Monoterpenes inhibit enzymes in the mevalonate-lipid metabolism pathway, including a selective inhibition of isoprenylation of 21–26 kDa small G proteins (13–15) and inhibition of ubiquinone and cholesterol synthesis (16). Inhibition of isoprenylation may affect signal transduction pathways because unprocessed or non-prenylated small G proteins are not properly localized within the cell and are thus non-functional (17). Furthermore, investigations into gene expression and protein level alterations associated with monoterpenes-mediated tumor regression revealed a increase in the mito-inhibitory transforming growth factor β1 (TGF-β1) and the mannose 6-phosphate/insulin-like growth factor II receptor (Man-6-P/IGF2R), which both facilitates latent-TGF-β1 activation and degrades the mammary mitogen insulin-like growth factor II (IGF2; Ref. 18). Interestingly, while Man-6-P/IGF2R RNA expression was significantly induced 2-fold in 10% limonene-treated carcinomas, the induced RNA expression of Man-6-P/IGF2R was not observed in the non-responsive 10% limonene-treated carcinomas (19). We hypothesized that biochemical events in the cytoplasm modify signal transduction leading to altered gene expression. Thus, in order to further elucidate the monoterpenes’ mechanism of cancer chemoprevention and chemotherapy, we screened for differentially expressed genes in 10% limonene-treated, DMBA-induced, advanced mammary carcinomas resected at mid-regression.

Complex tissues, such as mammary carcinomas, are composed of large heterogeneous cell populations and hence require sensitive gene expression screening methods. This is

Identifying Differential Gene Expression in Monoterpenetreated Mammary Carcinomas Using Subtractive Display*

(Received for publication, June 18, 1996, and in revised form, August 6, 1996)

Eric A. Ariazi‡ and Michael N. Gould§

From the Department of Human Oncology, University of Wisconsin-Madison, Madison, Wisconsin 53792

The monoterpenes limonene and perillyl alcohol, an emerging class of naturally occurring anti-cancer compounds, are highly effective against a variety of rodent organ-specific cancer models (reviewed in Refs. 1 and 2). Dietary administration of monoterpenes is effective for chemoprevention and chemotherapy of both 7,12-dimethylbenz[a]anthracene (DMBA)²-in-

duced and N-methyl-N-nitrosourea-induced rat mammary carcinomas (3–10). As a chemotherapeutic agent, dietary 10% limonene caused ~66% of advanced DMBA- or N-methyl-N-nitrosourea-induced carcinomas to completely regress and an additional ~23% to partially regress (6). Moreover, a 2% perillyl alcohol diet resulted in 50% complete and an additional 25% partial regression of advanced DMBA-induced mammary carcinomas (7). Treatment with both monoterpenes at these anti-cancer doses did not cause systemic toxicity. Monoterpenes are currently being tested in Phase I clinical trials on advanced cancer patients in the United Kingdom (11) and the United States.²

Several observations suggest that monoterpenes-mediated mammary carcinoma regression may involve a differentiation/remodeling process. Histopathology of monoterpenetreated, actively regressing tumors displays regions of dense anaplastic epithelium characteristic of mammary carcinomas and regions of a remodeled epithelial compartment. Regressing carcinomas do not show increased levels of lymphocyte infiltration, inflammation, or necrosis (6). Furthermore, monoterpenetreatment of neuro2A neuroblastoma cells causes morphologic differentiation within 4 h as characterized by neurite outgrowths (12). Monoterpenes inhibit enzymes in the mevalonate-lipid metabolism pathway, including a selective inhibition of isoprenylation of 21–26 kDa small G proteins (13–15) and inhibition of ubiquinone and cholesterol synthesis (16). Inhibition of isoprenylation may affect signal transduction pathways because unprocessed or non-prenylated small G proteins are not properly localized within the cell and are thus non-functional (17). Furthermore, investigations into gene expression and protein level alterations associated with monoterpenes-mediated tumor regression revealed a large increase of the mito-inhibitory transforming growth factor β1 (TGF-β1) and the mannose 6-phosphate/insulin-like growth factor II receptor (Man-6-P/IGF2R), which both facilitates latent-TGF-β1 activation and degrades the mammary mitogen insulin-like growth factor II (IGF2; Ref. 18). Interestingly, while Man-6-P/IGF2R RNA expression was significantly induced 2-fold in 10% limonene-treated carcinomas, the induced RNA expression of Man-6-P/IGF2R was not observed in the non-responsive 10% limonene-treated carcinomas (19).

We hypothesized that biochemical events in the cytoplasm modify signal transduction leading to altered gene expression. Thus, in order to further elucidate the monoterpenes’ mechanism of cancer chemoprevention and chemotherapy, we screened for differentially expressed genes in 10% limonene-treated, DMBA-induced, advanced mammary carcinomas resected at mid-regression.

Complex tissues, such as mammary carcinomas, are composed of large heterogeneous cell populations and hence require sensitive gene expression screening methods. This is

* This work was supported in part by United States Public Health Service, National Institutes of Health Grant R37-CA38128. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Supported by the Fellowship DAMD17-94-J-4041 granted from the U.S. Army Medical Research and Material Command.

§ To whom correspondence should be addressed: University of Wisconsin-Madison, Dept. of Human Oncology, K4/3 UW CSC, 600 Highland Ave., Madison, WI 53792. Tel.: 608-263-6615; Fax: 608-263-9947; E-mail: gould@humonc.wisc.edu.

† Supported by the Fellowship DAMD17-94-J-4041 granted from the U.S. Army Medical Research and Material Command.

² H. Bailey and M. N. Gould, personal communication.
because a particular gene may be induced or repressed only in a specific cell type, which would be diluted by all the other cell types. In order to achieve the required level of sensitivity, the subtraction method of Wang and Brown (20) was initially chosen because this approach uses polymerase chain reaction (PCR) amplification to generate a renewable source of cDNA for multiple rounds of subtraction, and the method efficiently removes commonly expressed cDNA from the experimental and control cDNA pools to allow for enrichment of differentially expressed genes. However, cloning the differentially expressed cDNAs involves cycles of classical probe hybridization experiments to isolate only a few clones at a time, followed by subtraction of the newly cloned cDNAs from the library. The cloning cycle is repeated until the library is completely screened, which may take 20–30 or more cloning cycles and thus renders this methodology inefficient.

