Genomics of medulloblastoma: from Giemsa-banding to next-generation sequencing in 20 years

Paul A. Northcott, M.Sc.,1–3 James T. Rutka, M.D., Ph.D.,1,3 and Michael D. Taylor, M.D., Ph.D.1–3

1Division of Neurosurgery, Arthur and Sonia Labatt Brain Tumour Research Centre; 2Program in Developmental and Stem Cell Biology, The Hospital for Sick Children; and 3Department of Laboratory Medicine and Pathobiology, University of Toronto, Ontario, Canada

Advances in the field of genomics have recently enabled the unprecedented characterization of the cancer genome, providing novel insight into the molecular mechanisms underlying malignancies in humans. The application of high-resolution microarray platforms to the study of medulloblastoma has revealed new oncogenes and tumor suppressors and has implicated changes in DNA copy number, gene expression, and methylation state in its etiology. Additionally, the integration of medulloblastoma genomics with patient clinical data has confirmed molecular markers of prognostic significance and highlighted the potential utility of molecular disease stratification. The advent of next-generation sequencing technologies promises to greatly transform our understanding of medulloblastoma pathogenesis in the next few years, permitting comprehensive analyses of all aspects of the genome and increasing the likelihood that genomic medicine will become part of the routine diagnosis and treatment of medulloblastoma.

(DOI: 10.3171/2009.10.FOCUS09218)

Key Words • medulloblastoma • genomics • microarrays • next-generation sequencing

Genomics involves the study of genes and their function, typically in the context of an organism, a tissue, or a particular cell type. Cancer is a genomic disease that accounted for an estimated ~640,000 deaths in the US and Canada in 2008.73,97 The goal of cancer genomics is to develop a comprehensive inventory of the full spectrum of mutations, whether inherited or acquired, that contribute to tumorigenesis. Ultimately, through a better understanding of the cancer genome, targeted treatment options can be developed and implemented so that deaths due to cancer can be reduced in the future.

The human genome consists of ~3 billion base pairs of DNA and encodes an estimated ~24–25,000 unique protein-coding genes.71,133 During tumorigenesis, a variety of somatic mutations arise at the level of the genome, collectively providing a selective growth advantage to cells harboring these mutations and promoting the onset of cancer. Some examples of the somatic mutations in the cancer genome include single base pair substitutions, insertions and deletions of DNA segments, structural rearrangements such as duplications, inversions, and translocations, as well as gene amplifications and deletions.149 Estimates from recent genome-wide sequencing efforts have suggested that a given cancer may contain anywhere from 40 to over 100 somatic mutations.78,125,173 Since these numbers do not directly account for genes affected by structural changes and copy number aberrations, the actual number of genes targeted for mutation in a given tumor is probably even higher. Beyond the genome, deregulation of the epigenome, including hypermethylation of gene promoters and changes to the histone code, also contributes to cellular transformation.23,75,76,156 Collectively, these abnormal genomic and epigenomic states in a cancer cell aberrantly impact gene expression, leading to the disruption of normal cellular processes, including cell division. Comprehensive cancer genomics, therefore, include studies at the level of the genome, epigenome, and transcriptome.

Medulloblastoma is an embryonal tumor of the cerebellum and the most common malignant brain tumor...
in childhood.55,101 Although 5-year overall survival rates have reached 60–80%, survivors often face a variety of long-term neurological, neuroendocrine, and social sequelae as a result of conventional treatment regimens (surgery, radiotherapy, and chemotherapy).30,46 It is therefore imperative for us to gain a better understanding of the genes driving medulloblastoma oncogenesis so that future targeted therapies that are more effective and less toxic can be made available.

Much of our current understanding of the molecular basis of medulloblastoma has been derived from insights into hereditary tumor syndromes151 and candidate gene approaches focused on developmental signaling pathways.60,96,170 For instance, individuals with Gorlin or Turcot syndrome possess germline mutations in the PTCH1 and APC tumor suppressor genes, respectively, and are predisposed to medulloblastoma, among other cancers.5,51,62,63 Studies of the PTCH1 gene in Gorlin syndrome and sporadic medulloblastomas, as well as knock-out studies of its mouse homolog, Ptc, have helped to establish a role for aberrant SHH signaling in 25–35% of medulloblastomas.55,96 Similarly, the identification of APC mutations in Turcot syndrome and more frequent mutations of CTNNB1 in sporadic cases have implicated the Wnt signaling cascade in 10–15% of patients with medulloblastomas.55,96 Furthermore, patients with Li-Fraumeni syndrome have germline TP53 mutations and can have a broad spectrum of cancer types, including medulloblastoma.95,148

Aside from what has been learned from the study of these familial tumor syndromes, the majority of additional oncogenes and tumor suppressor genes implicated in medulloblastoma have been discovered from a priori candidate gene selection. Mutational screening has further implicated additional SHH (SUFU and SMO) and Wnt (AXIN1) pathway genes.55,96 In addition, the Notch pathway is deregulated in a subset of human medulloblastomas and activated in certain mouse models.55,96 Furthermore, candidate epigenetic approaches have revealed hypermethylation of the promoter regions of known tumor suppressor genes: HIC1, RASSFIA, CASP8, and others.90 The relevance of several of these genes has been further validated in mouse models of medulloblastoma.80,129,134 Although these single-gene and/or candidate-gene studies have shed significant light on our understanding of medulloblastoma pathogenesis, the candidates identified to date very likely represent only a small piece of the genomic puzzle responsible for the onset and progression of this pediatric tumor. Indeed, recent data from whole-genome sequencing projects of multiple tumor types implicate as many as 100 mutated genes per genome.29,30,72,135,153 At present, cytogenetic aberrations affecting chromosome 17 remain the most common structural changes noted in medulloblastoma (Table 1);17,19,59 Of significant interest, both of these groups reported an isochromosome of the long arm of chromosome 17 [i(17q)] as the most frequent structural abnormality and, in at least a few cases, the only aberration observed. Isochromosome 17q is the most common iso-chromosome in human cancer,105 generating a net loss of 1 copy of the majority of 17p and a net gain of 1 copy of 17q (Fig. 2A). Frequent loss of heterozygosity on chromosome 17p in medulloblastoma was independently confirmed by multiple groups in the early 1990s, typically through deletion mapping by restriction fragment length polymorphism (RFLP) analysis.29,30,72,135,153 At present, cytogenetic aberrations affecting chromosome 17 remain the most prevalent targets of gene amplification and, in at least a few cases, the only aberration observed. Isochromosome 17q is the most common iso-chromosome in human cancer,105 generating a net loss of 1 copy of the majority of 17p and a net gain of 1 copy of 17q (Fig. 2A). Frequent loss of heterozygosity on chromosome 17p in medulloblastoma was independently confirmed by multiple groups in the early 1990s, typically through deletion mapping by restriction fragment length polymorphism (RFLP) analysis.29,30,72,135,153 At present, cytogenetic aberrations affecting chromosome 17 remain the most common structural changes noted in medulloblastoma (Table 1);17,19,59,126,138 However, insight into the individual gene or combination of genes on this chromosome that drive tumorigenesis has not significantly improved since these early findings.

