Isolation and characterization of the human aldehyde oxidase gene: conservation of intron/exon boundaries with the xanthine oxidoreductase gene indicates a common origin

Mineko TERAOKA, Mami KUROSAKI, Silvia DEMONTIS, Stefania ZANOTTA and Enrico GARATTINI¹

Laboratory of Molecular Biology, Centro Catullo e Daniela Borgomainerio, Istituto di Ricerche Farmacologiche 'Mario Negri', via Eritrea, 62, 20157 Italy

Aldehyde oxidase (AO) is a molybdo-flavo enzyme involved in the metabolism of various endogenous and exogenous N-heterocyclic compounds of pharmacological and toxicological importance. The enzyme is the product of a gene which is implicated in the aetio-pathogenesis of familial recessive amyotrophic lateral sclerosis. Here, we report the cloning and structural characterization of the human AO gene. AO is a single copy gene approximately 85 kb long with 35 transcribed exons. The transcription-initiation site and the sequence of the 5'-flanking region, containing several putative regulatory elements, were determined. The 5'-flanking region contains a functional promoter, as assessed by appropriate reporter constructs in transient

transfection experiments. Comparison of the AO gene structure shows conservation of the position and type of exon/intron junctions relative to those observed in the gene coding for another molybdo-flavoprotein, i.e. xanthine oxidoreductase (XOR). As the two genes code for proteins with a high level of amino acid identity, our results strongly suggest that the AO and XOR genetic loci arose as the consequence of a duplication event. Southern blot analysis conducted on genomic DNA from various animal species with specific cDNA probes indicates that the AO gene is less conserved than the XOR gene during evolution.

INTRODUCTION

Aldehyde oxidase (aldehyde: oxygen oxidoreductase, EC 1.2.3.1; AO) is an enzyme which belongs to the family of molybdo-flavoproteins along with xanthine oxidoreductase (XOR) and sulphite oxidase. These proteins are characterized by the presence of a molybdopterin cofactor which is essential for their catalytic activity [1–5]. AO plays an important role in the biotransformation of xenobiotics, since it is capable of reducing N-oxides, nitrosamines, hydroxamic acids, azo dyes, nitropolycyclic aromatic hydrocarbons [6] and sulphoxides [7]. In addition, the enzyme oxidizes a number of cancer chemotherapeutic agents, including methotrexate [8] and 6-methylthiopurine [9]. Consistent with its toxicological and pharmacological relevance, the enzyme is present at very high levels in the liver and lung [10]. More recently, we demonstrated that large amounts of the mouse AO transcript are also present in the epithelial component of the choroid plexus [11], the organ involved in the processes of absorption and secretion of endogenous and exogenous substances from and into the central nervous system. AO is also known to metabolize endogenous substrates like retinaldehyde [12,13] and dihydroxymandelaldehyde [14]. Retinaldehyde is the principal component of the visual pigments and a potential precursor of retinoic acid, the active metabolite of vitamin A, whereas dihydroxymandelaldehyde is a product of the catabolic pathway of adrenaline and noradrenaline. Thus, it is possible that AO plays a role in physiologically significant processes such as vision and synaptic transmission.

As to the significance of AO in pathological conditions, the protein is involved in the mechanisms of ethanol-generated hepatotoxicity [15] and may have importance in the generation

of some of the symptoms observed in the rare hereditary disease known as 'combined deficiency of molybdenum proteins' [16]. More recently, genetic evidence [17] has implicated AO in the genesis of the familial recessive form of amyotrophic lateral sclerosis (ALS), a rare and severe motor neuron disease characterized by progressive muscular paralysis leading to death. In this type of hereditary pathological state, the candidate gene maps on chromosome 2q33-q35, a very short distance from a genetic marker co-segregating with the disease [18]. This makes AO a likely candidate gene for familial recessive ALS [17]. In support of this hypothesis, we recently demonstrated, by *in situ* hybridization experiments, that mouse AO is selectively localized in the motor neurons of the cranio-facial and skeletal nerves [11]. Expression of the AO gene is limited to this cell type in the grey matter of the brain and spinal cord, as no other neuronal or glial cell populations contain the transcript at detectable levels [11].

From a structural point of view, the AO protein consists of two identical subunits of 150 kDa each. The two subunits contain two non-identical 2Fe/2S redox centres, a flavin and a molybdopterin binding site. The primary structure of the protein in cattle [10] and humans [19] has recently been deduced from the corresponding cDNAs. The amino acid sequence shows a relatively low level of similarity with AOs of plant origin [20,21]. In addition, human and bovine AOs show a remarkable level of amino acid identity with the XOR proteins [10,22–24]. There are no data on the genomic structure of AO in any organism of either plant or animal origin. In this report, we describe the isolation and structural characterization of the genetic locus coding for human AO, demonstrating that the intron/exon structure of the gene is complex and extremely similar to that of the gene coding

Abbreviations used: ALS, amyotrophic lateral sclerosis; AO, aldehyde oxidase; XOR, xanthine oxidoreductase.

¹ To whom correspondence should be addressed (e-mail eagarattini@irfmm.mnegr.it).

The nucleotide sequence data for the human AO gene exons and the 5'-flanking regions reported in this paper appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers Z99533–Z99567.

for XOR. In addition, we characterize from a structural and functional point of view the 5'-flanking region of the gene and we demonstrate that it contains a functional promoter.

MATERIALS AND METHODS

Cell lines

HepG-2, a human hepatoma cell line, and COS-7, a simian fibroblastic cell line, were obtained from the American Type Culture Collection (A.T.C.C., Rockville, MD, U.S.A.). Cell lines were cultured in Dulbecco's minimal essential medium (Gibco-BRL, Grand Island, NY, U.S.A.) supplemented with 10% (v/v) fetal calf serum (Gibco-BRL). Cells were negative for mycoplasma contamination as assessed by the Hoechst 33258 fluorescent dye system (Farbwerke Hoechst AG, Frankfurt, Germany).

