FEBS Letters 419 (1997) 186–190 FEBS 19602

The mouse relaxin-like factor gene and its promoter are located within the 3′ region of the JAK3 genomic sequence

Pasi Koskimies[a,*], Andrej-Nikolai Spiess[b], Petra Lahti[a], Ilpo Huhtaniemi[a], Richard Ivell[b]

[a]*Department of Physiology, University of Turku, 20520 Turku, Finland*
[b]*Institute for Hormone and Fertility Research, University of Hamburg, 22529 Hamburg, Germany*

Received 17 October 1997

Abstract Isolation and sequencing of a genomic clone encoding the mouse gene for the relaxin-like factor (RLF), which is endogenously expressed to a high level exclusively in Leydig cells, indicated that similar sequences were also present at the 3′ end of the mouse JAK3 gene, a gene expressed predominantly in lymphoid tissues. More extensive Southern blot, polymerase chain reaction and sequencing analyses showed that the published mouse sequence for exon 23 of the JAK3 gene in fact comprises two exons, 23A and 23B, separated by an additional novel intron of 2.2 kb, and that within this intron the promoter and exon 1 of the mouse RLF gene are encoded. The two overlapping transcripts appear to use different polyadenylation signals in the common 3′ untranslated region of exon 23B. Transient transfection of different RLF promoter reporter constructs into Leydig, Sertoli, granulosa and kidney cell lines indicate that as little as 0.7 kb of the region upstream of exon 1 of the RLF gene, and within the novel intron 22 of the JAK3 gene, is sufficient to account for cell-specific expression of the RLF gene. This promoter region is specifically hypomethylated in Leydig cells compared to non-expressing tissues.

Key words: JAK3; RLF; Leydig cell; Relaxin-like factor

1. Introduction

The relaxin-like factor (RLF), also referred to as the Leydig insulin-like (Ley-I-L) peptide, is a novel member of the insulin/IGF/relaxin family of hormones and growth factors [1]. The mouse RLF cDNA encodes a 122 amino acid precursor polypeptide, comprising a signal peptide followed by B-, C- and A-peptide domains [2], a structure similar to that of insulin and relaxin. Although the function of RLF has not yet been elucidated, in all species where the cDNA sequence has been determined, there is a conserved motif within the B-domain, -R-X-V-L-R-, which is similar to the defined receptor binding motif for relaxin [3]. Indeed, a chemically synthesized human RLF peptide has been shown to interact specifically with relaxin receptors [4]. In all species studied so far, RLF appears to be expressed to a high level almost exclusively within the Leydig cells of the testis [1,2,5–7]. Recent studies also indicate varying levels of RLF gene expression in the ovary [5,6,8]. In the mouse testis, RLF expression is increased during puberty [2,8,9], though RLF mRNA can be detected by reverse transcription polymerase chain reaction (RT-PCR) as early as fetal day 13.5 p.c. [9].

All members of this family share a similar genomic structure with two exons interrupted by a single intron, with the splice junction occurring within the open reading frame encoding the first part of the C- (connecting) peptide domain. In order to delimit and characterize the promoter region of the mouse RLF gene, and to understand the mechanisms ensuring the Leydig cell-specific pattern of expression, we have isolated and sequenced a genomic fragment cloned from a strain 129 mouse genomic library. During the course of this study, a database search indicated homology between exon 2 of the mouse RLF gene and a part of exon 23 of the mouse JAK3 genomic sequence (accession number U71201 [10]). Preliminary comparisons of the structural organization of the two genes were unable to provide a satisfactory explanation for this sequence identity. We therefore used a combination of Southern hybridization, PCR and sequencing strategies and showed that there is an additional intron within the published mouse JAK3 genomic sequence, which interrupts the published exon 23 sequence, and that the promoter region, exon 1 and intron of the mouse RLF gene are located within this novel intron. Thus the mouse RLF gene is located within the JAK3 cistronic domain and shares part of a downstream, protein-coding exon and 3′ untranslated region.