The establishment and cytogenetic characterization of permanent medulloblastoma cell lines and xenografts in the late 80s and early 90s also provided initial insight into the prevalence of oncogene amplification in medulloblastoma. Amplification of the MYC locus on 8q24, often in the form of double minutes, was reported in multiple cell lines and confirmed in primary tumors by several groups.8,18,48,130,157 The MYC family of protooncogenes (MYC, MYCN, and MYCLI) remain among the most prevalent targets of gene amplification in medulloblastoma (Fig. 2B).19,126

The application of CGH to the cytogenetic characterization of medulloblastoma in the late 90s resulted in a much greater appreciation of the degree of genomic imbalance present in this cancer. Using CGH to profile a panel of 27 primary medulloblastomas, Reardon and colleagues136 described frequent losses on chromosomes 10q, 11, 16q, 17p, and 8p as well as recurrent gains on chromosomes 7 and 17q. Several complementary follow-up studies based on a combination of G-banding, CGH, SKY, and FISH confirmed these now well-recognized regions of genomic instability in medulloblastoma and...
shed light on additional candidate regions. An innovative study involving members of our group retrospectively documented a series of 19 primary medulloblastomas (in addition to 5 sPNETs) using classic G-banding, FISH, CGH, and SKY. Spectral karyotyping is a multicolored FISH procedure that permits the identification of structural rearrangements and origins of marker chromosomes in the genome in a single experiment. This “chromosome painting” technique is particularly useful for detecting structural aberrations lacking a net change in copy number, such as balanced translocations. The use of SKY in this study enabled the comprehensive identification of recurrent structural rearrangements in medulloblastoma, including those on chromosomes 7, 17, 3, 14, 10, and 22—something not possible through the use of G-banding or CGH alone.

Although these cumulative efforts provided the pediatric brain tumor community with relatively detailed summaries of the medulloblastoma karyotype, not until the advent of new technologies capable of detecting copy number changes at a much higher resolution could novel candidate oncogenes and tumor suppressors in medulloblastoma be more efficiently identified through the use of genomics. Over the past 5 years, novel, high-resolution (that is, submegabase) genomic technologies have become available. The applications of some of these technologies in studies of the medulloblastoma genome are discussed in detail below.

**High-Resolution Analysis of Medulloblastoma: Digital Karyotyping, aCGH, and SNP Genotyping Arrays**

High-resolution genomic profiling of medulloblastoma has recently implicated multiple candidate oncogenes that are recurrently amplified in this malignancy (Table 2). In 2005, 2 very similar but independent studies led by Greg Riggins and Hai Yan used digital karyotyping to identify novel regions of copy number aberrations in the medulloblastoma genome. Digital karyotyping uses short sequence tags derived from specific genomic loci to provide a quantitative and relatively high-resolution profile...
of copy number aberrations throughout the genome. Boon et al. karyotyped 5 medulloblastoma cell lines by sequencing ~200,000 genomic tags per genome, identifying amplification of the OTX2 homeobox gene on chromosome 14q22 in the D487Med cell line. Using quantitative PCR, the authors confirmed recurrent amplification of OTX2 in both medulloblastoma cell lines (D425Med) and primary tumors. In the parallel study published by Di et al., OTX2 amplification was revealed in the D458Med cell line, also by digital karyotyping. In addition, by using data from SAGE libraries and quantitative RT-PCR, OTX2 (homeobox protein OTX2) was shown to be specifically overexpressed in medulloblastomas, especially lesions of anaplastic histology, as compared with a wide variety of other malignancies. Furthermore, the inhibition of OTX2 expression by siRNA-mediated knockdown or all-trans retinoic acid (ATRA) repressed medulloblastoma cell growth in vitro, suggesting that OTX2 may represent an attractive target for therapy, particularly in medulloblastomas of the anaplastic subtype.

Among the first authors to apply aCGH to medulloblastomas were Mendrzyk et al., who profiled 47 primary cases. These authors confirmed typical cytogenetic abnormalities including gains of chromosomes 17q, 7, and 1q as well as losses of 17p, 11p, 10q, and 8. Importantly, they also identified a minimal region of recurrent, high-level amplification targeting the CDK6 protooncogene on chromosome 7q21.2. They validated CDK6 copy number aberrations by using FISH and established a negative correlation between moderate-high CDK6 protein expression and overall survival by measuring CDK6 status on a medulloblastoma TMA.

Note that FOXG1 is another candidate gene implicated in medulloblastoma pathogenesis given its recurrent gain on 14q12 as revealed by aCGH analysis. Adesina et al. have analyzed a small panel of medulloblastomas...
Genomics of medulloblastoma

TABLE 1: Most prominent cytogenetic aberrations in medulloblastoma

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Gains Frequency (%)</th>
<th>Losses Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1q</td>
<td>19</td>
<td>8p</td>
</tr>
<tr>
<td>2p</td>
<td>20</td>
<td>9q</td>
</tr>
<tr>
<td>7</td>
<td>24</td>
<td>10q</td>
</tr>
<tr>
<td>17q</td>
<td>44</td>
<td>11p</td>
</tr>
<tr>
<td>i(17q)</td>
<td>30</td>
<td>16q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17p</td>
</tr>
</tbody>
</table>

* Frequency based on results of 500,000 SNP array profilings of 122 primary medulloblastomas.

An earlier aCGH study of medulloblastoma focused on a series of 16 primary cases and 3 medulloblastoma cell lines. The authors noted a novel region of homozygous deletion on chromosome 6q23 in the DAOY cell line that targeted only 2 previously uncharacterized genes, both of which exhibited reduced expression in a large percentage of primary medulloblastomas analyzed. Our group has since validated this region of homozygous deletion in DAOY and functionally confirmed L3MBTL3 as a putative medulloblastoma suppressor gene mapping to this locus.

Medulloblastomas can be histologically classified into 5 recognizable subtypes: classic, desmoplastic, anaplastic, large-cell, and medulloblastoma with extensive nodularity. The classic medulloblastoma is by far the most common, followed by the desmoplastic subtype, which makes up ~ 10–20% of cases, and large-cell and anaplastic tumors, which account for ~ 5–10% of cases. Although there is considerable variability in terms of patient outcome between the different histological subtypes and although histological staging has proven to be less than ideal method of stratification, there is a great deal of interest in defining their molecular basis. To gain an improved understanding of the genomics of desmoplastic medulloblastomas, Ehrbrecht et al. performed conventional CGH on a set of 22 sporadic cases of this subtype, followed by aCGH on a subset. In their analysis, novel regions of amplification were reported on chromosomes 9p and 17q22–24, implicating candidate oncogenes in these regions. Notably, JMJD2C was suggested as a putative oncogene mapping to the amplified region found on 9p, and we have since identified it as recurrently amplified and overexpressed in an independent sample cohort and affecting the state of methylation on histone lysines in normal progenitor cells of the developing cerebellum (that is, CGNPs).

