Chapter IV

Positional Cloning of Tumor Suppressor Genes from 3p21.3 Involved in Major Human Cancers

E. R. Zabarovsky1,2, V. N. Senchenko2, V. Loginov3, T. Pavlova1, V. I. Zabarovska1, A. Dmitriev3, M. L. Lung4, C.K. Panda5, V.I. Kashuba1,6, M.I. Lerman7 and E.A. Braga3

1MTC, Department of Clinical Science and Education, Sodersjukhuset, Karolinska Institute, Stockholm, 171 77, Sweden
2Engelhardt Institute of Molecular Biology, the Russian Academy of Sciences, Vavilova, 32, Moscow, 119991 Russia
3Russian State Genetics Center, 1-st Dorozhny proezd 1, Moscow, 117545 Russia
4Centre for Cancer Research and Department of Clinical Oncology, University of Hong Kong, Pokfulam, Hong Kong (SAR), People’s Republic of China
5Chittaranjan National Cancer Institute, 37 S. P. Mukherjee Road, Kolkata 700 026, India
6Institute of Molecular Biology and Genetics, Ukrainian Academy of Sciences, Kiev, Ukraine
7Cancer-Causing Genes Section, Laboratory of Immunobiology, Center for Cancer Research, National Cancer Institute-Frederick, Bldg 560, NCI-Frederick, Frederick, MD 21702; currently at NIH/NHLBI, Hematology Branch.10 CRC, Room 5E-3256 Bethesda, MD 20895, USA

1. Abstract

We have performed a comprehensive deletion survey of 3p on more than 400 lung, renal, breast, cervical and ovarian carcinomas (major epithelial cancers) using a defined set of markers, combining conventional LOH with quantitative real-time PCR (QPCR), NotI linking and jumping libraries, comparative genomic and NotI microarrays hybridizations. We identified two most
frequently affected 3p21.3 regions, LUCA (LUng CAncer) at the centromeric and AP20 at the telomeric border of 3p21.3. Aberrations of either region were detected in more than 90% of the studied tumors suggesting they harbor multiple tumor suppressor genes (TSG). Homozygous deletions (HD) were frequently detected in all tumors in both the LUCA and AP20 regions. To facilitate the identification of tumor suppressor genes (TSGs) in the chromosome 3p21.3 AP20 and LUCA sub-regions, we constructed physical and gene map of these segments. More than 30 genes were localized in these two regions and among them at least 12 TSGs were identified: RBSP3, ITGA9, MLH1, VILL, APRG1, RASSF1, HYAL1, HYAL2, SEMA3B, SEMA3F, NPRL2 and CACNA2D2. Among these TSGs were representatives of new types of TSGs: e.g. HYAL1 showing growth inhibiting activity only in vivo and RASSF1A and RASSF1C that are alternative forms of the same gene showing different tissue specificity. We found that several tumor suppressor genes in AP20 and LUCA 3p21.3 regions were co-regulated in tumors. These results supported the hypothesis on simultaneous inactivation of clusters cancer-causing genes in AP20 and LUCA regions during the development and progression of lung cancer and other epithelial tumors. Moreover we found an exceptionally high incidence of single-base mutations in the tumor suppressor genes RASSF1 and RBSP3 (CTDSPL) in major epithelial tumors. These mutations functionally inactivated tumor suppressor activity of these genes. Somatic hypermutations in tumor suppressor genes involved in major human malignancies offer a novel insight in cancer development, progression and spread. The data could be important for development of specific biomarker sets for early cancer diagnosis and new therapeutic approaches/strategies for cancer treatment. For example we selected a set of 23 markers (BHLHB2, FBLN2, EPHB1, GATA2, GORASP1, PRICKLE2, Hmm61490, ITGA9, LOC285205, LRRRC3B, MINA, MITF, MRPS17P3, NKIRAS1, PLCL2, TRH, UBE2E2, WNT7A, RARB2, p20-CGGBP, GNAI2, RPL32, THR2) that would discriminate/diagnose the majority of NSCLC cases with P > 95% and most complicated cases with a probability of more than 80%.

2. Introduction

Epithelial tumors are the most prevalent and lethal cancers in the world. They cause more than 80% of all cancer deaths. For example, only lung cancer kills >150,000 patients each year in the USA and many more around the world.

Carcinogenesis is a multi-step genetic process that involves mutational activation of proto-oncogenes and inactivation/silencing of tumor suppressor genes (TSGs). The localization of these genes to specific chromosome regions and identification of genes causally related to the neoplastic process constitutes one of the major efforts in cancer research today. It is also essential for elucidating signaling pathways in oncogenesis. Additionally, these genes and their products could be used to design new cancer biomarkers to improve early diagnosis, predict treatment strategies and disease course.

Deletions in Chr.3 are frequently associated with tumors, and more than 20 years ago it was suggested that this chromosome is probably the site of at least one, and probably several TSGs. Many studies showed abnormalities of the short arm of this chromosome in carcinomas of various organs: kidney (clear cell RCC), lung (SCLC and NSCLC), breast (BC), ovary, cervix (CC), testis, head and neck and others. (Kok et al., 1997; Mitelman et al., 1997).

Chromosome 3p deletions are detected in almost 100% of SCLC, RCC and in more than 90% of NSCLC (Braga et al., 1999, 2002; Zabarovsky et al., 2002; Imreh et al., 2003; Hesson et al., 2007). Chromosome transfer experiments indicated that different regions of 3p could
suppress the tumorigenic properties of cancer cell lines. These and other findings suggested the presence of several TSG in 3p.

These data appeared at the end of the 1980s, however the search for TSGs was hampered by the size of the region where they were probably located. This covers practically the whole short arm of Chr.3 (about 100 Mb). According to several studies, the most frequently affected regions (FARs) in sporadic RCC and SCLC are 3p13-p14 and 3p21.2-3p21.3 (Van den Berg & Buys, 1997; Erlandsson, 1998; Zabarovsky et al., 2002). Very little progress has been made over the first ten years. Candidate TSGs were found, but none of them produced good evidence of tumor suppressor activity. For example, the FHIT (3p14) and ARP (3p21.1) genes have been suggested as candidate TSGs but data from different groups were contradictory (Kok et al., 1997; Van den Berg & Buys, 1997; Erlandsson, 1998). The gene responsible for von Hippel Lindau (VHL) disease, located in 3p25-p26, was also cloned (Latif et al., 1993). The VHL gene is mutated in RCC cell lines and primary tumors, suggesting that it is involved in the development of RCC. This didn’t explain the frequent deletions in the 3p13-p21 region, however. Moreover, more and more data were found that VHL mutations were not necessary for the development of clear-cell RCC and most probably occurred during the progression phase of tumorigenesis. Similarly, confusing situations existed for other candidate genes (Van den Berg & Buys, 1997; Van den Berg et al., 1997; Zabarovsky et al., 2002).

Homozygous deletions frequently mark the position of TSGs. Our group and others described homozygous deletions on 3p in different places: 3p12-p13, 3p14, 3p21, 3p21-p22, and 3p25-p26 (Wei et al., 1996; Kok et al., 1997; Van den Berg et al., 1997). Such equivocal results prompted us to search for new options to study these cancers and to identify TSGs.

Our strategy to isolate TSG(s) was based on the hypothesis that the function of a TSG, localized to the most frequently deleted chromosomal region of a given tumor type, should be impaired or changed in those tumors. We looked, therefore, for a gene (or genes) that have at least one of the following features in tumors and cancer cell lines: a) deletion; b) rearrangement; c) mutation; d) anomalous mRNA size. A loss of expression can also be caused by methylation of the control regions of the gene. We have also suggested that the combined action of several genes located on Chr.3 and gene dosage may be important for the development of cancer. We thought that it is important to demonstrate tumor suppressor activity for candidate genes.

