Monoallelic Expression Determines Oncogenic Progression and Outcome in Benign and Malignant Brain Tumors

Erin J Walker¹², Cindy Zhang¹², Pedro Castelo-Branco¹², Cynthia Hawkins²³, Wes Wilson¹², Nataliya Zhukova¹², Noa Alon¹, Ana Novokmet¹, Berivan Baskin⁴, Peter Ray⁴, Christiane Knobbe⁵, Peter Dirks², Michael D Taylor², Sidney Croul²⁵, David Malkin¹, Uri Tabori¹²

¹Division of Haematology/Oncology, ²The Labatt Brain Tumor Research Centre, ³Department of Pathology, ⁴Department of Pediatric and Laboratory Medicine, The Hospital for Sick Children, Toronto, Canada; ⁵Department of Neuropathology, Toronto General Hospital, Toronto, Canada.

Running Title: Monoallelic expression in brain tumors

Keywords: Monoallelic gene expression, genome-wide, pediatric, adult, benign, malignant, brain tumors.

Financial Support: This work was supported by funds received from The Hospital for Sick Children Comprehensive Cancer Centre to UT, DM and EW, and in part by funds received from a grant of the Canadian Institutes of Health Research (#MOP-97834) to DM.
Abstract:

Although monoallelic expression (MAE) is a frequent genomic event in normal tissues, its role in tumorigenesis remains unclear. Here we performed SNP arrays on DNA and RNA from a large cohort of pediatric and adult brain tumor tissues to determine the genome-wide rate of MAE, its role in specific cancer-related genes, and the clinical consequences of MAE in brain tumors. We also used targeted genotyping to examine the role of tumor-related genes in brain tumor development, and specifically examined the clinical consequences of MAE at TP53 and IDH1. The genome-wide rate of tumor MAE was higher than in previously described normal tissue and increased with specific tumor grade. Oncogenes, but not tumor suppressors, exhibited significantly higher MAE in high-grade compared to low-grade tumors. This method identified 9 novel genes highly associated with MAE. Within cancer-related genes, MAE was gene specific; hTERT was most significantly affected, with a higher frequency of MAE in adult and advanced tumors. Clinically, MAE at TP53 exists only in mutated tumors and increases with tumor aggressiveness. MAE towards the normal allele at IDH1 conferred worse survival even in IDH1-mutated tumors. Taken together, our findings suggest that MAE is tumor and gene specific, frequent in brain tumor subtypes, and may be associated with tumor progression/aggressiveness. Further exploration of MAE at relevant genes may contribute to better understanding of tumor development and determine survival in brain tumor patients.
Introduction:

Loss of heterozygosity (LOH) of cancer-associated genes is a common and important mechanism in carcinogenesis. However LOH at the level of the RNA, or monoallelic gene expression (MAE), is a much less understood phenomenon. MAE is defined as expression of a gene that is restricted to one allele in the presence of a diploid, heterozygous genome, and is observed for imprinted genes and random X inactivation (lyonization). MAE has also been observed in the odorant receptor, immune receptor and immunoglobulin classes of genes (1, 2). For most other genes, expression is typically expected to occur from both parental alleles without preference. Recent studies, however, have uncovered a potential role for random MAE, as well as differential allele-specific expression (DAE) (i.e. a skew in allele-specific expression rather than complete monoallelic expression) in normal tissues as well as in some diseases, including cancer (3-9).

Several studies in normal tissues or derived cell lines have shown random MAE/DAE to be relatively frequent (5-20%), depending on the number of genes included in the study and the gene selection bias (3, 8, 10). These data together reveal the previously unappreciated random nature of MAE/DAE amongst non-imprinted genes, and provide a mechanism that may help to explain phenotypic differences between individuals, as well as their susceptibility to particular diseases and the disease course.

Cancer is a complex disease, the result of interplay between genetic, epigenetic alterations and environmental cofactors. We chose brain tumors as a platform to examine MAE for several reasons. Brain tumors are a major cause of morbidity and mortality in pediatric and adult oncology (11). The genetic basis of brain tumors has
been studied in depth over the past decade, leading to the identification of mutations/genetic alterations central to brain tumor development (12-15). However, while the role of alterations in the \textit{TP53}, \textit{RB} and \textit{RAS/MAPK} pathways in the development of specific brain tumors has been well characterized, research considering the widespread impact of MAE in these cancers is in its infancy.