A possible alternative method considered was Liang and Pardee’s (21) differential display, since it is also a PCR-based technique that selectively amplifies subsets of cDNAs based on primer design applied to two matched but non-subtracted cDNA populations. This technique clones genes quickly but only ~10% of the genes are differentially expressed when tested (22–24). Problems with reproducibility are probably due to the low stringency in annealing of primers, although recent modifications in primer design and experimental protocol have somewhat improved reproducibility (25–28). In addition, because the primers effectively anneal as 6- or 7-base oligomers of arbitrary sequence, a statistical argument is needed to determine when the libraries are sufficiently screened (21). Furthermore, there exists a high level of background noise in differential display due to the vast majority of competing heterologous RNA. Differential display shows a strong bias toward identifying abundant mRNAs (29). An efficient gene expression screen should identify rare transcripts, which account for ~90% of mRNA species in most cells (29).

In order to more efficiently identify differentially expressed monoterpene target genes, we developed and characterized an alternative approach, termed subtractive display (SD), which incorporates the strengths and minimizes the weaknesses of subtractive hybridization and differential display. We have applied SD to identify differentially expressed genes in limonene-treated advanced rat mammary carcinomas.

**EXPERIMENTAL PROCEDURES**

**Summary of SD Methodology**—A flowchart of the SD methodology is given in Fig. 1. cDNA was generated from two matched tissue sources or for this set of experiments, monoterpene-treated, regressing, mammary carcinomas (+cDNA) and control, non-treated carcinomas (−cDNA). ±cDNA was fragmented with restriction endonucleases followed by ligation of symmetrical linkers. The sequence at the cDNA fragment ends reflects the methods used to generate the cDNA libraries and restriction endonuclease recognition sites, and hence, are non-random sequences. The subtractive process involves multiple rounds of hybridization and PCR amplification, which renews the cDNA source. +cDNA was subtracted from −cDNA, and concurrently, −cDNA was subtracted from +cDNA, to derive induced and repressed genes, respectively. Subtracted cDNA was subcloned into a transforming plasmid and subjected to three cycles of precise, unidirectional deletions. The trimming process was monitored by displaying restriction enzyme digests that overlap the cDNA insert, the sequence selectively amplifies a subset of cDNAs. Following linker ligation, the fragmented cDNAs are subjected to PCR amplification using the 21-bp linkers as primers (sequences given in Ref. 20). It is assumed that cDNA fragments common to both +cDNA and −cDNA populations were amplified equally, while different cDNA fragments within a population may be differentially amplified. Depending on the direction of subtraction, whether up- or down-regulated cDNAs were being enriched, either +cDNA or −cDNA and their derived populations were tracer or driver cDNA. Thus in the following, driver or tracer cDNA is specified but not if +cDNA or −cDNA was used. 50 μg of driver cDNA was completely digested with EcoRI to suppress contaminating driver cDNA from being amplified. Driver cDNA was then photobiotinylated and combined with 2.5 μg of non-photobiotinylated tracer cDNA. The mixture was completely denatured by boiling and then cooled for a long hybridization (20 h). Addition of streptavidin led to formation of streptavidin-biotin-DNA complexes that were removed by several phenol-chloroform extractions. Subtracted tracer cDNA (termed 1cDNA) was again subtracted from 25 μg of biotinylated tracer cDNA, this time for a short hybridization of 2 h to yield 2cDNA. +2cDNA and −2cDNA were PCR amplified. 2.5 μg of non-photobiotinylated tracer 2cDNA was subtracted against 50 μg of EcoRI-treated biotinylated driver 2cDNA for a long hybridization followed by 25 μg of driver 1cDNA for a short hybridization, resulting in 4cDNA, which was then PCR amplified. Another similar cycle was performed giving 6cDNA. The subtractive hybridization process was monitored by displaying the cDNA libraries after each round of subtraction on a sequencing gel. The cDNA libraries were amplified using a primer derived from the original linker sequence ligated onto the cDNA fragment ends but modified by deletion of the two 5’ bases and the addition of the 2 mixed bases AT and GC on the 3’ end that overlapped the cDNA insert (5'-CTTGCTTGAATTCGGACTA(A/T)(C/G)-3’), giving 4 possible primers and 16 possible primer pairs. This mixed primer anneals at both ends of the cDNA fragments and because of the two mixed bases at the 3’ end that overlap the cDNA insert, the sequence selectively amplifies a representative cDNA subpopulation such that individual bands are visualized (Fig. 2). The PCR temperature profile used was 94°C (5 min), hot start followed by 94°C (1 min), 50°C (1 min), 72°C (2’’ min) for 30 cycles. The PCR products were resolved on a 5% polyacrylamide gel and visualized by exposure to X-Omat AR film (Kodak, Rochester, NY).

**Trimming Cycles**—The subtracted cDNA libraries were next recognized using a primer derived from the original linker sequence ligated onto the cDNA fragment ends but modified by deletion of the two 5’ bases at the 3’ end that overlap the cDNA insert, the sequence selectively amplifies a representative cDNA subpopulation such that individual bands are visualized (Fig. 2). The PCR temperature profile used was 94°C (5 min), hot start followed by 94°C (1 min), 50°C (1 min), 72°C (2’’ min) for 30 cycles. The PCR products were resolved on a 5% polyacrylamide gel and visualized by exposure to X-Omat AR film (Kodak, Rochester, NY).

**Preparation of cDNA from Mammary Carcinomas**—Regressing (monoterpane-treated) and non-regressing (control, not treated) carcinomas for all studies were generated by the following protocol. Virgin Wistar-Furth inbred rats at 50–55 days of age were given a single dose of DMBA (50 mg/kg body weight) by gastric intubation. Developing carcinomas were followed by palpation until they grew to ~10 mm in diameter, at which time rats were randomized to either a monoterpene diet of 10% (w/w) d-limonene or a control diet and pair-fed. Carcinomas from rats on monoterpene diet that regressed to 50% of their maximum diameter were resected. All remaining carcinomas were collected by 15 weeks after diet randomization. All animal use was in compliance with NIH guidelines for humane care and was approved by the University of Wisconsin-Madison Medical Center Animal Use Committee.

