Comparative expressed sequence hybridization to chromosomes for tumor classification and identification of genomic regions of differential gene expression

Yong-Jie Lu*†, Daniel Williamson*‡, Jeremy Clark‡, Rubin Wang*‡, Nicki Tiffin†, Lorraine Skelton§, Tony Gordon*†, Richard Williams†, Barry Allan§, Ann Jackman§, Colin Cooper‡, Kathy Pritchard-Jones†, and Janet Shipley*‡¶

*Molecular Cytogenetics, Institute of Cancer Research, Sutton, Surrey SM2 5NG, United Kingdom; †Section of Paediatric Oncology, Institute of Cancer Research and Royal Marsden National Health Science Trust, Sutton, Surrey SM2 5NG, United Kingdom; ‡Section of Molecular Carcinogenesis, Institute of Cancer Research, Sutton, Surrey SM2 5NG, United Kingdom; and §Cancer Research Campaign, Centre for Cancer Therapeutics, Institute of Cancer Research, Sutton, Surrey SM2 5NG, United Kingdom

Communicated by George F. Vande Woude, Van Andel Research Institute, Grand Rapids, MI, May 31, 2001 (received for review December 8, 2000)

Altered expression of genes can have phenotypic consequences in cancer development and treatment, developmental abnormalities, and differentiation processes. Here we describe a rapid approach, termed comparative expressed sequence hybridization (CESH), which gives a genome-wide view of relative expression patterns within tissues according to chromosomal location. No prior knowledge of genes or cloning is required, and minimal amounts of tissue can be used. Expression profiles are achieved in a manner similar to the identification of chromosomal imbalances by comparative genomic hybridization analysis. The approach is demonstrated to indicate a chromosomal region that harbors overexpressed genes that may be associated with a drug-resistant phenotype. In addition, known and new regions of differential gene expression in both normal tissues and tumor samples from the soft tissue sarcoma group of rhabdomyosarcoma (RMS) are indicated. These regions included 2p24; overexpression of *MYCN* at 2p24 was confirmed by quantitative reverse transcription–PCR for all of the alveolar RMS cases and did not necessarily correspond to genomic amplification. Evidence including region specific microarray analysis indicated that overexpression of several genes from a region may be required for detection by CESH. This evidence is consistent with clusters of functionally related genes and mechanisms that affect the expression of a number of genes at a particular genomic location. The distinctive CESH profiles demonstrated in different subtypes of RMS show potential for tumor classification.

The expression levels of genes in cells varies in different tissue types, during developmental processes, in response to internal and external signals, and in different disease states. An accumulation of genetic changes is associated with malignant transformation and cancer progression and can critically affect the expression and function of key genes. Response and resistance to cancer therapies have also been linked to genetic changes and/or the expression of key genes (1). However, the altered expression of genes in tumors may not necessarily be associated with a genetic alteration, and therefore approaches that directly assay the levels of expression can be an important step in identifying oncogenic mechanisms.

It is clear that the expression of many genes is altered during normal cellular processes and oncogenic transformation. Probably the most significant approach recently developed to screen and compare the expression patterns of many genes simultaneously has been microarray-based analysis. These analyses have been used to identify genes frequently under- or overexpressed and also to generate patterns of expression that can be analyzed to group or subdivide tumors into potentially useful clinical categories (2–5). Microarray methodologies generally require sequence information of the genes or known cloned cDNAs that are presented as a target for hybridization. Recently the approach has been suggested to be error-prone (6), and the downstream analysis and data collation are complex. Here we have explored the potential of a simple and rapid approach to expression profiling targeting banded chromosomes that we have termed comparative expressed sequence hybridization (CESH). The principle is to differentially label probes derived from the RNA of two tissues to be compared, cohybridize these to normal metaphase chromosomes, and measure the ratio of the normalized fluorescence intensities at points perpendicular to the chromosome axis in a manner analogous to the identification of chromosomal imbalances by using comparative genomic hybridization (CGH) (7). This approach could rapidly compare controls with tissues in normal, diseased, and experimental states to highlight regions of the genome differentially expressing genes as a starting point for identifying candidate genes from the region for higher-resolution analysis. As evidence from microarray data are emerging to support the suggestion that expression profiles of many genes are useful in tumor classification, we also wanted to explore the possibility that a simple rapid overview of expression patterns on metaphase chromosomes could be a useful tool in classifying tumors.

Materials and Methods

Samples. The cell line W1L2:R865 is resistant to the compound CB30865 and derived from the leukemic WIL2 cell line (8). Four rhabdomyosarcoma (RMS) cell lines were used in the study: SCMC-RM2 (9), RH30 (10), RMS (11), and RD (from the American Type Tissue Culture Collection). In addition, samples from 14 tumors with the histological classification of RMS were snap-frozen after surgical removal. The fusion gene status of cell lines and of these samples was determined (12). In total, 13 samples were classified as the alveolar subtype, 9 of which had the classical *PAX3-FKHR* fusion gene transcript including the cell lines SCMC-RM2, RH30, and RMS, and four samples had the variant *PAX7-FKHR* fusion gene transcript. The histological subtype of the remaining samples was embryonal, and no fusion gene transcripts were identified in these or the cell line RD. White blood cells separated from normal heparinized blood by using Ficol and tissue from normal muscle were stored frozen.

