FACE SHEET

Version Date Dec 12, 2009

Protocol Number(s)
- 09CN149

Protocol Title
A STUDY OF MOLECULAR AND GENETIC FACTORS FOR LIVER CANCER IN THE GREATER BALTIMORE AREA

Abbreviated title: Biomarkers for Liver Cancer

Principal Investigator and Affiliation
Xin Wei Wang, Ph.D.
Laboratory of Human Carcinogenesis,
Center of Cancer Research (CCR)
National Cancer Institute (NCI), Bethesda, MD
Phone number: 301-496-2099
Email: xw3u@mail.nih.gov

Lead Associate Investigator and Affiliation
Anuradha S. Budhu, Ph.D.
Laboratory of Human Carcinogenesis,
Center of Cancer Research (CCR)
National Cancer Institute (NCI), Bethesda, MD
Phone number: 301-496-5887
Email: budhua@mail.nih.gov

Associate Investigators and Affiliations
Nader Hanna, M.D.
Section of Surgical Oncology
University of Maryland School of Medicine, Baltimore, MD

Benjamin Philosophe, MD, Ph.D.
 Dept. of Transplant Surgery
University of Maryland School of Medicine, Baltimore, MD

Harris G. Yfantis, M.D.
Department of Pathology
Baltimore VA Medical Center
University of Maryland School of Medicine, Baltimore, MD
Sandeep Khurana, MBBS
Division of Gastroenterology
University of Maryland School of Medicine, Baltimore, MD

Sandeep Khurana, MBBS.
Division of Gastroenterology
University of Maryland, Baltimore, MD

Nilofer Azad, M.D.
Comprehensive Cancer Center
Johns Hopkins Medical Institute University, Baltimore, MD

Consulting Epidemiologists:
Christopher Loffredo, Ph.D.
Lombardi Comprehensive Cancer Center
Georgetown University, Washington, DC

Stefan Ambs, Ph.D., M.P.H.
Laboratory of Human Carcinogenesis,
Center of Cancer Research (CCR)
 National Cancer Institute (NCI), Bethesda, MD

Medical Advisory Investigator (MAI): Not applicable for this study

Accountable Investigator: Not applicable for this study

Adjunct Principal Investigator (API): Not applicable for this study

Research Contact
Donna Perlmutter
Field Supervisor
Phone number: 410-706-0720
Email: dperlmut@umaryland.edu

Study Sponsor: Not applicable for this study

IND/IDE number/name of investigational drugs and/or devices.
Not applicable for this study

For Multi-institutional protocols:
- *Overall Study PI:* Xin Wei Wang, Ph.D.
    Laboratory of Human Carcinogenesis,
    Center of Cancer Research (CCR)
    National Cancer Institute NCI), Bethesda, MD
    Phone number: 301-496-2099
    Email: xw3u@mail.nih.gov
- *Coordination Center:* NCI
- *Responsible DSMB or SMC:* Not applicable for this study

Multi-institutional Protocols

Our protocol is a multi-institutional protocol for which the NCI is the coordinating center. The University of Maryland School of Medicine Medical Center, the Veterans Affairs Medical Center, Baltimore and Johns Hopkins Medical Institute are participating institutions.

Contacts at the University of Maryland School of Medicine

NAME OF INSTITUTIONAL PI: Dean L. Mann, M.D.
INSTITUTION: University of Maryland School of Medicine, School of Medicine
COMPLETE ADDRESS: 22 South Greene Street, Baltimore, MD 21201
PHONE NUMBER: 410-328-5512
E-MAIL: dmann001@umaryland

NAME OF IRB CONTACT: Sue Buskirk
FWA NUMBER: 00007145
COMPLETE ADDRESS OF CONTACT: Human Research Protections Office
    UMB Biopark 1, 800 West Baltimore Street, Suite 100, Baltimore, MD 21201
PHONE NUMBER OF CONTACT: 410-706-5037
E-MAIL OF CONTACT: sbuskirk@som.umaryland.edu

NAME OF PHARMACY CONTACT: Not applicable for this study
COMPLETE ADDRESS OF CONTACT:
PHONE NUMBER OF CONTACT:
E-MAIL OF CONTACT:

Contacts at the Veterans Affairs Medical Center, Baltimore

NAME OF INSTITUTIONAL PI: Dong Lee, M.D.
INSTITUTION: Baltimore Affairs Medical Center
COMPLETE ADDRESS: 10 North Greene Street, Baltimore, MD 21201
PHONE NUMBER: 410-605-7251
E-MAIL: dong.lee@med.va.gov

NAME OF IRB CONTACT: Barbara Calabrese
FWA NUMBER: 00001483
COMPLETE ADDRESS OF CONTACT: Veterans Affairs Medical Center
    Research Office10 North Greene Street, Room 3C140, Baltimore, MD 21201
PHONE NUMBER OF CONTACT: 410-605-7131
E-MAIL OF CONTACT: barbara.calabrese@va.gov

NAME OF PHARMACY CONTACT: Not applicable for this study
COMPLETE ADDRESS OF CONTACT:
PHONE NUMBER OF CONTACT:
E-MAIL OF CONTACT:

Contacts at the Johns Hopkins Medical Institute:

        NAME OF INSTITUTIONAL PI: Rex Young, M.D.
        INSTITUTION: The Johns Hopkins Medical Institute
        COMPLETE ADDRESS: 1830 E. Monument Street, #538, Baltimore, MD 21205
        PHONE NUMBER: 410-502-0799
        E-MAIL: ryung@jhmi.edu

        NAME OF IRB CONTACT: Susan Trabing
        FWA NUMBER: 00000287
        COMPLETE ADDRESS OF CONTACT: JHSPH IRB Office
           615 N. Wolfe Street, Suite E1100, Baltimore, MD 21205
        PHONE NUMBER OF CONTACT: 410-614-3049
        E-MAIL OF CONTACT: strabing@jhsph.edu

        NAME OF PHARMACY CONTACT: Not applicable for this study
        COMPLETE ADDRESS OF CONTACT:
        PHONE NUMBER OF CONTACT:
        E-MAIL OF CONTACT:

PRECIS

Background:

Liver cancer is the third most deadly and fifth most common cancer worldwide. Hepatocellular carcinoma (HCC) is the most frequent primary cancer of the liver with rising incidence in the United States. HCC patients have a dismal outcome and are an underserved cancer population. We are proposing a liver cancer case-control and case-case study to be conducted in Baltimore, Maryland.

Objectives:

Our research is aimed to identify genetic and genomic changes that might explain the aggressive nature of liver cancer. We will test the primary hypothesis that HCC biomarkers exist in blood products (or mouth wash), urine and/or tissues that can be used for early diagnosis and prevention. A unique study factor is the use of urine samples for biomarker identification. A secondary goal will be to examine whether certain genetic and genomic factors modify liver cancer susceptibility using candidate and genome wide association study approaches. Although other studies have investigated genetic susceptibility, large scale studies have not been conducted for liver cancer. Our study will also be valuable due to access to frozen tumor specimens from pre-therapy resection, whose collection is required for gene expression analysis. If certain genes are related to HCC risk and progression, this tissue resource will be used to further validate our findings with laboratory investigations. The proposed study is designed to be exploratory and hypothesis generating.

Eligibility:

The participants will reside in Baltimore City and surrounding areas. Cases will have pathologically confirmed liver cancer or risk factors for HCC development. Risk factors include chronic liver disease due to hepatitis B and/or C viral infection, nonalcoholic steatohepatitis or alcoholic cirrhosis. The study will be supported by an epidemiological infrastructure developed by our University of Maryland School of Medicine resource contractor for ongoing lung and prostate cancer case-control studies. The enrollment of controls will begin concurrently with case accrual, and will continue for 5 years. Population-based controls will be recruited using a triple eligibility criterion, such that controls eligible for the lung/prostate study will be asked to complete a supplemental questionnaire for the liver study.

Design:

The study will include 250 primary HCC cases, 500 patients with chronic liver disease (high risk non-cancer cases) and a sample of 500 population-based controls. The HCC and high risk non-cancer cases will be recruited at three Baltimore hospitals, the Veterans Affairs Medical Center, the University of Maryland School of Medicine and the Johns Hopkins Medical Institute, over a period of 5 years. Controls will be identified through a Department of Motor Vehicle database and match cases by age, gender, race and county of residency. The study will involve administration of a questionnaire and collection of blood (or mouth wash) from all study subjects. Fresh-frozen tumor and tumor-surrounding non-cancerous specimens will be obtained from approximately 20 percent of the cancer patients. The cases will receive one questionnaire. The first 12 months will constitute a pilot study, during which we will evaluate and refine recruitment procedures.

TABLE of CONTENTS

APPENDICES

1) INFORMED CONSENT DOCUMENTS (ICD)
2) LIVER CANCER STUDY QUESTIONNARES
3) MD LUNG & PROSTATE CANCER STUDY & MULTI-ORGAN STUDIES
   QUESTIONNARE
4) ELIGIBILITY CRITERIA CHECKLISTS FOR HCC CASES, HIGH RISK NON-
   CANCER CASES AND POPULATION-BASED CONTROLS

1.0 INTRODUCTION

1.1 Study Objectives:

PRIMARY OBJECTIVE
- We will perform a comprehensive large-scale molecular profiling study to identify HCC biomarkers for early diagnosis and prevention. The objective will be accomplished with 1) proteomic and metabolomic analysis of blood products or urine from 250 HCC cases, 500 high risk non-cancer patients (patients with chronic liver disease due to hepatitis B and/or C viral infection, nonalcoholic steatohepatitis or alcoholic cirrhosis) and 500 population controls and 2) collection of data from questionnaires, medical and pathology records, and the National Death Index. Biomarkers will also be assessed from HCC patient tissues (tumor and nontumor; needle biopsy) that are related to HCC diagnosis, prognosis, metastasis and molecular subtypes. *We are testing the hypothesis that susceptibility and progression to HCC can be identified early through the measurement of blood product, urine or tissue-based biomarker expression.*

SECONDARY OBJECTIVES
- We will perform a candidate and genome-wide association studies to identify and validate genomic risk factors that are associated with HCC. The key aim is to identify susceptibility and progression genes for HCC. Within cases, we will examine the relation of genetic variants to morphologic, molecular, clinical, prognostic outcome endpoints. The objective will be accomplished with 1) genotype analysis of blood (or mouth wash) DNA from 250 HCC cases, 500 high risk non-cancer patients and 500 population controls and 2) collection of data from questionnaires, medical and pathology records, and the National Death Index. *We are testing the hypothesis that susceptibility and progression to HCC is modified by variant genes that modulate cancer development, metastasis and survival.*

1.2 Background and Rationale:

Liver cancer ranks as the third most deadly and fifth most common cancer worldwide (1). The prevalence of liver cancer is especially high in Asia and sub-Saharan Africa, while low prevalence occurs in Europe and North America. However, the incidence of liver cancer is substantially increasing in both the U.S and Europe in recent years (2,3).

Hepatocellular carcinoma (HCC) is the most frequent primary cancer of the liver (4). While hepatocarcinogenesis is a long-term process, HCC is considered a fatal disease because of its poor prognosis resulting from the asymptomatic nature of the early disease (4,5). Many individuals are diagnosed with advanced HCC and have a dismal outcome. A small group of patients are eligible for curative surgical treatment. However, these patients have a high rate of recurrence after surgery or intra-hepatic

metastases (6). Despite many studies of HCC, information regarding phenotypic and molecular changes associated with the development of this disease is still limited (7,8). Such information is needed to develop methodology for early detection of HCC. Despite routine screening by ultrasonography and serum α-fetoprotein (AFP) of individuals at high risk, most patients are diagnosed at late stages of HCC. Additional complications arise due the multinodular nature of HCC, portal vein invasion, intrahepatic metastasis as well as the recurrence of nodules at multiple distant sites of the liver. Due to the long waiting list for liver transplantation and the limited number of liver transplants, many patients fail the Milan criteria for transplant eligibility (9).

HCC risk factors and molecular pathogenesis

HCC is one of the few human cancers where the underlying etiology can often be identified. Some of the etiological factors positively associated with HCC include chronic viral hepatitis and metabolic disorders. Environmental causes also play a role in HCC development either alone or in synergy with viral infection, including exposure to aflatoxin B1 (AFB), nonalcoholic steatohepatitis (NASH), cigarette smoking, heavy alcohol consumption (10). Other factors associated with liver cancer include low retinol levels, low consumption of dark green vegetables, and vinyl chloride exposure (11). The overall risk of developing HCC under these conditions is approximately 1-6% annually. Whether these factors induce HCC directly or whether they act indirectly by causing chronic liver injury and cirrhosis is currently unclear. In addition, HCC occurs more frequently in individuals with certain genetic diseases including hemachromatosis (12), Wilson's disease (13), porphyria (14) and α-anti-trypsin deficiency (15). The dietary aflatoxin AFB1 is an agent found in mycotoxin contaminated foods and its uptake can lead to a G to T transversion at the third base of codon 249 of TP53. This mutation of TP53 causes inhibition of wild type p53 and thereby increases cell survival. AFB1 uptake is the only etiology known to cause a distinct gene mutation leading to HCC (reviewed in (16)).

The molecular mechanisms underlying the development of HCC are not well understood. Many genes have been implicated in the pathology of HCC, including those regulating DNA damage response pathways (e.g. p53), genes involved in regulating cell growth and apoptosis (e.g. TGF-β, SMAD2, SMAD4, M6P/IGF2R), cell cycle control genes (e.g. p16, Rb, cyclin D), and genes responsible for cell-cell interaction and signal transduction (e.g. e-cadherin, APC/β-catenin) (17-19). Aneuploidy and multiple genetic alterations are often present in HCC (7,19,20). These findings provide some clues about the mechanisms leading to HCC. However, the true portrait of HCC still remains to be determined.

Several gene expression profiling studies have been conducted to determine the effects on the host cell after infection with HBV, HCV or the combination of both viruses. These studies analyzed the global changes that occur in cell lines after infection with various hepatitis viral proteins or of hepatitis-infected tumor samples from patients and animal models of hepatitis-related HCC. These types of studies help to understand the intricate interplay between the host and virus. These analyses lead to a greater understanding of the molecular pathogenesis of HCC and allow for

the definition of molecular signatures that can be used to determine molecular markers for diagnosis and prognosis of HCC.

HCC diagnosis

AFP is the only widely used serum marker for HCC and allows for the identification of a subgroup of patients with small carcinomas. However, elevated serum AFP is only observed in 33-65% of patients with small HCC (21). Nonspecific elevation of serum AFP has been found in 15-58% of patients with chronic hepatitis and AFP levels highly vary between different ethnic backgrounds. Therefore, it is necessary to identify new serological HCC biomarkers that have a sufficient sensitivity and specificity for the diagnosis of HCC patients, especially in AFP normal and/or smaller-sized HCC cases. Several candidates, including des-γ-carboxy prothrombin by a revised enzyme immunoassay kit and AFP-L3 [the Lens culinaris agglutinin (LCA) bound fraction of AFP] have been reported as potential diagnostic HCC markers (22-24). However, AFP remains the only universally accepted HCC biomarker in clinical practice. Therefore, improvement of the current screening system of high risk patients is a major goal.

Most symptomatic HCC patients are diagnosed at an advanced stage, thus precluding their chance for surgical intervention (25). In contrast, HCC patients who were diagnosed at an early stage and received curative resection have a significantly improved survival time (26-28). Thus, early detection and resection have been generally recognized as the most important factors to achieve long term survival for HCC patients.

The diagnostic tools, treatment modalities and screening programs for HCC have improved in recent years, but early detection still remains a challenge. At the time of diagnosis, only about 20 percent of HCC are eligible for surgical resection and survival after this procedure is only 30-40 percent at 5 years. Staging systems have been created to define prognosis and treatment options for many diseases, including HCC (29). Staging is essential, particularly in malignant diseases, to select and improve treatment. The requirements of a good staging system includes- simplicity and ease of use; reproducibility; provision of reliable information on the natural history of the disease; and categorization of patients into various treatment groups (e.g. sorafenib; IFN/5FU).

Well-defined and generally accepted staging systems are available for almost all cancers. HCC is an exception as many different staging systems have been introduced around the world and currently there is no consensus on which one is best. This has led to considerable confusion in the literature. An improvement in treatment in association with better diagnostic techniques has changed the rather fatalistic approach to HCC. The main factors affecting the prognosis of HCC are tumor stage, aggressiveness and growth rate of the tumor, general health of the patient, liver function and choice of therapy. Many different prognostic models therefore need to be developed for each stratum of the disease because a single system cannot accurately establish the prognosis of all patients and help determine the efficacy of all available therapies.

Prognostic assessment and choice of treatment options in HCC is complex because they depend not only on the grade of cancer spread (tumor staging) but also on the grade of residual liver function (liver disease stage). A clinical staging system

for cancer patients provides guidance for patient assessment and making therapeutic decisions. It is useful in deciding whether to treat a patient aggressively and in avoiding the over-treatment of patients who would not tolerate the treatment or patients whose life expectancy rules out any chance of treatment. Clinical staging is also an essential tool for comparison between groups in therapeutic trials and for comparison between different studies. Scoring systems arise as a compromise between simplicity and discriminatory ability.

Many systems have been developed to stage HCC and each is based on various patient populations, tumor characteristics, treatment and inclusion criteria (30). Individual systems may be applicable to only patients after resection, after transplantation or with advanced tumors. The well defined tumor node metastasis (TNM) staging system has been widely utilized to stage liver cancer and has since been modified by several groups to fit subgroups of patients with HCC and cirrhosis. For those patients being evaluated for transplantation or resection, additional staging systems have been created. To confound matters, the definition of HCC itself seems to be inconsistent. Small well-differentiated HCCs identified in Asian countries are called regenerative nodules in the West. The lack of a consensus on the definition and staging of HCC combined with the wide heterogeneity of the disease has interfered with clinical recommendations and progress.

HCC develops in a previously diseased liver so that both HCC and cirrhosis deeply influence survival and simultaneously determine the applicability and efficacy of therapy. Uni-dimensional prognostic systems accounting for only one of these hepatic diseases such as Child-Pugh score (31), and TNM may result in inaccurate survival prediction of HCC patients. The HCC population is characterized by a great heterogeneity since both tumor and cirrhosis may be diagnosed at different evolutionary stages each with different therapeutic perspectives and survival probabilities. Therefore, a staging system must be able to stratify HCC patients at these different categories reflecting this large range of potential survival figures.

HCC metastasis

Primary human tumors can often be successfully treated with either surgery or chemotherapy. However, a major complication of such treatment regimens results from the development of metastases, the tumor colonies derived from the spread of cancer cells from a primary tumor through the blood or lymphatic system (32). Local invasion and distant metastases are attributable to 90% of human cancer deaths (33). Understanding the mechanisms involved in the process of tumor invasion and metastasis is a major challenge. The identification of biochemical factors augmented in invasive tumor cells may lead to improved methods to predict whether tumor cells have already metastasized.

A metastasis model has emerged over many years of study using animal and cell culture models (34). It is thought that cancer metastasis is a highly complex multistep process that involves alterations in growth, angiogenesis, dissemination, invasion and survival, which leads to subsequent attachment and growth of new cancer cell colonies (33,35). Principally, a successful metastasis involves tissue invasion, intravasation, survival through circulation, extravasation and outgrowth with angiogenesis. Interestingly, about 5000 to as much as 10^7 circulating cancer cells can

be detected in the peripheral blood at any given time in a cancer patient, suggesting a constant dissemination of tumor cells. However, animal model studies indicate that the metastatic process is rather inefficient, which leads to the assumption that only a few rare cancer cells from primary lesions have acquired all the necessary steps with the right combination to form a metastatic colony. The dominance of this view, together with the nature of a multistep process during cancer initiation and progression, has led to a systematic effort to identify molecular events associated with the multiple steps of this process. Such an approach, while much less successful than the identification of cancer initiation genes, has identified many genes attributable to metastasis (34). However, due to limited technologies, it is still difficult to understand how these genes contribute to the key steps necessary for metastasis, thus precluding their potential clinical use. Furthermore, these studies have been largely focused on individual candidate genes, an approach that may be insufficient to precisely define the molecular basis of metastatic cancer cells. Recently, the metastasis paradigm has been challenged by the facts that most of the genetic and epigenetic changes necessary for metastasis appear to be the hallmarks of cancer. This raises a debate as to whether the tendency to metastasize is largely determined by the identities of mutant alleles that are acquired relatively early during multistep tumorigenesis (36).

Global molecular profiling by high-density cDNA microarray has become increasingly useful in identifying signatures that reflect the specific natures of diseases, such as cancer (37). A recent study indicates that a molecular signature can be identified in primary tumors and can be used to predict breast cancer patient survival and metastasis (38). Similarly, by comparing metastatic cells to primary tumors of lung, breast and prostate, a molecular signature was uncovered from primary cancer lesions that could predict patient prognosis. (39). Since microarrays detect signals contributed by the bulk of the tissues examined, the results suggest that most of these primary tumor cells have acquired changes that favor metastasis. Recently, we found that the gene expression signature of primary HCCs with accompanying metastasis was very similar to that of their corresponding metastases (40). In contrast, the gene expression signature differs significantly between metastasis-free primary HCCs and HCCs with accompanying intra-hepatic metastases (40). These results are consistent with our findings that the HCC metastasis signature is independent of tumor size, tumor encapsulation and patient age. Our findings, together with others, support the hypothesis that genes favoring metastatic progression are initiated in the primary tumors. It also implies that one should not simply assume a favorable prognostic outcome if a small cancer lesion is detected. A clinical challenge is to be able to identify cancer patients with metastatic potential in advance so that an appropriate therapeutic regime can be applied. The identification of a molecular signature in primary tumors to predict metastasis and survival has provided an opportunity to classify these cancer patients in advance and thus allow clinicians to have ample time to manage them.

