

BTECH: A Platform to Integrate Genomic, Transcriptomic and Epigenomic Alterations in Brain Tumors

Min Wang · Hehuang Xie · Wendy Stellpflug · Veena Rajaram ·
Maria de Fatima Bonaldo · Stewart Goldman · Tadanori Tomita ·
Marcelo Bento Soares

Published online: 6 January 2011

© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract The identification of molecular signatures predictive of clinical behavior and outcome in brain tumors has been the focus of many studies in the recent years. Despite the wealth of data that are available in the public domain on alterations in the genome, epigenome and transcriptome of brain tumors, the underlying molecular mechanisms leading to tumor initiation and progression remain largely un-

known. Unfortunately, most of these data are scattered in multiple databases and supplementary materials of publications, thus making their retrieval, evaluation, comparison and visualization a rather arduous task. Here we report the development and implementation of an open access database (BTECH), a community resource for the deposition of a wide range of molecular data derived from brain tumor studies. This comprehensive database integrates multiple datasets, including transcript profiles, epigenomic CpG methylation data, DNA copy number alterations and structural chromosomal rearrangements, tumor-associated gene lists, SNPs, genomic features concerning Alu repeats and general genomic annotations. A genome browser has also been developed that allows for the simultaneous visualization of the different datasets and the various annotated features. Besides enabling an integrative view of diverse datasets through the genome browser, we also provide links to the original references for users to have a more accurate understanding of each specific dataset. This integrated platform will facilitate uncovering interactions among genetic and epigenetic factors associated with brain tumor development. BTECH is freely available at <http://cmbteg.childrensmemorial.org/>.

Electronic supplementary material The online version of this article (doi:10.1007/s12021-010-9091-9) contains supplementary material, which is available to authorized users.

M. Wang (✉) · H. Xie · M. de Fatima Bonaldo ·
M. B. Soares (✉)

Falk Brain Tumor Center, Cancer Biology and Epigenomics Program,
Children's Memorial Research Center, Department of Pediatrics,
Feinberg School of Medicine, Northwestern University,
2300 Children's Plaza, Box 220, Chicago, IL 60614-3394, USA
e-mail: mwang@childrensmemorial.org

M. B. Soares
e-mail: mboares@childrensmemorial.org

W. Stellpflug · T. Tomita
Department of Neurosurgery, Children's Memorial Hospital,
Feinberg School of Medicine, Northwestern University,
2300 Children's Plaza,
Chicago, IL 60614-3394, USA

V. Rajaram
Department of Pathology, Children's Memorial Hospital,
Feinberg School of Medicine, Northwestern University,
2300 Children's Plaza,
Chicago, IL 60614-3394, USA

S. Goldman
Department of Neuro-oncology, Children's Memorial Hospital,
Feinberg School of Medicine, Northwestern University,
2300 Children's Plaza,
Chicago, IL 60614-3394, USA

Keywords Database · Brain tumor · Genome browser ·
Genomics · Epigenomics · DNA methylation · Gene
expression

Introduction

Central nervous system (CNS) tumors, including brain and spinal cord tumors, have an estimated incidence rate of 18.71 cases per 100,000 person-years (American Brain

Tumor Association <http://www.abta.org/>). In children under age 15, CNS tumors are the leading cause of solid tumor death (Jemal et al. 2008). CNS tumors can be classified based on tumor location and histology. From the 2009 statistical reports of the Central Brain Tumor Registry of the United States (CBTRUS), the majority of primary CNS tumors are meningiomas (33.4%), glioblastomas (17.6%) and astrocytomas (7.4%). Currently, diagnosis and treatment of brain tumors are heavily dependent upon tumor histology, albeit brain tumors sharing the same histological features often display variable drug response and clinical behavior. These differences have been shown to be due to numerous genetic and epigenetic aberrations occurring during brain development (Zlatescu et al. 2001; Mueller et al. 2002; Brennan et al. 2009; Gravendeel et al. 2009). Since cancer development and progression result from an accumulation of genetic and/or epigenetic alterations (Vogelstein and Kinzler 1993), a comprehensive understanding of the molecular mechanisms underlying these events is needed for development of more effective treatment regimens. Much effort has been made to discover genetic and epigenetic events underlying brain tumor development, including a number of comprehensive studies aimed at uncovering alterations in gene expression, DNA methylation, DNA copy number alterations and/or chromosomal rearrangements.

Coordinated efforts from various centers and institutes have been taken to address the genomic changes involved in human cancers. The Cancer Genome Atlas (TCGA, <http://www.cancergenome.nih.gov>) has been providing a significant amount of datasets, both on gene expression and DNA methylation, for glioblastoma multiforme (GBM). The Repository of Molecular Brain Neoplasia Data (REMBRANDT, <http://caintegrator-info.nci.nih.gov/>) provides molecular research as well as clinical trials data on brain tumors including gliomas. The REMBRANDT site also provides tools to analyze clinical and experimental data across multiple clinical trials and studies.

In addition, a number of studies on brain tumors have generated large amounts of research data. For gene expression studies, both sequencing- and hybridization-based approaches have been utilized to identify molecular markers for the sub-classification of brain tumors, including Serial Analysis of Gene Expression (SAGE) and multiple expression microarray platforms. In addition, a number of gene expression studies have been conducted to correlate transcript profiles of particular brain tumors with their clinical behavior, response to particular drugs, or recurrence incidences (Freije et al. 2004; Bandres et al. 2005; Hegi et al. 2005; Hoelzinger et al. 2005; Rich et al. 2005; Verhaak et al. 2010). These studies have yielded a number of candidate genes that might prove valuable as prognostic markers and therapeutic targets.

Epigenetic modifications, and DNA methylation in particular, have been shown to be involved in tumor initiation, progression and response to drugs (Hegi et al. 2005; Smith et al. 2007). Methylation of DNA can suppress gene transcription either by preventing the binding of transcription factors (Kim et al. 2003), or by enabling the binding of chromatin repressors (Bird and Wolffe 1999). Large methylation datasets have been generated in an effort to develop the methylation landscape of normal human brain (Rollins et al. 2006; Schumacher et al. 2006; Wu et al. 2010).

