AmpliTaq and AmpliTaq Gold DNA Polymerase

The Most Referenced Brand of DNA Polymerase in the World

Date: **2005-05**

Notes: **Authors are listed alphabetically**

Gene (270)

Abe, M., K. Nata, et al. (2000). "Identification of a novel Reg family gene, Reg III[delta], and mapping of all three types of Reg family gene in a 75 kilobase mouse genomic region." Gene **246**(1-2): 111.
http://www.sciencedirect.com/science/article/B6T39-4019DM2-C/2/ef52f0cce86dc7252a55a1b46c319537

Adjaye, J., V. Bolton, et al. (1999). "Developmental expression of specific genes detected in high-quality cDNA libraries from single human preimplantation embryos." Gene **237**(2): 373.

http://www.sciencedirect.com/science/article/B6T39-3XD3S7F-9/2/3746e0b3da06dd1abc3764a254391645

Agata, Y., E. Matsuda, et al. (1998). "Rapid and efficient cloning of cDNAs encoding Kruppel-like zinc finger proteins by degenerate PCR." Gene **213**(1-2): 55.

http://www.sciencedirect.com/science/article/B6T39-3VGR20J-8/2/b37ec53646b50c5d06a8473649360686

To isolate cDNAs encoding Kruppel-like zinc finger proteins consisting of several hundred members, most of which are yet to be identified, from a limited number of available cells, we developed a rapid and efficient zinc finger gene cloning method based on reverse transcription-polymerase chain reaction (RT-PCR) using tagged, degenerate oligonucleotide primers corresponding to the conserved H/C link followed by the reverse blue selection to identify clones containing properly amplified fragments. More than 5 x 103 blue colonies were obtained from only 1 ng of total RNA. Eighty-eight out of 89 clones, which were randomly picked up from blue colonies and sequenced, encoded 60 different zinc fingers with the expected structure, and among them, only four have been previously described. Furthermore, it was possible to rapidly select clones that were differentially expressed in a tissue and stimulation-specific manner by a differential screening of the zinc-finger cDNA library using probes consisting of distinct sets of the zinc-finger PCR products. These results indicate that our PCR-based method is quite efficient and suitable for analyzing not only zinc finger genes but also other large gene families, especially when the available cells are very limited.

Akhmedov, N. B., N. I. Piriev, et al. (1997). "Structure and analysis of the transducin [beta]3-subunit gene, a candidate for inherited cone degeneration (cd) in the dog." Gene **194**(1): 47.

http://www.sciencedirect.com/science/article/B6T39-3V4BS2H-2R/2/f3cbb653c7c2bdce567fdbef0404c894

The cDNA for the [beta]3-subunit of cone-specific transducin (T[beta]3) was cloned and characterized from wild type dogs, and used in linkage studies as a candidate gene for cone degeneration. Sequence analysis of the T[beta]3 cDNA revealed an open reading frame of 1020 bp, potentially coding for a protein of 340 amino acids (aa). The deduced aa sequence of canine T[beta]3 shares 97% identity with the previously identified human T[beta]3, and 82% identity with bovine rod-specific transducin (T[beta]1). RT-PCR and sequencing of the amplified products demonstrated that the retinal canine T[beta]3 gene is expressed in two different transcripts which can be generated by alternative splicing of the intron in the 3'-untranslated region (UTR). The short and the long mRNAs differ in the length of their 3'-UTR by 456 nt. We have also determined the genomic organization of the canine T[beta]3 gene; it consists of ten exons and the first exon is in the 5'-UTR. The cDNA encoding T[beta]3 from cd-affected dogs was also cloned and sequenced. We found no differences at the nucleotide level between the cDNAs isolated from normal and diseased retinas. The level of transcription of T[beta]3 mRNA in the cd dog retina appeared to be normal. Linkage analysis of a crossbred informative pedigree showed five obligate recombinants out of nine informative off spring. These results suggest that T[beta]3 is not a candidate gene for the cone degeneration of the cd mutant.

