DEPARTMENT OF MICROBIOLOGY, TUMOR AND CELL BIOLOGY (MTC)

Karolinska Institutet, Stockholm, Sweden

NOTI GENOME SCANNING TO IDENTIFY UNKNOWN CANCER ASSOCIATED GENES IN MAJOR HUMAN EPITHELIAL MALIGNANCIES

Klas Haraldson

Stockholm 2010

Published by Karolinska Institutet. Printed by Universitetsservice, US-AB.

ISBN 978-91-7409-281-3

Natura nullibi magis perfecta, artificiosa magis quam in minimis.

Carl von Linné

To my parents

ABSTRACT

Epithelial cancers cause many deaths every year. Changes in the genes of human chromosome 3 are particularly common in epithelial cancers in several organs. Alterations in DNA methylation is one of the best known epigenetic changes in cancer. The abnormal epigenetic landscape of the cancer cell is characterized by a massive genomic hypomethylation and hypermethylation of CpG islands in the promoter regions of tumor suppressor genes.

Microarrays is a powerful tool for studying the molecular basis of diseases that are not possible with conventional methods. Being able to predict who will develop cancer and how the disease will behave and respond to treatment after diagnosis are some uses for this technology.

NotI microarrays is a novel technology that makes it possible to simultaneously detect changes in methylation, amplification and deletions in cancer. The NotI microarrays technology based on restriction enzyme NotI is methylation sensitive, and therefore makes it possible to detect if methylation causes a gene to be altered in cancer. Cancer samples are hybridized with NotI microarrays and are evaluated bioinformatically. Using NotI microarray technology, 181 NotI loci in human chromosome 3 have been analyzed for methylation, amplification and deletions in different epithelial cancers; lung cancer, renal cancer, breast cancer, cervical cancer, ovarian cancer, prostate cancer and colon cancer. The genes that were shown to be altered in cancer samples compared with normal tissue samples were analyzed further with more accurate methods; bisulfite sequencing of cloned PCR products was the method used to control and verify methylation status. NotI loci that were found to have relevant changes include genes *MINT24, BHLHB2, RPL15, RARbeta1, ITGA9, RBSP3, VHL, ZIC4, NKIRAS1, LRRC3B*, suggesting that they probably are involved in cancer development.

It was found that *NPRL2/G21* gene has growth inhibitory activity for renal and lung cell lines when tested under controlled physiologic conditions of gene expression both in vitro and in vivo. Mutations were found in experimental tumors and intragenic homozygous deletions in renal, lung, and other cancer cell lines. It was also observed that *NPRL2* could participate in mismatch repair.

Further, it was shown that somatic hypermutations in tumor suppressor genes involved in major human malignancies is a mechanism for the development, progression and spread of cancer. *RASSF1* and *RBSP3* were demonstrated to have a high incidence of somatic hypermutations in several cancer types.

Methylation and/or deletions of *LRRC3B* gene were detected in more than 50% of checked renal, cervical, ovarian, colon and breast cancer samples, and it was found to inhibit renal cell cancer cell line KRC/Y cell survival and replication in vitro.

Altered genes found with the NotI microarrays technology may potentially be used as biomarkers or targets for epigenetic therapy. A set of 18 markers was suggested (*BHLHB2, FBLN2, FLJ44898* (*EPHB1*), *GATA2, GORASP1, Hmm210782* (*PRICKLE2*), *Hmm61490, ITGA9, LOC285205, LRRC3B, MINA, MITF, MRPS17P3, NKIRAS1, PLCL2, TRH, UBE2E2, WNT7A*) that allow to discriminate/diagnose different types of lung cancer.

LIST OF PUBLICATIONS AND MANUSCRIPTS

I. Functional Characterization of the Candidate Tumor Suppressor Gene *NPRL2/G21* Located in 3p21.3C.
II. High Mutability of the Tumor Suppressor Genes *RASSF1* and *RBSP3* (*CTDSPL*) in Cancer.
III. Epigenetic analysis of non small cell lung cancer (NSCLC) using NotI microarrays.
IV. *LRRC3B* gene is frequently epigenetically inactivated in several epithelial malignancies and inhibit cell growth and replication.

CONTENTS

LIST OF ABBREVIATIONS

AP20	Alu-PCR clone region at the 3p21.3 telomeric border
ADC	Adenocarcinoma
BSP	Bisulfite sequencing PCR
Chr. 3p	Short arm of chromosome 3
Chr. 3q	Long arm of chromosome 3
CODE	Cloning of deleted sequences
CpG	Cytosine and guanine separated by a phosphate
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
EST	Expressed sequence tag
HD	Homozygous deletion
LOH	Loss of heterozygosity
LUCA	Lung cancer candidate genes region at the 3p21.3 centromeric border
MSP	Methylation specific PCR
NMA	NotI microarray
NSCLC	Non small cell lung carcinoma
PCR	Polymerase chain reaction
RT-PCR	Reverse transcription PCR
SNP	Single nucleotide polymorphism
SCC	Squamous cell carcinoma
TSG	Tumor suppressor gene

INTRODUCTION

1. General introduction to cancer

Cancer is a collective term for a large group of diseases that can affect every part of the body. One characteristic of cancer is the rapid creation of abnormal cells that grow beyond their usual boundaries, and which can then invade adjoining parts of the body and spread to other organs. This process is called metastasis. Metastases are the major cause of death from cancer. Cancer accounts for approximately 13% of all deaths worldwide. The main types of cancer leading to overall cancer mortality every year are lung, stomach, colorectal, liver and breast. The most common types of cancer worldwide (in order of the number of global deaths) are lung, stomach, liver, colon, esophagus, and prostate cancer for men, and breast, lung, stomach, colorectal and cervical cancer for women.

Cancer arises from one single cell. The transformation from a normal cell into a tumor cell is a multistage process, typically a progression from a precancerous lesion to malignant tumors. These changes are the result of the interaction between a person's genetic factors, and different types of external agents, including: physical carcinogens such as UV and ionizing radiation; chemical carcinogens such as asbestos, components of tobacco smoke, aflatoxin (a food contaminant) and arsenic (a drinking water contaminant); biological carcinogens, such as infections caused by certain viruses, bacteria or parasites. Some examples of infections are associated with certain types of cancer: viruses: hepatitis B and liver cancer, Human Papilloma Virus (HPV) and cervical cancer and human immunodeficiency virus (HIV) and Kaposi's sarcoma; bacteria: Helicobacter pylori and gastric cancer; parasites: schistosomiasis and bladder cancer. Aging is another important factor for the development of cancer. The incidence of cancer rises dramatically with age, most likely due to an accumulation of risks for specific types of cancer that increase with age. The cumulative risk is combined with the tendency for cellular repair mechanisms to be less effective as a person grows older. Tobacco use, alcohol consumption, low intake of fruits and vegetables, and chronic infections from hepatitis B (HBV), hepatitis C virus (HCV) and some types of human papilloma virus (HPV) are leading risk factors for cancer in low- and middle-income countries. In high-income countries, tobacco use, alcohol consumption, and being overweight or obese are major risk factors for cancer (World Health Organization, 2009).

In recent decades cancer research has generated a huge amount of knowledge, revealing cancer to be a disease involving dynamic changes in the genome. Several lines of evidence

indicate that tumorigenesis in humans is a multistep process and that these steps reflect genetic alterations that drive the progressive transformation of normal human cells into highly malignant derivatives. Many types of cancer are diagnosed in the human population with an age-dependent incidence that implicates four to seven rate-limiting, stochastic events. It has been argued that tumor development proceeds through an evolutionary process in which a series of genetic changes, each conferring one or another type of growth advantage, leads to the gradual conversion of normal human cells into cancer cells.

There are more than 100 different types of cancer, and subtypes of tumors can be found within specific organs. It was suggested that the vast catalog of cancer cell genotypes is a manifestation of six essential alterations in cell physiology, which together dictate malignant growth: self-sufficiency in growth signals, insensitivity to growth-inhibitory signals, evasion of programmed cell death (apoptosis), limitless replicative potential, sustained angiogenesis, and tissue invasion and metastasis (Hanahan and Weinberg, 2000).

The multistage progression model of colorectal tumorigenesis from adenoma to carcinoma has been well defined (Fig 1), which indicates that development of sporadic forms of the colorectal cancer requires six steps and different classes of interacting cancer causing genes including TSGs, oncogenes and mutator genes are involved (Fearon and Vogelstein, 1990; Armitage and Doll, 1954).

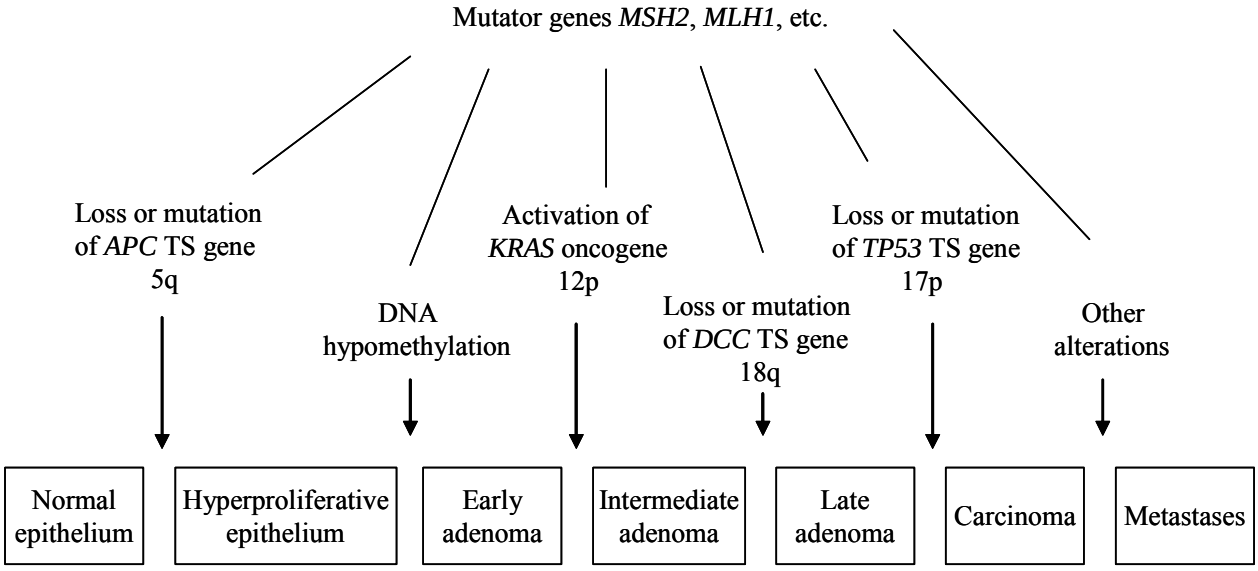

Figure 1. Fearon and Vogelstein's model for the development of colorectal cancer.

