed by trypsinizing, pelleting and resuspending all cells in 3 ml of 0.075 M KCl hypotonic solution. Dissolution of the plasma membranes and differential swelling of mitotic nuclei was achieved by 20 minutes exposure to hypotonic solution at room temperature with occasional gentle resuspension. Nuclei were pelleted and fixed in three changes of room temperature methanol-glacial acetic acid (3:1). Metaphase spreads were prepared by resuspending nuclei pellets in freshly prepared fix and dropping onto wet slides followed by air drying. Chromosome morphology was monitored during slide making by phase contrast microscopy using a Zeiss Phase 2 Planapo 16X objective.

R-banded chromosomes were produced by staining with Chromomycin A3 (0.5 mg/ml in McIlvane's solution, pH 7.0) for 15 minutes followed by counterstaining with Distamycin A (0.1 mg/ml in McIlvane's, pH 7.0) for 4 minutes (Donlon et al 1983). After mounting coverslips with glycerol, metaphase spreads were visualized through a standard fluorescein filter set on a Zeiss Universal Fluorescent microscope using a Planapo 63X objective.

In Situ hybridization

The chromosomal locations of some integrated SVX(neo) retroviruses were mapped by *in situ* hybridization of metaphase spreads using a tritium-labeled neo probe. The 1.4 kb Bam HI/Hind III band containing the neomycin phosphotransferase gene in pBR Neo plasmid was purified out of agarose by the NaI/glass bead technique

(Vogelstein and Gillespie 1979) and used as a template for the Klenow enzyme synthesis of radiolabeled strands (Feinberg and Vogelstein 1983) incorporating H^3-dCTP, H^3-dATP and occasionally H^3-dGTP to a specific activity of 2×10^7 to 1×10^8 cpm/ug. Unincorporated nucleotides were removed by spin-dialysis over a 1 ml 60% Sepharose CL6B/ 40% STE column. Labeled probes were stored at $^-20$ ^{o} C and used for up to 6 months after preparation.

A 300 bp Pst I fragment containing the human repetitive element Blur 11 was used as a probe to detect human chromosome fragments in some microcell hybrids. This fragment, purified from 1% agarose, was labeled to a specific activity of 6.5×10^8 cpm/μg and used as described above.

Slides were aged at least three days but less than one year before using for *in situ* hybridization. Hybridizations, washing and autoradiography were performed as described by Harper (1981). Briefly, RNA was removed from spreads by digestion with 200 μg/ml RNAse at 37° C for 1 hour. Metaphase chromosomes on rinsed and re-dehydrated slides were denatured by immersion in 70% formamide/2X SSC at 70° C for 2 minutes. Denaturation was stopped by immediate immersion in ice cold 70% ethanol and subsequent serial dehydration. Hybridizations were done in 50% formamide/20% dextran sulphate at 42° C for 12-16 hours. Hybridizations were always done with high (0.1 ng/ml) and low (0.01 ng/ml) probe concentrations. Stringent washing was achieved with 3 changes of 50% formamide/2X SSC followed by 4 changes of 2X SSC at 42° C. After serial dehydration, slides were dipped back-to-back in Kodak NT2B liquid emulsion diluted 1:1 with distilled water. Autoradiography was at 4° C for 3 to 14 days depending on the probe. Chromosomes were identified after autoradiography by R-band staining through the liquid emulsion (see above). For photography, these same slides were restained with Wright's as described (Harper and Saunders, 1981).

30-100 metaphase spreads with good banding, low background hybridization and adequate spreading were selected for analysis. For the single copy neo probe, the

location of every grain observed on a human chromosome was recorded on a previously prepared ideogram of the relevent human chromosomes. The number and distribution of these grains was compared to the average number of grains/chromosome/cell by recording the total number of grains, the total number of mouse plus human chromosomes and the number of each of the human chromosomes in each metaphase spread analyzed. The percent of grains which would be expected on each human chromosome if they were distributed randomly was calculated by adjusting the total number of a particular human chromosome according to its size relative to the average mouse chromosome (see legend to figure 1-4) and dividing this number by the total chromosome count. For example, the percentage of grains which would be expected on human chromosome 21 in hybrid HD-15, given a random distribution) is calculated as follows:

$$\frac{\text{163 total chromosomes 21 in 75 cells analyzed}}{\text{3 chromosomes 21 per average mouse chromosome}} = \text{54.33 mouse equivalents}$$

$$\frac{\text{54.33 mouse equivalent (= chromosomes 21)}}{\text{6263 total chromosomes}} = 0.86\%^{**}$$

*** i.e. human chromosomes 21 represent 0.86% of all chromosomes analyzed.* Therefore, *it is expected that if grains were distributed randomly, 0.86% of them would be on chromosome 21. i.e. this is the expected percentage of grains.*

RESULTS
CHAPTER 1

Dominant selectable markers were introduced into human chromosomes by infecting diploid human fibroblasts at passage ten with a high titer preparation of SVX(neo) retrovirus. G418 resistant cells from approximately 75 independent colonies were pooled to form a a mass population. It was necessary to start with cells at a very early passage, because primary human fibroblasts lose the ability to form micronuclei (the basis of the microcell transfer technique) long before their capacity to divide is exhausted. The mass population of infected cells was induced to form micronuclei by prolonged exposure to a high concentration of the mitotic inhibitor, colcemid. Micronuclei formed in 44% of the donor cells, with a few relatively large micronuclei in most of the micronucleate cells. The cells were enucleated by centrifugation in the presence of cytochalasin B and the resulting preparation of microcells was fused to mouse 3T6 cells. Hybrid cells were recovered following selection in G418.

Fifteen microcell hybrid clones were picked and screened for the presence of human chromosomes first by staining metaphase spreads with alkaline Giemsa, which differentiates human and rodent chromosomes. The method also permits the classification of human chromosomes into groups A,B,C,D,E or F (London conference, 1963) based on size and morphology. Six karyotypically simple clones were selected for further analysis to test for isozyme markers on chromosomes in the designated chromosome groups. For example, the HD-5 clone contained only a D group chromosome and so cell extracts from this clone were tested for isozyme markers corresponding to the three D group chromosomes. This revealed the presence of human nucleoside phosphorlyase (chromosome 14) and the absence of human esterase-10 (chromosome 13) and mannose phosphate isomerase (chromosome 15).

The other hybrids were more complex. HD-9 contains human chromosome 20 as its only intact human chromosome, but markers on chromosomes 14 (pAW101) and 21 (p267) were fortuitously also detected (see below) in the DNA of these cells. Apparently this clone also carries fragments of human chromosomes which were not detectable with the alkaline Giemsa stain.

Clones HD-4, HD-15, HD -18 and HD-20 were of particular interest because they contained chromosomes from the B group which includes human chromosome 4. Therefore these clones were more thoroughly characterized by R-band karyotyping. Table 1 reflects our finding that clones HD 4,18 and 20 each contain one human chromosome in the majority of cells with additional chromosomes in a lower percentage. HD-15 contains two human chromosomes. Chromosome 4 was not the prevalent chromosome in any of these hybrids. The initial isozyme marker analysis confirmed these cytogenetic data (Table 1).

Seeking more rigorous characterization of the human chromosome content of these hybrids, we performed Southern blot analysis on the DNA from hybrid cells using previously mapped DNA markers as probes. We first established that each hybrid contained only one copy of the SVX(neo) retrovirus by digesting genomic DNAs with Eco R1, an enzyme which does not cut within the retrovirus vector. Southern blots of agarose gels containing these DNAs were probed with a labeled 1.4kb fragment containing the neomycin phosphorylase (neo) gene. Since Eco R1 does not digest the SVX (neo) vector, the size of Eco RI fragments which are revealed by hybridization with neo reflect the position of Eco R1 sites in the host flanking DNA. The single neo-hybridizing band in each digest (Lanes 3-8, panel A ,Figure 1) indicates that there is only one copy of SVX(neo) in each clone. Furthermore, the different sizes of those fragments supports the hypothesis that SVX(neo) is integrated at different sites in each clone. We further hypothesized that SVX(neo) was integrated in the most prevalent human chromosome which was presumably maintained under selective pressure.

DNA markers which had been previously mapped to the particular chromosomes identified in the microcell hybrids were selected for Southern blot analysis of all six hybrids. We examined the hybrids with markers from: chromosome 14, pAW 101 and a cDNA for nucleoside phosphorylase, pNP ; from chromosome 21, p267; from chromosome 20, D20S2 and a cDNA for adenosine deaminase; from chromosome 11, a genomic fragment of the insulin gene and chromosome 4, D4S10 and 36/2.1. As the examples shown in panels B through D of Figure 1 show, each of these markers confirmed the cytogenetic and isozyme results with the exception of the pNP and p267 (not shown) markers from chromosomes 14 and 21, respectively. Hybrid HD-9 had previously been characterized as containing only human chromosome 20, but the Southern blot results indicated that DNA from these cells also contained fragments from chromosomes 14 and 21 since markers from those chromosomes were clearly present (eg., Lane 5, panel B).

The surprising results of marker analysis on hybrid HD-9 emphasized the need for other methods to detect human chromosomes or fragments which are not apparent by alkaline Giemsa or chromosome banding techniques. To approach this problem, we performed Southern blot analysis on Eco R1 digested DNA from each hybrid using a cDNA of argininosuccinate synthetase (pAS) as a probe. As described earlier, this probe detects 18 Eco R1 bands in human genomic DNA, 15 of which have been assigned to single human chromosomes. This experiment is thus designed to detect in each of the microcell hybrids the presence or absence of the 10 distinct human chromosomes containing pAS-related sequences. For example, the karyotype data predicted that in an Eco R1 digest of DNA from cell line HD-15, the pAS probe should reveal 1.6 and 4 kb bands, since these bands have been assigned to chromosome 11q. As the autoradiogram in Figure 2 (lane4, left panel) shows, the 1.6 kb band can be seen after 3 days of autoradiography. The 4 kb and also an expected 9 kb band from chromosome 4 co-migrate with cross hybridizing mouse bands. The DNA from cell line HD 20 should

contain a 3.4 kb Eco RI band located on chromosome 7, but this also comigrated with a cross-hybridizing mouse fragment of the same size (panel B). There are no other human-specific pAS bands seen in the DNA from any of the other cell lines, thus confirming the karyological data. Comparing panels A and B, however, we demonstrate the importance of doing short and long autoradiography after Southern blotting and hybridization: in the short exposeure, weakly hybridizing bands can be seen which upon longer exposure are occluded by closely migrating, more strongly hybridizing bands and conversely, only after long exposures will some weakly hybridizing bands (eg. 1.6 kb band in lane 4) be visible. In somatic cell hybrids, weakly hybridizing bands may frequently be the result of low copy number per cell of the chromosome or chromosome fragment containing that sequence.

Rather than simply inferring that the individual chromosomes which were most prevalent in each microcell hybrid actually carried the neo selectable marker, we wanted to verify the location of the introduced markers. For clones HD-4, 15 and 18 we were able to map the neo integration sites by *in situ* hybridization of a ^{3}H-labeled neo probe to preparations of metaphase chromosomes. In the first experiment, in clone HD-15 we observed 4.1% (16/395) of the total grains over human chromosome 21 in 75 metaphase spreads, although the chromosomes 21 represented only 0.6% of the total chromosome length in each cell counted. In the same experiment 4.3% (29/661) of the total grains were observed over human chromosome 11 in 90 metaphases examined from HD-18, although the chromosomes 11 represented only 1.3% of the total chromosome length in each cell counted. In contrast, we observed random background numbers of grains over the other human chromosomes present in each of these cell lines: 1% of grains over chromosome 11 in HD-15 (2% total chromosome length) and 1.2% of grains over chromosome 4 in HD-18 (1.6% of total chromosome length). In the second *in situ* hybridization experiment, on metaphase spreads from clone HD-4 we used a neo probe with higher specific activity, resulting in an increase in the background number

of hybridizing grains but with a concomitant increase in specific signal. We observed 10% (16/160) of grains over human chromosome 20, representing 0.4% of the total chromosome length, as opposed to 3.1% (5/160) of grains over chromosome 4, representing 0.5% of total chromosome length (Table 2). In both experiments a clear clustering of grains was seen over specific regions of individual human chromosomes (Figure 3), localizing the SVX(neo) integration sites to chromosome 21q2.2 in HD-15, to chromosome 11q1.4-2.2 in HD-18 and to chromsome 20p1.2 in HD-4.

In *in situ* hybridization experiments on HD-20 we were never successful at detecting specific neo signal over either human chromosome 4 or 7. Furthermore, segregation analysis testing for the co-segregation of neo with markers on chromosomes 4 or 7 showed that segregation of neo was not concordant with either human chromosome (not shown). These results suggested the presence of an additional human chromosome fragment containing the neo marker which had not been previously detected. To evaluate this possibility, we performed *in situ* hybridization on metaphase spreads from HD-20 using ^{3}H-labeled human genomic DNA as probe. Using probe with a specific activity of 1×10^{8} cpm/ug DNA, this more sensitive technique revealed a small human chromosome fragment translocated onto a mouse chromosome (Figure 4). This fragment presumably contains the integrated neo marker.

Still in pursuit of a microcell hybrid containing a human chromosome 4 carrying the neo selectable marker, we carried out a second round of microcell mediated chromosome transfers. The mass population of primary microcell hybrids was chosen for chromosome donors since this population should contain a representative sample of marked human chromsomes and the mouse host fibroblasts are significantly more permissive for micronucleation than diploid human fibroblasts. Again, G418 resistant clones were selected for initial characterization by alkaline Giemsa staining and limited isozyme analysis. Five clones were identified which contained B-group chromosomes and these were further characterized by R-banding. Four of these clones each had more

than one human chromosome per cell, but one clone was found to be a mixed population of cells some of which contained only human chromosome 4. Sub-cloning and R-band analysis of such sub-clones resulted in the retrieval of a microcell hybrid clone, HD-113/2B, which contains only human chromosome 4. Indeed, most cells in HD113/2B contain two to three copies of chromosome 4. Characterization of this hybrid cell line will be the focus of Chapter 2 and it is also the reagent for experiments described in Chapters 3 and 5.

Table 1. Summary of Karyotype and Human Chromosome Marker Analyses.

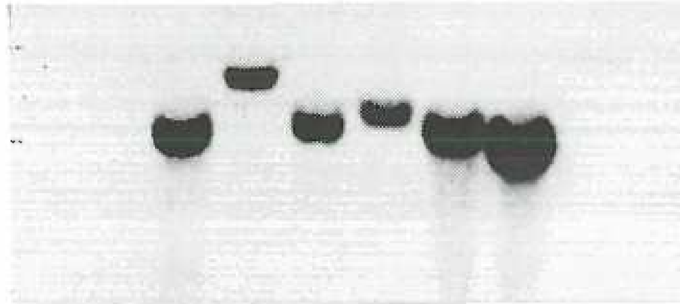
Clone	Human Chromosomes Retained (Fraction of Cells)	Human Markers Detected[d] (Corresponding Human Chromosome)	Site of neo Integration
HDm-4	#20 (0.84) #4 (0.28)	PGM-2 (4) G8 (4p) 3.6/1.2 (4p) D20S2 (20) pAS (9kb) (4)	20p1.2-2[a]
HDm-5	#14 (0.81) E group (0.25)	NP (14) pAW101 (14)	14[b]
HDm-9	#20 (0.90) #14 #21	p267 (21) ADA (20) D20S2 (20) pAW101 (14) pNP (14)	n.d.[c]
HDm-15	#21 (0.65) #11 (0.41) #4 (0.34)	p267 (21) PGM-2 (4) 3.6/1.2 (4p) Insulin (11p) pAS (1.6 and 4 kb) (11) pAS (9kb) (4)	21q2.2[a]
HDm-18	#11 (0.87) #4 (0.47)	pAS (9 kb) (4) PGM-2 (4) G8 (4p) 3.6/1.2 (4p) Insulin (11) pAS (1.6 and 4 kb) (11)	11q1.4-2.2[a]
HDm-20	#7 (1.00) #4 (0.17)	D7S8 (7) PGM-2 (4) G8 (4p) 3.6/1.2 (4p) pAS (9 kb) (4) pAS (3.4 kb) (7)	n.d.[b,c] (See results)

a. Determined by in situ hybridization
b. Determined by segregation analysis
c. n.d. - not determined
d. Cell extracts were tested for the following human isozymes as described in the text:
phosphoglucomutase-2 (PGM-2)
nucleoside phosphorylase (NP)
adenosine deaminase (ADA)
esterase-10 (chromosome 13 marker)
mannose phosphate isomerase (chromosome 15 marker)
glucose phosphate isomerase (chromosome 19 marker)

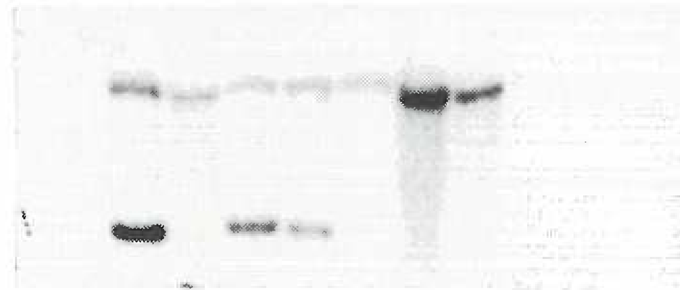
FIGURE 1: Autoradiogram of Southern blot analysis of six microcell hybrids. Eco RI digested DNAs from NIH3T3 mouse (lane 1); human skin diploid fibroblasts (lane 2); microcell hybrid HD 4 (lane 3); HD 5 (lane 4); HD 9 (lane 5); HD15 (lane 6); HD 18 (lane 7); and HD 20 (lane 8) cells were hybridized with restriction fragments containing the neo gene (Panel A); nucleoside phosphorylase (Panel B); the human insulin gene (Panel C); or the anonymous marker p267 (Panel D).

1 2 3 4 5 6 7 8

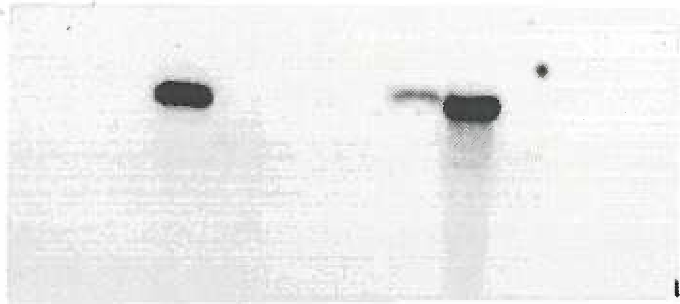
A



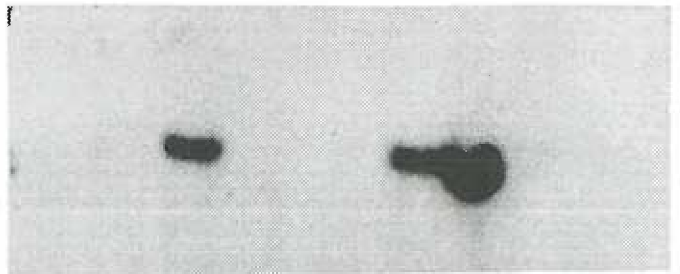
B



C

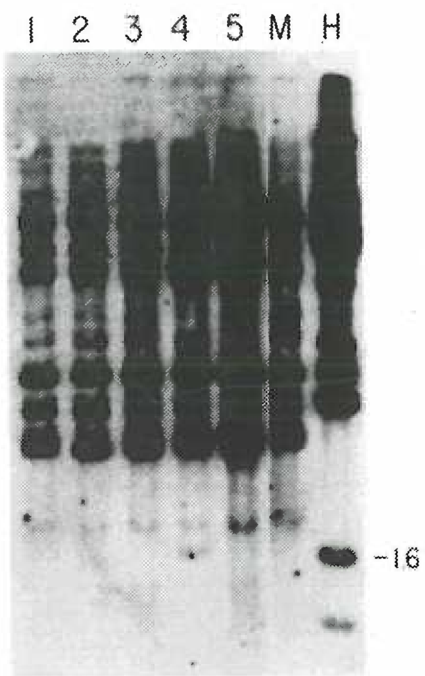


D



1 2 3 4 5 6 7 8

FIGURE 2: Autoradiograms of Southern blot analysis of five microcell hybrids. Eco RI digested DNAs (10ug) from HD4 (lane 1); HD 5 (lane 2); HD 9 (lane 3); HD 15 (lane 4); HD 20 (lane 5); NIH 3T3 mouse (lane M); and Hela (lane H) cells were probed with the cDNA of arginnosuccinate synthetase (pAS). Autoradiography was for 20 hours (right panel) to permit detection of closely migrating bands or for 3 days (left panel) to permit detection of poorly hybridizing bands (eg. 1.6 kb band lane 4)


1 2 3 4 5 M H
-16

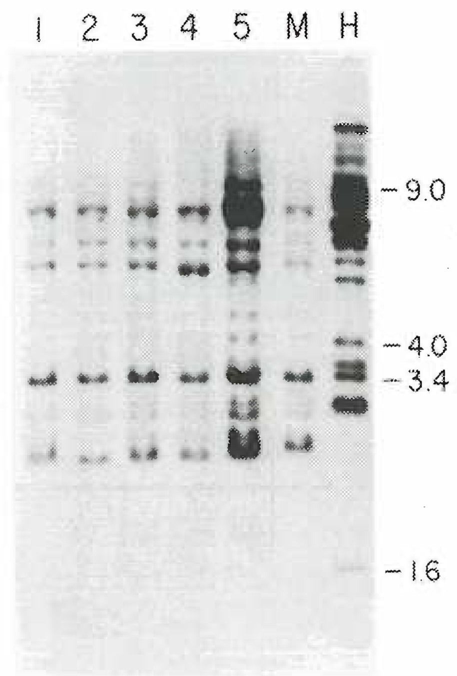
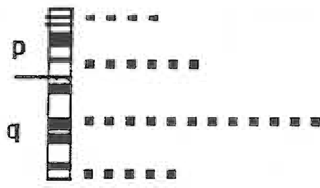
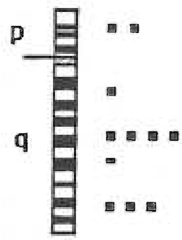

1 2 3 4 5 M H
-9.0
-4.0
-3.4
-1.6

FIGURE 3: Ideograms showing the localization of silver grains over human chromosomes after in situ hybridization of the neo gene fragment to metaphase chromosomes from cell lines HD18, HD 15 and HD 4. Each dot represents one grain. Total grains, cells and chromosomes scored in these experiments are given in Table 2.

HDm-18

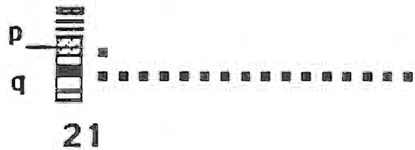


11

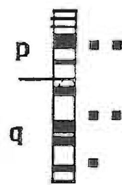


4

HDm-15

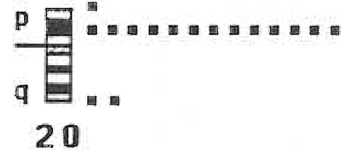


21

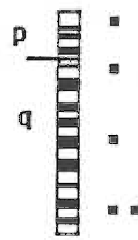


11

HDm-4



20



4

Table 2: Compiled data for total numbers of chromosomes, silver grains and cells counted after in situ hybridization of metaphase chromosomes from three microcell hybrids with the neo gene insert.

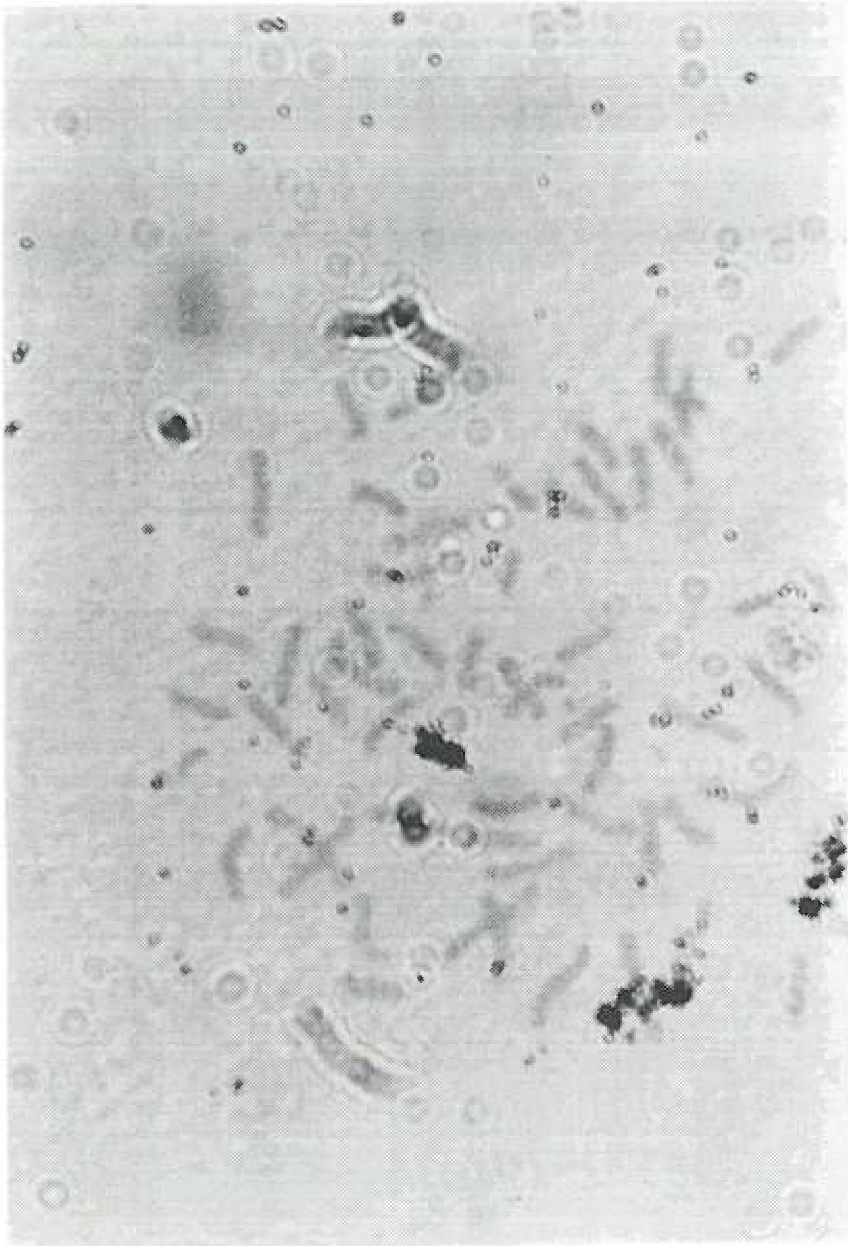
Cell line	Total grains/ Total chromosomes	# Human chromosomes	# grains on human chromosomes	% expected **/ % actual grains on human chromosomes
Experiment I				
HDm-15	395/6263	Chrom. 21: 163	16	0.9%/4.1%
	(75 cells)	Chrom. 11: 64	8	1.0%/2.0%
HDm-18	661/7345	Chrom 11: 98	29	1.3%/4.3%
	(90 cells)	Chrom 4: 58	11	~~1.1~~ 1.2%/1.6%
Experiment 2				
HDm-4	160/3912	Chrom. 20: 44	16	0.6%/10.0%
	(50 cells)	Chrom. 4: 14	5	0.5%/3.1%

** Human chromosomes were counted according to their size relative to the average mouse chromosome as:

Chromosome 11: 1.0 mouse equivalent
Chromosome 4: 1.5 mouse equivalent
Chromosome 21: 0.3 mouse equivalent

Example: % grains expected on human chromosome 21 in HDm-15 (if grains are randomly distributed over all chromosomes
163/3=54.33 mouse equivalents → 54.33/6263=0.86% of all chromosomes counted are chromosome 21
therefore, expect that randomly 0.86% of all grains will be on chromosome 21

FIGURE 4: Metaphase chromosome spread from cell line HD 20 after in situ hybridization with ^{3}H-labeled human genomic DNA. Chromosomes were initially identified by R-band staining and were subsequently stained with Wright's stain for photography (shown here). Arrow indicates the location of a small human fragment translocated onto a mouse chromosome. The intact human chromosomes 4 and 7 which were identifiable by fluorescnt microscopy with R-band staining are obliterated by silver grains here in bright field illumination.



DISCUSSION
CHAPTER 1

We have made a directed effort to create an interspecies hybrid cell line containing a single human chromosome 4 carrying a dominant selectable marker. In the process of obtaining such a cell line other useful hybrids have been isolated and characterized. Previously, interspecies hybrids made by fusing human and rodent cells have been characteristically unstable as human chromosomes randomly segregate. Maintaining desirable chromosomes through many passages was typically difficult and cultures had to be routinely re-characterized for human chromosome content. By introducing a dominant selectable marker (neo) into human chromosomes we have been able to develop hybrids which selectively maintain the marked chromosomes. These hybrids can be maintained indefinitely in selective media and do not require constant re-characterization.

The utility of any somatic cell hybrid is a function of how well its chromosome content is defined. We have used multiple isozyme and single copy DNA markers to characterize these six human/mouse hybrids. We have also visually inspected preparations of metaphase chromosomes from each, first with alkaline Giemsa to distinguish human from mouse chromosomes and then with R-band staining to identify specific chromosomes by their banding patterns. It is possible to identify most of the human chromosomes present in a hybrid using these four methods, but this is prohibitively tedious. We have demonstrated for the first time the use of a single DNA marker which enabled us to assay hybrid cells for the presence or absence of multiple human chromosomes . With the pAS probe, which detects a family of related sequences dispersed over 10 human chromosomes, we were able to confirm the presence of 3 human chromosomes as well as exclude 7 chromosomes in four hybrids and to exclude

the presence of all ten chromosomes from the other two hybrids. The appeal of this method is obvious: the DNA from many hybrids can be simulataneously evaluated on a few Southern blots and one hybridization with pAS replaces up to a dozen hybridizations with single copy probes.

Other interspecies hybrids have been developed by transferring pools of human chromosomes containing dominant selectable markers into rodent recipient cells (see Introduction). Since only one intact human chromosome was detected in these hybrids, it was inferred that the selectable marker must be in the apparent human chromosome. We wanted to specifically locate the integration site of the neo marker in these HD hybrids for two reasons: 1) identifying the location of the marker more completely characterizes the hybrid by clarifying which human chromosome (if any of the apparent intact ones) will be stably maintained, and 2) we are interested in using the selectable marker to transfer large fragments containing the marker. Our efforts to locate the neo integration site in hybrid HD-20 demonstrates the importance of this information; the neo marker in this hybrid is not in either of the expected, prevalent human chromosomes.

