NATURE|Vol 452|3 April 2008|doi:10.1038/nature06915

Enabling personalized cancer medicine through analysis of gene-expression patterns

Laura J. van 't Veer[1,2,3] & René Bernards[1,3,4,5]

Therapies for patients with cancer have changed gradually over the past decade, moving away from the administration of broadly acting cytotoxic drugs towards the use of more-specific therapies that are targeted to each tumour. To facilitate this shift, tests need to be developed to identify those individuals who require therapy and those who are most likely to benefit from certain therapies. In particular, tests that predict the clinical outcome for patients on the basis of the genes expressed by their tumours are likely to increasingly affect patient management, heralding a new era of personalized medicine.

The conventional approach to cancer therapy has been to provide treatment according to the organ or tissue in which the cancer originates. This approach was appropriate when there was only a rudimentary understanding of the molecular origins of cancer and the different intracellular signalling pathways that are perturbed in the various types of cancer (such as in breast cancer or lung cancer). In the past two to three decades, however, the genetic events that lead to cancer have been dissected, and it has become clear that cancer develops as a result of multiple genetic defects and that individuals with the same type of cancer often have dissimilar genetic defects in their tumours. This finding explains why patients who seem to have similar cancers respond in a heterogeneous manner to anticancer agents and shows clearly the huge obstacle to providing effective treatments for cancer.

In the past decade, cancer therapy has slowly but steadily begun to shift from a 'one size fits all' approach to a more personalized approach, in which each patient is treated according to the specific genetic defects in the tumour. Such an individualized approach requires the discovery and development of biomarkers (biological indicators) that help doctors to decide which patients to treat (known as prognostic biomarkers) and which therapy is most likely to be effective for a given patient (known as predictive biomarkers). More specifically, prognostic biomarkers predict the clinical outcome for a patient if no anticancer drugs are administered, whereas predictive biomarkers predict the outcome of a specific therapy for a patient. An example of why such biomarkers are needed to improve patient management is that, for some tumours, resection (that is, surgical removal) of the primary tumour might be curative; therefore, systemic therapy to eliminate any remaining tumour cells (also known as adjuvant therapy) would not be needed. By contrast, for more malignant primary tumours, aggressive systemic therapy, often chemotherapy, might be required after resection, in order to reduce the risk of the tumour recurring. However, the distinction between these is often unclear, so prognostic biomarkers that enable the likelihood of recurrence to be determined are urgently needed in the clinic.

In the case of breast cancer, large meta-analyses have shown that recurrence is likely in 20–30% of young women with early-stage (lymph-node-negative) breast cancer who undergo only surgery and localized radiation treatment[1]. But, in the United States, 85–95% of women with this type of cancer receive adjuvant chemotherapy, mostly because conventional clinicopathological parameters fail to identify reliably those patients who are likely to relapse. Therefore, 55–75% of women with early-stage breast cancer in the United States undergo a toxic therapy from which they will not benefit but will experience the side effects. So it is not surprising that the initial attempts to discover clinically relevant prognostic biomarkers have focused on breast cancer (discussed later).

The advent of DNA-microarray technology in the 1990s (refs 2, 3) made it possible to assess the expression of tens of thousands of genes in a single experiment. Systematic analysis of the gene-expression patterns of tumour samples enabled researchers to identify characteristic expression patterns of groups of genes that are associated with specific tumour traits. These patterns are known as gene-expression signatures.

In this review, we focus on gene-expression signatures as a new class of molecular diagnostic test for cancer. We discuss pitfalls in the discovery of gene-expression signatures, how such signatures can be used to develop clinically relevant tests and how these tests are likely to affect patient management and drug development in the future.

Building gene-expression profiles

The massive parallel quantification of messenger RNA abundance that is possible using DNA-microarray technology has enabled genome-wide gene-expression data to be collected for large numbers of biological specimens. Collecting this unprecedented amount of data (at least for biologists) has necessitated the development of new tools to analyse the large data sets. In principle, to find connections between the patterns of gene expression by tumour cells and the behaviour of these cells, there are three approaches: the data-driven approach, the knowledge-driven approach and the model-driven approach.

The most straightforward is the data-driven approach, in which a genome-wide analysis of gene expression is carried out, and then correlates between patterns of gene expression and certain tumour traits are searched for. The strength of this approach is that it is unbiased: there are no assumptions about which genes are likely to be involved in the process of interest. For example, in a data-driven study of the prognosis of patients with breast cancer, little was known about the function of 15 of the 70 genes that were found to constitute a prognostic gene-expression signature[4]. A drawback of this approach is that the outcome relies solely on the quality of the data (and the samples).

By contrast, using the knowledge-driven approach, genes that are thought to be relevant to a particular cancer trait are selected on the basis

[1]Agendia BV, Louwesweg 6, 1066 EC Amsterdam, the Netherlands. [2]Division of Pathology, [3]Cancer Genomics Centre, [4]Centre for Biomedical Genetics, [5]Division of Molecular Carcinogenesis, the Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, the Netherlands.

of the scientific literature. This approach is often used when only formalin-fixed paraffin-embedded tumour tissue is available. The RNA isolated from such tissue is fragmented, and such poor-quality RNA is far from ideal for genome-wide quantitative analysis using DNA microarrays. It can, however, be analysed by PCR with reverse transcription. In this approach, mRNA abundance is measured by using unique pairs of oligonucleotide primers that correspond to each gene, a labour-intensive process that precludes genome-scale analysis of gene expression. Hence, in studies involving formalin-fixed paraffin-embedded material, sets of 'likely suspect' genes are often tested. A drawback of this approach is that the outcome is only as good as the state of knowledge: genes that are not known to be involved in a process cannot be considered.

In the model-driven approach, the transcriptional responses of cells after exposure to specific stimuli are used to predict tumour traits. For example, a gene-expression signature for wound healing has been used to predict the survival of individuals with breast cancer[5]. Similarly, gene-expression signatures that reflect the activation of specific oncogenic pathways have been used both to determine prognosis and to predict responses to anticancer drugs[6,7]. This approach has the drawback that the experimental model used might not accurately reflect the processes that occur in tumours.