Akio, M., W. Shoji, et al. (1994). "A novel isoform of the neurofibromatosis type-1 mRNA and a switch of isoforms during murine cell differentiation and proliferation." Gene **148**(2): 245.

http://www.sciencedirect.com/science/article/B6T39-47PNWHM-1W/2/23b1c8b090601438894747b4ce247262

Four types of cDNAs encoding the GTPase-activating protein-related domain (GRD) of the mouse neurofibromatosis type-1 gene (NF1) have been cloned. One of these isoforms was a newly identified form termed type IV. Analysis of the genomic structure of the mouse NF1-GRD revealed two exons (23A and 23B) between exons 23 and 24, leading to the production of four types of NF1-GRD cDNAs by an alternative splicing mechanism. Amino-acid sequences encoded by NF1-GRD are highly conserved between human and mouse. Analysis of the expression of these transcripts in various tissues of adult mouse revealed that the type-1 transcript is predominantly expressed in neural tissues such as brain and spinal cord. Other forms, termed types II, III and IV, are also expressed in various tissues. The type-1 and type-II transcripts are expressed equivalently in undifferentiated P19 mouse teratocarcinoma cells, whereas type-1 expression becomes predominant during neuronal differentiation by retinoic acid treatment. Expression of type I is also shown to be correlated with cessation of cell proliferation in P19 cells, but not in NIH3T3 cells. These, together with other results, suggest that the four types of NF1-GRD transcripts generated by alternative splicing have some important biological roles in cell differentiation and proliferation.

Andersson, T., S. Borang, et al. (2003). "Shotgun sequencing and microarray analysis of RDA transcripts." Gene **310**: 39.

http://www.sciencedirect.com/science/article/B6T39-48KFP56-1/2/768bfa64b247d76014db2d4a7aea8507

Monitoring of differential gene expression is an important step towards understanding of gene function. We describe a comparison of the representational difference analysis (RDA) subtraction process with corresponding microarray analysis. The subtraction steps are followed in a

quantitative manner using a shotgun cloning and sequencing procedure that includes over 1900 gene sequences. In parallel, the enriched transcripts are spotted onto microarrays facilitating large scale hybridization analysis of the representations and the difference products. We show by the shotgun procedure that there is a high diversity of gene fragments represented in the iterative RDA products (92-67% singletons) with a low number of shared sequences (<9%) between subsequent subtraction cycles. A non redundant set of 1141 RDA clones were immobilized on glass slides and the majority of these clones (97%) gave repeated good fluorescent signals in a subsequent hybridization of the labelled and amplified original cDNA. We observed only a low number of false positives (<2%) and a more than twofold differential expression for 32% (363) of the immobilized RDA clones. In conclusion, we show that by random sequencing of the difference products we obtained an accurate transcript profile of the individual steps and that large-scale confirmation of the obtained transcripts can be achieved by microarray analysis.

Angiolillo, A., G. Russo, et al. (2002). "The human homologue of the mouse Surf5 gene encodes multiple alternatively spliced transcripts." Gene **284**(1-2): 169.

http://www.sciencedirect.com/science/article/B6T39-44YF890-7/2/4cc5217b609439ff067ae69a10204dc7

