9-10-2010

Two Methods to Detect Clonal Populations of Human Cells In Situ

Philip Scorza Hall
Yale University

Two Methods to Detect Clonal Populations of Human Cells *In Situ*

A Thesis Submitted to the
Yale University School of Medicine
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Medicine

By

Philip Scorza Hall

2010

Abstract

TWO METHODS TO DETECT CLONAL POPULATIONS OF HUMAN CELLS IN SITU. Philip Hall, Jonathan Murphy, and Jeffrey Sklar, Department of Pathology, Yale University, School of Medicine, New Haven, CT.

A molecular assay to detect clonal populations of human cells *in situ* would be potentially valuable for both investigational and diagnostic purposes. Two such methods are proposed, both utilizing fluorescence *in situ* hybridization (FISH). The first relies upon random monoallelic expression of genes (so-called allelic exclusion), in which a subset of human genes are normally expressed at a single allele in a fixed fraction of cells within a tissue, independent of the parental origin of the allele. It is hypothesized that application of FISH to assess the allelic expression patterns among one or more of these genes should be able to distinguish a monoclonal population of cells from a polyclonal one. The second method, specific for T-cells, relies upon VDJ segmental recombination at the T-cell receptor beta locus. With this method, our hypothesis is that analysis by FISH of the configuration of rearranged VDJ segments should be able to distinguish a monoclonal population of T cells from a polyclonal population. Both proposed assays were tested on benign tonsil and thymus tissue as well as on monoclonal cell pellets produced from neoplastic cell lines. In those analyses that could be completed, attempts to assess the expression pattern either of genes subject to random allelic exclusion or the determination of VDJ segmental recombination failed to distinguish monoclonality from polyclonality. Although unsuccessful, the failure of these attempts was due to technical limitations and not to fundamental problems with the underlying hypotheses.

Acknowledgements

Jeffrey Sklar, M.D.,Ph.D, Department of Pathology and Laboratory Medicine
Hui Li, Ph.D.
Xian-Yong Ma, Ph.D.
Jonathan Murphy
Jinglan Wang

Some research materials were made available through the generous donation of Peter Cresswell, Ph.D., Yale University, Department of Immunobiology, and George Miller, M.D., Yale University, Department of Pediatrics

Funding for this project came from the Yale University School of Medicine, Department of Pathology, Post-Sophomore Fellowship Grant

Table of Contents

Abstract	1
Acknowledgements	2
Introduction	4
Statement of Purpose	19
Methods	20
Results	30
Discussion	34
References	39
Figures	45

Introduction

The ability to determine whether a population of human cells is clonal, its lineage traceable to the division of a single cell, has been valuable for various types of basic biologic studies and for certain diagnostic applications. An important area for which considerations of clonality are relevant is neoplasia and cancer. Most, if not all neoplasms arise from the clonal proliferation of cells derived from a single transformed precursor cell (1-6). Moreover, the identification of multiple different types of preneoplastic or premalignant lesions has motivated the investigation of their clonality, especially in connection to the risk of potential transformation to true neoplasia. Indeed, several molecular methods now in widespread use for the diagnosis and detection of some human cancers rely upon the identification of a monoclonal population of cells (4, 7-9). These methods have all involved extraction of nucleic acids from cells that have been removed from tissues and are not broadly applicable to all cell types. In this thesis, we propose two methods of expanding the applicability of assays for monoclonality, notably using *in situ* techniques that preserve cell and tissue morphology, with a discussion of initial experience with these methods.

The difficulty of creating an assay for clonality derives from the fact that neoplastic cells share most of the same genetic makeup as non-neoplastic cells and display many of the same proteins. For those specific differences that occur, distinguishing neoplastic from normal, polyclonal cells by their genetic or protein composition alone requires prior knowledge of, or an extensive, often expensive search for how the two populations differ (1, 6, 10-11). In this manner, identifying

one or more genetic variants--such as a specific point mutation, as occurs frequently in neoplasia; a DNA deletion, as is associated with tumor suppressor genes; or a chromosomal translocation, as is found in several hematologic malignancies and sarcomas--has been beneficial not only for advancing our understanding of carcinogenesis but also for improving our ability to diagnose malignancy (6, 11). Unfortunately, for most cancers, a defining genetic variation either is not known, is present in only a fraction of cases of a given cancer, or is not a convenient marker for technical reasons (e.g., mutations that might occur anywhere over large regions of DNA within an oncogene). On these grounds, a generic, broadly applicable test for clonality would be of great diagnostic utility.

In the absence of a broadly applicable, objective test for clonality, current methods of diagnosing cancer continue to rely on the evaluation of morphologic and histopathologic features, which are subject to interobserver variability (12, 13). Nevertheless, because of the many decades of experience with the histopathology of tumors, much information has been accumulated about the histopathologic characteristics of neoplasms, and morphology-based diagnostic criteria have been established. A reliable and convenient method of detecting clonality would be valuable to both researchers and diagnosticians, particularly if clonality could be assessed *in situ* and directly correlated with morphology.

There have been several proposed assays to determine the clonality of a population of human cells, but each has been hampered by limitations related to specific tissues or to technical problems that restrict its use. Very few offer the potential for *in situ* analysis. The assays that can be applied to detect monoclonality

in more than one or a few specific tumor types have relied upon the determination of one of a small number of genetic events that are known to occur randomly during the differentiation of certain somatic cells. The guiding principle of such assays is that certain genetic or epigenetic events, whether dichotomous, as in the choice of a kappa or lambda light chain immunoglobulin gene for expression in B lymphocytes; or polychotomous, as in T-cell receptor or immunoglobulin gene rearrangement patterns, occur randomly in a polyclonal population of human cells; i.e., each cell in the population is (or is descended from) a cell in which the genetic or epigenetic event has occurred and has resulted in a fixed change (kappa expression vs. lambda, or a particular VDJ rearrangement) (4, 6-7, 9, 14-31). A monoclonal population, however, will display a decidedly, or statistically, non-random pattern, if the genetic or epigenetic event occurred prior to the clonal expansion of that particular cell population. As an example, the ratio of B lymphocytes making kappa to B lymphocytes making lambda light chains, within a given population of non-neoplastic cells, should fall within a statistically predictable range centered around 2:1. Deviation from that range is evidence of the presence of a clonal population (6). Similarly, the configuration of rearranged V, D, and J segments that have been joined together in the immunoglobulin heavy chain loci of genomic DNA (as well as the precise sequence of DNA at the junction of these segments and the length of that sequence) differs among different mature B lymphocytes. Gel electrophoresis of the products generated from the cellular DNA of a polyclonal population of B lymphocytes by the polymerase chain reaction (PCR) using primers that flank the VDJ junctions within the immunoglobulin heavy chain locus results in a smear of

similarly sized but diverse bands that differ in their precise lengths. In contrast, the finding of a discrete band by gel electrophoresis is indicative of the presence of a monoclonal population, i.e., that a significant fraction of the cells have an identical VDJ rearrangement, implying that those cells are all derived from a single precursor cell in which that particular VDJ rearrangement occurred (4, 8-10, 30, 32). Both of these examples, where specific, non-random events distinguish monoclonal from polyclonal cell populations, in which comparable events are random, have been used by clinicians to aid in the diagnosis of malignancy, notably lymphocytic cancers (2, 4, 7-9, 30, 33, 34). As illustrated by these examples, PCR, along with flow cytometry and immunohistochemistry, remain the diagnostic tools of choice for many lymphoproliferative disorders (35). However, these examples are limited to lymphocytes, the only cells known to undergo somatic recombination in a specific genetic locus during normal maturation and differentiation or to produce kappa or lambda light chains.

An assay for clonality must target a genetic or epigenetic event that occurs in the type of cell from which the neoplasm (or clone) arises. Thus, the most useful genetic/epigenetic event would be one that occurs in all somatic cells. An assay widely used in basic biologic studies has focused on X-chromosome inactivation, a random dichotomous event that occurs in female somatic cells during early embryologic development, whereby in each cell, either the maternally- or paternally-derived X-chromosome is inactivated, with stable transmission of the inactivated state to the same X chromosome in all of that cell's progeny (1, 6, 36). Assays that can distinguish whether the active X-chromosome in a female somatic cell is paternally-

or maternally-derived allow for clonal analysis of a population of human cells. Evidence for preponderance of inactivation of one or the other X-chromosome homologs is consistent with a monoclonal population of cells. However, such assessments depend in part on the characteristics of the background cell population and the ratio between activation of the two X-chromosomes in the particular tissue in which a population of cells is being analyzed, since skewing from the theoretical 1:1 ratio is sometimes observed in normal tissues. Initial attempts to distinguish the active X-chromosome relied on the analysis of gene products in female individuals heterozygous for a gene on the X-chromosome, commonly glucose-6-phosphate dehydrogenase (1, 3). While successful, these analyses were limited by the relative rarity of individuals heterozygous for those particular polymorphisms (3). Later analyses were based on the fact that DNA of the inactive X-chromosome is hypermethylated relative to the active X-chromosome. To distinguish the active from inactive X-chromosome DNA, the DNA extracted from the tissue sample was cleaved using methylation-sensitive endonucleases, followed by Southern blot hybridization, or, alternatively, DNA of a gene on the X-chromosome, typically the PGK or androgen receptor gene, was amplified using PCR, followed by analysis of the products by gel electrophoresis. In either of these strategies, X-chromosome loci were selected for analysis based on highly prevalent polymorphisms in DNA sequence that could be detected either by restriction fragment length polymorphism (RFLP) in Southern blot analysis or differently-sized PCR products (3, 5, 10, 37). Such polymorphism analysis enabled distinction of the two X-chromosomes within female tissues and permitted analysis of a larger percentage of females than were

heterozygous for G6PD. However, even using these methods, it was still possible to accomplish analysis in only a fraction of cases. For example, in one study, only thirteen of the fifteen tissue samples could be analyzed, either due to skewing in the normal surrounding tissues or, less frequently, to lack of heterozygosity in the locus used for analysis (3).

In addition to other technical limitations, the above strategies for analysis of clonality have been limited either to lymphocytes or to female somatic cells. Other strategies have been employed in particular settings. For example, variations in the stucture of Epstein-Barr virus (EBV) episomal DNA has been used as a clonal marker. Each cell infected with EBV has multiple identical circular episomes of viral DNA, formed by covalent ligation of the two ends of the double-stranded, linear genome upon cellular infection. Both ends of the linear genome are composed of the same tandem repeat sequences, about 500 basepairs in length, and the episomal form of the EBV genome in latently infected cells is generated by homologous recombination between the tandem repeats leaving variable numbers of such repeats in the episomal circles. The number of included repeats is stable and is passed on by precursor cells to their progeny without variation (6, 38). Analyzing the length of the region containing the repeats by Southern blot has helped identify clonality within several tumors, including nasopharyngeal carcinoma and Hodgkin's disease, but is obviously limited to EBV-infected cells (6, 38, 39). Other techniques have analyzed sites of integration into host cell DNA by Hepatitis B viral DNA, as a unique marker of an infected cell and its progeny. As with EBV, this marker is limited to only specific situations, such as hepatocellular carcinoma (6, 40).

The clonality assays already described have increased our understanding of carcinogenesis and enabled clonal analysis of multiple tumor types and preneoplastic processes. However, these assays have generally not been used routinely for diagnosis because of the difficulties or complexities of the technique or the lack of broad applicability to many tumor types (37, 41, 42). An ideal assay, then, would need to target a genetic event that occurs randomly within a given population of cells, one that occurs in multiple or most cell types and populations, one that produces some heritable, cellular change that is not affected by growth conditions or by neoplastic transformation, and one that can be easily assessed. Random monoallelic gene expression, also referred to as allelic exclusion, may provide exactly such an event, especially in light of the multiple recent advances in our understanding both of the process itself and of its prevalence within cells.

Monoallelic expression occurs when one of two alleles for a specific gene is activate and the other is inactive, such that only one of the two alleles is transcribed. This situation may arise through the random activation of one allele while the other allele remains inactive, or, conversely, through the random inactivation of one allele of a gene in which both alleles are initially active. On a chromosome-wide basis, a similar process gives rise to random X-chromosome inactivation in female somatic cells. So-called imprinting of autosomal genes also involves inactivation of only one allele in somatic cells, but differs fundamentally from random monoallelic expression because either the paternally- or maternally-inherited copy of an imprinted gene is consistently inactivated (36, 43-45). Imprinted genes have been implicated in several inherited human diseases, the best known being Prader-Willi and Angelman

Syndrome, which are both due to deletion of the same chromosomal region (15q11-13) but differ phenotypically depending upon the parental origin of the chromosome that suffers the deletion (43, 44, 46).

