Putative markers for the detection of breast carcinoma cells in blood

Elfatih Eltahir

Thesis submitted for the degree of Master of Science (Med. Sci) to the University of Glasgow, being an account of research conducted at the Beatson Institute for Cancer Research and the University Department of Surgery, Glasgow.

The Beatson Institute for Cancer Research
CRC Beatson Laboratories
Bearsden
Glasgow

Department of Surgery
Western Infirmary
Glasgow

February 1999

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Putative markers for the detection of breast carcinoma cells in blood

CONTENTS

LIST OF ABBREVIATIONS

bp	base pair
BCPT	Breast Cancer Prevention Trial
C	Centigrade
CK 19	cytokeratin 19
cDNA	complementary DNA
DEPC	diethyl pyrocarbonate
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
EBCTCG	Early Breast Cancer Trialists' Collaborative Group
ER	Oestrogen receptor
IBIS	International Breast Cancer Interventional Study
HRT	hormone replacement therapy
mRNA	messenger RNA
NPI	Nottingham Prognostic Index
PBS	phosphate-buffered saline
PEM	polymorphic epithelial mucin
RNA	ribonucleic acid
rt -PCR	reverse transcriptase polymerase chain reaction
SDS	sodium dodecyl sulphate
SSC	sodium chloride and sodium citrate
TBE	tris-borate buffer
Tris	trihydroxymethylaminomethane
USB	United States Biochemicals
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
USB	United States Biochemicals

LIST OF FIGURES Page

ACKNOWLEDGMENTS

I wish to thank Prof. D. W. George for accepting me as a research fellow.

My greatest gratitude to Mr Arnie Purushotham for his advice and constant encouragement and supervision throughout the course of this study.

My thanks also go to Dr George Birnie for the useful discussion and technical advice and for all my colleagues at the Beatson for help and support.

I would also like to thank the Bon Secours hospital for funding.

DECLARATION

The work described in this thesis was performed personally, unless otherwise acknowledged.

PUBLICATIONS

Eltahir EM, Mallinson DS, Birnie GD, Hagan C, George WD, Purushotham AD. Putative markers for the detection of breast carcinoma cells in blood. *Br J Cancer* 1998 Apr; 77(8): 1203-7.

SUMMARY

By using the DF3 antigen (MUC1) as a molecular marker, previous workers reported the detection of tumour cell dissemination by reverse transcriptase Polymerase Chain Reaction (rt-PCR) in a pilot study of patients undergoing surgery for breast cancer (Brown et al., 1994). The aim of this study was to extend these findings and also examine the role of other epithelial markers for their suitability as molecular markers for detection of circulating breast carcinoma cells using rt-PCR. It was envisaged that this work could be extended to the examination of bone marrow and lymph nodes for occult metastases. RNA was prepared from MCF-7 breast carcinoma cells and peripheral blood leucocytes of healthy female volunteers. This RNA was screened for mRNA of MUC1, CK19 and CD44 (Exon 8-11) by rt-PCR and the results validated by Southern blots. Variable degrees of expression of MUC1 and CD44 (Exon 8-11) were detected in normal peripheral blood, rendering these genes non-specific for epithelial cells and therefore unsuitable for use as markers to detect circulating breast carcinoma cells. Although CK19 mRNA was apparently specific, it was deemed unsuitable for use as a marker of circulating breast cancer cells in light of its poor sensitivity. These findings prompted an alternative approach to this problem to continue the search for the ideal epithelial marker. Therefore, included in this work is RNA fingerprinting, a modified Differential Display by the PCR technique which is conceived to allow the identification, molecular cloning and sequencing of differentially expressed genes. This technique is devised to amplify mRNA and display their 3' termini on polyacrylamide gels. The aim of RNA fingerprinting is to identify and characterise the differentially expressed genes in breast cancer samples in contrast with normal blood. In this study a total of 18 differentially expressed genes (using limited primer combinations) were analysed and found to have variable degrees of background expression when used as potential markers. However, further work is required to characterise the differentially expressed genes using all other primer combinations. This study has cast major doubts on the validity of previously published studies in this area. Until a suitable epithelial marker for breast cancer is identified, further work to identify tumour cells in blood will not be feasible.

1. Introduction

1. INTRODUCTION

1.1 EPIDEMIOLOGY OF BREAST CANCER

1.1.1 INCIDENCE

1.1.2 AETIOLOGY AND RISK FACTORS

1.2 PROGNOSTIC FACTORS OF BREAST CANCER

1.3. MANAGEMENT OF BREAST CANCER

1.4 DETECTION OF CIRCULATING CANCER CELLS

1.5 METHODS OF DETECTION OF CANCER CELLS

1.5.1. CYTOLOGY

1.5.2. FLOW CYTOMETRY

1.5.3. IMMUNOCYTOCHEMISTRY

1.5.4 PCR

1.6 CANDIDATE GENES

1.6.1 PEM

1.6.2 CK19

1.6.3 CD44

1.7 RNA FINGERPRINTING

1.8 AIMS

1. INTRODUCTION

1.1 EPIDEMIOLOGY OF BREAST CANCER

1.1.1 INCIDENCE

Breast cancer is the most common form of cancer among women in North America, almost all of Europe, Latin America and Australia (Lynn et al., 1995; Miller et al., 1991; Boring et al., 1993). It is estimated that the disease will affect five million women world-wide in the next decade (Corry et al., 1994). The incidence of breast cancer is increasing on average by 1% per year in almost all populations, in both industrial and developing countries (Miller et al., 1991; Levi et al., 1993).

For the global population as a whole, the incidence and prevalence of breast cancer increase with increasing age (Glass et al., 1988; Harris et al., 1992). The incidence increases rapidly during the fourth decade of life and continues to increase with age after menopause, but less dramatically than before. In Japan and other Far Eastern countries, there is a similar pattern of increase in breast cancer incidence with increasing age but absolute rates are lower at each age (Harris et al., 1992). Despite the increase in incidence rates, the age-adjusted mortality rates from breast cancer have proved remarkably stable in the USA. A number of factors such as earlier detection and advances in treatment have been suggested to account for this trend (Harris et al., 1992).

In Scotland breast cancer is the most common cancer among women. In 1993, there were 3004 registrations, accounting for almost 25% of the female cancer burden excluding non-melanoma skin cancer (ISD, 1998). For any woman, the theoretical risk of developing breast cancer up to the age of 74 is approximately

8% (Harris et al., 1998). It is also the second cause of death from cancer in women, responsible for 1244 deaths in Scotland in 1996.

Whilst in common with all cancers, the prognosis of breast cancer depends upon its stage at diagnosis, it is known that approximately 50% of women are still alive 5 years after their breast cancer was first diagnosed (Black et al., 1993).

1.1.2 AETIOLOGY AND RISK FACTORS

The aetiology of the majority of breast cancer remains unexplained, and it is possible that many of the factors responsible are as yet unrecognised. It has been observed that large increases in the rates of breast cancer occur among populations migrating from Japan to the USA, which indicates that breast cancer may be determined by factors other than genetic ones which suggests that the potential for prevention exists (Buell et al., 1973). The existence of non-genetic determinants of breast cancer is also indicated by the large variations of breast cancer among different countries and over time within countries (Armstrong et al., 1975; Prentice et al., 1990).

The known hormonal risk factors for breast cancer are related to the cumulative effect of oestrogen and possibly progesterone (Henderson, 1991), but controversy remains over the contribution of exogenous hormones to the development of breast cancer (Helzlsouer et al., 1995). The benefit of hormone replacement therapy (HRT) in reducing ischaemic heart disease and osteoporosis and improving menopausal symptoms suggest that for the first 10 years of use the overall effects of therapy on morbidity and mortality are likely to be beneficial (Helzlsouer et al., 1995), but concerns about HRT centre on the risk of

endometrial cancer and breast cancer, especially after long-term use (more than 10 years). There is also considerable uncertainty about the relationship between a woman's risk of developing breast cancer and the length of time she receives HRT (Grodstein et al., 1997, Brinton et al., 1997).

A family history of breast cancer is known to be an important risk for the disease (Claus et al., 1990) and although most of breast cancers are sporadic, approximately 5% are familial (Xu et al., 1996). The existence of a breast cancer susceptibility gene was demonstrated by genetic linkage analysis of breast cancer families (Hall et al., 1990). This gene was later named BRCA1 (Solomon et al., 1992) and preliminary reports suggest that it is a large gene and the mutations are not confined to particular regions. The susceptibility to breast cancer is generally inherited as an autosomal dominant with limited penetrance (Skolnick et al., 1992). About a third of familial cases are thought to be due to mutation in the BRCA1 gene on the long arm of chromosome 17. Some inherited BRCA1 mutations seem to be associated with only a modestly increased risk of breast cancer. A second hereditary breast cancer locus, BRCA2, has been identified in the long arm of chromosome 13 (Xu et al., 1996). In addition, a few cases arise from mutations in the p53 gene on the short arm of chromosome 17 (Hoskins et al., 1994).

Many families affected by breast cancer show an excess of ovarian, colon, prostatic and other cancers attributable to the same inherited mutation. Patients with bilateral breast cancer, those who develop a combination of breast cancer and another epithelial cancer, and women who get the disease at an early age are most likely to be carrying a genetic mutation that has predisposed them to developing breast cancer. Most breast cancers that are due to a genetic mutation

occur before the age of 65, and a woman with a strong family history of breast cancer who is still unaffected by the age of 65 has probably not inherited the genetic mutation.

A woman's risk of breast cancer is doubled if she has a first degree relative (mother, sister, or daughter) who developed the disease before the age of 50, and the younger the relative when she developed breast cancer the greater the risk (10% cumulative risk), this risk approaches 5% (close to the population risk) if the first-degree relative was older than 50 years at diagnosis. The risk increases by four to six times if two first-degree relatives develop the disease (McPherson et al., 1994).

Many other risk factors for breast cancer have also been identified, such as early menarche (Kampert et al., 1988), late menopause (Trichopoulos et al., 1972), nulliparity or late age at birth of first child (White, 1987; MacMahon et al., 1970) and a history of benign breast disease (Dupont et al., 1985). Women with severe atypical hyperplasia have a four to five times higher risk of developing breast cancer than women who do not have any proliferative changes in their breasts (Dupont et al., 1985). Women with this change and a family history of breast cancer have a nine fold increase in risk. Women with palpable cysts, duct papillomas, sclerosing adenosis, and moderate or florid epithelial hyperplasia have slightly higher risk of breast cancer (1.5-2 times) than women without these changes, but this increase is not thought to be clinically significant (Dupont et al., 1985).

Unfortunately it has not been possible, to date, to utilise the known risk factors to allow the identification of the small, high-risk group that accounts for a large proportion of women with the disease (Harris et al., 1992). Progress towards the

identification and subsequent modification of risk factors for the prevention of breast cancer has so far been disappointing, and prevention currently depends upon the artificial manipulation of hormones and growth regulators that underly the known risk factors. Although the preliminary results of the Breast Cancer Prevention Trial (BCPT) indicate a 45% reduction in breast cancer incidence with prophylactic use of tamoxifen (Pritchard, 1998), other similar chemoprevention trials (Veronesi et al., 1998; Powles et al., 1998) have failed to confirm BCPT results which are nevertheless consistent with 39% preventive effect of 5 years of tamoxifen reported by the Early Breast Cancer Trialists Collaborative Group (EBCTCG). The findings of The International Breast Cancer Intervention Study (IBIS) are still awaited to clarify the preventive benefits of tamoxifen.

1.2 PROGNOSTIC FACTORS OF BREAST CANCER

Prognostic factors are clinical and pathological features that give information in estimating the likely clinical outcome of patients suffering from cancer. Several histopathological, biochemical and immunological variables have been reported to be useful prognostic factors. Since breast cancer is a heterogenous disease and metastasis a complicated process, no single prognostic factor predicts whether a patient will develop recurrence. Most studies on prognostic factors are retrospective, with small numbers of patients and often insufficient follow-up to allow true determination of the effect of a prognostic factor on relapse or survival. A host of prognostic factors are being evaluated to determine their ability to predict recurrence in node-negative breast cancer . Tumour size, axillary lymph node status, histologic grade and the hormonal receptor status are established

prognostic factors for breast carcinoma (Schnurch, 1998).

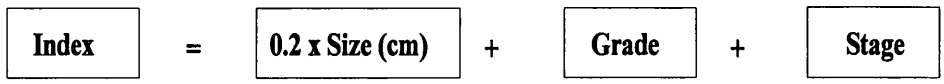
Clinical variants also have a bearing on prognosis and it is suggested that patients with small node-negative tumours have an excellent prognosis (Rosen et al., 1993) and the larger the tumour, the worse the prognosis (Fisher et al., 1969). Breast carcinoma occurring during pregnancy or lactation is generally considered to have a poor prognosis (Wile et al., 1989).

Sinus histocytosis or reactive hyperplasia of lymph nodes has been shown in some series to be associated with good outcome. The 5 year survival rate is likely to be 80% if three or less lymph nodes are involved, This figure drops to 30% when four or more lymph nodes are involved.

Nodal status remains the most powerful determinant of survival in breast cancer patients and a quantitative guide to prognosis from histopathology is best obtained by grading. Some tumour types have an especially good prognosis, particularly "medullary" carcinoma which frequently shows heavy infiltration of the tumour with lymphocytes.

The most widely used histological grading system for breast cancer in the UK is based on that of Bloom and Richardson (1957) and modified by Elston (1987). In this system a separate score of 1-3 is given for each of three features - tubular formation, nuclear pleomorphism and the number of mitoses per 10 high-power microscopic fields. A total score of 3-5 indicates grade I (well-differentiated carcinoma), 6-7 indicates grade II and 8-9 grade III (poorly-differentiated carcinoma). There is good correlation with prognosis: Grade I tumours are associated with significantly better survival than grade II (Elston et al., 1982). Figure 1 shows these prognostic factors have been combined in the Nottingham prognostic index (Haybittle et al., 1982), which, in addition to providing useful

Nottingham Prognostic Index (NPI)



Grade:
1 = tubular, cribriform, mucoid
2 = medullary, lobular

Stage: (refers to lymph node status)
1 = negative
2 = <4
3 = >4

Index:
<3.4 = low
3.41 - 5.4 = medium
>5.4 = high

Score of 3 or less is associated with excellent prognosis while a score of more than 6 is associated with less than 20% 5-year survival (Todd et al; 1986).

Figure I

prognostic information, also selects out high risk patients for adjuvant systemic treatment. Another prognostic index based on tumour size, lymph node status and mitotic activity index is advocated by Baak and colleagues (van der Linden et al., 1987).

The hormonal receptor status is the best known predictive factor with regard to response to hormonal treatment. It correlates with both longer disease-free interval and prolonged survival time (Preece et al., 1982; Rayter, 1991).

There are many biological factors such as kinetic indices, over-expressed growth factor receptors, mutated tumour-suppressor genes and enzymes that may play a role in the metastatic process (Silvestrini et al., 1996; Gilchrist et al., 1993; Seshadri et al., 1993; Allred et al., 1993; Friedrichs et al., 1993; Weidner et al., 1992; Maggiora et al., 1998).

S- phase fraction of the tumour appears to correlate well with the risk of relapse, but is subject to variation between laboratories and may not be independent of tumour grading by the pathologist (Hedley et al., 1987).

The lysosomal acidic protease cathepsin D is another potential prognostic indicator (Winstanley et al., 1993). This enzyme is present in high levels in cancer cells and its secretion from breast cancer cells is stimulated by oestrogen (Rochefort, 1990).

The prognostic value of over-expression of the HER2/neu gene product in both node-positive and node-negative breast cancers has been conflicting (Seshadri et al., 1993). In a study of 254 node-negative tumours, Gasparini (1994) demonstrated that microvessel density was a stronger predictor of relapse-free survival than tumour lymphatic invasion, p53 expression, or tumour size. Acenero (1998) established vascular density to be a significant prognostic indicator in both

node-negative and node-positive patients. However, p53 protein over expression was reported to be an independent predictor of reduced disease-free survival in node-positive and ER-positive patients but not in node-negative or ER-negative individuals (Levesque et al., 1998).

Determination of c-erbB-2 levels was reported to be an important biomarker to assess the extent of disease spread in the lymph nodes as well as a prognostic indicator for disease-free survival in patients who receive chemotherapy (Mehta et al., 1998).

DNA-RNA contents in breast cancer were found to be independent prognostic factors. Hara et al. (1998) found that high RNA index of more than 4.0 correlates significantly with poorer prognosis.

1.3. MANAGEMENT OF BREAST CANCER

The diagnosis of breast cancer is made by clinical examination, mammography and/or ultrasound, fine needle aspiration and core biopsy.

Local treatment to the breast

Breast-conserving surgery, combined with radiotherapy and systemic therapy, is now considered to be more appropriate than radical mastectomy for the majority of women with early stages of breast cancer, and is associated with survival rates comparable to those following mastectomy (Veronesi et al., 1990; Harris et al., 1992; Fisher et al., 1995).

Local treatment to the axilla

The axilla may either be surgically cleared of all nodes or alternatively node sampling may be performed. A minimum of four nodes sampled is required to adequately determine the pathological nodal status (Forrest et al., 1995). There is no need for postoperative radiotherapy if the axilla has been cleared. However, if sampled nodes are involved by tumour, radiotherapy may be indicated to control presumed residual disease in the axilla.

Adjuvant radiotherapy

Radiotherapy to the breast after conserving surgery significantly reduces the risk of local recurrence (Veronesi et al., 1993; Liljegren et al., 1994; Fisher et al., 1995). Although radiotherapy after mastectomy reduces the risk of local recurrence, it has very little effect on overall survival due to increased cardiac morbidity (Cuzick et al., 1994). However, in premenopausal women receiving adjuvant chemotherapy, postmastectomy radiotherapy appears to significantly improve survival (Overgaard et al., 1997; Ragaz et al., 1997).

After axillary sampling, the axilla should only be irradiated if node positive or inadequately sampled. If there is no axillary node involvement by tumour, radiotherapy confers no additional benefit in term of local recurrence and significantly increases the risk of morbidity (Dewar et al., 1987; Forrest et al., 1995).

Adjuvant systemic therapy

The objective of systemic therapy is to reduce the growth potential of micrometastases. The systemic treatment comprises cytotoxic and endocrine

therapies. Adjuvant systemic therapy improves relapse free survival and overall survival in all age groups (EBCTCG, 1988; 1996).

Cytotoxic chemotherapy is used as adjuvant treatment following primary surgery on the primary tumour and treatment of advanced or metastatic breast cancer.

Ovarian ablation improves long-term survival in women aged under 50 with early breast cancer (EBCTCG, 1996). Ovarian suppression can be carried out medically using LHRH-A or by radiotherapy or surgery. Oophorectomy was the first form of endocrine therapy to be used for breast cancer treatment, a century ago (Beatson et al., 1896), and adrenalectomy and hypophysectomy were introduced in the early 1950s (Huggins et al., 1952; Luft et al., 1953). Since then, systemic endocrine drugs, such as tamoxifen, have been successfully used in the treatment of advanced breast cancer (Boring et al., 1993).

Combination chemotherapy and tamoxifen are standard adjuvant therapies for most women with node-positive breast cancer and some women with node-negative breast cancer. However, prognostic factors that identify women with node-negative disease at high risk of recurrence are not yet available for the majority of patients (Gasparini et al., 1993).

Cytotoxic chemotherapy is associated with severe side-effects such as alopecia, weight gain, fatigue and can compromise the immune system (Hafstrom et al., 1990). However, the short-term use of chemotherapy in the adjuvant setting is generally well tolerated. Although some improvement in tolerability of chemotherapy has occurred since the introduction of intermittent, pulsed therapy and autologous bone marrow support techniques, the overall toxicity of chemotherapy is still considerable, particularly in postmenopausal women (Harris et al., 1992).

Despite major advances in adjuvant therapy, improvement in survival has been disappointingly small. Fisher (1997) reported a significantly better disease-free survival in patients with node-negative, ER-positive tumours who received chemotherapy plus tamoxifen than those who received tamoxifen only. Targeting patients appropriately for adjuvant therapy is currently based on clinico-pathological prognostic factors. Nodal status has been established as the most important prognostic indicator and it is suggested that a level I and level II axillary lymph node dissection be performed for all early invasive breast cancer (Coburn et al., 1995). The use of a prognostic score was recently proposed to replace routine axillary lymph node dissection as a means of identifying patients who benefit from adjuvant therapy (Menard et al., 1994). Sentinel node biopsy has been suggested as an alternative to axillary node dissection for clinical stage I breast cancer (Giuliano et al., 1994). This seems to be an attractive alternative, but its validity must be prospectively verified before it is universally adopted.

The use of systemic therapy early in the course of malignant disease has been shown to influence survival in some common cancers, such as those of breast (EBCTCG, 1994) and colon (Moertel et al., 1990; Riethmuller et al., 1994). The improvements are, however of limited scale. This results in exposure of some patients to unnecessary toxicity as well as making evaluation of the treatment difficult.

Recent interest has focused on developing laboratory methods to identify disseminated tumour cells in the circulation. This may identify a subgroup of patients suitable for adjuvant systemic chemotherapy. The identification of specific molecular markers may allow a more rational allocation of such adjuvant therapy. Apart from the prescription of therapy to a poor prognosis sub-group, the

molecular identification of residual disease may also prove useful in determining alternative end points. The practical problem of continuing with ineffective therapy over long periods may be more readily identified if reliable markers of tumour persistence can be identified.