A comparison of the data-driven approach and the knowledge-driven approach applied to a similar problem sheds more light on the advantages and disadvantages of these approaches. If the presence of a certain transcription factor is known to affect the prognosis of individuals with a particular cancer, then in the knowledge-driven approach to building a prognostic gene-expression signature, the gene encoding this transcription factor would be incorporated into a prognostic signature. In some cancers, however, this gene might be expressed, but its product non-functional (for example, as a result of a missense mutation). For this reason, in a data-driven approach, targets downstream of a transcription factor of interest are often found to be distinguishing features, rather than the gene encoding the transcription factor itself, because the expression of these targets provides more relevant information on the activity of the transcription factor. As an example, a 16-gene signature for the prognosis of breast cancer that was derived from 250 'candidate' genes selected on the basis of published studies includes *ESR1,* which encodes oestrogen receptor-α (ER-α; a transcription factor that is expressed by most breast cancers)[8]. By contrast, a set of 70 genes for assessing breast-cancer prognosis (discussed earlier) that was identified by a data-driven approach does not include *ESR1* itself but includes several genes that are targets of ER-α[4].

After large-scale gene-expression data sets have been collected, there are two fundamentally different ways to analyse them. One approach is to ask whether in a group of samples (for example, tissues from individuals diagnosed with a particular cancer), there are subgroups (or clusters) of samples with similar gene-expression patterns. These similarities in gene expression can be used to classify a cancer into subtypes that could have similarities in biological behaviour. This type of data analysis is called hierarchical clustering or unsupervised classification[9], and it has the advantage that additional clinical data are not required. For example, hierarchical clustering of breast-cancer specimens[10,11] identified five naturally occurring subtypes (referred to as 'intrinsic subtypes') of breast cancer that had not been observed previously. Some of these intrinsic subtypes differ markedly in their aggressiveness; for example, the prognosis for individuals with the subtype called basal-like breast cancer is significantly worse than that for individuals with another subtype, luminal-type breast cancer[11,12].

The second approach to data analysis is known as supervised classification. Samples are divided into groups that are known to have different clinical end points (for example, recurrence versus no recurrence, and drug response versus no drug response), and genes that can correctly identify the distinct groups are searched for (Fig. 1). One set of tumours (called the training set) is used to identify the genes that discriminate between the groups — the gene-expression signature — and then a second, independent, set of tumours (called the validation set) is used to test how well these genes can classify samples that have not been grouped. The situations in which these approaches have been used and their relative value are discussed in the next section.

Prognostic profiles

Initial studies of gene-expression profiling of cancer used simple hierarchical clustering to identify subtypes among apparently similar cancers. A landmark early study identified two distinct subtypes of diffuse large B-cell lymphoma (DLBCL) — germinal-centre B-like DLBCL, and activated B-like DLBCL — for which the overall survival of patients significantly differs[13]. This was the first in an avalanche of publications showing that the molecular classification of tumours on the basis of

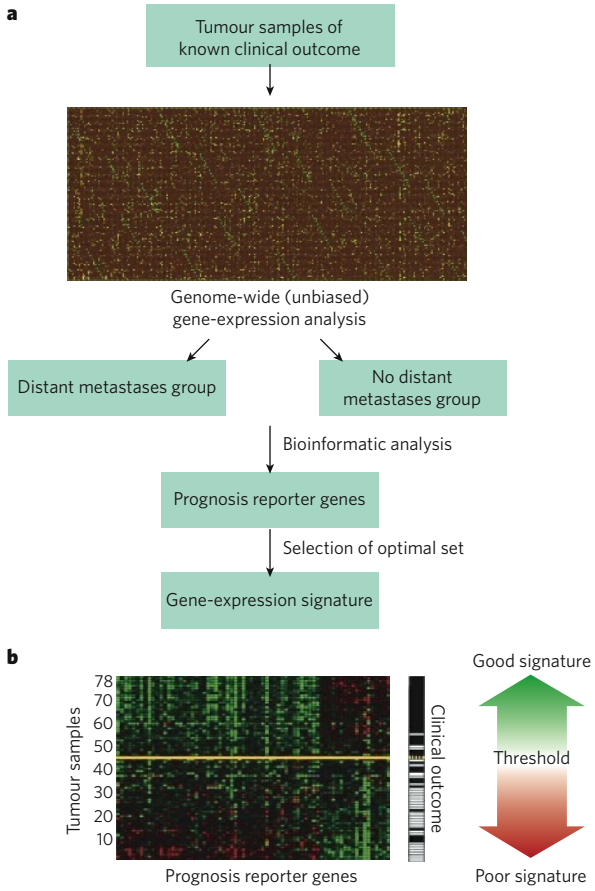

Figure 1 | Predicting disease outcome by using complex gene-expression tests. a, Generating a prognostic gene-expression signature by using supervised classification. The gene expression of cells in a set of tumours of known clinical outcome is analysed by using whole-genome microarrays. Colours indicate the level of expression for each gene: red, gene is more active than the average for tumours of this type; green, gene is less active than average; yellow, gene is equally active; and, black, gene is not expressed. The results for each tumour sample are then classified into two categories: tumours with a good outcome (no distant metastases developed), and tumours with a poor outcome (distant metastases developed). Using bioinformatic analysis, genes whose expression is significantly correlated with disease outcome are identified, and these are known as prognosis reporter genes. An optimal set of genes is then selected from the prognosis reporter genes by using bioinformatic algorithms, and the pattern of expression of this multigene set is known as a gene-expression signature (or classifier). **b**, The gene-expression signature generated in **a** is shown as a 'heat map'. The expression of the 70 prognosis reporter genes selected as the optimal set (vertical columns) is shown for 78 tumours (horizontal lines). So each of the 70 × 78 intersection points of the heat map shows how a particular gene is expressed in a given tumour. A red spot indicates that the gene is expressed at a higher level than average for tumours of this type, and a green spot that the gene is expressed at a lower level. The outcome of the disease is shown on the right: white indicates metastasis; black indicates no metastasis; and yellow indicates the threshold for metastasis. (Panel adapted, with permission, from ref. 4.)

