

DNA-MICROARRAY ANALYSIS OF BRAIN CANCER: MOLECULAR CLASSIFICATION FOR THERAPY

Paul S. Mischel*, Timothy F. Cloughesy[‡] and Stanley F. Nelson[§]

Abstract | Primary brain tumours are among the most lethal of all cancers, largely as a result of their lack of responsiveness to current therapy. Numerous new therapies hold great promise for the treatment of patients with brain cancer, but the main challenge is to determine which treatment is most likely to benefit an individual patient. DNA-microarray-based technologies, which allow simultaneous analysis of expression of thousands of genes, have already begun to uncover previously unrecognized patient subsets that differ in their survival. Here, we review the progress made so far in using DNA microarrays to optimize brain cancer therapy.

GLIOMA

Any brain tumour that originates from the glial cell lineage.

Departments of Pathology and Laboratory Medicine*, Neurology[‡], Human Genetics[§] and the Henry E. Singleton Brain Cancer Research Program at the David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California 90095, USA.
Correspondence to P.M.
e-mail: pmischel@mednet.ucla.edu
doi:10.1038/nrn1518

In *Anna Karenina*, Leo Tolstoy wrote: 'All happy families resemble one another, each unhappy family is unhappy in its own way'. Oddly enough, this might be a rather apt analogy for cancer. The highly regulated molecular events that are crucial for normal development and function are very similar between individuals, but in cancer, genetic and epigenetic alterations result in cascades of deregulated molecular events, which lead to genetically complex, highly individual tumours. The complexity is daunting, but finding consistencies that can be therapeutically exploited is vital for the development and clinical application of new treatments. Until recently, the tools that are required to attack this problem were not available, but the sequencing of the human genome and the development of global approaches for surveying virtually the entire expressed genome make this type of enquiry possible. The challenge has moved from deciphering the genetic code to understanding how it is used, or, in the case of cancer, misused.

DNA microarrays enable the acquisition of gene-expression data on a scale that was previously unimaginable. Computational methods for analysing vast amounts of data are being developed and quantitative tools for analysing networks are now available¹⁻³. This will facilitate the detection of meaningful patterns in these complex gene-expression signatures. Technology and bioinformatics have provided an unprecedented

opportunity to explore the development and function of the nervous system and to analyse diseases of the nervous system such as brain cancer.

The brain cancer problem

Let us begin by building a case for the importance of the brain cancer problem to the neuroscience community. First, there is a public-health imperative. Brain cancer is now the leading cause of death from cancer in children under the age of 15 and the second leading cause of death from cancer from age 15 to 34. In adults, brain cancer is proportionately less common than other cancers, yet it accounts for a disproportionate percentage of deaths from cancer⁴. At present, patients with glioblastoma (the most common form of GLIOMA in adults) have a median survival time of 12 months from the time of diagnosis, despite aggressive surgery, radiation and chemotherapy⁵. These numbers are not good by anyone's standards.

Second, there is a scientific imperative. Traditionally, neuroscientists have focused on degenerative diseases and developmental brain disorders, which provide valuable insights into the normal development and function of the nervous system. Brain tumours were considered to be too intractable, and not enough was known about them to provoke widespread interest. Recent advances in brain cancer genetics and the development of mouse brain cancer models show that many

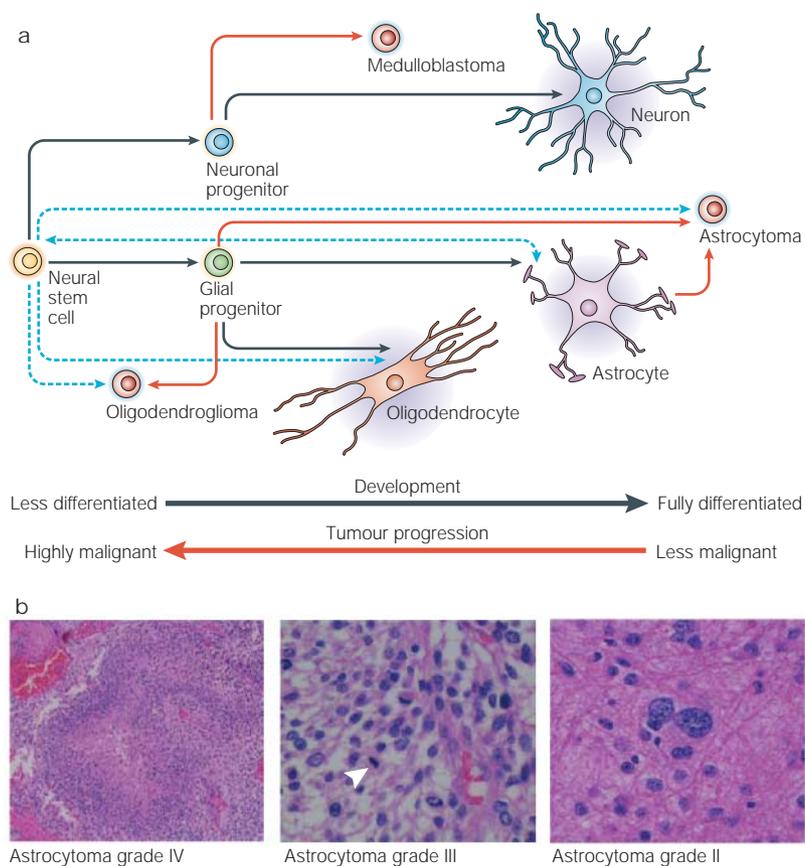


Figure 1 | Classification scheme for brain tumours. **a** | The present classification scheme for brain tumours. This classic model is based on the assumption that tumour cells of a specific lineage share microscopic similarity to a presumed neural or glial precursor. The black arrows indicate the hypothesized normal development and the red arrows indicate the hypothesized cell of origin of CNS tumours. It should be emphasized that recent work highlights a key role for neural stem cells in normal development and potentially in the formation of brain tumours (dotted blue arrows)^{7–9,11,12}. **b** | According to this scheme, less malignant tumours resemble their normal tissue counterparts; more malignant tumours resemble less differentiated precursor cells. Tumours are graded on the basis of the extent of anaplasia (de-differentiation) and other microscopic features that connote aggressive behaviour — such as mitotic activity, tumour necrosis and angiogenesis. Low-grade astrocytomas (grade II) have some anaplasia, but lack mitotic activity and necrosis. Intermediate-grade astrocytomas (grade III) have more anaplasia and readily detectable mitotic activity, but not necrosis. The white arrow points to a mitotic figure. Glioblastomas, which are the most malignant grade of astrocytoma, are highly anaplastic and contain mitotic activity and tumour necrosis.

of the pivotal mechanisms that are important for normal brain development are precisely those that have gone awry in brain cancer. Therefore, there is much to be learned about the development and function of the nervous system from studying brain cancer.

Biology of brain tumours

Primary brain tumours arise from the constituent cells of the CNS or their meningeal covering, whereas secondary brain tumours METASTASIZE from a distant site. In 1928, Bailey and Cushing suggested that brain tumours could be classified by their microscopic resemblance to a presumed CNS cell of origin or its developmental precursor⁶. Although recent work shows a more complex pattern, in which neural stem cells have an important role in both glial and neuronal development, and

potentially in the formation of oligodendroglial and astrocytic tumours, the Bailey and Cushing model has remained a guiding principle for brain tumour classification (FIG. 1)^{7–12}. Primary brain tumours that are composed of cells that resemble astrocytes are classified as astrocytomas. Similarly, tumours that resemble oligodendrocytes and ependymal cells are classified as oligodendrogliomas and ependymomas. Cerebellar tumours composed of small round cells that resemble the neuronal precursor cells of the external granule cell layer are classified as medulloblastomas. Immunohistochemical protein markers of astrocytic or neuronal differentiation, such as glial fibrillary acidic protein and synaptophysin, respectively, are used to corroborate the microscopic diagnosis¹³.