2. Tumor suppressor genes

Tumor suppressor genes are defined as recessive genes, i.e. they must sustain mutations or deletions of both alleles in order to contribute to cancer formation and progression. This definition implies that one functional allele of the tumor suppressor gene is sufficient for normal cell function. Patients with familial cancers frequently inherit one normal and one abnormal allele of the tumor suppressor gene from their parents. If the second, normal allele is lost, the protective effect of the gene product no longer exists. Therefore, introduction of a wild-type copy of the gene back into the tumor should inhibit further tumor growth.

2.1 Knudson's two-hit hypothesis

Knudson's two-hit hypothesis suggests that the mutation of one allele and the loss of the other (the loss of heterozygosity) give rise to cancer (Mungall, 2002). Mutations abnormally enhance function of oncogenes, or may lead to TSG losing function (Herman and Baylin, 2003). It was suggested that the two-hit hypothesis should be extended to include epigenetic mechanisms like DNA methylation (Jones and Laird, 1999; Dai et al., 2002) (Fig 2).

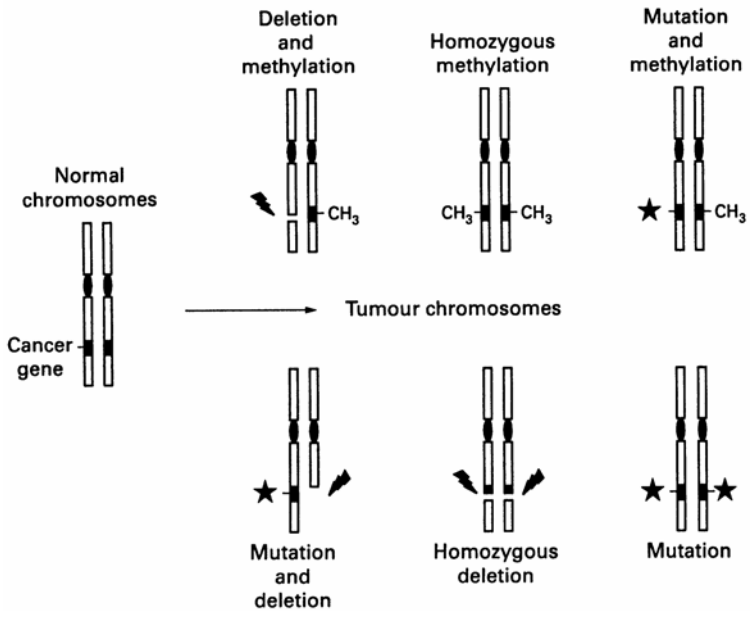

Figure 2. The combination of genetic and epigenetic events in cancer provides a mechanism for complete inactivation of both allelic locations (Costello and Christoph, 2001).

2.2 The tumor suppressor pRb

The first tumor suppressor gene identified was the *Rb* gene, which is associated with the childhood illness of retinoblastoma (Knudson, 1971). In an epidemiological study, Knudson and colleagues noticed that bilateral retinoblastoma occurred frequently within the same family, whereas unilateral retinoblastoma did not appear to be a genetically inherited disease. In families with bilateral retinoblastoma, karyotyping techniques were used to detect homozygous loss of chromosome 13q, a defect that was transmitted to offspring. Homozygous loss was found to be necessary but not sufficient for the formation of retinoblastoma, since not every family member with the loss of both alleles developed the disease. Later, the gene responsible for development of the disease was cloned and termed *Rb* for retinoblastoma. Reintroduction of this gene into cultured retinoblastoma tumor cells reversed the malignant phenotype, suggesting that the gene was indeed a tumor suppressor (Bookstein et al., 1989).

2.3 The tumor suppressor p53

The p53 protein is involved in sensing DNA damage and regulating cell death (Marx, 1993). In normal cells, when DNA damage is sensed by p53, the cell cycle is arrested to permit DNA repair. Upon completion of this process, the cell progresses through the mitotic cycle. If repair fails to occur, p53 initiates the process of apoptosis, or programmed cell death. Thus, normal cells with genetic defects die. If p53 is not present in the cell (via gene deletion) or is mutated to be nonfunctional, DNA damage is not repaired, and the cell progresses through the cell cycle, transmitting its damaged DNA to its progeny. p53 is so important to the maintenance of 'healthy' DNA that it is mutated or deleted in over 70% of human cancers, including osteosarcomas, rhabdomyosarcomas and carcinomas of the breast, colon, lung and prostate. Since the cloning of *Rb*, many other tumor suppressor genes have been identified (Table 1).

Table 1. Tumor suppressor genes and their function and associated cancers.

Name	**Protein function in normal cells**	**Associated cancers**
p53	Cell cycle regulator	Colon and others
BRCA1	Cell cycle regulator, genomic integrity and chromatin structure	Breast, ovarian, prostate and others
BRCA2	Genomic integrity	Breast, ovarian, prostate and others
PTEN	Tyrosine and lipid phosphatase	Prostate, glioblastomas
APC	Cell adhesion	Colon
DCC	Cell adhesion	Colon
MCC	Undetermined	Colon

p16-INK4A	Cell cycle regulator	Colon and others
MLH1	Mismatch repair	Colon and gastric cancers
MSH2	Mismatch repair	Colon and gastric cancers
DPC4	Cell death regulator	Pancreatic
Wt1	Cell death regulator	Wilms' tumor
NF1	Regulator of GTPases	Astrocytomas
NF2	Cell adhesion	Astrocytomas
VHL	Ubiquitination	Renal
PTC	Regulator of hedgehog signalling	Thyroid
TSC2	Cell cycle regulator	Breast and renal
TSG101	Cell cycle regulator	Renal and leukaemia

In normal cells, products of tumor suppressor genes have been shown to regulate negatively cell growth and proliferation. For example, the *Rb* gene product sequesters transcription factors that are required for normal cell cycle progression. The ability of Rb to function as a block to cell cycle progression is regulated by phosphorylation of the Rb protein on multiple serine residues (Harbour and Dean, 2000). In quiescence, Rb is hypo- or under-phosphorylated and binds members of the E2F transcription factor family. Upon growth factor stimulation of the cell, Rb becomes phosphorylated by cyclin-regulated kinases and releases E2F, which then induces gene transcription events necessary for cell division. In cancers, deletion or inactivation of Rb results in constitutively 'free' E2F, which in turn leads to unfettered gene transcription and oncogenic transformation. Similarly, as described above, p53 has been shown to sense DNA damage, cause cell cycle arrest, regulate transcription and stimulate apoptotic cell death pathways in normal cells (Marx, 1993). Loss of this function increases the chance of damaged DNA being transmitted to subsequent generations of cells. The exact role of BRCA1 is still unclear, but studies using mice that lack the *BRCA1* gene show that it is essential for cellular proliferation during early embryonic development (Zheng et al., 2000). BRCA1 may also regulate transcriptional events, since it is capable of acting as a coactivator of p53 and a corepressor of c-Myc. Recent studies also implicate a role for BRCA1 in chromatin remodelling, which is required for DNA transcriptional and replication events. PTEN regulates the phosphorylation status of phospholipids that are involved in regulating apoptotic pathways within the cell (Di Cristofano and Pandolfi, 2000). Taken together, these findings indicate that tumor suppressor gene products act by negatively controlling cell growth in normal cells and that their loss contributes to the unregulated cell growth seen in tumor cells.

2.4 Tumor suppressor genes on chromosome 3, especially on 3p21.3

It has been known for decades that cells can become malignant, after parts of chromosomes, which inhibit cell division, are eliminated. Deletions in tumor cells is therefore a first indication of potential sites for TSGs. The probability that a deleted region would contain a TSG is also supported, if it is located on a chromosome fragment that suppresses tumor formation when introduced into tumor cells.

Microcell hybrids that contain fragments of chromosome 3p have been designed and screened for tumorigenicity in athymic nude mice, and hybrid clones showed strong tumor suppression. The clones contained a 2-megabase fragment of human chromosome material in the region 3p21. Specific genes on 3p have been proposed to act as TSGs either because they are located in a common deleted region, have an obviously decreased expression or incidence of abnormal transcript, their ability to suppress tumorigenicity after transfection into tumor cells, the putative function of gene products, or a combination of several of these criteria (Kok et al., 1997; Killary et al., 1992). A careful mapping of homozygous deletions in the LUCA region has set the smallest HD region in 3p21.3C between markers D3S1568 and D3S4604. The smallest region of homozygous deletions in the AP20 region is located between markers D3S1298 and D3S3623 (Senchenko et al., 2004). The approximately 630-kb sized HD region of lung cancer has been examined in detail. The combination of molecular manual methods and computational techniques has made it possible to identify, isolate, characterize and describe the genes in this region (Lerman and Minna, 2000; Wei et al., 1996). Many of these genes have various growth inhibiting characteristics, but none of them is disabled by coding mutations in their remaining allele as in Knudson's two hit hypothesis. For some of these 3p21.3 TSGs the inactivation mechanism is epigenetic suppression by hypermethylation of the gene promoter (Oh et al., 2008). A summary of tumor suppressor genes on chromosome 3p is shown in table 2.