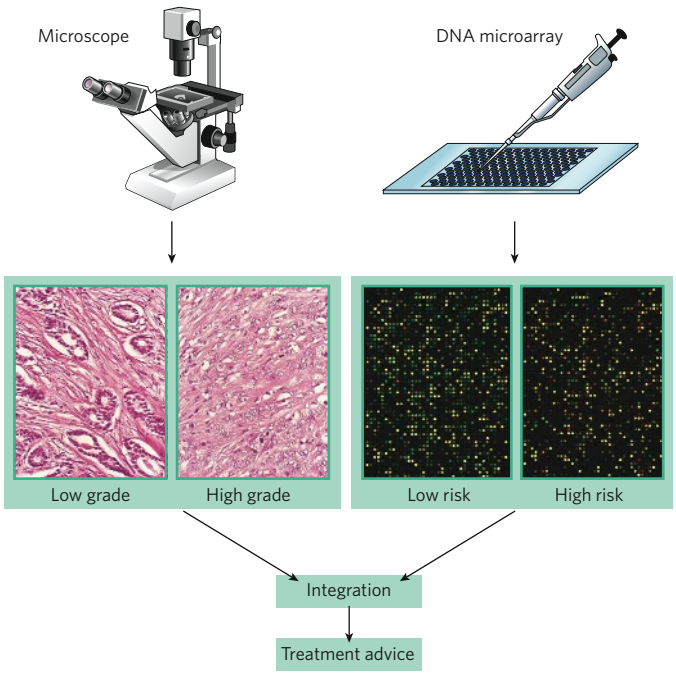

Figure 2 | Conventional and molecular diagnostic testing for cancer. Conventional diagnostic tests rely heavily on morphological criteria (that is, properties of cells that can be observed microscopically in tissue sections from a tumour) to judge the aggressiveness of cancer, a process known as grading. As an example, the differences between a 'low grade' adenocarcinoma of the breast (for which patients generally have a favourable prognosis) and a 'high grade' adenocarcinoma of the breast (for which patients have a worse prognosis) can be observed from the images. More recently, multigene-expression tests (also known as *in vitro* diagnostic multivariate index assays, IVDMIAs) have been shown to be powerful tools for predicting disease outcome and have become subject to scrutiny by the US Food and Drug Administration. As an example, a molecular description of low-risk and high-risk adenocarcinoma of the breast, as judged by a gene-expression signature consisting of 70 genes, is shown. One current challenge is how to integrate the knowledge obtained from these conventional tests and molecular diagnostic tests into a single recommendation for the oncologist treating the patient.

gene expression can identify previously undetected and clinically significant subtypes of cancer. Even though the gene-expression signatures found in these studies uncovered aspects of tumour-cell biology that had gone unnoticed, these studies were not designed to find signatures identifying subtypes of cancer that result in different prognoses. Their main purpose was to establish a molecular classification of tumours on the basis of their gene-expression patterns.

In a study that was designed to find a prognostic gene-expression signature for breast cancer, gene-expression data from breast-tumour samples of known clinical outcome were analysed by using supervised classification[4]. This study yielded the previously mentioned 70-gene signature for breast-cancer prognosis. Remarkably, when another research group independently used a similar approach[14], a 76-gene signature was identified, but this had only three genes in common with the 70-gene signature. This finding was interpreted by some to indicate that such gene-expression signatures are highly unstable[15]. But, because both of these gene-expression signatures could be independently validated in large groups of patients[16,17], it is more likely that the two signatures use different genes to monitor the same biological processes, an idea for which there is supporting evidence[18,19]. Using similar strategies, gene-expression signatures that assess the risk of recurrence of non-small-cell lung cancer[20] and several other cancer types have also been established.

A combination of a knowledge-based approach and a data-driven approach was used to identify the previously mentioned 16-gene signature for breast-cancer prognosis. First, a set of 250 candidate genes was identified and, after analysing the expression of the candidate genes in about 400 tumours, 16 genes were selected from this set[8]. Of all the gene-expression signatures for cancer that have been identified, only three are commercially available: the 70-gene signature for breast-cancer prognosis is available under the name MammaPrint (Agendia); the 16-gene signature as Oncotype DX (Genomic Health); and a 2-gene signature[21], which has recently been released, under the name the H/I test (AviaraDx). Now that such tests can be used in the clinic, one of the major challenges facing oncologists and pathologists is how to integrate the information obtained from conventional tests with that from these molecular tests (Fig. 2).

An important question when considering the development of new tests is whether prognostic gene-expression profiles are independent of the molecularly defined subtypes of cancer. As mentioned earlier, individuals with basal-like breast cancers generally have a worse prognosis than those with luminal-type breast cancers. Should a separate prognostic signature therefore be made for the basal-like subtype and the luminal-type subtype, or is a signature based on supervised classification of a diverse panel of breast tumours equally powerful for assessing the prognosis of patients with either subtype? Our studies of tumours from 295 patients with breast cancer indicate that within the group of patients who were predicted to be at high risk of relapse by using the 70-gene signature, the patients with the basal-like subtype have a comparable outcome to those with the luminal-type subtype (L.J.v.V., unpublished observations). This finding suggests that the prognostic value of the molecular subtyping of cancer, as was carried out in the initial gene-expression profiling studies, has been surpassed by that of prognostic gene-expression signatures such as the 70-gene signature and that the underlying molecular subtype does not contain additional crucial information for determining a patient's prognosis. It is possible, however, that patients with the various molecular subtypes of cancer respond differently to particular therapies.

Predictive profiles

In the area of predicting responses to particular therapies, gene-expression profiling studies have not yet delivered on their promise. It seems that responses to anticancer drugs are more difficult to predict by using molecular tests than prognosis is. One of the main reasons for this difficulty is that resistance to anticancer agents can result from a variety of mechanisms. Consequently, there might not be a gene-expression profile that correlates with resistance to a certain drug. This is true for both resistance to conventional chemotherapeutic agents, which are often pleiotropic, and to newer 'targeted therapies', which affect specific components of signalling pathways. It should be pointed out, however, that the differences between these types of anticancer drug are less marked than is generally assumed: some of the conventional chemotherapeutic agents (such as topoisomerase inhibitors) target specific enzymes, whereas some of the small-molecule-based targeted therapies (such as imatinib mesylate and lapatinib) are directed against more than one enzyme. In addition, resistance to drugs might result from subtle mutations that do not cause gross changes in gene expression, a process that is therefore undetectable by gene-expression profiling.