There have been an impressive number of inquiries into the relationship between developmental signal transduction pathways and their role in medulloblastoma. Mutations in the Wnt, SHH, and Notch pathways have been well described in the medulloblastoma literature. Nonetheless, a comprehensive understanding of how specific genomic events contribute to aberrant signaling of these pathways has not been established. An important finding relevant to deregulated Wnt signaling in medulloblastoma was reported in 2006 in 2 independent but related studies. Clifford et al. have profiled 19 primary medulloblastomas by aCGH, with the specific intent to genomically describe tumors with Wnt pathway activation (nuclear β-catenin; CTNNB1 or APC mutation). Interestingly, in both the initial cohort (19 cases) and a validation series (32 cases), a single copy deletion of chromosome 6 (monosomy 6) was found exclusively in the Wnt pathway tumors. Identical findings were reported by Thompson et al., who consistently observed a correlation between the Wnt pathway signature (Wnt pathway expression; CTNNB1 mutation) and markedly reduced expression of genes mapping to chromosome 6 because of deletion. Monosomy 6 is now widely accepted in the medulloblastoma community as a genomic marker of Wnt pathway tumors that is consistently associated with CTNNB1 mutation.

From a clinical perspective, monosomy 6/CTNNB1 mutation is among the most reliable genetic markers in medulloblastoma, correlating with a highly favorable prognosis. Indeed, 100% of patients found to belong to the Wnt immunohistochemical category in the recent St Jude Medulloblastoma-96 clinical trial were event free at 5-years, compared with only 65% of patients in the SHH category.

Neurosurg Focus / Volume 28 / January 2010
Very recently, Pfister and colleagues proposed a model for molecular risk stratification of pediatric medulloblastoma based on DNA copy number aberrations affecting chromosomes 6q, 17q, and MYC/MYCN loci. Using aCGH, the authors initially profiled 80 primary medulloblastomas in an attempt to identify genomic aberrations of prognostic value, and found a gain of chromosome 6q, amplification of MYC and MYCN, isolated gain of 17q, and i(17q) all to be associated with a poor clinical outcome. In contrast, the loss of chromosome 6q was indicative of an excellent prognosis, consistent with findings in the current literature. Validation of these prognostic markers in a nonoverlapping set of 260 primary cases by using interphase FISH on a medulloblastoma TMA, Pfister et al. were able to establish an elegant staging system whereby patient outcome could be predicted based on the genomic status of only 4 markers (arranged from worst to best outcome): MYC/MYCN amplification, 6q gain, 17q gain, 6q/17q balanced, and 6q loss.

Although several of the aCGH studies described above have been informative and have enhanced our understanding of the medulloblastoma genome, most have profiled relatively modest sample cohorts (median sample size: ~24 cases) using arrays that—although an improvement over classic CGH—are of insufficient density and thus resolution (median resolution ~ 500 kb) to detect very focal genetic events. To address these caveats, we retrospectively collected an unprecedented cohort of 201 fresh-frozen primary medulloblastomas and 11 medulloblastoma cell lines and analyzed their genomes using high-resolution SNP genotyping arrays. These oligonucleotide arrays consisted of 25mer probes designed to detect the genotype (that is, A or B allele) of known SNPs at loci distributed across the genome. The median intermarker distances (that is, resolution) of probes on the 100,000 and 500,000 arrays used in this study were 8.5 and 2.5 kb, respectively, which are at least an order of magnitude higher in terms of resolution than any previous array-based study of the medulloblastoma genome (Fig. 3). In addition to reporting cytogenetic gains and losses at frequencies already known in medulloblastoma (Fig. 4 and Table 1), this strategy of profiling a large number of samples on a high-resolution platform led to the identification of 191 high-level amplifications and 159 homozy-
Neurosurg Focus / Volume 28 / January 2010

Genomics of medulloblastoma

gous deletions, most of which had not been reported in medulloblastoma. Surprisingly, only 12 recurrent amplifications were identified, and a mere 6 homozygous deletions were found in more than 1 sample. Of the recurrent homozygous deletions, *EHMT1*, a euchromatic histone (H3K9) methyltransferase, was the lone gene mapping to the minimal common region of a deleted region on chromosome 9q34, suggesting that it may represent a novel tumor suppressor gene in medulloblastoma. Note that *EHMT1* functions as part of a transcriptional repressor complex that mediates gene silencing by promoting dimethylation of H3K9 (H3K9me2), a repressive epigenetic modification, in the promoter regions of target genes. The expression of *EHMT1* was shown to be significantly downregulated at both the mRNA and protein level, and staining for both *EHMT1* and H3K9me2 on a medulloblastoma TMA showed a significant correlation between *EHMT1* status and the H3K9 methylation state, consistent with a model in which the loss of *EHMT1* leads to H3K9 hypomethylation in medulloblastoma. In addition to *EHMT1*, 7 other genes with a putative role in the regulation of histone lysine methylation were also found to be the target of focal copy number aberrations in the data set, including *SMYD4*, *L3MBTL2*, *L3MBTL3*, *SCML2*, *JMJD2C*, *JMJD2B*, and *MYST3*. Recurrent targeting of genes sharing a common role in the modulation of histone lysine residues in medulloblastoma suggests that deregulation of the histone code, particularly histone lysine methylation, very likely contributes to the pathogenesis of at least some medulloblastomas.

**Medulloblastoma Transcriptome Profiling**

Typically, strategies aimed at the transcriptional profiling of cancer have involved the comparison of gene expression signatures obtained for normal and neoplastic tissues (Fig. 5). In one of the earliest studies of medulloblastoma gene expression profiling, Michiels et al. used SAGE to compare genes expressed in medulloblastoma...
Fig. 5. Heatmap showing gene expression data for 10 normal cerebellum (CB) samples (5 fetal and 5 adult) and 50 primary medulloblastomas analyzed using Affymetrix arrays. Genes exhibiting elevated expression are shown in red, whereas genes with reduced expression are depicted in green. Unsupervised hierarchical clustering of samples using the most differentially expressed genes results in a clear distinction between normal cerebellar samples and primary tumors.
Genomics of medulloblastoma

with those in the fetal brain (24.5 weeks). Serial analysis of gene expression uses DNA sequencing technology to digitally quantify mRNA abundance by counting the frequency with which a short sequence tag (that is, transcript) appears in a cDNA library. Serial analysis of gene expression technology is very effective in quantifying gene expression and in identifying novel genes/transcripts, as no a priori knowledge of the genome under study is required. In the study by Michiels and colleagues, ~ 10,000 tags were sequenced in both medulloblastoma and fetal brain, with ~ 6000 unique genes identified in each transcriptome. A comparison of the SAGE data revealed 138 genes with significant differential expression between the 2 sources, including ZIC1 (zinc finger protein ZIC 1) and OTX2, both showing significantly elevated expression in medulloblastoma that was confirmed by Northern blotting in multiple independent samples. As these genes are highly expressed in cerebellar germinal zones (that is, the external granule layer and subventricular zone), this study provided early clues into the origins of medulloblastoma.