Recent studies indicate that expression profiling with small non-coding RNA gene products (~22nt) known as microRNAs (miRNAs) is a superior method for cancer subtype classification and prognostication (41-43). miRNAs exist in many organisms and play key regulatory roles in mRNA translation and degradation by base pairing to partially complementary sites of the mRNA, predominantly in the 3' untranslated

region (44-46). miRNAs are expressed as long precursor RNAs that are processed by Drosha, a cellular nuclease, and subsequently transported to the cytoplasm by an Exportin-5-dependent mechanism (47,48). miRNAs are then cleaved by the DICER enzyme, resulting in ~17-24 nt mature miRNAs that associate with a RNA-induced silencing-like complex (RISC) (49,50). The expression patterns, function and regulation of miRNAs in normal and neoplastic human cells are largely unknown but emerging data and their frequent location at fragile sites, common break-points or regions of amplification or loss of heterozygosity reveal that they may play significant roles in human carcinogenesis. The abnormal expression of several miRNAs have been observed in Burkitt's lymphomas, B cell chronic lymphocytic leukemia (CLL) and in many solid cancer types, including breast, liver, lung, ovarian, cervical, colorectal and prostate (41,42,51-57). Functional analysis has revealed the downregulation of PTEN by miR-21, the tumor suppressor function of the let-7 family and the oncogenic function of the miR17-92 cluster (58-60). The biological and clinical relevance of miRNA expression patterns have been established in human B cell CLL and solid tumors, including breast cancers (42,54,61). Each miRNA has a distinct capability to potentially regulate the expression of hundreds of coding genes and thereby modulate several cellular pathways including proliferation, apoptosis and stress response (62). In addition, mature miRNAs are relatively stable. These phenomena make miRNAs superior molecular markers and targets for interrogation and as such, miRNA expression profiling can be utilized as a tool for cancer diagnosis (41,63). In addition, the study of miRNAs is advantageous in improving our understanding of the mechanisms of cancer progression.

In a recent study, we investigated the miRNA expression profile of HCC specimens from radical resection. We identified 20 miRNAs that are associated with HCC venous metastasis (64). In contrast to HCC staging systems, this 20-miRNA-based signature was capable of predicting survival and recurrence of HCC patients with multinodular or solitary tumors, including those with early-stage disease. Moreover, this signature was an independent and significant predictor of patient prognosis when compared to other available clinical parameters. Our study suggests that these 20 miRNAs can assist in HCC prognosis and may have clinical utility for the advanced identification of HCC patients with a propensity towards metastasis. Functional studies of these miRNAs may also help to elucidate the mechanism(s) leading to HCC metastasis.

Are metastatic signatures universal among cancer types? While the metastasis signature of adenocarcinomas of lung, breast and prostate is similar to the gene expression profiles of murine metastatic tumors (65), we have found no correlation between the signature of metastatic adenocarcinomas and the signature of metastatic HCC. These results suggest that the molecular program associated with HCC metastases may be unique to HCC patients.

Since Paget set forth the "seed vs. soil" hypothesis, most studies of metastasis have focused on the molecular analysis of the tumor and have largely ignored the heterogenic complexity of the surrounding tissue, which may also contribute to this process (66). Are pro-metastatic changes inherent to the tumor cell (seed) or are they acquired through time and/or influenced by environmental status such as the conditions of the metastasis sites (soil)? Cancer cells are capable of altering their

environment through a variety of mechanisms, including the production of growth factors and proteolytic enzymes, which allow for disruption of tissue homeostasis and the creation of pro-migratory and pro-invasive surroundings. However, tumors may also be directly affected by the tumor stroma itself, whereby tumor cells respond to certain factors present in the target organ that are permissive to and promote tumor extravasation, aggregation and metastasis (67). To determine whether metastatic ability is influenced by liver parenchyma, we have conducted gene expression profiling studies of non-cancerous liver parenchyma tissue from HCC patients with or without intrahepatic portal vein metastasis using cDNA microarray. We have identified a molecular signature that can significantly discriminate liver parenchyma from HCC patients with or without metastasis (68). The lead genes in the metastasis signature from liver parenchyma are involved in the cellular immune and inflammatory response, suggesting that the inflammatory status of tumor stroma may play an important role in promoting HCC tumor progression and metastasis. Thus, the condition of liver parenchyma, whether influenced by viral-hepatitis mediated liver damage or individual genetic constitution could significantly contribute to intrahepatic metastatic potential.

HCC genetics

Viral integration into the genome is thought to play a role in hepatocarcinogenesis. Though the HBV integration into the genomic loci of the oncogenes cyclin A2 , RARβ (retinoic acid receptor beta), hTERT (human telomerase reverse transcriptase) and SERCA1 encoding a sarco/endoplasmatic reticulum calcium ATPase has been reported, this seems to be a rare event and insertion is believed to occur randomly (69-72). Viral integration may rather affect a non-coding RNA or microRNA and thus further investigations are needed. In about 10% of HCC cases no insertion of HBV could be found, suggesting that integration is either not required or due to genetic instability (73).

In addition to viral integration, the HBV encoded transcriptional modulator HBx appears to be involved in liver carcinogenesis, since most integration events maintain HBx, suggesting an important role. The expression of HBx in transgenic mice demonstrated that HBx has an oncogenic function in the liver (74). However, it is unclear whether the expression of HBx alone is sufficient for carcinogenesis or an additional etiology, for example a dietary carcinogen, is required (75). In addition, ectopic expression of HBx has been implicated in chromosome instability, centrosome amplification, senescence, cell cycle progression and oxidative stress (76-78).

In addition to expression profiling analysis, karyotyping studies have also been performed using array-based comparative genomic hybridization (aCGH) and single nucleotide polymorphism (SNP) arrays to assess chromosome instability (CIN) in HCC (79-81). The existence of common chromosome aberrations is widely assumed to imply that cancer related genes exist within and around particular loci. Several studies have shown that gains of 1q, 5p, 5q, 6p, 7q, 8q, 17q and 20q and loss of heterozygosity (LOH) of 1p, 4q, 6q, 8p, 10q, 13q, 16p, 16q and 17p were associated with HCC. In addition, some of these copy number alterations and allelic imbalances

have been correlated with up-or down-regulated HCC related genes using Expression Imbalance Maps (EIM) (82).   Furthermore, several studies have demonstrated that certain chromosome abnormalities are correlated with clinicopathological features such as tumor size, advanced tumor grade and differentiation stage (83-85). A high number of chromosome abnormalities of a high fractional allelic imbalance (FAI) index has been correlated with a lower differentiation state, invasion and metastasis and a shorter survival period (86). These findings indicate that an increase in the overall chromosome abberations in HCC may be linked to an unfavorable prognosis and the extent of abberant chromosome regions could be useful for the prediction of prognosis and the selection of surgical therapy such as liver transplantation for HCC. Other studies have identified chromosome abnormalities in premalignant liver tissue such as chronic hepatitis and cirrhosis that may represent the early steps of hepatocarcinogenesis (87-90). Microsatellite instability (MSI), the mutation of short DNA repeats found evenly distributed throughout the genome normally repaired by the DNA repair mismatch system, has been observed in several types of cancers including breast, endometrial and gastric carcinomas. However, a comparatively low level of MSI has been observed in HCC and thus the biological and clinicopathological significance of MSI in HCC remains to be determined. These studies suggest that multiple genetic alterations accumulate during the development of HCC, however the exact sequence and number of events required for progression of this disease remains unclear.

In addition to HCC with obvious CIN, there are tumors with normal DNA content. In these cases, malignant transformation may occur through epigenetic mechanisms. One such mechanism is alterations in the methylation status of cancer-related genes such as tumor suppressors or oncogenes. Currently, the role of methylation in HCC carcinogenesis is not clear. Several studies also suggest that DNA hypermethylation of CpG islands is associated with HCC (91,92). Frequent promoter methylation and subsequent loss of protein expression has been demonstrated in HCC tumor suppressor genes p16, cadherin and 14-3-3o (93). However, these studies have assessed these genes in isolation and therefore global studies of the methylation phenotype in malignant, adjacent microenvironments and normal tissue need to be conducted to determine the role of methylation in hepatocarcinogenesis.   Global studies of the methylation phenotype are required to assess the significance of epigenetic silencing in hepatocarcinogenesis.

Allele variant genes belong to the category of low penetrance susceptibility genes (94). A mutation in a low penetrance gene does not cause cancer but contributes to cancer risk, and affects large segments of the population. The investigation of allele variant genes will provide clues to the role of genetic risk factors in nonhereditary cases. Further, the identification of low penetrance susceptibility genes can have important public health implications. The association between allele variant genes and cancer risk has been studied extensively (95,96). Only a small number of disease-associated allele variants have been identified at this point. As described above, HCC displays a series of gross genomic alterations, including the loss or gain of chromosome during cell division CpG island methylation, DNA rearrangements associated with HBV-related integration, DNA hypomethylation and microsatellite instability.

While genetic linkage-based studies have been effective in identifying causal genetic factors in Mendelian (single gene) diseases, most diseases do not exhibit Mendelian inheritance. An approach to identify genetic variants in common diseases circumventing this limitation is genome-wide association studies (GWAS). This approach involves a comprehensive search for genetic risk factors for a trait using a case-control-based association study that involves comparisons among hundreds of thousands of alleles between unrelated subjects with or without the trait of interest. GWAS of cancer based on common inherited SNPS are useful for the identification of germline risk alleles and have been conducted for prostate and colon cancers among others. In some cases, GWAS has identified molecular factors with utility for patient stratification, perhaps enabling personalized medicine. Due to the current limited case numbers for liver cancer in the United States, large scale GWAS studies have yet to be conducted.

An important goal is to classify HCC instances into subgroups based on the observed spectrum of genetic alterations and to determine treatment modalities that are most effective for each subgroup. Current high resolution platform limitations include studies of patients with different etiologies, but few analyze different time points during the progressive sequence of the disease and the limited size of the studies has undermined the capability to produce a consistent genomic profile. Thus, additional, well-designed genetic association studies are needed to enlighten our understanding of the variant genes affecting liver cancer.

RATIONALE FOR CONDUCTING THE STUDY IN THE GREATER BALTIMORE AREA

Our goal is to identify molecular and genetic risk factors for liver cancer. The three hospitals that we chose to recruit liver cancer patients are located in Baltimore City, Maryland. The study will be supported by an existing infrastructure with our contractor at the University of Maryland School of Medicine and an established relationship between our group and the Departments of Pathology at the two hospitals. The Laboratory of Human Carcinogenesis (LHC) is currently conducting lung and prostate cancer case-control studies at the same hospitals and is recruiting population-based controls to match those cases. The controls recruited for the lung and prostate studies may meet the eligibility criteria of the proposed liver cancer study. The mean and median age of cases in the lung and prostate cancer studies is 65-66 years. If the age distribution of the liver cancer patients does not change in the next 5 years, we expect about the same mean and median age for the liver cancer cases, as indicated by hospital records and cancer registry data (Tables 1 and 2). Such an age distribution will allow the matching of liver cancer cases with population-based controls using the established resources. Liver cancer cases will be identified primarily by liver panel tests including AFP and diagnostic imaging screening. We expect to recruit at least 10 patients per year (20 percent of expected recruited cases per year; n=50), who present with cancer and undergo radical resection, and/or transplantation. The remaining patients will have chemotherapy, radiation, or drug therapy, where only blood samples (or mouth wash) but no tumor specimens will be collected.

Table 1: Age Distribution of Liver Patients (HCC) at Two of the Participating Hospitals

Calendar Years Jan 2004 to Aug 2007			
Age (yrs)	University of Maryland School of Medicine Medical Center	Veterans Affairs Medical Center, Baltimore	Total
20-29	4	0	4
30-39	5	0	5
40-49	21	5	26
50-59	33	34	67
60-69	36	22	58
70-79	18	18	36
Total	117	79	206

Table 2: Age Distribution of High Risk Non-Cancer Patients at Two of the Participating Hospitals

Calendars Years Jan 2004 to Aug 2007			
Age (yrs)	University of Maryland School of Medicine	Veterans Affairs Medical Center, Baltimore	Total
20-29	68	4	72
30-39	177	24	201
40-49	422	417	839
50-59	426	1207	1633
60-69	119	196	315
70-79	35	46	81
Total	1247	1894	3141

RATIONALE FOR SELECTION OF SUBJECTS BY RACE

There is a large health disparity for HCC worldwide and in the United States (97). The causes of this health disparity are unknown. The incidence rates of HCC have tripled in the United States between 1975 and 2005. The age-adjusted trends examined in the Surveillance, Epidemiology and End Results (SEER) registries from 1975-2005 show a marked increase in HCC incidence among African-Americans compared to European Americans, even exceeding those of Asian/Pacific Islanders during 2003-2005 (98). The age-adjusted incidence of HCC between 2003 and 2005 was approximately 2.5 times higher in African-American men than European

American men. Also during this time, the age-adjusted incidence of HCC was approximately 2.3 times higher in African-American women than European American women. The prevalence of HCV is very high among African-Americans in the United States (approximately 3 fold higher than European Americans) and severity of liver disease and progression to HCC differs greatly from the European American population (99).   It is hypothesized that the genetic background of individuals and the tumor biology of HCC is different in the two ethnic groups.   In this vein, our recruiting strategy also includes racial matching between cases and population controls and among cases and high risk non-cancer cases taken the advantage that African American population is prevelant in the Great Baltimore area. However, our study does not exclude any ethnic group. Contingent upon the cohort characteristics at the end of the accrual period, we will examine the molecular and genetic differences among these ethnic groups in our collected cohort or after combination with other HCC-based studies conducted in the United States.

RATIONALE TO STUDY MOLECULAR PROFILES OF LIVER CANCER

HCC, endemic to Asia and Africa with a rising incidence in western countries (2,3,100), is one of the most common and aggressive cancers worldwide. It has been the third cancer killer worldwide and the second cancer killer in China since the 1990s (1,101). Globally, the five year survival rate of HCC is less than 5% and approximately 598,000 HCC patients die each year (102). The high mortality associated with this disease is mainly attributed to the inability to diagnose HCC patients at an early stage. In fact, most symptomatic HCC patients are diagnosed at an advanced stage, thus precluding their chance for surgical intervention (25). In contrast, HCC patients who were diagnosed at an early stage and received curative resection had a significantly improved survival time (26-28). Thus, early detection and resection have been generally recognized as the most important factors to achieve long term survival for HCC patients.

HCC represents an extremely poor prognostic cancer (7). The dismal outcome has been attributed to the highly vascular nature of HCC tumors which increases the propensity to spread and invade into neighboring or distant sites (6,103). Intra-hepatic metastases, especially venous metastases, are a major hallmark of metastatic HCC, with new tumor colonies frequently invading into the major branches of the portal vein, or to a lesser extent, the inferior vena cava and possibly other parts of the liver (104). Another feature of HCC is a high frequency of multiple nodules that occur in the same or different lobes. Many of these lesions can be multicentric, resulting from multiple de novo tumors, and thus may not be metastases. Recently, we developed a gene expression signature specific to primary HCC specimens to predict prognosis and venous metastases (40). The tumor signature provided 78% overall accuracy in predicting HCC patients with metastatic potential. The presence of a prognostic signature in primary HCC specimens was confirmed by several studies (105,106). We have also recently used high-throughput expression profiling to study the role of miRNAs in HCC (64). We identified 20 miRNAs that are associated with HCC venous metastasis and performed better than current HCC staging systems. Moreover, this signature was an independent and significant predictor of patient

prognosis when compared to other available clinical parameters. These 20 miRNAs may be clinically useful in identifying HCC patients with a propensity towards metastasis.

HCC is usually present in inflamed fibrotic and/or cirrhotic liver with extensive lymphocyte infiltration due to chronic hepatitis. Thus, it is possible that HCC metastatic propensity may be determined and/or influenced by the local tissue microenvironment of the host.

The accumulated evidence have clearly demonstrated that immune cells in tumor-surrounding regions can influence tumor progression (107). Our examination of the tumor microenvironment in HCC metastasis has indicated that the hepatic microenvironment from patients with HBV-positive metastatic HCC have a profound change in their gene expression profiles (68). We found that a dominant Th2-like cytokine profile, including an increase in IL-4, IL-8, IL-10 and IL-5 and a decrease in Th1-like cytokines including IL-1α, IL-1β, IL-2, IL-12p35, IL-12p40, IL-15, TNFα and IFNγ is associated with the metastatic phenotype. This immune-related signature also included an up-regulation of major histocompatibility class II genes, such as HLA-DR and HLA-DP, in patients with metastasis. A unique cytokine signature was capable of predicting patients with HCC metastasis in a cross-validated study. Moreover, the prognostic performance of this signature was independent of and superior to other available clinical parameters for determining HCC patient survival or recurrence.   We therefore suggest that the inflammatory status of the surrounding-tumor milieu, in addition to the metastatic potential of the tumor cells, may play a prominent role in promoting HCC tumor progression and metastasis.

HCC appears to be a heterogeneous disease since its clinical outcome can greatly vary. While the underlying molecular mechanisms of HCC pathogenesis remains largely unknown, multiple epigenetic and genetic changes have been associated with HCC, including the activation of oncogenes (e.g., N-ras, c-myc, c-fos) and inactivation of tumor suppressor genes (e.g., p53, p16, Rb) (7,76). Two signal transduction cascades that have been proposed to be critical in HCC are Wnt/Frizzled/β-catenin and insulin/IGF-1/IRS-1/MAPK pathways (108). In colorectal tumors, mutations in the Wnt/β-catenin signaling pathway such as in the adenomatous polyposis coli (APC) gene appear dominant (109). In contrast, various mutations in the Wnt components including APC, AXIN1, TCF4 and β-catenin have been found only in a small subset of HCC (110). To address the heterogeneity of HCC, new technologies such as microarray-based gene expression profiling and proteomic analyses have been applied to HCC classification (40,68,106,111,112) and several HCC subtypes have recently been revealed (113,114). One study suggested that the Wnt/β-catenin signaling pathway may be involved in a subtype of HCC with good prognosis (113). Functional analyses of these various signaling pathways offer promise for improving our understanding of HCC. The heterogeneous nature of HCC and the lack of appropriate biomarkers have hampered patient prognosis and treatment stratification. Recently, we have identified that a hepatic stem cell marker *TACSD1* (EpCAM) may serve as an early biomarker of HCC because its expression is highly elevated in premalignant hepatic tissues and in a subset of HCC. We then aimed to identify novel HCC subtypes that resemble certain stages of liver lineages by searching for EpCAM-coexpressed genes. A unique signature of EpCAM-positive

HCCs was identified by cDNA microarray analysis of 40 HCC cases and validated by oligonucleotide array analysis of 238 independent HCC cases which was further confirmed by immunohistochemical analysis of 101 HCC cases (115). Interaction network analysis and Kaplan-Meier survival analysis were conducted to reveal an association between prognosis and HCC subtypes that had unique molecular networks. EpCAM-positive HCC displayed a distinct molecular signature with features of hepatic progenitor cells including the presence of known stem/progenitor markers such as cKit, CK19, EPCAM and activated Wnt-β-catenin signaling, while EpCAM-negative HCC displayed features of mature hepatocytes. Moreover, HCC could be easily sub-classified into four groups with prognostic implication by determining the level of EpCAM and AFP. These four subtypes displayed distinct gene expression patterns that may resemble hepatic cell lineages, which include hepatic stem cell/hepatoblast-like HCC, bile duct epithelium-like HCC, hepatocytic progenitor-like HCC and mature hepatocyte-like HCC. Thus, this study allows us to establish an easy classification system to reveal HCC subtypes with distinct transcriptional profiles, which may enable prognostic stratification and assessment of HCC patients with adjuvant therapy and provide new insights into the potential cellular origin of HCC and its activated molecular pathways.

Further analysis indicates that EpCAM is a transcriptional target of Wnt-β-catenin signaling (116). We found that nucleur accumulation of β-catenin induced EpCAM gene expression, whereas the degradation of β-catenin or inhibition of Tcf/β-catenin complex formation reduced it in cultured normal human hepatocytes and HCC cell lines. We identified two Tcf binding elements in the EpCAM promoter that specifically bound to Tcf in an electrophoretic mobility shift assay. EpCAM promoter luciferase activity was down-regulated by the degradation of β-catenin or inhibition of Tcf/β-catenin complex formation. Furthermore, we found that EpCAM-positive HCC is more sensitive to Tcf/β-catenin binding inhibitors than EpCAM-negative HCC in vitro. Taken together, our data indicate that EpCAM is a Wnt-β-catenin signaling target gene and may be utilized to facilitate prognosis by enabling effective stratification of patients with predicted pharmaceutical responses to Wnt-β-catenin signaling antagonists. We suggest that the dissection of normal hepatogenesis may yield novel insights into the mechanism of hepatocarcinogenesis.