2. Materials and methods

2.1. Isolation and sequencing of genomic clones

Two RLF-specific primers, R1 (5′-CGCGCCGCTGCTACTGATGC) and R2 (5′-AGGGCCTGTGGTCTTGCTTACTGC) were designed based on the mouse RLF cDNA sequence [2], and corresponding to exons 1 and 2 of the human and porcine genes. These were used in a PCR reaction employing standard conditions, to amplify a mouse genomic sequence. The resulting 1.2 kb PCR amplicon was sequenced, labelled with [α-32P]dCTP (Amersham-Büchler, Braunschweig, Germany) using the random primer method (Prime-A-Gene Labelling Kit, Promega, Madison, WI), and used to screen approximately 5×10^5 clones of a mouse 129/SVJ genomic λ library (Lambda-FIX II, Stratagene, Palo Alto, CA). A single positive clone was isolated and purified. Digestion of the bacteriophage DNA with *Sal*I released an insert of approximately 5.5 kb, which was subcloned into the pBluescript KS II plasmid (Stratagene) for restriction mapping and sequence analysis. M13 and RLF-specific primers were used for cycle sequencing and the sequence data collated on an ABI 373 A automated sequencer (Perkin Elmer - Applied Biosystems, Foster City, CA). Sequence data were analysed using DNASTAR software, Madison, WI).

In order to check the structural relationship of the RLF gene with the JAK3 gene, further PCR analysis was undertaken using mouse genomic DNA as template. A forward primer (J5, see Fig. 1; 5′-CAGCCTTCGGCACCCTG) was designed based on a 5′ sequence of the published exon 23 of the JAK3 gene [10]. As reverse primers, sequences were used derived from exons 1 and 2 of the RLF gene (see Fig. 1) (R2, as above; R3, 5′-GTCTCCACAGGCTGCGT; R4, GTTTGGGACACAGGGAGGAG; R5, ACCAGCGCGTGCAGGAGATGT). PCR was carried out for 30 cycles with the following conditions: 95°C, 2 min denaturing, then 95°C for 20 s denaturing, 55°C for 40 s annealing, 72°C for 2 min 30 s extension, increasing by 10 s/cycle from cycles 11 to 30.

*Corresponding author. Fax: +358-2-2502610.
E-mail: pkoski@utu.fi

PII S0014-5793(97)01454-3

2.2. Southern analysis of genomic DNA

Twenty μg of genomic DNA derived from the kidney of a male mouse were exhaustively digested overnight with 100 units of *Bam*HI or *Sal*I. Samples of 10 μg were loaded in duplicate on to a 1.3% agarose gel, electrophoresed and transferred following standard procedures [11]. One of the duplicate membranes was probed with a 919 bp *Pst*I-specific fragment (Fig. 1) specific for exon 23 of the mouse JAK3 gene (nucleotides 8630–9549 [10]). The second duplicate membrane was probed with a 474 bp RLF-specific probe (Fig. 1), which contains ca. 300 bp of the promoter region and exon 1 of the RLF gene, and was generated by PCR using primers R301 (5′-CCAAGCTTACAACCTAAACTGCCTCCAAC, including an artificial *Hin*dIII site at the 5′ end) and R3 (see above), with the 5.5 kb RLF genomic fragment (see above) as template. Probes were labelled as above, and hybridization performed at 65°C overnight [12]. Membranes were washed under increasingly stringent conditions (final: 65°C in 0.1×SSC/0.1% SDS) and exposed to Biomax film (Eastman Kodak, Rochester, NY) for 48 h.

In order to detect whether the promoter region of the RLF gene might be differentially methylated in expressing compared to non-expressing tissues, 10 μg genomic DNA from either Percoll-purified mouse Leydig cells, spleen or liver, were digested overnight with 100 units each of *Bam*HI and *Hpa*II, or *Bam*HI and *Msp*I. Electrophoresis and blotting were performed as above. To compare the relative methylation patterns, the resulting blots were hybridized with the same 474 bp RLF-specific probe used above. Equal loading was verified by ethidium bromide staining of the completed gel (not shown).

