Cytogenet Genome Res 97:120–127 (2002)

Cytogenetic and Genome Research

Cloning, expression, and chromosomal localization of the mouse gene *(Scgb3a1*, alias *Ugrp2)* that encodes a member of the novel uteroglobin-related protein gene family

T. Niimi,[a] N.G. Copeland,[b] D.J. Gilbert,[b] N.A. Jenkins,[b] A. Srisodsai,[a] D.B. Zimonjic,[c] C.L. Keck-Waggoner,[c] N.C. Popescu,[c] and S. Kimura[a]

[a] Laboratory of Metabolism, [c] Laboratory of Experimental Carcinogenesis, National Cancer Institute, National Institutes of Health, Bethesda MD;
[b] Mouse Cancer Genetics Program, National Cancer Institute-Frederick, Frederick MD (USA)

Abstract. The mouse UGRP gene family consists of two genes, *Ugrp1* and *Ugrp2*. In this study, the genomic structure and expression patterns of *Ugrp2* and its alternative spliced form were characterized. The authentic *Ugrp2* gene has three exons and two introns, similar to the *Ugrp1* gene, which produces a secreted protein. The *Ugrp2* variant uses a sequence located between authentic exons 1 and 2, resulting in a cytoplasmic form due to a termination codon within the inserted sequence. Both mouse and human UGRP2 mRNAs are expressed in lung. In the case of human, the mRNA is expressed at the highest level in trachea, followed by salivary gland at a level similar to lung. Weak expression was also found in fetal lung and mammary gland. *Ugrp2* was mapped by fluorescence in situ hybridization to mouse chromosome 11A5–B1 and human chromosome 5q35. These regions are known to be homologous. Interspecific mouse backcross mapping was also performed to obtain further detailed localization of mouse *Ugrp1* and *Ugrp2*.

A mouse gene encoding the Uteroglobin-related protein (UGRP) 1, its homologous gene *Ugrp2* and their respective orthologous genes in human have recently been identified (Niimi et al., 2001). The mouse UGRP1 has 33, 41, and 81% amino acid identity to mouse and human UGRP2, and human UGRP1, respectively, suggesting that they comprise a new gene family (Niimi et al., 2001). Their nucleotide and amino acid sequences showed similarity to members of the uteroglobin/Clara cell secretory protein (UG/CCSP) gene superfamily, recently termed secretoglobins (SCGB) (Mukherjee and Chilton, 2000). For instance, mouse UGRP1 has an overall amino acid sequence identity of 25, 18, and 27% to mouse uteroglobin/clara cell secretory protein as the prototypical member of the family, human mammaglobin A, and rat prostatein C3, respectively (Niimi et al., 2001). Based on the HUGO standardized nomenclature, UGRP1 and 2 are members 2 and 1 of the secretoglobin family 3A, and the genes are designated as *Scgb3a2* and *Scgb3a1* for mouse and SCGB3A2 and SCGB3A1 for human, respectively (http://www.gene.ucl.ac.uk/nomenclature/genefamily/scgb.html).

The mouse *Ugrp1* gene spans approximately 3.0 kb on the C-D region of mouse chromosome 18 as determined by fluorescence in situ hybridization (FISH) (Niimi et al., 2001). This region is homologous to human chromosome 5q31→q34 (Searle et al., 1989; DeBry and Seldin, 1996), to which the human UGRP1 gene (SCGB3A2) has been assigned (Niimi et al., 2002). This region has been known to contain one of the asthma susceptibility genes (Postma et al., 1995; Ruffilli and

Supported in part by the National Cancer Institute, DHHS.

Received 20 February 2002; manuscript accepted 29 March 2002.

Request reprints from Shioko Kimura, PhD, Bldg. 37, Rm. 3E-24
National Institutes of Health, Bethesda, MD 20892-4258 (USA)
telephone: 301-496-0958; Fax: 301-496-8419; e-mail: shioko@helix.nih.gov

Current address of T.N.: Sekiguchi Biomatrix Signaling Project
ERATO Japan Science and Technology Corporation
c/o Aichi Medical University, Aichi 480-1195 (Japan).

Bonini, 1997; Cookson and Moffatt, 2000; Ober and Moffatt, 2000). The *Ugrp1* gene is composed of three exons and two introns, and alternative splicing produces two additional transcripts that either partially or completely retain the intron 2 sequence (Niimi et al., 2001). RT-PCR analysis revealed that these two splicing variants are expressed at very low levels. Promoter analysis demonstrated that a homeodomain transcription factor T/EBP (thyroid-specific enhancer-binding protein)/NKX2.1, also called TTF1 (thyroid transcription factor 1) binds two sites between –182 and –120 bp of the *Ugrp1* gene promoter, and activates transcription of the gene (Niimi et al., 2001).

UGRP1 is a homodimeric secretory protein of approximately 10 kDa and is mainly expressed in the lung and trachea (Niimi et al., 2001). The protein properties and the site of tissue-specific expression resemble those of UG/CCSP (Niimi et al., 2001), which is believed to function as an anti-inflammatory agent (Broeckaert and Bernard, 2000; Mukherjee and Chilton, 2000). Previous studies using a Th2 cytokine-based antigen model demonstrated decreased expression of *Ugrp1* and *Ugrp2* in inflamed mouse lungs, suggesting that UGRP1 and UGRP2 may also play a role in lung inflammation (Niimi et al., 2001). In fact, our recent studies demonstrated that the human UGRP1 gene is an asthma susceptibilty gene (Niimi et al., 2002).

In this study, the mouse *Ugrp2* gene was characterized by determination of gene structure, the presence of alternative splicing, expression pattern, and chromosomal localization by FISH and interspecific mouse backcross mapping.