Recent progress in whole genome expression profiling has generated prognosis-related molecular profiles for many cancer types (40,63,64,68,117-119). These include prognostic signatures for breast, prostate, lung and liver cancer that are related to metastasis, survival and recurrence. Comparisons between the value of molecular signatures versus current clinical parameters have been tested in such studies using multivariate approaches and in several instances, the molecular signatures perform comparably or significantly better than current clinical parameters in terms of outcome prediction. In liver cancer in particular, although investigations have shown that gene expression profiles can predict HCC outcome, there is no study in the literature that has attempted to identify the genes and pathways with relevance to etiology and aggressive behavior of HCC in certain ethnic groups or in males. Studies conducted in our laboratory predominantly consist males, HBV positivity and an Asian cohort and thus, it remains to be determined whether the identified mRNA

and miRNA signatures are also suitable for females, HCC patients with other underlying liver diseases such as those related to hepatitis C and/or alcohol, or other ethnic groups with HCC. Our objective is to investigate gene and microRNA expression patterns in both tumor tissue and the surrounding non-tumor tissue to identify signatures associated with diagnosis, prognosis and outcome of HCC. We hypothesize that gene products of both tumor and stromal cells drive tumor progression and the invasion of normal surrounding tissue and are relevant to metastasis and patient outcome. We will develop gene expression profiles of 50 microdissected tumors. We will use laser-captured microdissection to separate tumor from normal tissue, and to determine the molecular profiles that separate 1) HCC with metastasis or no metastasis and 2) good from poor outcome. Candidate marker genes for tumor diagnosis, progression and outcome will be validated by quantitative PCR, and will be further studied.

We expect to collect 50 (20% of all cases; n=250) fresh-frozen tumor and nontumorous specimens that 1) are primarily identified by diagnostic imaging and AFP screening, 2) are positive for cancer by pathology, and 3) are surgically removed by radical resection.

RATIONALE TO STUDY HEPATIC CANCER STEM CELLS IN HCC.

Although considered monoclonal in origin, tumor cells are heterogeneous in their morphology, clinical behavior, and molecular profiles (120,121). Recent evidence suggests that heterogeneity may be a derivation of endogenous stem/progenitor cells (122) or de-differentiation of a transformed cell (123). The notion of stem cells as targets of malignant transformation have brought forth the similarities between cancer cells and normal stem cells in their capacity to self-renew, produce heterogeneous progenies, and limitlessly divide (124). The cancer stem cell (CSC) (or Cancer Initiating Cell) concept is that a subset of cancer cells bears stem cell features that are indispensable for a tumor. Accumulating evidence suggests the involvement of CSCs in the perpetuation of various cancers including leukemia, breast cancer, brain cancer, prostate cancer and colon cancer (125-129). Experimentally, putative CSCs have been isolated using cell surface markers specific for normal stem cells. For example, leukemia-initiating cells in NOD/SCID mice are $CD34^{++}CD38^{-}$ (127). Breast cancer CSCs are $CD44^{+}CD24^{-/low}$ cells while tumor initiating cells of the liver, brain, colon and prostate are $CD133^{+}$ (126,128-130). Stem cell-like features of CSC have been confirmed by functional *in vitro* clonogenicity and *in vivo* tumorigenicity assays.

The liver is considered as a maturational lineage system similar to that in the bone marrow (131). Experimental evidence indicates that certain forms of hepatic stem cells (HpSC), present in human livers of all donor ages, are multipotent and can give rise to hepatoblasts (HB) (132,133), which are, in turn, bipotent progenitor cells that can progress either into the hepatocytic or biliary lineages (132,134). Recently, a population of CD90+ cells in HCC have displayed tumorigenic capacity with metastatic features (135). Although AFP is one of the earliest markers detected in the liver bud specified from the ventral foregut (136,137), its expression has only been found in HB and to a lesser extent in committed hepatocytic progenitors, not in later

lineages nor in normal human HpSC (132). Recent studies indicate that EpCAM (CD326) is a biomarker for HpSC as it is expressed in HpSCs and HBs (132-134).

We recently identified a novel HCC classification system based on EpCAM and AFP status (115). Gene expression profiles revealed that EpCAM$^+$ AFP$^+$ HCC (referred to as Hepatic Stem Cell-like HCC; HpSC-HCC) has progenitor features with poor prognosis (115). Wnt/β-catenin signaling, a critical player for maintaining embryonic stem cells (138), is activated in these HpSC-HCC, and EpCAM is a direct transcriptional target of Wnt/β-catenin signaling (116). Moreover, HpSC-HCC cells are more sensitive to β-catenin inhibitors than EpCAM^ꭅ HCC cells *in vitro* (116). Interestingly, a heterogeneous expression of EpCAM and AFP was observed in clinical tissues, a feature that may be attributed to the presence of a subset of CSCs. We have confirmed that isolated EpCAM$^⬲$ HCC cells from HpSC-HCC cell lines retain CSC features. These cells are highly invasive and tumorigenic, and have activated Wnt/β-catenin signaling, which may keep "stemness" possibly by inducing self-renewal and inhibiting asymmetric division. Consistently, EpCAM$^+$ HCC cells, but not EpCAM$^-$ HCC cells, isolated from HpSC clinical HCC samples can efficiently form spheroids *in vitro* and initiate tumors with serial dilution in NOD/SCID mice (139,140). We have also shown that EpCAM plays a crucial role of in the maintenance of HCC cells with CSC features.

A major hurdle in identifying the cellular origins of HCC is the lack of appropriate markers specific to certain stages of liver lineages. Identification of specific markers that can distinguish HpSC from committed progenitor cells may clarify the origin of HCC CSCs. It is likely that HCCs can derive from multiple stem/progenitor cell stages, a hypothesis that can now be tested.   Current results suggest that the expression patterns of various stem cell markers in CSCs may be different, possibly due to the heterogeneity of stem cell marker expression in normal stem/progenitor cells where the CSCs may originate. Therefore, it would be useful to comprehensively investigate the expression patterns of stem cell markers to characterize the population of CSC that may correlate with the activation of their distinct molecular pathways. In this vein, we aim to identify and characterize hepatic stem cell populations from HCC tissues and blood products, using stem cell markers such as CD133, CD90 and EpCAM. In addition, since CSCs are considered more metastatic and drug/radiation resistant than non-CSCs in the tumor, and are responsible for cancer relapse, the development of treatment strategies that can specifically eradicate CSCs is warranted (141,142). Therefore, we will also test the utility of combining antibodies against hepatic stem cell markers with conventional chemotherapy to target both CSCs and non-CSCs for the treatment of HCC. These data may shed new light on the pathogenesis of HCC and may open new avenues for therapeutic interventions for targeting hepatic CSCs.

RATIONALE TO STUDY ASSOCIATIONS BETWEEN SERUM/PLASMA CYTOKINES AND HCC DIAGNOSIS, PROGRESSION AND OUTCOME

Mounting evidence indicates the involvement of cytokines in hepatocarcinogenesis. Thus, various avenues have been taken to elucidate changes in

cytokine expression levels in patients with HCC. The mRNA and protein expression of cytokines in HCC and liver related diseases has been demonstrated by immunohistochemistry (IHC), quantitative real time polymerase chain reaction (qRT-PCR) and enzyme-linked immunosorbent assay (ELISA). To determine whether cytokine expression correlates with disease progression, many comparisons have been made including HCC tumor and non-tumor samples, chronic liver disease (CLD) patients and normal healthy individuals and cytokine measurements in response to various treatment regimens. Since cytokines are secreted and can be measured in the serum and plasma, several groups have analyzed certain cytokines for their predictive capacity in relation to HCC disease. These include IL-10, IL-1β, TNFα, IFNγ, IL-12, IL-2, IL-18 and IL-6. Several studies have demonstrated that changes in cytokine expression result from various treatment regimens for HCC including IL-6, TNFα, IL-1β, IL-10 and IL-12. Some studies have correlated changes in cytokine expression with HCC metastasis and/or recurrence including IL-6, IL-1β, TNFα, and IL-8.

In our laboratory, we have shown that patients with portal vein metastasis have higher levels of Th2-related anti-inflammatory cytokines than proinflammatory Th1-like cytokines. A global Th1/Th2-like cytokine shift in the venous metastases-associated liver microenvironment coincides with elevated expression of macrophage colony stimulating factor (CSF1). Moreover, a refined 17-gene signature was validated as a superior predictor of HCC venous metastases in an independent cohort, when compared to other clinical prognostic parameters. We suggest that a predominant humoral cytokine profile occurs in the metastatic liver milieu and a shift towards anti-inflammatory/immune-suppressive responses may promote HCC metastases. The HCC cytokine signature may be clinically useful for diagnosis and advance identification of HCC with a propensity towards metastasis/recurrence. However, the application of this technique is currently limited to post-surgical prognosis since samples must be supplied from patients who undergo invasive procedures and the analysis was performed using microarray platforms. It would therefore be advantageous, especially for early disease stage patients, to establish blood-borne methods and high-throughput methodology for diagnostic/prognostic marker detection, thus providing a less invasive, cost-effective technique with low procedural risk and the possibility of sequential examination.

We hypothesize that cytokines present in the blood products of patients differ between 1) those with HCC vs. without HCC, 2) those with HCC risk factors or without HCC and 3) those with HCC vs. those with HCC risk factors.

RATIONALE TO STUDY ASSOCIATIONS BETWEEN MICRORNAS AND HCC DIAGNOSIS, PROGRESSION AND OUTCOME

Recent studies indicate that expression profiling with small non-coding RNA gene products (~22nt) known as microRNAs (miRNAs) is a superior method for cancer subtype classification and prognostication (41-43,63). miRNAs exist in many organisms and play key regulatory roles in mRNA translation and degradation by base pairing to partially complementary sites of the mRNA, predominantly in the 3' untranslated region (44-46). miRNAs are expressed as long precursor RNAs that are processed by Drosha, a cellular nuclease, and subsequently transported to the

cytoplasm by an Exportin-5-dependent mechanism (47). miRNAs are then cleaved by the DICER enzyme, resulting in ~17-24 nt miRNAs that associate with a RNA-induced silencing-like complex (49). The expression patterns, function and regulation of miRNAs in normal and neoplastic human cells are largely unknown but emerging data and their frequent location at fragile sites, common break-points or regions of amplification or loss of heterozygosity reveal that they may play significant roles in human carcinogenesis. The enhanced expression of precursor miR-155 in Burkitt's lymphomas and the frequent deletion or downregulation of several miRNAs have been observed in B cell chronic lymphocytic leukemia (CLL) and in many cancer types, including breast, lung, ovarian, cervical, colorectal, prostate, and lymphoid (41,42,51-54,56,57,143). Functional analysis has also revealed the downregulation of PTEN by miR-21, the tumor suppressor function of the let-7 family and the oncogenic function of the miR17-92 cluster (58-60). The biological and clinical relevance of miRNA expression patterns have been shown in human B cell CLL and solid tumors, including breast cancers (42,54,61). Each miRNA has the unique capability to potentially regulate the expression of hundreds of coding genes and thereby modulate several cellular pathways including proliferation, apoptosis and stress response (62). This phenomenon makes miRNAs superior molecular markers and targets for interrogation and as such, miRNA expression profiling can be utilized as a tool for cancer diagnosis (41,63).

We have identified several miRNAs that could significantly distinguish paired hepatic tumor and non-tumor tissues from radical HCC resection with >90% accuracy, suggesting that miRNAs may be utilized for diagnostic purposes to identify the presence of HCC tumors or specific HCC tumor types. We have also investigated the miRNA expression profile of HCC specimens from radical resection (64). We developed a unique 20 miRNA signature that could significantly distinguish HCC venous metastasis from metastasis-free HCC. In contrast to HCC staging systems, this signature was capable of predicting survival and recurrence of HCC patients with multinodular or solitary tumors, including those with early-stage disease. Moreover, this signature was an independent and significant predictor of patient prognosis and relapse when compared to other available clinical parameters. This study suggested that these 20 miRNAs can enable HCC prognosis and may have clinical utility for the advance identification of HCC patients with a propensity towards metastasis/recurrence. The tumor and nontumor samples from HCC cases accrued in this study can be assessed for miRNA expression and relation to clinical parameters such as metastasis can be assessed.

Although the miRNA signatures identified in our previous studies may be clinically useful for diagnosis and advance identification of HCC with a propensity towards metastasis/recurrence, the application of this technique is currently limited to post-surgical prognosis since samples must be supplied from patients who undergo invasive procedures. It would therefore be advantageous, especially for early disease stage patients, to establish blood-borne methods for diagnostic/prognostic marker detection, thus providing a less invasive, cost-effective technique with low procedural risk and the possibility of sequential examination. Several recent studies have demonstrated that circulating cancer-related DNA/RNA can be detected in serum/plasma of cancer patients (144-147). Given this possibility, we propose that

serum miRNAs are superior diagnostic/prognostic markers for HCC. To test this hypothesis, we will investigate whether 1) miRNAs can be detected in HCC serum/plasma 2) miRNA serum/plasma profiling by qRT-PCR can significantly distinguish HCC tumor from nontumor samples and 3) miRNA qRT-PCR serum/plasma profiling can predict HCC prognosis. We hypothesize that miRNAs in the blood products of patients differ between 1) those with HCC vs. without HCC, 2) those with HCC risk factors or without and 3) those with HCC vs. those with HCC risk factors. These studies may afford the preliminary framework for a new high-throughput clinical platform that can supplement current HCC staging systems to enable diagnosis and prognosis by providing a well-defined selection criterion for rational treatment decisions. If these studies are successful, we can also explore whether serum/plasma miRNAs differ in context of HCC, metastasis, etiology, prognosis, gender or race in an expanded cohort (using a pooled dataset from other US based studies).

RATIONALE TO STUDY METABOLITES FROM URINE AS RISK FACTORS IN LIVER CANCER

Metabolomics (MS-based Mass Spectrometry) is a unique and global approach enabling the characterization of cancer through metabolic changes, including endogenous and exogenous small-molecular weight substances, whose regulations are tightly linked with phenotype (148,149). A metabolomics approach to better understand the biological mechanisms underlying this complex disease is ideal as it provides high resolution of thousands of low concentration metabolites that reflect physiological states, and that yield a unique snapshot of the end products of gene expression. Metabolomics allows for sensitive detection of metabolites that are present at low concentrations. Samples will be processed using cutting-edge metabolomics instrumentation, ultra-performance liquid chromatography coupled with quadruple time-of-flight mass spectrometry (UPLC-QTOFMS). Reliability of the metabolomic profiles, after peak extraction and retention time correction, will be established with adequate controls, including pooled samples, duplicate samples, blanks and endogenous control spike-ins. High-throughput metabolomic studies, combined with pattern recognition computational tools, have successfully used metabolic profiles to discern cancer samples from healthy controls (150,151). We plan to apply machine learning tools for sample classification using dependable metabolic profiles to unravel putative biomarkers predictive of HCC and to gain a deeper understanding of the mechanisms underlying HCC. These putative biomarkers will be studied in the context of their metabolic pathways, thus shedding light into the metabolic and biological mechanisms that orchestrate HCC carcinogenesis. We will use this method to identify metabolites in urine whose abundance patterns allow the discrimination between HCC and control patients, high-risk cases and control patients and HCC cases and high-risk non-cancer cases.

RATIONALE TO STUDY ALLELE VARIANT GENES AS RISK FACTORS IN LIVER CANCER

The association between allele variant genes and cancer risk has been studied extensively (95,96). There are various reports on HCC susceptibility as it relates to low penetrance susceptibility genes. Polymorphisms in HLA alleles have been demonstrated to be associated with HCC susceptibility, for example with hepatitis infection (152). The presentation of viral antigens to T cells occurs through HLA class I and class II molecules, making them key regulators of optimum immune responses. Specific combinations of polymorphic HLA alleles have been shown to affect vaccination responses and the ability to bind and present antigens. Like many cancer types, HCC is affected by elements of the immune response. In our laboratory, we have shown through molecular profiling studies that the major histocompatibility genes, HLA-DRA and HLA-DPA, two Class II HLA molecules, are significantly upregulated in the tissue microenvironment of HCC patients with a metastatic phenotype. We are thus interested in polymorphisms of these genes that affect HCC susceptibility and/or progression. A preliminary case-control study in our laboratory utilizing DNA from a Chinese population have shown that certain HLA-related alleles are indeed associated with risk and outcome.

In addition, we have shown that an alteration of cytokines, towards a predominance of an anti-inflammatory Th2-related response, occurs in the microenvironment of HCC patients with poor outcome. Several studies have identified cytokine SNPs which are functionally associated with liver disease and/or HCC. The alteration of cytokine expression levels and the functional consequences of these changes in HCC may be caused by the response of the immune system to the presence of a primary lesion. However, an individual's genetic constitution may also play an important role in affecting elements of the immune system and generating tumorigenic effects. Many studies have been conducted to analyze SNPs as genetic markers due to their high density and even distribution in the human genome. As such, many groups have utilized SNPs for fine mapping disease loci and for analyzing the association of genes with disease outcome to determine how or whether inherited germ-line mutations can contribute to the susceptibility of certain individuals to cancer development. In this line, several pro- and anti-inflammatory cytokine SNPs have been functionally associated with liver disease and/or HCC. These include TNFα, IL-1β, IFNγ, IL-10, IL-4, TGFβ, and IL6 . We have found through FACS experiments that the large shift towards Th2 cytokines occurs through the activity of T cells. A subpopulation of T cells, T-regulatory cells (Tregs), have been widely recognized as the primary mediators of immune tolerance, blocking beneficial immune responses and limiting antitumor immunity. The Treg population functions through a co-stimulatory molecule, cytotoxic T-lymphocyte antigen 4 (CTLA4) and in the absence of functional CTLA4, Treg-mediated suppression is reduced. Thus, the shift towards an anti-inflammatory phenotype may be due to an alteration in the Treg population in patients with poorly prognostic HCC. This alteration may be due to a molecular change or may occur through a genetic alteration. Our findings suggest that genetic alterations and inter-individual variations in the genetic code lead to changes in HLA and cytokine levels which may influence an individual's disease outcome.

Our goal is to study the modulation of HCC risk by allele variant genes that regulate inflammation-related pathways (Table 3). Taqman-based assays have been

developed, and we can genotype DNA samples either at the CGF or at our laboratory. We will analyze genotype-risk associations between HLA-DRA, HLA-DPA and CTLA-4 genotypes and HCC risk or outcome. An allele variant that is associated with cancer risk will be further validated to establish biological plausibility and mechanisms of causality (Figure 1). The additional verification steps, as shown in the figure, are important to minimize the possibility of a false positive relationship between a genotype and cancer risk. These experiments will provide important mechanistic information. The genes that are most significantly down regulated in our analysis of HCC will be further investigated for aberrant promoter methylation. If we can detect the presence of a tumor-specific methylation pattern, we will test serum DNA for the presence of the same methylation pattern. The feasibility of this approach has been shown (153,154).   If detected, a tumor-specific methylation pattern of serum DNA could potentially be used as a prognostic marker. Disease-associated genotypes will subsequently be validated in larger studies.

Table 3: Proposed Genes for HCC genetic association study

Gene	Pathway	Function
HLA-DRA	Inflammation	Antigen Presentation
HLA-DPA	Inflammation	Antigen Presentation
CTLA-4	Inflammation	T-reg function

Figure 1: Strategy to Identify and Validate Genetic Risk Factors in Human Cancer
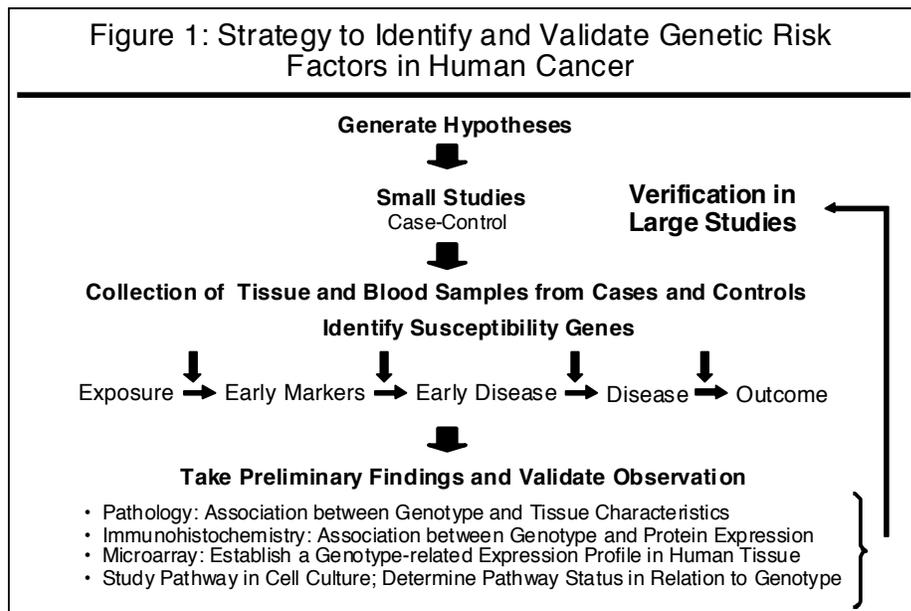
2.0   ELIGIBILITY ASSESSMENT AND ENROLLMENT

2.1   Eligibility Criteria

2.1.1 Inclusion Criteria

SELECTION OF CASE SUBJECTS (HCC AND HIGH RISK NON-CANCER CASES)

We will recruit incident cases of pathologically confirmed primary liver cancer or risk factors for liver cancer (patients with chronic liver disease due to hepatitis B and/or C viral infection, nonalcoholic steatohepatitis or alcoholic cirrhosis) at all stages of the disease that are age ≥18 years or ≤90 years. Treatment can be surgery or therapy. The following check list will be used to verify eligibility of a case subject.