Alu retrotransposons, the most prevalent repetitive element in the human genome at 1.2 million copies, have also been associated with the development of various tumors (Lin et al. 1988; Chen et al. 1995; Miki et al. 1996). In addition, Alu family sequences may disrupt the normal splicing pattern of a gene (Lev-Maor et al. 2003). Approximately 8,000 Alu elements have been predicted to have the potential to become exonized (Sorek et al. 2004). In addition, the continued amplification of Alu family sequences has given rise to insertion polymorphisms within human populations (Batzer et al. 1995; Xing et al. 2009).

Lastly, DNA copy number alterations, both losses and gains, are commonly observed in tumors and they are bound to affect gene expression (Taylor et al. 2005; Deshmukh et al. 2008).

Uncovering the molecular mechanisms underlying development and progression of brain tumors requires seamless integration of large genomic, epigenomic and transcriptomic datasets. Large amounts of data have been generated in numerous studies using various techniques, but the resulting datasets are typically in different formats and often not easily retrievable. Here we report the development of a database, **Brain Tumor Epigenomics/genomics Database at Children's Memorial Hospital (BTECH)**, designed to store and to facilitate integration of genomic, epigenomic and transcriptomic datasets derived from brain tumor studies. A genome browser has been installed to allow simultaneous visualization of genomic, epigenomic and transcriptomic features. The BTECH database contains manually curated datasets extracted from publications and various other public resources as well as up to date DNA methylation data generated in our laboratory from normal brain and brain tumor tissues. We anticipate that BTECH will prove invaluable to investigators conducting genomic, epigenomic and/or transcriptomic analyses of brain tumors in their pursuit of the molecular mechanisms underlying development and progression of these tumors.

Data Acquisition and Processing

BTECH harbors a wide range of molecular datasets on—or of relevance to—brain tumors, including gene expression, DNA

methylation, DNA copy number alterations (gains and losses) and structural chromosomal alterations, lists of genes of special interest such as cancer or brain tumor related, and differentially expressed in the brain or in brain tumors, manually curated from public datasets. In addition, BTECH includes general genomic annotations as well as specific data on subsets of Alu elements, including data on insertional polymorphism, transcription and methylation. These are listed in Table 1 along with pertinent information concerning the datasets.

(i) Gene Expression Datasets

To compile a collection of gene expression profiles from various brain tumor studies, queries were first conducted against either Entrez PubMed (<http://www.ncbi.nlm.nih.gov/Literature/>) or the NCBI Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/index.cgi>). The queries for GEO were

conducted with “brain tumor” as the keyword and “Homo sapiens” as the organism. Given the wide range of study purposes and experimental designs, only datasets using surgical tumor samples or drug-treated tumor cells lines were selected. In total, thirteen expression datasets were downloaded, derived from 105 different types or sub-types of brain tumors as well as from brain tumor cell lines (Freije et al. 2004; Bredel et al. 2005; Dong et al. 2005a; Ma et al. 2005; Pachiappan et al. 2005; Taylor et al. 2005; Bredel et al. 2006; Lee et al. 2006; Sun et al. 2006; Tso et al. 2006; Turkheimer et al. 2006; Donson et al. 2009). In addition, a large Serial Analysis of Gene Expression (SAGE) brain tumor dataset derived from seventy three SAGE libraries (Boon et al. 2002) was retrieved from GEO. Altogether these datasets encompass 937 brain-related biological samples (Table 1). Each downloaded SOFT file was further processed (Supplementary Table S1) and the final values were stored in BTECH.

Gene expression data on GBM from TCGA were downloaded from the TCGA data portal site (<http://www.cancergenome.nih.gov/dataportal/data/about/>, total 369 gene expression and 148 DNA methylation datasets as of September 30, 2010). For each dataset, three levels of data are available from TCGA (raw, normalized, and interpreted). Since the interpreted data were generated by merging the normalized data with the annotation data, which provide the most information, this level of data was downloaded. For the BTECH’s purpose of data integration, a total of 80 samples with both gene expression and DNA methylation data were selected. The associated gene expression and DNA methylation data are stored in BTECH, together with the available clinical information for these samples, which can be accessed from BTECH as well.

The REMBRANDT site (<https://caintegrator.nci.nih.gov/rembrandt/>) not only provides molecular research as well as clinical trials data on gliomas, but also provides tools to analyze clinical and experimental data across multiple clinical trials and studies. Because of the complication of the clinical information and the integrity of their analysis, datasets from this study are not included in BTECH. Instead, a hyperlink is provided for further reference.

We also obtained expression data from Gene Expression Atlas2 provided by the Genomics Institute of the Novartis Research Foundation (human reference sequence NCBI Build 36.1). In this study the expression patterns for thousands of protein-coding genes as well as poorly characterized genes were examined (Su et al. 2004). From this study we downloaded the expression profile of 79 different normal human tissues including brain tissues from

Table 1 Data contents in BTECH (as of October 2010)

Gene expression
• Data generated with Affymetrix platforms and SAGE
• 13 datasets
• 73 SAGE libraries
• 105 different types of brain tumors and cell lines
• 937 biological samples
• 79 human normal tissues
DNA methylation
• Data generated with microarray-based techniques, methylation sensitive enzyme-based sequencing, and bisulfate conversion
• Including normal brains and brain tumors
• Methylation levels of 38 Mb genomic regions
• 1.9 million CpG methylation values from 12 different tissues as reference
• 80 datasets with gene expression data from same GBM samples
• 31,178 Alu elements and their 5' flanking sequences from disease-free brain tissues, primary non-aggressive, primary aggressive, and recurrent pediatric ependymomas
Chromosomal alteration
• Data generated with aCGHs and SNPchips
• 68 ependymoma patients, 54 glial tumors, 178 gliomas
• DNA recombination rate from decode, Marshfield, Genethon
Genes of interests
• 2,722 tumor related genes
• 242 brain tumor related genes
• 321 brain tumor differentially expressed genes
Repetitive elements
• Annotation for 1,181,020 Alus
• 8,273 prone to exonized Alus
• 553 polymorphic Alus
General annotation
• Alu, SNP, CpG islands, Entrez genes, RefSeq mRNA, gene 3-frame translation, microRNA

such areas as cerebellum, cortex, medulla oblongata, pituitary, pons, and temporal lobe.

