UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 550

Application of Genomic and Expression Arrays for Identification of new Cancer Genes

HELENA NORD

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2010

ISSN 1651-6206
ISBN 978-91-554-7775-2
urn:nbn:se:uu:diva-121957

Dissertation presented at Uppsala University to be publicly examined in Rudbecksalen, Rudbecklaboratoriet, Dag Hammarskjölds väg 20, Uppsala, Tuesday, May 18, 2010 at 09:15 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in English.

Abstract
Nord, H. 2010. Application of Genomic and Expression Arrays for Identification of new Cancer Genes. Acta Universitatis Upsaliensis. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine* 550. 61 pp. Uppsala. ISBN 978-91-554-7775-2.

Copy number variation (CNV) comprises a recently discovered kind of variation involving deletion and duplication of DNA segments of variable size, ranging from a few hundred basepairs to several million. By altering gene dosage levels or disrupting proximal or distant regulatory elements CNVs create human diversity. They represent also an important factor in human evolution and play a role in many disorders including cancer. Array-based comparative genomic hybridization as well as expression arrays are powerful and suitable methods for determination of copy number variations or gene expression changes in the human genome. In paper I we established a 32K clone-based genomic array, covering 99% of the current assembly of the human genome with high resolution and applied it in the profiling of 71 healthy individuals from three ethnic groups. Novel and previously reported CNVs, involving ~3.5% of the genome, were identified. Interestingly, 87% of the detected CNV regions overlapped with known genes indicating that they probably have phenotypic consequences. In papers II through IV we applied this platform to different tumor types, namely two collections of brain tumors, glioblastoma (paper II) and medulloblastoma (paper III), and a set of bladder carcinoma (paper IV) to identify chromosomal alterations at the level of DNA copy number that could be related to tumor initiation/progression. Tumors of the central nervous system represent a heterogeneous group of both benign and malignant neoplasms that affect both children and adults. Glioblastoma and medulloblastoma are two malignant forms. Glioblastoma often affects adults while the embryonal tumor medulloblastoma is the most common malignant brain tumor among children. The detailed profiling of 78 glioblastomas, allowed us to identify a complex pattern of aberrations including frequent and high copy number amplicons (detected in 79% of samples) as well as a number of homozygously deleted loci. These regions encompassed not only previously reported oncogenes and tumor suppressor genes but also numerous novel genes. In paper III, a subset of 26 medulloblastomas was analyzed using the same genomic array. We observed that alterations involving chromosome 17, especially isochromosome 17q, were the most common genomic aberrations in this tumor type, but copy number alterations involving other chromosomes: 1, 7 and 8 were also frequent. Focal amplifications, on chromosome 1 and 3, not previously described, were also detected. These loci may encompass novel genes involved in medulloblastoma development. In paper IV we examined for the presence of DNA copy number alterations and their effect on gene expression in a subset of 21 well-characterized Ta bladder carcinomas, selected for the presence or absence of recurrences. We identified a number of novel genes as well as a significant association between amplifications and high-grade and recurrent tumors which might be clinically useful.

The results derived from these studies increase our understanding of the genetic alterations leading to the development of these tumor forms and point out candidate genes that may be used in future as targets for new diagnostic and therapeutic strategies.

Keywords: Array-CGH, Expression array, Copy number variation, Glioblastoma, Medulloblastoma, Bladder carcinoma, Oncogenes, Tumor suppressor genes

Helena Nord, Genomics, Rudbecklaboratoriet, Uppsala University, SE-75185 Uppsala, Sweden

ISSN 1651-6206
ISBN 978-91-554-7775-2
urn:nbn:se:uu:diva-121957 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-121957)

Till Patrik och Tindra

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I Díaz de Ståhl T, Sandgren J, Piotrowski A, **Nord H**, Andersson R, Menzel U, Bogdan A, Thuresson A-C, Poplawski A, von Tell D, Hansson C.M, Elshafie A.I, ElGhazali G, Imreh S, Nordenskjöld M, Upadhyaya M, Komorowski J, Bruder C.E.G, Dumanski J.P. Profiling of copy number variations (CNVs) in healthy individuals from three ethnic groups using a human genome 32 K BAC-clone-based array. *Hum Mutat. 2008 Mar;29(3):398-408.*

II **Nord H**, Hartmann C, Andersson R, Menzel U, Pfeifer S, Piotrowski A, Bogdan A, Kloc W, Sandgren J, Olofsson T, Hesselager G, Blomquist E, Komorowski J, von Deimling A, Bruder C.E.G , Dumanski J.P and Díaz de Ståhl T. Characterization of novel and complex genomic aberrations in glioblastoma using a 32K BAC array. *Neuro Oncol. 2009 Dec;11(6):803-18.*

III **Nord H**, Pfeifer S, Nistér M, Nilsson P, Strömberg B, Dumanski J.P, Alafuzoff I and Díaz de Ståhl T. Novel amplicons in pediatric medulloblastoma identified by high-resolution genomic analysis. *Manuscript*.

IV **Nord H**, Segersten U, Sandgren J, Wester K, Busch C, Menzel U, Komorowski J, Dumanski J.P, Malmström P-U and Díaz de Ståhl T. Focal amplifications correlate with high-grade and recurrences in stage Ta bladder carcinoma. *Int J Cancer. 2010 Mar 15;126(6):1390-402.*

Related papers by the author

i Andersson R, Bruder C.E.G , Piotrowski A, Menzel U, **Nord H**, Sandgren J, Hvidsten T.G, Díaz de Ståhl T, Dumanski J.P, Komorowski J. A Segmental Maximum A Posteriori Approach to Genome-wide Copy Number Profiling. *Bioinformatics 2008;24(6):751-8.*

ii Sandgren J, Díaz de Ståhl T, Andersson R, Menzel U, Piotrowski A, **Nord H**, Bäckdahl M, Kiss N, Braukhoff M, Dralle H, Hessman O, Larsson C, Åkerström G, Bruder C.E.G, Dumanski J.P, Westin G. Recurrent genomic alterations in sporadic benign and malignant pheochromocytomas and paragangliomas revealed by whole-genome array-CGH analysis. *Submitted manuscript*.

Contents

Abbreviations

Array-CGH	Array-comparative genomic hybridization
AT/RT	Atypical Teratoid/Rhabdoid tumor
BAC	Bacterial Artificial Chromosome
BER	Base-excision repair
bp	Base pair
cDNA	Complementary DNA
CNS	Central nervous system
CNV	Copy number variation
CSC	Cancer stem cell
DNA	Deoxyribonucleic acid
DWH	Data ware house
FISH	Fluorescent in situ hybridization
FoSTeS	Fork stalling and template switching
kb	Kilobase
LINE	Long interspersed nuclear element
LOH	Loss of heterozygosity
Mb	Megabase
MMR	Mismatch repair
MOR	Minimal overlapping region
mRNA	Messenger ribonucleic acid
NAHR	Non-allelic homologous recombination
NER	Nucleotide-excision repair
NHEJ	Non-homologous end joining
PCR	Polymerase chain reaction
PNET	Primitive neuro-ectodermal tumor
qPCR	Quantitative PCR
RT-PCR	Real-time PCR
SMAP	A segmental maximum a posteriori
SNP	Single nucleotide polymorphism
TSG	Tumor suppressor gene
WHO	World Health Organization

Introduction

Genetic variation

Human genetic variation is defined as the extent of genetic differences between human genomes, and this variation can be observed within a single subject, between individuals and between different human populations. The characterization of human genetic variation has advanced over the last few years thanks to improvements in DNA analysis technology. Techniques such as microarrays and massively parallel DNA sequencing has allowed for analysis of the entire human genome at very high resolution. Many studies have revealed the extent of genetic variation, and also mapped the exact position of the observed genetic differences. Genetic variations range from single nucleotide changes to large, microscopically visible chromosome anomalies (1-4).

One of the best explored types of variation is at the level of single nucleotides, so called single nucleotide polymorphisms (SNPs). This variation occurs when one base (A, T, C or G) in the DNA sequence differs between individuals, or between homologous chromosomes in an individual. In order to be categorized as a SNP, the sequence variant must present with a frequency of greater than 1% in the human population, otherwise the variant is described as a single nucleotide variant. Sequencing efforts have at present estimated that the human genome contains at least 10-11 million SNPs, encompassing ~0.3 % of the genome (4). It is assumed that the majority of SNPs do not have any phenotypic consequences. However, the impact of SNPs on the phenotype is dependent on the location of the polymorphism in the DNA sequence; whether it falls within the coding or non-coding regions, and numerous common known trait-associated SNPs have been identified (4-7). For example, a SNP variant in the *SLC24A4* (*solute carrier family 24, member 4* on 14q32.12) gene is associated with eye and hair color (8) and recent studies have identified SNPs in *FGFR2* (*fibroblast growth factor receptor 2* on 10q26), *TNRC9* (*trinucleotide repeat containing 9* on 16q12.1) and *MAP3K1* (*mitogen-activated protein kinase kinase kinase 1* on 5q11.2) genes associated with increased breast cancer risk in the general population (9).

Variants that distinguish two genomes in one species and include more than one single nucleotide are broadly defined as structural variations. They include insertions, deletions, translocations, duplications and inversions. Several studies suggest that these structural variants account for at least 70% of all variant bases in the human genome, and for any given individual, structural variations constitute between ~0.5 to 1% of the genome. All these variations very likely contribute to both human diversity and disease susceptibility due to altered gene dosage levels or by disruption of proximal or distant regulatory regions (3, 4, 7, 10, 11).

Copy number variations (CNVs) are usually defined as being larger than 1 kb in size and present in variable copy number on comparison with a reference genome. They include deletions, duplications and insertions. CNV can be inherited through the germline or occur sporadically as de novo aberrations (2, 3, 11-13). Improvement of microarray technologies as well as development of massively parallel DNA sequencing over the past five years has allowed the analysis of the human genome with unprecedented resolution, which has improved our knowledge about DNA variation and its connection to disease. In the 1980s, the frequency of CNVs was assumed to be low and directly related to specific genomic disorders (5, 14, 15). Today, it has been shown that CNVs encompass higher number of nucleotides and occur more frequently than SNPs (13). It is estimated that up to 12-13% of the human genome is subject to CNVs (2, 13). Larger copy number aberrations affecting numerous genes have been associated with many phenotypic traits and disease susceptibility, such as deletion of 22q11.2 which is associated with Di George syndrome and Prader-Willi syndrome caused by a 15q11-q13 deletion, to name but a few examples (2, 7, 15). There are also numerous examples of copy number aberrations encompassing only single genes, with a clear link to specific phenotypes or diseases. Deletion of the *NF1* (*neurofibromin 1* on 17q11.2) gene causes neurofibromatosis type 1 (16, 17), deletion of the *IRGM* (*immunity-related GTPase family, M* on 5q33.1) gene is associated with Crohn´s disease (18) and copy number differences of the *CCL3L1* (*chemokine (C-C motif) ligand 3-like 1* on 17q11.2) gene are associated with markedly enhanced HIV/acquired immunodeficiency syndrome (AIDS) susceptibility (19).

The mechanisms of CNV formation are not completely understood. It has been observed that CNVs often occur in regions carrying, or flanked by, large segmental duplications (1, 5, 11, 20). Segmental duplications (also called low-copy repeats) are blocks of repeated genomic DNA, often with more than 95% identity, that occur twice or more times in the haploid genome. They typically range in size between 1-400 kb and constitute ~5% of the sequence of the human genome (3, 21, 22). Several studies have noted that the presence of these segmental duplications predisposes these regions

to chromosomal instability and rearrangements (22-25). CNVs associated with segmental duplications are likely to arise by the mechanism of non-allelic homologous recombination (NAHR) (11, 13, 22, 26-28). In NAHR the recombination takes place between lengths of homologies at different genomic positions on a pair of chromosomes. NAHRs between homologies in direct orientation located on the same chromosome (intrachromosomal-NAHR) result in duplications or deletions, whereas NAHRs between inverted oriented homologies on the same chromosome lead to inversions. Recombination between homologies located on different chromosomes (interchromosomal-NAHR) result in reciprocal translocations (11, 29). The majority of rearrangements resulting from NAHR are recurrent and they can occur during meiosis as well as mitosis (11, 13, 22, 26-28).

CNVs can also be formed by non-homologous end joining (NHEJ), a mechanism also normally involved in the repair of double-strand breaks in DNA. In NHEJ, double strand brakes are bridged, modified and ligated. The product of repair often contains additional nucleotides at the DNA end junctions. In contrast to NAHR, this process is not dependent on segmental duplications to mediate the recombination. Instead, NHEJ has been observed in the vicinity of for instance identical *Alu* repeats, which are a family of short interspersed sequences common in the human genome, and between LINE elements (long interspersed nuclear element), which are a class of moderately repetitive transposable sequences lacking long terminal repeats. Non-recurrent rearrangements are thought to arise by NHEJ (5, 11, 13, 29, 30).

In addition to NAHR and NHEJ, a replication-error mechanism has recently been implicated in the formation of complex and non-recurrent chromosomal rearrangements. This mechanism is called fork stalling and template switching (FoSTeS). Rearrangements are formed when the DNA replication fork stalls - which leads to lagging strand disengagement from the original template and annealing, by micro homology at the 3´end, of another adjacent replication fork, and re-initiation of DNA synthesis. FoSTeS can result in deletions and/or duplications interrupted by either normal copy number or triplicates. The mechanism can create not only large genomic duplications of several Mb but also minor duplications/triplications and even rearrangements within single exons. FoSTeS has been implicated in gene duplication and exon shuffling which drives gene and genome evolution (13, 31-33). Characteristic features for each of the three rearrangement mechanisms are shown in Table 1.

Table 1. Comparison of the three major mechanisms that underlie human genetic rearrangements and CNV formation: non-allelic homologous recombination (NAHR), non-homologous end joining (NHEJ) and fork stalling and template switching (FoSTeS).

	NAHR	NHEJ	FoSTeS
Structural variation type	Duplications, Deletions, Inversions	Duplications, Deletions	Duplications, Deletions, Inversions, Complex
Homology flanking breakpoint (before rearrangement)?	Segmental duplications	*Alu* repeats, LINE elements	No
Breakpoint	Inside homology	Addition or deletion of basepairs, or microhomology	Microhomology

All of these mechanisms occur both in germ cells, where the rearrangements can be associated with genomic disorders, and in somatic cells, where the rearrangements can cause somatic diseases like cancer. Future development of assays with increased sensitivity will most certainly provide a more extensive overview of the structural variants in the human genome and increase our knowledge about their role in phenotypic variation and disease predisposition.

Cancer

Cancer is a large and heterogeneous group of genetic disorders that affect somatic cells. As a disease, it is responsible for one in eight deaths worldwide; it can develop at all ages; however, the risk increases with age. The definition cancer in fact includes over 100 distinct diseases with diverse risk and epidemiology and different tumor subtypes can even be found within specific tissues and organs. A tumor can be either benign, defined as a localized lesion that does not grow into adjacent tissues, or malignant, that invades nearby tissues and spreads colonies throughout the body, known as metastasis (34-36).

Cancer cells display an uncontrolled proliferation due to defects in the regulatory circuits that govern normal cell growth and differentiation. Cancer can originate from most of the cell types and organs of the body and thus the heterogeneity of cancer is at the level of cellularity, genomic complexity and diverse clinical behavior (34, 37). Ten years ago, Hanahan and Weinberg proposed six hallmarks of cancer (Figure 1, top). These capabilities are *i)* self-sufficiency in growth signals, *ii)* insensitivity to growth-inhibitory signals, *iii)* evasion of programmed cell death, *iv)* limitless replicative potential,

v) sustained angiogenesis and *vi)* tissue invasion and metastasis. These hallmarks were described as being abilities shared in common by most cancers; however the order in which these abilities are acquired by the cancer cells seems to differ between various tumors and their subtypes (34). Very recently, additional hallmarks have been proposed (38). These include *vii)* evasion of immune surveillance, *viii)* metabolic stress, *ix)* proteotoxic stress, *x)* mitotic stress, *xi)* oxidative stress and*, xii)* DNA damage stress (Figure 1, bottom). However, these hallmarks differ from the original hallmarks proposed by Hanahan and Weinberg in that they do not describe functional capabilities, but rather the state of cancer cells. In addition Negrini *et al.* recommended oxidative stress and proteotoxic stress to be secondary hallmarks of cancer, meaning that the proteotoxic stress may be secondary to aneuploidy which is manifested by genomic instability, and oxidative stress may be secondary to oncogenic signaling and metabolic stress (39).

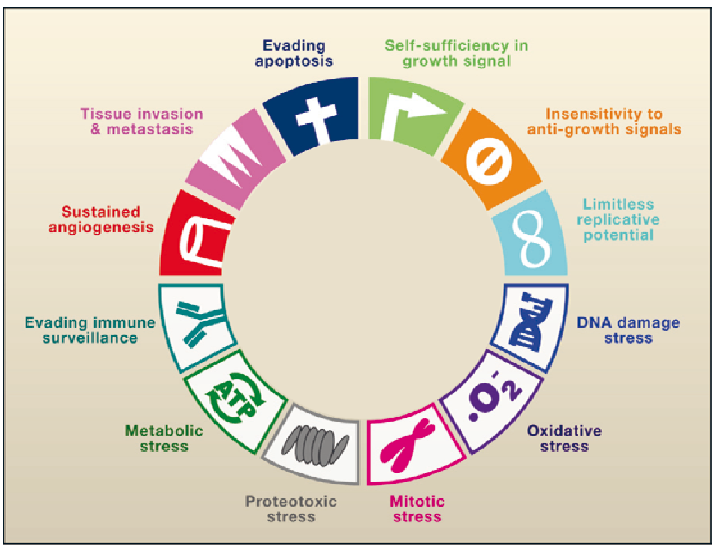

Figure 1. Capabilities and states acquired by cancer cells. In addition to the six hallmarks originally proposed by Hanahan and Weinberg (top half, (34)), the genomes of cancer cells obtain a set of additional hallmarks (lower half, (38)). Reprinted from Cell, *136*, Luo, J. *et al.*, Principles of Cancer Therapy: Oncogene and Non-oncogene Addiction, 823-837, 2009, with permission from Elsevier.