Table 2. Tumor suppressor genes on chromosome 3p.

Gene	Location (#)	Protein function in normal cells (*)	Associated Cancers (#)
ADAMTS9	3p14.3-p14.2	Cleaves the large aggregating proteoglycans, aggrecan and versican.	Nasopharyngeal, esophageal, prostate
BLU/ ZMYND10	3p21.3	-	Glioma, nasopharyngeal, lung, neuroblastoma
CACNA2D2	3p21.3	Subunit of voltage-dependent calcium channels.	Lung
DLEC1	3p22-p21.3	Inhibits cell proliferation.	Colon, gastric, lung, hepatocellular, nasopharyngeal,

			ovarian
FHIT	3p14.2	Involved in purine metabolism.	Lung, pancreatobiliary, gallbladder, esophageal, cervical, breast, prostate, hepatocellular, cholangiocarcinoma, renal, penile squamous cell, colorectal, gastric, leukemia, oral, bladder, AML, Burkitt's lymphoma, endometrial, diffuse large B-cell lymphoma, head and neck, anal, gastrointestinal, testicular germ cell, malignant mesothelioma, thyroid
FUS1/ TUSC2	3p21.3	Inhibits colony formation, causing G1 arrest and ultimately inducing apoptosis.	Lung
HYAL1	3p21.3-p21.2	May block the TGFB1-enhanced cell growth. Intracellularly degrade hyaluronan.	Breast, ovarian, lung, prostate, glioma, head and neck, laryngeal
HYAL2	3p21.3	Hydrolyzes high molecular weight hyaluronic acid.	Lung, glioma, brain
HYA22/ RBSP3/ CTDSPL	3p21.3	May function as a phosphatase involved in the regulation of cell growth and differentiation.	Lung, breast
ITGA9	3p21.3	Receptor for VCAM1, cytotactin and osteopontin.	Lung
LIMD1	3p21.3	Binds pRB and represses E2F-driven transcription.	Lung, breast
LRIG1	3p14	Act as a feedback negative regulator of signaling by receptor tyrosine kinases	Breast, malignant glioma, colorectal, cervical, renal
LTF	3p21.31	Serine protease of the peptidase S60 family that cuts arginine rich regions. Iron-binding protein in milk and body secretions with an antimicrobial activity. Anti-inflammatory activity. Regulation of cellular growth and differentiation.	Prostate, nasopharyngeal, renal, lung, breast
MLH1	3p21.3	Post-replicative DNA mismatch repair, implicated in DNA damage signaling, role in meiosis.	Hereditary nonpolyposis colorectal, gastric, endometrial, pancreatic, lung, ovarian, oral, renal, esophageal, prostate, leukemia, laryngeal, vulvar, melanoma, head and neck, lymphoma, stomach, testicular germ cell, AML, sarcoma, gallbladder, small bowel, breast
NPRL2/ G21/TUSC4	3p21.3	Suppresses Src-dependent tyrosine phosphorylation and activation of PDPK1 and its downstream signaling. Suppresses cell growth.	Hepatocellular, lung
PLCD1	3p22-p21.3	Trophoblast and placental development. Signal transducer that generates 2 second	Gastric, esophageal, breast

		messengers, diacylglycerol and inositol 1,4,5-trisphosphate, by hydrolyzing inositol phospholipids.	
PTPRG	3p21-p14	Regulate cell growth, differentiation, mitotic cycle receptor-type protein tyrosine phosphatase.	Gastric, breast, lymphoma
RARB	3p24	Receptor for retinoic acid. Controls cell function by directly regulating gene expression.	Lung, breast, nasopharyngeal, esophageal, endometrial, cervical, renal, ovarian, hepatocellular, prostate, papilloma, leukemia, head and neck, colorectal, bladder, gastric, AML, glioma
RASSF1	3p21.3	Required for death receptor-dependent apoptosis. Inhibition of APC activity and mitotic progression. Inhibits proliferation by negatively regulating cell cycle progression at the level of G1/S-phase transition by regulating accumulation of cyclin D1 protein.	Breast, multiple myeloma, ovarian, nasopharyngeal, Ewing sarcoma, Merkel cell carcinoma, lung, oral, renal, prostate, glioma, gastric, neuroblastoma, endometrial, hepatocellular, cervical, penile, bladder, mesothelioma, colorectal, Wilms tumors, gallbladder, melanoma, small bowel, meningioma, cholangiocarcinoma, Hodgkin's lymphoma, retinoblastoma, esophageal, thyroid, head and neck, osteosarcoma
ROBO1	3p12	Integral membrane protein that functions in axon guidance and neuronal precursor cell migration.	Head and neck, hepatocellular, lung, colorectal, medulloblastoma, glioma
SEMA3B	3p21.3	Inhibits axonal extension.	Prostate, renal, ovarian, colorectal, lung
SEMA3F	3p21.3	May play a role in cell motility and cell adhesion.	Prostate, lung
VHL	3p26-p25	Involved in the ubiquitination and subsequent proteasomal degradation. Involved in transcriptional repression.	Renal, pheochromocytoma, multiple myeloma, head and neck paraganglioma, acute myeloid leukaemia, tongue, pancreatic, breast, lung, oral, esophageal, meningioma

* Genecards database Feb 2010. Available at: http://www.genecards.org/

NCBI database "Gene" Feb 2010. Available at: http://www.ncbi.nlm.nih.gov/gene/

2.5 Examples of chromosome 3p tumor suppressor genes

NPRL2/G21 located in LUCA region codes for a soluble protein that consists of a protein binding domain, similar to the core domain of MutS, and a nitrogen permease regulator 2 domain. NPRL2/G21 suppresses tumor cell growth in vitro and in vivo. It is thought that NPRL2/G21 is involved in the regulation of the cell cycle and apoptosis and in DNA repair.

RASSF1 is a known tumor suppressor gene of the LUCA region and codes for several mRNAs. Isoform A is a major transcript and codes for a cytoplasmic protein whose C-terminal

region harbors a Ras association domain and a SARAH domain, which regulates the homo- and heterotypic interactions of SARAH-containing proteins. These functional domains indicate that RASSF1A plays an important role in signal transduction from the cell surface to the nucleus. In addition, RASSF1 is involved in regulating the cell cycle. Cells with a lower *RASSF1A* expression have a lower content of cyclin D1, one of the main regulators of the cell cycle. In complex with cyclin D1, cyclin-dependent kinases hyperphosphorylate and thereby inactivate pRB1, which prevents its binding with transcription factors and causes cell transformation.

RBSP3/CTDSPL from AP20 is expressed to yield two mRNA isoforms, A and B. Isoform B differs from isoform A in lacking one exon. The protein product of this gene belongs to the family of CTD (C-terminal domain) small serine phosphatases, which dephosphorylate serine residues in the C-terminal domain of the large RNA polymerase II subunit and other proteins. RBSP3/CTDSPL suppresses the growth of tumor cells both in vitro and in vivo. It is thought that RBSP3/CTDSPL dephosphorylates pRB1 and thus contributes to the negative regulation of the cell cycle.

ITGA9 is located in region AP20 and codes for a subunit of a membrane protein that is related to integrins, cell surface glycoproteins that are involved in cell–cell and cell–matrix adhesion. This gene encodes an alpha integrin. The protein encoded by this gene, when bound to the beta 1 chain, forms an integrin that is a receptor for VCAM1, cytotactin and osteopontin (Anedchenko et al., 2008).

3. Epigenetics

Epigenetic processes control the normal packaging and function of the human genome and contribute to pathological conditions. Three different mechanisms appear to play an important role in the initiation and maintenance of epigenetic modifications: DNA methylation, modification of histones (Fig 3) and RNA-associated silencing. Abnormal changes in DNA methylation (global hypomethylation and CpG island hypermethylation) were among the first events to be discovered (Ballestar and Esteller, 2005). Some studies over the past decades have shown that DNA methylation is involved in imprinting, gene regulation, chromatin structure, genome stability and diseases, especially cancer (Novik et al., 2002; Jones and Martienssen, 2005).

Increased expression of DNA-methyltransferase is a common characteristic of immortalized cell lines and human tumors. The addition of a methyl group to cytosine is mediated by three active DNMTs (DNMT1, DNMT-3a and DNMT-3b) (Esteller, 2008).

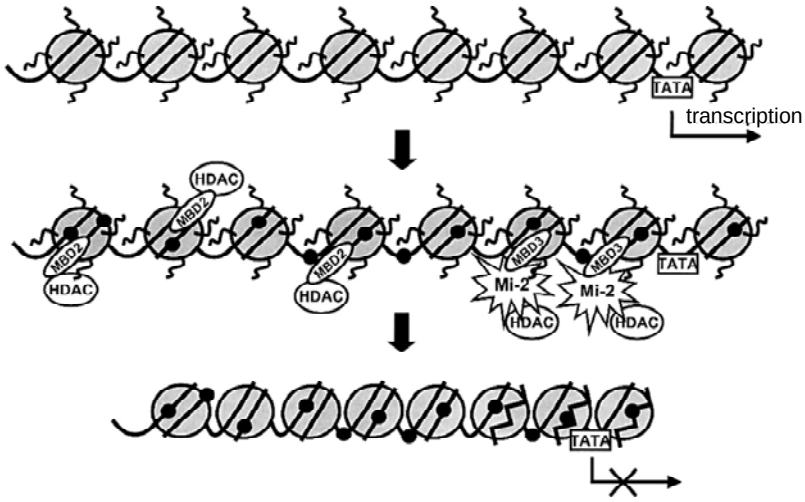

Figure 3. Epigenetic mechanisms: DNA methylation and histone acetylation.