Total tumor RNA was isolated using RNAzol B Reagent (Tel-Test, Friendswood, TX) and poly(A)^+ RNA was isolated using the PolyActract System 1000 (Promega, Madison, WI) following the manufacturer’s guidelines. The Superscript Plasmid System (Life Technologies, Inc., Gaithersburg, MD) was used for double-stranded cDNA synthesis from poly(A)^+ RNA. 5 μg of cDNA synthesized from 7 regressing and 7 non-regressing carcinomas were pooled and termed +cDNA and −cDNA, respectively. A portion of the +cDNA and −cDNA was re-sequenced to confirm the construction of non-subtracted cDNA. The vector pSport1 and transformed by electroporation into Electromax Efficiency DH12S Escherichia coli host cells (Life Technologies, Inc.).

**Subtraction**—We used the PCR-based subtractive hybridization method of Wang and Brown (20) described in detail elsewhere. In brief, the subtraction protocol operates by fragmenting the cDNA with restriction endonucleases (i.e., Alu I and Aci I plus EcoRI digestion to their termini) 21-bp oligodeoxynucleotide linkers having a 5’ blunt end and a 4-base 3’ overhang. The flush end contained an EcoRI site. Following linker ligation, the fragmented cDNAs are subjected to PCR amplification using the 21-bp linkers as primers (sequences given in Ref. 20). It is assumed that cDNA fragments common to both +cDNA and −cDNA populations were amplified equally, while different cDNA fragments within a population may be differentially amplified. Depending on the direction of subtraction, whether up- or down-regulated cDNAs were being enriched, either +cDNA or −cDNA and their derived populations were tracer or driver cDNA. Thus in the following, driver or tracer cDNA is specified but not if +cDNA or −cDNA was used. 50 μg of driver cDNA was completely digested with EcoRI to suppress contaminating driver cDNA from being amplified. Driver cDNA was then photobiotinylated and combined with 2.5 μg of non-photobiotinylated tracer cDNA. The mixture was completely denatured by boiling and then cooled for a long hybridization (20 h). Addition of streptavidin led to formation of streptavidin-biotin-DNA complexes that were removed by several phenol-chloroform extractions. Subtracted tracer cDNA (termed 1cDNA) was again subtracted from 25 μg of biotinylated tracer cDNA, this time for a short hybridization of 2 h to yield 2cDNA. +2cDNA and −2cDNA were PCR amplified. 2.5 μg of non-photobiotinylated tracer 2cDNA was subtracted against 50 μg of EcoRI-treated biotinylated driver 2cDNA for a long hybridization followed by 25 μg of driver 1cDNA for a short hybridization, resulting in 4cDNA, which was then PCR amplified. Another similar cycle was performed giving 6cDNA.

The subtractive hybridization process was monitored by displaying the cDNA libraries after each round of subtraction on a sequencing gel. The cDNA libraries were amplified using a primer derived from the original linker sequence ligated onto the cDNA fragment ends but modified by deletion of the two 5’ bases at the 3’ end that overlapped the cDNA insert (5’-CTTGCTTGAATTCGGACTA(A/T)(C/G)-3’), giving 4 possible primers and 16 possible primer pairs. This mixed primer anneals at both ends of the cDNA fragments and because of the two mixed bases at the 3’ end that overlap the cDNA insert, the sequence selectively amplifies a representative cDNA subpopulation such that individual bands are visualized (Fig. 2). The PCR temperature profile used was 94°C (5 min), hot start followed by 94°C (1 min), 50°C (1 min), 72°C (2’’ min) for 30 cycles. The PCR products were resolved on a 5% polyacrylamide gel and visualized by exposure to X-Omat AR film (Kodak, Rochester, NY).
Identifying Monoterpene-induced/repressed Genes

| TABLE I | Competitive RT-PCR primers and product sizes |
|---------|---------------------------------------------|
| Target DNA | Primers (5'-3') | Product size (bp) |
| LCI | 5'-CCCTACCCCTTCCCTCAATC | 745 |
| LCI MIMIC | 3'-TGCTCTTACACGTCTGCTCAG | |
| TGFβIIR | 5'-TTGTTACACGGGAGAGTAAAA | 414 |
| TGFβIIR MIMIC | 3'-TTGAGTCTAGGGAGGAGGATTT | |
| Neuroligin 1 | 5'-GGCCAAAGTAAAGCAGGAAATTT | 724 |
| Neuroligin 1 MIMIC | 3'-GGGGCCTATACTTCTGCTCCT | |
| GAPDH | 5'-CGCAAGTGAATACCTCCTCCG | 245 |
| GAPDH MIMIC 1 | 3'-TTTCTCCTCCCTOTATAAACA | 526 |
| GAPDH MIMIC 2 | 5'-CGCAAGTGAATACCTCCTCCG | 193 |
| GAPDH MIMIC 1 | 3'-CATGACGTCAACGAAATCTGGGG | 983 |
| GAPDH MIMIC 2 | 5'-CGCAAGTGAATACCTCCTCCG | 604 |
| LC1 | 3'-TTGAGTCTAGGGAGGAGGATTT | 450 |

sette contains BsrEI and BsgI recognition sites configured such that stepwise digestion first with BsgI and then BsrEI produces cleavage of DNA 16/14 nucleotides distal and 4/2 nucleotides proximal, respectively, relative to the 3' end of the cassette, thus conserving the trimming cassette but cutting into the 5' end of the cDNA fragment. The trimming cycle continues by treatment with mung bean nuclease to remove single-stranded ends of the linearized DNA and finally re-circularization of the plasmids with T4 DNA ligase. Noteworthy, cDNA fragments that may contain BsgI or BsrEI are not lost from the library since a portion of the cDNA fragment is retained. Additionally, this trimming procedure was shown to produce deletions with very high efficiency on a population of plasmids (30).