1.4 DETECTION OF CIRCULATING CANCER CELLS

There is increasing emphasis on the detection of circulating cancer cells in the peripheral blood. This has potentially significant therapeutic and prognostic implications. It may be feasible to select suitable patients to assess their response to systemic therapy, and also to identify residual or recurrent disease after potentially curative treatment.

Animal studies have demonstrated that manipulation of the primary tumour facilitates tumour cell spread via the blood stream and may promote metastases. The "no touch" isolation technique reported by Turnbull (1967) in patients with colon cancer has provided the only clinical evidence that surgical manipulation of primary tumours may precipitate tumour cell shedding. Tyzzer et al. (1913) demonstrated an increase in the number of pulmonary metastases by manipulating transplanted breast tumours in mice. The hypothesis that operative handling of malignant neoplasms facilitates tumour cell dissemination into the blood stream has been investigated by many workers using different methodologies.

1.5 METHODS OF DETECTION OF CANCER CELLS

Several workers have attempted to identify circulating tumour cells in various malignancies (Choy et al., 1996; McCulloch et al., 1995; Brown et al., 1995; Smith et al., 1991; Burchill et al., 1994; 1995). Various techniques including morphology, flow cytometry and cytogenetics have been used to detect disseminated tumour cells at a level of 1 in 100, but these have limited sensitivity and specificity and are time-consuming (Diel et al., 1992; Frank et al., 1990; Johnson et al., 1995).

1.5.1. CYTOLOGY

Historically, conventional cytology has been the method used to detect circulating tumour cells. Difficulties in distinguishing tumour cells from plasma cells and atypical mononuclear cells, occasionally found in normal blood, has lead to a variation in the interpretation of the morphological features of malignant cells.

1.5.2. FLOW CYTOMETRY

Flow cytometry is a method of quantitating components or structural features of cells primarily by optical means. Although it makes measurement on one cell at a time, it can process thousands of cells in a few seconds. Since different cells can be distinguished by quantitating structural features, flow cytometry can be used to count cells of different types in a mixture.

Flow cytometry provides valuable information regarding cellular DNA content and nuclei (ploidy), which are isolated from single cell suspensions and stained with fluorescent dye. The sample is run through a cytometer, which uses a laser or mercury light source to excite the fluorescently labelled nuclei. The emission is

recorded on a DNA histogram, which recognises aneuploid and diploid cells as separate peaks. Ploidy is quantified as the DNA index, a measure of the ratio of the mean DNA content of the test sample cell to that of the diploid controls. DNA index of greater than 1 is considered aneuploidy. Determination of ploidy status is useful prognostically in a variety of solid tumours such as ovarian cancer where ploidy status was found to be a significant independent prognostic factor. Aneuploidy also correlates with a worse prognosis in many solid tumours including melanoma and breast. Suzuki et al. (1997) suggested strict follow-up of aneuploid parathyroid adenomas, as there is increase in premalignant potential as determined by DNA flow cytometry.

Flow cytometry can also be used to estimate the tumour proliferative activity, which is a measure of the percentage of cells actively proliferating. The histogram is used to calculate the cell cycle distribution and the percentage of cells in the S (synthesis) phase.

Coustan-Smith et al. (1998) investigated minimal residual disease in bone-marrow aspirates from children with treated acute lymphoblastic leukaemia (ALL) and claimed a powerful prediction of relapse in children with ALL using multiparameter flow cytometry.

The value of flow cytometry in prognostic determination of breast cancer is unclear. Several studies have demonstrated a poorer prognosis for patients with node-negative aneuploid tumours as compared with diploid tumours, whereas others have been unable to assign ploidy a role as an independent prognostic factor. However, in patients with stage II and III disease, aneuploidy and a high S-phase fraction have strongly correlated with diminished disease-free survival (Hurwitz et al., 1992).

1.5.3. IMMUNOCYTOCHEMISTRY

This technique allows cell-specific (as opposed to Western blot) localisation of protein in thin tissue sections, using antibodies directed against specific peptide sequence. The technique can be used alone, but is a much more powerful tool when employed in conjunction with other techniques such as in situ hybridisation. Historically, immunocytochemistry started in 1941 when Coon identified pneumococci using a direct fluorescent method. This was followed by the indirect method, the addition of 'horsedish' peroxidase anti-peroxidase technique of 1979 and the use of Avidin and Biotin complex in the early 1980s (Burnet et al., 1997). Initial applications were directed primarily to the identification of cell and sub-type (lineage related markers). Most recently immunohistochemical markers have been described that show great promise in determining tumour prognosis at a very early stage of tumour development, independent of stage and grade (Taylor et al., 1997).

Immunocytochemistry which may detect one tumour cell in 10^5 normal cells (Molino et al., 1991; Osborne et al., 1991) has also been used in an attempt to detect circulating breast carcinoma cells peri-operatively (Glaves et al., 1988; Choy et al., 1996; McCulloch et al., 1995). This technique is, however dependent upon the availability of antibodies to tumour-associated cell-surface antigens and may be subject to false positives when antibodies cross-react or tumour antigens are presented on host immune cells (Hydermann and McCartney, 1985; Leather et al., 1993) combined immunocytochemistry and conventional cytology to detect circulating malignant epithelial cells in patients with colorectal cancer in the peri-operative period.

The K-ras gene mutation has been used to detect circulating colorectal cancer

cells by an immunobead-PCR assay, but this approach is limited because of the low frequency of the gene mutations (30%) in these tumours (Hardingham et al., 1995). No similar specific mutations are known for breast cancer.

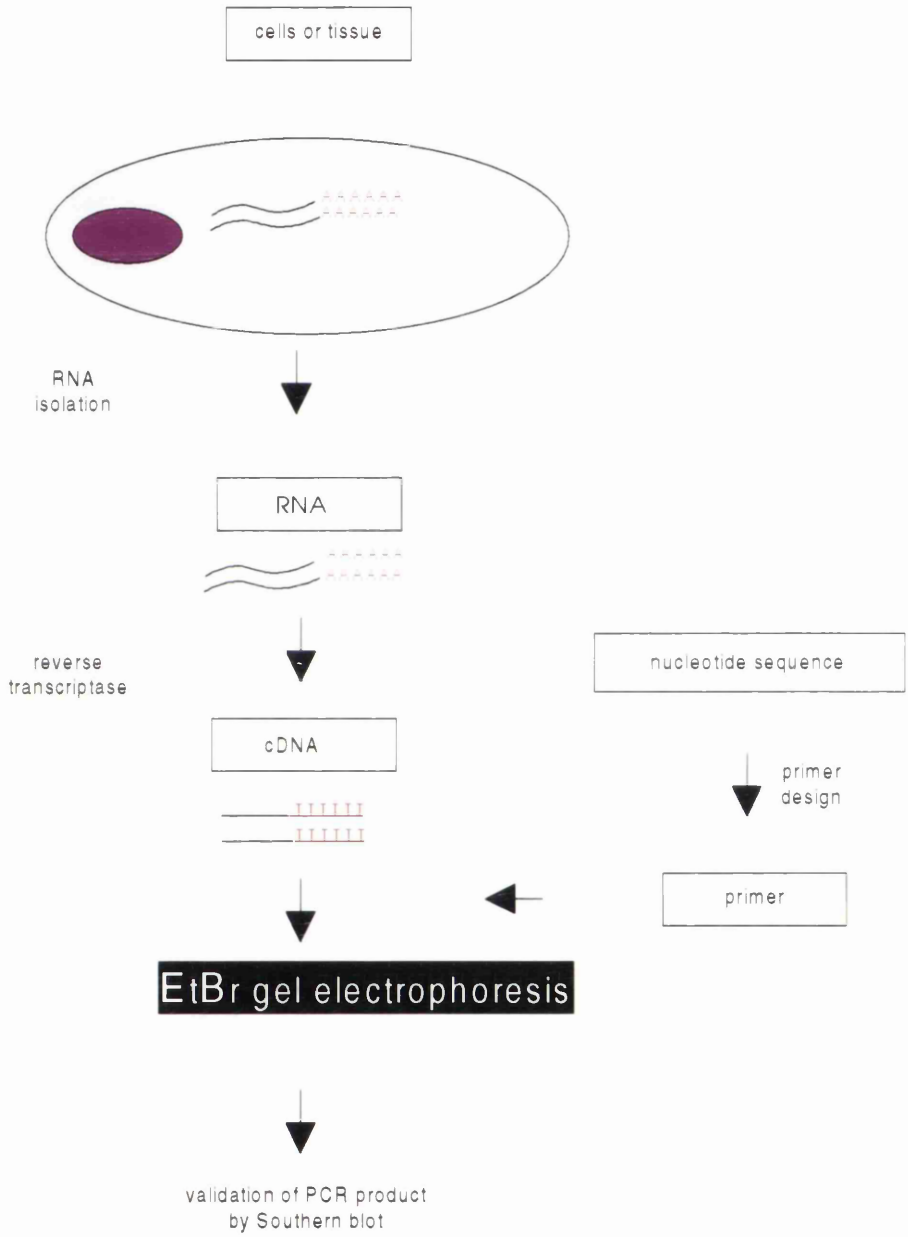


Figure 2. Schematic diagram of the rt-PCR method

1.5.4 PCR

The polymerase chain reaction (PCR) is a technique for *in vitro* amplification of specific sequence of either DNA or RNA using two oligonucleotide primers that hybridise to opposite strands and flank the region of interest in the target DNA. A repetitive series of cycles involving template denaturation, primer annealing, and the extension of the annealed primers by DNA polymerase results in the exponential accumulation of a specific fragment whose termini are defined by the 5' ends of the primers. Because the primer extension product synthesised in one cycle can serve as a template in the next, the number of target DNA copies approximately doubles at every cycle. Thus, 20 cycles of PCR yields about a million-fold (2^{20}) amplification. The PCR method was described by Kary Mullis and colleagues at the Cetus Corporation (Mullis et al., 1986; 1987) and originally applied by a group in the Human Genetics Department at Cetus to amplify human-globulin DNA and in the prenatal diagnosis of sickle-cell anaemia (Saiki et al., 1985; 1986; Embury et al., 1987). A schematic diagram of the rt-PCR method is shown in Figure 2.

PCR can potentially detect one cell in 10^6 normal cells (Mullis et al., 1986; Saiki et al., 1985). It therefore has the necessary sensitivity to detect a small number of tumour cells in peripheral blood by amplification of specific DNA sequences.

The first tissue-specific enzyme used as a molecular marker was tyrosinase, expressed by melanocytes as part of the melanin synthesis pathway. Since normal melanocytes are not normally found in peripheral blood, detection of tyrosinase transcription in peripheral blood by RT-PCR should demonstrate the presence of circulating melanoma cells (Smith et al., 1991). Using this technique it was possible detect one melanoma cell in at least 10^5 normal cells.

Tumour-specific DNA sequence abnormalities have been identified mainly in haematological malignancies such as acute lymphoblastic leukaemia (Yamada et al., 1990) where RT-PCR has been used to aid in the diagnosis or follow-up of patients with the disease. When there is a tumour specific sequence, PCR has been shown to detect one malignant cell in up to 10^7 normal cells (Mattano et al., 1992; Fabrega et al., 1993; Alkan et al., 1993; Cross al., 1993; Datta et al., 1994; Gerhard et al., 1994; Negrin and Peasando, 1994; Luppi et al., 1996; Mori et al., 1996). This increases sensitivity of detection by an order of magnitude when compared with immunocytochemistry.

For solid tumours however, tumour-specific DNA sequence abnormalities are uncommon and tissue specific RNA has been used as a molecular marker for rt-PCR (Smith et al., 1991; Burchill et al., 1994) since epithelial tissues are not normally present in peripheral circulation. The sensitivity of tumour cell detection by rt-PCR was greater than that of immunocytochemistry for CEA or cytokeratins (Gerhard et al., 1994).

1.6 CANDIDATE GENES

1.6.1 PEM

The human polymorphic epithelial mucin (PEM) is an extensively glycosylated high molecular weight mucin glycoprotein (250 - 500 kDa). Purified mucins from different normal and malignant tissues demonstrate heterogeneity in carbohydrate composition and structure (Shimizu and Yamauchi, 1982; Gendler et al., 1991; Seregni et al., 1997). Monoclonal antibodies have been used to identify many of the clinically useful epitopes and to develop a serum assay that detects and quantifies mucin levels in blood and other body fluids and have been proven to be of clinical diagnostic value. Some of these antibodies recognise epitopes exposed on tumour mucin but not on corresponding normal cell populations (SM-3 and B72.3), whereas others recognise epitopes exposed on mucin produced by normal and/or malignant cells of several organ sites (DnPan 2 and CA 19-19).

The human DF3 breast carcinoma - associated antigen (MUC1) gene is a core protein of PEM and is uniformly and highly expressed apically by glandular epithelium of normal and malignant human mammary epithelium (Lancaster et al., 1990; Ho et al., 1993) and rarely expressed in non-epithelial tissues (Zotter et al., 1988). It is also found in the cytosol and on the cell membrane of adenocarcinomas of ovary and lung and mesothelial cells (Ho et al., 1993; Hilkens, 1988).

The splicing pattern and nucleotide sequence of MUC1 mRNA expressed in ovarian cell lines (CaOV-3, OVCAR-3 and SKOV) is the same as that of breast and pancreatic cell lines with a single nucleotide exception (Stern et al., 1992).

Members of the MUC1 family of mucins isolated from different tissue types have been shown to differ in biochemical properties and immunological reactivity.

However, the molecular basis of the alterations that occur in mucins during the pathogenesis of different diseases is poorly understood.

Studies that demonstrate expression of more than one mucin core protein by different adenocarcinomas and their normal tissue counterparts suggest that expression of distinct arrays of core proteins may be associated with tumours from distinct organ sites (Yonezawa et al., 1991; Devine et al., 1991; Devine and Mckenzie, 1992; Ogata et al., 1992; Seregni et al., 1996; Kim, 1998). Hollingsworth (1994) reported that most pancreatic tumour cell lines express MUCI and/or MUC4, but extremely low levels of MUC2 and almost undetectable levels of MUC3 mRNA whereas colon tumour cell lines generally express high levels of MUC2 and/or MUC3, but rarely express significant levels of MUC4.

These studies suggest that tumour-derived mucin detected in serum may show a core protein composition distinct from serum mucin derived from normal tissues.

MUCI is expressed in the tissue of 90% of cases of breast cancer (Papadimitriou et al., 1993). It has been detected in patients' serum and has been shown to be increased in metastatic disease (Abe et al., 1993).

1.6.2 CYTOKERATINS

Intermediate filaments (diameter, about 10 nm) are primary components of mammalian cell cytoskeleton and constitute a multigene family of related proteins distinguished by cell type-specific expression (Nagel et al., 1988).

The cytokeratins (CKs) comprise some 20 different isotypes and are predominantly expressed in epithelial cells where they show strict lineage- and differentiation- associated patterns of expression (Moll et al., 1982; Sun et al., 1984). They have also been detected in many normal and neoplastic non-epithelial cell types using immunohistochemical techniques (Traweek et al., 1993).

Malignant cells generally retain the intermediate filaments of their progenitor cell type and consequently CKs have been used to characterise neoplastic cells of epithelial origin (Osborn and Weber, 1983; Cooper et al., 1985; Lane et al., 1990). Carcinomas are characterised by cytokeratins, sarcomas of muscle cells by desmin, non muscle sarcomas by vimentin and gliomas by glial fibrillary acidic protein (Osborn and Weber, 1983).

Cytokeratin 19 (CK19) is synthesised mainly in embryonic and adult simple epithelia, but has also been found in stratified epithelia as well. CK19 is the smallest known keratin and is remarkable in that, unlike all other keratins, it does not have a designated partner for the formation of filaments, implying that regulation of its expression is different from other keratin-encoding genes (Lussier et al., 1990).

The cytokeratins are primarily expressed in epithelial cells (Moll et al., 1982). As the specificity of expression of intermediate filaments proteins is retained in malignant tumours, they are suitable as tumour markers (Moll, 1993). CK19 has

been reported as a specific and sensitive marker for detection of breast carcinoma cells in peripheral blood and the bone marrow of patients with breast cancer by rt-PCR (Datta et al., 1994). CK19 has been shown to be expressed in the tissue of 90% of invasive breast cancer (Papadimitriou et al., 1993).

1.6.3 CD44

CD44 is a widely expressed cell surface glycoprotein that serves as an adhesion molecule in cell-substrate and cell-cell interactions, including lymphocyte homing, cell migration, and metastasis (Haynes et al., 1991; Matsumura and Tarin, 1992; MacKay et al., 1994). It is thought to be associated with aggressive histological features in breast cancer. It has also been implicated in conferring metastatic potential to rat carcinoma cells (de la Torre et al., 1995).

CD44 is a transmembrane glycoprotein occurring in several isoforms with different extracellular regions. The various transcripts are encoded by one gene locus containing 20 exons, of which at least 10 can be alternatively spliced in nascent RNA (Fox et al., 1994). Isoforms encoded by the variant exons (termed CD44v) are highly restricted in their distribution in non-malignant tissue as opposed to the standard form of CD44 (CD44s) abundant in many tissues.

MacKay et al. (1994) used antibodies specific for variants CD44v4, CD44v6 and CD44v9. They found that most epithelia expressed CD44v9, whereas expression of CD44v4 and CD44v6 was more tightly restricted. These variant isoforms were most abundantly expressed in the generative cells of epithelia such as the myoepithelial cells of the mammary gland.

Specific variant isoforms containing exon 6v have been shown to render non-

metastatic rat tumour cells to have metastatic potential (Hofmann et al., 1991). Based on the prominent role in rat metastasis formation, CD44v isoforms were suggested to be involved in human tumour progression. Correlations between prognosis and expression of CD44 have been reported for gastric and colon carcinoma, non-Hodgkin's lymphoma, and more recently for breast carcinoma (Kaufmann et al., 1995).

CD44 (Exon 8-11) was used as a potential marker since a CD44 variant encoding these sequences has been shown to be expressed in human colorectal and breast cancers and has been implicated in the induction of the metastastatic phenotype in rat pancreatic tumour cells (Gunthert et al., 1991; Hofmann et al., 1991; Delatorre et al., 1995). CD44 splice variants specific to breast malignant tumour cells have also been described by Matsumura (1992) who reported that peripheral blood leucocytes contained only the standard form of CD44 mRNA, although others have reported traces of these splice variants in blood (Fox et al., 1994).

1.7 RNA FINGERPRINTING

This is a modified differential display technique, which has been developed as a tool to detect and characterise altered gene expression in eukaryotic cells. The basic principle is to systematically amplify messenger RNA and then distribute their 3' termini on a denaturing polyacrylamide gel. The technique was applied to compare mRNA from normal blood and tumour-derived human mammary epithelial cells. Complementary DNA fragments corresponding to apparently differentially expressed mRNA are recovered and sequenced.

1.8 AIMS

The aim of the work presented here was to examine the DF3 antigen, CK19 and CD44 (E8-11) as suitable molecular markers for the detection of circulating breast carcinoma cells.

This work has also included the RNA fingerprinting technique which is conceived to allow the identification and molecular cloning and sequencing of differentially expressed genes. The aim of the latter part of this work was to isolate sequence tags which are expressed in all or a majority of malignant breast tumours but not in any normal peripheral blood samples.

2. Methods

2. METHODS

2.1 Tissue culture

2.1.1 Cell and maintenance

2.1.2 Mycoplasma testing

2.2 DNA methodoloy

2.2.1 Agarose gel electrophoresis

2.2.2 Southern blot analysis of DNA

2.2.3 Preparation and quantitation of oligonucleotides

2.3 RNA methodology

2.3.1 Collection of samples

2.3.2 RNA extraction, purification and quantification

2.3.3 Reverse Transcriptase Polymerase Chain Reaction (rt-PCR)

2.3.4 Primer sequences for rt-PCR

2.3.5 Seeding experiments

2.4 RNA Fingerprinting

2.5 Molecular cloning

2 METHODS

2.1 Tissue culture

Aseptic manipulation was observed using sterilised glassware in a class II microbiological safety cabinet with vertical airflow.

2.1.1 Cell lines and maintenance

MDA and MCF7 human breast carcinoma cells express CK19 (Moll et al., 1982) Both lines were maintained in RPMI-1640 (GIBCO BRL) supplemented with 10% foetal calf serum, 0.2% sodium bicarbonate, 2 mM sodium pyruvate and 4 mM glutamine. Cells were grown as monolayers and were passaged every 4 days at 5 x 10^5 cells/ml in Nunclon tissue culture grade flasks and kept at 37 ^{o}C in 5% CO_2 , 95% O_2. Cells were stored by freezing in RPMI with 10% di-methyl sulphoxide (DMSO) in 1ml cryotubes at -70^{o} C and were then maintained in liquid nitrogen.