Expression of tumor suppressors and oncogenes is tightly regulated during an individual’s life, and it has been established that aberrant gene expression, such as that observed from loss of imprinting, can predispose individuals to cancer development (16). It is therefore reasonable to propose that aberrant gene expression caused by lack of one allele, MAE, could also lead to the development of cancer; indeed preliminary evidence exists to support this notion (7-9).

Previous studies have used Affymetrix and Illumina genome-wide SNP arrays to identify MAE in normal and tumor tissues (8, 9), however, the majority of studies investigating MAE/DAE have interrogated either cell lines or germline material from patients, and few have examined primary tumor tissues. Thus the role of MAE/DAE in cancer, and more specifically in brain tumor development, is as yet undetermined. Furthermore, the mechanisms and functional consequences of MAE in primary tumors have not been studied. In order to examine the role of MAE in brain tumors, we chose specific tumors (pediatric gliomas and choroid plexus tumors) that have distinct clinical and biological subgroups, benign and malignant, and do not routinely transform from the low- to high–grade state. In addition, we analyzed adult gliomas, since these cancers tend to transform from low to high-grade and therefore represent a clinically mixed group. We also included multiple phases of tumor development, such as autopsy samples, to examine changes in MAE over time. We achieved this by
genotyping DNA and RNA (cDNA) samples, and identifying SNPs in which DNA genotypes were heterozygous and corresponding RNA genotypes were homozygous. We used Affymetrix SNP 6.0 arrays and high-throughput genotyping techniques to interrogate the extent of MAE and to uncover novel genes displaying MAE. We then examined the role of MAE in specific brain cancer related genes. We find MAE to be a frequent event in brain tumors; further, it is gene and tumor specific and may precede and predict tumor behavior.

Methods:

Patients and tissues:

This study was approved by the Research Ethics Board at The Hospital for Sick Children. Tumor and blood samples from pediatric brain tumor patients were collected from the Pathology Department or through the Arthur and Sonia Labatt Brain Tumor Research Centre at The Hospital for Sick Children. Adult brain tumor tissues were collected from the Canadian Virtual Brain Tumor Bank. The number of tissues analyzed in each group is shown in Table 1. In order to define the differences between low and high grade brain tumors, we analyzed two tumor subtypes which are known to have these subgroups: gliomas and choroid plexus tumors. We also compared low and high grade gliomas between children and adults, since these exhibit different clinical behaviors. In total, we analyzed 102 samples from 51 tumors by SNP array and 236 brain tumor, blood and cell line samples (Table 1) for presence of MAE, by genotyping 34 SNPs across 21 candidate genes by Sequenom genotyping (Supplementary Table 1). We also used direct sequencing on a cohort of patients
(Table 1) to assess the TP53 (all exons) and IDH1 (exon 4) genes for mutations (Supplementary Table 2).

**Nucleic acid extraction methods.**

DNA and RNA were extracted from fresh frozen tumor tissues using standard phenol/chloroform extraction for DNA and Trizol (Invitrogen) for RNA (see supplementary methods for details). The concentration of each nucleic acid was obtained using a NanoDrop spectrophotometer (Thermo Fisher Scientific, DE, USA). In addition, DNA samples were analyzed using a 1% TBE-agarose gel to assess DNA integrity, and RNA samples were analyzed using the Agilent 2100 bioanalyzer (Agilent Technologies, CA, USA). Nuclear and cytoplasmic RNA was extracted from cell lines using a commercial kit (Norgen Biotek Corporation, Canada), following the manufacturers instructions.

**Microarray analysis.**

All nucleic acids were assessed using Affymetrix Human SNP 6.0 arrays (Affymetrix, CA, USA) and all microarrays were processed by The Centre for Applied Genomics (TCAG) at The Hospital for Sick Children, Toronto. DNA was processed according to standard Affymetrix protocols. RNA samples were processed according to the GeneChip Whole Transcript (WT) Double Stranded Target Assay protocol, with the following exceptions: 1) the starting amount of RNA was 2µg; and 2) the cRNA sample generated by the first synthesis cycle was divided into two reactions rather than three, to ensure the volume of double stranded cDNA would be sufficient for hybridization.

**Cell culture, media and drug treatment.**
Cell lines were maintained as follows: Med8a and D283 cells in DMEM (Wisent, St-Bruno, Quebec, CANADA), UW228, Daoy and fibroblasts 158FB and 1604FB in AMEM (Wisent), D425 and D458 in IMEM with 20% FBS and 1.1g bicarbonate (Wisent), ONS76 and RES256 (2% FBS) in DMEM/F12 (Wisent) and U87 and U251 in DMEM media (Wisent) with 1% L-glutamine, 1% sodium pyruvate and 1% non-essential amino acids. All media were supplemented with 10% FBS (Invitrogen) unless indicated. Fetal neural stem HF240 cells, and glioma TIC lines G179 and GLINS1 were maintained as previously described (17).