**PCR—**All PCR reactions were carried out using the following conditions unless otherwise stated. In a final reaction volume of 50 μl, PCR was performed using 200 μM of each dNTP, 400 nM of each primer, and 10 μM of each dNTP plus 10 μM of (α-32P)SdATP, and the PCR temperature profile used was 94°C (5 min) hot start followed with 94°C (1 min), 59°C (1 min), 72°C (1 min) for 30 cycles and lastly 72°C (10 min) to ensure double-stranded cDNA. Labeled PCR products were resolved on a 5% polyacrylamide sequencing gel and visualized by a PhosphorImager (Molecular Dynamics, Sunnyvale, CA). These experiments were also performed in parallel by omitting the radiolabel and instead silver staining the gel (31) to visualize the bands directly, which greatly improves efficiency of band isolation. Unique bands were excised, eluted from the gel, and re-amplified using the same fixed PCR primer and a modified subpopulation primer, or pan primer, that amplifies the entire cDNA population. The pan primer has two C bases and an EcoRI site (underlined) added to the 5' end and the two terminal variable 3' bases removed from the subpopulation primer (5'-CGGAAATTCCGGAGGAGGCAGTA-3'). Isolated cDNA fragments were digested with EcoRI, subcloned into the multiple cloning site of the pSPORT1 plasmid (Life Technologies, Inc.), and transformed by electroporation into Electromax Efficiency DH12S E. coli host cells (Life Technologies, Inc.).

**Sequence Analysis—**All subcloned cDNA fragments were sequenced using the M13/pUC forward primer and the PRISM Ready Reaction Dye Deoxy Terminator Cycle Sequencing Kit (Perkin Elmer), following the manufacturer's protocol. The sequencing reaction products were resolved on an ABI PRISM Automated DNA Sequencer (Advanced Biotechnologies Inc., Columbia, MD). The monoterpene-induced gene (MIG) and monoterpene-repressed gene (MRG) sequences were compared to a non-redundant nucleotide sequence data base that includes sequences from the Brookhaven Protein Data Bank, GenBank, GenBank updates, EMBL, and EMBL updates using the BLAST algorithm at the National Center for Biotechnology Information (NCBI) (32).

**Competitive Reverse Transcriptase-PCR (RT-PCR)—**Relative RNA expression levels between regressing (10% limonene-treated) and non-regressing (control, not treated) mammary carcinomas were determined by competitive RT-PCR (33-38). The competitive RT-PCR assay involved PCR of a target gene (i.e. a MIG/MRG cDNA) competing for amplification with its exogenously added MIMIC, and amplification of an internal standard (i.e. glyceraldehyde 3-phosphate dehydrogenase (GAPDH)) competing for amplification with its exogenously added MIMIC, for a total of 4 PCR products per sample. A MIMIC is a piece of heterologous DNA with flanking sequences identical to the primer sequences used for the target gene, such that a known amount of MIMIC is exogenously added, which competes for simultaneous co-amplification with a target gene. In essence, the competitive RT-PCR assay normalizes the MIG/MRG cDNA to an external control (MIG/MRG MIMIC) and to an internal GAPDH control, which itself was normalized to an external control (GAPDH MIMIC). After the target MIG/MRG expression for each sample was normalized, its expression level in each sample of the control and regressing carcinoma groups was averaged and compared in order to determine fold-induction or fold-repression. Statistical significance was calculated using the two-tailed Student's t test.

Relative RNA expression levels for lipocortin 1 (LC1), transforming growth factor β type II receptor (TGF-βIIR), and neuroligin 1 were determined by competitive RT-PCR. RNA was prepared from five individual regressing (10% limonene-treated) and five individual non-regressing (control, not treated) mammary carcinomas as described above. cDNA was generated using the 1-Stand cDNA Synthesis Kit (Clontech Laboratories, Palo Alto, CA). Primers were designed using Oligo 5.0 Primer Analysis software (National Biosciences, Plymouth, MN), with the exception of the GAPDH primers (Clontech Laboratories). MIMICs were generated using the PCR MIMIC Construction kit (Clontech Laboratories). Generation of MIMICs involves a 1° PCR reaction using a composite primer, consisting of target gene or gene-specific primer sequence at the 5' end and heterologous DNA primer sequence on the 3' end. First, an aliquot of the 1° PCR reaction was used as template in a 2° PCR reaction using only target gene or gene-specific primers. Primers and dNTPs were removed by passing the 2° PCR reaction through a Chromatin-100 column (Clontech Laboratories) and the MIMIC was quantified by spectrophotometric analysis. Primer sequences and product sizes are given in Table I. In a reaction volume of 50 μl, the final concentration of exogenously added PCR MIMICs were 4.0 × 10⁻² amol/μl of LC1 MIMIC, 8.0 × 10⁻⁵ amol/μl of TGF-βIIR MIMIC, 4.0 × 10⁻⁶ amol/μl of neuroligin 1 MIMIC, and 4.0 × 10⁻⁵ amol/μl GAPDH MIMIC 1 or GAPDH MIMIC 2. The PCR temperature
amplification of the cDNA libraries after each round of subtraction. We monitored the subtractive process by PCR ing long (20 h) and short (2 h) hybridizations for a total of 6 MRG. The cDNA libraries were subjected to cycles of alternating endonucleases and ligation of 21-bp oligodeoxynucleotide linkers onto the cDNA fragment ends, thus allowing for a randomization of DNA sequence. Subtracted, reconfigured cDNA was further tested by isolating 2 prominent bands from both subtracted cDNA lanes (cross-hatched box) or asymmetric linkers. Subtracted, reconfigured cDNA libraries are PCR displayed using 16 sets of primers including 1 fixed primer and 1 of 16 subpopulation primers. The subpopulation primer overlaps the cDNA insert by 2 bases (designated by NN), which selects for amplification of a subset of cDNAs.

Results
An overview of the SD methodology is presented in Fig. 1. An evaluation of each aspect of this technology is presented below.

Subtraction of cDNA Libraries—The subtractive hybridization procedure developed by Wang and Brown (20) served to remove commonly expressed cDNA fragments and enrich those that are differentially expressed. The subtraction process involved initial fragmentation of the cDNA libraries with restriction endonucleases and ligation of 21-bp oligodeoxynucleotide linkers onto the cDNA fragment ends, thus allowing for a renewable source of cDNA by PCR. The monoterpane-treated derived cDNA (+cDNA) was subtracted against control, non-treated derived cDNA (−cDNA) to enrich for MIG. The inverse subtraction, control, non-treated cDNA subtracted against monoterpene-treated cDNA, was also performed to enrich for repressed genes. Each round of subtraction involved fragmentation of the cDNA libraries with restriction endonucleases and ligation to symmetrical linkers (boxed boxes) for PCR amplification. +cDNA and −cDNA are subjected to multiple rounds of subtractive hybridization and PCR amplification to enrich for both induced and repressed genes. Subtracted libraries are reconfigured by subcloning into a trimming vector and subjected to 3 cycles of precise, unidirectional trimming which serves to randomize DNA sequence (shaded box) and provide a different linker on one end of the cDNA fragments (cross-hatched box) or asymmetric linkers. Subtracted, reconfigured cDNA libraries are PCR displayed using 16 sets of primers including 1 fixed primer and 1 of 16 subpopulation primers. The subpopulation primer overlaps the cDNA insert by 2 bases (designated by NN), which selects for amplification of a subset of cDNAs.

profile used for all competitive RT-PCR experiments was 94 °C for 1 min, 60 °C for 1 min, 72 °C for 2 min for 30 cycles, with the exception of LCl, which required an annealing temperature of 58 °C. PCR products were resolved on a 2% NuSieve GTG-agarose gel (FMC BioProducts, Rockland, ME). Primers were 5' end-labeled with fluorescein during their synthesis, thereby allowing PCR product quantitation using a FluorImager (Molecular Dynamics).