A third type of monoallelic expression involves random transcriptional inactivation of a single allele in cells in which both alleles are initially active; in this type of monoallelic expression, the choice of which allele is expressed is independent of its parental origin (47-52). In fact, it seems that random monoallelic expression of a given gene often occurs in only a minority of the cells within a tissue in which that gene is active, but it appears to be a stable percentage (50). In 1994, Chess *et al*. described random monoallelic expression of olfactory receptor genes in sensory neurons, and in 2007, Gimelbrant *et al*. analyzed 4,000 somatic human genes in clonal cell lines and identified 300 as demonstrating random monoallelic expression (48, 50). The expression patterns indicated that monoallelic expression was frequently present in a consistent minority of cells regardless of the specific tissue from which the cells were derived.

Compared to the markers for clonality so far described, random monoallelic expression has several theoretical advantages. It occurs in genes on multiple autosomal chromosomes, making it usable in both male and female cells. For several genes identified in the Gimelbrant study, evidence was provided that random monoallelic expression was present in multiple tissue types in a measurable and consistent percentage of cells within each tissue type, but not in all cells. Finally, random monoallelic gene expression is observed in a significant minority of expressed genes, each of which could be a promising target for a clonality assay

depending upon the particular tissue and the extent to which that gene is expressed in a monoallelic fashion within that tissue (i.e., the fraction of cells that exhibit monoallelic expression).

The difficulty in creating an assay for clonality arises in creating a molecular technique to identify the pattern of monoallelic expression within a population of cells. The necessary and important assumption, to be confirmed or rejected in this and future studies, is that the pattern of random monoallelic expression is passed from progenitor cells to their progeny in a stable fashion. However, as stated above, X-chromosome inactivation is the biologic process most analogous to random monoallelic gene expression and the stability of inactivation in that process is well known. Given the success of assays targeting X-chromosome inactivation, random allelic inactivation would seem a promising target for investigation for this experiment. Although it would be possible to use a similar technique to that used in X-chromosome inactivation, namely the analysis of polymorphisms within a gene subject to random monoallelic expression, the limited association of suitable and predictable polymorphisms within appropriate genes greatly complicates this approach. But even more importantly, random allelic inactivation occurs in only a percentage of cells, instead of in every cell as with X-chromosome inactivation. Upon initial consideration, this observation presents a difficult challenge for distinguishing a monoclonal population of cells from a polyclonal population based on patterns of allelic expression, at least if an approach similar to that used for assessing X-chromosome inactivation were employed. However, this feature of monoallelic gene expression in subtotal fractions of cell populations within tissues

actually offers a special opportunity for its use as a marker for *in situ* detection of clonality.

Despite the advantages of *in situ* assays for the determination of monoclonality in preneoplastic or neoplastic tissues in histologic sections, very few such assays currently exist. The one *in situ* clonality assay widely used at present in clinical settings has employed antibody staining for kappa and lambda immunoglobulin light chains (6). Besides being limited to B cells, this method has various disadvantages, principally problems involving its application to formalin-fixed, paraffin-embedded tissues - the standard material from which histologic sections are prepared. *In situ* hybridization for kappa and lambda immunoglobulin light chain mRNA has also been utilized (53), but this approach is generally considered too difficult for routine clinical use. In 2001, Nuovo *et al.* described the use of reverse transcriptase *in situ* PCR to detect T-cell-receptor beta rearrangements in T cell populations, which enabled them to distinguish clonal T cell populations in lymphomas from polyclonal populations in reactive lymph nodes (54). However, this method has significant drawbacks that hinder its clinical adoption, notably its requirement for multiple tissue sections and sequential experimental repetitions to distinguish among different V segments using the twenty-five forward primers necessary to detect the uniquely expressed TCR-beta rearrangements. Furthermore, *in situ* PCR is a very problematic technique that is not in general use because of limited reproducibility.

On the other hand, *in situ* genetic examination has been greatly augmented by advances in fluorescence *in situ* hybridization (FISH) (27, 55-57). FISH has been

widely adopted clinically in the diagnosis and characterization of many malignancies, largely based upon its ability to detect genetic translocations, deletions, and other chromosomal abnormalities (56, 58). All of these applications involve hybridization to DNA targets. Additionally, a simple modification of DNA FISH--RNA FISH, which targets RNA sequences--has enabled the study of gene expression by identifying sites of RNA transcription at the cellular level (43, 57, 59-66). Importantly, the technology can be used with formalin-fixed, paraffin-embedded tissue, rendering it promising for pathologic research and clinical diagnosis (67). In the Gimelbrant study, the use of RNA FISH was coupled with DNA FISH to confirm monoallelic or biallelic expression of two sample genes, death-associated protein kinase 1 (*DAPK1*) and early B cell factor (*EBF*), in peripheral blood monocytes (PBMCs) and clonal lymphoblast lines (50). In this study, 36% and 38% of PBMCs showed monoallelic expression for the two genes, respectively, and in the clonal lymphoblast lines known to display monoallelic expression of one or both genes, 97% and 98% showed monoallelic expression with FISH. However, only 77% and 78% of cells from clonal lymphoblast lines known to have biallelic expression of the two genes displayed biallelic expression with FISH. The reasons for this seem likely to be technical.

Despite the reduced sensitivity of combined RNA and DNA FISH for detecting biallelic expression, use of these methods still seems to offer a promising method for clonality analysis. If RNA and DNA FISH were used to assay whether one or both alleles for a given gene are expressed in a collection of human cells, the ratio of monoallelic to biallelic expression should fall within a statistically predictable

ratio, depending upon the gene chosen for analysis, the tissue in question, and the sensitivity of the assay. To increase the sensitivity of the assay, it should also be possible to investigate the expression patterns of multiple genes simultaneously with probes tagged with multiple fluorophors to distinguish the different genes. This is the strategy we chose for our studies.

We began by selecting two genes for analysis, based upon their expression and prevalence of random monoallelic inactivation in many tissues. These genes were in fact the same two used for FISH analysis in the Gimelbrant study: *DAPK1*, on chromosome 9q34, and early B cell factor (*EBF*), on chromosome 5q34 (50). We then designed a fluorescent probe for hybridization to intronic sequences of these genes. The same probe was selected for both RNA and DNA FISH, with the method of hybridization distinguishing between the RNA and DNA targets. For each gene, DNA FISH serves as an internal control, as it should identify two alleles in each cell. RNA FISH tests whether transcription is monoallelic or biallelic, based on the presence or absence of a nascent, unspliced mRNA (59, 60, 65). The use of an intronic probe assures that only nascent mRNA will be visualized, since mature mRNA lacks the introns to which the probe is designed to hybridize. Because splicing of exons occurs soon after transcription, or perhaps even co-transcriptionally, nascent, precursor mRNA is localized near or at the site of transcription, close to the gene encoding that mRNA.

While the published data forming the rationale and methodologic bases for these studies for *in situ* assessment of clonality seemed fairly strong, we also decided to pursue a parallel, more limited but still potentially useful approach to the *in situ*

assessment of clonality for clinical purposes. This approach also involved FISH and was directed at assessment of the clonality of T cells. As described above, Nuovo *et al.* used RT *in situ* PCR for their assay because FISH was thought to be too insensitive at the time to identify the short regions of DNA involved in TCR beta recombination and because of the need to use multiple probes on several tissue sections to distinguish among the myriad possible combinations of variable (V), diversity (D), or joining (J) segments (54). However, we decided to attempt an approach markedly different from that of Nuovo, *et a*l. by utilizing FISH directed at a specific region in the TCR beta gene and a statistical feature of TCR beta gene rearrangement that characterizes the TCR beta gene in normal T lymphocytes.

In a review of T cell development published in 1992, Malissen *et al.* discussed the frequency of various configurations of TCR beta gene rearrangements among T cells. Within developing T cells, both alleles undergo D-J joining, but whether or not only one or both alleles undergo subsequent V-DJ joining depends upon the outcome of first V-DJ rearrangement with respect to production of a functional protein product. If rearrangement in that first allele is successful, the second allele does not undergo rearrangement (68). Great diversity in the sequence of DNA across the VDJ junction occurs among fully rearranged alleles due to the combined effects of exonucleolytic digestion removing varying numbers of basepairs at the ends of recombining segments plus the addition of variable numbers of random basepairs at the ends of the segments by the enzyme terminal transferase prior to ligation of V, D, and J segments. Non-functional VDJ rearrangements are relatively common, since two-thirds of the sequences generated at the VDJ junctions will be out

of frame in the J and C (constant region) coding sequences. Malissen *et al.* predicted accordingly a two-thirds failure rate for the first allelic rearrangement and a two-thirds failure rate for the second allelic rearrangement. As a viable T cell needs at least one successful V-DJ rearrangement of the TCR beta allele, the ratio of cells with one rearranged locus and one non-rearranged locus to those with two rearranged loci should be about 3:2, as summarized in Figure 4 of their paper. In an analysis of ten T cell clones, they found a ratio of 5:5, with one clone interestingly possessing two successful V-DJ recombinations.

Although their experiment was done in murine T cells and we could find no analogous analysis of human T cells, we thought that while the ratio might be somewhat different in human T cells, the principle should still allow us to distinguish a polyclonal population with a mix of biallelic and monoallelic recombination patterns from a monoclonal population with a predominance of one or the other pattern.

To distinguish the recombination patterns *in situ*, we developed FISH probes targeting the 64kb region of DNA between V29-1, the second-most 3' variable region of the TCR beta locus on chromosome 7, and D1, the most 5' diversity region (69). An additional, rarely used variable region, V30, is located 3' of the C region (69). Control probes were developed targeting the DNA regions immediately 5' and 3' of the TCR beta locus, so as not to be involved in genetic recombination. As successful V-DJ recombination involving any of the variable regions other than V30 causes the excision and rapid destruction of the 64kb region of the TCR beta locus, DNA FISH performed with probes for this test region should be able to distinguish a monoclonal

population of T cells from a polyclonal population (70). Specifically, in a polyclonal population, every cell should display the control probe signal on both alleles, but there should be a mix of cells with one allele having only D-J recombination, and therefore one test probe signal, and cells with both alleles having V-DJ recombination, and no test probe signals. A monoclonal population should have one type of cells or the other. In an actual tissue sample containing a monoclonal sample, there are still always polyclonal T cells and non-lymphoid cells intermixed with monoclonal T cells, and in this situation, there should be a preponderance of cells with either one test probe signal or no test probe signal.

In situ clonality assays, for which random monoallelic gene expression and genetic recombination within TCR genes represent two distinct and promising approaches, have the potential to add to our understanding of carcinogenesis and facilitate clinical diagnosis of malignancy. Random monoallelic gene expression, if it turns out to be a successful target, could be used in a variety of tissue and tumor types. Further research will continue to characterize the large variety of human genes known to be subject to random monoallelic expression and the extent to which this occurs in various tissue types. Genetic recombination within the T cell receptor beta locus may be a technically easier assay, requiring DNA FISH alone without the need for RNA FISH, but as described in this thesis, it is limited to the study of T cell proliferations. Nevertheless, if the principle works for T cells, it should be possible to develop an analogous test for B cells through analysis of recombination within the immunoglobulin heavy chain locus.

Statement of Purpose

The aim of this thesis is to describe and develop two methods of *in situ* analysis of the clonality of populations of human cells, both of which rely upon fluorescence *in situ* hybridization (FISH). In the first method, a combination of DNA FISH and RNA FISH will be used to characterize the expression patterns in lymphoid tissues of two genes which are known to be subject to random monoallelic gene expression. In the second method, DNA FISH will be used to characterize the genetic recombination patterns of the T cell receptor beta locus in T cells. Our hypothesis is that both of these methods should be able to distinguish statistically a monoclonal population of cells from a polyclonal population within tissue sections on a microscope slide.

Methods

Probe Generation:

DAPK1 and *EBF* probes: DNA and RNA FISH probes were generated using an adaptation of the method used by Gimelbrant *et al*., chosen because of their successful use for both DNA and RNA FISH in that paper (50). BAC clones were obtained from the BACPAC Resource Center, available at bacpac.chori.org, containing between 200-400 kb of the two genes in question. The great majority of DNA within these BACs consists of introns but some exon sequence is included. For *DAPK1*, the clone RP11-107G16 was obtained, and for *EBF*, RP11-155P16. These were the same BAC clones used by Gimelbrant *et al*. To amplify BAC DNA, the clones in *E. coli* were incubated in LB media containing 12.5 (m)g/ml chloramphenicol. DNA was isolated from 5ml bacterial cultures using the Qiagen QiaPrep Spin Miniprep kit, eluated from the spin column with 50 (m)l of Qiagen elution buffer (EB) in the final step.