The presence of chromosome fragments translocated onto rodent chromosomes is a common problem with somatic cell hybrids of all kinds. The alkaline Giemsa differential staining method for discriminating between human and rodent chromosomes frequently fails to detect small human fragments - eg. the fragment found in HD-20. *In situ* hybridization with either radiolabeled or non-isotopically tagged DNA fragments containing reitterated sequences like the Blur 8 representative of the Alu family is more sensitive and reliable. Total genomic human DNA can also be used as a probe as we show here. The technique relies on the ubiquitous distribution of repetitive DNA in the human genome. An estimated 20-60% of human DNA sequences are repeated 500 to 100,000 times (McDougall et al1973) and represent the large fraction of genomic DNA which re-anneals under low to moderate C_0t conditions. Interestingly, in stringent

hybridization conditions these elements sharply discriminate between human and non-human DNA. Thus, when used as probes in *in situ* hybridizations of metaphase chromosomes from interspecies hybrids, these elements hybridize specifically to human chromosomes. Furthermore, because many repetitive elements are so ubiquitously distributed in every region of the genome, even small chromosome fragments contain sufficient copies of them to be detected by this method (Gerhard et al, 1987). Other non-isotopic methods for attaching reporter molecules to probes have been developed which increase the sensitivity and clarity of this method (Manueledis and Ward 1984; Langer-Safer et al, 1982) These have been used in elegant experiments to identify small interspecies translocations and insertions (Durnam et al, 1985).

Hybrids containing single human chromosomes were difficult to obtain throughout these experiments. This was due to the tendency of both the human diploid fibroblasts and the primary microcell hybrid fibroblasts to form large micronuclei. The large micronuclei each contain and transfer several chromosomes, resulting in the formation of hybrids that are rather complex. This resistance to micronucleation is typical of human diploid fibroblasts after passage 17 (Lugo, pers. comm) but it was a surprising characteristic of the primary microcell hybrids which seemed to behave more like their human than rodent parents.

Hamster cells would have been a better choice for fusion partners for these microcell mediated chromosome transfers since they micronucleate well. However, in view of our long range goal to insert a marker near the HD region, we anticipated that it would be necessary to infect primary hybrids containing chromosome 4 with secondary and possibly tertiary retroviral vectors to successfully target that region. As mentioned previously, hamster cells lack receptors for ecotropic and amphotropic retroviruses and so they were not, in fact suitable cells for these experiments.

Using this procedure, however, we were able to obtain a monochromosomal hybrid (HD-113/2B) containing human chromosome 4. This was the first step toward

selectively isolating the region of chromosome 4 containing the putative Huntington's disease gene.

INTRODUCTION
CHAPTER TWO

Chapter 1 described the derivation of a human/mouse hybrid cell line (HD 113/2B) retaining only human chromosome(s) 4. HD 113/2B was the result of microcell mediated chromosome transfer of human chromosomes containing exogenously introduced 'neo' markers into recipient mouse NIH 3T6 cells. These cells are G418 resistant and we show in this chapter that they contain one copy of the neo marker; we also seek to demonstrate that chromosome 4 is retained by these cells due to the integration of the neo marker into that chromosome. These data mark the first step toward our final goal of introducing a dominant selectable marker into the 4p region surrounding the putative HD gene. However, we also show the use of such a mono-chromosomal hybrid to rapidly map a newly cloned human gene, the mineralocorticoid receptor gene and to map the several copies of a low order repetitive element, p11L26, which reside on chromosome 4. Finally, we have used HD 113/2B to identify large restriction fragments of chromosome 4 containing moderately repetitive sequences, a result which provides new information about the number and distribution of a repetitive element on a single human chromosome.

Several methods for mapping the integration sites of retroviral vectors carrying dominant selectable markers have been developed. As described in the Introduction to this thesis, Weis et al mapped the integration site of a neo marker to the histocompatability locus in mouse chromosome 17. Mapping the integration site was a side benefit of the method which had been used to select the chromosome transferrent containing this neo marker. Essentially, by selecting for the co-transfer of histocompatability antigen genes and for neo, they selected for the random integration of a retroviral vector into the histocompatability locus. Weis et al were thus able to map this integration site because these experiments were directed at introducing a marker

into a particular region, and markers from the region were available to test for co-transfer; this method would not be applicable to mapping a randomly chosen integration site.

Another method for regionally mapping an inserted selectable marker is by testing for co-segregation of the inserted marker with an endogenous marker in the target chromosome. An example of this is the mapping of an integration into an X/11 translocation chromosome performed by Tom Glaser (unpublished). Fusion hybrids were made between lymphocytes from a patient carrying balanced X/11 translocation chromosomes and CHO cells. These hybrids were selected in HAT medium to select for retention of the X/11 translocation chromosome; some hybrids also retained the normal X. One hybrid, which by cytogenetic analysis carried only the translocation chromosome, was then infected with SVX(neo) virus to introduce the neo marker. Approximately 150 G418 resistant clones were picked and split into two parallel cultures; one culture was frozen and the other was allowed to grow for many generations in the absence of selection to permit segregation of the human translocation chromosome. The segregating cultures were then split further into two cultures; 6-thioguanine was added to one to select for the loss of the HPRT locus on the X portion of the translocation chromosome. Resistant clones were then challenged with G418 to test for co-segregation of the neo marker. 6-thioguanine resistant/G418 sensitive clones were presumed to have lost both markers by segregation of an X/11 translocation chromosome carrying a neo marker in addition to the endogenous HPRT marker. Although this was a convincing experiment to map an introduced marker, the large number of clones which had to be maintained and assayed made it overwhelmingly tedious. Furthermore, the protocol depended on the availability of a negative selection for loss of the syntenic marker (6-thioguianine).

More direct methods for analyzing the integration site of a retroviral vector have involved molecularly cloning the integrated retrovirus plus adjacent host flanking

sequences. If single copy fragments can be derived from the cloned flanking sequences, then these can be used to map the integration site. Such fragments can be used as hybridization probes on a panel of DNAs from somatic cell hybrids containing an overlapping representation of the 24 human chromosomes. Concordance of the presence of the fragment with each chromosome can be scored as for any somatic cell mapping experiment to identify the chromosome from which the fragment was derived. An integration into the mouse X chromosome has been mapped using this protocol (Geissler, unpublished).

Clones containing the integrated virus can be generated by the COS cell fusion method of Cepko described in the Introduction or by selecting a phage or cosmid clone from a library made with DNA from the infected cell line. For the SVX(neo) and (gpt) generation of vectors, selecting a clone from a phage library requires screening the library by hybridization with a neo or gpt probe. A clone containing the neo marker can sometimes be selected from a cosmid library by simply allowing a kanamycin (neomycin analogue) resistant clone to grow out of a complete library. This latter method was used by Geissler (unpublished) to clone an integration site in a mouse X chromosome and we employed the COS cell fusion method to map the neo integration in HD 113/2B as described in this chapter. However new vectors, including the one described in this thesis (Chapter 4), have been developed to improve the facility of 'rescuing' (cloning) integrated vectors.

Perhaps the most direct mapping tool available to date for any DNA marker is *in situ* hybridization of metaphase chromosomes. Chapter 1 detailed the use of this tool to map three integration sites in microcell hybrids. This is an elegant mapping method, but some chromosome regions may be refractive to hybridization under standard conditions thus preventing loci in those locations from being detected. Reasons for this are purely speculative but it is likely that local chromatin structure would dictate the (un)availablility of a particular sequence for hybridization to a homologous probe. In

the event that *in situ* hybridization does not give conclusive mapping results, alternative methods are desirable. This is exemplified by the results presented in this chapter for mapping the neo marker in HD113/2B.

The HD113/2B monochromosomal hybrid has also been an useful reagent for experiments other than those directed at inserting a dominant marker near the HD gene. For instance, in this chapter I briefly describes how HD113/2B was used to unequivocally map a cDNA encoding the human mineralocorticoid receptor to chromosome 4 (Arriza et al 1987). This map assignment illustrates the effectiveness of monochromosomal hybrids for mapping single copy genetic loci - an expected usage of such a hybrid. However, the isolated chromosomes 4 in HD113/2B have also been useful for investigating the distribution of low and moderately repeated sequences on this particular chromosome.

There are two main classes of repetitive elements in the human genome. Satellite DNA is the most highly repetitive class of elements, and fittingly gained its name from the fact that it forms a shoulder or satellite band off the main band of sheared genomic DNA on a CsCl gradient. The tendency of satellite elements to form a seperate - either lighter or heavier - band cannot be explained by predicting their buoyant density from base composition. Rather, this anomalous buoyant density is somehow determined by the characteristic tandem repetition of short sequences in these elements. The distribution of satellite elements on chromosomes is highly clustered both on 'satellite' arms of acrocentric chromosomes (Pardue and Gall 1970) and around centromeres of non-acrocentric chromosomes. Indeed, satellite DNA is the major component of the constitutive heterochromatin, or C bands, which are revealed by alkaline and SSC treatment of metaphase chromosomes (Arrighi and Hsu, 1971; Chen and Ruddle, 1971; and Yunis and Yasmineh 1971).

The distribution of the other class of repetitive elements, the interspersed elements, is not so tightly clustered. Short interspersed elements (SINES), which are

present 300,000-500,000 times per human genome, are so ubiquitously distributed that greater than 95% of human DNA fragments 18 kb in length contain at least one SINE element (Tashima et al 1981). The SINE family of repeats is typified by the Alu sequence which was originally defined by the band of 300bp fragments seen on an ethidium bromide stained gel of Alu I digested human DNA. These sequences comprise 3-6% of the genome, are highly species limited (Houck et al 1979), but are functionally cryptic.

Long interspersed elements (LINES) are present in less than 10,000 copies per genome and are typically longer sequences. The Kpn I family of LINES are 6 kb in length and occur some 10,000 times per genome, representing 2% of the total genome. The Kpn I family contains short inverted repeats demarking the ends of each element; this is one of several characteristics which indicate that these elements may be truncated versions of an archetypical retrovirus (Hattori et al 1986; Loeb et al 1986). The distribution of LINES over chromosomal distances is unknown.

The two repetitive sequences which were examined here are present in fewer copy number than even the LINE class of elements. There are a few thousand copies per haploid genome of sequences detected by a fragment (A36Fc) isolated from the G_{γ} globin gene and a few hundred copies represented by the cloned fragment, p11L26. The chromosomal distribution and organization of these elements has remained largely unknown.

By *in situ* hybridization, we show here that the distribution of the p11L26 fragment on chromosome 4 is limited to 4q. In addition, by pulsed field gel electrophoresis, sequences detected by A36Fc can be localised to individual megabase restriction fragments of chromosome 4. The data in this thesis indicate how the use of such probes in conjunction with monochromsomal hybrids can make a significant contribution to deriving a macromolecular map of each human chromosome.

MATERIALS AND METHODS
CHAPTER 2

Cell lines

The microcell hybrid cell lines HD113 and HD113/2B are described in Chapter 1. JS4, a microcell hybrid containing the dominant selectable marker gpt and a single human chromosome 4 was a gift of E. Stanbridge (Saxon, 1986). DNAs used to construct the panel of chromosome 4 and chromosome 5-retaining hybrids in addition to hybrids containing fragments of chromosome 4 were gifts of M. McDonald and most are described (MacDonald et al 1987). 'Wally' DNA is from a hamster/human hybrid which retains a human chromosome 4 in 20-30% of cells (M. MacDonald, pers. comm.).

In Situ Hybridization

In situ hybridization was used to map the integration site of SVX(neo) retrovirus in chromosomes of the cell line HD113/2B. The experimental protocol, collection of data and analysis of the results was exactly as described in Chapter 1. In other experiments, designed to detect fragments of human chromosomes in HD113/2B, total human DNA was used as probe in identical conditions as already described. Total human DNA was labeled with ^{3}H-dCTP (New England Nuclear #NET601) and ^{3}H-dATP(NEN #NET657) by the random primer method in a standard reaction containing 10ul of each radionuclide previously lyophilized and resuspended in an equal volume of water before use in the reaction. Probe at a specific activity of 4.3 x 10^{7}cpm/µg was used in the 50% formamide hybridization mixture at a final concentration of 0.1ng/ml. Autoradiography was for 5 days at 4° C.

A low order repetetive sequence, p11L26 (Shaw et al 1986), was labeled in an identical reaction to a specific activity of 1.7 x 10^7cpm/µg and was used at a final concentration 0.1ng/ml. Autoradiography was for 7 days at 4° C.

An Xba I fragment from the COS fusion rescue clone (see below) was also labeled in an identical reaction to a specific activity of 1.5 x 10^7cpm/µg. Before being used in the hybridization reaction, these fragments were pre-annealed with a large excess of unlabeled human and mouse DNA by the method of Litt and White (1985) for 2 hours at 65° C to block repetetive sequences from participating in subsequent hybridization with metaphase chromosomes. Otherwise hybridization was performed as usual and autoradiography was for 8 days at 4° C.

COS Cell Fusion 'Rescue' of Integrated Provirus

HD113/2B and COS 7 cells were co-plated at a 1:1 ratio by re-plating recently confluent plates at 1:2 dilution. Fully confluent cultures were used for COS cell fusion rescue exactly as described by Cepko et al (1984). Briefly, the two cell types were fused by 60 seconds exposure to 50% PEG/DME and closed circular DNAs were harvested 18-24 hours later from heterokaryons by the method of HIRT. These DNAs were used to transform competent HB101 E. Coli (Bethesda Research Laboratories, Catalogue #8260 SA) as per the manufacturer. Plasmids prepared from an overnight kanamycin resistant mass culture were size fractionated by agarose gel electrophoresis. Plasmid molecules larger than 9 kb (large enough to include flanking DNA in addition to the SVX provirus) were purified out of agarose and used to retransform HB101 bacteria. Individual kanamycin resistant colonies were picked and used to inoculate small cultures from which plasmids were again harvested. Restriction enzyme digests of these plasmids were analyzed first by ethidium bromide staining and direct visualization, and then blotted onto Zetabind membrane for Southern blot analysis, as described (Southern 1975).

Pulsed Field Gel Electrophoresis

Fragments of DNA from HD113/2B generated by the restriction enzymes Nru I and Sfi I were size seperated by pulsed field gel electrophoresis (PFG). Preparation of the DNA, restriction enzyme digestions and electrophoresis were performed essentially as described (Smith and Cantor 1987). In particular, electrophoresis was for 45 hours at 15 C in TBE buffer using an LKB orthogonal field electrophoresis unit with electrode settings at positions '30', '120' and '210' on the south and east electrodes and at position '60' for the north and west. 330 volts generating approximately 220 mAmps were alternately switched between opposite poles (NS vs EW) every 100 seconds. At the end of the run the gel was stained with ethidium bromide (0.5 μg/ml), photographed and processed for Southern blotting as usual. Yeast chromosome standards were also prepared as described. Phage lambda chromosome contatamers were prepared from a 1 liter culture of EMBL3 lambda phage which had been concentrated and purified by equilibrium centrifugation in CsCl. Various amounts (2-25 μl) of the banded phage were used directly to make agarose plugs by adding sterile PBS to a volume of 50μl and mixing with 50μl of 0.5% agarose. Concentration of phage particles was critical to achieving concatamerization of many (more than 10) phage chromosomes and since the number of phage particles present in the CsCl band used to make the plugs was variable between preparations, the exact aliquot of any phage preparation which was sufficient had to be determined empirically. Plugs were incubated in ESP (0.5M EDTA, 1.0% Sarcosyl lauroylsulphate, 1 mg/ml Proteinase K) buffer at 50° C for 50 hours before being used directly for electrophoresis. Adding 1M MgCl to the ESP buffer to a final concentration of 10mM was often helpful for stabilizing concatamers.

Southern blots of PFGs were hybridized in standard 50% formamide solution containing 2-3 x10^6 cpm/ml for 36 hours at 42°. Autoradiography was for 5-7 days at $^-70°$ C.

RESULTS
CHAPTER 2

Initial cytogenetic characterization of the microcell hybrid HD113/2B, as described in Chapter 1, determined that it contained chromosome 4 as the only human chromosome. To verify the cytogenetic description, we tested for the presence of four markers previously mapped to chromosome 4 in the DNA from HD113/2B. Southern blots of Eco RI, Hind III and Bam HI digested DNA were probed with the anonymous fragments D4S10 , C4H (Gilliam, unpublished) and 3.6/1.2 plus the expressed marker RAF2. Example autoradiograms of hybridization results for D4S10 and C4H shown in Figure 6 (lane 4) particularly verifies the presence of this marker which is closely linked to the HD gene. The parent clone HD113 (lane 3) and another hybrid which selectively retains chromosome 4, JS4 (lane 6) also contain the D4S10 marker; the probe detects a polymorphism at the D4S10 locus in JS4 cells, as defined by the appearance of two bands in each enzyme digest. The 3.6/1.2 and RAF2 and D4S10 markers are also present in HD113/2B (not shown).

The same Southern blots were used to establish that HD113/2B carries a single SVX(neo) provirus. There were no Eco RI nor Bam HI restriction sites in the SVX(neo) vector; therefore,if neo were hybridized to Southern blots of DNA from HD113/2B which had been digested with these enzymes, the size of the fragment detected by neo would be determined by the position of sites in the host flanking sequence. Thus the single Eco RI and Bam HI fragments seen on the autoradiogram shown in Figure 7 (lane 4) demonstrated that only one SVX(neo) virus had become integrated into a chromosome of HD113/2B. The same sized Eco RI and Bam HI fragments were also seen in Lanes 3 containing DNA from hybrid HD113 - the hybrid from which HD113/2B was subcloned. HD113 was a more complex hybrid carrying several other human chromosomes in most

cells; therefore, co-retention of human chromosome 4 and the same SVX(neo) integration site (as determined by the unique restriction fragments) during the subcloning of HD113/2B was evidence that this integration site was in chromosome 4.

To further characterize HD113/2B as a monochromosomal hybrid we sought confirmation that there were no other human chromosome fragments. *In situ* hybridization using ^{3}H-labeled genomic human DNA as probe on preparations of metaphase chromosomes from HD113/2B failed to detect human sequences other than the cytologically obvious chromosomes 4 (Figure 8). We would not expect fragments smaller than 10^{7} kb (see Gerhard et al 1987) to be detected by this method and therefore could not eliminate the possibility that such small fragments were present in HD113/2B.

Since the SVX(neo) provirus seemed to be repeatedly associated with the presence of chromosome 4, we hypothesized that the provirus was integrated into that chromosome. To test this hypothesis we performed two experiments to directly map the integration site. First, slide preparations of metaphase chromosomes were hybridized *in situ* with a ^{3}H-labeled neo fragment under standard hybridization, washing and autoradiography conditions (Harper 1981). 85 R-banded metaphase spreads were chosen for analysis. The location of every silver grain which occurred over a chromosome 4 was recorded on an ideogram and the total number of grains and chromosomes were also scored. 6.4% of the total grains occurred over 175 chromosomes 4 scored, although these represented 3.6% of the total chromosome lengths. As depicted in Figure 9, 54% of the grains on chromosomes 4 were clustered over 4p1.4-1.6 with a secondary cluster (26%) over 4q3.2-3.5. These results suggested that the SVX(neo) provirus was integrated in the 4p1.4-1.6 region. An alternative or second site at 4q could not be excluded.

We had demonstrated by Southern analysis that there was only one copy of the SVX(neo) provirus in the chromosomes of HD113/2B. To map the integration site of the

retrovirus by a complementary method, we sought to clone host DNA sequences flanking the integrated provirus. Following the COS cell fusion protocol of Cepko et al (1984) we were able to 'rescue' SVX(neo) plus approximately 2kb of flanking DNA as a plasmid molecule. The flanking DNA was found to contain mostly highly repetitive sequences, based on their avid hybridization to total genomic human DNA (not shown). However, one 2kb Xba I fragment was noted to have only weak hybridization to total human DNA and was tested further for the presence of a single copy fragment sequence which could be used to map the chromosome origin of this integration site.

The labeled Xba I fragment was prehybridized with a vast excess of cold human and mouse DNA fragments in 125mM phosphate buffer at 65 C for either 1.5 or 4 hours prior to being used in subsequent Southern hybridizations. The autoradiogram shown in Figure 10, panels A and B, are the results of such Southern hybridizations to a panel of Hind III digested, somatic cell hybrid DNAs each containing a single human chromosome 4 or 5 on a hamster or mouse genetic background (gift of M. McDonald). The multiple, but resolvable, number of Hind III fragments detected in every lane of panel A reflect hybridization of repetitive human, mouse and cross-hybridizing hamster sequences to portions of the probe which remain unblocked by cold competitor DNA. Hybridization to mouse sequences is expected because the Xba I fragment contains approximately 180 bp of the retroviral LTR sequence; these sequences are present in many copies of endogenous retroviruses residing in the mouse genome. Lanes 8 and 19 (panel A) contain total mouse DNA which was poorly digested by Hind III but which nevertheless show hybridization to this briefly pre-annealed probe. The lane containing human DNA was unfortunately removed from this blot before hybridization, but the blot shown in panel B does include a human DNA lane, to which even the 4 hour prehybridized probe still hybridizes with great intensity.

Preannealing the Xba I fragment for 4 hours (panel B) did block most of the mouse repetitive sequences, leaving only the two strongly hybridizing bands in Lane 10

containing DNA from HD113/2B. These two bands were also seen in panel A and correspond to the two Hind III fragments of SVX(neo) which were detected by a neo (Figure 3) probe. The presence of both of these bands on the autoradiogram, representing both ends of the provirus, indicated that the Xba I fragment probe contained sequences from both flanking sequences. Since under these prehybridization and hybridization conditions, the Xba I fragment probe still detects many human sequences (Lane 1), but no specific mouse sequences (Lane 2) we also concluded that the Xba I fragment originated from a human chromosome. Furthermore, the presence of single fragments in lanes 8 and 9 containing DNA from two hybrids containing only human chromosome 4 indicated that the integration site of the SVX(neo) provirus were consistent with an assignment to chromosome 4. However, the absence of similar fragments from lane 7, which contains a separate preparation of DNA from the hybrid represented in lane 8, and from lane 11, which contains DNA from another hybrid carrying a single chromosome 4, were not consistent with assignment to chromosome 4.

We also used the 4 hour pre-annealed Xba I fragment as a probe on a standard mapping panel of Eco RI digested DNAs from 15 somatic cell hybrids. Although the repetitive sequences in the Xba I fragment obscured single copy bands in total human DNA (lane H, Figure 11), single bands were detected in 5 of the somatic cell hybrids. Concordancy analysis with human chromosome content of each hybrid indicated that human chromosomes 4, 9, 11 and 22 presented the least discordancies (Table 3) compared to all other candidates.

The results from these two experiments are most consistent with human chromosome 4 in HD113/2B as the integration site of SVX(neo) since it is the only chromosome common to all three sets of data.. Experiments described in Chapter 3 further investigate this question.

HD113/2B, having been well characterized as a monochromosomal hybrid, was used as a gene mapping reagent in several experiments. For instance, a cDNA encoding the human mineralocorticoid receptor had been preliminarily assigned to chromosome 4 by hybridization to a standard somatic cell hybrid mapping panel (Rettig et al 1984). We subsequently hybridized an 800 bp and a 1 kb fragment from this cDNA to a Southern blot containing Eco RI digests of DNA from HD-4, 5, 9, 15 and HD113/2B. These probes detect both human and mouse fragments as seen in Lanes 1 and 2 of Figure 12. All of the human specific bands are also detected in the DNA of HD 20, HD-15, HD113/2B . Since chromosome 4 is the only human chromosome common to all three of these hybrids and in particular since it is the only human chromosome in HD113/2B, the mineralocorticoid receptor gene was assigned to chromosome 4 (Arriza et al 1987).

A low order repetitive family of sequences represented by a cloned fragment, p11L26, had been previously identified and some Eco RI fragments containing those sequences had been mapped to specific human chromosomes(Shaw 1986). For instance, and as shown in Figure 17 of Chapter 3, 10 Eco RI fragments were mapped to chromosome 4. However, as the distribution of those sequences on chromosome 4 was unknown, we sought to map them by *in situ* hybridization of p11L26 to the isolated chromosomes 4 in HD113/2B. Approximately 30 metaphase spreads were scored for the number and distribution of grains occurring over chromosomes 4. These are reported in the ideogram of Figure 13 and show that the family of repetitive sequences represented by p11L26 is distributed only on the long arm of chromosome 4.

We also investigated the number and distribution of another repeated sequence represented by the clone A36Fc. A36Fc was isolated from a genomic clone of the human β-globin region. It is represented by a 2.2 kb Eco RI-Pst I fragment located 5 kb upstream of the Gγ gene. This family of repeats is limited to the human genome; thus the A36 clone does not cross-hybridize with any rodent sequences. Previous estimates for

copy number of A36-related sequences were between 1000-5000 copies per genome(Gusella 1982). If these copies were spread evenly over the genome, there would be an A36 repeat every 600,000 to 3,000,000 base pairs. Since chromosome 4 represents between 4-6% of the human genome, 60 to 300 copies of A36 would be expected to be on each chromosome 4.

To explore the distribution of A36 repeats on chromosome 4, we digested DNA from HD113/2B with the restriction enzymes Nru I and Sfi I and separated the restriction fragments by pulsed field gel electrophoresis (Smith and Cantor 1987). Nru I cut the DNA into a broad range of sizes from 50 to 1000 kb, as estimated by comparison with a ladder of λ concatamers (lower panel, Figure 14). Sfi I fragments, on the other hand, range in size from 50-700 kb. A Southern blot of these DNAs, probed with A36 revealed that A36 repeats were not distributed randomly over the entire range of Nru I fragments. Rather, they were limited to fragments ranging in size from 200 kb to 750 kb as illustrated by the autoradiogram in Figure 14 (upper panel). Furthermore, since 10-12 distinct bands could be seen in the Nru I digests, we could estimate that there are only 10-30 copies of A36-related sequences on chromosome 4. In contrast, no distinct bands could be resolved in the Sfi I digest, presumably because there are too many closely migrating bands detected by the probe.

FIGURE 6: Autoradiogram of Southern blot analysis of Bam HI, Hind III, or Eco RI digested DNAs from Hela (lane 1); NIH 3T3 (lane 2); microcell HD113 (lane 3); HD113/2B (lane 4); HD 5 (lane 5); microcell JS4 (lane 6) cells hybridized with the D4S10 anonymous marker.

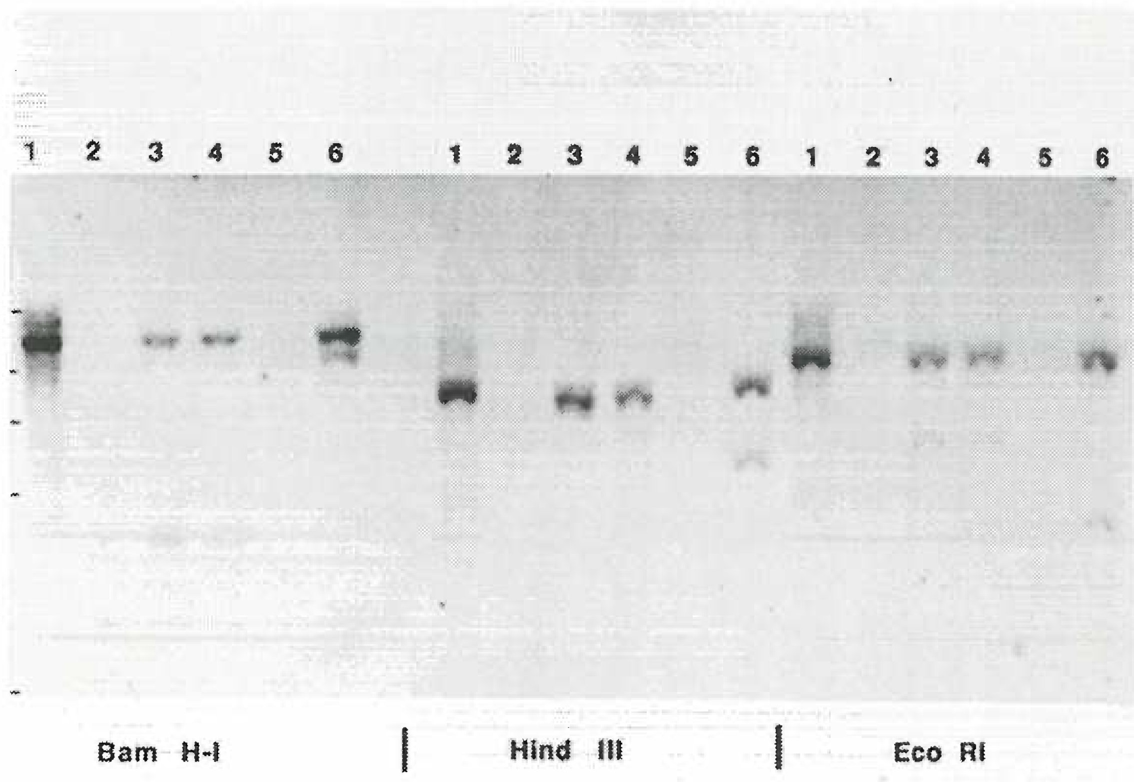

1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6
Bam H-I
Hind III
Eco RI

FIGURE 7: Autoradiogram of Southern blot analysis of Bam HI, Hind III or Eco RI digested DNAs from Hela (lane 1); NIH3T3 (lane 2); microcell hybrid HD113 (lane 3); HD113/2B (lane 4); HD 5 (lane 5) and microcell hybrid JS4 (lane 6) cells hybridized to the neo gene insert.