The presence of disseminated disease at diagnosis is a strong, independent indicator of a poor prognosis for medulloblastoma, occurring in ~ 1 of every 3 cases. Therefore, understanding the molecular basis of metastatic medulloblastoma is extremely important clinically. The first study to specifically compare metastatic (M+) with nonmetastatic (M0) medulloblastoma at a gene expression level was performed by MacDonald and colleagues in 2001. Twenty-three primary medulloblastomas designated as either M+ or M0 were analyzed using Affymetrix G110 cancer arrays, which identified 85 genes as differentially expressed between the 2 classes. Using a supervised class prediction algorithm, this 85-gene signature classified the M+ and M0 tumors with 72% accuracy. Interestingly, PDGFR and members of the Ras/MAPK signaling cascade were reported to be significantly upregulated in metastatic versus nonmetastatic cases. The overexpression of PDGFR in metastatic disease was confirmed by immunohistochemistry in an independent set of tumors. In vitro assays performed in the DAOY medulloblastoma cell line showed that the
PDGF ligand activated the Ras/MAPK pathway and promoted cell migration in this system, whereas neutralizing antibodies against PDGFR attenuated MAPK signaling and prevented ligand-mediated migration. In a follow-up study by Gilbertson and Clifford,24 the association between PDGFR overexpression and metastatic medulloblastoma was confirmed, further supporting the validity of this pathway as a candidate for targeted therapy.

Undoubtedly, one of the most seminal papers in medulloblastoma genomics was written by Pomeroy and colleagues,132 who surveyed the expression profiles of a large series of primary brain tumors and made at least 3 findings of clinical significance. Initially, 42 samples consisting of medulloblastomas (10 patients), atypical teratoid rhabdoid tumors (5 patients), renal and extra-renal rhabdoid tumors (5 patients), sPNETs (8 patients), and nonembryonal brain tumors (malignant glioma; 10 patients), and normal human cerebellum (4 patients) were analyzed using Affymetrix HuGeneFL arrays containing ~6000 known genes. Based on differentially expressed transcripts, these authors showed a clear distinction between the different tumor types, establishing that histologically similar tumors such as medulloblastomas and sPNETs were molecularly distinct. The molecular distinction between medulloblastomas and sPNETs has important clinical implications; because of their similar histology, these tumors were historically classified under the same broad category of PNET and thus were treated with the same protocols.141 In recent years, however, it has become evident that medulloblastomas and sPNETs are both molecularly and biologically distinct, with sPNETs typically exhibiting a worse prognosis.31,89,100,127,137,155

These same authors also compared the expression profiles in a set of 34 medulloblastomas of either classic (25 cases) or desmoplastic (9 cases) histology, revealing a notable degree of statistically significant differential expression between the 2 subtypes. Genes signifying desmoplastic medulloblastoma included PTCH, GLI, MYCN, and IGFI, all of which are now well-described targets of SHH signaling. Although a link between mutations in the SHH pathway and medulloblastoma pathogenesis had already been discovered, this study was among the first to document an association between aberrant SHH signaling and sporadic desmoplastic medulloblastoma.

A third key finding in this study stemmed from the authors’ use of gene expression data to predict the outcome of 60 patients with medulloblastoma for whom clinical follow-up data were available. Using a class prediction algorithm, an 8-gene classification model was generated and successfully predicted the survival status for 47 of the 60 patients profiled. Genes correlated with a favorable outcome included markers of cerebellar differentiation (β-NAP, NSCL1, and TRKC) as well as genes encoding components of the extracellular matrix (lysyl hydroxylase [PLOD], collagen Type Vα1, and elastin). In contrast, genes associated with a poor outcome included those with a role in cell proliferation and metabolism (MYBL2, enolase 1, LDH, HMGI[Y], and cytochrome C oxidase) as well as ribosomal protein-encoding genes. Much like the study by MacDonald et al.,48 described earlier, the report by Pomeroy and colleagues132 demonstrates the utility of correlating gene expression profiles in medulloblastoma with a particular phenotype (that is, favorable vs poor outcome) and provides a rationale for the incorporation of similar molecular profiling strategies in the future diagnosis and treatment of patients with medulloblastoma.

Following the Pomeroy study, a number of independent groups, including our own, have engaged in medulloblastoma transcriptome profiling efforts using a variety of technologies.22,79,115,123,175 Boon et al.22 used SAGE to analyze 20 primary medulloblastomas, identifying 30 transcripts exhibiting elevated expression in tumors compared with normal cerebellum and additional regions of the brain. The cancer-testis antigen PRAME, CD24, PRL, TOP2A, MYCN, and BARHL1 were all overexpressed in this study. Data from more recent array-based studies by our group have confirmed the aberrant expression of cancer-testis antigens from the MAGE and GAGE families in medulloblastoma cell lines and in some primary samples, suggesting that these genes may be important in medulloblastoma.79

In an earlier study we used suppression subtractive hybridization to identify genes deregulated in both human and mouse (Ptc+/−) medulloblastoma compared with normal, species-matched cerebellum.135 In suppression subtractive hybridization, double-stranded cDNA libraries are first prepared from tester (that is, medulloblastoma) and driver (that is, normal cerebellum) RNA samples.864 Heat-denatured tester cDNA is subsequently digested, adapter ligated, and then hybridized with the denatured driver cDNA to generate a subtracted cDNA library that is PCR amplified and cloned into a recipient plasmid for bacterial transformation and sequencing of clones for gene identification. In this effort, over 100 upregulated cDNA fragments were identified in the human library, including ULIP (also known as DPYSL3), SOX4, NNAT, and the previously implicated BARHL1 and OTX2 genes. In addition, genes identified as upregulated in medulloblastomas from Ptc+/− mice included CCND2, TEMPO, Musashi-1, and others.

Another informative expression profile of medulloblastoma was generated by Neben et al.,152 who analyzed ~4200 genes in 35 primary medulloblastomas in an attempt to identify those associated with patient outcome. Based on mRNA levels, 54 genes were shown to be markers of poor outcome, and a subset of these (9 genes) was further evaluated by immunohistochemistry in a nonoverlapping set of 180 cases on a medulloblastoma TMA. Of these candidate genes, STK15 positivity was identified as a negative prognostic marker of overall survival, whereas other putative markers implicated in the study (that is, cyclin D1 and stathmin 1) were not.