These capabilities are acquired by changes occurring in the DNA sequence of cancer cells. Theodor Boveri suggested more than 100 years ago that abnormalities of the chromosomes have a central role in tumor formation (40, 41). The cancer genome is often said to be aneuploid, which means that it contains an abnormal number of chromosomes (41-43). Other alterations that can be found in cancer are: point mutations, inter- or intrachromosomal rearrangements, copy number changes, including gains and deletions but

also so-called epigenetic alterations (40, 44). Epigenetics is the study of heritable changes in gene expression that are not due to any alteration in the DNA sequence (45). DNA methylation, when a methyl group is added to a cytosine that precedes a guanine, constitutes one of the best known epigenetic modifications in the human genome (46). Hypermethylation of CpG islands in the promoters of genes is associated with transcriptional silencing of tumor suppressor genes in many cancers. Loss of methylation (hypomethylation), in contrast, can lead to gene activation and chromosomal instability (47). Other epigenetic changes that can contribute to carcinogenesis include the modifications of histones, involving the addition of different chemical groups e.g. methyl, acetyl and phosphate. These modifications can have a variety of effects on the cell, such as alteration of gene transcription and DNA repair (46). In addition to these alterations, the cancer cell can acquire completely new DNA sequences from exogenous sources, notably from viruses. Viruses can modify the gene expression in the host cell and thereby promote cell proliferation and contribute to several types of cancer (40, 48).

Today it is believed that cancer is caused by an accumulation of somatic mutations in a series of genes over time rather than a single mutation in one gene (40, 49). The exact number of such critical mutations is not known and is likely cancer type-specific. Some of these mutations may be acquired by the ancestor of the cancer cell yet remain biologically normal, and lacking any specific phenotype. Subsequent mutations may be induced by mutagens of both internal and external origin. These mutations are usually repaired by different DNA repair processes, namely the earlier described process of NHEJ. However sometimes these processes can fail and mutations go unrepaired in the DNA. Some examples of external mutagenic agents are; tobacco smoke carcinogens which contribute to both lung and urinary bladder cancer, naturally occurring compounds, such as aflatoxins produced by fungi, which are associated with liver cancer and finally, ultraviolet light which is strongly associated with skin cancer (40, 50, 51).

Nevertheless, cancers do not always arise sporadically; cancer predisposition can also be inherited through the germline, by inheritance of a mutated cancer associated gene from either parent. It is estimated that 5-10% of all cases of cancer are hereditary (36). The most common alterations occurring in germline are point mutations or small deletions or insertions. These mutations do not cause cancer *per se*, but they confer an increased relative risk of cancer among carriers. Individuals with germline mutations often develop multiple tumors early in life compared to persons whose mutations have occurred somatically (43, 52). One example of cancer with hereditary linkage is breast cancer, in which a defective copy of *BRCA1* (*breast cancer 1* on 17q21.31) or *BRCA2* (*breast cancer 2* on 13q13.1) genes confer a >90% risk of breast cancer development among carriers (53). Other examples in-

clude syndromes like Li-Fraumeni and Gorlin, linked to germline mutations in the *TP53* tumor suppressor gene and *PTCH* (the human homolog of the Drosophila "patched" gene PTC), respectively, where affected individuals have a greatly increased susceptibility to several kinds of cancer (54, 55).

Mutations in cancer may be broadly classified into two categories; driver or passenger mutations. Driver mutations are said to confer selective growth advantage for the cancer cell. These mutations are by definition found in cancer genes of which approximately 400 have been identified to date (http://www.sanger.ac.uk/genetics/CGP/Census/). Driver mutations are those mutations which are positively selected during the evolution of the cancer. Passenger mutations on the other hand, do not really contribute to cancer development because they have not been subject of selection. Neither do they confer a growth advantage. Nonetheless, passenger mutations are frequently found within cancer genomes and are 'carried along' in the clonal expansion process (40, 44, 50). The genes frequently found to be involved in cancer are generally divided into two major categories; oncogenes and tumor suppressor genes. Oncogenes and tumor suppressor genes operate in the same physiologic manner: they drive tumorigenesis by increasing the tumor cell number through stimulation or inactivation of different pathways regulating cell division, cell death or cell cycle arrest (34, 40, 43, 49, 52, 56).

Oncogenes

Oncogenes are forms of proto-oncogenes whose normal activity control growth signaling and anti-apoptotic pathways, yet due to mutation are no longer capable of responding to normal regulatory signals. They form a very heterogeneous group of genes that can be divided into five classes: secreted growth factors (e.g. *PDGF*), cell surface receptors (e.g. *EGFR*), tyrosine kinases (e.g. *SRC*), membrane associated G-proteins (e.g. *RAS* genes) and nuclear transcription factors (e.g. *MYC* genes). Mutations leading to oncogene activation are dominant, meaning that one single hit in one allele is generally enough to confer a selective growth advantage to the cell. Oncogene activation can result from chromosomal translocations, gene amplifications, mutations or by hypomethylation of CpG islands in the promoter region (43, 57). Either of these mechanisms can result in an alteration of proto-oncogene structure or an increase of proto-oncogene expression, or alternatively, both mechanisms might collaborate to create an oncogene (43).

Chromosomal translocations can lead to transcriptional activation of proto-oncogenes or to the creation of aberrant fusion proteins. The fusion of *abl* (*Abelson murine leukemia viral oncogene homolog 1, v-abl* on 9q34.12) and *bcr* (*breakpoint cluster region* on 22q11.23) proto-oncogenes creates the Philadelphia chromosome, which is implicated in chronic myeloid leukemia

(58). Gene amplification is defined as the increase in copy number of a gene within the genome of a cell and leads to increased gene expression. Gene amplification produces karyotypic abnormalities such as double-minute chromosomes and homogeneous staining regions. Double-minute chromosomes are characteristic minichromosome structures without centromeres. Homogeneous staining regions are segments of chromosomes uniformly stained after G banding. Both double-minute chromosomes and homogeneous staining regions contain up to several hundred copies of a gene. An example of oncogene activation by amplification is the case of *EGFR* (*epidermal growth factor receptor* on 7p11.2), an aberration that is frequently observed in glioblastoma (59).

Mutations may activate proto-oncogenes through structural alterations in their encoded proteins. These modifications usually involve important protein regulatory regions and often lead to the uncontrolled, constitutive activity of the defective protein. Several types of mutations, such as point mutations (changing a single amino acid within the protein), deletions, and insertions, are capable of activating proto-oncogenes. Point mutations are for example frequently detected in the ras family of proto-oncogenes (K-ras, H-ras, and N-ras) (60). It has been estimated that as many as 15% to 20 % of human tumors may contain a ras mutation. Ras mutations have been linked to carcinogen exposure, and point mutations within the 12th codon of H-*Ras* (*Harvey rat sarcoma viral oncogene homolog, v-Ha-ras* on 11p15.5), are a common event in urinary bladder carcinoma (61). In contrast, oncogene activation by DNA hypomethylation has been identified in gastric cancer, where the R-*Ras* (*related RAS vira,l r-ras, oncogene homolog* on 19q13.3-q13.4) gene is found to be hypomethylated (62).

Tumor suppressor genes (TSGs)

The TSGs also control cell-growth and proliferation but they have the opposite role of oncogenes in that they inhibit or tightly control cell growth in their normal state. Mutations in TSGs reduce the activity of the gene product or result in a complete loss of the protein. Inactivation can occur by point mutations striking many sites of the coding sequence of the gene, by random deletions, insertions or by epigenetic changes such as hypermethylation of CpG islands in the gene promoters that silence transcription of the gene (43, 46, 49). Once these growth-suppressing genes are inactivated, the proliferation of the cancer cells may be accelerated, no longer being held back by the actions of their transcripts (43). TSGs generally follow the Knudson two-hit hypothesis that requires mutations on each of the two alleles of the gene to promote tumor development. For individuals with a germline mutation on one of the alleles, often only one hit on the other allele is sufficient for tumorigenesis (Figure 2) (63). However haploinsufficiency, when the total level

of a gene product produced by the cell is about half of the normal level (due to one copy being inactivated by mutation), may in some cases not be sufficient to permit the cell to function normally and thereby contribute to tumor development (43, 64). Haploinsufficiency *PTEN* (*phosphatase and tensin homolog* on 10q23.31) is known to accelerate the transformation of astrocytomas from low-grade to high grade (65).

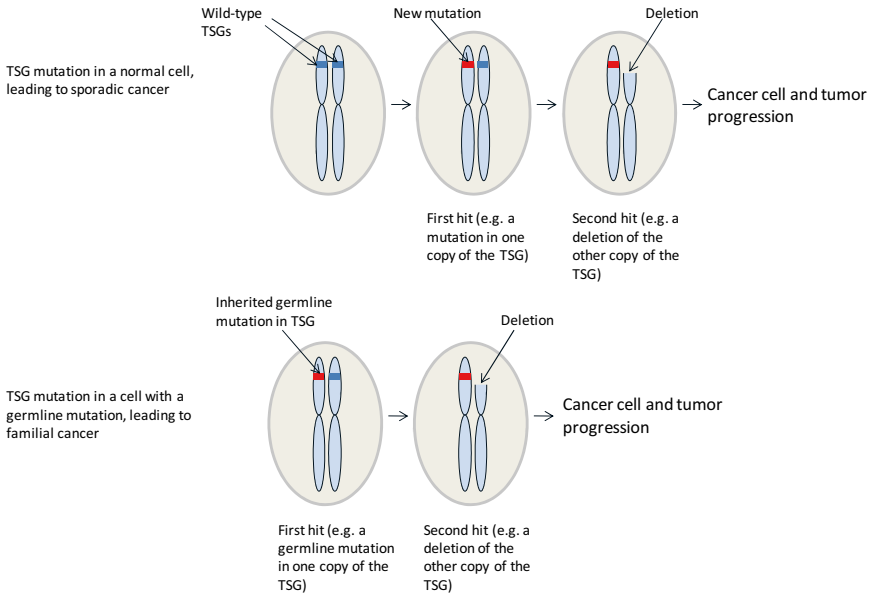

Figure 2. Knudsons two-hit hypothesis for tumorigenesis involving a tumor suppressor gene.

TSGs are broadly divided into two types: gatekeeper and caretaker genes. The genes acting as gatekeepers directly control cellular proliferation by inducing cell death or cell cycle arrest. The *RB1* (*retinoblastoma 1* on 13q14.2), *CDKN2A* (*cyclin-dependent kinase inhibitor 2A* on 9p21.3) and *TP53* (*tumor protein p53* on 17p13.1) are three examples of gatekeeper genes, whose malfunction is frequently involved in many types of cancer. Caretaker genes in contrast, do not directly regulate proliferation; they act to maintain the integrity of the genome and thereby prevent tumor development. Owing to the intrinsic chemical instability of DNA, all cells are under constant threat of acquiring mutations caused by factors such as exposure to genotoxic metabolic or environmental agents. Caretaker genes keep genetic alterations to a minimum, but when present in a defective form the mutation rate is increased, affecting all genes including gatekeeper genes, which can directly regulate tumor growth. Caretaker genes include mismatch repair (MMR), nucleotide-excision repair (NER) and base-excision (BER) genes

which are responsible for the repair of DNA replication errors and the genes controlling processes such as mitotic recombination and chromosomal segregation. *BRCA1* and *BRCA2*, which are strongly associated with breast cancer, and *MLH1* (*mutL homolog 1, colon cancer, nonpolyposis type 2* on 3p22.2) involved in colon and uterus cancer are both examples of deregulated caretaker genes playing a significant role in tumor development (39, 43, 66, 67).

CNS tumors

Tumors of the central nervous system (CNS) are a heterogeneous group of both benign and malignant neoplasms. Brain tumors occur in children as well as adults (68, 69). Their classification, according to the World Health Organization (WHO 2007) system, is based on cell morphology and the primary site of the affected organ. CNS tumors are named after the cell type that tumor cells resemble. Based on this, the tumors can be divided in three distinct groups: ***a)*** gliomas (astrocytoma, oligodendroglial tumors and ependymoma); ***b)*** meningiomas; and ***c)*** embryonal tumors (medulloblastoma, primitive neuro-ectodermal tumors (PNET) and Atypical Teratoid/Rhabdoid Tumor (AT/RT)). Furthermore, brain tumors are also graded according to malignancy. Grade I tumors are applied to neoplasms with low proliferation potential and are often correlated with good prognosis whereas grade IV tumors are highly-proliferative and mitotically active and these tumors are often fatal (55).

a) Gliomas present with a morphology and gene-expression characteristics similar to glia, astrocyte and oligodendrocyte cells, which together constitute the supporting tissue in the brain. This large group of brain tumors is further divided into three groups: astrocytomas, oligodendroglial tumors and ependymomas. Patients with gliomas often develop symptoms that include headaches and seizures. Speech and visual loss can also occur. The majority of gliomas arise sporadically and a predisposition does not usually run within families, however, there are examples of families with more than one affected member (55, 70).

Astrocytomas are the most common type of all gliomas and account for approximately 60% of all primary tumors of the CNS. They are comprised of cells which resemble the supporting astrocyte cells of the brain. The cellular origin of astrocytomas is still unknown, although it is believed that they arise from astrocyte precursors or stem cells (70, 71). Astrocytomas show different degrees of malignancy and can be classified as; pilocytic astrocytoma (grade I), mainly occurring in children and young adults and low grade diffuse astrocytoma (grade II), typically affecting young adults. In contrast,

anaplastic astrocytoma (grade III) and glioblastoma (grade IV) primarily occur in adults (55). Grade I tumors are almost always curable by surgery. Grade II and III patients have a mean survival of 10-15 and 2-3 years respectively while glioblastomas, being the most malignant grade of tumour, grow rapidly and patients generally have a mean survival of about 1 year (55, 70). Glioblastomas can be divided into two groups: primary and secondary glioblastoma, depending on the clinical history of the patient (72). Glioblastoma will be discussed in greater detail in a subsequent section, since it is the focus tumor type studied in paper II. Several different genetic aberrations are frequently observed in astrocytomas. Gains of chromosome 5 and 7 are often found in pilocytic astrocytoma, *TP53* mutations, overexpression of *PDGFRA* (*platelet-derived growth factor receptor alpha* on 4q12), gain of 7q and 8q, loss of heterozygosity (LOH) on 22q and deletion of chromosome 6 are often seen in low grade diffuse astrocytoma. *TP53* and *PTEN* mutations, LOH of 6q, 10q, 17p, 19q and 22q are detected in anaplastic astrocytoma while gain of chromosome 7, loss of chromosome 10 and deletion of 9p are frequent events in glioblastoma.

Oligodendroglial tumors include oligodendrogliomas and oligoastrocytomas. Oligodendrogliomas are divided into two categories; grade II and anaplastic grade III (55). Oligodendrogliomas are more common in adults and account for only 2% of all brain tumors in children. The tumors are composed of cells that resemble oligodendroglial cells. These tumors are frequently observed with combined loss of 1p and 19q, which is strongly associated with favorable outcome for the patient (73). Up to 90% of all oligodendrogliomas carry this alteration (55, 74). Oligoastrocytomas are similarly divided into the two categories of grade II and anaplastic grade III. They are composed of a mixture of oligodendroglioma and diffuse astrocytoma cells and usually develop in middle-aged individuals (55). The genetic picture of these tumors is similar to that of oligodendrogliomas, the combination of 1p and 19q loss being found in approximately 50% of tumors (55, 74).

Ependymal tumors, as their name suggests, resemble the ependymal cells of the brain and can occur both in the spinal canal and in the ventricular system. While such tumors may develop at all ages, tumors located in the spinal canal develop more frequently in adults, whereas intracranial tumors predominate in children. Ependymal tumors are further subdivided into subependymoma (grade I), myxopapillary ependymoma (grade I), ependymoma (grade II) and anaplastic ependymoma (grade III) (55). The most frequent genetic alteration in sporadic ependymoma is monosomy 22 and the *NF2* (neurofibromin 2) gene on chromosome 22 is clearly involved in ependymoma tumorgenesis (55, 75-77). Gain of chromosomes 1, 7 and 9p as well as losses of chromosomes 17, 6q and 9q are also common (78, 79).

b) Meningiomas develop from the meninges which is the membrane that surrounds the brain and spinal cord. This tumor is considered to be derived from neoplastic meningothelial or arachnoidal cells. The tumors are usually benign (grade I) and generally slowly growing. However, meningiomas with a greater likelihood of recurrence and/or aggressive behavior are graded as grade II or III. Meningiomas typically occur sporadically in adults, although they can also develop in association with the Neurofribromatosis type 2 (NF2) syndrome. The most common genetic aberration associated with both sporadic and NF2-associated meningiomas is deletion of chromosome 22. Biallelic inactivation of the *NF2* gene is found in meningiomas of all grades and is thought to be an early event in tumorgenesis. Other genetic aberrations associated with meningioma are deletion of 1p and loss of chromosomes 6, 10, 14, 18 and 19 (55).

c) Embryonal tumors of the CNS are the most common malignant brain tumors affecting children. Embryonal tumors derive from the embryonic (fetal) tissue and are most common in children or young adults. This group of brain tumors includes medulloblastoma (grade IV), primitive neuroectodermal tumors (PNET) (grade IV) and atypical teratoid/rhabdoid tumor (AT/RT) (grade IV) (55).

Medulloblastoma is a highly malignant, invasive tumor of the cerebellum affecting mainly children and adolescents. This is the most common malignant brain tumor in children (80, 81). The most frequent cytogenetic abnormality in medulloblastoma is isochromosome 17q (55). Medulloblastoma will also be discussed more detail further below as it is the tumor type studied in paper III of this thesis.

PNETs are located in the cerebrum, but they can also be encountered in the spinal cord or suprasellar region. These tumors also occur predominantly in children or adolescents and they are composed of undifferentiated or poorly differentiated neuroepithelial cells. Loss of 4q, 9p, 14q and 19q has been identified in PNETs (55, 82).

AT/RT tumors are highly malignant and predominantly (50%) located in the posterior fossa. The remaining AT/RT tumors can arise in the supratentoral, pineal, multifocal or spinal locations. This tumor is most common in younger children (<3 years) and composed of neoplastic rhabdoid cells. Mutation or loss of *INI1* (*integrase interactor 1* on 22q11.23) is the genetic hallmark of AT/RT tumors (55).