Materials and methods

Cloning and DNA sequencing

The mouse *Ugrp2* gene was identified by searching the GenBank nucleotide sequence databases including expressed sequence tags (ESTs) for sequences similar to the mouse *Ugrp1* gene using the BLAST program (National Center for Biotechnology Information, Bethesda, MD). An EST clone AI391046 containing an apparent entire coding sequence that exhibited a nucleic acid sequence similar to the *Ugrp1* gene, was designated *Ugrp2*. A partial cDNA sequence of the *Ugrp2* (EST AI391046) was obtained by RT-PCR using mouse adult lung RNA as a template. This was then used to isolate a full-length cDNA by screening a mouse lung cDNA library in the λZAPII vector (Stratagene, La Jolla, CA). Hybridization was carried out at 65 °C in 6× SSC, 0.5 % SDS, 5× Denhardt's, 0.1 mg/ml of denatured salmon sperm DNA for 16 h. The membrane was washed twice with 2× SSC containing 0.1 % SDS at room temperature for 10 min and once with 0.1× SSC containing 0.1 % SDS at 55 °C for 30 min. A positive plaque was subjected to secondary and tertiary screening. A cDNA encoding human UGRP2 was isolated by RT-PCR using RNA prepared from adult human lung (Ambion, Austin, TX) and primers designed based on EST sequence AW974727 that exhibited similarities to the mouse *Ugrp2* cDNA sequence. Full-length mouse and human UGRP2 cDNA sequences are available from GenBank with the accession nos. AF313456 and AF313458, respectively.

Mouse and human genomic DNAs were isolated from a mouse and a human BAC genomic library (Incyte Genomics, St. Louis, MO) using labeled mouse and human cDNAs as probes, respectively. Mouse and human UGRP2 cDNAs and genomic DNAs were digested with restriction enzymes, subcloned into pBluescript II (Stratagene), and sequenced using an ABI prism dye terminator cycle sequencing ready reaction kit and a model 377 DNA sequencer (PE Applied Biosystems, Foster City, CA).

DNA genomic sequence analysis was carried out using the Human Genome database (National Center for Biotechnology Information) and Celera Discover System and Celera's associated databases. The nucleotide sequences reported in this paper appear in the GenBank nucleotide sequence database with the accession numbers AF313456, AF313457 and AF313458 for mouse *Ugrp2* type A and B mRNAs, and human UGRP2, respectively.

RNA analyses

A multiple mouse tissue Northern blot (Clontech Laboratories, Palo Alto, CA) and a human multiple tissue expression (MTETM) array (Clontech) were hybridized with a full-length mouse and human UGRP2 cDNA as probes, respectively. Hybridization was performed in ExpressHyb™ Hybridization Solution (Clontech) at 68 °C for 2 h. The membrane was washed twice with 2× SSC containing 0.1 % SDS at room temperature for 10 min and twice with 0.1× SSC containing 0.1 % SDS at 55 °C for 20 min, followed by exposure to X-ray film at –80 °C.

For reverse transcription of mRNAs, 2 μg of total RNA was pretreated with DNase I, incubated for 10 min at 70 °C, and chilled on ice. The reactions were carried out in a final volume of 20 μl containing RNA, 4 μl of 5× first strand synthesis buffer (Invitrogen Life Technologies, Carlsbad, CA), 1 μl of a mixture of four dNTPs (2.5 mM each), 2 μl of 0.1 M dithiothreitol (DTT) and 100 ng of random primers. After incubation at 37 °C for 2 min, 200 units of SuperScript II reverse transcriptase (Life Technologies) was added and the incubation continued for 60 min at 37 °C. Single stranded cDNAs in 0.1 μl of the reaction mixture were amplified by PCR using AmpliTaq DNA polymerase (PE Applied Biosystems) under the following conditions; denaturation at 94 °C for 30 s, annealing at 60 °C for 30 s, and extension at 72 °C for 1 min, for 30 or 25 cycles when tissue RNAs or plasmids were used as template, respectively. The oligonucleotide primers used for RT-PCR were as follows (see Fig. 1); P1: 5′-GAGACTCATTCTACCATGAAG-3′ (nt 37–57), P2: 5′-CTCGGTGACACACTTCCTGG-3′ (nt 408–389).

Fluorescence in situ hybridization

Mouse and human UGRP2 probes of entire BAC clone genomic DNAs labeled with biotin or digoxigenin were used for fluorescence in situ hybridization (FISH) of chromosomes derived from methotrexate-synchronized normal peripheral lymphocytes and from mouse spleen cultures, respectively. The conditions of hybridization, detection of fluorescence signals, digital-image acquisition, processing and analysis, direct localization of signals on banded chromosomes were carried out as previously described (Popescu et al., 1994; Zimonjic et al., 1995). To confirm the identity of mouse chromosomes, preparations were rehybridized with mouse chromosome painting probes (Cambio) and previously observed labeled metaphases were rerecorded.

Interspecific mouse backcross mapping

Interspecific backcross progeny were generated by mating (C57BL/6J × *M. spretus*) F_1 females and C57BL/6J males as described (Copeland and Jenkins, 1991). A total of 205 N_2 mice were used to map the *Ugrp1* and *Ugrp2* loci (see text for details). DNA isolation, restriction enzyme digestion, agarose gel electrophoresis, Southern blot transfer and hybridization were performed essentially as described (Jenkins et al., 1982). All blots were prepared with Hybond N+ nylon membrane (Amersham Pharmacia Biotech, Piscataway, NJ). The *Ugrp1* probe, an ~500-bp fragment of mouse cDNA, was labeled with [α^{32}P]dCTP using a random primed labeling kit (Stratagene); washing was done at a final stringency of 1.0× SSCP, 0.1 % SDS, 65 °C. A fragment of 7.9 kb was detected in *Eco*RV digested C57BL/6J DNA and a fragment of 6.8 kb was detected in *Eco*RV digested *M. spretus* DNA. The *Ugrp2* probe, an ~500-bp fragment of mouse cDNA, detected an 11.0-kb *Eco*RI fragment in C57BL/6J DNA and a 14.0-kb *Eco*RI fragment in *M. spretus* DNA. The presence or absence of the *M. spretus*-specific fragments was followed in backcross mice.

A description of the probes and RFLPs for the loci linked to *Ugrp1* including *Nr3c1*, *Mcc,* and *Lmnb1* has been reported previously (Justice et al., 1992); those linked to *Ugrp2* include *Sox30* and *Il13* (McKenzie et al., 1993; Osaki et al., 1999). Recombination distances were calculated using Map Manager, version 2.6.5. Gene order was determined by minimizing the number of recombination events required to explain the allele distribution patterns.