Fig. 1. Flowchart of the subtractive display methodology. cDNA libraries are generated from matched tissues and designated +cDNA (monoterpane-treated) and −cDNA (control, non-treated) cDNA (solid box) is fragmented with restriction endonucleases and ligated to symmetrical linkers (boxed boxes) for PCR amplification. +cDNA and −cDNA are subjected to multiple rounds of subtractive hybridization and PCR amplification to enrich for both induced and repressed genes. Subtracted libraries are reconfigured by subcloning into a trimming vector and subjected to 3 cycles of precise, unidirectional trimming which serves to randomize DNA sequence (shaded box) and provide a different linker on one end of the cDNA fragments (cross-hatched box) or asymmetric linkers. Subtracted, reconfigured cDNA libraries are PCR displayed using 16 sets of primers including 1 fixed primer and 1 of 16 subpopulation primers. The subpopulation primer overlaps the cDNA insert by 2 bases (designated by NN), which selects for amplification of a subset of cDNAs.

Fig. 2. PCR product display during the subtractive hybridization process. A subpopulation of the cDNA libraries was selectively amplified using a primer of sequence 5' -CTTGCTTGAAATTGGGAC-TA/AT/GC-3', which anneals at both ends of the cDNA fragments and overlaps two bases of the cDNA insert. Lanes alternate monoterpene-treated regressing (+) and control non-regressing (−) carcinoma-derived cDNA. Beginning on the left, lanes labeled + and − are cDNAs before subtraction, and the numbered lanes toward the right are cDNAs after each consecutive round of subtraction. The arrow indicates the band identified as the YWK-II gene. PCR products were labeled with [α-35S]dATP and resolved on a 5% polyacrylamide sequencing gel and visualized on x-ray film.

identification, before the libraries were reconfigured using pTRIM14, and displayed them on a sequencing gel. A representative cDNA subpopulation was PCR amplified by modifying the 21-bp primer, which anneals to the linkers initially ligated onto the fragment ends, such that the two 5' bases were removed and two mixed bases were added to the 3' end, thus overlapping the cDNA insert, as described in detail under “Experimental Procedures.” Fig. 2 demonstrates the efficiency of the subtractions. The unsubtracted +cDNA and −cDNA lanes showed smears, while with each additional cycle of subtraction, individual bands became increasingly prominent. Furthermore, the banding patterns between the subtracted +cDNA and the subtracted −cDNA lanes were reproducible, indicating that specific cDNA fragments were increasingly enriched while others were driven out of the population. Efficiency of subtraction was further tested by isolating 2 prominent bands from both the +cDNA and −cDNA lanes, followed by subcloning and sequencing the cDNAs by similar methods of clone characterization discussed below. One of the derived +cDNA fragments was identified by a nucleotide data base search using the BLAST algorithm (32) as the YWK-II gene, which was again identified at the end of the SD process (indicated by an arrow; compare Figs. 2 and 5) and shown to be differentially expressed as discussed below.
Identifying Monoterpane-induced/repressed Genes

Subtracted, Reconfigured cDNA Library + Subpopulation PCR Primer Set

![Diagram of cDNA fragment and primers used in subtractive display](Image)

**FIG. 3.** Schematic of a cDNA fragment and primers used in subtractive display. The schematic represents a cDNA fragment processed through 6 rounds of subtractive hybridization and 3 cycles of trimming (±16 bp cDNA). The unique linker sequences at each end of the cDNA fragment are boxed; the bold sequence in the right linker is an EcoRI recognition site. The shaded region of the cDNA insert represents 2 randomized base pairs, after trimming, which overlap the 2 bases (NN) at the 3' end of the subpopulation primer.

Reconfiguration and Display of the Subtracted Libraries—
The subtracted cDNA libraries were next reconfigured for the display step. The libraries were reconfigured because during their construction both ends of the cDNA fragments were ligated to the same linker, i.e. the ends are symmetrical. Symmetrical primer sites limit the ability to display the subtracted cDNA libraries because a primer designed to overlap the cDNA insert by 2 bases will anneal on both ends of the cDNA fragment, thus not permitting all 2-base independent combinations at both cDNA insert ends for complete screening of the libraries. In addition, the first few bases at both ends of the cDNA fragments in the subtracted libraries are not random but reflect both the methods used to make the cDNA (priming the poly(A) tail with a sequence containing a NotI site) and that used to fragment the cDNA (restriction digestion with two 4-base cutters). One end of each cDNA fragment needs to be a random sequence, such that a subpopulation primer will selectively amplify a small number of distinct bands during the PCR-display step, as discussed below. Reconfiguration of the subtracted cDNA libraries was accomplished by cloning the subtracted libraries into pTRIM14 (30) and processing the cDNA libraries through 3 cycles of precise, unidirectional, 14 bp trimming, as discussed under “Experimental Procedures.”

pTRIM14 was constructed with a trimming cassette composed of BsgI and BseRI recognition sites. A trimming cycle operates by sequential digestion with BsgI and then BseRI to produce cuts 16/14 nucleotides distal and 4/2 nucleotides proximal, respectively, relative to the 3' end of the trimming cassette, thereby producing a deletion onto the 5' end of the cDNA fragment but retaining the trimming cassette. Single-stranded DNA ends are removed by digestion with mung bean nuclease followed by ligation with T4 DNA ligase. The subtracted cDNA libraries were subjected to three trimming cycles, generating 6/4 bp trimming, as discussed under “Experimental Procedures.”