Glioblastoma

Glioblastoma is the most frequent primary brain tumor affecting adults and the disease is invariably fatal (55, 70). The peak incidence occurs in the age range of 45-75 years and the incidence in Europe and North America is 3-4 new cases per 100 000 population per year. The tumor is most often located in the subcortical white matter of the cerebral hemispheres and rarely metastasizes via the cerebrospinal fluid. However, the tumor has a very high proliferative rate and widespread microvascular proliferation. Necrotic areas are also often observed. The clinical symptoms for glioblastoma include often headache, nausea/vomiting due to raised intracranial pressure and one third of the patients experience epileptic seizures (55, 83). Besides surgical resection, the current standard care for glioblastoma patients is treatment by adjuvant radiation and chemotherapy with the use of alkylating agents such as Temozolomide. However, the prognosis for glioblastoma is very poor and little improvement in the median survival has been observed over the past 25 years (84, 85). Patients without surgical resection have a median survival of only 2.5 months from the time of diagnosis. Those who undergo surgical resection demonstrate a median survival of 7.9 months and the addition of Temozolomide to adjuvant radiotherapy after surgery extends median survival to 14.6 months (72). Thus, the inclusion of Temozolomide in treatment regimes was shown to improve the two year survival rate from 10.4 to 26.5%, representing the most important therapy advancement to date (86). The poor outcome observed in glioblastoma patients is largely due to tumor recurrence. Two alternative models (accounting for all cancers) have been proposed to explain the chemo- and radiotherapy resistance developed by these tumors (87): *i)* the clonal evolution hypothesis (88) and *ii)* the cancer stem cell hypothesis (89-91). The clonal evolution hypothesis suggests that subpopulations of tumor cells with different mutations continuously arise during tumor evolution. These subpopulations are selected by the tumor microenvironment during the progression and treatment of the tumor. Each subpopulation can therefore become predominant after radiotherapy and/or chemotherapy due to the selection pressure. The subpopulations which are most resistant to treatment will be selected for and can potentially re-initiate the tumor. Alternatively, the cancer stem cell (CSC) hypothesis suggests that the CSCs which exhibit stem cell-like characteristics such as self-renewal and multipotency, can propagate and or re-initiate the tumor. Several properties allow CSCs to survive conventional chemo- and radiotherapies and re-establish tumor growth after treatment. However, the CSC hypothesis remains controversial (87).

The majority of glioblastomas are sporadic although there are cases of occurrence in more than one family member. This is most often seen within inherited tumor syndromes including Turcot and Li-Fraumeni syndromes, neuro-

fibromatosis type 1 and multiple enchondromatosis (55, 92). Most glioblastomas (>90%) develop *de novo* without evidence of pre-malignant precursor lesions and tend to have a short clinical history. These tumors are classified as primary glioblastoma and they are typically observed in older patients, with a mean age of diagnosis of 62 years. Secondary glioblastomas develop more slowly, often by progression from diffuse astrocytoma grade II or anaplastic astrocytoma grade III and are typically identified in younger patients, mean age of diagnosis, 45 years (72, 93, 94). The genetic pathways leading to glioblastoma differ between these two groups (Figure 3).

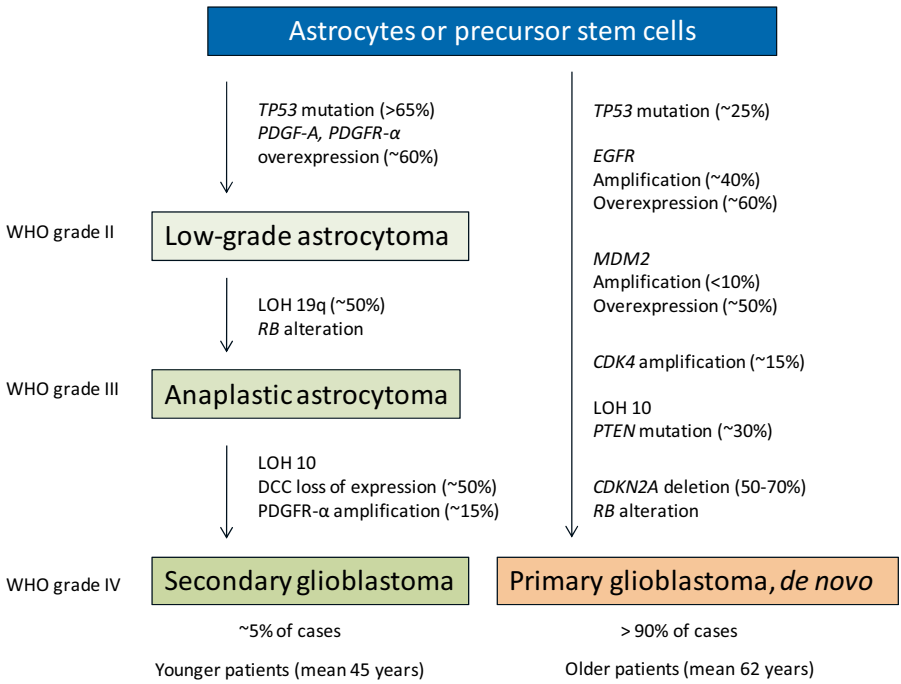

Figure 3. Genetic pathways operative in the evolution of secondary and primary (*de novo*) glioblastomas.

Primary glioblastoma is characterized by amplification and over-expression of the *EGFR* and *MDM2* (*mouse double minute 2* on 12q15) genes, deletion of the *CDKN2A* gene, LOH on 10p and 10q and *PTEN* mutation. In secondary glioblastoma, the major alterations observed are over-expression of the *PDGFRA* gene, *TP53* and *PTEN* mutations, *DCC* (*deleted in colorectal carcinoma* on 18q21.1), promoter methylation of the *RB1* gene and finally loss of expression and/or LOH of 19q and 10q (55, 68, 72, 95, 96). Recently, mutations in the *IDH1* (*isocitrate dehydrogenase 1* on 2q33.3) gene have also been shown to be implicated in the development of secondary glioblas-

toma (97). Almost all glioblastomas that develop in children are primary glioblastoma. Pediatric glioblastoma however, presents with different genetic profiles than that observed in adults; they often display a high frequency of *TP53* mutations (40%), low frequency of *EGFR* amplifications (6%) or *CDKN2A* deletions (19%) and absence of *MDM2* amplifications (55).

Genetic pathways in glioblastoma

Different pathways, such as the RAS, TP53, PI3K (*phosphoinositide-3-kinase*) and the cell cycle control pathways are known to be disrupted in this tumor type (98, 99). The RAS pathway is frequently activated by amplification of growth factor receptor genes such as *EGFR* (7p11) or *PDGFRA* (4q12). Between 30-50% of glioblastoma tumors show *EGFR* amplification (59, 99). *PDGFRA* amplification is less common and is identified in approximately 15 % of glioblastomas. The TP53 pathway is often disrupted by *TP53* (17p13.1) mutations which is the main genetic hallmark of secondary glioblastoma (identified in >65% of cases). *TP53* mutations are also observed, although less frequently (25% of cases), in primary glioblastoma. The TP53 pathway can also be disrupted by *MDM2* (12q15) amplification which is observed in approximately 10% of glioblastomas, or by loss of *CDKN2A* (9p21) which is frequently (76%) observed in this tumor. Both *MDM2* amplification and *CDKN2A* loss are observed in primary as well as secondary glioblastomas. The PI3K/Akt pathway is also involved in glioblastoma development through aberrant growth factor signaling and loss of *PTEN* (10q23) which is mutated in 15-40% of primary glioblastoma. The cell cycle control pathways can also be disrupted in glioblastoma by loss of *CDKN2A* (9p21), amplification of *CDK4* (*cyclin-dependent kinase 4* on12q14) or *RB1* alterations (13q14). Amplification of *CDK4* is observed in 15% of tumors and LOH at 13q including the *RB1* locus is detected in 12% of primary glioblastomas and 38% of secondary glioblastomas (55, 70, 97, 99). Among the many aberrations found in glioblastoma, hypermethylation of the *MGMT* (*O-6-methylguanine-DNA methyltransferase* on 10q26.3) promoter is also a common alteration in both primary and secondary glioblastomas. The *MGMT* gene codes for a DNA repair protein involved in cellular defense against mutagens (100). Promoter methylation of *MGMT* is then associated with longer survival of glioblastoma patients since the tumor cells have a suppressed ability to protect themselves against alkylating agents such as Temozolomide, frequently used in glioblastoma treatments (95, 101, 102).

Medulloblastoma

Medulloblastoma is an invasive embryonal tumor that predominantly manifests in children and adolescents. Seventy percent of medulloblastomas occur in children younger than 16 years with a peak incidence at the age of 7. The incidence has been estimated to be about ~1 new case per 100 000 population per year. Approximately 65% of patients affected by medulloblastoma are males. The tumor arises in the cerebellum from cerebellar stem and precursor cells, including immature granule cells, and has a tendency to metastasize via cerebrospinal fluid pathways (55, 103-105). The majority of medulloblastomas arise in the cerebellar vermis, but a few are located in the cerebellar hemispheres. The clinical symptoms of medulloblastoma include truncal ataxia, disturbed gait, intracranial hypertension, headache and vomiting (55). Almost all medulloblastomas are sporadic, although a few cases of familiar medulloblastomas have been reported, again associated within familial cancer predisposition syndromes like Li-Fraumeni, Gorlin, Turcot and Rubinstein-Taybi syndromes (54, 55). Medulloblastoma are divided into three histological subtypes; classic-, desmoplastic/nodular- and anaplastic/large cell medulloblastoma(55, 106, 107). The classic medulloblastomas are more frequent than the other subtypes and account for approximately 65% of all cases. This subtype consists of sheets of small round cells and displays neuronal differentiation (55, 108). The desmoplastic/nodular suptype is characterized by nodular, reticulin-free zones surrounded by densely packed highly proliferative cells. This subtype account for 25% of all medulloblastoma tumors and is more commonly seen in patients with Gorlin syndrome. Patients with desmoplastic/nodular medulloblastoma display better prognosis than classical and anaplastic/large cell subtypes, and in particular, younger children with this subtype have a better outcome (106, 109, 110). Anaplastic/large cell medulloblastoma is the most undifferentiated subtype and the cells display significant nuclear pleomorphism, prominent nucleoli and abundant mitoses. This subtype has been reported to occur in 5-10% of tumors. Individuals with the anaplastic subtype usually have a worse outcome than the other two subtypes (55, 108).

A 5-year survival rate of 60 to 80% can be achieved with current management strategies, which involve maximal surgical tumor resection, adjuvant chemotherapy and craniospinal irradiation (111-113). Unfortunately, a significant proportion of survivors suffer severe long-term neurocognitive sequelae due to the intensive chemotherapeutic and radiotherapy regimes employed in treatment (114). Patients are today classified in standard- or high-risk groups, with the high-risk including patients younger than 3 years with incomplete tumor resection and/or with evidence of metastasis (55). However, different reports indicate that these clinical variables are insufficient in

defining disease risk and additional biomarkers for improvement of outcome prediction are desirable (112, 115).

Genetic pathways in medulloblastoma

Several cellular signaling pathways are reported to be frequently disrupted in medulloblastoma (Table 2). The Hedgehog signaling pathway, which plays an important role in cerebellar development, is often involved in medulloblastoma tumorigenesis. The most common alterations hitting this pathway are mutations of *Ptch1* (*patched homolog 1* on 9q22.1-q31) gene, which have been described in up to 20% of medulloblastomas. Other genetic alterations affecting the Hedgehog pathway include activating mutations of the *Smoh* (*smoothened homolog* on 7q31) and *Sufu* (*suppressor of fused homolog* on 10q24.32) genes and deletion of the tumor suppressor gene *REN* (*renin* on 17p13.2), which maps to a region commonly affected by deletions in medulloblastoma. Activation of Wnt and APC signaling pathways have also been observed, mainly due to mutations of the *β-catenin* (on 3p21), *APC* (*adenomatous polyposis coli* on 5q21-q22) and *AXIN2* (on 17q23-q24) genes, and also by deletions of the *AXIN1* (on 16p13.3) gene. The Notch pathway may also play a role in medulloblastoma development. Overexpression of Notch1 (on 9q34.3) and Notch2 (on 1p13-p11), which are receptors involved in normal development, cell-fate determination, proliferation and survival, have been observed. Several other signaling 'pathways' have also been reported to be involved in medulloblastoma development and progression: c-myc signaling and receptor tyrosine kinase signaling by ErbB, c-met, IGF-R and TrkC (55, 116-118).

The most frequent chromosomal aberration detected in medulloblastoma is isochromosome 17q, which is present in 30-50% of tumors (55, 119). Loss of 17p and gain of 17q occurring independently are also common as well as trisomy 7, gain of 1q and amplification of *MYCN* (*v-myc myelocytomatosis viral related oncogene* on 2p24.1), *ERBB2* (*v-erb-b2 erythroblastic leukemia viral oncogene homolog 2* on 17q11.2-q12) or *hTERT* (*telomerase reverse transcriptase* on 5p15.33) (68, 80, 116, 120). Additionally, loss of chromosomes 6, 7q, 8, 9q, 10q, 11 and 16q have also been reported in medulloblastoma (120-122).

Table 2. Summary of the important signaling pathways in medulloblastoma development.

Pathway	Deregulation	Association with medulloblastoma subtype	Functional effects
Hedgehog	Activating mutation of *Ptch1* (20%), *Sufu* (9%), *Smoh* (<5%), and deletion of *REN*	Desmoplastic/nodular	Cell cycle progression and induction of growth factors
Notch	Over-expression of Notch1 and Notch2	Not known	Progenitor cell survival
Wnt	Activating mutation of *β-catenin* (9%), *APC* (4%), *AXIN1* (12%)	Classic	Not known
c-met	Over-expression	Anaplastic/large cell	Tumor cell proliferation, cell cycle progression, cell survival, migration, invasion, cell size
erbB2	Over-expression	Not known	Cell migration, invasion and metastasis
TrkC	Over-expression	Not known	Apoptosis and maybe invasion
IGF-R	Over-expression and phosphorylation	Not known	Progenitor cell proliferation
c-Myc	Over-expression (64%) and amplification (15%)	Anaplastic/large cell	Cell proliferation, cell cycle progression and cell size regulation

Urinary bladder carcinoma

Bladder cancer is the fourth most common malignant disease in Europe, with 91 000 new cases diagnosed each year and causing approximately 37 000 deaths (123). The risk of developing bladder cancer increases with age, with a peak incidence between 50 and 70 years. The disease is also three times more common among men than women (124, 125). Occupational exposures, tobacco use and pharmaceutical drug use are all risk factors for developing cancer of the urinary bladder (126). Bladder cancer is a heterogeneous disease and is classified into several stages, Ta, T1-T4 and Carcinoma in situ (CIS), based upon the extent of invasion into the surrounding tissues (127). Approximately 70% of these patients present with papillary non-invasive, stage Ta tumors, with low risk for progression and death. However, as many as 50-70% of patients diagnosed with stage Ta will suffer from tumor recurrences after initial resection of the lesion. This high percentage of patients affected by tumor recurrences makes this tumor type one of the most prevalent neoplasms. The prevalence is estimated to be three to eight times higher than the incidence which poses this disease as a major burden on the health care system. More importantly, the high risk of recurrence also affects the

patient´s quality of life. Therefore, it is necessary to identify patients with a high likelihood of tumor recurrence and progression in order to promote more aggressive therapy regimes for these patients (124, 128). T1 tumors invade into the subepithelial connective tissue and represent approximately 20% of all newly diagnosed cases. CIS is a flat type of tumor, often found together with T1 tumors. This tumor spreads along the surface of the bladder and is associated with an increased risk for progression into invasive bladder cancer (127, 128). T2-T4 tumors are solid muscle infiltrating tumors and constitute a minority of cases. Patients with these tumors have a high risk for metastasis and death (127-129). It is difficult to predict at diagnosis which patients will progress from non-muscle-invasive tumors, primarily T1 tumors, to solid muscle infiltrating tumors (T2-T4). However, when comparing the progression-free survival of non-muscle-invasive tumors, it was observed that invasion of the subepithelial connective tissue is a prognostic indicator for risk of disease progression and reduced survival (124). Bladder cancer tumors are also divided into histological grades (low-grade and high-grade). Low-grade tumors are mostly common among Ta tumors, whereas the majority of T1-tumors are high-grade. Advanced tumors (T2-4) are only high-grade.

The treatment regimes with chemo- or immunotherapy depend on the tumor grade and recurrence status (124). Patients with superficial Ta and T1 tumors undergo initial transurethral resection in addition to treatment with chemo- or immunotherapy by intravesical instillations, in an attempt to prevent recurrences. Mitomycin-C (MMC), Bacillus Calmette Guerin (BCG) and interferon alfa are frequently used agents for this purpose (124, 128, 130). Radical cystectomy is the standard care for patients with muscle invasive bladder cancer (T2-T4), but the 5-year survival rate is only in between 40-60% and has not improved significantly over recent years (124, 128, 131).

Genetic pathways in urinary bladder carcinoma

The most common reported genetic alteration among the noninvasive bladder tumors are deletion of chromosome 9 and point mutations of *FGFR3* (*fibroblast growth factor receptor 3* on 4p16) (125, 131-134). Activating mutation of *FGFR3* leads to MAPK (mitogen activated protein kinase) pathway stimulation which regulates various cellular activities such as gene expression, mitosis, differentiation, cell survival and apoptosis (133, 134). *FGFR3* mutations are far more frequent in low-grade Ta tumors (up to 80%), with only 10-20% of muscle invasive carcinoma showing this mutation (133-136). In addition to these alterations, loss of chromosome Y is observed to be a frequent event in the noninvasive bladder tumors (137). Occasionally, small amplicons or homozygous deletions have also been identified (138).