For verification that the BAC DNA contained the appropriate genetic sequences of *DAPK1* or *EBF*, PCR was performed using primers chosen from the NCBI sequences for both genes (www.ncbi.nlm.nih.gov/gene) using MacVector software. For *DAPK1*, the primers chosen were TCTGTGTCCATCCCCCCGAT, forward, and CCATCTATTCCCTTTCCTTTCCGT, reverse, both presented in 5' to 3' direction. These primers are complementary to DNA sequences approximately 3.5 kb apart, within an intron of *DAPK1* according to the NCBI database, and were chosen to be that far apart in case they might be used for generation of FISH probe instead of BAC clone verification. For *EBF*, the primers chosen were

TTCTCTCTCTTGGCTAAGCGG, forward, and CAATGAGCGGAAAAGCGAGG, reverse, also presented 5' to 3'. These primers are complementary to DNA sequences approximately 3.3 kb apart within an intron of EBF, according to the NCBI database. Verification PCR was performed as follows: 17.3 µl H_2O, 1 µl forward primer, 1 µl reverse primer, 2 µl clone DNA or H_20 for control, 0.5 µl Mg2+ 50x buffer, 0.5 µl 50x dNTP mix, 0.2 µl Taq polymerase, were mixed and subjected to 35 cycles of PCR at a 64˚C annealing temperature. The PCR products were separated by 1% agarose gel electrophoresis at 120 mV for 30 minutes, with the results shown in Figure 1.

TCR beta probes: In a similar manner, BAC clones were selected containing regions of the TCR beta (*TCRB*) locus. According to the NCBI sequence, there is a 64 kb region between V29-1 and D1, as referred to above. Comparing the NCBI Clone Registry and Map Viewer (available at www.ncbi.nlm.nih.gov/projects/genome/clone/) with the sequence map of *TCRB*, the BAC CTD-3217E23 was found to contain a 16kb region within the target 64 kb. This BAC was acquired from Invitrogen (clones.invitrogen.com), and is abbreviated VDR, for V-D region. The BAC RP11-368I15 was obtained from the BACPAC Resource Center containing a 250 kb region 3' to the *TCRB* locus for use as a control, here abbreviated 3CR, for 3' control region. This clone had been used by Soulier *et al*. in an experiment using DNA FISH to identify the *TCRB* locus (71). An additional BAC, RP11-10L5 acquired from the BACPAC Resource Center contains approximately 250kb of DNA 5' to the *TCRB* locus as an additional control, and is

abbreviated 5CR, for 5' control region. Each of these BAC clones in *E. coli* were cultured in LB media containing 12.5 µg/ml chloramphenicol, and the DNA was extracted from the bacteria using the Qiagen QiaPrep Spin Miniprep kit as above. The clones were verified by PCR to contain the appropriate region of DNA, using the same PCR parameters as for *DAPK1* and *EBF* above. Primers were chosen using the MacVector program from the NCBI genetic sequence of *TCRB* and the surrounding DNA regions to be 150-350 bp apart, as follows, all listed 5' to 3':

VDR: CCACTAAATGATGTTGTC, forward;
TGTGCTCGTTAAGGATTTC, reverse.
3CR: TTTGGGGAGCACCCTTTG, forward;
CAGGAAGGACAGCTCCT, reverse
5CR: GTTAAAACTTACCTCATTAG, forward;
GTGTGGCAAACAGACAG, reverse

Gel electrophoresis separation of the PCR products for the *TCRB* probes with water controls (PCR amplification performed without template added) is presented in Figure 2.

In subsequent experiments conducted by Jonathan Murphy, an attempt was made to increase the amount of signal obtained from the VDR region of the *TCRB* loci by generating twenty-two non-overlapping probes from the VDR using PCR of VDR DNA from whole cell DNA. These probes ranged in size from 2.8 kb to 3.5 kb in length, and cumulatively covered approximately 55 kb of the 64 kb region between V29-1 and D1. Together, these probes covered about 39 kb of DNA beyond that within the CTD-3217E23 BAC. (BACs covering the entire 64 kb region are available, but all of these contain considerable amounts of DNA outside the 64 kb

region, rendering those BACs unsuitable for our purposes; i.e., all DNA complementary to the probe is deleted as a result of V-DJ joining within the *TCRB* locus.

Additionally, twenty-three probes, ranging in size from 3.0 kb to 3.6 kb in length, were generated from the 3' control region. The PCR to generate the probes was performed as follows: 18 µl H_2O, 0.25 µl forward primer, 0.25 µl reverse primer, 1.5 µl clone DNA, 2.0 µl Mg2+, 2.5 µl 10x buffer, 0.5 µl 50x dNTP mix, 0.2 µl Taq polymerase, were mixed and subjected to 37 cycles of PCR at a 47.5˚C annealing temperature. The PCR products were separated by 0.7% agarose gel electrophoresis and purified from the gel using the Qiagen Gel Purification kit, eluting with 30 ul EB buffer (Qiagen).

Probe Labeling:

Probes were labeled using the Vysis Nick Translation Labeling Kit (Vysis No. 32-801300, now Abbott 07J00-001) as follows: 17.5 µl Miniprep clone DNA were mixed with 5 µl 0.1mM dTTP, 10 µl 0.1 mM dNTP, 5 µl 10x nick translation buffer, 10 µl nick translation enzyme, and 2.5 µl of dUTP tagged with either 0.2 mM Spectrum Red (now Abbott No. 02N34-050) or 0.2 mM Spectrum Green (now Abbott No. 02N32-050). The resulting mixture was vortexed, briefly centrifuged, and incubated for fourteen hours at 15˚C, before the nick translation reaction was stopped by heating the samples to 70˚C for ten minutes. At this point, 10 µl of Cot-1 DNA (Abbott No. 06J31-001) and 150 µl 100% ethanol were added, the mixture was again vortexed and briefly centrifuged before being placed at -80˚C for 30 minutes. The

resulting sample was centrifuged at 4˚C at 12,000 rpm for 30 minutes, the supernatant was discarded, and the pellet was resuspended in 20 µl hybridization buffer (500 µl formamide, 100 µl 20xSSC, 200 µl dextran sulfate, 200 µl dH_2O). The labeled probes were stored at -20˚C until use.

Tissue Preparation:

Both polyclonal tissue and monoclonal cell lines were used as material for FISH analysis with both sets of probes. Benign human tonsil and thymus tissue, fixed in formalin and embedded in paraffin, were obtained from the Molecular Diagnostics Service of the Yale Department of Pathology, where these tissues and cells are frequently used as controls in their clinical assays. 10 µm-thick sections were cut from the paraffin blocks and transferred to slides by the Yale Pathology Tissue Services, Research Histology Service.

The clonal epithelial cell lines HESC (a human endometrial line) and HCT116 (a human colon cancer line) were obtained and cultured in DMEM and McCoy's 5A Modified Media, respectively, with 10% FBS, 1% penicillin/streptomycin, and 1% glutarate added. These cell lines were grown directly on eight chambered slides, which were then washed with PBS, fixed with a 1:1 methanol:acetone mixture for five minutes, air-dried, and washed again in 2xSSC for sixty minutes before being stored at 4˚C.

The T cell lines Jurkat, HSB, HSC, SKW3, SUPT1, HUT78, and HPB-ALL were obtained by generous donation from the laboratory of Peter Cresswell. Ph.D., Yale University, Department of Immunobiology. The B cell line Raji was acquired

from the laboratory of George Miller, M.D., Yale University, Department of Pediatrics. These clonal lymphocyte lines were cultured in RPMI media with 10% FBS, 1% pen/strep, 1% glutarate. Cell pellets for each cell line were created by spinning 32 ml of each culture at 1500 rpm for three minutes, removing the supernatant, resuspending in 10 ml PBS, vortexing, centrifuging again at 1500 rpm for three minutes, removing the supernatant and resuspending in 10 ml of 10% buffered formalin. After fixing for one hour in formalin, the cultures were centrifuged again at 1500 rpm for three minutes, dehydrated with successive resuspensions in 10 ml 70%, 95%, and 100% ethanol before finally being resuspended in 40 µl Histogel (available at www.labstore.com, No. HG-4000). These cell pellets were cut into 10 µm sections with a microtome and placed on slides by Research Histology.

In the subsequent experiments conducted by Jonathan Murphy, the clonal B cell line Namalwa and the clonal T cell line Jurkat were cultured in the same fashion as the T cell lines above. 1 ml of the cell lines in RPMI media was centrifuged at 300 G for five minutes and the pellet was resuspended in 300 µl PBS. 200 µl of this suspension was added to a single well of an eight-well poly-D-lysine coated slide. The slide was centrifuged for five minutes at 200 G. The polystyrene vessel was removed from the slide, and the slide was submerged in 0.075 M KCl hypotonic solution for twelve minutes at 37˚C, followed by submersion in a 70% ethanol solution for five minutes. The slide was allowed to air dry, and was stored at 4 ˚C prior to use.

Sequential RNA and DNA FISH:

Only *DAPK1* and *EBF* probes were used for sequential RNA and DNA FISH, which was carried out on tonsil, thymus, the Jurkat and HUT78 cell pellets, and the HESC and HCT116 glass slides. All but the HESC and HCT116 glass slides required deparaffinization prior to hybridization.

Deparaffinization: Slides requiring deparaffinization were heated at 56˚C for 2-3 hours, before being immersed in xylene three times for ten minutes each at room temperature. Two final washes in 100% ethanol for five minutes at a time followed, before the slides were allowed to air dry.

Probe Hybridization: Following deparaffinization, if necessary, slides were pretreated in 2xSSC at 37˚C for one hour before immersion in 1:25 dilution of the protease Digestall3 (Zymed, now Invitrogen) for five minutes at room temperature. After protease digestion, slides were washed in PBS-T for two minutes, then fixed in 10% buffered formalin for one minute before a second wash in PBS-T for two minutes, all at room temperature. Following this washing, slides underwent sequential dehydration in 70%, 95%, and 100% ethanol for two minutes each at room temperature, and were allowed to air dry. At this point, 10 µl of either *DAPK1* or *EBF* probe solution, having been denatured previously at 73˚C for five minutes, were applied. A coverslip was sealed on with rubber cement, and the slides were placed on a plate heater at 37˚C, where they incubated for fourteen hours to allow RNA hybridization to occur. Following hybridization, the coverslips were removed, and

the slides underwent stringency washing in 0.5xSSC at 72˚C two times for five minutes each, followed by immersion in PBS-T for two minutes at room temperature. They were again fixed in 10% buffered formalin for one minute before a wash in PBS-T and sequential dehydration in 70%, 95%, and 100% ethanol for two minutes each at room temperature. Following air-drying, another 10 µl of the same probe solution, either *EBF* or *DAPK1*, was applied, this time labeled with the opposite color, either Spectrum Red or Spectrum Green. A coverslip was sealed on with rubber cement, and this time, the slides were incubated at the denaturation temperature of 73˚C on a plate heater for five minutes before incubation for fourteen hours at 37˚C to allow DNA hybridization to occur. A second stringency wash followed, with 0.5xSSC at 72˚C two times for five minutes each, followed by immersion in PBS-T for two minutes at room temperature. 15 µl of DAPI was added to stain the nuclei, a coverslip was applied, and the slides were stored at 4˚C. Microscopic examination was performed with an Olympus fluorescence microscope having polarized filters able to detect Spectrum Red and Spectrum Green.

DNA FISH Alone:

For experiments requiring DNA FISH alone, we used the VP 2000 (Abbott Molecular No. 02J11-060) slide processor utilized by the Molecular Diagnostics Service for clinical FISH assays. This instrument is calibrated to process formalin-fixed, paraffin-embedded tissue, so we were able to use slides prepared from tonsil, thymus, and the T-cell and B-cell pellets. Both *DAPK1* and *EBF* (as a control) and the TCR beta probes were used. On the VP 2000 Processor, slides were incubated

for sixty minutes at 55˚C, before immersion in xylene three times for ten minutes each at room temperature, and in 100% ethanol two times for five minutes each, also at room temperature. Slides then dried for five minutes at 55˚C, and were placed in a 0.2N HCl bath for fifteen minutes at room temperature. Next they were washed in a water bath for three minutes at room temperature, before immersion in pretreatment solution (VP 2000 reagent, Abbott Molecular 30-801250) for thirty minutes at 80˚C. Following pretreatment, slides were immersed in protease solution (Abbott Molecular 30-801255, 0.1 N HCl) for thirty minutes at 37˚C. They were then washed in 2xSSC for five minutes, and finally 70%, 85%, and 100% ethanol for one minute each, all at room temperature. After drying for five minutes at 55˚C, 10 µl of either *EBF* or *DAPK1* probes, or a mixture of 7µl VDR probes with 7 µl of either 3CR or 5CR were applied to a slide. A coverslip was placed on the slides, sealed with rubber cement, and slides were placed on a plate heater at 73˚C for five minutes before incubating at 39˚C for sixteen hours. Following hybridization, the coverslips were removed and the slides were placed in posthybridization buffer (67 ml 20xSSC, 547 ml ddH_2O, 2 ml Igepal CA 630) for 120 minutes at 74˚C. At this point, 10 µl DAPI was added, a coverslip was placed on the slides, and the slides were stored at 4˚C until microscopic examination.