Another important impediment to the discovery of predictive gene-expression signatures is that, for genome-wide gene-expression studies, large numbers of tumour samples are required (of the order of 100), to reduce the probability that associations between gene-expression signatures and therapy outcomes are spurious. An additional consideration is that drug-response gene-expression profiles cannot be constructed from tumour samples from patients who have undergone adjuvant therapy. Patients who do not relapse after adjuvant therapy could have had a tumour with either a 'good outcome' (that is, a tumour that did not metastasize, for which adjuvant therapy was unnecessary) or a 'poor outcome' (that is, a tumour that metastasized and, in this case, responded to therapy). Hence, prediction cannot be separated from prognosis in this type of study. In principle, tumour samples from patients who are known to have metastatic disease could be used to develop drug-response profiles (all are, by definition, of the poor-outcome type), but such samples are not readily available, because patients who are treated for metastatic cancer

do not usually have their tumours biopsied first. Furthermore, patients with metastatic cancer often receive combination therapies, making it difficult to determine which drug a patient responded to. One way to circumvent some of these issues is to use gene-expression profiling in a 'neo-adjuvant' setting, in which patients are treated systemically before the resection of a primary tumour (often carried out when the tumour is large). Short-term responses to anticancer drugs can be determined in this setting by using imaging technologies, and primary tumours can often be sampled by needle biopsy[22,23].

As a result of all of these factors, few drug-response gene-expression signatures have been published, and those that have are not clinically useful in their current form because they have not been properly validated against an independent series of tumours. It should be noted, however, that many clinical trials exploring the neo-adjuvant approach are underway, a well-documented example being the I-SPY trial, sponsored by the National Cancer Institute (http://tr.nci.nih.gov/iSpy).

A more general problem with predictive gene-expression signatures is that doctors often do not have a suitable replacement for the first-line therapy and are unlikely to withhold treatment entirely just because a biomarker indicates that the treatment probably will not be effective. Thus, the sensitivity (that is, the percentage of patients with a certain trait that test positive) and the specificity (that is, the percentage of patients without a certain trait that test negative) of predictive gene-expression signatures will need to improve markedly before they will be useful in the clinic.

Potential short cuts to predictive biomarker development

Given the problems with generating predictive gene-expression profiles, several groups have attempted to generate such profiles through identifying genes or pathways that potentially affect how a cell responds to a drug, often by using models based on cell lines. A small number of human tumour samples can then be tested for the expression of these *in vitro*-generated sets of candidate genes. The approaches that have been explored so far are discussed in this section and illustrated in Fig. 3.

In vitro-generated predictive profiles using cell-line models

Gene-expression profiling of a series of human cancer cell lines with known drug sensitivity has been used to identify patterns of gene expression that correlate with responses to drugs *in vitro*[24–26]. These studies used a heterogeneous set of cell lines (a panel of 60 human cancer cell lines of different tissue origins, called NCI-60) with known sensitivity profiles to several anticancer drugs. The gene-expression profiles generated are therefore designed to be independent of the tissue of origin, but it is questionable whether such tissue-independent profiles are reliable predictors of drug responses. The validity of this approach has been challenged recently[27]; however, an independent validation of such *in vitro*-generated gene-expression signatures was also published around the same time[28]. It might be more fruitful to carry out such gene-expression profiling studies on panels of cell lines derived from a single type of cancer. Using a more homogeneous panel of cell lines might allow the identification of drug-resistance mechanisms that are missed when a heterogeneous set of cancer cell lines is used. Indeed, for breast cancer, the first results obtained by using this approach seem to be promising[29]. But there is no direct evidence that using a more homogeneous cell-line panel yields a better profile. An example of the general approach is shown in Fig. 3a.

Signalling pathways

In the past, the most effective therapies for cancer were based on empirically derived evidence from large clinical studies. This approach has resulted in a standard therapeutic regimen in which drug combination A is given for breast cancer, for example, but combination B is given for lung cancer, and so on. The continued use of such standard protocols to treat each type of cancer ignores one of the most important lessons from the past two decades of molecular genetics research on cancer: namely, that each tumour has a complex and unique set of genetic alterations that drive the oncogenic proliferation of the cells. Nevertheless, the fact that distinct drug combinations show activity against specific types of cancer reflects differences in the molecular pathways that are predominantly