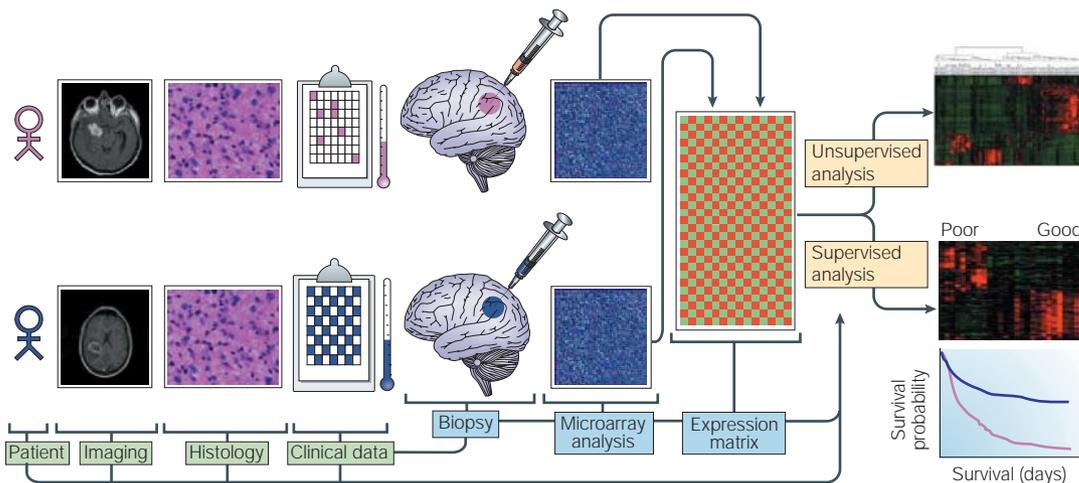
According to this scheme, less malignant tumours resemble their normal tissue counterparts, whereas more malignant tumours resemble less differentiated precursor cells. So, anaplasia ('de-differentiated appearance') implies biological aggressiveness. Tumours are graded according to the extent of anaplasia that they show, and the presence and extent of other microscopic features that connote aggressive behaviour, such as mitotic activity, tumour necrosis and angiogenesis. Low-grade astrocytomas have some anaplasia, but lack mitotic activity, necrosis and vascular proliferation. Intermediate-grade astrocytomas (anaplastic astrocytomas) have more anaplasia and readily detectable mitotic activity, but no necrosis or angiogenesis. Glioblastomas, the most malignant grade of astrocytoma, are highly anaplastic and contain mitotic activity, tumour necrosis and/or vascular proliferation. Medulloblastomas, also a highly malignant type of tumour, are like glioblastomas with regard to their 'blastic' (or highly anaplastic) appearance. This classification and grading system has proved useful for predicting the overall survival for groups of patients with brain tumours. However, it provides relatively limited insight into the underlying molecular lesions. Furthermore, clinically relevant subsets that might differ significantly in their clinical course and response to therapy cannot be identified by the current classification system.

Recent work shows that chronic activation of key intracellular signalling pathways might be crucial for the formation and progression of brain cancer. Chronic activation of the phosphatidylinositol 3-kinase and the Ras–MAPK (mitogen-activated protein kinase) signalling pathways, which arise from a range of upstream genetic lesions, promotes glioma formation and progression in mouse genetic models¹⁴. Similarly, unopposed sonic hedgehog (SHH) pathway activation (on the basis of patched (PTC) haploinsufficiency), disruption of DNA repair owing to deficiency of DNA ligase IV, and combined cell-cycle dysregulation and p53 dysfunction all promote the formation of medulloblastoma in mouse genetic models^{15–19}. Correlative studies of patient samples show chronic activation of many of these pathways in tumour samples^{20,21}. The same pathways that promote tumour formation and progression might actually sensitize cancer cells to targeted pathway inhibitors^{22–24}. Therefore, new classification systems for brain cancer need to be developed that can identify alterations in the

METASTASIS

The spread of cancer cells from one organ or tissue to another, usually though the blood stream or the lymphatic system.

Box 1 | DNA microarrays



DNA-microarray analysis is most useful when it can be integrated with clinical, imaging and histological data. Substantial effort is required to develop appropriate databases that contain key clinical information, including patient characteristics such as age and sex. Brain imaging is routinely undertaken and images are housed in a central database. Histological photomicrographs document cellular morphology, and clinical data are entered in real time through wireless input devices to ensure accurate and up-to-date information. Biopsy material is preserved for future analyses, linked to clinical data and used to extract RNA for large-scale expression analysis using microarrays.

DNA microarrays can survey virtually the entire expressed genome. A small amount of high quality RNA from tumour (or non-tumour) tissue is labelled and hybridized on the surface of a chip, which is composed of spotted cDNA clones or probes spotted or synthesized on the surface of the chip (oligonucleotide arrays), providing a relatively reproducible and affordable way to analyse thousands of genes simultaneously (see figure). The availability of high quality, high-density microarrays, coupled with improved methods for RNA extraction, preservation and labelling, have reduced many of the inherent technical challenges in microarray studies. The primary hurdles now lie in the interpretation, rather than the acquisition, of the data.

Interpretation of DNA-microarray data is challenging because of its potential for noise and because of the complexity involved in analysing a data matrix of thousands of elements (typically over 40,000 transcripts in hundreds of tumours). First, the data are normalized so that gene-expression profiles can be compared between samples (individual chips). Then genes for which expression does not vary meaningfully throughout the experiment, but which can confuse data interpretation (the 'noise'), are filtered out. To find meaningful patterns, computational methods are used, which might help to define relevant groups of tumours and/or genes. Hierarchical clustering can be used to identify groups of tumours or genes with similar global gene-expression profiles. These transcriptionally defined groups can then be probed for correlations with biological, histological or survival-associated distinctions. This type of analysis, in which groups are defined entirely on the basis of gene-expression profiles without reference to tumour type or grade, is considered to be 'unsupervised'. Alternatively, it is possible to identify groups of genes for which expression correlates with a biological, histological or survival-associated parameter ('supervised analysis')¹⁰¹. Supervised and unsupervised analyses provide different and often complementary types of information, so most microarray studies use combinations of these approaches. Unsupervised analysis provides global portraits about the predominant grouping of the data, but data can be grouped in many ways and the predominant grouping might not be the most biologically relevant structure. Supervised analysis, by identifying groups of genes that correlate with a relevant parameter, can provide relevant lists of differentially expressed genes that might highlight important biological differences.

SINGLE NUCLEOTIDE POLYMORPHISMS
Bi-allelic (typically) base pair substitutions, which are the most common forms of genetic polymorphism.

ALTERNATIVE SPLICING
During splicing, introns are excised from RNA after transcription and the cut ends are rejoined to form a continuous message. Alternative splicing allows the production of different messages from the same DNA molecule.

pathways and networks that drive their progression, and which could also potentially be used to select them for targeted therapy.

Predictive molecular diagnostics
A revolution in clinical medicine. The genomic revolution is transforming clinical medicine. Instead of the current model of population risk assessment and empirical treatment, we will move to one of predictive individualized care based on molecular classification and targeted therapy²⁵. High-throughput genomic techniques will accelerate this process. Screening for gene polymorphisms and

loss of heterozygosity by SINGLE NUCLEOTIDE POLYMORPHISM microarrays; analysing chromosomal gains and losses by comparative genomic-hybridization arrays; determining global patterns of methylation, acetylation and ALTERNATIVE SPLICING on microarrays; and identifying characteristic proteomic profiles will probably all play a part in the new molecular diagnostics²⁶⁻³¹. Gene-expression profiling will probably be central to this effort. There are many ways to globally survey the expressed genome, but, in this review, we focus on the role of DNA microarrays in developing predictive molecular diagnostics for patients with brain tumours (BOX 1).

Targeting specific pathways. Why are predictive molecular diagnostics so important? Largely, this is a function of new and improved therapeutic options. Radiation and cytotoxic chemotherapy, the standard forms of traditional cancer treatment, gain their therapeutic advantage predominantly from the increased susceptibility to treatment of rapidly dividing tumour cells relative to non-neoplastic cells. However, there is also considerable damage to rapidly proliferating non-neoplastic cells such as bone marrow and gastrointestinal epithelium. New approaches are being developed to specifically target proteins or pathways that are altered in tumour cells, thereby potentially providing more effective and less toxic therapies. These new pathway inhibitors, alone or in synergistic combination with other pathway inhibitors and/or traditional chemotherapeutic agents, are likely to become standard therapy²².

The high frequency of signalling-pathway alterations in brain tumours, such as chronic phosphoinositide 3-kinase pathway activation in malignant gliomas and chronic SHH signalling in medulloblastomas, enables identification of potential therapeutic targets for small-molecule inhibitors^{22–24}. High-throughput screens of small-molecule inhibitors have generated many pathway/oncogene-specific drugs that could potentially be used for the treatment of patients with cancer. However, as promising as these inhibitors are, they are likely to fail if they are not directed to the right patients. Importantly, traditional pathological examination is unlikely to provide sufficient insight into the underlying molecular alterations that might guide the selection of patients for these molecularly targeted therapies. Morphologically identical tumours can be distinct in their mutational patterns, signalling-pathway alterations and gene-expression profiles, and, most importantly, in their response to a range of therapies³². Therefore, new predictive molecular diagnostics need to be developed and integrated with drug development and clinical-trial design.