In the subsequent experiments conducted by Jonathan Murphy, the probes generated by PCR were added to the 8-well slide containing clonal Namalwa or Jurkat cells, following the VP2000 protocol as described above. Following hybridization, the coverslips were removed and slides were placed in 0.4XSSC/0.3% NP-40 for two minutes at 73˚C, followed by 2XSSC/0.1% NP-40 at room

temperature for one minute. Slides were allowed to air dry for ten minutes at room temperature, at which point 10 µl DAPI was added, a coverslip was applied, and the slides were stored at 4˚C until microscopic examination.

Results

We tested both methods designed to detect cell clonality *in situ*, the first to assess allelic expression patterns of *DAPK1* and *EBF*, both of which are known to be subject to random monallelic gene expression, and the second to assess the status of allelic recombination in the T cell receptor beta locus.

To test the first method, probes corresponding to introns of the *DAPK1* and *EBF* genes were hybridized to benign tonsil and thymus tissue as well as to clonal T cell pellets and clonal epithelial cells. These experiments either involved sequential RNA and DNA FISH with distinct probes labeled with different fluorophors to distinguish DNA from RNA, or involved DNA FISH alone with a single fluorophor. The DNA FISH served to identify the two alleles in each cell and provided a control for accessibility of each allele for hybridization. RNA FISH allowed the determination of whether a cell expressed *DAPK1* or *EBF* at one or both alleles. Only cells containing two distinct DNA FISH hybridization loci and at least one RNA FISH hybridization locus could be used for clonal assessment. Cells not meeting these criteria were disregarded.

Under the experimental conditions used, initial attempts at sequential RNA and DNA FISH resulted in no cells that could be counted for clonal assessment. A representative image is shown in Figure 3, which demonstrates DAPI signal identifying lymphocyte nuclei, and no appreciable green signal identifying the *DAPK1*-encoding DNA alleles or red signal identifying nascent RNA being transcribed at one or both alleles.

In view of these results, RNA FISH was set aside, and DNA FISH was performed alone on tonsil tissue using the VP 2000 Processor, with green-labeled probes to identify the *DAPK1* or *EBF* alleles in separate hybridizations. These hybridizations were more successful. Representative images are shown in Figures 4 and 5, which illustrate DNA FISH on tonsil tissue using *DAPK1* and *EBF* probes, respectively.

To test the second method, which involved analysis of recombination in the TCR beta gene and required only DNA FISH without RNA FISH, probes corresponding to the VDR region between V29-1 and D1 of the TCR beta locus along with probes for control regions both 5' and 3' of the *TCRB* locus were hybridized to tonsil, thymus, and monoclonal T-cell pellets using the VP 2000 Processor. The VDR probes were labeled with either red or green fluorophor, using a different color from the control region probes to distinguish the test signals from the control signals. Only cells containing two control region signals identifying the TCR beta locus could be used for clonal assessment. Successful analyses should reveal a VDR signal present next to one or neither of the control region signals, depending on whether one or both alleles had undergone complete V-DJ rearrangement and lost the region of DNA to which the VDR probe should hybridize.

Under the experimental conditions used, no cells demonstrated VDR signal next to a control region signal in a cell containing two control region signals. Very few cells contained two control region signals, and the presence of nonspecific background signal rendered most experiment iterations unusable despite stringency washings. A representative image is shown in Figure 6, which is a composite of

green, red, and blue filtered images of benign thymus tissue, to which red-labeled VDR and green-labeled 3CR probes have been hybridized. While there are multiple red signals, their size and distribution indicates that they most likely represent nonspecific background signal instead of true VDR hybridization. As a comparison, Figure 7 is a green-filtered image of thymus tissue to which green-labeled 3CR probes have been hybridized. While there are certainly some larger signals that may be non-specific, the presence of several smaller signals, often in pairs, suggests that at least some of the green signals reflect true hybridization. To be sure of this, it would be helpful to see at least some cells with red VDR signals, but under these experimental conditions none were identified.

In a subsequent experiment, conducted by Jonathan Murphy, the second method was again tested, this time using PCR-generated probes from the 64 kb test region between V29-1 and D1 at the *TCRB* locus and from the 3CR control region. The test region probes were labeled green, and the control region probes were labeled red. These probes were hybridized to clonal Namalwa B cells and to clonal Jurkat T cells. Figures 8 and 9 show Namalwa B cells photographed in the fluorescence microscope using a red filter and a green filter, respectively. The red-filtered image demonstrated two red signals in most of the Namalwa cells, most likely identifying the control regions located just 3' to the two alleles of the T-cell receptor beta locus. The green-filtered image demonstrated two green signals in identical locations, with occasional extraneous signals that most likely represent artifact. The green signals that overlap with the red signals identified on the red filter most likely represent true hybridization to the T cell receptor beta locus. The test region is intact at both alleles

in these B cells, which have not undergone recombination within the *TCRB* loci. By comparison, Figures 10 and 11 show Jurkat T cells photographed using a red filter and a green filter, respectively. The red-filtered image again demonstrated two red signals in most cells, likely identifying the control region 3' to the *TCRB* locus, but the green-filtered image demonstrates only one green signal that overlaps with the red signals in each cell, along with occasional artifactual signals that do not overlap with red signals. The presence of only one overlapping signal in these T-cells is consistent with the fact that one of the *TCRB* alleles has successfully undergone VDJ recombination, and the second allele has not undergone V-DJ joining. While these images represent analyses of only two clonal cell lines, they suggest that the use of DNA FISH to determine the status of T cell receptor beta recombination may, with additional refinements, be used to assess clonality of T cells *in situ*.

Discussion

In this thesis, two promising methods of detecting a monoclonal population of human cells have been described, although their feasibility and usefulness cannot yet be determined as the molecular techniques are still being refined. The promise of both methods relies upon several necessary assumptions, which could not be affirmed or rejected in this study. Unless these assumptions are rejected by subsequent work, the clonality assays proposed here have the potential to become valuable clinical tools to aid in the diagnosis of malignancy, or at the very least useful methods for studying malignant transformation and preneoplastic lesions.

Potential pitfalls to the use of random monoallelic gene expression as an assay for monoclonality would include the finding that this phenomenon occurs in too few cells within a tissue, that the phenomenon is limited to only a small set of tissues, that the progeny of a dividing cell do not retain the same pattern of allelic inactivation, or that even a single cell can change its pattern of allelic inactivation during its lifetime. While any of these pitfalls could turn out to be real, the experiments described in this thesis did not fail because of any of them.

Similarly, using FISH to analyze T cell receptor gene rearrangement as an assay for clonality of T cell populations would be problematic if the ratio of human T cells with rearrangements of one TCR beta allele to those with rearrangements of both is either too small or too large to allow statistically significant analysis of variations within cell populations, or that a cell with only one allelic rearrangement early in its development can frequently undergo a second rearrangement during its later life. The experiments described in this thesis found no evidence to suggest that

these problems applied, but neither did they rule out the possibility of these problems potentially complicating the assay.

The inability of this experiment to support or reject the necessary assumptions underlying both proposed methods relies solely on the failure of our experimental technique. In further experiments, we will attempt to increase the sensitivity and specificity of the hybridization assays by pursuing one or several of the following methods. First, the use of BAC clones as sources of probe, while successful in the Gimelbrant *et al.* study, may result in the labeling of too much bacterial DNA and not enough target human DNA to cause enough specific hybridization versus nonspecific binding of probe. Instead, we can either purify the BAC DNA away from bacterial genomic DNA, subclone it, or use pooled, labeled PCR products as probes that specifically target the test regions. The resulting reduction of background signal may allow more successful visualization of true hybridization. As described in the Results section, our initial experiments with pooled, labeled PCR products have already been very promising. Additionally, the fluorophors we used may not be bright enough to be visible when hybridized to such short DNA regions as the VDR probe, which is only 16 kb long. Using so-called indirectly-labeled probes rather than probes directly labeled with fluorescent tags would allow the use of signal amplification steps to increase the signal of the shorter regions and allow better visualization under the microscope. For example, biotin-conjugated nucleotides incorporated into the probes can be detected by incubating the hybridized probe with FITC-avidin, followed by biotinylated anti-avidin, and a final round of FITC-avidin (65). One or several of

these methods may improve the assays so that we can assess clonality on a routine basis.

One of the more immediate areas where either method of *in situ* clonality analysis would be useful is for very small biopsy specimens in which there are few cells available for analysis. As the use of minimally invasive biopsy techniques become more prevalent and biopsy samples become smaller, the challenge of diagnosis on low numbers of cells has grown. Additionally, some neoplastic disorders contain very few neoplastic cells, the bulk of the mass in these tumors being made up of tissue reacting to the presence of the neoplastic cells. Hodgkin's disease is an example of such a disorder, in that the neoplastic Reed-Sternberg cell usually makes up less than one percent of the cells in the mass. Finally, certain cancers are being identified at earlier and earlier stages, when the total number of malignant cells may be quite low.

A disease that illustrates all of the above diagnostic problems and for which *in situ* analysis of clonality would be very useful is the assessment and diagnosis of cutaneous T-cell lymphoma (CTCL). CTCL is a relatively rare malignancy, with an estimated annual incidence of 1:100,000; treatment options are limited and prognosis is often poor, depending upon the particular type (72). The diagnosis is often difficult, relying primarily on histologic and immunophenotypic features that can be difficult to distinguish from reactive lymphocytic responses, such as chronic dermatoses, which are quite common (12, 13). Because of the difficulty in making the distinction between CTCL and chronic dermatoses, the actual number of skin biopsy specimens for which CTCL enters the differential diagnoses is very high. As

the amount of tissue available for examination is often very small, an *in situ* assay that could distinguish a reactive process from a monoclonal proliferation would be valuable to clinicians and researchers. To this end, in 2003 Magro *et al*. published their work using *in situ* RT-PCR to examine the TCR beta rearrangement patterns, with primers distinguishing the 25 possible V region recombinations (26). They examined the tissue of 28 patients with cutaneous T-cell infiltrates including benign lesions as well as CTCL, diagnosed by clinical, histologic, and immunophenotypic analysis. Of the eight cases of primary CTCL lymphoma, seven were identified as monoclonal by *in situ* RT-PCR, while the other case was found to be "biclonal."

As discussed earlier, there are several limitations to the potential for clinical adoption of *in situ* RT-PCR as currently performed-–limitations that a DNA FISH-based assay would overcome. The amount of tissue required for DNA FISH is much less, and DNA FISH does not require sequential iterations to distinguish the 25 possible V-DJ rearrangements. However, the DNA FISH-based assay as we have described it could not identify the particular VDJ rearrangement of a monoclonal population, as the RT *in situ* PCR-based assay can. Nevertheless, this property is probably not necessary for a useful assay of clonality in T cell disorders. This opinion is based on the diagnostic utility of immunohistochemical detection of restricted kappa versus lambda light chain immunoglobulin associated with monoclonal B cell processes. Therefore, if our proposed assay for clonal TCR beta gene rearrangements can be performed successfully, it could facilitate clinical examination of cutaneous T cell proliferations with the possibility of emerging as an important diagnostic assay. In the long run, if the *in situ* assay using random

monoallelic gene expression can be perfected, this too would be applicable to cutaneous T cell disorders.

We conclude that although limits of time prevented the full development of the *in situ* assays to identify clonality proposed here, both have the potential to become valuable tools in research and clinical practice. They certainly merit further investigation.

References

1. Fialkow PJ. "Clonal origin of human tumors." *Biochim Biophys Acta* 458:283-321, 1976.

2. Levy R, Warnke R, Dorfman RF, Haimovich J. "The monoclonality of human B-cell lymphomas." *J Exp Med* 145:1014-1028, 1977.

3. Mashal RD, Lester SC, Sklar J. "Clonal analysis by study of X chromosome inactivation in formalin-fixed paraffin-embedded tissue." *Cancer Res* 53:4676-4679, 1993.