**Molecular Classification of Medulloblastoma**

Over the past decade, significant progress has been made in how we study the cancer genome. Indeed, gene expression profiling has proven to be an effective tool for the molecular classification of cancer, including brain tumors.57,109,120,132 Following the studies of MacDonald and Pomeroy, Thompson et al.155 were the first to truly...
Genomics of medulloblastoma

establish the existence of unique molecular subgroups of medulloblastoma using gene expression data. Profiling a series of 46 primary medulloblastomas, these authors performed unsupervised hierarchical clustering with the most informative genes in the data set, identifying 5 molecular subgroups of medulloblastoma. By integrating immunohistochemistry, FISH, and mutational screening data generated from these samples, it was shown that molecular subgroups of medulloblastoma have specific genomic and clinical features. Importantly, this study was the first to demonstrate that Wnt (that is, monosomy 6 and CTNNB1 mutation) and SHH tumors (that is, PTCH1 and SUFU mutation) are closely related and marked by elevated expression.

In a recent study, Kool et al. utilized a similar integrative genomics approach to further characterize molecular subgroups of medulloblastoma. By combining array-based gene expression and copy number profiles for 52 primary cases, Kool and colleagues recapitulated the 5 molecular subgroups described by Thompson et al. and correlated the different subgroups with specific genomic and clinical features. Importantly, the authors furthered our knowledge of non-Wnt/SHH tumors (subgroups A and B), showing that the 3 remaining subgroups (C, D, and E) are closely related and marked by elevated expression of neuronal differentiation (subgroups C and D) and retinal (subgroups D and E) genes. Furthermore, metastatic disease was shown to be more highly associated with subgroups C, D, and E, providing further support for the potential stratification of patients based on molecular subgrouping.

Beyond Protein-Coding Genes: miRNAs in Medulloblastoma

Over the course of the past 5 or so years, there has been a literal explosion in the miRNA field, especially with respect to elucidating their role in human disease—in particular, cancer. Small, noncoding, single-stranded RNA molecules, miRNAs posttranscriptionally regulate gene expression through their interaction with complementary sequences in the 3’ untranslated regions of target miRNAs. Target miRNAs are either degraded or translationally repressed by specific miRNAs, depending on the degree of complementarity between the miRNA and its target. Despite an intense amount of investigation into the involvement of miRNAs in a variety of cancer types, knowledge of their role in medulloblastoma pathogenesis is still in its infancy. The few studies of the entire miRNAsome that have been conducted to date are discussed below.

Ferretti and colleagues recently performed a TaqMan quantitative RT-PCR–based profiling of 248 miRNAs in a panel of medulloblastomas (14 cases) and normal cerebellar controls (7 cases), reporting an overwhelming bias toward downregulation of miRNAs in tumors versus controls. A subset of 86 miRNAs previously reported to be expressed in neuronal tissues and/or implicated in cancer were further analyzed in a larger cohort of tumors (34 cases), with the authors selecting miR-9 and miR-125a as 2 neuronal candidates downregulated in medulloblastoma for functional studies. The expression of both miR-9 and miR-125a was induced by retinoic acid treatment of D283 medulloblastoma cells, an agent known to inhibit medulloblastoma cell proliferation. In addition, ectopic overexpression of miR-9 and miR-125a inhibited proliferation, impaired anchorage-independent growth, and promoted apoptosis of D283 cells. Truncated trkC was identified as a target for posttranscriptional repression by both miR-9 and miR-125a in this study, suggesting a possible correlation between the loss of miR-9/miR-125a and the upregulation of the proproliferative truncated trkC in medulloblastoma.

In an effort to discriminate miRNAs deregulated in SHH-driven medulloblastomas from non-SHH cases, the gene set of authors used the Gli1 expression status to stratify a panel of 31 medulloblastomas into 2 classes (Gli1high and Gli1low) before profiling a set of 250 miRNAs using TaqMan-based quantitative RT-PCR. This approach revealed a set of 34 miRNAs exhibiting significant differential expression between the 2 classes. Three candidates exhibiting reduced expression in Gli1high tumors—miR-125b, miR-324–5p, and miR-326—were chosen for functional analysis based on their predicted capacity to target and repress the SHH family members, Smo and Gli1. Indeed, all 3 candidates were shown to repress Smo mRNA levels when overexpressed in DAOY medulloblastoma cells. Additionally, the expression of these candidate miRNAs correlated with the differentiation state of cultured CGNPs, presumed cells of origin for SHH-driven medulloblastomas, and their ectopic expression reduced SHH-mediated proliferation and promoted neurite outgrowth in the same cell type.

In 2 distinct but parallel comprehensive analyses of the human and mouse medulloblastoma miRNAsomes, the miR-17/92 polycistron was identified as a putative medulloblastoma oncogene. A bona fide oncogene in B-cell lymphoma, miR-17/92 has been reported to be aberrantly expressed in a variety of human tumors. To identify miRNAs deregulated in mouse models of medulloblastoma, Uziel et al. performed unbiased NGS to quantify miRNA abundance in medulloblastoma cells isolated from spontaneous tumors of Ink4c+/−Ptc+/− or Ink4c−/−Ptc+/− and p53+/− mice as compared with wild-type control cerebellum (1 month old) and CGNPs (6 days old). This strategy revealed 26 miRNAs with elevated expression and 24 with reduced expression in the tumor models. Among upregulated miRNAs in murine medulloblastoma cells, miR-17/92 and related paralogs accounted for 9 of 26. In addition, the authors provided evidence that miR-17/92 might cooperate with SHH signaling in medulloblastoma, showing preferential upregulation of miR-17/92 in the SHH subtype by quantitative RT-PCR profiling of a small panel of human tumors (5 SHH lesions and 5 non-SHH lesions). To evaluate its oncogenic potential in a context relevant to SHH-driven medulloblastoma, miR-17/92 was retrovirally transduced into 6-day-old CGNPs isolated from both Ink4c+/−Ptc+/− and Ink4c−/− mice prior to orthotopic transplantation of miR-17/92–expressing CGNPs into immunocompromised mice. Notably, only cells derived from the Ptc−/− background developed medulloblastoma (9 of 9 cases) in this model. Furthermore, tumor cells were sensitive to
the Smo inhibitor cyclopamine, exhibited elevated Math1 and Gli1 mRNA levels, and lost expression of the wild-type Ptc allele—all markers of activated SHH signaling and supportive of a synergistic connection between miR-17/92 and SHH in these tumors.

As detailed earlier, we recently performed high-resolution SNP array profiling on a group of > 200 medulloblastomas.\(^\text{118}\) This effort revealed multiple regions of previously unreported copy number aberrations in the medulloblastoma genome, including recurrent, high-level amplification of miR-17/92 on chromosome 13q31.\(^\text{118,119}\) Subsequent interphase FISH performed on a medulloblastoma TMA consisting of a nonoverlapping series of tumors confirmed miR-17/92 amplification in ~ 6% of cases. To gain further insight into the role of miRNAs in medulloblastomas, we next used miRNA microarrays to globally profile the human medulloblastoma miRNAome in a series of 90 primary medulloblastomas and 10 normal cerebellar controls (5 fetal and 5 adult samples). Remarkably, miR-17/92 and related paralogs (miR-106a/363 and miR-106b/25) were identified as the most highly upregulated miRNAs in medulloblastoma when compared with normal cerebellum in this analysis. The combination of miR-17/92 amplification and consistent overexpression suggested miR-17/92 as a key player in medulloblastoma pathogenesis.