Apart from these alterations, the low grade noninvasive bladder tumors display few molecular alterations (139). Whole chromosome 9 deletions are common in all grades and stages of bladder cancer with more than half of tumors presenting this aberration. However, several studies have shown that 9q loss is more common in low-grade and low-stage tumors compared to muscle invasive tumors, which in turn more frequently display 9p deletions (140-143). Four different tumor-suppressor regions have been mapped to chromosome 9. High-grade and high-stage tumors are often characterized by homozygous deletion of the tumor suppressor genes *CDKN2A* and *CDKN2B* located on 9p21 which is one of four different minimal tumors suppressor candidate regions (144). The other three regions are located on 9q22, 9q32-33 and 9q34 (125, 131, 145, 146). These candidate regions encompass the tumor suppressor genes *Ptch1* (147, 148), *DBC1* (*deleted in bladder cancer-1* on 9q32-q33) (149, 150) and *TSC1* (*tumor suppressor candidate-1* on 9q34.13) (151). CIS and muscle invasive tumors show frequent alterations of the *TP53* and *RB* genes and pathways and particularly muscle invasive tumors display a wide range of genomic alterations, including losses of 2q, 5q, 8p, 10q, 11q, 13q, 17p, 18q and gains of 1q, 5p, 6p, 8q, 10p, 17q and 20q (139, 152). A model for bladder cancer initiation and progression with genes frequently activated or inactivated are shown in Figure 4.

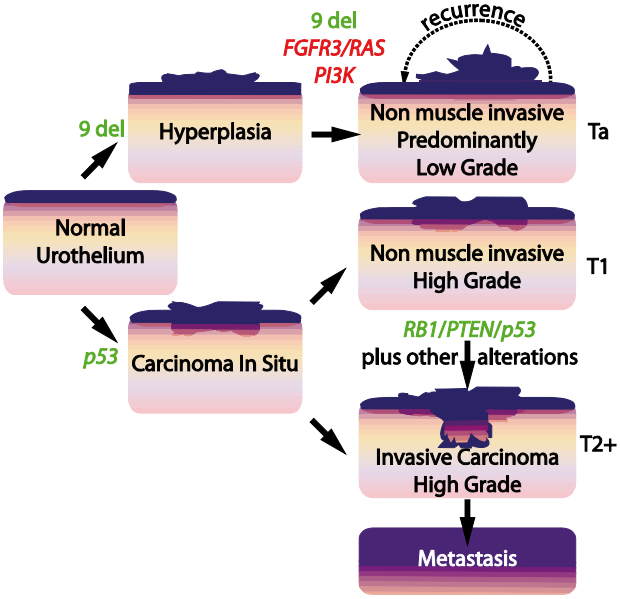

Figure 4. Schematic model for bladder cancer initiation and progression. Genes frequently altered by activating or inactivating mutations are shown in red and green respectively.

Present investigation

Aims

- To explore the extent of copy number variations in a large set of phenotypically normal individuals by applying a full-coverage 32K BAC-based array.

- To identify DNA copy number aberrations and comprehensively search for novel candidate gene loci involved in glioblastoma and medulloblastoma development, using a full-coverage 32K BAC-based array.

- To correlate DNA copy number and gene expression status with clinical behavior on a clinically well-characterized cohort of Ta-stage bladder carcinomas selected by the presence or absence of recurrences.

Methods

Microarray-based genomic hybridization (array-CGH)

This method is an advanced technique first described in 1997 by Solinas-Toldo and colleagues and it was first referred to as matrix-CGH (153) or array-CGH (154). This technique has emerged from conventional metaphase-CGH, first developed by Kallioniemi and colleagues in 1992 (155), in which test (tumor sample) and reference DNA (control sample) are differentially labeled with fluorochromes and competitively hybridized to normal metaphase chromosome spreads on a glass slide. Regions of gain or loss of DNA are then detected as changes in the ratio of the intensities of the two fluorochromes along the chromosome. In array-CGH the chromosomes are substituted with an ordered set of targeted sequences with known chromosomal position spotted on a glass slide. Arrays can be made of genomic clones, cDNA, PCR fragments or oligonucleotides (156-158).

One of the main approaches for identifying unbalanced structural variations such as CNVs is array-CGH. This is a powerful and robust method for carrying out genome-wide scans and to find novel CNVs. Furthermore, one way to find candidate genes involved in tumorgenesis is to identify recurrent DNA copy number alterations in the tumor genome. Gain or loss of genomic material may change the level of expression of those genes encompassed in the aberration, which then modifies normal growth control and cellular pathways. Genetic alterations such as amplifications and deletions may activate proto-oncogenes or inactivate tumor suppressor genes. Characterization of these DNA copy number alterations is then important for the understanding of cancer. The advantages of array-CGH compared to metaphase-CGH, LOH studies (using microsatellite or restriction fragment length polymorphism (RFLP) markers) and FISH (fluorescent *in situ* hybridization) are the higher resolution of analysis and the power of the method allowing the investigation of the whole genome in a single experiment. Schematic illustration of array-CGH is given in Figure 5.

In the papers included in this thesis we applied an array (32K-array) established in our lab. This platform is a clone-based array composed of 32,396 bacterial artificial chromosomes (BACs) covering 99% of the current assembly of the human genome with an average resolution of up to 60 kb, which is

~100x better resolution compared to metaphase-CGH. The clone library was (159) purchased from BACPAC Resources Center at Children's Hospital Oakland Research Institute (Oakland, CA, USA; (http://bacpac.chori.org/pHumanMinSet.htm). The set of clones were amplified using three different degenerate-oligonucleotide–primed PCR primers (160) and then reamplified with a universal primer labeled with an amino group, which allows the attachment of the DNA to a glass slide. The DNA was printed onto a Codelink HD microarray slide with a high-throughput microarray printer constructed by the Lawrence Berkeley National Laboratory.

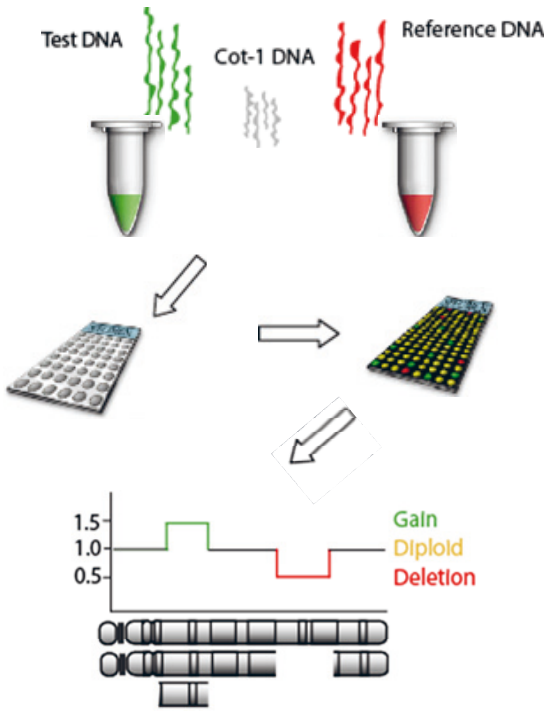

Figure 5. Array-based comparative genomic hybridization. The ratio of signal intensities detected for each spot is indicative of the relative DNA copy number in test versus reference DNA and quantitative evaluation of fluorescence intensity ratios allows the identification of loss or gain of chromosomal material in the test DNA.

Several steps including DNA labeling, testing of hybridization temperature, and washing conditions were carefully optimized to obtain the most optimal signal to noise ratio. We used 1µg of test and reference DNA which were labeled with different fluorochromes; Cy3 and Cy5 respectively. Labeled samples were then mixed with unlabeled human Cot-1 DNA to block repetitive sequences. Cot-1 DNA is enriched in repetitive sequences and binds complementary repetitive sequences in the test DNA. Subsequently, the mix-

ture was vaccum evaporated and resuspended in hybridization solution. The mixture was then hybridized to the microarray for 20 hours. Following washing steps, the microarray was scanned and the fluorescence intensity was analyzed using image analysis software. The raw data files were then uploaded to a laboratory information management system database for storage, hosted by the Linnaeus Centre for Bioinformatics (LCB; http://base.lcb.uu.se). LCB also provides tools for filtering and statistical analysis of microarray data within the LCB Data WareHouse (LCB-DWH) (161); we applied several filters to the hybridization raw data files. These filters remove oversaturated spots, spots with low signal-to-noise ratio in channels and spots either manually or automatically flagged as bad, absent or not found in the image analysis software. To remove possible dye bias or spatial effects, we also normalized all data using print-tip locally weighted scatter-plot smoothing (162). Clones were classified as balanced, gained or deleted using the open source software SMAP (163), available from Bioconductor (http://www.bioconductor.org) and within LCB-DWH (161) which was developed in parallel to our experiments. For visualization of the results a graphical viewing tool that plots all clones according to their chromosomal positions was employed.

Expression microarray

Microarrays can also be used to simultaneously measure the expression level of many thousands of genes in the genome. Expression microarrays are a very effective tool in evaluating differences in gene expression levels and identifying specific genes whose expression is altered. This method can reflect changes in the genome which are not detectable by sequencing methods or copy number analysis. Identification of these altered genes can provide clues that aid in the identification of aberrant molecular pathways underlying the disease of interest. One of the most studied human diseases by expression arrays is cancer (164-167). In 1995 a microarray-based method for high-throughput monitoring of plant gene expression was described and one year later the same group described a microarray containing over 1000 human cDNAs (168, 169). Today, multiple commercial expression microarrays are available and the market leader is the Affymetrix GeneChip which is the most frequently used microarray for gene expression. Over 3,000 scientific publications describe results from this platform (170-172). Expression arrays were applied in both paper III and paper IV and the Affymetrix Human Genome U133Plus2.0 array used provides coverage of over 47 000 transcripts. This array consists of thousands of oligonucleotide probes that are synthesized *in situ* and covalently attached to a solid matrix. Each transcript is represented by 11 pairs of 25mer oligonucleotides that serve as unique, sequence-specific detectors. Six major steps are required in this method; *i)*

preparation of RNA from the test sample, *ii)* synthesis of cDNA from the RNA sample using reverse transcriptase and an oligo-dT primer, *iii)* amplification of biotin-labeled antisense mRNA (also referred to as cRNA) using an *in vitro* transcription reaction with the cDNA as a template, *iv)* addition of cRNA to the hybridization mix and injection into the GeneChip array followed by hybridization for 16 hours, *v)* several washing and staining steps including a fluorphore conjugated to avidin that binds to biotin followed by a signal amplification step using specific antibodies, and finally *vi)* scanning of the GeneChip array with a confocal laser scanner and recording of fluorescent signals on the array. Prior to analysis, the data is preprocessed by non-specific signal correction, normalization and filtering. The intensities for each sequence are converted to a quantitative read-out of relative gene-expression levels when compared to cRNA controls of known concentration (earlier added to the hybridization mixture). A number of software packages that implement algorithms can be used for calculating the signal intensity from the array. We used packages from the Bioconductor project and the raw data was normalized using the robust multi-array average (RMA) method (173). An empirical Bayes moderated t-test was applied to search for differentially expressed genes between groups (174). By comparing two or more hybridization patterns produced on separate arrays of the same array type, one can determine differences in mRNA levels between samples (171, 175-177). This type of array has a one-color system design, which means that a single sample is hybridized to the array. Using gene expression data from the same tumors already analyzed for copy number changes can aid in the identification of candidate genes within the altered regions in a cancer genome.

Quantitative real-time polymerase chain reaction (q-PCR)

Real-time PCR (RT-PCR), also known as quantitative PCR (qPCR) is a quantitative PCR method for the determination of products generated during each cycle in a PCR reaction which is directly proportional to the amount of template prior to the start of amplification (178). The first documentation of RT-PCR was in 1993 by Higuchi *et al.* (179). This technique can be applied for measuring mRNA expression levels, DNA copy number, transgene copy number and expression analysis, allelic discrimination, and viral titers. RT-PCR is the method of choice for validation of detected copy number alterations in targeted regions because of the low cost and fast turnaround time (180). The two most common methods of performing of RT-PCR are the probe-based method and the intercalate-based method. The probe-based RT-PCR, known as TaqMan RT-PCR requires a pair of PCR primers and an additional fluorogenic probe which emits fluorescence when bound to the newly synthesized DNA. The intercalate-based method, known as SYBR-Green method, has intercalating dyes which bind to newly synthesized DNA and

emit fluorescence. In both cases the fluorescent signal in each PCR-reaction is captured at frequent intervals with a highly-sensitive camera (178).

The RT-PCR method was applied in paper II to validate selected regions which displayed altered gene copy number in the tumor tissue as detected by array-CGH. The relative copy numbers were determined using the SYBR Green method and the gene copy number in the tumor DNA was normalized to a reference gene and calibrated to normal DNA.

Fluorescent *in situ* hybridization (FISH)

FISH is a molecular cytogenetic technique used to visualize specific labeled DNA probes hybridized to chromosomal structures. The first application of this technique came in 1980 by Bauman *et al.* (181). FISH is widely used today as a diagnostic tool to identify numerical and structural abnormalities in the DNA. By this method one can detect genetic aberrations such as deletions, inversions, translocations, and gene amplifications. However, certain genetic alterations, for instance, insertions or smaller alterations may be difficult to detect by FISH (182, 183). FISH can be performed on interphase nuclei or metaphase chromosomes. The fact that FISH can be used on nondividing nuclei allows for retrospective analysis of formalin-fixed, paraffin embedded tissue which is the most common form of archived material. This technique also allows for the use of small biopsies, sometimes with as few as 40-100 cells. The specimen is first firmly attached to a microscope slide. Thereafter a number of deparaffinization (for formalin-fixed, paraffin embedded tissues) and enzyme treatment steps are implemented and the DNA within the specimen is denaturized and hybridized to a specific fluorescent probe which only binds to the region of interest. Different kinds of probes can be used but all probes must be designed to be complementary to the chromosomal region one wishes to detect. Centromeric probes, derived from repetitive centromeric sequences, are used to detect losses or gains (mostly monosomies or trisomies) of specific chromosomes. Locus specific probes are specific for individual genes and are usually used to detect translocations, inversions, amplifications or deletions of genes. After hybridization and washing, fluorescence microscopy is used to detect the fluorescent signals (182, 183). This method is currently being optimized to be used in paper III for validation of selected copy number alterations detected by array-CGH.

Results and Discussion

Paper I: Profiling of copy number variations (CNVs) in healthy individuals from three ethnic groups using a human genome 32K BAC-clone-based array

Copy number variations are recognized as a common source of genetic differences between individuals. Many studies continue to show that CNVs are found in all humans as well as in other mammals. Investigations on the occurrence of CNVs are important in providing deeper insight into the impact of CNVs on human phenotypic variability, complex behavioral traits, disease susceptibility and evolution. In this study, we first established and carefully validated a high-resolution microarray containing over 32,000 BACs and then applied this array in the profiling of CNVs in the 'healthy' human genome. The set of clones (159) covers 99% of the current assembly of the human genome with an average resolution of up to 60 kb. The clones were degenerate-oligonucleotide-primed-PCR (DOP-PCR) amplified prior to in-house printing onto microarray slides. Validation of the array included several self-self hybridizations of a healthy female reference, which we used in all subsequent hybridizations, as well as hybridizations of this female reference against different pools of healthy individuals. As part of the validation process we also performed hybridizations of different cell lines and tumor samples with previously well-characterized genotypes. We could correctly identify the expected aberrations known to be present in the cell lines and tumors samples, and we also detected a number of small aberrations which had not previously been detected. Furthermore, the results showed a very high degree of reproducibility.

In this study a series of 71 healthy individuals (44 men and 27 women) from three different ethnic groups were analyzed using the 32K-BAC array with the aim of establishing a baseline for CNVs and in addition, identifying novel CNVs. The use of the term 'baseline' in this context refers to the normal frequency and genomic distribution of CNVs in normal individuals. Thirty-three Europeans, 24 Africans and 14 Asian subjects were included and all individuals were hybridized against the same female reference. DNA was isolated from peripheral blood and profiled; the advantage of this strategy being that a risk of false positive results due to cell culturing artifacts is

completely avoided. After analyzing all hybridizations, 1,078 autosomal CNVs were detected which involved at least two neighboring BAC clones. The average number of CNVs per individual was 15.1 (minimum = 3, maximum = 36) with an average size of 357.925 bp. The average size of the CNVs was similar between gains and losses, however, a larger number of gains than deletions was observed (835 vs. 243) which suggests a lower tolerance of deletions in the human genome, as compared to gains. We also grouped all individual overlapping CNVs into CNV regions (CNVRs). We found 315 distinct regions (52 deleted, 237 gained, and 26 either gained or deleted regions). These encompassed 118.1 Mb which means that at least ~3.5% of the human genome is involved in CNVs. A proportion of the identified CNVRs overlapped with previously reported regions annotated in the Database of Genomic Variants (62.5%), but a considerable number of them (37.5%) represented new variants. Interestingly, a total of 87% of the detected CNVRs overlap with known genes verifying that they might have phenotypic consequences. Furthermore, when looking at the reported regions in detail, it could be concluded that segmental duplications (SDs) were overrepresented in deleted regions (60% vs. 46% for deletions and gains, respectively) and in the larger deleted regions we could detect as much as 70% of SD content which indicates that SDs are hotspots for chromosomal rearrangements and the formation of CNVs. We also investigated if it was possible to identify any population-specific CNV for any of the ethnic groups studied. For this we performed complete unsupervised hierarchical clustering, based on the Euclidean distances between samples (using the log2-ratios of 1,145 clones identified within CNVs). We could not observe however, any ethnic-specific clustering, in that we found no cluster containing individuals of only one population. It is possible that the resolution of the platform or number of cases studied was not high enough to detect any existing differences. An alternative explanation is that CNVs do not vary between the ethnic groups studied.

In summary, we established a 32K BAC array which is a powerful and reliable tool for detection of copy number alterations in the genome, we validated and applied it in the determination of CNVs in healthy individuals. This study contributes to the establishment of the common baseline for CNVs, which is an important resource in studies addressing predisposition to different human diseases.

Paper II: Characterization of novel and complex genomic aberrations in glioblastoma using a 32K BAC array

Glioblastoma is a devastating disease and despite treatment less than half of patients are alive 1 year from diagnosis. The genetic factors underlying glioblastoma development are still poorly understood. In this study we aimed to identify chromosomal alterations at the level of DNA copy number that may be related to glioblastoma initiation and progression. A cohort of 78 tumor samples, in addition to 46 matched peripheral blood samples, was analyzed using a high resolution 32K BAC array. This array is a very powerful tool for detection of copy number imbalances genome-wide. The regions of loss or gain identified in either constitutional or tumor-derived DNA can potentially indicate the position of putative gene(s) involved in the development of glioblastoma tumors. By profiling both tumor and blood DNA we were able to distinguish between genetic events that are tumor-specific and those present in constitutional DNA. The importance of discovering these aberrant genes is that they could serve as diagnostic and/or prognostic markers for improved early diagnosis and/or even identify targets for new therapeutic strategies.