(ii) DNA Methylation Dataset

Several large-scale DNA methylation datasets on human normal brain or brain tumors have been made public in the past few years (Rollins et al. 2006; Schumacher et al. 2006; Wu et al. 2010). The methylation datasets were retrieved from the publications of these studies. Moreover, as part of the Human Epigenome Project, the DNA methylation profiles of chromosomes 6, 20, and 22 were also generated; these were obtained from multiple samples derived from a total of 12 different tissue types (Eckhardt et al. 2006). Although brain tissue was not included in this study, the methylation data might prove valuable as reference for future studies. Therefore, their datasets were downloaded as well (2006 release). As these datasets were generated using different techniques, four sets of *perl* scripts were written in order to make the methylation datasets compatible with one another and also with the BTECH genome browser (Supplementary Table S2).

As aforementioned, methylation data from TCGA on GBM were downloaded from the TCGA data portal site. A total of 80 samples with both DNA methylation and gene expression data were selected. The corresponding DNA methylation and gene expression data, together with the clinical information for these samples, are available from BTECH.

In addition, a high-throughput sequence-based strategy was developed that enables simultaneous determination of the methylation status of thousands of CpG sites localized within and in the 5' flanking genomic region of over 30,000 Alu retrotransposons dispersed throughout the genome (Xie et al. 2009). This is of special relevance given that hypomethylation of repetitive elements has been shown to be strongly associated with cancer progression and poor clinical outcome (Cho et al. 2007; Estecio et al. 2007; Roman-Gomez et al. 2008). Hence, Alu repeats may prove invaluable as reporters of epigenomic alterations during tumor development and progression. Several ependymoma tumor specimens—including primary non-aggressive, primary aggressive, and recurrent tumors—and disease-free brain tissue samples were investigated using the technique discussed above (Xie et al. 2010). The resulting methylation datasets from all these samples are stored in BTECH, hence easily retrievable. To preserve the integrity of the data and the sample set, they are listed under a separate category, namely “Soares’ Lab Methylation Data”, with separate tracks for individual tumor samples. The results from this dataset can be visualized through the genome browser (Fig. 1).

(iii) Chromosomal Alteration Dataset

DNA copy number alterations including both gains and losses are frequently associated with tumor initiation and/or tumor progression (Brito-Babapulle and Atkin 1981; Whang-Peng et al. 1982). Such alterations might result in decreased expression of a tumor suppressor gene or in increased expression of an oncogene. Three such datasets have been generated for brain tumors in recent years. One study used matrix-based comparative genomic DNA microarray to analyze chromosomal alterations in 68 ependymomas (Mendrzyk et al. 2006). Another study used cDNA microarray-based comparative genomic hybridization technology to profile copy number alterations in a series of 54 gliomas encompassing a wide range of histopathological types and tumor grades (Bredel et al. 2005). The third study analyzed 178 tumors for genomic alterations using the Affymetrix 100 K single-nucleotide polymorphism (SNP) arrays (Kotliarov et al. 2006). For each alteration identified in the aforementioned studies, BTECH contains information on genomic localization and, whenever possible, the number of times it has been observed (Supplementary Table S3).