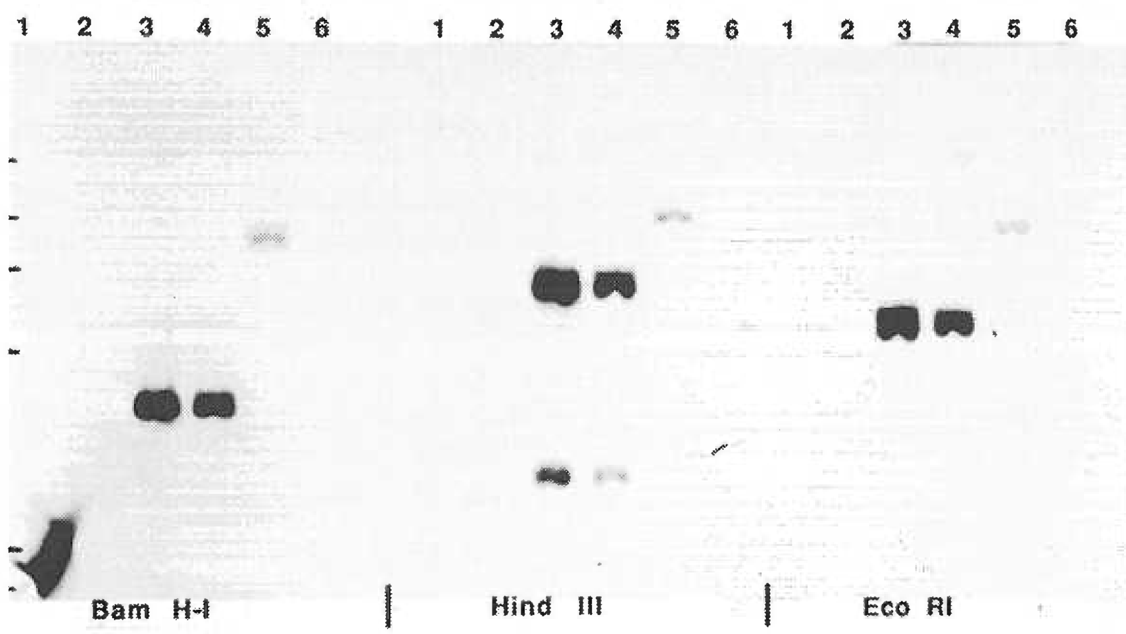

1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6
Bam H-I
Hind III
Eco RI

FIGURE 8: Metaphase chromosome spread from HD113/2B after in situ hybridization with ^{3}H-labeled total human DNA. Spreads were stained for R-banding first and then restained with Wright's stain for photography. Note that this cell has four chromosomes 4 which are obliterated by the density of silver grains.

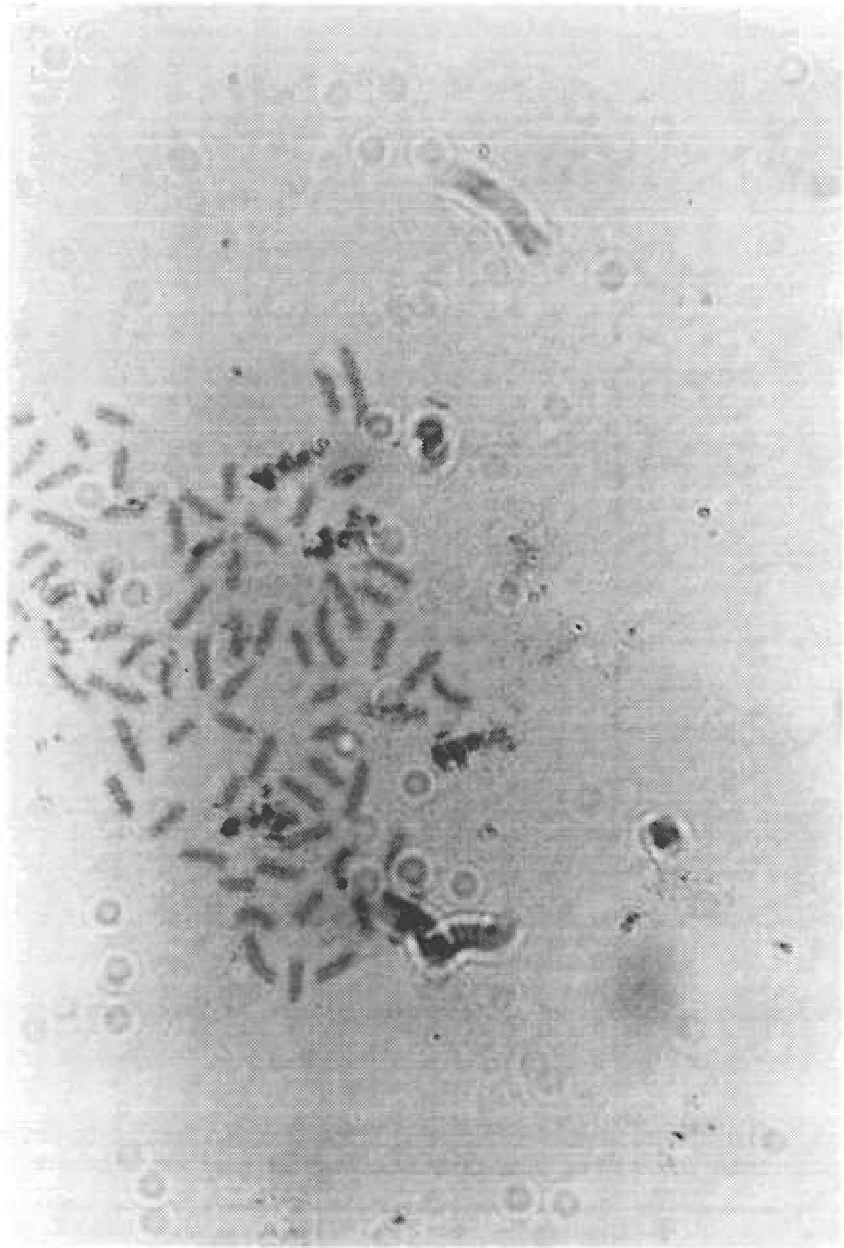


FIGURE 9: Ideogram showing distribution of silver grains over human chromosome 4 in cell line HD113/2B after in situ hybridization with neo insert probe. Each dot equals one grain.

HD 113 2B

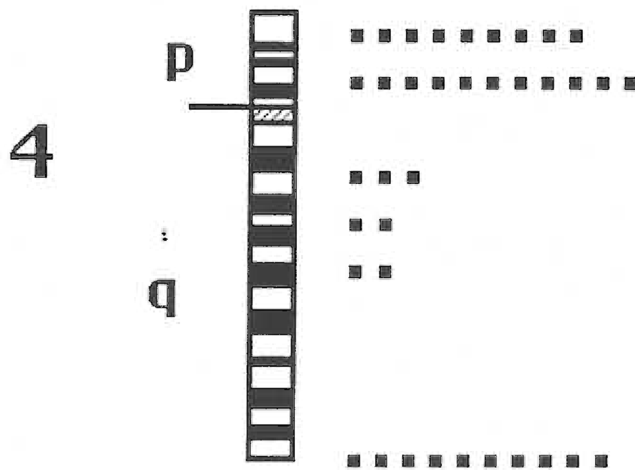
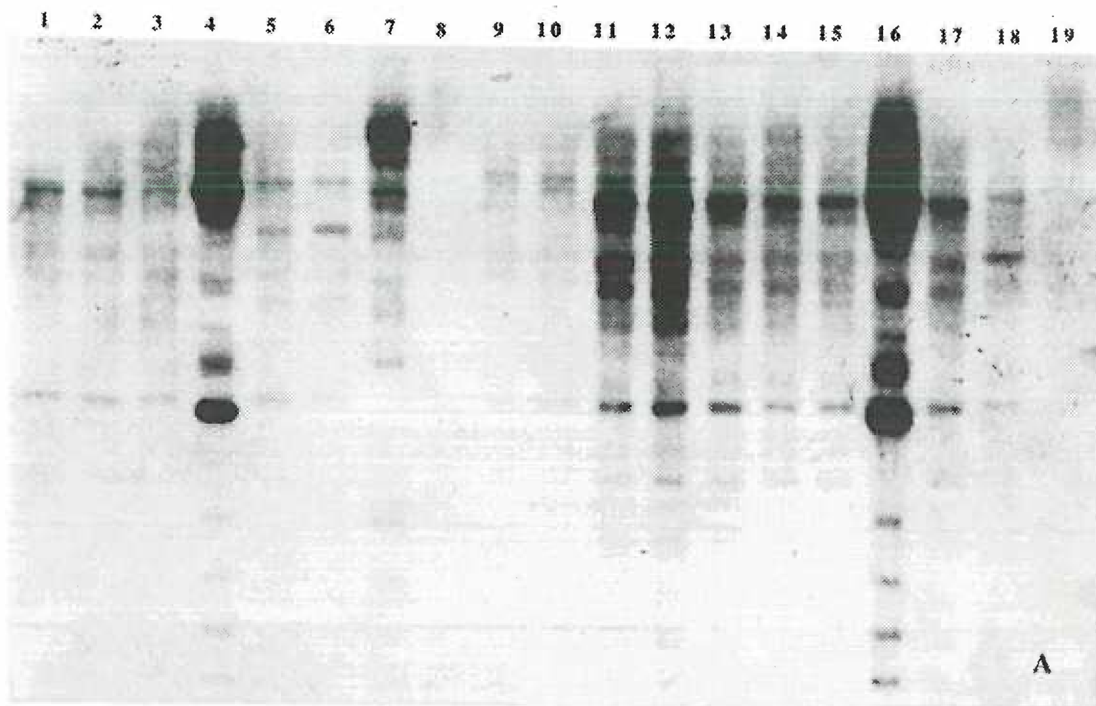


FIGURE 10: Upper panel; autoradiogram of Southern blot analysis of DNAs digested with Hind III from HHW599 (lane 1); HHW693 (lane 2); HHW416 (lane 3); HD 113/2B (lane 4); deletion hybrid from JS4 (lanes 5 and 6); JS4 (lane 7); mouse A9 (lane 8); Ade A fusion hybrid 8C13 (lane 9); Ade A fusion hybrid 8C10 (lane 10); 'Wally' (lanes 11 and 12); HHW 693 (lane 13); HHW 416 (lane 14); HHW599 (lane 15); HD 113/2B (lane 16); 'Wally' (lane 17); hamster recipient Ade A H395A23 (lane 18); mouse A9 (lane 19) cells hybridized with Xba I fragment from fusion rescue plasmid preannealed for 2.5 hours with excess cold human and mouse DNAs.

Lower panel: autoradiogram of Southern blot anlaysis of DNAs digested with Hind III from Hela (lane 1); mouse A9 (lane 2); deletion hybrids from JS4 (lanes 3 and 4); Ade A fusion hybrid 8C10 (lane 5); Ade A fusion hybrid 8C13 (lane 6); 'Wally' (lane 7); 'Wally' (lane 8); JS 4 (lane 9); HD 113/2B (lane 10); HHW416 (lane 11); HHW599 (lane 12); Ade A H395A23 (lane 13); 'Wally' (lane 14); Ade A fusion hybrids 8C10 and 13 mixed (lane 15); 'Wally' (lane 16); HHW416 (lane 17); HHW693 (lane 18 and 19); HHW599 (lane 20) hybridized with Xba I fragment from fusion rescue plasmid prennealed for 4 hours with excess cold human and mouse DNAs. Lanes 14-20 were loaded with only 1ug of DNA each.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
A

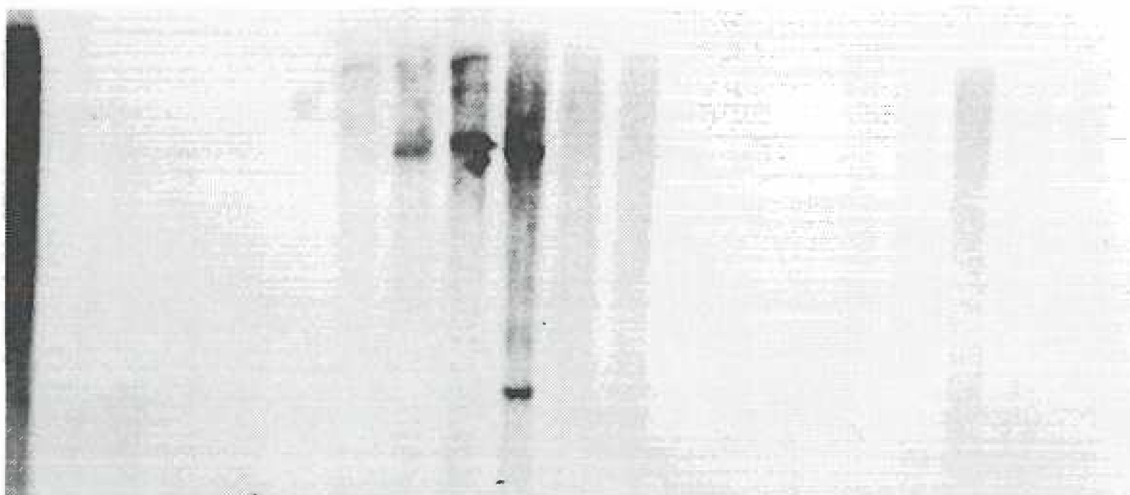

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
B

FIGURE 11: Autoradiogram of Southern blot anlaysis of Eco RI digested DNAs from a panel of somatic cell hybrids hybridized to the Xba I rescue fragment containing flanking sequences next to SVX(neo) provirus in HD113/2B. The fragment was preannealed for 4 hours before hybridization. M= mouse DNA; H= human Hela DNA. The lanes which were scored as containing Xba I fragment are indicated above.

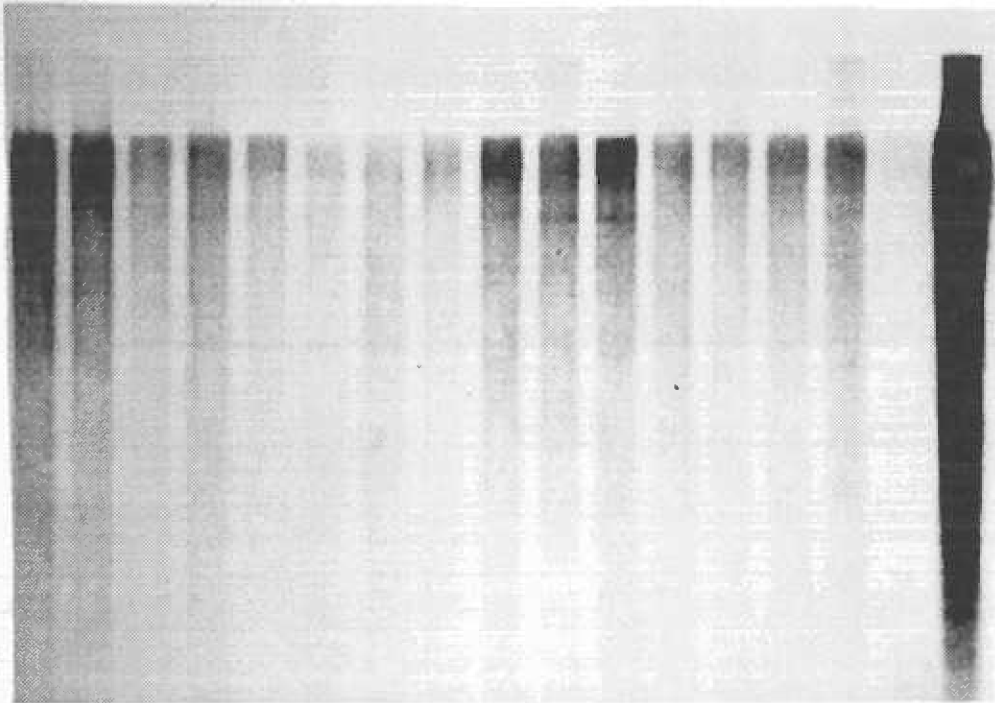

Fragment Present
+ + + + + +
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 M H

FIGURE 12: Autoradiogram of Southern blot analysis of DNAs from human Hela (lane 1); mouse NIH 3T3 (lane 2); microcell hybrid HD 4A (lane 3); HD 5 (lane 4); HD 9 (lane 5); HD 15 (lane 6); HD 113/2B (lane 7) cells digested with Eco RI and hybridized with 800 bp and 1000 bp fragments from hMR cDNA. Only 1 ug of human Hela DNA was loaded in lane 1 whereas 10 ug was loaded in all other lanes. HD 4A is a later passage of microcell hybrid HD 4 which retains human chromosome 4 in <20% of cells, as verified by re-hybridizing this same filter with the chromosome 4 anonymous probe <u>D4S10.</u>

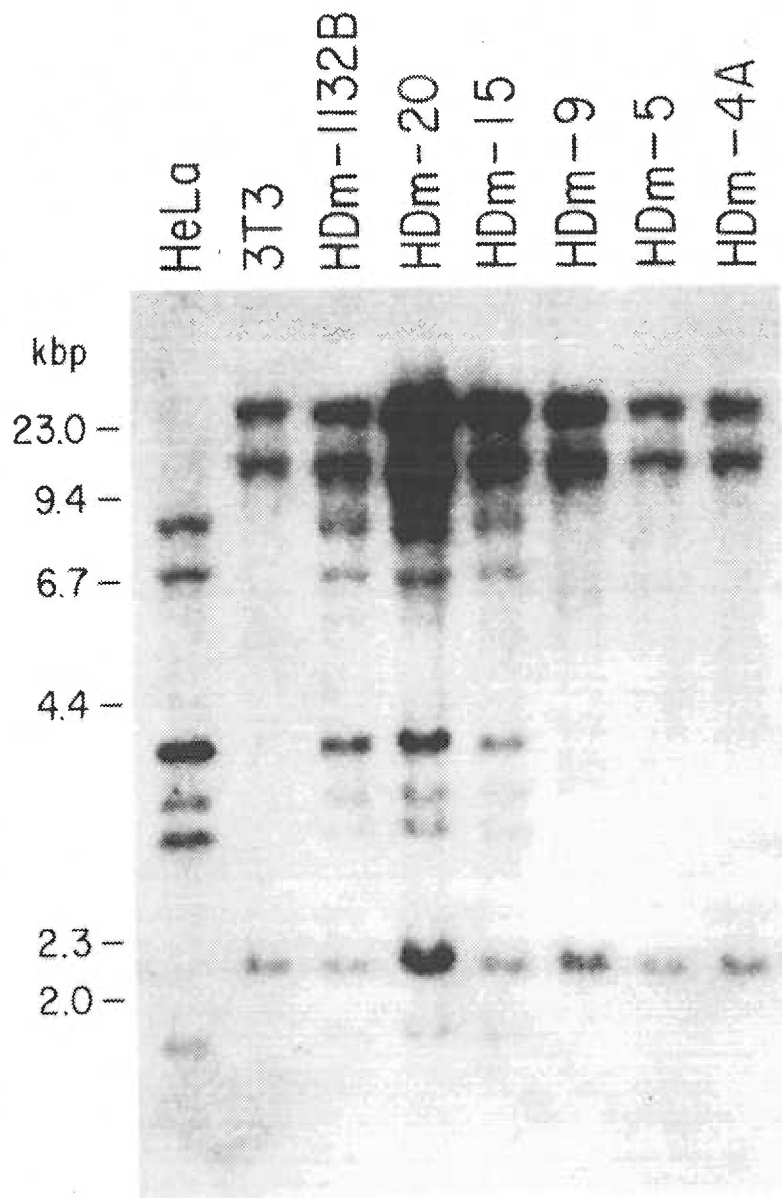
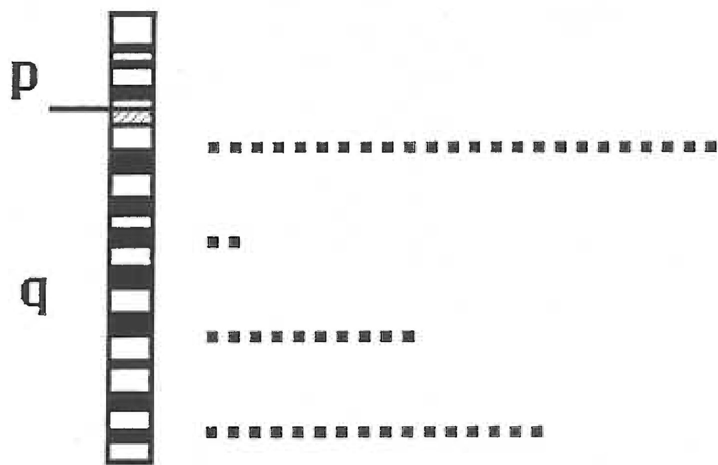
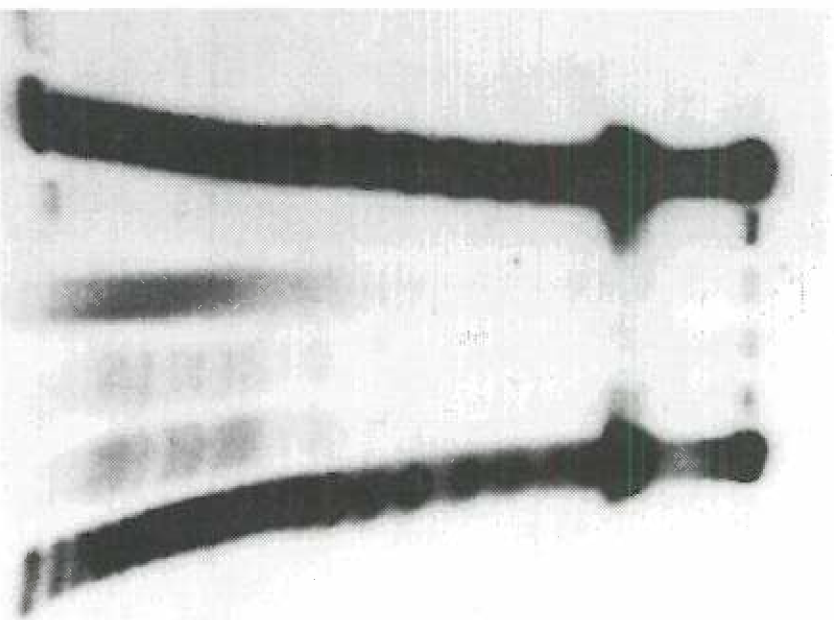

HeLa
3T3
HDm-1132B
HDm-20
HDm-15
HDm-9
HDm-5
HDm-4A
kbp
23.0
9.4
6.7
4.4
2.3
2.0

FIGURE 13: Ideogram depicting the distribution of silver grains over human chromosome 4 in microcell hybrid HD113/2B after in situ hybridization with a fragment containing the low order repetetive sequence p11L26.



4

FIGURE 14: Pulsed field gel electrophoresis of DNA from HD 113/2B cells digested with either Nru I or Sfi I and probed with the human specific repetetive sequence A36 Fc (upper panel). Lower panel: Ethidium bromide stained gel before Southern blotting showing differences in the range of sizes of restriction fragments produced by each enzyme. Yeast chromosome markers (outermost lanes) and a ladder of λ concatamers (next lanes inward) serve as size markers; a shorter autoradiography exposure was used to distinguish individual λ bands.


λ Concatamers
Sfi I
Sfi I
Nru I
Nru I
λ Concatamers

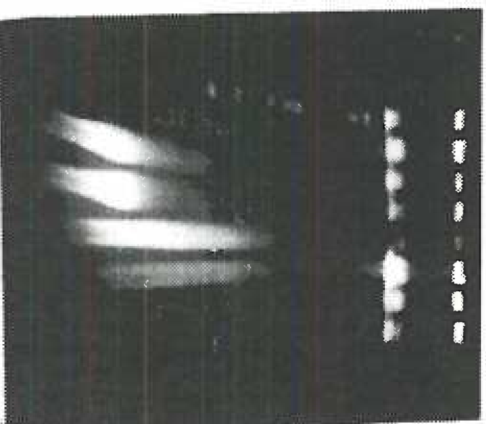


Table 3: Discordanciy analysis of Xba I fragment from Cos-fusion rescued plasmid containing integrated SVX(neo). with 15 somatic cell hybrids containing human chromosomes. Xba I fragment is flanking human DNA. Somatic cell hybrids are listed across the top and human chromosomes on the left. + indicates the presence of an intact chromosome and **o** indicates the presence of a chromosome fragment or rearranged chromosome. These hybrids (**o**) were not included in the concordancy analysis which is why not all of the discordancies plus concordancies totals do not equal 15. * *'s along the top indicate the presence of Xba I fragment in DNA from each hybrid.

	NSK-1 *	NSK-2 *	NSK-3	NSK-4 *	NSK-5	NSK-6	NSK-7	NBE-C2	NBE-D4 *	NBE-E1 *	NBE-G2 *	NBE-H1	NBE-J2	NBE-K1	NBE-M1	#DISCORDANCIES	#CONCORDANCIES
1	+	o	o	o	o	o	o	o	o	+	o	+	o	o	+	2	2
2															+	6	9
3					+	o										6	6
4	o			+		+	+		+	+						*4*	*10*
5	+						+		+							5	10
6	+		+	+	+					+	+				+	5	10
7	+	o			+		+				+					6	8
8			+	+												6	9
9	+	+														*4*	*11*
10	+				+		+		+	+	o			+	+	6	8
11	+	+		+	+		+	o	o	+	o	o	o	o	o	*2*	*7*
12	+		+	+	+	+			+	+						5	9
13																	
14	+	+	+	+	+	o	+		+	+	+	+		+	+	6	8
15	+	o	+	+	+	o		+	+	+		+	+	+	+	8	5
16									+		+					3	8
17	+		o		o	o						+				6	6
18																	
19			+	+		+		o	o	o		o	+		o	7	3
20	+		+		+		+	+	+	+	+	+	+	+	+	10	5
21		+	+	+	+	o	+	+		+	+	+	+		+	9	5
22									+		+					*4*	*11*
X	+		+				+	+	+	+	+	+				6	9

DISCUSSION
CHAPTER 2

Having derived a somatic cell hybrid containing a single human chromosome and a single copy of integrated SVX(neo) retrovirus vector, we first confirmed that the chromosome 4 in HD 113/2B contained the known DNA markers which are linked to the HD gene, D4S10 and C4H. We also confirmed that the neo marker was integrated into the human chromosome 4 and, in view of our long range goal to insert a dominant selectable marker in the region of 4p containing the Huntington's disease gene, we also determined where in chromosome 4 the provirus was integrated. A further, long term goal of our laboratory has been to be able to introduce and map a sufficient number of selectable markers in human chromosomes to create a library of insertions at known locations; such a library would be a valuable resource for somatic cell approaches to many gene mapping and gene expressions studies. For this reason - to reach the larger goal - we explored various ways to map the integration sites of retroviral vectors which have been used to insert selectable markers. Several of these were used to map the marker present in HD 113/2B (this chapter) and in Chapter 4 we describe the construction of a new vector with features which facilitate mapping the integrated provirus vector.

In situ hybridization using neo as a hybridization probe indicated that neo was integrated at the end of 4p, with some additional grains clustering at the terminus of 4q. Since we knew that there is a single copy of neo in HD113/2B, we sought to resolve this conflict by cloning the integrated SVX(neo) plus its flanking host DNA sequences so that we could determine the chromosomal origin of the flanking sequences.

Following the COS-cell fusion rescue protocol of Cepko (1984), we were able to rescue SVX(neo) plus approximately 2 kb of flanking sequences. The small size of flanking DNA which is included in this plasmid is typical for the COS fusion method of

retrieving integrated proviruses and is a reflection of proposed mechanism for how this method works (Conrad et al 1982). The ubiquitous number of SV40 T-antigen molecules available for binding to the origin sequences in SVX(neo) induce multiple, simultaneous rounds of replication from this one origin. An 'onion skin' of many strands of newly replicated DNA builds around the origin . These strands apparently are substrate for recombination enzymes which primarily promote homologous recombination; hence it is not surprising that the majority of plasmid molecules recovered by this method are circularized SVX(neo) which has become circularized via recombination in the long stretch of perfectly homologous sequences of the LTR sequences. Only occassionally do homologous sequences occur on either side of the provirus, providing substrate for recombination enzymes, and even more occcassionally does non-homologous recombination occur to the same end. Size selecting the mass population of circular molecules harvested from heterokaryons is helpful in retrieving plasmids which contain flanking sequences, but a very small percentage of even that fraction contain flanking fragments larger than 2 kb.

The cloned flanking sequence was mostly highly repetitive. However, a labeled 2.3 kb Xba I fragment including 250 bp of the retroviral LTR was prehybridized with an excess of cold human and mouse DNA and then used to detect single copy sequences in a panel of somatic cell hybrids containing human chromosomes 4. This Xba I fragment probe revealed one approximately 5.8 kb Hind III fragment in the DNA from each of two hybrids (JS4 and 'Wally') which contain only chromosome 4. Although this result would indicate that the Sba I fragment originated from chromosome 4 in HD113/2B, this conclusion is challenged by the absence of a 5.8 kb fragment in the DNA of another hybrid also containing chromosome 4 (Lane 11). Such a fragment is also absent from the DNA from 'Wally' loaded in lane 14 (Figure 3), but this is not surprising given that only 1 ug of DNA was loaded in this lane.

Finally, the Xba I fragment was ued as probe on a standard somatic cell hybrid mapping panel and was seen to be present in concordance with human chromosome 4 in 4 of 10 hybrid clones. Similar degrees of concordancy were demonstrated with chromosomes 9, 11 and 22 and so these chromosomes could not be excluded as possible integreation sites. Thus, the evidence presented in this chapter only suggested that the SVX(neo) provirus in HD113/2B was integrated in the most distal portion of human chromosome 4. Since we wanted to derive subchromsomal hybrids as part of the next step in the attempt to isolate the region around HD, we chose to further verify the location of the SVX(neo) inseeertion site by the method of Goss and Harris (1975); that is, by analyzing the sequences which could be cotransfered with the neo marker from HD113/2B to recipient hamster cells. These hybrids are the subject of Chapter 3.

The utility of a monochromosomal hybrid like HD113/2B in conventional somatic cell mapping was demonstrated by the rapid assignment of a new cDNA for human mineralocorticoid receptor gene to human chromosome 4. The presence of human specific fragments detected by hMR in the DNA of HD113/2B obviated the need for condordance analysis to make the map assignment. Clearly, a panel of 24 monochromosomal hybrids is highly desirable to so simplify chromosome assignments of newly isolated markers.