We observed a variety of both known and novel tumor-specific copy number alterations, including not only single losses or gains but also high-copy number amplifications and homozygous deletions. The most frequent genetic aberration involving a whole chromosome was monosomy 10 followed by trisomy 7 and 83% of tumors were identified with a combination of these aberrations. Other common alterations affecting whole chromosomes were monosomy 22, trisomy 19 and 20. The most common copy number alterations involving whole p and/or q arms, and interstitial and/or terminal gains or deletions was the entire or interstitial deletion of 9p, followed by interstitial loss of 1p.

We also identified a complex pattern of aberrations including high and narrow copy number amplicons. A total of 63 cases (81% of patients) presented with amplicons, varying in number from 1 to 19 per sample (three on average). The most frequent amplicon encompassed the *EGFR* locus and was identified in 39 samples (50%). For 19 of these samples the normalized ratio for clone CTD-2026N22, which lies within the EGFR amplicon, was higher than 10, which is consistent with the presence of at least 20 DNA copies of this locus. The second most common amplicon was an amplification of the *CDK4* locus, this event being observed in 16 cases (20.5%). Amplifications of *PDGFRA* and *MDM2* were observed in 15.4% and 11.5% of tumors, respectively. In addition to these amplicons, which are known events in glioblastoma, several novel amplicons were identified that encompassed interesting candidate genes such as; *GRB10* (*growth factor receptor-bound protein*

10 on 7p12.2), *MKLN1* (*muskelin 1, intracellular mediator containing kelch motifs* on 7q32.3), *PPARGC1A* (*peroxisome proliferative activated receptor gamma, coactivator 1 alpha* on 4p15.2), *HGF* (*hepatocyte growth factor* on 7q21.11), *NAV3* (*neuron navigator 3* on 12q14.3- q21.1), *CNTN1* (*contactin 1* on 12q12), *SYT1* (*synaptotagmin I* on 12q21.2) and *ADAMTSL3* (*ADAMTS-like 3* on 15q25.1). Moreover, extremely complex amplifier genotypes with numerous high and low copy number amplicons were specifically found on chromosomes 7, 12, and 22. One sample with multiple amplicons on chromosome 7 had a maximum ratio of 16.05 for the *EGFR* locus which indicates 32 DNA copies of the *EGFR* region. Furthermore, another sample presented with high and multiple amplicons on chromosome 12. One of them, encompassing the *MDM2* locus, displayed a maximum ratio of 27, which is consistent with 54 gene copies.

This analysis also allowed us to identify numerous narrow homozygous deletions. These regions included several previously reported tumor suppressor genes known to be involved in glioblastoma development, as well as novel candidate genes. Homozygous deletion of 9p21.3, encompassing *CDKN2A*, was identified in 31 samples (40%). Interestingly, several different loci affected by biallelic deletion on chromosome 9 were also identified, which strengthens the notion that genes other than *CDKN2A* located on this chromosome may be involved in tumor development. Notably, homozygous deletion of *ELAVL2* (*embryonic lethal abnormal vision, Drosophila,-like 2* on 9p21) was observed in 16 patients, an aberration that has also been observed in pediatric gliomas (184). Several other homozygously deleted loci encompassing interesting candidate tumor suppressor genes such as *BNC2* (*basonuclin 2* on 9p22.2), *PTPLAD2* (*protein-tyrosine phosphatase-like A domain-containing protein 2* on 9p21.*3*) and *PTPRE* (*protein tyrosine phosphatase, receptor type, E* on 10q26) were also identified.

Furthermore, we also defined minimal overlapping regions (MORs) of gains and deletions. These regions are of special interest as they are likely to harbor candidate oncogenes or tumor suppressor genes. We identified 185 different regions of deletion and gain involving at least 3 tumor samples. To identify possible candidate genes within these regions, we used publically available expression data and determined the top significantly up-or down-regulated genes in comparison to normal brain. Several novel genes not previously associated with glioblastoma, in addition to known genes with a previous correlation to glioblastoma, were detected.

Finally, we compared the copy number variations detected in blood to publicly available data (Database of Genomic Variants) and a large series of healthy individuals using the same platform (185) and concluded that all

observed alterations in blood were previously categorized as normal disease-unrelated variations.

By this large comprehensive analysis of glioblastomas we concluded that not a single pair of tumors presented with identical genomic profiles. This large individual variation of genetic alterations demonstrates the underlying complexity of the disease. Whole genome profiling by array-CGH is an effective tool to highlight candidate tumorigenic genes. These genes may be used in future as targets for new diagnostic and therapeutic strategies.

Paper III: Novel amplicons in pediatric medulloblastoma identified by high-resolution genomic analysis

Patients with medulloblastoma are currently divided into high- or standard risk groups. The high risk group includes patients younger than 3 years, with incomplete surgical resection, and/or evidence of metastasis. However, the clinical parameters defining risk stratification have been shown to be insufficient (112, 186-188) and a significant proportion of patients die from the disease. In addition, the majority of survivors suffer from long-term neuroendocrine and cognitive dysfunctions due to the intensive treatment regimes required to treat their disease (114). Currently, the treatment protocol is based only on risk-stratification with over- or under-treatment as common outcomes. It is essential to increase the understanding of the biology underlying the development of medulloblastoma and also find biomarkers that could help to differentiate between high and standard risk patients.

We therefore analyzed a collection of 25 Swedish medulloblastoma samples collected over a period of more than 25 years, using 32K BAC array. We were able to generate a detailed map of aberrations across the genome. The most frequent DNA copy number aberration observed in medulloblastoma was the combination of 17p loss and 17q gain, indicative of an isochromosome 17q (i(17q)). This alteration was identified in 40 % of tumors. The second most common aberration was the entire or partial loss of 1p, identified in 32% of cases. Furthermore, monosomy 8, trisomy 7 and 19, as well as interstitial deletions of 10q and partial gains of 8p and 17q were also relatively common.

Since we had access to clinical data for all cases, we studied the correlation between chromosomal aberrations and patient outcome. We observed that all but one of the patients affected with tumors presenting with i(17q), were alive at the point of manuscript completion, and we also could determine a significant correlation (Kaplan-Meier) for the presence of this aberration with long-

er survival time, *p*-value 0.047. These results are in disagreement with previous studies showing a correlation between this abnormality and poor outcome (189, 190). It is possible that the difference between studies may be due to random sample selection. We think that further investigations in larger series of samples are needed before the presence of this aberration can be used in the clinic as a marker of disease progression and survival. We also identified three cases with very interesting profiles displaying high copy number amplicons, samples that were selected for validation using SNP-based comparative genomic hybridization (Illumina 610Q beadchips). These arrays cover the genome with over 590 000 SNPs and 21 000 markers for analyzing common CNV regions. Two of these samples were also analyzed with U133Plus2.0 expression arrays to identify candidate genes within the amplified regions. One tumor presented with two independent amplicons on chromosome 1. The narrower and most distal one encompassed 42 genes, 28 of which were up-regulated. Among them we detected *MYCL1* (*v-myc myelocytomatosis viral oncogene homolog 1, lung carcinoma derived, avian* on 1p34.2) which is known to be amplified in medulloblastoma tumors (119). The other amplified region encompassed 38 genes, 24 of which were up-regulated, and possible candidate oncogenes in this region are: *SSX2IP* (*synovial sarcoma, X breakpoint 2 interacting protein* on 1p22.3), *LMO4* (*LIM domain only 4* on 1p22.3*), PKN2* (*protein kinase N2* on 1p22.1) and *CYR61* (*cysteine-rich, angiogenic inducer, 61* on 1p22.3). In the other tumor we also identified two independent amplicons, but on chromosome 3. The more distal one included only one gene, *EOMES* (*eomesodermin homolog* on 3p21.3-p21.2), and the other encompassed 56 genes, 37 of which were up-regulated. *DCLK3* (*doublecortin-like kinase 3* on 3p22.2), *RPSA* (*ribosomal protein SA* on 3p21.3) and *CTNNB1* (*β-catenin* on 3p22.1) are potential candidate genes within this region. In the third tumor we observed an amplicon on chromosome 8 encompassing 69 genes. This tumor could not be analyzed for gene expression levels. *MYC* (*v-myc myelocytomatosis viral oncogene homolog* on 8q24.21) and *DDEF1* (*development and differentiation enhancing factor 1* on 8q24.21) are two proto-oncogenes that mapped to this region. *MYC* is known to be involved in medulloblastoma development (115, 191). Two of these three patients with high copy number amplicons are deceased but the third patient is still alive after 19.4 years. Interestingly, the patient who remains alive was shown to have amplification of the *CTNNB1* locus, and it has been shown that mutations in this gene are associated with better survival (192).

We also defined the MORs of aberrations in the series and identified 34 regions of loss or gain. These regions were distributed throughout the genome and to identify possible candidate genes within these regions, we used publically available expression data to determine the most significantly up-or down-regulated genes. Several genes already associated with medulloblastoma were mapped to these regions, but also novel candidate genes.

In conclusion, by profiling a series of medulloblastomas we could identify a number of novel genomic aberrations and various candidate genes within these regions.

Paper IV: Focal amplifications correlate with high-grade and recurrences in stage Ta bladder carcinoma

Patients with the lowest malignancy state and the most common type of bladder carcinoma (Ta) have a low risk of progression and death. However, the risk for recurrences is relatively high and between 50-70% of patients will suffer from tumor recurrence. This is clearly a problem for both the individual patient as well as for the health care system (124, 128). The prediction of tumor recurrence is a challenge and reliable biomarkers are needed in clinical practice. Today, a number of common genomic aberrations have been identified in Ta bladder cancer. These aberrations include 9q/ 9p deletions, mutation of the *fibroblast growth factor receptor 3* (*FGFR3*), loss of chromosome Y and mutation or amplification of cyclin D1 on 11q13 (137, 193, 194). However, no single predictive biomarker is currently used for molecular diagnostic tests. By analyzing the DNA copy number alterations, as well as the gene expression levels, it is possible to characterize the genetic pathways behind tumor development.

The aim of this study was to evaluate DNA copy number as well as the gene expression levels of a clinically well-characterized subset of 21 Ta bladder carcinomas. These primary Ta samples included 2 LMP (papillary urothelial neoplasm of low malignant potential), 10 low-grade and 9 high-grade tumors from 14 males and 7 females with an average age of onset of 72.3 and 64.6 years respectively. All samples were selected according to their status of recurrence. Ten of the patients had several recurrent tumors, 2-15 within 10 months to 14 years, whereas 11 had no recurrences in a follow-up time of minimum 4 years. All cases were profiled with 32K BAC array to identify copy number alterations that could be important in tumor development and to differentiate recurrent from non-recurrent cases.

The array-CGH analysis revealed losses and gains of whole chromosomes in addition to interstitial and/or terminal copy number aberrations. The most frequent alteration (detected in 45% of cases) involving a whole chromosome was heterozygous deletion of chromosome 9. This aberration was observed in both the recurrence group and in the recurrence-free group. Other chromosomal imbalances such as gain at 1q, 5p, 17q and 20q were also relatively common, but more frequently observed in the recurrence group. Interestingly, high and narrow copy number amplicons were significantly over-

represented in the recurrent group, but also among the higher grade tumors. We identified in total 18 amplicons, thus, 33% of tumors presented with amplicons. The amplicons varied in size (0.25-11.5Mb) and in number (1-7 per tumor) and included known oncogenes (e.g. *FGFR3*, *CCND1* (*cyclin D1* on 11q13.2), *MYC* and *MDM2* as well as novel candidate oncogenes such as *MYBL2* (*v-myb myeloblastosis viral oncogene homolog (avian)-like 2* on 20q13.1), *YWHAB* (*tyrosine 3-monooxygenase/tryptophan* on 20q13.1) and *SDC4* (*syndecan 4* on 20q13.12). We also identified a few small homozygous deletions (0.65 to 3.10 Mb in size). These biallelic deletions encompassed several tumor suppressor genes, such as *CDKN2A/B*, *PTEN*, *RB1* which are known to be involved in bladder cancer, in addition to a novel gene *RNASEH2B* (*ribonuclease H2, subunit B* on 13q14 .13). Furthermore, 23 MORs involving gains or losses, of varying size (0.45-53.81Mb), were identified, and oncogenes and tumor suppressor genes known to be involved in bladder cancer as well as other types of cancer were mapped to these regions. The MORs that encompassed most samples were 9p24.3-21.1, 9q22.32-34.11 and 2q.37.1 for deletions and 17q.25.1-25.3 and 20q11.21-13.33 for gains.

Gene expression levels were evaluated in 12 samples (5 cases without recurrences and 7 cases with recurrences) with the U133Plus2.0 expression array from Affymetrix and compared in recurrence versus non-recurrence groups. We did not observe any significant difference in gene expression between recurrence versus non-recurrence groups. The gene expression data was also used to identify candidate genes within the novel and narrow amplicons as well as the homozygous deleted loci when compared to gene expression data of normal bladder tissue. We identified several potential candidate genes, namely *MYBL2*, a proto-oncogene involved in cell cycle progression, *YWHAB*, which encodes an antiapoptotic protein, *SDC4*, an important component of focal adhesions, and *TOMM34*, involved in the import of precursor proteins into mitochondria. These were genes for which DNA amplification was linked to transcript up-regulation. Two of the genes encoding TOMM34 and SDC4 proteins were selected for immunohistochemistry staining. The results confirmed that protein expression also is up-regulated in the tumor with amplification.

To summarize, the whole genome profiling of Ta bladder carcinomas revealed a large individual variation in terms of copy number alterations. In addition, the significant correlation between high copy number amplicons and high-grade and recurrence cases was an interesting finding which may be useful in the identification of patients who would benefit from a more aggressive therapy.

Concluding remarks and future perspectives

Array-based comparative genomic hybridization and gene expression arrays are powerful and relevant methods for the determination of copy number variations or gene expression changes in the human genome. The establishment of the 32K BAC array covering the whole genome and the identification of a baseline for common CNVs in healthy individuals in paper I was fundamental for subsequent studies of tumor-specific CNVs. In paper II, III and IV we applied this platform in the profiling of glioblastoma, medulloblastoma and bladder cancer with the primary objective of identifying copy number imbalances that could indicate faulty genes and the pathways altered by CNVs. Numerous known and novel MORs of gains and losses were identified in all tumor types. Moreover, in paper III and IV, we explored the gene expression levels in a number of samples, to aid in the identification of possible candidate tumor suppressor genes and/or oncogenes within the regions of interest. The results derived from these studies have increased our understanding of the genetic alterations leading to the development of these different tumor types and demonstrate the potential prognostic value of array-CGH. Both the discovery of new, and confirmation of recurrent genomic alterations within cancer genomes is important as it represent the first step guiding the development of new anti-cancer agents, directed to specific proteins encoded by these genes. Targeted cancer therapy targets specific molecular elements that are essential for survival of the tumor. One example of a revolutionizing drug which targets a specific genetic aberration is Gleevec (imatinib), which is now extensively used in the treatment of chronic myeloid leukemia. By inhibiting the major driving factor of the tumor, the *bcr-abl* aberrant fusion protein, Gleevec proved to be an extremely effective treatment option. The obvious success of Gleevec underscores the importance of investigating recurrent changes in the cancer genome in order to improve future cancer therapies.

To comprehensively characterize a cancer genome at a deeper level, high-resolution and genome-wide investigations must be undertaken. The optimal cancer genome analysis would involve a combination of complementary methods such as microarrays and sequencing, as well as the use of integrative approaches, which study changes in the transcriptome, and epigenome. Such unbiased and multi-dimensional data sets would provide a more complete picture of the complex alterations that occur prior to and during tumor

development. Massively parallel DNA sequencing provides a rapid and effective way to sequence the whole or selected regions of the genome. This method has dramatically improved on earlier techniques in terms of its ability to read massive amounts of sequences in parallel, in a fast and cost-effective way. There are four different technologies currently available: 454 (Roche), SOLiD (Applied Biosystems), Solexa (Illumina) and Helicos (195-198). The workflow and principles involved in these technologies are relatively similar. They use a combination of chemical and enzymatic reactions to sequence fragments in a highly parallel fashion. The read-out is then captured by image analysis and thereafter mapped to a reference genome. The sequenced genomic regions can then provide information about genetic alterations other than copy number aberrations, such as point mutations, inversions and translocations. The knowledge obtained from sequencing data can be added to that on copy number alterations and gene expression to improve the probability of identifying candidate cancer genes that are most likely to be drivers. Driver genes in turn are those that are worthy of deeper investigation, using functional studies such as cell-lines or animal models for instance. The discrimination between driver and passenger mutations is a critical dilemma in cancer genomic studies, because not all alterations in the cancer genome actually contribute to carcinogenesis. Furthermore, the frequency of mutations in a gene is not the only relevant factor in evaluating the importance of a mutation, the type of mutation can provide useful information regarding its potential role in disease. Massively parallel DNA sequencing can also be used for whole-genome transcriptome profiling. This permits the quantification of transcript abundance and identification of novel genes and splice-variants (199).

Abnormal epigenomic states, such as DNA hyper- or hypomethylation in a cancer cell are also relevant to study as they can impact on gene expression considerably. Many cancer types have been associated with changes in the epigenome that dysregulate normal transcription. The investigation of DNA methylation could for example reveal inactivated tumor suppressor genes which cause disruption of multiple pathways important for tumor development. Additionally, the understanding of how DNA methylation influences biological processes is still quite limited, and deeper knowledge of the methylation status within different tumor types could allow for enhanced diagnosis or tailor-made treatments in the future. Recent technology development has enabled analysis of DNA methylation in a genome-wide manner. DNA methylation studies can employ array-based or non-array-based methods and the procedure can be divided into two steps. The first step is the identification and enrichment of methyl-cytosines in the DNA sample. Common methods are: restriction enzyme-based method, chromatin immunoprecipitation (ChIP) and bisulfate conversion. The second step involves capturing of the enriched or modified DNA. In the array-based category, the

arrays are designed to analyze bisulfate-converted DNA or instead employ a restriction-enzyme-based methylation analysis. Non-microarray-based experimental design includes for example, massively parallel DNA sequencing after bisulfate conversion (200, 201).