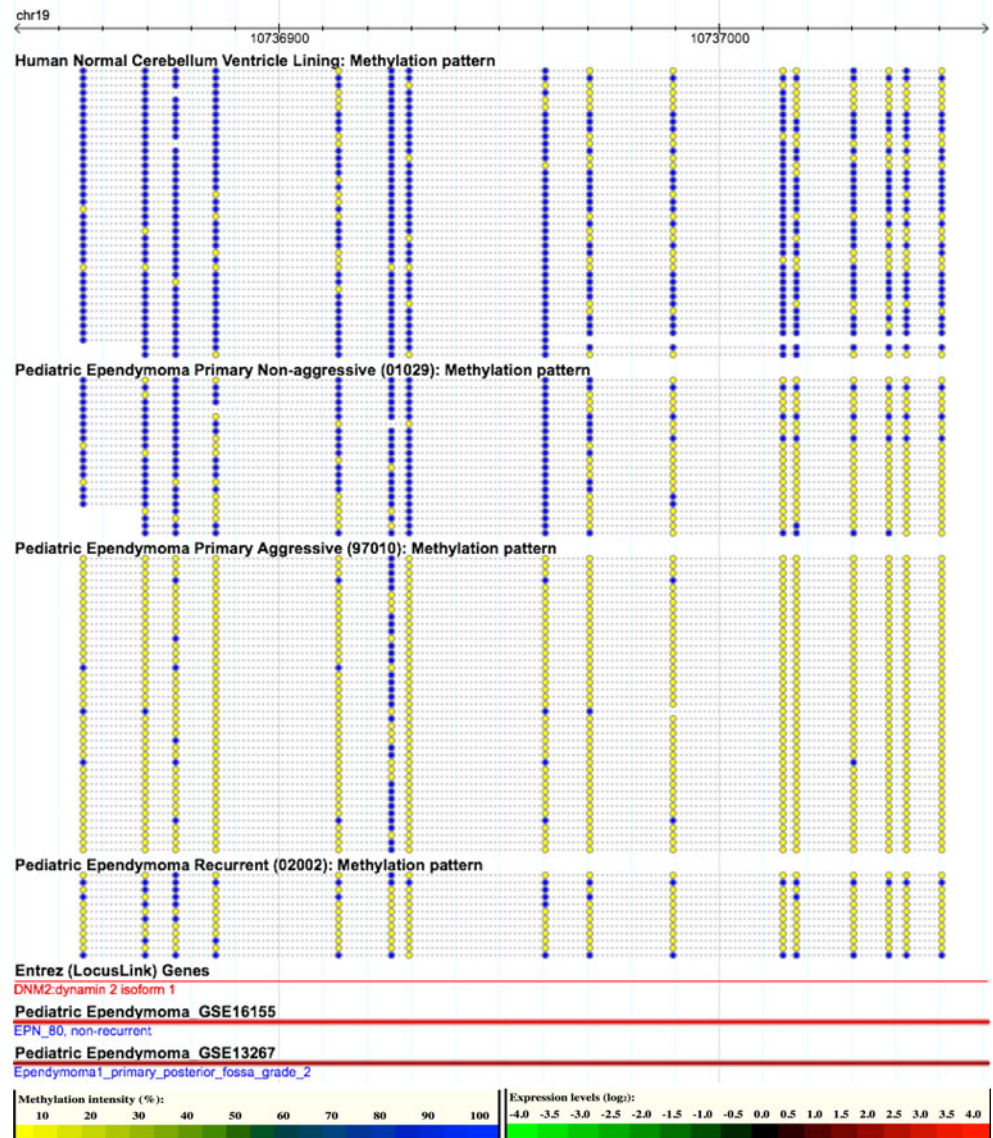
(iv) Genes of Interests

a. Tumor-related Genes

A series of genetic and epigenetic alterations occur during tumor initiation and progression, some conferring a growth advantage. Genes that are directly or indirectly affected by such alterations are referred to as tumor-related genes. We compiled a list of 3,432 tumor-related genes for inclusion in BTECH. These include oncogenes, tumor suppressor genes and genes whose altered expression has been associated with tumor development, progression and/or metastasis.

The compiled list was derived from two resources: the NCBI database and the Affymetrix Human Cancer Array HC-G110 (Affymetrix, <http://www.affymetrix.com/>). First, we searched the “gene” category of the NCBI database for the keyword “tumor” or “cancer”. The NCBI database contains information from the **O**nline **M**endelian **I**nheritance in **M**an (OMIM), and provides a summary of known functions for the genes within the database. This resulted list contained a total of 3,223 genes of which 214 were oncogenes, 57 were tumor suppressor genes and the remainders were genes with altered expression in tumors (as of September 30, 2010). Next, we included 365 genes from the Affymetrix Human Cancer Array HC-G110 (Affymetrix, <http://www.affymetrix.com/>) that were represented in the “known gene” table in the UCSC database.

Fig. 1 An example of an Alu element that shows lower methylation (yellow) in pediatric ependymomas compared to normal brain controls, and the lowest in recurrent tumors. This Alu element resides in an intron of the DNMT2 gene. Datasets from gene expression studies on pediatric ependymomas indicate a higher expression (red) of this gene in tumor samples