Abbreviations: CGH, comparative genomic hybridization; CESH, comparative expressed sequence hybridization; DOP-PCR, degenerate oligonucleotide primed-PCR; RMS, rhabdomyosarcoma; YAC, yeast artificial chromosome.

¶To whom reprint requests should be addressed at: Molecular Cytogenetics, Institute of Cancer Research, Haddow Laboratories, 15 Cotswold Road, Sutton, Surrey SM2 5NG, United Kingdom. E-mail: shipley@icr.ac.uk.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

GENETICS

A cell line, U20S, had been demonstrated to express metallothionein genes on exposure to zinc. Zinc-treated and untreated cells were available for analysis (data not shown).

RNA Preparation and Reverse Transcription. RNA was extracted from tissue samples and cultured cells by using Trizol (Life Technologies, Grand Island, NY) according to the manufacturer's instructions. The concentration of RNA was measured by using a spectra-imager before RNase-free DNase (Ambion, Austin, TX) treatment. Two units of DNase I was used to treat 25 μg of RNA for 1 h, as in the manufacturer's instructions. DNA-free RNA (0.1–1 μg) was reverse transcribed by using a random hexamer and Superscript II (Life Technologies), as in the manufacturer's instructions.

Second-Strand Labeling. Second-strand labeling was carried out in the same tube after reverse transcription. The 20 μl of reverse transcription reaction was diluted to 150 μl containing 36.8 units of DNA polymerase I (Promega), 2 units of RNase H (Life Technologies), 0.02 mM dATP, dCTP, and dGTP, 0.01 mM dTTP, 0.01 mM FluoroRed (Amersham Pharmacia), 50 mM Tris·HCl, 5 mM $MgCl_2$, 10 mM β-mercaptoethanol, and 10 μg/ml BSA. After 3-h labeling at 15°C, 0.06 units of DNase I was added and incubated for another 1.5 h to reduce the probe size. After the reaction was stopped by EDTA, unincorporated nucleotides were removed by passing through Sephadex G-50 columns.

cDNA Library. A cDNA library from W1L2:R865 was made by using a cDNA Synthesis Kit (Stratagene) according to the manufacturer's instructions and ligated into a pBluescript vector.

PCR. D2S123 microsatellite marker primers (13) were used to test for genomic DNA contamination in DNase-treated RNA samples. cDNA reverse transcribed from 0.1 μg of RNA was amplified by degenerate oligonucleotide primed (DOP)-PCR in a 25-μl reaction for 7 + 35 cycles according to Telenius *et al.* (14). Five microliters of these products was subsequently labeled in a 50-μl PCR reaction for 25 cycles by using the same 6MW primer, a reduced dTTP concentration of 0.1 mM and FluoroGreen or FluoroRed nucleotides (Amersham Pharmacia) at a final concentration of 0.02 mM. cDNAs from the cDNA library was subjected to PCR by using the M13 forward and reverse primer sequences within the pBluescript vector (Stratagene). They were subsequently labeled for 25 cycles of PCR by reducing the dTTP concentration and adding FluoroRed into the PCR buffer. DOP-PCR resulted in a probe size in the region of 250–1,500 bp in length with a higher labeling efficiency than any of the other methods as judged by agarose gel electrophoresis.

Hybridization. Normal blood metaphases were made according to standard cytogenetic methods. Before hybridization, slides were treated with 70% acetic acid for 10 min, washed with PBS, treated with RNase A for 30 min (0.1 μg/ml) and then rinsed in 2 × SSC. After dehydration, slides were denatured in 70% formamide/2 × SSC at 72°C for 1–3 min. Three micrograms of each of two differentially labeled probes, estimated to be derived from 10 ng of RNA, were denatured together with 0.1 μg of fragmented but unlabeled human genomic DNA and 10 μg Cot1 DNA (Life Technologies) in 10 μl of hybridization buffer (50% formamide/2 × SSC/10% dextran sulfate) at 75°C for 6 min. The probes were applied to the slide and sealed under a 22 × 22-mm coverslip. Hybridization was carried out in a humid box at 37°C for 48 h. After hybridization, slides were washed in 0.3% Igepal (Sigma)/0.4 × SSC at 72°C for 2 min, then in 0.1% Igepal/2 × SSC at 45°C for 2 min and finally 0.1%/Igepal 2 × SSC for 5 min at room temperature. After dehydration, slides were mounted in an antifade solution containing 4′,6-diamidino-2-phenylindole.

CESH Analysis. Images were captured by using a cooled charge-coupled device camera mounted onto a Zeiss fluorescence microscope and controlled by a Macintosh computer by using SMARTCAPTURE software (Digital Scientific, Cambridge, U.K.). Five good-quality metaphases with 320- to 450-band resolution were analyzed for each case by using the QUIPS CGH analysis software (Vysis, Chicago, IL) to produce an average profile of the fluorescence intensity ratios. Unlike CGH, the CESH ratios of fluorescence frequently change within a small chromosomal region. Misalignment in averaged profiles can reduce resolution, and therefore two or three metaphases with less condensed chromosomes (around 700-band resolution) were also examined individually to help define regions with differentially expressing genes.