Loss of heterozygosity (LOH) analysis of a number of lung cancer and RCC biopsies resulted in the identification of several FARs in 3p. The problem was that different studies defined different FARs. Different groups applied different subsets of 3p polymorphic markers and comparison of these data was seriously complicated (Kok et al., 1997; Zabarovsky et al., 2002). Moreover, many markers were mapped either wrongly or imprecisely. Most publications described partial 3p losses (Van den Berg et al., 1997), but other studies (Wilhelm et al., 1995) suggested that the interstitial deletions found in other works were in fact spurious, due to contamination of tumors with normal cells. If different markers are differently affected by this admixture, artificial LOH patterns may be obtained. We were alerted to this problem in the course of our study on nasopharyngeal carcinoma (NPC; Hu et al., 1996). Multiple interstitial deletions were found on the short arm of Chr.3 in several tumors. Different markers were lost with different frequencies among the cases, ranging from 9% to 60%. We referred to this effect as the "zebra pattern". The question arose whether this pattern was an artifact, reflecting differences in the sensitivity of detection with different
markers, or whether it reflected true cases of interstitial deletions. In that study (Hu et al., 1996), we compared three selected microsatellite markers in terms of the sensitivity of their detection by a polymerase chain reaction (PCR) assay. The comparison was carried out by serially diluting human DNA with mouse DNA for PCR amplification using different primers. There were no major differences in sensitivity.

Later, we performed a more detailed study of this problem (Liu et al., 1999). LOH analysis was carried out on DNA samples containing different alleles, which were mixed in different proportions to mimic possible contamination of tumor cells with normal cells in biopsies. These experiments did not reveal significant differences in the sensitivity of detection for different markers, but there was a systematic difference between the low (L) and high (H) molecular weight alleles of the same marker. The absence of H was detected with higher sensitivity than the absence of L allele. This asymmetry introduces a potential source of error. In the presence of contaminating normal cells, the same marker in the same tumor may be considered as deleted or retained, depending on which allele (H or L) is deleted. Random screening of 50 papers published in 1994-1996 revealed that the loss of L allele is detected only at 56% frequency of the loss of H allele (Liu et al., 1999).

To avoid bias in detection of LOH with H and L alleles we developed the following rules.

1. More weight must be given to LOH found with the L than with the H allele ("rule of allele L"). For instance, markers that appear to be retained were considered as such only in cases where the suspected allele is close to a locus showing LOH with respect to allele L. Otherwise, if no independent supporting data were available, we considered such markers as non-informative.

For example, analysis of the films for contiguous informative markers A-B-C revealed that the H allele is deleted in A and/or C, and B seemed to be undeleted. In this case, we consider that B is uninformative. If the L allele were deleted in both A and C, then we would consider B to be retained.

2. The number of deleted H and L alleles should be established for each of the normal/tumor pairs. These numbers should be similar.

3. Comparative genome hybridization, quantitative real-time PCR or another method of analysis independent from CA-repeat should be used to confirm interstitial deletions.

We decided also to use an identical marker set to compare genomic alterations of 3p in different tumors. We compared chromosome 3p in five cancer types: clear cell renal cell carcinoma (RCC), breast (BC), non-small cell lung (NSCLC), cervical (CC) and epithelial ovarian carcinoma (EOC) (311 patients in total).

In summary, our results have unequivocally proved the existence of interstitial deletions in all studied tumors and have identified at least three FARs common for five cancers.

1. 3p21.3-22 (AP20 region, around D3S1298 and D3S1260 (Kashuba et al., 1995). Homozygous deletions in the AP20 region were reported in five SCLC cell lines and three tumor biopsies (Murata et al., 1994; Roche et al., 1996).

2. 3p21.31 (between D3S2409 and D3S3667). This 600kb region was called major epithelial cancer region No. 3 (MECA3).
3. 3p21.2 (Lung Cancer, LUCA region, around D3S1235 and D3S1289). Several groups have reported homozygous deletions in this region for SCLC (Daly et al., 1993; Wei et al., 1996).

Several directions of our previous studies were associated with NotI jumping and linking clones. We suggested an approach to combine physical and gene mapping methods to characterize large regions of human chromosomes (Zabarovsky et al., 1990, 1994; Allikmets et al., 1994), and NotI linking and jumping clones were proposed as framework markers.

We developed simplified new procedures for jumping and linking library construction and constructed numerous linking and jumping libraries with different restriction enzymes (Zabarovsky et al., 1994) in an attempt to generate representative NotI linking libraries, covering the whole human genome. We experimentally confirmed that there is a direct connection between CpG islands, NotI sites and genes in the human genome. We created high-density grids that contained 50,000 clones and sequenced these clones (Kutsenko et al., 2002). We estimated that sequences of these clones will provide information for up to 10%-20% of all human genes, and will lead to the identification of thousands of new genes. We generated 100,000 NotI flanking sequences and identified among them approximately 22,000 unique NotI sequences containing 17 Mb information. Using this information we identified and mapped numerous new genes. Even now more than 4% of these NotI sequences are still absent in public databases. Thus at least 1000 novel genes can be discovered using these clones.

The presence of genetic alterations in tumors is now widely accepted, and explains the irreversible nature of tumors. However, observations on tissue differentiation indicated that it shares something in common with carcinogenesis, that is, epigenetic changes. Now, DNA methylation in CpG sites is known to be precisely regulated in tissue differentiation, and is supposed to be playing a key role in the control of gene expression in mammalian cells. The genes involved include tumor suppressor genes, genes that suppress metastasis and angiogenesis, and genes that repair DNA suggesting that epigenetics plays an important role in tumorigenesis (Goldberg et al. 2006). It becomes clear that methylation is a basic, vital feature/mechanism in mammalian cells. Inactivation of DNA methyltransferase is lethal for mice. It is involved in hereditary and somatic cancers, hereditary and somatic diseases, apoptosis, replication, recombination, immune response, etc.

It is believed that epigenetics can be used for diagnosis, prognosis, prediction and even for direct treatment of cancer. It was shown that methylation changes in the developing tumors could be detected 1-3 years before malignancy can be revealed by any other methods (Pal misano et al., 2000). Detection of cancer one year earlier means enormously increased chances for the successful medical treatment.

Based on the growing understanding of the roles of DNA methylation, several new methodologies were developed to make a genome-wide search for changes in DNA methylation. There are several main methods (see Sugimura and Ushijima, 2000; Shames et al., 2007) for testing methylation in human genome, e.g. restriction landmark genomic scanning (RLGS), methylation-sensitive-representational difference analysis (MS-RDA), methylation-specific AP-PCR (MS-AP-PCR), methyl-CpG binding domain column/segregation of partly melted molecules (MBD/SPM), real-time qPCR, expression microarrays (Shames et al., 2006) and CpG islands microarrays (CGI, Yan et al., 2001). Although each of them has its own advantages none of them is suited for large-scale screening as all are rather inefficient, complicated or technically challenging. For example,
after analysis of 1000 clones isolated using MBD/SPM, nine DNA fragments were identified as CpG islands and only one was specifically methylated in tumor DNA. Maximum resolution of RLGS is 1 000-2 000 NotI boundary clones (i.e. appr. 700-800 NotI sites) and we have isolated more than 20 000. Several companies produce now microarrays or kits for detection methylated genes, e.g. Applied Biosystems, Agilent Technologies, SA Biosciences, Affymetrix and others. However they are based on the principles described above with all their advantages and drawbacks. Recently Methylated DNA Immunoprecipitation (MeDIP) technology was developed in different variants but is too early to make conclusion about efficiency and robustness of this procedure.

New type of microarrays recently developed in our group open new possibilities for large scale study of methylation patterns in normal and pathological cells (Li et al., 2002). The main objective is to prepare and to use NotI microarrays (NMA, i.e. glass microarrays with attached NotI flanking DNA fragments) for comparison of normal and malignant cells at the genomic level and finding genes differently methylated in normal and tumor cells.

The main advantages of NMA compared to other methylation microarrays: all clones represent genes; incomplete digestion doesn’t generate artificial results; only 0.01-0.05% of the total DNA is labeled increasing sensitivity several thousands times compared to other approaches where total human DNA is labeled; NotI representation (NR) probe contains 10-fold less repeats than total human DNA; homo- and hemizygous methylation and demethylation can be detected.

In this chapter we shall summarize our data on positional cloning of cancer associated genes involved in major epithelial malignancies from AP20, MECA3 and LUCA regions. For this positional cloning we used not only genetic but also genomic and epigenetic mapping.