Hu-Surf5 is included within the Surfeit locus, a cluster of six genes originally identified in mouse. In the present study, we have cloned and characterized the Hu-Surf5 gene and its mRNA multiple transcripts. Comparison of the most abundant cDNA and genomic sequence shows that the Hu-Surf5 is spread over a region of approximately 7.5 kb and consists of five exons separated by four introns. The nucleotide sequence of the genomic region flanking the 3'-end of the Hu-Surf5 gene revealed the presence of a processed pseudogene of human ribosomal protein L21 followed by Hu-Surf6 gene. Only 110 bp separate the transcription start site of Hu-Surf5 and Hu-Surf3/L7a gene and the transcription direction is divergent. Earlier studies defined the 110 bp region essential for promoter activity of Hu-Surf3/L7a. Here, we show that this region stimulates transcription with a slightly different efficiency in both directions. The bidirectional promoter lacks an identifiable TATA box and is characterized by a CpG island that extends through the first exon into the first intron of both genes. These features are characteristic of housekeeping genes and are consistent with the wide tissue distribution observed for Hu-Surf5 expression. Hu-Surf5 encodes three different transcripts, Surf-5a, Surf-5b, and Surf-5c, which result from alternative splicing. Two protein products, SURF-5A and SURF-5B have been characterized. Production of chimaeras between the full-length SURF-5A or SURF-5B and the green fluorescent protein (GFP) allowed to localize both proteins in the cytoplasm.

Barany, F. and D. H. Gelfand (1991). "Cloning, overexpression and nucleotide sequence of a thermostable DNA ligase-encoding gene." Gene **109**(1): 1.

http://www.sciencedirect.com/science/article/B6T39-47T1JJN-36/2/dcd93bc3906a9ee77ce75e67965fb71e

Thermostable DNA ligase has been harnessed for the detection of single-base genetic diseases using the ligase chain reaction [Barany, Proc. Natl. Acad. Sci. USA 88 (1991) 189-193]. The Thermus thermophilus (Tth) DNA ligase-encoding gene (ligT) was cloned in Escherichia coli by genetic complementation of a ligts7 defect in an E. coli host. Nucleotide sequence analysis of the gene revealed a single chain of 676 amino acid residues with 47% identity to the E. coli ligase. Under phoA promoter control, Tth ligase was overproduced to greater than 10% of E. coli cellular proteins. Adenylated and deadenylated forms of the purified enzyme were distinguished by apparent molecular weights of 81 kDa and 78 kDa, respectively, after separation via sodium dodecyl sulfate-polyacrylamide-gel electrophoresis.

Barcham, G. J., A. E. Andrews, et al. (1995). "Cloning and expression of a cDNA encoding ovine interleukin 7." Gene **154**(2): 265.

http://www.sciencedirect.com/science/article/B6T39-3Y5MNGT-2V/2/c711a4b871f86599651bb0508338c89e

Using the polymerase chain reaction (PCR) and primers based on regions of homology between the human and murine interleukin 7 (IL-7)-encoding cDNAs, we have amplified an ovine (ov) IL-7 cDNA from reverse-transcribed RNA extracted from concanavilin A (Con A)-activated ovine lymph-node cells. The nucleotide sequence of the cDNA and the predicted amino acid (aa) sequence showed significant homology to those of the human and murine molecules. The ovIL-7 cDNA encodes a 176-aa polypeptide that, based on analysis of murine IL-7, is processed to a protein of 151 aa. The cDNA was demonstrated to encode a protein with IL-7 biological activity. Supernatants from COS or CHO-K1 cells transfected with an expression vector containing the ovIL-7 cDNA were able to synergise with a suboptimal level of Con A to induce proliferation of ovine thymocytes. In addition, both supernatants were able to induce thymocyte proliferation, albeit at a reduced level, in the absence of Con A. Further experiments demonstrated that for induction of ovine thymocyte proliferation, recombinant (re)-ovIL-7 was able to synergise with re-human (h) IL-2 but not re-hIL-6 or tumor necrosis factor-[alpha] (re-hTNF[alpha]).