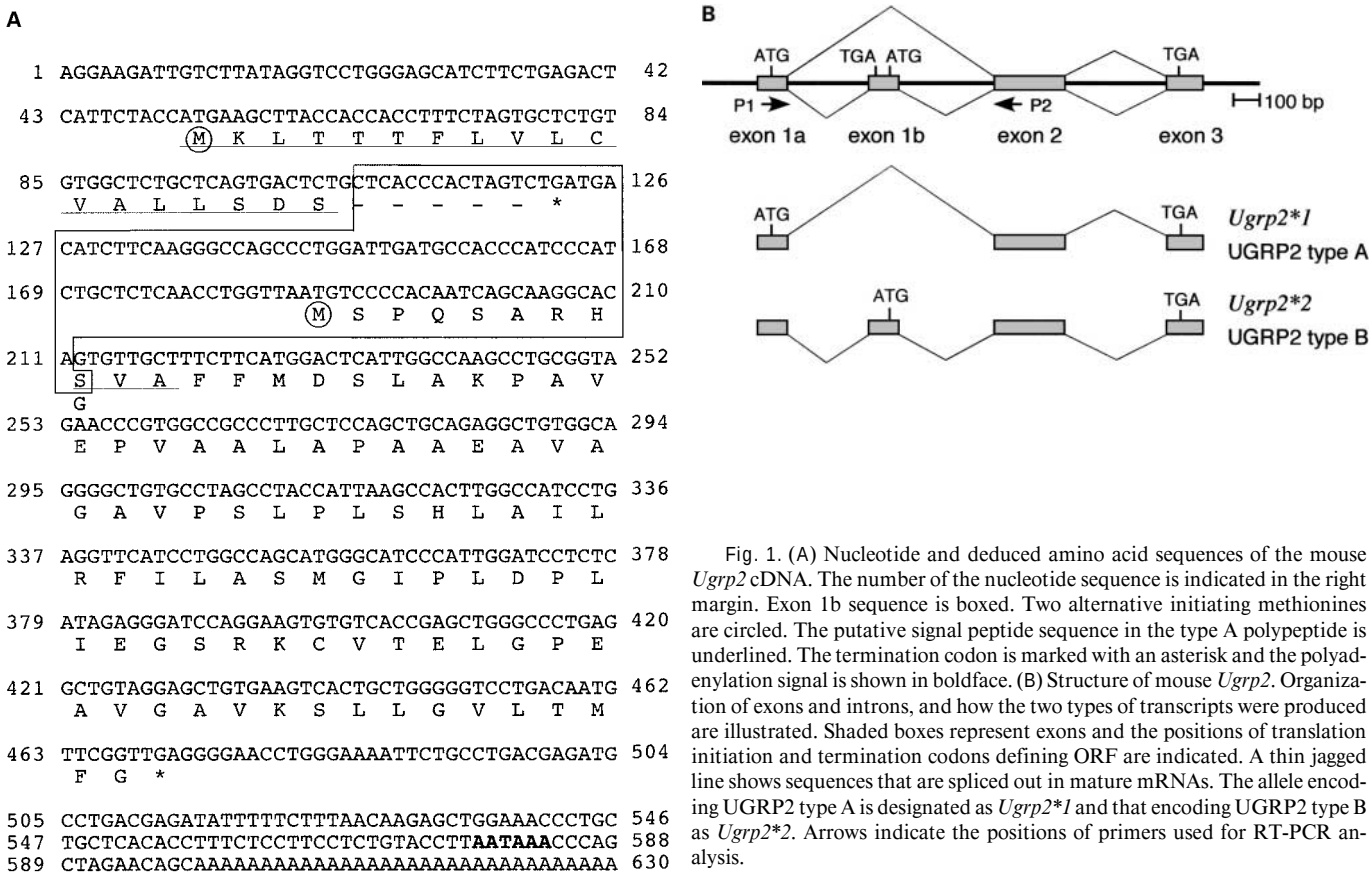

Fig. 1. (A) Nucleotide and deduced amino acid sequences of the mouse *Ugrp2* cDNA. The number of the nucleotide sequence is indicated in the right margin. Exon 1b sequence is boxed. Two alternative initiating methionines are circled. The putative signal peptide sequence in the type A polypeptide is underlined. The termination codon is marked with an asterisk and the polyadenylation signal is shown in boldface. (B) Structure of mouse *Ugrp2*. Organization of exons and introns, and how the two types of transcripts were produced are illustrated. Shaded boxes represent exons and the positions of translation initiation and termination codons defining ORF are indicated. A thin jagged line shows sequences that are spliced out in mature mRNAs. The allele encoding UGRP2 type A is designated as *Ugrp2*1* and that encoding UGRP2 type B as *Ugrp2*2*. Arrows indicate the positions of primers used for RT-PCR analysis.

Results and discussion

Cloning of a novel Ugrp1-related cDNA

The mouse Expressed Sequence Tags database (dbEST) was searched for sequences similar to the mouse *Ugrp1* cDNA using the BLAST algorithm software. An extensive set of sequences was identified that had an entire or a partial open reading frame (ORF) encoding a protein exhibiting significant similarity to UGRP1. A clone (GenBank Accession No. AI391046) that contained an entire ORF sequence was designated as mouse *Ugrp2*. RT-PCR was performed to obtain a *Ugrp2* partial cDNA clone using mouse adult lung RNAs as template and two regions of the EST sequence as primers, which was then used to isolate a full-length mouse *Ugrp2* cDNA by screening a mouse adult lung cDNA library. Eight clones with positive hybridization signal were identified in 1×10^6 recombinant phage. Among them, seven appeared to contain an entire ORF and the 5′- and 3′-noncoding region of *Ugrp2*. Interestingly, one clone contained an additional 105-bp sequence inserted within the coding region. A stop and a start codon were present at positions 15 and 81 of the inserted sequence, respectively. As a result, this cDNA encodes a UGRP2 variant with a shorter amino acid sequence. We refer to the two *Ugrp2* cDNAs as types A and B, encoding of 104 and 94 amino acids, respectively (Fig. 1A). Computer analysis revealed that the first 21 residues of the UGRP2 type A polypeptide may function as a signal sequence for targeting the protein to a secretory pathway. The UGRP2 type B polypeptide has nine unique amino acids at the N-terminus that replace the first 19 amino acids of the type A polypeptide, suggesting that the type B polypeptide is likely to be a cytoplasmic protein.