2.1.2 Mycoplasma testing

Cells were tested periodically for the presence of mycoplasma. The cells to be tested were pelleted by centrifugation at 1500 rev/min for 5 min in an MSE bench centrifuge. The cell-free supernatant was then placed in a sterile tissue culture grade petri dish (Nunclon), 2 ml of fresh medium was added, and 2 x 10^5 NRK49 fibroblast were seeded into the dish. These cells are known to be mycoplasma-free. The dish was incubated at 37 ^{o}C in a humidified incubator at 5% CO_2, 95% O_2 for 3 days. At the end of the incubation period the cells were fixed with a fixative (1 volume of glacial acetic acid : 3 volumes of ethanol) for 5 min. The fixative was then removed and the cells air dried. Hoechst 33258 stain was

prepared from a 1 mg/ml stock by diluting 1:20000 in phosphate buffered saline. This solution was poured onto the fixed cells and incubated at room temperature for 10 min. The stain was removed and cells were washed two times in water. The cells were viewed by fluorescence microscopy using a water immersion lens. Hoechst 33258 is a fluorescent stain for DNA. If the cells were mycoplasma-free, only the cell nucleus fluoresced; however if the original cells were mycoplasma infected fluorescence was also detected in the cytoplasm of the NRK49 cells.

2.2 DNA methodology

2.2.1 Gel electrophoresis of PCR products

rt-PCR products were analyzed by 2% agarose gel electrophoresis (5 volts/cm for 2-3 hours) stained with ethidium bromide (0.5 μg/ml). A 100 bp DNA ladder or ϕX174 RF DNA/Hae III Fragments (Gibco BRL) were used as markers for comparison. The gel was then photographed under ultraviolet illumination using a polaroid camera.

2.2.2 Southern blot analysis of DNA

PCR products were transferred from ethidium-stained agarose gels onto nylon membrane (Hybond N^+, Amersham,UK.) by overnight capillary transfer using 0.4M NaOH (Maniatis et al., 1989). Filters were then placed in roller bottles (Hybaid, UK.) and prehybridised at 65°C for 4-6 hours in 20 ml of prehybridisation buffer containing 6X SSC (0.9M NaCl, 0.9M Na citrate), 10X Denhardts (0.2% (w/v) each of ficol, polyvinylpyrrolidone and bovine serum albumin) and 100 μl/ml denatured sonicated salmon sperm DNA.

Prehybridisation buffer was removed and hybridisation fluid containing 1X 10^6

cpm/ml of [^{32}p] γ ATP labelled oligonucleotide probe complementary to an internal sequence in the PCR product for MUCI, and hybridisation was continued for further 12-18 h (For CK19 and CD44 the downstream primers were used as probes). Oligonucleotide hybridisation temperature was calculated from an imperical formula Ti + 4X (G or C) + 2X (A or T) - 5°C and hybridisation was continued for further 12-18 hours.

Filters were then washed at Ti for 60 minutes in a buffer containing 6X SSC, 0.1% SDS and for 5 minutes in the same buffer at Ti + 5°C. At the completion of washing, filters were placed in polythene bags and exposed to X-ray film (Kodak XAR-2 or equivalent) in cassettes with intensifying screen and exposed at -70°C for 12-36 hours. Films were developed in a Kodak M7A automatic X-ray processor.

2.2.3 Preparation and quantitation of oligonucleotides

Primers were designed on an Applied Biosystem model 381A DNA synthesizer. 5' trityl groups were removed on the synthesizer before the oligonucleotide was eluted from the column in concentrated (29%) ammonia solution for 1.5h at room temperature. Oligonucleotide primers were deprotected by incubation in ammonia in a sealed glass vial overnight at 55°C. The deprotected primers were then purified using the COP cartridge method (Cruachem Ltd., Glasgow, Scotland). Oligonucleotides were suspended in water and quantified spectrophotometrically. An A260 value of 1 was taken to equal 33mg/ml of oligonucleotide. A ratio of A260/A280 was approximately 1.8 when sufficently purified.

2.3 RNA METHODOLOGY

2.3.1 Collection of samples

Peripheral blood samples were obtained from healthy female volunteers aged between 18 and 48 years. The first 5 ml were discarded and the second 5 ml were collected in potassium EDTA bottles and promptly transported to the laboratory at 4^0C for immediate processing.

2.3.2 RNA extraction, purification & quantification

In preparation and handling of RNA, scrupulous precautionary steps were taken in order to avoid degradation by contaminating Rnases which may yield spurious results. All techniques were carried out wearing gloves, tubes and solutions were pre-cooled on ice and centrifugation carried out at 4^0C and aerosol-resistant, Dnase- and Rnase-free tips were used for all manipulations. All tubes were also pre-treated with diethylpyrocarbonate (DEPC), an irreversible inhibitor of RNases. The tubes were immersed in a solution of 0.1% (v/v) DEPC for a minimum of 2 hours in a fume hood, the solution decanted off and the tubes autoclaved followed by drying in an oven at 80^oC. The tubes were subsequently stored in sealed DEPC-treated containers. Distilled water was also DEPC-treated: 0.1% (v/v) solution was prepared, incubated overnight and autoclaved.

Total cellular RNA was extracted from cell lines and normal whole blood using RNAzol^{TM} B (Biogenesis Ltd.) according to the manufacturer's instructions. The concentration of aqueous solutions of RNA was measured spectrophotometrically. An A260 value of 1 was taken to equal 40 μg/ml RNA and A260 /280 ratio of 1.9 represented essentially pure RNA.

2.3.3 Reverse transcriptase-polymerase chain reaction (rt-PCR)

Reverse transcription of RNA and PCR amplification of cDNA was carried out using RNA PCR Kit (GeneAmp RNA PCR kit, Perkin-Elmer Cetus, Norwalk, Connecticut, USA) following the manufacturer's instructions, modified to facilitate treatment of samples with DNase 1.

The Dnase 1 treatment was carried out in the same reaction mix prior to addition of the reverse transcriptase.

The oligonucleotide primers for rt-PCR amplification were synthesised on the Applied Biosystem DNA Synthesiser as described in section 2.2.2.3.

One microgramme of total RNA in 2μl of DEPC-treated water was mixed with 16.5 μl of reverse transcriptase master mix excluding reverse transcriptase [final concentrations in 20μl = 5 mM $MgCl_2$, 1X PCR buffer (50 mM KCl, 10mM Tris-HCl, pH 8.3), 1 mM each dNTPs, 20 units of RNase inhibitor and 2.5μM random hexamers]

Contaminating DNA was removed by adding 0.4μl of 0.25 μg/ μl of RNase-free DNase 1 (USB) and incubating at 37^0C for 30 minutes. The reaction was then stopped by heating to 95^0C for 5 minutes and then cooled to room temperature. One microlitre of 50 U/μl of maloney murine leukaemia virus reverse transcriptase was then added and the rt-PCR procedure followed according to the manufacturer's instructions, including one round of reverse transcription (42^0C for 15 minutes, 99^0C for 5 minutes and soaked at 5^0C for 5 minutes) in a total volume of 20 μl, followed by addition of 80μl of PCR master mix containing Taq polymerase and primers specific for the genes of interest (MUC1, K19 or CD44) (final concentration in 100 μl = 2mM MgCl2, 1X PCR buffer, 2.5 units of

AmpliTaq DNA polymerase, and 100 ng of each upstream and downstream specific primers).

The PCR reaction was carried out in a Perkin-Elmer Cetus 480 DNA thermal cycler; the PCR conditions for each set of primers are given below. A preceding single cycle was the same for all sets of primers and consisted of denaturing cycle at 93^0C for 5 min, primer annealing at 55-60^0C for 5 min and polymerization at 72^0C for 15 min.

The PCR conditions for MUC1 were 30 cycles: 93^0C for 1 min, 60^0C for 1 min and 72^0C for 5 min and a final extension cycle at 72^0C for 15 min. The MUC1 primers produced a 288 bp PCR product.

For CK19, the PCR conditions were, 40 cycles of a single round were carried out as described above to yield a 1097 bp PCR product. Where two rounds of PCR were performed, nested primers were used as described by Datta et al. (1994).

For CD44, the PCR conditions were, a first round of 20 cycles (using the primers for the standard form of CD44 which produced a fragment of 486 bp) and a second round of 30 cycles where 1-2 µl of first round PCR product was used as a template in a 100µl reaction containing either Exon 8-11 or Exon 12-14 primers.

For all samples, the quality of RNA was routinely checked by running the standard form of CD44(S) as an internal control. Positive controls were RNA extracted from MCF-7 cell lines and 'negative controls' contained all components of the rt-PCR reaction but no target RNA template. 'Positive controls' of RNA extracted from MCF-7 and/or MDA cell lines were routinely included.

2.3.4 Primer sequences for rt-PCR

Prepared as previously described in 2.2.2.3

MUC 1:

MUC1 - 5' primer	5' CGTCGTGGACATTGATGGTACC - 3'
MUC1 - 3' primer	5' GGTACCTCCTCTCACCTCCTCCAA - 3'

CK19:

CK19 - 5' primer	5'- TTATTGGCAGGTCAGGAGAAGAGCC -3'
CK19 - 3' primer	5'- AGCTAACCATGCAGAACCTCAACGACCGC -3'

CD44:

CD44 (S) - 5' primer	5' AGTCACAGACCTGCCCAATGCCTTTG - 3'
CD44 (S) - 3' primer	5' CACCTTCTTGACTCCCATGTGAGTGT - 3
CD44 (Exon 8-11) - 5' primer	5' CTTGGATCCAACCACACCACGGGCTTTTGACCACA -3'
CD44 (Exon 8-11) - 3' primer	5' CTTGGATCCTTCTTCCTGCTTGATGACCTCGTCCC -3'
CD44 (Exon 12-14) - 5' primer	5' ATATGGACTCCAGTCATAGTACAACGCTTCAGC -3'
CD44 (Exon 12-14) - 3' primer	5' CTGATAAGGAACGATTGACATTAGAGTTGGAAT -3'

2.3.5. Seeding experiments

To assay the sensitivity of the technique, seeding experiments were carried out by adding varying concentrations of MCF-7 cells (0 -10^6 cells) to 10 mls of normal blood. RNA was then extracted and quantified as described in section 2.2.3.2 and rt-PCR for MUC1 and CK19 performed as described in section 2.2.3.3.

2.4 RNA Fingerprinting

RNA Fingerprinting reactions were carried out using Delta™ RNA Fingerprinting Kit (Clontech Laboratories, Inc., CA94303-4607, USA) following the manufacturer's instructions.

2.5 Molecular cloning

Molecular cloning reactions were carried out using TA Cloning Kit and INV α F'

One Shot™ Kit (Invitrogen Corporation, San Diego, USA) following the manufacturer's instructions.

3. Results

For all reactions PCR optimization was performed under the conditions described in section 2.2.3. Those initial steps involved adjusting the annealing temperature and all the PCR components to determine if the conditions for amplifications were optimal and to eliminate the possibility of non-specific amplification by the primers. This step was time consuming because of the many variables involved in different experiments.

There was no detectable difference in PCR amplification products in samples collected in bottles containing Lithium heparin or EDTA demonstrating no significant inhibitory effect of heparin anticoagulant on PCR (Figure 6).

The seeding experiments demonstrated a decreasing signal from MUC1 mRNA corresponding to a decreasing number of seeded tumour cells, reflecting an apparent sensitivity of 1 tumour cell per ml of blood. However, the unseeded blood sample in lane 1 showed a faint band, casting doubt on the specificity expression of MUC1 mRNA (figure 1A). Further samples of normal peripheral blood showed variable expression of MUC1 mRNA in 4 of 6 volunteers (figure 1B). A further experiment was performed using the standard form of CD44 (CD44s) mRNA as internal control. Similar background expression was demonstrated in peripheral blood samples from healthy human volunteers (figure 1C). This variation in expression was not due to degradation of RNA in some specimens since CD44s was equally represented in each (figure 1D). In total, peripheral blood from 21 of 23 volunteers showed positive bands when assayed for MUC1 mRNA.

The results of the assay for variants of CD44 in peripheral blood from 10 healthy volunteers is depicted in figure 2. The metastasis-associated variant (Exons 8-11) mRNA is present in 4 of 10 samples.

In contrast, CK19 mRNA was not detected in 48 samples of normal peripheral blood (figure 3 and 4). However, under the experimental conditions used, the seeding experiments demonstrated that it was impossible to detect tumour cells at concentrations less than 10^4 cells per 10 ml of blood i.e. one tumour cell per 10^4 white cells (figure 5). This failure of detection did not appear to be due to degradation of RNA since CD44s mRNA was detected in all these samples by rt-PCR (figures 3A and 5A). In order to increase the sensitivity of detection, we attempted to use nested primers as described by Datta et al (1994), but abandoned this in light of frequency of detection of CK19 mRNA in blood from healthy volunteers.

Figure 1

Figure 1. Sensitivity and specificity of MUC1 mRNA

A: Autoradiographic detection of MUC1 mRNA in blood samples from a healthy subject mixed with serial dilutions of MCF-7 cells after 30 cycles of PCR amplification and hybridization with 32p-labelled MUC1 oligonucleotide.

Lanes 1 to 6 = 0, 10, 10^2, 10^3, 10^4 and 10^5 MCF-7 cells/10 ml of blood respectively; lane 7 = MCF-7 mRNA.

B: Autoradiograph of Southern blot showing hybridization of 32p-labelled MUC1 oligonucleotide to 30 cycles PCR product. It demonstrates 288 bp DNA in the MCF7 mRNA (lane 8) and in 4 out of 6 normal peripheral blood samples (lanes 2,5,6 and7).

Lane 1= "No"RT control.

C: Ethidium bromide-stained agarose gel of rt-PCR product from MUC1 mRNA in 5 other normal bloods (lanes 2-6). Lane 1= "No" RT control, Lane 7 = MCF-7 mRNA.

D: Ethidium bromide-stained agarose gel of rt-PCR product for CD44s mRNA as internal control for the integrity of RNA of all samples of normal blood and MCF-7 mRNA in Fig 1C.

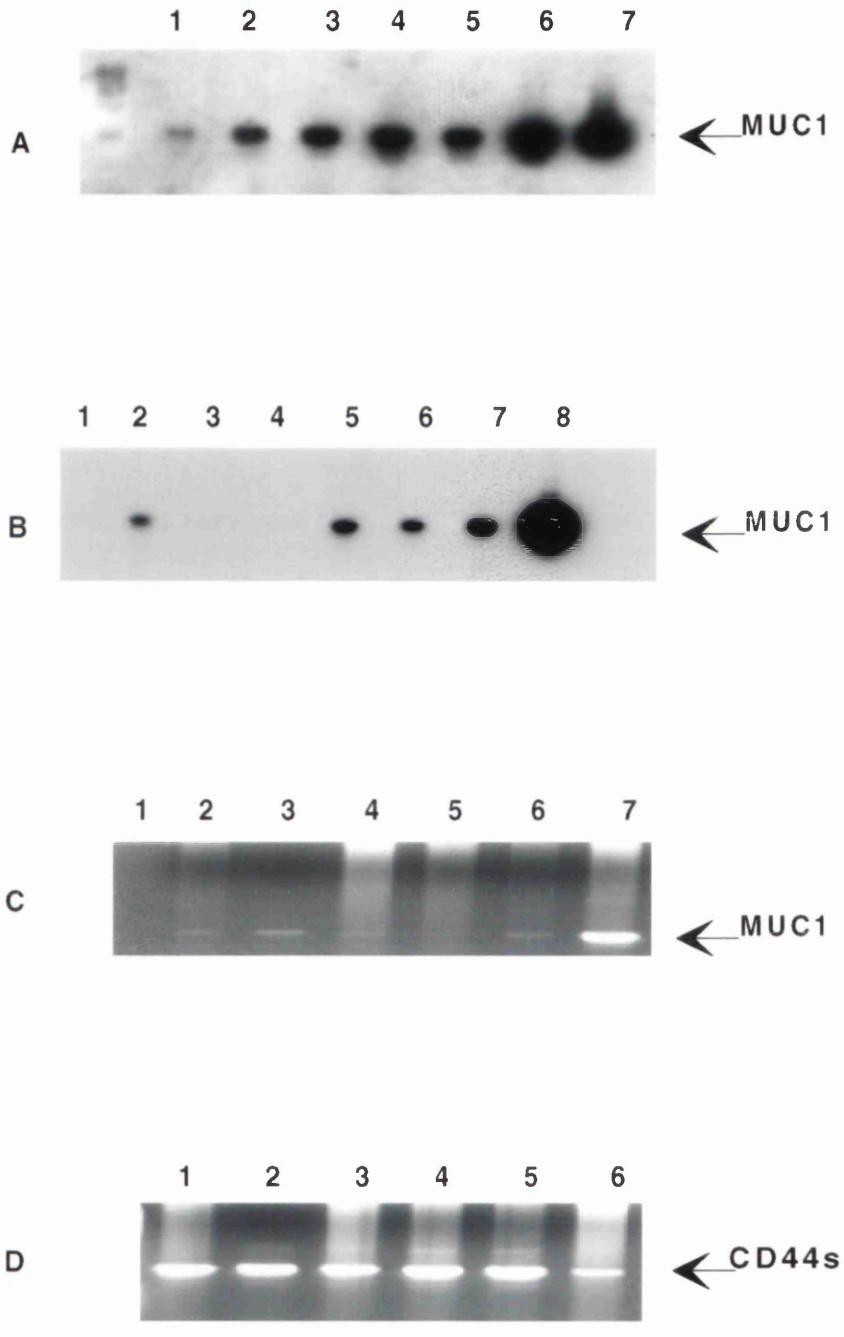

1 2 3 4 5 6 7
A
MUC1
1 2 3 4 5 6 7 8
B
MUC1
1 2 3 4 5 6 7
C
MUC1
1 2 3 4 5 6
D
CD44s

Figure 2

Figure 2. Detection of CD44 (Exon 8-11) mRNA in normal blood

Ethidium bromide-stained agarose gel of rt-PCR product for CD44 mRNA showing expression of exon 8-11 in 4 out of 10 samples of peripheral blood from healthy volunteers (lanes 4,5,8 and 10). The rt-PCR product of exon 12-14 is used here as internal control.

Lane 11 = MCF-7 mRNA and Lane 12 = "No"RT control.

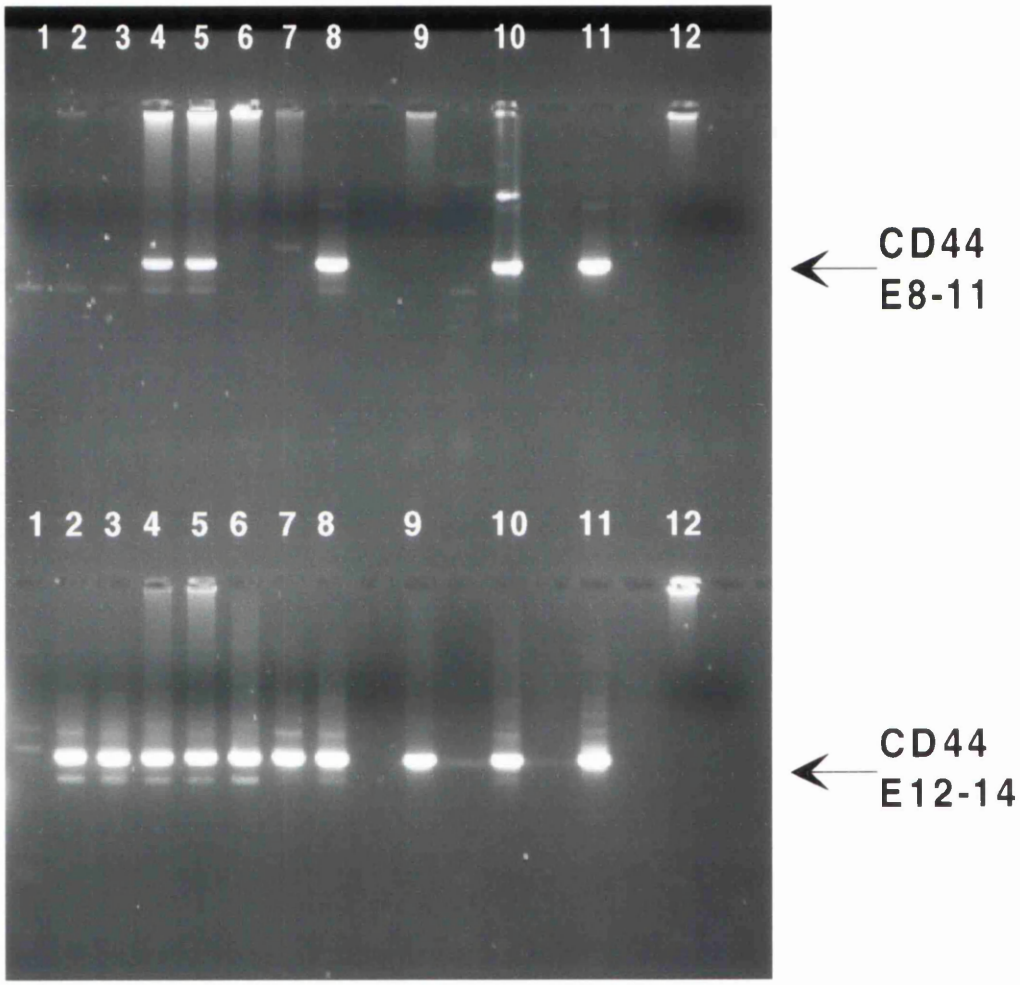

1 2 3 4 5 6 7 8 9 10 11 12
CD44
E8-11
1 2 3 4 5 6 7 8 9 10 11 12
CD44
E12-14

Figure 3

Figure 3. Specificity of CK19 mRNA

A: Ethidium bromide-stained agarose gel showing 486 bp bands of CD44s as internal control for 5 blood samples (lanes 2-6) and MCF-7 cells (lane 7).