The integration of copy number-, gene expression-, mutation-, and DNA methylation data will provide an overview of the pathways and networks altered during tumor development. Such information will be important for therapeutic decisions in clinics since a more personalized treatment can be achieved for each patient based on their genetic and epigenetic profile. Furthermore, identified candidate genes may in future be harnessed as targets for new diagnostic and therapeutic strategies.

Acknowledgements

This work was carried out at the Department of Genetics and Pathology, Uppsala University. Financial support provided by The Swedish Cancer Society, The Swedish Children's Cancer Foundation, The Lions Cancer Research Foundation, The Department of Genetics and Pathology and The Medical Faculty of Uppsala University is gratefully acknowledged.

I would like to take this opportunity to express my warmest gratitude to several people:

First of all I would like to thank my hard-working supervisor **Teresita Díaz de Ståhl**, for first accepting me as an exam-work student and afterwards a PhD student. I am so grateful for all your support, scientific knowledge and for always taking time to discuss and helping me with all kinds of things.

My co-supervisor **Jan Dumanski**, for your great knowledge about genomic microarrays, and for returning my phone-call five years ago when I wanted to do my exam-work in your group. My co-supervisors **Susan Pfeifer** and **Erik Blomquist**, for your clinical expertise concerning brain tumors. Extra thanks to Susan for teaching me the FISH method. My co-supervisor **Ulrika Segersten**, for an interesting collaboration in the bladder carcinoma project.

My co-authors and collaborators, especially **Robin Andersson,** for creating SMAP, and for making new plug-ins in the DWH after our needs. Thanks also to **Hanna Göransson Kultima** and **Maria Rydåker** at the Affymetrix platform. **Per-Uno Malmström**, for the bladder carcinoma project.

Former and present members of the Dumanski and Díaz de Ståhl groups: **Arek**, **Adam**, **Birgitta**, **Caisa**, **Calle**, **Cecilia**, **Johanna, Johan, Magda**, **Uwe**, **Ricardo**, **Cihan**, **Hamid**, **Lars** and **Mehdi**.

I would like give my warmest special thanks to **Johanna Sandgren** for being such a nice and caring friend during these years as a PhD student. We have shared a couple of PhD courses and conferences together, and you made it all more fun. I have really enjoyed having you in our group. I also want to give **Cihan Cetinkaya** an extra 'thank you' for being an open and caring person, and also for teaching me how to handle cells… and students ☺

All other friendly people at the Rudbeck laboratory, especially on the 3rd floor. No one mentioned and no one forgotten.

Simin Tahmasebpoor, at the Uppsala Biobank, for help with tumor samples.

Fiona Murray, for careful and excellent proof-reading.

Alla gamla kursare på Bioteknologiprogrammet på SLU, speciellt **Lena**, **Pernilla**, **Hung** och **Peter** för att ni helt enkelt är så sköna människor och för att ni gjorde tråkiga datalabbar eller grupparbeten så mycket roligare!

Mina närmsta vänner **Helena**, **Karin**, **Pernilla**, **Walle** och **Camilla**. Även fast vi inte ses så ofta nuförtiden så vet jag att ni finns där och att vi aldrig kommer att tappa kontakten. Ni har alla berikat mitt liv med fantastiska och roliga minnen och jag hoppas att vi får många fler. **Helena**, du har betytt mycket i mitt liv då vi har känt varandra i nästan 25 år! Vår 'förmåga' att veta vad den andra tänker är nästan läskig ibland ☺. Jag önskar att inte var 30 mil mellan oss.

Mina föräldrar för att ni alltid har stöttat mig och gett mig en kärleksfull och trygg uppväxt. Älskade **mamma**, jag beundrar dig för din styrka och för att du baske mig kan allt! Tack för de veckor Tindra fått varit hos sin "Mammo". Älskade **pappa**, att komma hem till dig i Hammaren ger mig kraft och ett inre lugn. Jag älskar er!

Mina älskade syskon. **Anita**, för att du är en riktig kämpe som alltid är glad och omtänksam. **Lars**, för att du helt enkelt är världens snällaste storebror och för att du inspirerat mig att plugga och doktorera i Uppsala. Det har kännts tryggt att ha dig nära, och skönt att ha dig att prata plugg och jobb med när ingen annan riktigt förstår. Jag är så stolta över er båda två! Mina syskonbarn, **Linda**, **Josefine**, **Jacob** och **Linnea**. Linda, du och jag har speciella band och du är en av mina bästa vänner, löv u!

Övrig släkt, bland andra **Elisabet**, **Ola**, **Johan** och **Lena**.

Patrik, min kärlek, jag ser fram emot vad som komma skall. Tack för att du är en sån snäll, kärleksfull, lugn och trygg person som stöttar och peppar mig när det känns tungt. **Tindra**, älskade lilla gounge! Att komma hem till dig efter jobbet när du kommer och slänger dig runt halsen på mig och ropar "Mamma!" fyller mig med sån lycka och energi! Sist, och faktiskt minst (storleksmässigt) vill jag tacka den lilla i magen, för att du vart med mig varje stund under arbetet av denna avhandling. Vem du nu än är, så älskar jag dig redan.

References

1. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, Lee C. Detection of large-scale variation in the human genome. *Nat Genet* 2004;36(9):949-51.
2. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Zhang J, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME. Global variation in copy number in the human genome. *Nature* 2006;444(7118):444-54.
3. Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. *Nat Rev Genet* 2006;7(2):85-97.
4. Frazer KA, Murray SS, Schork NJ, Topol EJ. Human genetic variation and its contribution to complex traits. *Nat Rev Genet* 2009;10(4):241-51.
5. Freeman JL, Perry GH, Feuk L, Redon R, McCarroll SA, Altshuler DM, Aburatani H, Jones KW, Tyler-Smith C, Hurles ME, Carter NP, Scherer SW, Lee C. Copy number variation: new insights in genome diversity. *Genome Res* 2006;16(8):949-61.
6. Korn JM, Kuruvilla FG, McCarroll SA, Wysoker A, Nemesh J, Cawley S, Hubbell E, Veitch J, Collins PJ, Darvishi K, Lee C, Nizzari MM, Gabriel SB, Purcell S, Daly MJ, Altshuler D. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat Genet* 2008;40(10):1253-60.
7. Shlien A, Malkin D. Copy number variations and cancer susceptibility. *Curr Opin Oncol* 2010;22(1):55-63.
8. Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Magnusson KP, Manolescu A, Karason A, Palsson A, Thorleifsson G, Jakobsdottir M, Steinberg S, Palsson S, Jonasson F, Sigurgeirsson B, Thorisdottir K, Ragnarsson R, Benediktsdottir KR, Aben KK, Kiemeney LA, Olafsson JH, Gulcher J, Kong A, Thorsteinsdottir U, Stefansson K. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet* 2007;39(12):1443-52.
9. Antoniou AC, Spurdle AB, Sinilnikova OM, Healey S, Pooley KA, Schmutzler RK, Versmold B, Engel C, Meindl A, Arnold N, Hofmann W, Sutter C, Niederacher D, Deissler H, Caldes T, Kampjarvi K, Nevanlinna H, Simard J, Beesley J, Chen X, Neuhausen SL, Rebbeck TR, Wagner T, Lynch HT, Isaacs C, Weitzel J, Ganz PA, Daly MB, Tomlinson G, Olopade OI, Blum JL, Couch FJ, Peterlongo P, Manoukian S, Barile M, Radice P, Szabo CI, Pereira LH, Greene MH, Rennert G, Lejbkowicz F, Barnett-Griness O, Andrulis IL, Ozcelik H, Gerdes AM, Caligo MA, Laitman Y, Kaufman B, Milgrom R, Friedman E, Domchek SM, Nathanson KL, Osorio A, Llort G, Milne RL, Benitez J, Hamann U, Hogervorst FB, Manders P, Ligtenberg MJ, van den Ouweland AM,

Peock S, Cook M, Platte R, Evans DG, Eeles R, Pichert G, Chu C, Eccles D, Davidson R, Douglas F, Godwin AK, Barjhoux L, Mazoyer S, Sobol H, Bourdon V, Eisinger F, Chompret A, Capoulade C, Bressac-de Paillerets B, Lenoir GM, Gauthier-Villars M, Houdayer C, Stoppa-Lyonnet D, Chenevix-Trench G, Easton DF. Common breast cancer-predisposition alleles are associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers. *Am J Hum Genet* 2008;82(4):937-48.

10. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Maner S, Massa H, Walker M, Chi M, Navin N, Lucito R, Healy J, Hicks J, Ye K, Reiner A, Gilliam TC, Trask B, Patterson N, Zetterberg A, Wigler M. Large-scale copy number polymorphism in the human genome. *Science* 2004;305(5683):525-8.
11. Hastings PJ, Lupski JR, Rosenberg SM, Ira G. Mechanisms of change in gene copy number. *Nat Rev Genet* 2009;10(8):551-64.
12. Shlien A, Malkin D. Copy number variations and cancer. *Genome Med* 2009;1(6):62.
13. Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. *Annu Rev Med* 2010;61:437-55.
14. Ghanem N, Uring-Lambert B, Abbal M, Hauptmann G, Lefranc MP, Lefranc G. Polymorphism of MHC class III genes: definition of restriction fragment linkage groups and evidence for frequent deletions and duplications. *Hum Genet* 1988;79(3):209-18.
15. Fanciulli M, Petretto E, Aitman TJ. Gene copy number variation and common human disease. *Clin Genet* 2009.
16. Riva P, Corrado L, Natacci F, Castorina P, Wu BL, Schneider GH, Clementi M, Tenconi R, Korf BR, Larizza L. NF1 microdeletion syndrome: refined FISH characterization of sporadic and familial deletions with locus-specific probes. *Am J Hum Genet* 2000;66(1):100-9.
17. Bausch B, Borozdin W, Neumann HP. Clinical and genetic characteristics of patients with neurofibromatosis type 1 and pheochromocytoma. *N Engl J Med* 2006;354(25):2729-31.
18. McCarroll SA, Huett A, Kuballa P, Chilewski SD, Landry A, Goyette P, Zody MC, Hall JL, Brant SR, Cho JH, Duerr RH, Silverberg MS, Taylor KD, Rioux JD, Altshuler D, Daly MJ, Xavier RJ. Deletion polymorphism upstream of IRGM associated with altered IRGM expression and Crohn's disease. *Nat Genet* 2008.
19. Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R, Catano G, Nibbs RJ, Freedman BI, Quinones MP, Bamshad MJ, Murthy KK, Rovin BH, Bradley W, Clark RA, Anderson SA, O'Connell R J, Agan BK, Ahuja SS, Bologna R, Sen L, Dolan MJ, Ahuja SK. The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science* 2005;307(5714):1434-40.
20. Tuzun E, Sharp AJ, Bailey JA, Kaul R, Morrison VA, Pertz LM, Haugen E, Hayden H, Albertson D, Pinkel D, Olson MV, Eichler EE. Fine-scale structural variation of the human genome. *Nat Genet* 2005;37(7):727-32.
21. Bailey JA, Yavor AM, Massa HF, Trask BJ, Eichler EE. Segmental duplications: organization and impact within the current human genome project assembly. *Genome Res* 2001;11(6):1005-17.
22. Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, Pertz LM, Clark RA, Schwartz S, Segraves R, Oseroff VV, Albertson DG, Pinkel D, Eichler EE. Segmental Duplications and Copy-Number Variation in the Human Genome. *Am J Hum Genet* 2005;77(1).

23. Chance PF, Abbas N, Lensch MW, Pentao L, Roa BB, Patel PI, Lupski JR. Two autosomal dominant neuropathies result from reciprocal DNA duplication/deletion of a region on chromosome 17. *Hum Mol Genet* 1994;3(2):223-8.
24. Shaw CJ, Bi W, Lupski JR. Genetic proof of unequal meiotic crossovers in reciprocal deletion and duplication of 17p11.2. *Am J Hum Genet* 2002;71(5):1072-81.
25. Bailey JA, Baertsch R, Kent WJ, Haussler D, Eichler EE. Hotspots of mammalian chromosomal evolution. *Genome Biol* 2004;5(4):R23.
26. Lupski JR. Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet* 1998;14(10):417-22.
27. Inoue K, Lupski JR. Molecular mechanisms for genomic disorders. *Annu Rev Genomics Hum Genet* 2002;3:199-242.
28. Kidd JM, Cooper GM, Donahue WF, Hayden HS, Sampas N, Graves T, Hansen N, Teague B, Alkan C, Antonacci F, Haugen E, Zerr T, Yamada NA, Tsang P, Newman TL, Tuzun E, Cheng Z, Ebling HM, Tusneem N, David R, Gillett W, Phelps KA, Weaver M, Saranga D, Brand A, Tao W, Gustafson E, McKernan K, Chen L, Malig M, Smith JD, Korn JM, McCarroll SA, Altshuler DA, Peiffer DA, Dorschner M, Stamatoyannopoulos J, Schwartz D, Nickerson DA, Mullikin JC, Wilson RK, Bruhn L, Olson MV, Kaul R, Smith DR, Eichler EE. Mapping and sequencing of structural variation from eight human genomes. *Nature* 2008;453(7191):56-64.
29. Shaw CJ, Lupski JR. Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease. *Hum Mol Genet* 2004;13 Suppl 1:R57-64.
30. Elliott B, Richardson C, Jasin M. Chromosomal translocation mechanisms at intronic alu elements in mammalian cells. *Mol Cell* 2005;17(6):885-94.
31. Slack A, Thornton PC, Magner DB, Rosenberg SM, Hastings PJ. On the mechanism of gene amplification induced under stress in Escherichia coli. *PLoS Genet* 2006;2(4):e48.
32. Lee JA, Carvalho CM, Lupski JR. A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell* 2007;131(7):1235-47.
33. Zhang F, Khajavi M, Connolly AM, Towne CF, Batish SD, Lupski JR. The DNA replication FoSTeS/MMBIR mechanism can generate genomic, genic and exonic complex rearrangements in humans. *Nat Genet* 2009;41(7):849-53.
34. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000;100(1):57-70.
35. Nguyen DX, Massague J. Genetic determinants of cancer metastasis. *Nat Rev Genet* 2007;8(5):341-52.
36. Garcia M, Jemal A, Ward E, Center M, Hao Y, Siegel R, Thun M. Global Cancer Facts & Figures 2007. Atlanta; 2007.
37. Chung CH, Bernard PS, Perou CM. Molecular portraits and the family tree of cancer. *Nat Genet* 2002;32 Suppl:533-40.
38. Luo J, Solimini NL, Elledge SJ. Principles of cancer therapy: oncogene and non-oncogene addiction. *Cell* 2009;136(5):823-37.
39. Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability - an evolving hallmark of cancer. *Nat Rev Mol Cell Biol* 2010;11(3):220-8.
40. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature* 2009;458(7239):719-24.
41. Schvartzman JM, Sotillo R, Benezra R. Mitotic chromosomal instability and cancer: mouse modelling of the human disease. *Nat Rev Cancer* 2010;10(2):102-15.

42. Storchova Z, Pellman D. From polyploidy to aneuploidy, genome instability and cancer. *Nat Rev Mol Cell Biol* 2004;5(1):45-54.
43. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat Med* 2004;10(8):789-99.
44. Pleasance ED, Cheetham RK, Stephens PJ, McBride DJ, Humphray SJ, Greenman CD, Varela I, Lin ML, Ordonez GR, Bignell GR, Ye K, Alipaz J, Bauer MJ, Beare D, Butler A, Carter RJ, Chen L, Cox AJ, Edkins S, Kokko-Gonzales PI, Gormley NA, Grocock RJ, Haudenschild CD, Hims MM, James T, Jia M, Kingsbury Z, Leroy C, Marshall J, Menzies A, Mudie LJ, Ning Z, Royce T, Schulz-Trieglaff OB, Spiridou A, Stebbings LA, Szajkowski L, Teague J, Williamson D, Chin L, Ross MT, Campbell PJ, Bentley DR, Futreal PA, Stratton MR. A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature* 2010;463(7278):191-6.
45. Holliday R. The inheritance of epigenetic defects. *Science* 1987;238(4824):163-70.
46. Esteller M. Epigenetics in cancer. *N Engl J Med* 2008;358(11):1148-59.
47. Eden A, Gaudet F, Waghmare A, Jaenisch R. Chromosomal instability and tumors promoted by DNA hypomethylation. *Science* 2003;300(5618):455.
48. Talbot SJ, Crawford DH. Viruses and tumours--an update. *Eur J Cancer* 2004;40(13):1998-2005.
49. Weber BL. Cancer genomics. *Cancer Cell* 2002;1(1):37-47.
50. Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, Davies H, Teague J, Butler A, Stevens C, Edkins S, O'Meara S, Vastrik I, Schmidt EE, Avis T, Barthorpe S, Bhamra G, Buck G, Choudhury B, Clements J, Cole J, Dicks E, Forbes S, Gray K, Halliday K, Harrison R, Hills K, Hinton J, Jenkinson A, Jones D, Menzies A, Mironenko T, Perry J, Raine K, Richardson D, Shepherd R, Small A, Tofts C, Varian J, Webb T, West S, Widaa S, Yates A, Cahill DP, Louis DN, Goldstraw P, Nicholson AG, Brasseur F, Looijenga L, Weber BL, Chiew YE, DeFazio A, Greaves MF, Green AR, Campbell P, Birney E, Easton DF, Chenevix-Trench G, Tan MH, Khoo SK, Teh BT, Yuen ST, Leung SY, Wooster R, Futreal PA, Stratton MR. Patterns of somatic mutation in human cancer genomes. *Nature* 2007;446(7132):153-8.
51. Pfeifer GP, Besaratinia A. Mutational spectra of human cancer. *Hum Genet* 2009;125(5-6):493-506.
52. Knudson AG. Cancer genetics. *Am J Med Genet* 2002;111(1):96-102.
53. Foulkes WD. Inherited susceptibility to common cancers. *N Engl J Med* 2008;359(20):2143-53.
54. von Koch CS, Gulati M, Aldape K, Berger MS. Familial medulloblastoma: case report of one family and review of the literature. *Neurosurgery* 2002;51(1):227-33; discussion 33.
55. Louis DN OH, Wiestler OD, Cavenee WK. WHO classification of tumours of the central nervous system. Lyon: International Agency for Research on Cancer (IARC) 2007.
56. Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, Rahman N, Stratton MR. A census of human cancer genes. *Nat Rev Cancer* 2004;4(3):177-83.
57. Feinberg AP, Ohlsson R, Henikoff S. The epigenetic progenitor origin of human cancer. *Nat Rev Genet* 2006;7(1):21-33.
58. Nowell PC, Hungerford DA. Chromosome studies on normal and leukemic human leukocytes. *J Natl Cancer Inst* 1960;25:85-109.
59. Gan HK, Kaye AH, Luwor RB. The EGFRvIII variant in glioblastoma multiforme. *J Clin Neurosci* 2009;16(6):748-54.