Characteristics of the Ugrp2 gene and the encoded polypeptide

In order to analyze the *Ugrp2* gene structure, a mouse BAC genomic library was screened using a full-length cDNA as a probe. The mouse *Ugrp2* gene encoding type A polypepetide is composed of three exons and two introns, and thus resembles the structure of the *Ugrp1* gene (Fig. 1B) (Niimi et al., 2001). The exon/intron splice sites are well conserved between *Ugrp1* and *Ugrp2*, suggesting that they constitute a gene family and share a common ancestral gene. The 105-bp insertion sequence (exon 1b) was found 300 bp downstream of the authentic exon 1, named exon 1a, indicating that UGRP2 type A and B polypeptides are encoded by splicing variation of a transcript from the same gene. Based on the gene nomenclature system, the type A is encoded by the *Ugrp2*1* allele, and the type B by the *Ugrp2*2* allele. The presence of splicing variants has also been reported for the *Ugrp1* gene (Niimi et al., 2001). Thus, the *Ugrp1* gene encodes three types of transcripts generated through alternative splicing, which in addition to normal splicing, either partially or entirely retains the second intron sequence.

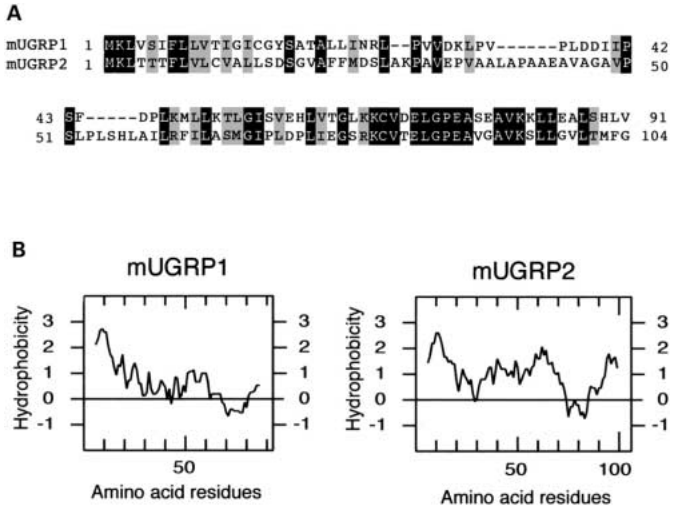

Fig. 2. Alignment of amino acid sequences and Kyte-Doolittle hydrophobicity plots of UGRP family proteins. (A) Mouse UGRP2 type A amino acid sequence is aligned with that of mouse UGRP1 type A. The identical and conserved residues are shown in black and shaded boxes, respectively. (B) Analysis of UGRP family protein structures by Kyte-Doolittle hydrophobicity plot.

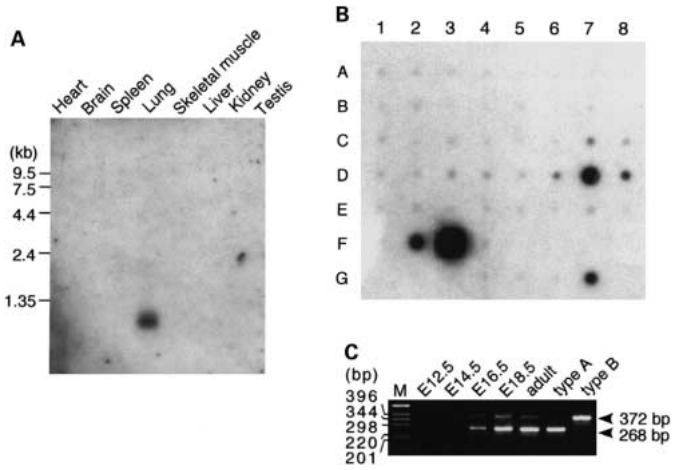

Fig. 3. Expression of *Ugrp2* transcripts. (A) Northern blot analysis of adult mouse tissues probed with ^{32}P-labeled mouse *Ugrp2* cDNA. (B) Dot blot anlysis of human tissues probed with ^{32}P-labeled human UGRP2 cDNA. RNA sources are as follows; A1: whole brain, A2: amygdala, A3: caudate nucleus, A4: cerebellum, A5: cerebral cortex, A6: frontal lobe, A7: hippocampus, A8: medulla oblongata, B1: occipital lobe, B2: putamen, B3: substantia nigra, B4: temporal lobe, B5: thalamus, B6: subthalamic nucleus, B7: spinal cord, C1: heart, C2: aorta, C3: skeletal muscle, C4: colon, C5: bladder, C6: uterus, C7: prostate, C8: stomach, D1: testis, D2: ovary, D3: pancreas, D4: pituitary gland, D5: adrenal gland, D6: thyroid gland, D7: salivary gland, D8: mammary gland, E1: kidney, E2: liver, E3: small intestine, E4: spleen, E5: thymus, E6: peripheral leukocyte, E7: lymph node, E8: bone marrow, F1: appendix, F2: lung, F3: trachea, F4: placenta, G1: fetal brain, G2: fetal heart, G3: fetal kidney, G4: fetal liver, G5: fetal spleen, G6: fetal thymus, and G7: fetal lung. (C) RT-PCR for the expression of two types of transcript. The positions of primers are indicated in Fig.1B. RT-PCR was carried out using RNAs prepared from lungs of E12.5 to E18.5 mouse embryos and adult. The product sizes are shown on the right. The results obtained with type A and B cDNA clones as a control template are indicated as type A and B, respectively. One-kb DNA ladder (Invitrogen Life Technology) was used as a size marker (M) and the size is indicated on the left margin.

The mouse UGRP1 type A, the most abundant form, and UGRP2 type A polypeptides exhibit 47% similarity in amino acid sequences (Fig. 2A). Similarity is especially high in the C-terminal one third of the polypeptide sequences. In the middle, many hydrophobic amino acid residues are present in the UGRP2 polypeptide, which do not align with any of the UGRP1 amino acid residues. This results in UGRP2 being more hydrophobic than UGRP1 (Fig. 2B).

Since a high level of *Ugrp2* expression was observed in mouse lung (see below) and the *Ugrp1* gene promoter is known to be transactivated by a homeodomain transcription factor T/EBP/NKX2.1 (Niimi et al., 2001), the mouse *Ugrp2* gene promoter was examined to determine whether it is regulated by the same transcription factor. Sequence analysis using TF (transcription factor) Search (Heinemeyer et al., 1998) revealed that several T/EBP/NKX2.1 binding consensus sites are present within 400 bp of the mouse *Ugrp2* gene promoter. However, no transactivation was observed when a co-transfection experiment was performed using *Ugrp2* gene promoter-luciferase reporter constructs and a T/EBP/NKX2.1 expression plasmid (data not shown). Expression of lung-specific genes such as surfactant proteins-A (Bruno et al., 1995; Bruno et al., 2000), -B (Bohinski et al., 1994; Yan et al., 1995), and -C (Kelly et al., 1996), and UG/CCSP (Sawaya et al., 1993; Ray et al., 1996; Braun and Suske, 1998) has been demonstrated to be controlled by a combination of transcription factors including T/EBP/NKX2.1, HNF-3α and β, and GATA-6 (also see Mendelson, 2000; Costa et al., 2001). Further analysis of the mouse *Ugrp2* gene promoter is required to understand the regulation of lung-specific expression of this gene.