Early experience of pathway inhibitors in clinical trials has highlighted the importance of molecular diagnostics. The success story for targeted therapy in clinical trials comes from a specific tumour type — chronic myelogenous leukaemia (CML) — in which the target (the constitutively active BCR–ABL (breakpoint cluster region–Abelson murine leukaemia viral oncogene homologue) kinase fusion protein) is nearly always expressed. The ABL-kinase inhibitor imatinib mesylate (Gleevec (STI-571), Novartis) promotes remission in up to 95% of patients with CML^{33,34}. For patients with CML, the morphological diagnosis indicates the presence of the molecular target (the BCR–ABL fusion protein). However, most solid tumours, including brain tumours, are not like CML. The molecular target might be present in only a relatively small subset of morphologically identical tumours. This unrecognized molecular heterogeneity with regard to the drug target is likely to complicate clinical trials³⁵. Furthermore, multiple molecular lesions can collaborate to activate pathways in such a way that targeting a specific molecule might be ineffective if there is a downstream mutation that also activates that pathway³⁶. Rational application of specific

therapies will require the detection of relevant molecular subsets and biomarkers of these subsets that can guide appropriate patient selection. DNA microarrays (BOX 1), by virtue of their ability to detect global snapshots of transcriptional changes that provide a ‘readout’ of multiple upstream pathway alterations, might provide an important new tool for predictive molecular diagnostics.

Detecting meaningful genomic signatures. The first generation of DNA-microarray studies in human cancer focused on detecting differences in gene-expression profiles between tumours of different types and grades. The result was clear: tumours that arise from cells of different origin and that show different grades of aggressiveness have distinctive global transcriptional signatures^{37–42}. The formation of cytoskeletal and nuclear structures and the assembly of complex architectural tissue patterns requires expression of specific proteins — which is largely a function of gene transcription. A similar principle applies when pathologists use morphological assessment as a reflection of the underlying cell biology^{43–47}. That DNA microarrays can be used to recapitulate well-known distinctions in tumour type indicates that gene-expression profiles that are detected by genomic analysis will be biologically relevant.

Detecting global transcriptional differences between tumours that look different is reassuring, but it might not be clinically useful. Pathologists are already skilled at distinguishing tumours of different types and grades. The value that is added by genomic analysis is likely to arise from the identification of previously unrecognized, clinically relevant molecular subsets and the development of predictive profiles and specific biomarkers that will guide therapy. Furthermore, the discovery capabilities of DNA microarrays could potentially yield new therapeutic targets. For a variety of cancers, including leukaemias, lymphomas and epithelial cancers, DNA microarrays have enabled detection of patient subsets that differ significantly in survival, findings which could not be detected by traditional pathological examination^{48–56}. These data indicate that DNA microarrays will provide a powerful tool for the molecular dissection of clinically relevant tumour subsets.

It is important to recognize that differential expression of genes does not necessarily imply causality and is only a first step towards target identification. Potentially important therapeutic targets might be differentially expressed at the protein but not the RNA level, so that they might not be detected by DNA-microarray analysis. The activation state of some proteins might also be more important than their expression level for determining their suitability as a target. This point is highlighted by the observation that, in patients with non-small-cell lung cancer, activating mutations in the kinase domain of the epidermal growth factor receptor (EGFR) are strongly associated with an enhanced clinical response to the EGFR inhibitor gefitinib^{57,58}. Therefore, for the purpose of target identification, DNA microarrays provide an important first step that must be followed by extensive functional analyses.

Molecular subsets of brain tumours

Medulloblastomas. Medulloblastomas have distinctive global gene-expression profiles that readily distinguish them from some morphological mimics, including supratentorial primitive neuroectodermal tumours (which look exactly like medulloblastomas but arise in the cerebral hemispheres instead of the cerebellum) and atypical teratoid/rhabdoid tumours, as well as malignant gliomas⁵⁹. There is also evidence for molecular subsets of medulloblastomas. The transcriptional pattern of DESMOPLASTIC medulloblastomas (a common variant of medulloblastoma) differs from classic medulloblastomas, particularly in the elevated expression of several components of the SHH signalling pathway, which supports a role for deregulated SHH signalling in these tumours⁵⁹.

DNA microarrays can also detect molecular subsets of medulloblastoma cases that differ in terms of patient survival. Complementary DNA (cDNA)-microarray analysis of 60 medulloblastoma samples that were taken before treatment yielded a class-predictor model composed of as few as eight genes that could accurately predict the survival of the patients with medulloblastoma⁵⁹. Therefore, the gene-expression profiles contained vital prognostic information that could potentially be detected and captured in models incorporating small numbers of genes for clinical screening. For medulloblastoma, patients have traditionally been stratified into 'average-risk' or 'high-risk' groups, largely on the basis of age, extent of post-surgical residual disease and metastasis. Pomeroy and colleagues demonstrated that the gene-expression data could predict patient outcome independently of these clinical variables⁶⁰.

Taking a different approach to the problem, DNA-microarray analysis was used to identify gene-expression differences between metastatic and non-metastatic medulloblastomas. A relatively small number of genes (85) could accurately classify a sample as being metastatic or non-metastatic⁶¹. Platelet-derived growth factor receptor α (*PDGFR α*) was one of the genes that correlated most strongly with metastasis at both the mRNA and protein level, and PDGFR signalling promoted medulloblastoma adhesion and CHEMOTAXIS in a MAPK1/2-dependent fashion in assays that were carried out *in vitro*. These results imply a functional role for the PDGFR-signalling pathway in medulloblastoma invasion and indicate its potential as a biomarker of aggressiveness⁶¹.

In addition to their value for analysis of patient samples, DNA microarrays can also provide important insights into disease when they are applied to experimental models that can be genetically or pharmacologically manipulated. Lee *et al.* analysed the gene-expression profiles of medulloblastomas that were derived from a set of genetically defined mouse crosses¹⁵. They compared the transcriptional profiles of medulloblastomas that were derived from *PTC1*^{+/-} (patched homologue 1^{+/-}) mice (15% medulloblastoma occurrence), *PTC*^{+/-};*p53*^{-/-} mice (100% medulloblastoma occurrence), *LIG4*^{-/-} (DNA ligase IV^{-/-}); *p53*^{-/-} mice (100% medulloblastoma

occurrence), and *p53*^{-/-} mice treated with cyclin-dependent kinase (CDK) inhibitors (which cause a relatively low frequency of medulloblastomas). The gene-expression pattern across medulloblastomas was similar regardless of the genetic background in which they formed¹⁵. The global transcriptional pattern of medulloblastomas was similar to that of the developing cerebellum, but not the mature cerebellum, which supports the concept that medulloblastomas might arise from immature cerebellar precursor cells such as the external granule-cell layer.

To identify downstream genes that might be involved in medulloblastoma development in the background of deregulated SHH activity, Oliver *et al.* applied DNA-microarray analysis to cerebellar granule-cell precursors isolated from wild-type or *PTC1*^{+/-} mice⁶². They identified a set of SHH transcriptional targets, including cyclin D1 and neuroblastoma *Myc*-related oncogene (*NMYC*), the expression of which has been implicated in medulloblastoma development and progression. Taken together, these studies further highlight the role of deregulated SHH signalling in neural precursor cells during medulloblastoma formation. Importantly, these data are very much in line with the gene-expression profiles that have been detected in patient samples⁵⁹.

Gliomas. Low-grade astrocytomas, oligodendrogliomas and glioblastomas have distinctive global gene-expression profiles, which are clearly separable from each other and from the profiles of normal brain tissue (FIG. 2)⁶³⁻⁶⁸. Furthermore, these tumour subtypes can be accurately distinguished from each other by a relatively small number of genes, which are heavily weighted towards genes encoding proteins that are involved in such crucial processes as cell proliferation, proteosomal function, energy metabolism and signal transduction⁶³. Using a similar approach, Khatua *et al.* observed elevated expression of components of the EGFR-FKBP12-HIF2 α (EGFR-FK506 binding protein 1A-hypoxia-inducible factor 2, α subunit) pathway in high-grade childhood astrocytomas relative to low-grade childhood astrocytomas⁶⁹. So, differences in gene-expression patterns might help in the development of new therapeutic targets. But can DNA microarrays provide insights into diagnostically challenging gliomas, and can they detect subsets of morphologically identical tumours? Several recent studies shed light on these questions.