The reconfigured, subtracted cDNA libraries were displayed by consecutive amplification of cDNA subpopulations using 16 sets of PCR primers (Figs. 3 and 5). In each primer set, one primer is always the same or fixed and anneals at the unmodified or right (3') end of the cDNA fragment. This fixed primer is a 17-mer with a sequence of 5'-ATTAGAATTCCGACTA-3', and comprised in the 5'-3' direction of pTRIM14 sequence, an EcoRI site (underlined), and original linker sequence. The second primer in each primer set anneals on the randomized or left (5') end of the cDNA fragment and is termed the subpopulation primer. The 5' end of the subpopulation primer anneals to 14 bp of pTRIM14 sequence, and its 3' end overlaps the cDNA fragment by 2 bp. These two 3' bases in each individual subpopulation primer will consist of 1 of the 16 possible 2-base combinations, resulting in a subpopulation primer sequence of 5'-GAGGAGGTGCAGTANN-3' (Fig. 2). Therefore, 1 of a total of 16 sets of PCR primers results in selective annealing and amplification of a subpopulation of the cDNA, and consecutive amplification with each of the 16 sets of primers allows complete screening of the libraries.

In order to maximize subpopulation selective amplification and minimize redundancy of clone isolation, the dNTP concentration and annealing temperature of the SD reactions were optimized. Optimization of dNTP concentration and annealing temperature increases the number of unique bands in one subpopulation and reduces the number of bands that appear multiple times across various subpopulations due to mispriming of the two terminal 3' bases of the subpopulation primer. Initial SD optimization experiments used representative sets of SD primers including the fixed primer and the subpopulation primers CC, AC, AG, and GA, as specified by the variable bases of the subpopulation primer (or the two 3' penultimate bases of the sequence 5'-GAGGAGGTGCAGTANN-3'), at dNTP concentrations of 200, 20, 10, and 2 μM, all at an annealing temperature of 59°C (data not shown). The 20 μM SD reaction amplified the cDNA fragments efficiently with moderate band redundancy, while the 10 μM SD reaction amplified the cDNA fragments with lower efficiency but less band redundancy. SD conditions were further refined using the subpopulation primer sets AC, AG, and GA because of their similar sequence. These three primer sets were used in SD experiments under annealing conditions of 57, 59, 61, and 63°C using 10 and 20 μM dNTP (Fig. 4). SD reaction conditions of 20 μM dNTP and an annealing temperature of 59°C resulted in reproducibly efficient amplification with relatively low redundancy of bands.

SD reactions were performed with all 16 sets of subpopulation primers using both 10 μM dNTP (data not shown) and 20 μM dNTP (Fig. 5). Both experiments were performed in parallel, with radiolabel and without radiolabel in order to isolate bands by silver staining (31). Unique bands from the SD reactions were excised, eluted from the gel, reamplified using the same fixed PCR primer and a pan primer that amplifies all cDNA fragments (sequence given under “Experimental Procedures”), and PCR products were subcloned into pSport1. Individual subcloned cDNAs were designated as a MIG or MRG by virtue of the cloning process (i.e. whether the cDNA originated from the monoterpane-treated +cDNA or the untreated control –cDNA, respectively).

**Characterization of Differentially Expressed Clones**—We isolated 42 MIG and 58 MRG rat cDNAs (Table II), all of which were sequenced and compared against sequence data bases using the BLAST algorithm for possible identification. A MIG/MRG clone was defined to be a known gene if it shows ≥95% sequence identity over the entire cDNA fragment with a known
gene from the sequence data base. The cDNA insert sizes ranged from 45 to 312 bp due to the methods used for library generation, modification, and cloning. Considering the MIG cDNAs as one group, 9 were known, of which 3 were identified by more than 1 MIG cDNA. Additionally, 33 MIG cDNAs were not identified in the sequence data base. Considering the MRG cDNAs as one group, 1 was a known gene and 57 MRG cDNAs were not identified in the sequence data base. The identification of genes was limited due to our strict criteria for sequence identity plus the fact that our sequences were from rat whose representation in the data bases is limited.

More than 1 cDNA fragment, varying in molecular weight, identified cytochrome c oxidase subunit II, sperm membrane protein YWK-II (39), and Y89 (59 ends similar to YWK-II) (Table II). The multiple cDNA fragments that identified each of these genes had common sequences but differed in the size of the cDNA fragment; for example, of the 6 cDNAs that identified cytochrome c oxidase subunit II, all the clones shared 70 bp while the largest clone (MIG-27) had an additional 32 bp not found in the smallest clone (MIG-3).

The potential of the SD screen to identify differentially expressed genes was suggested by the isolation of Man-6-P/IGF2R as MIG-42. We have previously shown that the Man-6-P/IGF2R was induced 2.0-fold (p < 0.0002) at the RNA level in 10% limonene-treated, regressing mammary carcinomas using an RNase Protection Assay (Table III; Ref. 19).

We next tested the differential expression of other identified MIG/MRG clones. We initially confirmed that YWK-II, a gene identified by 7 separate MIG clones (Table II), was induced 2.9-fold (p < 0.00006) in a panel of 10% limonene-treated regressing carcinomas relative to control non-regressing carcinomas (Table III). Expression was quantitated by Northern blot analysis and a PhosphorImager (Molecular Dynamics). YWK-II was also found to be most highly expressed in rat brain, moderately expressed in normal mammary gland, liver, and kidney, and at low levels in the spleen (data not shown). However, Northern blots and RNase Protection Assay use ~1 µg of poly(A)+ RNA or 10 µg of total RNA per sample. In order to conserve regressing tumor RNA, which is quite limited, we used competitive RT-PCR (48–53) to determine relative fold-induction or fold-repression of the remaining MIG/MRG cDNAs.

We tested the differential expression of MIG-12, identified as lipocortin 1 (LC1; Ref. 40), using the competitive RT-PCR assay described under "Experimental Procedures" (Fig. 6A). LC1 was induced 2.9-fold (p < 0.00003) in 10% limonene-treated regressing carcinomas relative to control non-regressing carcinomas (Fig. 6B, Table III). We then examined LC1 RNA expression in the normal rat mammary gland by the same assay and found that LC1 was induced 3.1-fold (p < 0.0005) in the involuting mammary gland relative to the virgin mammary gland (data not shown). LC1 has been shown to be a marker for apoptosis in the involuting rat mammary gland (40).