ELIGIBILITY CRITERIA- HCC CANCER CASES
1. Diagnosed with primary liver cancer (HCC) within the last one year
2. Physician diagnosis based on AFP and ultrasound/imaging or Pathological diagnosis of primary liver cancer made at the local hospital pathology department
3. Resides in Maryland or surrounding states
4. Has a residential working phone within his/her home (*The reason for a participant having a residential working phone is that the matching population controls are selected through the Motor Vehicle Administration records and must have a residential telephone number to be selected as part of the random sample.)
5. Age is ≥18 years and ≤90 years
6. A non-objection statement by the physician from the hospital where the patient was identified, or listed as the treating physician by the tumor registry or surgical pathology report, to contact the patient is obtained
7. Is not currently residing in an institution, such as a prison, nursing home, or shelter
8. Is not a severely ill patient in the intensive care unit
9. Is able to give informed consent
10. Is physically and mentally capable of performing the interview
11. Must understand English well enough to be interviewed
12. Has never been interviewed as a control for this study
13. Subject provides informed consent and signs form.

ELIGIBITY CRITERIA– HIGH RISK NON-CANCER CASES
1. Has been diagnosed with chronic Hepatitis B or C, NASH or alcoholic liver disease with chronicity present at least for a 6 month period
2. Resides in Maryland or surrounding states
3. Has a residential working phone within his/her home
4. Age ≥18 years and ≤90 years

5. A non-objection statement by the physician from the hospital where the patient was identified or is being treated or by the surgical pathology report, to contact the patient is obtained
6. Is not currently residing in an institution, such as a prison, nursing home, or shelter
7. Is not a severely ill patient in the intensive care unit
8. Is able to give informed consent
9. Is physically and mentally capable of performing the interview
10. Must understand English well enough to be interviewed
11. Has never been interviewed as a control for this study
12. Subject provides informed consent and signs form.

SELECTION OF POPULATION-BASED CONTROLS

Population-based controls will be identified through the Motor Vehicle Administration (MVA), and matched on age (year of birth), race, gender and geography to cases. We sample controls in proportion to the population size of their county of residence. Recruitment of controls will start concurrently with case accrual, using the age, race and gender frequency distribution of cancer patients in previous years. The sampling frame is continually updated as we accrue information on the case distribution of age, race and gender. We will exclude controls that do not have a listed home phone number. The following check list will be used to verify eligibility of a control subject.

ELIGIBILITY CRITERIA - POPULATION-BASED CONTROLS
1. Resides in Baltimore City, Anne Arundel, Baltimore, Caroline, Carroll, Cecil, Dorchester, Harford, Howard, Kent, Queen Anne's, Somerset, Talbot, Wicomica and Worchester County.
2. Has a residential working phone within his/her home
3. Age ≥ 18 years and ≤ 90 years
4. Have never had radiation therapy or chemotherapy
5. Is not currently residing in an institution, such as a prison, nursing home, or shelter
6. Is physically and mentally capable of performing the interview
7. Must understand English well enough to be interviewed
8. Has never been interviewed as a control for the study
9. Subject provides informed consent and signs form.

2.1.2 Exclusion Criteria

Exclusion Criteria-HCC CANCER CASES

- Liver cancer patients who are older than 90 years are excluded because of co-morbidity considerations and the difficulty of matching those cases with population-based controls.
- Children and institutionalized patients are excluded from the proposed study. Very few HCC patients at the 3 participating hospitals are younger than 18 years of age and HCC incidence mainly occurs in adults.
- Liver cancer patients diagnosed with HIV infection

Exclusion Criteria-HCC HIGH-RISK NON-CANCER CASES

- High Risk Non-Cancer patients who are older than 90 years are excluded because of co-morbidity considerations and the difficulty of matching those cases with population-based controls.
- Children and institutionalized patients are excluded from the proposed study. Very few HCC patients at the 3 participating hospitals are younger than 18 years of age and HCC incidence mainly occurs in adults. High-Risk Non-Cancer cases who are younger than 18 years are excluded due to the need of matching with HCC cases and population controls.
- High Risk Non-Cancer cases diagnosed with HIV infection

Exclusion Criteria-POPULATION BASED CONTROLS

- Population based controls who are older than 90 years are excluded because of co-morbidity considerations and the difficulty of matching with HCC cases and high risk non-cancer cases.
- Population controls who are younger than 18 years and are institutionalized are excluded due to the need of matching with HCC cases and high-risk non-cancer cases.
- Population controls with prior cancer with the exception of non-melanomic skin cancer (squamous or basal cell)

2.2   Screening Evaluation

The verification of eligibility and enrollment of subjects will follow the procedures that are outlined above and summarized in the two flow diagrams at the end of the section (Figures 2, 3 and 4). The interviewer will assess the eligibility of liver cancer patients, and an interviewer will determine the eligibility of population-based controls. Eligible candidates will be offered an incentive with a monetary value of up to $25 to participate in the study.

ENROLLMENT OF HCC CASES AND HIGH RISK NON-CANCER CASES
Cases will be identified through resources including, but not limited to, daily afternoon visits, or phone calls, to the pathology and surgery departments (UMMC GI division and Departments at the VA and JHU) to identify all cases diagnosed that day with liver cancer or presenting with risk factors associated with liver cancer

development. We will also review the weekly lists of scheduled surgeries. We have established active collaborations with the Departments of Pathology and Surgery at the University of Maryland School of Medicine, the Veterans Affairs Medical Center, Baltimore and Johns Hopkins Medical Institute University Medical Center to maximize our ability for recruitment of new HCC or high risk non-cancer cases.

Case recruitment will follow the new HIPAA regulations ("HIPAA Privacy Rule"). However, to better assess eligibility, we intend to review medical and pathology records, cancer center registries and hospital databases. Our contractor has been granted a HIPAA waiver to perform the aforementioned study activities for our protocol titled "Resource for the collection and evaluation of human tissues and cells from donors with an epidemiology profile", University of Maryland School of Medicine IRB #HP-00042163. It is our plan to seek the same waiver for the proposed study.

We will seek an agreement with the Departments of Pathology and Surgery and the treating physicians that they will put a note into the medical record of severely ill liver cancer patients. The note will indicate that we should not contact this patient because of health concerns. If no such note is found in the medical record, we will proceed and contact eligible patients. If a note is found, the interviewer will record that this note was the reason for ineligibility on the eligibility record, in addition to age, race, and gender.

After the eligibility has been confirmed, an interviewer will contact patients to get both informed consent and authorization to obtain, use and disclose protected health information for research. The interviewer will administer the questionnaire to those who consented and blood (or mouth wash) and urine will be collected (all recruiters are trained in phlebotomy). The investigators will preferably obtain informed consent and administer the questionnaire well in advance of either a scheduled surgery or an alternative treatment such as radiation or cytokine therapy. However, informed consent and administration of the questionnaire can proceed after surgery or treatment. Some of the patients will not be rushed to surgery. The disease is detected by imaging and liver panel levels including AFP, and patients may seek a second or third opinion before they decide on a treatment option. Thus, we may have the opportunity of recruiting most patients weeks or months before either a surgery is scheduled or an alternative treatment will start. It is always at the discretion of the patient to give informed consent on one day but to have the interview or the blood collection (or mouth wash) on another day that is convenient for him. In the situation where we cannot get access to an eligible patient well in advance of surgery, we will approach the patient up to one day before surgery or the day of surgery to seek consent. We will offer to administer the questionnaires and obtain blood (or mouth wash) after the surgery. However, if the patient wishes to be interviewed and have the blood drawn (or mouth wash) at the same day when he gives informed consent, we will do so.

PROCEDURES
1. Interviewer obtains patient's name, location and physician's name

2. Interviewer screens medical record for eligibility. If ineligible, record age, race and gender. Record the patient's name, date of birth, county of residence, medical number and reasons for ineligibility.
3. The treating physician will receive and sign a blanket approval from the Maryland contract personnel for his/her patients' consent (to ensure that it is appropriate to contact their patients). If blanket approval is not received, physician consent will be obtained for his/her individual patient. If physician does not want the patient to be contacted, the interviewer records reason for ineligibility on eligibility record, in addition to age, race, and gender
4. Interviewer contacts case and obtains informed consent and authorization to obtain, use and disclose protected health information for research
5. If patient refuses (either right away or repeatedly delaying consent [failure to make commitment in response to 3 requests]), then record brief information from medical record about age, race and gender, on eligibility record. If subject requests that no such information should be kept on file, we will not record this information and will delete any information that already exists on file. However, we will keep the necessary information needed from the medical record dependent upon IRB approval or hospital approval to avoid re-contacting those subjects that refused to participate in our study.
6. Perform interview and complete questionnaire for those consented. The primary liver questionnaire will be administered. The interviewer will give a copy of the questionnaire to the patient before the interview begins. The patient will have the opportunity to read a question while being interviewed
7. Collect about 35cc of blood from HCC cases or high risk non-cancer cases. 35cc of blood will be collected and divided into two 7 ml red top tubes (serum) and two 10ml green top tubes (plasma). Aliquots (1.5ml) wil be frozen and batch shipped to LHC every three to four weeks. If a blood draw is not possible, a mouth wash is obtained.
8. Collect urine (50 ml)
9. Provide incentive of $25 for completion of the questionnaire.
10. Blood (separation of serum and blood clot; buffy coat; plasma and red blood cells) and urine to be processed within 8-24 hours at the University of Maryland School of Medicine, Department of Pathology. If a blood draw is not possible, a mouth wash is obtained. For mouthwash collection: Pre-rinse mouth with 20mL water. Rinse with 30ml Listerine for 30 sec and spit into a collection tube. Process mouthwash in 50mL centrifuge tube with 10mL 1XTE buffer. Resuspend pellet in 100ml 1XTE and aliquot (1.8ml) for storage at -80ºC.
11. Review pathology report at University of Maryland School of Medicine Department of Pathology to confirm diagnosis and abstract pathological data.
12. Request confirmation of the pathology for the specimens that are collected (performed at University of Maryland School of Medicine Department of Pathology).
13. Request part of tumor and nontumorous block and frozen or fresh tumor and nontumorous tissue (or needle biopsy), if available, for storage.

14. If subject is re-contacted at a later time for additional phlebotomy or urine sample, then an additional $25 incentive is given. If subject is contacted for additional questionnaire information, then no additional award is given.

ENROLLMENT OF POPULATION-BASED CONTROLS

Identify subjects through Motor Vehicle Administration database, frequency-matched on age, gender, race, and county of residence. Search the Internet for the home phone number. After mailing a letter that introduces the study, we will contact the person by phone one week later to verify eligibility and the willingness to participate in the study. We will obtain both informed consent and authorization to obtain, use and disclose protected health information for research at a location (usually their home) that is convenient to the subject. Medical records will not be collected from population controls. A trained interviewer/phlebotomist will contact the person and administer the questionnaire to those who consented. Blood (or mouth wash) and urine will be collected either at the interview or at a separate appointment.

PROCEDURES
1. An introductory letter is sent to a prospective subject to notify him/her of the study.
2. The subject is contacted by phone one week later to explain the study and how he/she was selected, and to request participation. The permission to interview the person is obtained. If person refuses to participate (either right away or repeatedly delaying consent [failure to make commitment for 3 requests]), then record brief information about age, race and gender. If subject refuses to give the information and also requests that no such information will be kept on file, we will not record this information and will delete any information that already exists on file. However, we will keep a MVA ID number to avoid re-contacting those subjects that refused to participate in our study.
3. Visit the subject to obtain informed consent and authorization to obtain, use and disclose protected health information for research. This is done either at the study office, or other locations such as person's home or person's work location.
4. Perform interview and complete the questionnaire. The primary general questionnaire will be administered first followed by the liver supplement (and prostate supplemental (if male)). The interviewer will give a copy of the questionnaire to the person before the interview begins. The person should be able to read the questions while being interviewed.
5. Collect about 55 cc of blood from high risk non-cancer cases for serum, plasma, leukocyte and lymphocyte collection. 55cc of blood will be collected and divided into two 7 ml red top tubes (serum) and four 10ml green top tubes (plasma). Aliquots (1.5ml) wil be frozen and batch shipped to LHC every three to four weeks. If a blood draw is not possible, a mouth wash is obtained. For mouthwash collection: Pre-rinse mouth with 20mL water. Rinse with 30ml Listerine for 30 sec and spit into a collection tube. Process mouthwash

in 50mL centrifuge tube with 10mL 1XTE buffer. Resuspend pellet in 100ml 1XTE and aliquot (1.8ml) for storage at -80°C.
6. Collect urine (50 ml).
7. Provide incentive of $25.
8. Blood (separation of serum and blood clot; buffy coat; plasma and red blood cells) and urine to be processed within 8-24 hours at the University of Maryland School of Medicine, Department of Pathology.
9. If subject is re-contacted at a later time for additional phlebotomy or urine sample, then an additional $25 incentive is given. If subject is contacted for additional questionnaire information, then no incentive is given.

RE-CONTACTING SUBJECTS FOR ADDITIONAL INFORMATION OR BLOOD

The interviewer who administered the questionnaire will be the person who will re-contact the subject. An alternative interviewer will re-contact the patient if the original interviewer is not available. This re-contact will not involve a script because of the large number of possible circumstances. An additional $25 incentive will be provided to a patient if they are re-contacted for blood or urine samples, but not for clarification of any questionnaire responses. Subjects will be re-contacted only under certain conditions. The conditions are as follows:

1. If there is missing information in the questionnaire
2. If the provided information is illegible
3. If there is a discrepancy between questionnaire and medical record information
4. If there is an obvious inconsistency in the answers to two similar questions in the questionnaires
5. If an insufficient volume of blood (or mouth wash)/urine has been collected
6. If there has been mishandling of a blood/urine sample in the laboratory
7. If there has been an accident leading to the loss of a blood (or mouth wash)/urine sample
8. If the analysis of a blood (or mouth wash)/urine sample returns a highly unusual result

We will not contact subjects, who refused to answer questions during the interview process, for responses to those questions they refused. Missing information due to refusal will be marked as such in the questionnaire.

PROCEDURES
1. Subject is contacted by phone notifying him/her of the need for additional information or blood (or mouth wash)/urine. If notified by letter, we will include a postage paid envelope and form to be returned if the subject does not want to be contacted.

2. If only additional information is requested, we will ask the subject by phone for the information. This will be done at the first phone contact, or two weeks after the notification letter was sent out.
3. If subject agrees, then additional phlebotomy or urine collection can be done at the home or in the hospital, whichever is convenient and an additional $25 incentive is given.
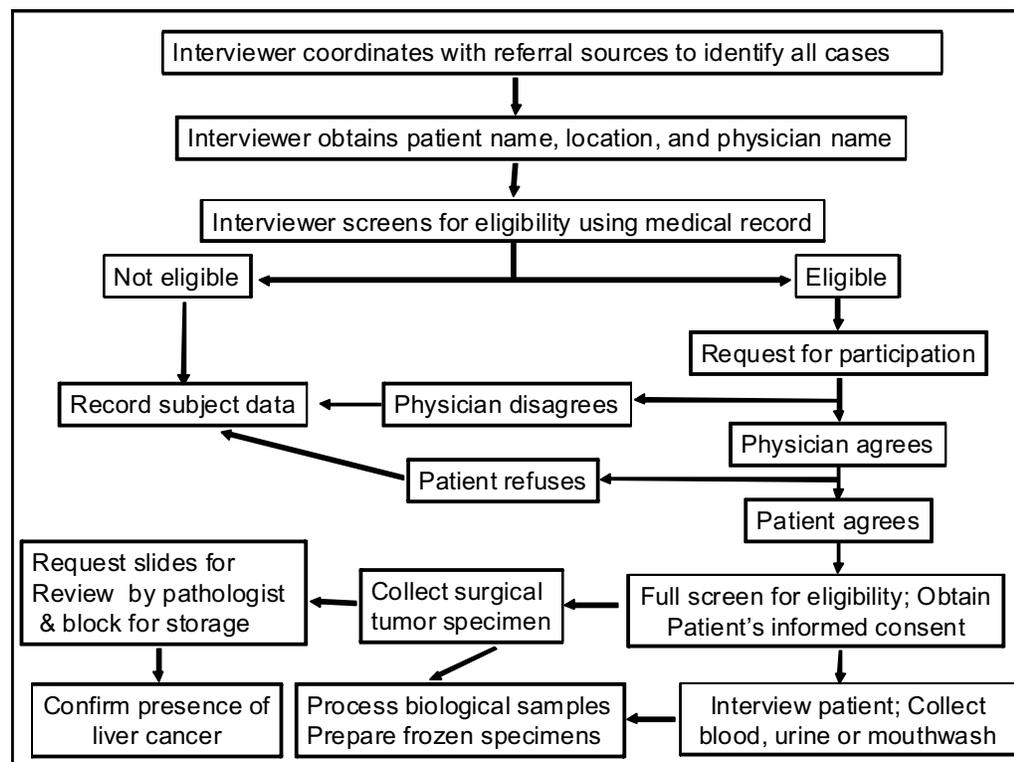
Figure 2: Flow Diagram for the Accrual of HCC Cases

```
┌─────────────────────────────────────────────────────────────┐
│   Interviewer coordinates with referral sources to identify  │
│                         all cases                            │
│                            │                                 │
│   Interviewer obtains patient name, location, and physician  │
│                          name                                │
│                            │                                 │
│   Interviewer screens for eligibility using medical record   │
│                            │                                 │
│   ┌──────────────┐                      ┌──────────────┐     │
│   │ Not eligible │ ◄────────────────────│  Eligible    │     │
│   └──────────────┘                      └──────────────┘     │
│          │                                     │             │
│          │                        ┌─────────────────────┐    │
│          │                        │ Request for          │    │
│          │                        │ participation        │    │
│          ▼                        └─────────────────────┘    │
│   ┌──────────────┐  ┌───────────────────┐      │             │
│   │ Record       │◄─│ Physician disagrees│◄────┤             │
│   │ subject data │  └───────────────────┘ ┌──────────────┐   │
│   └──────────────┘                        │ Physician    │   │
│        ▲        ┌───────────────────┐     │ agrees       │   │
│        └────────│ Patient refuses   │◄────└──────────────┘   │
│                 └───────────────────┘     ┌──────────────┐   │
│                                           │ Patient agrees│  │
│                                           └──────────────┘   │
│  ┌────────────┐  ┌─────────────┐     ┌───────────────────┐   │
│  │ Request    │  │ Collect     │     │ Full screen for   │   │
│  │ slides for │◄─│ surgical    │◄────│ eligibility;      │   │
│  │ Review by  │  │ tumor       │     │ Obtain Patient's  │   │
│  │ pathologist│  │ specimen    │     │ informed consent  │   │
│  │ & block    │  └─────────────┘     └───────────────────┘   │
│  │ for storage│                              │               │
│  └────────────┘                              ▼               │
│        │     ┌──────────────────┐  ┌──────────────────────┐  │
│        ▼     │ Process biological│  │ Interview patient;   │  │
│  ┌──────────┐│ samples Prepare  │◄─│ Collect blood, urine │  │
│  │ Confirm  ││ frozen specimens │  │ or mouthwash         │  │
│  │ presence ││                  │  └──────────────────────┘  │
│  │ of liver ││                  │                            │
│  │ cancer   │└──────────────────┘                            │
│  └──────────┘                                                │
└─────────────────────────────────────────────────────────────┘
```

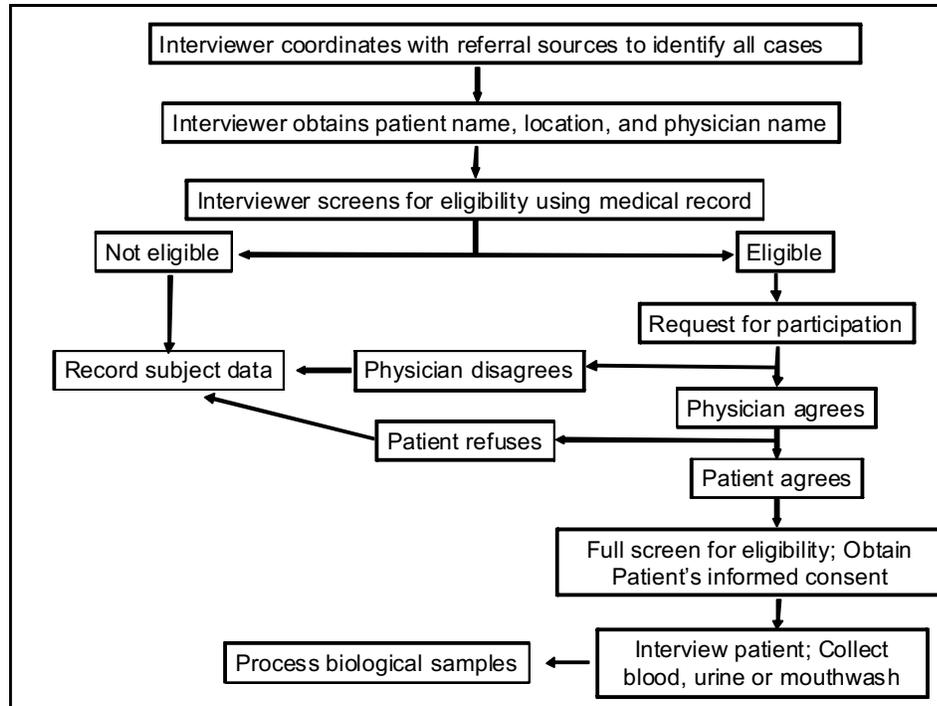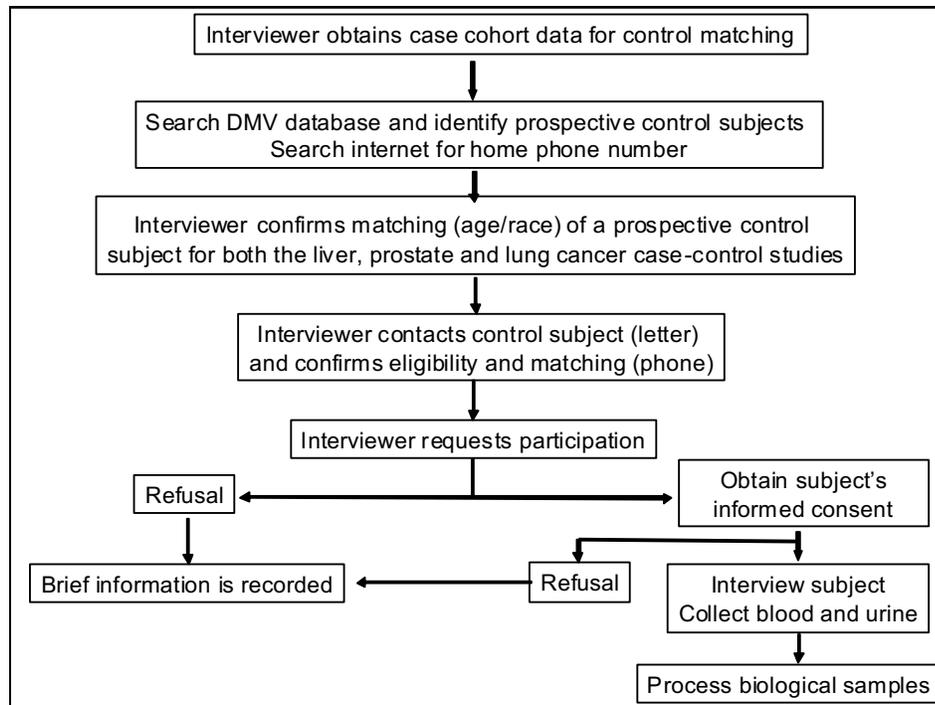**Figure 3: Flow Diagram for the Accrual of High Risk Non-Cancer Cases**

Figure 4: Flow Diagram for the Accrual of Population Controls

2.3 Registration Procedures

The proposed study is not a clinical investigation but an observational study. Subjects will not receive any type of therapy under the protocol. The collection of blood is the only medical procedure that is performed for this protocol. Patient care is not modified by the protocol. Cancer patients will receive their usual care and are asked to donate tumor specimens that are surgically removed as part of their routine treatment.