Barnes, W. M. (1992). "The fidelity of Taq polymerase catalyzing PCR is improved by an N-terminal deletion." Gene **112**(1): 29.

http://www.sciencedirect.com/science/article/B6T39-47T2TG2-1R/2/4bce72e1a879ba4095c931aac05ceab0

KlenTaq DNA polymerase is an N-terminally truncated Thermus aquaticus (Taq) DNA polymerase I. As expressed from a gene construct in Escherichia coli, translation initiates at Met236, bypassing the 5'->3' exonuclease domain of the DNA polymerase-encoding gene. A sensitive forward mutation assay was used to measure the relative number of mutations introduced into the entire lacZ gene by the polymerase chain reaction (PCR) under various conditions which allow the amplification of such a large DNA span. Two selectable markers, one at each end of the test lacZ fragment, were employed to avoid the plating and scoring of PCR artefacts such as primer initiation in the midst of the lacZ gene, and cloning artefacts such as empty vector plasmid. The measured relative mutation rate was twofold lower for KlenTaq as compared to the full-length Taq DNA polymerase.

Bartkiewicz, M., N. Hernando, et al. (1995). "Characterization of the osteoclast vacuolar H+-ATPase B-subunit." Gene **160**(2): 157.

http://www.sciencedirect.com/science/article/B6T39-3Y6HGWH-3C/2/97da7b07f2dc5b484b9e3fec6617491f

During bone resorption, osteoclasts acidify the extracellular bone resorbing compartment via a vacuolar H+-ATPase (V-ATPase), which resides in the ruffled-border membrane. In an effort to characterize the composition of the osteoclast V-ATPase catalytic domain, we have isolated a cDNA clone that encodes the V-ATPase B-subunit from a cDNA library constructed from highly purified chicken osteoclasts. Comparison of the predicted amino-acid sequence with the published sequences of isoforms of V-ATPase B-subunits from other sources revealed that the

chicken osteoclast B-subunit is brain type and not kidney type. Furthermore, only clones encoding the brain type isoform of subunit B could be generated by PCR from a cDNA library prepared from human osteoclastoma osteoclast-like cells. Northern blot analysis revealed that two B-subunit mRNAs, approx. 1.7 and 3.5 kb in length, are expressed in chicken bone marrow mononuclear cells, brain and kidney, although the relative amounts of these two transcripts were different in each tissue. In brain, the 3.5-kb mRNA was predominantly expressed. In bone marrow cells, the levels of the 1.7-kb mRNA were higher than in other tissues and expression of this message was increased by 1,25-dihydroxyvitamin -3, suggesting that this mRNA is specifically upregulated during osteoclast differentiation. These results indicate that the B-subunit isoforms present in the catalytic domains of the osteoclast and kidney V-H+-ATPases are different and further suggest that selective expression of isoforms of the B-subunit in these two tissues could provide a structural basis for some of the differences we have reported in the pharmacology and catalytic properties of these two enzymes.

Basnakian, A. G., A. B. Singh, et al. (2002). "Identification and expression of deoxyribonuclease (DNase) I alternative transcripts in the rat." Gene **289**(1-2): 87.

http://www.sciencedirect.com/science/article/B6T39-45CNFV2-1/2/883b408ef0205a8fd87ee2ead8c6391b

Deoxyribonuclease (DNase) I has been implicated in the induction of DNA fragmentation and cell death, however little is known about its regulation in vivo. In the present study, we describe that DNase I messenger RNA (mRNA) is alternatively spliced in rat kidney, and the activity of the DNase I correlates with the alternative splicing during the course of renal ischemia/reperfusion. Northern blot analysis with mRNA from control rat kidneys and kidneys subjected to ischemia/reperfusion in vivo yielded two bands of approximately 1.3 and 1.5 kb, suggesting the possibility of alternative splicing. However, prolonged reperfusion up to 16 h resulted in the predominant expression of 1.3 kb transcript. The disappearance of the 1.5 kDa band was associated with the increased DNase I activity in the kidney during ischemia/reperfusion. To study the alternative splicing of the DNase I mRNA, rat kidney cortex DNA complementary to RNA library was screened using rat DNase I probe. Twenty-one positive clones were obtained and were compared with the reported DNase I mRNA transcript cloned from rat parotid gland. All clones showed 100% homology with the reported DNase I coding sequence and part of 5'-untranslated region (5'-UTR), named exon 1a by us. Twelve out of 21 isolated clones had longer 5'-UTR than previously described, and DNase I pre-mRNA was alternatively spliced in this region. Six out of these 12 clones contained extra up to 153 bp in extreme 5' end, whereas, in six other clones, an internal 132 bp segment (exon 1b) of this additional sequence was absent, and only the extreme 5'UTR sequence (exon 1c) was found in these clones. The nucleotide analysis showed that alternating exon 1b has the possibility of a secondary structure with high internal homology and potential for at least one major stable stem-loop. Both newly identified segments, exons 1b and 1c, were also identified in genomic DNA. The long splice variant, containing exon 1b, is expressed only in the kidney among different tissues tested. Exon 1b inhibited translational activity of DNase I mRNA in vitro. Our data suggest that alternative splicing in 5'-UTR in the kidney may provide a prompt DNA-independent regulation of DNase I activity when DNA is damaged during ischemic injury.