The pathological distinction between glioblastoma and anaplastic oligodendroglioma is an important one: the prognosis for glioblastoma patients is substantially worse, and the treatment options are different. For 'classic' examples of each tumour, the distinction is not difficult. However, many malignant gliomas are more challenging. Because the distinction is purely morphological, a patient's diagnosis is largely the subjective assessment of the pathologist based on relatively slim data. To determine whether gene-expression profiles could accurately classify these tumours, Nutt and colleagues performed DNA-microarray analysis⁷⁰. They compared the gene-expression patterns of tumours that

DESMOPLASTIC

A term that refers to the growth of dense fibrous tissue around a tumour.

cDNA

Complementary DNA that is produced from an RNA template by an RNA-dependent DNA polymerase.

CHEMOTAXIS

The movement of cells in response to a chemical gradient that is provided by chemotactic agents.

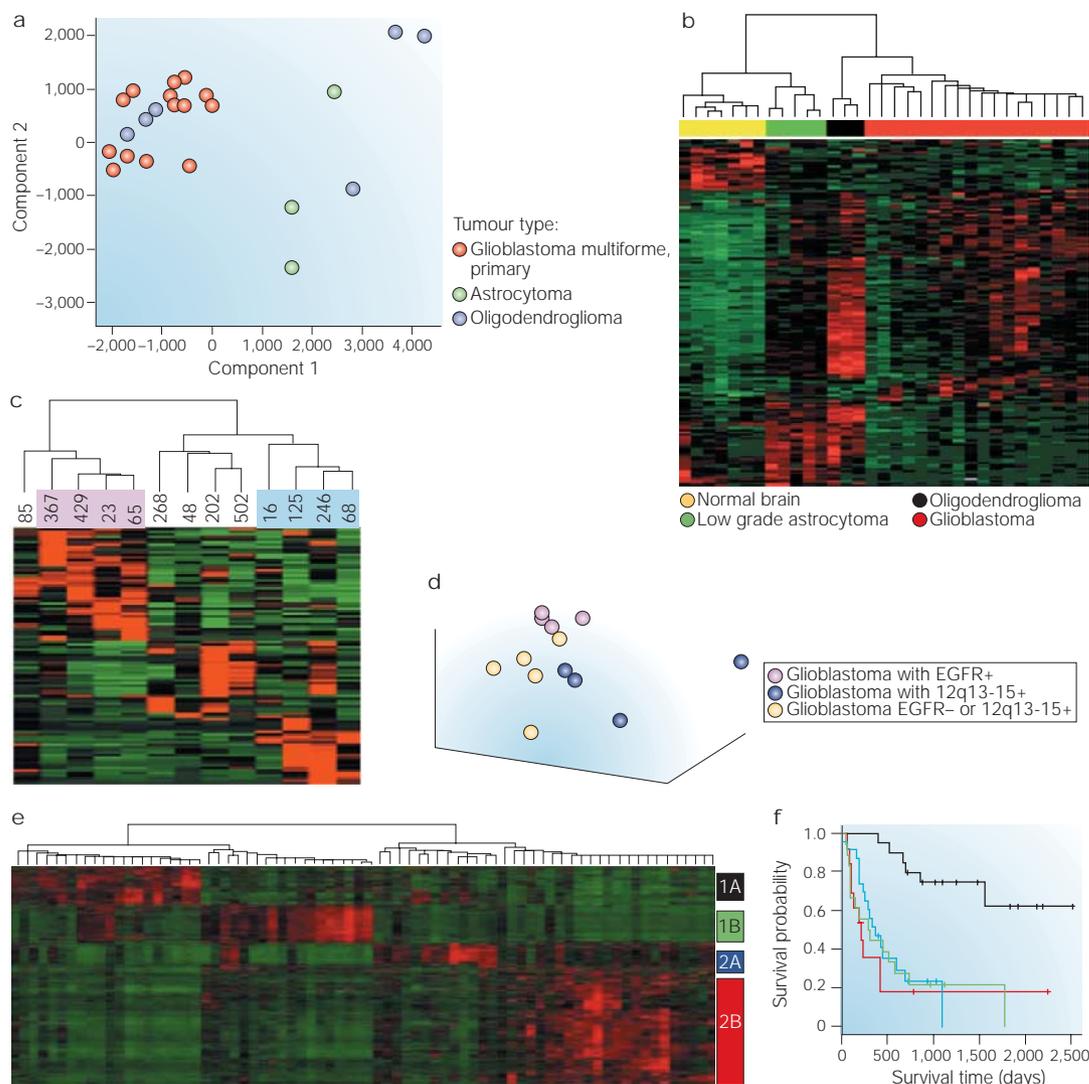


Figure 2 | DNA-microarray analyses can identify relevant clinical subsets of gliomas. **a** and **b** show that different subtypes of gliomas have distinct gene-expression profiles. **a** | The gene-expression patterns of gliomas of different types and grades. Multidimensional scaling on the basis of expression of 12,555 genes shows that gliomas of different histological type and grade have distinct transcriptional profiles. **b** | Hierarchical clustering shows that gliomas of different type and grade can be readily distinguished from each other and from normal brain by a relatively small number of genes (170 genes). Figure adapted, with permission, from REF. 63 © (2003) Macmillan Magazines Ltd. **c** and **d** show identification of molecular subsets of microscopically identical glioblastomas. **c** | Hierarchical clustering identifies three molecular subsets of primary glioblastomas on the basis of the differential expression of 90 genes. One subset is associated with epidermal growth factor receptor (EGFR) overexpression (pink), one is associated with overexpression of a contiguous set of genes on chromosome 12q13-15 (blue), and the third lacks either alteration. **d** | Multidimensional scaling shows that these subsets have distinct global transcriptional profiles. Figure adapted, with permission from REF. 71 © (2003) Macmillan Magazines Ltd. **e** and **f** show the detection of clinically relevant, previously undetected subsets of patients with high-grade gliomas that have significantly different survival times. **e** | Hierarchical clustering of 85 high-grade glioma samples on the basis of the expression of 595 genes that are highly differentially expressed in patients with relatively good survival times versus those with shorter survival times⁷². Four subsets of patients are detected. **f** | Kaplan–Meier survival analysis shows that these genes can identify the subset of patients who are most likely to have prolonged survival times (cluster 1A, black).

RT-PCR
Reverse
transcriptase–polymerase chain
reaction (PCR) — a reaction in
which messenger RNA is
converted into DNA (reverse
transcription), which is then
amplified by PCR.

were clearly and easily pathologically classified as either ‘classic’ glioblastomas or ‘classic’ anaplastic oligodendrogliomas, and developed a gene-expression-based classifier. When they applied this classifier to a set of diagnostically challenging specimens, they found that gene-expression profiling was a more reliable method of predicting survival than pathological assessment. Practically, this analysis might help in the development

of specific immunohistochemical or reverse transcriptase–polymerase chain reaction (RT-PCR) markers that could be translated into clinical practice. However, this study points to an issue of greater importance: transcriptional information contains more data about outcome than does pathological examination, and this could potentially be used to develop a predictive molecular diagnostic procedure.

As well as helping the classification of diagnostically challenging tumours, DNA microarrays might also enable the detection of molecular subsets within 'classic' pathological types. One recent study used global gene-expression analysis to uncover new molecular subsets of morphologically identical tumours⁷¹. EGFR expression is common in primary glioblastomas (those that arise *de novo* as high-grade tumours), detected in approximately two-thirds of cases. Until recently it has been unclear whether EGFR-expressing glioblastomas are a distinct molecular subset, and the biological and transcriptional consequences of EGFR overexpression in glioblastomas have not been clarified. To address this issue, primary glioblastomas were stratified as being either EGFR protein-expressing or EGFR protein-negative, and differences in the transcriptional profiles were analysed. EGFR-expressing glioblastomas had a globally distinctive pattern of gene expression compared with non-EGFR-expressing primary glioblastomas, indicating that they are a biologically relevant subset. Furthermore, a relatively small number of genes could be used to distinguish between EGFR-expressing and EGFR-negative primary glioblastomas, and this list of genes was highly enriched for signalling molecules, many of which could potentially provide therapeutic targets. Not surprisingly, the EGFR-negative primary glioblastomas were not a uniform subclass: at least two further subsets were detected, including one in which a set of contiguous genes on chromosome 12q13-15 was overexpressed, which is consistent with a chromosomal amplification. These data indicate that patterns of gene expression can uncover biologically relevant molecular subsets of morphologically identical glioblastomas (FIG. 2)⁷¹. The short survival time of patients with glioblastoma (12 month median survival) has made the detection of differences in patient survival difficult.