Quantitative PCR. Primers for the *MYCN* gene within exon3 and across the exon2/exon3 boundary were designed according to the Applied Biosystem's instructions to quantify relative genomic and expression levels, respectively. TaqMan analysis was carried out according to the manufacturer's instructions by using an Applied Biosystems 7700 Sequence detector. The genomic level of *MYCN* was measured relative to *hsRPB4* at 2q21 and the relative expression levels measured relative to *GAPDH* expression. Each assay was multiplexed and determined by a comparative method validated by Applied Biosystems with the formula $2^{-\Delta\Delta Ct}$.

Microarray Analysis. Fifteen I.M.A.G.E. (Integrated Molecular Analysis of Genomes and their Expression) cDNA clones from the 2p24 region obtained from the Human Genome Mapping Project Resource Centre (www.hgmp.mrc.ac.uk) and a gene set of 5,500 clones were gridded onto polylysine-coated microscope slides. Cy3-labeled RNA from 10 pooled normal muscle samples and Cy5-labeled RNA from cases 31, 281, RD, 82, and 93 were hybridized to the slides essentially as described (15). Digital images were captured by using a genepix 4000 scanner and software (Axon Instruments, Foster City, CA) and analysis basically carried out as previously described (4, 16). The median localized background signal for each clone was subtracted and each ratio normalized by a single scalar factor such that the mean log ratio = 0. The log ratios for each array showed a normal distribution. Ninety-five percent confidence limits were calculated on the basis of self/self control hybridizations such that there was a 95% probability that clones with measured ratios outside of the confidence interval did not have an actual ratio of 1.

Results and Discussion

To test whether differences in relative expression levels could be detected by using chromosomes as the target, a model system was used. The cell line W1L2:R865 is resistant to a lipophilic quinazoline-based aminomethyl pyridine compound, CB30865, and is associated with amplification at 7q22, as determined by CGH (8). Hybridization of cDNA to chromosomes generally results in signals that are difficult to detect because of the small target footprint of coding sequence in the genome. Therefore, various approaches to label reverse-transcribed RNA were compared. First, cDNA reverse transcribed from 3 μg of total RNA from W1L2:R865, the parent cell line W1L2, and normal tissues was labeled during second-strand cDNA synthesis, as described in *Materials and Methods* (fragment size range 200–2,000 bp). Second, cDNA from a W1L2:R865 cDNA library was amplified and labeled by PCR by using the vector M13 forward and reverse primer sites. Both these methods had a lower labeling efficiency and produced a significantly weaker chromosomal painting signal in comparison to a probe derived from amplifying and