However, the isolation of chromosome 4 on a rodent genetic background also provided the opportunity to study for the first time the number and distribution of repetitive sequences over an entire chromosome. Since the repetition frequency of the repetitive element A36Fc predicted that these elements would be present every 600 to 3,000 kb it was necessary to digest the chromosome 4 in HD113/2B cells into fragments in this size range. We found that the restriction enzyme Nru I produces fragments ranging from 250 kb to >1 megabase in size and these fragments could be separated by pulsed field gel electrophoresis. A Southern blot of such digested DNAs probed with the A36 element revealed that there are only 10-30 copies of A36 on

chromosome 4 rather than the expected 60-300 copies. Moreover, these copies are preferentially found on relatively small (<750 kb) Nru 1 fragments. This last observation may indicate that A36Fc repetitive elements are paritcularly associated with chromosome regions containing expressed sequences since the Nru 1 enzyme is sensitive to methylation of the cytosine residues in its recognition sequence. The enzyme will only digest recognition sites which are not methlyated and hypomethylation has been associated with chromosome regions which are available for transcription by polymerase II (i.e. expressed).

The discovery that A36 Fc elements define a resolvable number of Nru I fragments on chromosome 4 provides a new method for detecting rearrangements and deletions of this and probably other chromosomes in hybrid cell lines. For example, the number of Nru I fragments represented in the subchromosomal hybrids described in Chapter 3 could be determined by this method. The deletion hybrids of chromosomes 11 and 12, described in the Introduction to this thesis could also be analyzed in this way and in conjunction with mapping data for specific markers on those chromosomes, a large scale restriction/genetic map of these chromosomes could be generated. In short, the technique of pulsed field gel electrophoresis, which bridges the resolution gap between cytogenetics and standard restriction enzyme mapping, combined with the use of moderately repetitive elements will give the first composite picture of the macromolecular organization of human chromosomes.

INTRODUCTION
CHAPTER 3

Our goal had been to isolate the region of human chromosome 4 surrounding the HD gene in an interspecies hybrid cell line. The first step toward this goal was to introduce a selectable marker sufficiently near to the HD gene to be able to subsequently cotransfer the HD gene with that marker. We had preliminary evidence that the neo marker in HD113/2B was integrated in the tip of 4p - i.e. in the appropriate region - although the mapping of neo was thus far equivocal. Therefore, we proposed to take the second step toward possibly isolating the HD region of chromosome 4 by transferring large chromosome fragments surrounding the HD113/2B neo marker into secondary rodent recipient cells. The creation of these sub-chromosomal hybrids would achieve two goals: (1) by analyzing the chromosomal origin of human DNA which was co-transferred with the neo marker we could confirm the location of the integration site of neo by inference, as previously done by Weis and (2) if the integration site was indeed in 4p, we would have isolated the HD region, as intended.

We anticipated that to directly determine the chromosome origin of cotransferred sequences it would be necessary to construct a phage library with DNA from one of the hybrids which could be screened for clones containing human DNA . Single copy sequences from such human clones could be used to probe a panel of DNAs from human/rodent somatic cell hybrids; the hybridization pattern across the panel would permit assignment of the probe to a single human chromosome by standard concordancy analysis. By inference, the integration site of the neo marker could also be assigned to the same human chromosome.

The phage library was also desirable for achieving a third goal, which was to generate an enriched source of molecular clones from the region immediately surrounding the neo dominant selectable marker. This would be the end point of our

directed effort to selectively isolate sequences from a chromosome region of interest. All of the human sequences cloned in such a library would be exclusively from that region.

In this chapter we describe the construction of five independent hybrids containing various sized fragments of human DNA which were selectively transferred and maintained by their association (synteny) with the neo marker of HD113/2B. We further constructed a phage library from one of these hybrids and describe here mapping of a human clone isolated from the library.

MATERIALS AND METHODS
CHAPTER 3

Goss-Harris hybrids

Deletion hybrids retaining only sub-chromosomal fragments surrounding the neo marker of HD113/2B were obtained by the method of Goss and Harris (1975). For each fusion experiment 1.2 x 10^6 HD113/2B cells were harvested by trypsinization, counted, rinsed in PBS and resuspended in 10ml DME without serum. The cell suspension was distributed into a100mm bacterial petri dish (to prevent re-attachment of cells) which was placed on ice. Cells were then exposed to either 4000 or 8000 rads of γ irradiation in a Gammacell Cs source (40 or 80 minutes at 100 rads/min). The length of time required to achieve these doses is the reason the cells were kept at 0-4° C; that is, to prevent cellular recombinase/repair activity during the radiation insult. The irradiated donor cells were used immediately for fusion with 3 x 10^6 CHTG49 hamster cells. The two populations were gently pelleted by centrifugation in a clinical tabletop centrifuge (setting 3) in a Falcon 1059 tube. After thoroughly draining the pellet, the cells were resuspended by tapping. Fusion was induced by adding 0.3 ml of 50% polyethylene glycol (Ave. molecular weight 1000, Baker #6218-7)in DME and immediately centrifuging for three minutes in a clinical tabletop centrifuge (setting 2). Toxic exposure of the cells to PEG was controlled by quickly diluting it with 10 ml of DME (w/o serum) followed by re-pelleting. The pellet was rinsed two more times with 10 ml DME (w/o serum). Finally the cells were resuspended in 10 ml DME plus 10% FCS and 0.5 ml of the suspension was distributed into each of twenty 60mm tissue culture dishes. The next day the culture media was changed to DME (10% FCS) plus 1 mg/ml G418 to select hybrids containing the neo marker. Five G418 resistant

colonies were picked from seperate dishes after 11 days in selective media, expanded into large cultures and frozen or analyzed as described.

Construction of EMBL phage library

A recombinant library of genomic DNA from the hybrid GH8000R#3 was constructed in the λ phage vector EMBL3. 9-20 kb sized inserts were prepared with high molecular weight DNA from GH8000R#3 by partial digestion with Sau 3A (NEB #169). To ensure that all Sau 3A sites would be represented in the library, enzyme digests were performed for short, medium or long time periods as follows: 60 ug of DNA was prewarmed to 37° C with appropriate 1X buffer (as recommended by the manufacturer) in a total reaction volume of 400ul. 3 units of Sau 3A were added and the reaction was incubated at 37° C for 10 minutes at which time 100 ul were removed and added to 10ul of 100mM EDTA in a 65° C water bath. The enzyme in this aliquot was heat denatured for 10 minutes at 65° C. After 20 minutes another 200 ul of the enzyme reaction was removed and heat treated as above. The last 100 ul was similarly treated after 30 minutes incubation. These DNA aliquots were pooled, precipitated in ethanol, analyzed for degree of digestion by agarose gel electrophoresis and then treated with units of calf intestinal phosphatase (Boehringer Manheim # 713 023) to remove 5' phosphates as recommended by the manufacturer. Phosphatase was heat denatured at 68° C for 15 minutes. The DNA was loaded immediately onto a 12 ml 1.25M-5M NaCl gradient and size fractionated by centrifugation at 4° C in a Beckman SW41 rotor for 4 hours at 36,000 rpm. 1 ml fractions were collected from the top of the tube using a P1000 Pipetteman and cut-off blue tips to prevent shearing of the DNA. Odd numbered fractions were assayed by gel electrophoresis and fractions containing DNA in the 9-25kb size range were selected. Fractions were pooled, dialyzed and the DNA recovered

by ethanol precipitation. The resulting DNA precipitates were resuspended at a final concentration of approximately 0.5 ug/ul in TE.

The EMBL vectors are λ replacement vectors with a large cloning capacity and pUC-derived polylinker sequences flanking the middle fragment. EMBL3 vector 'arms' were prepared as recommended (Frischauf et al 1983). Briefly, 20ug of phage DNA was digested to completion with Bam HI. After confirming the completeness of this first digest, the phage DNA was digested with Eco RI. This second digest was done to ensure that the middle fragments could not re-ligate to form non-recombinant phage during construction of the library. The DNA was precipitated with isopropanol to selectively recover large fragments (thereby lose the polylinker fragments) and this pellet was resuspended at a final concentration of 0.5 ug/ul in TE.

Test ligations were performed using various ratios of arms to insert DNAs. Final ligations for the library were done with approximately 2ug of EMBL3 DNA and 1 ug of GH8000R#3 in a final reaction volume of 10ul containing 50mM TRIS (pH 7.5), 10mM $MgCl_2$, 10mM dithiothreitol, 1 mM spermidine (pH 7.0), 1 mM ATP and 0.1 mg/ml BSA plus 1 unit of T4 DNA ligase (Bethesda Research Labs #5224SC) at 16° C overnight.

Ligation mixes were heat inactivated for 10 minutes at 65° C and then used directly in in vitro packaging reactions with the commercially prepared extracts, Gigapack Gold (Promega Biotec), as per the manufacturer. The resulting library was plated on a fresh overnight culture of LE 392 grown in 0.2% maltose.

RESULTS
CHAPTER 3

We transferred large chromosome fragments surrounding the neo marker in HD113/2B into CHTG 49 hamster cells by the method of Goss and Harris (1975). Chromosomes were fragmented with either 4000 or 8000 rad doses of γ-irradiation and transferred to CHTG 49 cells by cell-cell fusion in 50% PEG (Figure 15). Five hybrid clones (GH4000R #'s 1 and 2; GH8000R#1,2 and 3) containing fragments surrounding the neo marker were selected after 14 days growth in G418-containing medium. Each of these hybrids was isolated from separate culture dishes, indicating that they were not sibling clones.

If the neo marker in HD113/2B had integrated near to the markers most closely linked to the HD locus, then these markers would be co-transferred with neo into the GH hybrids. Therefore, DNA from each of the GH hybrids, digested with Eco RI, Bam HI and Hind III, was tested by Southern blotting for the presence of three such markers. A plasmid containing RAF2, the human pseudogene of the raf/mil oncogene, (Bonner et al, 1985) was used in its entirety as a probe to detect both the pBR 322 component of the neo marker and the RAF pseudogene. The resulting autoradiogram (Figure 16) showed that Hind III, Bam HI and Eco RI bands detected by RAF2 were present in the HD113/2B donor DNA (lanes 1, indicated by arrows) but were not present in any of the five GH hybrids DNAs. The neo marker was present in DNA from HD113/2B (lanes 1) and was also transferred to each of the GH hybrids (lanes 2-7). DNA from the hamster recipient did not contain either marker, as expected (lanes 2). We noted that the intensity of the neo marker band was consistently greater in one of the GH hybrids (8000R#3, lanes 7) than any of the others, although the amount of DNA loaded in each lane was the same. Indeed, assuming that the intensity of the neo bands in the Goss Harris hybrids other than GH 8000R#3 represent single copies of neo, then comparing the intensity of this

marker in the hybrids relative to the intensity of the neo-hybridizing fragment in HD113/2B indicated that HD113/2B may contain more than one copy of SVX(neo). However, because there is only one integration site (as reflected by the unique Eco RI, Bam HI and Hind III neo-hybridizing bands in DNA from HD113/2B), these extra copies of SVX(neo) would have to have been generated by duplication of the entire region containing the integration site.

We also tested for the cotransfer of the C4H and D4S10 markers. Again, DNA from HD113/2B contained bands detected by both the C4H and D4S10 probes (Figure 17, lanes 6, panels A and B respectively) but DNA from all the the GH hybrids did not (lanes 1-5). The C4H marker cross hybridizes with a sequence in hamster DNA which is seen in each of the GH hybrid DNAs as well as the parent CHTG49 DNA. In fact, this cross hybridizing fragment acts as an internal positive control for hybridization, verifying that the absence of the human C4H sequence in each of the hybrids is not due to limitations of detection with this probe. Two other recently isolated markers from this region of chromosome 4 were also not detected in any of the hybrid DNAs (M. McDonald, personal communication). The failure to co-transfer any of these three markers excludes the integration site of SVX(neo) in HD113/2B from the region immediately surrounding these markers in 4p16.

We had mapped the chromosome 4-specific fragments detected by a low order repeat probe, p11L26, to multiple sites on 4q (Chapter 2). Therefore, using p11L26 as a probe we could test for the presence of many fragments from 4q in DNAs from the GH hybrids with one hybridization experiment. The results of such an experiment, shown in Figure 18, demonstrate that none of the p11L26 fragments in HD113/2B (lanes 7) were transferred to the GH hybrids. These results exclude multiple regions of 4q from containing the integrated neo marker including one in 4q3 (see Figure 9, Chapter 2) near the secondary in-situ mapping site of SVX(neo).

To assess the quantity and species origin of DNA transferred into the GH hybrids, the same Southern blots shown in Figure 4 were re-hybridized with two human-specific moderately repeated sequence probes, Blur 8 and TεGC. Blur 8 detects sequences in the Alu family of repeated elements which are present in 300-500,000 copies per genome. Human chromosome 4, analyzed alone in DNA from HD113/2B, contained so many Alu-related fragments that they could not be resolved on a standard Southern blot (Lanes 2 of Figure 19). However, each of the GH hybrids contained a limited (20-50) number of Blur 8-related fragments. The pattern of these fragments was very similar in each of the hybrids, as expected for the co-transfer of a contiguous region immediately surrounding the neo marker. We estimated from the number of Blur-8 fragments and their expected frequency in the genome every 15-20 kb , that each hybrid contained between 1000 and 2000 kb of human DNA; this represented less than one interphase G-band.

A human specific sequence which detects a family of repeated elements present in only 10-20,000 copies per genome was also used to analyze the human DNA content of the GH hybrids. This cloned sequence, TεGC, was derived from the same region of the human beta globin gene cluster as was the A36 repeated element used in Chapter 2. In fact, these representatives of each family are only 1.5 kb apart in the beta globin cluster (Gusella et al 1982). From its frequency in human DNA we estimated that, if TεGC is distributed randomly in the genome, then there should be one copy every 150-300 kb. Thus a hybrid carrying 1-2000kb of human DNA should contain approximately 6-14 copies of TεGC-related elements.

This prediction is very nearly met by the results seen after hybridizing TεGC to the GH hybrid DNAs. For example, in the brief autoradiography exposure shown in panel A of Figure 20, at least 15-20 fragments are detected by the probe in each restriction digest of DNA from GH8000#3 (lanes 7). As noted previously, this hybrid consistently contains more human DNA per ug of total DNA loaded perlane, and so

represents the largest fragment co-transferred with the neo marker in these experiments. The longer exposure shown in panel B permits detection of all the $T_{\varepsilon}GC$ related fragments in each hybrid and again, a common set of fragments which must reside close to the selected marker is revealed in each of the hybrids. This probe also detects differences in the pattern of transferred fragments, confirming that the hybrids are the result of individual transfer events. From the number of $T_{\varepsilon}GC$ fragments detected in each hybrid and the repetition frequency of this element, we estimate that approximately 3000-5000 kb of human DNA has been transferred to each hybrid. This is a 2-fold increase over the estimate made from the number of Blur 8 fragments in the hybrids, but these estimates are based on two assumptions: (1) that each repetitive element is randomly distributed and (2) that each band seen on the autoradiogram represents one or two copies of an element. Both assumptions are undoubtedly gross simplifications and therefore estimates based on them would be expected to vary by at least 2-3 fold.

The cotransfer of a common set of human DNA fragments with the neo marker into each of the five GH hybrids confirmed that the neo integration was in a human chromosome. Indeed, there was very little mouse DNA transferred, as determined by probing the same Southern blot with a mouse specific repeated element (Figure 21). However, none of the markers from human chromosome 4 which were tested had been cotransferred - in particular, the markers closely linked to the HD locus.

To establish the chromosome origin of the transferred human fragments, a phage library was constructed in the EMBL3 lambda vector using Sau 3A partially digested DNA from hybrid GH8000R#3. In 10^4 phage plated from this library, 50 were selected on a first screening with labeled human genomic DNA as containing human inserts. Of those 50, only 2 (411 and 415) actually contained human inserts, based on second and third screenings. From the estimates of the human DNA content of GH8000R#3, two clones is in the range of the expected frequency of phage containing

human inserts in a library made from this DNA. A Sau 3a fragment lacking repeated elements was isolated from clone 411 and used as a probe on a panel of DNAs from various somatic cell hybrids containing human chromosomes. The concordancy analysis with human chromosomes is underway.

FIGURE 15: Schematic diagram depicting the method used to create deletion hybrids of chromosome 4, by the method of Goss and Harris .

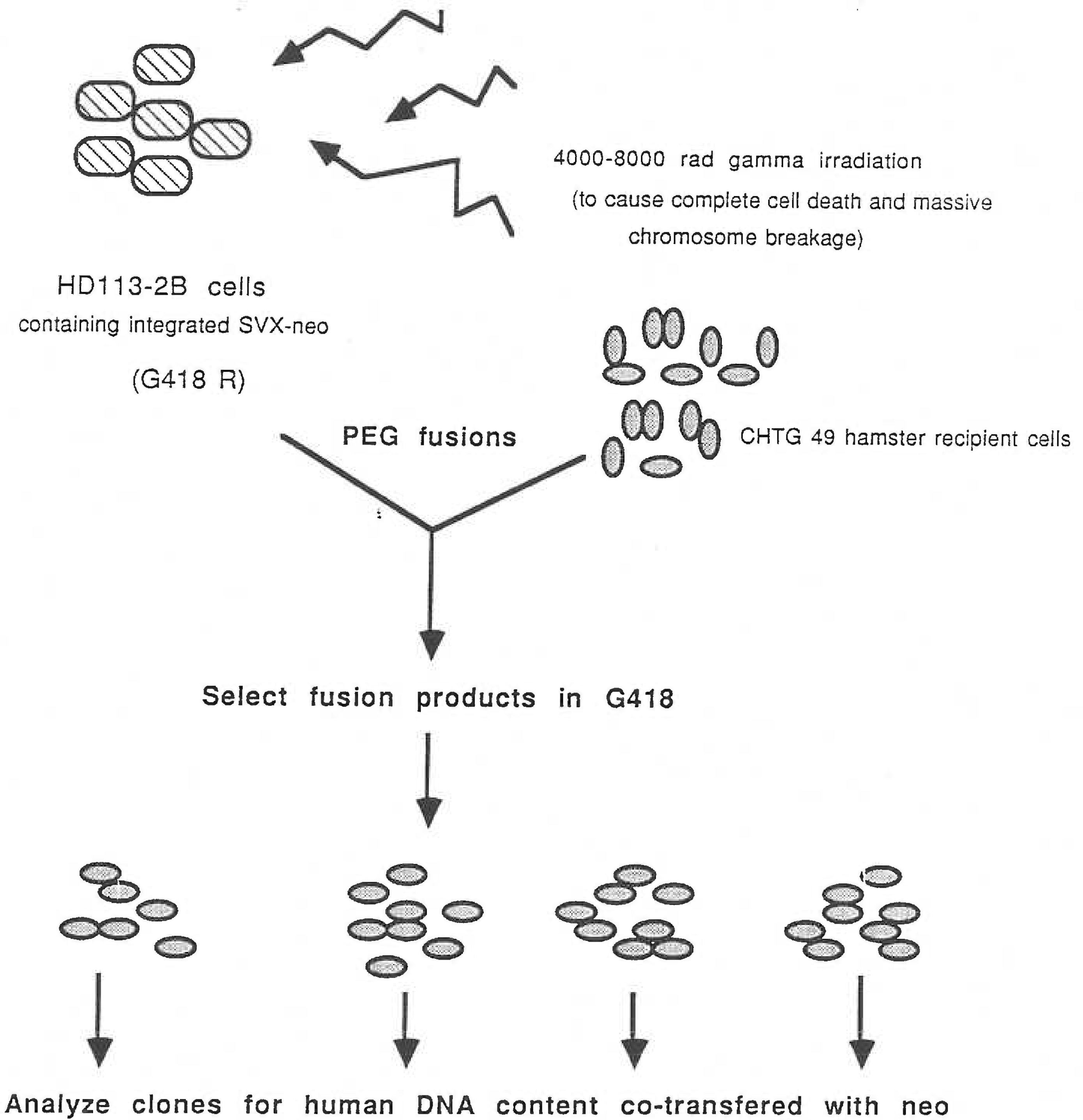

Goss-Harris hybrid construction
4000-8000 rad gamma irradiation
(to cause complete cell death and massive
chromosome breakage)
HD113-2B cells
containing integrated SVX-neo
(G418 R)
PEG fusions
CHTG 49 hamster recipient cells
Select fusion products in G418
Analyze clones for human DNA content co-transfered with neo

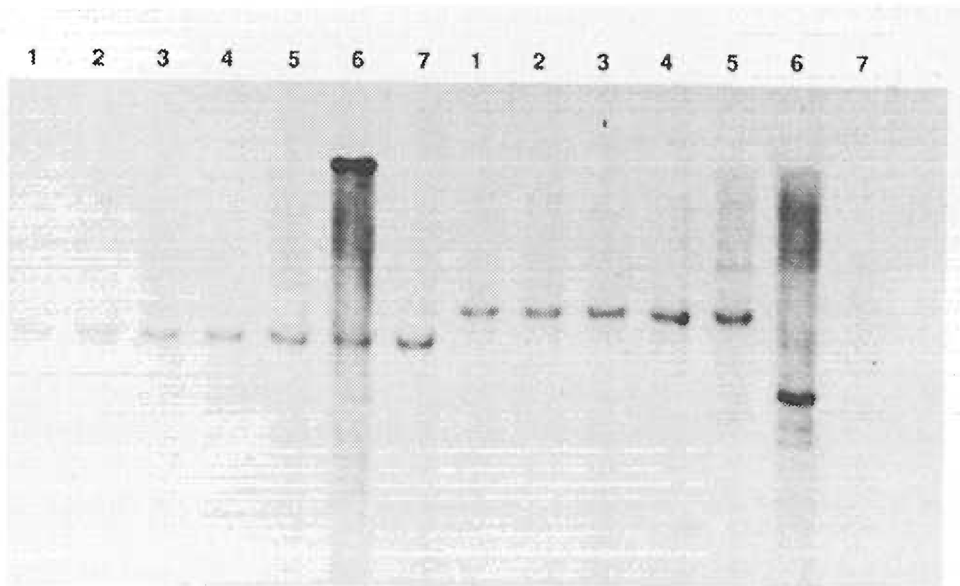
FIGURE 16: Autoradiogram of Southern blot analysis of DNAs digested with Hind III, Bam HI or Eco RI from donor HD113/2B (lane 1); recipient CHTG49 hamster (lane 2); and deletion hybrids GH4000R#1 (lane 3); GH4000R#2 (lane 4); GH4000R#3 (lane 5); GH8000R#2 (lane 6); and GH8000R#3 (lane 7) cells hybridized with the RAF-2 cDNA in pBR322 plasmid. The RAF2 specific bands are denoted with arrows - Hind III fragment (lower arrow), Bam HI fragment (upper arow) and Eco RI fragment (middle arrow). All other bands are the SVX(neo) proviral fragments detect by pBR322. λ Hind III size markers are at left margin.

1 2 3 4 5 6 7 1 2 3 4 5 6 7 1 2 3 4 5 6 7

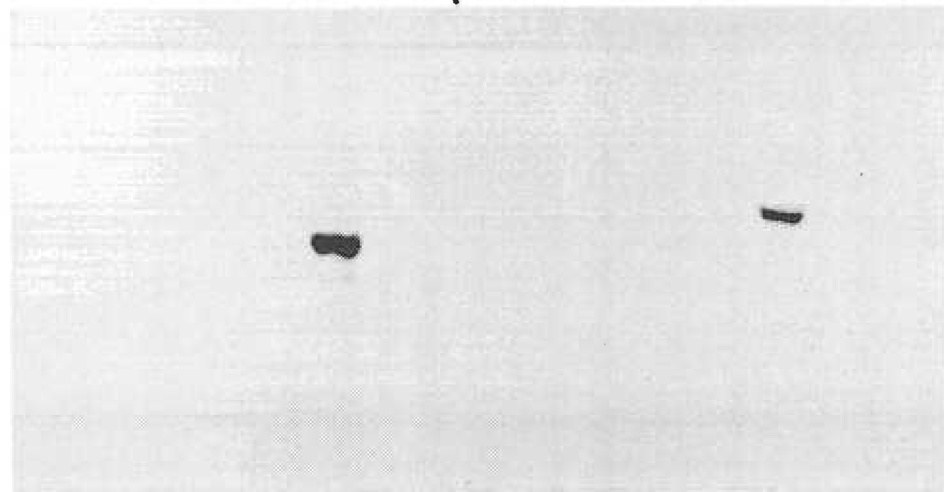
Hind III | Bam H-I | Eco RI

FIGURE 17: Autoradiogram of Southern blot analysis of DNAs digested with Hind III or Eco RI from deletion hybrids GH4000R#1 (lane 1); GH4000R#2 (lane 2); GH4000R#3 (lane 3); GH8000R#2 (lane 4); GH8000R#3 (lane 5); donor HD113/2B (lane 6) and recipient CHTG49 (lane 7) cells hybridizied to anonymous probes C4H (panel A) or D4S10 (panel B).

A



B



Hind III | Eco RI

FIGURE 18: Autoradiogram of Southern blot analysis of DNAs from deletion hybrids hybridized with the low order repeat p11L26. Content of lanes is exactly as in Figure 16.

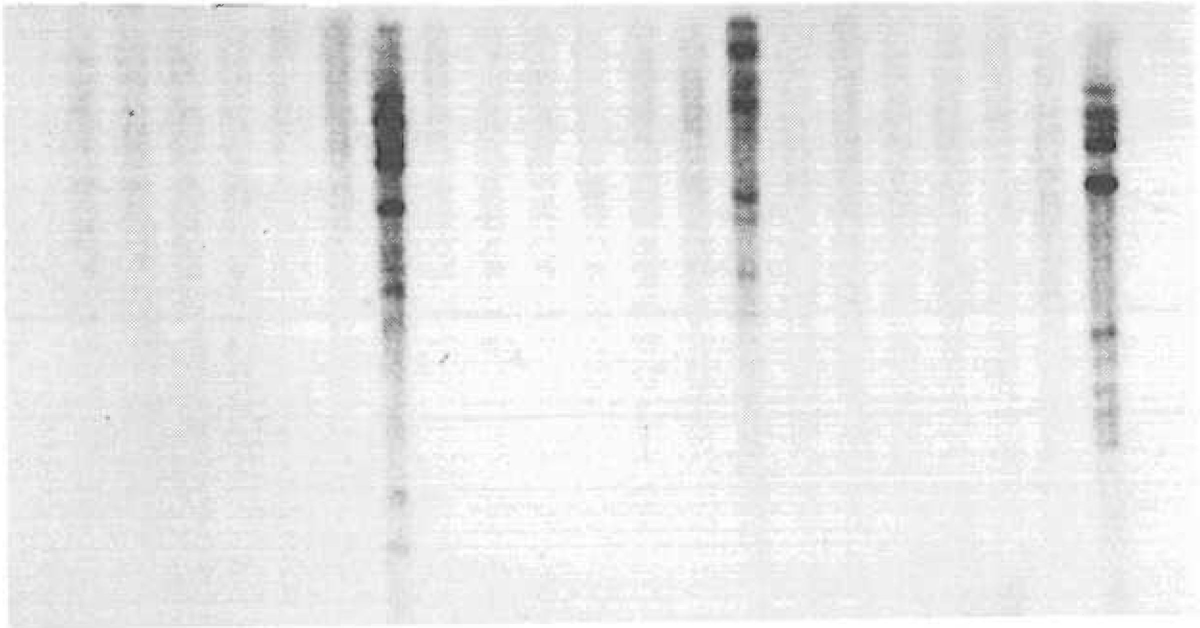

1 2 3 4 5 6 7 1 2 3 4 5 6 7 1 2 3 4 5 6 7
Hind III
Bam HI
Eco RI

FIGURE 19: Autoradiogram of Southern blot analysis of DNAs from recipient CHTG49 (lane 1); donor HD113/2B (lane 2); deletion hybrids GH4000R#1 (lane 3); GH4000R#2 (lane 4); GH4000R#3 (lane 5); GH8000R#2 (lane 6); GH8000R#3 (lane 7) cells hybridized to a 300 bp fragment detecting BLUR repetetive sequences. Only 1 ug of HD113/2B was loaded in lanes 1.