The accrual to study will not begin until the patient has been deemed eligible. Once the patient is considered eligible, a research study coordinator will contact the potential participant to schedule an interview. The Research Study Coordinator will register the patient at screening in the database where a unique study id number will be assigned (For patients at the VA, information is stored in a VA survey database until consent is obtained from the patient). Basic demographics are entered into the system for tracking purposes.
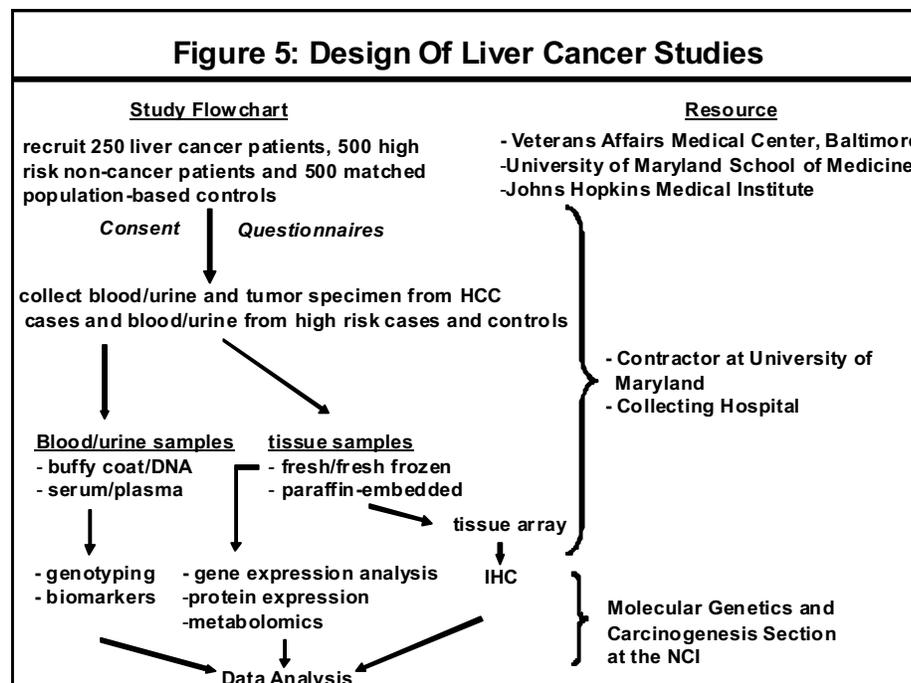
2.4 Randomization Procedures:

Not applicable for this study. The proposed study is not a clinical investigation but an observational study.

3.0 STUDY IMPLEMENTATION

3.1 Study Design:

We are proposing a case-case and a case-control study in which we collect 1) questionnaire information and 2) blood (or mouth wash) and urine samples from all subjects, and 3) tumor and nontumorous specimens from approximately 20% of the cases. Additional information for cases will be abstracted from medical and pathology records. The duration of the recruitment phase is 5 years. Research laboratory studies will begin at the end of the recruitment phase. The design of the study is illustrated in Figure 5.

Figure 5: Design Of Liver Cancer Studies

Study Flowchart

recruit 250 liver cancer patients, 500 high risk non-cancer patients and 500 matched population-based controls

Consent | *Questionnaires*

collect blood/urine and tumor specimen from HCC cases and blood/urine from high risk cases and controls

Blood/urine samples
- buffy coat/DNA
- serum/plasma

tissue samples
- fresh/fresh frozen
- paraffin-embedded

tissue array

- genotyping
- biomarkers

- gene expression analysis
- protein expression
- metabolomics

IHC

Data Analysis

Resource

- Veterans Affairs Medical Center, Baltimore
- University of Maryland School of Medicine
- Johns Hopkins Medical Institute

- Contractor at University of Maryland
- Collecting Hospital

Molecular Genetics and Carcinogenesis Section at the NCI

Questionnare

Trained interviewers will administer the primary liver questionnaire to HCC patients and high risk non-cancer patients. Population controls will receive the general questionnaire designed for the ongoing lung and prostate studies along with the supplemental liver questionnare. The interview will last approximately one hour. The questionnaire will be administered first and assesses prior medical and cancer history, tobacco use, family medical history, and exposure to known risk factors for liver cancer. It assesses anthropometry, medical history and family medical history.

Interviewer Training

The interviewers will receive a procedure manual. They will be trained in how to identify eligible subjects, how to provide informed consent, how to administer and properly complete the questionnaires, how to perform phlebotomy, and how to properly process blood (or mouth wash) and urine samples. The field supervisor will provide the training. Newly hired interviewers will practice the administration of the questionnaires to office volunteers. Interviewers will then administer the questionnaires and draw blood from subjects under the supervision of the field supervisor, who will provide feedback after the interview.

Assessment of Disease Stage and Outcome

For HCC cases, followup information, such as disease stage, metastasis, outcome (relapse and survival) and treatment, will be assessed from information gathered on an annual basis from the tumor registry and the National Death Index. For high risk cases, an annual followup will assess whether the patient has developed HCC through examination of the medical record.

Personnel

The Principal Investigator, Dr. Xin Wei Wang, a Staff Scientist, Dr. Anuradha Budhu and the NCI Project Officer, Dr. Glenwood Trivers, all at the LHC, NCI, will be responsible for the overall monitoring of the study. The associate investigators of the study, who are at the LHC, Veterans Affairs Medical Center, Baltimore, the University of Maryland School of Medicine and Johns Hokins Medical Institute, support the study with their expertise in pathology, clinical and laboratory research aspects of liver cancer. The study will be facilitated through a service contract to our laboratory. The contractor is the University of Maryland School of Medicine Pathology Department. The contract supports an established epidemiological infrastructure at the 3 participating hospitals. Contract personnel include a PI for the contract and a case-control supervisor, a nurse and two interviewers, staff to collect and process biological specimens, data entry personnel, and a biostatistician and epidemiologist for data analysis and quality control assessment. Leoni Leondaridis, a software consultant to the LHC, will assist with database development and management.

Study Personnel Supported by Our Current Contract

Employee	Position	Percent Effort*
D. Mann	P.I.	60
D. Perlmutter	Case-Control Supervisor	100
David Carmean	Field Supervisor	90
J. Cottrell	Tissue Collection Supervisor	70
TBA	Nurse	50
Tisha Trinh	Res. Study Coord.	100
Vanessa	Res. Study Coord.	100

Frederick		
Jane Kim	Res. Study Coord.	50
A. Salabes	Research Assistant	100
TBA*	Research Assistant	50
Daniel Harris	Res. Study Coord.	100
L. Hall	Res. Study Coord.	100
B. Williams	Data Entry	65
C. Loffredo	Epidemiologist	10

*for lung, prostate and liver studies

Previous Experience of the University of Maryland School of Medicine Contractor With Case-Control Studies

The first case-control study conducted by the University of Maryland School of Medicine at Baltimore contractor began in 1986, using the University of Maryland School of Medicine and the Veterans Affairs Medical Center, Baltimore to recruit cases and hospital-based controls. It was a lung cancer study to investigate the association between genotypes and lung cancer risk. A second, follow-up study was completed in 1998, which was designed to replicate the first study with some minor study modifications. A third study began in 1999. This lung cancer case-control study has a target accrual of 1,200 subjects, separated into 400 cases, 400 frequency-matched hospital controls, and 400 frequency-matched population-based controls. The aim of this study is to investigate lung cancer risk with respect to gender, race, and inherited susceptibility as it relates to mutagen sensitivity, p53 mutation load, and allele variant genes. Mutagen sensitivity is assayed in blood lymphocytes, and compared between cases and controls. Preliminary results confirm the hypothesis that lung cancer patients are more sensitive to mutagen-induced DNA damage. This study (NCI protocol OH98-C-N027) is scheduled to continue for several more years, as the request for recruitment above target accrual has been approved.

Another study aims to investigate prostate cancer risk with respect to race, diet, lifestyle and inherited susceptibility as it relates to genetic variants and morphologic, molecular, clinical, prognostic and outcome endpoints. This prostate cancer case-control study has a target accrual of 1200 subjects, separated into 600 cases and 600 frequency-matched population-based controls recruited from the University of Maryland School of Medicine and the Veterans Affairs Medical Center, Baltimore.

Limitations of the Case-Control Design

A frequent limitation of case-control studies is the heterogeneity of the population as it relates to race, age, and catchment area. If not addressed, the study may yield differences in genetic frequencies that are unrelated to the case status. To correct for the problem, we have restricted the study to a narrow geographic area, and will match for age and gender by frequency. We have

conducted a survey to determine the expected age range for cases. We anticipate a mean and median age such that it will permit an age matching with population-based controls using the established resources. Cases could be rendered more similar to controls with a selection criterion for controls that would achieve matching of cases based on census tract. This alternative would not eliminate a possible selection bias but would make the study substantially more difficult to perform. Also, our experience in Baltimore is that population control accrual is not skewed to one part of Baltimore. A 2002 survey for the ongoing lung cancer case-control study indicated that the recruited cases and population-based controls have a similar residential history. Thirty-eight percent of the cases were from Baltimore City versus 54% of the controls, and 62% of the cases lived in Baltimore suburbs or Eastern Shore counties versus 46% of the controls. We do not expect a selection bias based on genotype that would be confounded by behavior or education, as the genotype is unknown to the subjects and there is no plausible relation to either behavior or other factors that might affect the selection. A limitation common to other case-control studies is the possibility of differences between cases and the control population that are caused by a confounding disease, and not the disease under investigation (case identification among a hospitalized cohort, hospital-based controls). We are limiting the bias by matching cases with population-based, but not hospital-based, controls.

Another limitation relates to problems associated with rapid subject accrual. We expect to identify all cases through hospital pathology and oncology departments. This will be verified through tumor registry records at 6-months intervals. Based on our experience in the ongoing lung and prostate cancer case-control studies, we expect to contact most of the subjects prior to treatment or surgery, but will confirm this during our pilot study. We will allow for accrual of subjects before and after surgery, which will give us more flexibility for the accrual process.

3.1.1 Dose Limiting Toxicity
Not applicable for this study

3.1.2 Dose Escalation
Not applicable for this study

3.2   Study Drug Administration:
Not applicable for this study

3.3   Dose Modification for Toxicities:
Not applicable for this study

3.4 Correlative Studies for Research/Pharmacokinetic Studies:

All specimens will be given a study identification number at the contract site. Specific specimens will be sent to NIH investigators based on queries sent to the Maryland Contractor. The handling procedures and use of the specimens is outlined below.

Genotyping at the NCI
- Isolate DNA from whole blood or buffy coat using a previously tested DNA isolation kit
- Determine DNA quality with Taqman-assays and $OD_{260/280}$ readings >1.8
- Perform genotype analysis with quality-controlled Taqman assays, using 384-well plates. Assays will be performed at the NCI Core Genotyping Facility. Each assay will contain negative and positive controls, and 10% blinded duplicates across and within Taqman plates. Greater than 90% concordance among duplicates will be required for inclusion in data analysis. A >90% completion rate, >98% concordance rate and hardy-weinberg equilibrium tests will be used for inclusion. Odds ratios for cases and controls will be assessed using logistical regression while association with survival among cases will be assessed by cox hazard regression using STATA. Our survival assessment will be survival after surgery date.
- Data analysis will be performed at the LHC

Gene expression analysis at the NCI
- Cut fresh-frozen tissue into serial 8μm sections
- Microdissect tissue into tumor and non-tumor tissue, using laser-captured microdissection (LCM). A pathologist will help with the procedure. The equipment for LCM is provided by the NCI core facility
- Isolate, amplify and label RNA, using existing protocols. Currently, our laboratory uses a QIAGEN protocol for the isolation of total RNA from microdissected tissue and the small sample labeling protocol by Affymetrix to produce biotin-labeled cRNA
- Hybridize cRNA onto the Affymetrix GeneChip Plus 2 arrays (gene expression)
- Array quality assessment will include array image inspection and in-house software packages developed by our bioinformatics staff
- Data analysis will be performed at the LHC

Stem cell marker analysis at the NCI
- Fresh blood will be collected (~10mL) in green top tubes containing sodium heparin and will be sent to NCI for immediate processing or stored at 4°c overnight. Cancer stem cells will also be monitored in available fresh/fresh frozen tissue by IHC analysis of stem cell markers.

- Red blood cells will be removed with lysis buffer (25mL lysis buffer/1mL blood) with 5 minute incubation followed by centrifugation at 300g.
- Circulating tumor cells (CTCs) will be isolated via negative enrichment to remove normal blood cells
- CTCs will be labeled for stem cell marker expression such as EpCAM, CD133 etc using commercial antibodies and relative amounts will be determined by FACS

MicroRNA expression analysis at the NCI
Tissue analysis:
- Cut fresh-frozen tissue into serial 8μm sections
- Microdissect tissue into tumor and non-tumor tissue, using laser-captured microdissection (LCM). A pathologist will help with the procedure. The equipment for LCM is provided by the NCI core facility
- Isolate, amplify and label RNA, using existing protocols. Currently, our laboratory uses a QIAGEN protocol for the isolation of total RNA from microdissected tissue and the small sample labeling protocol by Affymetrix to produce biotin-labeled cRNA
- Data analysis will be performed at the LHC in a similar fashion as described above for gene expression analysis

Blood Product analysis:
- Isolate small RNA from serum or plasma samples using commercially available kit (MirVANA, Invitrogen)
- Perform MicroRNA expression analysis using pre-made Taqman probes for specific microRNA purchased from Applied Biosystems; Each sample will be performed in triplicate and those with greater than 5% error in expression readings will be excluded from further analyses. Cycle readings under 36 will be considered good. Differences between high and low RNA readings will be based upon the median, tertile or quartile cutoffs.
- Data analysis will be performed at the LHC

Cytokine expression analysis at the NCI
- Incubate serum or plasma samples with cytokine probes using commercially available multiplexed human cytokine plate (MesoScale)
- Each assay will contain negative and positive controls, blanks and 10% blinded duplicates across and within Taqman plates. Greater than 95% concordance among duplicates will be required for inclusion in data analysis. Differences between high and low RNA readings will be based upon the median, tertile or quartile cutoffs.
- Data analysis will be performed at the LHC

Metabolomics analysis at the NCI

- Urine samples from HCC cases, high risk non-cancer cases and population controls (~150uL) will be prepared with neat HPLC grade acetonitrile (1:1) and centrifuged followed by processing by MS-based mass spectrometry in both positive and negative-ion mode.
- Quality will be assessed using control samples including pooled samples, blanks, duplicates and endogenous "spike-ins"
- Data analysis will be performed at the LHC

3.5   Study Calendar:

The study will be implemented in two phases. The first phase is conducted as a pilot study to evaluate and refine recruitment procedures. The second phase is the main study that will be closed when the projected 250 HCC cases, 500 high risk non-cancer cases and 500 controls have been enrolled. Participants in the pilot study will count toward the total accrual.

TIME SCHEDULE OF THE STUDY

Month 1   Begin pilot study

Month 13   Review and revise protocol if needed
               Resubmit to IRB if needed

Month 14   Begin accrual for the study at the projected rate of 50 HCC cases, 100 high risk non-cancer cases and 100 population-based controls per year.

Month 62   Close enrollment if target accruals are met

PILOT STUDY

During the 12-month pilot phase, we will recruit about 100 patients who are either diagnosed with HCC or have risk factors associated with liver cancer development. We will also start recruiting controls that match the cases on age (5-year intervals), race and gender. We will recruit the cases at the 3 participating hospitals.

We will assess the availability of liver cancer patients, who 1) are willing to participate in the study, 2) provide blood specimen (or mouth wash), 3) provide tumor specimens. We will also assess the availability of controls that are eligible for both the lung, prostate and the liver cancer case-control studies and frequency-match the liver cases by age, race and gender.

The criteria that are used to determine if the pilot is successful are 1) the accrual of 50 HCC cases, 100 high risk non-cancer cases and 100 matched controls, 2) the collection of 250 blood samples (or mouth wash), and 3) the collection of 10 fresh-frozen tumor and nontumorous specimens (20% of 50 cases). If major modifications are required to improve accrual rates or other

aspects of the study, then the pilot study will be extended to address these issues. Otherwise, we will start recruiting for the main study at the projected rate of 50 HCC cases, 100 high risk non-cancer cases and 100 controls per year.

3.6   Concurrent Therapies:
Not applicable for this study.

3.7   Surgical Guidelines:
Not applicable for this study.

3.8   Radiation Therapy Guidelines:
Not applicable for this study.

3.9   Criteria for Removal from Protocol Therapy and Off Study Criteria

3.9.1 Criteria for removal from protocol therapy:
Not applicable for this study. We are proposing a study that is not a clinical investigation. Participant subjects will not receive treatment under protocol.

3.9.2 Off-study Criteria
The participant completes the questionnaire, blood (or mouth wash) and urine specimen collection. If the participant goes to surgery, tissue collection is also a part of participation. If the patient refuses part of the procedures once enrolled, he/she can still participate. If the patient refuses all parts of the enrollment after signing the consent form, he/she will be withdrawn from the study. The date and reason of patient withdrawal will be documented. Subjects will remain on-study in case they have to be re-contacted for questions regarding their questionnaire data or for collection of additional blood (or mouth wash) or urine as described in Section2.2 "RE-CONTACTING SUBJECTS FOR ADDITIONAL INFORMATION OR BLOOD". Once a subject is taken off study, no further data can be collected.

4.0   SUPPORTIVE CARE
Any investigator on the team involved in this questionnaire will arrange a social work consult if the patient requests.

5.0   DATA COLLECTION AND EVALUATION

5.1   Data Collection

The University of Maryland School of Medicine contractor is responsible for the collection and evaluation of the study data. All data will be kept secure. Personal identifiers will not be used when collecting and storing data. An

enrollment log will be maintained with the Maryland Contractor which is the only location of personal identifiers with unique subject identification number.

All data will be entered and stored in a centralized case-control study database at the University of Maryland School of Medicine. Leoni Leondaridis, a software consultant to the LHC, is assisting the contractor with database maintenance and upgrades. The centralized database consists of three databases, a tissue, a survey, and a tracking database, that are linked through unique identifiers. The tracking database contains the personal information of each study participants. Access to this database is strictly controlled, and the database is not accessible to NCI researchers. The tissue database is a tracking and inventory database for fixed and fresh-frozen tissues, and for blood samples. Updated files of the database are available to the NCI researcher. The survey database contains the demographic and epidemiological information of the study participants. Access to information in this database has to be requested. Information from the tissue and survey databases cannot be linked to any personal identifiers.

SURVEY DATA
Initially, all forms including the eligibility screener forms, consents, HIPAA authorization forms, interview forms, specimen collection forms, pathology and medical reports, etc., will be assigned a study ID number, which is common only to all forms and specimens that belong to one particular participant. All forms are then entered into the centralized case-control study database for tracking and storage. The forms are filed in locked file cabinets in the secured study office.