3. Results

a) Localization of Critical Regions on the Short Arm of Chromosome 3 Using Allelotyping of the Dense Set of Polymorphic Markers

In this study 311 patients were analyzed: 80 cases of RCC; 95 of BC, mainly (65) ductal; 47 NSCLC mainly squamous cell lung cancer (38 cases) and adenocarcinoma (9 cases); 50 EOC of diverse histological types, including 32 cases of serous adenocarcinoma; 39 CC of squamous cell histological type only. The order of polymorphic markers on 3p, used in this work, was based on NCBI data (Build 33) and UCSC Browser data. Primers for microsatellite markers were from GenBank Amplicon databases. Location of marker NL1-024/D3S4285 and primers for this locus were determined in our previous studies (Braga et al., 1997; Kashuba et al., 1999; Sulimova et al., 2002, 2005).

Multiple interstitial deletions were found in prevalence to terminal deletions in 3p in five epithelial tumor types

The representative subsets of RCC, BC, NSCLC, CC and EOC samples were examined using identical 3p markers. Proportion of tumor cases with allelic losses was high: 93% (74/80) in RCC, 88% (84/95) in BC, 84% (42/50) in EOC, 79% (37/47) in NSCLC, and 79% (31/39) in CC. Terminal deletions are thought to result from mitotic recombination (De
Interstitial deletions are more useful and can indicate location of TSG(s). Terminal deletions were frequent in RCC (29%) and NSCLC (34%) only. Multiple interstitial deletions were found in prevalence to terminal deletions in each tumor type under the study: 64% vs. 29% in RCC, 45% vs. 34% in NSCLC, 85% vs. 3% in BC, 82% vs. 2% in EOC and 77% vs. 2% in CC (P<<0.05). These data probably indicated the location of several putative TSG(s) in chromosome 3p.

**Chromosome 3p "hot spots“ affected in major epithelial tumors**

To assess the contribution of different regions of 3p to various tumor types and to localize novel region(s) potentially harboring TSG(s) or protooncogene(s), the allelic imbalance (AI) frequencies were scored in five tumor types for the 3p markers. AI profile of the 3p markers displayed several peaks for each cancer type.

The markers were divided into nine groups, and the average AI frequencies for each region in each tumor type were counted (Figure 1). Four regions, namely # 3, 4, 6 and 7 showed the highest frequencies of AI in all five tumor types. The average values of AI frequencies in these regions were fairly higher in comparison with the average value of AI in 3p for all markers for a given tumor and exceeded the SD, CI 95% (Figure 1).

Region #3 in 3p23-p22.3 harbors RARB2, THRB, and TGFBR2 genes (Kok et al., 1997). Three others are located in 3p21.3. Two of them, #4 represent AP20 (marked by NL1-024/D3S4285 and D3S1298) and #7 or LUCA (marked by D3S4614) described by us in (Lerman and Minna, 2000; Protopopov et al., 2003).

![Figure 1. Allelic imbalance frequencies in nine regions of 3p. Twenty three markers were grouped in nine regions of 3p. Average values of AI frequencies for these regions were displayed on Levels Plot using the S-PLUS 2000 Professional Edition program (http://www.mathsoft.com/splus). Color intensity (white-gray-black) indicates AI frequency](attachment:image.png)
The third frequently affected region in 3p21.3 (#6, marked by D3S2409, D3S2456, and D3S3667), was localized for the first time, and importance of this novel site was proved using Fisher exact test ($P<10^{-3}$; Figure 1).

The new region (600kb size, 49.3-49.9 Mb according to the NCBI and UCSC databases), was localized in 3p21.31 close to the LUCA region that harbors a cluster of putative and genuine TSG(s) (Lerman and Minna, 2000; Zabarovsky et al., 2002; Imreh et al., 2003; Dammann et al., 2003; Vos et al., 2004; Dreijerink et al., 2001; Li et al., 2004a, 2004b). According to the UCSC Browser data, the D3S2409-D3S3667 MECA3 interval contains numerous UniGene clusters (~30 genes/600 kb). Some of the genes residing in this region (e.g. RHOA/ARHA, GPX1, MST1, MSTR1/RON, USP4, DAG1, Figure 1) were reported to be involved in vital cell functions (such as apoptosis, adhesion, signal transduction, angiogenesis and others) and suggested to be involved in the neoplastic transformation or tumor progression (Pille et al., 2005; Hu et al., 2003; Lee et al., 2001; Peace et al., 2001; DeSalle et al., 2001; Sgambato et al., 2003).

Figure 2. The use of multiplex PCR to confirm amplification events. (A). Allelic imbalance patterns of polymorphic markers in some RCC, BC and EOC cases examined by multiplex PCR (LOH – loss of one allele; RET – retention of both alleles; LOH +A – loss of one allele, associated with amplification of a second allele; NI – non-informative events). Polyacrylamide gel (10%) electrophoresis and silver staining was used. (B). Representative examples of multiplication of RHOA in some RCC and BC cases; 5′-region of the RHOA and exon 3 of the RARβ2 was examined.
The D3S1578 marker was additionally analyzed in EOC because earlier tumor suppressing activity was mapped to this locus using a functional assay (Rimessi et al., 1994). However, AI frequencies in D3S1578 in EOC were only 41% (14/34), even less than the average value for the 3p markers tested in EOC (43%, 263/616).

Thus, four regions #3, 4, 6 and 7 (Figure 1) were frequently affected in all five tested major epithelial tumors. Three other regions #1, 5 and 9 (Figure 1) were also found affected in some tumors. Region around VHL gene (#1, Figure 1) showed AI mainly in RCC, NSCLC and CC. Common eliminated region (CER or #5, Figure 1) localized by functional assay previously (Kholodnyuk et al., 1997, 2002), was predominantly affected in RCC and EOC. Region #9 in 3p14.2-p13 (Figure 1) was mainly involved in CC.

**Contribution of amplification events in allelic alterations of 3p microsatellites**

Pattern of allelic alterations was usually represented by loss of one allele, and in some cases allelic losses were accompanied by amplifications of the second allele. Amplifications were detected in 13% (141 amplifications of 1052 informative events) of RCC, 10% (130/1290) of BC, 8% (52/643) of NSCLC, 9% (60/657) of EOC and in 2% (11/594) of CC.

Such amplification events can unmask activation of protooncogene(s) happened simultaneously with loss or inactivation of adjacent TSG(s). Multiplex PCR was applied to confirm LOH associated with amplification events (Figure 2).

The non-random distribution of amplification events along 3p was especially pronounced in BC within the MECA3. Amplification events in the MECA3 constituted 16.5% (31/187, for three markers D3S2409, D3S2456 and D3S3667). This value was significantly higher than the average value (10%, 130/1290) for all 3p markers analyzed (P=0.03). We suggested that 3p and especially the MECA3 harbor putative oncogene(s) possibly located in the close proximity to TSG(s) loci. Indeed, 5’-region of the putative protooncogene RHOB showed also increased copy numbers in some RCC and BC cases.

**Correlation of 3p allelic alterations with tumor progression**

To examine a possible association of 3p allelic alterations with tumor progression, the most common histological subtypes (clear cell RCC, ductal and lobular BC, serous ovarian adenocarcinoma, squamous cell lung cancer and squamous cell CC) were analyzed. Two approaches were used. In one approach the number of affected tumor samples in a given region was divided by the total number of tumor samples and compared between groups of less or more advanced tumor stage/malignancy grade (for BC and RCC) or tumor differentiation (for EOC, and NSCLC). In the second approach, the percentage of AI events within the certain clinical groups relative to the total informative events in a particular 3p region was compared. For example, altogether 41 AI events from 60 informative events were detected in RCC in the LUCA region and among them 18 AI from 32 informative events were from tumors Stage I-II and 23 AI vs. 28 informative events were from Stages III+IV (Figure 3, C). Comparison of the main clinical variants using these two calculation approaches for RCC and serous EOC was demonstrated in Figure 3, A-D. Correlations were shown to be significant for AP20 and LUCA regions in clear cell RCC (Figure 3, A, C), serous EOC (Figure 3, B, D), and ductal BC (not shown), and also for AP20 – in squamous cell NSCLC (not shown). The MECA3 was associated with advanced clinical stages and poor differentiation of serous EOC (Figure 3, B, D). The region #3 in 3p23-p22.3 (D3S1283, D3S3047), harboring THRBI, RARβ2 and TGFBRII, also showed correlation with the
progression of clear cell RCC (Figure 3, C), serous EOC (Figure 3, D) and ductal BC (not shown).

In summary, AI in all four most important 3p regions (AP20, LUCA, MECA3, and around RARβ2) was correlated (P<0.05) with tumor progression of certain histological subtypes.