A recent large-scale DNA-microarray analysis by Freije *et al.* showed that gene-expression-based grouping of tumours is a more powerful predictor of survival than pathological type, grade or age⁷². The authors constructed a gene-expression-based classifier that detected distinct molecular subsets of morphologically identical gliomas, including glioblastomas, that differed significantly in terms of patient survival (FIG. 2). This classifier was validated on an additional external and independent dataset from another institution^{70,72}. The study shows that DNA-microarray analysis can identify previously unrecognized, clinically relevant subsets of patients with glioblastoma in a robust and reproducible fashion.

DNA microarrays to predict therapeutic response
Genomic correlates of chemosensitivity. That DNA microarrays can be used to detect molecular subsets that differ in terms of survival time indicates that it will soon be possible to develop gene-based predictors of therapeutic responses. Many model-system studies are already addressing this question. The NCI60, a panel of cancer cell lines that is maintained by the National Cancer Institute, has been screened for chemosensitivity to a large number of anti-cancer agents. Overlaying gene-expression-profiling data on the functional-response

data in these cell lines indicates that it will be possible to develop DNA-microarray-based approaches to predicting chemosensitivity⁷³⁻⁷⁶. The clinical application of this strategy will depend on the design of well-coordinated clinical trials, in which predictors that have been developed in model systems and in retrospective analyses of patient samples can then be tested in prospective clinical trials. Several promising studies indicate that this will be possible; for example, retrospective analyses of gene-expression profiles that correlate with the response to chemotherapy for a few types of cancer (non-brain tumours) have begun to emerge^{77,78}.

DNA microarrays might also prove to be a powerful tool for the rational application of combination therapy. By comparing pre- and post-treatment gene-expression profiles in patients with acute lymphocytic leukaemia who were treated with two different chemotherapies (methotrexate and mercaptopurine), alone or in combination, Cheok and colleagues showed that the effects of single-agent versus combination therapy on gene expression were largely non-overlapping⁷⁹. That is, combination therapy did not result in an additive gene-expression response; it induced a profoundly different transcriptional response in comparison to either agent when administered alone. These data warrant a re-consideration of the potential effects of combination chemotherapy, and also a reinterpretation as to the mechanism of its efficacy.

As predictive molecular diagnostics are being developed, new anti-cancer compounds are being screened. Chemical genomics, in which chemical libraries are tested for the ability to modulate cellular states, is central to this process. These screens usually rely on a functional readout; that is, the ability of a chemical to induce growth arrest or differentiation. Stegmaier *et al.* showed that DNA-microarray analysis can be used to identify a small set of genes that could serve as a surrogate marker of the desired cellular state (differentiation)⁸⁰. This small set of genes could then serve as a molecular diagnostic tool that could easily be assayed with an RT-PCR reaction, thereby facilitating the functional analysis of compounds. This approach, which is known as gene-expression-based high-throughput screening, is likely to have a substantial impact on the screening of chemical libraries and the development of new drugs⁸⁰.

A new concept for biomarkers. The use of DNA microarrays is set to revolutionize the development and use of tumour biomarkers. The few brain tumour biomarkers that are currently available include chromosomal loss of 1p and 19q for oligodendrogliomas⁸¹, and neurotrophic tyrosine kinase, receptor, type 3 (TRKC), NMYC, CMYC (the related MYC family member) and v-erb-b2 erythroblastic leukaemia viral oncogene homologue 2 (ErbB2) for medulloblastoma⁸². Analysis of the association between any of these single biomarkers and response to therapy requires large numbers of patients to obtain sufficient statistical power. A single biomarker has limited predictive power if many other genes or proteins are important for determining outcome. By contrast, DNA

microarrays can be used to detect groups of genes that, in the aggregate, contain significantly more predictive information than does any individual biomarker. The fact that an eight-gene model can be used to predict survival in medulloblastoma⁵⁹, and that a six-gene model can be used to predict the outcome in patients with breast cancer⁴⁸, indicates that chemotherapy-response predictors can be modelled using relatively small panels of genes that can be screened by RT-PCR or immunohistochemistry. This will allow more robust conclusions to be drawn from a series of much smaller, more streamlined studies. Early data indicate that this approach, as well as being easier to carry out and more economical, will yield better methods of patient stratification for therapy⁸³. By carefully analysing toxicities, it might be possible to use the same array data to identify predictors of adverse therapeutic responses.

DNA microarrays might also have an important role in the identification of novel serum biomarkers and molecular-imaging probes. Efficient biophysical-separation methods for detecting differentially expressed mRNAs that encode secreted and membrane-associated gene products have been developed⁸⁴, and bioinformatic approaches for detecting secreted proteins can also be applied to DNA-microarray data. This approach has already led to the identification of the secreted glycoprotein **YKL40** as a potential biomarker for glioblastoma⁸⁵. DNA microarrays could probably also be used in the identification of a new generation of molecular-imaging probes⁸⁶, which will be useful for diagnosis and for monitoring responses to therapy. For example, DNA-microarray studies can identify genes, the expression of which might be a surrogate marker of pathway activation. *In vivo* imaging using such a probe would allow direct, repeatable and quantitative non-invasive monitoring of the effect of targeted pathway inhibitors in patients.

Using microarrays to analyse pathways

Recently, there has been an increasing recognition that detecting gene-expression patterns that are conserved across species can highlight key functional networks^{87,88}, and identifying gene-expression patterns that are common to many types of cancer (and/or other biological processes) might also be enlightening. Recent studies have identified a stereotyped fibroblast serum-response gene-expression pattern that is shared by wound healing and many types of cancer, and which is a predictor of metastases and short patient survival⁸⁹. Similarly, Whitfield *et al.* identified a cell-cycle-specific pattern of gene co-expression in HeLa cervical cancer cells in culture, which is detected in various cancers but not in normal proliferative tissues⁹⁰. Studies in which gene-expression signatures are compared across biological processes and/or tumour types should yield important new insights into gene expression in cancer.

Assembling gene lists into pathways. Genes do not act as individual units, they collaborate in overlapping pathways, the deregulation of which is a hallmark of cancer. New bioinformatics tools are being developed that will

allow the projection of potential pathway alterations on the basis of gene-expression data. Gene-ontology databases, which allow for dynamic mapping of gene-expression data into potential pathways on the basis of their functional annotation and known molecular interactions, are central to this effort^{91,92}. In addition, integrated pathway-analysis tools such as Ingenuity Pathway Analysis and Cytoscape have been developed, which can integrate gene-expression data with other molecular databases — such as protein-interaction databases — to facilitate the development of new and more complete pathway maps^{91,93}. As well as providing convenient ways to analyse existing data, there is potential for important information to be discovered as a consequence of the iterative nature of this process. Empirical DNA-microarray data from tissue samples or experimental models can be placed in the context of present knowledge about pathways, and new and expanded pathway connections or specific gene-gene interactions can potentially be inferred, which can be functionally analysed and used to build on the existing pathway knowledge base.

Refining pathway maps might have important implications for many areas of neuroscience, not just brain cancer. Cancer cells do not 'invent' new pathways; they use pre-existing pathways in different ways or they combine components of these pathways in a new fashion. By mapping, expanding and refining pathway maps in brain cancer, DNA-microarray studies might provide insight into the connectivity of these pathways in the developing and normally functioning brain. One only needs to consider how much has been learned about normal brain function by analysing signalling pathways in a relatively small number of cancer cell lines, such as the rat pheochromocytoma line PC12 (REFS 94,95), to recognize the potential value of these studies.