Baysal, B. E., J. E. Farr, et al. (1998). "Genomic organization and precise physical location of protein phosphatase 2A regulatory subunit A beta isoform gene on chromosome band 11q23." Gene **217**(1-2): 107.

http://www.sciencedirect.com/science/article/B6T39-3V7WV9H-F/2/0831c88d386e5dc0e95d45b6f57e934a

Protein phosphatase 2A (PP2A) holoenzyme plays a critical role in cell-cycle control and growth-factor signaling, and is implicated in tumorigenesis. Because the protein phosphatase 2 regulatory subunit A beta isoform gene (PPP2R1B) maps within the critical region of hereditary paraganglioma (PGL1) on chromosomal band 11q23, we characterized its genomic structure and evaluated it as a candidate gene for PGL1. PPP2R1B has 15 exons spanning approx. 27 kb genomic distance. We placed the exons on genomic EcoRI fragments and identified their flanking intronic sequences. The gene was oriented from telomere to centromere. Splice acceptor and donor sites of all introns conformed to the GT/AG rule. Northern analysis with a cDNA probe identified 2.5 kb and 5.0 kb transcript sizes. We identified an ATG initiation codon in a favorable context and mapped two transcription start sites 15 bp and 66 bp upstream of it. We also mapped a 3'-polyadenylation site 504 bp downstream of the TGA stop codon, consistent with the 2.5 kb transcript size. We did not detect germ-line mutations by single-stranded conformational polymorphism (SSCP) analysis or major rearrangements by Southern analysis in a set of PGL1 patients. In conclusion, we precisely mapped and characterized the structure of PPP2R1B and evaluated it as a candidate gene for PGL1.

Bellizzi, D., M. A. Losso, et al. (2001). "A model for the involvement of Okazaki fragments maturation in the expansion of short tandem repeats." Gene **276**(1-2): 153.

http://www.sciencedirect.com/science/article/B6T39-4441RPS-M/2/0ecf7daadc2d6eb9dc8a662c2e6b056e

We propose a model for the expansion of short tandem repeats (ESTR), a phenomenon which has been found to occur in human DNA and is associated with a dozen of neuromuscular diseases. The model is based mainly on theoretical considerations and recovers experimental data from the literature; it also finds support in preliminary results obtained by us in multiprimed polymerase chain reactions designed to assess the effects of a downstream primer on the fidelity of the elongation of an upstream one. The model links the occurrence of the ESTR to a defective maturation of the Okazaki fragments (OF), and in particular to an improper processing of their 3' termini. This may occur when the last OF approaches the 5' terminus of the previous one in a susceptible region of the template. It is postulated here that when a growing OF has progressed past the priming region and its main portion has been synthesized, upon approaching its conclusion, the final elongation may take place in a region of the template where certain triplets are repeated: in that case a series of aberrations on the elongation mechanism may occur. These aberrations could involve (a) the displacement of the 5' terminus of the penultimate, properly matured OF, enacted by the incoming 3' terminus of the last OF, (b) the switch of the latter to the displaced strand of the former as template, (c) the fold-back on itself of the growing 3' terminus of the last OF, (d) its assumption of an unusual structure because of the repetition, and (e) some impairment of its removal by structure-specific exo-endonuclease(s). Derangements of this last part of the process may trigger the ESTR.