Figure 3. Correlation of allelic alterations in 3p regions with tumor progression in clear cell RCC and serous ovarian adenocarcinoma (serous EOC) estimated by two approaches. The percentage of tumor samples with AI in a given region relative to the total tumor samples is shown for early and advanced clinical and histological groups (A and B). The percentage of AI events for early and advanced clinical and histological groups relative to all informative events in a given region is shown (C and D). Diff-
tion, a level of tumor differentiation in abbreviated form: high (H), medium (M), and low (L) level of tumor differentiation; 1, 2, 3 designate tumor malignancy grade. The main TNM stages (I+II vs. III+IV), tumor malignancy grade (1+2 vs. 3) and tumor differentiation level (H + M vs. L) were compared.
b) Careful Mapping of LUCA and AP20 Regions Using Real Time Quantitative PCR

Importantly significant correlation was observed between DNA copy numbers for two non-polymorphic STS markers and LOH data for adjacent polymorphic loci (Braga et al., 2002; Senchenko et al., 2003, 2004).

**QPCR for detection copy number changes**

This approach is very precise and permits for example detection of X chromosome copy number differences between males (one copy) and females (two copies) (Senchenko et al., 2003). However, QPCR should be used with caution as cancer cells usually show numerous chromosomal abnormalities leading to copy number changes of many genes including reference genes. That is why the most important cases including samples with homozygous deletions used for fine localization of the minimal homozygously deleted regions were tested with different reference genes: ACBT, PF2K and NotI linking clone 924-021 containing an unknown gene. In the majority of cases, results with different reference genes were very similar. Cases showing different relative DNA copy number with different reference genes were not considered in the paper.

**QPCR results for RCC, BC and SCLC tumor samples and cell lines**

Schematic map of AP20 and LUCA regions is shown in Figure 4. Summary of these data (Senchenko et al., 2003; 2004) are shown in Table 1. Using only two markers, NLJ-003 and NL3-001, aberrations can be detected in more than 90% of studied cancers (percents/cases): 93% (49 cases) in RCC, 91% (20 cases) in BC, 96% (22 cases) in SCLC, 91% (29 cases) in CC samples. These results are even more remarkable as published LOH level for BC and CC is lower even using many microsatellite markers (80% and 75%, respectively, (Braga et al., 2002). These results strongly support the suggestion that these two regions are real hot spots for rearrangements in MEC (Major Epithelial Carcinomas). In general, both AP20 and LUCA regions showed similar patterns and frequencies of aberrations.

Still some differences were clear. For instance, level of hemizygous deletions in AP20 region was a little bit higher than in LUCA region but it is easy to understand as many deletions are terminal and AP20 is located closer to the telomere. The highest level of hemizygous deletions was detected in SCLC: 60.9% in NLJ-003 and 60% in NL3-001. Most likely this results from the use of tumor cell lines. The highest fraction of homozygous deletions in NLJ-003 was detected in CC (15.6%) and for NL3-001 it happened in 18% of BC cases. The highest fraction of amplification was detected for both loci in RCC: 34% in NLJ-003 and 42.5% in NL3-001. To confirm that amplification is a common event in RCC we performed additional experiments with 5 renal cell carcinoma cell lines (RCC) showing LOH and considered to have 3p-deletions (Alimov et al., 2000). However, cytogenetic analysis demonstrated that among them only A498 had 2 copies of chromosome 3 and four lines had from 3 to 5 copies (KRC/Y, HN51, CAKI1 and TK164). Nevertheless, all these 5 lines revealed only one marker in LOH analysis arguing that one 3p arm was lost.
Figure 4. Schematic maps of the LUCA and AP20 regions: pointed arrows, indicating the orientations of transcription, represent the genes. At the right side homozygous deletions are shown.

Table 1. Frequency of genomic aberrations in two 3p regions

<table>
<thead>
<tr>
<th>Cancer</th>
<th>Aberrations</th>
<th>Amplification or multiplication</th>
<th>Homozygous deletions</th>
<th>Hemizygous deletions</th>
<th>Total copy number changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ap20, probe NLJ-003, cases/percent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCC</td>
<td>18/34</td>
<td>8/15</td>
<td>18/34</td>
<td>44/83</td>
<td></td>
</tr>
<tr>
<td>BC</td>
<td>6/29</td>
<td>3/14</td>
<td>7/33</td>
<td>16/73</td>
<td></td>
</tr>
<tr>
<td>SCLC</td>
<td>4/17</td>
<td>3/13</td>
<td>14/61</td>
<td>21/91</td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>7/22</td>
<td>5/16</td>
<td>14/44</td>
<td>26/81</td>
<td></td>
</tr>
<tr>
<td>LUCA, probe NL3-001, cases/percent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCC</td>
<td>17/43</td>
<td>4/10</td>
<td>8/20</td>
<td>28/73</td>
<td></td>
</tr>
<tr>
<td>BC</td>
<td>5/23</td>
<td>4/18</td>
<td>9/42</td>
<td>18/82</td>
<td></td>
</tr>
<tr>
<td>SCLC</td>
<td>3/15</td>
<td>2/10</td>
<td>12/60</td>
<td>17/85</td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>9/28</td>
<td>5/16</td>
<td>10/31</td>
<td>24/75</td>
<td></td>
</tr>
</tbody>
</table>
Due to limited amount of tumor DNA it is hard to test amplified samples by Southern hybridization. However, multiplex PCR with microsatellite markers clearly proved that amplification indeed occurred in some samples and could be even seen in the same samples where other regions of 3p are deleted (Senchenko et al., 2004).

**Comparison of LOH and QPCR data. fine deletion mapping of homozygous deletions**

Real-time PCR data for NLJ-003 and NL3-001 loci were compared with allelic alterations observed for neighboring (and other 3p) polymorphic markers for RCC and BC samples. Twenty polymorphic markers were used in this work and correlation over 85% between real-time PCR and LOH data was observed in both loci. We considered results as inconsistent if nearest microsatellite markers from both sides showed results incompatible with QPCR data. These inconsistent results can be explained by different factors (Senchenko et al., 2003; 2004). For example, it could be explained by duplication of both chromosomes. The most important reason for disagreements is different locations of microsatellite and QPCR markers and occurrence of chromosomal breakpoints between the markers.

Cases with homozygous deletions were further mapped. Nineteen tumors having HD in AP20 region were additionally mapped with D3S1611, D3S2417 and D3S3623 and INTR1H22 (see Figure 4). The smallest region homozygously deleted in AP20 is flanked by D3S1298 and D3S3623. RCC cases ##27 and 35 showed a breakpoint at D3S1298. RCC sample #46 had a breakpoint at D3S2417 and #33 at D3S3623. BC cases 4 and 6 had telomeric breakpoint at D3S2417 and centromeric breakpoint at D3S1298. Lung cancer cell lines H1693 and H1688 had centromeric breakpoints at D3S1298.

Fifteen tumors with homozygous deletions in NL3-001 were mapped using additional markers: D3S1568, D3S4614, D3S2968, D3S4597, D3S4604, 5’SEMA5 and D3S1573 (see Figure 4). The smallest region homozygously deleted in LUCA is bordered by D3S1568 and D3S4604. Several samples, e.g. RCC ##33 and 8 proved that centromeric border is inside the CACNA2D2 gene. The telomeric border is inside the SEMA3F gene (D3S4604) for RCC #33 and BC ## 4 and 6.

**Complex 3p rearrangements in major epithelial cancers**

The study again highlighted the complex character of chromosomal rearrangements in 3p during development of epithelial tumors. In addition to homo- and hemizygous deletions, duplication/amplification of the whole chromosome 3 or some regions were frequently (15%-42.5%) detected. Thus, amplification of 3p is very common in studied cancers and probably in all major epithelial cancers. Therefore, microsatellite deletion data should be evaluated carefully as AI means not only LOH but also amplification. It is important to mention that not all cases of multiplication will be detected by microsatellite analysis as AI. Allele titration experiments revealed significant bias for the detection of LOH for the H (high molecular weight) allele compared to L (low molecular weight). In general, AI in cases of L-allele deletions can be detected only if the content of L-allele in DNA is less than 30% (Liu et al., 1999). Simple calculations for the chromosome duplication (two copies of L- and H-alleles) and triplication (three copies of each allele) cases as it was done in Liu et al. (Liu et al., 1999) showed that AI could not be revealed for the majority of cases if one or two copies of L-allele would be deleted respectively. One tentative explanation of these phenomena might be that inactivated TSG(s) were amplified together with neighboring proto-oncogenes/oncogenes.
Human chromosome 3p contains many genes that can play a dominant oncogenic role. MST1 receptor (RON) and its ligand MST1 gene are located in 3p21 and their overproduction can result in autocrine stimulation and uncontrolled proliferation (Angeloni et al., 2001). Therefore, in addition to the deletion of TSGs amplification of some 3p21.3 regions could reflect increased activity of resident potential oncogenes (Senchenko et al., 2003).