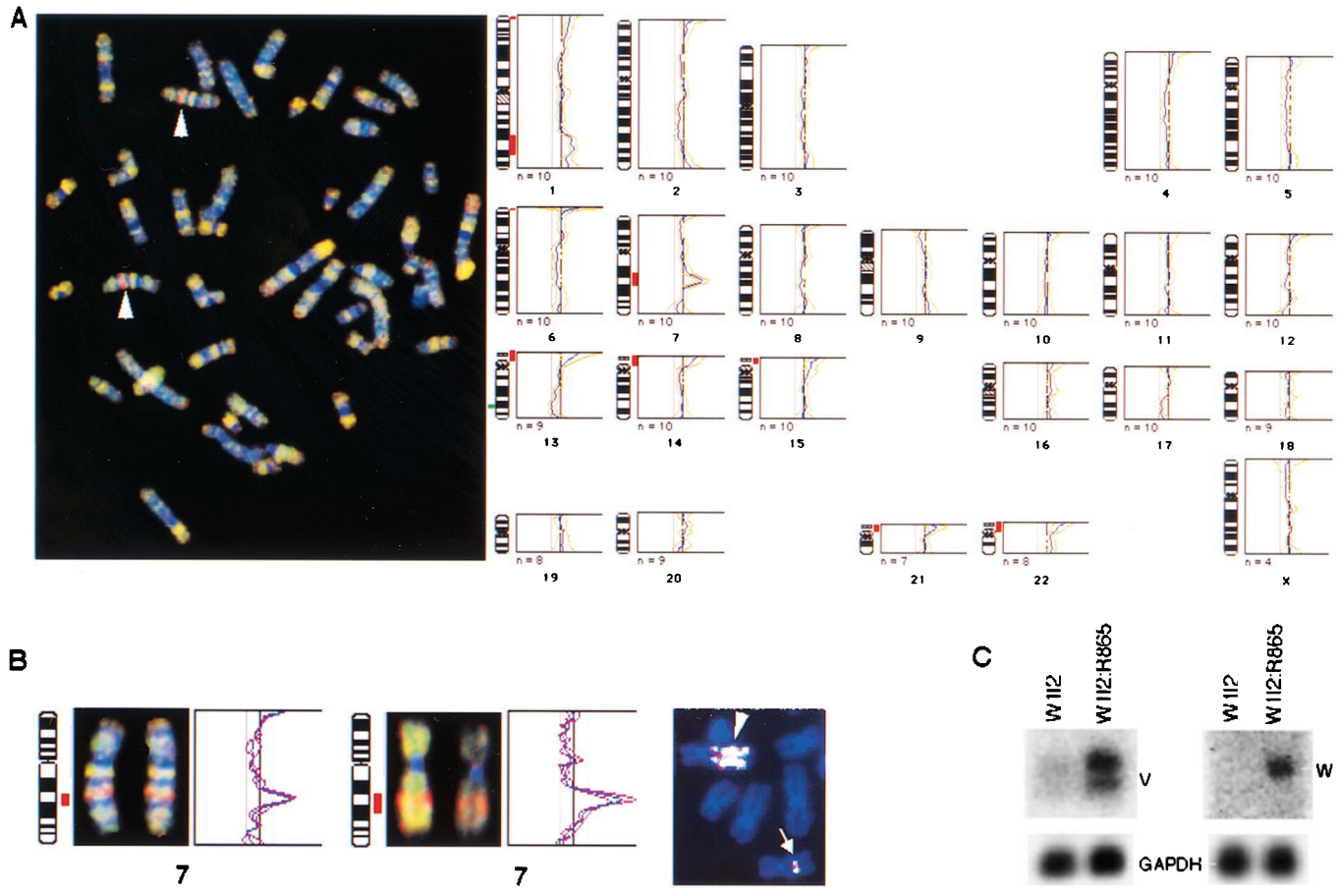

Fig. 1. Detection of overexpression of genes within the 7q22 amplicon of the drug-resistant cell line W1L2:R865 by CESH. (*A*) An example of CESH by using cDNA from the drug-resistant cell line W1L2:R865 (red) cohybridized with cDNA from the parent line W1L2 (green). A distinct red band on 7q can be seen, indicating overexpression of genes from this region (arrowhead, *Left*). The profile (*Right*) shows that the average red/green fluorescence intensity ratio exceeds the 1.2 limit on 1q and 7q and reaches the 0.8 limit on 13q. (*B*) Comparison of CESH (*Left*) and CGH (*Center*) for W1L2: R865 vs. W1L2 for chromosome 7. Although the profiles look similar, the hybridization patterns are very different. Genomic amplification at 7q22 was also demonstrated by hybridization of YACs yWSS642 and yWSS3307 to the homogenously staining region (HSR) on a marker chromosome from the W1L2:R865 cell line (*Right*, arrowhead, signals on the HSR; arrow, signal on a normal chromosome 7). (*C*) Northern results showing overexpression of gene sequences V and W contained within the YACs yWSS642 and yWSS3307 in W1L2:R865 compared with W1L2. The housekeeping gene *GAPDH* was used to control for differences in RNA loading.

labeling single-stranded cDNA through DOP-PCR (14). The hybridization patterns of the probes labeled by the three different techniques were similar, suggesting that any biases introduced by nonlinear amplification as a result of PCR are minimal at this resolution.

DOP-PCR has been extensively used to amplify test and control genomic DNA for CGH experiments (17, 18). DOP-PCR of DNA from a single cell has been demonstrated to amplify at around 90% efficiency at 10 randomly selected loci (17). It has been suggested that only partial hybridization at the 3′ end of the 6-bp nondegenerate region of the primers is sufficient to generate tagged DNA (19). Our investigation of DOP-PCR amplification of an actin probe that had no specific sites for the 6-bp and 4 sites for the 5-bp recognition site produced 7 or 8 discrete bands (data not shown). Therefore, we estimate that the majority of cDNAs are amplified by DOP-PCR.

As genomic DNA contamination could potentially interfere with results, the method used includes a DNase treatment of the RNA before reverse transcription. We confirmed the lack of genomic sequences in the treated RNA product by comparing PCR analysis of untreated and treated RNA by using microsatellite primers. Only the former produced a PCR product.

In control experiments, single RNA samples were reverse transcribed and aliquots differentially labeled by DOP-PCR and cohybridized to metaphase chromosomes. The average fluorescence intensity ratios and standard deviation did not exceed 1.0 $\pm$ 0.2 at any chromosomal location outside the centromeres, telomeres and heterochromatic regions (see Fig. 4, which is published as supplemental data on the PNAS web site, www.pnas.org). This is similar to CGH analysis (7, 20). Cohybridization of the differentially labeled probes for W1L2 and the drug-resistant line, W1L2:R865, produced the average profile shown in Fig. 1*A*. An average fluorescence intensity outside the 1.2 limit was reproducibly found at the 7q22 region, indicating overexpression of genes at this location in the resistant line. CGH has indicated genomic amplification at 7q22 in W1L2:R865 that was confirmed by fluorescence *in situ* hybridization by using two yeast artificial chromosomes (YACs) (CEPH megayacs yWSS642 and yWSS3307), which showed multiple copies in W1L2:R865 but not in the parent cell line (Fig. 1*B*). Two expressed sequences corresponding to these YACs have been shown to be overexpressed relative to the parent line by Northern analysis (Fig. 1*C*). The relative excess fluorescence at 7q22 detected by CESH therefore correlates with overexpression of genes at this locus in W1L2:R865 relative to the parent cell line.