B: Ethidium bromide-stained agarose gel of rt-PCR product from CK19 mRNA. Lanes 7 = MCF-7 mRNA. Lanes 2-6 = Normal blood samples showing no visible bands.

C: Southern blot showing hybridization of 32p-labelled CK19 oligonucleotide to 40 cycles PCR product. This shows a 1069 bp band in the MCF-7 cells and confirms its absence in 5 normal blood samples. Lane 1 = "No" RT control.

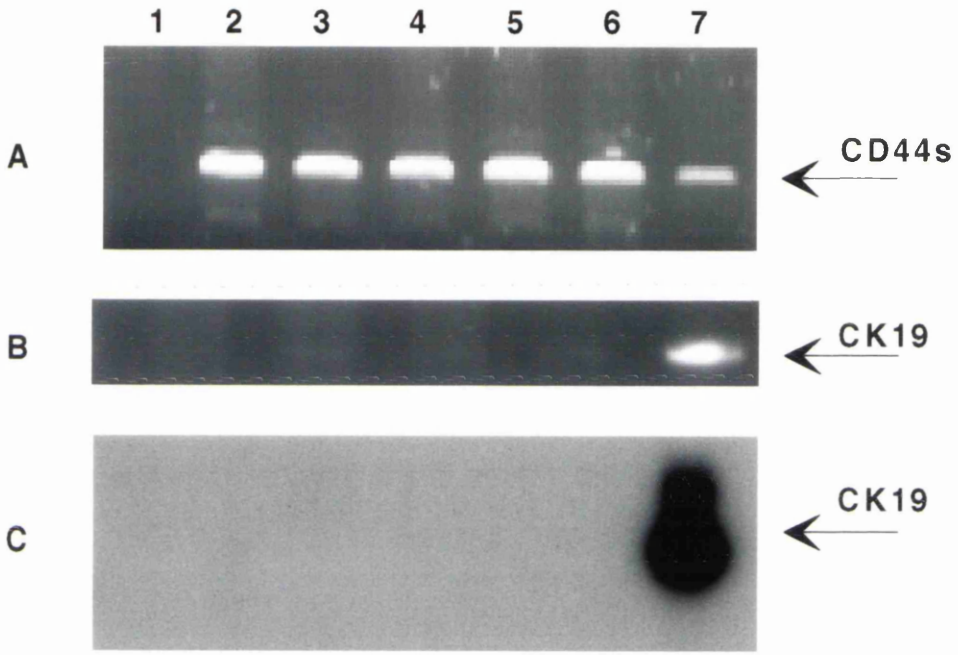

1
2
3
4
5
6
7
A
CD44s
B
CK19
C
CK19

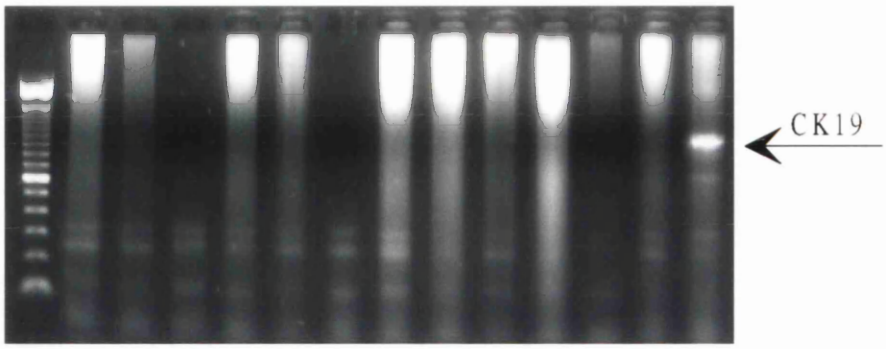
Figure 4

Figure 4. Specificity of CK19 mRNA

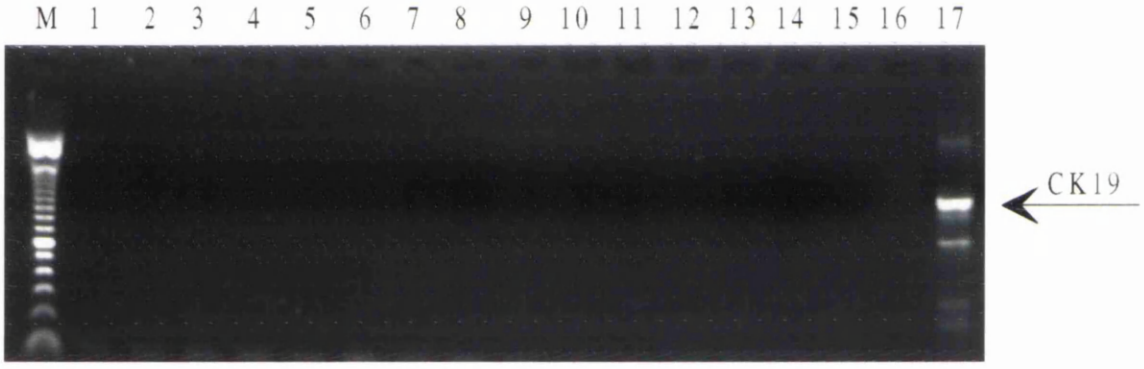
A: Ethidium bromide-stained agarose gel of rt-PCR product from CK19 mRNA. Lane 13 = MCF-7 mRNA. Lanes 2-12 = Normal blood samples showing no visible bands. Lane 1= "No" RT control.

B: Ethidium bromide-stained agarose gel of rt-PCR product from CK19 mRNA. in 15 other normal bloods (lanes 2-16). Lane 1= "No" RT control, Lane 17 = MCF-7 mRNA.

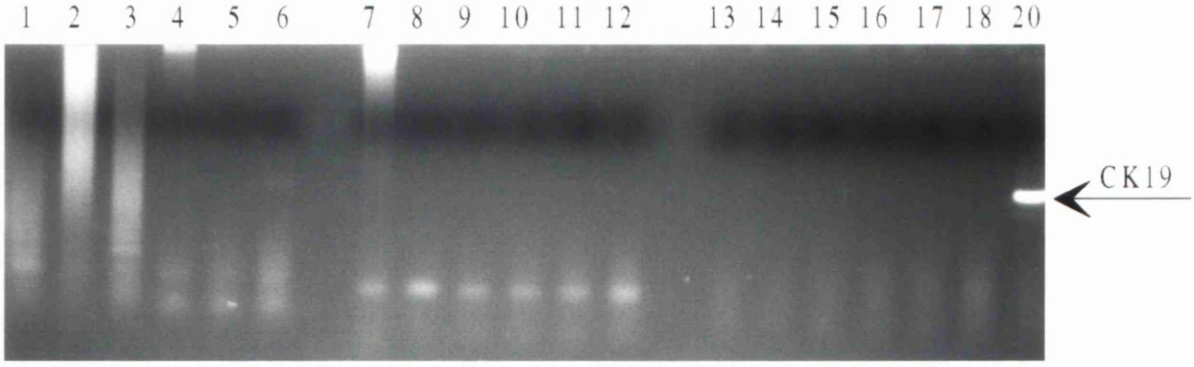
C: Ethidium bromide-stained agarose gel of rt-PCR product from CK19 mRNA. in 17 other normal bloods (lanes 2-18). Lane 1= "No" RT control, Lane 20 = MCF-7 mRNA.


CK19

A


M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
CK19

B


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 20
CK19

C

Figure 5

Figure 5. Sensitivity of CK19 mRNA

A: Ethidium bromide-stained agarose gel showing bands of CD44s as internal control for the integrity of RNA.

B: rt-PCR detection of CK19 mRNA in blood samples from a healthy subject mixed with serial dilutions of MCF-7 cells detected on ethidium bromide-stained agarose gel after 40 cycles of PCR amplification.

C: Southern blot and hybridization of 32p-labelled CK19 oligonucleotide to 40 cycles PCR product of Fig-4B.

Lane1= "No" RT control; lanes 2 to 8 = 0, 10, 10^2, 10^3, 10^4, 10^5 and 10^6 MCF-7 cells/10 ml of blood, respectively.

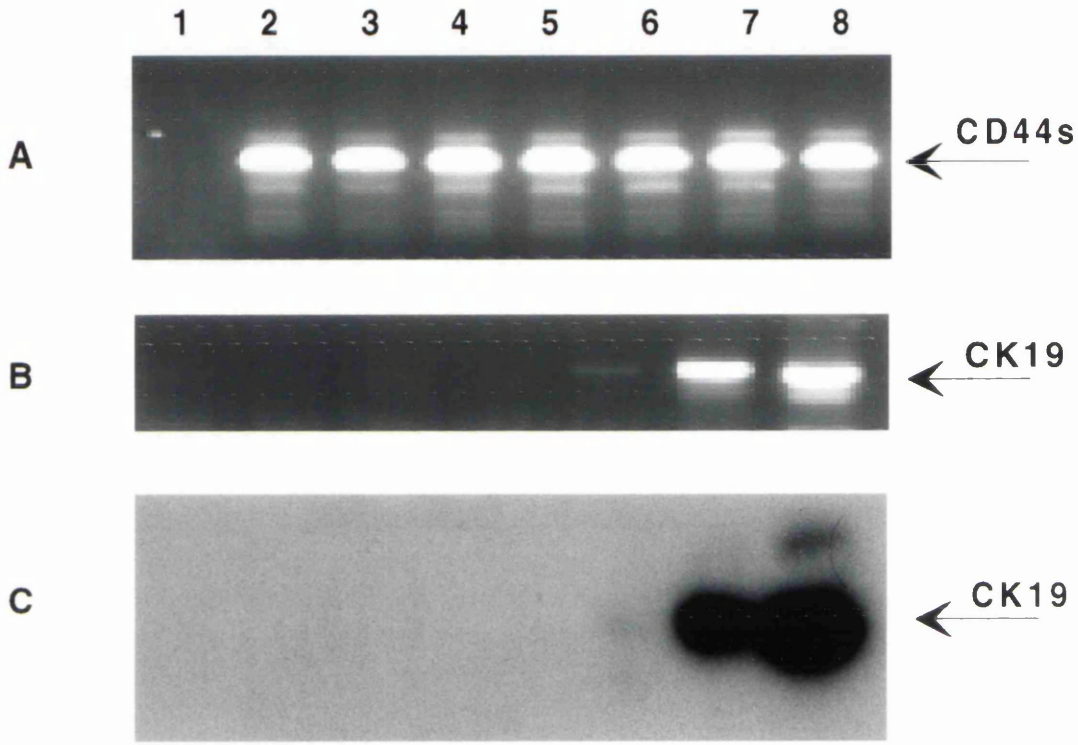
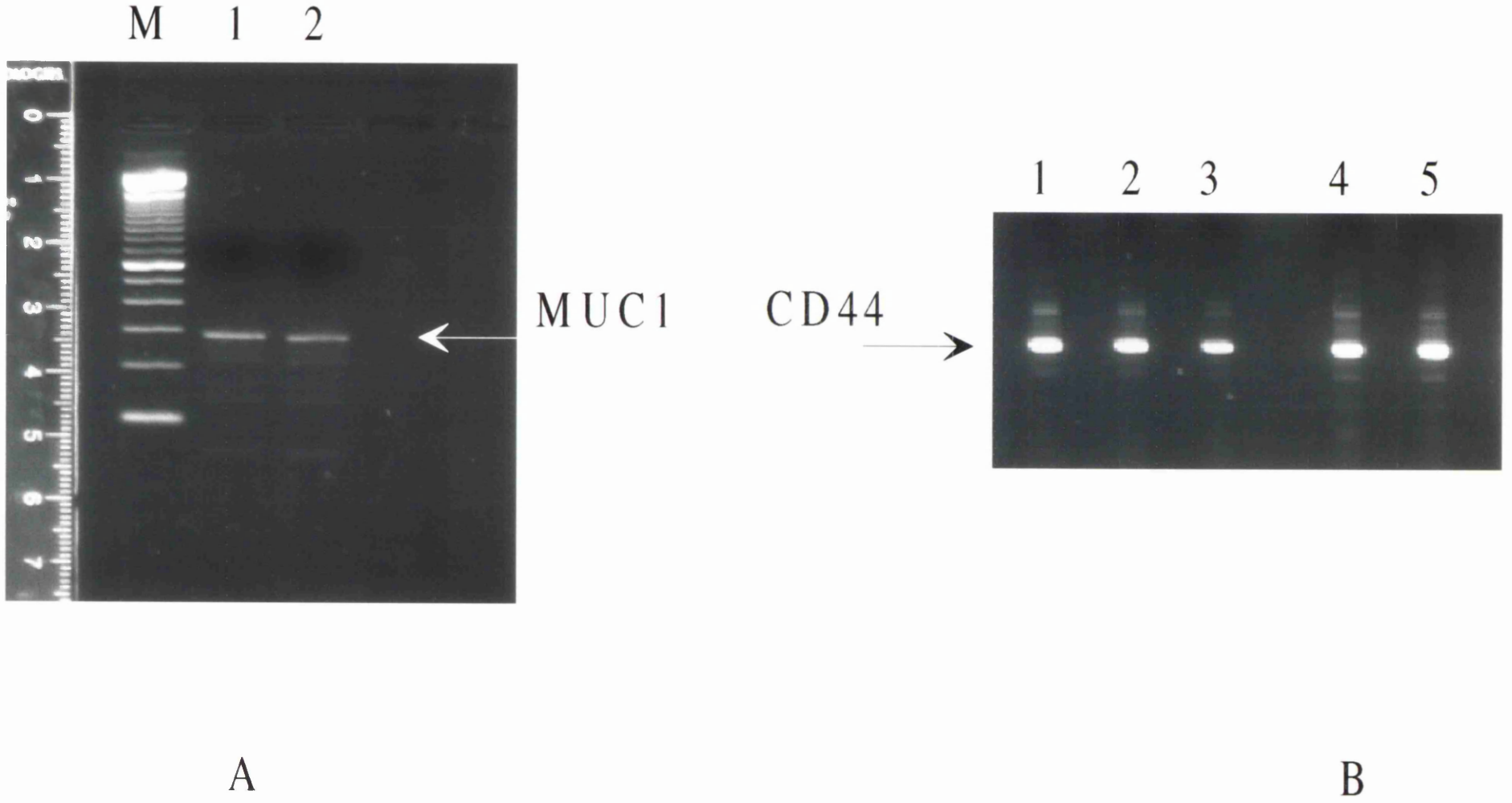

1
2
3
4
5
6
7
8
A
CD44s
B
CK19
C
CK19

Figure 6

Figure 6. Heparin vs EDTA

A: Ethidium bromide-stained agarose gel of rt-PCR product from MUC1 mRNA in the same sample of normal blood (lanes 1 = sample collected in bottle containing Lithium heparin-2 = sample collected in bottle containing Potassium EDTA). Lane M= 100 bp ladder marker.

B: Ethidium bromide-stained agarose gel of rt-PCR product for CD44s mRNA (lanes 1-6). Lanes 1-3 = samples collected from three normal volunteers in bottles containing Lithium heparin. Lanes 4-6 = corresponding sample collected in bottle containing Potassium EDTA .


M
1
2
MUC1
CD44
1
2
3
4
5
A
B

RNA Fingerprinting

Differential display reactions were performed to compare the pattern of expression of breast carcinoma with that of normal blood. For these reactions multiple tumour and blood samples were used. RNA was prepared from ten malignant breast tumours and ten blood samples from healthy female volunteers. RNA was extracted as previously described in section 2.2.3. RNase-free DNase was used to remove contaminating DNA as this may interfere with the identification of differentially expressed bands (Liang et al., 1995). The initial fingerprinting experiments demonstrated that all contaminating DNA had been removed .

A 1μg aliquot of each tumour RNA was taken and combined to give the tumour mix. The same was done for the normal blood samples. Each RNA mix was reverse-transcribed into cDNA.

The generation of fingerprints is sensitive to template quality and small changes in template concentration. It is therefore necessary to control for the possibility that an observed difference between two RNA samples is due to slight differences in RNA quality or concentration. Repeated fingerprinting at only one concentration does not control for these effects. The simplest way to ensure the authenticity of a differentially amplified product is to fingerprint a titration of the RNA. Only products that occur at two or more concentrations in one sample and not at all in the other sample should to be considered (McClelland et al., 1994). Therefore each cDNA was diluted to give the "A" dilution (1 in 10) and the "B" dilution (1 in 40) and fingerprinted.

A total of 30 anchored and arbitrary primer combinations were used (T1-9 and P4-6). The number and intensity of the bands in a fingerprint depended on the primer combination used. A small number produced fingerprints which were too faint to be of use. Using this limited primer combination, a total 18 of such differentially expressed

bands were recovered for further characterization. These bands ranged in size from approximately 300-700 bp.

As shown in figure 7 many PCR products were found to be expressed in both breast cancer and normal blood. Whereas other products were expressed in breast cancer only.

These differentially expressed bands were cut out from the dried down gel, reamplified and cloned into the TA cloning vector pCRII (Invitrogen).

Reamplification of PCR products under the conditions described in section 2.3 has confirmed the size of the differentially expressed bands (Figure 8).

Before ligation into a vector, the reamplified bands were further treated with phenol/chloroform and figure 9.B demonstrate the 450 bp band after phenol/chloroform extraction.

Plasmid minipreps of 2 to 4 clones from each band were digested with *Eco*R1 and run on an agarose gel to determine the size of the insert. Figure 10 shows the 450 bp fragment after digestion of the insert of nine clones with *Eco*R1.

A selection of the clones were partially sequenced (seq1-18) and the sequences compared against the GenBank/EMBL nucleic acid data bases .

The results of sequencing of multiple clones produced from one band have shown that the inserts were often heterogenous with respect to size, but also the sequence of the clones with apparently the same size insert were different. This observation has been subsequently confirmed with all other bands as they have been processed further (data not shown).

When tested as a potential marker, the 450 bp band was found to be expressed in 4 normal blood samples as figure 11 shows. Similar experiments showed all of 18

differentially expressed bands studied, demonstrated variable degree of expression in normal blood rendering them non-specific.

Figure 7

Figure 7. PCR Fingerprinting reactions

Autoradiograph of polyacrylamide gel electrophoresis showing differential display of mRNA from normal blood versus breast carcinoma using three different primer combinations.

Lane	cDNA sample	Primers
1	1A	P5T7
2	2A	P5T7
3	1B	P5T7
4	2B	P5T7
5	1A	P5T8
6	2A	P5T8
7	1B	P5T8
8	2B	P5T8
9	1A	P5T9
10	2A	P5T9
11	1B	P5T9
12	2B	P5T9

RNA from breast carcinoma were compared to RNA of normal blood,
1A: dilution A of cDNA made from breast cancer RNA
1B: dilution B of cDNA made from normal blood RNA
2A: dilution A of cDNA made from breast cancer RNA
2B: dilution B of cDNA made from normal blood RNA

Lane 8 shows 4 differentially expressed band fragments (arrows) which were excised for further analysis,
Band 1 ~ 750bp
Band 2 ~ 500bp
Band 3 ~ 450bp
Band 4 ~ 390bp

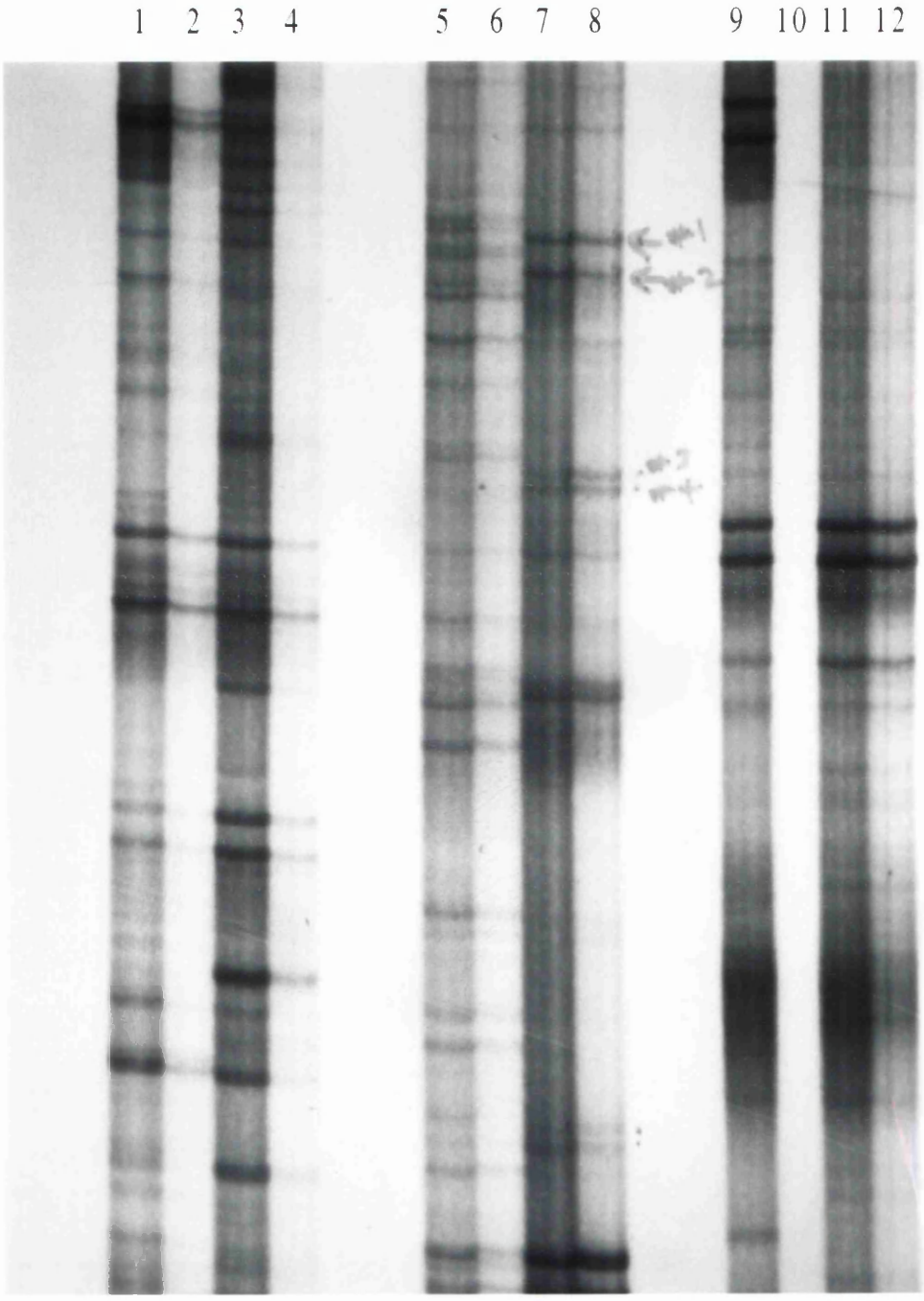

1 2 3 4
5 6 7 8
9 10 11 12

Figure 8

Figure 8. Reamplification of differentially expressed bands

2% Ethidium bromide gel after 20 cycles of PCR amplification of eluted bands with the primer sets used in the original fingerprinting reaction. Lane M = 100 bp ladder marker. Lane 1= 800 bp band. Lane 2 = 450 bp band. Lanes 3-8 = other 5 bands between 400- 600 bp. Lane 9 = "No" RT control.