Isolation and characterization of AO genomic clones

Lambda EMBL3 and EMBL4 libraries from human leucocyte DNA and from human placental DNA respectively were purchased from Clontech Laboratories, Inc. (Palo Alto, CA, U.S.A.). The phages containing the human AO gene were isolated by *in situ* plaque hybridization [25], using different human AO cDNA fragments as probes. Human AO cDNA fragments were PCR-amplified with specific amplimers designed on the basis of the published sequence [19], and subcloned in pBluescript (Stratagene, La Jolla, CA, U.S.A.) or in pMOS using the T/A cloning kit (Pharmacia, Uppsala, Sweden) according to the instruction of the manufacturer. The nucleotide sequence of each fragment was confirmed on both strands. Probes were radiolabelled with [α - 32 P]dCTP (Amersham, Little Chalfont, U.K.) to a specific activity of 1×10^9 c.p.m./ μ g according to the random priming method [26]. After overnight hybridization, the filters were washed twice in $2 \times$ SSC/0.1% SDS at room temperature for 10 min, followed by two washes in $0.1 \times$ SSC/0.1% SDS for 30 min each at 65 °C ($1 \times$ SSC is 0.15 M NaCl/0.015 M sodium citrate). Positive phages were isolated by two or three rounds of screening and characterized by restriction mapping, Southern blot analysis and subcloning of exon-containing DNA fragments. Overlapping lambda phages represent the entire structure of the human AO gene, except for a region containing exons 21–25. This region was amplified as eight overlapping fragments from human placental high-molecular-mass DNA by long-range PCR, using thermostable rTth DNA polymerase following the protocol suggested by the manufacturer (Cetus-Perkin Elmer, Branchburg, NJ, U.S.A.). The following fragments were amplified: AO-PCR1, containing part of exon 20, intron 20, exon 21, intron 21 and part of exon 22, which was amplified with the forward amplimer 5'-CAACACAACCTCTCTCAAGCCA3' (AO-oligo1), corresponding to nt 2128–2151, and the reverse amplimer 5'-GTTAGCTGGGAGCTTCAAGGTTGA3' (AO-oligo2) complementary to nt 2353–2376 of the published human AO cDNA sequence [19]; AO-PCR2, containing part of exon 21, intron 21, exon 22, intron 22 and part of exon 23, which was amplified with the forward amplimer 5'-AAAGCATGCTTGT-CGTTCCCAAGGG3' (AO-oligo3), corresponding to nt 2262–2285, and the reverse amplimer 5'-AGGATGGCGGCCTCC-AGTTATTAA3' (AO-oligo4) complementary to nt 2521–2544 of the AO cDNA; AO-PCR3, containing part of exon 22, intron 22, exon 23, intron 23 and part of exon 24, which was amplified with the forward amplimer 5'-TCAACCTGAAGCTCCAG-CTAAC3' (AO-oligo5), corresponding to nt 2353–2376, and the reverse amplimer 5'-TAATGCTCCATGTCCAGGGCCAAG3'

(AO-oligo6) complementary to nt 2589–2612 of the AO cDNA; AO-PCR4, containing part of exon 24, intron 24, exon 25, intron 25 and part of exon 26, which was amplified with the forward amplimer 5'-CTTGCCCTGGACATGGAGCATTAA3' (AO-oligo9), corresponding to nt 2589–2612, and the reverse amplimer 5'-CATGGCCAGTCCTTCTCTTCCA3' (AO-oligo10) complementary to nt 2998–3021 of the AO cDNA. In addition, four other fragments named AO-PCR15, AO-PCR16, AO-PCR17 and AO-PCR18 containing respectively part of exons 21–22 and intron 21, part of exons 22–23 and intron 22, part of exons 23–24 and intron 23, part of exons 24–25 and intron 25, were amplified with the following couples of oligonucleotides: AO-oligo3/AO-oligo2; AO-oligo5/AO-oligo4; 5'-GCAGTTCCG-CTGTGTTCTGGAACGA3' corresponding to nt 2485–2508 (AO-oligo7)/AO-oligo6; AO-oligo9/5'-GGCTGCAACTCCG-TGATACAAGA3' complementary to nt 2794–2817 (AO-oligo8). Relevant exons in each PCR DNA fragment were sequenced in both directions using appropriate oligonucleotide primers, either by direct sequencing of the PCR products using the Sequenase PCR product sequencing kit (United States Biochemical, Cleveland, OH, U.S.A.), or following cloning in pBluescript.

DNA sequence analysis

Appropriate DNA fragments were subcloned into the plasmid vector pBluescript and sequenced by the dideoxynucleotide-termination method [27] using double-stranded DNA as templates and T7 DNA polymerase (Pharmacia) according to the instructions of the manufacturer. Appropriate oligodeoxynucleotide primers were custom-synthesized by Duotech (Milan, Italy). Computer analysis of the DNA sequences was performed using the GeneWorks sequence analysis system (Intelligenetics, San Diego, CA, U.S.A.). A search for potential binding sites for transcriptional factors in the 5'-flanking region of the AO gene was performed using the MatInspector algorithm and the TRANSFAC database [28].

Southern blot analysis on genomic DNA

Southern blot analysis was performed according to standard procedures [25] on high-molecular-mass genomic DNA extracted from human placenta, bovine, mouse, rat, chicken and lizard liver, as well as *Drosophila melanogaster* total body and lizard tail. The probe used was either human AO cDNA or the entire coding region of bovine AO [10] and bovine XOR [23,29].

Determination of the 5'-end of the AO mRNA by primer extension and RNase mapping analysis

Total RNA was extracted from a human liver biopsy sample and from HepG2 cells. The poly(A)⁺ fraction of HepG2 RNA was selected according to standard protocols [25]. Primer extension analysis was conducted with two synthetic oligonucleotides (5'-TGTTGGTGTCCGCGCTGGTTCGAGG3' and 5'-TGGGG-CGGGCTTGCTGGAGCCTAAAG3') complementary to nt –1/–24 and –95/–121 of the human AO cDNA (Figure 6) labelled at their 5'-ends with [γ - 32 P]ATP by T4 polynucleotide kinase [25]. The primer (approximately 0.5 pmol, 3×10^6 c.p.m.) was mixed with 5 μ g of poly(A)⁺ RNA isolated from HepG2 cells and precipitated with ethanol. The pellet was resuspended in 20 μ l of extension buffer containing 0.01 M Tris/HCl (pH 8.3)/0.09 M KCl/0.002 M Mn(OAc)₂/0.0002 M NTP. Reverse

Figure 1 Organization of the human AO gene locus

(A) The middle diagram represents the AO gene locus, with transcription proceeding from left to right. Exons are shown as numbered vertical boxes. Horizontal lines above the Figure represent inserts of recombinant phages or PCR-amplified genomic DNA fragments and are labelled with their respective code numbers. Restriction endonuclease sites for *KpnI*, *SacI* and *XbaI* are indicated at the bottom of the Figure. (B) Subclones of the gene fragments containing exons (indicated as boxes) are shown. Black areas within the boxes represent translated regions, whereas white areas represent untranslated regions of the AO mRNA. Restriction sites are indicated as follows: B, *Bam*HI; E, *Eco*RI; Eg, *Eag*I; Hc, *Hinc*II; H, *Hind*III; K, *Kpn*I; P, *Pvu*II; Pt, *Pst*I; S, *Sac*I; X, *Xba*I. Not all of the restriction sites are listed.

transcription was started by the addition of 5 units of rTth DNA polymerase (Cetus-Perkin-Elmer). After incubation for 1 h at 70 °C, 250 µg/ml of DNase-free RNase A was added and further incubated at 37 °C for 30 min. The reaction was stopped by the addition of 4 µl of 0.5 M NaOH and heated at 100 °C for 3 min. The primer-extended products were neutralized by the addition of 4 µl of 0.5 M HCl and 2 µl of 2 M Tris/HCl (pH 7.5) and precipitated with ethanol. The pellets were resuspended in 95 % formamide/0.02 M EDTA/0.05 % Bromophenol Blue/0.05 % Xylene Cyanol FF, heated to 95 °C and electrophoresed on a 6 % (w/v) polyacrylamide/7 M urea sequencing gel.