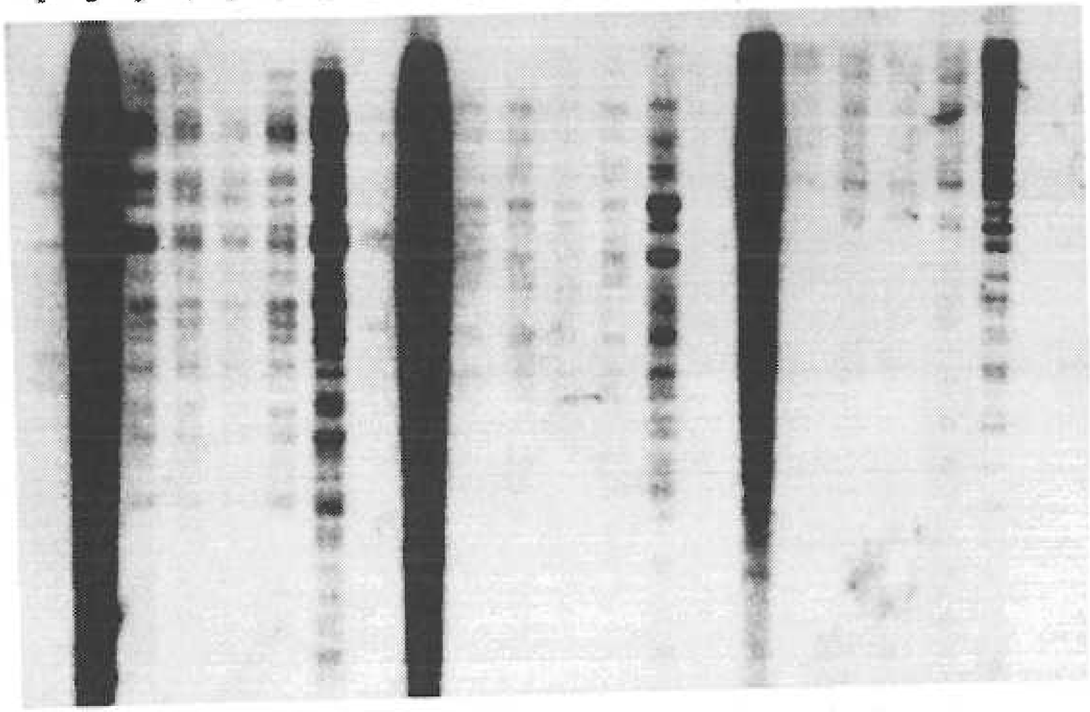
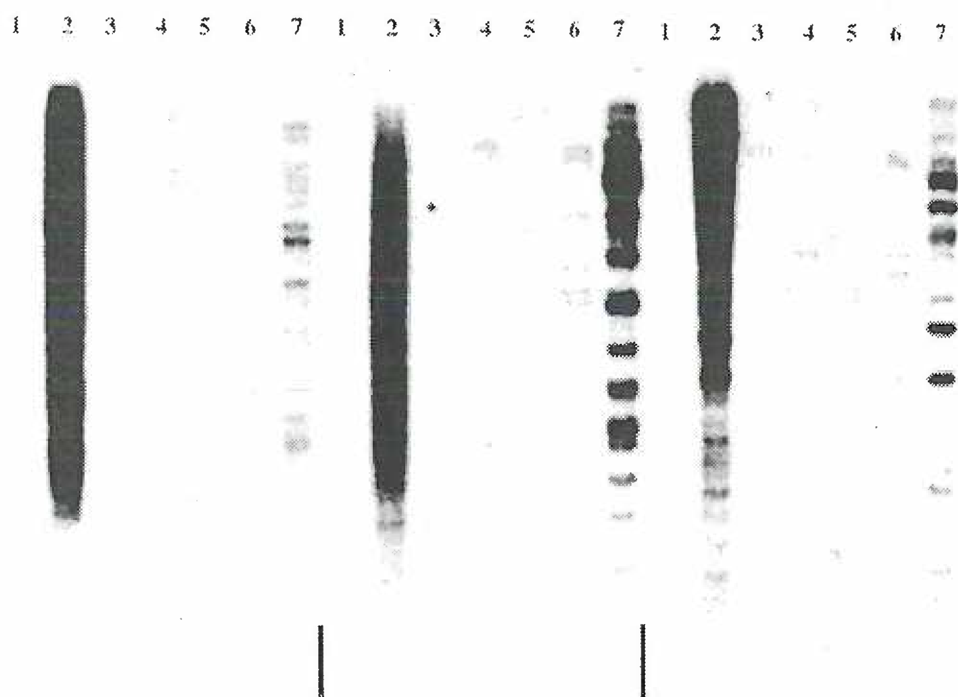
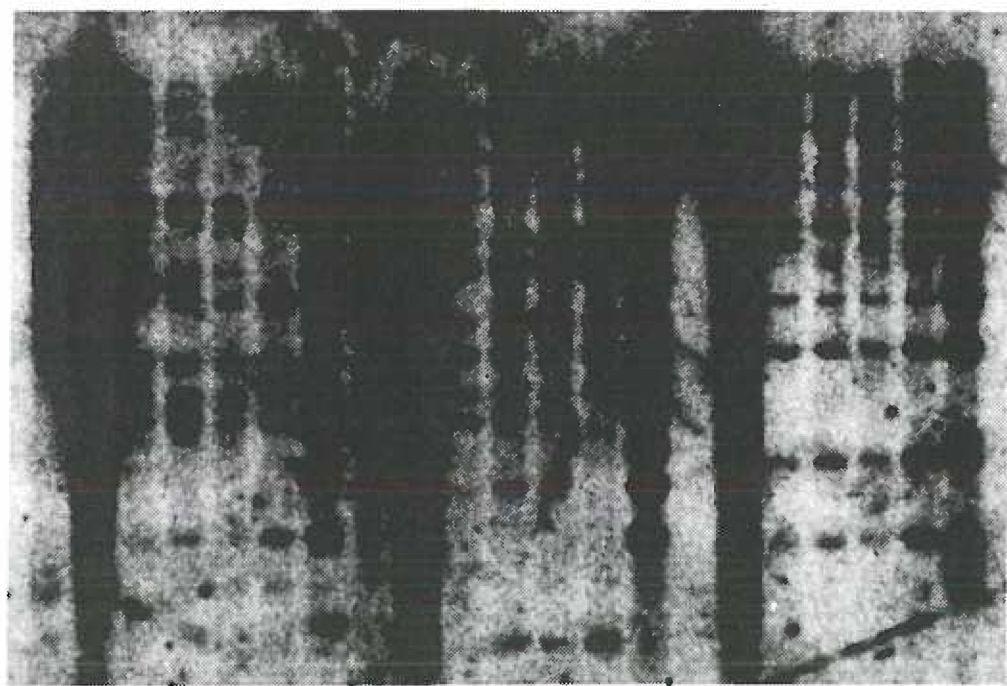

1 2 3 4 5 6 7 1 2 3 4 5 6 7 1 2 3 4 5 6 7
Hind III
Bam HI
Eco RI

FIGURE 20: Autoradiogram of Southern blot analysis of DNAs digested with Hind III, Bam HI and Eco RI from recipient CHTG49 (lane 1); donor HD 113/2B (lane 2); deletion hybrids GH4000R#1 (lane 3); GH4000R#2 (lane 4); GH4000R#3 (lane 5); GH8000R#2 (lane 6); and GH8000R#3 (lane 7) cells hybridized to $T_{\varepsilon}GC$ repetetive sequence. Short autoradiography (panel A) permits detection of individual bands in DNA from donor HD 113/2B containing intact chromosome 4 (1ug/lane) whereas long exposure (panel B) permits detection of less strongly hybridizing bands which characterize each hybrid.



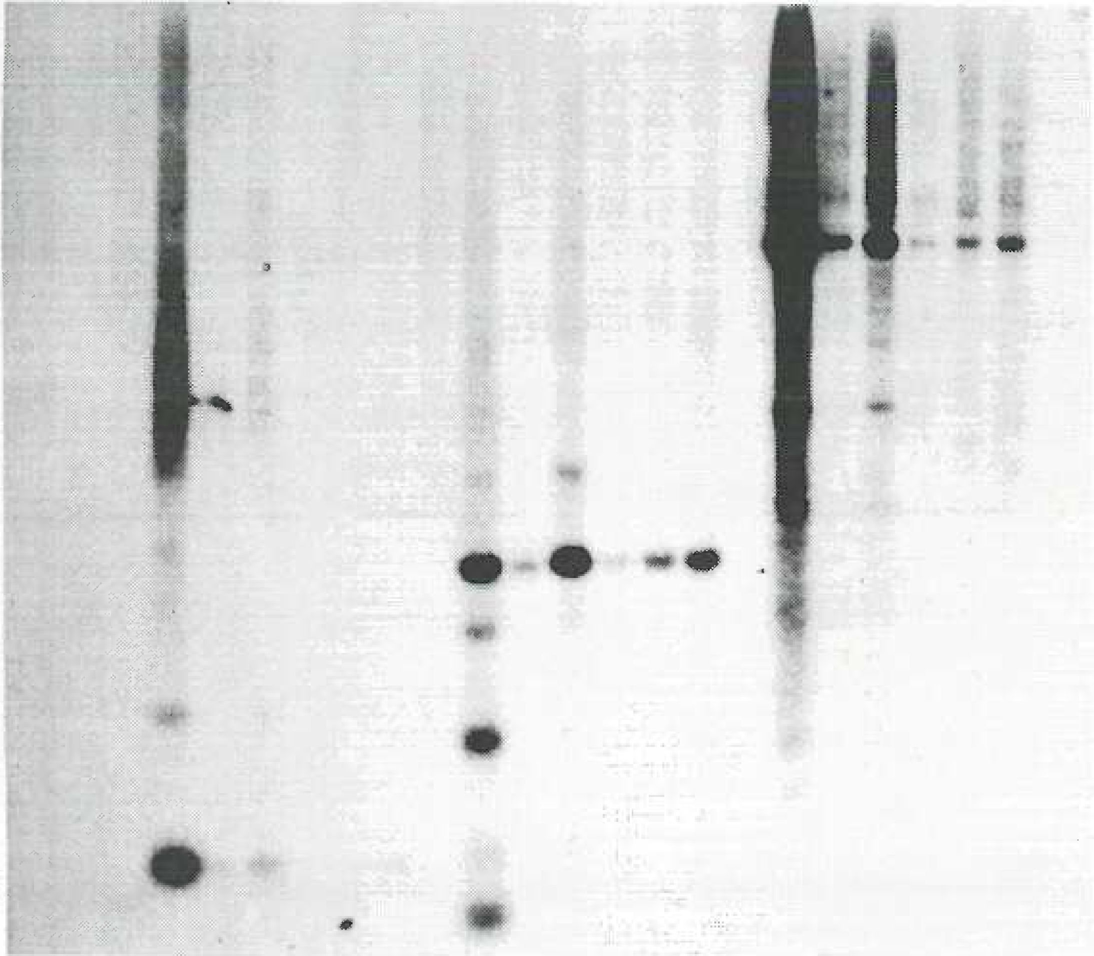
A



B

Hind III | Bam HI | Eco RI

FIGURE 21: Autoradiogram of Southern blot analysis of identical blot depicted in Figure 18 hybridized with mouse specific repetetive element Tu96. Only 1ug of HD113/2B DNA was loaded per lane whereas 10ug/lane of DNA was loaded from each hybrid.


1 2 3 4 5 6 7 1 2 3 4 5 6 7 1 2 3 4 5 6 7
Eco RI
Bam HI
Hind III

DISCUSSION
CHAPTER 3

In this chapter we describe the construction of sub-chromosomal hybrids using the method of Goss and Harris. Southern blot analysis of the DNAs from these hybrids revealed that each hybrid contained the selected neo marker plus cotransferred human DNA. However, this human DNA did not include the known markers D4S10 or C4H which are linked to the HD gene. To establish the origin of the cotransferred sequences, a phage library was made from DNA from one hybrid and phage clones containing human DNA were isolated. We were able to map the origin of one clone by hybridizing a single copy fragment from it to a panel of DNAs from somatic cell hybrids.

One observation was that all five of the hybrids described here characteristically contain a rather limited amount of human DNA as assessed by hybridization to two human specific repetitive elements. Although we had expected to induce relatively infrequent chromosome breaks to produce large fragments for transfer, the actual outcome was consistent transfer of relatively small fragments containing the neo marker. Furthermore the size of the transferred fragments did not correlate with the dose of γ irradiation used in a particular experiment; the sizes of the fragments transferred after 4000 rad exposure were equivalent to those transferred after 8000 rad. In the original experiments by Goss and Harris using this method (1977), the PGK gene which has subsequently been mapped to Xq13 was frequently transferred with the selectable marker, HPRT, at q26-27 and, occasionally, the G6PD gene at q28 was also included. These fragments represented nearly a third of the long arm of chromosome 11. Indeed, the ability to fragment chromosomes into these rather large fragments and transfer them intact to recipient cells was the underlying principle on which Goss and Harris based the method of mapping contiguous genes by cotransfer. Other sub-chromosomal

hybrids made by this method to transfer portions of chromosome 11 also have been found to contain large portions of that chromosome, although hybrids containing very small fragments have also been observed (Lugo, T and C. Jones, personal communication). In this latter case, the dose of radiation correlated with the size of the fragments so generated; the greater the dose of gamma irradiation, the smaller the fragments which were generated and subsequently transferred.

Several explanations for this unexpected observation are possible. It is possible that the HD113/2B cell line is particularly sensitive to chromosome breakage by radiation. Then, although the 4000 and 8000 rad doses were the recommended doses for radiating rodent donor cells, these may have been sufficiently high doses to induce chromosome breaks very frequently in these cells. It is also possible that the calibration of independent Cs radiation sources is not standardized so that doses cannot be compared between machines. A final alternative explanation is that the neo marker in HD113/2B is located in a chromosome region which is unusually sensitive to chromosome breaking agents or more specifically, perhaps it is near one site which is particularly unstable. The similarity in the size of the fragments seen in all five hybrids supports this last explanation since it would predict that the most likely break in the vicinity of the marker would be at one sensitive site.

If the neo marker were in an unstable region this could also explain the secondary site of in situ hybridization seen in experiments using neo as a probe on chromosomes from HD113/2B . Since the fragment containing the neo marker is located on the very tip of 4p perhaps it is prone to translocation. Under these conditions the fragment could become associated with the other telomere (4q) in a large enough percentage of cell to be detected as a site hybridizing with neo during in situ hybridization experiments.

Regardless of the mechanism which generated the small fragments transferred into these GH hybrids, however, this fortuitous event has resulted in the isolation of a

limited region of chromosome 4 - that very near the telomere. In summary, the in situ and cos fusion rescue mapping show that the SVX(neo) provirus is located in 4p14-16. Therefore, the phage library constructed from GH8000R#3 hybrid should be a rich source of clones derived from this region of chromosome 4. Since no marker has been identified on the telomere side of the HD gene (Gilliam et al, 1987), and since a flanking marker for the HD gene is highly desirable for diagnostic purposes, this library may provide the reagent for achieving this goal.

INTRODUCTION
CHAPTER 4

Our cumulative experience with the SVX(neo) vector as a tool for introducing and mapping dominant selectable markers clarified the need for a vector possessing characteristics which would make 'rescuing' and mapping of the provirus much easier. Additionally, we anticipated the need to introduce a second selectable marker into neo^R cells. Therefore, we designed and constructed a new vector, pSP-1, which met both of these needs. pSP-1 also contains a polylinker defining recognition sequences for two restriction enzymes which cut mammalian DNA into very large (greater than 100 kb) fragments; the rationale for this is discussed below. In this chapter we describe the construction and use of pSP-1.

The dominant selectable marker carried by SP-1 virus is auxotrophism for the essential amino acid, histidine. Specifically, a 1350 bp fragment containing the His D gene of *S. typhimurium* was retrieved from SV2-His (Hartman unpublished) for use in SP-1. The His D gene is one of10 structural genes in the His operon of *S. typhimurium* which encode proteins responsible for the synthesis of histidine from PRPP (phosphoribosyl-pyrophosphate), a pathway induced only by histidine starvation conditions (Ames and Hartman, 1963; Roth et al, 1966). His D encodes histidinol dehydrogenase, the enzyme catalyzing the final step in this pathway, that is, the conversion of histidinol to histidine.

Since eukaryotic cells lack the his operon, cells starved for histidine absolutely fail to survive longer than 3 days in culture. However, eukaryotic cells can absorb histidinol, and if they also possess the His D gene under eukaryotic gene expression promoters, they can metabolize histidinol to histidine, thereby becoming auxotrophic in His-deficient culture media (Hartman unpublished). The selections forHis-

transformed cells can be performed even in media containing histidine if a sufficient concentration of histidinol is included; histidinol is toxic to cells lacking His D due to the efficient, yet erroneous, charging of tRNAs with histidinol by $tRNA^{his}$ synthetase. The $tRNA^{his}$ synthetase actually has a higher affinity for histidinol than for its native substrate. Selection for cells transformed with His D is thus rather complete; unlike selection regimens using toxic substances which can be escaped by cellular mutations affecting uptake of the toxin, the His selection, performed in the absence of histidine, cannot be complemented by any cellular mutations because histidine is absolutely required for survival.

To facilitate the cloning of integrated proviral SP-1 plus associated flanking host sequences, we included the amber suppressor tRNA gene of Mo-$MuLV^{sup}$ (Reik et al, 1985) in the 3' LTR of pPS-1. This gene (su III) has allowed facile recovery of proviruses containing it in cloning vectors marked with amber mutations, such as Charon 4A (Lobel et al, 1985 and Reik et al, 1985). Recovery is achieved by constructing a complete genomic library of DNA from infected cells in phage vectors requiring amber suppressors for lytic growth. When such a library is plated on su^- host E. Coli (eg MC1061) only those phage containing inserts with suppressor tRNA genes will produce plaques. In practice, other phage plaques are observed, presumably due to spontaneous reversion mutations in the phage vector, but the enrichment for clones containing provirus and the ease of the selection protocol make this a very effective rescue method. Host flanking sequences are co-cloned with the provirus and these can be searched for restriction fragments containing unique copy sequences. These single copy fragments can then be used to map the integration site of SP-1 provirus by hybridization to an appropriate mapping panel of somatic cell hybrids. We demonstrate in this chapter the function of su III in the recovery of proviral SP-1.

The recent advent of pulsed field gel electrophoresis for separating large DNA fragments prompted us to also include in pSP-1 a synthetic oligonucleotide sequence

defining the recognition sites for two enzymes cleaving DNA rarely: Sfi I and Nru I. We anticipated that the ability to introduce rare restriction sites into host DNA would be useful for several reasons. First, it would be possible to determine directly whether an integration of SP-1 had occured within several hundred kilobases of a target gene. This could be detected by hybridizing a Southern blot of Nru I-digested DNA from uninfected and infected cells with the target gene probe. If SP-1 integrated within the same Nru I fragment containing the target gene, then the size of this fragment, as detected by the target gene probe, would be reduced in the DNA from infected cells relative to that from uninfected cells. Secondly, if an SP-1 provirus were found to be located close enough to a gene of interest - eg. the Huntington's disease gene - to be cotransferred in Goss-Harris fragment hybrids, then teh polylinker would serve to identify an Nru I fragment which may also contain the target gene, or at least provide a starting site for chromosome jumping or linking libraries.

Third, since these recognition sequences contain several CpG dinucleotides and since Nru I in particular is sensitive to methylation of such cytosine residues, we reasoned that we would be able to assess the methylation status of SP-1 integration sites. By introducing an exogenous Nru I site into host DNA and assuming that this site would be methlyated de novo, we would be able to observe the frequency with which a retrovirus vector like SP-1 integrates into a methylated region of host chromosomes. Previous work has indicated that retrovirus integration occurs preferentially in regions of DNAse I hypersenstivity, regions associated with transcriptional activity (Breindl et al, 1984; Robinson and Gagnon, 1986; Conklin and Groudine, 1986; Rohdewohld et al, 1987; Rijaya et al, 1986). Since hypomethylation over large distances has also been shown to be a prerequisite for transcription (reviewed by Bird, 1986), the ability to observe the methylation state of proviral SP-1 would provide additional evidence for assessing the randomness of retroviral integration. We present preliminary evidence here that

proviral integration in fact occurs in hypomethylated regions of the host chromosome much more frequently than would be expected if integration were random.

MATERIALS AND METHODS
CHAPTER 4

Construction of pSP-1

pSP-1 was derived from pZIP-ZIP vector (a gift of R.C. Mulligan) and Mo-MULV^{sup} (a gift of Reik et al). pZIP-ZIP vector is pZIP-SVX(neo) vector (Cepko et al, 1984) without the neo gene fragment. The construction of pSP-1 is as described in Results and as depicted in Figure 22. Restriction digests were performed as recommended by the manufacturers. Purification of isolated DNA fragments from agarose was performed essentially as described using NaI and silica glass powder (Cepko et al, 1984). The polylinker containing recognition sequences for Nru I and Sfi I enzymes was synthesized at Integrated Genetics (Framingham, Mass.) as per specifications.

Derivation of ψ-2 Producer Cell Lines

The final pSP-1 plasmid was transfected into ψ-2 cells by first coprecipitating 1 µg of pSP-1 and 9 µg of carrier NIH 3T3 DNA with 0.125 M $CaCl_2$ in transfection cocktail. This precipitate was incubated on ψ-2 cells in 10ml DME (+ 10% CS) for 4 hours at 37°C. Negative controls were performed exactly in parallel except that only NIH 3T3 DNA was used in the transfection cocktail. DMSO shock was performed with 10% DMSO in DME (no CS) for 20 minutes at 37°C followed by two washes with DME and finally feeding transfected cells with complete media for 12-18 hours. After recovery feeding, cells were split 1:20 into selective media (DME + 10% CS + 1 mM, 2 mM or 5mM histidinol). We observed that 2 mM histidinol was optimal for initial selection of NIH-derived cell lines such as ψ-2 cells; no non-transfected cells survive this concentration of the compound and after 6-7 days in selection, resistant colonies

could be further cultured in 5 mM histidinol. Producer clones were picked after 10-14 days growth in selection.

To establish the titers of infectious SP-1 virus produced by ψ-2 clones, 18 hr culture media from 70-80% confluent ψ-2 (SP-1) clutures was harvested, filtered through Gelman #4184 (0.2 m) sterile filter units, and applied to sub-confluent HD113/2B cells after serial dilutions from 10^{-1} to 10^{-3}. Infections were done in 3 ml of DME (+ 10% CS) in the presence of 0.8 μg/ml polybrene, for 2 hours at 37° C. Infected recipient cultures were split 1:10 24 hours later into DME (10% CS) containing 5 mM histidinol and his-auxotrophic colonies were counted after 10-14 days in selective culture.

Analysis of SP-1 Infected HD113/2B Clones

DNA was extracted from infected HD113/2B cells, digested with either Eco RI alone or double digested with Eco RI plus either Nru I or Sfi I and the resulting fragments were size separated by gel electrophoresis as described in General Materials and Methods. Southern blots of these samples were probed also as described with the 1350 bp Bam HI fragment containing the His D gene of *S. typhimurium* which had been previously subcloned into the Bam HI site of Puc 18. The His fragment could not be isolated directly from pSP-1 for use as a probe because other contaminating fragments (eg. the LTR sequences) from pSP-1 hybridize to many other sequences in NIH mouse DNA.

Construction of Genomic Library in Phage EMBL3B Vector

Proviral SP-1 and its flanking host sequences were selectively cloned via construction of a genomic library of DNA from HD113/2B (SP-1) infected clone #32 in the replacement phage vector EMBL3B (Frischauf et al 1983). The library was constructed exactly as in Chapter 3 except that vector EMBL3B replaced EMBL3.

EMBL3B vector is marked with two amber mutations in phage genes A and B, both of which are required for lytic phage growth. After titering the library on LE392 bacteria, 1 x 10^5 phage were plated on a su$^-$ baterial strain, MC1061 (Casadaban et al, 1980) to select for phage carrying proviral SP-1 and its su III (amber supressor) tRNA gene. Thirteen plaques were picked from 13 separate plates and the phage DNA from each was analyzed by restriction digest and Southern blotting using His D as a probe for phage containing SP-1 provirus.

RESULTS
CHAPTER 4

The retroviral vector pSP-1 was derived from the previously described pZIP vector (Cepko et al, 1984). pZIP is a murine leukemia virus (MuLV) transcription unit derived from an integrated Moloney MuLV provirus (Hoffman et al 1982) plus pBR322 sequences necessary for the propagation of the plasmid vector DNA in E. coli. The specific sequences retained in pZIP are as described by Cepko . In place of the retroviral sequences encoding the gag-pol and env polypeptides, two unique restriction endonuclease cleavage sites (Bam HI and Xho I) had been previously inserted, to permit the expression of sequences introduced into the vector from either the full-length or spliced retroviral transcripts (Weiss et al 1982).

As schematically depicted in Figure 1, a 1350 bp Bam HI fragment containing the His D gene of *Salmonella typhimurium* *(S. typhimurium)* was inserted into the unique Bam HI site of pZIP, just 5' of the 3' splice site for subgenomic spliced messenger RNAs. Thus, the His D gene transcript is generated as a full length mRNA from the transcription start site in the 5' LTR.

The Xho 1/Sac 1 fragment of pZIP which includes the p15E gene of M-MuLV through the Sac 1 site in the 3'LTR was replaced with the comparable fragment from a Mo-MuLV^{sup} vector constructed by Reik et al (1985). A 200 bp synthetic bacterial suppressor tRNA gene had been inserted in the 3' LTR of Mo-MULV^{sup} included in this fragment. The supressor gene was excised originally from πVX plasmid with Eco RI and ligated with Bam H1 linkers. The derivative Bam H1 fragment was then inserted into a Bam H1-modified Sau 3A site in the 3' LTR.

The vector thus far described, pSP-1*, satisfied two of the three requirements for an optimal retroviral vector for somatic cell genetic usage. To meet the third requirement of introducing recognition sites for rare cutter restriction enzymes a polylinker was designed, synthesized and cloned into the unique Xho I site in the 3' end of pZIP. The polylinker contains the recognition sequences for Sfi I and Nru I restriction enzymes flanked by Bgl II sites as cloning ends. pSP-1* was linearized with Xho I and the Xho I sticky ends were flushed to blunt ends by Klenow polymerase enzyme. Bgl II linkers were then blunt end ligated onto the ends and the plasmid was recircularized in the presence of an excess of polylinker. The resulting pSP-1 construct containing the polylinker is depicted in Figure 2.

Plasmid pSP-1 DNA was introduced into ψ-2 producer cells by $CaPO_4$ mediated gene transfer and producer clones were picked after 14 days growth in 5mM histidinol. The virus-containing culture media from these clones was used to infect HD113/2B cells. Serial dilutions of the virus media were tested to establish virus titers for each clone as described in Materials and Methods. The titers from 7 producer ψ-2 (SP-1)cell clones ranged from 2×10^4 to $>5 \times 10^6$ infectious particles per ml of culture media (Table 4). Producer ψ-Am clones were created by this same method although these virus stocks will not be described further here.

To demonstrate the efficacy of the special features of SP-1 as a retroviral vector, HD113/2B cells were infected with high titer SP-1 stocks and, after selection in 5mM histidinol, 60 independent clones were picked, expanded into large cultures, frozen and a subset were further analyzed. We were interested again in inserting the vector into the human chromosome 4 in HD113/2B cells. To this end, metaphase chromosome spreads were prepared and an initial13 of these preparations were R-band stained as described. By analyzing approximately 10 metaphase spreads from each clone, the percentage of cells which retained a human chromosome

4 was scored. Since these clones were grown in the absence of G418, the human chromosome was expected to have been randomly segregating through the many cell doublings which were passed during selection for His-auxotrophism, expansion and culturing of these clones for chromosome harvests. As presented in Table 5, half of the clones retained chromosome 4 in less than 50% of cells and only two clones retained chromosome 4 in greater than 85% of cells: clones K and 32. Clone 32 retained 2-3 copies of chromosome 4 in most cells, as seen in the parent HD113/2B cell line. This clone was therefore selected as a good candidate for harboring an integrated SP-1 provirus vector in chromosome 4.

To map the integration site of SP-1 in clone 32, DNA from this clone was used to construct a phage library in the EMBL3B cloning vector (Frischauf et al 1983). As demonstrated previously, the Su III tRNA gene in the LTR of SP-1 permits the selective cloning of integrated provirus plus flanking host sequences in a phage vector such as EMBL3B. EMBL3B contains two amber mutations in the A and B lambda genes which are required for lytic phage growth in E.Coli strains lacking supressor genes (eg. MC1061). However, these mutations can be suppressed in recombinant phage containing proviral SP-1 via its Su III gene. Thus, the only plaques which will be formed on a lawn of MC1061 (su^-) from an EMBL3B library of DNA from an SP-1 infected clone will be those containing proviral SP-1. Since EMBL3B is a replacement vector requiring 9-20 kb inserts for phage packaging, additional host flanking sequences will also be included with the 4.0 kb of SP-1 provirus (Figure 24).

From an initial 1 x 10^5 phage particles plated on MC1061, 13 plaques were picked from 13 seperate plates. DNA from these clones was digested with several restriction enzymes, size seperated by gel electrophoresis and blotted onto Zetabind membrane. However, none of these clones were subsequently found to contain the SP-1 provirus DNA, as ascertained by hybridization with the 1350 bp His D gene

fragment. Apparently, the rate of spontaneously ocurring mutations which suppress the amber mutations in EMBL3B is rather high: at least 1.2 x 10^{-4} mutations per phage clone.

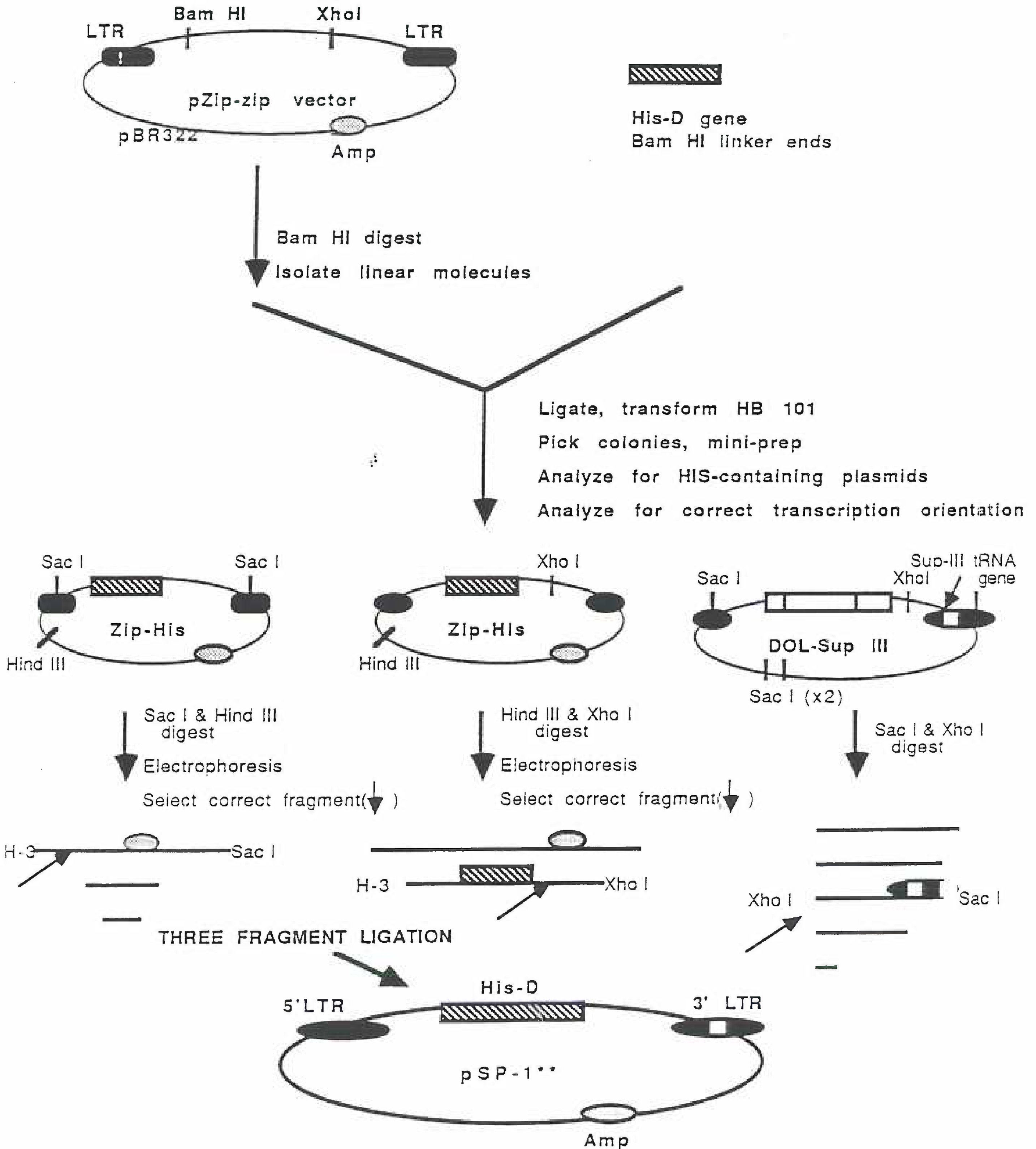
The ability of SP-1 to introduce recognition sequences for Nru I and Sfi I restriction enzymes was tested by digesting DNAs from 22 SP-1 infected HD113/2B clones with Eco RI, and Eco RI plus either Nru I or Sfi I. The resulting fragments were isze separated by gel electrophoresis and anlyzed by Southern blotting using the His D gene as a probe. In each clone, the Eco RI fragment detected by His (lane a) is reduced in size after double digestion with Sfi I (lane b), indicating that the polylinker in SP-1 does indeed introduce a new Sfi I site into host DNA by integration of the provirus (lanes c, Figure 25). Similarly, double digestion with Nru I also reduced the size of the fragment detected with His in most clones, with the exception of clones J and 9. In both of these clones, although the Sfi I site in the polylinker was completely digested, the Nru I site was completely resistant to digestion. This experiment was repeated with the same results indicating that the Nru I digestion is complete in each sample. Southern analysis of another two clones gave unexpected results. Double digestion of DNA from clone 29 with either Nru I or Sfi I generated two fragments detectable with the His D gene probe rather than the usual one fragment (Figure 25, clone 29, lanes b and c). Furthermore, these two fragments appear to be present in equal molar amounts; that being approximately half of the molar amount in the original Eco RI fragment. The most plausible explanation for these results is that clone 29 contains two tandem copies of SP-1, alligned head to tail, on an approximately 20 kb Eco RI fragment (Figure 26A). The smaller, 4.5 kb fragment is thus an Nru I/Nru I or Sfi I/Sfi I fragment which spans the distance between the two polylinkers plus 600 bp of intervening host DNA, and the 7.0 kb fragment is an Eco RI/Nru I or Eco RI/Sfi I

fragment from the 5' Eco RI site in host DNA to the first polylinker site in the 5' copy of SP-1.