In addition, we generated a list of 242 genes that have been reported to be specifically associated with brain tumors, once again using the NCBI database. In this case we searched the “gene” category using the brain tumor names from the WHO grading system. An annotation table for these 242 genes was generated using the UCSC genome database and subsequently included in BTECH.

b. Brain Tumor Differentially Expressed Genes

An initial search of PubMed for brain tumor related large-scale gene expression studies revealed twelve publications, eleven of which involving microarrays (Ljubimova et al. 2001; Mariani et al. 2001; Park et al. 2003; Raza et al. 2004; Dong et al. 2005b; Fathallah-Shaykh 2005; Hoelzinger et al. 2005; Mehrian Shai et al. 2005; Kikuchi et al. 2006; Bozinov et al. 2008; Rand et al. 2008), and one

based on SAGE analysis (Boon et al. 2002). All genes identified in these studies as differentially expressed were extracted for inclusion in BTECH. For the SAGE-based study we extracted the genes that were found to have a *p*-value of less than 0.05 after pair-wise comparisons between normal brain and different types of brain tumors. A total of 544 differentially expressed genes were thus extracted, of which 321 could be unambiguously mapped to specific genomic loci. These 321 human genes were included in the BTECH database, along with the respective annotations regarding conditions under which they were identified as differentially expressed.

(v) Alu Elements

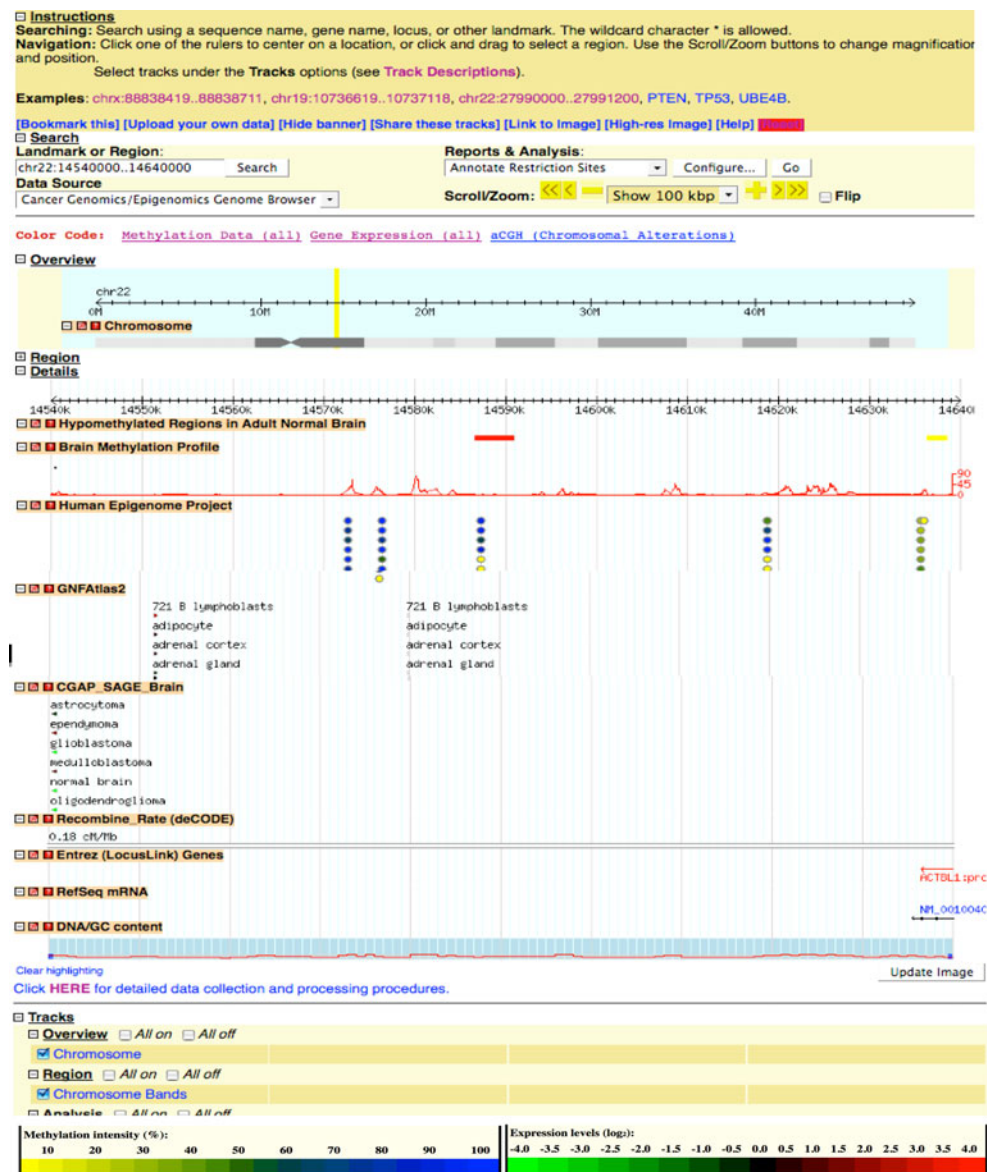
Almost half of the human genome is composed of repetitive elements, which consist of interspersed