QUESTIONNAIRE DATA
The questionnaire data are entered into the survey database. If a discrepancy occurs, a supervisor will verify the correct entry. Prior to entry, all questionnaires are coded and edited using a standardized code for occupation, illness, and medication. A coder will highlight all sections in a questionnaire that contain either an illogical response or missing data. The highlighted sections will be reviewed by a supervisor, or the Research Study Coordinator where needed. A patient may be re-contacted to clarify an answer or to complete a missing response. The questionnaires are then entered into the database and verified by the Epi Info software. A contractor to the University of Maryland School of Medicine developed the Epi Info software. The software is programmed to include range checks, skip patterns, and codes for missing data and don't know answers. If errors are detected, a supervisor will assist with the correction. Monthly reports on coding and editing, data entry and the verification status are prepared and presented at monthly staff meetings and site visits.
During the study, quality control will consist of data comparisons among interviewers to determine the quantity and quality of information that they have gathered, by evaluating characteristics such as interview duration, number of interview problems reported, number of refusals, distribution of subject answers, and number of missing and incomplete answers. In addition, a small random sample of subjects will be re-contacted to inquire about the interview experience,

to note any problems, and to confirm that the interview actually took place. Quality assessment will follow procedures that are already in place for the lung and prostate case-control studies.

OUTCOME DATA

For HCC cases, followup information, such as disease stage, metastasis, outcome (relapse and survival) and treatment, will be assessed from information gathered on an annual basis from the tumor registry and the National Death Index. For high risk cases, an annual followup will assess whether the patient has developed HCC through examination of the medical record.

DATABASE SETUP

The current system is designed using Access 97. It is housed on a secure Windows Server NT with Windows authentication and only authorized users have access to the shared folder. In addition, the database itself has another layer of security that lets only registered users log on to the application. The database was specifically designed to be a tracking system for all data from hospital cases and the controls recruited through the MVA database. The database is backed up by the School of Medicine server at the University of Maryland School of Medicine, and also locally by the study data manager. The database has an incremental daily backup, and a regular full back up each weekend. This system is constantly updated.

The database is also used to perform the frequency matching of cases with population-based controls from the MVA database. The MVA data sets are imported into the database and can be searched for matching controls but the contract statistician performs the actual matching. The system has a built-in tool that loads the names and/or addresses into search engines like Yahoo, and if the home phone is found for the candidate control subject, a unique registration number is generated for the person. This study ID number is then used in all subsequent steps. Once a study ID number has been generated, the subjects can be contacted for study eligibility and consent.

5.2   Response Criteria

Not applicable for this study.

5.3   Toxicity Criteria

Not applicable for this study.

5.4   Sample Storage, Tracking, and Disposition

COLLECTION AND PROCESSING OF BIOLOGICAL SAMPLES

SAMPLE TYPE AND USE:

We will collect blood (or mouth wash), urine, tumor and nontumorous tissue specimens. The collected blood (or mouth wash) will be used for the analysis of genotypes, single nucleotide polymorphisms, methylation status, and the measurement of cytokine and microRNA concentrations in serum and plasma. Urine samples are collected to study the association between metabolites in the liver and risk or progression of the disease. The tissue specimens are collected for whole-genome gene expression analysis and the preparation of tissue arrays for immunohistochemical analysis of protein expression. All samples will be collected by the Maryland Contractor and assigned a study ID number. Personal identifiers will not be used on stored research samples.

PROCEDURES FOR COLLECTION OF BLOOD (OR MOUTH WASH) AND URINE

The collection of blood, mouth wash and urine will follow procedures that have been established for our protocol titled "Resource for the collection and evaluation of human tissues and cells from donors with an epidemiology profile", University of Maryland School of Medicine at Baltimore IRB #0298229. The same procedures are used in the lung and prostate cancer case-control study. The procedures are as follows.

Protocol for blood collection:
1. Observe universal precautions for prevention of transmission of blood-borne pathogens
2. Clean skin with alcohol wipe and wait to dry
3. Collect about 35-55cc of blood for serum and plasma extraction dependent on patient type. Apply pressure and band-aid

Protocol for mouthwash collection:
1. Rinse mouth with about 20ml water to wash out any food in mouth. Discard.
2. Rinse mouth with 30ml Blue color Listerine for 30 seconds and spit to a collection container.
3. Cover with lid and send to lab with ID label.
4. Complete Intake Form.
5. Fill out Specimen Collection Form.

Protocol for urine collection:
1. Collect approximately 50ml urine from each patient.

PROCEDURES FOR PROCESSING OF BLOOD AND URINE AT THE UNIVERSITY OF MARYLAND SCHOOL OF MEDICINE DEPARTMENT OF PATHOLOGY

Processing of Blood Samples:
   a. Centrifuge green top tube(s) and red top tube(s), approximately 850 g (2000 rpm on Sorval T 6000) for 10 minutes to separate components.

b. Withdraw plasma (green tops) and aliquot into sterile cryovials at 1.5 ml per vial; withdraw serum (red tops) and aliquot into sterile cryovials at 0.5 ml per vial.

c. Label each vial appropriately with printed bar code label (i.e., plasma normal [from green top tube], serum normal [from red top tube]).

d. From the green top tubes, remove the buffy coat from the red blood cell pellet; this appears as a white/pink layer on top of the RBCs, and can be done with a serological pipette or a transfer pipette. Place buffy coat in a cryovial, wash once with Phosphate Buffered Saline (PBS), and centrifuge as above to precipitate cells into a pellet. Aspirate supernatant, label vial appropriately (as buffy coat normal) with printed bar code label.

e. From the red top tube, pour the RBC/clot into a 13 ml plastic tube, label appropriately with a bar code label, and place in NIH cryobox (manilla colored cardboard 3 inch cryobox) in -80 C freezer.

f. For collection of RBCs, add PBS to green top tubes to resuspend the RBC pellet, and centrifuge for 5 min. at 850 g (2000 rpm). Aspirate PBS and repeat this wash step. After the second wash, re-suspend the cells in an equal volume of PBS, aliquot 1.5 ml each into up to 10 cryovials, and centrifuge as above. Aspirate PBS, and label vials appropriately (as RBC normal).

g. Place all plasma, serum, buffy coat, and RBC vials in NIH cryobox (white, Sarstedt cardboard cryobox) in -140 C freezer.

Processing of mouthwash:
1. Transfer specimen from the container to a 50 ml centrifuge tube.
2. Rinse container with about 10 ml of Listerine or PBS (Phosphate Buffered Saline) and add to 50 ml tube.
3. Centrifuge tube at 2700 rpm for 15 minutes to pellet cells.
4. Discard the supernatant carefully so as not to disturb pellet.
5. Wash the cell pellet in 25 ml of 1 X TE buffer and again centrifuge at 2700 rpm for 15 minutes to pellet cells.
6. Discard the supernatant and re-suspend the pellet in 100 ul 1 X TE buffer facilitated by vortexing.
7. Aliquot the sample into two 1.8 ml labeled NUNC Cryovials and store samples in cryobox for NIH (white, Sarstedt cardboard cryobox in -140 C freezer).
8. Batch ship to NIH with other NIH tissues.

Processing of urine:
1. Using a 25 ml serological pipet, aliquot urine into 13 ml plastic tubes, with up to 10 ml into each tube, using up to 5 tubes.
2. Label tubes with printed bar code labels and place in NIH cryobox (manila colored cardboard 3 inch cryobox) in -80C freezer.

PROCEDURES FOR TISSUE COLLECTION

The collection of tissue specimens will follow procedures that have been established for our protocol titled "Resource for the collection and evaluation of human tissues and cells from donors with an epidemiology profile", University of Maryland School of Medicine at Baltimore IRB #0298229. The same procedures are used in the lung and prostate cancer case-control study. For surgical specimens, each late afternoon one of the lab personnel checks the schedule for possible tissues. A note that reminds the operating room (OR) personnel to send the resected specimen to the pathology fresh, not fixed, is given to the OR. When the resection of a tumor is done, the OR will call the University of Maryland School of Medicine Department of Pathology. A designated person will transport the tissue to the pathology, where the tissue is procured as soon as possible. All biological samples will be stored at the pathology until further notice. The procedures are as follows.

1. Tissues are collected at the time of liver resection or transplant
2. Notify surgeon and the pathology department of tissue collection
3. A designated person will take the tumor and nontumorous specimen(s) as fast as possible, but within 10min of resection, to the pathology room
4. A pathology technician will immediately process the sample.
5. Within 20min of receipt, place 3/4 of the specimen in a pre-labeled container. Of the ¾ specimen, immediately flash-freeze ½ of this specimen and store the container at -70° C. For the remaining ½ specimen, place in a 15ml centrifuge tube containing stem cell media and antibiotics. If the fresh sample is procured before 3pm, the fresh specimen will be sent to LHC with colpack in a styrofoam box by global messenger service. If the fresh sample is procured after 3pm, the fresh specimen, an email will be sent to NIH investigators notifying them of the specimen to determine whether they want the specimen sent the same day.
6. Within 20min of receipt, prepare a paraffin block of remaining tissue section
7. Prepare a H/E slide for diagnosis by pathologist
8. Send samples to LHC at approximately two week intervals where they are inventoried and sent for long term storage at the Frederick repository.

TISSUE DATA AND TRACKING OF BIOLOGICAL SPECIMENS

Tissues and blood samples (or mouth wash) are entered into the database under the patient's unique study ID number. The information on tissue, blood (or mouth wash) and urine samples includes date procured, type of tissue, type of preservation, serum, plasma, buffy coat, amount, number of vials, date shipped. Different tissue preservation methods, e.g., formaldehyde-fixed, OCT-fixed, flash-frozen, will generate different records. Frozen tissues are intermediately stored at the University of Maryland School of Medicine and shipped in batches, on dry ice, to the NCI. Paraffin-embedded blocks are kept at the University of Maryland School of Medicine Department Of Pathology. The tissues, blood (or mouth wash) and urine samples that are received at the NCI will be entered into an inventory database at the LHC and will be stored long-term at a contracting

facility in Frederick, MD during the study and following completion of the study. The PI will report any loss or destruction of samples to the IRB.

6.0   STATISTICAL SECTION

STUDY POPULATION
This study proposes recruitment of HCC cases, high-risk non-cancer cases and population controls. Both men and women and members of all races and ethnic groups are eligible as cases or high-risk cases for this study. Population controls are currently limited to Caucasians, African-Americans and Asians.

SUMMARY OF STUDY DESIGN AND OBJECTIVES
We are proposing a liver cancer case-control study that will include 250 liver cancer cases, 500 high risk non-cancer cases and a sample of 500 population-based controls (Table 4). The aim of the study is to identify HCC biomarkers for early diagnosis and progression and to identify genetic risk factors associated with HCC risk and outcome. The accrual time is projected to be 5 years, averaging an annual accrual of 50 HCC cases, 100 high risk non-cancer cases and 100 frequency-matched controls. Liver cancer patients, who are older than 90 years, are excluded because of co-morbidity considerations.

The primary objective of the study is to identify HCC biomarkers. The objective will be accomplished with 1) microarray analysis (cDNA and miRNA) of tumor and nontumor tissues from 50 HCC cases; 2) blood product profiling (serum, plasma) to assess cytokine and miRNA in 250 HCC cases, 500 high risk non-cancer cases and 500 population controls and 3) metabolite profiling (urine) to assess metabolites in 250 HCC cases, 500 high risk non-cancer cases and 500 population controls. A secondary objective is to examine the associations of variant genes with liver cancer risk and outcome. This secondary objective will be accomplished by genotype analysis of blood (or mouth wash) DNA from 250 HCC cases, 500 high risk non-cancer cases and 500 controls. Both objectives will involve collection of data from questionnaires, medical and pathology records, and the National Death Index. All statistical analysis will be performed at the LHC.

Table 4: Target Accrual Numbers for the Proposed Study

Target Accrual Numbers of Cases and Controls			
	Hospital 1	**Hospital 2**	**Total**
HCC Cases	125	125	250
High Risk Non-Cancer Cases	250	250	500
Population Controls	250	250	500
Total	625	625	1250

AVAILABILITY OF CASES AND JUSTIFICATION OF SAMPLE SIZE

A review of cancer registry data shows that 2 of the participating hospitals (UMMS and VA) treated 206 liver cancer cases and 3141 high risk non-cancer cases from 2004-2007 (Tables 1 and 2). Most of the cases had residency in the greater Baltimore area. We expect to accrue 250 HCC cases, 500 high risk non-cancer cases and 500 population controls over a 5 year period. Among the 250 HCC cases, approximately 10% may undergo resection and allow collection of tissue samples for microarray-based experiments.

We will conduct genome-wide gene expression analysis of liver carcinomas and nontumorous liver tissues to compare expression profiles among HCC cases The minimum number of tumors needed to generate a molecular profile that differentiates groups cannot be determined for the proposed experiments. A best estimate is 13 per group based on the following assumptions: A significance level of 0.001 and a power of 0.05, assuming the data is base 2log transformed and that the variance in expression for a gene of interest is 0.5, and that one wishes to be able to detect genes for which one of the two groups exhibits a two-fold change in expression at the specified power. However, recent gene expression studies used fewer tissues than we are proposing for our experiments and yielded expression signatures that predicted HCC metastasis and survival (40). Thus, the proposed sample size for our investigations appears to be adequate and is still suitable to do laser-captured microdissection.

We will use the case-control design to detect statistically significant differences in the frequency of genotypes between liver cancer patients, high risk non-cancer patients and randomly selected population-based controls that are matched by age and gender to the cases. The study is designed to detect risk associations that have odds ratios of ≥ 1.5. Associations between liver cancer and genotypes of the inflammation and p53 pathways have been found to have odds ratios that are in the range of 1.5-4.4 (155-159). We are not aiming to detect weak risk associations that are reflected by odd ratios between 1 and 1.5. Such associations are prone to a selection bias and confounding, and would require a larger sample. Power calculations for the proposed study are provided below. Our proposed study is exploratory and designed to detect potentially novel genotype–cancer risk associations. The finding of a new disease-associated genotype can subsequently be validated in a larger study.

POWER CALCULATIONS

Power calculations were performed with the Power and Sample Size (PS) software, Version 2.1.30. William D. Dupont and Walton D. Plummer at the Vanderbilt University Medical Center have developed this software to conduct power calculations for clinical and epidemiological studies. The program is available on the Internet at http://www.mc.vanderbilt.edu/prevmed/ps/.

ANALYSIS OF VARIANT GENES WITH ALL CASES GROUPED TOGETHER

We assume 1) a study population of 250 HCC cases, 500 high risk non-cancer cases and 500 population controls and 2) two-sided alpha = 0.05, and 3) a prevalence of 5%, 10%, 20% or 40% for the disease-associated allele for odds ratios of 1.5, 1.8, 2.0 or 2.5. We selected the candidate allele variants using the following criteria (and/or): 1) the allele variant gene is functionally important (literature), and 2) the polymorphism will likely affect protein expression/stability/activity or mRNA splicing/stability, and 3) the frequency of the rare allele is greater than 5% in the general population. The power to detect an association between an allele variant and cancer risk at a given odds ratio is shown in Table 5 (HCC vs High risk non-cancer cases or HCC vs Population control) and Table 6 (High risk non-cancer cases vs Population Control). For a multiple comparison study, a Bonferroni correction (0.05/3) is used to generate an alpha of 0.017. The power to detect an association between an allele variant and cancer risk at a given odds ratio using multiple variables is shown in Table 7 (HCC vs High risk non-cancer cases or HCC vs Population control) and Table 8 (High risk non-cancer cases vs Population Control).

Table 5: Power Calculations for Proposed Case-Control Studies (HCC vs High risk non-cancer cases or HCC vs Population control)

Odds ratio	5% Prevalence Power (%)	10% Prevalence Power (%)	20% Prevalence Power (%)	40% Prevalence Power (%)
1.5	26	42	61	74
1.8	50	73	91	97
2.0	64	87	97	99
2.5	89	99	100	100

Table 6: Power Calculations for Proposed Case-Control Study (High risk non-cancer cases vs Population Control)

Odds ratio	5% Prevalence Power (%)	10% Prevalence Power (%)	20% Prevalence Power (%)	40% Prevalence Power (%)
1.5	33	55	77	89
1.8	63	87	98	100
2	79	96	100	100
2.5	97	100	100	100

Table 7: Power Calculations with Multiple Comparison Correction (HCC vs High risk non-cancer cases or HCC vs Population control)

Odds ratio	5% Prevalence Power (%)	10% Prevalence Power (%)	20% Prevalence Power (%)	40% Prevalence Power (%)
1.5	15	27	45	59
1.8	34	59	82	92
2	49	77	94	98
2.5	79	97	100	100

Table 8: Power Calculations with Multiple Comparison Correction (High risk non-cancer cases vs Population Control)

Odds ratio	5% Prevalence Power (%)	10% Prevalence Power (%)	20% Prevalence Power (%)	40% Prevalence Power (%)
1.5	19	38	63	79
1.8	46	76	95	99
2	64	91	99	100
2.5	92	100	100	100

DATA ANALYSIS PLAN

ANALYSIS OF GENE EXPRESSION PROFILES

The data from the Affymetrix HG-U133A Plus2 array will be estimated using the RMA procedure from Bioconductor (www.bioconductor.org) and global median centering. Previous analysis on raw data from replicate samples with RMA indicates that the mean standard deviation of log2 intensities is 0.1 across all intensity levels. This corresponds to a typical technical error of 5%-10% (160). Good quality normalized array data will be imported to BRB ArrayTools for further analysis.

The data from microRNA arrays (Ohio State University in-house microRNA chip, version 4) will be assessed through an R program package developed in-house to determine quality (LHCChip) and to perform normalization (loess or median). Good quality normalized array data will be imported to BRB ArrayTools for further analysis.

We will select differentially expressed genes for contrasts between HCC and high risk non-cancer cases, HCC and controls as well as high risk non-cancer cases and controls. Among HCC cases, we will select differentially expressed genes for contrasts between tumor and nontumor groups, metastasis groups,

outcome groups and subtype groups. All comparisons will be performed using the statistical packages provided in BRB ArrayTools such as class comparison, class prediction and survival analysis. Genes will be chosen based on statistical p value (p001 or p01) and FDR cutoff (<5%) To address multiple comparisons, we will use a Benjamini and Hochberg procedure (161).

ANALYSIS OF SERUM/PLASMA PROFILES

The data from cytokine or microRNA levels in serum/plasma will be first be assessed for quality based on blanks, intra and inter-plate sample replicates. A minimum of 10% error (coefficient of variation) will be allowed. High or low levels of cytokines or microRNAs will be determined based on the median across all samples tested or based on tertile or quartile cutoffs. The association with diagnosis (HCC case vs population control), disease progression (High risk non-cancer case vs HCC case or High risk non-cancer case vs Population Control) or prognosis will be graphically and statistically assessed using PrismGraph (Student's T tests or Kaplan-Meier Curves/Log-rank Pvalue). P values below 0.05 will be considered significant.

ANALYSIS OF URINE PROFILES

The data from urine metabolomics studies will be analyzed using R to extract relevant peaks (by controlling signal to noise ratios) and retention time alignments using the statistical package XCMS. The detection of relevant metabolites, contributing to cancer or other clinical variable classifications, will be assessed via ANOVA (Analysis of Variance) and random forest (RF). For each RF model, the contribution of each variable to sample classification is determined by their Gini importance measure. Multiple iterations of RF will be performed and metabolites that consistently contribute to the classifications will be selected. Differences in abundance of the metabolites in the predictive panels will be validated in an independent platform using a triple quadrupole MS. The identity of these compounds will be verified by comparison with standard curves generated from relevant standards, which are commercially available. Fragmentation patterns will also be compared with those of appropriate standards. The metabolites will then be compared among HCC and high risk non-cancer cases, HCC and controls as well as high risk non-cancer cases and controls.

ANALYSIS OF GENOTYPES

We will analyze genotype-risk associations in the context of known risk factors of HCC, and will investigate relationships between genotypes and poor outcome. If we investigate the association using a priori hypotheses, we will use an unadjusted alpha = 0.05 as the cutoff for significance. If we do not use a priori hypotheses, we will adjust alpha appropriately. An a priori hypothesis would be based on existing knowledge that associates the genotype with cancer risk.

Data analysis of the case-control studies will examine HCC occurrence and relation to allele variant genes as well as cirrhosis and relation to allele variant genes. Data analysis of the case-case study will examine risk factors for HCC and relation to allele variant genes in the context of patient outcome. Allele variant genes will be categorized (homozygous wild-type, wild-type/mutant, homozygous mutant) for analysis. Other risk factors will also be examined as categorized variables, separated into subgroups of exposure (e.g., dichotomous, tertiles, quartiles). The relationship between HCC and risk factors will be examined by the chi-square test, and by odds ratios with 95% confidence intervals, using the homozygous wild-type genotype (major allele), or the lowest exposure subgroup, as the reference group. Odds ratios and 95% confidence intervals will be calculated by unconditional logistic regression. Regression analysis will be performed in univariate and multivariate models to assess associations between risk factors and cancer susceptibility, and to adjust for possible confounding factors. Analyses correlating tissue marker expression with, cytokine concentrations, miRNAs, metabolites or genotypes, will be done by ANOVA, with the student's t-test, or a nonparametric test, such as the Mann-Whitney U rank sum test. To look at trends, the Mantel-Haenszel test for trend will be used to assess the correlation between marker expression/frequency and categorized (e.g., low, medium, high) variables. All analyses will be performed with the STATA statistical software package.

Ethnicity and gender are also known risk factors of HCC. The foreseen HCC cases collected in our proposed study will not currently provide statistical power to address the association of these factors with HCC disease and risk due to the necessity of sub-stratification and thus, a direct impact on study power. However, we will attempt to expand our study by pooling our data with that of other studies performed in the United States (US Cohort Consortium, MD Anderson/Boston) to address power for such stratifications and to allow for the identification of allele variants with frequencies lower than 5% (GWAS analysis). If sufficiently powered after cohort pooling, HCC cases, high risk non-cancer cases and controls will be examined separately by race and gender or any other available clinical parameters such as stage or treatment. Dependent upon their size and composition, these additional cohorts may also be used to validate findings from our study.