Six pairs of tumor/normal DNA samples (four BC and two RCC) that showed amplification of NL3-001 were analyzed with the SSCP method to look for mutations in the kinase domain of the receptor tyrosine kinase RON (Angeloni et al., 2003). This gene is located closely telomeric to the LUCA region. No mutations were found but single nucleotide polymorphisms (SNP) in intron 18 and exon 20. One sample was found heterozygous for an intronic C to A transversion of a SNP located 10 bp upstream the intron/exon splicing site between intron 18 and exon 19 consistent with amplification.

To understand the possible effect of this base change on the splicing process, we looked for possible splicing variants of exon 19 in a cell line homozygous for the C to A mutation. The lung cancer cell line NCI-H1650, out of several analyzed by SSCP, was found homozygous; mRNA extracted from these cells was retrotranscribed into cDNA and amplified with PCR primers specific for the region of interest (primers: ex18Fw and 5Rv, see (Angeloni et al., 2003)). The PCR gave two bands, one of the expected size (356 bp) and one between the 200 and 300 bp markers. This band was excised from gel, cloned and sequenced. This band had a sequence corresponding to an alternative splicing of RON cDNA lacking exon 19. One possible effect of this alternative splicing could consist in bringing Tyr1353 and Tyr1360 (the receptor docking site) closer to the activation loop. Samples 2, 7 and 11 carry a SNP in exon 20 (A4031G). Thus it is possible to suggest that amplified allele has some different features than the wild type allele and this could facilitate malignant progression.

Indeed, it was recently found that activated RON is involved in some lung cancers (Danilkovitch-Miagkova et al., 2003).

Another striking observation made in this study is the finding of very high level of homozygous deletions in LUCA and AP20 regions, amounting to more than 10%. As it was shown in our previous study the results of QPCR were highly reliable (Senchenko et al., 2003; 2004) and are concordant to the PCR and Southern results. QPCR data for NL3-003 and NL3-001 loci were combined with LOH data to precisely localize positions of candidate TSGs (Figure 4).

As it was mentioned above the smallest region homozygously deleted in LUCA is bordered by D3S1568 and D3S4604. Therefore, the centromeric border is inside the CACNA2D2 gene. The telomeric border is inside the SEMA3F gene for all 4 different types of tested carcinomas. This implies that genes located centromeric to CACNA2D2 and telomeric to SEMA3F (Lerman and Minna, 2000) could be excluded from the list of candidate TSG(s).

The results confirmed earlier suggestion that both SEMAF and SEMAB genes should be considered as candidate TSGs (Zabarovsky et al., 2002, Senchenko et al., 2003).

The smallest region homozygously deleted in AP20 was flanked by D3S1298 and D3S3623. Thus, only four known genes could still be considered as candidate TSGs: APRG1, ITGA9, HYA22 and VILL. This analysis excludes DLEC1 and MYD88 (Protopopov et al., 2003, Daigo et al., 1999) as candidate TSG(s) involved in the studied cancers.

Comparing results of this and a previous study (Senchenko et al., 2003; 2004) with that of Daigo et al. (Daigo et al., 1999), who studied homozygous deletion in SCLC cell line
ACC-LC5, several conclusions can be drawn. Homozygous deletions around NLJ-003 are common in different MEC like SCLC, NSCLC, RCC, BC and CC. After deletion mapping of ACC-LC5 three genes were suggested as candidate TSGs: GOLGA4, ITGA9 and HYA22, however none of them had TSG function demonstrated. Our data excluded GOLGA4 from the candidates but add APRG1 and VILL genes (Protopopov et al., 2003). Our preliminary data suggested that sequences for HYA22 and VILL genes were determined incorrectly and some of the splicing forms of HYA22 gene have growth inhibiting activity both in vitro and in vivo. Additional studies are needed to clarify which of the candidate genes has TSG activity.

We showed earlier (Protopopov et al., 2003) that the AP20 region is heavily methylated in all studied RCC cell lines suggesting that hypermethylation of TSG(s) in this region may play a critical role similar to the situation in the LUCA region. Moreover, genes in the region have multiple splicing forms and both mutation or functional analyses of such genes represent a difficult task. Thus AP20 region may be similar to LUCA region and contains several TSGs.

It was suggested that aberrations in the LUCA and AP20 regions could be causatively linked (Senchenko et al., 2003). Indeed, homozygous deletions in both regions often happened in the same tumor (P<3x10^{-7}). The estimation of possible interdependency between all aberrations in loci NLJ-003 and NL3-001 as different events was carried out using a permutation test for all four types of studied carcinomas. This test also revealed a significant correlation between different aberrations in these two loci (P<10^{-6}). The same results were obtained using Pearson correlation for numeric values of copy number changes of these loci. Thus these analyses demonstrated that aberrations in both LUCA and AP20 regions occurred simultaneously in the same tumor with high probability.

c) NotI-Microarrays to Identify Cancer Causing Genes and Develop Biomarkers for Early Diagnosis, Prognosis and Prediction of Major Types of Cancer

The fundamental problems for genome wide screening using NotI clones are: (i) the size and complexity of the human genome; (ii) the number of repetitive sequences; and (iii) the small size of the NotI oligonucleotides (55 bp). To solve this problem, the special procedure was developed to amplify only regions surrounding NotI sites, so called NotI representation (NR). Other DNA fragments were not amplified. We suggested the use of NMA for genome screening in combination with this new method for labeling genomic DNA where only sequences surrounding NotI sites are labeled. Experiments demonstrated high sensitivity of the method. For example, we tested epithelial/mesenchymal transition of vulval A431 cells. Changes in methylation pattern were already detected after 48 hrs DOX+ activation of transcriptional factor SIP1. These changes were confirmed by bisulphite sequencing and MSP of 3 genes: THRB, GATA2 and MINT24. Thus for the first time we demonstrated that changes in methylation of genes can happen during 48hrs. The labeling procedure is as follows. DNA is digested with NotI and linkers labeled with biotin are ligated to the NotI sticky ends. DNA is digested with Sau3A (or/and other enzymes) and NotI-Sau3A fragments are purified from other genomic fragments with magnetic beads. Afterwards linkers to Sau3A ends are added and NotI-Sau3A fragments are amplified, labeled and hybridized to microarrays containing NotI DNA fragments.
The main advantages of NMA compared to other methylation microarrays: all clones represent genes; incomplete digestion doesn’t generate artificial results; only 0.01-0.05% of the total DNA is labeled increasing sensitivity several thousands times compared to other approaches where total human DNA is labeled; NR probe contains 10-fold less repeats than total human DNA; homo- and hemizygous deletions can be detected; homo- and hemizygous methylation and demethylation can be detected.

Another main difference of our approach is that we label only digested (unmethylated) DNA and in other approaches only undigested (methylated) DNA is labeled. Moreover, we can detect methylated genes without normal control using only original tumor DNA and tumor DNA after the whole genome amplification (WGA) with Phi29 when methylation is removed. NotI microarrays are the only existing microarrays affording the opportunity to detect methylation and copy number changes simultaneously or differentially (Li et al., 2002). As it was mentioned above there is no reason why NMA cannot be used to study histone modifications and we are currently performing these experiments.

Our estimation is that human genome contains 10,000-15,000 NotI sites and 5,000-9,000 of them are unmethylated in a particular cell. At present we use NMA for human chr.3 (190 NotI containing genes).