In subsequent experiments, cDNAs from lymphocytes from different individuals were differentially labeled by DOP-PCR

GENETICS

Table 1. Chromosome regions containing differentially expressed genes in normal human tissues as indicated by CESH analysis

Tissue pairs	Region expressing genes at relatively high level (CESH ratio > 1.2)	Region expressing genes at relatively low level (CESH ratio < 0.8)
Lymphocyte/fibroblast	6p21–23, 10p13, 3p21-ter	5p15, 10q25–26, 12q24, 13q31–34
Muscle/lymphocyte	2q23, 2q24, 3q27-ter, 16p13, 18p, Xp11.3-21	6p21–23
Muscle/fibroblast	2q23, 2q24, 3p21-ter, Xp11.3-21	5p13-ter, 18q21

and cohybridized as described. The average fluorescence intensities and their standard deviation did not exceed 1 ± 0.2. Cohybridization of differentially labeled muscle tissue from the same and different individuals also produced similar results. In sex-mismatched samples, the profiles for chromosome X remained within the limits, which is consistent with classical X inactivation theory. For lymphocytes, muscle, and normal human fibroblasts, the regions that labeled most intensely were 1p32–36, 3p14–21, 5q31, 6p21, 9q33–34, 11q13, 12q11–13, 12q24, 15, 16p, 17, 19, 20q, and 22. Signals on chromosomes 4, 13, 18, and X were significantly less intense (see Fig. 4). These common regions of strong hybridization are known to be gene rich from SAGE analysis (21) and indirectly from studies by using hybridization of DNA with a specific base composition (isochores) (22). Therefore, these intensely painting regions are direct evidence for gene-rich regions of the genome.

In addition to the strongly hybridizing regions common to tissues, highly reproducible differences were observed in experiments directly comparing the different tissues (Table 1). The overexpression indicated at 6p in lymphocytes when compared with fibroblasts and muscle tissue is likely to reflect the differential expression of genes in the MHC cluster at this location. The most striking difference corresponded to higher-level expression of genes in the 2q23–32 region in muscle, which was composed of two separate peaks at 2q23 and 2q31–32 (Fig. 2).

CESH analysis of RNA from the RMS samples compared with normal muscle RNA was then performed to further explore the ability of the technique to identify regions of the genome harboring differentially expressed genes and assess its potential for producing profiles useful for tumor classification. Generally, the CESH profiles were more consistent than those observed by CGH (Fig. 3) (20, 23). A higher-level gene expression from the 2q23–32 region was consistently indicated in the control terminally differentiated muscle compared with the RMS and similar to that seen in comparison of muscle to lymphocytes and fibroblasts (Table 1; Figs. 2 and 3). Overexpression of genes in the 2q23 and 2q31-q32 regions as well as relative overexpression of genes from the Xp11.2–21 region may be functionally related to terminally differentiated muscle tissue.

The two main histological subtypes of RMS are known as alveolar and embryonal (24). The alveolar subtype usually has either the *PAX3-FKHR* fusion gene transcript associated with the chromosome translocation t(2;13) (q35;q14) or the variant *PAX7-FKHR* fusion gene transcript associated with the t(1;13) (p36;q14) (24). Regions of differential gene expression, such as relative underexpression from 6q and overexpression from 8p23, were frequently indicated in all subtypes of RMS by CESH. However, specific differences in the CESH profiles of samples from the alveolar subtype known to have either the *PAX3-FKHR* ($n = 9$) or variant *PAX7-FKHR* ($n = 4$) fusion gene transcripts and the embryonal ($n = 5$) subtype were also found, which shows potential for tumor classification (Fig. 3). Alveolar cases with the