60. Rodenhuis S. ras and human tumors. *Semin Cancer Biol* 1992;3(4):241-7.
61. Boulalas I, Zaravinos A, Karyotis I, Delakas D, Spandidos DA. Activation of RAS family genes in urothelial carcinoma. *J Urol* 2009;181(5):2312-9.
62. Nishigaki M, Aoyagi K, Danjoh I, Fukaya M, Yanagihara K, Sakamoto H, Yoshida T, Sasaki H. Discovery of aberrant expression of R-RAS by cancer-linked DNA hypomethylation in gastric cancer using microarrays. *Cancer Res* 2005;65(6):2115-24.
63. Knudson AG, Jr. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A* 1971;68(4):820-3.
64. Veitia RA, Birchler JA. Dominance and gene dosage balance in health and disease: why levels matter! *J Pathol* 2010;220(2):174-85.
65. Kwon CH, Zhao D, Chen J, Alcantara S, Li Y, Burns DK, Mason RP, Lee EY, Wu H, Parada LF. Pten haploinsufficiency accelerates formation of high-grade astrocytomas. *Cancer Res* 2008;68(9):3286-94.
66. Levitt NC, Hickson ID. Caretaker tumour suppressor genes that defend genome integrity. *Trends Mol Med* 2002;8(4):179-86.
67. van Heemst D, den Reijer PM, Westendorp RG. Ageing or cancer: a review on the role of caretakers and gatekeepers. *Eur J Cancer* 2007;43(15):2144-52.
68. Kleihues P, Cavenee W. Pathology and genetics of tumours of the nervous system. 2000.
69. Kufe DW, Pollock RE, Weichselbaum RR. Cancer Medicine. 6th ed. London: Decker; 2003.
70. Holland EC. Gliomagenesis: genetic alterations and mouse models. *Nat Rev Genet* 2001;2(2):120-9.
71. Lindberg N, Kastemar M, Olofsson T, Smits A, Uhrbom L. Oligodendrocyte progenitor cells can act as cell of origin for experimental glioma. *Oncogene* 2009;28(23):2266-75.
72. Ohgaki H, Dessen P, Jourde B, Horstmann S, Nishikawa T, Di Patre PL, Burkhard C, Schuler D, Probst-Hensch NM, Maiorka PC, Baeza N, Pisani P, Yonekawa Y, Yasargil MG, Lutolf UM, Kleihues P. Genetic pathways to glioblastoma: a population-based study. *Cancer Res* 2004;64(19):6892-9.
73. Cairncross JG, Ueki K, Zlatescu MC, Lisle DK, Finkelstein DM, Hammond RR, Silver JS, Stark PC, Macdonald DR, Ino Y, Ramsay DA, Louis DN. Specific genetic predictors of chemotherapeutic response and survival in patients with anaplastic oligodendrogliomas. *J Natl Cancer Inst* 1998;90(19):1473-9.
74. Hartmann C, von Deimling A. Molecular pathology of oligodendroglial tumors. *Recent Results Cancer Res* 2009;171:25-49.
75. Ebert C, von Haken M, Meyer-Puttlitz B, Wiestler OD, Reifenberger G, Pietsch T, von Deimling A. Molecular genetic analysis of ependymal tumors. NF2 mutations and chromosome 22q loss occur preferentially in intramedullary spinal ependymomas. *Am J Pathol* 1999;155(2):627-32.
76. Hulsebos TJ, Oskam NT, Bijleveld EH, Westerveld A, Hermsen MA, van den Ouweland AM, Hamel BC, Tijssen CC. Evidence for an ependymoma tumour suppressor gene in chromosome region 22pter-22q11.2. *Br J Cancer* 1999;81(7):1150-4.
77. Huang B, Starostik P, Kuhl J, Tonn JC, Roggendorf W. Loss of heterozygosity on chromosome 22 in human ependymomas. *Acta Neuropathol (Berl)* 2002;103(4):415-20.
78. Ward S, Harding B, Wilkins P, Harkness W, Hayward R, Darling JL, Thomas DG, Warr T. Gain of 1q and loss of 22 are the most common changes detected by comparative genomic hybridisation in paediatric ependymoma. *Genes Chromosomes Cancer* 2001;32(1):59-66.

79. Hirose Y, Aldape K, Bollen A, James CD, Brat D, Lamborn K, Berger M, Feuerstein BG. Chromosomal abnormalities subdivide ependymal tumors into clinically relevant groups. *Am J Pathol* 2001;158(3):1137-43.
80. Aldosari N, Rasheed BK, McLendon RE, Friedman HS, Bigner DD, Bigner SH. Characterization of chromosome 17 abnormalities in medulloblastomas. *Acta Neuropathol* 2000;99(4):345-51.
81. Sarkar C, Deb P, Sharma MC. Recent advances in embryonal tumours of the central nervous system. *Childs Nerv Syst* 2005;21(4):272-93.
82. Russo C, Pellarin M, Tingby O, Bollen AW, Lamborn KR, Mohapatra G, Collins VP, Feuerstein BG. Comparative genomic hybridization in patients with supratentorial and infratentorial primitive neuroectodermal tumors. *Cancer* 1999;86(2):331-9.
83. Wen PY, Kesari S. Malignant gliomas in adults. *N Engl J Med* 2008;359(5):492-507.
84. Morokoff AP, Novak U. Targeted therapy for malignant gliomas. *J Clin Neurosci* 2004;11(8):807-18.
85. Wong ML, Kaye AH, Hovens CM. Targeting malignant glioma survival signalling to improve clinical outcomes. *J Clin Neurosci* 2007;14(4):301-8.
86. Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJ, Belanger K, Brandes AA, Marosi C, Bogdahn U, Curschmann J, Janzer RC, Ludwin SK, Gorlia T, Allgeier A, Lacombe D, Cairncross JG, Eisenhauer E, Mirimanoff RO. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med* 2005;352(10):987-96.
87. Xie Z. Brain tumor stem cells. *Neurochem Res* 2009;34(12):2055-66.
88. Nowell PC. The clonal evolution of tumor cell populations. *Science* 1976;194(4260):23-8.
89. Dean M, Fojo T, Bates S. Tumour stem cells and drug resistance. *Nat Rev Cancer* 2005;5(4):275-84.
90. Salmaggi A, Boiardi A, Gelati M, Russo A, Calatozzolo C, Ciusani E, Sciacca FL, Ottolina A, Parati EA, La Porta C, Alessandri G, Marras C, Croci D, De Rossi M. Glioblastoma-derived tumorospheres identify a population of tumor stem-like cells with angiogenic potential and enhanced multidrug resistance phenotype. *Glia* 2006;54(8):850-60.
91. Dick JE. Stem cell concepts renew cancer research. *Blood* 2008;112(13):4793-807.
92. Farrell CJ, Plotkin SR. Genetic causes of brain tumors: neurofibromatosis, tuberous sclerosis, von Hippel-Lindau, and other syndromes. *Neurol Clin* 2007;25(4):925-46, viii.
93. Kleihues P, Ohgaki H. Primary and secondary glioblastomas: from concept to clinical diagnosis. *Neuro Oncol* 1999;1(1):44-51.
94. Ohgaki H, Kleihues P. Genetic pathways to primary and secondary glioblastoma. *Am J Pathol* 2007;170(5):1445-53.
95. Nakamura M, Watanabe T, Yonekawa Y, Kleihues P, Ohgaki H. Promoter methylation of the DNA repair gene MGMT in astrocytomas is frequently associated with G:C --> A:T mutations of the TP53 tumor suppressor gene. *Carcinogenesis* 2001;22(10):1715-9.
96. Maher EA, Brennan C, Wen PY, Durso L, Ligon KL, Richardson A, Khatry D, Feng B, Sinha R, Louis DN, Quackenbush J, Black PM, Chin L, DePinho RA. Marked genomic differences characterize primary and secondary glioblastoma subtypes and identify two distinct molecular and clinical secondary glioblastoma entities. *Cancer Res* 2006;66(23):11502-13.

97. Parsons DW, Jones S, Zhang X, Lin JC, Leary RJ, Angenendt P, Mankoo P, Carter H, Siu IM, Gallia GL, Olivi A, McLendon R, Rasheed BA, Keir S, Nikolskaya T, Nikolsky Y, Busam DA, Tekleab H, Diaz LA, Jr., Hartigan J, Smith DR, Strausberg RL, Marie SK, Shinjo SM, Yan H, Riggins GJ, Bigner DD, Karchin R, Papadopoulos N, Parmigiani G, Vogelstein B, Velculescu VE, Kinzler KW. An integrated genomic analysis of human glioblastoma multiforme. *Science* 2008;321(5897):1807-12.
98. Furnari FB, Fenton T, Bachoo RM, Mukasa A, Stommel JM, Stegh A, Hahn WC, Ligon KL, Louis DN, Brennan C, Chin L, DePinho RA, Cavenee WK. Malignant astrocytic glioma: genetics, biology, and paths to treatment. *Genes Dev* 2007;21(21):2683-710.
99. Network TCGAR. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 2008;455(7216):1061-8.
100. Hawkins NJ, Lee JH, Wong JJ, Kwok CT, Ward RL, Hitchins MP. MGMT methylation is associated primarily with the germline C>T SNP (rs16906252) in colorectal cancer and normal colonic mucosa. *Mod Pathol* 2009;22(12):1588-99.
101. Esteller M, Garcia-Foncillas J, Andion E, Goodman SN, Hidalgo OF, Vanaclocha V, Baylin SB, Herman JG. Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *N Engl J Med* 2000;343(19):1350-4.
102. Hegi ME, Diserens AC, Gorlia T, Hamou MF, de Tribolet N, Weller M, Kros JM, Hainfellner JA, Mason W, Mariani L, Bromberg JE, Hau P, Mirimanoff RO, Cairncross JG, Janzer RC, Stupp R. MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med* 2005;352(10):997-1003.
103. Read TA, Hegedus B, Wechsler-Reya R, Gutmann DH. The neurobiology of neurooncology. *Ann Neurol* 2006;60(1):3-11.
104. Fan X, Eberhart CG. Medulloblastoma stem cells. *J Clin Oncol* 2008;26(17):2821-7.
105. Sutter R, Shakhova O, Bhagat H, Behesti H, Sutter C, Penkar S, Santuccione A, Bernays R, Heppner FL, Schuller U, Grotzer M, Moch H, Schraml P, Marino S. Cerebellar stem cells act as medulloblastoma-initiating cells in a mouse model and a neural stem cell signature characterizes a subset of human medulloblastomas. *Oncogene* 2010.
106. Jenkin D, Shabanah MA, Shail EA, Gray A, Hassounah M, Khafaga Y, Kofide A, Mustafa M, Schultz H. Prognostic factors for medulloblastoma. *Int J Radiat Oncol Biol Phys* 2000;47(3):573-84.
107. Dhall G. Medulloblastoma. *J Child Neurol* 2009;24(11):1418-30.
108. Polkinghorn WR, Tarbell NJ. Medulloblastoma: tumorigenesis, current clinical paradigm, and efforts to improve risk stratification. *Nat Clin Pract Oncol* 2007;4(5):295-304.
109. Rutkowski S, Bode U, Deinlein F, Ottensmeier H, Warmuth-Metz M, Soerensen N, Graf N, Emser A, Pietsch T, Wolff JE, Kortmann RD, Kuehl J. Treatment of early childhood medulloblastoma by postoperative chemotherapy alone. *N Engl J Med* 2005;352(10):978-86.
110. Ehrbrecht A, Muller U, Wolter M, Hoischen A, Koch A, Radlwimmer B, Actor B, Mincheva A, Pietsch T, Lichter P, Reifenberger G, Weber RG. Comprehensive genomic analysis of desmoplastic medulloblastomas: identification of novel amplified genes and separate evaluation of the different histological components. *J Pathol* 2006;208(4):554-63.
111. Zeltzer PM, Boyett JM, Finlay JL, Albright AL, Rorke LB, Milstein JM, Allen JC, Stevens KR, Stanley P, Li H, Wisoff JH, Geyer JR, McGuire-Cullen P,

Stehbens JA, Shurin SB, Packer RJ. Metastasis stage, adjuvant treatment, and residual tumor are prognostic factors for medulloblastoma in children: conclusions from the Children's Cancer Group 921 randomized phase III study. *J Clin Oncol* 1999;17(3):832-45.
112. Fernandez-Teijeiro A, Betensky RA, Sturla LM, Kim JY, Tamayo P, Pomeroy SL. Combining gene expression profiles and clinical parameters for risk stratification in medulloblastomas. *J Clin Oncol* 2004;22(6):994-8.
113. Lo KC, Ma C, Bundy BN, Pomeroy SL, Eberhart CG, Cowell JK. Gain of 1q is a potential univariate negative prognostic marker for survival in medulloblastoma. *Clin Cancer Res* 2007;13(23):7022-8.
114. Mulhern RK, Palmer SL, Merchant TE, Wallace D, Kocak M, Brouwers P, Krull K, Chintagumpala M, Stargatt R, Ashley DM, Tyc VL, Kun L, Boyett J, Gajjar A. Neurocognitive consequences of risk-adapted therapy for childhood medulloblastoma. *J Clin Oncol* 2005;23(24):5511-9.
115. Pfister S, Remke M, Benner A, Mendrzyk F, Toedt G, Felsberg J, Wittmann A, Devens F, Gerber NU, Joos S, Kulozik A, Reifenberger G, Rutkowski S, Wiestler OD, Radlwimmer B, Scheurlen W, Lichter P, Korshunov A. Outcome prediction in pediatric medulloblastoma based on DNA copy-number aberrations of chromosomes 6q and 17q and the MYC and MYCN loci. *J Clin Oncol* 2009;27(10):1627-36.
116. Marino S. Medulloblastoma: developmental mechanisms out of control. *Trends Mol Med* 2005;11(1):17-22.
117. Guessous F, Li Y, Abounader R. Signaling pathways in medulloblastoma. *J Cell Physiol* 2008;217(3):577-83.
118. de Bont JM, Packer RJ, Michiels EM, den Boer ML, Pieters R. Biological background of pediatric medulloblastoma and ependymoma: a review from a translational research perspective. *Neuro Oncol* 2008;10(6):1040-60.
119. McCabe MG, Ichimura K, Liu L, Plant K, Backlund LM, Pearson DM, Collins VP. High-resolution array-based comparative genomic hybridization of medulloblastomas and supratentorial primitive neuroectodermal tumors. *J Neuropathol Exp Neurol* 2006;65(6):549-61.
120. Northcott PA, Nakahara Y, Wu X, Feuk L, Ellison DW, Croul S, Mack S, Kongkham PN, Peacock J, Dubuc A, Ra YS, Zilberberg K, McLeod J, Scherer SW, Sunil Rao J, Eberhart CG, Grajkowska W, Gillespie Y, Lach B, Grundy R, Pollack IF, Hamilton RL, Van Meter T, Carlotti CG, Boop F, Bigner D, Gilbertson RJ, Rutka JT, Taylor MD. Multiple recurrent genetic events converge on control of histone lysine methylation in medulloblastoma. *Nat Genet* 2009;41(4):465-72.
121. Reardon DA, Michalkiewicz E, Boyett JM, Sublett JE, Entrekin RE, Ragsdale ST, Valentine MB, Behm FG, Li H, Heideman RL, Kun LE, Shapiro DN, Look AT. Extensive genomic abnormalities in childhood medulloblastoma by comparative genomic hybridization. *Cancer Res* 1997;57(18):4042-7.
122. Yin XL, Pang JC, Ng HK. Identification of a region of homozygous deletion on 8p22-23.1 in medulloblastoma. *Oncogene* 2002;21(9):1461-8.
123. Boyle P, Ferlay J. Cancer incidence and mortality in Europe, 2004. *Ann Oncol* 2005;16(3):481-8.
124. Kaufman DS, Shipley WU, Feldman AS. Bladder cancer. *Lancet* 2009;374(9685):239-49.
125. Mitra AP, Cote RJ. Molecular pathogenesis and diagnostics of bladder cancer. *Annu Rev Pathol* 2009;4:251-85.
126. Gallagher DJ, Milowsky MI. Bladder cancer. *Curr Treat Options Oncol* 2009;10(3-4):205-15.