Network analysis. In addition to analysing pathways by integrating gene-expression data with functional-annotation databases, it will be important to begin to analyse gene-expression networks without any *a priori* assumptions. This is particularly important if in the diseased state the genes interact with each other in different pathways or networks than they do in health. The last few years have seen a substantial growth in the recognition of the importance of networks and the development of quantitative tools for analysing them². Complex systems, which range from non-biological to cellular networks, adhere to universal organizational principles. Modularity, in which cellular functions are carried out by groups of interacting molecules, is an important feature of these networks^{2,96,97}. For example, modularity can be detected in the metabolic protein-interaction networks of many organisms⁹⁷. It can also be detected at the level of gene-expression networks^{87,88,98}. Clusters of genes with related functions show correlated expression patterns^{87,88}. Furthermore, modules of interconnected genes with shared biological functions are conserved across the gene-expression networks of various species. High-level self-organization can also be detected in the gene-expression networks of

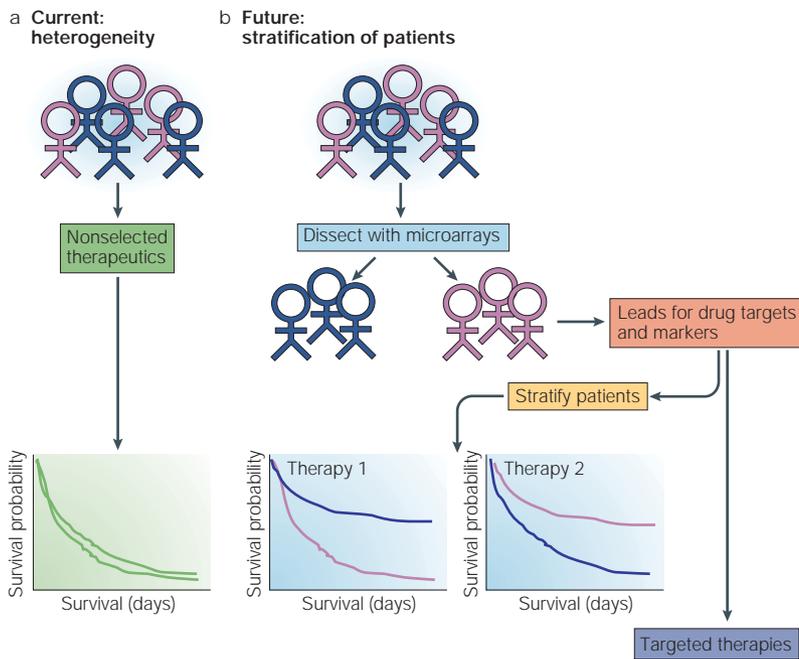


Figure 3 | Prospects for integrating genomic analysis of brain tumours with clinical-trial development. **a** | Currently, patient inclusion in clinical trials is highly reliant on histological classification, which provides only limited insight into the molecular heterogeneity of brain tumours. Therefore, potentially effective treatments that might be of benefit to specific patient subsets (which are not detectable by histology) will not be recognized. **b** | With the integration of genomic analyses such as microarrays, heterogeneous groups can be identified to allow patient stratification. In addition, the detection and characterization of the molecular heterogeneity provides direct and indirect insights into probable targets for inhibition therapy. These inhibitors can be used more efficiently by targeting them to groups of patients whose tumours are more likely to respond to the specific therapies. This approach might also identify biomarkers that can then be used to stratify patients for targeted therapy.

cancer cells⁹⁹. Analysing the network properties of gene-expression data might reveal the organizational pattern of gene expression in cancer, which might, in turn, help us to identify new potential drug targets.

Moving into the clinic: challenges ahead

Reproducibility and data sharing. The recent generation of commercially available DNA microarrays are proving to be technically robust and reliable. However, subtle differences in sample preparation and RNA extraction and labelling can have a profound impact on the gene-expression data. More importantly, the bioinformatics strategies that are used to analyse the data are far from uniform. Therefore, reproducibility between laboratories and DNA-microarray platforms is an important issue. The key to addressing this problem is the public availability of raw data. Groups that identify clinically relevant gene-expression signatures need to be able to validate them on independent data sets from other institutions. Furthermore, the validity of gene-expression patterns for predicting the outcome or response to therapy needs to be independent of the array platform that is used for analysis. Cooperation between investigators, larger validation studies and improved data-analysis

methods across array platforms will be vital for the integrity and translational value of DNA-microarray studies. For example, the recent work of Freije *et al.* showed that a DNA-microarray-based survival predictor was robust across sample sets, institutions and DNA-microarray platforms^{70,72}.

Arrays for everyone, or marker subsets? If further analysis continues to support the usefulness of DNA microarrays as molecular diagnostics of response and outcome, should each cancer patient receive a DNA-microarray analysis of their tumour? Universal DNA-microarray screening of cancer patients would require standardized laboratories with standardized procedures for performance and analysis of the gene-expression data¹⁰⁰. Alternatively, would it be better to distil these gene-expression differences into small genetic marker sets that remain highly predictive, but which can be easily assayed? Furthermore, will the screening be carried out by quantitative RT-PCR or immunohistochemistry? Can we develop methods for carrying out these assays such as RT-PCR reactions on routinely processed paraffin-embedded biopsies? Clearly, we have much work to do.

Towards a personalized medicine. At the beginning of this review, we described the shift away from population risk assessment and empirical treatment of patients with brain tumours to one of predictive individualized care based on molecular classification and targeted therapy²⁵. The tools for this are already close at hand. It is now possible to imagine a day in the not-too-distant future when serum biomarkers and molecular imaging probes that are identified by DNA microarrays will be used for screening or early detection. Tumours will undergo global DNA-microarray analysis (or analysis of a subset of markers) to identify pathway alterations that point to the most beneficial therapy or combination of therapies (FIG. 3). Responses to therapy will be quantitatively and reproducibly monitored in a minimally invasive fashion using molecular-imaging probes and/or serum biomarkers to detect the biological effect of drugs on their intended target gene or pathway. In parallel, new therapies will emerge as DNA microarrays and other global genomic and proteomic technologies probe networks and pathways for their Achilles' heel. New molecular diagnostics will emerge that combine global gene-expression analysis with advances in activation-specific antibodies²⁰, proteomic analysis²⁶ and other genomic techniques to look at polymorphisms, chromosomal gains and losses, point mutations and methylation patterns²⁸⁻³⁰. The iterative nature of this discovery process and the ability to test these new 'biomarkers' in smaller and more streamlined clinical trials will continue to refine molecular diagnostics and will also provide insights into the underlying biology of brain cancer. In the process we might provide new hope for brain cancer patients and shed more light on our understanding of normal brain development and function.