Preparation of the riboprobes and RNase mapping analysis were performed according to the method of Melton et al. [30] with minor modifications. A *KpnI*-*EagI* genomic fragment containing 1.5 kb of the 5' non-coding region and the first exon derived from the phage clone 10A (Figure 1) was subcloned in pBluescript KS⁺ (pAO-D15Eag). This plasmid was linearized with *Pvu*II and used as a template for *in vitro* transcription with T7 RNA polymerase (Promega Inc., Madison, WI, U.S.A.) in the presence of [α -³²P]UTP, according to the manufacturer's instructions. The probe used for the detection of the AO transcript was a 475 nt antisense RNA, consisting of 403 nt of the 5' non-coding region, 42 nt of the coding sequence of exon 1 and 30 nt

of the polylinker sequence of pBluescript. Hybridization was carried out at 45 °C for 16 h and RNase digestion was at 37 °C for 30 min using a mixture of RNase A and RNase T1 in appropriate buffers (RPA kit, Ambion, Austin, TX, U.S.A.). RNase-protected fragments were electrophoresed on a 6 % polyacrylamide/7 M urea sequencing gel.

Analysis of the promoter activity of the human AO gene 5'-flanking region

To construct pAOs-Luc and penhAOs-Luc, which contain approximately 1.5 kb of 5'-flanking region of the human AO gene in the sense orientation in front of the luciferase reporter gene, the plasmid pAO-D15Eag was cleaved with *EagI*, the resulting sticky ends were blunted with the Klenow fragment of DNA polymerase I, and the AO gene 5'-flanking region and exon 1-containing sequence was released with *KpnI*. The resulting fragment was inserted into the plasmid pGL-Basic vector (Promega Inc.; renamed p0-Luc in this report) or pGL-Enhancer vector (penh0-Luc), which were modified to have one blunt end at the *XhoI* site and a *KpnI* sticky end. To construct pAOas-Luc and penhAOas-Luc containing the 1.5 kb of the AO gene 5'-

flanking region and exon 1 sequence in the opposite orientation relative to the reporter gene, pAO-D15Eag was digested with *KpnI* and *EagI*, blunt-ended by the Klenow fragment of DNA polymerase I and T4 DNA polymerase. The resulting fragment was inserted in pGL-Basic vector or in pGL-Enhancer vector blunt-ended at the *XhoI* site.

The plasmids pAAOs-Luc and penhAAOs-Luc were generated by the insertion of the *KpnI*–*SacI* fragment derived from pAO-D15Eag into the *KpnI*–*SacI* sites of pGL-Basic vector and pGL-Enhancer vector respectively. To construct the plasmids pAAOas-Luc and penhAAOas-Luc, pAO-D15Eag was cleaved by *KpnI*, blunt ended and digested with *SacI*. The resulting fragment was inserted in the *XhoI*-blunt-ended and *SacI*-digested pGL-Basic vector and pGL-Enhancer vector respectively. The plasmid pRSVLuc contains the luciferase gene under the control of the Rous sarcoma virus promoter and served as a positive control for the transfections. The COS-7 and HepG2 cells were transfected with circular plasmid DNA [$1.6 \mu\text{g}/(3\text{--}4) \times 10^5$ cells] using DMRIE-C[®] cationic liposomes (Gibco-BRL, Grand Island, NY, U.S.A.) according to the instructions of the manufacturer. To monitor transfection efficiencies, $0.4 \mu\text{g}$ of the β -galactosidase expression vector pCH110 (Pharmacia) was cotransfected with the luciferase constructs. Forty-eight hours after transfection, cells were harvested and lysed and aliquots were assayed for luciferase activity using a commercial kit (Promega) and for β -galactosidase activity, as detailed in a previous paper [31].

RESULTS AND DISCUSSION

Structure of the human AO gene

To determine the structure of the human AO gene, we isolated a series of overlapping lambda phage clones and PCR-derived DNA fragments that encompass the complete gene (Figure 1). The gene extends for approximately 85 kb and consists of 35 exons. The positions of exons were located by a combination of restriction mapping of the genomic clones shown in Figure 1 and Southern blot analysis using appropriate probes from the human AO cDNA [19]. The length of the intervening intronic sequences was verified by long-range PCR analysis using specific exonic amplimers. The physical map of the gene obtained from the various lambda clones was confirmed by Southern blot analysis of genomic DNA extracted from human placenta using various regions of the cDNA as probes (Figure 2). All the bands in the Southern blot analysis are accounted for by the physical map shown in Figure 1. Our data demonstrate that there is a single copy of the AO gene in humans. This is different from what was reported in the mouse, where classical genetic analysis suggested the presence of two genetic loci (*AOX-1* and *AOX-2*) [32], which possibly code for two structurally related AO isoenzymes differentially regulated in their expression by androgenic steroids [33].

All of the exon/intron junctions, determined by comparison of the genomic and cDNA sequences, conform to the consensus sequences established for other eukaryotic genes: all introns start with GT and end with AG (Table 1). The first exon contains the entire 5' untranslated region and encodes the putative start codon and the first 15 amino acids. The coding region extends from exon 1 to exon 35. The last 16 amino acids are located on exon 35. Except for the last exon, which is 861 bp long and contains mainly the 3' untranslated region, all the other exons are relatively short and range in length from 53 bp (exon 30) to 228 bp (exon 26).