4. Sklar J, Longtine J. "The clinical significance of antigen receptor gene rearrangements in lymphoid neoplasia." *Cancer* 70:1710-1718, 1992.

5. Vogelstein B, Fearon ER, Hamilton SR, Preisinger AC, *et al.* "Clonal analysis using recombinant DNA probes from the X-chromosome." *Cancer Res* 47:4806-4813, 1987.

6. Wainscoat JS, Fey MF. "Assessment of clonality in human tumors: a review."*Cancer Res* 50:1355-1360, 1990.

7. Aisenberg AC, Wilkes B, Jacobson J, Harris N. "*Immunoglobulin gene rearrangements in adult non*-Hodgkin's lymphoma." *Am J Med* 82: 738-44, 1987.

8. Bourguin A, Tung R, Galili N, Sklar J. "Rapid, nonradioactive detection of clonal T-cell receptor gene rearrangements in lymphoid neoplasms." *Proc Natl Acad Sci USA* 87:8536-8540, 1990.

9. Longtine J, Fox E, Reynolds C, Sklar J. "Molecular analysis of DNA rearrangements in leukemias and non-Hodgkin's lymphomas." *Curr Protoc Hum Genet* Chapter 10:Unit 4.1-28, 2000.

10. Diaz-Cano SJ, Blanes A, Wolfe HJ. "PCR techniques for clonality assays." *Diagn Mol Pathol* 10:24-33, 2001.

11. Levy R, Levy S, Cleary ML, Carroll W, Kon S, Bird J, Sklar J. "Somatic mutation in human B-cell tumors." *Immunol Rev* 96:43-58, 1987.

12. Burg G, Kempf W, Cozzio A, Feit J, Willemze R, *et al*. "WHO/EORTC classification of cutaneous lymphomas 2005: histological and molecular aspects." *J Cutan Pathol* 32:647-674, 2005.

13. Smoller BR, Santucci M, Wood GS, Whittaker SJ. "Histopathology and genetics of cutaneous T-cell lymphoma." *Hematol Oncol Clin North Am* 17:1277-1311, 2003.

14. Aifantis I, Buer J, von Boehmer H, Azogui O. “Essential role of the pre-T cell receptor in allelic exclusion of the T cell receptor beta locus.” *Immunity* 5:601-607, 1997.

15. Bagg A. “Immunoglobulin and T-cell receptor gene rearrangements: minding your B's and T's in assessing lineage and clonality in neoplastic lymphoproliferative disorders.” *J Mol Diagn* 8:426-429, 2006.

16. ten Boekel E, Melchers F, Rolink A. “The status of Ig loci rearrangements in single cells from different stages of B cell development.” *Int Immunol* 7:1013-1019, 1995.

17. Born W, White J, Kappler J, Marrack P. “Rearrangement of IgH genes in normal thymocyte development.” *J Immunol* 140:3228-3232, 1988.

18. Cleary ML, Meeker TC, Levy S, Lee E, Trela M, Sklar J, Levy R. “Clustering of extensive somatic mutations in the variable region of an immunoglobulin heavy chain gene from a human B cell lymphoma.” *Cell* 44:97-106, 1986.

19. Ghia P, ten Boekel E, Sanz E, de la Hera A, Rolink A, Melchers F. “Ordering of human bone marrow B lymphocyte precursors by single-cell polymerase chain reaction analyses of the rearrangement status of the immunoglobulin H and L chain gene loci.” *J Exp Med* 184:2217-2229, 1996.

20. Glusman G, Rowen L, Lee I, Boysen C, Roach JC, Smit AF, Wang K, Koop BF, Hood L “Comparative genomics of the human and mouse T-cell receptor loci.” *Immunity* 15(3): 337-349, 2001.

21. Hesslein DG, Schatz DG. “Factors and forces controlling V(D)J recombination.” *Adv Immunol* 78:169-232, 2001.

22. Joachims ML, Chain JL, Hooker SW, Knott-Craig CJ, Thompson LF. “Human alpha beta and gamma delta thymocyte development: TCR gene rearrangements, intracellular TCR beta expression, and gamma delta developmental potential--differences between men and mice.” *J Immunol* 176:1543-1552, 2006.

23. Korsmeyer SJ, Hieter PA, Ravetch JV, Poplack DG, Waldmann TA, Leder P. “Developmental hierarchy of immunoglobulin gene rearrangements in human leukemic pre-B-cells.” *Proc Natl Acad Sci USA* 78:7096-7100, 1981.

24. Krangel MS. “Gene segment selection in V(D)J recombination: accessibility and beyond.” *Nat Immunol* 4:624-630, 2003.

25. LeBien TW, Elstrom RL, Moseley M, Kersey JH, Griesinger F. "Analysis of immunoglobulin and T-cell receptor gene rearrangements in human fetal bone marrow B lineage cells." *Blood* 76:1196-1200, 1990.

26. Magro CM, Nuovo GJ, Crowson AN. "The utility of the *in situ* detection of T-cell receptor beta rearrangements in cutaneous T-cell-dominant infiltrates." *Diagn Mol Pathol* 12(3):133-141, 2003.

27. Pernis B, Chiappino G, Kelus AS, Gell PG. "Cellular localization of immunoglobulins with different allotypic specificities in rabbit lymphoid tissues." *J Exp Med* 122:853-876, 1965.

28. Petrie HT, Livak F, Burtrum D, Mazel S. "T cell receptor gene recombination patterns and mechanisms: cell death, rescue, and T cell production." *J Exp Med* 182:121-127, 1995.

29. Rajewsky K. "Clonal selection and learning in the antibody system." *Nature* 381:751-758, 1996.

30. Trainor KJ, Brisco MJ, Wan JH, Neoh S, Grist S, Morley AA. "Gene rearrangement in B- and T-lymphoproliferative disease detected by the polymerase chain reaction." *Blood* 78:192-196, 1991.

31. Volpe JM, Kepler TB. "Large-scale analysis of human heavy chain V(D)J recombination patterns." *Immunome Res*, 4:3, 2008.

32. Theodorou I, Delfau-Larue MH, Bigorgne C, Lahet C, Cochet G, *et al*. "Cutaneous T-cell infiltrates: analysis of T-cell receptor gamma gene rearrangement by polymerase chain reaction and denaturing gradient gel electrophoresis." *Blood* 86(1):305-10, 1995.

33. Korsmeyer SJ, Waldmann TA. "Immunoglobulin genes: rearrangement and trans-location in human lymphoid malignancy." *J Clin Immunol*. 4:1-11, 1983.

34. Rodriguez J, Pugh WC, Cabanillas F. "T-cell-rich B-cell lymphoma." *Blood* 82:1586-1589, 1993.

35. Hunt JL. "Molecular pathology in anatomic pathology practice: a review of basic principles." *Arch Pathol Lab Med* 132:248-260, 2008.

36. Wutz A, Gribnau J. "X inactivation Xplained." *Curr Opin Genet Dev* 17:387-393, 2007.

37. Nomura S, Kaminishi M, Sugiyama K, Oohara T, Esumi H. "Clonal analysis of isolated intestinal metaplastic glands of stomach using X-linked polymorphism." *Gut* 42:663-8, 1998.

38. Raab-Traub N., Flynn K. "The structure of the termini of the Epstein-Barr virus as a marker of clonal cellular proliferation." *Cell* 47: 883-889, 1986.

39. Weiss LM, Strickler JG, Warnke RA, Purtilo DT, Sklar J. "Epstein-Barr viral DNA in tissue of Hodgkin's disease." *Am J Pathol* 129:86-91, 1987.

40. Brechot C, Hadchouel M, Scotto J, Fonck M, Potedt F, Vyas GN, Tiollais P. "State of hepatitis B virus DNA in hepatocytes of patients with hepatitis B surface antigen-positive and -negative liver diseases." *Proc Nati Acad Sci USA* 78:3906-3910, 1981.

41. Leedham SJ, Preston SL, McDonald SAC, Elia G, Bhandari P, Poller D, Harrison R, Novelli MR, Jankowski JA, Wright NA. "Individual crypt genetic heterogeneity and the origin of metaplastic glandular epithelium in human Barrett's esophagus." *Gut* 57:1041-1048, 2008.

42. Zhuang Z, Vortmeyer AO, Mark EJ, Odze R, Emmert-Buck MR, Merino MJ, Moon H, Liotta LA, and Duray PH. "Barrett's esophagus: metaplastic cells with loss of heterozygosity at the APC gene locus are clonal precursors to invasive adenocarcinoma." *Cancer Res* 56:1961–1964, 1996.

43. Kohda A, Taguchi H, Okumura K. "Visualization of biallelic expression of the imprinted SNRPN gene induced by inhibitors of DNA methylation and histone deacetylation." *Biosci Biotechnol Biochem*. 65(5):1236-9, 2001.

44. Morison IM, Ramsay JP, Spencer HG. "A census of mammalian imprinting." *Trends Genet* 21:457-465, 2005.

45. O'Neill MJ. "The influence of non-coding RNAs on allele-specific gene expression in mammals." *Hum Mol Genet* 14(1):R113-120, 2005.

46. Reik W, Walter J. "Genomic imprinting: parental influence on the genome." *Nat Rev Genet* 2:21-32, 2001.

47. Bix M, Locksley RM. "Independent and epigenetic regulation of the interleukin-4 alleles in CD4+ T cells." *Science* 281:1352-1354, 1998.

48. Chess A, Simon I, Cedar H, Axel R. "Allelic inactivation regulates olfactory receptor gene expression." *Cell* 78:823-834, 1994.

49. Gimelbrant A, Ensminger AW, Qi P, Zucker J, Chess A. "Monoallelic expression and asynchronous replication of p120 catenin in mouse and human cells." *J Biol Chem* 280:1354-1359, 2005.

50. Gimelbrant A, Hutchinson JN, Thompson BR, Chess A. "Widespread monoallelic expression on human autosomes." *Science* 318:1136-1140, 2007.

51. Hollander GA, Zuklys S, Morel C, Mizoguchi E, *et al*. "Monoallelic expression of the interleukin-2 locus." *Science* 279:2118-2121, 1998.

52. Rhoades KL, Singh N, Simon I, Glidden B, Cedar H, Chess A. "Allele-specific expression patterns of interleukin-2 and Pax-5 revealed by a sensitive single-cell RT-PCR analysis." *Curr Biol* 10:789-792, 2000.

53. Magro C, Crowson AN, Porcu P, Nuovo GJ. "Automated kappa and lambda light chain mRNA expression for the assessment of B-cell clonality in cutaneous B-cell infiltrates: its utility and diagnostic application." *J Cutan Pathol.* 30(8): 504-11, 2003.

54. Nuovo GJ, Morrison C, Porcu P, Caligiuri MA, Suster S. "*In situ* determination of T cell receptor beta expression patterns." *J Histochem Cytochem* 49(2):139-145, 2001.

55. Lichter P, Cremer T, Borden J, Manuelidis L, Ward DC. "Delineation of individual human chromosomes in metaphase and interphase cells by *in situ* suppression hybridization using recombinant DNA libraries." *Hum Genet* 80:224-234, 1988.

56. Pinkel D, Landegent J, Collins C, Fuscoe J, Segraves R, Lucas J, Gray J. "Fluorescence *in situ* hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4." *Proc Natl Acad Sci USA* 85:9138-9142, 1988.

57. Takizawa T, Gudla PR, Guo L, Lockett S, Misteli T. "Allele-specific nuclear positioning of the monoallelically expressed astrocyte marker GFAP." *Genes Dev* 22:489-498, 2008.

58. Meiers I, Singh H, Hossian D, Lang K, Qian J, Verhest AP, Bostwick DG. "Improved filter method for urine sediment detection of urothelial carcinoma by fluorescence *in situ* hybridization." *Arch Pathol Lab Med.* 131(10):1574-7, 2007.

59. Braidotti G. "RNA-FISH to analyze allele-specific expression." *Methods Mol Biol.* 181:169-80, 2001.

60. Capodieci P, Donovan M, Buchinsky H, Jeffers Y, Cordon-Cardo C, Gerald W, Edelson J, Shenoy SM, Singer RH. "Gene expression profiling in single cells within tissue." *Nat Methods* 2:663-665, 2005.

61. Chakalova L, Carter D, Fraser P. "RNA fluorescence in situ hybridization tagging and recovery of associated proteins to analyze in vivo chromatin interactions." *Methods Enzymol* 375**:** 479-493, 2004.

62. Kislauskis EH, Li Z, Singer RH, Taneja KL. “Isoform-specific 3'-untranslated sequences sort alpha-cardiac and beta-cytoplasmic and actin messenger RNAs to different cytoplasmic compartments.” *J Cell Biol* 123:165-172, 1993.