2.3. Transfection of cell lines with RLF promoter constructs

Mouse Leydig (BLT-1 [13]), Sertoli (MSC-1 [14]), granulosa (KK-1 [15]) and monkey kidney (COS-7) cells were cultured in Dulbecco's modified Eagle's medium/F12 (DMEM/F12; Gibco, Paisley, Scotland), supplemented with 10% heat-inactivated fetal calf serum (Bioclear, Berks, UK) and 0.1 g/l gentamycin (Biological Industries, Bet-Haemek, Israel). A second mouse Leydig cell line (mLTC-1 [16]) was cultured as described previously [17]. Firefly luciferase reporter gene constructs were ligated to 4.3 kb and 0.7 kb fragments of the putative mRLF promoter in the promoterless luciferase vector. Cell lines were transiently transfected in triplicate by electroporation using equal molarity (10 pmol) of promoterless, 0.7 kb RLF and 4.3 kb RLF luciferase expression constructs. pBluescript plasmid DNA was used as carrier to equalize the amount of transfected DNA [18]. A total of approximately 5×10^6 cells in serum- and antibiotic-free medium were electroporated in a Genepulser cuvette (0.4 cm gap, BioRad, Richmond, CA) by using 350 V/960 μF for BLT-1 cells; 320 V/960 μF for LTC, KK-1 and MSC-1 cells; and 300 V/960 μF for COS-7 cells. The electroporated cells were incubated for 24 h on 10 cm plates, before harvesting and analysis of luciferase activity [21]. Luciferase activity was measured using the 125I luminometer (BioOrbit, Turku, Finland), and results corrected for transfection efficiency by relating to values for a cotransfected CMV-β-galactosidase construct.

3. Results

3.1. Structural organization of the RLF/JAK3 gene locus

By comparison with the cDNA sequence [2], and as indicated in Fig. 1, the 5.5 kb *Sal*I restriction fragment isolated from the RLF-positive bacteriophage clone of mouse genomic DNA included approximately 4.3 kb of the region upstream of exon 1 of the RLF gene. This is followed by an intron of 740 bp and most of exon 2 up to, but not including the polyadenylation signal within the 3′ untranslated region (UTR) of the RLF mRNA. Within exonic regions, there is full agreement with the cDNA sequence, with the 175 bp exon 1 encoding the signal and B-peptide domains, as well as the first 7 amino acids of the C-domain. Exon 2 encodes the remainder of the C-domain, together with the A-peptide domain and part of the 3′ UTR.

Approximately 1500 bp of the 5′ UTR and putative promoter region upstream of the first ATG translation initiator codon was sequenced. There is a TATA-box 36 bp upstream of the ATG codon, and an Sp1 site about 15 nucleotides further upstream. Three putative binding sites for the transcription factor steroidogenic factor −1 (SF1/Ad4BP [19,20]) are located at positions −64 to −57, −144 to −137 and −479 to −472, relative to the ATG codon (Fig. 1). Additionally, there is a recognition motif for the sex determining region Y transcription factor (SRY [21,22]) at position −211 to −217.

In addition to the immediate region of the RLF gene, also ca. 600 bp of the extreme 5′ end of the *Sal*I restriction fragment were sequenced. A homology search showed this region to be 96% identical to nucleotides 6627–7219 of the published mouse JAK3 genomic sequence [10], corresponding to exon 17, intron 17, and part of exon 18 of this gene. Further homology mapping showed that the sequence −1515 to −1095 relative to the initiator codon of the RLF gene corresponded with 94% homology to nucleotides 9409–9829 of the JAK3 gene, equivalent to part of exon 23. Further, exon 2 of the RLF gene shows 96.6% identity to the immediately following nucleotides 9830–10230 of the JAK3 gene, corresponding to the subsequent part of exon 23. There were no corresponding homologies to other parts of the mouse RLF gene sequence, such as exon 1 or the immediate upstream promoter. Assuming the sequences of the mouse RLF and mouse JAK3 genes are indeed syngeneic, then the only plausible explanation is that the published JAK3 genomic sequence is incorrect, and