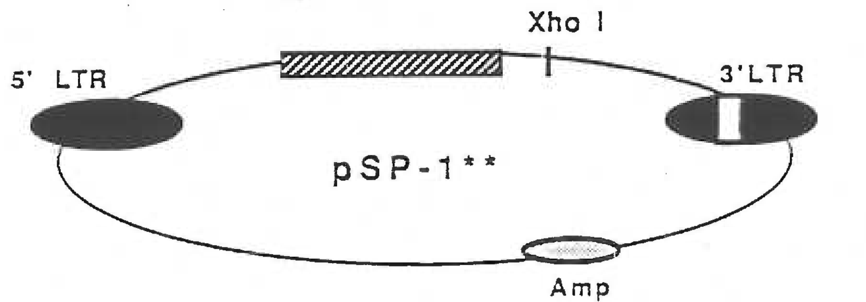
Finally, double digestion of DNA from clone 25 with either Sfi I/Eco RI or Nru I/Eco RI each resulted in single His-containing fragments smaller than their original Eco RI fragments, as in most clones,but they were of different sizes: the Eco RI/Nru I fragment was 5.0 kb and the Eco RI/Sfi I fragment was 4.3 kb. Apparently, and as schematically depicted in Figure 26B, SP-1 integrated within 1.6 kb of an endogenous Sfi I site so that the His-containing fragment in lane b of Figure 3 (clone 25) is an Sfi I/Sfi I fragment rather than the usual Eco RI/Sfi I fragment.

FIGURE 22: Scheme for construction of pSP-1 retroviral vector. See Results and Materials and Methods for details.

Construction scheme for pSP-1



Construction of pSP-1 (cont)



Xho I digest
Isolate linear molecules

Ligate to Bgl II linkers

Ligate to polylinker designed to contain restriction sites as follows:
Bgl II - Mlu I - Not I - Sac II - Sfi I - Nru I - Bgl II

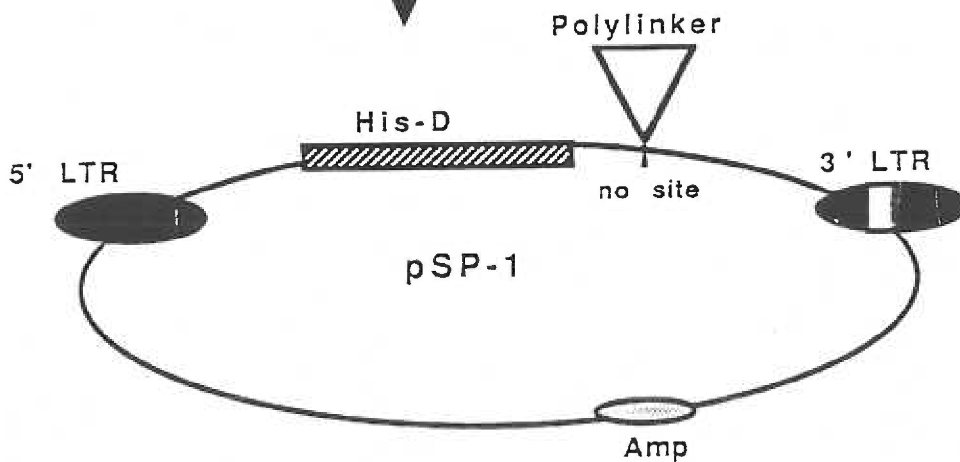
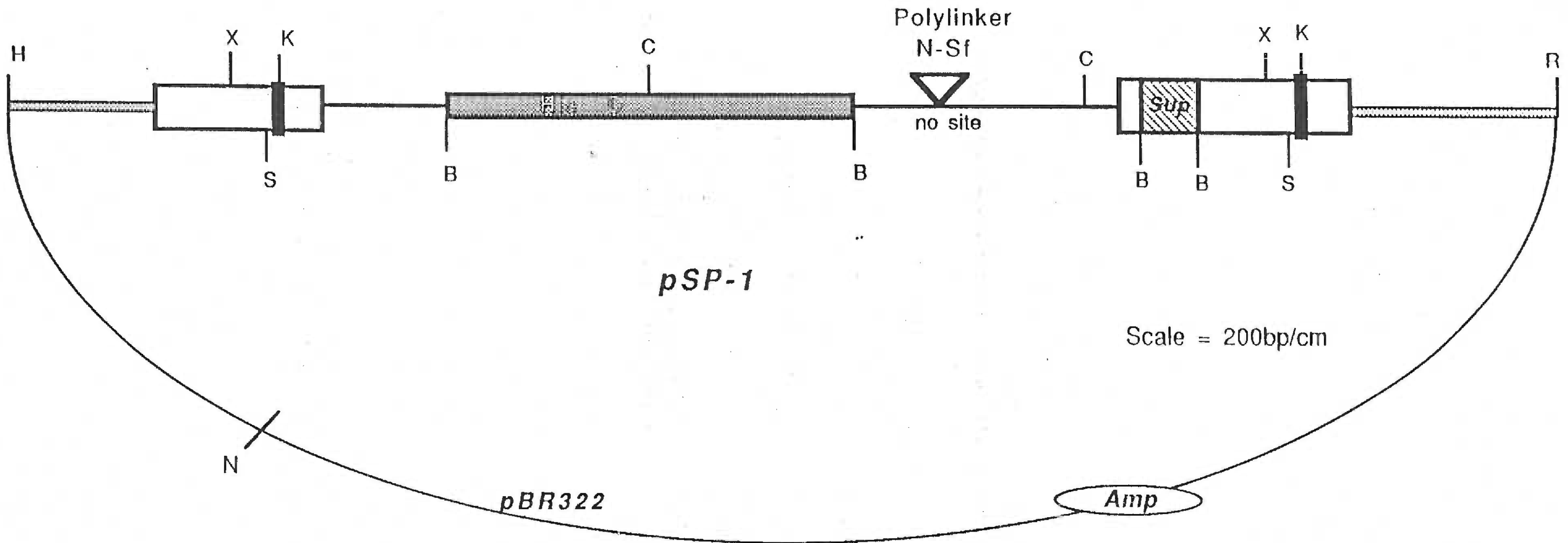


FIGURE 23: Restriction map of pSP-1 retrovirus vector.



C = Cla I
B = Bam HI
X = Xho I
S = Sac I
H = Hind III
R = Eco RI
K = Kpn I
Xb = Xba I
N = Nru I
Sf = Sfi I

Table 4: Virus titers produced by clones of ψ-2 cells transfected with pSP-1. Titers were determined on HD113/2B cells and calculated from the number of His-auxotrophic colonies surviving in 5mM histidinol 14 days after infection.

ψ-2 CLONE	TITER (infectious particles/ml)
F	1.5×10^6
G	2.4×10^6
H	1.9×10^6
I	2.0×10^4
J	3.6×10^6
K	4.6×10^6
L	5.0×10^6
:	

TABLE 5: Karyotype analysis of SP-1 infected clones of HD113/2B hybrid. Chromosomes were stained for R-banding and observed by fluoresence microscopy for enumerating the human chromosome content of each metaphase spread.

Karyotype analysis of SP-1 infected clones of HD113/2B hybrid

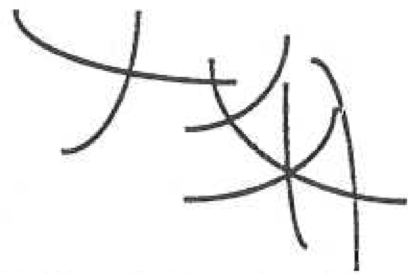
Clone Name	# Cells w/o human chromosome 4	# Cells w/ human chromosome 4	% Cells w/4
HD113/2B-SP-K	1	9	90%
-J	3	10	77
-I	10	0	0
-H	2	8	80
-G	2	8	80
-39	2	10	83
-28	14	1	6
-33	6	1	14
-32	2	13*	86
-5	4	11	73
-3	12	3	20
-23	19	1	5
-22	6	9	60
-4	8	7	47

* Most cells have 2 or 3 copies of human chromosome 4

FIGURE 24: Selective cloning of proviral SP-1 DNA plus flanking host DNA sequences via su III is achieved by generating a genomic library in the EMBL3B phage vector and plating it on host E.Coli which are su^-.

Rescue of Integrated SP-1 from Infected Cells via Suppressor Selection

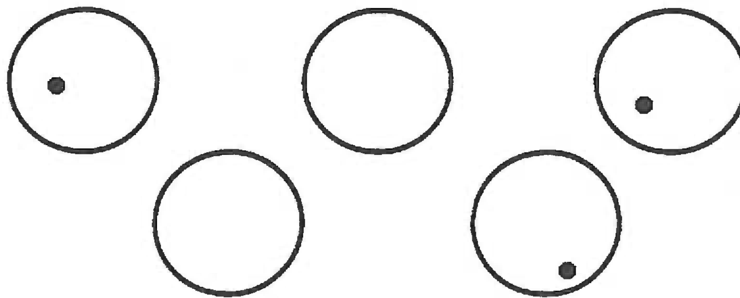
EMBL3B vector DNA cut with Bam HI
(Note A and B amber mutations in
required host genes)



DNA from SP-1 infected cells partially digested
with Sau3A and size selected

Ligate vector to inserts, package into phage particles

Plate on E. Coli strain MC 1061 (Suppressor minus)

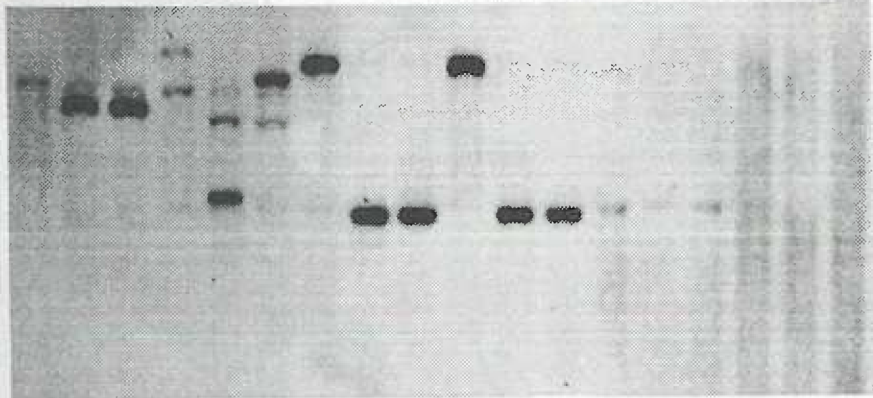


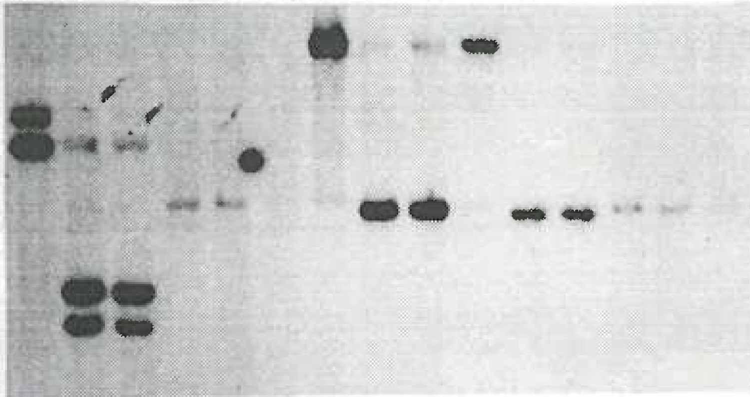
Only phage containing SP-I with Su-III to suppress amber mutations
will be able to form plaques on MC 1061 lawn

Pick plaques and analyze for authenticity (i.e. not spontaneous suppressor mutations)

Select phage clones which contain Histidine marker (i.e. containing SP-1)

FIGURE 25: Autoradiograms of Southern blot analysis which demonstrates the introduction of new Sfi I and Nru I sites into host chromosomes by integration of SP-1 provirus. Southern blots of DNAs from infected clones of HD113/2B digested with either Eco RI (lane a), Eco RI and Sfi I (lane b) or Eco RI and Nru I (lane c) were hybridized to the His D gene fragment.


a b c a b c a b c a b c a b c a b c
CLONE: I J S W HD113/2B HELA


a b c a b c a b c a b c a b c
CLONE: 30 32 37 39 G

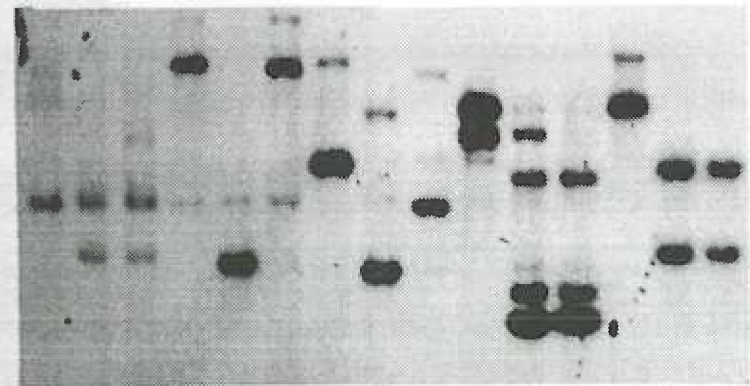
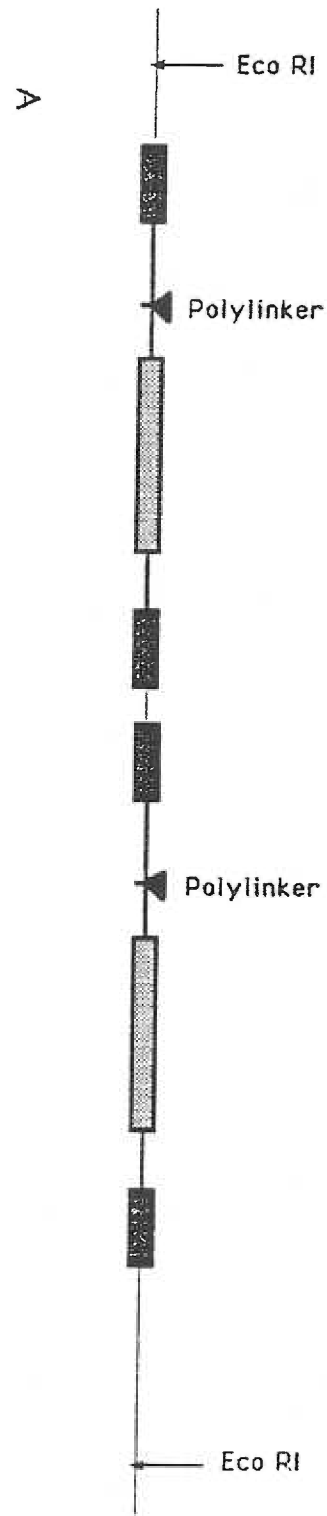
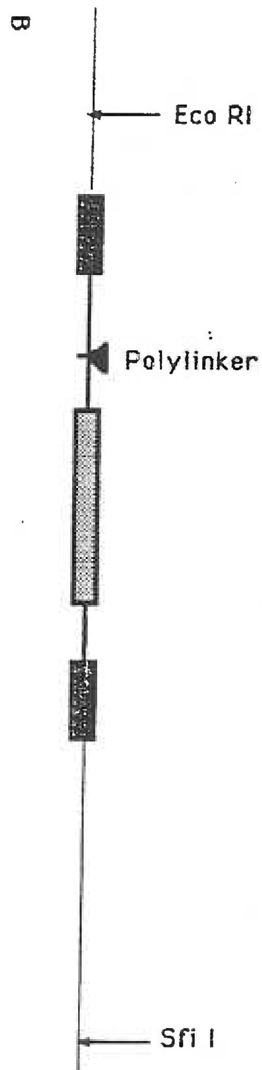

a b c a b c a b c a b c a b c
CLONE: 8 9 25 27 29

FIGURE 26: Proposed models of anomalous integrations of SP-1 proviruses in two seperate infected cell lines, as described in text. Diagram A depicts head to tail double integration in clone 29 and diagram B depicts integration of SP-1 very close to endogenous Sfi I site in clone 25.


A
Eco RI
Polylinker
Polylinker
Eco RI


B
Eco RI
Polylinker
Sfi I

DISCUSSION
CHAPTER 4

We have constructed a new retroviral vector for use in somatic cell genetics. SP-1 virus, a replication defective derivative of the pZIP vector, has three unique characteristics: it carries the dominant selectable marker, His D, which confers auxotrophism for histidine on infected host cells; an su III tRNA gene in the 3'LTR of SP-1 allows selective cloning of the provirus plus host flanking sequences; and a polylinker in the 3' end of SP-1 introduces recognition sequences for the rarely cutting restriction enzymes Sfi I and Nru I.

We observed a relatively high incidence of clones which were able to grow in 5mM histidinol but which were subsequently found not to contain proviral SP-1. This result demonstrates the need to do the selection regime in histidine deficient medium because cells which become resistant to histidinol by a mechanism other than infection by SP-1 are not selected against. Although the clone 32 was originally selected as the substrate for making a rescue library in EMBL3B because it retained human chromosome 4 in a high percentage of cells suggesting that chromosome as the site of SP-1 integration, we were of course unable to rescue any provirus since none was actually present in this cell line. Karyotyping done on a later passage of clone 32 revealed the loss of human chromosome 4 from nearly 50% of cells, as would be expected by random segregation in the absence of selection. Previous observations of the HD113/2B cell line grown in the absence of G418 had indicated that the human chromosome 4 segregated slowly from these parental cells, so it is not surprising to have randomly selected a clone like clone 32 which segregated chromosome 4 particularly slowly. Other clones such as clone K are being analyzed further for possible integrations of SP-1 into chromosome 4.

We demonstrated that SP-1 introduces sites for rarely cutting restriction enzymes (Sfi I and Nru I)into host chromosomes by analyzing the DNA from 12 infected clones which had been digested with each enzyme. Southern blots of these DNAs hybridized with His D gene probe revealed that, in 12/12 clones, digestion with Sfi I cleaved the Eco RI fragment containing proviral SP-1 into two fragments. Digestion with Nru I also cleaved the Eco RI fragment in 10/12 clones, verifying that the recognition sequence for this enzyme is also faithfully transmitted by SP-1.

The restriction sites for Sfi I and Nru I in SP-1 virus provided an opportunity to assess the randomness of retroviral integration. Two observations support previous suggestions that retroviral integration occurs preferentially in regions associated with gene expression : (1) 10/12 Nru I sites introduced by SP-1 were digested by Nru I, despite the sensitivity of Nru I to methlyation of CpG residues in its recognition sequence, and (2) in only 12 integration sites analyzed, one site was within 1.6 kb of an endogenous Sfi I site although Sfi I sites are rare and are generally seperated by several hundred kilobases in mammalian DNA.

The mechanism of retrovirus integration is largely unknown. Shortly after the appearance of circular forms of the DNA genome in host nuclei, recombination between the host and viral genomes is promoted by products of the viral pol and gag genes (reviewed by Goff) and also in part by unidentified host gene products. Integration results in duplication of 3 to 6 base pairs of the target host DNA, the number of bases being characteristic of different virues (Hughes et al, 1981a; Major and Varmus, 1981; Varmus and Swanstrom, 1982). Recombination between the host and viral genomes occurs at a specific site within the long terminal repeat of the viral genome, usually within 2 base pairs of the termini (Colicelli and Goff, 1985). In contrast, acceptor sites in the host genome appear to be locally random, as demonstrated by the failure to identify a consensus sequence in multiple cloned integration sites (Hughes et al 1981b). However, because many retroviral integrations have been observed to

occur near DNAse I hypersensitive sites of expressed genes, several investigators have hypothesized that integration occurs preferentially in transcribed regions (   ). Indeed, of 13 independent integration sites examined in mouse germ line and tissue culture cells, all but one integration was discovered to have occured within a few hundred base pairs of a DNAse I hypersensitive site (See Introduction, this chapter). DNAse I hypersensitive sites have been well correlated with regions of open chromatin conformation and this conformation is thought to be a prerequisite for transcription (Conklin and Groudine, 1984). Notably, four of the 13 sites just described were non-selected for phenotypic expression, i.e. not selected by association with increased or decreased expression of a cellular gene.

Retroviral vectors have been used for insertional mutagenesis, most thoroughly by Jaenisch and coworkers to isolate host genetic loci controlling embryonic development in the mouse (Soriano et al,1987 ; Jaenish et al, 1983). In analyzing 48 proviral insertions into germ line DNA they noted that fully 4% of insertions resulted in lethal mutant phenotypes, a percentage remarkably similar to that observed in similar experiments using other transposable elements in yeast and Drosophila. This is a surprising observation because if insertion of transposable elements were truly random then the expected frequency of insertional mutations would be determined by the size of the target, i.e, .the percentage of the genome occupied by coding sequence. Since these three organisms have widely disparate amounts of non-coding sequence, this similarity in frequencies of insertional mutations seems to preclude a random insertion mechanism.

Thus, the cumulative data strongly suggests that integration of retroviruses occurs preferentially in regions of open chromatin conformation and perhaps more precisely, in transcriptionally active regions. Another characteristic of transcriptionally active regions is that they contain HTF (Hpa II Tiny Fragments) islands which are highly enriched for unmethylated CpG residues. HTF islands have been

found at the 5' and 3' ends of most housekeeping genes (reviewed by Bird 1986) and also near some tissue specific genes . Recently, sites for some rare-cutting enzymes have been discovered in clusters, or islands, near transcribed genes (Brown and Bird, 1986; Burmeister and Lehrach, 1986; Estivill et al, 1987). Indeed, 3 of 4 random cosmid clones containing sites for one of the rare cutter enzymes, Sac II, were found to also contain transcribed sequences (Lindsay and Bird, 1987) indicating that these sites are preferentially located near transcription units.

We selected two rare-cutter enzyme sites to be transmitted and inserted by SP-1, those for Nru I and Sfi I. Nru I enzyme does not cleave recognition sites containing methylated CpG residues, whereas Sfi I enzyme recognizes a sequence which does not contain CpG residues for methylation. Sfi I sites occur with much greater frequency than would be predicted from the base composition of its recognition sequence (see complete digest in Figure 14) and are extremely non-randomly distributed in the genome with several sites occurring in one 'island' and these clusters being seperated by an average of 500 kb (C. Smith unpublished). The frequency of Nru I sites is approximately 2.7 x 10^{-4} sites/genome in bulk DNA and 1 x 10^{-4} sites/genome in island DNA. Only 27% of Nru I sites occur in islands (Linday and Bird, 1987). These unmethylated Nru I sites (cleavable) are apparently distributed randomly in the genome as reflected in the regular array of fragment sizes ranging from > 1 megabase to < 100 kb. in a complete Nru I digest of genomic DNA (Figure 14).

The choice of these two recognition sites therefore allowed us to address three issues: (1) are these sites, introduced by retroviral integration, methylated de novo?, (2) with what frequency are they de novo methylated? and (3) with what frequency do integrations occur near other Sfi I sites, these being most likely to occur in islands?

We observed that 2/12 Nru I sites introduced by SP-1 integration were resistant to digestion, although the same site was digested to completion with Sfi I. We conclude that these two sites were methylated de novo. To exclude the possibility that

the proviruses could have induced methylation or de-methylation of sequences in and around their integration sites we would have to examine the pre-integration sites with probes isolated from each site. Other proviral insertions have induced changes in chromatin organization very near their integration sites (Breindl et al 1984, Schubach and Groudine, 1984), and proviruses introduced into embryonic EC mouse cells have been observed to be frequently methylated and thereby not expressed (Jahner et al, 1982; Jahner and Jaenisch, 1985). Additionally, the prevalence of unmethylated state of SP-1 integration sites could possibly be explained by the fact that we selected for expression of the His D gene in isolating these integrations. However, this apparently does not systematically select against integrations in non-expressed, or non-methylated sites since the two clones containing methylated Nru I sites were selected in the same way (clones 9 and J).

Finally, the observation of one SP-1 integration, among only 12 integrations examined, within 1.6 kb of an endogenous Sfi I site also suggests that integration may have occurred preferentially near HTF islands. Since Sfi I sites occur in clusters every 300-500 kb (1×10^4 sites per genome) and most frequently in HTF islands (Hereditary Disease Foundation, 1986), we would have expected only one in 10^4 random integrations to occur near an Sfi I site.

These data are consistent with a model for retrovirus integration which includes a preference for host chromosome sites which are locally hypomethylated, and concordantly also contain HTF islands. As both of these characteristics have been associated with transcribed sequences, as have DNAse I sites observed near other retroviral integrations, we would predict that further experiments will demonstrate a prevalence of trancribed sequences in the vicinity of these integrated retroviruses.

INTRODUCTION
CHAPTER 5

In Chapters 1 through 4 we described the development of tools which permit selective and directed isolation of any particular chromosome region of interest. We have demonstrated the feasibility of isolating a chromosome region containing a target gne that is identifiable only by its phenotypic expression; this is done by first inserting a dominant selectable marker near the target and then cotransferring the selectable marker and its surrounding region into genetically heterologous cells. Because the transfer of large chromosome fragments is likely to preserve native chromatin structure and also regulatory elements which may be quite distal to the gene, this approach should be applicable to the isolation of genes which are refractory to isolation by DNA-mediated gene transfer. If the target gene encodes a cell surface marker this approach is particularly facile, as described previously .

One such gene which has failed to be reproducibly transferred to heterologous cells by DNA mediated gene transfer is the mouse gene encoding the ecotropic murine leukemia virus (MuLV) receptor protein (Nelson, D.L.; Handelin B.L.; Cunningham, J.M.; Baltimore D.; personal communication). As discussed in the Introduction, penetration of mouse cells by MuLVs is promoted by cell surface receptors which specifically bind viral gp70 envelope protein . The tropism of MuLVs are determined by the presence or absence of these receptors on host cells; ecotropic viruses bind to receptors present only on mouse cells, xenotropic viruses to receptors on non-mouse cells, amphotropic viruses to a distinct receptor on both mouse and non-mouse cells, and dual tropic viruses (laboratory recombinants) bind to a fourth type of receptor also found on both mouse and non-mouse cells.

The ecotropic receptor has been most studied. These receptors are saturable and are expressed on fibroblasts and lymphocytes (DeLarco and Todaro, 1976; Bishayee et

al, 1978; Choppin et al 1981; Kalyanaraman et al, 1978). That the ecotropic receptor is a protein has been demonstrated by the ability to obliterate binding activity with chymotrypsin, trypsin and pronase but not with DNAse, RNAse or neuraminidase (Bishayee et al, 1978; Kalyanaraman et al, 1978). Although the receptor is prevalent on fibroblasts (approximately 5 $x10^5$ per cell) (Bishayee et al, 1978; DeLarco and Todaro, 1976), it has been refractory to purification by cross-linking (Besmer and Baltimore, 1971), affinity chromatography (Landen and Fox, 1980) or immunoprecipitation (Robinson et al, 1980). Crude preparatoins by these methods have identified moieties which range in size from 10,000 to 190,000 molecular weight. Partial solubilization in deoxycholic acid has recently revealed that the receptor is an intrinsic membrane protein with an affinity constant for gp70 of 4 $x10^8$ M^{-1}(Johnson and Rosner, 1986).

The genetic determinant for the ecotropic receptor has been mapped to mouse chromosome 5 (Kozak and Rowe, 1979) and that for the amphotropic and dual tropic receptors to chromosomes 8 and 1, respectively (Hilkens et al, 1979; Kozak, 1983). Attempts to isolate receptor mutants in mouse fibroblasts by selecting against cell surface expression of the protein indicate that mutations are rare ($<10^{-10}$) or perhaps that the receptor is required for cell viability in vitro (see this chapter).

Since the gene for ecotropic receptor protein could be mapped to a single genetic locus it was reasonable to expect that it could be genetically isolated by transfering DNA from receptor bearing mouse cells to non-receptor-bearing, non-mouse cells (eg. hamster or human). As stated above, however, such experiments utilizing $CaPO_4$ mediated gene transfer methodologies have not been convincingly successful. Because DNA mediated gene transfer is characterized by fragmentation and recombination of transfected DNA, the attempts using this method may have failed because the receptor gene is very large or because expression of it requires distal regulatory elements. For example, the hamster gene which confers resistance to

multiple cancer therapy drugs (MDR) had never been successfully transfered by standard transfection, but was finally transfered to and expressed in mouse cells after chromosome mediated gene transfer (Gros et al, 1985). The MDR gene was subsequently found to be coded for by sequences covering over100 kb - a gene far too large to be transfered intact by standard transfection.