Benito, E. P., J. M. Diaz-Minguez, et al. (1992). "Cloning and sequence analysis of the Mucor circinelloides pyrG gene encoding orotidine-5'-monophosphate decarboxylase: use of pyrG for homologous transformation." Gene **116**(1): 59.

http://www.sciencedirect.com/science/article/B6T39-47V9HV7-14/2/d69d558222f4e09d03c59b5ea0e4a342

A 3.2-kb BamHI genomic DNA fragment containing the pyrG gene of Mucor circinelloides was isolated by heterologous hybridization using a pyrG cDNA clone of Phycomyces blakesleeanus as the probe. The complete nucleotide sequence of the M. circinelloides pyrG gene encoding orotidine-5'-monophosphate decarboxylase (OMPD) was determined and the transcription start

points (tsp) were mapped by primer extension analysis. The predicted amino acid sequence showed homology with the OMPD sequences reported from other filamentous fungi, with 96% similarity with the OMPD of P. blakesleeanus. Analysis of the sequence revealed the presence of two short introns whose length and location were confirmed by sequencing a cDNA clone and comparing this with its genomic counterpart. The intron splice sites and the 5'- and 3'-noncoding flanking regions show general features of fungal genes. Northern-blot hybridization revealed the pyrG transcript to be approx. 1.0 kb. The M. circinelloides pyrG cDNA clone was able to complement the pyrF:Mu-1 mutation of Escherichia coli when inserted between bacterial expression signals. Additionally, the genomic clone complemented the M. circinelloides pyrG4 mutation. When an M. circinelloides autonomous replication sequence was included in the transforming plasmid, the average transformation frequency obtained was 600 to 800 transformants per [mu]g DNA and per 106 viable protoplasts.

Benoit De Coignac, A., C. Bisbal, et al. (1998). "cDNA cloning and expression analysis of the murine ribonuclease L inhibitor." Gene **209**(1-2): 149.

http://www.sciencedirect.com/science/article/B6T39-3SRBHWR-N/2/a75a7860bd17bc65aac2c81915613e13

The 2-5A/RNase L system is one of the pathways induced by interferon (IFN). It plays a major role in the antiviral and antiproliferative activities of IFNs. Recently, we have shown that the activity of the RNase L could be inhibited by a proteic inhibitor, the RNase L Inhibitor (RLI). Human RLI (Hu-RLI) was cloned and characterized. We describe here the isolation and characterization of the cDNA encoding the murine RLI (Mu-RLI). Hu-RLI and Mu-RLI protein have 98% amino acid identity. Mu-RLI is functionally homologous to Hu-RLI, and all the structural features and amino acid sequence motifs of Hu-RLI are conserved in Mu-RLI. Moreover, reticulocyte lysate translated Mu-RLI protein is also able to inhibit 2-5A binding on 2-5A-dependent RNAse-L. Northern blot analysis revealed that Mu-RLI cDNA hybridizes with one mRNA of 3.5 kb except for the testis where two mRNA of 3.5 and 2.1 kb, respectively, are detected, suggesting a tissue-specific regulation.