3.1 CpG islands and methylation

Clusters of unmethylated CpG dinucleotides (better known as CpG islands) exist in the promoter and exon sequences of about 40% of mammalian genes, while other parts of the genomes of mammals contain few CpG dinucleotides, which are largely methylated. CpG islands were originally defined as regions of 200 bases or more with a (G + C)-content of at least 50% and a ratio of observed to expected CpG frequency of at least 0.6. A more strict definition of CpG islands, the Takai-Jones criteria, gives a better relation of CpG islands with 5' regions of genes and excludes most Alu repeats (Takai and Jones, 2002). The abnormal

epigenetic landscape of the cancer cell is characterized by a massive genomic hypomethylation and hypermethylation of CpG island promoters of TSGs (Esteller, 2006). Hypermethylation of the CpG-island promoter can affect genes involved in the cell cycle, DNA repair, the metabolism of carcinogens, cell-to-cell interaction, apoptosis and angiogenesis, all of which are involved in the development of cancer (Esteller, 2007) (Fig 4).

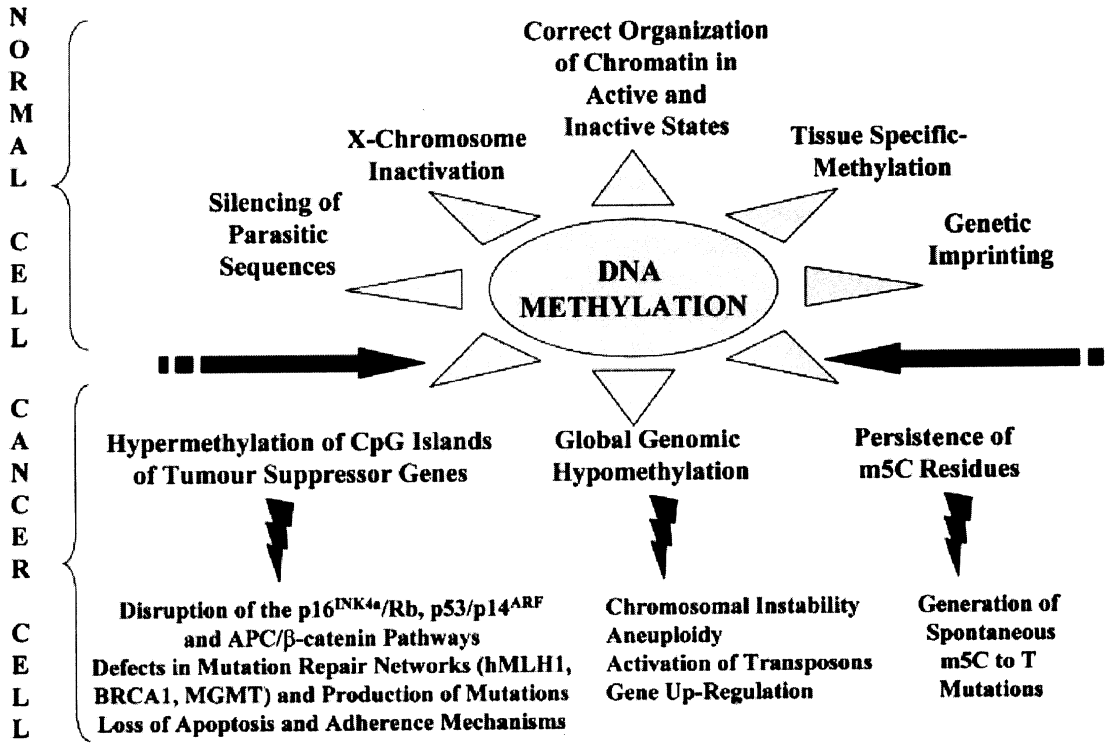

Figure 4. DNA methylation at the centre of the normal and malignant behavior of the cell (Esteller and Herman, 2002).

Almost half of all tumor suppressor genes are methylated in their promoter regions, preventing gene transcription (Baylin, 1997). Abnormally high levels of methylation appear in cancer cells that have a loss in the *p21/WAF1* gene. In normal cells, p21 protein negatively regulates the ability of DNA-methyltransferase to add a methyl group to CpG islands, thereby protecting these sites in the DNA from methylation. Inactivation or loss of p21 allows these sites to be methylated and transcriptionally silenced. Some tumor suppressor genes shown to be methylated in tumors include *BRCA1*, *VHL*, and *p16INK4A*.

3.2 Environmental influence on methylation

The organism is constantly exposed to environmental stress caused by air pollution, lifestyle and working conditions. The toxic components of the environment such as nickel, cadmium, arsenic, aflatoxin, ionizing radiation, tobacco smoke and different infectious agents (HPV, HBV, and Helicobacter pylori) are suspected to have a potential effect on DNA methylation, chromatin organization and histones (Herceg, 2007).

3.3 Approaches to determine DNA methylation status

The main approaches for determination of CpG dinucleotides methylation status include: bisulfite sequencing, bisulfite pyrosequencing, MS-REA (methylation-sensitive restriction endonuclease assay), MSP (methylation-specific PCR), COBRA (combined bisulfite restriction analysis), MS-SnuPE (methylation-sensitive single nucleotide primer extension), MS-SSCA (methylation-sensitive single-strand conformation analysis), MethyLight, HeavyMethyl, MALDI-TOF MS (matrix-assisted laser desorption/ionization time-to-flight mass spectrometry), FMCA (fluorescence melting curve analysis), restriction genome scanning, use of microarrays for determining genes with increased expression after DNA methylation inhibition, DMH (differential methylation hybridization), use of capability of methylCpG-binding domain of MeCP2 protein to bind methylated DNA, immunoprecipitation of methylated DNA with antibodies specific for methylated cytosines (Kvasha, 2008).

3.4 DNA methylation diagnostics and therapy

Several studies have examined methylation status for a number of known cancer genes in a certain type of cancer. Defining methylation profiles may be useful for the development of marker panels for early detection or prediction of the risk of cancer precursors (Wang et al., 2008a). Extensive profiles of DNA changes that characterize all forms of human cancer would provide valuable insight into the mechanisms behind the development of each tumor type and will provide molecular markers, which may radically improve the detection of cancer (Esteller et al., 2001).

Genetic analysis has shown that cell-free circulating DNA in plasma or serum of cancer patients have similar genetic changes found in the corresponding tumor, allowing for the detection of mutations and microsatellite changes with the help of biomarkers (Esteller et al., 1999; Ramirez et al., 2003). DNA methylation can easily be detected in DNA derived from neoplastic and pre-neoplastic lesions in serum, urine and sputum, and promoter

hypermethylation as biomarkers in various body fluids for early detection of cancer has been proposed for various types of cancer (Worm and Guldberg, 2002).

Epigenetic therapy attempts to reverse the abnormalities by the use of natural and synthetic compounds, molecules that are active on specific epigenetic targets (Mai and Altucci, 2009). Different classes of chemical compounds including nucleotide analogues, adenosine analogues, aminobenzoic derivatives, polyphenols, hydrazines, phthalides, disulfides and antisenses have been detected and evaluated as potential drugs (Yu and Wang, 2008).

4. Somatic hypermutation

Cancer is characterized by an accumulation of somatic mutations. The five to seven causative mutations estimated to accumulate in many cancers can be explained not only by the frequency of several independent mutational events. Accumulation of mutations in certain cancers can be triggered by a mutator phenotype, i.e. a high frequency of mutations in tumors because of frequent genetic defects or deficiencies in DNA repair. It has been shown that so-called mutation showers exist and that they constitute at least 0.2%, perhaps 1% or more of the mutational events in a specific transgenic mouse system (Wang et al., 2007).

In another study, it was shown that cancer ESTs have greater variations than normal ESTs, and these varations can not be explained by known and putative SNPs. In addition, these EST variations in cancer were not random, but could be determined by the composition of the substituted base and the composition of the closely located bases (Brulliard et al., 2007).

Moreover, a high incidence of RASSF1A mutation in primary NPC tumor tissues have been identified, including transitions, transversions and deletions. Most of these mutations lead to amino acid changes (Pan et al., 2005).

It has further been shown that an abnormal hypermutation activity affect various loci, in large-cell lymphomas (DLCLs), which are tumors derived from germinal centers. These mutations are located in the 5'-untranslated and coding sequences, are independent of chromosomal translocations and have similar characteristics as the V-region-associated somatic hypermutations of the immune system. In contrast to mutations in the V-regions, however, these mutations are not detectable in normal B cells from germinal centers or in other lymphomas from germinal centers, suggesting a DLCL-associated defect in somatic hypermutation. The hypermutable genes are susceptible to chromosomal translocations in the same area, which is in line with a role for hypermutation in generating translocations by DNA double-strand breaks (Pasqualucci et al., 2001).

5. Principles for the NotI microarrays technology

Microarrays is a powerful tool for investigating the molecular basis for interactions to an extent not possible with conventional methods. The technology will lead to improvements in the development of rational methods of treatment and improvements in diagnosis and prognosis of cancer. Microarrays will be used more in clinical practice in specialized centers and clinics. The ability to predict who will develop cancer and how the disease will behave and respond to treatment after diagnosis is considered one of the potential advantages of this technology (Russo, Zegar and Giordano, 2003).

The restriction endonuclease NotI recognizes the base pair sequence 5'-GCGGCCGC-3' (Lambert et al., 2008). The cleavage and methylation sensitive properties of NotI were described in detail (Kessler and Manta, 1990; Nelson and McClelland, 1991; Roberts and Macelis, 1991).

For genome-wide analysis of methylation, deletions and amplifications, NotI clone microarrays and a genomic subtraction approach may be used, because NotI recognition sites are associated with CpG islands and genes. The genomic subtraction procedure CODE (cloning of deleted sequences), has been adapted to NotI flanking sequences and CpG islands and therefore can be applied to prepare probes from cancer/normal sample pairs for NotI microarrays (Li et al., 2002; Li et al., 2001).