Expression of Ugrp2 transcripts

Ugrp2 expression was examined by Northern blot analysis in adult mouse tissues. A single 0.6-kb transcript was detected only in the lung among all tissues examined (Fig. 3A). Human multiple tissue expression arrays were next examined for the expression of UGRP2 using RNA dot blot analysis (Fig. 3B). The strongest signal was obtained with trachea, followed by salivary gland and lung. Mammary gland and fetal lung also had a weak, but clear positive signal. Contrary to this, the expression of *Ugrp2* was not detected in mouse salivary gland and mammary gland as determined by RT-PCR (data not shown).

In order to confirm the presence of two types of transcripts, RT-PCR was performed using mouse embryonic and adult lung mRNAs as templates (Fig. 3C). Initially, PCR reactions were carried out using exon 1a (P1) and 2 (P2)-specific primer pair (see Fig. 1B) and an individual type of cDNA clone as template. These produced fragments of 268 and 372 bp that correspond to type A and B transcripts, respectively. Both RNAs

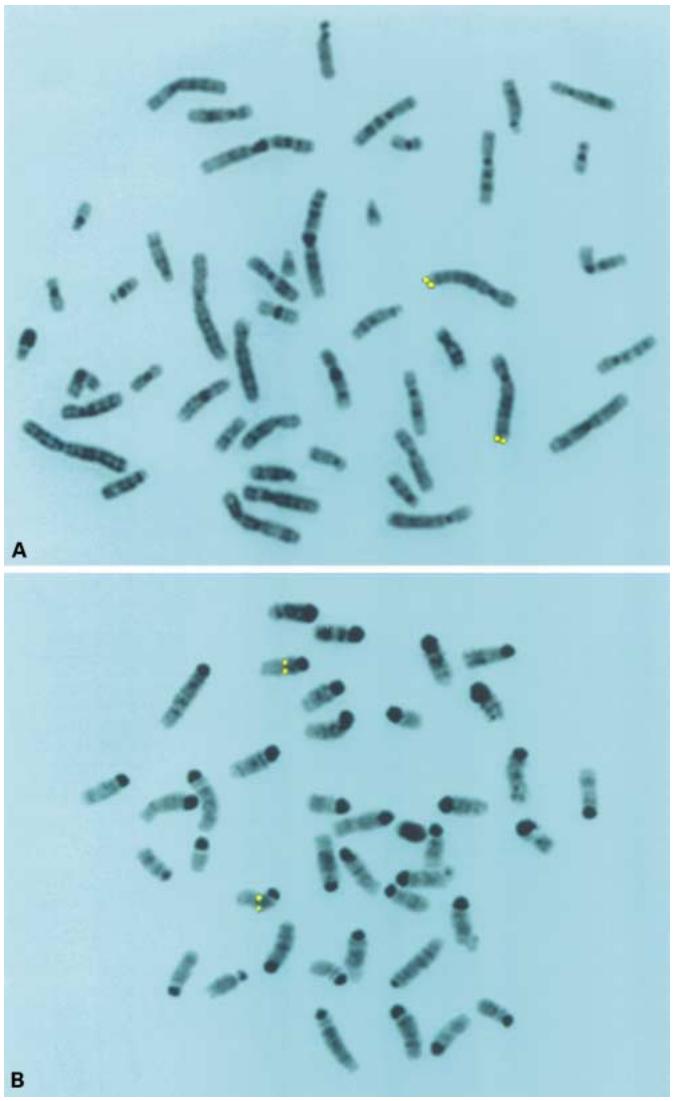

Fig. 4. FISH analyses for *Ugrp2* chromosomal localization. Human (A) and mouse (B) chromosome spreads after hybridization with species-specific biotin-labeled *Ugrp2* gene probes are shown. Both chromosome 5 in human and chromosome 11 in mouse exhibit symmetrical fluorescent signals at regions 5q35 and 11A5–B1, respectively, as depicted by enhanced DAPI-induced chromosome banding.