It is not known whether the exons encode discrete functional and structural domains of the protein. However, as shown in Figure 3, the sequence corresponding to the two [2Fe-2S] redox

Figure 2 Southern blot hybridization of human genomic DNA

Genomic DNA obtained from human placenta was digested with the indicated enzymes, size fractionated ($20 \mu\text{g}/\text{lane}$) on a 1.0% agarose gel, transferred to a nylon membrane and hybridized to radiolabelled full-length cDNA (A) or parts of human AO cDNA (B) amplified by RT/PCR using oligonucleotides designed on the basis of the published sequence of human AO cDNA [19]. The positions of molecular-mass-markers (kb) are shown on the left side of each panel. The probes used for the experiment shown in (B) are regions: 1, 1–995; 2, 984–2333; 3, 2243–3033; 4, 3033–4011; 5, 44011–4820 of the published sequence of the human AO cDNA.

centres, as defined by similarity with the structure of the molybdenum-containing protein from *Desulfovibrio gigas* (DvgMOP), crystallized by Romao et al. [34], are split by exon/intron junctions 3/4 and 5/6 respectively. In addition, the junction between exons 22 and 23 is located within the signature sequence present in all the molybdenum-cofactor containing proteins [35]. Finally, positioning of exon/intron boundaries relative to a low-resolution three-dimensional structure of the AO protein obtained by modelling against DvgMOP (L. De Gioia, personal communication) demonstrates several inter-

Table 1 Exon-intron organization of the human AO gene

Exon sequences are shown in uppercase letters, and intron sequences in lowercase letters. The positions of nucleotides that border the 5' and 3'-ends of the exons are numbered below the DNA sequence. Numbering starts at the A of the first methionine codon of the published human AO cDNA sequence [19] which is designated 1 and takes into account the six extra nucleotides (amino acid residues 354 and 355) which are present in the deposited EMBL nucleotide sequence (accession number L11005), but not in the originally published sequence. Nucleotides within introns are not numbered. Amino acids bordering the splice junctions are shown above the nucleotide sequence. The region of the mRNA cleavage/polyadenylation signal at the 3'-end of exon 35 is underlined. The length of the first exon is calculated on the basis of the downstream transcription-initiation site (-56 from the first ATG, Figure 6).

Exon	Exon size (bp)	5' Splice site	Intron size (kb)	3' Splice site	Exon	Exon size (bp)	5' Splice site	Intron size (kb)	3' Splice site
1	101	..Gly Arg Lys ..GCC CGC AAGTgagccccc.. 37	6.5	..tttggtatagGTG ATA GAA.. 46	19	123	..Thr Ile Glu ..ACA ATT GAGtaatgcacg.. 2116	3.0	..ctctttctagGAA AGT ATA.. 2125
2	58	..Lys Lys L ..AAG AAG Ggtatcctttt.. 97	2.2	..cctcttcagATT CCA CTC.. 104	20	97	..Ile Leu Glu G ..ATT CTT GAA Gtataaaagta.. 2212	6.0	..ctgtaacagagGT GAA ATA CAT.. 2222
3	97	..Arg Ile Ar ..AGG ATA AGgtaccagca.. 193	2.2	..cattctgaagG CAT CAC CCA.. 201	21	125	..Tyr Ile Gln ..TAT ATA CAGgtacaatggg.. 2338	2.3	..ttgttccacagCAC ATT GTT.. 2347
4	109	..Pro Val Gln ..CCT GTT CAGgtgaggaatt.. 301	1.8	..gtgcttccagGAG AGG ATT.. 310	22	134	..Ala Ala Asn Lys ..GCC GCA AAC AAgtaagtggag.. 2470	1.5	..tctgttttttAGA CAT GCC CGT.. 2481
5	127	..Ala Leu Gly G ..GCC CTT GGT Ggtagttata.. 427	2.5	..tctcttcagATT AAC CTG.. 437	23	88	..Lys Tyr Lys ..AAG TAC AAGgtgagtacaa.. 2560	2.8	..tcttcccccagGCT GGA TTC.. 2569
6	62	..Phe Cys Lys ..TTC TGT AAGtaagtggaa.. 490	1.0	..tgtccactagACT TCG GGC.. 499	24	87	..Ser Leu Phe ..TCA TTA TTCgttaagtgttt.. 2649	1.5	..atttctccagGTG ATA GAA.. 2658
7	90	..Gly Ser Lys ..GGA AGT AAGtctcagtaaa.. 580	1.3	..tttttttaagACA AGT CCA.. 589	25	192	..Pro Glu Lys ..CTT GAG AAGgtcaatactaa.. 2839	6.5	..ttcttccagGTG CGA ATC.. 2848
8	81	..Glu Leu Met ..GAG CTA ATGgtgagtaaa.. 661	0.7	..tttttcaatagATA ATG GCT.. 670	26	228	..Ala Gly Gln ..GCT GGT CAGgtgagttctc 3067	5.0	..tctgtccagagGCT GCT GCC.. 3076
9	144	..Thr Ser Val G ..ACC TCT GTG Ggtatgtgacc.. 805	3.0	..gatcctttagGG CCT GAA.. 815	27	96	..Met Ile Gln ..ATG ATT CAGgttaagaatgc.. 3163	2.5	..taaaattccagGTG GTC AGC.. 3172
10	93	..Ala Tyr Asn G ..GCA TAT AAT Ggtgagttctc.. 898	3.0	..cttccatttcaagCA CTC ACC.. 908	28	129	..Ala Val Lys ..GCA GTA AAGgttaacagta.. 3292	0.5	..taatctcttagCAT GCC TGT.. 3301
11	152	..Asn Met Ala ..AAC ATG CCGtatgtatct.. 1051	0.3	..ttttttaaagTCT TTA GGG.. 1060	29	75	..Lys Asp Trp ..AAA GAC TGGgtgaggaatc.. 3367	1.5	..tcttccacagGCA CAG ACT.. 3376
12	94	..Leu Ser Lys G ..CTA TCA AAA Ggttaagtaca.. 1144	2.5	..ctttccacagAAA GGA AAA.. 1154	30	53	..Gly Tyr Phe Ar ..GGA TAC TTC AGgttaatactc.. 3418	1.5	..tttttccagA GGT GAG.. 3429
13	110	..Ser Arg Lys ..TCA AGG AAGgtgagaacct.. 1255	1.2	..tttttcttcagTGG GAA TTT.. 1264	31	115	..Asp His Lys ..GAT CAT AAGgttcagtaacc.. 3535	3.5	..ttaataactagAAC ATC GAA.. 3544
14	185	..Leu Ile Gly Ar ..CTC ATT GGA AGgttaagcatt.. 1438	1.2	..gaatttccagG CAC TGG.. 1449	32	66	..Ile Gly Gln ..ATA GGC CAGgtcagtgtaa.. 3601	2.5	..cttgcctcagATT GAA GGT.. 3610
15	163	..Lys Lys Met ..AAA AAG ATGtaaacaaact.. 1603	2.5	..tacttttccagCAT CCA GGT.. 1612	33	189	..Ser Ser Lys ..TCA TCT AAGgttaagtaag.. 3790	1.0	..tcttcttcagGCT CTG GGA.. 3799
16	93	..Lys Tyr Gln ..AAG TAC CAGgtgagcggtat.. 1696	3.5	..tatttaaacagAAT ATA GGC.. 1705	34	168	..Thr Lys Met ..ACA AAA ATGgtcagtttgg.. 3958	0.9	..tttatttttagATT CCG AGA.. 3967
17	170	..Ala Lys Ile Va ..GCT AAG ATT Ggttaagtggta.. 1864	0.7	..tatcatctagG TCT ATT.. 1875	35	861	..ATTAATAATTTTAAATGGcaaaaggtattgtagatgtagtcttccaaagaagataataaagtcgtagcaagcattaaaaa tatgctcaaacctcattagctcatttaaggaatgcaatcaaaccaatcaggtgagccaccattttt....		
18	127	..Thr Asp Lys ..ACA GAT AAGgtactgcatt.. 1993	4.0	..tgtgtctgtagGTG TTC TGT.. 2002					