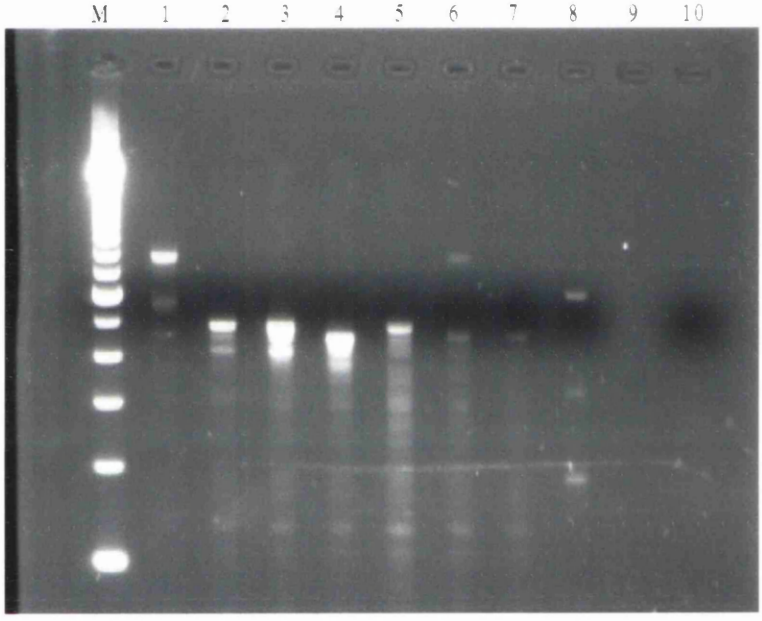
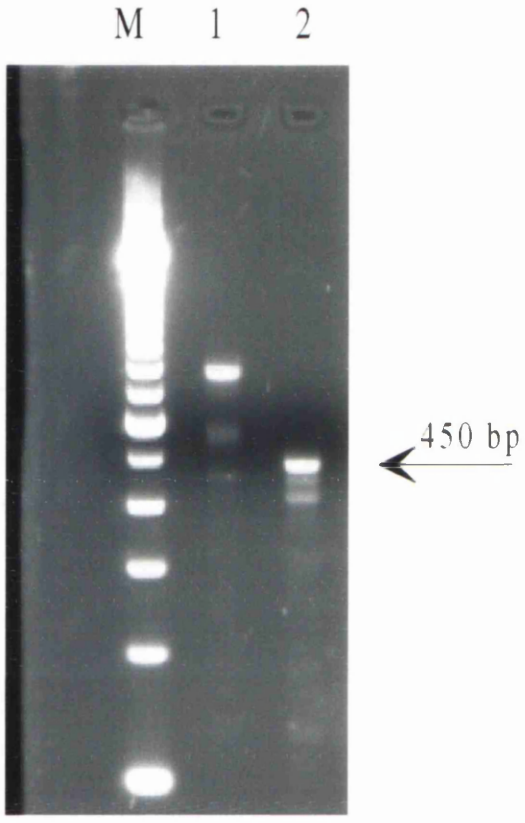

M
1
2
3
4
5
6
7
8
9
10

Figure 9

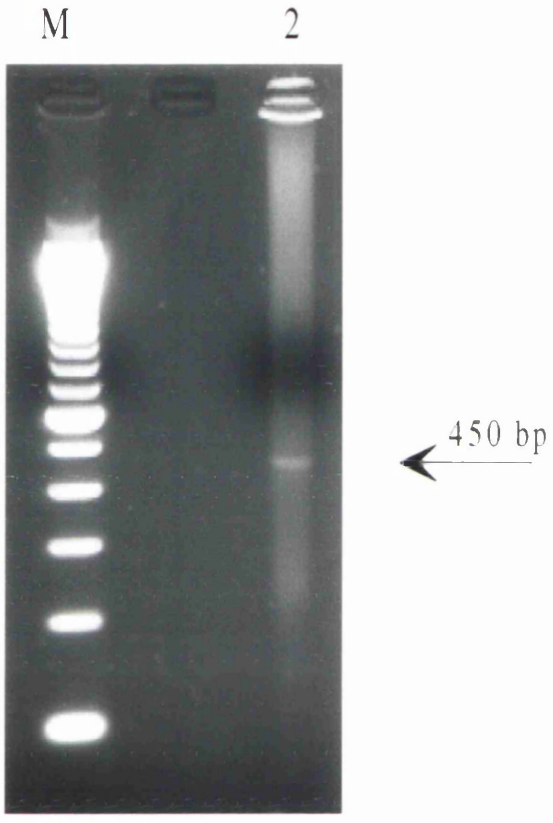
Figure 9. Reamplification of differentially expressed bands

A: 2% Ethidium bromide gel after 20 cycles of PCR amplification of eluted bands Lane M = 100 bp ladder marker. Lane 1= 800 bp band. Lane 2 = 450 bp band.

B: Lane 2 = DNA amplification product after phenol/chloroform extraction for the 450 bp band shown in A, which is now ready for ligation into pCR™II.



A



B

Figure 10

Figure 10. Digestion of inserts

2% Ethidium bromide-stained agarose gel

Lane M = 100 bp ladder marker, lanes 1-9 = digestion of the inserts in from 9 white colonies showing the 450 bp bands (bottom arrow) and the rest of plasmid (top arrow), lane 10 = a blue colony which does not contain the 450 bp insert.

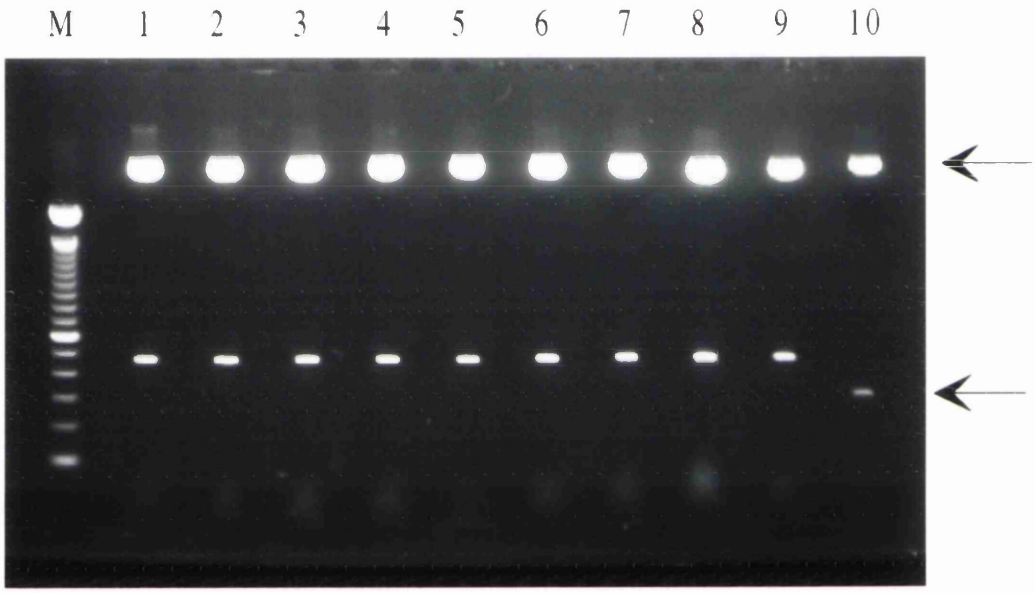
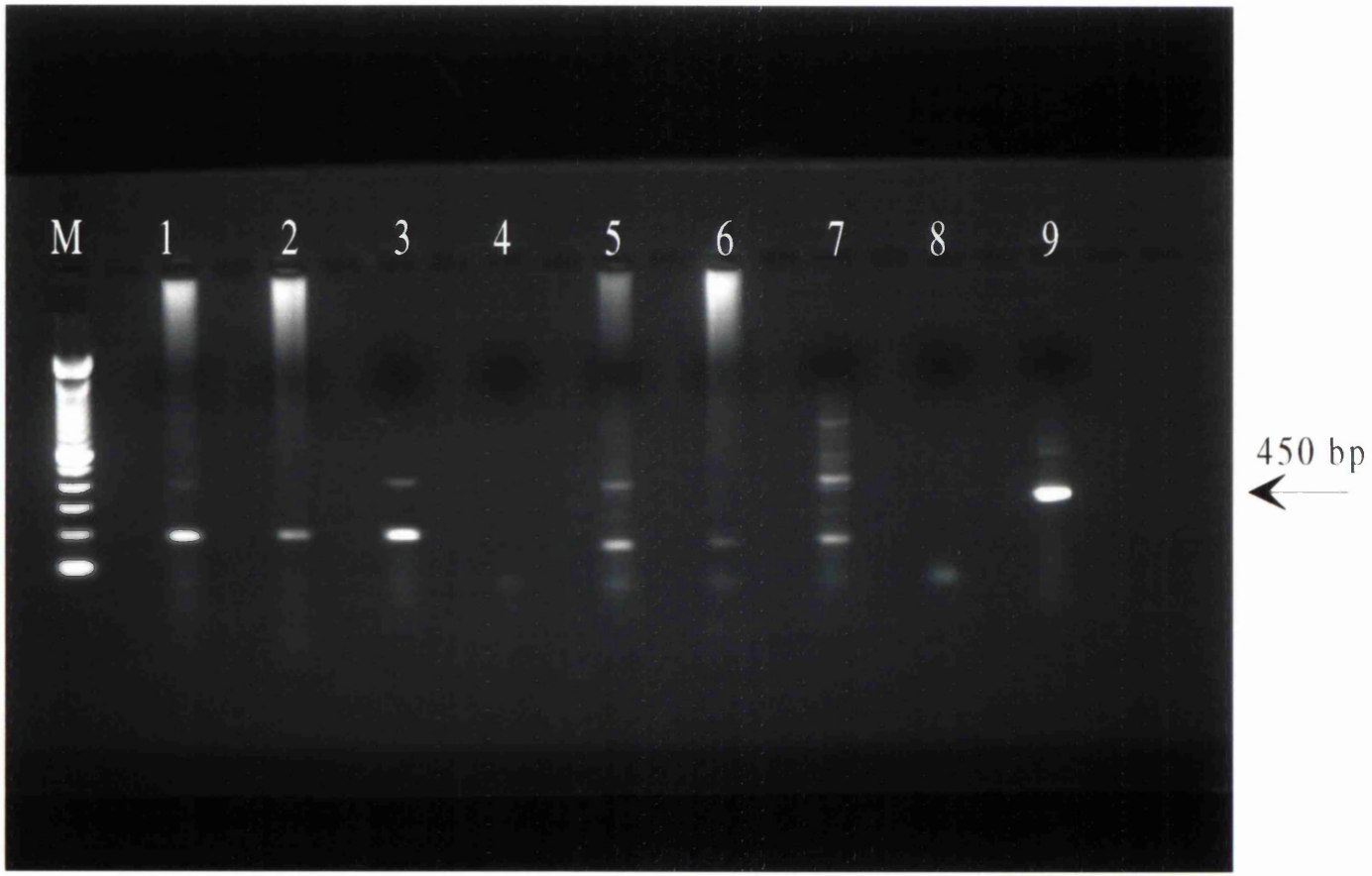

M
1
2
3
4
5
6
7
8
9
10

Figure 11

Figure 11. Specificity of the 450 bp differentially expressed band

2%Ethidium bromide-stained agarose gel of rt-PCR product from (450 bp band) mRNA. Lane 1 = 100 bp ladder marker. Lane 8 = "No" RT control . Lane 9 = MCF-7 mRNA l, Lanes 1-7 = Normal blood samples showing no visible bands in lanes 2,4 and 6.


M
1
2
3
4
5
6
7
8
9
450 bp

4. Discussion

The previously reported results of a study by Brown et al. (1995) concluded that the DF3 (MUC1) antigen may be used as a specific marker for detecting circulating tumour cells in patients undergoing surgery for breast cancer. Under similar experimental conditions we have demonstrated almost ubiquitous background expression of MUC1 in the peripheral blood of 21 of 23 healthy female volunteers. In support of these findings, Hoon et al. (1995) also detected MUC1 expression in 7 of 8 normal donor peripheral blood leucocytes and four normal lymph nodes.

The specific occurrence of variant CD44 mRNA in tumour cells has been described by Matsumura et al. (1992). They reported that peripheral blood leucocytes expressed the standard form of CD44 without the splice variants. CD44 (Exon 8-11) is implicated in the metastasis of rat pancreatic tumour cells (Gunthert et al., 1991; Hofmann et al., 1991) and the sequences spliced into a CD44 variant believed to be expressed in colonic tumour and breast cancer (Delatorre et al., 1995). Our results also raise serious doubts about the value of using CD44 (Exons 8-11) to detect circulating tumour cells because of the frequency with which it is found in normal peripheral blood. It is not known whether expression of these two markers in normal peripheral blood is due to low level of expression of genes in a proportion of peripheral blood cells, or due to presence of a few epithelial cells in the blood of some individuals.

CK19 was not expressed in the peripheral blood of healthy volunteers and in theory could be a specific epithelial marker. However, the sensitivity of detection of CK19 in a single round of amplification was limited to greater than 10^4 cells per 10 ml of blood. Because of the

frequency with which CK19 mRNA was detected in normal blood after two amplification rounds, we abandoned using the nested primers described by Datta et al. (1994) which were designed to increase sensitivity. We also found increasing the number of cycles in a single amplification round above 40 cycles enhanced sensitivity but reduced specificity as previously reported by Schoenfield et al.(1994). Several reptitions produced the same results. The failure to detect lower concentrations of tumour cells did not appear due to degradation of RNA since undegraded CD44(S) mRNA was detected in these samples by rt-PCR.

Our studies clearly demonstrate that, while rt-PCR has the potential to detect minute quantities of a specific RNA sequence against a background of a vast excess of other RNA, there are dangers in assuming that the technique can be used to detect minute numbers of specific cells because of the well-established phenomenon of illegitimate transcription which results in very low levels but non-specific expression of irrelevant genes in many cell types (Chelly et al., 1989). Illegitimate transcription limits the usefulness of rt-PCR to detect another putative epithelial cell-specific mRNA (De Graaf et al., 1997). Nevertheless, we used CK19 primers which were designed to preclude the amplification of the pseudogene sequence (Datta et al., 1994). We also routinely treated all RNA preparations with RNase-free DNase1, and ran 'no-RT' controls with all reactions.

The data on CK19 in the literature are contradictory. Several authors have reported no transcription of CK19 mRNA in normal blood (Datta et al.,

1994; Traweek et al., 1993; Kruger et al., 1996) whilst transcription of CK19 mRNA in normal blood has been reported by Burchill et al., 1994 and Krismann et al., 1995). Noguchi et al. (1996) reported no transcription of CK19 mRNA in lymph nodes, while transcription in lymph nodes was observed by Schoenfield et al., (1994).

It is clear that rt-PCR can be a highly sensitive and specific method of detecting minute quantities of a specific RNA. However, this study raises serious concern about the background expression detected in peripheral blood of healthy female volunteers.

The technique of immunobead-PCR may be a solution to the problem if combined with detection of an irrefutable genetic marker of a tumour such as K-ras mutations (Hardingham et al., 1993; 1995). It may be applicable in those tumours that carry the mutated gene. Such a specific mutation is as yet unknown in breast cancer.

The PCR is difficult to quantitate, nevertheless, serial dilutions can be performed by comparing the amplification over a range of RNA or DNA concentrations for a target gene with a control gene during the exponantial phase of amplification. As the level of gene expression and transcription rate between individuals vary according to the prevalence of the encoded protein, this semi-quantitative way may not provide distinct information regarding tumour cell number. It may also fail to give sound and accurate information despite the continued presence of tumour cells when transcription is momentarily down regulated due to the effect of chemotherapy. Further more, because of intermittent tumour cell

shedding, blood samples may not contain the adequate number of circulating tumour cells.

Immunofluorescent detection using a multi-colour FACS technique are alternative approaches for detection of circulating tumour cells, but sample size is a limiting factor for the utilisation of immunofluorescent detection using a multi-colour FACS technique.

Further work is essential to identify a suitable epithelial marker that has the required specificity and sensitivity to be used reliably in the detection of circulating tumour cells from solid tumours and meticulous attention to methodology is required to overcome the technical problems with this method before its broader use in clinical practice can be recommended.

The primer combination used in the latter part of the work performed in the RNA fingerprinting is a representative sample of a total of 90 studied. One disadvantages of the differential display that a large number of bands are generated which do not represent differentially expressed mRNA.

Potentially expressed clones may be made up of more than one cDNA species (Liang et al., 1995) and hence the resulting clones produced are heterogenous in nature. This necessitates a method for screening large numbers of clones quickly to determine which are differentially expressed before they can be characterized further.

Due to the restricted amounts of RNA that we obtained from the blood and tumour samples it was necessary to use another method to initially determine differential expression of those 18 bands and save the limited RNA for Northern blots for more promising bands. Upstream and down stream primers from the sequences of the differentially expressed bands

were used as potential markers to screen normal blood of healthy voulnteers. This is less than ideal. Nevertheless, the non-specificity of those 18 bands determined that they were not genuinely differentially expressed, and therefore were not characterized further.

For the differential display to be fully evaluated, the whole range of arbitrary and anchored primer combination must be used in a set up which has the necessary expertise and enough cDNA stocks for characterization of the differentially expressed bands.

Future work in RNA fingerprinting should encompass a more comprehensive study to utilise the potentials of the technique, with an extension of the approach to compare malignant breast tumours with benign tumours, normal breast epithelial cells in addition to normal blood. When a suitable epithelial marker has been identified, it should be tested in similar experiments described in this thesis, prior to use in clinical practice.

5. References

Abe M, Kufe D. Characterization of cis-acting elements regulating transcription of the human DF3 breast carcinoma-associated antigen (MUC1) gene. *Proc Natl Acad Sci USA* 1993; **90:** 282-286.

Acenero MJ, Gallego MG, Ballesteros PA, Gonzalez JF. Vascular density as a prognostic indicator for invasive ductal breast carcinoma. *Virchows Arch* 1998; **432:** 113-117.

Alkan S, Ross CW, Siddiqui J, Sheldon S, Hansos CA. Polymerase chain reaction (PCR) detection of myl/rar-alpha in acute promyelocytic leukemia (AML) using nested primers. *Lab Invest* 1993; **68:** 1.

Allred DC, Clark GM, Elledge R, Fuqua SA, Brown RW, Chamness GC, Osborne CK, McGuire WL. Association of p53 protein expression with tumor cell proliferation rate and clinical outcome in node-negative breast cancer. *J Natl Cancer Inst* 1993; **85:** 200-6.

Armstrong B, Doll R. Environmental factors and cancer incidence and mortality in different countries, with special reference to dietary practices. *Int J Cancer* 1975; **15:** 617-631.

Beatson GT. On the treatment of inoperable cases of carcinoma in mamma; suggestion for new method of treatment with illustrative cases. *Lancet* 1896; **2:** 104-107.

Black RJ, Sharp L, Kendrick SW. Trends in cancer survival in Scotland 1968-1990. Edinburgh: Information and statistics division, National health service in Scotland; 1993.

Bloom HJ & Richardson WW. Histological grading and prognosis in breast cancer. *Br J Cancer* 1957; **11:** 359-377.