As shown by the aforementioned studies of Thompson\(^\text{154}\) and Koo,\(^\text{82}\) medulloblastomas can be classified into unique molecular subgroups based on distinct gene expression signatures and specific genomic and genetic features.\(^\text{82,154}\) Using Affymetrix exon arrays to comprehensively profile the transcriptome of the same 90 primary medulloblastomas analyzed by miRNA microarray, we described 4 distinct molecular subgroups of medulloblastoma.\(^\text{118}\) These subgroups include the well-characterized Wnt and SHH subgroups described earlier, as well as 2 independent subgroups we have designated Groups C and D. Through the integration of genomics (copy number), mRNA, and miRNA expression, we found that miR-17/92 was most highly expressed in SHH-driven medulloblastomas, in agreement with the observations reported by Uziel and colleagues.\(^\text{158}\) Additionally, we showed elevated miR-17/92 levels in tumors exhibiting high MYCN (SHH) and MYC (Group C, Wnt) expression, indicative of miR-17/92 transcriptional upregulation by N-Myc and Myc and confirming miR-17/92 aberrancy in a large percentage of human medulloblastomas (~ 60%). Deregulation of miR-17/92 was conserved in well-characterized, SHH-driven mouse models of medulloblastoma—Ptc\(^{+/−}\) and SmoAl—also in concordance with the findings mentioned above. Finally, using CGNPs isolated from wild-type mice, we showed that miR-17/92 is transcriptionally induced by SHH through N-Myc, maintains CGNPs in a proliferative state in the absence of SHH, and synergizes with SHH to enhance CGNP cell growth. Cumulatively, the results of Uziel et al. and our own strongly support miR-17/92 as a legitimate medulloblastoma oncogene that cooperates with SHH signaling to promote and/or enhance CGNP proliferation.

#### Beyond Genomics: the Medulloblastoma Epigenome

Until recently, the majority of cancer research efforts had focused on describing the genetic basis of cancer, studying everything from large cytogenetic aberrations to SNPs and mutations. Over the past few years, however, there has been an ever-growing volume of literature linking the deregulation of epigenetics to malignancy.\(^\text{23,42,75,76,156}\) Epigenetics is defined as “mitotically heritable changes in gene expression that are not accompanied by modifications in primary DNA sequence.” Epigenetic modifications include DNA methylation on cytosine residues, most often in the context of CpG dinucleotides, as well as posttranslational modification of histone proteins, such as methylation, acetylation, phosphorylation, and ubiquitination.\(^\text{10,64}\) Hypermethylation of CpG islands located at the 5‘ end of genes has been reported in most cancers and, either alone or in combination with genetic mechanisms (that is, gene deletion or mutation), can contribute to tumor suppressor gene silencing. Although a handful of known tumor suppressors can be silenced by promoter methylation in medulloblastoma by using candidate gene approaches (that is, HIC1, RASSF1A, and CASP8),\(^\text{58,61,65,68,91,93,140,166}\) the application of unbiased, whole-epigenome strategies to identify novel candidates have been scant to date, consisting of only those instances described in the few published reports mentioned below.

Among the earliest studies to implicate aberrant promoter methylation in medulloblastoma on a global scale was an effort led by Frühwald and colleagues\(^\text{90}\) who used the technique of restriction landmark genomic scanning to analyze DNA methylation patterns in 17 primary medulloblastomas and 5 medulloblastoma cell lines. Using this method, the authors identified methylation in up to 1% of all CpG islands in primary tumors and up to 6% in medulloblastoma cell lines. In addition, an association between hypermethylated sequences in medulloblastoma and a poor prognosis was implied. Collectively, these findings provided early evidence that epigenetic events are likely to play a role in medulloblastoma pathogenesis.

In a study using microarray-based differential methylation hybridization, Waha et al.\(^\text{167}\) identified hypermethylation of the SGC5 (secretory granule, neuroendocrine protein 1 [7B2 protein] gene) in 16 (~ 70%) of 23 primary medulloblastomas and 7 (~ 87%) of 8 medulloblastoma cell lines. Differential methylation hybridization involves a series of enzymatic digestions with methylation-sensitive restriction enzymes, and uncut (methylated) fragments are PCR-amplified before hybridization to microarrays containing probes designed to interrogate CpG islands throughout the genome.\(^\text{174}\) The expression of SGC5 was found to be downregulated in the majority of primary samples and cell lines as compared with normal cerebellar controls, and SGC5 transcription was restored in cell lines treated with the demethylating agent, 5-aza (5-aza-2′-deoxycytidine). Furthermore, the reexpression of SGC5 in the D283Med cell line resulted in growth suppression and reduced colony formation, suggesting that SGC5 may be a putative tumor suppressor gene in medulloblastoma.

P. A. Northcott, J. T. Rutka, and M. D. Taylor

Neurosurg Focus / Volume 28 / January 2010
Genomics of medulloblastoma

Pfister and colleagues developed and applied a technique known as array-based profiling of reference-independent methylation status (aPRIMES) to globally survey DNA methylation patterns in the medulloblastoma genome. This technique compares 2 differentially digested (methylation-sensitive and methylation-specific) aliquots from the same sample genome by competitive hybridization to a CpG island microarray. The advantage of using test-versus-test as opposed to test-versus-control hybridization is the avoidance of both the influences of tissue-specific methylation that may be present in the control sample and the genomic aberrations that may exist in the test sample and not in the control genome. Using this methodology, Pfister et al. showed a striking association between samples classified as either “low methylators” or “high methylators” and patient outcome, with the high-methylator group exhibiting reduced overall survival. In addition, the GLI C2H2-type zinc-finger protein family member ZIC2 was identified as a hypermethylated candidate using aPRIMES and was subsequently confirmed to be epigenetically silenced in a panel of primary medulloblastomas by using a combination of pyrosequencing and quantitative RT-PCR analysis.

In 2 technically similar yet independent genome-wide methylation studies conducted by Anderton et al. and Kongkham et al., 5-aza-treated medulloblastoma cell lines were profiled on Affymetrix expression arrays in an effort to uncover novel tumor suppressor genes silenced by aberrant promoter methylation. In the report by Anderton and colleagues, 3 medulloblastoma cell lines (D245Med, D283Med, and MED8A) were either left untreated or exposed to 5-aza, and transcripts showing increased expression by microarray in response to the DNA methyltransferase inhibitor were investigated further by bioinformatically confirming the presence of a 5’ CpG island and assessing methylation status through bisulfite sequencing. This approach combined with gene expression analysis identified COL1A2 as an epigenetically silenced candidate in medulloblastoma that is preferentially inactivated in desmoplastic and noninfant (> 3 years) desmoplastic cases.