NotI linking clones are clones that contain the regions flanking NotI sites. The relative order of NotI clones and their relationship to known chromosomal markers have been identified. Most of the NotI linking clones are located in CpG islands in the genome. A majority of these clones contain transcribed sequences, which provides an effective link between the physical and functional genetic maps (Allikmets et al., 1994; Talmadge et al., 1995).

The method of preparation of NotI probes of genomic DNA for comparative genomic hybridization has two major advantages. First, the infrequently cutting NotI restrictase, whose recognition sequence includes two CpG dinucleotides. Cytosine methylation in these dinucleotides prevents cleavage of the sequence. Hence, this site, entirely or partially, disappears from the labeled probe in the case of homo- or hemizygous methylation of the NotI recognition sequence. Second, use of the frequently cutting Sau3AI restrictase in combination with NotI yields 100–2000 bp long labeled DNA fragments. The biotinylated NotI linker makes it possible to get rid of DNA fragments not bound to magnetic beads. Therefore, the probe contains mainly short NotI–Sau3A DNA fragments with minor amounts of repeats caused by inevitable procedural inaccuracies. Thus, the probe is enriched in nucleotide sequences

homologous to DNA from NotI-binding clones immobilized on the microarray (Pavlova et al., 2009) (Fig 5).

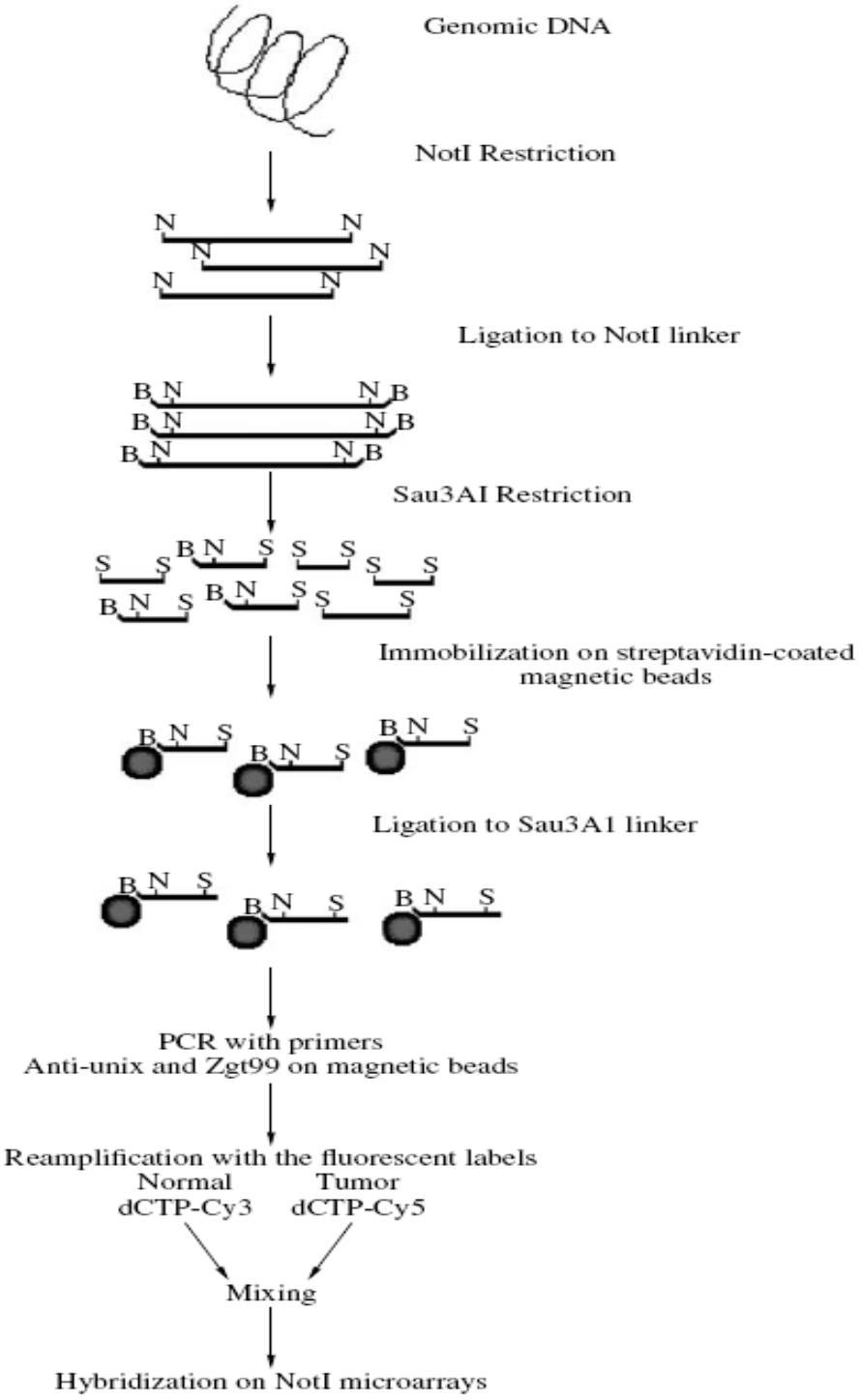

Figure 5. Preparation of DNA probes for hybridization on NotI microarrays. N, NotI site; B, biotin; S, Sau3AI. (Pavlova et al., 2009).

THE SCOPE OF THIS THESIS

The aim of this thesis is to present my results obtained from work with the NotI microarrays technology. Results form this work is presented in paper III and paper IV. I have also participated in work to finely characterize the structure and function of cancer tumor suppressor genes on chromosome 3. Results for 6 of these genes are described in the papers of this thesis: NPRL2/G21 (paper I), RASSF1 (paper II), RBSP3 (paper II and paper III), VHL (paper III), ITGA9 (paper III), and LRRC3B (paper IV).

METHODS

Bisulphite sequencing of cloned PCR products (paper III, paper IV)

Treatment of DNA with bisulfite converts cytosine residues to uracil while leaving 5-methylcytosine residues unaffected (Fig 6). Thus, bisulfite treatment introduces specific changes in the DNA sequence that depend on the methylation status of individual cytosine residues, yielding single-nucleotide resolution information about the methylation status of a DNA segment.

All sites of unmethylated cytosines are displayed as thymines in the resulting amplified sequence of the sense strand, and as adenines in the amplified antisense strand. Bisulphite sequencing requires cloning of PCR products prior to sequencing for adequate sensitivity (Frommer et al., 1992).

To design primers for bisulphite sequencing is a challenge because of the extreme DNA sequence composition after bisulfite modification and the specific requirements for primers and their location on the DNA template. MethPrimer, a program on the basis of Primer 3, is designed for construction of PCR primers for the detection of methylation. A DNA sequence is entered into the program and the program searches the sequence for CpG islands. Primers are constructed around found CpG islands, or in specified regions. MethPrimer can design primers for both BSP and MSP. The results of the primer search is delivered via a Web browser, both in text and graphics (Li and Dahiya, 2002; Li, 2007). For most genes checked, NCBI database "Gene" (http://www.ncbi.nlm.nih.gov/gene) was used to get the molecular location of the gene and then UCSB genome browser (http://genome.ucsc.edu) was used to get the DNA sequence and the RefSeq Genes filter was used to mark the exons. Then the region containing the NotI site(s) was used as input sequence in the MethPrimer program (http://www.urogene.org/methprimer/index1.html) and primers were picked using the bisulphite

sequencing PCR option and optimized by modifying the default search parameters of the program. PCR was performed with the designed primers on samples suggested to be methylated after analyzing the NMA results. PCR products were cloned and at least 6 (to obtain statistical significance) clones of each PCR product were sequenced and the methylation status of the NotI sites was analyzed.

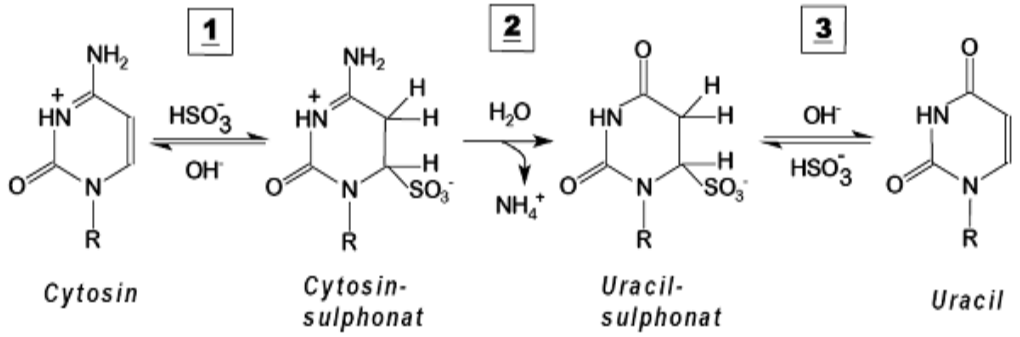

Figure 6. Conversion of unmethylated cytosine to uracil by a treatment with bisulfite.

Growth inhibition experiments with candidate TSGs in vitro (all papers)

The gene to be checked for growth inhibition in vitro is cloned into the episomal tetracycline regulated pETE-Hyg or pETE-Bsd vector and used for transfection of a cancer cell line (KRC/Y and U2020) cells. The empty pETE vector is used as a negative control. The inhibition of colony formation is later compared for the vector-transgene and empty vector. As a positive control a known strong TSG cloned into the pETE-vector can be used.