Therefore it seemed likely that the ecotropic receptor gene could be isolated by the methods described thus far in this thesis. Dominant selectable markers could be introduced into a mass population of mouse cells and chromosomes from this population used as donors in irradiation hybrid experiments (Goss and Harris). Since the receptor is a cell surface marker, hybrids containing both the introduced marker and the receptor gene could be easily selected as had been done previously to isolate the H-2 antigen region (Weis et al, 1984). However, no antisera specific for the receptor has been available. Therefore, the intial step toward the goal of isolating the receptor gene in a limited somatic cell hybrid was to develop assays which would easily discriminate receptor-bearing cells from non-receptor-bearing cells. We report in this chapter two such assays; one permits positive selection for receptor bearing cells and the other permits selection against those cells.

VIROLOGY 140, 183–187 (1985)

Cell Surface Receptors for Murine Leukemia Viruses: Two Assays and Their Implications

BARBARA L. HANDELIN[1] AND DAVID KABAT

Department of Biochemistry, School of Medicine, Oregon Health Sciences University, Portland, Oregon 97201

Received June 21, 1984; accepted September 10, 1984

Two assays were developed for identifying individual cells which bear murine leukemia virus receptors: an erythrocyte rosette assay for ecotropic receptors, and an efficient immune cytotoxic assay for cells with ecotropic or amphotropic receptors. Both assays indicate that ecotropic MuLV adsorbed to its cell surface receptor only slowly becomes internalized. Furthermore, attempts to isolate murine fibroblast variants lacking these ecotropic MuLV receptors were unsuccessful, suggesting either that mutations in the receptor gene are rare (less than 10^{-7} per generation) or that the receptor is required for cell viability. These assays are rapid and can be used to identify receptor-bearing cells in mixed populations, a prerequisite for molecular genetic studies. © 1985 Academic Press, Inc.

Penetration of cells by murine leukemia viruses (MuLVs) requires binding of the MuLV membrane glycoprotein (gp70) to specific host membrane receptors (*1*). This receptor specificity determines the host range of MuLVs as ecotropic, xenotropic, amphotropic, or dual tropic. Numerous studies indicate that the receptors are proteins (*2–9*) encoded by separate genetic loci (*10–14*). Yet despite the potential importance of MuLV receptors for pathogenesis (*15*) and their critical role in the use of MuLVs as gene transfer vehicles (*16–20*), there has been no consistent evidence concerning their structures or concerning the mechanisms by which they mediate viral penetration.

It would be useful to study these receptors by molecular genetic methods. For this purpose it is necessary to develop methods which rapidly identify individual cells which either contain or lack particular MuLV receptors. In this report we describe two techniques which are useful for this purpose and we employ these methods to characterize the receptors which mediate ecotropic MuLV penetration. Our results suggest that binding of gp70 to the ecotropic receptor is only slowly followed by removal of gp70 from the cell surface and that this receptor may be required for viability of cultured murine fibroblasts.

A freeze–thaw lysate of ecotropic (Friend) MuLV was prepared by suspending 6 mg of virus particles purified from the culture medium of Eveline cells (*21*) in 1 ml TSE buffer (20 m*M* Tris, pH 7.4; 100 m*M* NaCl; 1 m*M* EDTA, pH 8.0) and freezing the suspension to −70° followed by thawing to 30°, three times. The assay was done using the supernatant retrieved after centrifugation at 100K *g* in a Beckman type 65 rotor for 35 min. This supernatant contained approximately 750 μg protein/ml which was highly enriched for gp70, consistent with previous results (*22*).

The medium in a 25-cm² flask of cultured cells was replaced by 2 ml of rosetting medium (Dulbecco's modified Eagle's medium plus 1% bovine serum albumin, pH 7.3) plus 7.5 μg of the gp70 containing virus extract. After 45 min incubation at 37°, the medium was replaced with fresh rosetting medium plus 2 μl of a previously described goat antiserum to Rauscher MuLV gp70 (*23*). Twenty minutes later, erythrocytes coupled with *Staphylcoccus aureus* protein A (*24*) were added, followed

[1] Author to whom requests for reprints should be addressed.

0042-6822/85 $3.00
Copyright © 1985 by Academic Press, Inc.
All rights of reproduction in any form reserved.

by 2 hr incubation at 37°. Cell monolayers were washed gently with rosetting medium until nonspecific background erythrocyte binding was negligible as determined by observation with a phase contrast microscope.

All tested murine cells (i.e., Balb/C 3T3, SC-1, and NIH 3T3 fibroblasts) rosette heavily in these conditions, whereas all nonmurine cells (i.e., cat Fe S+L−, monkey CV-1, mink CCL64, and hamster CRL 1640) fail to rosette, suggesting that the assay is specific for ecotropic MuLV receptors. Figure 1 illustrates the specificity of this technique for detecting receptor-bearing murine fibroblasts in a background of CCL64 mink lung fibroblasts.

The protocol for this assay (i.e., erythrocyte attachment occurring substantially after gp70 adsorption) suggests that the ecotropic MuLV gp70 which is adsorbed to its receptor is not rapidly internalized or otherwise removed from the cell surface. Consistent with this conclusion, we have observed extensive specific rosetting of murine cells even when addition of antiserum is delayed until 4 hr after removing the gp70 containing medium from the cells.

Replacing the protein A coupled erythrocytes in the rosetting procedure described above with GIBCO (Grand Island, N. Y.) rabbit complement (1:12 dilution) results in specific lysis of cells bearing ecotropic receptors. Balb/c 3T3 and NIH 3T3 cells are completely lysed whereas all nonmurine cells listed above are completely resistant. As expected, this killing requires prior adsorption of ecotropic MuLV or of a gp70-containing extract plus use of appropriate cytotoxic antiserum.

We attempted to develop similar cytotoxicity assays for nonecotropic MuLV receptors using amphotropic, xenotropic, and dual tropic virus preparations. As expected, amphotropic MuLV (4070A) plus appropriate cytotoxic antisera sensitized both murine and nonmurine cells. However, the xeno (Balb virus 2) and dual tropic (HIX and MCF-247) virus preparations did not sensitize any cells, although they were infectious to their target cells and although the antisera used were

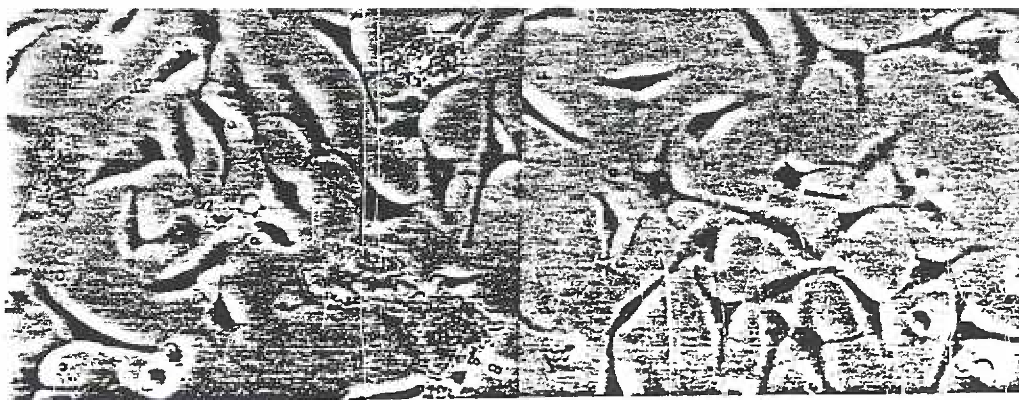


FIG. 1. Erythrocyte rosette assay for detecting cells which have ectropic MuLV receptors. Murine Balb/c 3T3 and mink lung CCL64 fibroblasts were cocultured in a ratio of 1:10. Parallel cultures were then subjected to the rosetting assay (see text). The left frame shows the rosetting of murine Balb/c 3T3 fibroblasts in the presence of gp70, goat antiserum to gp70, and protein A-coated erythrocytes in a background of nonrosetted CCL64 fibroblasts. Observation of many fields indicated that the ratio of rosetted mouse to nonrosetted mink cells was consistent with the input ratios. The right frame is the control assay in which gp70 was omitted. Rosetting also failed to occur when the specific viral antiserum was replaced with preimmune goat serum. Magnification is 190×.

highly cytotoxic to these infected cells (unpublished observations). The latter negative results may reflect low receptor concentrations or avidities, rapid internalizations of bound gp70, or steric hindrances to cytotoxic sensitization.

A quantitative study which illustrates the specificity and killing efficiency of the assay for ecotropic and amphotropic MuLV receptors is shown in Fig. 2. Balb/c and CCL64 cells were prelabeled with L-[^{35}S]methionine and release of radioactivity into their culture medium was analyzed during the cytotoxic assay. The results substantiate the specificity described above.

The complement-mediated killing of Balb/c 3T3 fibroblasts by ecotropic MuLV gp70 plus cytotoxic antiserum is extremely rapid and efficient (greater than 99.9% killing after 3 hr). Therefore, we used this technique in an effort to isolate resistant murine Balb/c 3T3 fibroblast mutants which lack receptors. For this purpose, a culture of 10^6 cells was subjected to 11 cycles of killing followed by periods allowing for recovery and regrowth to a population size of 10^6 cells. However, increased resistance to killing of the cell population was never observed. Mathematical analysis *(25)* of these results suggests that the frequency of receptor-negative Balb/c 3T3 mutants must be less than 10^{-7} per generation or that such mutants are inviable.

As with the rosette assay, the protocol for this cytotoxic sensitization assay im-

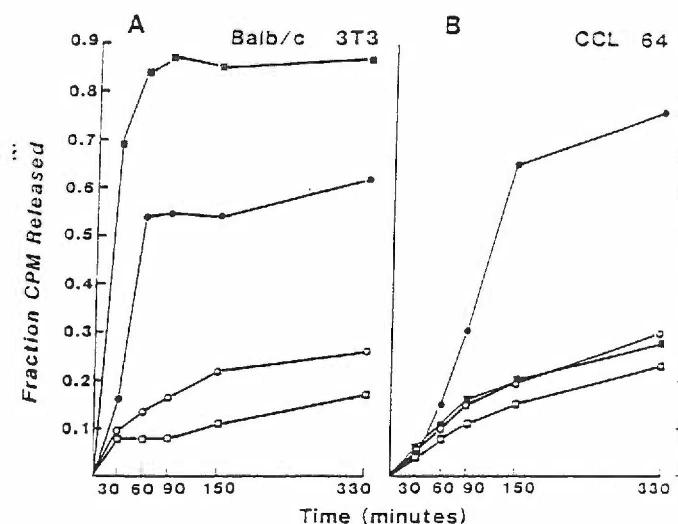


FIG. 2. Cytotoxic assay for cells with receptors for ecotropic and amphotropic MuLVs. Cultures (25 cm^2) containing 5×10^{-5} murine Balb/c 3T3 or mink lung CCL64 fibroblasts were labeled by incorporation of L-[^{35}S]methionine for 2 hr followed by a 2-hr chase with nonradioactive complete medium as described previously *(29)*. The cultures were then treated sequentially, as described in the text, with ecotropic MuLV or with amphotropic MuLV (4070-A) followed by incubation with medium supplemented with 9% rabbit complement (GIBCO, Grand Island, N. Y.) and either 2 μl cytotoxic antiserum specific for gp70 or with preimmune goat serum, as described previously *(29)*. Radioactivity released into the culture medium was periodically assayed in 5-μl samples. The total radioactivity in each culture was assayed at the end of the experiment after solubilization with 0.1% sodium dodecyl sulfate. (A) Balb/c 3T3 cells' data after incubation with ecotropic MuLV gp70 and either antibody (■) or preimmune serum (□) and after incubation with amphotropic MuLV gp70 and either antibody (●) or preimmune serum (○). Eighty-five percent release of radioactivity is 100% killing as indicated by cell viability studies. Also, viable cells slowly released some radioactivity. (B) Results obtained with CCL-64 mink cells using the same procedures.

plies that ecotropic MuLV gp70 which is adsorbed to its receptor is not rapidly internalized or otherwise removed from the cell surfaces. Furthermore, the quantitative study (Fig. 2) was done with intact virions, demonstrating that multivalent particles are also only slowly internalized. Indeed, Balb/c 3T3 fibroblasts remain highly sensitive to cytotoxic antiserum added as long as 8 hr after virus adsorption (unpublished observations). However, since our assay is not quantitative, we conclude only that gp70 is not completely internalized by this time. In contrast, some other ligands become internalized and degraded within minutes of binding to their receptors (*26*). One possible interpretation is that the outer membranes of ecotropic and amphotropic MuLVs may fuse directly with the plasma membranes of receptor-bearing cells, resulting in penetration of only the virion cores. According to this model, gp70 entry into intracellular organelles would not occur during infection. However, it is also possible that the intact virions are internalized by slow pinocytosis or by rapid pinocytosis followed by recycling of gp70 to the cell surfaces. Earlier electron microscopy studies are consistent with all of these models (*27*).

Our results call into question the interpretation of earlier work using a cytotoxic assay to investigate penetration by other enveloped viruses (*28*). In particular, the authors assumed that enveloped viruses which cause cytotoxic sensitization of infectable cells must penetrate by fusion of the virion and cell surface membranes, thereby resulting in deposition of the virion glycoproteins into the plasma membranes. However, our results establish that sensitization can be accomplished simply by adsorption of a viral glycoprotein. Clearly, therefore, bystander virus or glycoprotein adsorbed onto cell receptors might in some cases suffice for cytotoxic sensitization.

We are currently using these techniques in efforts to molecularly clone the MuLV receptor genes. These methods may also be applicable to the study of other cell surface receptors.

ACKNOWLEDGMENTS

We thank Ethel Polonoff for excellent technical assistance. B.L.H. is a graduate student in the Department of Medical Genetics and is a Tartar Trust Fellow. This research was supported by Grant PCM 8218398 from the National Science Foundation.

REFERENCES

1. GAZDAR, A. F., OIE, H., LALLEY, P., MOSS, W. W., MINNA, J. D., and FRANCKE, U., *Cell* 11, 949–956 (1977).
2. BISHAYEE, S., STRAND, M., and AUGUST, J. T., *Arch. Biochem. Biophys.* 189, 161–171 (1978).
3. DELARCO, J., and TODARO, G., *Cell* 8, 365–371 (1976).
4. KALYANARAMAN, V. W., SARNGADHARAN, M. G., and GALLO, R. C., *J. Virol.* 28(3), 686–696 (1978).
5. LANDEN, B., and FOX, C. F., *Proc. Natl. Acad. Sci. USA* 77(8), 4988–4992 (1980).
6. ROBINSON, P. J., HUNSMANN, G., SCHNEIDER, J., and SCHIRRMACHER, V., *J. Virol.* 36(1), 291–294 (1980).
7. SCHAFER, W., FISCHINGER, P. J., COLLINS, J. J., and BOLOGNESI, D. P., *J. Virol.* 21, 35–40 (1977).
8. SCHAFFAR-DESHAYES, L., CHOPPIN, J., and LEVY, J. P., *J. Immunol.* 126(6), 2352–2354 (1981).
9. TWARDZIK, D. R., FOWLER, A. K., WEISLOW, O. S., HEGAMYER, G. A., and HELLMAN, A., *Proc. Soc. Exp. Biol. Med.* 162, 304–309 (1979).
10. HILKENS, J., COLOMBATTI, A., STRAND, M., NICHOLS, E., RUDDLE, F. A., and HILGERS, J., *Som. Cell Genet.* 5(1), 39–49 (1979).
11. OIE, H. K., GAZDAR, A. F., LALLEY, P. A., RUSSELL, E. K., MINNA, J. D., DELARCO, J., TODARO, G. J., and FRANCKE, U., *Nature (London)* 274, 60–63 (1978)
12. REIN, A., *Virology* 120, 251–257 (1982).
13. KOZAK, C. A., *J. Virol.* 98(1), 300–303 (1983).
14. RUDDLE, N. H., CONTA, B. S., LEINWAND, L., KOZAK, C., RUDDLE, F., BESMER, P., and BALTIMORE, D., *J. Exp. Med.* 148, 451–465 (1978).
15. WEISS, R., TEICH, N., VARMUS, H., and COFFIN, J., *In* "RNA Tumor Viruses, Molecular Biology of Tumor Viruses," 2nd ed., pp. 236–246. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y., 1982.
16. JOYNER, A., KELLER, G., PHILLIPS, R. A., and BERNSTEIN, A., *Nature (London)* 305(6), 556–558 (1983).

17. MANN, R., MULLIGAN, R. C., and BALTIMORE, D., *Cell* 33, 153-159 (1983).

18. WEI, C., GIBSON, M., SPEAR, P. G., and SCOLNICK, E. M., *J. Virol.* 39(3), 935-944 (1981).

19. TABIN, J. C., HOFFMAN, J. W., GOFF, S. P., and WEINBERG, R. A., *Mol. Cell. Biol.* 2(4), 426-436 (1982).

20. MILLER, A. D., JOLLY, D. J., FRIEDMAN, T., and VERMA, I. M., *Proc. Natl. Acad. Sci. USA* 80, 4709-4713 (1983).

21. SEIFERT, E., CLAVIEZ, M., FRANK, H., HUNSMAN, G., SCHWARZ, H., and SCHAFER, W., *Z. Naturforsch.* 30(5), 689-700 (1975).

22. STRAND, M., and AUGUST, J. T., *J. Biol. Chem.* 251, 559-564 (1976).

23. EVANS, L. H., DRESLER, S., and KABAT, D., *J. Virol.* 24, 865-874 (1977).

24. GODING, J. W., *J. Immunol. Method* 10, 61-66 (1976).

25. BESTWICK, R. K., MACHIDA, C. A., and KABAT, D., *Mol. Cell Biol.* 4, 908-914 (1984).

26. BROWN, M. S., ANDERSON, R. G. W., and GOLDSTEIN, J. L., *Cell* 32, 663-667 (1983).

27. ANDERSEN, K. B., and NEXO, B. A., *Virology* 125, 85-98 (1983).

28. FAN, D. P., and SEFTON, B. M., *Cell* 15, 985-992 (1978).

29. KABAT, D., RUTA, M., MURRAY, M. J. and POLONOFF, E., *Proc. Natl. Acad. Sci. USA* 77(1), 57-61 (1980).

SUMMARY AND CONCLUSIONS

Our goal was to demonstrate the feasibility of isolating a given region of the human genome containing a gene of interest by inserting a dominant selectable marker near enough to the gene so that the entire region surrounding it could be selectively mobilized. We chose as a model gene the genetic determinant for Huntington's Disease which had previously been mapped to the tip of human chromosome 4. There were two main motivations behind this choice: (1) if a dominant selectable marker could be inserted near the putative HD gene so that this unknown HD determinant could be transferred into heterologous cells, then the function of the gene could be studied for the first time in vitro. This might shed light on the biochemical or physiological action of its gene product; and (2) a goal for the Collaborative Agreement has been to identify a marker which would flank the HD locus on the other side of HD from the D4S10 marker. Such a marker is highly desirable for diagnostic purposes and it also would delineate a region (between the two makers) to be searched for the HD gene. If we could insert a marker near enough to the HD gene to be able to transfer small fragments around it, this would represent a significantly enrichmented source of possible flanking markers.

We proceeded stepwise. First, a monochromosomal hybrid, HD113/2B, which had one proviral SVX(neo) as a dominant selectable marker was derived by microcell fusion. We mapped the integration site of SVX(neo) by in situ hybridization and by rescuing the provirus plus host flanking sequences. The host flanking sequences could be mapped provisionally to human chromosome 4 and the in situ hybridization detected neo sequences at 4p1.4-1.6.

Limited hybrids were then created by fragmenting the chromosomes of HD113/2B and transferring human sequences with the neo marker into CHTG recipient hamster cells. These $G418^R$ hybrids were tested for cotransfer of the markers known to be linked to HD and were all negative. However, a phage library made from one hybrid, GH8000R#3, was screened for human clones and among several human positive clones picked, one has been preliminarily mapped to

chromosome 4. If confirmed, this result will establish the integration site of SVX(neo) as 4p1.4-1.6.

In summary, the SVX(neo) virus integration site is probably located in the tip of human chromosome 4. The marker transfers only a very small fragment of the human chromosome after chromosome breakage with irradiation, perhaps indicating that it is located in an unstable region which is susceptible to fragmentation. Together, these data are consistent with the hypothesis that the integration of SVX(neo) is in telomeric sequences which, as understood from studies of other organisms and other human chromosomes, undergo recurrent recombination and reassortment. This hypothesis is testable by mapping more clones from the GH8000R#3 library, especially using the panel of somatic cell hybrids derived by MacDonald et al (1987). Such a location for the neo marker would provide the opportunity to meet both of the goals for this project - that is, to identify a flanking marker and to be able to mobilize the HD gene.

Anticipating less fortuitous results and also other projects, we designed, constructed and utilized a new retrovirus vector, SP-1. SP-1 carries a new dominant selectable marker, His, and is rescuable by the suppressor gene protocol used by others (Reik et al, 1985 and Lobel et al, 1985). In addition, it introduces recognition sequences for two rarely cutting restriction enzymes, Nru I and Sfi I. We showed that we could infect HD113/2B cells with ecotropic SP-1 virus and that the provirus faithfully inserted new restriction sites into the host DNA. These restriction sites, by their hypomethylation after integration into the host chromosome, permitted us to observe that the proviruses carrying them may preferentially integrate near hypomethylated host chromosome regions. Further experiments are being pursued to fully explore this question on the randomness of proviral integration.

Finally, we returned to a gene cloning challenge that has resisted success by several investigators: tha t is, to clone the murine cellular ecotropic receptor gene. The tools we had developed for targeting a given chromosome region presented an opportunity to attempt to isolate that receptor gene in a limited somatic cell hybrid, as a first step toward cloning it. Having acknowledged the advanatages of utilizing cell surface markers in this somatic cell genetics

protocol, we capitalized on the expression of ecotropic receptors on host cell membranes in vitro. Two assays for cell surface expression of the receptor gene have been developed as the necessary tools for discriminating between future hybrid cells which have retained the receptor gene from those which have not.

REFERENCES

Ames, B.N., and Hartman, P.E. (1963) The histidine operon. Cold Spring Harbor Symp. Quant. Biol. **28**, 349-356. 4

Arrighi, F.E. and Hsu, T.C. (1971) Localization of heterchromatin in human chromosomes. Cytogeneticis **10**, 81-86.

Arriza, J.L., Weinberher, C., Cerelli, G., Glaser, T.M., Handelin, B.L., Housman, D.E., and Evans, R.M. (1987) Cloning of human mineralocorticoid receptor cDNA, a structural and functional kinship with the gludocorticoid receptor. Science **237**, 268-275.

Athwal, R.S., Searle, B.M., Jansons, V.K. (1985b) Diphtheria toxin sensitivity in a monochromosomal hybrid containing human chromosome 5. J. of Hered. **76**, 3299-334.

Athwal, R.S., Smarsh, M., Searle, B.M., Deo, S.S. (1985a) Integration of a dominant selectable marker into human chromosomes adn transfer of marked chromosomes to mouse cells by microcell fusion. Som. Cell Mol. Genet. **11**, 177-187.

Barker, D., et al (1987) Gene for von Recklinghausen Neurofibromatosis is in the perientromeric region of chromosome 17. Science **236**, 1100-1102.

Barski, G. Sorieul, S., and Cornefert, F. (1960) Production dans des cultures in vitro de deux souches cellularies en association, de cellules de caractere "hybride". CR. Hebd. Seances Acad. Sci. **251**, 1825-1827.

Beaudet, A.L., Su, T.S., O'Brien, W.E., D'Eustachio, P., Barker,P.E. and Ruddle, F.H. (1982) Dispersion of argininosuccinate-synthetase-like human genes to multiple autosomes and the X chromosome. Cell **30**, 287-293.

Bell, J. (1934) Huntington's Chorea. Treasury of Human Inheritance, vol IV. Cambridge Univ. Press (Cambridge, Mass.).

Besmer, P., and Baltimore, D. (1977) Mechanism of restriction of ecotropic and xenotropic murine elukemia viruses and formation of psuedotypes between the two viruses. J. Virol. **21** 965-973.

Bird, A.P. (1986) CpG-rich islands and the function of DNA methylation. Review. Nature **321**, 209-213.

Bishayee, S., Strand, M., and August, J.T. (1978) Cellular membrane receptors for oncovirus envelope glycoprotein. Arch. Biochem. Biophys. **189**, 161-171.

Bonner, T.I., Kerby, S.B., Sutrave, P., Gunnell, M.A., Mark, G., and Rapp, U.R. (1985) Structure and biological activity of human homologs of the raf/mil oncogene. Mol. Cell. Biol. **5**, 1400-1407.

Botstein, D., White, R.L., Skolnick, M., Davis, R.W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. **32**, 314-331.

Breindl, M., Harbers, K. and Jaenisch, R. (1984) Retrovirus-induced lethal mutation in collagen I gene of mice is associated with an altered chromatin structure. Cell **38**, 9-16.

Bruns, G., Gusella, J., Keys, C., Housman, D., and Gerald, P. (1982) Isolation of DNA segments from the human X. Cytogenet. and Cell Genet. **32**, 256.

Calos, M.P., Legkowski, J.S. and Botchan, M.R. (1983) High mutation frequency in DNA transfected into mammalian cells. Proc. Natl. Acad., Sci. USA **80**, 3015-3019.

Capecchi, M.R. (1980) High efficiency trasformation by direct microinjection of DNA into cultured mamalian cells. Cell **22**, 479-488.

Carle, G.F. and Olson, M.V. (1984) Separation of chromosomal DNA molecules from yeast by orthogonal-field-alternation gel electrophoresis. Nuc. Acids Res. **12** (14), 5647-5664.

Caspersson, T., Zech, L., and Johanson, C. (1970) Differential banding of alkylating fluorochromes in human chromosomes. Exp. Cell Res. **60**, 3.

Cavanee, W., Leach, R., Mohandas, T., Pearson, P., and White, R. (1984) Isolation and regional localization of DNA segments revealing polymorphic loci from human chromosome 13. Am.J. Hum. Genet. **36**, 10-24.

Cepko, C.L., Roberts, B.E., and Mulligan, R.C. (1984) Construction and applications of a highly transmissible murine retrovirus shuttle vector. Cell **37**, 1053-1062.

Chasin, L.A., Feldman, A., Konstam, M., Urlaub, G. (1974) Reversion of a Chinese hamster cell auxotrophic mutant. Proc. Nat. Acad. Sci. USA **71**, 718-22.

Chen, T.R. and Ruddle, F.H. (1971) Karyotype analysis utilizing differentially stained constitutive heterochromatin of human and murine chromosomes. Chromosoma **34**, 51-72.

Colicelli, J., and Goff, S.P. (1985) Mutants and pseudorevertants of Moloney murine leukemia virus with alterations at the integration site. Cell **42**, 573-580.

Conklin, C.F., and Groudine, M. (1984) Chromatin structure and gene expression, in DNA Methylation (Springer-Verlag, New York).

Conklin, K.F. and Groudine, M. (1986) Varied interactions between proviruses and adjacent host chromatin. Mol. Cell. Biol. **6**, 3999-4007.

Conrad, S.E., Jiu, C-P, Botchan, M.R. (1982) Fragment spanning the SV40 replication origin is the only DNA sequence required in cis for viral excision. Science **218**, 1223-1225.

DeLarco, J., and Todaro, G.J. (1976) Membrane receptors for murine leukemia viruses, characterization using the purified envelope glycoprotein, gp 71. Cell **8**, 365-371.

Donlon, T.A., Litt, M., Newcom, S.R. and Magenis, R.E. (1983) Localization of the restriction fragment length polymorphism D14S1 (pAW-101) to chromosome 14q32.1-32.2 by in situ hybridization. Am. J. Hum. Genet. **35**, 1097-1106.

Drets, M.E. and Shaw, M.W. (1971) Specific banding patterns of human chromosomes. Proc. Natl., Acad. Sci. USA **68**, 2073-2077.

Durnam, D.M., Gelinas, R.E. and Myerson, D. (1985) Somat. Cell Mol. Genet. **11**, 571-577.

Egeland, J.A., Gerhard, D.S., Pauls, D.L., Sussex, J.N., Kidd, K.K., Allen, C.R., Hosteter, A.M. and Housman, D.E. (1987) Bipolar affective disorders linked to DNA markers on chromosome 11. Nature **325**, 7783-787.

Ege, T., Ringertz, N.R., Hamberg, H. and Sidebottom, E.(1977) Preparation of microcells. Methods Cell Biol. **15**, 339-357.

Eighth International Workshop on Human Gene Mapping (HMG 8). Cytogenet. Cell Genet. **40**.

Elston, R.C., Pericak-Vance, M.A., Meyers, D.A., Conneally, P.M. (1980) Pedigree analysis of Huntington disease allowing for variable age of onset. Am. J. Hum. Genet. **32**, 142A.

Ephrussi, B., and Weiss, M.C. (1965) Interspecific hybridization of somatic cells. Proc. Nat. Acad. Sci. USA, **52**, 1940-1042.

Estivill, Xavier, Farrall, M., Scambler, P.J., et al (1987) A candidate for the cystic fibrosis locus isolated by selection for methylation-free islands. Nature **326**, 840-845.

Farrer, L.A., Conneally, P.M., Yu, P.L. (1984) Natural history of Huntington's disease - possible role of "aging genes". Am. J. Hum. Genet. **18**, 115-123.

Fearon,E.R., Vogelstein, B., and Feinberg, A.P. (1984) Somatic deletion and duplication of genes on chromosome 11 in Wilms' tumours. Nature **309**, 176-178.

Feinberg, A.P. and Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**, 6-13.

Ford, C.E. and Hamerton, J.L. (1956) The chromosomes of man. Nature **178**, 1020-1023.

Fournier, R.E.K., and Rudle, F.H. (1977) Microcell-mediated transfer of murine chromosomes into mouse, Chinese hamster, and human somatic cells. Proc. Natl. Acad. Sci. USA **74**, 319-323.

Friend, K.K., Chen, S., and Ruddle, F.H. (1976) Differential staining of interspecific chromosomes in somatic cell hybrids by alkaline giemsa stain. Somatic Cell Genet. **2**, 183-188.