Kongkham et al. performed a similar genome-wide 5-aza screen in a larger group of 9 medulloblastoma cell lines but incorporated multiple additional criteria when filtering identified candidate genes as compared with those in the Anderton et al. study. These authors selected for further analysis those genes that demonstrated in at least 2 cell lines a ≥ 2-fold upregulation in expression following 5-aza treatment, contained a predicted CpG island in their promoter region, and were identified as targets of loss of heterozygosity based on SNP genotyping studies. Under these criteria, SPINT2, a negative regulator of the HGF/Met signaling pathway, was identified, exhibiting robust reexpression in 6 of the 9 medulloblastoma cell lines profiled. The authors confirmed the downregulation of SPINT2 in a significant percentage of primary medulloblastomas (> 2-fold in 41 of 56 samples) analyzed by quantitative RT-PCR, and, importantly, that aberrant promoter methylation (assessed by methylation-specific PCR) correlated with the observed reduction in gene expression in most cases. Stable reexpression of SPINT2 in medulloblastoma cell lines resulted in the attenuation of the malignant phenotype, inhibiting cell proliferation, anchorage-independent growth in soft agar, and cell motility. Furthermore, the orthotopic transplantation of D283 cells stably reexpressing SPINT2 into recipient nude mice significantly delayed the time to death compared with empty vector control cells in an intracerebellar xenograft model. These data strongly implicate SPINT2 as a putative tumor suppressor gene in medulloblastoma and shed further light on the apparent role of aberrant HGF/Met signaling in medulloblastoma etiology.

Collectively, these recent studies of the medulloblastoma epigenome have proven informative and have further implicated epigenetic gene silencing as an important mechanism of tumor suppressor gene inactivation in medulloblastoma. The future application of strategies that enrich for epigenetic modifications (that is, methylation-dependent immunoprecipitation [MedIP] or chromatin immunoprecipitation [ChIP]) combined with high-resolution microarrays or NGS technologies will probably lead to an improved appreciation of the role that epigenetics plays in medulloblastoma.

Next-Generation Genomics of Medulloblastoma

Over the past few years, microarray technologies have significantly increased our understanding of the medulloblastoma genome, transcriptome, and, to some extent, epigenome. Moving forward, array platforms will undoubtedly continue to be used in genome-wide profiling of medulloblastoma, especially as the resolution and coverage of these methods continue to improve and the cost of these screens remains affordable. However, recent breakthroughs in DNA sequencing technologies have taken the genomics community by storm, and their application in medulloblastoma research is, without question, imminent.

Since the early 1990s, the capillary-based Sanger method of DNA sequencing has been the mainstay for most applications in molecular biology, even the first drafts of the human genome published in 2000. More recently, conventional sequencing has been successfully used in large-scale resequencing efforts, profiling anywhere from a few hundred genes to all known protein-coding genes in a single cancer genome. Indeed, initial exon resequencing of the colorectal, breast, pancreatic, and glioblastoma multiforme genomes has revealed new genes and pathways involved in the pathogenesis of the respective cancer types. However, these studies relied on the PCR-mediated amplification of literally hundreds of thousands of exons combined with an enormous workload of conventional sequencing, unrealistic tasks for most of the cancer genomics community.

Fortunately, over the last few years a revolution in DNA sequencing technology has occurred and is rapidly changing the field of cancer genomics. Next-generation (also known as “next-gen” or “deep”) sequencing biochemistries now permit the parallel acquisition of up to 10s of gigabases (Gb) of DNA sequence of variable “read” length in a single experiment. Multiple NGS options are currently available—454 genome sequencer...
(Roche), gene analyzer (Solexa/Illumina), SOLiD system (Applied Biosystems)5,146—each with its own strengths and weaknesses. The repertoire and capabilities of these platforms are continually improving. For example, the SOLiD 3 System currently boasts > 20 Gb of DNA sequence per run, compared with ~ 750–1000 bp generated using traditional Sanger sequencing. Although the cost/base ratio is significantly lower for next-generation technologies, the current cost of a single SOLiD 3 run is in the neighborhood of $15,000 (vs ~ $5 for conventional sequencing), making next-generation tools prohibitively expensive for a large percentage of the research community, especially when considering the sequencing of large numbers of patient samples.

The biochemistry and real-time imaging-based data acquisition involved in NGS are what allow these technologies the parallel, high-throughput capacity that is not feasible with conventional Sanger sequencing; detailed reviews on the technical principles of conventional sequencing and NGS have been reported.5,106,111,146 In the Sanger method, DNA to be sequenced is either randomly fragmented and cloned into a high-copy number plasmid prior to bacterial transformation (shotgun sequencing) or PCR-amplified using target-specific primers (gene-specific sequencing). The amplified template is then subjected to a series of sequencing cycles whereby the template is denatured, primers are annealed, and a new complementary strand is synthesized in the presence of fluorescein-labeled dideoxynucleotide triphosphates (ddNTPs), 1 unique color for each of the 4. End-labeled DNA fragments are subsequently separated using high-resolution capillary electrophoresis followed by laser excitation of labeled fragment ends, with the emission spectra producing a 4-color chromatogram that can be translated into DNA sequence. Modern Sanger sequencing units can simultaneously run samples in 96- or 384-well format, generating reads of up to 1000 bp in length per sample and providing some degree of throughput.

Despite differences in template amplification and sequencing biochemistry associated with current next-generation platforms, the principles and workflow involved in these technologies are relatively similar. Initially, a template library is prepared through random DNA fragmentation and adapter ligation. Ligated fragments are then bound to micron-sized beads (Roche and Applied Biosystems) or a planar substrate (Solexa/Illumina) and PCR amplified (that is, emulsion PCR or bridge PCR) as clusters or colonies consisting of thousands of clonal “features” to be sequenced. An array or flow-cell can consist of literally millions of clustered features, enabling massively parallel downstream sequencing. So-called sequencing-by-synthesis is then performed with either a polymerase (Roche and Solexa/Illumina) or ligase (Applied Biosystems) that serially extends primed templates. Light (pyrosequencing with Roche) or fluorescence (Solexa/Illumina and Applied Biosystems) emitted following the incorporation of dNTPs (deoxynucleoside triphosphates) or oligonucleotides is then captured by imaging the full array of synthesized features at the end of each cycle. The final result of a full next-generation run currently ranges from up to ~ 1,000,000,000 reads of ~ 400 bp with Roche, ~ 150,000,000 reads of 35 bp with Solexa/Illumina, and ~ 300,000,000 reads of 50 bp with the Applied Biosystems. These figures translate into 10s of Gb of DNA sequence in a single run (1 human genome comprises ~ 3 Gb), allowing for high-coverage, whole-genome sequencing on 1 machine within a few weeks—this in sharp contrast to the sequencing “factories” and several years required for the initial Sanger-based sequencing of the human genome not more than a decade ago.