Growth curve inhibition experiments can then be performed using e.g. KRC/Y cells stably transformed with the gene to be tested. The pETE vectors carrying wild type alleles of the tested gene is transfected into KRC/Y expressing tTA and clonal cell lines are selected. Expression of the transgene can be tested by Northern hybridization. The best tetracycline regulated clones for the gene are then selected. Growth curves are prepared for these, and growth inhibition is observed and compared to the colony assays. All clones expressing the transgene in KRC/Y cells can further be tested with cell proliferation assay based on measurement of cellular DNA content via fluorescent dye binding.

Growth inhibition experiments with candidate TSGs in vivo (paper I)

The gene can be further investigated in vivo by using the gene inactivation test (GIT). The test mimics the inactivation of TSGs during tumor growth in vivo. Clonal cell lines

expressing transgenes are inoculated into SCID mice and the expression of the transgene is controlled by tetracycline administered ad libitum in the drinking water. In this setting TSGs suppress tumor formation in SCID mice, unless they are eliminated or mutated. All grown tumors are analysed for the presence and expression of the transgene. Cell line e.g. KRC/Y derived transgene cell clones are inoculated into SCID mice. For each SCID mouse, $5x10^6$ cells are inoculated (one inoculation per mouse). In control mice, KRC/Y cells transfected with empty pETE vector are used. Half of the SCID mice are then given drinking water containing 1 mg/ml tetracycline. The inhibition of tumor growth is observed for all clones expressing the transgene. All grown tumors are explanted and tested for the presence of pETE-transgene constructs by PCR (Wang et al., 2008b).

Work with cell lines (all papers)

Small cell lung cancer cell lines used include U2020, ACC-LC5, GLC20 and N417. ACC-LC5 carried a deletion in 3p21.3 and GLC20 has a homozygous deletion in LUCA region. Renal clear cell cancer lines used include A498, Caki1, Caki2, KRC/Y, ACHN, TK164, HN4, TK10 and KH-39. Other cell lines used were lymphoblastoid cell line CBMI-Ral-STO, osteosarcoma cell line Saos-2, cervical cancer cell line HeLa and non-small cell lung carcinoma line A549.

Bioinformatics (paper I, paper II)

Publicly available EST sequence were examined for the genes RASSF1A and RBSP3 (paper II). Analysis of EST databases also revealed the occurrence of nonsense and missense mutations in NPRL2 clones from various cancers (paper I). The gene sequences were searched against GenBank EST division, collection of expressed sequence tags. An additional manual refinement against low quality sequences was performed. Nucleotide similarity searches were performed with BLAST 2.2. Probabilities of mutation frequency differences were calculated using Poisson distribution. NPRL2 was also bioinformatically compared to its yeast ortholog and also the functional protein domains were predicted from the DNA sequence.

Deletion mapping (paper I)

Detection of homozygous deletions was performed with single and double PCR. To reduce the effect of normal cell contamination only 28 PCR cycles were used. The primer pairs produced bands with normal control DNA.

Quantitative PCR (paper III)

The reactions described were performed using ABI 7000 PRISM™ SDS (Applied Biosystems) with RQ software (PCR program: 10 min at 95°C, then 40 two-step cycles 15 s at 95°C and 60 s at 60°C) in total volume 25 µl in triplicate. All probes contained the dye FAM at 5'-end and RTQ1 at 3'-end. QPCR data were analyzed using the relative quantification or $\Delta\Delta C_t$-method based on mRNA (or DNA) copy number ratio (R) of target gene versus reference gene in a given tumor sample relative to another reference sample. All preliminary validation steps were done: standardization of all assays, reproducibility of the qPCRs in parallel and in independent runs, selection of reference samples and testing of reference genes.

NotI microarrays (paper III, paper IV)

Microarrays were prepared on 76 × 26 mm glass. One hundred eighty one NotI-linking clones from human chromosome 3 with inserts up to 15 kb were immobilized on the slides in six replications. DNA from *E. coli* was used as negative hybridization control. Plasmid DNA for immobilization on the glasses was isolated and printed on the silanized glasses at a concentration of 0.25 µg/ml.

Microarrays were immersed in the blocking solution (1% bovine serum albumin, 5 × SSC, and 0.1% SDS) at 42°C for 5–7 min and washed twice in water. The DNA on microarrays was denatured by boiling in water for 2 min. The microarrays were then washed with isopropyl alcohol and dried in centrifuge at 1000 rpm in 50 ml tubes (Falcon) for 2 min.

Preparation of NotI probes: Genomic DNA was isolated from tumor tissues with a DNA Extraction kit. DNA samples (1-3 µg) were digested with 10 U of NotI at 37°C for 4hrs. The enzyme was inactivated by three freezing– thawing cycles. Then DNA was ligated to a NotI linker with 5U of T4 DNA ligase at 4°C for 12 hrs. After it DNA was digested with 15U Sau3AI at 37°C for 3 hrs and immobilized on magnetic beads. The unbound excess of DNA was removed by three washings with B&W Buffer. Immobilized DNA samples were ligated to Sau3A linker at 4°C for 12 h with 5U of T4 DNA ligase.

PCR amplification of ligated DNA fragments. PCR was performed in the following reaction mixture: DNA immobilized on magnetic beads (1 µl), primer Anti-Univ (20 µM, 1 µl), primer Zgt99 (20 µM, 1 µl), 5×polymerase buffer (6 µl), 5U of Taq DNA polymerase (1 µl), dNTPs (2 mM, 3 µl, and 17 µl of water). PCR program: start cycle 95°C , 2 min then 35 cycles: 95°C for 45s, 64°C for 40s, and 72°C for 140s; final extension at 72°C for 4 min. Additional

PCR was made for labeling DNA with fluorophores (dNTP-Cy 5 for tumor cells and dNTP-Cy 3 for normal cells). PCR was done as described above but only 1.8 µl of 0.8 mM dCTP was added. Also 1.2 µl of dNTP-Cy 5 for tumor cells and dNTP-Cy 3 for normal cells (1.25 nmol solution) was used.

Oligonucleotides for NotI linker:

NotAntBio: 5'-Biotin-CAGGACTGACCCTTTTGGGACCGC-3';

NotAntComp: 5'-GGCCGCGGTCCCAAAAGGGTCAGTGCTG- 3'.

Oligonucleotides for Sau3A linker:

SauZgt- Block: 5'-GATCCTCAAACGCGT-block3', where block is the propyl group, not recognized by most enzymes during elongation, and

SauZgtComp: 5'-GGCGATCTATCCTAGAGCCCGTACGCGTTTGAG- 3'.

Primers for PCR:

Anti-univ 5'-CAGCACTGACCCTTTTGGGACC- 3' and

Zgt99 5'-GGCGATCTATCCTAGAGCCCGT- 3'.

The amplified DNA was precipitated with ethanol and dissolved in 20 µl of water. All manipulations with labeled DNA were carried out in darkness. Labeled DNA samples were mixed with 10 µl of water, 100 µl of formamide, and 100 µl of 2.5 × hybridization buffer (5 × SSC, 5×Denhardt solution, 1 mM sodium pyrophosphate, and 50 mM Tris HCl pH 7.5). Then DNA samples were denatured by heating at 96°C for 5 min and cooled on ice. Two microliters of 0.1% SDS were added to each sample.

Hybridization of coupled NotI samples was carried out at 42°C for 15 h in a Lucidea Base device. Automatic washing of the microarrays was performed in the same device.

Microarrays were scanned in a GenePix 4000A. The results were processed with GenePix Pro 6.0 software.

Statistical analysis (paper III)

Nonparametric Wilcoxon test was used to compare mRNA expression differences of target genes and reference gene GAPDH in NSCLC samples. Nonparametric Kruskal-Wallis and Mann-Whitney rank-sum tests were used for analysis of genomic DNA copy and methylation changes in groups of SCC and ADC samples with and without metastases. All statistical procedures were performed using the BioStat software.

RESULTS

Paper I

Colony formation assays showed aberrant expression of NPRL2. It was found that the NPRL2/G21 gene has growth inhibitory activity for renal cell carcinoma, small cell lung cancer, and non-small cell lung cancer cell lines when tested under controlled physiological conditions of gene expression both in vitro and in vivo in SCID mice. In addition, mutations were found in experimental tumors and intragenic homozygous deletions were found in renal cell carcinoma, small cell lung cancer, non-small cell lung cancer and other cancer cell lines. Analysis of EST databases also revealed the presence of nonsense and missense mutations in NPRL2 clones obtained from different types of cancer. All these features are consistent with the conclusion that NPRL2/G21 is a multiple tumor suppressor gene. The inactivation or loss of it can lead to the development of breast and cervical cancers.

It was further suggested that the nuclear NPRL2/G21 protein may be involved in mismatch repair and signaling to cell cycle checkpoints that activate apoptotic pathways.

Paper II

This was the first report of high frequencies of somatic mutations in RASSF1A and RBSP3 in various cancers. Suggesting it may underlay the mutator phenotype of cancer, somatic hypermutations in tumor suppressor genes involved in major human malignancies, gives new knowledge about development, spread and progression of cancer.

The frequency of mutations was similar to other reported cases of somatic hypermutations found in three genes in large-cell lymphomas, but it was significantly lower than for immunoglobulin genes. However, for the first time high frequency of somatic mutations in various tissues, including non-hematopoietic and in tumor suppressor genes contrary to previous reports, which investigated oncogenes.

By sequencing 327 RASSF1A and RBSP3 clones, 364 mutations were detected at frequencies reaching 0.70 per 100 base pairs, and many clones contained more than 1 mutation. At present, the nature of the mechanism responsible for this hypermutability is not known. The results also suggest that mutations are not completely random.

Paper III

Chromosome 3 specific NotI microarrays containing 181 NotI linking clones associated with genes were hybridized with 40 paired normal/tumor DNA samples of non –small cell lung

cancer (NSCLC), including 28 squamous cell carcinoma (SCC) and 12 adenocarcinoma (ADC) primary tumors. Forty one genes showed methylation/deletion frequencies in more than 20% of NSCLC samples. In general SCC samples were more frequently methylated as compared to ADC. Moreover, in SCC methylation was observed already at Stage I of tumor development whereas many genes showed tumor progression specific methylation in ADC.