Chromosome 3 specific NMA containing 181 NotI linking clones associated with genes were hybridized with more than 300 samples representing various cancers: breast, kidney, cervical, colon, ovarian, prostate, leukemia and lung. The experiments showed that all cancers have specific methylation pattern and still each tumor sample had unique pattern. For all studied cancers we found genes specifically methylated in malignant cells. Many genes were found to be methylated in a very high fraction of cancer samples (more than 80%). These genes can be divided in two classes: cancer specific and common for several types of cancer. Some examples of genes/EST involved in several cancers: MINT24 (AF135524), BHLHB2, LOC285205, NBEAL2 (KIAA0540), FLJ44898 (AK126846), GATA-2, RARbeta1, RBSP3 (CTDSPL), VHL. In addition to known methylated genes (like VHL, RARB1) novel candidate genes were found (CACNA2D3, ITGA9, GORASP1, other) that were unknown previously to be involved in the development of cancer. Methylation of 25 genes was confirmed by methyl specific PCR and bisulfite sequencing in more than 100 tumor samples. Methylation status of the genes correlated with expression (10 genes were confirmed using qPCR). Some genes were tested functionally and demonstrated growth inhibiting activity proving that NMA are efficient instrument to discover cancer causing genes and especially tumor suppressor genes (see below). Genes specifically demethylated in tumors were also found (e.g. in breast cancer: SH3BP5, PRO2730, LOC131973, LOC131961, LOC440953). They could represent oncogene-like genes.

We tested NMA also in other model systems. The aim was to test DNA (cytosine-5-) methyltransferases action. We compared colon cancer cell line HCT116 wt with: (1) HCT116 DNMT1 knock out; (2) HCT116 DNMT3b knock out and (3) HCT116 DNMT1 and DNMT3b double knock out (DKO) cell lines and found demethylated genes in all experiments (done in three replicates). For DKO cell lines we could compare our data with SAGE data (Dr. J. Li, collaborator from NCI, Frederick, USA; personal communication) The result showed the demethylation is correlated with the increased expression.

We started to perform more deep analysis of our data trying to find good markers for cancer and progression detection. The first results with cervical cancer look very promising and we can see different methylation patterns at different tumor stages. We decided to check
one of the genes (RBSP3) showed increasing methylation during the progression using qPCR and expression data confirmed this suggestion. With lung cancer we obtained similar data.

We tested 20 prostate samples and detected genes that were not only cancer specifically methylated but also allow to discriminate aggressive from non-aggressive forms of prostate cancer (e.g., LRRC3B, RBSP3, GLCE, ITGA9, other). This discrimination is one of the major problems with prostate cancer (Figure 5).

Previously we suggested, that both RASSF1 (TSG from the 3p21.3 LUCA region) and RBSP3 (TSG from 3p21.3 AP20 region) could cooperate with each other in induction of cell cycle arrest: the former by inhibiting cyclin D1 and the latter by activating RB protein. This could explain frequent homozygous deletions both in LUCA and in AP20 regions in the same tumor (P<3x10^{-7}) and support the hypothesis that TSGs in these two regions may have a synergistic effect (Senchenko et al., 2003, 2004, Kashuba et al., 2004).

With NMA we again detected that these regions showed similar results and were frequently methylated in the same tumor samples.

![Figure 5. Representative picture showing hybridization of prostate samples to NMA. Green boxes show increased methylation in cancers. Yellow, no change and white are not informative. AD – adenoma, N/agr – non-aggressive cancer, Agr – aggressive cancer. Highlighted genes shows good discrimination between these 3 types of prostate cancer development](image-url)
Quantitative real-time PCR was used to measure the mRNA level for tumor-suppressor and candidate genes of 3p21.3 (RBSP3/CTDSPL, NPR2/G21, RASSF1A, ITGA9, HYAL1 and HYAL2) in major types of non-small cell lung cancer (NSCLC): squamous cell lung cancer (SCC) and lung adenocarcinoma (AC). A significant (2- to 100-fold) and frequent (44–100%) decrease in mRNA levels was observed in NSCLC. The mRNA level decrease and its frequency depended on the histological type of NSCLC for all genes. The downregulation of RASSF1A and ITGA9 was significantly associated with AC progression; the same tendency was observed for RBSP3/CTDSPL, NPR2/G21, HYAL1, and HYAL2. In SCC, the downregulation of all genes was not associated with the clinical stage, tumor cells differentiation, and metastasis in lymph nodes. The RBSP3/CTDSPL, NPR2/G21, ITGA9, HYAL1 and HYAL2 mRNA levels significantly (5- to 13-fold on average) decreased at a high frequency (83–100%) as early as SCC Stage I. Simultaneous downregulation of all six genes was observed in majority of tumor samples and was independent of the gene position in 3p21.3 and the functions of the protein products. The Spearman correlation coefficient rs was 0.63–0.91, p < 0.001. The highest rs values were obtained for gene pairs ITGA9–HYAL2 and HYAL1–HYAL2, whose products mediate cell–cell adhesion and cell–matrix interactions.

Coregulation of the genes was assumed on this basis. Both genetic and epigenetic mechanisms proved to be important for downregulation of RBSP3/CTDSPL and ITGA9. This finding supported the hypothesis that the clusters of cancer related genes in the extended 3p21.3 locus are simultaneously inactivated during the development and progression of lung cancer and other epithelial tumors. A significant and frequent decrease in the mRNA level of the six genes in SCC could be important for developing specific biomarker sets for early SCC diagnosis and new approaches to cure NSCLC.

In summary we selected set of 23 markers (BHLHB2, FBLN2, EPHB1, GATA2, GORASP1, PRICKLE2, Hmm61490, ITGA9, LOC285205, LRRC3B, MINA, MITF, MRPS17P3, NKIRAS1, PLCL2, TRH, UBE2E2, WNT7A, RARB2, p20-CGGBP, GNAI2, RPL32, THRB) that would allow to discriminate/diagnose the majority of NSCLC cases with probability more than 95% and most complicated cases with probability more than 80%.

d) Identification of Novel Tumor Suppressor Genes

The most widely used approach to designate a candidate gene as being a TSG, is by positive demonstration of inactivating mutations of this gene in tumour biopsies. Even such a simple requirement for mutational inactivation can be complicated in cases with dominant negative mutations, LOH or heterozygous mutations in genes whose dosage is critical, and in genes in which one allele is imprinted (Haber & Harlow, 1997; Nicolaides et al., 1998). In such cases, only functional approaches that can demonstrate tumour suppression activity for one of the candidate genes can help to solve the problem. Evidence of the ability to suppress cellular proliferation cannot always be used for the identification of a tumor suppressor. In many tumours carrying multiple genetic alterations, reconstituting a tumor suppressor gene is not sufficient to reverse completely their malignant phenotype. For example, in the study of Zhou et al. (1994), it was demonstrated that Rb-reconstituted osteosarcoma Saos 2 cells were still tumorigenic in nude mice. It must be stressed that reconstitution experiments designed to demonstrate direct growth inhibition of tumors are difficult to perform and analyse.

In collaboration with Dr. Imreh's group, we have previously observed the non-random loss of human 3p21-p22 fragments from mouse-human microcell hybrids following
progressive growth in SCID (Imreh et al., 1994). We have taken this to indicate that tumor growth antagonizing or suppressor gene(s) were located in the deleted region. We decided to address the question by asking whether a known suppressor gene, RB, would show corresponding behaviour and whether this feature can be developed into a functional test for the identification of TSGs. Wild type and mutated RB genes were built into appropriate constructs that permitted the expression of the gene in the absence but not the presence of tetracycline. To achieve this, a new retroviral vector, pLNCtTA, constitutively expressing tetracycline responsive transcriptional activator protein (tTA) was constructed and an A9 cell line expressing tTA was created. A new vector, pETI (Elimination Test Integrating vector), that provided hygromycin resistance in the transfected cells, and allowed tetracycline regulated expression of the inserted gene was also made. Transfected A9 mouse fibrosarcoma cells were grown in SCID mice in the presence or absence of tetracycline. Wild type RB but not deleted RB was deleted/inactivated during tumour growth (Li et al., 1999).

From these experiments, we suggested a new functional test for TSG identification - GIT. This test is based on the functional inactivation of the analysed genes in contrast to existing tests based on growth suppression (Killary et al., 1992; Kok et al., 1997; Protopopov et al., 2002).

We have also tested p53 and 3PK genes (Sithanandam et al., 1996) in this system. Again inactivation of p53 was observed but the 3PK gene that was used as a control was expressed and present after two passages in SCID mice.

We selected a set of genes to be tested using GIT by introducing them within tetracycline regulatable constructs into human cancer cell lines. Cells carrying the construct were selected and tested for growth in vitro and in vivo, under conditions when the gene of interest was expressed or silenced.