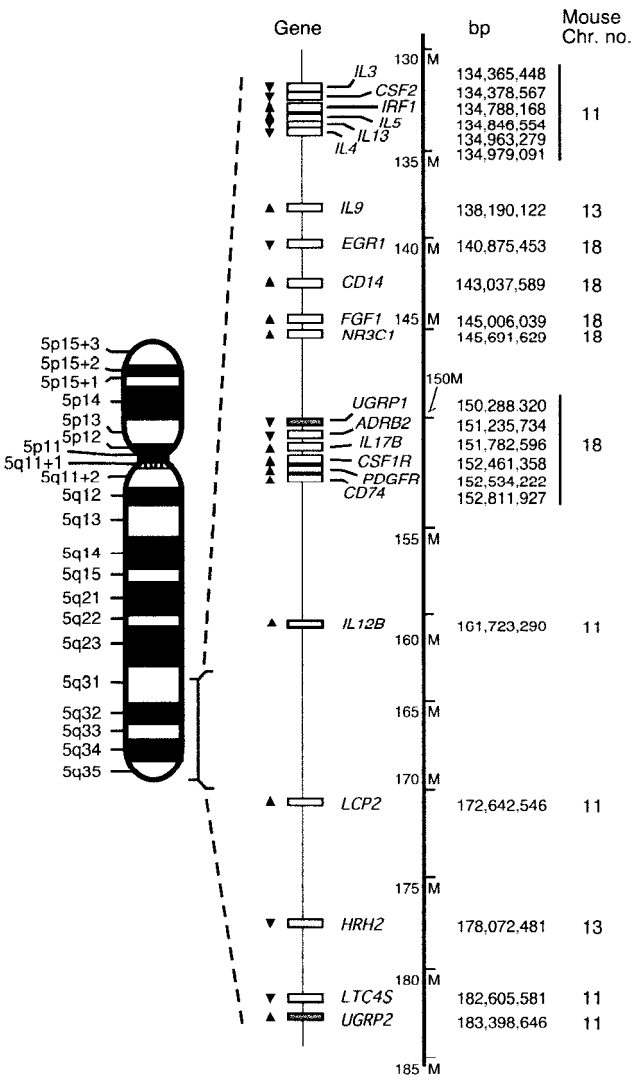

Fig. 5. Diagram of gene localization in human chromosome 5q31→q35. Genes listed include those that may be involved in inflammation. They are shown in the order of which they are localized from those close to the centromere (top) to those close to the telomere (bottom). Abbreviation for genes; IL3: interleukin 3, CSF2: colony stimulating factor 2, IRF1: interferon regulatory factor 1, IL5: interleukin 5, IL13: interleukin 13, IL4: interleukin 4, IL9: interleukin 9, EGR1: early growth response 1, CD14: CD14 antigen, FGF1: fibroblast growth factor 1, NR3C1: glucocorticoid receptor 1, ADRB2: adrenergic receptor β 2, IL17B: interleukin 17B, CSF1R: colony stimulating factor 1 receptor, PDGFR: platelet-derived growth factor receptor, CD74: CD74 antigen, IL12B: interleukin 12B, LCP2: lymphocyte cytosolic protein 2, HRH2: histamine receptor, LCT4: leukotrien C4 synthase. The number shown to the right of the gene is bp (M: mega) starting from the tip of the telomere of 5p used in the Celera Discovery System. Triangle indicates direction of gene. The mouse chromosome number for each gene is shown on the far right. This figure was generated through the use of the Celera Discovery System and Celera's associated databases.

prepared from mouse adult and embryo lungs demonstrated the expression of both types of transcripts although the band corresponding to the type B transcript was faint, suggesting that the type B transcript seems to be expressed at low levels. The expression of both type A and B transcripts was detected around E16.5, which increased toward the end of gestation and stayed at a similar level thereafter. Although this type of RT-PCR is not quantitative, the ratio in the expression levels between types A and B appears to be similar regardless of developmental stages. This is somewhat different from what was found for the three types of *Ugrp1* transcripts, in which type A is most abundant and two other forms were barely detected. These data suggest that the two forms of UGRP2 may possess distinct functions.

Previously, it was demonstrated that the expression of *Ugrp1* decreased in lungs of the TH2 cytokine-based antigen mouse model (Niimi et al., 2001). The expression of *Ugrp2* appeared also to be decreased in these lungs although the level of reduction was not as large as that for *Ugrp1* (Niimi et al.,

2001). These results suggested that UGRP1 and 2 may be involved in TH2 cytokine-based inflammation. In fact, our recent results demonstrated that human UGRP1 is involved in asthma (Niimi et al., 2002). It would be interesting to examine the role of the two forms of UGRP2 in lung inflammation. Further, the function of UGRP2 in mammary gland and salivary gland remains to be understood. In this regard it was recently reported that HIN-1, which is identical to UGRP2 is a candidate tumor suppressor gene in mammary gland (Krop et al., 2001).

Chromosomal localization of Ugrp2 gene by FISH

The human and mouse UGRP2/*Ugrp2* genes were mapped by FISH using chromosomes prepared from normal human peripheral leukocytes and mouse spleen cultures. In human cells, symmetrical fluorescent signals on sister chromatids were observed at the telomeric region of the long arm of chromosome 5 in 25 out of 30 metaphases. The probe had high specificity for this site as a symmetrical signal was not observed on other chromosomes. In ten metaphases with good resolution DAPI G-like banding the FISH signal was localized on chromosome 5 at region q35 where we assign the UGRP2 gene (Fig. 4A). Similarly, mouse *Ugrp2* probe showed label specificity for a single medium size chromosome identified by banding and chromosome painting as chromosome 11 and precisely localized at region 11A5–B1 (Fig. 4B). Region 5q35 in human is homologous with region 11A5–B1 in mouse (Searle et al., 1989; DeBry and Seldin, 1996).

Mouse *Ugrp1* gene was previously mapped to chromosome 18C–D (Niimi et al., 2001), a region that is homologous to human chromosome 5q31→q32 where the human UGRP1 gene was assigned (Niimi et al., 2002). Human chromosome 5q31→q35 has homologous regions in three mouse chromosomes; chromosomes 11, 13, and 18 (Searle et al., 1989; DeBry and Seldin, 1996). Thus, the UGRP1 and UGRP2 genes in human are located on chromosome 5q31→q35 whereas they are split into two mouse chromosomes 18 and 11, respectively (Fig. 5).

Human chromosome 5q31→q34 has been reported to contain one of the asthma susceptibility genes (Postma et al., 1995; Ruffilli and Bonini, 1997; Cookson and Moffatt, 2000; Ober and Moffatt, 2000). In addition to the UGRP1 gene, this region contains numerous gene candidates that may potentially be involved in the airway inflammation associated with atopic asthma, including a number of proinflammatory cytokines such as interleukin (IL)-3, 4, 5, 9, and 13, granulocyte macrophage colony-stimulating factor (CSF), and the β2-adrenergic receptor (ADRB2). Relative locations of these genes to the human UGRP1 and UGRP2 genes were determined using human genomic sequence databases (Fig. 5). The distance between UGRP1 and UGRP2 was calculated to be approximately 30 Mbp. Thus, it may be possible that the expression of UGRP1 and/or UGRP2 is directly or indirectly regulated by one of these cytokines, which affects the inflammatory status of the lung involved in asthma. Further experiments are required to determine whether this is the case.

Chromosomal localization of Ugrp1 and Ugrp2 genes by interspecific mouse backcross mapping

The mouse chromosomal locations of *Ugrp1* and *Ugrp2* genes were further determined by interspecific backcross analysis using progeny derived from matings of (C57BL/6J × *Mus spretus*)F_1 × C57BL/6J mice. This interspecific backcross-mapping panel has been typed for over 3100 loci that are well distributed among all autosomes as well as the X chromosome (Copeland and Jenkins, 1991). C57BL/6J and *M. spretus* DNAs were digested with several enzymes and analyzed by Southern blot hybridization for informative restriction fragment length polymorphisms (RFLPs) using mouse cDNA probes specific for each gene. The 6.8-kb *Eco*RV *M. spretus* RFLP (see Materials and methods) was used to follow the segregation of the *Ugrp1* locus in backcross mice. The mapping results indicated that *Ugrp1* is located in the central region of mouse chromosome 18 linked to *Nr3c1*, *Mcc*, and *Lmnb1*. Although 144 mice were analyzed for every marker and are shown in the segregation analysis (Fig. 6A), up to 179 mice were typed for some pairs of markers. Each locus was analyzed in pairwise combinations for recombination frequencies using the additional data. The ratios of the total number of mice exhibiting recombinant chromosomes to the total number of mice analyzed for each pair of loci and the most likely gene order are: centromere-*Nr3c1* – 4/179 – *Mcc* – 0/154 – *Ugrp1* – 14/152 – *Lmnb1*. The recombination frequencies (expressed as genetic distances in centiMorgans [cM] ± the standard error) are: centromere – *Nr3c1* – 2.2 ± 1.1 – (*Mcc*, *Ugrp1*) – 9.2 ± 2.4 – *Lmnb1*. No recombinants were detected between *Mcc* and *Ugrp1* in 154 animals typed in common, suggesting that the two loci are within 1.9 cM of each other (upper 95% confidence limit).