**Data Analysis.**

Affymetrix Genotyping Console 4.0 was used to genotype the DNA and cDNA samples, using default parameters. Genotypes for UTR and exonic SNPs were then exported to allow comparison between DNA and cDNA pairs from the same samples. The genotypes for paired gDNA and cDNA samples at each SNP were then compared using a php script, which scored the SNP, depending on whether the gDNA/cDNA genotypes were identical, the cDNA was homozygous when the gDNA was heterozygous, or if there was no result for either gDNA or cDNA. This data was used to determine genome wide rates of homozygous and MAE genotypes. In addition, gene-based analysis of the SNP data was also performed for individual samples and groups of samples; genes were only included in this analysis if they were genotyped by at least two SNPs. We also compared our overall MAE rates with previously published data regarding genome-wide predictions of MAE in normal tissues (3). To determine if the MAE we observed in the genome-wide data was potentially due to allele-specific copy number changes, we examined whether any or our shortlisted genes were located within regions of copy-number change. We found this not to be the case for these samples and genes.
The MassARRAY iPlex system (Sequenom) was used to genotype DNA and cDNA samples. The Sequenom software was used to determine gDNA and cDNA genotypes, and a trained technician then confirmed individual sample genotypes to ensure genotype accuracy. Samples for which the allele ratio was greater than 0.7:0.3 were considered to show differential allelic expression. The Sequenom system is also able to show changes in gDNA copy number, however we observed no such changes in the SNPs analyzed using this system.

All statistical analyses were performed using SPSS v15.

Results:

Genomic rate of MAE in brain tumors.

In order to assess the magnitude of MAE in brain tumors in an unbiased manner and to uncover genes that utilize MAE and contribute to tumorigenesis, we performed a genome wide analysis of MAE using Affymetrix SNP 6.0 arrays in a series of 51 low and high grade brain tumors (Table 1). For this analysis, we hybridized paired DNA and cDNA from primary tumors to Affymetrix SNP 6.0 chips, to perform genome wide genotyping of the DNA and cDNA. In order to detect MAE, we only considered SNPs for which the DNA genotype was heterozygous and we removed all homozygous DNA genotypes prior to any analysis. Thus, any cDNA genotype that was subsequently homozygous was considered to display MAE.

To confirm the validity of the Affymetrix SNP 6.0 array in detecting MAE in a genome-wide manner, we examined heterozygous DNA genotypes from the X chromosome of female patients. As expected from the lyonization principle, we observed complete LOH in the cDNA samples, indicating X inactivation (Supplementary Figure 1). In order to further validate the utility of this platform to
diagnose MAE, we analyzed the frequency of MAE in a subset of imprinted and olfactory receptor genes, which are known to display MAE (1, 2). We observed a median of 82% of informative olfactory receptor genotypes and 50% of informative imprinted gene genotypes were MAE. We confirmed by sequencing the genotypes for SNPs within the \textit{CPA4} gene, which is known to be imprinted, with 100% accuracy.

In order to determine the role of MAE in low and high grade tumors, we investigated the genome wide rates of DNA homozygosity and MAE within each brain tumor type. As expected, the percentage of SNPs with homozygous DNA genotypes was very high (>80%) in most tumors and higher (albeit not significantly) in high grade tumors vs. low grade tumors (Table 2). Importantly, for the remaining informative (heterozygote) SNPs, MAE was significantly greater in the pediatric high grade gliomas (pHGG) versus low grade gliomas (pLGG) (\(p=0.012\)) and the choroid plexus carcinomas (CPC) versus choroid plexus papillomas (CPP) (\(p=0.0011\)) (Table 2).

Genome wide rates of SNP MAE varied between 28-45% of informative genotypes. Additionally, the gene-based analysis demonstrated the number of genes involved in MAE was higher in the pediatric high grade tumors, though this was only significant in the choroid plexus tumors (CPC vs. CPP \(p=0.006\), pHGG vs. pLGG \(p=0.063\), Figure 1a). Taken together, homozygous expression of alleles (either through DNA LOH or MAE) was greater in pediatric high grade tumors compared to low-grade tumors, suggesting an important role for MAE in identifying genes involved in the clinical and biological differences in these tumor subtypes.