repeats and tandem repeats (Jordan et al. 2003). Alu elements, as the most abundant short interspersed nuclear elements, account for approximately 10% of the human genome sequence (Kazazian 2004). Rather than being passive bystanders in the human genome, Alu elements play an active role in cancer predisposition and development. Regions that have a high density of Alu elements have been reported to be associated with cancer-inducing chromosomal alterations such as deletions, duplications, and translocations (Mauillon et al. 1996; Strout et al. 1998; Kolomietz et al. 2001; Gad et al. 2002; Teugels et al. 2005; Abo-Dalo et al. 2010; Iskow et al. 2010). A number of transcriptionally competent elements continue to amplify in the human genome by retrotransposition, thus leading to insertional polymorphism

(Batzer et al. 1995; Cordaux et al. 2007). Moreover, Alu elements were shown to be associated with 5% of the alternatively spliced internal exons in the human genome (Lev-Maor et al. 2003). Approximately 8,000 Alus were predicted to have the potential to exonize (Sorek et al. 2004). In the BTECH database, we provide the annotations for all Alu elements in the human genome (approximately 1.2 million) based on the UCSC's repeat masker table. In addition, we also extracted the list of Alu elements that are polymorphic with respect to integration in human populations (Batzer et al. 1995; Cordaux et al. 2007), as well as the list of Alu elements that are prone to becoming exonized (Sorek et al. 2004).

In addition to the genetic alterations that Alu elements may cause, hypomethylation of Alu elements has been

Fig. 2 An integrative view of brain tumor genomics, epigenomics, and transcriptomic data using genome browser. Detailed view can be found at: <http://cmbteg.childrensmemorial.org/cgi-bin/gbrowse/btech/> with querying region: chr22:14540000..14640000



shown to be strongly associated with cancer progression and poor clinical outcome (Riggs and Jones 1983; Feinberg and Tycko 2004; Cho et al. 2007; Estecio et al. 2007; Roman-Gomez et al. 2008). In a recent publication using a high-throughput sequencing-based technique, we reported the methylation status of 31,178 Alu elements and their 5' flanking sequences in normal brain and in ependymomas of different grades (Xie et al. 2010). These data have been included in the BTECH database, and retrievable from the genome browser listed under category "Soares' Lab Methylation Data" with separate tracks for individual tumor samples (Fig. 1).

(vi) General Annotations

In addition to the molecular data and the selected gene lists discussed above, BTECH also harbors relevant genome annotations. The following genome annotations were downloaded from the databases indicated below:

- RefSeq mRNA, Entrez genes, three-frame translation, DNA recombination rates were downloaded from UCSC genome database (human reference sequence NCBI Build 36.1).
- CpG islands were identified by the program described by *Takai et al.* (Takai and Jones 2002) without masking the repetitive elements.
- SNP annotations were from dbSNP (<http://www.ncbi.nlm.nih.gov/SNP/>, build 127)
- All microRNA elements were downloaded and extracted from the Sanger Institute (<http://www.sanger.ac.uk/>, release 14).
- Chromosome ideogram information was from NCBI.
- Genomic sequences were based on human reference sequence (NCBI Build 36.1).

Database Design and Implementation

The BTECH database has been installed on a Mac OS X Server and is freely accessible through <http://cmbteg.childrensmemorial.org>. The BTECH database consists of a relational database hosted by the MySQL database management system (<http://www.mysql.com/>). In addition, a genome browser has been installed to provide an integrated interface for displaying experimental datasets and annotations using the genome annotation viewer software components from GMOD (Generic Model Organism Database project) (Stein et al. 2002). The datasets displayed on the Genome Browser were represented by different tracks. These tracks are laid on top of the human genomic sequence, thus enabling simultaneous visualization of all corresponding molecular data, sequence features and annotations (Fig. 2).

From the genome browser, users can query the database using gene names, symbols, Genbank accession numbers or genomic co-ordinates.

Future Directions

The BTECH database is the first integrated molecular database that has been designed specifically for brain tumor studies. The database will be continuously updated and expanded. Future development will be focusing on the design of query engines for the database. Such query engines would allow users to search against the vast amount of brain tumor study datasets and provide efficient data mining. In addition, we are also going to analyze various kinds of alterations and develop hypotheses for a better understanding of the molecular mechanisms underlying brain tumor development and progression.

Information Sharing Statement

BTECH and its data contents are freely available to the public from this website: <http://cmbteg.childrensmemorial.org>. The visualization of these data and the corresponding data resources can be accessed from: <http://cmbteg.childrensmemorial.org/cgi-bin/gbrowse/btech/>

Acknowledgements We thank Dr. Elio Vanin (Children's Memorial Research Center, Chicago) for critical reading and valuable comments of the manuscript. This study was supported by The Everett/O'Connor Charitable Trust; Dr. Ralph & Marian C. Falk Medical Research Trust; Gus Foundation; The Maeve McNicholas Memorial Foundation and the Medical Research Institute Council. Last but not least, we thank all members of our community who contributed their data to public databases.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Abo-Dalo, B., Kutsche, K., Mautner, V., & Kluwe, L. (2010). Large intragenic deletions of the NF2 gene: breakpoints and associated phenotypes. *Genes Chromosom Cancer*, 49(2), 171–175.
- Bandres, E., Andion, E., Escalada, A., Honorato, B., Catalan, V., Cubedo, E., et al. (2005). Gene expression profile induced by BCNU in human glioma cell lines with differential MGMT expression. *J Neurooncol*, 73(3), 189–198.
- Batzler, M. A., Rubin, C. M., Hellmann-Blumberg, U., Alegria-Hartman, M., Leeftang, E. P., Stern, J. D., et al. (1995). Dispersion and insertion polymorphism in two small subfamilies of recently amplified human Alu repeats. *J Mol Biol*, 247(3), 418–427.
- Bird, A. P., & Wolffe, A. P. (1999). Methylation-induced repression—belts, braces, and chromatin. *Cell*, 99(5), 451–454.