Using this test we have tested 16 genes from 3p21.3. Eight genes (TCEA1, MLH1, RHOA, 3PK, PL6, 101F6, BLU, TGFB) did not show any effect in the tested cell lines. Six genes (RBSP3, NPRL2/G21, RASSF1A, RASSF1C, SEMA3B, SEMA3F) had strong inhibitory activity, both in vitro and in vivo. Two genes, HYAL1 and HYAL2 showed strong inhibitory activity in vivo in SCID mice but not in vitro, as judged by colony formation inhibition and growth curve assays.

**Conclusion**

Altogether using GIT we identified 12 cancer causing genes in AP20 and LUCA regions (see Figure 4): RBSP3, ITGA9, MLH1, VILL, APRG1, RASSF1, HYAL1, HYAL2, SEMA3B, SEMA3F, NPRL2 and CACNA2D2.

Moreover we found an exceptionally high incidence of single-base mutations in the tumor suppressor genes RASSF1 and RBSP3 (CTDSPL) in major epithelial tumors. These mutations functionally inactivated tumor suppressor activity of these genes. Somatic hypermutations in tumor suppressor genes involved in major human malignancies offer a novel insight in cancer development, progression and spread (Kashuba et al., 2009). Thus our results demonstrated that positional cloning strategy involving genetic, genomic, epigenetic mapping and functional analysis is a powerful tool to find methylated genes/tumor suppressor genes and resulted in identification of many novel genes/biomarkers that can be important for
the development of specific biomarker sets for early diagnosis, prognosis and new approaches to therapy.

**Acknowledgments**

The results summarized in this review were obtained under the support of the Russian Foundation for Basic Research (EAB 10-04-01213; VNS 08-04-01577) and INTAS (03-51-4983). ERZ was supported by research grants from the Swedish Cancer Society, the Swedish Research Council, the Swedish Foundation for International Cooperation in Research and Higher Education (STINT), the Swedish Institute, and Karolinska Institute. The work was also supported by State Contracts Nos. 02.740.11.5227 and 16.740.11.0173, Ministry of Education and Science of the Russian Federation (2010-2011)

**References**

Allikmets, RL; Kashuba, VI; Bannikov, VM; Petrov, N; Pettersson, B; Lebedeva, T; Gizatullin, R; Zakharyev, VM; Winberg, G; Dean, M; Uhlen, M; Kisselev, LL; Klein, G; Zabarovsky, ER. NotI linking clones as tools to join physical and genetic mapping of the human genome. *Genomics*, 1994, 19, 303-309.

Alimov, A; Kost-Alimova, M; Liu, J; Li, C; Bergerheim, U; Imreh, S; Klein, G; Zabarovsky, ER. Combined LOH/CGH analysis proves the existence of interstitial 3p deletions in renal cell carcinoma. *Oncogene*, 2000, 19, 1392-1399.

Angeloni, D; Duh, FM; Moody, M; Dean, M; Zabarovsky, ER; Sentchenko, V; Braga, E; Lerman, MI. C to A single nucleotide polymorphism in intron 18 of the human MST1R (RON) gene that maps at 3p21.3. *Mol. Cell Probes*, 2003, 17, 55-57.


Braga, EA; Kotova, EJ; Pugacheva, EM; Gizatullin, RZ; Kashuba, VI; Lerman, MI; Aksenova, MG; Efremov, IA; Zabarovsky, ER; Kisselev, LL. Human chromosome 3: detection of microsatellites of di-, tri-, and tetra-motifs, analysis of its allocation and markers formation. *Mol. Biol.*, (Moscow), 1997, 31, 985-996.


Braga, E; Senchenko, V; Bazov, I; Loginov, W; Liu, J; Ermilova, V; Kazubskaya, T; Garkavtseva, R; Mazurenko, N; Kisseliov, F; Lerman, M; Klein, G; Kisselev, L; Zabarovsky, ER. Critical tumor-suppressor gene regions on chromosome 3P in major human epithelial malignancies: allelotyping and quantitative real-time PCR. *Int. J. Cancer*, 2002, 100, 534-541.

Daigo, Y; Isomura, M; Nishiwaki, T; Tamari, M; Ishikawa, S; Kai, M; Murata, Y; Takeuchi, K; Yamane, Y; Hayashi, R; Minami, M; Fujino, MA; Hojo, Y; Uchiyama, I; Takagi, T;

Daly, MC; Xiang, RH; Buchhagen, D; Hensel, CH; Garcia, DK; Killary, AM; Minna, JD; Naylor, SL. A homozygous deletion on chromosome 3 in a small cell lung cancer cell line correlates with a region of tumor suppressor activity. *Oncogene*, 1993, 8, 1721-1729.

Dammann, R; Schagdarsurengin, U; Strunnikova, M; Rastetter, M; Seidel, C; Liu, L; Tommasi, S; Pfeifer, GP. Epigenetic inactivation of the Ras-association domain family 1 (RASSF1A) gene and its function in human carcinogenesis. *Histol. Histopathol.*, 2003, 18, 665-677.

Danilkovitch-Miagkova, A; Duh, FM; Kuzmin, I; Angeloni, D; Liu, SL; Miller, AD; Lerman, MI. Hyaluronidase 2 negatively regulates RON receptor tyrosine kinase and mediates transformation of epithelial cells by jaagsiekte sheep retrovirus. *Proc. Natl. Acad. Sci.*, USA, 2003, 100, 4580-4585.


DeSalle, LM; Latres, E; Lin, D; Graner, E; Montagnoli, A; Baker, RT; Pagano, M; Loda, M. The de-ubiquitinating enzyme Unp interacts with the retinoblastoma protein. *Oncogene*, 2001, 20, 5538-5542.

Dreijerink, K; Braga, E; Kuzmin, I; Geil, L; Duh, FM; Angeloni, D; Zbar, B; Lerman, MI; Stanbridge, EJ; Minna, JD; Protopopov, A; Li, J; Kashuba, V; Klein, G; Zabarovsky, E. The candidate tumor suppressor gene, RASSF1A, from human chromosome 3p21.3 is involved in kidney tumorigenesis. *Proc. Natl. Acad. Sci.*, USA, 2001, 98, 7504-7509.


Hu, Li-Fu; Eiriksdottir, G; Lebedeva, T; Kholodnyuk, I; Alimov, A; Chen, F; Luo, Y; Zabarovsky, ER; Ingvarsson, S; Klein, G; Emborg, I. Allelic deletion on chromosome 3 in primary nasopharyngeal carcinoma. *Genes, Chromosomes & Cancer*, 1996, 17, 118-126.


Kashuba, VI; Gizatullin, RZ; Protopopov, AI; Li, J; Vorobieva, NV; Fedorova, L; Zabarovska, VI; Muravenko, OV; Kost-Alimova, M; Dominsky, DA; Kiss, C; Allikmets, R. Analysis of NotI linking clones isolated from human chromosome 3 specific libraries. *Gene*, 1999, 239, 259-271.

Kashuba, VI; Li, J; Wang, F; Senchenko, VN; Protopopov, A; Malyskova, A; Kutsenko AS; Kadyrova, E; Zabarovska, V; Muravenko, O; et al. RBSP3 (HYA22) is a tumor suppressor gene implicated in major epithelial malignancies. *Proc Natl Acad Sci U S A.*, 2004, 101, 4906-4911.

Kholodnyuk, I; Kost-Alimova, M; Kashuba, VI; Gizatullin, RZ; Stanbridge, E; Szeles, A; Zabarovsky, ER; Klein, G; Imreh, S. The region of 3p22-p21.3 is non-randomly eliminated from mouse human microcell hybrids during tumor growth in SCID mice. *Genes Chromosomes Cancer*, 1997, 18, 200-211.


Kutsenko, AS; Gizatullin, RZ; Al-Amin, AN; Wang, F; Kvasha, SM; Podowski, RM; Matushkin, YG; Gyanchandani, A; Muravenko, OV; Levitsky, VG; et al. NotI flanking sequences: a tool for gene discovery and verification of the human genome. *Nuc Acids Res.*, 2002, 30, 3163-3170.