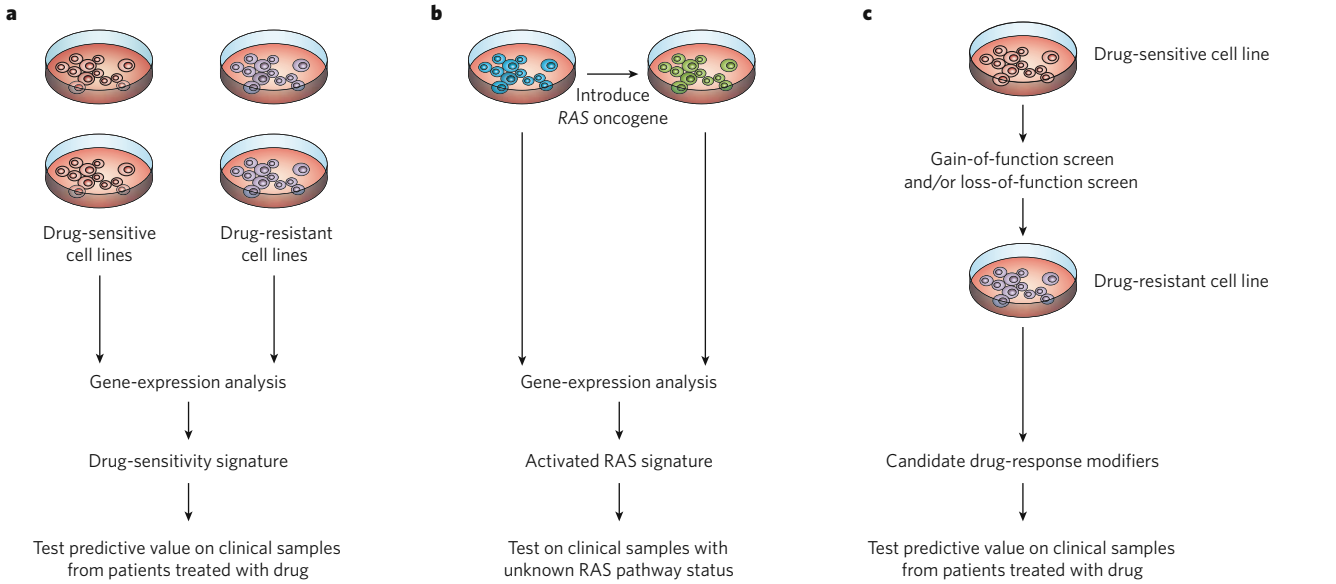

Figure 3 | Short cuts to the development of drug-response biomarkers. **a**, Collections of tumour cell lines of known drug sensitivity can be used to build gene-expression signatures that discriminate between sensitive and resistant cell lines. Such *in vitro*-generated drug-sensitivity signatures can be validated on tumour samples from patients treated with the same drugs. **b**, Gene-expression signatures for signalling pathways can be constructed *in vitro* by introducing the gene of interest (a mutant *RAS* gene that is constitutively active in the example here) into tumour cell lines and studying the effect of the presence of the oncogene on genome-wide gene expression. Tumour samples for which the status of the RAS pathway is unknown can then be assessed by comparing their gene-expression patterns with that of the 'activated RAS pathway' identified *in vitro*. If a drug that targets the RAS pathway is available, then similarity between the gene-expression profile of the tumour and a RAS pathway signature could be used to guide the choice of therapy. **c**, Functional genetic approaches can be used *in vitro* to uncover which genes can contribute to drug resistance in tumour cell lines. More specifically, using these approaches — genome-scale gain-of-function screens or RNA-interference-based loss-of-function screens — full-length complementary DNAs or small interfering RNAs are introduced to change the abundance of gene products, turning drug-sensitive cell lines into drug-resistant cell lines. The predictive ability of the genes that are candidates for modifying drug responses can then be examined by assessing their expression levels in a relatively small number of clinical samples from patients treated with the same drug.

deregulated in these cancers. As such, these empirically derived treatment protocols can be viewed as primitive forms of targeted therapy, supporting the idea of stratifying tumours mainly according to their signalling-pathway perturbations rather than their tissue of origin. The challenge in developing molecular tests to predict drug responses is to identify the altered pathways in each tumour so that each patient receives the optimal targeted therapy. To facilitate this, there needs to be a shift away from describing cancers according to their tissue and cell type of origin (for example, adenocarcinoma of the breast) towards describing them by the main pathways that drive tumour-cell proliferation (for example, phosphatidylinositol-3-OH kinase (PI(3)K)-driven cancer or WNT-driven cancer).

Indeed, for more than 30 years, it is has been routine practice to classify breast tumours as ER positive or ER negative, and this classification is commonly used to decide a patient's eligibility for hormonal therapy, which is one of the main forms of adjuvant therapy for breast cancer (involving inhibition of the growth-stimulating effects of the female hormone oestrogen on the cancer). Therefore, neither the concept of targeted therapy nor the concept of naming tumours by the pathways that drive their proliferation is new. The main reason why it has not been adopted on a larger scale is because the tools to measure the activation of various signalling pathways have been lacking. Moreover, for most cancers, it is unclear which pathway drives the oncogenic process.

One important recent advance that should aid the development of tests to predict drug responses is the finding that activation of a signalling pathway leads to characteristic changes in gene expression, which can be identified using gene-expression analysis. Joseph Nevins's research group[6,7,30], for example, has established several 'pathway gene-expression signatures' (the RAS signature, SRC signature, MYC signature and E2F signature) by experimentally manipulating cell lines to activate certain pathways, and these signatures can be used both to determine prognosis and to select specific therapies that target the activated pathways. This approach is illustrated in Fig. 3b. Whether such pathway signatures are similar in different tissue types remains to be investigated. It might be necessary to determine pathway signatures for each cancer type so that the optimal treatment can be chosen for each patient.

Such cell-line models are useful for building pathway gene-expression signatures, but they are often not an accurate representation of the *in vivo* situation in human cancer. To overcome this problem, Lao Saal *et al.*[31] analysed human breast-cancer samples by using immunohistochemistry to identify a set of primary tumours that expressed the tumour-suppressor gene *PTEN* (phosphatase and tensin homologue, which is a component of the PI(3)K-signalling pathway) at a low level and then linked this 'PTEN-low' phenotype to a specific gene-expression signature of these tumours. Searching for this signature was shown to be a more sensitive way to detect tumours in which the PTEN–PI(3)K pathway is activated than using immunohistochemistry to detect PTEN itself (because events other than loss of, or a reduction in, *PTEN* expression can also modulate the pathway). Such a 'pathway-integrative' signature might therefore be an effective tool to guide the use of therapies that target this pathway[31].

Alternatively, instead of starting from a cancer specimen, the process can be inverted: first, a database of gene-expression patterns that result from well-defined perturbations of specific pathways can be established, and then a test gene-expression data set (for example, from a given cancer specimen) can be assessed for how well it matches any of the gene-expression profiles in the database[32,33]. Such gene-expression compendia (also known as connectivity maps) can be powerful tools for deciding how to treat patients. This approach uncovered a similarity between the effects of exposure to heat-shock protein 90 inhibitors and inhibition of androgen-receptor signalling, a signalling pathway that is central to prostate cancer[34]. Similarly (and perhaps of greater clinical relevance), searching a database of gene-expression signatures induced in response to drug treatment revealed that rapamycin (which inhibits mTOR) induces a signature that overlaps with a signature for sensitivity to glucocorticoids that is found in samples from patients with acute lymphoblastic leukaemia (ALL)[35]. Indeed, when tested, rapamycin was able to induce glucocorticoid sensitivity in ALL samples. These data underscore the power of this type of approach to identify connections between apparently unrelated biological perturbations, which can lead to important insights into the factors that mediate drug sensitivity.