The human and bovine AO proteins share a remarkable degree of sequence identity with mammalian XORs [10]. Furthermore, AO and XOR have similar molecular mass, subunit composition, redox centre distribution as well as an overlapping set of substrates [9,10,35–37]. We therefore compared the intron/exon organization of the human AO and XOR genes and found them to be almost identical (Figure 3). The exceptions were the boundaries between exons 7/8 and 15/16, and the suppression of one intron (AO exon 26 is the exact sum of XOR exons 26 and

Figure 4 Southern blot hybridization of DNA from various animal species

Genomic DNA obtained from the indicated animal species was digested with *Bam*HI, size fractionated (20 μ g/lane) on a 1.0% agarose gel and transferred to a nylon membrane. The same filter was sequentially hybridized with a fragment corresponding to the entire coding region of bovine AO [10] and bovine XOR [29]. The hybridization and washing were performed under high-stringency conditions. The data presented are representative of two independent experiments giving identical results. The positions of DNA size markers are shown on the left.

27 due to suppression of intron 26). Not only are the positions of 33 out of 35 exons the same in the AO and XOR genes, but also the type of intron/exon junction is identical. Whenever an exon/intron junction in the AO gene is placed after the first (type I), second (type II) or third nucleotide (type O) of a codon, an identical type of junction is correspondingly observed in the XOR gene. If this analysis is extended to the comparison between human AO and mouse XOR [31], a similar situation is observed (results not shown), since the mouse XOR protein and respective gene are extremely similar to the human counterparts. As conservation in the position of exon/intron boundaries is a characteristic of multigene families, whose members originate by successive duplication events [38,39], our data suggest that mammalian AO and XOR also derived from gene duplication. It is likely that this duplication event is a relatively recent one and follows the remarkable exon multiplication phenomenon observed in the XOR gene during phylogenesis. In fact, it is known that in *Aspergillus nidulans* [35], *Drosophila* [40,41] and *Calliphora* [42], the XOR gene has a simple structure consisting of 3–5 exons, as opposed to the 36 exon structure of the human and mouse counterparts [31,43]. Gene duplications are generated by a number of different mechanisms. In the case of AO and XOR, localization of the two human genetic loci on the same chromosome (Chr. 2) at a short distance from each other would be consistent with a tandem duplication event. In the mouse, the AO and XOR genetic loci map to different chromosomes (Chr. 1 [32] and Chr. 17 [31] respectively). However, these two chromosomes contain clusters of genes whose human homologues are located on chromosome 2. Thus it is possible that the AO/XOR gene couple is linked on the same chromosome in animal species that are evolutionarily more primitive than the mouse.

To gain insight into the structural conservation of the AO/XOR gene couple during phylogenesis, we performed Southern blot analysis on DNA extracted from different organisms representing distinct steps in the process of evolution. Experiments were performed under the most stringent conditions

with probes consisting of the entire coding sequence of bovine AO and XOR cDNAs [10,29]. As shown in Figure 4, the AO cDNA probe recognizes a number of genomic bands after hybridization with DNA from all the mammalian species analysed as well as from chicken and lizard. However, cross-hybridization of AO was not observed in *Drosophila melanogaster* or toad, even upon much longer exposures of the autoradiogram. When the AO probe was stripped off and the same filter was rehybridized with the XOR cDNA, clear signals of similar intensity were observed in all the animal species, including *D. melanogaster* and toad. In *Drosophila*, the number and the size of bands recognized are consistent with those expected on the basis of the available physical map of the XOR gene [40]. In chicken, the XOR cDNA probe gave rise to a smeared signal, probably because it recognizes highly repetitive sequences within the genome of this avian species. Notice that the pattern of bands recognized by XOR in each species is different from that highlighted by the AO cDNA, illustrating the specificity of hybridization.

Our results indicate that the structure of the mammalian AO gene is very similar to that of reptiles and birds, whereas it diverges from that of insects and amphibians. Thus, relative to AO, the XOR gene shows a much higher degree of conservation during phylogenesis. The presence of molybdo-flavoproteins with AO enzymic activity has been reported in bacteria [44] and plants [20,21], organisms that are known to express XOR enzymic activity as well. Plant AOs show an overall level of identity with mammalian AO and XOR of approximately 30% ([20,21]; M. Terao, M. Kurosaki, S. Demontis, S. Zanotta and E. Garattini, unpublished work). In contrast, identity between *A. nidulans* XOR, the most primitive organism for which an amino acid sequence of the protein is available [35], and mammalian XORs is 40–46%. This is consistent with our Southern blot analysis and reinforces the idea that the level of conservation among AO proteins during phylogenesis from lower organisms to mammals is significantly lower than that of XORs. In addition, the AO/XOR couple in cattle and humans [10] is more than 50% identical. Based on this evidence, we propose that the AOs present in lower organisms have a different evolutionary origin from mammalian AOs and the two types of enzymes serve a fundamentally different biological function.