7.0   HUMAN SUBJECTS PROTECTIONS

7.1   Rationale for Subject Selection:

We will recruit study subjects between the ages of 18 to 90 years. Children and institutionalized patients, are excluded from the proposed study. Very few HCC patients at the 3 participating hospitals are younger than 18 years of age. We will not recruit cases that are older than 90 years of age because of co-morbidity considerations, and the additional difficulty of matching those cases with population-based controls.

The protection of study subjects from research risk will be achieved by several mechanisms. Prior to enrolling subjects, we will obtain approval of the study from the IRB at the NCI, followed by the IRBs at the University of Maryland School of Medicine, Johns Hopkins Medical Institute and the Veterans Administration Office for Research. Written informed consent will be required from the study subjects for participation. The form will state that individual results will not be provided to the participants. Only overall study results and progress can be provided if requested by a subject. According to new HIPAA regulations ("HIPAA Privacy Rule"), a separate written authorization to obtain, use and disclose protected health information for research will be required from the study subjects for participation in the study.

Study subjects' confidentiality will be maintained at all times. Subjects will be assigned unique study ID numbers. These unique study numbers will be linked to the subject's identifier information in a database, and to the hard copy of the identifier sheet. This information will be secured at the University of Maryland School of Medicine at Baltimore. The database has two levels of password security, which will only allow authorized individuals to access the information. A log will automatically record who assessed the information, and what information was obtained. Biological samples will be labeled with the unique study number but no other identifier information. Thus, biological samples and results from the analyses of biological samples can neither be linked back to the participant nor be used in any way to identify the participants.

Identifier information for non-participants will also be recorded to avoid a re-contact. This information will be stored in a database with two levels of password security, which will allow only authorized individuals to access the information. A log will automatically record who assessed the information and what was assessed. Non-participants will be assigned a unique study number. This number will be used for tracking of reasons for non-participation, and for available demographic information. The PI will ensure appropriate IRB review and approval from each site (see Section 9.0).

7.2   Participation of Children:

Children (subjects under 18yrs) are excluded from this study. Children rarely develop HCC. It is a cancer type mainly occuring in adult populations (greater than 18 years).

7.3   Evaluation of Benefits and Risks/Discomforts:

We are proposing a study that is not a clinical investigation. Participating subjects will not receive treatment under the protocol. The only medical procedure is the collection of blood conducted by a nurse or trained phlebotomist for which subjects will be monetarily compensated ($25). Thus, the treatment-related risk under the protocol is minimal. The main potential risk to participants is a loss of confidentiality about their disease or about their susceptibility to develop the disease. There will be no direct benefits for study participants.

STRATEGIES TO MINIMIZE RISK

- Study subjects' confidentiality will be maintained at all times. The guidelines are described in paragraph "Human Subject Protections". Our consent form will also state that individual results will not be provided to the participants.
- All interviewers will be trained personnel. During the study, characteristics such as interview duration, number of interview problems reported, number of refusals, distribution of subject answers, and number of missing and incomplete answers will be evaluated. The reviews are aimed to improve the performance of the interviewers and to minimize the stress that is experienced by an interviewed person.

7.4   Risks/Benefits Analysis:

As described in section 7.3, there are minimal risks to the subjects of this proposed study that do impact on their treatment protocol or significantly affect their health. The small risk associated with blood draws are no greater than those ordinarily associated with routine laboratory tests. It is hoped that the research data will improve our understanding of liver cancer biology with respect to causes of the aggressiveness and poor survival associated with this disease. Our research is also aimed to identify genetic markers that are useful diagnostic and prognostic markers for liver cancer.

- ***For pediatric patients:*** Children (subjects under 18yrs) are excluded from this study.

- ***For cognitively or physically impaired patients:*** These patients will be excluded from this study.

7.5   Consent and Assent Process and Documentation:

CONSENT PROCEDURE FOR CASES

A candidate participant is identified through hospital records and contacts with the Departments of Pathology and Surgery. After a non-objection statement by the treating physician has been obtained, an interviewer will contact the patient to get both informed consent and authorization to obtain, use and disclose protected health information for research. Severely ill patients in the intensive care unit will not be approached. If a patient is found to be unable to give informed consent, the consent procedure and the enrollment into the study will be stopped. The nurse will explain the purpose of the research study to assure that everything is understood. The patient will be given the option that either he/she reads the form by himself/herself or that the nurse will read the form to him/her. The nurse will answer questions. The consent form describes the purpose of the study, procedures, risks and potential discomforts, benefits, and the independence of the quality of medical care from the decision to participate in the study. The form explains the confidentiality of the study, the right to withdraw from the

study at anytime, and the protection of privacy as it relates to genetic testing. Consent for participation in the study has been obtained when the patient and the research study coordinator have both signed 2 identical consent forms. One form will be given to the patient and the original signed consent goes to Medical Records; the original is placed in research record. At the VA, consents will be scanned into the medical record. The proposed consent form for cases is attached to this protocol.

CONSENT PROCEDURE FOR CONTROLS

A candidate participant is identified through a MVA database. After a home phone number is confirmed, an interviewer will contact the subject to verify eligibility and the willingness to participate in the study. If the two criteria are fulfilled, the interviewer will schedule a meeting at a location that is convenient for the subject. If the subject is found to be unable to give informed consent, the consent procedure and the enrollment into the study will be stopped. The interviewer will explain the purpose of the research study to assure that everything is understood. The patient will be given the option that either he/she reads the form by himself/herself or that the interviewer will read the form to him/her. The interviewer will answer questions. The consent form describes the purpose of the study, procedures, risks and potential discomforts, benefits, and the independence of the quality of medical care from the decision to participate in the study. The form explains the confidentiality of the study, the right to withdraw at anytime, and the protection of privacy as it relates to genetic testing. Consent for participation in the study has been obtained when the patient and the research study coordinator have both signed 2 identical consent forms. One form will be given to the patient and copy placed in research record. The proposed consent forms for population-based is attached to the protocol.

8.0   SAFETY REPORTING REQUIREMENTS/DATA AND SAFETY MONITORING PLAN

8.1   Adverse Event Definitions

Administration of questionnaires and blood draws are considered minimal risk and adverse events are therefore not anticipated. Should they occur adverse events will be reported to the NCI IRB within 7 days of notification of the adverse event. The NCI PI is responsible for reporting to the NCI IRB events that occur in participants enrolled at all participating sites and that meet the requirements for expedited reporting within 7 days of notification of the adverse event.

8.1.1 Adverse Event

An adverse event is defined as any reaction, side effect, or untoward event that occurs during the course of the clinical trial, whether or not the event is considered related to the treatment or clinically significant. For this study, AEs

will include events reported by the patient, as well as clinically significant abnormal findings on physical examination or laboratory evaluation. A new illness, symptom, sign or clinically significant laboratory abnormality or worsening of a pre-existing condition or abnormality is considered an AE. All AEs must be recorded on the AE case report form.

All AEs, including clinically significant abnormal findings on laboratory evaluations, regardless of severity, will be followed until satisfactory resolution. AEs should be reported up to 30 days following the last dose of study drug. AEs that are considered treatment related, expected, continuing, but not resolvable by 30 days after treatment completion (e.g., alopecia) will not be followed after the 30-day period.

8.1.2 Serious Adverse Event (SAE)

Serious adverse drug experience (or SAE): Any adverse drug experience occurring during any study phase (treatment or follow-up) and at any dose that results in any of the following outcomes:
• Death,
• A life-threatening adverse drug experience
• Inpatient hospitalization or prolongation of existing hospitalization
• A persistent or significant disability/incapacity
• A congenital anomaly/birth defect.
• Important medical events that may not result in death, be life-threatening, or require hospitalization may be considered a serious adverse drug experience when, based upon appropriate medical judgment, they may jeopardize the patient or subject and may require medical or surgical intervention to prevent one of the outcomes listed in this definition.

Disability: A substantial disruption of a person's ability to conduct normal life functions.
Life-threatening adverse drug experience: Any adverse drug experience that places the patient or subject, in the view of the investigator, at immediate risk of death from the reaction as it occurred, i.e., it does not include a reaction that had it occurred in a more severe form, might have caused death.

8.1.3 Unexpected adverse drug experience:

Any adverse drug experience, the specificity or severity of which is not consistent with the current **investigator brochure**. ''Unexpected,'' as used in this definition, refers to an adverse drug experience that has not been previously observed (e.g., included in the investigator brochure) rather than from the perspective of such experience not being anticipated from the pharmacological properties of the pharmaceutical product.

8.2   NCI-IRB Adverse Event Reporting

8.2.1 NCI-IRB Expedited Reporting of and Adverse Events and Deaths

The Protocol PI will report to the NCI-IRB:

- All serious adverse events (SAEs) that are **not** in the consent form, but are possibly, probably or definitely related to the research. A SAE is defined as an untoward medical occurrence that
  - resulted in a death;
  - was life-threatening;
  - required or prolonged hospitalization;
  - caused persistent or significant disability/incapacity;
  - resulted in congenital anomalies or birth defects; or
  - required intervention to prevent permanent impairment or death.
- All other deaths not included in the SAE category above, except deaths due to progressive disease.
- All deaths that occur within 30 days of the last dose of study drug or treatment, except deaths due to progressive disease.
- All grade 3 and 4 (CTCAE) events that are not in the consent and that are possibly, probably or definitely related to the research.

Reports must be received by the NCI-IRB within 7 working days via iRIS.

8.2. 2 NCI-IRB Requirements for PI Reporting of Expected (In Consent) and Unexpected (Not in Consent) Adverse Events at Continuing Review

The protocol PI will report to the NCI-IRB:

- All Grade 2 events that are not in the consent form, but are possibly, probably or definitely related to the research;
- All Grade 3 and 4 events that are possibly, probably or definitely related to the research;
- All Grade 5 events regardless of attribution;
- All Serious Events regardless of attribution.
- **NOTE**: Grade 1 events are not required to be reported.

8.2.3 NCI IRB Reporting of IND Safety Reports
Only IND Safety Reports that require a sponsor recommended change to the protocol or the consent form or in the opinion of the PI increases risks to study participants will need to be reported to the NCI IRB.

8.3 IND Sponsor, FDA, and the NIH Office and the NIH Office of Biotechnology Activities (OBA) Reporting Criteria

Not applicable for this study

Investigator-Sponsor Expedited Adverse Event Reporting Criteria

Not applicable for this study

IND Safety Reports to the FDA (Refer to 21 CFR 312.32)

Not applicable for this study

FDA Annual Reports (Refer to 21 CFR 312.33)

Not applicable for this study

Expedited Adverse Event Reporting Criteria to the IND Manufacturer

Not applicable for this study

8.4 Data and Safety Monitoring Plan

The resource contractor at University of Maryland School of Medicine will conduct the study and is responsible for data collection and storage. The collected data will be monitored as follows. The contractor is required to quality assess the questionnaire data, as described under 5.1. A small random sample of subjects will be re-contacted to inquire about their interview experience, to note any problems, and to confirm that the interview actually took place. The PI, Dr. Xin Wei Wang, a Staff Scientist, Dr. Anuradha Budhu and the NCI Project Officer, Dr. Glenwood Trivers, together with the consulting epidemiologists under the contract, Dr. Christopher Loffredo and Dr. Stefan Ambs, will monitor the study through weekly conference calls, regular meetings, and an annual site visit. The contractor will be required to provide a written monthly update, a six-month report, and a year-end report. The data that are particularly monitored are the accrual rates, participation rates, age, race and gender distribution, the status of frequency matching, and the number and stage distribution of fresh-frozen tumor specimens. The PI, the associate investigators, and the consulting epidemiologists, will thoroughly review the study once per year at the annual site visit for the contract.

The resource contractor will collect all data that pertain to the proposed study. The data will be stored in a central database at the University of Maryland School of Medicine, as described under 5.1. The contractor will provide data summaries in the semi-annual and annual reports. The database can be queried upon request from NCI investigators. The contractor will provide query results and statistical analysis. The contractor may also provide data in a file format to the NCI researcher for review and analysis at NCI.

9.0 MULTI-INSTITUTIONAL GUIDELINES:

Our protocol is a multi-institutional protocol for which the NCI is the coordinating center. The proposed study will be conducted through a service

contract to our laboratory. The contractor is the University of Maryland School of Medicine Pathology Department. The University of Maryland School of Medicine, the Veterans Affairs Medical Center, Baltimore and Johns Hopkins Medical Institute University are participating hospitals. The MPA number of 2 of the participating hospitals (University of Maryland School of Medicine and the Veterans Affairs Medical Center, Baltimore) is M1174. Contact addresses at the participating hospitals are provided on page two of this protocol.

The PI of the study, Dr. Xin Wei Wang, will seek Institutional Review Board (IRB) approval of the study from the NCI IRB, followed by IRBs at the University of Maryland School of Medicine and the Veterans Administration Office for Research and Johns Hopkins Medical Institute University. He will provide the NCI IRB with a copy of the participating institution's IRB approval prior to enrolling subjects. He will also provide the NCI IRB with a copy of the participating institution's approved yearly continuing review and with copies of all amendments, consents and approvals from each participating institution.

9.1   IRB Approvals:
The PI will provide the NCI IRB and Central Registration Office with a copy of the participating institution's approved yearly continuing review. Registration will be halted at any participating institution in which a current continuing approval is not on file at the NCI IRB.

9.2   Amendments and Consents:
The CCR PI will provide the NCI IRB with copies of all amendments, consents and approvals from each participating institution.

9.3   Data and Specimen Collection Procedures:
The PI will provide specific guidelines for quality assurance, data collection and format, and data receipt by the coordinating institution (recommend at least quarterly). It is recommended that data collection forms/system be consistent for all institutions. All adverse events from participating institutions must be submitted to the NCI IRB according to current NCI IRB policy.

9.4   NCI Guidance for Reporting Serious Adverse Events for Multi-Center Trials:
The reporting requirements for adverse events in multi-center trials when the NCI PI is responsible for the research and the coordination of the other research sites is the same as with any NCI intramural research protocol. Serious adverse event reports should be submitted from the participating centers along with a PI assessment of the event in the same time frame as adverse event reports that occur at the NCI. The review of the event should be submitted when available. If the NCI is a participating site, rather than the responsible site, then the serious adverse event report should be submitted per protocol or as IND Safety Reports if generated from the sponsor.

9.5 **Data Center Audits:**

The resource contractor at University of Maryland School of Medicine will conduct the study and is responsible for data collection and storage. The PI, Dr. Xin Wei Wang, a Staff Scientist, Dr. Anuradha Budhu and the NCI Project Officer, Dr. Glenwood Trivers, together with the consulting epidemiologists under the contract, Dr. Christopher Loffredo and Dr. Stefan Ambs, will monitor the study through weekly conference calls, regular meetings, and an annual site visit. The contractor will be required to provide a written monthly update, a six-month report, and a year-end report including accrual rates, participation rates, age, race and gender distribution, the status of frequency matching, and the number and stage distribution of fresh-frozen tumor specimens. The PI, the associate investigators, and the consulting epidemiologists, will thoroughly review the study once per year at the annual site visit for the contract.

10.0 **PHARMACEUTICAL AND INVESTIGATIONAL DEVICE INFORMATION**
Not applicable for this study

10.1 **Source:** Not applicable for this study

10.2 **Toxicity:** Not applicable for this study

10.3 **Formulation and preparation:** Not applicable for this study

10.4 **Stability and storage:** Not applicable for this study

10.5 **Administration procedures:** Not applicable for this study

10.6 **Incompatibilities:** Not applicable for this study

11.0 REFERENCES

1. Parkin DM, Pisani P, Ferlay J. Global cancer statistics. CA Cancer J Clin 1999;49:33-64, 2.

2. El-Serag HB, Mason AC. Rising incidence of hepatocellular carcinoma in the United States. N Engl J Med 1999;340:745-50.

3. Taylor-Robinson SD, Foster GR, Arora S, Hargreaves S, Thomas HC. Increase in primary liver cancer in the UK, 1979-94. Lancet 1997;350:1142-3.

4. Carr BI, Flickinger JC, Lotze MT. Hepatobiliary Cancers: Cancer of the Liver. In: DeVita JrVT, Hellman S, Rosenberg SA, editors. Cancer Principles & Practice of Oncology. Philadelphia: Lippincott-Raven; 1997. p.1087-114.

5. Curley SA, Izzo F, Gallipoli A, de Bellis M, Cremona F, Parisi V. Identification and screening of 416 patients with chronic hepatitis at high risk to develop hepatocellular cancer. Ann Surg 1995;222:375-80.

6. Nakakura EK, Choti MA. Management of hepatocellular carcinoma. Oncology (Huntingt) 2000;14:1085-98.

7. Thorgeirsson SS, Grisham JW. Molecular pathogenesis of human hepatocellular carcinoma. Nat Genet 2002;31:339-46.

8. Kim JW, Wang XW. Gene expression profiling of preneoplastic liver disease and liver cancer: a new era for improved early detection and treatment of these deadly diseases? Carcinogenesis 2003;24:363-9.

9. Llovet JM, Schwartz M, Mazzaferro V. Resection and liver transplantation for hepatocellular carcinoma. Semin Liver Dis 2005;25:181-200.

10. Kensler TW, Qian GS, Chen JG, Groopman JD. Translational strategies for cancer prevention in liver. Nat Rev Cancer 2003;3:321-9.

11. Forman D, Bennett B, Stafford J, Doll R. Exposure to vinyl chloride and angiosarcoma of the liver: a report of the register of cases. Br J Ind Med 1985;42:750-3.

12. Powell LW, Burt MJ, Halliday JW, Jazwinska EC. Hemochromatosis: genetics and pathogenesis. Semin Liver Dis 1996;16:55-63.

13. Berman JJ. Cell proliferation and the aetiology of hepatocellular carcinoma. J Hepatol 1988;7:305-9.

14. Huang WS, Liao LY, Wang CS, Chen PH. Hepatocellular carcinoma presenting with acquired porphyria: a case report and review of the literature. Chang Keng I Hsueh Tsa Chih 1999;22:111-6.

15. Elzouki AN, Eriksson S. Risk of hepatobiliary disease in adults with severe alpha 1-antitrypsin deficiency (PiZZ): is chronic viral hepatitis B or C an additional risk factor for cirrhosis and hepatocellular carcinoma? Eur J Gastroenterol Hepatol 1996;8:989-94.

16. Hussain SP, Schwank J, Staib F, Wang XW, Harris CC. TP53 mutations and hepatocellular carcinoma: insights into the etiology and pathogenesis of liver cancer. Oncogene 2007;26:2166-76.

17. Staib F, Hussain SP, Hofseth LJ, Wang XW, Harris CC. TP53 and liver carcinogenesis. Hum Mutat 2003;21:201-16.

18. Koike K, Tsutsumi T, Fujie H, Shintani Y, Kyoji M. Molecular mechanism of viral hepatocarcinogenesis. Oncology 2002;62 Suppl 1:29-37.

19. Levy L, Renard CA, Wei Y, Buendia MA. Genetic alterations and oncogenic pathways in hepatocellular carcinoma. Ann N Y Acad Sci 2002;963:21-36.

20. Feitelson MA, Sun B, Satiroglu Tufan NL, Liu J, Pan J, Lian Z. Genetic mechanisms of hepatocarcinogenesis. Oncogene 2002;21:2593-604.

21. Taketa K. Alpha-fetoprotein: reevaluation in hepatology. Hepatology 1990;12:1420-32.

22. Okuda H, Nakanishi T, Takatsu K et al. Measurement of serum levels of des-gamma-carboxy prothrombin in patients with hepatocellular carcinoma by a revised enzyme immunoassay kit with increased sensitivity. Cancer 1999;85:812-8.

23. Shiraki K, Takase K, Tameda Y, Hamada M, Kosaka Y, Nakano T. A clinical study of lectin-reactive alpha-fetoprotein as an early indicator of hepatocellular carcinoma in the follow-up of cirrhotic patients. Hepatology 1995;22:802-7.

24. Li D, Mallory T, Satomura S. AFP-L3: a new generation of tumor marker for hepatocellular carcinoma. Clin Chim Acta 2001;313:15-9.

25. Yuen MF, Cheng CC, Lauder IJ, Lam SK, Ooi CG, Lai CL. Early detection of hepatocellular carcinoma increases the chance of treatment: Hong Kong experience. Hepatology 2000;31:330-5.

26. Zhou XD, Tang ZY, Yang BH et al. Experience of 1000 patients who underwent hepatectomy for small hepatocellular carcinoma. Cancer 2001;91:1479-86.

27. Yamamoto J, Okada S, Shimada K et al. Treatment strategy for small hepatocellular carcinoma: comparison of long-term results after percutaneous ethanol injection therapy and surgical resection. Hepatology 2001;34:707-13.

28. Poon RT, Fan ST, Lo CM, Liu CL, Wong J. Long-term survival and pattern of recurrence after resection of small hepatocellular carcinoma in patients with preserved liver function: implications for a strategy of salvage transplantation. Ann Surg 2002;235:373-82.

29. Wildi S, Pestalozzi BC, McCormack L, Clavien PA. Critical evaluation of the different staging systems for hepatocellular carcinoma. Br J Surg 2004;91:400-8.

30. Cillo U, Bassanello M, Vitale A et al. The critical issue of hepatocellular carcinoma prognostic classification: which is the best tool available? J Hepatol 2004;40:124-31.

31. Ueno S, Tanabe G, Nuruki K et al. Prognostic performance of the new classification of primary liver cancer of Japan (4th edition) for patients with hepatocellular carcinoma: a validation analysis. Hepatol Res 2002;24:395-403.