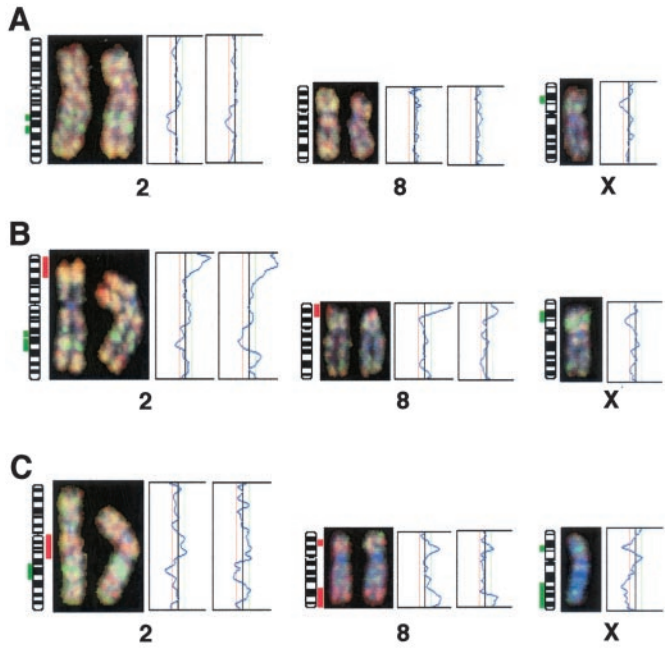

Fig. 2. CESH analysis and the red-to-green ratios of fluorescence intensities for chromosomes 2, 8, and X. In all cases, the same cDNA from a skeletal muscle sample of a male individual was labeled in green. (*A*) Cohybridization with cDNA from short-term culture of normal fibroblasts from a male individual (red). Lower level of expression of genes in the 2q23, 2q31, and Xp11.3–21 regions is seen in normal fibroblast compared with normal muscle. (*B*) Hybridization with cDNA from a female alveolar RMS with a *PAX3-FKHR* fusion gene (red). A relative reduced level of expression of genes from 2q23, 2q31, and Xp11.3–21 and overexpression of genes from 2p21-ter and 8p21-ter is indicated in this tumor sample. (*C*) Hybridization with cDNA from a female embryonal RMS primary sample (red) shows relative reduced expression of genes from 2q23–31, Xp11.3–21, and Xq and overexpression of genes from 2q11–21, 8p21, and 8q23-ter.

PAX3-FKHR showed lack of expression from the 10q22–23 region in contrast to the *PAX7-FKHR* cases. It is likely that expression differences in addition to the fusion gene transcripts are important and account for the clinical differences associated with tumors carrying these different fusion genes (25). CESH indicated overexpression from a gene or genes located at 8q24 in the embryonal subtype, which may include *MYCC*. Embryonal tumors are heterogeneous in their clinical behavior, and the differences in the expression profiles, particularly from genes in the 2p24, 12q13–15, and 10q22–23 regions, may begin to subdivide this more frequent subtype of RMS.

Involvement of the 2p24 region was identified by CESH, indicating overexpression of a gene or genes at this location in all 13 alveolar cases. Overexpression of *MYCN* at this location was demonstrated by quantitative reverse transcription–PCR relative to normal muscle tissue in all of the alveolar cases with either fusion gene (in excess of 90-fold). Overexpression was irrespective of genomic amplification of *MYCN*. Amplification was identified in 5 of the 13 cases by using quantitative PCR and was consistent with the CGH results for the 2p24 region (Fig. 3 *A* and *B*). Overexpression of *MYCN* has been noted in a number of alveolar RMS cell lines by microarray analysis (26), and we have previously suggested involvement of the *MYCN* gene in alveolar RMS (24). Our results suggest that overexpression of *MYCN* may be associated with all primary alveolar RMS, irrespective of genomic amplification, and may be of key significance.

Only two embryonal cases (281, RD) showed overexpression from 2p24 by CESH analysis. However, *MYCN* was not shown to be overexpressed by quantitative reverse transcription–PCR