The transforming growth factor β type II receptor (TGF-βIIIR; 41–44) was identified by MIG-33 and was induced 3.1-fold (p < 0.0002) in 10% limonene-treated regressing carcino-
mas relative to control non-regressing carcinomas (Table III) by the competitive RT-PCR assay. The SD screen identified neuroligin 1 as the clone MRG-31 (45). The competitive RT-PCR assay demonstrated neuroligin 1 expression to be repressed 8.8-fold in one of five or not detectable in four of five of 10% limonene-treated regressing carcinomas as compared to control carcinomas (Table III).

DISCUSSION

The Subtractive Display Gene Expression Screen—In order to better understand the mechanism by which monoterpenes mediate tumor regression, we sought to identify differential gene expression patterns of actively regressing, advanced DMBA-induced rat mammary carcinomas treated with limonene relative to control (not treated) carcinomas. Due to the complexity of regressing tumor tissue, we required a very sensitive gene expression screening method that could detect relatively small alterations in gene expression levels and/or genes which may be expressed in low abundance. Because of these concerns, we integrated the strengths of subtractive hybridization and differential display into a new methodology termed SD.

Wang and Brown’s (20) subtractive hybridization method generates cDNA for subtraction using PCR; however, clones are isolated by classical probe-hybridization techniques that require an undefined number of probing cycles. In contrast, SD cloning of enriched cDNAs is PCR-based and is therefore faster, more efficient, and more sensitive. Unlike Liang and Pardee’s (21) differential display where the 59 primer preferentially anneals upstream on the cDNA at a low temperature, the SD method uses a 16-base subpopulation primer and a 17-base fixed primer, each annealing at defined left and right linkers, respectively, on the cDNA fragment ends. In addition, annealing is performed at the highest temperature that still permits PCR to proceed (59 °C). The combination of larger primers annealing at defined sites and a very high annealing temperature both contribute to high specificity and stringency, leading to high reproducibility between identical reactions. In addition, by primer design, the libraries are completely screened in 16 sets of reactions. Furthermore, because the cDNA libraries are subtracted before the display step, background noise is significantly reduced, thereby increasing reproducibility and detection of rare cDNAs. Moreover, there exists a potential for noise with differential display because differential display depends on precise registry of cDNA fragments in paired lanes such that unique bands are isolated; yet identical cDNA fragments in treated and non-treated populations may both be amplified but differ in size by a few base pairs. Alternatively, with SD, since all cDNA fragments are putatively differentially expressed because of the subtraction step, the precise registry of cDNA fragments in paired gel lanes is not necessary to identify the differentially expressed genes.

The SD screen identified the Man-6-P/IGF2R, which was already known to be induced 2-fold at an RNA level in the epithelium of regressing carcinomas (19). This initially suggested that SD had high sensitivity. Additionally, every MIG/
MRG clone tested for differential gene expression exhibited consistent differential expression across a panel of regressing carcinomas relative to non-regressing carcinomas, thus demonstrating the ability of SD to identify a high percentage of true differentially expressed clones. It should be noted that identification of a single gene by multiple clones versus one clone does not indicate the degree of differential expression; YWK-II was identified by 7 separate bands varying in molecular weight but showed a degree of induction similar to Man-6-P/IGF2R, TGF-βIIR, and LC1, all of which were identified by one clone (Table II). However, YWK-II may be more abundantly expressed. The multiple isolation of cDNA fragments that identify the same gene may have resulted from restriction digestion of the original cDNA before subtraction with AluI and AluI plus RsaI, and/or cloning artifacts. Furthermore, band intensity does not indicate degree of differential expression, as there was no discernible difference in band intensity which identified neuroligin 1 and bands that identified the other known clones.

The Process of Monoterpene-mediated Mammary Carcinoma Regression—Based upon histopathological analysis of monoterpene-treated regressing mammary carcinomas, the phenomenon of tumor regression was postulated to involve a differentiation/remodeling process (6). The identified MIG/MRG cDNAs are consistent with such a mechanism of regression. The present study is consistent with the hypothesized involvement of the TGF-β signaling pathway with tumor regression. The Man-6-P/IGF2R and the TGF-βIIR were identified as up-regulated. The Man-6-P/IGF2R facilitates latent-TGF-β1 activation (18, 46) and trafficks IGF2, a potent mammary carcinoma mitogen, into lysosomes for degradation (reviewed in Ref. 47). Also the Man-6-P/IGF2R has been reported to be a tumor suppressor gene in both liver and breast cancer (48, 49). Interestingly, the SD screen identified the TGF-βIIR, since it is postulated that breast tumors may not respond to TGF-β because the tumors express low levels TGF-βII receptors (50, reviewed in Ref. 51).

The SD screen isolated LC1, which is a marker for apoptosis during mammary gland involution. At the protein level, LC1 was induced 10-fold only in the alveolar epithelium, but remained at basal levels in ductal epithelium and stroma during involution, as determined by immunohistochemistry (40). Because RNA expression of LC1 in the context of the entire involuting mammary gland was not reported, we determined LC1 RNA expression during involution relative to the virgin mammary gland and found LC1 was induced 3.1-fold. It should be noted that most overexpression of genes thus far analyzed was in the 2–3-fold range. In some cases this could be an underestimate due to the fact that our analysis methods averaged gene expression over cell types in a very heterogeneous tissue. For example, LC1 was 3.1-fold overexpressed in total involuting mammary gland (RNA) but was 10-fold (protein) overexpressed in the alveolar cell lineage and not overexpressed in the ductal cell lineage. In other cases it is likely that a 2-fold increase could be biologically meaningful in vivo. For

### Table III

| Gene          | Fold induction/repression | Quantitation method |
|---------------|---------------------------|---------------------|
| Man-6-P/IGF2R | 2.0-fold induced (p ≤ 0.002) | RPA (data from Ref. 19) |
| YWK-II        | 2.9-fold induced (p ≤ 0.00006) | Northern blot |
| LC1           | 2.9-fold induced (p ≤ 0.00003) | Competitive RT-PCR |
| TGF-βIIR      | 3.1-fold induced (p ≤ 0.0002) | Competitive RT-PCR |
| Neuroligin 1  | 8.8-fold repressed (1/5); ND (4/5) | Competitive RT-PCR |