Functional genetic approaches

Another way to identify genes that can serve as biomarkers to predict drug responses is to take suitable cell-line models and use genetic screens to find genes or pathways whose altered activity modulates sensitivity to anticancer drugs *in vitro* (Fig. 3c). After such candidate genes have been identified, their expression can be measured in tumour samples from patients with cancer who have been treated with the same drug, and the pattern is then correlated with resistance to the drug. The validation of a defined set of candidate genes requires fewer tumour samples than an unbiased (genome-wide) search for predictive biomarkers, thereby bypassing one of the largest bottlenecks in the discovery of robust biomarkers: the availability of suitable tumour samples.

This approach might be useful to improve responses to trastuzumab (Herceptin), for example. This antibody-based drug is effective, particularly in combination with chemotherapy, for treating patients with breast cancers that produce the epidermal growth-factor receptor ERBB2 (also known as HER2 and NEU). But less than 35% of patients with *ERBB2*-expressing metastatic breast cancer respond when treated with trastuzumab alone, and it is largely unclear why this is the case[36]. Moreover, gene-expression studies of samples from patient who were treated in the neo-adjuvant setting with trastuzumab have failed to find a clear drug-response profile[22]. To uncover trastuzumab-response modifiers, we have used loss-of-function RNA-interference screens, in which gene activity is suppressed on a large scale by using short duplex RNAs, to identify genes that confer resistance to trastuzumab in *ERBB2*-expressing breast-cancer cell lines. Of a set of 8,000 genes tested, only loss of expression of the tumour-suppressor gene *PTEN* conferred resistance to trastuzumab[37], a finding in close agreement with an earlier study implicating this gene in resistance to trastuzumab[38]. Gene-expression signatures for the loss of *PTEN* expression in breast cancer are available[31], and these could be suitable for predicting responses to trastuzumab, as well as to other therapies (such as gefitinib and lapatinib) targeted to ERBB2 or the various members of the epidermal growth-factor receptor family.

Using a similar functional genetic approach, Charles Swanton *et al.*[39] identified regulators of mitotic arrest and ceramide metabolism as determinants of sensitivity to paclitaxel and other chemotherapeutic drugs, and Angelique Whitehurst *et al.*[40] identified genes whose suppression increases the efficacy of paclitaxel against non-small-cell lung cancer. Such synthetic lethal interactions (that is, a combination of two non-lethal events that together result in cell death) could be used to suggest valuable combination therapies for cancer. Along similar lines, two groups recently found that breast tumours with a mutation in *BRCA1* or *BRCA2* are hypersensitive to inhibitors of poly(ADP-ribose) polymerase 1 (PARP1)[41,42]. Gene-expression signatures that identify tumours with defects in the BRCA1 pathway are available[4], so these could be used to determine which patients will respond to PARP1 inhibitors.

Implementing gene-expression profiles in the clinic

Translating biomarker research into clinically useful tests has often been a frustrating activity. Many of the biomarkers identified in the initial tumour studies, which were retrospective, failed to be validated in subsequent studies. One of the main reasons for these failures was that early biomarker discovery was knowledge-driven, but the knowledge was often of poor quality. By contrast, some of the more recent gene-expression signatures were derived from large data-driven, genome-wide studies with excellent data quality, so these biomarkers are far more likely to be validated than previously identified biomarkers from knowledge-driven studies.

Both the US Food and Drug Administration (FDA) and the medical community have recognized that multigene signatures are better biomarkers than single molecules, so why are so few gene-expression signatures available in the clinic? First, on the basis of past failures, doctors are often reluctant to use biomarkers that have been validated only

by retrospective studies; they insist on validation by prospective studies before biomarkers are used in routine clinical practice. A second impediment is that DNA-microarray technology was initially not very robust (at least around the year 2000) and, to many scientists and doctors, it still has a poor reputation, which has been unfounded since industry became involved in production[43]. Third, the correct regulatory path for using multigene tests in clinical practice is unclear. For two of the three multigene tests that are commercially available, large prospective validation studies are in progress: the study TAILORx (ref. 44) is validating the 16-gene signature marketed as Oncotype DX (ref. 8); and the study MINDACT (ref. 45) is validating the 70-gene signature marketed as MammaPrint[4,46]. But the results of these large studies, which require thousands of patients, are at least five years away, and the medical community seems to be divided about when to start using such tests in routine clinical practice[47]. This dilemma was exacerbated recently when the FDA cleared the first multigene assay (MammaPrint) on the basis of only retrospective validation (http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfPMN/PMNSimpleSearch.cfm?db=PMN&id=K062694). It should also be considered that the cost of large prospective trials is prohibitive for developing the 'average' molecular diagnostic test. Last, it is often forgotten that purely prognostic gene-expression profiles based on archival tumour samples from patients who were not given adjuvant therapy (which, at least for early-stage breast cancer, is a relatively recent inclusion in the treatment regimen) cannot be validated in current prospective clinical trials, because most patients now receive some form of adjuvant therapy. And, as mentioned earlier, adjuvant therapy 'contaminates' the validation of a purely prognostic profile, so clinical studies in which adjuvant therapy is used are suboptimal for the validation of prognostic gene-expression signatures. For all of these reasons, we expect that robust retrospective studies will become the norm for validating prognostic gene-expression signatures.

The FDA's involvement in regulating molecular diagnostic tests is contested by some. But it seems logical for several reasons that the FDA take an active role in the market approval of these tests. First, molecular diagnostic tests are likely to increasingly affect patient management as tests begin to show that treatment with a particular drug will benefit groups of patients whose tumours have a certain gene-expression pattern. However, if the predictions of such tests are incorrect, then patients could be given an inappropriate drug. Second, these molecular tests are complex, and their accuracy and reproducibility are unlikely to be understood by most doctors. It is therefore improbable that the 'average' doctor would know whether a gene-expression test reliably predicts responses to an anticancer drug or reliably assigns a tumour as having a high risk of recurrence. For these reasons, the FDA has expressed its intention to regulate these molecular diagnostic tests similarly to how it regulates anticancer drugs.