The 14.0-kb *Eco*RI *M. spretus* RFLP (see Materials and methods) was used to follow the segregation of the *Ugrp2* locus in backcross mice. *Ugrp2* mapped to the proximal region of mouse chromosome 11 linked to *Sox30* and *Il13*. In this case, 112 mice were analyzed for every marker and are shown in the segregation analysis (Fig. 6B) and up to 141 mice were typed for some pairs of markers. Again, each locus was analyzed in pairwise combinations for recombination frequencies using the additional data. The ratios of the total number of mice exhibiting recombinant chromosomes to the total number of mice analyzed for each pair of loci and the most likely gene order are: centromere – *Sox30* – 6/141 – *Ugrp2* – 3/115 – *Il13*. The recombination frequencies (expressed as genetic distances in centiMorgans [cM] ± the standard error) are: centromere – *Sox30* – 4.3 ± 1.7 – *Ugrp2* – 2.6 ± 1.5 – *Il13*.

We have compared our interspecific map of chromosome 11 and 18 with a composite mouse linkage map that reports the map location of many uncloned mouse mutations (provided from the Mouse Genome Database, http://www.informatics.jax.org/). *Ugrp1* and *Ugrp2* mapped in regions of the composite map that lack mouse mutations with a phenotype that might be expected for an alteration in these loci (data not shown).

By both FISH and interspecific backcross mapping, mouse *Ugrp1* and *Ugrp2* genes were mapped to chromosome 18 and 11, respectively, both of which have homology with human chromosome 5q31→q35 (Searle et al., 1989; DeBry and Sel-

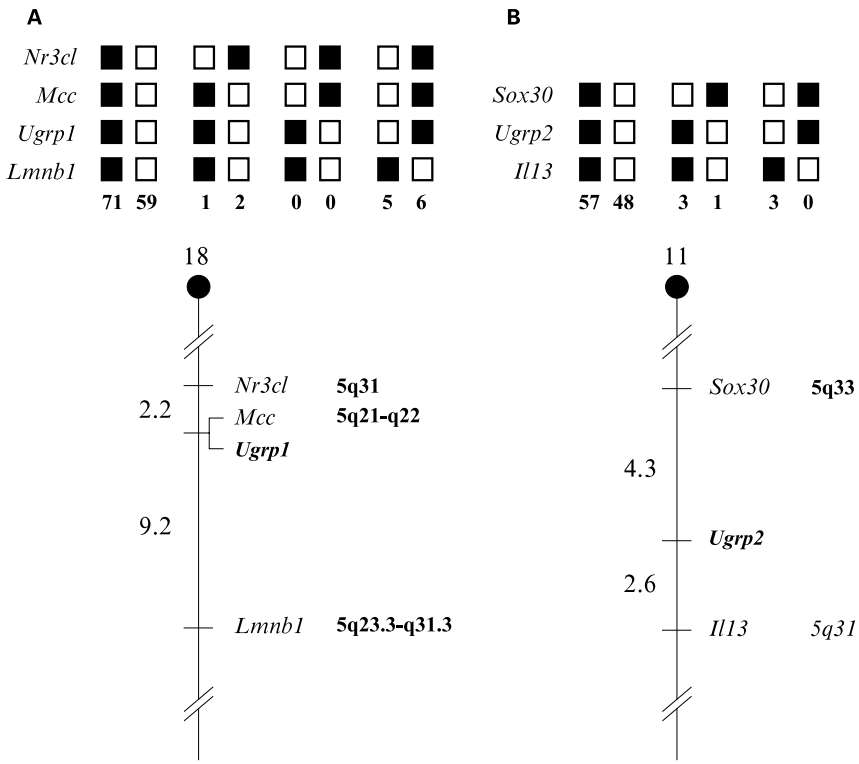

Fig. 6. Murine chromosomal location of *Ugrp1* and *Ugrp2*. The segregation patterns of *Ugrp1* (A) and *Ugrp2* (B) and their flanking genes in 144 and 112 backcross animals, respectively, that were typed for all loci are shown at the top of the figure. For individual pairs of loci, more animals were typed (see text). Each column represents the chromosome identified in the backcross progeny that was inherited from the (C57BL/6J × *M. spretus*) F_1 parent. The shaded boxes represent the presence of a C57BL/6J allele and white boxes represent the presence of an *M. spretus* allele. The number of offspring inheriting each type of chromosome is listed at the bottom of each column. Partial chromosome 18 and 11 linkage maps showing the location of *Ugrp1* (A) and *Ugrp2* (B) in relation to linked genes are shown at the bottom of the figure. Recombination distances between loci in centimorgans are shown to the left of the chromosome and the positions of loci in human chromosomes, where known, are shown to the right. References for the human map positions of loci cited in this study can be obtained from the GDB (http://www.gdb.org/).

din, 1996) where human UGRP1 and UGRP2 genes were localized. In mouse and human genomic sequence databases, no other gene(s) was found that belongs to the same gene family by exhibiting significant sequence similarity to these two genes.

Acknowledgements

We thank Dr Frank J. Gonzalez (NCI, Bethesda, MD) and Deborah B. Householder (NCI-Frederick, Frederick, MD) for helpful discussions and critical review of the manuscript, and for excellent technical assistance.

References

Bohinski RJ, Di Lauro R, Whitsett JA: The lung-specific surfactant protein B gene promoter is a target for thyroid transcription factor 1 and hepatocyte nuclear factor 3, indicating common factors for organ-specific gene expression along the foregut axis. Mol Cell Biol 14:5671–5681 (1994).