Among genes frequently methylated/deleted in NSCLC only a few genes were already known tumor suppressor genes, like *RBSP3/CTDSPL*, *RARB*, *VHL* and *THRB*. The majority of found genes were previously not shown to be involved in lung carcinogenesis, like *TRH, PPP2R3A, RPL32, LOC285205, FGD5, ROPN1, p20-CGGBP, NBEAL2* and *LOC440946*.

Two genes were additionally tested for methylation by bisulfite sequencing: *VHL* (2 tumor samples) and *ITGA9* (4 tumor samples). In all tested cases methylation was confirmed.

A significant downregulation of *RBSP3* and *ITGA9* was observed in early SCC and prevailed in advanced SCC and was independent of clinical stage, extent of cell differentiation, and metastasis to regional lymph nodes. In ADC, an increase in frequency and extent of downregulation of *ITGA9* was statistically significant and correlated with tumor progression. The same tendency was observed for the *RBSP3* gene. Methylation data correlated with expression of these genes.

A decrease in *RBSP3* mRNA was due to deletions (not excluding methylation) in 33% and promoter methylation without deletions in 33% of ADC cases. The downregulation of *ITGA9* was associated with deletions and/or methylation in 60% of ADC cases. Methylation without deletions was observed in 64% for *RBSP3* and 55% for *ITGA9* in SCC samples. Deletions (not excluding methylation) were found in 27% for *RBSP3* and 36% for *ITGA9*. Altogether downregulation of *RBSP3* in 80% of NSCLC cases resulted from deletions and/or methylation. For *ITGA9* these factors were responsible for decrease of expression in 77% of NSCLC samples.

To test for the growth suppressing effect of *ITGA9* and *RBSP3,* colony formation assays using the KRC/Y renal cancer and U2020 lung cancer cells were performed. *RBSP3* and *ITGA9* strongly inhibited colony formation both in renal and lung cancer cell lines.

Using obtained data a set of 18 markers (*BHLHB2, FBLN2, FLJ44898* (*EPHB1*), *GATA2, GORASP1, Hmm210782* (*PRICKLE2*), *Hmm61490, ITGA9, LOC285205, LRRC3B, MINA, MITF, MRPS17P3, NKIRAS1, PLCL2, TRH, UBE2E2, WNT7A*) that allow to discriminate/diagnose SCC and ADC, two main types of NSCLC (early detection, progression, metastases) with probability more than 95% and most complicated cases with probability more than 80% was suggested.

The study demonstrated that NotI microarrays are powerful tools to find methylated genes/tumor suppressor genes and resulted in identification of many novel genes/biomarkers that can be important for the development of more specific biomarker sets for early diagnosis and new approaches to therapy of NSCLC.

Paper IV

Chromosome 3 specific NotI microarrays containing 180 genes were hybridized to NotI representation probes prepared using matched tumor/normal samples from major epithelial cancers: breast, cervical, kidney, colon, ovarian and prostate cancers.
Methylation and/or deletions of LRRC3B was found in all tested primary tumors compared to normal controls ranging from 50% in prostate samples to 78% in cervical samples. To confirm these results bisulfite sequencing of cloned PCR products (6-14 clones for each cancer sample, a total of 87 clones were sequenced) from two renal, one cervical, two ovarian, two colon and two breast cancer samples was carried out. In all cases methylation of the LRRC3B NotI site region was detected (two NotI sites in PCR-fragment). Colony formation assays using KRC/Y cells were performed and LRRC3B showed very strong colony formation inhibition in vitro. When LRRC3B was expressed the cloning efficiency was less than 95%, compared to the empty vector. In addition, colonies grown in KRC/Y cells expressing LRRC3B were much smaller than in KRC/Y cells transfected with empty pETE vector.

GENERAL CONCLUSIONS (NotI microarrays)

Hypothesis testing

By checking the data obtained with NotI-microarrays with the control method *bisulfite sequencing of cloned PCR products*, the hypothesis that NotI microarrays technology achieves the results it seeks; to detect methylation changes in NotI sites in the cancer genome, has not been disproved. Therefore the hypothesis is still provisional (according to the strict scientific definition of a falsifiable hypothesis), but it is a hypothesis that has been rigorously tested and not falsified and therefore can form a reasonable basis for action, i.e. a reliable methodology to produce new objective scientific data.

According to the definition of an empirical hypothesis, since it is a well-crafted, well-controlled experiment, a lack of falsification does count as verification, since such an experiment ranges over the full scope of possibilities in the problem domain. The initial hypothesis is verified but incomplete.

Choosing control method

One problem we had to solve was how to control and verify the microarray results. We considered both the MSP and *Bisulphite sequencing of cloned PCR products* methods. We chose *Bisulphite sequencing of cloned PCR products* because it is the most accurate method and it can detect all individual CpGs in the PCR-region flanking the NotI-site, although it is more time- and labor consuming than MSP. With the control results obtained with *Bisulphite sequencing of cloned PCR products* much more can be concluded about the sensitivity of the NotI-microarrays compared with control results obtained by MSP. *Bisulphite sequencing of cloned PCR products* is widely recognized in the scientific community as the most exact and detailed method to study methylation, so therefore it is a reliable, objective and wellknown control method.

Discovering biomarkers and targets for epigenetic therapy

Altered genes found with the NotI microarrays technology may potentially be used as biomarkers and targets for epigenetic therapy. Sets of NotI site markers can be designed that allow to discriminate different types of cancer. Such marker sets have been produced already with NotI sites on chromosome 3, and applying NotI microarrays on all human chromosomes has the potential to produce many more NotI site markers to be used in diagnosis, prediction, prognosis and treatment of cancer.

ACKNOWLEDGEMENTS

This thesis work was performed at the Department of Microbiology, Tumor and Cell Biology (MTC), Karolinska Institute, Stockholm, Sweden. I wish to express my sincere gratitude to all who have helped me in any way during these studies.

Eugene Zabarovsky, my supervisor, for giving me the opportunity to work in his lab, for sharing his extensive knowledge in molecular biology, brilliant scientific guidance and writing skills about our results and for support and encouragement.

Vladimir Kashuba, my co-supervisor, for constructive scientific discussions and for guiding me through lab works. For sharing his extensive knowledge in molecular biology.

Georg Klein, for sharing his extensive knowledge in cancer biology.

Ingemar Ernberg, for being a brilliant teacher in cancer biology.

Yihai Cao, for valuable philosophical discussions and general advice.

Anna Krook and Francesca Chiodi, for kindly being my mentors.

Present and former colleagues of RCC group: Veronika Zabarovska, Tatiana Pavlova, Fuli Wang, Tatiana Ivanova, Lev Petrenko, Elvira Grigorieva, Alena Malyukova, Elian Rakhmanaliev, Jinfeng Li, Rinat Gizatullin, Julia Koblyakova, Alexei Protopopov, Alexey Kutsenko, Surya Pavan Yenamandra, Igor Bazov, Inessa Skrypkina, Sergey Kvasha, Tatiana Prudnikova, Viktor Shevchuk, Olga Kharchenko.

My dear parents Ingemar and Eva, for your unconditional love, continuous support and encouragement. My brothers and sisters Emma, Jakob, Elis, Olof, Sigrid and Erik for understanding me and supporting me. My friends and relatives.

All friends and personnel at MTC during the years, for creating a friendly atmosphere.

REFERENCES

Allikmets RL, Kashuba VI, Pettersson B, Gizatullin R, Lebedeva T, Kholodnyuk ID, Bannikov VM, Petrov N, Zakharyev VM, Winberg G, Modi W, Dean M, Uhlén M, Kisselev LL, Klein G, Zabarovsky ER. NotI linking clones as a tool for joining physical and genetic maps of the human genome. Genomics, 19: 303-9, 1994.

Anedchenko EA, Dmitriev AA, Krasnov GS, Kondrat'eva TT, Kopantsev EP, Vinogradova TV, Zinov'eva MV, Zborovskaia IB, Polotski BE, Sakharova OV, Kashuba VI, Zabarovsky ER, Senchenko VN. Down-regulation of RBSP3/CTDSPL, NPRL2/G21, RASSF1A, ITGA9, HYAL1 and HYAL2 genes in non-small cell lung cancer. Molecular Biology, 42: 859-869, 2008.

Armitage P, Doll R. The age distribution of cancer and a multi-stage theory of carcinogenesis. Br J Cancer, 8: 1-12, 1954.

Ballestar E, Esteller M. The epigenetic breakdown of cancer cells: from DNA methylation to histone modifications. Prog Mol Subcell Biol, 38: 169-81, 2005.

Baylin SB. Tying it all together: epigenetics, genetics, cell cycle, and cancer. Science, 277: 1948–1949, 1997.

Bookstein R, Lee EY, Peccei A, Lee WH. Human retinoblastoma gene: long-range mapping and analysis of its deletion in a breast cancer cell line. Molecular and Cellular Biology, 9: 1628–1634, 1989.

Brulliard M, Lorphelin D, Collignon O, Lorphelin W, Thouvenot B, Gothié E, Jacquenet S, Ogier V, Roitel O, Monnez JM, Vallois P, Yen FT, Poch O, Guenneuques M, Karcher G, Oudet P, Bihain BE. Nonrandom variations in human cancer ESTs indicate that mRNA heterogeneity increases during carcinogenesis. Proc Natl Acad Sci U S A, 104: 7522–7527, 2007.

Dai Z, Weichenhan D, Wu YZ, Hall JL, Rush LJ, Smith LT, Raval A, Yu L, Kroll D, Muehlisch J, Frühwald MC, de Jong P, Catanese J, Davuluri RV, Smiraglia DJ, Plass C. An AscI boundary library for the studies of genetic and epigenetic alterations in CpG islands. Genome Res, 12: 1591-8, 2002.