Determination of the transcription-initiation site, sequence analysis and functional characterization of the 5'-flanking region of the AO gene

To define the 5'-end of exon 1 of the human AO gene, RNase mapping analysis was performed on total RNA obtained from human liver and the HepG2 hepatoma cell line (Figure 5A). The RNA probe used was 475 nt long, containing 403 nt of 5' non-coding region and 42 nt of coding region. In human liver, four prominent protected bands of 96–97 nt and 104–105 nt were observed, based on the mobility of the RNA fragments relative to the DNA sequencing ladders. This locates two major transcription initiation sites at 54–55 and 62–63 nt upstream of the putative first ATG codon. Upon longer exposure of the autoradiogram, at least two other larger RNA fragments of approximately 138 and 211 nt were protected, suggesting the presence of two further minor transcription initiation sites at approximate positions –96 and –169 from the first ATG codon. In HepG2 cells, two other protected RNA fragments were detected further upstream around positions –170 and –250. The presence of the HepG2-specific RNase-protected fragments, as well as the difference in the intensity of other fragments commonly observed

Figure 5 Mapping of the 5' end of human AO mRNA

(A) RNase protection analysis. The radiolabelled riboprobe (5×10^5 c.p.m./reaction) consisted of 403 nt of the 5' non-coding region, 42 nt of the coding region of exon 1 and 30 nt of pBluescript. RNase-protected bands were analysed on a 6% polyacrylamide/urea sequencing gel in parallel with a DNA sequencing ladder shown on the left side. The position of the major protected band is indicated by symbols on the right. Lane 1, 10 μ g of yeast tRNA; lane 2, 10 μ g of total liver RNA; lanes 3 and 4, longer exposure of the autoradiogram shown in lanes 1 and 2; lane 5, 20 μ g of yeast tRNA; lane 6, 20 μ g of total HepG2 RNA; lane 7, 2 μ g of HepG2 poly(A)⁺ RNA. RNase-protected bands common to liver and HepG2 are indicated by identical symbols on the right side of each panel. (B) Primer extension analysis. HepG2 poly(A)⁺ RNA (5 μ g) (lanes 2 and 4) or an equivalent amount of yeast tRNA (lanes 1 and 3) were mixed with the ³²P-labelled oligonucleotide primer A (complementary to nt -1/-24 of the human AO gene; lanes 1 and 2) or primer B (complementary to nt -95/-121 of the human AO gene; lanes 3 and 4) and reverse transcribed with thermostable rTth DNA polymerase. Extended products were analysed on a 6% polyacrylamide/urea gel. The positions of DNA size markers of known molecular mass (*Hpa*II-digested pBluescript) are shown on the left. Extended bands which correspond to RNase-protected fragments are indicated on the right side with symbols identical with those used in (A). The experiments are representative of at least two others giving essentially the same results.

in liver and HepG2 cells, suggests that transcription initiation may be guided by cell-specific processes.

To confirm the location of the transcription-initiation sites defined by RNase mapping, primer-extension analysis was performed on poly(A)⁺ RNA isolated from the HepG2 cells using an oligonucleotide primer (primer A) complementary to nt -1/-24 of the human AO gene (Figure 5B). Two downstream transcription-initiation sites were identified as 49 nt and 55-57 nt extended fragments. Considering that, under our experimental conditions, the electrophoretic mobility of RNA bands is approximately 10% slower than that of DNA bands of corresponding length [17], the position of the two transcription-initiation sites is coincident with that defined by RNase mapping analysis. Primer extension defines other upstream transcription-initiation sites, some of which are consistent with those determined by RNase mapping analysis. Two very long extension products (larger than 320 bp) were consistently observed with two distinct oligonucleotide primers (Figure 5B, lanes 2 and 4). These bands are

prominent in HepG2 cells, although they are also detectable in liver (results not shown). In conclusion, the data demonstrate the presence of multiple transcription-initiation sites throughout the first 403 bases upstream of the first ATG, although the most active in liver are those located in the vicinity of the putative translation-initiation site.

The 5'-flanking region of the hAO gene up to nt -1510 was sequenced in both directions (Figure 6). The 5'-flanking region does not contain any canonical TATA box in the vicinity of the two major transcription-initiation sites, in line with the presence of RNA species with different 5' ends. Four ATG sequences were found between the putative first methionine and the most upstream transcription-initiation sites identified by primer-extension analysis in both liver and hepatoma cells. However, all of them are contained within the context of unfavourable translation start sites, suggesting that the only translationally active ATG codon is the one already defined by cloning of the human AO cDNA [19]. In fact, this is the only ATG surrounded

The nucleotide of the first methionine codon is numbered as +1. Upstream sequences are indicated by negative numbers. The 5'-flanking as well as intron sequences are shown by lowercase letters, whereas exon 1 sequences are shown by uppercase letters. Consensus sequences for the binding of known transcriptional factors are indicated by arrows above the sequence. The position of endonuclease restriction sites (*KpnI*, *SacI* and *EagI*), which were used for the generation of luciferase reporter gene constructs (see Table 2) are underlined. The positions of the major downstream transcription-initiation sites are indicated by dots below the sequence.

by an optimal Kozak's consensus sequence [45]. Multiple sequences, similar to those binding various types of transcription factors were determined by computer-assisted analysis. Given that the AO gene is expressed at high levels in the liver [10], the presence of multiple sites for the binding of liver-specific transcription factors like CCAAT-enhancer-binding protein (C/EBP) and HNF-3b is of particular interest. In addition, the AO gene 5'-flanking region contains several binding sites for homeobox-containing transcriptional factors such as Nkx-2 and Ftz, which raises the possibility that AO is a developmentally regulated gene. Finally, the two cAMP-response element binding protein (CREB) binding sites and the signal transduction and activators of transcription (STAT) binding site suggest cAMP-dependent and cytokine-dependent regulation respectively. It remains to be established whether the sites are functionally active. The 5'-flanking region does not show any resemblance to the corresponding region of human [43], mouse [31] or rat [46] XOR, although all these genes, like AO, are characterized by multiple transcription-initiation sites, consistent with the absence of canonical TATA boxes. This suggests that the two genes are controlled by different transcription factors, which is consistent with a different tissue-specific distribution of the two enzymes [29] and regulation by different endogenous and exogenous stimuli ([37]; M.Terao and E. Garattini, unpublished work).

To test whether the AO gene 5'-flanking region has promoter activity, transient transfection experiments with plasmid constructs containing the firefly luciferase as a reporter gene were performed (Table 2). HepG2 hepatoma and COS-7 fibroblastic cell lines were selected to evaluate potential cell-specific promoter activity of the AO gene 5'-flanking region, since the former expresses significant amounts of the AO transcript, whereas the latter does not synthesize the mRNA at detectable levels (results not shown). The 5'-flanking region of the AO gene is characterized by the presence of promoter activity, as demonstrated by the strong increase in luciferase enzymic activity relative to what was observed following transfection of the corresponding promoterless reporter gene. As expected for a functional promoter element, the transcription-enhancing effect of the AO gene 5'-flanking region is orientation-dependent (pAOs-Luc vs pAOas-Luc). If the level of luciferase activity is normalized for the expression of a cotransfected plasmid containing the β -galactosidase reporter under the control of a strong constitutive viral promoter, the AO promoter activity is more evident in HepG2 relative to COS-7 cells. This suggests that the 5'-flanking region contains sequences that are responsible for the cell-specific expression of the AO gene. In HepG2 cells, the promoter activity is stimulated 3–4-fold in an orientation-specific manner in the presence of an SV40 enhancer element positioned downstream of the reporter gene (penhAOs-Luc).