**Fig. 6.** LC1 (MIG-12) RNA expression using competitive RT-PCR. A, the differential expression of LC1 between a panel of control non-regressing (lanes labeled CON1-CON5) and 10% limonene-treated regressing (lanes labeled LIM1-LIM5) carcinomas was demonstrated by competitive RT-PCR. PCR products corresponding to GAPDH, LC1, GAPDH MIMIC, and LC1 MIMIC, and their respective sizes (in bp) are indicated. PCR primers were 5′-labeled with fluorescein; products were resolved on a 2% agarose gel and visualized on a FluorImager. B, LC1 was induced 2.9-fold in 10% limonene-treated regressing carcinomas compared to control non-regressing carcinomas. The LC1 expression was quantified from the gel in panel A by first normalizing the LC1 and GAPDH bands to their respective MIMICs and second by normalizing LC1 to GAPDH.
example, we detected a 2.0-fold increase in the Man-6-P/IGF2R, an imprinted gene in rodents (52), suggesting that a 2-fold difference in its expression has been evolutionarily selected and conserved.

Both the induced YWK-II expression and repressed neurologin 1 expression could play a role in tumor regression through a differentiation process. YWK-II, a transmembrane protein, is expressed in the mammary gland (data not shown) and has been shown to be a marker for differentiation of spermatogonia (53). It modulates cell-cell adhesion, as in sperm-egg adhesion during fertilization (54). Also, the polypeptide sequences of the YWK-II transmembrane and cytoplasmic domains are 70.6% homologous to the same domains of the human A4 amyloid protein found in brain plaques of Alzheimer disease patients (39). Differential gene expression of the A4 amyloid protein is associated with retinoic acid-induced morphologic differentiation of neureit processes in two model systems; the A4 amyloid mRNA is induced 34-fold in retinoid-treated P19 embryonal carcinoma cells (55) and 10-fold in retinoid-treated SH-SY5Y neuroblastoma cells (56). Neurologin 1 is a neuronal cell surface protein found enriched in synaptic plasma membranes and binds to brain-specific β-neurexins as a ligand. Expression of neurologin 1 was tested in various tissues, but was restricted to brain tissue (45). Therefore, neurologin 1 expression in mammary carcinomas may reflect deregulated gene expression. Interestingly as carcinomas regress, neurologin 1 expression was repressed or turned off, which is consistent with regulated gene expression in most normal tissue.

The SD screen was applied to monoterpene-treated mammary carcinomas resected at mid-regression. Therefore, the identified MIG/MRG cDNAs reflect differential gene expression patterns consistent with apoptosis and differentiation of actively regressing carcinomas. In order to better understand the process by which monoterpenes first initiate tumor regression, we are currently identifying early MIG/MRG cDNAs using SD on mammary carcinomas treated with monoterpenes for 24 h.

In summary, the observed alterations in gene expression help to better define the process associated with monoterpenemediated tumor regression. Although the data presented correlate monoterpene-treatment, induction of TGF-β signaling components, and tumor regression, the data do not show that TGF-β signaling is a causal event in monoterpenemediated tumor regression. It is, however, consistent with our working hypothesis that monoterpenes promote the up-regulation of the Man-6-P/IGF2R, thereby causing increased levels of activated TGF-β1 available for ligand binding to the up-regulated TGF-βRII and initiating mitogenicatory and apoptotic signaling. We are currently investigating this hypothesis by evaluating whether the Man-6-P/IGF2R up-regulation is the central causal event that triggers tumor regression.

Acknowledgments—We thank Wendy S. Kennan and Jill D. Haag for assistance with generation and collection of tissue, Dr. Laurie A. Shepel for advice during the course of this investigation and for critical review of this manuscript, and Jennifer L. Tymrak for critical review of this manuscript.

REFERENCES

1. Crowell, P. L., and Gould, M. N. (1994) Crit. Rev. Oncog. 5, 1–22
2. Gould, M. N. (1995) J. Cell Biochem. Suppl. 22, 139–144
3. Crowell, P. L., Kennan, W. S., Haag, J. D., Ahmad, S., Vedejs, E., and Gould, M. N. (1992) Carcinogenesis 13, 1261–1264
4. Eglehede, J. A., Elson, C. E., Qureshi, A., Tanner, M. A., and Gould, M. N. (1994) Carcinogenesis 15, 641–646
5. Elson, C. E., Maltzman, T. H., Boston, J. L., Tanner, M. A., and Gould, M. N. (1989) Carcinogenesis 9, 331–332
6. Haag, J. D., Lindstrom, M. J., and Gould, M. N. (1992) Cancer Res. 52, 4021–4028
7. Haag, J. D., and Gould, M. N. (1994) Cancer Chemother. Pharmacol. 34, 477–483
8. Maltzman, T. H., Hurt, L. M., Elson, C. E., Tanner, M. A., and Gould, M. N. (1989) Carcinogenesis 10, 781–783
9. Gould, M. N., Moore, C. J., Zhang, R., Wang, B., Kennan, W. S., and Haag, J. D. (1994) Cancer Res. 54, 3540–3443
10. Russin, W. A., Eoesty, J. D., Elson, C. E., Tanner, M. A., and Gould, M. N. (1989) Carcinogenesis 10, 2161–2164
11. McNamee, D. (1993) Lancet 342, 801
12. Shi, W., and Gould, M. N. (1995) Cancer Lett. 95, 1–6
13. Crowell, P. L., Chang, B. R., Haag, J. D., Elson, C. E., and Gould, M. N. (1991) J. Biol. Chem. 266, 17679–17685
14. Crowell, P. L., Ren, Z., Lin, S., Vedejs, E., and Gould, M. N. (1994) Biochem. Pharmacol. 47, 1405–1413
15. Gelb, M. H., Tamao, F., Yokoyama, K., Ghomashchi, F., Eoesty, J. D., and Gould, M. N. (1995) Cancer Lett. 91, 169–175
16. Roy, Z., and Gould, M. N. (1994) Cancer Lett. 76, 185–190
17. Konig, G., Masters, C. L., and Beyreuther, K. (1990) J. Neurochem. 58, 1863–1873
18. Konig, G., Masters, C. L., and Beyreuther, K. (1990) FEBS Lett. 269, 305–310