Braun H, Suske G: Combinatorial action of HNF3 and Sp family transcription factors in the activation of the rabbit uteroglobin/CC10 promoter. J biol Chem 273:9821–9828 (1998).

Broeckaert F, Bernard A: Clara cell secretory protein (CC16): characteristics and perspectives as lung peripheral biomarker. Clin Exp Allergy 30:469–475 (2000).

Bruno MD, Bohinski RJ, Huelsman KM, Whitsett JA, Korfhagen TR: Lung cell-specific expression of the murine surfactant protein A (SP-A) gene is mediated by interactions between the SP-A promoter and thyroid transcription factor-1. J biol Chem 270:6531–6536 (1995).

Bruno MD, Korfhagen TR, Liu C, Morrisey EE, Whitsett JA: GATA-6 activates transcription of surfactant protein A. J biol Chem 275:1043–1049 (2000).

Cookson WO, Moffatt MF: Genetics of asthma and allergic disease. Hum molec Genet 9:2359–2364 (2000).

Copeland NG, Jenkins NA: Development and applications of a molecular genetic linkage map of the mouse genome. Trends Genet 7:113–118 (1991).

Costa RH, Kalinichenko VV, Lim L: Transcription factors in mouse lung development and function. Am J Physiol Lung Cell Mol Physiol 280:L823–838 (2001).

DeBry RW, Seldin MF: Human/mouse homology relationships. Genomics 33:337–351 (1996).

Heinemeyer T, Wingender E, Reuter I, Hermjakob H, Kel AE, Kel OV, Ignatieva EV, Ananko EA, Podkolodnaya OA, Kolpakov FA, Podkolodny NL, Kolchanov NA: Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. Nucl Acids Res 26:362–367 (1998).

Jenkins NA, Copeland NG, Taylor BA, Lee BK: Organization, distribution, and stability of endogenous ecotropic murine leukemia virus DNA sequences in chromosomes of *Mus musculus*. J Virol 43:26–36 (1982).

Justice MJ, Gilbert DJ, Kinzler KW, Vogelstein B, Buchberg AM, Ceci JD, Matsuda Y, Chapman VM, Patriotis C, Makris A, et al: A molecular genetic linkage map of mouse chromosome 18 reveals extensive linkage conservation with human chromosomes 5 and 18. Genomics 13:1281–1288 (1992).

Kelly SE, Bachurski CJ, Burhans MS, Glasser SW: Transcription of the lung-specific surfactant protein C gene is mediated by thyroid transcription factor 1. J biol Chem 271:6881–6888 (1996).

Krop IE, Sgroi D, Porter DA, Lunetta KL, LeVangie R, Seth P, Kaelin CM, Rhei E, Bosenberg M, Schnitt S, Marks JR, Pagon Z, Belina D, Razumovic J, Polyak K: HIN-1, a putative cytokine highly expressed in normal but not cancerous mammary epithelial cells. Proc natl Acad Sci USA 98:9796–9801 (2001).

McKenzie AN, Li X, Largaespada DA, Sato A, Kaneda A, Zurawski SM, Doyle EL, Milatovich A, Francke U, Copeland NG: Structural comparison and chromosomal localization of the human and mouse IL-13 genes. J Immun 150:5436–5444 (1993).

Mendelson CR: Role of transcription factors in fetal lung development and surfactant protein gene expression. A Rev Physiol 62:875–915 (2000).

Mukherjee AB, Chilton BS: The uteroglobin/clara cell protein family. Ann NY Acad Sci 923:1–356 (2000).

Niimi T, Keck-Waggoner CL, Popescu NC, Zhou Y, Levitt RC, Kimura S: UGRP1, a uteroglobin/clara cell secretory protein-related protein, is a novel lung-enriched downstream target gene for the T/EBP/NKX2.1 homeodomain transcription factor. Mol Endocrinol 15:2021–2036 (2001).

Niimi T, Munakata M, Keck-Waggoner CL, Popescu NC, Levitt RC, Hisada M, Kimura S: A polymorphism in the human *UGRP1* gene promoter that regulates transcription is associated with an increased risk of asthma. Am J hum Genet 70:718–725 (2002).

Ober C, Moffatt MF: Contributing factors to the pathobiology. The genetics of asthma. Clin Chest Med 21:245–261 (2000).

Osaki E, Nishina Y, Inazawa J, Copeland NG, Gilbert DJ, Jenkins NA, Ohsugi M, Tezuka T, Yoshida M, Semba K: Identification of a novel Sry-related gene and its germ cell-specific expression. Nucl Acids Res 27:2503–2510 (1999).

Popescu N, Zimonjic D, Hatch C, Bonner W: Chromosomal mapping of the human histone gene H2AZ to 4q24 by fluorescence in situ hybridization. Genomics 20:333–335. (1994).

Postma DS, Bleecker ER, Amelung PJ, Holroyd KJ, Xu J, Panhuysen CI, Meyers DA, Levitt RC: Genetic susceptibility to asthma-bronchial hyperresponsiveness coinherited with a major gene for atopy. New Engl J Med 333:894–900 (1995).

Ray MK, Chen CY, Schwartz RJ, DeMayo FJ: Transcriptional regulation of a mouse Clara cell-specific protein (mCC10) gene by the NKx transcription factor family members thyroid transciption factor 1 and cardiac muscle-specific homeobox protein (CSX). Mol cell Biol 16:2056–2064 (1996).

Ruffilli A, Bonini S: Susceptibility genes for allergy and asthma. Allergy 52:256–273 (1997).

Sawaya PL, Stripp BR, Whitsett JA, Luse DS: The lung-specific CC10 gene is regulated by transcription factors from the AP-1, octamer, and hepatocyte nuclear factor 3 families. Mol cell Biol 13:3860–3871 (1993).

Searle AG, Peters J, Lyon MF, Hall JG, Evans EP, Edwards JH, Buckle VJ: Chromosome maps of man and mouse. IV. Ann hum Genet 53:89–140 (1989).

Yan C, Sever Z, Whitsett JA: Upstream enhancer activity in the human surfactant protein B gene is mediated by thyroid transcription factor 1. J biol Chem 270:24852–24857 (1995).

Zimonjic DB, Rezanka L, DiPaolo JA, Popescu NC: Refined localization of the erbB-3 proto-oncogene by direct visualization of FISH signals on LUT-inverted and contrast-enhanced digital images of DAPI-banded chromosomes. Cancer Genet Cytogenet 80:100–102 (1995).