Boring CC, Squires TS, Tong T, Cancer statistics, 1993. *CA Cancer J Clin* 1993; **43:** 7-26.

Brinton LA, Schairer C. Postmenopausal hormone-replacement therapy: time for a reappraisal? *N Engl J Med* 1997; **336:** 1821-1822.

Brown DC, Purushotham AD, Birnie GD, George WD. Detection of intraoperative tumour-cell dissemination in patients with breast-cancer by use of reverse transcription and polymerase chain- reaction. *Surgery* 1995; **117:** 96-101.

Buell P. Changing incidence of breast cancer in Japanese-American women. *J Natl Cancer Inst* 1973; **51:** 1479-83.

Burchill SA, Bradbury FM, Smith B, Lewis IJ, Selby P. Neuroblastoma cell detection by reverse transcriptase-polymerase chain reaction (RT-PCR) for tyrosine hydroxylase mRNA. *Int J Cancer* 1994; **57:** 671-675.

Burchill SA, Bradbury FM, Pittman K, Southgate J, Smith B, Selby P. Detection of epithelial cancer cells in peripheral blood by reverse transcriptase-polymerase chain reaction. *Br J Cancer* 1995; **71:** 278-281.

Burnet R, Guichard Y, Barale E. Immunohistochemistry for light microscopy in safety evaluation of therapeutic agents: an overview. *Toxicology* 1997; **119:** 83-93.

Chelly J, Concordet JP, Kaplan JC, Kahn. A Illegitimate transcription: transcription of any gene in any cell type. *Proc Natl Acad Sci USA* 1989; **86:** 2617-21.

Choy A, McCulloch P. Induction of tumour cell shedding into effluent venous blood breast cancer surgery. *Br J Cancer* 1996; **73:** 79-82.

Claus EB, Risch NJ, Thompson WD. Age at onset as an indicator of familial risk of breast cancer. *Am J Epidemiol* 1990; **131:** 961-972.

Coburn MC, Bland KI. Surgery for early and minimally invasive breast cancer. *Curr Opin Oncol* 1995; **7:** 506-10.

Cooper D, Schermer A, Sun TT. Classification of human epithelia and their neoplasms using monoclonal antibodies to keratins: strategies, applications and limitations. *Lab Invest* 1985; **52:** 243-256.

Corry JF, Lonning PE. Systemic therapy in breast cancer: efficacy and cost utility. *Pharmaco Econom* 1994; **5:** 198-212.

Coustan-Smith E, Behm FG, Sanchez J, Boyett JM, Hancock ML, Raimondi SC, Rubnitz JE, Rivera GK, Sandlund JT, Pui CH, Campana D. Immunological detection of minimal residual disease in children with acute lymphoblastic leukaemia. *Lancet* 1998; **351:** 550-554.

Cross NCP, Feng L, Chase A, Bungey J, Hughes TP, Goldman JM. Competitive polymerase chain reaction to estimate the numbers of bcr-abl transcripts in chronic myeloid leukemia patients after bone marrow transplantation. *Blood* 1993, **82:** 1929-36.

Cuzick J, Stewart H, Rutqvist L, Houghton J, Edwards R, Redmond C, Peto R, Baum M, Fisher B, Host H. Cause-specific mortality in long-term survivors of breast cancer who participated in trials of radiotherapy. *J Clin Oncol* 1994; **12:** 447-453.

Datta YH, Adams PT, Drobyski WR, Ethier SP, Terry VH, Roth MS. Sensitive detection of occult breast cancer by the reverse transcriptase polymerase chain reaction. *J Clin Oncol* 1994; 12**:** 475-482.

Debus E, Moll R, Franke WW, Weber K, Osborn M. Immunohistochemical distinction of human carcinomas by cytokeratin typing with monoclonal antibodies. *Am J Clin Pathol* 1984; **114:** 121-129

de Graaf H, Maelandsmo GM, Ruud P, Forus A, Oyjord T, Fodstad O, Hovig E. Ectopic expression of target genes may represent an inherent limitation of RT-PCR assays used for micrometastasis detection: studies on the epithelial glycoprotein gene EGP-2. *Int J Cancer* 1997; **72:** 191-196.

Delatorre M, Heldin P, Bergh J. Expression of the CD44 glycoprotein (lymphocyte-homing receptor) in untreated human breast-cancer and its relationship to prognostic markers. *Anticancer Res* 1995; **15:** 2791-95.

Devine PL, Layton GT, Clark BA,Birrell GW, Ward BG, King PX, McKenzie IF. Production of MUC1 and MUC2 mucins by human tumour cell lines. *Biochem Biophys Res Comm* 1991; **178:** 593-599.

Devine PL, McKenzie IF. Mucins: structure, function, and associations with malignancy. *Bioessays* 1992; **14:** 619-25.

Dewar JA, Sarrazin D, Benhamou E, Petit JY, Benhamou S, Arriagada R, Fontaine F, Castaigne D, Contesso G. Management of the axilla in conservatively treated breast cancer: 592 patients treated at Institut Gustave-Roussy. *Int J Radiat Oncol Biol Phys* 1987; **13:** 475-481.

Diel IJ, Kaufmann M, Goerner R, Costa SD, Kaul S, Bastert G. Detection of tumour cells in bone marrow of patients with primary breast cancer: a prognostic factor for distant metastasis. *J Clin Oncol* 1992; **10:** 1534-39.

Dilworth D, McCarrey J. Single-step elimination of contaminating DNA prior to reverse transcriptase PCR. *PCR Methods and Applications* 1992; **1:** 279-282.

Dupont WD, Page DL. Risk factors for breast cancer in women with proliferative disease. *N Engl J Med* 1985; **312:** 146-151.

Early Breast Cancer Trialists' Collaborative Group. Effects of adjuvant tamoxifen and of cytotoxic therapy on mortality in early breast cancer: An overview of 61 randomized trials among 28,896 women. *N Engl J Med* 1988; **319:** 1681-92.

Early Breast Cancer Trialists' Collaborative Group. Systemic treatment of early breast cancer by hormonal, cytotoxic or immune therapy: *Lancet* 1992; **339:** 71-85.

Early Breast Cancer Trialists' Collaborative Group. Ovarian ablation in early breast cancer: overview of the randamised trials. *Lancet* 1996; **348:** 1189-96.

Elston CW, Gresham GA, Rao GS, Zebro T, Haybittle JL, Houghton J, Kearney G. The cancer research campaign (King's /Cambridge) trial for early breast cancer: clinico-pathological aspects. *Br J Cancer* 1982; **45:** 655-669.

Elston CW. Grading of invasive breast cancer. In: Diagnostic histopathology of the breast. Page DL & Anderson TJ (eds), Edinburgh: Churchill Livingstone 1987.

Embury SH, Scharf SJ, Saiki RK, Gholson MA, Golbus M, Arnheim N, Erlich HA. Rapid prenatal diagnosis of sickle cell anemia by a new method of DNA analysis. *N Engl J Med* 1987; **316:** 656-661.

Eschwege P, Dumas F, Blanchet P, Le Maire V, Benoit G, Jardin A. Haematogenous dissemination of prostatic epithelial cells during radical prostatectomy. *Lancet* 1995; **346:** 1528-30.

Fabrega S, Laporte JP, Giarratana MC, Douay L, Fouillard L, Da WM, Perrot C, Barbu V, Gorin NC. Polymerase chain reaction: a method for monitoring tumor cell purge by long-term culture in BCR/ABL positive acute lymphoblastic leukemia. *Bone Marrow Transplant* 1993; **11:** 169-73.

Feuer EJ, Wun LM, Boring CC, Flanders WD, Timmel MJ, Tong T. The lifetime risk of developing breast cancer. *J Natl Cancer Inst* 1993; **85:** 892-897.

Fisher B, Slack NH, Bross ID. Cancer of the breast: size of neoplasm and prognosis. *Cancer* 1969; **24:** 1071-80.

Fisher B, Anderson S, Redmond CK, Wolmark N, Wickerham DL, Cronin WM. Reanalysis and results after 12 years of follow-up in a randomized clinical trial comparing total mastectomy with lumpectomy with or without irradiation in the treatment of breast cancer. *N Engl J Med* 1995; **333:**1456-61.

Fisher B, Dignam J, Wolmark N, DeCillis A, Emir B, Wickerham DL, Bryant J, Dimitrov NV, Abramson N, Atkins JN, Shibata H, Deschenes L, Margolese RG. Tamoxifen and chemotherapy for lymph node-negative, estrogen receptor-positive breast cancer. *J Natl Cancer Inst* 1997; **89:** 1673-82.

Forrest AP, Everington D, McDonald CC, Steele RJ, Chetty U, Stewart HJ. The Edinburgh randomized trial of axillary sampling or clearance after mastectomy. *Br J Surg* 1995; **82:**1504-8.

Fox SB, Fawcett J, Jackson DG, Collins I, Gatter KC, Harris AL, Gearing A, Simmons DL. Normal human tissues, in addition to some tumours, express multiple different CD44 isoforms. *Cancer Res* 1994; **54:** 4539-46.

Frank JA, Ling A, Patronans NJ, Carrasquillo JA, Horvath K, Hickey AM, Dwyer AJ. Detection of malignant bone tumors: MR imaging vs scintigraphy. *Am J Roentgenol* 1990; **155:** 1043-48.

Friedrichs K, Gluba S, Eidtmann H, Jonat W. Overexpression of p53 and prognosis in breast cancer. *Cancer* 1993; **72:** 3641-7.

Gasparini G, Pozza F, Harris AL. Evaluating the potential usefulness of new prognostic and predictive indicators in node-negative breast cancer patients. *J Natl Cancer Inst* 1993; **85:** 1206-19.

Gasparini G, Weidner N, Bevilacqua P, Maluta S, Dalla Palma P, Caffo O, Barbareschi M, Boracchi E, Pozza F. Tumor microvessel density, p53 expression, tumor size and peritumoral lymphatic vessel invasion are relevant prognostic markers in node-negative breast carcinoma. *J Clin Oncol* 1994; **12:** 456-466.

Gendler SJ, Spicer AP, Lalani EN, Duhig T, Peat N, Burchell J, Pemberton L, Boshell M, Taylor-Papadimitriou J. Structure and biology of a carcinoma-associated mucin, MUC1. *Am Rev Respir Dis* 1991; **144:** 42-7.

Gerhard M, Juhl H, Kalthoff H, Schreiber HW, Wagener C, Neumaier M. Specific detection of carcinoembryonic antigen-expressing tumor cells in bone marrow aspirates by polymerase chain reaction. *J Clin Oncol* 1994; **12:** 725-9.

Gilchrist KW, Gray R, van Driel-Kulker AM, Mesker WE, Ploem-Zaaijer JJ, Ploem JS, Taylor SG, Tormey DC. High DNA content and prognosis in lymph node positive breast cancer. A case control study by the University of Leiden and ECOG. (Eastern Cooperative Oncology Group). *Breast Cancer Res Treat* 1993 ; **28:**1-8.

Giuliano AE, Kirgan DM, Guenther JM, Morton DL. Lymphatic mapping and sentinel lymphadenectomy for breast cancer. *Ann Surg* 1994; **220:** 398-401.

Glass A, Hoover RN. Changing incidence of breast cancer. *J Natl Cancer Inst* 1988; **80:** 1076-77.

Glaves D, Huben RP, Weiss L. Haematogenous dissemination of cells from human renal adenocarcinoma. *Br J Cancer* 1988; **57:** 32-35.

Grodstein F, Stampfer MJ, Colditz GA, Willett WC, Manson JE, Joffe M, Rosner B, Fuchs C, Hankinson SE, Hunter DJ, Hennekens CH, Speizer FE. Postmenopausal hormone therapy and mortality. *N Engl J Med* 1997; **336:** 1769-77.

Gross HJ, Verwer B, Houck D, Hoffman RA, Rectenwald D. Model study detecting breast cancer cells in peripheral blood mononuclear cells at frequencies as low as 10^{-7}. *Proc Natl Acad Sci USA* 1995; **92:** 537-451.

Gunthert U, Hofmann M, Rudy W, Reber S, Zoller M, HauBmann I. A new variant of glycoprotein CD44 confers metastatic potential to rat carcinoma cells. *Cell* 1991; **65:** 13-24.

Hafstrom L, Domellof L, Rudenstam CM. Adjuvant chemotherapy with 5-fluorouracil, vincristine and CCNU for patients with Dukes' C colorectal cancer. *Br J Surg* 1990; **77:** 1345-48.

Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B, King MC. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* 1990; **250:** 1684-89.

Hames BD, Higgins SJ. Gene probe 1: a practical approach. Oxford University Press Inc., New York 1995.

Hara S, Kondo M, Tagawa Y, Itoyanagi N, Kishimoto K, Matuo S, Akamine S, Takahashi T, Oka T, Ayabe H. Long-term survival in terms of DNA-RNA contents in breast cancer. *Gan To Kagaku Ryoho* 1998; **25:** 422-5.

Hardingham JE, Kotasek D, Farmer B, Butler R, Mi JX, Sage RE, Dobrovic A. Immunobead-PCR: A technique for the detection of circulating tumour cells using immunomagnetic beads and the polymerase chain reaction. *Cancer Res* 1993; **53:** 3455-58.

Hardingham JE, Kotasek D, Sage RE, Eaton MC, Pascoe VH, Dobrovic A. Detection of circulating tumour cells in colorectal cancer by immunobead-PCR is a sensitive prognostic marker for relapse of disease. *Mol Med* 1995; **1:** 789-794.

Harris JR, Lippman ME, Veronesi U, Willett W. Breast cancer, part 3. *N Engl J Med* 1992; **327:** 473-480.

Harris V, Sandridge AL, Black RJ, Brewster DH, Gould A. Cancer registration statistics Scotland, 1968-1995. Edinburgh: ISD Scotland publication, 1998.

Haybittle JL, Blamey RW, Elston CW, Johnson J, Doyle PJ, Campbell FC, Nicholson RI, Griffiths K. Aprognostic index in primary breast cancer. *Br J Cancer* 1982; **45:** 361-366.

Haynes BF, Liao HX, Patton KL. The transmembrane hyaluronate receptor (CD44): multiple functions, multiple forms. *Cancer Cells* 1991; **3:** 347-50.

Hedley DW, Rugg CA, Gelber RD. Association of DNA index and s-phase fraction with prognosis of nodes positive early breast cancer. *Cancer Res* 1987; **47:** 4729-35.

Helzlsouer KJ. Epidemiology, prevention and early detection of breast cancer. *Curr Opin Oncol* 1995; **7:** 489-494.

Henderson BE, Bernstein L. The role of endogenous and exogenous hormones in the etiology of breast cancer. In: Breast Diseases. Harris JR, Hellman S, Henderson IC, Kinne DW (eds) , New York: Lippincott, 1991.

Heyderman E, McCartney JC. Epithelial membrane antigen and lymphoid cells. *Lancet* 1985; **1:** 109.

Hilkens J, Biochemistry and function of mucins in malignant disease. *Cancer Rev* 1988; **11-12:** 25-54.

Ho SB, Niehans GA, Lyftogt C, Yan PS, Cherwitz DL, Gum ET. Heterogeneity of mucin gene expression in normal and neoplastic tissues. *Cancer Res* 1993; **53:** 641-51.

Hofmann M, Rudy W, Zoller M, Tolg C, Ponta H, Herrlich P, Gunther U. CD44 splice variant confer metastatic behaviour in rats: homologous sequences are expressed in human tumour cell lines. *Cancer Res* 1991; **51:** 5292-97.

Hollingsworth MA, Strawhecker JM, Caffrey TC, Mack DR. Expression of MUC1, MUC2, MUC3 and MUC4 mucin mRNAs in human pancreatic and intestinal tumor cell lines. *Int J Cancer* 1994; **57:** 198-203.

Hoon D, Doi F, Giuliano A, Schmid P, Conrad A.The detection of breast carcinoma micrometastases in axillary lymph nodes by means of reverse transcriptase-polymerase chain reaction. *Cancer* 1995; **76:** 533-535.

Hoskins K, Weber BL. The biology of breast cancer. *Curr Opin Oncol* 1994; **6:** 554-559.

Huggins C, Bergenstal DM. Inhibition of human mammary and prostate cancer by adrenalectomy. *Cancer Res* 1952; **1:** 134-141.

Hurwitz M, Sawicki M, Samara G, Passaro E Jr. Diagnostic and prognostic molecular markers in cancer. *Am J Surg* 1992;**164:**299-306.

Information and Statistics Division, Scottish health statistics 1997. Edinburgh: ISD, National health service in Scotland; 1998.

Jensen OM, Esteve J, Moller H, Renard H. Cancer in the European community and its Member States. *Eur J Cancer.* 1990; **26:** 1167-256.

Johnson PWM, Burchill SA, Selby PJ. The molecular detection of circulating tumour cells. *Br J Cancer* 1995; **72:** 268-276.

Kampert JB, Whittemore AS, Paffenbarger RS. Combined effect of childbearing, menstural events, and body size on age-specific breast cancer risk. *Am J Epidemiol* 1988; **128:** 962-979.

Kaufmann M, Heider KH, Sinn HP, von Minckwitz G, Ponta H, Herrlich P. CD44 variant exon epitopes in primary breast cancer and length of survival. *Lancet* 1995; **345:** 615-9.

Kim YS. Mucin glycoproteins in colonic neoplasia. *Keio J Med* 1998; **47:** 10-8.

Krismann M, Todt B, Schroder J, Gareis D, Muller KM, Seeber S, Schutte J. Low specificity of cytokeratin 19 reverse transcriptase-polymerase chain reaction analyses for detection of Haematogenous lung cancer dissemination. *J Clin Oncol* 1995; **13:** 2769-75.

Kruger W, Krzizanowski C, Ilolweg M, Stockschlader M, Kroger N, Jung R, Mross K, Jonat W, Zander AR. Reverse transcriptase/polymerase chain reaction detection of cytokeratin-19 mRNA in bone marrow and blood of breast cancer patients. *J Cancer Res Clin Oncol* 1996; **122:** 679-686.

Lancaster CA, Peat N, Duhig T, Wilson D, Taylor-Papadimitriou J, Gendler SJ. Structure and expression of the human polymorphic epithelial mucin gene: an expressed VNTR unit. *Biochem Biophys Res Commun* 1990; **173:** 1019-29.

Lane EB, Alexander CM. Use of keratin antibodies in tumor diagnosis. *Sem Cancer Biol* 1990; **1:** 165-179.

Leather AJ, Gallegos NC, Kocjan G, Savage F, Smales CS, Hu W, Boulos PB, Northover JM, Phillips RK. Detection and enumeration of circulating tumour cells in colorectal cancer. *Br J Surg* 1993; **80:** 777-780.

Levesque MA, Katsaros D, Yu H, Giai M, Genta F, Roagna R, Ponzone R, Massobrio M, Sismondi P, Diamandis EP. Immunofluorometrically determined p53 accumulation as a prognostic indicator in Italian breast cancer patients. *Int J Cancer* 1998; **79:** 147-52.

Levi F. Cancer incidence in five continents, Vol VI. *Eur J Cancer* 1993; **29A:** 2315-19.

Liang P, Pardee AB. Recent advances in differential display. *Curr Opin Immunol* 1995; **7:** 274- 280.

Liljegren G, Holmberg L, Adami HO, Westman G, Graffman S, Bergh J. Sector resection with or without postoperative radiotherapy for stage I breast cancer:

five-year results of a randomized trial. Uppsala-Orebro Breast Cancer Study Group. *J Natl Cancer Inst* 1994; **86:** 717-22.

Luft R, Olivecrona H. Experience with hypophysectomy in man. *J Neurosurg* 1953; **10:** 310-316.

Luppi M, Morselli M, Bandieri E, Federico M, Marasca R, Barozzi P, Ferrari MG, Savarino M, Frassoldati A. Sensitive detection of circulating breast cancer cells by reverse-transcriptase polymerase chain reaction of maspin gene. *Ann Oncol* 1996; **7:** 619-24.

Lussier M, Filion M, Compton JG, Nadeau JH, Lapointe L, Royal A. The mouse keratin 19-encoding gene: sequence, structure and chromosomal assignment. *Gene* 1990; **95:** 203-13.

Lynn AG, Ries MS, Top 5 cancers for females and males in the US. *J Nat Cancer Inst* 1995; **87:** 867.

Mackay CR, Terpe HJ, Stauder R, Marston WL, Stark H, Gunthert U. Expression and modulation of CD44 variant isoforms in humans. *J Cell Biol* 1994; **124:** 71-82.

MacMahon B, Cole P, Lin TM, Lowe CR, Mirra AP, Ravnihar B, Salber EJ, Valaoras VG, Yuasa S. Age at first birth and breast cancer risk. *Bull WHO* 1970;

43: 209-221.

Maggiora P, Marchio S, Stella MC, Giai M, Belfiore A, De Bortoli M, Di Renzo MF, Costantino A, Sismondi P, Comoglio PM. Overexpression of the RON gene in human breast carcinoma. *Oncogene* 1998; **16:** 2927-33.

Matsumura Y, Tarin D. Significance of CD44 gene products for cancer diagnosis and disease evaluation. *Lancet* 1992; **340:** 1053- 58.

Mattano LA Jr, Moss TJ, Emerson SG. Sensitive detection of rare circulating neuroblastoma cells by the reverse transcriptase-polymerase chain reaction. *Cancer Res* 1992; **52:** 4701-5.