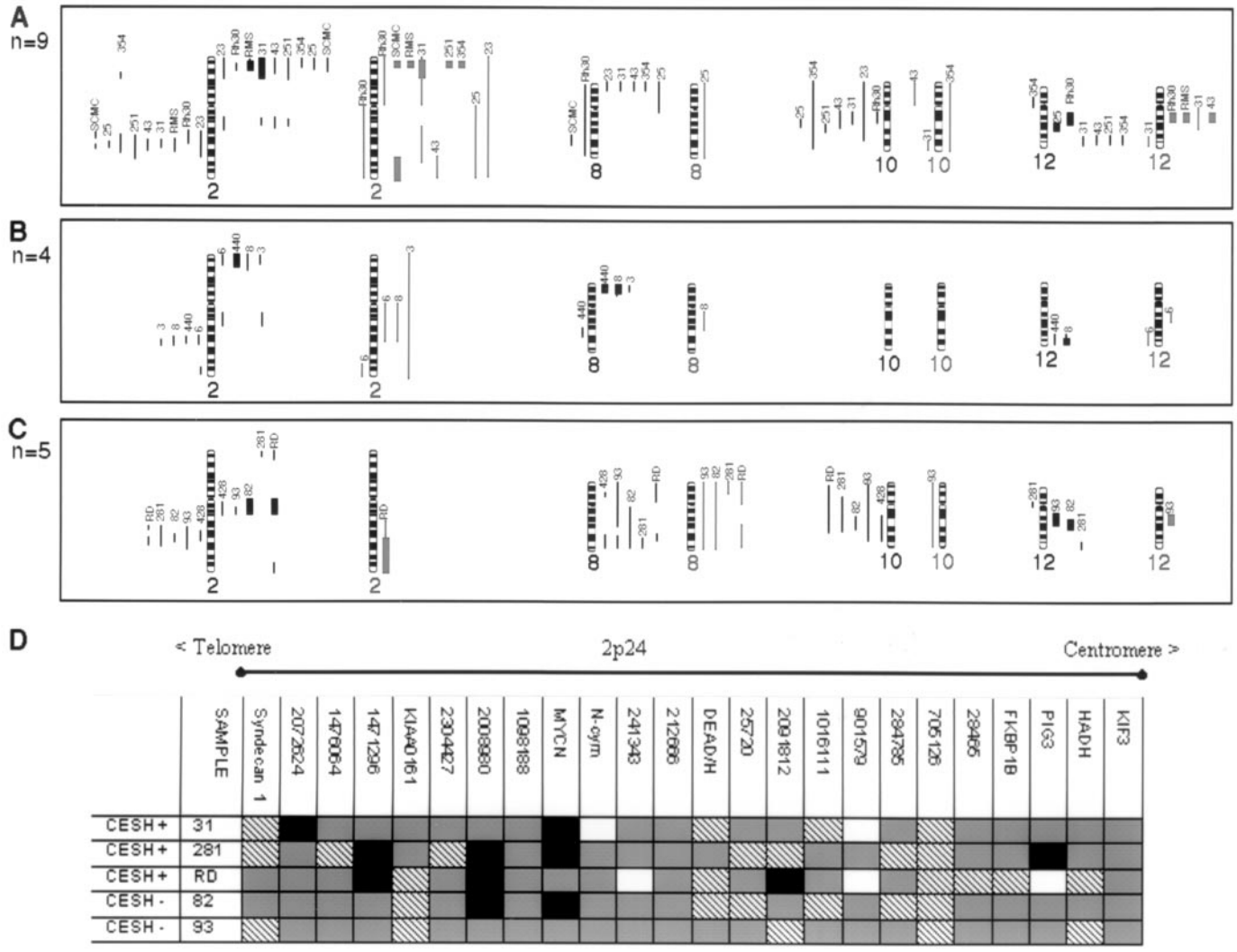

Fig. 3. Summary of CESH and CGH analysis for specific chromosomes in RMS and the corresponding microarray expression analysis for the 2p24 region. The first chromosome of a particular number shows the CESH analysis with vertical black lines on the left side of each chromosome, indicating reduced expression levels, and vertical black lines on the right side, indicating overexpression compared with control muscle. The second chromosome summarizes the CGH results, which include some previously published cases (Rh30, RMS, 31, 43, 251, 354, 25, SCMC, 6, 440, 8, 93), with gray lines to the left- and right-hand side indicating loss and gain of chromosome material, respectively (20, 23). High-level overexpressions/gains are highlighted by bold lines where the ratios exceeded 1.5. (*A*) Results from six primary tumor samples and three cell lines derived from alveolar RMS with *PAX3-FKHR* fusion genes. (*B*) Results from four primary tumor samples of alveolar RMS with *PAX7-FKHR*. (*C*) Results from four primary samples and one cell line of embryonal RMS. (*D*) Summary of microarray expression analysis indicating 2p24 clones significantly overexpressed, black; significantly underexpressed, white; not significantly different, gray; or expression levels insufficient for analysis, stippled.

analysis relative to muscle in RD. Furthermore, case 82, which by CESH analysis did not indicate overexpression from the 2p24 region, showed *MYCN* overexpression. Region-specific microarray analysis of these embryonal cases plus an alveolar case (case 31) was used to explore this region further and determine the limits of the CESH approach. The results are summarized in Fig. 3*D* and display the clones significantly under- or overexpressed by using a 95% confidence interval. It is notable that, although RD showed no *MYCN* overexpression, three other clones from the region did, indicating the involvement of genes other than *MYCN* in producing a CESH response in this region. Although two genes were overexpressed in case 82, this was not detected by CESH. However, together the microarrays suggest that at least two genes must be significantly overexpressed within this region to be detected by CESH.

Consistent with this, involvement of two expressed sequences at 7q21 in the drug-resistant cell line WIL2:R865 has been shown (Fig. 1), and a further seven cDNA clones from the region were also indicated by microarray analysis (data not shown). Cells from the osteosarcoma cell line U2OS have been shown to overexpress by at least a 2-fold level several metallothionein genes that cluster at 16q13 when treated with zinc compared with untreated cells (data not shown). This region was not detected by CESH, and in this case the level of expression and 0.3- to 0.4-kb transcript size may be below the resolution of CESH. The transcript size, level of change, and genomic location that may mask changes in gene-rich regions will also influence detection by CESH. Therefore, although regions harboring differentially expressed genes may be missed, anything indicated by CESH is likely to be real and involve a number of genes. This is consistent with evidence for clusters of functionally related genes at particular chromosomal locations and mechanisms that may affect the expression of a number of genes at a particular genomic location (5, 27).

The CESH profiles for the RMS cases studied were very different from those from the lymphoblastoid cell line W1L2, which is a cell line derived from a leukemia, and other tumor-derived samples that have been investigated (data not shown). Therefore, expression profiling of tumors at this resolution rapidly produces distinct patterns that demonstrate potential as a tool in classifying tumors. CESH may also be useful to study genes involved in tissue differentiation and embryogenesis, as we have detected expression differences between different normal tissues, and the approach can be applied to very small, even microdissected, tissue samples. The identification of the location of under- or overexpressed genes potentially important in tumorigenesis and other processes is appropriate to combine with techniques to retrieve region-specific cDNAs by using microdissected chromosome material or genomic clones (28, 29). We suggest that region-specific microarray analysis using such selected clones and/or available markers, as demonstrated here for 2p24, is an appropriate followup approach to CESH analysis. Also, now that many genes are positioned to chromosomal locations, candidate genes for investigation can be readily selected.

We thank S. Hing and J. Anderson for technical assistance and R. Wooster for access to the Institute of Cancer Research Geneset. This work is supported by the Cancer Research Campaign.