127. Wolff DJ. The genetics of bladder cancer: a cytogeneticist's perspective. *Cytogenet Genome Res* 2007;118(2-4):177-81.
128. Díaz De Ståhl T, Segersten U, Malmström PU. Molecular genetics of bladder cancer: an update. *Minerva Urol Nefrol* 2008;60(4):205-16.
129. Reuter VE. The pathology of bladder cancer. *Urology* 2006;67(3 Suppl 1):11-7; discussion 7-8.
130. Gardmark T, Jahnson S, Wahlquist R, Wijkstrom H, Malmstrom PU. Analysis of progression and survival after 10 years of a randomized prospective study comparing mitomycin-C and bacillus Calmette-Guerin in patients with high-risk bladder cancer. *BJU Int* 2007;99(4):817-20.
131. Knowles MA. Molecular pathogenesis of bladder cancer. *Int J Clin Oncol* 2008;13(4):287-97.
132. Orlow I, LaRue H, Osman I, Lacombe L, Moore L, Rabbani F, Meyer F, Fradet Y, Cordon-Cardo C. Deletions of the INK4A gene in superficial bladder tumors. Association with recurrence. *Am J Pathol* 1999;155(1):105-13.
133. Bakkar AA, Wallerand H, Radvanyi F, Lahaye JB, Pissard S, Lecerf L, Kouyoumdjian JC, Abbou CC, Pairon JC, Jaurand MC, Thiery JP, Chopin DK, de Medina SG. FGFR3 and TP53 gene mutations define two distinct pathways in urothelial cell carcinoma of the bladder. *Cancer Res* 2003;63(23):8108-12.
134. van Rhijn BW, van der Kwast TH, Vis AN, Kirkels WJ, Boeve ER, Jobsis AC, Zwarthoff EC. FGFR3 and P53 characterize alternative genetic pathways in the pathogenesis of urothelial cell carcinoma. *Cancer Res* 2004;64(6):1911-4.
135. Billerey C, Chopin D, Aubriot-Lorton MH, Ricol D, Gil Diez de Medina S, Van Rhijn B, Bralet MP, Lefrere-Belda MA, Lahaye JB, Abbou CC, Bonaventure J, Zafrani ES, van der Kwast T, Thiery JP, Radvanyi F. Frequent FGFR3 mutations in papillary non-invasive bladder (pTa) tumors. *Am J Pathol* 2001;158(6):1955-9.
136. Rieger-Christ KM, Mourtzinos A, Lee PJ, Zagha RM, Cain J, Silverman M, Libertino JA, Summerhayes IC. Identification of fibroblast growth factor receptor 3 mutations in urine sediment DNA samples complements cytology in bladder tumor detection. *Cancer* 2003;98(4):737-44.
137. Sauter G, Moch H, Wagner U, Novotna H, Gasser TC, Mattarelli G, Mihatsch MJ, Waldman FM. Y chromosome loss detected by FISH in bladder cancer. *Cancer Genet Cytogenet* 1995;82(2):163-9.
138. Veltman JA, Fridlyand J, Pejavar S, Olshen AB, Korkola JE, DeVries S, Carroll P, Kuo WL, Pinkel D, Albertson D, Cordon-Cardo C, Jain AN, Waldman FM. Array-based comparative genomic hybridization for genome-wide screening of DNA copy number in bladder tumors. *Cancer Res* 2003;63(11):2872-80.
139. Knowles MA. Molecular subtypes of bladder cancer: Jekyll and Hyde or chalk and cheese? *Carcinogenesis* 2006;27(3):361-73.
140. Ruppert JM, Tokino K, Sidransky D. Evidence for two bladder cancer suppressor loci on human chromosome 9. *Cancer Res* 1993;53(21):5093-5.
141. Keen AJ, Knowles MA. Definition of two regions of deletion on chromosome 9 in carcinoma of the bladder. *Oncogene* 1994;9(7):2083-8.
142. Simoneau M, LaRue H, Aboulkassim TO, Meyer F, Moore L, Fradet Y. Chromosome 9 deletions and recurrence of superficial bladder cancer: identification of four regions of prognostic interest. *Oncogene* 2000;19(54):6317-23.
143. Lindgren D, Liedberg F, Andersson A, Chebil G, Gudjonsson S, Borg A, Mansson W, Fioretos T, Hoglund M. Molecular characterization of early-stage bladder carcinomas by expression profiles, FGFR3 mutation status, and loss of 9q. *Oncogene* 2006;25(18):2685-96.

144. Berggren de Verdier PJ, Kumar R, Adolfsson J, Larsson P, Norming U, Onelov E, Wijkstrom H, Steineck G, Hemminki K. Prognostic significance of homozygous deletions and multiple duplications at the CDKN2A (p16INK4a)/ARF (p14ARF) locus in urinary bladder cancer. *Scand J Urol Nephrol* 2006;40(5):363-9.
145. Czerniak B, Chaturvedi V, Li L, Hodges S, Johnston D, Roy JY, Luthra R, Logothetis C, Von Eschenbach AC, Grossman HB, Benedict WF, Batsakis JG. Superimposed histologic and genetic mapping of chromosome 9 in progression of human urinary bladder neoplasia: implications for a genetic model of multistep urothelial carcinogenesis and early detection of urinary bladder cancer. *Oncogene* 1999;18(5):1185-96.
146. Wada T, Berggren P, Steineck G, Adolfsson J, Wijkstrom H, Norming U, Hansson J, Hemminki K, Larsson P. Bladder neoplasms--regions at chromosome 9 with putative tumour suppressor genes. *Scand J Urol Nephrol* 2003;37(2):106-11.
147. McGarvey TW, Maruta Y, Tomaszewski JE, Linnenbach AJ, Malkowicz SB. PTCH gene mutations in invasive transitional cell carcinoma of the bladder. *Oncogene* 1998;17(9):1167-72.
148. Aboulkassim TO, LaRue H, Lemieux P, Rousseau F, Fradet Y. Alteration of the PATCHED locus in superficial bladder cancer. *Oncogene* 2003;22(19):2967-71.
149. Habuchi T, Luscombe M, Elder PA, Knowles MA. Structure and methylation-based silencing of a gene (DBCCR1) within a candidate bladder cancer tumor suppressor region at 9q32-q33. *Genomics* 1998;48(3):277-88.
150. Nishiyama H, Takahashi T, Kakehi Y, Habuchi T, Knowles MA. Homozygous deletion at the 9q32-33 candidate tumor suppressor locus in primary human bladder cancer. *Genes Chromosomes Cancer* 1999;26(2):171-5.
151. Knowles MA, Habuchi T, Kennedy W, Cuthbert-Heavens D. Mutation spectrum of the 9q34 tuberous sclerosis gene TSC1 in transitional cell carcinoma of the bladder. *Cancer Res* 2003;63(22):7652-6.
152. Mitra AP, Datar RH, Cote RJ. Molecular pathways in invasive bladder cancer: new insights into mechanisms, progression, and target identification. *J Clin Oncol* 2006;24(35):5552-64.
153. Solinas-Toldo S, Lampel S, Stilgenbauer S, Nickolenko J, Benner A, Dohner H, Cremer T, Lichter P. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* 1997;20(4):399-407.
154. Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, Dairkee SH, Ljung BM, Gray JW, Albertson DG. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* 1998;20(2):207-11.
155. Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors; 1992 Oct 30.
156. Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A, Williams CF, Jeffrey SS, Botstein D, Brown PO. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet* 1999;23(1):41-6.
157. Mantripragada KK, Buckley PG, Jarbo C, Menzel U, Dumanski JP. Development of NF2 gene specific, strictly sequence defined diagnostic microarray for deletion detection. *J Mol Med* 2003;81(7):443-51.
158. Bignell GR, Huang J, Greshock J, Watt S, Butler A, West S, Grigorova M, Jones KW, Wei W, Stratton MR, Futreal PA, Weber B, Shapero MH, Wooster

R. High-resolution analysis of DNA copy number using oligonucleotide microarrays. *Genome Res* 2004;14(2):287-95.
159. Ishkanian AS, Malloff CA, Watson SK, DeLeeuw RJ, Chi B, Coe BP, Snijders A, Albertson DG, Pinkel D, Marra MA, Ling V, MacAulay C, Lam WL. A tiling resolution DNA microarray with complete coverage of the human genome. *Nat Genet* 2004;36(3):299-303.
160. Fiegler H, Carr P, Douglas EJ, Burford DC, Hunt S, Smith J, Vetrie D, Gorman P, Tomlinson IP, Carter NP. DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones. *Genes Chromosomes Cancer* 2003;36(4):361-74.
161. Ameur A, Yankovski V, Enroth S, Spjuth O, Komorowski J. The LCB Data Warehouse. *Bioinformatics* 2006;22(8):1024-6.
162. Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, Speed TP. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* 2002;30(4):e15.
163. Andersson R, Bruder CE, Piotrowski A, Menzel U, Nord H, Sandgren J, Hvidsten TR, de Stahl TD, Dumanski JP, Komorowski J. A Segmental Maximum A Posteriori Approach to Genome-wide Copy Number Profiling. *Bioinformatics* 2008;24(6):751-8.
164. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999;286(5439):531-7.
165. Holloway AJ, van Laar RK, Tothill RW, Bowtell DD. Options available--from start to finish--for obtaining data from DNA microarrays II. *Nat Genet* 2002;32 Suppl:481-9.
166. Tan PK, Downey TJ, Spitznagel EL, Jr., Xu P, Fu D, Dimitrov DS, Lempicki RA, Raaka BM, Cam MC. Evaluation of gene expression measurements from commercial microarray platforms. *Nucleic Acids Res* 2003;31(19):5676-84.
167. Bhattacharya S, Mariani TJ. Array of hope: expression profiling identifies disease biomarkers and mechanism. *Biochem Soc Trans* 2009;37(Pt 4):855-62.
168. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 1995;270(5235):467-70.
169. Schena M, Shalon D, Heller R, Chai A, Brown PO, Davis RW. Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc Natl Acad Sci U S A* 1996;93(20):10614-9.
170. Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang C, Kobayashi M, Horton H, Brown EL. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol* 1996;14(13):1675-80.
171. Dalma-Weiszhausz DD, Warrington J, Tanimoto EY, Miyada CG. The affymetrix GeneChip platform: an overview. *Methods Enzymol* 2006;410:3-28.
172. Auer H, Newsom DL, Kornacker K. Expression Profiling Using Affymetrix GeneChip Microarrays. *Methods Mol Biol* 2009;509:35-46.
173. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003;4(2):249-64.
174. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004;3:Article3.

175. Harrington CA, Rosenow C, Retief J. Monitoring gene expression using DNA microarrays. *Curr Opin Microbiol* 2000;3(3):285-91.
176. Schinke-Braun M, Couget JA. Expression profiling using affymetrix genechip probe arrays. *Methods Mol Biol* 2007;366:13-40.
177. Roberts PC. Gene expression microarray data analysis demystified. *Biotechnol Annu Rev* 2008;14:29-61.
178. Ginzinger DG. Gene quantification using real-time quantitative PCR: an emerging technology hits the mainstream. *Exp Hematol* 2002;30(6):503-12.
179. Higuchi R, Fockler C, Dollinger G, Watson R. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Biotechnology (N Y)* 1993;11(9):1026-30.
180. D'Haene B, Vandesompele J, Hellemans J. Accurate and objective copy number profiling using real-time quantitative PCR. *Methods* 2010;50(4):262-70.
181. Bauman JG, Wiegant J, Borst P, van Duijn P. A new method for fluorescence microscopical localization of specific DNA sequences by in situ hybridization of fluorochromelabelled RNA. *Exp Cell Res* 1980;128(2):485-90.
182. Levsky JM, Singer RH. Fluorescence in situ hybridization: past, present and future. *J Cell Sci* 2003;116(Pt 14):2833-8.
183. Tanas MR, Goldblum JR. Fluorescence in situ hybridization in the diagnosis of soft tissue neoplasms: a review. *Adv Anat Pathol* 2009;16(6):383-91.
184. Wong KK, Tsang YT, Chang YM, Su J, Di Francesco AM, Meco D, Riccardi R, Perlaky L, Dauser RC, Adesina A, Bhattacharjee M, Chintagumpala M, Lau CC. Genome-wide allelic imbalance analysis of pediatric gliomas by single nucleotide polymorphic allele array. *Cancer Res* 2006;66(23):11172-8.
185. de Ståhl TD, Sandgren J, Piotrowski A, Nord H, Andersson R, Menzel U, Bogdan A, Thuresson AC, Poplawski A, von Tell D, Hansson CM, Elshafie AI, Elghazali G, Imreh S, Nordenskjöld M, Upadhyaya M, Komorowski J, Bruder CE, Dumanski JP. Profiling of copy number variations (CNVs) in healthy individuals from three ethnic groups using a human genome 32 K BAC-clone-based array. *Hum Mutat* 2008;29(3):398-408.
186. Packer RJ, Goldwein J, Nicholson HS, Vezina LG, Allen JC, Ris MD, Muraszko K, Rorke LB, Wara WM, Cohen BH, Boyett JM. Treatment of children with medulloblastomas with reduced-dose craniospinal radiation therapy and adjuvant chemotherapy: A Children's Cancer Group Study. *J Clin Oncol* 1999;17(7):2127-36.
187. Thomas PR, Deutsch M, Kepner JL, Boyett JM, Krischer J, Aronin P, Albright L, Allen JC, Packer RJ, Linggood R, Mulhern R, Stehbens JA, Langston J, Stanley P, Duffner P, Rorke L, Cherlow J, Friedman HS, Finlay JL, Vietti TJ, Kun LE. Low-stage medulloblastoma: final analysis of trial comparing standard-dose with reduced-dose neuraxis irradiation. *J Clin Oncol* 2000;18(16):3004-11.
188. Mueller S, Chang S. Pediatric brain tumors: current treatment strategies and future therapeutic approaches. *Neurotherapeutics* 2009;6(3):570-86.
189. Lamont JM, McManamy CS, Pearson AD, Clifford SC, Ellison DW. Combined histopathological and molecular cytogenetic stratification of medulloblastoma patients. *Clin Cancer Res* 2004;10(16):5482-93.
190. Pan E, Pellarin M, Holmes E, Smirnov I, Misra A, Eberhart CG, Burger PC, Biegel JA, Feuerstein BG. Isochromosome 17q is a negative prognostic factor in poor-risk childhood medulloblastoma patients. *Clin Cancer Res* 2005;11(13):4733-40.
191. Aldosari N, Bigner SH, Burger PC, Becker L, Kepner JL, Friedman HS, McLendon RE. MYCC and MYCN oncogene amplification in medulloblasto-

ma. A fluorescence in situ hybridization study on paraffin sections from the Children's Oncology Group. *Arch Pathol Lab Med* 2002;126(5):540-4.
192. Ellison DW, Onilude OE, Lindsey JC, Lusher ME, Weston CL, Taylor RE, Pearson AD, Clifford SC. beta-Catenin status predicts a favorable outcome in childhood medulloblastoma: the United Kingdom Children's Cancer Study Group Brain Tumour Committee. *J Clin Oncol* 2005;23(31):7951-7.
193. Sauter G, Moch H, Carroll P, Kerschmann R, Mihatsch MJ, Waldman FM. Chromosome-9 loss detected by fluorescence in situ hybridization in bladder cancer. *Int J Cancer* 1995;64(2):99-103.
194. Zaharieva BM, Simon R, Diener PA, Ackermann D, Maurer R, Alund G, Knonagel H, Rist M, Wilber K, Hering F, Schonenberger A, Flury R, Jager P, Fehr JL, Mihatsch MJ, Gasser T, Sauter G, Toncheva DI. High-throughput tissue microarray analysis of 11q13 gene amplification (CCND1, FGF3, FGF4, EMS1) in urinary bladder cancer. *J Pathol* 2003;201(4):603-8.
195. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 2005;437(7057):376-80.
196. Valouev A, Ichikawa J, Tonthat T, Stuart J, Ranade S, Peckham H, Zeng K, Malek JA, Costa G, McKernan K, Sidow A, Fire A, Johnson SM. A high-resolution, nucleosome position map of C. elegans reveals a lack of universal sequence-dictated positioning. *Genome Res* 2008;18(7):1051-63.
197. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Keira Cheetham R, Cox AJ, Ellis DJ, Flatbush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS, Kirk SM, Li H, Liu X, Maisinger KS, Murray LJ, Obradovic B, Ost T, Parkinson ML, Pratt MR, Rasolonjatovo IM, Reed MT, Rigatti R, Rodighiero C, Ross MT, Sabot A, Sankar SV, Scally A, Schroth GP, Smith ME, Smith VP, Spiridou A, Torrance PE, Tzonev SS, Vermaas EH, Walter K, Wu X, Zhang L, Alam MD, Anastasi C, Aniebo IC, Bailey DM, Bancarz IR, Banerjee S, Barbour SG, Baybayan PA, Benoit VA, Benson KF, Bevis C, Black PJ, Boodhun A, Brennan JS, Bridgham JA, Brown RC, Brown AA, Buermann DH, Bundu AA, Burrows JC, Carter NP, Castillo N, Chiara ECM, Chang S, Neil Cooley R, Crake NR, Dada OO, Diakoumakos KD, Dominguez-Fernandez B, Earnshaw DJ, Egbujor UC, Elmore DW, Etchin SS, Ewan MR, Fedurco M, Fraser LJ, Fuentes Fajardo KV, Scott Furey W, George D, Gietzen KJ, Goddard CP, Golda GS, Granieri PA, Green DE, Gustafson DL, Hansen NF, Harnish K, Haudenschild CD, Heyer NI, Hims MM, Ho JT, Horgan AM, Hoschler K, Hurwitz S, Ivanov DV, Johnson MQ, James T, Huw Jones TA, Kang GD, Kerelska TH, Kersey AD, Khrebtukova I, Kindwall AP, Kingsbury Z, Kokko-Gonzales PI, Kumar A, Laurent MA, Lawley CT, Lee SE, Lee X, Liao AK, Loch JA, Lok M, Luo S, Mammen RM, Martin JW, McCauley PG, McNitt P, Mehta P, Moon KW, Mullens JW, Newington T, Ning Z, Ling Ng B, Novo SM, O'Neill MJ, Osborne MA, Osnowski A, Ostadan O, Paraschos LL, Pickering L, Pike AC, Chris Pinkard D, Pliskin DP, Podhasky J, Quijano VJ, Raczy C, Rae VH, Rawlings SR, Chiva Rodriguez A, Roe

PM, Rogers J, Rogert Bacigalupo MC, Romanov N, Romieu A, Roth RK, Rourke NJ, Ruediger ST, Rusman E, Sanches-Kuiper RM, Schenker MR, Seoane JM, Shaw RJ, Shiver MK, Short SW, Sizto NL, Sluis JP, Smith MA, Ernest Sohna Sohna J, Spence EJ, Stevens K, Sutton N, Szajkowski L, Tregidgo CL, Turcatti G, Vandevondele S, Verhovsky Y, Virk SM, Wakelin S, Walcott GC, Wang J, Worsley GJ, Yan J, Yau L, Zuerlein M, Mullikin JC, Hurles ME, McCooke NJ, West JS, Oaks FL, Lundberg PL, Klenerman D, Durbin R, Smith AJ. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 2008;456(7218):53-9.

198. Pushkarev D, Neff NF, Quake SR. Single-molecule sequencing of an individual human genome. *Nat Biotechnol* 2009;27(9):847-52.
199. Northcott PA, Rutka JT, Taylor MD. Genomics of medulloblastoma: from Giemsa-banding to next-generation sequencing in 20 years. *Neurosurg Focus* 2010;28(1):E6.
200. Chen X, Jorgenson E, Cheung ST. New tools for functional genomic analysis. *Drug Discov Today* 2009;14(15-16):754-60.
201. Cheung HH, Lee TL, Rennert OM, Chan WY. DNA methylation of cancer genome. *Birth Defects Res C Embryo Today* 2009;87(4):335-50.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 550

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-121957

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2010