63. Levsky JM, Shenoy SM, Chubb JR, Hall CB, Capodieci P, Singer RH. “The spatial order of transcription in mammalian cells.” *J Cell Biochem.* 102(3):609-17, 2007.

64. Levsky JM, Shenoy SM, Pezo RC, Singer RH. “Single-cell gene expression profiling.” *Science* 297:836-840, 2002.

65. van Raamsdonk CD, Tilghman SM. “Optimizing the detection of nascent transcripts by RNA fluorescence in situ hybridization.” *Nucleic Acids Res* 29:E42-42, 2001.

66. Smith KP, Byron M, Johnson C, Xing Y, Lawrence JB. “Defining early steps in mRNA transport: mutant mRNA in myotonic dystrophy type I is blocked at entry into SC-35 domains.” *J Cell Biol* 178:951-964, 2007.

67. Nishio J, Althof PA, Bailey JM, Zhou M, Neff JR, Barr FG, Parham DM, Teot L, Qualman SJ, Bridge JA. “Use of a novel FISH assay on paraffin-embedded tissues as an adjunct to diagnosis of alveolar rhabdomyosarcoma.” *Lab Invest* 86:547-556, 2006.

68. Malissen M, Trucy J, Jouvin-Marche E, Cazenave PA, Scollay R, Malissen B. “Regulation of TCR alpha and beta gene allelic exclusion during T-cell development.” *Immunol Today* 13:315-322, 1992.

69. Lefranc, MP, Lefranc G. “The T cell receptor FactsBook.” London, UK: Academic Press, p. 398, 2001.

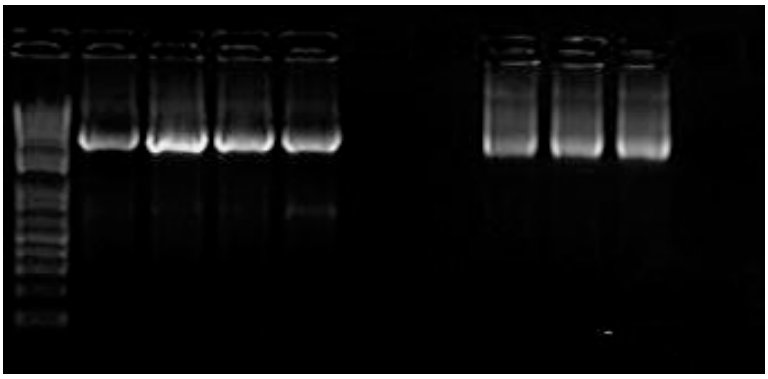
70. Hazenberg MD, Verschuren MCM, Hamann D, *et al*. “T cell receptor excision circles as markers for recent thymic emigrants: basic aspects, technical approach, and guidelines for interpretation.” *J Mol Med* 79:631-640, 2001.

71. Soulier J, Clappier E, Cayuela J-M, Regnault A, Garcia-Peydro M, et al. “HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL).” *Blood* 106: 274-286, 2005

72. Willemze R, Jaffe ES, Burg G, Cerroni L, Berti E, *et al*. “WHO-EORTC classification for cutaneous lymphomas.” *Blood* 105:3768-3785, 2005.

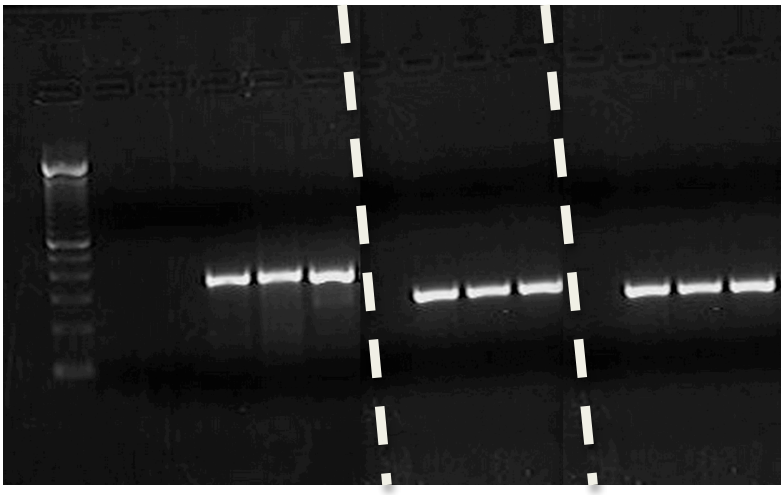
Figures:

Figure 1.



BAC Clone Verification: 2% agarose gel electrophoresis with PCR products of BAC clone DNA using probes complementary to *DAPK1* and *EBF* sequences. Lane 1, 1 kb ladder of size markers. Lanes 2-6, PCR products of DNA extracted from BAC clone RP11-107G16 with primers targeting *DAPK1* reveals ~3.5kb product, with a water control in Lane 6. Lane 7, blank. Lanes 8-11, PCR products of DNA extracted from BAC clone RP11-155P16 with primers targeting *EBF* reveals ~3.3kb product, with a water control in Lane 11.

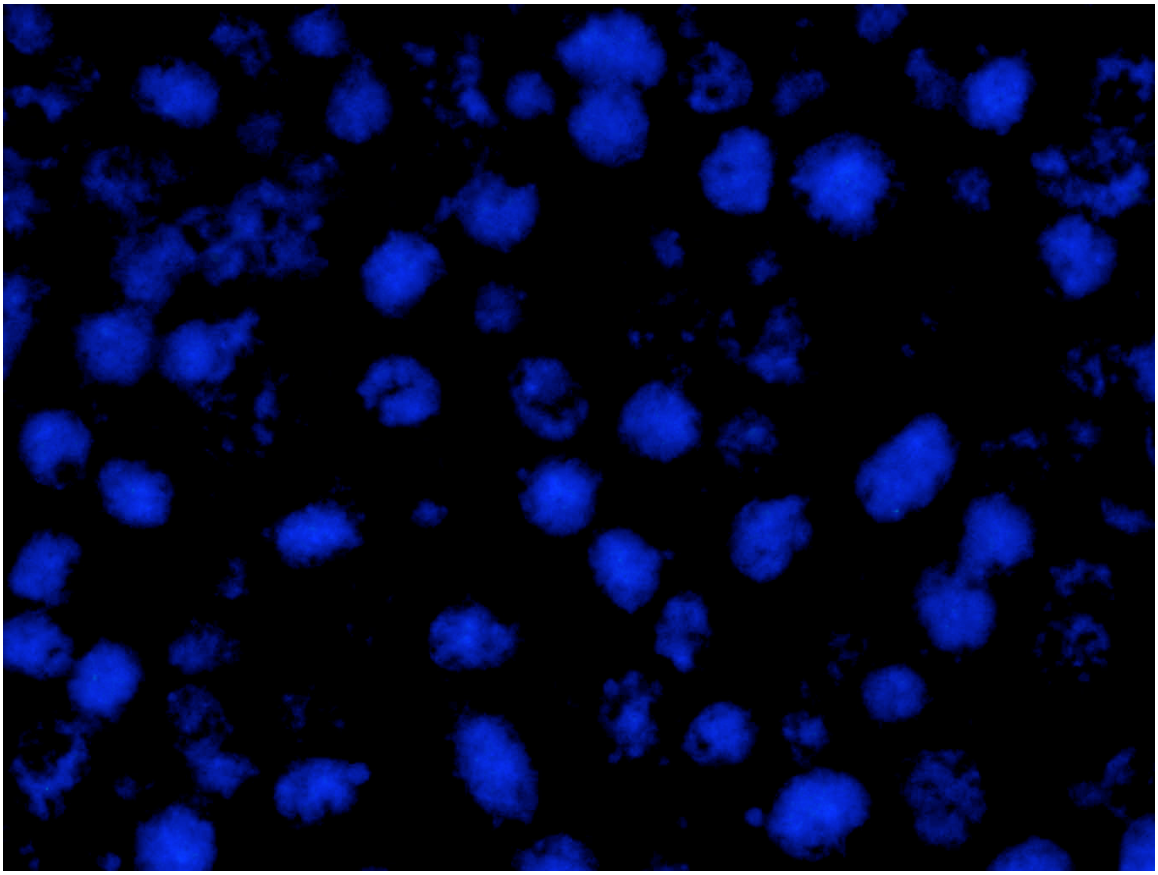
Figure 2.



BAC Clone Verification, cont'd: Composite 2% agarose gel electrophoresis of PCR products of BAC clone DNA using probes complementary to VDR, 3CR, and 5CR sequences. Lane 1, 100 bp ladder. Lane 2, blank. Lanes 3-6, PCR products of DNA extracted from BAC clone CTD-3217E23 with primers targeting VDR reveals ~300 bp product, with a water control in Lane 3. Lanes 7-10, PCR products of DNA extracted from BAC clone RP11-368I15 with primers targeting 3CR reveals ~250 bp

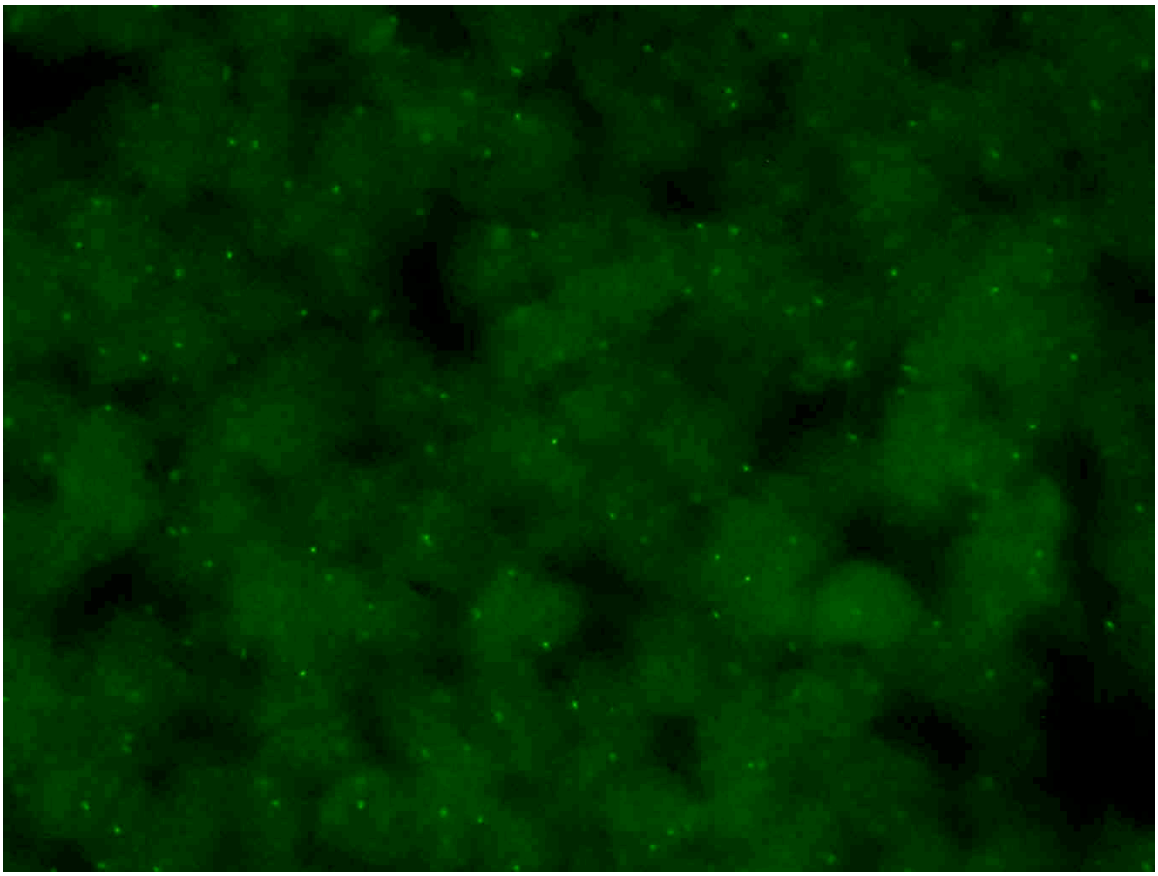
product, with a water control in Lane 7. Lanes 11-14, PCR products of DNA extracted from BAC clone RP11-10L5 with primers targeting 5CR reveals ~250 bp product, with a water control in Lane 11.

Figure 3.