1. Slonim, D. K. From patterns to pathways: gene expression data analysis comes of age. *Nature Genet.* **32** (Suppl.), 502–508 (2002).
2. Barabasi, A. L. & Oltvai, Z. N. Network biology: understanding the cell's functional organization. *Nature Rev. Genet.* **5**, 101–113 (2004).
This paper is an excellent review of the concept of network biology, and describes the quantitative tools that can be used to analyse networks.
3. Ideker, T. Systems biology 101 — what you need to know. *Nature Biotechnol.* **22**, 473–475 (2004).
4. Legler, J. M. *et al.* RESPONSE: re: brain and other central nervous system cancers: recent trends in incidence and mortality. *J. Natl Cancer Inst.* **92**, 77–78 (2000).
5. Mischel, P. S. & Cloughesy, T. F. Targeted molecular therapy of GBM. *Brain Pathol.* **13**, 52–61 (2003).
6. Bailey, P. & Cushing, H. A *Classification of the Tumors of the Glioma Group on a Histogenic Basis with a Correlated Study of Prognosis* (Lippincott, Philadelphia, 1928).
7. Lu, Q. R. *et al.* Common developmental requirement for Olig function indicates a motor neuron/oligodendrocyte connection. *Cell* **109**, 75–86 (2002).
8. Doetsch, F. The glial identity of neural stem cells. *Nature Neurosci.* **6**, 1127–1134 (2003).
9. Doetsch, F., Caille, I., Lim, D. A., Garcia-Verdugo, J. M. & Alvarez-Buylla, A. Subventricular zone astrocytes are neural stem cells in the adult mammalian brain. *Cell* **97**, 703–716 (1999).
10. Zhou, Q. & Anderson, D. J. The bHLH transcription factors OLIG2 and OLIG1 couple neuronal and glial subtype specification. *Cell* **109**, 61–73 (2002).
11. Hemmati, H. D. *et al.* Cancerous stem cells can arise from pediatric brain tumors. *Proc. Natl Acad. Sci. USA* **100**, 15178–15183 (2003).
12. Bachoo, R. M. *et al.* Epidermal growth factor receptor and Ink4a/Arf: convergent mechanisms governing terminal differentiation and transformation along the neural stem cell to astrocyte axis. *Cancer Cell* **1**, 269–277 (2002).
13. Kleihues, P. *et al.* The WHO classification of tumors of the nervous system. *J. Neuropathol. Exp. Neurol.* **61**, 215–225 (2002).
14. Holland, E. C. *et al.* Combined activation of Ras and Akt in neural progenitors induces glioblastoma formation in mice. *Nature Genet.* **25**, 55–57 (2000).
15. Lee, Y. & McKinnon, P. J. DNA ligase IV suppresses medulloblastoma formation. *Cancer Res.* **62**, 6395–6399 (2002).
In this paper, the authors compare the global gene-expression profiles of medulloblastomas that are derived from a set of genetically defined mouse crosses, thereby identifying the contribution of a number of genes, including PTC1, LIG4 and p53, in the development of medulloblastoma.
16. Lee, Y. *et al.* A molecular fingerprint for medulloblastoma. *Cancer Res.* **63**, 5428–5437 (2003).
17. Weiner, H. L. *et al.* Induction of medulloblastomas in mice by sonic hedgehog, independent of Gli1. *Cancer Res.* **62**, 6385–6389 (2002).
18. Rao, G., Pedone, C. A., Coffin, C. M., Holland, E. C. & Fults, D. W. c-Myc enhances sonic hedgehog-induced medulloblastoma formation from nestin-expressing neural progenitors in mice. *Neoplasia* **5**, 198–204 (2003).
19. Rao, G. *et al.* Sonic hedgehog and insulin-like growth factor signaling synergize to induce medulloblastoma formation from nestin-expressing neural progenitors in mice. *Oncogene* **23**, 6156–6162 (2004).
20. Choe, G. *et al.* Analysis of the phosphatidylinositol 3'-kinase signaling pathway in glioblastoma patients *in vivo*. *Cancer Res.* **63**, 2742–2746 (2003).
21. Ermoian, R. P. *et al.* Dysregulation of PTEN and protein kinase B is associated with glioma histology and patient survival. *Clin. Cancer Res.* **8**, 1100–1106 (2002).
22. Vivanco, I. & Sawyers, C. L. The phosphatidylinositol 3-kinase AKT pathway in human cancer. *Nature Rev. Cancer* **2**, 489–501 (2002).
23. Berman, D. M. *et al.* Medulloblastoma growth inhibition by hedgehog pathway blockade. *Science* **297**, 1559–1561 (2002).
This paper was crucial in defining the role of the hedgehog pathway in the genesis of medulloblastoma, and in identifying inhibition of the hedgehog pathway as a potential therapy.
24. Wechsler-Reya, R. & Scott, M. P. The developmental biology of brain tumors. *Annu. Rev. Neurosci.* **24**, 385–428 (2001).
This outstanding review highlights the genetic mechanisms that are known to be involved in medulloblastoma, as well as other paediatric brain tumours.
25. Ramaswamy, S. Translating cancer genomics into clinical oncology. *N. Engl. J. Med.* **350**, 1814–1816 (2004).
26. Liotta, L. A. *et al.* Protein microarrays: meeting analytical challenges for clinical applications. *Cancer Cell* **3**, 317–325 (2003).
27. Lindblad-Toh, K. *et al.* Loss-of-heterozygosity analysis of small-cell lung carcinomas using single-nucleotide polymorphism arrays. *Nature Biotechnol.* **18**, 1001–1005 (2000).
28. Hoque, M. O., Lee, C. C., Cairns, P., Schoenberg, M. & Sidransky, D. Genome-wide genetic characterization of bladder cancer: a comparison of high-density single-nucleotide polymorphism arrays and PCR-based microsatellite analysis. *Cancer Res.* **63**, 2216–2222 (2003).
29. Albertson, D. G. & Pinkel, D. Genomic microarrays in human genetic disease and cancer. *Hum. Mol. Genet.* **12**, R145–R152 (2003).
30. Shi, H. *et al.* Triple analysis of the cancer epigenome: an integrated microarray system for assessing gene expression, DNA methylation, and histone acetylation. *Cancer Res.* **63**, 2164–2171 (2003).
31. Yeakley, J. M. *et al.* Profiling alternative splicing on fiber-optic arrays. *Nature Biotechnol.* **20**, 353–358 (2002).
32. Mischel, P. S., Nelson, S. F. & Cloughesy, T. F. Molecular analysis of glioblastoma: pathway profiling and its implications for patient therapy. *Cancer Biol. Ther.* **2**, 242–247 (2003).
This review provides an overview of the gene-expression and signal-transduction alterations in glioblastoma and suggests potential therapeutic strategies.
33. Druker, B. J. Perspectives on the development of a molecularly targeted agent. *Cancer Cell* **1**, 31–36 (2002).
34. Sawyers, C. L. Disabling Abl-perspectives on Abl kinase regulation and cancer therapeutics. *Cancer Cell* **1**, 13–15 (2002).
35. Betensky, R. A., Louis, D. N. & Cairncross, J. G. Influence of unrecognized molecular heterogeneity on randomized clinical trials. *J. Clin. Oncol.* **20**, 2495–2499 (2002).
36. Bianco, R. *et al.* Loss of PTEN/MMAC1/TEP in EGF receptor-expressing tumor cells counteracts the antitumor action of EGF tyrosine kinase inhibitors. *Oncogene* **22**, 2812–2822 (2003).
37. Golub, T. R. *et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**, 531–537 (1999).
38. Armstrong, S. A. *et al.* MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nature Genet.* **30**, 41–47 (2002).
39. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
40. Sorlie, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl Acad. Sci. USA* **98**, 10869–10874 (2001).
41. Sorlie, T. *et al.* Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc. Natl Acad. Sci. USA* **100**, 8418–8423 (2003).
42. Ramaswamy, S. *et al.* Multiclass cancer diagnosis using tumor gene expression signatures. *Proc. Natl Acad. Sci. USA* **98**, 15149–15154 (2001).
43. Shia, J. *et al.* Value of histopathology in predicting microsatellite instability in hereditary nonpolyposis colorectal cancer and sporadic colorectal cancer. *Am. J. Surg. Pathol.* **27**, 1407–1417 (2003).
44. Cardiff, R. D. *et al.* Validation: the new challenge for pathology. *Toxicol. Pathol.* **32**, 31–39 (2004).
45. Boorman, G. A. *et al.* Toxicogenomics, drug discovery, and the pathologist. *Toxicol. Pathol.* **30**, 15–27 (2002).
46. Dhanasekaran, S. M. *et al.* Delineation of prognostic biomarkers in prostate cancer. *Nature* **412**, 822–826 (2001).
47. Shappell, S. B. *et al.* Prostate pathology of genetically engineered mice: definitions and classification. The consensus report from the Bar Harbor meeting of the Mouse Models of Human Cancer Consortium Prostate Pathology Committee. *Cancer Res.* **64**, 2270–2305 (2004).
48. Lossos, I. S. *et al.* Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. *N. Engl. J. Med.* **350**, 1828–1837 (2004).
49. Bullinger, L. *et al.* Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N. Engl. J. Med.* **350**, 1605–1616 (2004).
50. Valk, P. J. *et al.* Prognostically useful gene-expression profiles in acute myeloid leukemia. *N. Engl. J. Med.* **350**, 1617–1628 (2004).
51. Rosenwald, A. *et al.* The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N. Engl. J. Med.* **346**, 1937–1947 (2002).
52. Glinisky, G. V., Higashiyama, T. & Glinisky, A. B. Classification of human breast cancer using gene expression profiling as a component of the survival predictor algorithm. *Clin. Cancer Res.* **10**, 2272–2283 (2004).
53. Glinisky, G. V., Glinisky, A. B., Stephenson, A. J., Hoffman, R. M. & Gerald, W. L. Gene expression profiling predicts clinical outcome of prostate cancer. *J. Clin. Invest.* **113**, 913–923 (2004).
54. Hedenfalk, I. *et al.* Molecular classification of familial non-BRCA1/BRCA2 breast cancer. *Proc. Natl Acad. Sci. USA* **100**, 2532–2537 (2003).
55. van de Vijver, M. J. *et al.* A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* **347**, 1999–2009 (2002).
56. Lapointe, J. *et al.* Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proc. Natl Acad. Sci. USA* **101**, 811–816 (2004).
57. Lynch, T. J. *et al.* Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* **350**, 2129–2139 (2004).
58. Paez, J. G. *et al.* EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **304**, 1497–1500 (2004).
59. Pomeroy, S. L. *et al.* Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature* **415**, 436–442 (2002).
This important paper showed that medulloblastomas can be readily distinguished from other brain tumours, including morphological mimics, on the basis of gene-expression profiling. It also showed that a relatively small number of genes could accurately predict patient survival and response to therapy.
60. Fernandez-Tejedor, A. *et al.* Combining gene expression profiles and clinical parameters for risk stratification in medulloblastomas. *J. Clin. Oncol.* **22**, 994–998 (2004).
In a logical continuation of the work described in reference 59, the authors showed that gene-expression data can predict the outcome for patients with medulloblastoma, independent of clinical variables.
61. MacDonald, T. J. *et al.* Expression profiling of medulloblastoma: PDGFRA and the RAS/MAPK pathway as therapeutic targets for metastatic disease. *Nature Genet.* **29**, 143–152 (2001).
In this paper, the authors used cDNA-microarray analysis to identify a potentially crucial and therapeutically targetable pathway that might promote metastasis in medulloblastoma.
62. Oliver, T. G. *et al.* Transcriptional profiling of the sonic hedgehog response: a critical role for N-myc in proliferation of neuronal precursors. *Proc. Natl Acad. Sci. USA* **100**, 7331–7336 (2003).
63. Shai, R. *et al.* Gene expression profiling identifies molecular subtypes of gliomas. *Oncogene* **22**, 4918–4923 (2003).
64. Rickman, D. S. *et al.* Distinctive molecular profiles of high-grade and low-grade gliomas based on oligonucleotide microarray analysis. *Cancer Res.* **61**, 6885–6891 (2001).
65. Sallinen, S. L. *et al.* Identification of differentially expressed genes in human gliomas by DNA microarray and tissue chip techniques. *Cancer Res.* **60**, 6617–6622 (2000).
66. Huang, H. *et al.* Gene expression profiling of low-grade diffuse astrocytomas by cDNA arrays. *Cancer Res.* **60**, 6868–6874 (2000).
67. Fuller, G. N. *et al.* Reactivation of insulin-like growth factor binding protein 2 expression in glioblastoma multiforme: a revelation by parallel gene expression profiling. *Cancer Res.* **59**, 4228–4232 (1999).
68. Godard, S. *et al.* Classification of human astrocytic gliomas on the basis of gene expression: a correlated group of genes with angiogenic activity emerges as a strong predictor of subtypes. *Cancer Res.* **63**, 6613–6625 (2003).
69. Khatua, S. *et al.* Overexpression of the EGFR/FKBP12/HIF-2 α pathway identified in childhood astrocytomas by angiogenesis gene profiling. *Cancer Res.* **63**, 1865–1870 (2003).
70. Nutt, C. L. *et al.* Gene expression-based classification of malignant gliomas correlates better with survival than histological classification. *Cancer Res.* **63**, 1602–1607 (2003).
This paper showed that cDNA-microarray technology could potentially be used to address diagnostically confusing gliomas, and that transcriptional information contains more data about outcome than does pathological examination.
71. Mischel, P. S. *et al.* Identification of molecular subtypes of glioblastoma by gene expression profiling. *Oncogene* **22**, 2361–2373 (2003).
This paper highlights the potential of cDNA microarrays to detect molecular subsets of morphologically identical glioblastomas.
72. Freije, W. A. *et al.* Gene expression profiling of gliomas strongly predicts survival. *Cancer Res.* [in the press].
In this paper, the authors demonstrate that gene-expression-based grouping of malignant gliomas is a more powerful predictor of survival than pathological type, grade or age.
73. Scherf, U. *et al.* A gene expression database for the molecular pharmacology of cancer. *Nature Genet.* **24**, 236–244 (2000).