Multiple “proof-of-principle” studies using next-generation technology have now been performed to analyze the various aspects of both normal and cancer genomes. Published reports have described array-based, targeted capture, and NGS of ~ 200,000 protein-coding exons in the human genome, allowing specific identification of both common and rare sequence variants.95,116 Unbiased, whole-genome NGS has also been reported for multiple normal human genomes15,80,107,171 and, recently, the acute myeloid leukemia genome from a single individual.88 Producing nearly 100 billion bases of sequence, authors of the acute myeloid leukemia study described 10 genes with acquired somatic mutations that were not present in the patient-matched genome from normal skin cells.

Perhaps equally impressive, NGS has not been limited to studies aimed at the identification of sequence variants and mutations.5,114 Structural aberrations, including inter- and intrachromosomal rearrangements (inversions, inverted/tandem duplications, and translocations) as well as copy number aberrations (amplifications and deletions), have been identified in human cancer cells using NGS, with improved specificity and sensitivity compared with array-based methods.25,27 Whole-transcriptome profiling (also known as RNA-Seq) has also been described using next-generation approaches, permitting the quantification of transcript abundance (mRNA, miRNA, and so forth) and the identification of novel genes and isoforms in an unbiased manner.110,113,114,122,169,172 In contrast to the array-based technologies used in gene expression analyses, RNA-Seq requires no a priori knowledge of the transcriptome under investigation—thus enabling full-transcriptome characterization—and eliminates biases associated with array content. Similarly, unbiased NGS has been extended to studies of the mammalian epigenome, including genome-wide analyses of DNA methylomes,37,102 mapping of histone modifications,118 and detailing of the locations of DNA-binding proteins.74,119

Collectively, these emerging NGS-based approaches for studying the cancer genome hold great promise for comprehensive analyses of medulloblastoma. As the cost and bioinformatics involved in next-generation become more mainstream, next-generation–based profiling of the medulloblastoma genome, transcriptome, and epigenome will surely be the priority of several investigative groups (Fig. 6). Undoubtedly, these efforts will lead to a more complete understanding of the genes and pathways involved in the initiation, maintenance, and progression of medulloblastoma. Moreover, as larger patient cohorts are gathered and profiled using these advanced methods, more specific and reliable molecular classification of medulloblastoma will probably be possible. Finally, the correlation of genomic data with patient clinical data,
Genomics of medulloblastoma

such as the presence of metastatic disease and overall survival, will undoubtedly be improved.

Conclusions

Twenty years of studying the medulloblastoma genome has facilitated a detailed description of the medulloblastoma karyotype, led to the identification and validation of bona fide oncogenes and tumor suppressors, and implicated key signaling pathways and networks that are recurrently deregulated. Much of the recent progress in this field is owed to improvements in the technologies available for analyzing the genome. To comprehensively appreciate a cancer genome such as medulloblastoma, unbiased, high-resolution, genome-wide investigations must be undertaken, ideally using a combination of complementary (microarrays and NGS) and integrative (genome, transcriptome, and epigenome) technologies. Of course, an adequate sample size is critical to such studies if a full range of both common and rare genomic changes is to be captured. Large, coordinated multiinstitutional consortia such as The Cancer Genome Atlas (TCGA) are now applying this philosophy to study the genomes of brain (glioblastoma multiforme), lung (squamous carcinoma), and ovarian (serous cystadenocarcinoma) cancers.26 As part of their pilot project, TCGA is profiling large numbers of tumors of these origins to assess aberrations in DNA sequence (substitutions and indels), genomic copy number (amplifications and deletions), chromosomal composition (rearrangements), DNA methylation (hyper- and hypomethylation), and gene expression (aberrant expression and splicing). Results from TCGA and other similar large-scale collaborative efforts utilizing this type of broad approach to cancer genomics have recently demonstrated successes.26,36,78,125 To make progress in the fine genomic mapping of medulloblastoma, analogous strategies are warranted.

The generation of large genomic data sets of medulloblastoma using multiple orthogonal technologies will have its challenges. A logistical issue facing potential large-scale genomic projects on medulloblastoma relates to the immensity of the bioinformatics involved, as profiling large sample cohorts inevitably produces large amounts of data of varying complexity. The integration of multiple genomic data sets for maximal extraction of biological information will be a major priority if these new technologies are to be used to their full potential. In addition, the use of NGS technology for cancer genomics studies is still in its infancy, and the cost as well as the bioinformatics and computing challenges related to this promising technology remain a significant hurdle. Before NGS can be efficiently applied to large-scale medulloblastoma projects, financial feasibility and informatics considerations must be addressed.

Discriminating between so-called driver and passenger mutations is also a common yet critical dilemma in cancer genomics studies.124,149 In other words, not all somatic alterations in a cancer genome actually contribute to cancer development and confer a clonal advantage to a tumor cell. Distinguishing genes that provide a clonal advantage (drivers) in tumorigenesis from those that do not (passengers) requires the integration of multiple data types (copy number, sequencing, and expression) and, ideally, functional validation—neither of which is a simple task when considering possibly hundreds of candidate cancer genes. Thus, the use of genomic studies alone will be insufficient for determining the genes “driving” medulloblastomas, and follow-up functional studies, including mouse models, will be a necessity.

A third issue to consider is the reality of intratumoral heterogeneity and how global profiling strategies are typically performed using nucleic acid extracted from bulk tumor tissue. Methods such as array-based copy number and gene expression profiling essentially produce an average for a genomic region or gene in a given template. Since medulloblastomas are neither histologically nor molecularly monoclonal, genomic events that are present in only a small percentage of cells will be under-called or missed outright when bulk tumor is the source of the template queried. Indeed, the amplification of known medulloblastoma oncogenes, such as MYC and MYCN, is often found in only 10–20% of cells as determined by FISH (S. Pfister, personal communication, 2009), suggesting that these events are under-called by array profiling. Strategies that involve the analysis of distinct cell populations (that is, laser capture microdissection and cell sorting) or even single cells within a tumor will be required to avoid losing potentially valuable genetic information present only in tumor cell subpopulations.

Despite these relevant technical and logistical considerations, the next few years promise to be an exciting period for the community studying medulloblastoma. Large-scale, collaborative genomics projects will provide a more detailed characterization of this genome than ever before, and, optimistically, many new candidates will be uncovered. These efforts should continue to improve our ability to diagnose, stratify, and treat medulloblastoma, eventually leading to decreased deaths and improved quality of life for patients.

Disclosure

The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper.

Acknowledgments

The authors sincerely thank Paul Paroutis and Christian Smith for assistance with artwork and figure preparation.

References


P. A. Northcott, J. T. Rutka, and M. D. Taylor
Genomics of medulloblastoma


P. A. Northcott, J. T. Rutka, and M. D. Taylor
Genomics of medulloblastoma


P. A. Northcott, J. T. Rutka, and M. D. Taylor


Manuscript submitted September 18, 2009.
Accepted October 22, 2009.
Address correspondence to: Michael D. Taylor, M.D., Ph.D., Division of Neurosurgery, The Hospital for Sick Children, 555 University Avenue, Suite 1504, Toronto, Ontario, M5G 1X8, Canada. email: mdtaylor@sickkids.ca.