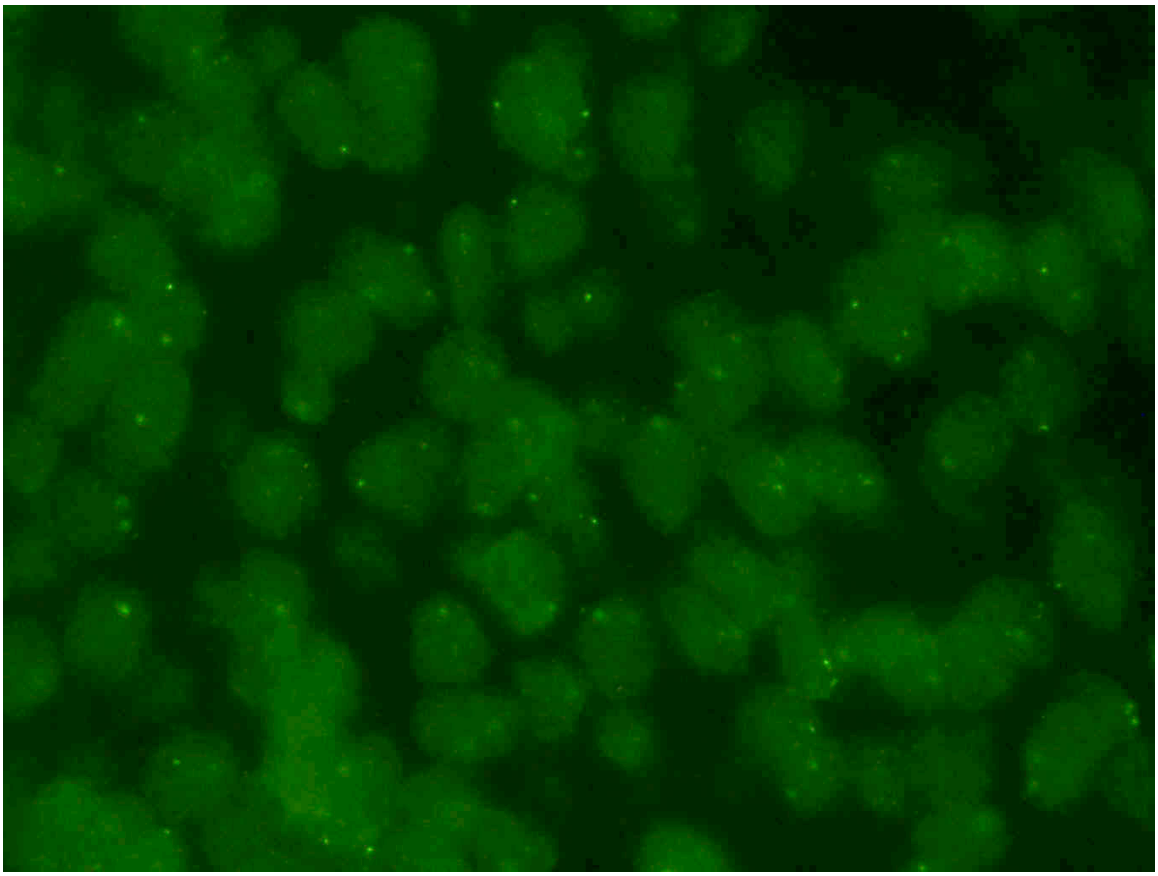
DAPK1 Sequential RNA and DNA FISH: Representative composite image of blue-, red-, and green-filtered microscopy of a section from a monoclonal HSB T cell pellet hybridized first to red-labeled *DAPK1* probe in RNA FISH, followed by green-labeled *DAPK1* probe in DNA FISH. Nuclei have been stained blue with DAPI, which binds to DNA and therefore marks the entire nucleus of interphase cells. Neither red nor green signals are visible, suggesting that little, if any hybridization had occurred. A significant amount of nonspecific and non-nuclear red and green signal was present, not visible in this image. The only signals suggestive of true hybridization occurred with DNA FISH using the VP 2000 processor, and not with sequential RNA and DNA FISH (see below). (x600 magnification)

Figure 4.



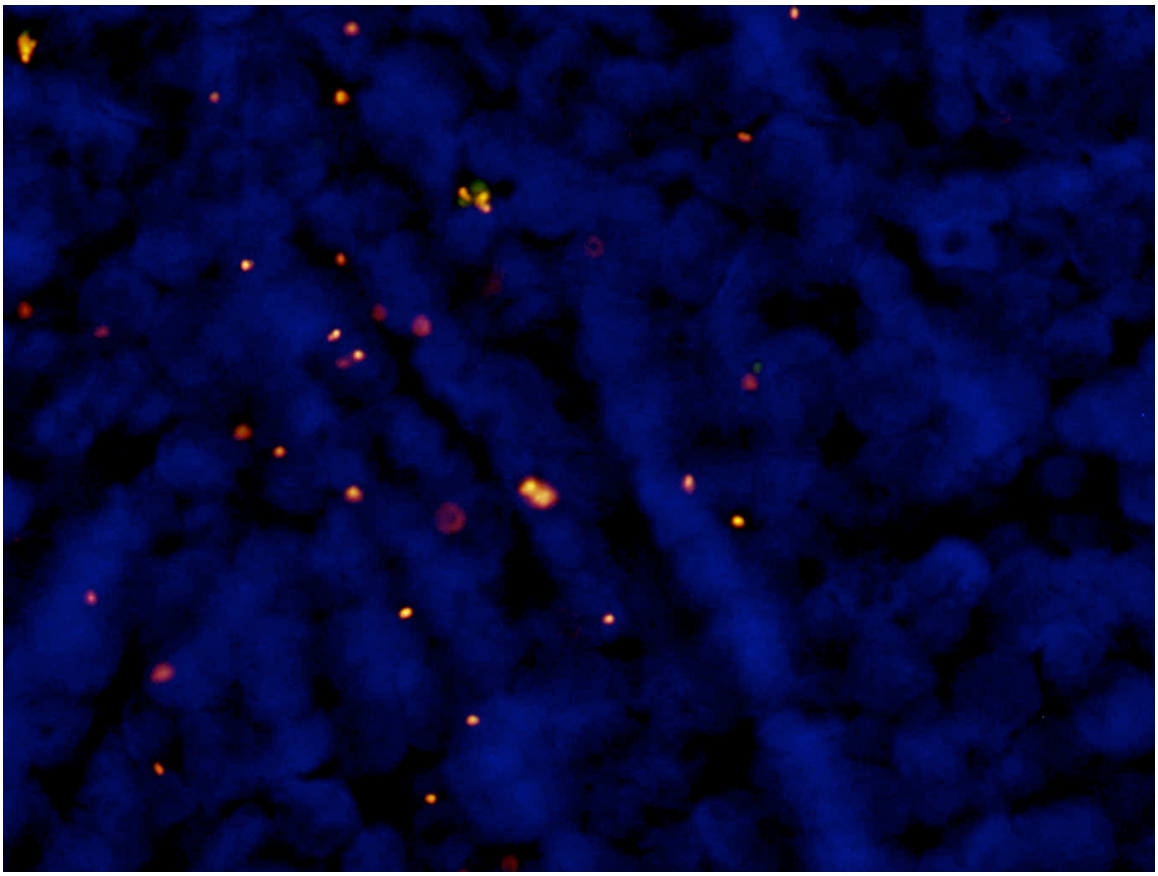
DAPK1 DNA FISH Control: Green-filter microscopy of benign tonsil tissue hybridized to *DAPK1* probes in DNA FISH using the VP 2000 processor. Multiple cells display two green signals suggestive of *bone fide* hybridization, although simultaneous RNA FISH would help to confirm this. The extent of hybridization is better determined by direct inspection under the microscope, rather than in photographs, because of the hybridization signals lying in several focal plains. (x600 magnification)

Figure 5.



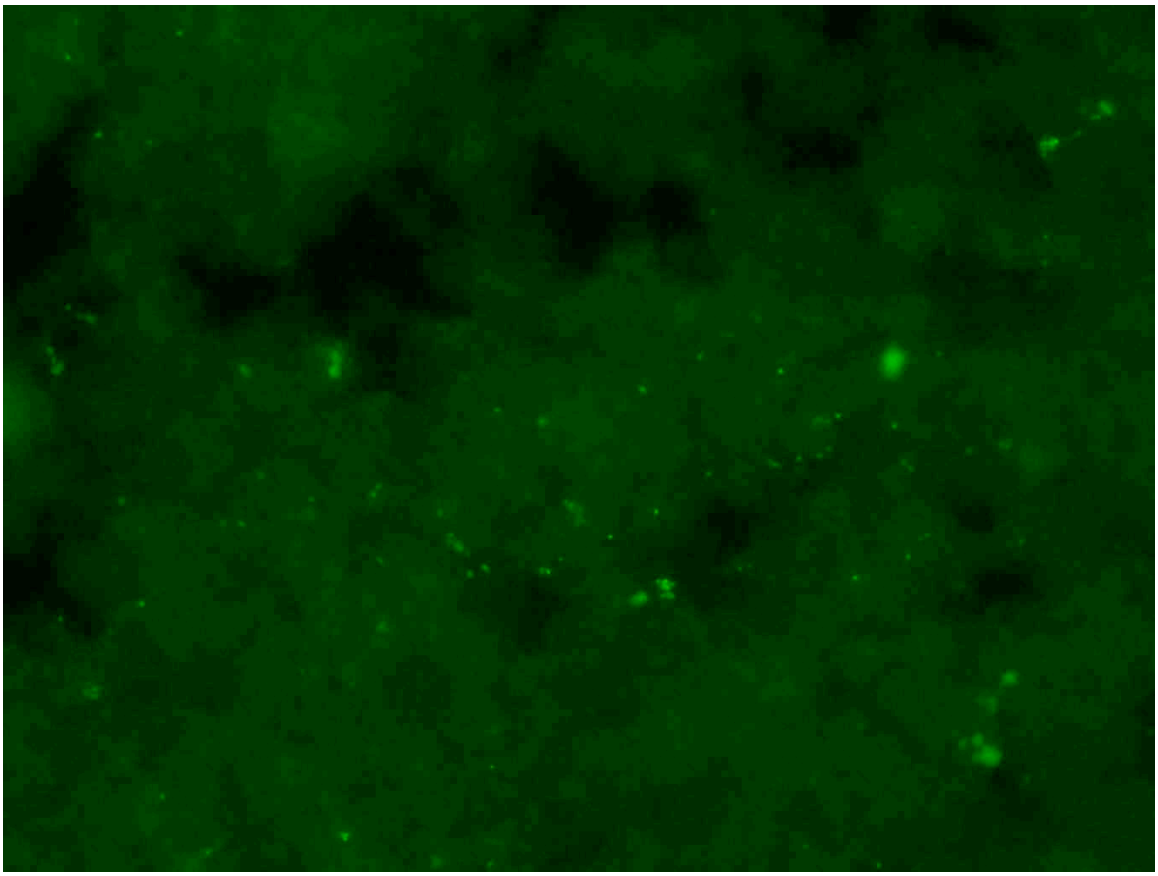
EBF DNA FISH Control: Green-filter microscopy of benign tonsil tissue hybridized to *EBF* probes in DNA FISH using the VP 2000 processor. Again, multiple cells display two green signals suggesting *bona fide* hybridization. (x600 magnification)

Figure 6.