32. Chambers AF, Groom AC, MacDonald IC. Dissemination and growth of cancer cells in metastatic sites. Nat Rev Cancer 2002;2:563-72.

33. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000;100:57-70.

34. Steeg PS. Metastasis suppressors alter the signal transduction of cancer cells. Nat Rev Cancer 2003;3:55-63.

35. Liotta LA. Mechanisms of cancer invasion and metastasis. Important Adv Oncol 198528-41.

36. Bernards R, Weinberg RA. A progression puzzle. Nature 2002;418:823.

37. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 1995;270:467-70.

38. Van de Vijver MJ, He YD, Van't Veer LJ et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 2002;347:1999-2009.

39. Ramaswamy S, Ross KN, Lander ES, Golub TR. A molecular signature of metastasis in primary solid tumors. Nat Genet 2003;33:49-54.

40. Ye QH, Qin LX, Forgues M et al. Predicting hepatitis B virus-positive metastatic hepatocellular carcinomas using gene expression profiling and supervised machine learning. Nat Med 2003;9:416-23.

41. Calin GA, Ferracin M, Cimmino A et al. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. N Engl J Med 2005;353:1793-801.

42. Calin GA, Liu CG, Sevignani C et al. MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. Proc Natl Acad Sci U S A 2004;101:11755-60.

43. Lu J, Getz G, Miska EA et al. MicroRNA expression profiles classify human cancers. Nature 2005;435:834-8.

44. Lee RC, Ambros V. An extensive class of small RNAs in Caenorhabditis elegans. Science 2001;294:862-4.

45. Lau NC, Lim LP, Weinstein EG, Bartel DP. An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science 2001;294:858-62.

46. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of novel genes coding for small expressed RNAs. Science 2001;294:853-8.

47. Yi R, Qin Y, Macara IG, Cullen BR. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev 2003;17:3011-6.

48. Gregory RI, Shiekhattar R. MicroRNA biogenesis and cancer. Cancer Res 2005;65:3509-12.

49. Lee Y, Jeon K, Lee JT, Kim S, Kim VN. MicroRNA maturation: stepwise processing and subcellular localization. EMBO J 2002;21:4663-70.

50. Hutvagner G, Zamore PD. A microRNA in a multiple-turnover RNAi enzyme complex. Science 2002;297:2056-60.

51. Metzler M, Wilda M, Busch K, Viehmann S, Borkhardt A. High expression of precursor microRNA-155/BIC RNA in children with Burkitt lymphoma. Genes Chromosomes Cancer 2004;39:167-9.

52. Takamizawa J, Konishi H, Yanagisawa K et al. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. Cancer Res 2004;64:3753-6.

53. Michael MZ, O' Connor SM, Holst Pellekaan NG, Young GP, James RJ. Reduced accumulation of specific microRNAs in colorectal neoplasia. Mol Cancer Res 2003;1:882-91.

54. Iorio MV, Ferracin M, Liu CG et al. MicroRNA gene expression deregulation in human breast cancer. Cancer Res 2005;65:7065-70.

55. Calin GA, Dumitru CD, Shimizu M et al. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A 2002;99:15524-9.

56. Sonoki T, Iwanaga E, Mitsuya H, Asou N. Insertion of microRNA-125b-1, a human homologue of lin-4, into a rearranged immunoglobulin heavy chain gene locus in a patient with precursor B-cell acute lymphoblastic leukemia. Leukemia 2005;19:2009-10.

57. Chan JA, Krichevsky AM, Kosik KS. MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. Cancer Res 2005;65:6029-33.

58. Johnson SM, Grosshans H, Shingara J et al. RAS is regulated by the let-7 microRNA family. Cell 2005;120:635-47.

59. Hayashita Y, Osada H, Tatematsu Y et al. A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. Cancer Res 2005;65:9628-32.

60. Meng F, Henson R, Lang M et al. Involvement of human micro-RNA in growth and response to chemotherapy in human cholangiocarcinoma cell lines. Gastroenterology 2006;130:2113-29.

61. Volinia S, Calin GA, Liu CG et al. A microRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci U S A 2006;103:2257-61.

62. Ambros V. MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. Cell 2003;113:673-6.

63. Yanaihara N, Caplen N, Bowman E et al. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. Cancer Cell 2006;9:189-98.

64. Budhu A, Jia HL, Forgues M et al. Identification of metastasis-related microRNAs in hepatocellular carcinoma. Hepatology 2008;47:897-907.

65. Hunter K, Welch DR, Liu ET. Genetic background is an important determinant of metastatic potential. Nat Genet 2003;34:23-4.

66. Paget S. The distribution of secondary growths in cancer of the breast. 1889. Cancer Metastasis Rev 1989;8:98-101.

67. Mueller MM, Fusenig NE. Friends or foes - bipolar effects of the tumour stroma in cancer. Nat Rev Cancer 2004;4:839-49.

68. Budhu A, Forgues M, Ye QH et al. Prediction of venous metastases, recurrence and prognosis in hepatocellular carcinoma based on a unique immune response signature of the liver microenvironment. Cancer Cell 2006;10:99-111.

69. Horikawa I, Barrett JC. cis-Activation of the human telomerase gene (hTERT) by the hepatitis B virus genome. J Natl Cancer Inst 2001;93:1171-3.

70. Hytiroglou P, Theise ND. Telomerase activation in human hepatocarcinogenesis. Am J Gastroenterol 2006;101:839-41.

71. Chami M, Gozuacik D, Saigo K et al. Hepatitis B virus-related insertional mutagenesis implicates SERCA1 gene in the control of apoptosis. Oncogene 2000;19:2877-86.

72. Murakami Y, Saigo K, Takashima H et al. Large scaled analysis of hepatitis B virus (HBV) DNA integration in HBV related hepatocellular carcinomas. Gut 2005;54:1162-8.

73. Feitelson MA, Lee J. Hepatitis B virus integration, fragile sites, and hepatocarcinogenesis. Cancer Lett 2006.

74. Kim CM, Koike K, Saito I, Miyamura T, Jay G. HBx gene of hepatitis B virus induces liver cancer in transgenic mice. Nature 1991;351:317-20.

75. Zheng Y, Chen WL, Louie SG, Yen TS, Ou JH. Hepatitis B virus promotes hepatocarcinogenesis in transgenic mice. Hepatology 2007;45:16-21.

76. Wang XW, Hussain SP, Huo TI et al. Molecular pathogenesis of human hepatocellular carcinoma. Toxicology 2002;181-182:43-7.

77. Forgues M, Difilippantonio MJ, Linke SP et al. Involvement of Crm1 in hepatitis B virus X protein-induced aberrant centriole replication and abnormal mitotic spindles. Mol Cell Biol 2003;23:5282-92.

78. Wang W, Budhu A, Forgues M, Wang XW. Temporal and spatial control of nucleophosmin by the Ran-Crm1 complex in centrosome duplication. Nat Cell Biol 2005;7:823-30.

79. Nagai H, Pineau P, Tiollais P, Buendia MA, Dejean A. Comprehensive allelotyping of human hepatocellular carcinoma. Oncogene 1997;14:2927-33.

80. Boige V, Laurent-Puig P, Fouchet P et al. Concerted nonsyntenic allelic losses in hyperploid hepatocellular carcinoma as determined by a high-resolution allelotype. Cancer Res 1997;57:1986-90.

81. Kim YJ, Lee HS. Single nucleotide polymorphisms associated with hepatocellular carcinoma in patients with chronic hepatitis B virus infection. Intervirology 2005;48:10-5.

82. Midorikawa Y, Tsutsumi S, Nishimura K et al. Distinct chromosomal bias of gene expression signatures in the progression of hepatocellular carcinoma. Cancer Res 2004;64:7263-70.

83. Ohsawa N, Sakamoto M, Saito T, Kobayashi M, Hirohashi S. Numerical chromosome aberrations in hepatocellular carcinoma detected by fluorescence in situ hybridization. J Hepatol 1996;25:655-62.

84. Tamura S, Nakamori S, Kuroki T et al. Association of cumulative allelic losses with tumor aggressiveness in hepatocellular carcinoma. J Hepatol 1997;27:669-76.

85. Bando K, Nagai H, Matsumoto S et al. Identification of a 1-cM region of common deletion on 4q35 associated with progression of hepatocellular carcinoma. Genes Chromosomes Cancer 1999;25:284-9.

86. Nishida N, Fukuda Y, Kokuryu H et al. Accumulation of allelic loss on arms of chromosomes 13q, 16q and 17p in the advanced stages of human hepatocellular carcinoma. Int J Cancer 1992;51:862-8.

87. Tsopanomichalou M, Kouroumalis E, Ergazaki M, Spandidos DA. Loss of heterozygosity and microsatellite instability in human non-neoplastic hepatic lesions. Liver 1999;19:305-11.

88. Yamada T, De Souza AT, Finkelstein S, Jirtle RL. Loss of the gene encoding mannose 6-phosphate/insulin-like growth factor II receptor is an early event in liver carcinogenesis. Proc Natl Acad Sci U S A 1997;94:10351-5.

89. Yeh SH, Chen PJ, Shau WY et al. Chromosomal allelic imbalance evolving from liver cirrhosis to hepatocellular carcinoma. Gastroenterology 2001;121:699-709.

90. Roncalli M, Bianchi P, Grimaldi GC et al. Fractional allelic loss in non-end-stage cirrhosis: correlations with hepatocellular carcinoma development during follow-up. Hepatology 2000;31:846-50.

91. Kanai Y, Ushijima S, Tsuda H, Sakamoto M, Hirohashi S. Aberrant DNA methylation precedes loss of heterozygosity on chromosome 16 in chronic hepatitis and liver cirrhosis. Cancer Lett 2000;148:73-80.

92. Kondo Y, Kanai Y, Sakamoto M, Mizokami M, Ueda R, Hirohashi S. Genetic instability and aberrant DNA methylation in chronic hepatitis and cirrhosis--A comprehensive study of loss of heterozygosity and microsatellite instability at 39 loci and DNA hypermethylation on 8 CpG islands in microdissected specimens from patients with hepatocellular carcinoma. Hepatology 2000;32:970-9.

93. Herath NI, Leggett BA, MacDonald GA. Review of genetic and epigenetic alterations in hepatocarcinogenesis. J Gastroenterol Hepatol 2006;21:15-21.

94. Shields PG, Harris CC. Cancer risk and low-penetrance susceptibility genes in gene-environment interactions. J Clin Oncol 2000;18:2309-15.

95. Williams JA. Single nucleotide polymorphisms, metabolic activation and environmental carcinogenesis: why molecular epidemiologists should think about enzyme expression. Carcinogenesis 2001;22:209-14.

96. Goode EL, Ulrich CM, Potter JD. Polymorphisms in DNA repair genes and associations with cancer risk. Cancer Epidemiol Biomarkers Prev 2002;11:1513-30.

97. Brawley OW. Prostate cancer and black men. Semin Urol Oncol 1998;16:184-6.

98. Altekruse SF, McGlynn KA, Reichman ME. Hepatocellular Carcinoma Incidence, Mortality, and Survival Trends in the United States From 1975 to 2005. J Clin Oncol 2009.

99. Di Bisceglie AM, Lyra AC, Schwartz M et al. Hepatitis C-related hepatocellular carcinoma in the United States: influence of ethnic status. Am J Gastroenterol 2003;98:2060-3.

100. Deuffic S, Poynard T, Buffat L, Valleron AJ. Trends in primary liver cancer. Lancet 1998;351:214-5.

101. Yang L, Parkin DM, Ferlay J, Li L, Chen Y. Estimates of cancer incidence in China for 2000 and projections for 2005. Cancer Epidemiol Biomarkers Prev 2005;14:243-50.

102. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin 2005;55:74-108.

103. Tang ZY. Hepatocellular carcinoma-Cause, treatment and metastasis. World J Gastroenterol 2001;7:445-54.

104. Yuki K, Hirohashi S, Sakamoto M, Kanai T, Shimosato Y. Growth and spread of hepatocellular carcinoma. A review of 240 consecutive autopsy cases. Cancer 1990;66:2174-9.

105. Iizuka N, Oka M, Yamada-Okabe H et al. Oligonucleotide microarray for prediction of early intrahepatic recurrence of hepatocellular carcinoma after curative resection. Lancet 2003;361:923-9.

106. Lee JS, Chu IS, Heo J et al. Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. Hepatology 2004;40:667-76.

107. Pollard JW. Tumour-educated macrophages promote tumour progression and metastasis. Nat Rev Cancer 2004;4:71-8.

108. Branda M, Wands JR. Signal transduction cascades and hepatitis B and C related hepatocellular carcinoma. Hepatology 2006;43:891-902.

109. Segditsas S, Tomlinson I. Colorectal cancer and genetic alterations in the Wnt pathway. Oncogene 2006;25:7531-7.

110. Lee HC, Kim M, Wands JR. Wnt/Frizzled signaling in hepatocellular carcinoma. Front Biosci 2006;11:1901-15.

111. Ding SJ, Li Y, Tan YX et al. From proteomic analysis to clinical significance: overexpression of cytokeratin 19 correlates with hepatocellular carcinoma metastasis. Mol Cell Proteomics 2004;3:73-81.

112. Boyault S, Rickman DS, de Reynies A et al. Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. Hepatology 2007;45:42-52.

113. Laurent-Puig P, Legoix P, Bluteau O et al. Genetic alterations associated with hepatocellular carcinomas define distinct pathways of hepatocarcinogenesis. Gastroenterology 2001;120:1763-73.

114. Lee JS, Heo J, Libbrecht L et al. A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. Nat Med 2006;12:410-6.

115. Yamashita T, Forgues M, Wang W et al. EpCAM and alpha-fetoprotein expression defines novel prognostic subtypes of hepatocellular carcinoma. Cancer Res 2008;68:1451-61.

116. Yamashita T, Budhu A, Forgues M, Wang XW. Activation of hepatic stem cell marker EpCAM by Wnt-ß-catenin signaling in hepatocellular carcinoma. Cancer Research 2007;67:10831-9.

117. Perou CM, Sorlie T, Eisen MB et al. Molecular portraits of human breast tumours. Nature 2000;406:747-52.

118. Henshall SM, Afar DE, Hiller J et al. Survival analysis of genome-wide gene expression profiles of prostate cancers identifies new prognostic targets of disease relapse. Cancer Res 2003;63:4196-203.

119. Seike M, Yanaihara N, Bowman ED et al. A cytokine gene signature of the lung adenocarcinoma and its tissue environment predicts prognosis. J Natl Cancer Inst 2007;99:1257-69.

120. Fialkow PJ. Clonal origin of human tumors. Biochim Biophys Acta 1976;458:283-321.

121. Heppner GH. Tumor heterogeneity. Cancer Res 1984;44:2259-65.

122. Jordan CT, Guzman ML, Noble M. Cancer stem cells. N Engl J Med 2006;355:1253-61.

123. Clarke MF, Dick JE, Dirks PB et al. Cancer stem cells--perspectives on current status and future directions: AACR Workshop on cancer stem cells. Cancer Res 2006;66:9339-44.

124. Wicha MS, Liu S, Dontu G. Cancer stem cells: an old idea--a paradigm shift. Cancer Res 2006;66:1883-90.

125. Al Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A 2003;100:3983-8.

126. Singh SK, Hawkins C, Clarke ID et al. Identification of human brain tumour initiating cells. Nature 2004;432:396-401.

127. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat Med 1997;3:730-7.

128. Ricci-Vitiani L, Lombardi DG, Pilozzi E et al. Identification and expansion of human colon-cancer-initiating cells. Nature 2007;445:111-5.

129. O'Brien CA, Pollett A, Gallinger S, Dick JE. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. Nature 2007;445:106-10.

130. Ma S, Chan KW, Hu L et al. Identification and characterization of tumorigenic liver cancer stem/progenitor cells. Gastroenterology 2007;132:2542-56.

131. Sigal SH, Brill S, Fiorino AS, Reid LM. The liver as a stem cell and lineage system. Am J Physiol 1992;263:G139-G148.

132. Schmelzer E, Wauthier E, Reid LM. The Phenotypes of Pluripotent Human Hepatic Progenitors. Stem Cells 2006;24:1852-8.

133. Schmelzer E, Zhang L, Bruce A et al. Human hepatic stem cells from fetal and postnatal donors. J Exp Med 2007;204:1973-87.

134. Dan YY, Riehle KJ, Lazaro C et al. Isolation of multipotent progenitor cells from human fetal liver capable of differentiating into liver and mesenchymal lineages. Proc Natl Acad Sci U S A 2006;103:9912-7.

135. Yang ZF, Ho DW, Ng MN et al. Significance of CD90(+) Cancer Stem Cells in Human Liver Cancer. Cancer Cell 2008;13:153-66.

136. Zaret KS. Regulatory phases of early liver development: paradigms of organogenesis. Nat Rev Genet 2002;3:499-512.

137. Shafritz DA, Oertel M, Menthena A, Nierhoff D, Dabeva MD. Liver stem cells and prospects for liver reconstitution by transplanted cells. Hepatology 2006;43:S89-S98.

138. Reya T, Clevers H. Wnt signalling in stem cells and cancer. Nature 2005;434:843-50.

139. Yamashita T, Ji J, Budhu A et al. EpCAM-positive hepatocellular carcinoma cells are tumor-initiating cells with stem/progenitor cell features. Gastroenterology 2009;136:1012-24.

140. Ji J, Yamashita T, Budhu A et al. Identification of microRNA-181 by genome-wide screening as a critical player in EpCAM-positive hepatic cancer stem cells. Hepatology 2009;50:472-80.

141. Dean M, Fojo T, Bates S. Tumour stem cells and drug resistance. Nat Rev Cancer 2005;5:275-84.

142. Rich JN. Cancer stem cells in radiation resistance. Cancer Res 2007;67:8980-4.

143. Cimmino A, Calin GA, Fabbri M et al. miR-15 and miR-16 induce apoptosis by targeting BCL2. Proc Natl Acad Sci U S A 2005;102:13944-9.

144. Fleischhacker M, Beinert T, Ermitsch M et al. Detection of amplifiable messenger RNA in the serum of patients with lung cancer. Ann N Y Acad Sci 2001;945:179-88.

145. Kopreski MS, Benko FA, Kwak LW, Gocke CD. Detection of tumor messenger RNA in the serum of patients with malignant melanoma. Clin Cancer Res 1999;5:1961-5.

146. El Hefnawy T, Raja S, Kelly L et al. Characterization of amplifiable, circulating RNA in plasma and its potential as a tool for cancer diagnostics. Clin Chem 2004;50:564-73.

147. Chen XQ, Bonnefoi H, Pelte MF et al. Telomerase RNA as a detection marker in the serum of breast cancer patients. Clin Cancer Res 2000;6:3823-6.

148. Kaddurah-Daouk R, Kristal BS, Weinshilboum RM. Metabolomics: a global biochemical approach to drug response and disease. Annu Rev Pharmacol Toxicol 2008;48:653-83.

149. Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics. Mass Spectrom Rev 2006.

150. Kind T, Tolstikov V, Fiehn O, Weiss RH. A comprehensive urinary metabolomic approach for identifying kidney cancerr. Anal Biochem 2007;363:185-95.

151. Kim YS, Maruvada P, Milner JA. Metabolomics in biomarker discovery: future uses for cancer prevention. Future Oncol 2008;4:93-102.

152. Singh R, Kaul R, Kaul A, Khan K. A comparative review of HLA associations with hepatitis B and C viral infections across global populations. World J Gastroenterol 2007;13:1770-87.

153. Calvisi DF, Ladu S, Gorden A et al. Mechanistic and prognostic significance of aberrant methylation in the molecular pathogenesis of human hepatocellular carcinoma. J Clin Invest 2007;117:2713-22.

154. Yamada S, Nomoto S, Fujii T et al. Frequent promoter methylation of M-cadherin in hepatocellular carcinoma is associated with poor prognosis. Anticancer Res 2007;27:2269-74.

155. Jeng JE, Tsai JF, Chuang LY et al. Tumor necrosis factor-alpha 308.2 polymorphism is associated with advanced hepatic fibrosis and higher risk for hepatocellular carcinoma. Neoplasia 2007;9:987-92.

156. Ezzikouri S, El Feydi AE, Chafik A et al. The Pro variant of the p53 codon 72 polymorphism is associated with hepatocellular carcinoma in Moroccan population. Hepatol Res 2007;37:748-54.

157. Tseng LH, Lin MT, Shau WY et al. Correlation of interleukin-10 gene haplotype with hepatocellular carcinoma in Taiwan. Tissue Antigens 2006;67:127-33.

158. Ho SY, Wang YJ, Chen HL et al. Increased risk of developing hepatocellular carcinoma associated with carriage of the TNF2 allele of the -308 tumor necrosis factor-alpha promoter gene. Cancer Causes Control 2004;15:657-63.

159. Wang Y, Kato N, Hoshida Y et al. Interleukin-1beta gene polymorphisms associated with hepatocellular carcinoma in hepatitis C virus infection. Hepatology 2003;37:65-71.

160. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 2003;31:e15.

161. Reiner A, Yekutieli D, Benjamini Y. Identifying differentially expressed genes using false discovery rate controlling procedures. Bioinformatics 2003;19:368-75.

APPENDICES

1) INFORMED CONSENT DOCUMENTS (ICD)
2) LIVER CANCER STUDY QUESTIONNARE

3) MD LUNG & PROSTATE CANCER STUDY & MULTI-ORGAN STUDIES QUESTIONNARE
4) ELIGIBILITY CRITERIA CHECKLISTS FOR HCC CASES, HIGH RISK NON-CANCER CASES AND POPULATION-BASED CONTROLS