Berchtold, M. W. and M. C. Berger (1991). "Isolation and analysis of a human cDNA highly homologous to the yeast gene encoding L17A ribosomal protein." Gene **102**(2): 283.

http://www.sciencedirect.com/science/article/B6T39-47PNXV0-PV/2/0223bbb53c03db0da991e3cda4ef34ea

A cDNA from human brain poly(A) + RNA with significant similarity to the gene encoding yeast L17A large subunit ribosomal (r) protein (L17A) was isolated using the polymerase chain reaction. The deduced amino acid (aa) sequence of 140 aa (calculated pI of 10.79) exhibits a 78% similarity to that of the yeast L17A r protein (88% when conservative aa replacements are considered as well). This indicates that L17A is one of the best conserved r-proteins and therefore may play a critical role in ribosome function. In contrast to its eubacterial and chloroplast counterparts, human L17A contains an N-terminal extension of 19 aa which may be involved in nuclear targeting of the r-protein. Approximately five to seven genes in mammalian genomes give strong hybridization signals when probed with the human L17A homologue cDNA. Whereas the LIT A homologue was found to be expressed at similar levels in several human tissues as a transcript of 600 nucleotides, a several-fold higher transcript level was detected in the rapidly growing neuroblastoma cell line, SK-N-BE.

Bergeron, H., D. Labbe, et al. (1998). "Cloning, sequence and expression of a linear plasmid-based and a chromosomal homolog of chloroacetaldehyde dehydrogenase-encoding genes in Xanthobacter autotrophicus GJ10." Gene **207**(1): 9.

http://www.sciencedirect.com/science/article/B6T39-3S1PY3W-2/2/54284a1e0ab23bda4055cf17a9644c1c

The degradation of 1,2-dichloroethane (DCE) by Xanthobacter autotrophicus GJ10 proceeds via chloroacetaldehyde (CAA), a toxic intermediate in the cells if it is not metabolized further by the NAD+-dependent CAA dehydrogenases. Here, we describe the cloning, sequence and expression in Escherichia coli of aldA, a plasmid-located CAA dehydrogenase-encoding gene of GJ10 as well as a chromosomal homolog, designated aldB. The DNA-predicted amino acid (aa) sequences of the two proteins (505 aa in AldA and 506 aa in AldB) are 84% identical. The cloned aldA and aldB genes were verified by their expression in the E. coli T7 polymerase/promoter and the pUC lac promoter systems. The expression level of AldA and its enzymatic activity towards CAA were both higher than those of AldB. In a hybrid construct, the 3'end of aldB was able to complement, although not completely, the corresponding portion of aldA to produce a functional gene. Both AldA and AldB proteins of GJ10 share the highest degree of sequence identity with an acetaldehyde dehydrogenase (ALDH) encoded by acoD of Alcaligenes eutrophus (77.3-78% identity). Together with at least three other ALDHs of prokaryotic origin, these proteins apparently form a special class of ALDHs whose expressions are dependent on RpoN factors. By pulsed-field gel electrophoresis the 225-kb pXAU1 plasmid encoding aldA was shown to be linear.

Berquin, I. M., L. Cao, et al. (1995). "Identification of two new exons and multiple transcription start points in the 5'-untranslated region of the human cathepsin-B-encoding gene." Gene **159**(2): 143.

http://www.sciencedirect.com/science/article/B6T39-3Y5FPYJ-2/2/e55e3250fdd68f7989c13c26e85f50fc

Transcripts for cysteine protease cathepsin B (CTSB) were found to be highly variable in the 5'-UTR (untranslated region). In cDNA clones from a human gastric adenocarcinoma cDNA library, we have identified two new exons (designated 2a and 2b) between exons 2 and 3 in the 5'-UTR of the gene. All of the exons of the 5'-UTR could be alternatively spliced to produce several transcript species. In addition, transcription was initiated from more than one promoter region. Using RT-PCR (reverse transcription-polymerase chain reaction) and primer extension assays, CTSB mRNA species were found to differ among tissues and between a glioblastoma sample and a cell line derived from it. Exons 2a and 2b were detected more frequently in tumor samples than in matched normal samples. Thus, factors related to the cell differentiation and environment seem likely to determine the types of transcripts that are expressed which in turn could influence the overall steady-state level of CTSB mRNAs and their rate of translation. Interestingly, at least three upstream translation initiation codons were observed and could constitute a means of controlling translation initiation.