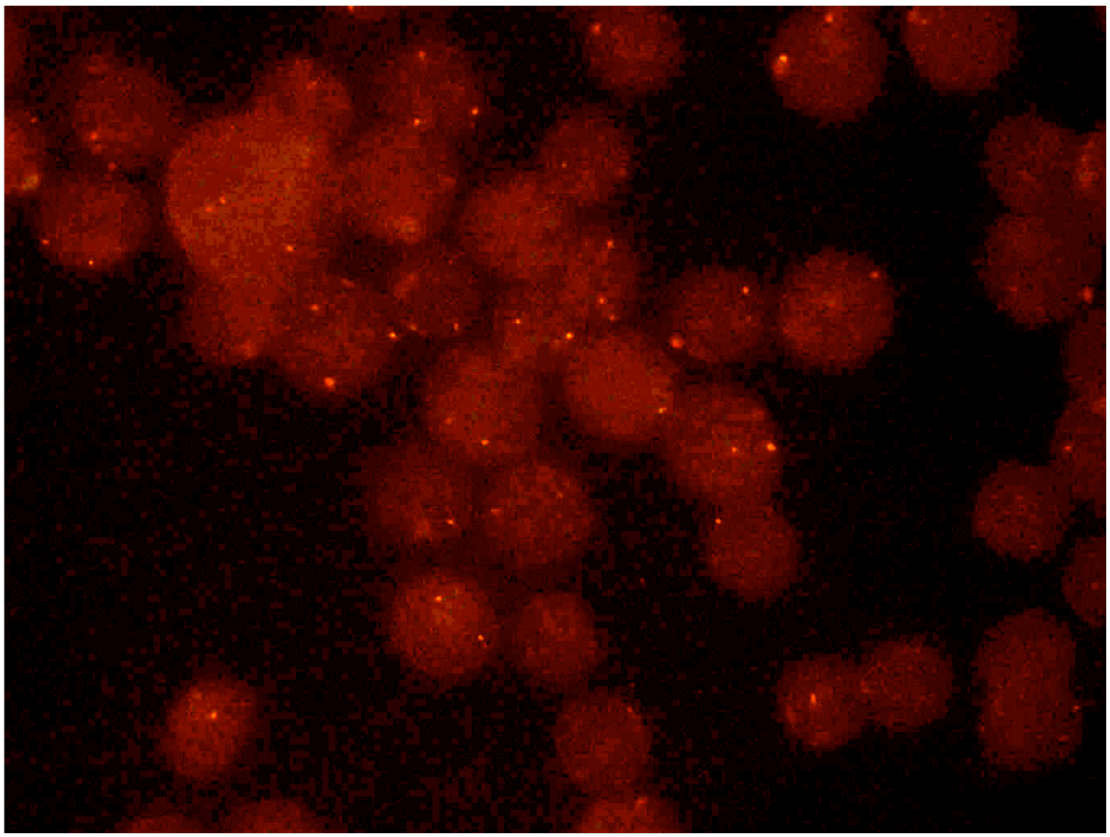
TCR Beta DNA FISH: Composite image of blue-, green-, and red-filtered microscopy of benign thymus tissue hybridized to red-labeled VDR probes and green-labeled 3CR probes by DNA FISH on the VP 2000 processor. Multiple red signals are present, but the lack of cells with two green (control) signals and the large size of several of the red signals indicate that these signals are unlikely to represent true hybridization.

Figure 7.



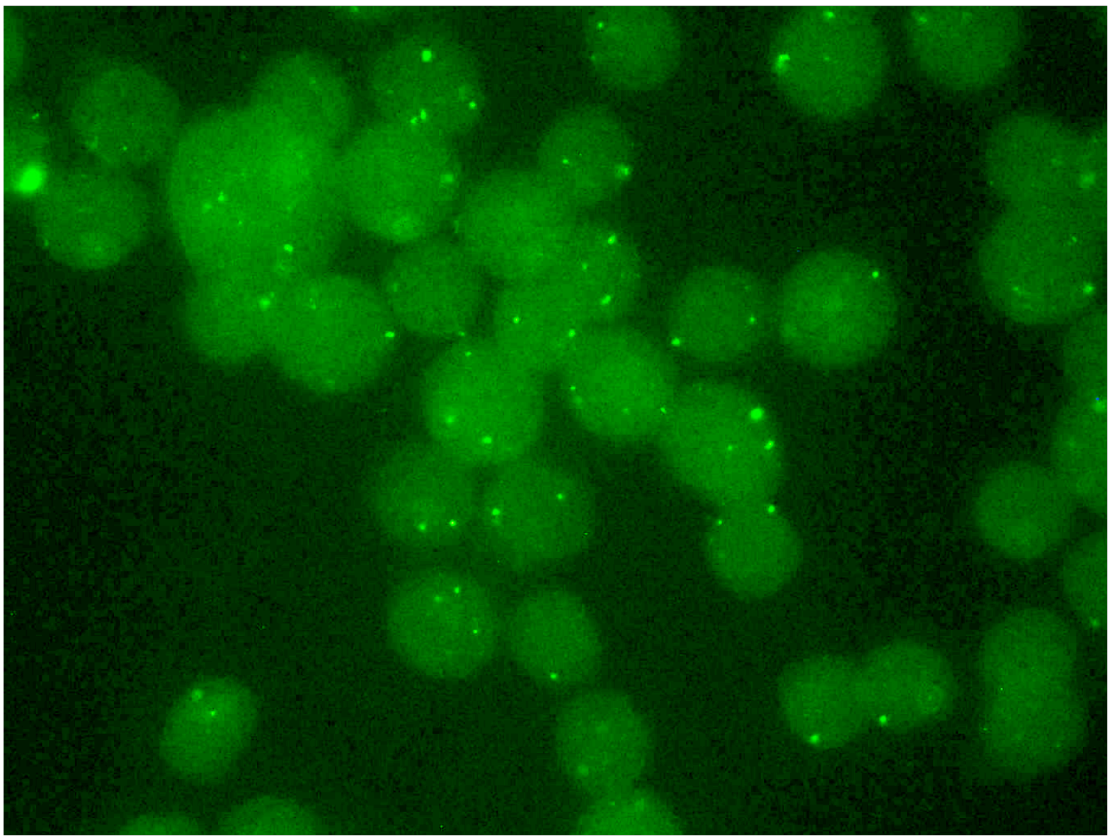
3CR DNA FISH Control: By comparison, green-filtered microscopy of benign thymus tissue hybridized to green-labeled 3CR probe in DNA FISH on the VP 2000 processor indicates smaller signals, often with two signals in a cell, more suggestive of true hybridization than the red signals in Figure 6. However, most cells contain no green signals or obviously nonspecific signal (large, bright spots).

Figure 8.



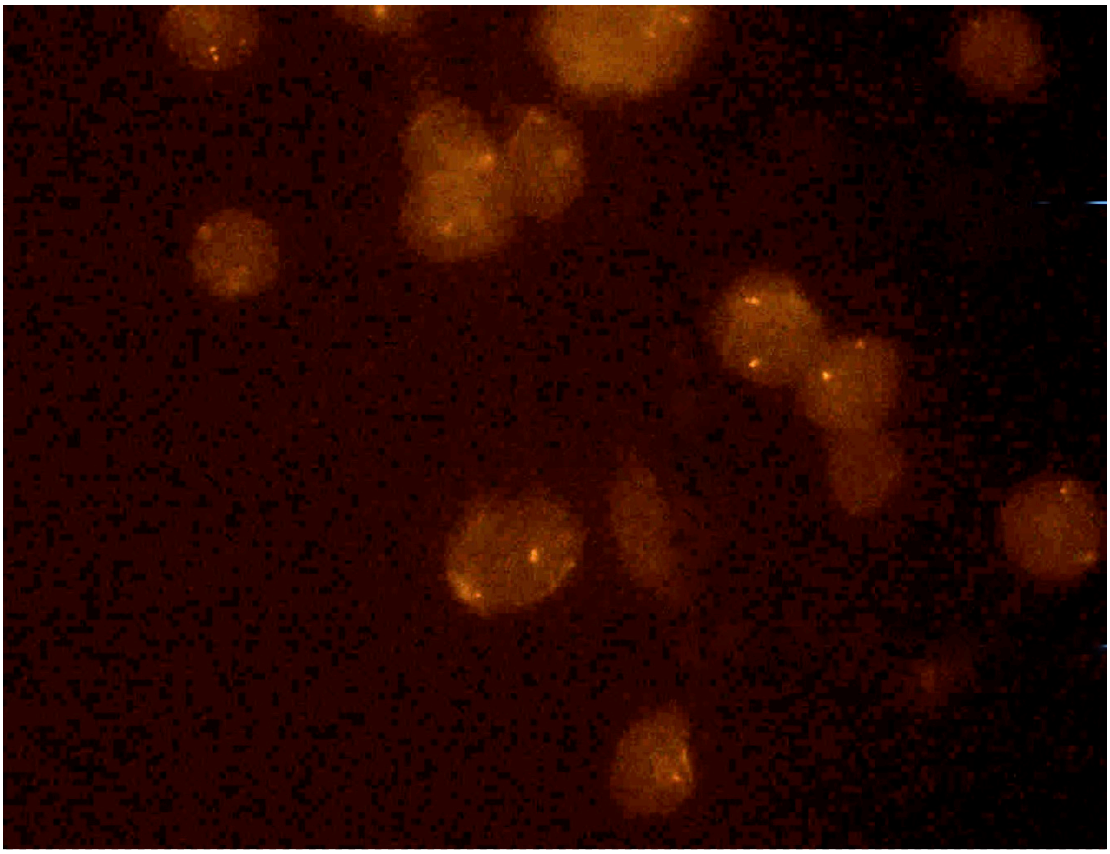
B Cell TCR beta DNA FISH Control: Red-filter microscopy of clonal Namalwa B cells hybridized to pooled PCR-generated probes targeting the 3CR region, in DNA FISH performed with the VP 2000 processor. Multiple cells display two red signals suggestive of *bona fide* hybridization. This experiment was conducted by Jonathan Murphy.

Figure 9.



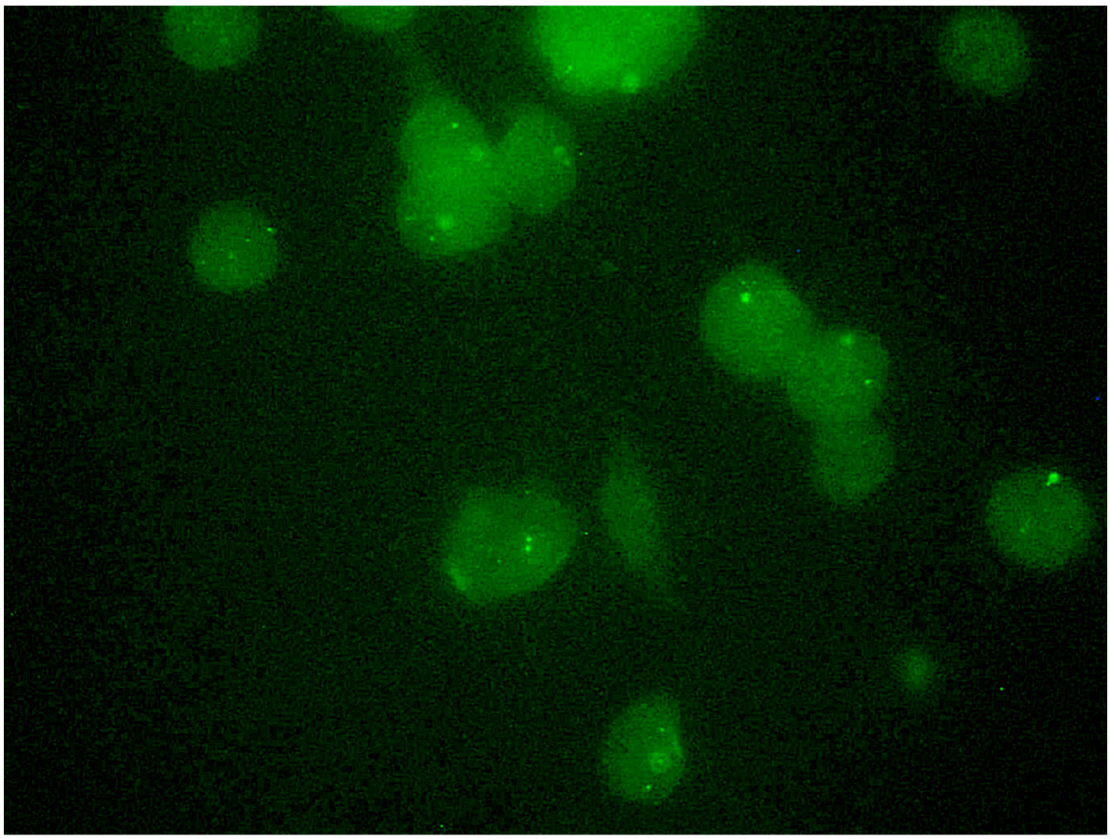
B Cell TCR beta DNA FISH: Green-filtered photograph of the same microscopic view as Figure 8. Green-labeled PCR-generated probes specific to the 64 kb VDR test region deleted in successful V-DJ recombination have been hybridized to clonal Namalwa B-cells in DNA FISH. Again, multiple cells display two green signals suggestive of *bona fide* hybridization to the *TCRB* test region, with occasional artifactual signals.

Figure 10.



T Cell TCR beta DNA FISH Control: Red-filter microscopy of clonal Jurkat T cells hybridized to pooled PCR-generated probes specific for the 3CR region, in DNA FISH performed with the VP 2000 processor. Multiple cells display two red signals likely indicating the control region adjacent to the *TCRB* locus. This experiment was conducted by Jonathan Murphy.

Figure 11.



T Cell TCR beta DNA FISH: Green-filtered photograph of same microscopic view as in Figure 10. Green-labeled pooled PCR probes targeting the test region between V29-1 and D1 have been hybridized to clonal Jurkat T-cells in DNA FISH performed with the VP2000 processor. Most cells have a single green signal overlapping one of the two red signals in Figure 10, suggestive of a single non-rearranged *TCRB* allele. The other green signals most likely represent artifact.