As primer extension and RNase mapping analysis indicated the presence of multiple upstream transcription-initiation sites, we performed transient transfection experiments with a series of other constructs containing a shorter version of the AO gene 5'-flanking region. Though these constructs do not contain the major downstream transcription-initiation sites (p Δ AOs-Luc, p Δ AOas-Luc, penh Δ AOs-Luc and penh Δ AOas-Luc), they maintain a functional promoter acting in an orientation-specific manner in both HepG2 and COS-7 cells. However, the level of promoter activity is much lower than that observed with the intact 5'-flanking region. This indicates that the first 120 bp sequence of the AO 5'-flanking region is essential for strong basal promoter activity. The region acts in a cell-context-independent way, allowing the expression of luciferase in both HepG2 and COS-7 cells. On the basis of these results, we conclude that the downstream transcription-initiation sites are the most relevant

Table 2 Functional activity of the hAO 5'-flanking region

The plasmid p0-Luc contains a promoterless luciferase reporter gene. The plasmid penh0-Luc is the same as p0-Luc except for the presence of an SV40 enhancer sequence located downstream of the luciferase reporter gene. The plasmids pAOs-Luc and pAOas-Luc contain a 1.5 kb *KpnI*–*EagI* fragment of the AO gene 5'-flanking region (see Figure 6 for the sequence) inserted upstream of the luciferase reporter gene in the sense (pAOs-Luc) or antisense (pAOas-Luc) orientation. The plasmids p Δ AOs-Luc and p Δ AOas-Luc contain a 1.2 kb *KpnI*–*SacI* fragment of the AO gene 5'-flanking region (see Figure 6 for the sequence) inserted upstream of the luciferase reporter gene in the sense (p Δ AOs-Luc) or antisense (p Δ AOas-Luc) orientation. The plasmids penhAOs-Luc, penhAOas-Luc, penh Δ AOs-Luc and penh Δ AOas-Luc are identical with pAOs-Luc, pAOas-Luc, p Δ AOs-Luc and p Δ AOas-Luc respectively, except for the presence of an SV40 enhancer sequence located downstream of the luciferase reporter gene. COS-7 or HepG-2 cells were transfected with the indicated plasmid along with pCH110 (containing the bacterial β -galactosidase gene under the control of the SV40 early promoter). Forty-eight hours after transfection, aliquots of cell lysates were subjected to the measurement of luciferase and β -galactosidase activity. Luciferase activity is defined in arbitrary units following normalization for the efficiency of transfection through measurement of β -galactosidase activity. The level of expression of each construct, relative to that of the promoterless p0Luc or penh0Luc, taken as 1, is indicated in parentheses (fractions of each value are rounded to the nearest whole integer). Results are the means $\pm$ S.D. of three different culture dishes. The results are representative of three experiments, which gave essentially the same values.

Plasmid	10 ³ $\times$ Luciferase activity (chemiluminescence arbitrary units)	
	HepG2	COS-7
p0-Luc	2.0 $\pm$ 0.3 (1)	24.6 $\pm$ 0.5 (1)
pAOs-Luc	1282.9 $\pm$ 221.6 (641)	301.2 $\pm$ 59.5 (12)
pAOas-Luc	95.9 $\pm$ 19.3 (48)	14.2 $\pm$ 2.1 (1)
p Δ AOs-Luc	89.3 $\pm$ 9.5 (45)	36.5 $\pm$ 8.2 (1)
p Δ AOas-Luc	< 0.5	< 0.5
penh0-Luc	1.8 $\pm$ 0.1 (1)	12.3 $\pm$ 1.8 (1)
penhAOs-Luc	4203.1 $\pm$ 177.9 (2335)	1149.8 $\pm$ 169.2 (93)
penhAOas-Luc	46.0 $\pm$ 16.1 (26)	58.5 $\pm$ 1.4 (5)
penh Δ AOs-Luc	389.6 $\pm$ 29.3 (216)	101.2 $\pm$ 11.5 (8)
penh Δ AOas-Luc	6.6 $\pm$ 0.2 (4)	6.6 $\pm$ 0.1 (1)
pRSV-Luc	40678.3 $\pm$ 3749.1	3805.9 $\pm$ 862.0

for the expression of AO in liver and in HepG2 cells. It remains to be established whether these initiators are also involved in the expression of the gene in other cell types, such as the motoneuron and the epithelial component of the choroid plexus in the central nervous system [11].

Besides its significance from an evolutionary point of view, the molecular cloning of the gene coding for AO is an important first step in the definition of the regulatory sequences and factors involved in the tissue-specific expression of the enzyme in humans. In addition, it is relevant for the definition of the role that AO plays in the genesis of familial recessive [17,18], and perhaps other forms of, sporadic ALS. As to this last point, Southern blot experiments conducted on DNA extracted from several patients suffering from the sporadic form of ALS failed to demonstrate gross rearrangement in the structure of the AO gene (results not shown). However, definition of the AO exonic structure should allow the application of the single-strand chain polymorphism technique for the determination of potential point mutations in the coding region of the AO gene of familial and sporadic cases of ALS. Studies with this aim are in progress.

This work was supported in part by grants from the Telethon-Italy (grant no. 839, to M.T.) and from the Associazione per la Ricerca contro il Cancro (AIRC). S.D. is the recipient of a fellowship from the Regione Sardegna. We are grateful to Dr. Gregory Goodall, Professor Silvio Garattini, Dr. Mario Salmons and Dr. Luca De Gioia for useful suggestions and critical reading of the manuscript.