In contrast, the frequency of MAE genes was not significantly different between the adult LGG and HGG, which in addition have a very broad range of MAE rates.
(Supplementary Table 4). This is consistent with the nature of adult gliomas, which invariably progress from low to high grade, whereas pediatric tumors rarely do so.

As the global, genome-wide MAE rate was greater than the MAE rate observed in the gene-based analysis, we hypothesized that MAE is gene specific. To further investigate this hypothesis, we compared MAE rates in particular gene subgroups (tumor suppressors and oncogenes), and also utilized a gene-specific approach to analyze known brain-tumor associated genes (see below).

In order to determine whether MAE plays a differential role in tumor suppressors vs. oncogenes, we investigated the relative frequency of MAE in known tumor suppressors (TS) and oncogenes, as described on the CancerGenes website and via literature searches (Supplementary Table 3). While the rate of MAE in oncogenes and TS was similar within the low and high grade tumor cohorts (Figure 1b), oncogenes had significantly higher MAE in high grade tumors than in low grade tumors (31% MAE vs. 18% MAE respectively; p=0.010). There was no significant difference in the rate of MAE in tumor suppressors between low and high grade tumors (p=0.122). This suggests that high grade tumors utilize allele specific expression of oncogenes during transformation, potentially augmenting mutant or altered alleles.

**MAE to detect novel targets.**

In order to identify novel tumor-associated genes, we looked specifically at genes with high rates of MAE in each of the tumor types. Using the genome-wide data generated from the Affymetrix SNP 6.0 arrays, we initially restricted our gene list by including only those genes which were informative in >75% of tumors within at least one sample group; this yielded a short list of 56 genes. We then identified genes that displayed MAE in 100% of these samples, to identify the genes most affected by
MAE. We identified 9 genes that displayed MAE in all informative samples (Table 3). Strikingly, 7 of these 9 genes were reported in the literature to be associated with cancer. While the proportion of informative cases for these genes in each tumor group varied, all informative samples showed complete MAE at these loci.

Specific analysis of candidate genes reveals high MAE in *hTERT*. Having identified genome-wide differences between different tumor types, we investigated the extent of MAE in known brain tumor-associated genes. To achieve this, we used a semi-quantitative genotyping method (Sequenom) to analyze a large cohort of tumor and normal samples for multiple coding SNPs within genes of interest (Supplementary Table 1).

We observed a global increase in MAE/DAE with increasing tumor phase, from normal brain (~7%), through tumors at time of biopsy/resection (~11%), tumors sampled at autopsy (~15%) to cell lines (~27%) (Figure 2a). Interestingly, MAE was observed in only a minority of genes (Figure 2b) and only 5 genes had more than 10% MAE/DAE; *TERT*, the gene with the highest rate of MAE was recently identified in genome-wide association studies as a risk factor for development of gliomas (18, 19). All other genes, except *MSH3* (p=0.2) and *cMYC* (p=0.06), had significantly lower MAE/DAE rates compared to *TERT* (Figure 2b).

Higher rate of *hTERT* MAE in adult and advanced brain tumors. The highest rate of MAE/DAE at 35% was observed in *hTERT*. MAE/DAE at *hTERT* varied according to tumor type and phase (Figure 2c). Specifically, a striking difference was observed between adult and pediatric tumors; over 50% of adult gliomas showed aberrant *hTERT* expression compared to ~30% of pediatric gliomas.
(p=0.0224, Fisher’s Exact test). Similarly, the cell line/autopsy tumor samples which represent late phases of tumor evolution had even higher MAE/DAE where >80% of informative cell lines had aberrant hTERT expression. The presence of MAE in the tumors and cell lines was confirmed by direct sequencing. Interestingly, we observed a tendency towards one allele in samples displaying MAE (67% of MAE samples expressed the G allele rather than the A allele for rs2736098), however there was no difference in allele expression according to tumor grade.

In an attempt to elucidate the mechanisms underlying the observed allele-specific expression at the hTERT locus, we first sought to localize the changes in expression, to either pre- or post-transcription. We examined nuclear and cytoplasmic RNA extracts from informative cell lines, and found no differences in allelic expression between the two cellular compartments, suggesting the allele-specific alterations in hTERT expression are due to pre-transcriptional mechanisms.