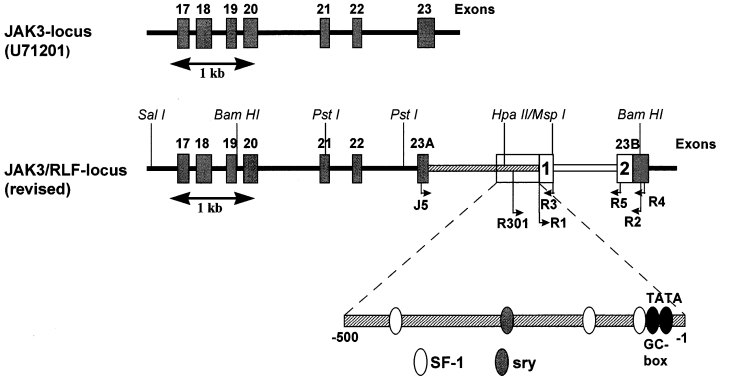

Fig. 1. Schematic organization of the mouse JAK3/RLF gene locus, indicating relative positions of the JAK3 (hatched boxes) and RLF (open boxes) exons, positions of PCR primers used, and locations of potential transcription factor binding sites. Also shown above is the original JAK3 genomic structure (accession number U71201). For details, see text.

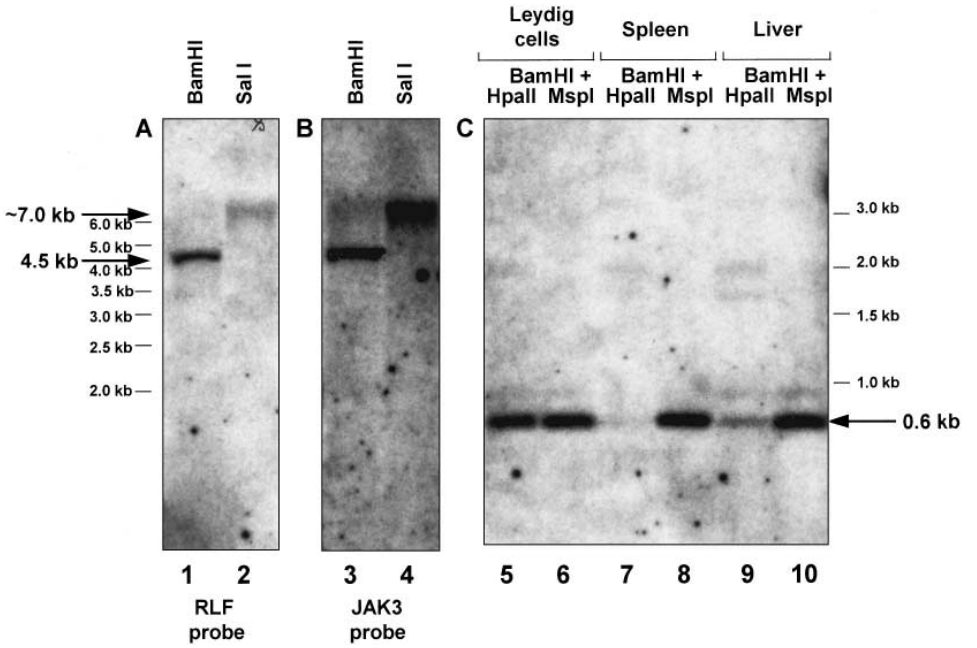

Fig. 2. Southern hybridization of mouse genomic DNA. A and B represent duplicate blots hybridized against RLF and JAK3 specific probes respectively. C: Genomic DNA from Leydig cells, spleen or liver, exhaustively digested with *Bam*HI and either *Hpa*II or *Msp*I, as indicated. The undigested 5.5 kb *Bam*HI precursor fragment did not resolve from the slots in this short run gel (not included). The positions of DNA size markers are shown beside the panels.

that the proposed exon 23 is in fact a composite of two exons separated by a substantial novel intron, which would also include exon 1 of the RLF gene, as reflected in the structure of our RLF genomic clone (Fig. 1).