- Boon, K., Osorio, E. C., Greenhut, S. F., Schaefer, C. F., Shoemaker, J., Polyak, K., et al. (2002). An anatomy of normal and malignant gene expression. *Proc Natl Acad Sci USA*, *99*(17), 11287–11292.
- Bozinov, O., Kohler, S., Samans, B., Benes, L., Miller, D., Ritter, M., et al. (2008). Candidate genes for the progression of malignant gliomas identified by microarray analysis. *Neurosurg Rev*, *31*(1), 83–89. discussion 89–90.
- Bredel, M., Bredel, C., Juric, D., Duran, G. E., Yu, R. X., Harsh, G. R., et al. (2006). Tumor necrosis factor-alpha-induced protein 3 as a putative regulator of nuclear factor-kappaB-mediated resistance to O6-alkylating agents in human glioblastomas. *J Clin Oncol*, *24*(2), 274–287.
- Bredel, M., Bredel, C., Juric, D., Harsh, G. R., Vogel, H., Recht, L. D., et al. (2005). High-resolution genome-wide mapping of genetic alterations in human glial brain tumors. *Cancer Res*, *65*(10), 4088–4096.
- Brennan, C., Momota, H., Hambarzumyan, D., Ozawa, T., Tandon, A., Pedraza, A., et al. (2009). Glioblastoma subclasses can be defined by activity among signal transduction pathways and associated genomic alterations. *PLoS ONE*, *4*(11), e7752.
- Brito-Babapulle, V., & Atkin, N. B. (1981). Break points in chromosome #1 abnormalities of 218 human neoplasms. *Cancer Genet Cytogenet*, *4*(3), 215–225.
- Chen, J., Heerdt, B. G., & Augenlicht, L. H. (1995). Presence and instability of repetitive elements in sequences the altered expression of which characterizes risk for colonic cancer. *Cancer Res*, *55*(1), 174–180.
- Cho, N. Y., Kim, B. H., Choi, M., Yoo, E. J., Moon, K. C., Cho, Y. M., et al. (2007). Hypermethylation of CpG island loci and hypomethylation of LINE-1 and Alu repeats in prostate adenocarcinoma and their relationship to clinicopathological features. *J Pathol*, *211*(3), 269–277.
- Cordaux, R., Srikanta, D., Lee, J., Stoneking, M., & Batzer, M. A. (2007). In search of polymorphic Alu insertions with restricted geographic distributions. *Genomics*, *90*(1), 154–158.
- Deshmukh, H., Yeh, T. H., Yu, J., Sharma, M. K., Perry, A., Leonard, J. R., et al. (2008). High-resolution, dual-platform aCGH analysis reveals frequent HIPK2 amplification and increased expression in pilocytic astrocytomas. *Oncogene*, *27*(34), 4745–4751.
- Dong, S., Nutt, C. L., Betensky, R. A., Stemmer-Rachamimov, A. O., Denko, N. C., Ligon, K. L., et al. (2005a). Histology-based expression profiling yields novel prognostic markers in human glioblastoma. *J Neuropathol Exp Neurol*, *64*(11), 948–955.
- Dong, H., Wade, M., Williams, A., Lee, A., Douglas, G. R., & Yauk, C. (2005b). Molecular insight into the effects of hypothyroidism on the developing cerebellum. *Biochem Biophys Res Commun*, *330*(4), 1182–1193.
- Donson, A. M., Birks, D. K., Barton, V. N., Wei, Q., Kleinschmidt-Demasters, B. K., Handler, M. H., et al. (2009). Immune gene and cell enrichment is associated with a good prognosis in ependymoma. *J Immunol*, *183*(11), 7428–7440.
- Eckhardt, F., Lewin, J., Cortese, R., Rakyan, V. K., Attwood, J., Burger, M., et al. (2006). DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet*, *38*(12), 1378–1385.
- Estecio, M. R., Yan, P. S., Ibrahim, A. E., Tellez, C. S., Shen, L., Huang, T. H., et al. (2007). High-throughput methylation profiling by MCA coupled to CpG island microarray. *Genome Res*, *17*(10), 1529–1536.
- Fathallah-Shaykh, H. M. (2005). Logical networks inferred from highly specific discovery of transcriptionally regulated genes predict protein states in cultured gliomas. *Biochem Biophys Res Commun*, *336*(4), 1278–1284.
- Feinberg, A. P., & Tycko, B. (2004). The history of cancer epigenetics. *Nat Rev Cancer*, *4*(2), 143–153.
- Freije, W. A., Castro-Vargas, F. E., Fang, Z., Horvath, S., Cloughesy, T., Liao, L. M., et al. (2004). Gene expression profiling of gliomas strongly predicts survival. *Cancer Res*, *64*(18), 6503–6510.
- Gad, S., Caux-Moncoutier, V., Pages-Berhouet, S., Gauthier-Villars, M., Coupier, I., Pujol, P., et al. (2002). Significant contribution of large BRCA1 gene rearrangements in 120 French breast and ovarian cancer families. *Oncogene*, *21*(44), 6841–6847.
- Gravendeel, L. A., Kouwenhoven, M. C., Gevaert, O., de Rooi, J. J., Stubbs, A. P., Duijm, J. E., et al. (2009). Intrinsic gene expression profiles of gliomas are a better predictor of survival than histology. *Cancer Res*, *69*(23), 9065–9072.
- Hegi, M. E., Diserens, A. C., Gorlia, T., Hamou, M. F., de Tribolet, N., Weller, M., et al. (2005). MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med*, *352*(10), 997–1003.
- Hoelzinger, D. B., Mariani, L., Weis, J., Woyke, T., Berens, T. J., McDonough, W. S., et al. (2005). Gene expression profile of glioblastoma multiforme invasive phenotype points to new therapeutic targets. *Neoplasia*, *7*(1), 7–16.
- Iskow, R. C., McCabe, M. T., Mills, R. E., Torene, S., Pittard, W. S., Neuwald, A. F., et al. (2010). Natural mutagenesis of human genomes by endogenous retrotransposons. *Cell*, *141*(7), 1253–1261.
- Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., Murray, T., et al. (2008). Cancer statistics, 2008. *CA Cancer J Clin*, *58*(2), 71–96.
- Jordan, I. K., Rogozin, I. B., Glazko, G. V., & Koonin, E. V. (2003). Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends Genet*, *19*(2), 68–72.
- Kazazian, H. H., Jr. (2004). Mobile elements: drivers of genome evolution. *Science*, *303*(5664), 1626–1632.
- Kikuchi, T., Daigo, Y., Ishikawa, N., Katagiri, T., Tsunoda, T., Yoshida, S., et al. (2006). Expression profiles of metastatic brain tumor from lung adenocarcinomas on cDNA microarray. *Int J Oncol*, *28*(4), 799–805.
- Kim, J., Kollhoff, A., Bergmann, A., & Stubbs, L. (2003). Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, Peg3. *Hum Mol Genet*, *12*(3), 233–245.
- Kolomietz, E., Al-Maghrabi, J., Brennan, S., Karaskova, J., Minkin, S., Lipton, J., et al. (2001). Primary chromosomal rearrangements of leukemia are frequently accompanied by extensive submicroscopic deletions and may lead to altered prognosis. *Blood*, *97*(11), 3581–3588.
- Kotliarov, Y., Steed, M. E., Christopher, N., Walling, J., Su, Q., Center, A., et al. (2006). High-resolution Global Genomic Survey of 178 Gliomas Reveals Novel Regions of Copy Number Alteration and Allelic Imbalances. *Cancer Res*, *66*(19), 9428–9436.
- Lee, J., Kotliarova, S., Kotliarov, Y., Li, A., Su, Q., Donin, N. M., et al. (2006). Tumor stem cells derived from glioblastomas cultured in bFGF and EGF more closely mirror the phenotype and genotype of primary tumors than do serum-cultured cell lines. *Cancer Cell*, *9*(5), 391–403.
- Lev-Maor, G., Sorek, R., Shomron, N., & Ast, G. (2003). The birth of an alternatively spliced exon: 3' splice-site selection in Alu exons. *Science*, *300*(5623), 1288–1291.
- Lin, C. S., Goldthwait, D. A., & Samols, D. (1988). Identification of Alu transposition in human lung carcinoma cells. *Cell*, *54*(2), 153–159.
- Ljubimova, J. Y., Lakhter, A. J., Loksh, A., Yong, W. H., Riedinger, M. S., Miner, J. H., et al. (2001). Overexpression of alpha4 chain-containing laminins in human glial tumors identified by gene microarray analysis. *Cancer Res*, *61*(14), 5601–5610.
- Ma, D., Nutt, C. L., Shanehsaz, P., Peng, X., Louis, D. N., & Kaetzel, D. M. (2005). Autocrine platelet-derived growth factor-dependent gene expression in glioblastoma cells is mediated largely by activation of the transcription factor sterol regulatory element binding protein and is associated with altered genotype and patient survival in human brain tumors. *Cancer Res*, *65*(13), 5523–5534.