Di Cristofano A, Pandolfi PP. The multiple roles of PTEN in tumor suppression. Cell, 100: 387–390, 2000.

Esteller M, Sanchez-Cespedes M, Rosell R, Sidransky D, Baylin SB, Herman JG. Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. Cancer Res, 59: 67-70, 1999.

Esteller M, Corn PG, Baylin SB, Herman JG. A gene hypermethylation profile of human cancer. Cancer Res, 61: 3225-3229, 2001.

Esteller M. The necessity of a human epigenome project. Carcinogenesis, 27: 1121-5, 2006.

Esteller M. Cancer epigenomics: DNA methylomes and histone-modification maps. Nat Rev Genet, 8:286-298, 2007.

Esteller M. Epigenetics in Cancer. N Engl J Med, 358: 1148, 2008.

Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell, 61: 759-767, 1990.

Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CL. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci, 89: 1827-1831, 1992.

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

Harbour JW, Dean DC. Rb function in cell-cycle regulation and apoptosis. Nature Cell Biology, 2: E65–E67, 2000.

Herceg Z. Epigenetics and cancer: towards an evaluation of the impact of environmental and dietary factors. Mutagenesis, 22: 91–103, 2007.

Herman JG, Baylin SB. Gene silencing in cancer in association with promoter hypermethylation. N Engl J Med, 349: 2042-2054, 2003.

Jones PA, Laird PW. Cancer epigenetics comes of age. Nature Genet, 21: 163-167, 1999.

Jones PA, Martienssen R. A blueprint for a Human Epigenome Project: the AACR Human Epigenome Workshop. Cancer Res, 65: 11241-6, 2005.

Kessler C, Manta V. Specificity of restriction endonucleases and DNA modification methyltransferases a review. Gene, 92: 1-248, 1990.

Killary AM, Wolf ME, Giambernardi TA, Naylor SL. Definition of a tumor suppressor locus within human chromosome 3p21-p22. Proc Natl Acad Sci USA, 89: 10877-10881, 1992.

Kok K, Naylor SL, Buys CH. Deletions of the short arm of chromosome 3 in solid tumors and the search for suppressor genes. Adv Cancer Res, 71: 27-92. 1997.

Knudson AG Jr. Mutation and cancer: statistical study of retinoblastoma. Proceedings of the National Academy of Sciences of the USA, 68: 820–823, 1971.

Kuzmin I, Geil L, Ge H, Bengtsson U, Duh FM, Stanbridge EJ, Lerman MI. Oncogene, 18: 5672–5679, 1999.

Kvasha SM. Modern methodical approaches to determining the DNA methylation status and their use in oncology [Article in Ukrainian]. Ukr Biokhim Zh, 80: 5-15, 2008.

Lambert AR, Sussman D, Shen B, Maunus R, Nix J, Samuelson J, Xu SY, Stoddard BL. Structures of the rare-cutting restriction endonuclease NotI reveal a unique metal binding fold involved in DNA binding. Structure, 16: 558-69, 2008.

Lerman MI, Minna JD. The 630-kb lung cancer homozygous deletion region on human chromosome 3p21.3: identification and evaluation of the resident candidate tumor suppressor genes. Cancer Res, 60: 6116-6133, 2000.

Li J, Wang F, Kashuba VI, Wahlestedt C, Zabarovsky E. Cloning of deleted sequences (CODE): A genomic subtraction method for enriching and cloning deleted sequences. Biotechniques, 31: 788, 790, 792-3, 2001.

Li J, Protopopov A, Wang F, Senchenko V, Petushkov V, Vorontsova O, Petrenko L, Zabarovska V, Muravenko O, Braga E, Kisselev L, Lerman MI, Kashuba V, Klein G, Ernberg I, Wahlestedt C, Zabarovsky ER. NotI subtraction and NotI-specific microarrays to detect copy number and methylation changes in whole genomes. Proc Natl Acad Sci U S A, 99: 10724-9, 2002.

Li LC. Designing PCR primer for DNA methylation mapping. Methods Mol Biol, 402: 371-84, 2007.

Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics, 18: 1427-31, 2002.

Mai A, Altucci L. Epi-drugs to fight cancer: From chemistry to cancer treatment, the road ahead, Int J Biochem Cell Biol, 41: 199-213, 2009.

Marx J. How p53 suppresses cell growth. Science, 262: 1644–1645, 1993.

Mungall AJ. Meeting review: epigenetics in development and disease. Comp Funct Genomics, 3: 277-81, 2002.

Nelson M, McClelland M. Site-specific methylation: effect on DNA modification methyltransferases and restriction endonucleases. Nucleic Acids Res, 19 Suppl: 2045-71, 1991.

Novik KL, Nimmrich I, Genc B, Maier S, Piepenbrock C, Olek A, Beck S. Epigenomics: genome-wide study of methylation phenomena. Curr Issues Mol Biol, 4: 111-28, 2002.

Oh JJ, Boctor BN, Jimenez CA, Lopez R, Koegel AK, Taschereau EO, Phan DT, Jacobsen SE, Slamon DJ. Promoter methylation study of the H37/RBM5 tumor suppressor gene from the 3p21.3 human lung cancer tumor suppressor locus. Hum Genet, 123: 55-64, 2008.

Pasqualucci L, Neumeister P, Goossens T, Nanjangud G, Chaganti RS, Kuppers R, Dalla-Favera R. Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. Nature, 412: 341–346, 2001.

Pan ZG, Kashuba VI, Liu XQ, Shao JY, Zhang RH, Jiang JH, Guo C, Zabarovsky E, Ernberg I, Zeng YX. High frequency somatic mutations in RASSF1A in nasopharyngeal carcinoma. Cancer Biol Ther, 4: 1116–1122, 2005.

Pavlova TV, Kashuba VI, Muravenko OV, Yenamandra SP, Ivanova TA, Zabarovskaia VI, Rakhmanaliev ER, Petrenko LA, Pronina IV, Loginov VI, Iurkevich OIu, Kiselev LL, Zelenin AV, Zabarovskiĭ ER. Use of NotI microarrays in analysis of epigenetic and structural changes in epithelial tumor genomes by the example of human chromosome 3. [Article in Russian]. Mol Biol (Mosk), 43: 339-47, 2009.

Ramirez JL, Taron M, Balana C, Sarries C, Mendez P, de Aguirre I, Nunez L, Roig B, Queralt C, Botia M, Rosell R. Serum DNA as a tool for cancer patient management. Rocz Akad Med Bialymst, 48: 34-41, 2003.

Roberts RJ, Macelis D. Restriction enzymes and their isoschizomers, Nucleic Acids Res, 19 Suppl: 2077-109, 1991.

Russo G, Zegar C, Giordano A. Advantages and limitations of microarray technology in human cancer. Oncogene, 22: 6497-507, 2003.

Senchenko VN, Liu J, Loginov W, Bazov I, Angeloni D, Seryogin Y, Ermilova V, Kazubskaya T, Garkavtseva R, Zabarovska VI, Kashuba VI, Kisselev LL, Minna JD, Lerman MI, Klein G, Braga EA, Zabarovsky ER. Discovery of frequent homozygous deletions in chromosome 3p21.3 LUCA and AP20 regions in renal, lung and breast carcinomas. Oncogene, 23: 5719–5728, 2004.

Takai D, Jones PA. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc Natl Acad Sci USA, 99: 3740–3745, 2002.

Talmadge CB, Zhen DK, Wang JY, Berglund P, Li BF, Weston MD, Kimberling WJ, Zabarovsky ER, Stanbridge EJ, Klein G, Sumegi J. Construction and characterization of a NotI linking library from human chromosome region 1q25-qter. Genomics, 29: 105-14, 1995.

Wang J, Gonzalez KD, Scaringe WA, Tsai K, Liu N, Gu D, Li W, Hill KA, Sommer SS. Evidence for mutation showers. Proc Natl Acad Sci U S A, 104: 8403–8408, 2007.

Wang SS, Smiraglia DJ, Wu YZ, Ghosh S, Rader JS, Cho KR, Bonfiglio TA, Nayar R, Plass C, Sherman ME. Identification of novel methylation markers in cervical cancer using restriction landmark genomic scanning. Cancer Res, 68: 2489-97, 2008a.

Wang F, Grigorieva EV, Li J, Senchenko VN, Pavlova TV, Anedchenko EA, Kudryavtseva AV, Tsimanis A, Angeloni D, Lerman MI, Kashuba VI, Klein G, Zabarovsky ER. HYAL1 and HYAL2 inhibit tumor growth in vivo but not in vitro. PLoS ONE, 3: e3031, 2008b.

Wei MH, Latif F, Bader S, Kashuba V, Chen JY, Duh FM, Sekido Y, Lee CC, Geil L, Kuzmin I, Zabarovsky E, Klein G, Zbar B, Minna JD, Lerman MI. Construction of a 600-kilobase cosmid clone contig and generation of a transcriptional map surrounding the lung cancer tumor suppressor gene (TSG) locus on human chromosome 3p21.3: progress toward the isolation of a lung cancer TSG. Cancer Res, 56: 1487-1492, 1996.

World Health Organization (February 2009). "Cancer". Available at: http://www.who.int/mediacentre/factsheets/fs297/en/index.html. Retrieved 28.10.2009.

Worm J, Guldberg P. DNA methylation: an epigenetic pathway to cancer and a promising target for anticancer therapy. J Oral Pathol Med, 31: 443–449, 2002.

Yu N, Wang M. Anticancer drug discovery targeting DNA hypermethylation. Curr Med Chem, 15: 1350-75, 2008.

Zheng L, Li S, Boyer TG, Lee WH. Lessons learned from BRCA1 and BRCA2. Oncogene, 19: 6159–6175, 2000.