In September 2006, the FDA issued a draft guidance document for the use of a new type of molecular diagnostic test called an *in vitro* diagnostic multivariate index assay (IVDMIA). These tests use complex mathematical algorithms to interpret large amounts of gene- or protein-expression data for the purpose of guiding medical decision-making. After a period open for comment by the diagnostic industry and clinical laboratories, the FDA issued a second version of this draft document in July 2007 (http://www.fda.gov/cdrh/oivd/guidance/1610.pdf). It is expected that the final version of this document will become the basis for regulatory enforcement of the molecular diagnostic industry in the United States. Unfortunately, Europe (once again) lags behind. The European counterpart of the FDA, the European Medicines Agency, has not announced any plans to regulate complex molecular diagnostic tests other than requiring the Conformité Européene (CE) marking (a generic mark that is mandatory for many products on the market in the European Union, indicating that manufacturers have conformed to EU legislation). The regulation of molecular diagnostic tests is viewed by some as being a burden and impeding innovation. Others, however, think that the FDA's approval of molecular diagnostic tests will improve the acceptance of these tests in the clinic. Acceptance might not be far off, however, considering that one in eight women with early-stage breast cancer in the United States is likely to have a molecular diagnostic gene-expression test in 2008 (based on sales of Oncotype DX). It seems that not all doctors are going to wait for the results of the prospective validation studies that are underway.

Future perspectives

The intention of the FDA to regulate molecular diagnostic tests reflects the increased effect that these tests are likely to have on patient management. Traditionally, diagnostic tests for cancer have been carried out in local hospitals. The quality control for these tests has been poor though. A large study that investigated interlaboratory variance in the immunohistochemical detection of ER in breast cancer, across 200 laboratories in 26 countries, showed a false-negative rate of 30–60% (ref. 48), yet this type of test is used routinely to decide whether a patient should receive hormonal therapy. By contrast, a recent study investigating interlaboratory variance of a DNA-microarray-based test for breast-cancer prognosis found an extremely high concordance between laboratories[49], which is consistent with other studies of the reproducibility of using microarray platforms[43]. Given their prognostic power and reliability, molecular diagnostic tests are expected to become increasingly relevant tools in tailoring care to each patient. Diagnostic tests have traditionally been thought of as low-cost items in the health-care chain; however, the new generation of molecular diagnostic tests will be more drug-like in terms of their effects on patient care, their oversight by regulatory authorities and, consequently, their cost. The increased expenditure on molecular diagnostics could, however, be mitigated by the subsequent reduction in use of costly (molecularly targeted) therapies[50].

Molecular diagnostic tests will also be pivotal in identifying patients who respond to experimental anticancer drugs in clinical trials. Increasingly, drugs will be developed together with a dedicated companion diagnostic test that identifies responders to the drug in question. This co-development of drug and biomarker underscores the need for regulatory authorities to control both the drug and the companion diagnostic test. There seems to be no return from this new path to drug development. At first glance, stratifying patients in this way might seem unattractive to the pharmaceutical industry, because it reduces the size of the market for each anticancer drug. Conversely, however, molecular profiling might uncover commonalities between seemingly different tumours, potentially expanding the market for a candidate drug. When a drug has been developed and is available together with a companion diagnostic test that correctly identifies patients who benefit from the drug, there is no longer a place for a similar drug for which patients cannot be adequately stratified. Or, as one of our colleagues in the drug development industry said recently: "Pharmacogenomics, you either do it, or it is done to you." For once, the choice seems simple. ■

1. Early Breast Cancer Trialists' Collaborative Group. Polychemotherapy for early breast cancer: an overview of the randomised trials. *Lancet* **352,** 930–942 (1998).
2. Schena, M., Shalon, D., Davis, R. W. & Brown, P. O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270,** 467–470 (1995).
3. Fodor, S. P., Read, J. L., Pirrung, M. C., Lu, A. T. & Solas, D. Light-directed, spatially addressable parallel chemical synthesis. *Science* **251,** 767–773 (1991).
4. van 't Veer, L. J. *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415,** 530–536 (2002).
5. Chang, H. Y. *et al.* Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. *Proc. Natl Acad. Sci. USA* **102,** 3738–3743 (2005).
6. Bild, A. H., Potti, A. & Nevins, J. R. Linking oncogenic pathways with therapeutic opportunities. *Nature Rev. Cancer* **6,** 735–741 (2006).
7. Bild, A. H. *et al.* Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* **439,** 353–357 (2006).
   This paper shows that the activation status of oncogenic pathways is associated with a characteristic gene-expression signature and that these signatures can be used both to determine a patient's prognosis and to predict his or her response to particular drugs.
8. Paik, S. *et al.* A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N. Engl. J. Med.* **351,** 2817–2826 (2004).
9. Quackenbush, J. Computational analysis of microarray data. *Nature Rev. Genet.* **2,** 418–427 (2001).
10. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406,** 747–752 (2000).
11. 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).
12. 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).