REFERENCES

- 1 Hille, R. and Massey, V. (1985) in *Molybdenum Enzymes* (Spiro, T. G., ed.), pp. 443–518, Wiley Interscience, New York
- 2 Barber, M. J., Coughlan, M. P., Rajagopalan, K. V. and Siegel, L. M. (1982) *Biochemistry* **21**, 3561–3568
- 3 Rajagopalan, K. V. and Johnson, J. L. (1992) *J. Biol. Chem.* **267**, 10199–10202
- 4 Johnson, J. L., Rajagopalan, K. V. and Meyer, O. (1990) *Arch. Biochem. Biophys.* **283**, 542–545
- 5 Krenitsky, T. A. (1978) *Biochem. Pharmacol.* **27**, 2763–2764
- 6 Yoshihara, S. and Tatsumi, K. (1985) *Arch. Biochem. Biophys.* **242**, 213–224
- 7 Kitamura, S., Narai, N., Hayashi, M. and Tatsumi, K. (1983) *Chem. Pharm. Bull. (Tokyo)* **31**, 776–779
- 8 Jacobs, S. A., Stoller, R. G., Chabner, B. A. and Johnes, D. G. (1976) *J. Clin. Invest.* **57**, 534–538
- 9 Krenitsky, T. A., Neil, S. M., Elion, G. B. and Hitchings, G. H. (1972) *Arch. Biochem. Biophys.* **150**, 585–599
- 10 LiCalzi, M., Raviolo, C., Ghibaudi, E., De Gioia, L., Salmona, M., Cazzaniga, G., Kurosaki, M., Terao, M. and Garattini, E. (1995) *J. Biol. Chem.* **270**, 31037–31045
- 11 Bendotti, C., Kurosaki, M., Garattini, E. and Terao, M. (1997) *Neuroreport* **8**, 2343–2349
- 12 Tomita, S., Tsujita, M. and Ichikawa, Y. (1993) *FEBS Lett.* **336**, 272–274
- 13 Huang, D.-Y. and Ichikawa, Y. (1994) *Biochem. Biophys. Res. Commun.* **205**, 1278–1283
- 14 McGeer, P. L., Eccles, J. C. and McGeer, E. G. (1978) in *Molecular Neurobiology of the Mammalian Brain*, pp. 233–293, Plenum Press, New York
- 15 Mira, L., Maia, L., Varreira, L. and Manso, C. F. (1995) *Arch. Biochem. Biophys.* **318**, 53–58
- 16 Holmes, E. W. and Wyngaarden, J. B. (1989) *Hereditary Xanthinuria in The Metabolic Bases of Inherited Disease*, 6th Edn., (Scriver C. V., Beaudet, A. L., Sly, W. S. and Valle D., eds.), McGraw-Hill, New York
- 17 Berger, R., Mezey, E., Clancy, K. P., Harta, G., Wright, R. M., Repine, J. E., Brown, R. H., Brownstein, M. and Patterson, D. (1995) *Somatic Cell Mol. Genet.* **21**, 121–131
- 18 Hentati, A., Bejaoui, K., Pericak-Vance, M. A., Hentati, F., Speer, M. C., Hung, W.-Y., Figlewicz, D. A., Haines, J., Rimmeler, J., Hamida, C.-B., Hamida, M.-B., Brown, R. H. and Siddique, T. (1994) *Nature Genet.* **7**, 425–428
- 19 Wright, R. M., Vaitaitis, G. M., Wilson, C. M., Repine, T. B., Terada, L. S. and Repine, J. E. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10690–10694
- 20 Sekimoto, H., Seo, M., Dohmae, N., Takio, K., Kamiya, Y. and Koshihara, T. (1997) *J. Biol. Chem.* **272**, 15280–15285
- 21 Ori, N., Eshed, Y., Pinto, P., Parani, I., Zamir, D. and Fluhr, R. (1997) *J. Biol. Chem.* **272**, 1019–1025
- 22 Amaya, Y., Yamazaki, K., Sato, M., Noda, K., Nishino, T. and Nishino, T. (1990) *J. Biol. Chem.* **265**, 14170–14175
- 23 Berglund, R., Rasmussen, J. T., Anderson, M. D., Rasmussen, M. S. and Petersen, T. A. (1996) *J. Dairy Sci.* **79**, 198–204
- 24 Ichida, K., Amaya, Y., Noda, K., Minoshima, S., Hosoya, T., Sakai, O., Shimizu, N. and Nishino, T. (1993) *Gene* **133**, 279–284
- 25 Maniatis, T., Fritsch, E. F. and Sambrook, J. (1989) in *Molecular Cloning: A Laboratory Manual*, 2nd Edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- 26 Feinberg, A. P. and Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13
- 27 Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- 28 Wingender, E., Kel, A. E., Kel, O. V., Karas, H., Heinemeyer, T., Dietze, P., Knueppel, R., Romashenko, A. G. and Kolchanov, N. A. (1997) *Nucleic Acids Res.* **25**, 265–268
- 29 Terao, M., Kurosaki, S., Zanotta, S. and Garattini, E. (1997) *Biochem. Soc. Trans.* **25**, 791–796
- 30 Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K. and Green, M. R. (1984) *Nucleic Acids Res.* **12**, 7035–7056
- 31 Cazzaniga, G., Terao, M., Lo Schiavo, P., Galbiati, F., Segalla, F., Seldin, M. F. and Garattini, E. (1994) *Genomics* **23**, 390–402
- 32 Holmes, R. S. (1979) *Biochem. Genet.* **17**, 517–527
- 33 Yoshihara, S. and Tatsumi, K. (1997) *Arch. Biochem. Biophys.* **338**, 29–34
- 34 Romao, M. J., Archer, M., Moura, I., Moura, J. J. G., LeGall, J., Engh, R., Schneider, M., Hof, P. and Huber, R. (1995) *Science* **270**, 1170–1176
- 35 Glatigny, A. and Scazzocchio, C. (1995) *J. Biol. Chem.* **270**, 3534–3550
- 36 Carpani G., Racchi M., Ghezzi P., Terao M. and Garattini, E. (1990) *Arch. Biochem. Biophys.* **279**, 237–241
- 37 Terao, M., Cazzaniga, G., Ghezzi, P., Bianchi, M., Falciani, F., Perani, P. and Garattini, E. (1992) *Biochem. J.* **283**, 863–870
- 38 Rzhetsky, A., Ayala, F. J., Hsu, L. C., Chang, C. and Yoshida, A. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 6820–6825
- 39 Copeland, N. G., Gilbert, D. J., Schindler, C., Zhong, Z., Wen, Z., Darnell, Jr., J. E., Mui, A. L.-F., Miyajima, A., Quelle, F. W., Ihle, J. N. and Jenkins, N. A. (1995) *Genomics* **29**, 225–228
- 40 Keith, T. P., Riley, M. A., Kreitman, M., Lewontin, R. C., Curtis, D. and Chambers, G. (1987) *Genetics* **116**, 67–73
- 41 Riley, M. A. (1989) *Mol. Biol. Evol.* **6**, 33–52
- 42 Houde, M., Tiveron, M. C. and Bregegere, F. (1989) *Gene* **85**, 391–402
- 43 Xu, P., Huecksteadt, T. P. and Hoidal, J. R. (1996) *Genomics* **34**, 173–180
- 44 Joshi, M. S. and Rajagopalan, K. V. (1994) *Arch. Biochem. Biophys.* **308**, 331–334
- 45 Kozak, M. (1996) *Mamm. Genome* **7**, 563–574
- 46 Chow, C.-W., Clark, M., Rinaldo, J. and Chalkley, R. (1994) *Nucleic Acids Res.* **22**, 1846–1854