**LOH and MAE of TP53 correlate with advanced tumor stage.**

Over 50% of cancers and a significant number of advanced gliomas (20) and CPC (21) harbor either a heterozygous or monozygous somatic TP53 mutation with LOH. We examined the rate of TP53 MAE using the common codon 72 polymorphism (rs1042522) and TP53 mutation status in peripheral blood lymphocytes and brain tumors from different phases (Figure 3a). We found no MAE in blood samples, (including those from Li-Fraumeni syndrome (LFS) patients, who harbor germline heterozygous TP53 coding mutations). Furthermore, pediatric low-grade gliomas that do not harbor TP53 mutations did not reveal MAE. In contrast, adult LGGs had an MAE rate of ~20%, in addition to the 43% of samples with homozygous TP53 mutations, resulting in a total of 63% LOH at expressed alleles. For high grade
tumors, which showed the same trend in pediatric and adult samples, ~ 20% of heterozygous samples demonstrated MAE, in addition to the 58% of tumors with homozygous TP53 mutations, for a total of 78% LOH of expressed alleles. Furthermore, we observed a high rate of homozygous TP53 mutations in both autopsy tumors (83%) and established long term cell lines (100%). Finally, in order to determine if MAE in a mutant TP53 tumor could be induced over time, we examined the effect of implanting a TP53 double heterozygote (c. C455T and G797A) patient-derived glioblastoma cell line into mouse cortex. The xenograft tumor, which previously expressed both alleles, displayed a shift in the allele ratios revealing DAE towards the mutant G797A allele and the normal C455T allele (Figure 3b). Together, these findings demonstrate association between increasing frequency of homozygous TP53 mutation at the expression level with increasing glioma grade and disease stage. The phenomenon of MAE at TP53 mutation loci was tumor specific since paired blood and tumor DNA/RNA from LFS patients who had MAE in their tumor did not reveal MAE in their blood samples. Further, adult low grade tumors with MAE predominantly expressed only the normal allele (66% vs. 37% with MAE to the mutant allele), while the high grade tumors predominantly expressed the mutated allele (66% vs. 37% with MAE to the normal allele)(data not shown).

In order to identify the mechanism for MAE at the TP53 locus, we considered two explanations: 1) the presence of promoter polymorphisms that affect transcription factor binding sites, thus leading to allele-specific expression; and 2) polymorphisms within the 3’UTR that might affect transcript degradation by altering microRNA attachment, leading to reduced degradation of one allele over the other. We sequenced the reported 2024 bp of the promoter region (Figure 3c) in 6 samples with and without MAE. While we identified a number of known polymorphisms within
some of the samples, we did not identify promoter polymorphisms in any MAE sample. We also sequenced the 3’UTR of known microRNA attachment sites but did not find any mutation or allelic change.

**MAE in IDH1 predicts survival in adult gliomas**

We then investigated the role of MAE at another frequently mutated gene in adult gliomas, the *IDH1* (exon 4). We sequenced exon 4 of *IDH1* in an unbiased cohort of 31 adult low grade tumors and 46 adult high grade tumors. Consistent with previous data (22), we found a mutation rate of 71% (n=22) in the adult low grade tumors and 23% (n=11) in the adult high grade tumors. Of the mutated tumors, 5 (15%) demonstrated MAE. We then investigated the role of MAE at the *IDH1* locus in a second cohort of *IDH1* mutated adult gliomas (42 LGG and 25 HGG), with clinical data available for analysis. Within the clinical cohort, we identified *IDH1* MAE in 12/67 tumors (17.9%). In 10/12 cases, MAE was specific for the normal allele, which is associated with reduced survival. Clinical correlation within the low grade glioma cohort revealed significantly higher patient survival (p=0.04) for mutated tumors that did not have MAE, compared to patients whose tumors had MAE towards the normal allele (Figure 4). This data agrees biologically with previous data, which has demonstrated a survival advantage in patients with mutations, compared to those without *IDH1* mutations (23, 24).

*Discussion:*

Monoallelic expression of genes by allele-specific silencing has been described and is a common mechanism in imprinted genes, immune receptors and olfactory receptors.
The existence of genome wide random MAE in normal tissues has recently been described (3, 8, 10). However, the extent of MAE in brain tumors, and its biological and clinical significance are still unknown. This study demonstrates that MAE is relatively common in brain tumors. It is a gene and tumor specific event rather than a general mechanism and may have a cardinal role in tumor progression, as well as clinical implications.