McClelland M, Welsh J. RNA fingerprinting by arbitrarily primed PCR. *PCR Methods Appl* 1994; **4:** S66-81.

McCulloch P, Choy A, Martin L. Association between tumour angiogenesis and tumour cell shedding into effluent venous blood during breast cancer surgery. *Lancet* 1995; **346:** 1334-35.

Mcpherson K, Steel CM, Dixon JM. Breast cancer- epidemiolgy, risk factors and genetics. *Br Med J* 1994; **309:** 1003-6.

Mehta RR, McDermott JH, Hieken TJ, Marler KC, Patel MK, Wild LD, Das Gupta TK. Plasma c-erbB-2 levels in breast cancer patients: prognostic

significance in predicting response to chemotherapy. *J Clin Oncol* 1998; **16:** 2409-16.

Menard S, Bufalino R, Rilke F, Cascinelli N, Veronesi U, Colnaghi MI. Prognosis based on primary breast carcinoma instead of pathological nodal status. *Br J Cancer* 1994; **70:** 709-712.

Miller BA, Fever EJ, Hankey BF, The increasing incidence of breast cancer since 1982: relevance of early detection. *Cancer Causes Control* 1991; **2:** 67-74.

Moertel CG, Fleming TR, Macdonald JS, Haller DG, Laurie JA, Goodman PJ, Ungerleider JS, Emerson WA, Tormey DC, Glick JH. Levamisole and fluorouracil for adjuvant therapy of resected colon carcinoma. *N Engl J Med* 1990; **322:** 352-8.

Molino A, Colombatti M, Bonetti F, Zardini M, Pasini F, Perini A, Pelosi G, Tridente G, Veneri D, Cetto GL. A comparative analysis of three different techniques for the detection of breast cancer cells in bone marrow. *Cancer* 1991; **67:** 1033-6.

Moll R, Franke WW, Schiller DL, Geiger B. The catalogue of human cytokeratins: Patterns of expression in normal epithelia, tumours and cultured cells. *Cell* 1982; **31:** 11-24.

Moll R. Cytokeratins as markers of differentiation. Expression profiles in epithelia and epithelial tumours. *Veroffentlichungen aus der pathologie* 1993; **142:** 1-197.

Mori M, Mimori K, Ueo H, Karimine N, Barnard GF, Sugimachi K, Akiyoshi T. Molecular detection of circulating solid carcinoma cells in the peripheral blood: the concept of early systemic disease. *Int J Cancer* 1996; **68:** 739-43.

Mullis, K.B., Faloona, F., Scharf, S.J., Saiaki, RK, Horn,GT, Erlich, HA. Supp. Quant. Biol. 51: 263-273. Cold Spring Harbor, New York 1986.

Mullis KB, Faloona F. Method enzymol. 155: 335. Academic Press, New York and London 1987.

Nagel RB, Intermediate filaments: a review of basic biology. *Am J Sur Pathol* 1988; **12:** 4-16.

Negrin RS, Pesando J. Detection of tumor cells in purged bone marrow and peripheral-blood mononuclear cells by polymerase chain reaction amplification of bcl-2 translocations. *J Clin Oncol* 1994; **12:** 1021-7.

Noguchi S, Hiratsuka M, Furukawa H, Aihara T, Kasugai T, Tamura S, Imaoka S, Koyama H, Iwanaga T. Detection of gastric cancer micrometastases in lymph nodes by amplification of keratin 19 mRNA with reverse transcriptase-polymerase chain reaction. *Jap J Cancer Res* 1996; **87:** 650-4.

Ogata S, Uehara H, Chen A, Itzkowitz SH. Mucin gene expression in colonic tissues and cell lines. *Cancer Res* 1992; **52:** 5971-8.

Overgaard M, Hansen PS, Overgaard J, Rose C, Andersson M, Bach F, Kjaer M, Gadeberg CC, Mouridsen HT, Jensen MB, Zedeler K. Postoperative radiotherapy in high-risk premenopausal women with breast cancer who receive adjuvant chemotherapy. Danish Breast Cancer Cooperative Group 82b Trial. *N Engl J Med* 1997; **337:** 949-55.

Osborn M, Weber K.Tumor diagnosis by intermediate filament typing: a novel tool for surgical pathology. *Lab Invest* 1983; **48:** 372-394.

Osborne MP, Wong GY, Asina S, Old LJ, Cote RJ, Rosen PP. Sensitivity of immunocytochemical detection of breast cancer cells in human bone marrow. *Cancer Res* 1991; **51:** 2706-9.

Pantel K, Izbicki J, Passlick B, Angstwurm M, Haussinger K, Thetter O, Riethmuller G. Frequency and prognostic significance of isolated tumour cells in bone marrow of patients with non-small-cell lung cancer without overt metastases. *Lancet* 1996; **347:** 649-652.

Papadimitriou JT, Berdichevsky F, D'Souza B, Burchell J. Human models of breast cancer. *Cancer Surveys* 1993; **16:** 59-69.

Powles T, Eeles R, Ashley S, Easton D, Chang J, Dowsett M, Tidy A, Viggers J, Davey J. Interim analysis of the incidence of breast cancer in the Royal Marsden Hospital tamoxifen randomised chemoprevention trial. *Lancet* 1998; **352:** 98-101.

Preece PE, Wood RA, Mackie CR, Cuschieri A. Tamoxifen as initial sole treatment of localised breast cancer in elderly women: a pilot study. *Br Med J* 1982; **248:** 869-70.

Prentice RL, Sheppard L. Dietary fat and cancer: consistency of the epidemiologic data and disease prevention that may follow from a practical reduction in fat consumption. *Cancer Causes Control* 1990; **1:** 81-97.

Pritchard KI. Is tamoxifen effective in prevention of breast cancer? *Lancet* 1998; **352:** 80-81.

Ragaz J, Jackson SM, Le N, Plenderleith IH, Spinelli JJ, Basco VE, Wilson KS, Knowling MA, Coppin CM, Paradis M, Coldman AJ, Olivotto IA. Adjuvant radiotherapy and chemotherapy in node-positive premenopausal women with breast cancer. *N Engl J Med* 1997; **337:** 956-962.

Rayter Z. Steroid receptors in breast cancer. *Br J Surg* 1991; **78:** 528-35.

Riethmuller G, Schneidergadicke E, Schlimok G, Schmiegel W, Raab R, Hoffken K, Gruber R, Pichlmaier H, Hirche H, Pichlmayr R. Randomized trial of

monocolonal antibody for adjuvant therapy of resected Dukes C colorectal carcinoma. *Lancet* 1994; **343:** 1177-83.

Rochefort H. Cathepsin D in breast cancer. *Breast Cancer Res Treat* 1990; **16:** 3-13.

Rosen PP, Groshen S, Kinne DW, Norton L. Factors influencing prognosis in node-negative breast carcinoma: Analysis of 767 T1N0M0/T2N0M0 patients with long-term follow-up. *J Clin Oncol* 1993; **11:** 2090-100.

Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA, Arnheim N. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 1985; **230:** 1350-4.

Sambrook J, Fritsch EF, Maniatis T. Molecular Cloning a laboratory manual. Cold Spring Harbour: Cold Spring Harbour Laboratory Press, 1989.

Schoenfeld A, Luqmani Y, Smith D, O'Reily S, Shousha S, Sinnett HD, Coombes C. Detection of breast cancer micrometastases in axillary lymph nodes by using polymerase chain reaction. *Cancer Res* 1994; **54:** 2986-90.

Schnurch HG. Relative indicators for disease outcome in breast carcinoma. *Schweiz Rundsch Med Prax* 1998; **87:** 504-510.

Screaton GR, Bell MV, Jackson DG, Cornelis FB, Gerth U, Bell JI. Genomic structure of DNA coding the lymphocyte homing receptor CD44 reveals at least 12 alternatively spliced exons. *Proc Natl Acad Sci USA* 1992; **89:** 12160-64.

Seregni E, Botti C, Lombardo C, Cantoni A, Bogni A, Cataldo I, Bombardieri E. Pattern of mucin gene expression in normal and neoplastic lung tissues. *Anticancer Res* 1996; **16:** 2209-13.

Seregni E, Botti C, Massaron S, Lombardo C, Capobianco A, Bogni A, Bombardieri E. Structure, function and gene expression of epithelial mucins. *Tumori* 1997; **83:** 625-32.

Seshadri R, Firgaira FA, Horsfall DJ, McCaul K, Setlur V, Kitchen P. Clinical significance of HER-2/neu oncogene amplification in primary breast cancer. The South Australian Breast Cancer Study Group. *J Clin Oncol* 1993; **11:** 1936-42.

Silvestrini R, Daidone MG, Benini E, Faranda A, Tomasic G, Boracchi P, Salvadori B, Veronesi U. Validation of p53 accumulation as a predictor of distant metastasis at 10 years of follow-up in 1400 node-negative breast cancers. *Clin Cancer Res* 1996; **12:** 2007-13.

Shimizu M, Yamauchi K. Isolation and characterization of mucin-like glycoprotein in human milk fat globule. *J Biochem* 1982; **91:** 515- 519.

Smith B, Selby P, Southgate J, Pittman K, Bradley C, Blair G. Detection of melanoma cells in peripheral blood by means of reverse transcriptase and polymerase chain reaction. *Lancet* 1991; **338:** 1227-29.

Skolnick MH, Cannon-Albright LA. Genetic predisposition to breast cancer. *Cancer* 1992; **70:** 1747-54.

Solomon E, Ledbetter D. Report of the committee on the genetic constitution of chromosome 17. *Cytogenet Cell Genet* 1992; **58:** 686-701.

Stamenkovic I, Amiot M, Pseudo J, Seed B. A lymphocyte molecule implicated in lymph node homing is a member of the cartilage link protein family. *Cell* 1989; **56:** 1057-62.

Stern L, Palatsides M, de Kretser T, Ford M. Expression of the tumor-associated mucin MUC1 in an ovarian tumor cell line. *Int J Cancer* 1992; **50:** 783-90.

Steward CG, Goulden NJ, Katz F, Baines D, Martin PG, Langlands K, Potter MN, Chessells JM, Oakhill A. A polymerase chain reaction study of the stability of Ig heavy-chain and T-cell receptor delta-gene rearrangements between presentation of childhood B-lineage acute lymphoblastic leukemia. *Blood* 1994; **83:** 1355-62.

Sun TT, Eichner R, Schermer A, Cooper D, Nelson WG, Weiss RA. Classification, expression and possible mechanisms of evolution of mammalian

epithelial keratins: a unifying model. In *Cancer Cells*: The transformed phenotype. Levine AJ, Van de woude GF, Topp WC, Watson JD (ed) pp. 169-176. Cold Spring Harbour Laboratory Press: Cold Spring Harbour, New York 1984.

Suzuki S, Tsuchiya A, Hara K, Abe R. Nuclear DNA content, S-phase fraction analysis and expression of proliferating cell nuclear antigen (PCNA) of normal and hyperplastic/neoplastic parathyroid glands. *Fukushima J Med Sci* 1997; **43:** 41-50.

Taylor CR, Cote RJ. Immunohistochemical markers of prognostic value in surgical pathology. *Histol Histopathol* 1997; **12:** 1039-55.

Todd JH, Dowle C, Williams MR. Confirmation of a prognostic index in primary breast cancer. *Br J Cancer* 1986; **56:** 489-492.

Traweek ST, Liu J, Battifora H. Keratin gene expression in non-epithelial tissues. Detection with polymerase chain reaction. *Am J Pathol* 1993; **142:** 1111-18.

Trichopoulos D, MacMahon B, Cole P. Menopause and breast cancer risk. *J Natl Cancer Inst* 1972; **48:** 605-613.

Turnbull RB, Kyle K, Watson FR, Spratt J. Cancer of the colon: the influence of the non-touch isolation technique on survival rates. *Ann Surg* 1967; **166:** 420-425.

Tyzzer EE. Factors in the production and growth of tumour metastases. *J Med Res* 1913; **28:** 309-331.

van der Linden JC, Baak JP, Lindeman J, Hermans J, Meyer CJ. Prospective evaluation of prognostic value of morphometry in patients with primary breast cancer. *J Clin Pathol* 1987; **40:** 302-306.

Veronesi U, Banfi A, Salvadori B, Luini A, Saccozzi R, Zucali R, Marubini E, Del Vecchio M, Boracchi P, Marchini S. Breast conservation is the treatment of choice in small breast cancer: long-term results of a randomized trial. *Eur J Cancer* 1990; **26:** 668-70.

Veronesi U, Luini A, Del Vecchio M, Greco M, Galimberti V, Merson M, Rilke F, Sacchini V, Saccozzi R, Savio T. Radiotherapy after breast-preserving surgery in women with localized cancer of the breast. *N Engl J Med* 1993; **328:** 1587-91.

Veronesi U, Maisonneuve P, Costa A, Sacchini V, Maltoni C, Robertson C, Rotmensz N, Boyle P. Prevention of breast cancer with tamoxifen: preliminary findings from the Italian randomised trial among hysterectomised women. *Lancet* 1998; **352:** 93-97.

Weidner N, Folkman J, Pozza F, Bevilacqua P, Allred EN, Moore DH, Meli S, Gasparini G. Tumor angiogenesis: a new significant and independent prognostic indicator in early-stage breast carcinoma. *J Natl Cancer Inst* 1992; **84:** 1875-87.

White E. Projected changes in breast cancer incidence due to the trend towards delayed childbearing. *Am J Public Health* 1987; **77:** 495-497.

Wile AG, DiSaia PJ. Hormones and breast cancer. *Am J Surg* 1989; **157:** 438-42.

Winstanley JHR, Leinster SJ, Cooke TG, Westley BR, Platt-Higgins AM, Rudland PS. Prognostic significance of Cathepsin-D in patients with breast cancer. *Br J Cancer.* 1993; **67:** 767- 772.

Xu CF, Solomon E. Mutations of the BRCA1 gene in human cancer. *Semin Cancer Biol* 1996; **7:** 33-40.

Yamada M, Wassermann R, Lange B, Reichard BA, Womer RB, Rovera G. Minimal residual disease in childhood B-lineage lymphoblastic leukemia: persistance of leukemic cells during the first 18 months of treatment. *N Engl J Med* 1990; **323:** 448-455.

Yonezawa S, Byrd JC, Dahiya R, Ho JJ, Gum JR, Griffiths B, Swallow DM, Kim YS. Differential mucin gene expression in human pancreatic and colon cancer cells. *Biochem J* 1991; **276:** 599-605.

Zotter S, Hagman C, Lossnitzer A, Mooi WJ, Hilgers J, Tissue and tumour distribution of human polymorphic epithelial mucin. *Cancer Rev* 1988; **11:** 55-101.

APPENDIX 1: MATERIALS

1 Cell lines, media and tissue culture supplies

2 Chemicals

3 Enzymes

4 KITS

5 Markers

6 Membranes, paper, TLC plates and X-ray film

7 Water

8 Blood

9 Solutions

MATERIALS

1 Cell lines, media and tissue culture supplies

MCF7 breast carcinoma cell lines were obtained from laboratory stocks.

Supplier: *Beatson Institute Central Service*

- sterile water
- sterile PBS
- sterile glassware and glass pipettes

Supplier: *Costar, 1 Alewife Ctr, Cambridge, MA 02140, USA*

- sterile disposable cell scrapers

Supplier: Fisons Europe Life Technologies Ltd., Paisley, U.K.

- dimethylsulphoxide (DMSO)

Supplier: *Gibco Europe Life Technologies Ltd., Paisley, U.K.*

- Foetal calf serum (FCS)
- 200mM glutamine
- 100mM sodium pyruvate
- 7.5% (w/v) sodium bicarbonate
- 2.5% (w/v) trypsin

Supplier: *A/S Nunc, Roskilde, Denmark*

-Tissue culture flasks
- Nunc tubes

Supplier: *Worthington, New Jersey, USA*

-Trypsin

2 Chemicals

Supplier: *British Drud House Chemicals Ltd.,Poole, Dorset,U.K.*

Unless otherwise stated, all chemicals used were obtained from BDH and were generally AnalaR Grade or better.

Supplier: *Amersham International plc, Amersham, Bucks., U.K.*

- [γ-^{32}P] ATP ~ 5000Ci/mmol
- [α ^{35}S] dATP ~ 600Ci/mmol

Supplier: *Bethesda Research Laboratory, Life Technologies, Inc., USA*

- agarose, ultra pure electrophoresis grade
- LMP agarose, ultra pure electrophoresis grade

Supplier: *Cinna/Biotecx Laboratories Inc., Houston, Texas, USA*

- RNAzolB

Supplier: *Difco laboratories, Detroit, Michigan, USA*

- bactro-tryptone and agar

Supplier: *Fison Scientific Equipment, Loughborough, U.K.*

- acetic acid
- chlorophorm
- glycerol

Supplier: *James Burrough Ltd., Witham, Essex, U.K.*

- ethanol

Supplier: *Koch-Light Ltd., Haverhill. Suffolk, U.K.*

- isoamyl alcohol

Supplier: *National Diagnostics, Manville, New Jersey, USA*

- ecoscint A

Supplier: *Rathburn Chemicals Ltd., Wakeburn, U.K.*

- phenol

Supplier: *Sigma Chemical Co., Ltd., Poole, Dorset, U.K.*

- Ampicillin
- bromophenol blue
- bovine serum albumin (BSA) (fraction V)
- ethidium bromide
- methylene blue
- salmon sperm DNA
- TMED

Supplier: *Unipath Ltd., Basinstoke, Hants., U.K.*

- PBS (Dulbecco A, tablet form)

3 Enzymes

Supplier: *Nothumbria Biologicals Ltd., Cramlington, Northumberland, U.K.*

- RNase A
- T4 polynucleotide kinase

Supplier: *Promega, Southhampton, Hampshire, U.K.*

- Taq DNA Polymerase + reaction buffer

Supplier: *Sigma Chemical Co., Ltd., Poole, Dorset, U.K.*

- lysozyme
- Phenol : Chloroform

4 Kits

Supplier: *Clontech Laboratories, Inc., CA94303-4607, USA*

- Delta ^{TM} RNA Fingerprinting Kit

Supplier: *Hybaid Limited, Middlesex, UK*

- Hybaid Recovery ™ Plasmid Mini Prep Kit

Supplier: *Invitrogen Corporation, San Diego, USA*

- TA Cloning Kit
- INV α F' One Shot™ Kit

Supplier: *Perkin Elmer Cetus, Norwalk, CT 06859, USA*

- Gene Amp PCR reagent kit

Supplier: *United States Biochemicals, Cleveland, Ohio, USA*

- Sequenase version 2.0 kit

5 Markers

Supplier: *Bethesda Laborotaries Research, Gibco Ltd., Paisley, U.K.*

- ϕX174 RF DNA / Hind III fragment
- 100bp DNA ladder

6 Membranes, paper, TLC plates and X-ray film

Supplier: *Amersham International plc, Amersham, Bucks., U.K.*

- Hybond N+

Supplier: Camlab Ltd., Cambridge, U.K.

- Polygram TLC plates

Supplier: *Eastman Kodak Co., Rochester, New York, USA*

- X-ray film (XAR)

Supplier: *Sartorius Ltd., Epsom, Surrey, U.K.*

- nitrocellulose membranes

7 Water

Distilled water for buffers and all general solutions were purified through a Millipore MilliRO 15 system. Water for procedures which required further purification was obtained from a Millipore MilliQ system at 18 MΩ cm.

1.8 Blood

Normal blood, normal and malignant breast tissue were used to provide material for rt-PCR and RNA fingerprinting and were obtained by informed consent from volunteers and patients attending the Department of Surgery at the Western Infirmary-Glasgow.