Two analyses were carried out to check this hypothesis. Firstly, PCR reactions were performed, using mouse genomic DNA as template and specific primers from the 5′ portion of exon 23 of the JAK3 gene (J5; forward primer) and exons 1 and 2 of the RLF gene (R2, R3, R4 and R5; reverse primers), R2 and R4 being equivalent to the 3′ portion of exon 23 of the JAK3 gene. The resulting products (not shown) of 2.4, 1.3, 2.45 and 2.13 kb, respectively, imply the presence of ca. 2 kb of additional intronic sequence as predicted from the organization of the bacteriophage clone (Fig. 1). The validity of the PCR products was checked by DNA sequencing. Secondly, Southern hybridizations of genomic blots were performed using RLF- and JAK3-specific probes (Fig. 2A and B). For the restriction enzymes employed (*Bam*HI and *Sal*I), an identical pattern was found for the two probes. This shows firstly that the restriction map predicted from the organization of the bacteriophage clone reflects the true genomic structure, the *Bam*HI fragment being approximately 2 kb larger than is predicted from the published JAK3 genomic sequence, and that both genes are syngeneic as anticipated.

3.2. *RLF gene promoter specificity and activity*

Transient transfection of the RLF promoter luciferase constructs showed that both the 4.3 kb and the 0.7 kb promoter constructs, compared to the promoterless construct, gave high (60-fold) specific luciferase activity in the two mouse Leydig cell lines, but none in the Sertoli cell line (Fig. 3). A similar high level of expression to that in the Leydig cell lines was also seen in the KK-1 murine granulosa cell line. Only a low basal activity (5–10-fold) was evident in the transfected COS-7 cells. The 4.3 kb promoter construct indicated a marginally lower luciferase activity than the 0.7 kb fragment. These results indicate that apparent full activity and specificity of the mouse RLF promoter probably reside within less than 0.7 kb of the upstream promoter region, that is, within the region defined above as intronic between exons 23A and 23B of the JAK3 gene (Fig. 1).

3.3. *Methylation analysis of genomic DNA in the RLF promoter region*

Genomic regions requiring access by transcription factors are characterized by a low level of methylation in tissues where a gene is being expressed by comparison with tissues where the gene is not expressed. Thus, in order to confirm that

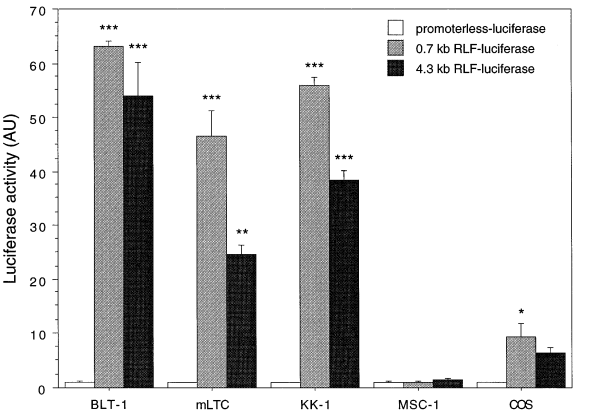

Fig. 3. Effect of two different RLF promoter constructs on the activity of the luciferase reporter gene in different cell lines. The data shown represent the mean of triplicates from two independent experiments (±S.E.M.). Post hoc statistical analysis used Fisher's PLSD, with the significance levels of differences from the expression of the promoterless luciferase construct set at $*P<0.05$, $**P<0.=1$, $***P<0.001$.

the immediate upstream region from exon 1 of the RLF gene represents a probable gene promoter, tissue-specific differential methylation was assessed by comparing the restriction pattern obtained with the methylation-sensitive endonuclease *Hpa*II, with that from its methylation-insensitive isoschizomer *Msp*I (Fig. 2C). The sequence data indicate that there are two *Hpa*II/*Msp*I sites within the region covered by the specific probe used, yielding upon complete digestion a major fragment of 578 bp, following preliminary cleavage with *Bam*HI. In Leydig cells, *Hpa*II digestion was as efficient as *Msp*I digestion, showing that this region is relatively unmethylated compared to tissues which do not express the RLF gene, such as the spleen or liver where *Hpa*II digestion was inefficient.