Friend, S.H., Bernards, R., Rogelj, S., Weinberg, R.S., Rapaport, J.M., Albert, D.M., and Dryja, T.P. (1986) A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. Nature **323**, 643-646.

Frischauf, A-M, Lehrack, H, Poustka, A., and Murray, N. (1983) Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. **170**, 827-842.

Garrod, A.E. (1902) The incidence of alcaptonuria - a study in chemical individuality. Lancet. **ii**, 1616-1620.

Gerhard, D.S., Jones, C., Morse, H.G., Handelin, B., Weeks, V. and Housman, D. (1987) Analysis of human chromosome 11 by somatic cell genetics - reexamination of derivatives of human-hamster cell line J1. Som. Cell. Mol. Genet. **13**, 293-304.

Gilliam, J.C., Tanzi, R.E., Haines, J.L., Bonner, T.I., Faryniarz, A.G., Hobbs, W.J., MacDonald, M.E., Cheng, S.V., Folstein, S.E., Conneally, P.M., Wexler, N.S., and Gusella, J.F. (1987) Localization of the Huntington's disease gene to a small segment of chromosome 4 flanked by D4S10 and the telomere. Cell, in press.

Gilliam, T.C., Scambler, P., Robbins, T., Ingle, C., Williamson, R., and Davies, K.E. (1984) The positions of three restriction fragment length polymorphisms on chromosome 4 relative to known genetic markers. Hum. Genet. **68**, 154-158.

Gilliam, T.C., Healey, S.T., MacDonald, M.E., Stewart, G.D., Wasmuth, J.J., Roy, J.C., Gusella, J.F. (1987a) Isolation of polymorphic DNA fragments from human chromosome 4. Nuc. Acid Res. in press.

Goddard, J.M., Caput, D., Williams, S.R. and Martin, D.W. (1983) Cloning of human purine-nucleoside phosphorylase cDNA sequences by complementation in E. coli Proc. Natl. Acad. Sci. USA **80** 4281-4285.

Goff, S.P., (1986) Retroviral integration and insertional mutagenesis (Book in DH office)

Goldstein, J.L., Brown, M.S., Anderson, R.G.W., Russell, D.W. and Schneider, W.J. (1985) Annu. Rev. Cell. Biol. **1**, 1-39.

Goldstein, J.L., and Brown, M.S. (1973) Familial hypercholesterolemia - identification of a defect in the regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase activity asociated with over production of cholesterol. Proc. Natl. Acad. Sci. USA **70**, 2804-2808.

Goodfellow, P.N., Banting, G., Trowsdale, J., Chambers,S., and Solomon, E. (1982) Introduction of a human X-6 translocation chromosome into a mouse teratocarcinoma - investigation of control of HLA, B, C expression. Proc. Natl. Acad. Sci. USA **79**, 1190-1194.

Goss, S.J. and Harris, H. (1975) New method for mapping genes in human chromosomes. Nature **255**, 680-684.

Goss, S.J. and Harris, H. (1977) Gene transfer by means of cell fusion. II. Statistical mapping of 8 loci on human chromosome 1 by statistical analysis of gene assortment in somatic cell hybrids. J. Cell. Sci. **25**, 39-57.

Goss, S.J. and Harris, H. (1977) Gene transfer by means of cell fusion. I. Statistical mapping of the human X-chromosome by analysis of radiation-induced gene segregation. J. Cell. Sci. **25**, 17-37.

Graham, F.L., and van der Eb, A.J. (1973) A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52**, 456-467.

Gros, P., Fallows, D.A., Croop, J.M. and Housman, D.E. (1986) Chromosome-mediated gene transfer of multidrug resistance. Mol. Cell. Biol. **6**, 3785-3790.

Gusella, J.F., Jones, C., Kao, F-T., Housman, D.E., Puck, T.T. (1982) Genetic fine structure mapping in human chromosome 11 by use of repetitive DNA sequences. Proc. Natl. Acad. Sci. USA **79**, 7804-7808.

Gusella,J.F., Wexler, N.S., Conneally, P.M., Naylor, S.L., Anderson, M.A., Tanzi, R.E., Watkins, P.C., Ottina, K., Wallace, M.R., Sakaguchi, A.Y., Young, A.B., Shoulson, I., Bonilla, E., and Martin, J.B. (1983) A polymorphic DNA marker genetically linked to Huntington's disease. Nature **306**, 234-238.

Gusella, J., Varsanyi-Breiner, A, Kao, T.T., Jones, C., Puck, T.T., Keys, C., Orkin, S. and Housman, D. (1979) Isolation and localization of DNA segments from specific human chromosomes. Proc. Natl. Acad. Sci USA **76**, 5239-5243.

Gusella, J.F., Keys, C., Varsanyi-Breiner, A., Kao, F-T., Jones, C., Puck, T.T. and Housman, D. (1980) Isolation and localization of DNA segments from specific human chromosomes. Proc. Natl. Acad. Sci. USA **77**, 2829-2833.

Harper, M.E. and Saunders, G.F. (1981) Localization of single copy DNA sequences on G-banded human chromosomes by in situ hybridization. Chromosoma **83**, 431-439.

Harris, H. and Watkins, J.F. (1965) Hybrid cells derived from mouse and man - artificial heterodaryons of mammalian cells from different species. Nature **205**, 640-646.

Hattori, M.S., Kuhara, O., Takenaka, and Sakai, Y. (1986) L1 family of repetitive DNa sequences in primates may be derived from a sequence encoding a reverse transcriptase-related protein. Nature. **321**, 625-628.

Hereditary Disease Foundation, Colaborative Research Agreement Meeting Report , prepared by Barbara L. Handelin. October, 1986.

Hilkens, J., Colombatti, A., Strand, M., Nicols, E., Ruddle, F.A. and Hilger, J. (1979) Som. Cell. Genet. **5**, 39-49.

Hirt, B. (1967) Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **26**, 365-369.

Hoffman, J.W., Steffen, D., Gusella, J. Tabin, C., Bird, S., Cowing, D., and Weinberg, R.A. (1982) DNA methylation affecting the expression of murine leukemia proviruses. J. Virol. **44**, 144-157.

Houck, C.M., Rinehart, F.P., Schmid, C.W. (1979) A ubiquitous family of repeated DNA sequences in the human genome. J. Mol. Biol. **132**, 289-306

Hsu, T.C. (I179) Human and Mammalian Cytogenetics - An Historical Perspective (Springer-Verlag, New York).

Hsu T.C. and Pomerat, C.M. (1953) Mammalian chromosomes in vitro. II. A method for spreading the chromosomes of cell in tissue culture. J. Hered. **44**, 23-29.

Hughes, S.H., Mutschler, A., Bishop, J.M., and Varmus, H.E. (1981a) A Rous sarcoma virus provirus is flanked by short direct repeats of a cellular DNa sequence present in only one copy prior to integration. Proc. Natl. Acad. Sci. USA **78**, 4299-4303.

Hughes, S.H., Vogt, P.K., Stubblefield, E., Bishop, J.M., and Varmus, H. E. (1981b) Integration of avian sarcoma virus DNA in chicken cells. Virology **108**, 208-221.

Jaenisch, R., Harbers, K., Schnieke, A., Lohler, J., Chumakov, I., Jahner, D., Grotkopp, D., and Hoffman, E. (1983) Germline integration of Moloney murine leukemia virus at the Mov 13 locus leads to recessive lethal mutation and early embryonic death. Cell, **32**, 209-216.

Johnson, P.A. and Rosner, M.R. (1986) Characterization of murine-specific leukemia virus receptor from L cells. J. Virol. **58**, 900-908.

Jones, C. (1975) Synteny betweeen the pro+ marker and human glutamate oxaloacetate transaminase. Somatic Cell Genet., 1, 345-354.

Jones, M.B. and Phillips, C.R. (1970) Affected parent and age of onset in Huntinton's chorea. J. Med. Genet. 7, 20-21.

Jones, C. and Kao, F.T. (1978) Regional mapping of the gene for human lysosomal acid phosphatase (ACP2) using a hybrid clone panel containing segemnts of human chromosome 11. Hum. Genet. **45**, 1-10.

Jones, C., Kao, F.T., Taylor, R.T. (1980) Chromosomal assignment of the gene for folylpolyglutamate synthetase to human chromosome 9. Cytogenet. Cell Genet. **28**, 181-194.

Jones, C., Wuthier, P., Kao, F.T., Puck, T.T. (1972) Genetics of somatic mammalian cells. XV. Evidence for linkage between human genes for lactic dehydrogenase B and serine hydroxymethylase. J.Cell. Physiol. **80**, 291-98.

Jones, C. Patterson, D., Kao, F.T. (1981) Assignment of the gene coding for phosphoriboslyglycineamide formyltransferase to human chromosome 14. Somatic Cell Genet. **7**, 399-409.

Kalyanaraman, V.S., Sarngadharan, M.G. and Gallo, R.C. (1978) Characterization of Rauscher murine leukemia virus envelope glycoprotein receptor in membranes from murine fibroibasts. J. Virol. **28**, 686-696.

Kao, F.T., Jones, C., Puck, T.T. (1977) Genetics of cell-surface antigens - regional mapping of three components of the human cell-surface antigen complex A_L on chromosome 11. Somat. Cell. Genet. **3**, 421-429.

Kao, F.T. and Puck, T.T. (1970) Genetics of somatic mammalian cells - linkage studies with human-Chinese hamster cell hybrids. Nature **228**, 329-332.

Klobutcher, L.A., and Ruddle, F.H. (1979) Phenotype stabiliization and integration of transferred material in chromosome-mediated gene transfer. Nature **280**, 657-660.

Knowlton, R.G., Cohen-Haguenauer, O., Van Cong, N., Rezal, J., Brown, V.A., Barker, D., Braman, J.C., Schumm, J.W., Tsui, L.C., Buchwald, M., and Donis-Keller, H. (1985) A polymorphic DNA marker linked to cystic fibrosis is located on chromosome 7. Nature **318**, 380-382.

Koenig, M., Hoffman, E.P., Bertelson, C.J., Monaco, A.P. Feener, C., Kunkel, L.M. (1987) Complete cloning of the Duchenne Muscualr Dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. Cell **50**, 509-517.

Kozak, C.A., and Rowe, W.P. (1979) Genetic mapping of the ecotropic murine leukemia virus-inducing locus of BLAB/c mouse to chromosome 5. Science **204**, 69-71.

Kozak, C.A. (1983) Genetic mapping of a mouse chromsomal locus requried for mink cell focus-transforming virus replication. J. Virol. **48**, 300-303.

Kunkel, L.M., Monaco, A.P. Middlesworth, W., Ochs, H.D. and Latt, S.A. (1985) Specific cloning of DNA fragments absent from the DNA of a male patient with an X chromosome deletion. Proc. Natl. Acad. Sci. USA **82**, 4778- 4782.

Landegent, J.E., Jansen in de Wal, N., Fisser-Groen, Y.M., Bakker, E., Van der Ploeg, M., Pearson, P.L. (1986) Fine mapping of the Huntington disease linked D4S10 locus by non-radioactive in situ hybridization. Hum. Genet. **73**, 354-357.

Landen, B., and Fox, F.C. (1980) Isolation of BPgp70, a fibroblast receptor for the envelope antigen of Rauscher murine elukemia virus. Proc. Natl. Acad. Sci. USA **77**, 4988-4992.

Langer-Safer, P.R., Levine, M., and Ward, D.C. (1982) Immunological method for mapping genes on Drosophila polytene chromosomes Proc. Natl. Acad. Sci. USA **79**, 4381-4385.

Lawrance, S.K., Smith, C.L., Srivastava, R., Cantor, C.R., Weissman, S.M. (1987) Megabase-scale mapping of the HLA gene complex by pulsed field gel electrophoresis. Science **235**, 1387-1390.

Law, M.L., Davidson, J.N., Kao, F-T. (1982) Isolation of a human repetitive sequence and its application to regional chromosome mapping. Proc. Natl. Acad. Sci. USA **79**, 7390-7394.

Law, M.L., and Kao, F.T. (1978) Induced segregation of human syntenic genes by 5-bromodeoxyuridine and near-visible light. Somatic Cell Genet. **4**, 465-476.

Lee, W-H., Bookstein, R., Hong, F., Young, L-J., Shew, J-Y., and Lee, E. Y-H. (1987) Human retinoblastoma susceptibility gene - cloning, identificatin and sequencing. Science **285**, 1394-1399.

Littlefield, J.W. (1964) Selection of hybrids from matings of fibroblasts in vitro and their presumed recombinants. Science **145**, 709-710.

Litt, M and White, R. (1985) A highly polymorphic locus in human DNA revealed by cosmid-derived probes. Proc. Natl. Acad. Sci. USA **82**, 6206-6210.

Lobel, L.I., Patel, m., King, W., Nguyen-Huu, M.C., Goff, S.P. (1985) Construction and recovery of viable retroviral genomes carrying a bacterial suppressor transfer RNA gene. Science 228, 329-331.

Lugo, T.G., Handelin, B., Killary, A.M., Housman, D. and Fournier R.E.K. (1987) Isolation of microcell hybrid clones containing retroviral vector insertions into specific human chromosomes. Mol. Cell. Biol. in press.

Lugo, T.G. and Baker, R.M. (1983) Chromosome-mediated gene transfer of HPRT and APRT in an intraspecific human cell system. Somat. Cell Genet. 9, 175-188.

MacDonald, M.E., Anderson, M.A., Gilliam, T.C., Lisbeth Tranebjaergo, Carpenter, N.J., Magenis, E., Hayden, M.R., Healey, S.T., Bonner, T.I., and Gusella, J.F. (1987) A somatic cell hybrid panel for localizing DNA segments near the Huntington's disease gene. Genomics, in press.

Magenis, R.E., Gusella, J.F., Weliky, K., Olson, S., Haight, G., Toth-Fijel, S., Sheehy, R. (1986) Huntington disease-linked restriction fragment length polymorphism localized within band p16.1 of chromosome 4 by in sity hybridization. Am. J. Hum. Genet. 39, 383-391.

Majors, J.E., and Varmus, H.E., (1981) Nucleotide sequences of host-proviral junctions for mouse mammary tumour virus. Nature **289**, 253-258.

Maniatis,T., Fritsch, E.F., Smabrook, J. (1982) Molecular Cloning - A Laboratory Manual (Cold Spring Harbor Laboratories, Cold Spring Harbor, N.Y.)

Mann, R., Mulligan, R.C. and Baltimore, D. (1983) Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell **33**, 153-159.

Manuelidis, L. and Ward, D.C. (1984) Chromosoma **91** 28-38.

Martin, J.B. (1982) Huntington's disease - genetically programmed cell death in the human central nervous system. Nature **229**, 205-206.

McBride, O.W. and Ozer, H.L. (1973) Transfer of genetic information by purified metaphase chromosomes. Proc. Natl. Acad. Sci. USA **70**, 1259-1262.

McDougall, J.K., Kucherlapati, R., and Ruddle, F.H. (1973) Nature **245**, 172-175.

Merritt, A.D., Conneally, P.M., Rahman, N.F., Drew, A.L. (1969) Juvenile Huntington's chorea, in Progress in Neurogenetics, edited by Barbeau, A., Brunette, T.R., Excerpta Medica Foundation (Amsterdam), 645-650.

Migeon, B.R., Grown, T.R., Axelman, J., Migeon, C.J. (1981) Studies of the locus for androgen receptors - localization on the human X chromosoem and evidence for homology with the locus in the mouse. Proc. Natl. Acad. Sci. USA. **78**, 6339-6343.

Miller, C.K., and Temin, H.M. (1983) High-efficiency ligation and recombination of DNA fragments by vertebrate cells. Science **220**, 607-609.

Miller, C.L., Ruddle, F.H. (1978) Co-transfer of human X-linked markers into murine somatic cells via isolated metaphase chromosomes. Proc. Natl, Acad. Sci. USA **75**, 3987-3990.

Monaco, A.P., Neve, R.L., Colletti-Feener, Cl, Bertelson, C.J., Kurnit, D.M., Kunkel, L.M. (1986) Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. Nature **323**, 646-650.

Monaco, A.P., Bertelson, C.J., Middlesworth, W., Colletti, C.A., Aldridge, J., Fischbeck, K.H., Bartlett, R., Perical-Vnace, A., Rose, A.D., and Kunkel, L.M. (1985) Detection of deletions spanning the Duchenne muscular dystrophy locus using a tightly linked DNA segment. Nature **316**, 842-845.

Moore, E.E., Jones, C., Kao, F.T., Oates, D.C. (1977) Synteny between glycinamide ribonucleotide synthetase and superoxide dismutase (soluble). Am. J. Hum, Genet. **29**, 389-396.

Mueller, G.A., Graessmann, A., and Graessmann, G. (1978) Mapping of early SV40 specific functions by microinjection of different early viral DNA fragments. Cell **15**, 579-585.

Mulligan, R.C. and Berg, P.(1980) Expression of a bacterial gene in mammalian cells. Science **209**, 1422-1427.

Murray, J.M., Davies, K.E., Harper, P.S., Meredith, L., Mueller, C.R., and Williamson, R. (1982) Linkage relationship of a cloned DNA sequence on th eshort arm of the X chromosome to Duchenne muscular dystrophy. Nature **300**, 69-71.

Murray, M.J., Shih, Cl, Cowing, D., Finkelstein, R. and Weinberg, R. A. (1981) Three different human tumor cell lines contain different oncogenes. Cell **25**, 355-361.

Myers, R.H., Goldman, D., Bird, E.D., et al (1983) Maternal transmission in Huntington's disease. Lancet **i**, 208-210.

Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E., and White, R. (1987) Variable number of tandem repeat (VNTR) markers for human gene mapping. Science **235**, 1616-1622.

Nelson, D.L., Weis, J.H., Przyborski, M.J., Mulligan, R.C., Seidman, J.G>, and Housman, D.E. (1984) Metaphase chromosome transfer of introduced selectable markers. J. Mol. and Appl. Genet. **2**, 563-577.

Nelson, D.L. (1984) Metaphase chromosome transfer of introduced selectable markers - use in gene mapping. PhD thesis, Massachusetts Institute of Technology (Cambridge, Mass.)

Olsen, A.S., McBride, O.W., Moore, D.E. (1981) Number and size of human X chromsome fragments transferred to mouse cells by chromosome-mediated gene transfer. Mol. Cell. Biol. **1**, 439-448.

Orkin, S.H., Daddona, P.E., Shewach, D.S., Markham, A.F., Bruns, G.A., Goff, S.C. and Kelley, W.N. (1983) Molecular cloning of human adenosine deaminase gene sequences. J. Biol. Chem. **258**, 82-84.

Owerbach, D., Bell, G.I., Rutter, W.J. and Shows, T.B. (1980) The insulin gene is located on chromosome 11 in humans. Nature **286**, 82-84.

Pardue,M.L. and Gall, J.G. (1970) Chromosomal localization of mouse satellite DNA. Science **168**, 1356-1358.

Patil, S.R., Merrick, S., and Lubs, H.A. (1971) Identification of each human chromosome with a modified Giemsa stain. Science **173**, 821-822.

Peter, P.M., Kamarck, M.E., Hemler, M.E., Strominger, J.L., and Ruddle, F.H. (1984) Genetic and biochemical characterization of human lymphocyte cell surface antigens - the A-1A5 and A-3A4 determinants. J.Exp.Med. **159**, 1441-1454.

Pritchard, C.A. and Goodfellow, P.N. (1986) Development of new methods in human gene mapping - selection for fragments fo the human Y chromosome after chromosome-mediated gene transfer. EMBO J. **5** (5), 979-985.

Pritchard, C.A. and Goodfellow, P.N. (1987) Investigation of chromosome-mediated gene transfer using the HPRT region of the human X chromosome as a model. Genes and Devel. **1**, 172-178.

Puck, T.T. and Kao, F.T. (1982) Somatic cell genetics and its application to medicine. Ann. Rev. Genet. **16**, 225-271.

Puck, T.T., Marcus, P.I., and Cieciura, S.J. (1956) Clonal growth of mammalian cells in vitro. Growth characteristics of colonies from single HeLa cells with and without a "feeder" layer. J. Exp. Med. **108**, 273-284.

Ray, P.N., Belfall, B., Duff, Cl, Logan, Cl, Kean, Vl, Thompson, M.W.m, Sylvester, J.E., Gorski, J. L., Schmickel, R.D., and Worton, R.G. (1985) Cloning of the breakpoint of an X/21 translocation associated with Duchennes muscular dystrophy. Nature **318**, 672-675.

Reeders, S.T., Breuning, M.H., Davies, K.E., Nickolis, R.D., Jarman, A.P., Higgs, D.R., Pearson, P.L., Weatherall, D.J. (1985) A highly polymorphic DNA marker linked to adult polycystic kidney disease on chromosome 16. Nature **317**, 542-544.

Reed, TE, Chandler, J.H. (1958) Huntington's chorea in Michigan. I. Demography and genetics. Am. J. Hum. Genet. **10**, 201-225.

Reik, W., Weiher, H., and Jaenisch, R. (1985) Replication-competent Moloney murine leukemia virus carrying a bacterial suppressor tRNA gene - selective cloning of proviral and flanking host sequences. Proc. Nat. Acad. Sci. USA. **82**, 1141-1145.

Rein, A. (1982) Virology **120**, 251-257.

Rettig, W.J., Dracopoli, N.C., Goetzger, T.A., Spengler, B.A., Biedler, J.L., Oettgen, H.F., and Old, L.J. (1984) Somatic cell genetic analysis of human cell surface antigens, chromosomal assignments and regulaton of expression in rodent-human hybrid cells. Proc. Natl. Acad. Sci. USA. **81**, 6437-6441.

Robinson, H.L. and Gagnon, G.C. (1986) Patterns of proviral insertion and deletion in avian leukosis virus-induced lymphomas. J. Virol. **57**, 28-36.

Rohdewohld, H., Weiher, H., Reik, W., Jaenisch, R., and Breindl, M. (1987) Retrovirus integration and chromatin structure , Moloney murine leukemia proviral integration sites map near DNAse I-hypersensitive sites. J. Virol. **61**, 336-343.

Rothfels, K.H. and Siminovitch, L. (1958) An air drying technique for flattening chromosomes in mammalian cells grown in vitro. Stain Technol. **38**, 73-77.

Roth, J.R., Anton, D.R. and Hartman, P.E. (1966) Histidine regulatory mutants in S. typhimurium. I. Isolation and general properties. J. Mol. Biol. **22**, 305-323.

Ruddle, LF.H. and Creagan, R.P. (1975) Parasexual approaches to the genetics of man. Ann. Rev. Genet. **9**, 407-486.

Saxon, P.J., Srivatsan, E.S., Leipzig, G.V., Sameshima, J.H., and Stanbridsge, E.J. (1985) Selective transfer of individual human chromosomes to recipient cells. Mol. Cell. Biol. **5**, 140-146.

Schubach, W., and Groudine, M. (1984) Alteration of c-myc chromatin structure by avian leukosis virus integration. Nature **307**, 702-708.

Schwartz, D.C. and Cantor, C.R. (1984) Separation of yeast chromosome-sized DNAs by Pulsed Field Gradient Gel Electrophoresis. Cell **37**, 67-75.

Seabright, M. (1971) A rapid banding technique for human chromsomes. Lancet **ii**, 971-972.

Seizinger, B.R. et al (1987) Genetic linkage of von Recklinghausen Neurofibromatosis to the nerve growth factor receptor gene. Cell **49**, 589-594.

Shaw, D.J., Brook, J.D., Brown, D.S. and Thmas, N.S.T. (1986) Mapping human chromosomes in somatic cell hybrids using a low-copy-number repetitive sequence. Somat. Cell Mol. Genet. **12**, 333-337.

Smith, C.L. and Cantor, C.R. (1987) Purification, specific fragmentation, and separation of large DNA molecules. Methods in Enzymology, in press.

Smith, C.L., Econome, J.G., Schutt, A., Klco, S., and Cantor, C.R. (1987) A physical map of the Escherichia coli K12 genome. Science **236**, 1448-1453.

Soriano, P,, Szabo, P and Bernardi, G. (1982) The scattered distribution of actin genes in the mouse and human genomes. EMBO J. 1, 579-583.

Soriano, P., Gridley, T., and Jaenishch, R. (1987) Retroviruses and insertional mutagenesis in mice - proviral integration at the Mov 34 locus leads to early embryonic death. Genes and Devel. **1**, 3656-3675.

Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol.Biol. **98**, 503-518.

Southern, P.J. and Berg, P. (1982) Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. J. Mol. Appl. Genet. 1, 327-341.

St. George-Hyslop, P.H., et al (1987) The genetic defect causing Familial Alzheimer's disease maps on chromosome 21. Science **235**, 885-890.

Sumner, A.T., Evans, H.J. and Buckland, R.A. (1971) A new technique for distinguishing between human chromosomes. Nature New Biol. **232**, 31-32.

Su, T-S, Bock, H-G.O., O'Brien, W.E. and Beaudet, A.L.(1981) Cloning of cDNA for argininosuccinate synthetase mRNA and study of enzyme overproduction in a human cell line. J. Biol. Chem. **256**, 11826-31.

Szybalska, E. Hl, Szybalski, W. (1962) Genetics of human cell lines. IV. DNA mediated heritable transofrmation of a biochemical trait. Proc. Nat. Acad. Sci. USA **48**, 2026-2034.

Tanzi, R.E., Gusella, J.F., Watkins, P.C., Bruns, G.A.P., St. George-Hyslop, P., Van Keuren, M.L., Patterson, D., Pagan, S., Kurnit, D.M., Neve, R.L. (1987) Amyloid b protein gene - cDNA, mRNA distributin, and genetic linkage near the Alzheimer locus. Science **235**, 880-884.

Tashima, M., Clabretta, B., Torelli, G., Scofield, M., Maizel, A., and Saunders, G.F. (1981) Presence of a highly repetitive and widely dispersed DNa sequences in the human genome. Proc. Natl. Acad. Sci. USA **78**, 1508-1512.

Thayer, R.E., Harper, M.E., Sawyer, J., Singer, M.F., McBride, O.W. (1987) Localization of a unique DNA segment in the telomeric region of human chromosome 4p. Cytogenet. Cell Genet., in press.

Tjio, J.H., and Levan, A. (1956) The chromosome number of man. Hereditas **42**, 1-6.

Tunnacliffe, A., Parkar,M., Povey, S., Bengtsson, B.O.,Stanley, K., Solomon, E. and Goodfellow, P. (1983) Integration of Ecogpt and SV40 early region sequences into human chromosome 17, a dominant selection system in whole cell and microcell human-mouse hybrids. MEBO J. **9**, 1572-1584.

Varmus, H., and Swanstrom, R. (1982) Replication of retroviruses, in RNA Tumor Viruses (Cold Spring Harbor Laboratory, Cold Sptring Harbor, N.Y.).

Vijaya, S., Steffen, D.L., and Robinson, H.L. (1986) Acceptor sites for retroviral integrations are near DNAse I-hypersensitive sites in chromatin. J. Virol. 60, 683k-692.

Vogelstein, B. and Gillespie, D. (1979) Preparative and analytical purification of DNA from agarose. Proc. Natl. Acad. Sci. USA 76,615-619.

Wake, C.T., Gudewicz, T., Porter, T., White, A. and Wilson, J.H. (1984) How damaged is the biologically active subpopulation of transfected DNA? Mol. Cell. Biol. 4, 387-398.

Wang, H.S., Greenberg, C.R., Hewitt, J., Kalousek, D., Hayden, M.R. (1986) Subregional assignment of the linked marker G8 (D4S10) for Huntington's disease to chromosome 4p16.1-16.3. Am. J. Hum. Genet. 39, 397-403.

Wang, H.C. and Fedoroff, S. (1972) Banding in human chromosoems treated with trypsin. Nature New Biol. 235, 52-53.

Wasmuth, J.J., Carlock, L.R., Smith, B., and Immken, L.L. (1986) A cell hybrid and recombinant DNA library that facilitate identification of polymorphic loci in the vicinity of the Huntington Disease gene. Am. J. Hum. Genet. 39, 397-403.

Watkins, P., Tanzi, R., Landes, G., Tricoli, J.V., Shows, T.B., and Gusella, J.F. (1984) DNA segments for human chromosome 21. Cytogenet. Cell Genet. 37, 602.

Weiss, J.H., Nelson, D.L. , Prxyborski, M.J., Chaplin, D.D., Mulligan, R.C., Housman, D.E. and Seidman, J.G. (1984) Eukaryotic chromosome transfer - linkage of the murine major histocompatability complex to an inserted dominant selectable marker. Proc. Natl. Acad. Sci. USA 81, 4879-4883.

Weiss, M.C. and Green, H. (1967) Huamn-mouse hybrid cell lines containing partial complements of human chromosomes and functioning human genes. Proc. Nat. Acad. Sci. USA 58, 1104-1111.

Weiss, R., Teich, N., Varmus, H., and Coffin, J. (1982) RNA Tumor Viruses (Cold Spring Harbor, New York).

Wexler, N.S., Gusella, J.F., Conneally, P.M. and Housman, D.E. (1975) Huntington's Disease and the new genetics.

White, R.S., Woodwar, S., Leppert, M., O'Connell, P., Hoff, M., Herbst, J., Lalouel, J.M., Dean, M., and Vande Woude, G. (1985) A closely linked genetic marker for cystic fibrosis. Nature 318, 382-384.

Wigler, M., Silverstein, S., Lee, L.S., Pellicer, A., Cheng, Y., and Axel, R. (1977) Transfer of purified herpes virus thymidine kinase gene to cultured mouse cells. Cell 14, 725-731.

Worton, R.G., Duff, C., Sylvester, J.E., Schmickel, R.D., and Willard, H.F. (1984) Duchenne muscular dystrophy involving translocation of the dmd gene next to ribosomal RNA genes. Science 224, 1447-1449.

Wyman, A.R., and White, R. (1980) A highly polymorphic locus in human DNA. Proc. Natl. Acad. Sci. USA 77, 6754-6758.

Yunis, G. and Hasmineh, W.G. (1975) Heterochromatin, satellite DNA and cell function. Science **174**, 1200-1209.

Zabel, B.U., Naylor, S.L., Sakaguchi, A.Y., and Gusella, J.F. (1986) Mapping of the DNA locus D4S10 and the linked Huntington's disease gene to 4p16-p15. Cytogenet. Cell. Genet. **42**, 187-190.