13. Alizadeh, A. A. *et al.* Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403,** 503–511 (2000).
This study shows that previously undetected and clinically significant subtypes of cancer can be identified by molecular classification of tumours on the basis of gene expression.
14. Wang, Y. *et al.* Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* **365,** 671–679 (2005).
15. Ein-Dor, L., Zuk, O. & Domany, E. Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer. *Proc. Natl Acad. Sci. USA* **103,** 5923–5928 (2006).
16. Desmedt, C. *et al.* Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. *Clin. Cancer Res.* **13,** 3207–3214 (2007).
17. Buyse, M. *et al.* Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J. Natl Cancer Inst.* **98,** 1183–1192 (2006).
18. Fan, C. *et al.* Concordance among gene-expression-based predictors for breast cancer. *N. Engl. J. Med.* **355,** 560–569 (2006).
19. Yu, J. X. *et al.* Pathway analysis of gene signatures predicting metastasis of node-negative primary breast cancer. *BMC Cancer* **7,** 182, doi:10.1186/1471-2407-7-182 (2007).
20. Potti, A. *et al.* A genomic strategy to refine prognosis in early-stage non-small-cell lung cancer. *N. Engl. J. Med.* **355,** 570–580 (2006).
21. Ma, X.-J. *et al.* The *HOXB13:IL17BR* expression index is a prognostic factor in early-stage breast cancer. *J. Clin. Oncol.* **24,** 4611–4619 (2006).
22. Harris, L. N. *et al.* Predictors of resistance to preoperative trastuzumab and vinorelbine for HER2-positive early breast cancer. *Clin. Cancer Res.* **13,** 1198–1207 (2007).
23. Dressman, H. K. *et al.* Gene expression profiles of multiple breast cancer phenotypes and response to neoadjuvant chemotherapy. *Clin. Cancer Res.* **12,** 819–826 (2006).
24. Potti, A. *et al.* Genomic signatures to guide the use of chemotherapeutics. *Nature Med.* **12,** 1294–1300 (2006).
25. Staunton, J. E. *et al.* Chemosensitivity prediction by transcriptional profiling. *Proc. Natl Acad. Sci. USA* **98,** 10787–10792 (2001).
26. Lee, J. K. *et al.* A strategy for predicting the chemosensitivity of human cancers and its application to drug discovery. *Proc. Natl Acad. Sci. USA* **104,** 13086–13091 (2007).
27. Coombes, K. R., Wang, J. & Baggerly, K. A. Microarrays: retracing steps. *Nature Med.* **13,** 1276–1277 (2007).
28. Bonnefoi, H. *et al.* Validation of gene signatures that predict the response of breast cancer to neoadjuvant chemotherapy: a substudy of the EORTC 10994/BIG 00-01 clinical trial. *Lancet Oncol.* **8,** 1071–1078 (2007).
29. Neve, R. M. *et al.* A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10,** 515–527 (2006).
30. Huang, E. *et al.* Gene expression phenotypic models that predict the activity of oncogenic pathways. *Nature Genet.* **34,** 226–230 (2003).
31. Saal, L. H. *et al.* Poor prognosis in carcinoma is associated with a gene expression signature of aberrant PTEN tumor suppressor pathway activity. *Proc. Natl Acad. Sci. USA* **104,** 7564–7569 (2007).
32. Hughes, T. R. *et al.* Functional discovery via a compendium of expression profiles. *Cell* **102,** 109–126 (2000).
33. Lamb, J. *et al.* The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313,** 1929–1935 (2006).
In this study, a compendium of gene-expression profiles of cells subjected to known perturbations was used to find connections among small molecules that share a mechanism of action, between chemicals and physiological processes, and between drugs and diseases.
34. Hieronymus, H. *et al.* Gene expression signature-based chemical genomic prediction identifies a novel class of HSP90 pathway modulators. *Cancer Cell* **10,** 321–330 (2006).
35. Wei, G. *et al.* Gene expression-based chemical genomics identifies rapamycin as a modulator of MCL1 and glucocorticoid resistance. *Cancer Cell* **10,** 331–342 (2006).
36. Vogel, C. L. *et al.* Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. *J. Clin. Oncol.* **20,** 719–726 (2002).
37. Berns, K. *et al.* A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. *Cancer Cell* **12,** 395–402 (2007).
This study used large-scale RNA-interference screens to identify oncogenic pathways that cause resistance to anticancer drugs in the clinic.
38. Nagata, Y. *et al.* PTEN activation contributes to tumor inhibition by trastuzumab, and loss of PTEN predicts trastuzumab resistance in patients. *Cancer Cell* **6,** 117–127 (2004).
39. Swanton, C. *et al.* Regulators of mitotic arrest and ceramide metabolism are determinants of sensitivity to paclitaxel and other chemotherapeutic drugs. *Cancer Cell* **11,** 498–512 (2007).
40. Whitehurst, A. W. *et al.* Synthetic lethal screen identification of chemosensitizer loci in cancer cells. *Nature* **446,** 815–819 (2007).
41. Farmer, H. *et al.* Targeting the DNA repair defect in *BRCA* mutant cells as a therapeutic strategy. *Nature* **434,** 917–921 (2005).
42. Bryant, H. E. *et al.* Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* **434,** 913–917 (2005).
43. Shi, L. *et al.* The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nature Biotechnol.* **24,** 1151–1161 (2006).
44. Sparano, J. A. TAILORx: trial assigning individualized options for treatment (Rx). *Clin. Breast Cancer* **7,** 347–350 (2006).
45. Mook, S. *et al.* Individualization of therapy using MammaPrint: from development to the MINDACT Trial. *Cancer Genomics Proteomics* **4,** 147–155 (2007).
46. 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).
47. Harris, L. *et al.* American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *J. Clin. Oncol.* **25,** 5287–5312 (2007).
48. Rhodes, A. *et al.* Reliability of immunohistochemical demonstration of oestrogen receptors in routine practice: interlaboratory variance in the sensitivity of detection and evaluation of scoring systems. *J. Clin. Pathol.* **53,** 125–130 (2000).
49. Ach, R. A. *et al.* Robust interlaboratory reproducibility of a gene expression signature measurement consistent with the needs of a new generation of diagnostic tools. *BMC Genomics* **8,** 148, doi:10.1186/1471-2164-8-148 (2007).
50. Hornberger, J., Cosler, L. E. & Lyman, G. H. Economic analysis of targeting chemotherapy using a 21-gene RT-PCR assay in lymph-node-negative, estrogen-receptor-positive, early-stage breast cancer. *Am. J. Manag. Care* **11,** 313–324 (2005).

Acknowledgements We thank L. Wessels and P. Borst for discussions. Our work was supported by grants from the Centre for Biomedical Genetics, the Cancer Genomics Centre and the Dutch Cancer Society.

Author Information Reprints and permissions information is available at npg.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Correspondence should be addressed to R.B. (r.bernards@nki.nl).