4. Discussion

The mouse JAK3 gene encompasses ca. 12 kb of genomic sequence, reportedly comprising 23 exons and 22 introns. In the mouse, there would appear to be several splice variants, resulting from alternative splicing in the 3′ region [23]. To date, JAK3 expression has been described as predominantly in lymphoid tissue and is related to signal transduction in NK and T cells [24]. The RLF gene in contrast is small, comprising two exons, with a short intervening sequence, and has a length, excluding the upstream promoter of less than 2 kb. The RLF gene is characterized by a very specific tissue distribution, in the mouse being almost exclusively restricted to the testicular Leydig cells [2]. Other tissues in the mouse only indicate levels of expression detectable by RT-PCR techniques. In the course of characterizing the mouse RLF gene, we identified regions of homology to segments of the 3′ region of the mouse JAK3 gene. However, the experimentally determined organization of this genomic region did not tally with the published genomic sequence for mouse JAK3. Instead, we have been able to show that the published exon 23 of the JAK3 gene in fact includes a further intron of 2 kb, resulting in exons 23A and 23B. Exon 23B is syngeneic with exon 2 of the mouse RLF gene, using the same acceptor splice junction. This exon encompasses therefore both parts of the protein-coding region of the RLF transcript and downstream 3′ UTR, as well as the 3′ UTR of the JAK3 transcript. There would appear to be a differential use of polyadenylation signals, since the RLF and JAK3 sequences defined by cDNA cloning have differing 3′ UTRs. Exon 1 of the mouse RLF gene is located in the middle of the new intron 23 of the JAK3 gene and appears not be used in any of the splice variants published for JAK3 in the mouse or rat.

From the literature JAK3 and RLF appear to have discrete tissue distributions, the former predominantly in lymphoid tissues, with additional low expression in lung, kidney and intestine [25], the latter in testicular Leydig cells [2]. Both genes are expressed maximally in the adult organism.

Transient transfection analysis of RLF promoter constructs indicated that sufficient information is contained within approximately only 0.7 kb of the 5′ region upstream of the RLF exon 1, to encode both Leydig cell specificity and regulatory elements. Thus also the complete promoter for the mouse RLF gene would appear to be contained within intron 23 of the JAK3 gene. In general agreement with the mouse genomic sequence encoding RLF recently published by Zimmerman et al. [9], we also find motifs in the promoter region corresponding to the transcription factor SF-1, which is known to mediate regulation of several genes in steroidogenic tissues [20]. Further experiments are necessary to determine which of these putative *cis* elements are used in vivo, and whether these alone are sufficient to explain both cell specificity and upregulation. However, the promoter region used for JAK3 expression would appear to be discrete from that for RLF, since the general pattern of tissue specificity is quite different. Furthermore, in Leydig cells, where the RLF gene is highly expressed, the RLF promoter sequence is markedly hypomethylated, whereas in tissues where this gene is not expressed the respective segment is considerably methylated (liver and spleen). The spleen, however, is a lymphoid tissue with moderate JAK3 expression [25], suggesting that hypomethylation in intron 22 is not needed for JAK3 gene expression.

Recently, Safford et al. [26] reported that a similar colocalization of the RLF (Ley-I-L) and JAK3 genes occurs on human chromosome 19p12-13.1. Thus it would appear that this colocalization is not a uniquely murine event but is a feature which has been conserved through evolution.

Acknowledgements: We should like to thank Rasoul Zarreh-Hoshyari-Khah for helpful discussions. This research was supported by grant Iv7/4-2-7 from the Deutsche Forschungsgemeinschaft, and by grants from the Foundation of the Finnish Cancer Societies, and from the University Foundation of Turku, Finland. Some of the results form part of the doctoral thesis of A.N.S. in the Faculty of Biology at the University of Hamburg.