The extent of MAE in normal tissues ranges from 5-20% (3, 8, 10), with approximately 10% of genes demonstrating MAE when analyzed using Affymetrix SNP arrays. This is in agreement with the levels of MAE we observed for most genes in the candidate gene study, and the levels of MAE found in blood samples. We observed a continuous trend of low to high MAE from normal tissue, low grade and high grade tumors originating from the same tissue, and late established tumor cells such as autopsy biopsies and cell lines. This finding was consistent on a single gene level (TP53 and hTERT, Figures 2c & 3a) as well as on a genome wide level, where the average level of MAE within the brain tumor samples was ~20% (Figure 1a) compared to ~10% in normal cell lines (3). Furthermore, since DNA LOH increases with tumor progression, the addition of MAE will be associated with a dramatic increase on the rate of allelic expression throughout the process of tumorigenesis.

Vogelstein described the process of progression from premalignant lesions to high grade tumors in colon cancer through stochastic acquisition of mutations in multiple cancer related genes (25). Our observations suggest that MAE may precede or enhance a mutation by expression of only the mutant or disease related allele and by that may add a new dimension to Vogelstein’s prediction.
Our investigation of MAE within specific classes of genes revealed a significant
difference in the MAE frequency between low and high grade tumors, with fewer
oncogenes demonstrating MAE in low grade tumors. These data suggest that MAE in
the context of brain tumors is predominantly advantageous to the tumor, if one
considers the biological situation that predicts that low grade tumors will have fewer
oncogenes disrupted than their high grade counterparts.

Recent studies have highlighted the potential role for microRNA in cancer
development (26, 27), and TP53 has a number of microRNA binding sites within the
3’UTR. Although we sequenced the entire promoter, 5’UTR and 3’UTR regions, we
were able to discard polymorphisms in regulatory regions as the basis for MAE at the
TP53 locus. However, in investigating MAE in the TP53 gene, we show that MAE
can contribute to tumor progression within a single gene. We observed no MAE in
tumor samples that do not have TP53 mutations and that do not progress, but high
rates (~20%) of MAE in tumors that have TP53 mutations and are progressive. While
one could argue that the mutated transcript is quickly degraded, thus leaving only one
allelic form to be detected, we found MAE of both the normal and mutated alleles.
Indeed we observed that low grade tumors predominantly display MAE of the normal
allele while high grade tumors predominantly express only the mutated allele. This
data suggests a tantalizing notion that more malignant tumors are “using” MAE at the
RNA level to generate homozygously expressed TP53 mutations while still
possessing both alleles at the DNA level. This hypothesis agrees with the genome-
wide framework, where we observe more MAE within high grade tumors compared
to non-progressing low grade tumors (ie. pediatric low grade gliomas and CPPs).
The clinical implications of MAE are yet to be realized. Nevertheless, our preliminary findings suggest that MAE can be used to predict survival in tumors harboring mutations that are of prognostic relevance. Our analysis of IDH1 mutations revealed a clinically significant distinction between mutated-MAE tumors and mutated-non-MAE tumors. Specifically, mutated IDH1 gliomas with MAE towards the normal allele have worse survival than non-MAE tumors (Figure 4). This adds a new dimension to the recent studies which demonstrated that mutations at IDH1 are predictors of better survival in gliomas (24, 28) and can help tailor therapies for these patients.

In summary, we have demonstrated that in brain tumors, a high level of MAE is present and is tumor and gene specific. Furthermore, MAE may be a mechanism utilized by cancers to achieve a more aggressive phenotype. Finally, genome wide search may identify novel genes involved in these cancers. Additional studies are required to determine the mechanisms that govern MAE and allow for a full appreciation of the role of MAE in disease development.

**Acknowledgements:**

We would like to thank all of the patients and their families who were involved with the study, as well as the Canadian Virtual Brain Tumor Bank and its affiliate members for the adult tumor tissues. We also thank staff at TCAG for assistance and advice with the microarray processing, and staff at AGTC for assistance with processing and analysis of the Sequenom genotype data.

Grant Support: This work was supported by funds received from The Hospital for Sick Children Comprehensive Cancer Centre to UT, DM and EW, and in part by
funds received from a grant of the Canadian Institutes of Health Research (#MOP-97834) to DM.
References:

Table 1: Sample composition for the genome-wide analysis, candidate gene analysis and direct sequencing components.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Affymetrix SNP 6.0 arrays</th>
<th>Sequenom Genotyping (candidate gene analysis)</th>
<th>Direct Sequencing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>TP53</td>
</tr>
<tr>
<td>Pediatric low grade glioma (WHO I/II)</td>
<td>8</td>
<td>61</td>
<td>20</td>
</tr>
<tr>
<td>Pediatric high grade glioma (WHO III/IV)</td>
<td>6</td>
<td>27</td>
<td>11</td>
</tr>
<tr>
<td>Pediatric choroid plexus papilloma</td>
<td>11</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Pediatric choroid plexus carcinoma</td>
<td>8</td>
<td></td>
<td>17</td>
</tr>
<tr>
<td>Adult low grade glioma (WHO I/II)</td>
<td>9</td>
<td>30</td>
<td>14</td>
</tr>
<tr>
<td>Adult high grade glioma (WHO IV)</td>
<td>9</td>
<td>44</td>
<td>17</td>
</tr>
<tr>
<td>Blood samples</td>
<td></td>
<td>46</td>
<td>27</td>
</tr>
<tr>
<td>Control normal brain</td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Cell lines</td>
<td></td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>Autopsy brain tumors</td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>51</td>
<td>236</td>
<td>136</td>
</tr>
</tbody>
</table>
Table 2: Average percentage of SNPs with homozygous and MAE genotypes.
aHGG=adult high grade glioma, aLGG=adult low grade glioma, pHGG=pediatric high grade glioma, pLGG=pediatric low grade glioma, CPC=choroid plexus carcinoma, CPP=choroid plexus papilloma.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>% homozygous genotypes</th>
<th>% MAE genotypes*</th>
<th>p-value^</th>
</tr>
</thead>
<tbody>
<tr>
<td>aHGG</td>
<td>88.2</td>
<td>37.7</td>
<td>NS</td>
</tr>
<tr>
<td>aLGG</td>
<td>87.3</td>
<td>35.9</td>
<td></td>
</tr>
<tr>
<td>pHGG</td>
<td>89.3</td>
<td>39.5</td>
<td>0.012</td>
</tr>
<tr>
<td>pLGG</td>
<td>76.0</td>
<td>28.9</td>
<td></td>
</tr>
<tr>
<td>CPC</td>
<td>90.0</td>
<td>45.4</td>
<td>0.001</td>
</tr>
<tr>
<td>CPP</td>
<td>87.7</td>
<td>35.5</td>
<td></td>
</tr>
</tbody>
</table>

* The percentage of MAE SNPs, as a proportion of informative genotypes; ^ p-value derived from Mann-Whitney test of independent samples for MAE genotypes.
Table 3: List of genes with highest rate of MAE in each tumor subtype.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Association with cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADAMTS5</td>
<td>Over-expressed in glioblastomas [29]</td>
</tr>
<tr>
<td>ST13</td>
<td>Leads to apoptosis via ASK1-JNK pathway [30]</td>
</tr>
<tr>
<td>MUC16</td>
<td>Over-expressed in ovarian cancer [31]</td>
</tr>
<tr>
<td>ALDH1A2</td>
<td>Candidate tumor suppressor gene in prostate cancer [32]</td>
</tr>
<tr>
<td>SLC38A1</td>
<td>Over-expressed in hepatocellular carcinoma [33]</td>
</tr>
<tr>
<td>TIMELESS</td>
<td>Required for CHK2 activation and checkpoint control [34]</td>
</tr>
<tr>
<td>NTRK3</td>
<td>NTRK3 production promotes neuroblastoma cell survival [35]</td>
</tr>
<tr>
<td>IYD</td>
<td>Not yet described</td>
</tr>
<tr>
<td>TULP4</td>
<td>Not yet described</td>
</tr>
</tbody>
</table>
FIGURE LEGENDS:

Figure 1. A: Average gene MAE in tumor subtypes. The average percentage of genes that have > 2 SNPs with MAE out of genes with > 2 heterozygous SNPs is shown for each tumor subtype, +/- 1SD. aHGG=adult high grade glioma, aLGG=adult low grade glioma, pHGG=pediatric high grade glioma, pLGG=pediatric low grade glioma, CPC=choroid plexus carcinoma, CPP=choroid plexus papilloma. B: The percentage of known oncogenes and tumor suppressors demonstrating MAE. Box and whiskers plots, where the box represents the middle 50% of samples, the whiskers show the upper 35% of samples, and the solid lines in the middle of the boxes represent the median value. Data is shown by tumor grade.

Figure 2. Sequenom analysis of MAE/DAE in normal and malignant brain tissues. Semi-quantitative genotyping was performed on 236 tissues, using 34 SNPs. Since not all SNPs are informative for each tissue, the data is presented as a percentage of the number of informative genotypes within each group. A. Rate of MAE/DAE in normal brain, tumors and cell lines. B: The rate of MAE/DAE at each gene locus. * p<0.05, ** p<0.01, *** p<0.0001. C: The rate of MAE/DAE at the TERT locus. * p=0.0224. For all graphs, the black bars represent MAE, while the white bars represent DAE.