- Mariani, L., Beaudry, C., McDonough, W. S., Hoelzinger, D. B., Demuth, T., Ross, K. R., et al. (2001). Glioma cell motility is associated with reduced transcription of proapoptotic and proliferation genes: a cDNA microarray analysis. *J Neurooncol*, *53*(2), 161–176.
- Mauillon, J. L., Michel, P., Limacher, J. M., Latouche, J. B., Dechelotte, P., Charbonnier, F., et al. (1996). Identification of novel germline hMLH1 mutations including a 22 kb Alu-mediated deletion in patients with familial colorectal cancer. *Cancer Res*, *56*(24), 5728–5733.
- Mehrian Shai, R., Reichardt, J. K., Ya-Hsuan, H., Kremen, T. J., Liao, L. M., Cloughesy, T. F., et al. (2005). Robustness of gene expression profiling in glioma specimen samplings and derived cell lines. *Brain Res Mol Brain Res*, *136*(1–2), 99–103.
- Mendrzyk, F., Korshunov, A., Benner, A., Toedt, G., Pfister, S., Radlwimmer, B., et al. (2006). Identification of gains on 1q and epidermal growth factor receptor overexpression as independent prognostic markers in intracranial ependymoma. *Clin Cancer Res*, *12*(7 Pt 1), 2070–2079.
- Miki, Y., Katagiri, T., Kasumi, F., Yoshimoto, T., & Nakamura, Y. (1996). Mutation analysis in the BRCA2 gene in primary breast cancers. *Nat Genet*, *13*(2), 245–247.
- Mueller, W., Hartmann, C., Hoffmann, A., Lanksch, W., Kiwit, J., Tonn, J., et al. (2002). Genetic signature of oligoastrocytomas correlates with tumor location and denotes distinct molecular subsets. *Am J Pathol*, *161*(1), 313–319.
- Pachiappan, A., Thwin, M. M., Manikandan, J., & Gopalakrishnakone, P. (2005). Glial inflammation and neurodegeneration induced by candoxin, a novel neurotoxin from Bungarus candidus venom: global gene expression analysis using microarray. *Toxicol*, *46*(8), 883–899.
- Park, P. C., Taylor, M. D., Mainprize, T. G., Becker, L. E., Ho, M., Dura, W. T., et al. (2003). Transcriptional profiling of medulloblastoma in children. *J Neurosurg*, *99*(3), 534–541.
- Rand, V., Prebble, E., Ridley, L., Howard, M., Wei, W., Brundler, M. A., et al. (2008). Investigation of chromosome 1q reveals differential expression of members of the S100 family in clinical subgroups of intracranial paediatric ependymoma. *Br J Cancer*, *99*(7), 1136–1143.
- Raza, S. M., Fuller, G. N., Rhee, C. H., Huang, S., Hess, K., Zhang, W., et al. (2004). Identification of necrosis-associated genes in glioblastoma by cDNA microarray analysis. *Clin Cancer Res*, *10* (1 Pt 1), 212–221.
- Rich, J. N., Hans, C., Jones, B., Iversen, E. S., McLendon, R. E., Rasheed, B. K., et al. (2005). Gene expression profiling and genetic markers in glioblastoma survival. *Cancer Res*, *65*(10), 4051–4058.
- Riggs, A. D., & Jones, P. A. (1983). 5-methylcytosine, gene regulation, and cancer. *Adv Cancer Res*, *40*, 1–30.
- Rollins, R. A., Haghghi, F., Edwards, J. R., Das, R., Zhang, M. Q., Ju, J., et al. (2006). Large-scale structure of genomic methylation patterns. *Genome Res*, *16*(2), 157–163.
- Roman-Gomez, J., Jimenez-Velasco, A., Agirre, X., Castillejo, J. A., Navarro, G., San Jose-Eneriz, E., et al. (2008). Repetitive DNA hypomethylation in the advanced phase of chronic myeloid leukemia. *Leuk Res*, *32*(3), 487–490.
- Schumacher, A., Kapranov, P., Kaminsky, Z., Flanagan, J., Assadzadeh, A., Yau, P., et al. (2006). Microarray-based DNA methylation profiling: technology and applications. *Nucleic Acids Res*, *34*(2), 528–542.
- Smith, L. T., Otterson, G. A., & Plass, C. (2007). Unraveling the epigenetic code of cancer for therapy. *Trends Genet*, *23*(9), 449–456.
- Sorek, R., Lev-Maor, G., Reznik, M., Dagan, T., Belinky, F., Graur, D., et al. (2004). Minimal conditions for exonization of intronic sequences: 5' splice site formation in alu exons. *Mol Cell*, *14*(2), 221–231.
- Stein, L. D., Mungall, C., Shu, S., Caudy, M., Mangone, M., Day, A., et al. (2002). The generic genome browser: a building block for a model organism system database. *Genome Res*, *12*(10), 1599–1610.
- Strout, M. P., Marcucci, G., Bloomfield, C. D., & Caligiuri, M. A. (1998). The partial tandem duplication of ALL1 (MLL) is consistently generated by Alu-mediated homologous recombination in acute myeloid leukemia. *Proc Natl Acad Sci USA*, *95*(5), 2390–2395.
- Su, A. I., Wiltshire, T., Batalov, S., Lapp, H., Ching, K. A., Block, D., et al. (2004). A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc Natl Acad Sci USA*, *101*(16), 6062–6067.
- Sun, L., Hui, A. M., Su, Q., Vortmeyer, A., Kotliarov, Y., Pastorino, S., et al. (2006). Neuronal and glioma-derived stem cell factor induces angiogenesis within the brain. *Cancer Cell*, *9*(4), 287–300.
- Takai, D., & Jones, P. A. (2002). Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc Natl Acad Sci USA*, *99*(6), 3740–3745.
- Taylor, M. D., Poppleton, H., Fuller, C., Su, X., Liu, Y., Jensen, P., et al. (2005). Radial glia cells are candidate stem cells of ependymoma. *Cancer Cell*, *8*(4), 323–335.
- Teugels, E., De Brakeleer, S., Goelen, G., Lissens, W., Sermijn, E., & De Greve, J. (2005). De novo Alu element insertions targeted to a sequence common to the BRCA1 and BRCA2 genes. *Hum Mutat*, *26*(3), 284.
- Tso, C. L., Shintaku, P., Chen, J., Liu, Q., Liu, J., Chen, Z., et al. (2006). Primary glioblastomas express mesenchymal stem-like properties. *Mol Cancer Res*, *4*(9), 607–619.
- Turkheimer, F. E., Roncaroli, F., Hennuy, B., Herens, C., Nguyen, M., Martin, D., et al. (2006). Chromosomal patterns of gene expression from microarray data: methodology, validation and clinical relevance in gliomas. *BMC Bioinform*, *7*, 526.
- Verhaak, R. G., Hoadley, K. A., Purdom, E., Wang, V., Qi, Y., Wilkerson, M. D., et al. (2010). Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell*, *17*(1), 98–110.
- Vogelstein, B., & Kinzler, K. W. (1993). The multistep nature of cancer. *Trends Genet*, *9*(4), 138–141.
- Whang-Peng, J., Kao-Shan, C. S., Lee, E. C., Bunn, P. A., Carney, D. N., Gazdar, A. F., et al. (1982). Specific chromosome defect associated with human small-cell lung cancer; deletion 3p(14–23). *Science*, *215*(4529), 181–182.
- Wu, X., Rauch, T. A., Zhong, X., Bennett, W. P., Latif, F., Krex, D., et al. (2010). CpG island hypermethylation in human astrocytomas. *Cancer Res*, *70*(7), 2718–2727.
- Xie H, Wang M, Bonaldo MD, Smith C, Rajaram V, Goldman S, Tomita T, Soares MB (2009) High-throughput sequence-based epigenomic analysis of Alu repeats in human cerebellum. *Nucleic Acids Res*
- Xie, H., Wang, M., Bonaldo Mde, F., Rajaram, V., Stellpflug, W., Smith, C., et al. (2010). Epigenomic analysis of Alu repeats in human ependymomas. *Proc Natl Acad Sci USA*, *107*(15), 6952–6957.
- Xing, J., Zhang, Y., Han, K., Salem, A. H., Sen, S. K., Huff, C. D., et al. (2009). Mobile elements create structural variation: analysis of a complete human genome. *Genome Res*, *19*(9), 1516–1526.
- Zlatescu, M. C., TehraniYazdi, A., Sasaki, H., Megyesi, J. F., Betensky, R. A., Louis, D. N., et al. (2001). Tumor location and growth pattern correlate with genetic signature in oligodendroglial neoplasms. *Cancer Res*, *61*(18), 6713–6715.