References

[1] Adham, I.M., Burkhardt, E., Benahmed, M. and Engel, W. (1993) J. Biol. Chem. 268, 26668–26672.
[2] Pusch, W., Balvers, M. and Ivell, R. (1996) Endocrinology 137, 3009–3013.
[3] Ivell, R. (1997) Rev. Reprod. 2, 133–138.
[4] Büllesbach, E. and Schwabe, C. (1995) J. Biol. Chem. 270, 16011–16015.
[5] Bathgate, R., Balvers, M., Hunt, N. and Ivell, R. (1996) Biol. Reprod. 55, 1452–1457.
[6] Roche, P.J., Butkus, E., Wintour, E.M. and Tregear, G. (1996) Mol. Cell. Endocrinol. 121, 171–178.
[7] Ivell, R., Balvers, M., Domagalski, R., Ungefroren, H., Hunt, N. and Schulze, W. (1997) Mol. Hum. Reprod. 3, 101–108.
[8] Ivell, R., Pusch, W., Balvers, M., Kilic, E., Spiess, A.-N., Charlton, H. and Hunt, N. (1997) Biol. Reprod. 56, (Suppl. 1) 181.
[9] Zimmermann, S., Schöttler, P. and Adham, I. (1997) Mol. Reprod. Dev. 47, 30–38.
[10] Kumar, A., Toscani, A., Rane, S. and Reddy, E.P. (1996) Oncogene 13, 2009–2014.
[11] Sambrook, J., Fritsch, E. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[12] Church, G.M. and Gilbert, W. (1984) Proc. Natl. Acad. Sci. USA 81, 1991–1995.
[13] Kananen, K., Markkula, M., El-Hefnawy, T., Zhang, F.-P., Paukku, T., Su, J.-G.J., Hsueh, A.J.W. and Huhtaniemi, I. (1995) Mol. Endocrinol. 9, 616–627.
[14] Pechon, J.J., Behringer, R.L., Cate, K.A., Harwood, R.L. and Palmiter, R.D. (1992) Mol. Endocrinol. 6, 1403–1411.
[15] Kananen, K., Markkula, M., Rainio, E., Su, J.-G.J., Hsueh, A.J.W. and Huhtaniemi, I. (1996) Mol. Cell. Endocrinol. 119, 135–146.
[16] Rebois, R.V. (1982) J. Cell Biol. 94, 70–76.
[17] El-Hefnawy, T., Krawczyk, Z., Nikula, H., Viherä, I. and Huhtaniemi, I. (1996) Mol. Cell. Endocrinol. 119, 207–217.
[18] Showe, M., Williams, D. and Showe, L. (1992) Nucleic Acids Res. 12, 3153–3175.
[19] Lala, D., Rice, D. and Parker, K. (1992) Mol. Endocrinol. 6, 1249–1258.

[20] Parker, K. and Schimmer, B. (1997) Endocrinol. Rev. 18, 361–377.
[21] Denny, P., Swift, S., Connor, F. and Ashworth, A. (1993) EMBO J. 11, 3705–3712.
[22] Van Houte, L., Chuprina, V., Van der Wetering, M., Boelens, R., Kaptein, R. and Clevers, H. (1995) J. Biol. Chem. 270, 30516–30524.
[23] Gurniak, B.G., Thomis, D.C. and Berg, L.J. (1996) Blood 87, 3151–3160.
[24] Witthuhn, B.A., Silvennoinen, O., Miura, O., Lai, K.S., Cwik, C., Liu, E.T. and Ihle, J.N. (1994) Nature 370, 153–157.
[25] Takahashi, T. and Shirasawa, T. (1994) FEBS Lett. 342, 124–128.
[26] Safford, M.G., Levenstein, M., Tsifrina, E., Amin, S., Hawkins, A.L., Griffin, C.A., Civin, C.I. and Small, D. (1997) Exp. Hematol. 25, 374–386.