9 Solutions

DEPC Treated Water (0.1%)

999 ml double distilled water

1ml DEPC

Autoclave solution

Denhardt's Solution (100x)

2% (w/v) BSA fraction V

2% (w/v) Ficol™

2% (w/v) PVP

made up with DEPC treated water

DNA dye Mixture

30% v/v Glycerol

0.25% w/v Bromophenol blue

0.25% w/v Xylene cyanol

0.5M EDTA (pH 8)

186.1 g Na ethylenediamine tetraacetate.$2H_2O$

800 ml Water

Adjust pH to 8 with NaOH

3M NaAc (pH 7)

408.1 g $3NaAc.3H_2O$

800ml Water

Phosphate Buffered Saline (PBS)

0.8% NaCL

0.115% $Na_2 HPO_4$

0.02% KCL

0.02% KH_2PO_4

Prehybridisation buffer (100ml)

48ml Water

20ml 50x Denhardt's solution

30ml 20x SSC

02ml sDNA (5mg/ml stock)

RPMI

800ml sterile distilled water

26.6ml 7.5% $NaHCO_3$

10ml 100mM Na Pyruvate

10ml 200mM L-Glutamine

1ml 1M NaOH

100ml Foetal Calf Serum

100ml RPMI 1640 (10x)

R.B.C. Lysis Buffer (10x)

8.3% NH_4 CL

0.37g/l Na_2 EDTA

10.0g/l $KHCO_3$

10%SDS (pH 7.2)

100 g SDS

900ml Water

SSC (20x) pH 7

175.3 g NaCL

88.2 g Na citrate

800ml H_2O

Sequencing loading solution

950 µl deionized Formamide

50 µl TBE buffer (20x)

0.1 mg Bromophenol blue

0.1 mg Xylene cyanol

TBE(1x) pH8

89mM Tris borate

89mM boric acid

2.5mM EDTA

X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactosidase)
prepared in dimethyl formamide

APPENDIX 2: RNA Fingerprinting sequences

BAND 1.Seq (P5T7)

LENGTH: 452

1 TTGNTTACGG CCANTGGGCT TGTAATACTA CTCACTATAG GGCGAATTGG

51 GCCCTCTAGA TGCATGCTCG AGCGGCCGCC AGTGTGATGG ATATCTGCAG

101 AATTCGGCTT ATTAACCCTC ACTAAAGATC TGACTGAAAA ATAAAGTTTT

151 TGAACAATTT CATGAAGGTT GGGTTGTCGA GATCATGAGT TCTGTCTCTA

201 ATGACAACTC TGTGGGAAAT GGATAAACTC TGATAAACTC TATCGCACTG

251 CACACAAATG TGAGGCTGGA AGGCAGTACG TAATTTATGA GGCCTTTTGT

301 TGTAAAAATG AGAGTTGGCA GGATTCCGAG TAAATATCTT TTTTCCTAAG

351 GGTGGAAGAT AGGACTTGGA GAGATCAATA TGTTGGTTGA CAGCATTCGA

401 GAATGTTTGA TTTGAGCTTC AGCATCATAG GTGGGGAACA TGTNGCTTTA

451 TC

Forward Primer : 111 - 136

BAND 2.Seq (P5T8)

LENGTH: 455

1 NCCTTAGNTG NNTAAGGCCA CTGGNCCCGT TATACTTTTC ACTATAGGGC

51 GAATTGGGCC CTCTAGATGC ATGCTCGAGC GGCCGCCAGT GTGATGGATA

101 TCTGCAGAAT TCGGCTTATT AACCCTCACT AAAGATCTGA CTGCAGAAAA

151 CTGGTCAATA ATTTATCCTA GGGTAGAAAA AGCTAAGGAA CTANAAAAAG

201 AAAATATAAC ATNTACAGAT GTGTNGGTCA CTANTCTNGT CCACTTCATT

251 GTGGATNCTG ANGCACCTCT GAGCCCGTGT AAAGGCCACG GGATCAGTGA

301 TGAAGANAAT CNATTTCGTC TCAGCTGCCA CGTGCCCTGC AGAGGTGGCC

351 AGCATTAGAA GGTAACGCAT ATGGTATACA AATCCTCTCA TCTCCTTCNG

401 TCAGATCTTT AGTGAGGGTT GGTAAGCCGA ACTCCAGCCA CTGGCGGGCG

451 NTACC

Forward Primer : 118 - 143

BAND 3.Seq (P5T8)

LENGTH: 449

1 NGNNANNNCT TAAGNNCCGT GACACTATTC ACTATAGGGC GAATTGGGCC

51 CTCTAGATGC ATGCTCGAGC GGCCGCCAGT GTGATGGATA TCTGCAGAAT

101 TCGGCTTATT AACCCTCACT AAAGATCTGA CTGCCACTGT TGGCTTTATT

151 GCACTAGCCA GTATCACAGA GAGTTCCACT AACAGCAGGG GGATGAAACT

201 GTTGATGCTG TTATCAACAG AAATTGTCAA TTACCTGACA GCAGACACTA

251 GGTGCTACTT GTTAGAGGCA AAAAAAAAGA TATCACTCAG CATAATAAGC

301 CGAATTCCAG CACACTGGCG GGCCGTTACT AGTGGATCCG AGCTCGGTAC

351 CAAGCTTGGC GTAATCATGG TCATAGCTGT TTCCTGTGTG AAATTGTTAT

401 CCGCTCACAA TTCCACACAA CATACGAGCC GGAAGCATAA AGTGTAAAG

Forward Primer : 108 - 133

BAND 4.Seq (P5T8)

LENGTH: 457

1 NCNNGTNTAG ANNCAGCCCT GACCCGTAAA ACTATTCACN ANAGGGCGAA

51 TCGGGGCCCT CTAAATGCAT GCTCGAGCGG CCGCCAGTGT GATGGATATC

101 TGCAGAATTC GGCTTATTAA CCCTCACTAA AGATCTGACT GGATGGGGTT

151 ACCGGGTTAT CGCTTTGCAG TATCCAGTTN ATCGGGACCA TCTCGAGNNN

201 NGTGATGGAT CCAGAAAACT TNTAGACCAT TTACAATTGG ATAAAGTTCA

251 TCTTTTTGGC GCTTCTTTGG GAGGCTTTTT GGCCCAGAAA TTTGCTGAAT

301 ACACTCACAA ATCTCCTAGA GTCCATTCCC TAATCCNTCT GCAATTCCCT

351 CAGTGACACC TCTATCTTCA ACCAAACTTG GAACTGCNAA CAGGATTAAA

401 CCTCCATTTC AGCTAATCAT GGGAACAATT TAAGTCCCCT CCCGATTATN

451 ACTGGTC

Forward Primer : 116 - 141

BAND 5.Seq (P5T8)

LENGTH: 449

1 CGNCTANCGT TAATGGCTTG TGACACGACT CACTATAGGG CGAATTGGGC

51 CCTCTAGATG CATGCTCGAG CGGCCGCCAG TGTGATGGAT ATCTGCAGAA

101 TTCGGCTTAT TAACCCTCAC TAAAGATCTG ACTGAAGGCA AGCATAATTT

151 AGAGGAAAAG GCAAAAAAAA AGATATCACT CAGCTAATGA GTCCGAATTC

201 CAGCACACTG GCGGCCGTTA CTAGTGGATC CGAGCTCGGT ACCAAGCTTG

251 GCGTAATCAT GGTCATAGCT GTTTCCTGTG TGAAATTGTT ATCCGCTCAC

301 AATTCCACAC AACATACGAG CCGGAAGCAT AAAGTGTAAA GCCTGGGGTG

351 CCTAATGAGT GAGCTAACTC ACATTAATTG CGTTGCGCTC ACTGCCCGCT

401 TTCCAGTCGG GAAACTGTCG TGCCAGCTGC ATTAATGAAT CGGCCAACG

Forward Primer : 109 - 134

BAND 6.Seq (P5T8)

LENGTH: 450

1 TCTTNGNCCT GCCGGCCAGT GGAATTGTAA TACGACTCAC TATAGGGCGA

51 ATTGGGCCCT CTAGATGCAT GCTCGAGCGG CCGCCAGTGT GATGGATATC

101 TGCAGAATTC GGCTTATTAA CCCTCACTAA AGATCTGACT GGAGCAGCAA

151 AAAAAAAGAT ATCACTCAGC ATAAAAGCCG AATTCCAGCA CACTGGCGGC

201 CGTTACTAGT GGATCCGAGC TCGGTACCAA GCTTGGCGTA ATCATGGTCA

251 TAGCTGTTTC CTGTGTGAAA TTGTTATCCG CTCACAATTC CACACAACAT

301 ACGAGCCGGA AGCATAAAGT GTAAAGCCTG GGGTGCCTAA TGAGTGAGCT

351 AACTCACATT AATTGCGTTG CGCTCACTGC CCGCTTTCCA GTCGGGAAAC

401 TGTCGTGCCA CCTGCATTAA TGAATCNGCC ACCCCCGGGA AAAGCGGTTG

Forward Primer : 116 - 141

BAND 7.Seq (P5T9)

LENGTH: 452

1 NCGTATNACG CCAGTGGGCT TGTAATACGA CTCACTATAG GGCGAATTGG

51 GCCCTCTAGA TGCATGCTCG AGCGGCCGCC AGTGTGATGG ATATCTGCAG

101 AATTCGGCTT ATTAACCCTC ACTAAAGATC TGACTGATGT GATCTTATTT

151 GGACAAAGAA GTGTTGTTGG AAATATCAGT CATGTGAATT AATTAACAAT

201 TTCTTCCTTC CTAGCGAGAG TAAGACCATG GAGTAGAGCA GATATATTAG

251 TCCATTCCAT GTGGAAATCC GGCCTGTAGT TGTATATTCG TATAGCTATA

301 TGTTGGTAGG ATGGACCTAC ATGGTTTAAT GAAAGTGGAA GGAAAATCTA

351 TTTTCAGTCA GATCTTTAGT GAGGGTTAAT AAGCCCGAAT TCCAGCACAC

401 TGGCGGGCCG TTACTAGTGG ATCCGAGCTC GGTACCAAGC TTGGCGTTAT

451 CA

Forward Primer : 111 - 136

BAND 8.Seq (P6T7)

LENGTH: 455 Mon,

1 NTNCNTNAGA NNAAGCCCTG NNCCGTAAAA CTATTCACNA NAGGGCGAAT

51 CGGGCCCTCT AAATGCATGC TCGAGCGGCC GCCAGTGTGA TGGATATCTG

101 CAGAATTCGG CTTATTAACC CTCACTAAAT GCTGGGTGTG GTGGCGCACA

151 TCTGTAGTCC CTGCTACTGG AGAGGCTGAG GTGGGAGGAC TGCNNGAGCC

201 TGAGGCTGCA GACAGCCATG ATCATGCCAC AGCTCTGGCA TGGGTGACAG

251 AAAGAGACCC NATCTCAAAA AAAAAAAAAA ACAACCAGGT GTGCCGGATG

301 CAGTGGTGCA TGCATGTAGT CCCAGCTACT TGGCTGAGGC AGGGACCATC

351 ACCCGAGCCT CAGGAATTCC AAGCCGCTTT AAGCTANGAT GCACCCATGC

401 CCCNGTGACT AAACCAANAC CCTTNCNCCN NAAAAAAANA ANNNCCCCCC

451 CCCTC

Forward Primer : 114 - 138

BAND 9.Seq (P5T3)

LENGTH: 451

1 GTCTTNAGCC TAGTGGCATG TTATACGACT CACTATAGGG CGAATAGGGC

51 CCTCTAGATG CATGCTCGAG CGGCCGCCAG TGTGATGGAT ATCTGCAGAA

101 CTCGGCTATT AACCCTCAC TAAATGCTGG GTGTNAAAAN CNACCGNACT

151 CGTGAATTAT GTNCTNNANT NTACAATAGC GGATNCTAGT ACTANTATAT

201 NTGANTTGAA TNCATTCCCT GAATCAGANG TACCAACGAG GTCAGAGTGT

251 TNGTGTNCAN TCGGGANGCT TCGTGTCTCN CTATACCNCA TTTANATCTG

301 ANNANTGTTA CTNATAGTGA TCAATGGGAA GGTNAGTTTA CCCCCCCGCT

351 ATACTTGCCG NAACNANGTC ACANGTCNCC CGANCTNACC NCAAGCACNG

401 ATGCTACCAC AACCCTTTAC GNGANNAACA TANCCTTCNC CATCCTCTGA

451 CG

Forward Primer : 107 - 133

BAND 10.Seq (P4T3)

LENGTH: 459

1 GTGCTGNCTA AAAGCCANTG GNCCNGTAAA AACTTTTTCA CTATAGGGCG

51 AATTGGGCCC TCTAGATGCA TGCTCGAGCG GCCGCCAGTG TGATGGATAT

101 CTGCAGAATT CGGCTTATTA ACCCTCACTA AATGCTGGTA GGGATGTAAA

151 TTAATTCAAC CACTATGGAA AGCTGTTTGG AGATTTCTCT AAGANCGGAA

201 AACAGACCTA CCAACTGACC CAGCAATCCT ATTGATGGGT ATATACCCAA

251 AAGAAAAGAA ATGATTCTAC CAAAAAGACA CATGCACTCT ATATTCATCA

301 CAGTGCTATT CACAATAGCA AAAACATGGA CTAAACCTAA GTGCCCATCC

351 ACAGTAGACT GGATGAAGAA AATATGGTAC ATACACTTCA TNGAATACTA

401 TGCANCCCNT AAAAAGAATT AANTCATGTC CCTTGCNNCC ACCCCAATNC

451 CNGGTGGAT

Forward Primer : 117 - 141

BAND 11.Seq (P4T3)

LENGTH: 444

1 TGTGGCNTTG ACGGCCAGTG AATTGTAATA CGACTCACTA TAGGGCGAAT

51 TGGGCCCTCT AGATGCATGC TCGAGCGGCC GCCAGTGTGA TGGATATCTG

101 CAGAATTCGG CTTATTAACC CTCACTAAAT GCTGGTAGGA AGGTNAAANG

151 GTGTAACNGC TGTGAATAAA AGTCTGTTGC TTCNNCTNAA AGCNAACNCN

201 CCAGNGCCAC TCCCGCGGAN NTNGCAGTAC AATTGAAAAG GTTCTCANAC

251 ACTTGTATGC CACTATTCAC TNCGGAAATA TNAGGCAAAN GGTGGACACA

301 ATCCACGTAT CTTCCAACAG ATGAATGGAN ANACAATATG TGGTCTATAA

351 ATACAACANT NCGTAATTCA GTAATAACAA GNAGTATGTG CTNCACTATG

401 ATGAACCTTG AAANCCTCTA NTACCCACCT ACCGAAATTT TGAN

Forward Primer : 113 - 138

BAND 12.Seq (P4T4)

LENGTH: 450

1 NTGCTTNACG GCTAGTGGNC CTGTAATACT ACTCACTATA GGGCGAATTG

51 GGCCCTCTAG ATGCATGCTC GAGCGGCCGC CAGTGTGATG GATATCTGCA

101 GAATTCGGCT TATTAACCCT CACTAAATGC TGGTAGGAGG TTATCCNATT

151 ANCGANGGTA CCACATATAG GTGCANGGTG CCTGTCATTC ACTGTNTTAT

201 CTGGTNTANA TCGGNCTAAN NCTGGGGGAA GCTATNCTCT TTCAGTGGAT

251 AATAAAATTG GTATCTCAAT CGTGAAACAT NTCANNGGTG TGTGAAGAAA

301 AANCANCCNA TCTGTGGGTG TTGANATCTN GACCNTACTG TGTANGTTAC

351 GTGCCTNTGT GGATGTGCAC TACCAGCATT TAGTGANGGT TAANAANCCG

401 AATTCCACCA CACTGGCGGG CCGTTACTNN TGGATCCNAG CTCGGTACCT

Forward Primer : 112 - 137

BAND 13.Seq (P4T4)

LENGTH: 446

1 TGNTGNNTTA AGGCCAGTGG GCCTGTAATA CGACTCACTA TAGGGCGAAT

51 TGGGCCCTCT AGATGCATGC TCGAGCGGCC GCCAGTGTGA TGGATATCTG

101 CAGAATTCGG CTTATTAACC CTCACTAAAT GCTGGTAGNN CCACCGACNG

151 NATCNTTCNC NAAAGACCTG CTTATCTACC ATNAAGATAN NTTCCTANCC

201 GGCTCTGANC CTCAATCTCN CTGTGTGCCT GCATCTGNTT CNGNNCTGAA

251 TATNCCTCTA CGGTGGANGG ANACATANGN NCTGGAGNTG TGCACNGATC

301 CTCTTCGACT GGGCCTGGNT GGNGGTGTNT AAAGGCNAGG TNATACTTTC

351 CCCCCCNTCN NGCATNTACT CCCACTTTCT CCTGAACTAN GGCCACCTAC

401 TNANGGTGAA GGTTCNGAAT CCGAATTCCC TGGCNGNNGN GGCTGN

Forward Primer : 113 - 138

BAND 14.Seq (P4T4)

LENGTH: 450

1 GTNNGNCTTA GNCTTAGTGG GCCTGTAATA CGACTCACTA TAGGGCGAAT

51 TGGGCCCTCT AGATGCATGC TCGAGCGGCC GCCAGTGTGA TGGATATCTG

101 CAGAATTCGG CTTATTAACC CTCACTAAAT GCTGGTAGNG TACATATGCT

151 GTCCCCTACC CCCCANGTGC NANAAGNAAG CAGATGCTAT NTATNGGGGT

201 TTNCTCACAN NGTNAGANGA CATTNGCCAG NGTCAATTGT NATTGACTGN

251 NGCTATCNAG TAANTCCAAN GGAAACAGGT NTTTATGCCA CATTAANNTT

301 NCNGAAACCA CCCATGCNNN GANNAGTTGC TACAGGTTAT GGGANAANAG

351 AGTGCTGACA CCACTCGGTA TANGTTTACC ATGGAACTGT AATNACACTT

401 TGTACTCTCA NAATCNNGTG CACATTAGAC TTCCTGTTTA TCTGTGATAN

Forward Primer : 114 - 138

BAND 15.Seq (P5T3)

LENGTH: 486

1 TGTGNTGNCT TAAGGCCAGT GGGCCTGTAA TACGACTCAC TATAGGGCGA

51 ATTGGGCCCT CTAGATGCAT GCTCGAGCGG CCGCCAGTGT GATGGATATC

101 TGCAGAATTC GGCTTATTAA CCCTCACTAA AGATCTGACT GANNTCNNAA

151 ATNANNCCTT AGGAGGTTGG ACNGCTCGGG CCCTNGTGTG AGNNAANNAN

201 CCCTTNAGNA TGCCTTNTGG ACNNTACGGT GTCACNNTCA GACCATNGAC

251 AGNGGGTNAN TCCCACATCT GTCNGTCTNA GAGGACTTGT CCNAACGCAC

301 TTCAATANTC GTGGTTCTGN ACACAAAAGA TTTTGNTCAT NGAGAAAGCA

351 GCTANGTTTA AGGTACCTAN CACATCGNAG GAAACGATGG CTNCNCACCC

401 CCCNCCNATG NCCAGGNATT ANCCCTGTAT NGACTGGANC NCTTTCAAGG

451 GCANGGANGA TCNTCCGCTG TNGGTGTCNN GTNCTT

Forward Primer : 116 - 141

BAND 16.Seq (P6T3)

LENGTH: 452

1 GNCTTNAGCC TAGTGGCNTG TAATACGACT CACTATAGGG CGAATTGGGC

51 CCTCTAGATG CATGCTCGAG CGGCCGCCAG TGTGATGGAT ATCTGCAGAA

101 TTCGGCTTAT TAACCCTCAC TAAATGCTGG GTGTNAAAAN CNACCGNACT

151 CGTGAATTAT GTNCTNNANT NTACAATAGC GGATNCTAGT ACTANTATAT

201 NTGANTTGAA TNCCTTCCCT GAATCAGANG TACCAACGAG GTCAGAGTGT

251 TNGTGTNCAN TCGGGANGCA TCGTGTGTCN CTATACCNCA TTTANATCTG

301 ANNANTGTTA CTNATNGTGA TCAATGGGAA GGTNAGTTTA CCCCCCCGCT

351 ANACTTGCCG NAACNANGTC ACANGTCNCC CGANCTNACC NCAAGCACNG

401 ATGCTACCAC AACCCTTTAC GNGANNAACN TANCCTTCNC CATCCTCTGT

451 CN

Forward Primer : 108 - 133

BAND 17.Seq (P6T3)

LENGTH: 471

1 TGNGNTGNNT AAGGCCAGTG GGCCTGTAAT ACNTCTCACT ATAGGGCGAA

51 TTGGGCCCTC TAGATGCATG CTCGAGCGGC CGCCAGTGTG ATGGATATCT

101 GCAGAATTCG GCTTATTAAC CCTCACTAAA TGCTGGGTGN CTATNAAATA

151 AAATACCTTA TAACATCCNT CTNAAAATCC GTAGATGGAG GAGACAAGAA

201 CGACTAATNA ATATCAAATN CGGGCCNGTC NTCTCNATGT GTCTNACTCA

251 NCANNGGCCA CTTGCTTCAT GCNAACNGGT CTGGGAGANG TGAANATCAC

301 AGATGGACTC CAGGTCTCCA CACTAAGGAA GGCTGTCGCC CTCTGTCTAC

351 NCATAAGATA TTCATCGGGT CTCCCTTNCA CATGGNCCAT GAATCATNGA

401 GCCTGTGGGT AATNCCCCCG TTCACNGCCT ANAACCNANT TTANNCNNAC

451 TTCTAANCTA NCTTCCTCAA T

Forward Primer : 115 - 139

BAND 18.Seq (P6T4)

LENGTH: 450

1 NCNGAGNNAT TGNCGTCGTA CAACTCTTCA CTATAGGGCG AATTGGGCCC

51 TCTAAATGCA TGCTCGAGCG GCCGCCAGTG TGATGGATAT CTGCAGAATT

101 CGGCTTATTA ACCCTCACTA AATGCTGGGT GNTATGTTTN TGGATGCCTC

151 TCACGGNNCG NAGTNNCATN AAANCAANTG TGTGGAACGN NCGCCGAACN

201 TNGTNGGTAA GGTTACCGAT NTNCNAGAAN GATNAANACT NGNNNNTTGN

251 CAGGANCACN NCACCNNNNA CNGACCAGGG CCTTATACAT NGGAGNTCTC

301 NNNTNTGAAA ACAGAGGGAC TGGCAATTGA CGTGATTGTG CTANATGCTA

351 AATTGGTNCG TTNNAAACTG GGGGGTGGTG ANTGCTTATN CCTTCCGAGC

401 ANANACGTGT TANATTTTTA ACTGGAAGAT CGCCGTCTCG ANTAATTGAN

Forward Primer : 107 – 131