Figure 3. A: Progression of TP53 mutation and monoallelic expression in tumor tissues. This graph shows the percentage of samples tested within each group that have heterozygous TP53 mutations or heterozygous at codon 72 (solid white), homozygous TP53mutations (solid black) or MAE for TP53mutations or at the codon 72 polymorphism (pattern). B: Allele shift in TP53 mutations after implantation. GliNS cells were implanted into mice, and the subsequent tumors were analysed for mutation status. Representative cDNA sequence chromatograms of exon 5 and 8 heterozygous mutations. The panel on the left shows equal allele heights as observed in 3/4 cases, while the panel on the left shows representative results from one tumor, which demonstrated a shift in allele expression. This shift was observed in the forward and reverse direction, it was observed for both the exon 5 and exon 8 mutations, and was not observed in the corresponding gDNA samples. C: Representative picture of the regions of TP53 covered by sequencing. At the 5’ end of the gene, we sequenced 2024bp 5’ of exon 1, reported to include the entire promoter region, as well as exon 1 and 303bp of intron 1. At the 3’ end of the gene, we sequenced exon 11, which includes the 3’UTR, and 141bp into the on-coding region. We did not sequence 111bp within the 3’UTR, as this region contained a large poly-T region, but also contained no miRNA binding sites. miRNA binding sites predicted using TargetScanHuman (www.targetscan.org/) and microRNA.org (www.microRNA.org).

Figure 4. Survival curves for low grade glioma patients with IDH1 mutations. All patients had heterozygote (informative) mutations. Patients with IDH1 mutations and MAE (purple) to the normal allele had significantly shorter survival than patients with IDH1 mutations but no MAE (expressing both the mutant and normal alleles, green) (p=0.04).
Figure 1

A

B

% of Oncogenes with MAE

% of Tumor suppressors with MAE

0 10 20 30 40 50 60 70 80 90 100

0 10 20 30 40 50 60 70 80 90 100

high grade low-grade

high-grade low-grade

p=0.063

p=0.006

23

28

*^14

*
Figure 2

A

% of informative genotypes with allele shift

Normal brain (n=134)
Low grade (n=912)
High grade (n=648)
Autopsy (n=83)
Cell lines (n=133)

B

% of informative genotypes with allele shift

TERT (n=127)
MSH3 (n=46)
CDKN2A (n=49)
TP53 (n=184)
EGFR (n=170)
MLH3 (n=100)
cMYC (n=13)
MDM2 (n=28)
BRAF (n=73)
IDH1 (n=19)
MTH1 (n=111)
SPARC (n=24)
KIAA1549 (n=108)
FAT2 (n=31)
PIK3R1 (n=97)
PTCH1 (n=98)
NF1 (n=79)
RTEL (n=100)

C

% of informative samples with allele shift

All bloods (26/43)
Normal brain (7/13)
pLGG (34/55)
pHGG (15/26)
aLGG (17/30)
aHGG (27/44)
Cell lines + autopsy (16/23)
Figure 3

A

![Bar chart showing percentage of cases across different categories](image)

- Blood samples (n=27)
- pLGG (n=32)
- aLGG (n=14)
- HGG (n=45)
- PM tumors (n=6)
- Cell lines (n=12)

B

![Sequence analysis of Exons 5 and 8](image)

C

![Diagram of TP53 coding region and miRNA targets](image)
Figure 4

Overall Survival

MAE n=11

No MAE n=31

Months

0 60 120 180 240
## Monoallelic Expression Determines Oncogenic Progression and Outcome in Benign and Malignant Brain Tumors.

Erin J Walker, Cindy Zhang, Pedro Castelo-Branco, et al.

*Cancer Res* Published OnlineFirst December 5, 2011.

<table>
<thead>
<tr>
<th>Updated version</th>
<th>Access the most recent version of this article at: doi:10.1158/0008-5472.CAN-11-2266</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supplementary Material</td>
<td>Access the most recent supplemental material at: <a href="http://cancerres.aacrjournals.org/content/suppl/2012/01/23/0008-5472.CAN-11-2266.DC1">http://cancerres.aacrjournals.org/content/suppl/2012/01/23/0008-5472.CAN-11-2266.DC1</a></td>
</tr>
<tr>
<td>Author Manuscript</td>
<td>Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.</td>
</tr>
</tbody>
</table>

**E-mail alerts** Sign up to receive free email-alerts related to this article or journal.

**Reprints and Subscriptions** To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

**Permissions** To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.