74. Staunton, J. E. *et al.* Chemosensitivity prediction by transcriptional profiling. *Proc. Natl Acad. Sci. USA* **98**, 10787–10792 (2001).
75. Wallqvist, A. *et al.* Mining the NCI screening database: explorations of agents involved in cell cycle regulation. *Prog. Cell Cycle Res.* **5**, 173–179 (2003).
76. Butte, A. J., Tamayo, P., Slonim, D., Golub, T. R. & Kohane, I. S. Discovering functional relationships between RNA expression and chemotherapeutic susceptibility using relevance networks. *Proc. Natl Acad. Sci. USA* **97**, 12182–12186 (2000).
77. Bohen, S. P. *et al.* Variation in gene expression patterns in follicular lymphoma and the response to rituximab. *Proc. Natl Acad. Sci. USA* **100**, 1926–1930 (2003).
78. Sotiriou, C. *et al.* Gene expression profiles derived from fine needle aspiration correlate with response to systemic chemotherapy in breast cancer. *Breast Cancer Res.* **4**, R3 (2002).
79. Cheok, M. H. *et al.* Treatment-specific changes in gene expression discriminate *in vivo* drug response in human leukemia cells. *Nature Genet.* **34**, 85–90 (2003).
80. Stegmaier, K. *et al.* Gene expression-based high-throughput screening (GE-HTS) and application to leukemia differentiation. *Nature Genet.* **36**, 257–263 (2004).
81. Reifenberger, G. & Louis, D. N. Oligodendroglioma: toward molecular definitions in diagnostic neuro-oncology. *J. Neuropathol. Exp. Neurol.* **62**, 111–126 (2003).
82. Gajjar, A. *et al.* Clinical, histopathologic, and molecular markers of prognosis: toward a new disease risk stratification system for medulloblastoma. *J. Clin. Oncol.* **22**, 984–993 (2004).
83. Liu, E. T. & Karuturi, K. R. Microarrays and clinical investigations. *N. Engl. J. Med.* **350**, 1595–1597 (2004).
84. Diehn, M., Eisen, M. B., Botstein, D. & Brown, P. O. Large-scale identification of secreted and membrane-associated gene products using DNA microarrays. *Nature Genet.* **25**, 58–62 (2000).
85. Tanwar, M. K., Gilbert, M. R. & Holland, E. C. Gene expression microarray analysis reveals YKL-40 to be a potential serum marker for malignant character in human glioma. *Cancer Res.* **62**, 4364–4368 (2002).
86. Herschman, H. R. Molecular imaging: looking at problems, seeing solutions. *Science* **302**, 605–608 (2003).
87. Bergmann, S., Ihmels, J. & Barkai, N. Similarities and differences in genome-wide expression data of six organisms. *PLoS Biol.* **2**, E9 (2004).
88. Stuart, J. M., Segal, E., Koller, D. & Kim, S. K. A gene-coexpression network for global discovery of conserved genetic modules. *Science* **302**, 249–255 (2003).
89. Chang, H. Y. *et al.* Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds. *PLoS Biol.* **2**, E7 (2004).
90. Whitfield, M. L. *et al.* Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol. Biol. Cell* **13**, 1977–2000 (2002).
91. Ficenc, D. *et al.* Computational knowledge integration in biopharmaceutical research. *Brief. Bioinform.* **4**, 260–278 (2003).
92. Buckingham, S. Bioinformatics: programmed for success. *Nature* **425**, 209–215 (2003).
93. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
94. Angelastro, J. M. *et al.* Identification of diverse nerve growth factor-regulated genes by serial analysis of gene expression (SAGE) profiling. *Proc. Natl Acad. Sci. USA* **97**, 10424–10429 (2000).
95. Liu, D. X. & Greene, L. A. Regulation of neuronal survival and death by E2F-dependent gene repression and derepression. *Neuron* **32**, 425–438 (2001).
96. Hartwell, L. H., Hopfield, J. J., Leibler, S. & Murray, A. W. From molecular to modular cell biology. *Nature* **402** (Suppl.), C47–C52 (1999).
97. Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N. & Barabasi, A. L. The large-scale organization of metabolic networks. *Nature* **407**, 651–654 (2000).
98. Alon, U. Biological networks: the tinkerer as an engineer. *Science* **301**, 1866–1867 (2003).
99. Agrawal, H. Extreme self-organization in networks constructed from gene expression data. *Phys. Rev. Lett.* **89**, 268702 (2002).
100. Chen, J. *et al.* The PEPR GeneChip data warehouse, and implementation of a dynamic time series query tool (SGOT) with graphical interface. *Nucleic Acids Res.* **32**, D578–D581 (2004).
101. Ramaswamy, S. & Golub, T. R. DNA microarrays in clinical oncology. *J. Clin. Oncol.* **20**, 1932–1941 (2002).

Acknowledgements

We wish to thank S. Horvath, M. Carlson, W. Freije and Z. Fang for their contribution to this work, and we thank D. Geschwind and H. Kornblum for helpful discussions about this review. The authors are supported by the NINDS and NCI and by Accelerate Brain Cancer Cure, the Packard Foundation, the Harry Allgauer Foundation through The Doris R. Ullmann Fund for Brain Tumor Research Technologies, the Henry E. Singleton Brain Tumor Endowment, Art of the Brain and the Ziering Family Foundation in memory of Sigi Ziering.

Competing interests statement

The authors declare no competing financial interests.

Online links

DATABASES

The following terms in this article are linked online to:

Entrez Gene:

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene>
EGFR | ErbB2 | FKBP12 | p53 | SHH | TRKC | YKL40

FURTHER INFORMATION

Encyclopedia of Life Sciences: <http://www.els.net/>

Brain cancers

Mischel homepage: <http://www.neuroscience.ucla.edu/faculty-page.asp?key=729>

Cloughesy homepage: <http://www.neurology.ucla.edu/faculty/CloughesyT.htm>

Nelson homepage: <http://research.arc2.ucla.edu/Departments/humgen/faculty.cfm?FacultyKey=144>

National Cancer Institute:

<http://genome-www.stanford.edu/nci60/index.shtml>

Access to this interactive links box is free online.