Latif, F; Tory, K; Gnarra, J; Yao, M; Duh, FM; Orcutt, ML; Stackhouse, T; Kuzmin, I; Modi, W; Geil, L; et al. Identification of the von Hippel-Lindau disease tumor suppressor gene. *Science*, 1993, 260, 1317-1320.


Lee, KK; Ohyama, T; Yajima, N; Tsubuki, S; Yonehara, S. MST, a physiological caspase substrate, highly sensitizes apoptosis both upstream and downstream of caspase activation. *J. Biol. Chem.*, 2001, 276, 19276-19285.


Li, J; Protopopov, A; Wang, F; Sentchenko, 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, E. NotI subtraction and NotI-specific microarrays to detect copy number and methylation changes in whole genomes. *Proc. Natl. Acad. Sci. USA*, 2002, 99, 10724-10729.

Li, J; Wang, F; Protopopov, A; Malyukova, A; Kashuba, V; Minna, JD; Lerman, MI; Klein, G; Zabarovsky, E. Inactivation of RASSF1C during in vivo tumor growth identifies it as a tumor suppressor gene. Oncogene, 2004a, 23, 5941-5949.

Liu, J; Zabarovska, VI; Braga, E; Alimov, A; Klein, G; Zabarovsky, ER. Loss of heterozygosity in tumor cells requires re-evaluation: the data are biased by the size-dependent differential sensitivity of allele detection. FEBS Lett., 1999, 462, 121-128.


Murata, Y; Tamari, M; Takahashi, T; Hori, Y; Hibi, K; Yokoyama, S; Inazawa, J; Yamakawa, K; Ogawa, A; Takahashi, T; Nakamura, Y. Characterization of an 800 kb region at 3p22-p21.3 that was homozygously deleted in a lung cancer cell line. Hum Mol Genet., 1994, 3, 1341-1344.


Palmisano, WA; Divine, KK; Saccomanno, G; Gilliland, FD; Baylin, SB; Herman, JG; Belinsky, SA. Predicting lung cancer by detecting aberrant promoter methylation in sputum. Cancer Res., 2000, 60, 5954-5958.

Peace, BE; Hughes, MJ; Degen, SJF; Waltz, SE. Point mutations and overexpression of Ron induce transformation, tumor formation, and metastasis. Oncogene, 2001, 20, 6142-6151.

Pille, JY; Denoyelle, C; Varet, J; Bertrand, JR; Soria, J; Opolon, P; Lu, H; Pritchard, LL; Vannier, JP; Malvy, C; Soria, C; Li, H. Anti-RhoA and anti-RhoC siRNAs inhibit the proliferation and invasiveness of MDA-MB-231 breast cancer cells in vitro and in vivo. Mol Ther., 2005, 11, 267-274.


Protopopov, A; Kashuba, V; Zabarovska, VI; Muravenko, OV; Lerman, MI; Klein, G; Zabarovsky, ER. An integrated physical and gene map of the 3.5-Mb chromosome 3p21.3 (AP20) region implicated in major human epithelial malignancies. Cancer Res., 2003, 63, 404-412.

Rimessi, P; Gualandi, F; Morelli, C; Trabanelli, C; Wu, Q; Possati, L; Montesi, M; Barrett, JC; Barbanti-Brodano, G. Transfer of human chromosome 3 to an ovarian carcinoma cell line identifies three regions on 3p involved in ovarian cancer. Oncogene, 1994, 9, 3467-3474.

Roche, J; Boldog, F; Robinson, M; Robinson, L; Varella-Garcia, M; Swanton, M; Waggoner, B; Fishel, R; Franklin, W; Gemmill, R; Drabkin, H. Distinct 3p21.3 deletions in lung cancer and identification of a new human semaphorin. Oncogene, 1996, 12, 1289-1297.

Sgambato, A; Migaldi, M; Montanari, M; Camerini, A; Brancaccio, A; Rossi, G; Cangiano, R; Losasso, C; Capelli, G; Trentini, GP; Cittadini, A. Dystroglycan expression is frequently reduced in human breast and colon cancers and is associated with tumor progression. Am. J. Pathol., 2003, 162, 849-860.
Senchenko, V; Liu, J; Braga, E; Mazurenko, N; Loginov, W; Seryogin, Y; Bazov, I; Protopopov, A; Kisseljov, FL; Kashuba, V; Lerman, MI; Klein, G; Zabarovsky, ER. Deletion mapping using quantitative real-time PCR identifies two distinct 3p21.3 regions affected in most cervical carcinomas. Oncogene, 2003, 22, 2984-2992.

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

Shames, DS; Girard, L; Gao, B; Sato, M; Lewis, CM; Shivapurkar, N; Jiang, A; Perou, CM; Kim, YH; Pollack, JR; Fong, KM; Lam, CL; Wong, M; Shyr, Y; Nanda, R; Olopade, OI; Gerald, W; Euhus, DM; Shay, JW; Gazdar, AF; Minna, JD. A genome-wide screen for promoter methylation in lung cancer identifies novel methylation markers for multiple malignancies. PLoS Med., 2006, 3, 1-20.


Sithanandam, G; Latif, F; Duh, FM; Bernal, R; Smola, U; Li, H; Kuzmin, I; Wixler, V; Geil, L; Shrestha, S; Lloyd, PA; Bader, S; Sekido, Y; Tartof, KD; Kashuba, VI; Zarabovsky, ER; Dean, M; Klein, G; Lerman, MI; Minna, JD; Rapp, UR; Allikmets, R. 3PK, a new mitogen-activated protein kinase - activated protein kinase located in the small cell lung cancer tumor suppressor gene region. Molec. Cell Biology, 1996, 16, 868-876.


Sulimova, GE; Kutsenko, AS; Rakhmanaliev, ER; Udina, IG; Kompaniytsev, AA; Protopopov, AI; Moisjak, EV; Klimov, EA; Muravenko, OV; Zelenin, AV; Braga, EA; Kashuba, VI; et al. Human chromosome 3: integration of 60 NotI clones into a physical and gene map. Cytogenet Genome Res., 2002, 98, 177-183.

Sulimova, GE; Rakhmanaliev, ER; Klimov, EA; Kompaniytsev, AA; Udina, IG; Zarabovsky, ER; Kisselev, LL. Notl-STS markers for human chromosome 3 are gene markers. Mol. Biol. (Moscow), 2005, 39, 593-607.


van den Berg, A; Dijkstra, T; Draaijers, TG; Hulsbeek, MM; Maher, ER; van den Berg, E; Störkel, S; Buys, CH. Analysis of multiple renal cell adenomas and carcinomas suggests allelic loss at 3p21 to be a prerequisite for malignant development. Genes Chromosomes Cancer, 1997, 19, 228-232.

Vos, MD; Martinez, A; Elam, C; Dallol, A; Taylor, BJ; Latif, F; Clark, GJ. A role for RASSF1A tumor suppressor in the regulation of tubulin polymerization and genomic stability. Cancer Res., 2004, 64, 4244-4250.

Wei, MH; Latif, F; Bader, F; Kashuba, V; Chen, JJ; Duh, FM; Sekido, Y; Lee, CC; Geil, L; Kuzmin, I; Zarabovsky, 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


Yan, PS; Chen, CM; Shi, H; Rahmatpanah, F; Wei, SH; Caldwell, CW; Huang, TH. Dissecting complex epigenetic alterations in breast cancer using CpG island microarrays. *Cancer Res.*, 2001, 61, 8375-8380.

Zabarovsky, ER; Boldog, F; Thompson, T; Scanlon, D; Winberg, G; Marcsek, Z; Erlandsson, R; Stanbridge, EJ; Klein, G; Sumegi, J. Construction of a human chromosome 3 specific NotI linking library using a novel cloning procedure. *Nucl. Acids Res.*, 1990, 18, 6319.

Zabarovsky, ER; Lerman, MI; Minna, JD. Tumor suppressor genes on chromosome 3p involved in the pathogenesis of lung and other cancers. *Oncogene*, 2002, 21, 6915-6935.

Zabarovsky, ER; Kashuba, VI; Zakharyev, VM; Petrov, N; Pettersson, B; Lebedeva, T; Gizatullin, R; Pokrovskaya, ES; Bannikov, VM; Zabarovska, VI; Allikmets, RL; Stanbridge, EJ; Winberg, G; Uhlen, M; Kisselev, LL; Klein, G. Shot-gun sequencing strategy for long range genome mapping: first results. *Genomics*, 1994, 21, 495-500.