1. Schloenlein, P. V. (1993) *Cytotechnology* **12,** 63–89.
2. Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., *et al.* (1999) *Science* **286,** 531–537.
3. Alizadeh, A. A., Eisen, M., Davis, R., Ma, C., Lossos, I., Rosenwald, A., Boldrick, J., Sabet, H., Tran, T. R., Yu, X., *et al*. (2000) *Nature (London)* **403,** 503–511.
4. Bittner, M., Meltzer, P., Chen, Y., Jiang, Y., Seftor, E., Hendrix, M., Radmacher, M., Simon, R., Yakhini, Z., Ben-Dor, A., *et al*. (2000) *Nature (London)* **406,** 536–540.
5. Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, M. A., *et al*. (2000) *Nature (London)* **406,** 747–752.
6. Knight, J. (2001) *Nature (London)* **410,** 860–861.
7. Kallioniemi, A., Kallioniemi, O.-P., Sudar, D., Rutovitz, D., Gray, J. W., Waldman, F. & Pinkel, D. (1992) *Science* **258,** 818–821.
8. Skelton, L. A., Ormerod, M. G., Titley, J., Kimbell, R., Brunton, L. A. & Jackman, A. L. (1999) *Br. J. Cancer* **79,** 1692–1701.
9. Hayashi, Y., Sugimoto, T., Horii, Y., Hosoi, H., Inazawa, J., Kemshead, J. T., Inaba, T., Hanada, R., Yamamoto, K., Gown, A. M., *et al*. (1990) *Int. J. Cancer* **45,** 705–711.
10. Hazelton, B. J., Houghton, J. A., Parham, D. M., Douglass, E. C., Torrance, P. M., Holt, H. & Houghton, P. J. (1987) *Cancer Res.* **47,** 4501–4507.
11. Garvin, J. A., Stanley, W. S., Bennett, D. D., Sullivan, J. L. & Sens, D. A. (1986) *Am. J. Pathol.* **125,** 208–217.
12. Anderson, J., Renshaw, J., McManus, A., Carter, A., Mitchell, C., Adams, S. & Pritchard-Jones, K. (1997) *Am. J. Pathol.* **150,** 477–482.
13. Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millasseau, P., Vaysseix, G. & Lathrop, M. (1992) *Nature (London)* **359,** 794–801.
14. Telenius, H., Carter, N., Bebb, C. E., Nordenskjold, M., Ponder, B. A. J. & Tunnacliffe, A. (1992) *Genomics* **13,** 718–725.
15. Eisen, M. & Brown, P. (1999) *Methods Enzymol.* **303,** 179–205.
16. Stratowa, C., Loffler, G., Lichter, P., Stilgenbauer, S., Haberl, P., Schweifer, N. & Dohner, H. (2001) *Int. J. Cancer* **91,** 474–480.
17. Wells, D., Sherlock, J. K., Handyside, A. H. & Delhanty, J. D. (1999) *Nucleic Acids Res.* **27,** 1214–1218.
18. Lu, Y. J., Osin, P., Lakhani, S. R., Di Palma, S., Gusterson, B. A. & Shipley, J. M. (1998) *Cancer Res.* **58,** 4721–4727.
19. Cheung, V. G. & Nelson, S. F. (1996) *Proc. Natl Acad. Sci. USA* **93,** 14676–14679.
20. Weber-Hall, S., Anderson, J., McManus, A., Abe, S., Nojima, T., Pinkerton, R., Pritchard-Jones, K. & Shipley, J. (1996) *Cancer Res.* **56,** 3220–3224.
21. Caron, H., van Schaik, B., van der Mee, M., Baas, F., Riggins, G., van Sluis, P., Hermus, M.-C., van Asperen, R., Boon, K., Voute, P. A., *et al*. (2001) *Science* **291,** 1289–1292.
22. Cardiner, K. (1996) *Trends Genet.* **12,** 519–524.
23. Gordon, A. T., Brinkschmidt, C., Anderson, J., Coleman, N., Dockhorn-Dworniczak, B., Pritchard-Jones, K. & Shipley, J. (2000) *Genes Chromosomes Cancer* **28,** 220–226.
24. Anderson, J., Gordon, A., Pritchard-Jones, K. & Shipley, J. (1999) *Genes Chromosomes Cancer* **26,** 275–285.
25. Kelly, K., Womer, R., Sorensen, P., Xiong, Q. & Barr, F. (1997) *J. Clin. Oncol.* **15,** 1831–1836.
26. Khan, J., Simon, R., Bittner, M., Chen, Y., Leighton, S. B., Pohida, T., Smith, P. D., Jiang, Y., Gooden, G. C., Trent, J. M., *et al*. (1998) *Cancer Res.* **58,** 5009–5013.
27. Trowsdale, J. (1993) *Trends Genet.* **9,** 117–122.
28. Rouquier, S., Trask, B. J., Taviaux, S., van den Engh, G., Diriong, S., Lennon, G. G. & Giorgi, D. (1995) *Nucleic Acids Res.* **23,** 4415–4420.
29. Gracia, E., Fischer, U., elKahloun, A., Trent, J. M., Meese, E. & Meltzer, P. S. (1996) *Hum. Mol. Genet.* **5,** 595–600.