Beyer, K. S., S. M. Klauck, et al. (2001). "Construction of a physical map of an autism susceptibility region in 7q32.3-q33." Gene **272**(1-2): 85.

http://www.sciencedirect.com/science/article/B6T39-43HT1F4-B/2/b158eb08f395045df02ebad8925c210f

The fast evolving progress of the human genome mapping and sequencing efforts facilitate the detection of genes also for complex traits. We focus on the detection of susceptibility loci for

autism, a prototypical pervasive developmental disorder. Five genome screens worldwide have identified several putative locations of susceptibility genes thus far, with the most common region on chromosome 7q. In order to identify new candidate genes for infantile autism we constructed a physical map of bacterial artificial chromosome, P1-derived artificial chromosome and yeast artificial chromosome clones of a 3 Mb region between D7S1575 and D7S500, including a complete contig of the ~1.2 Mb region around D7S2533, the marker with the most significant association result. We developed 16 novel sequence tag sites and mapped 23 genes/expressed sequence tags to the contigs. As this map contains a putative autistic disorder locus this integrated physical and transcript map provides a valuable resource for identification of candidate gene(s).

Bhat, K. S. (1993). "Generation of a plasmid vector for deletion cloning by rapid multiple site-directed mutagenesis." Gene **134**(1): 83.

http://www.sciencedirect.com/science/article/B6T39-47PH7G4-1NX/2/53188ea07d65237b15a35070153f0b51

The construction of a new plasmid vector, devoid of all MboI (GATC) and TspEI (AATT) restriction sites, is described. The lack of these two frequent-cutting restriction sites is a unique feature among plasmids. This new plasmid, pBRkanf1-, allows selective fragmentation of a cloned insert. As a result, the vector offers an alternative strategy to create overlapping and sequentially deleted subclones. In addition, the construction of the new plasmid required the development of a rapid and accurate multiple site-directed mutagenesis procedure. The mutagenesis method uses a combination of DNA amplification and chain extension by DNA polymerase. By this method, mutations are created progressively from one end of a DNA molecule to the other.

Bindra, P. S., R. Knowles, et al. (1993). "Conservation of the amino acid sequence of SV2, a transmembrane transporter in synaptic vesicles and endocrine cells." Gene **137**(2): 299.

http://www.sciencedirect.com/science/article/B6T39-47PH64N-13N/2/1557a3bff5af73d5c381a8e3c3b58523

SV2 is a secretory vesicle-specific protein produced by all neurons and by endocrine cells. The deduced amino acid (aa) sequence of this protein indicates that it is a transmembrane transporter [Bajjalieh et al., Science 257 (1992) 1211-1213; Feany et al., Cell 70 (1992) 861-867; Gingrich et al, FEBS Lett. 312 (1992) 115-122]. To determine the regions of the protein that are the most highly conserved throughout evolution, and might therefore be essential for the function of SV2, we isolated a cDNA clone encoding SV2 from the elasmobranch fish, Discopyge ommata, and compared the deduced aa sequence to two isoforms from rat, SV2A and SV2B [Bajjalieh et al., Proc. Natl. Acad. Sci. USA 90 (1993) 2150-2154]. The comparison indicates that although the N-terminal cytoplasmic domain of SV2 is the most divergent region, it contains a highly conserved sequence that is predicted to be the epitope for a monoclonal antibody that crossreacts with all species and two isoforms of SV2 [Buckley and Kelly J. Cell Biol. 100 (1985) 1284-1294; Bajjalieh et al., Proc. Natl. Acad. Sci. USA 90 (1993) 2150-2154]. The remainder of the protein is highly conserved: 62% of the aa in SV2 from D. ommata are identical to the rat SV2A sequence, and 12% are conservative substitutions. The high degree of conservation of this protein throughout evolution and across species indicates that it mediates a critical function of synaptic vesicles.

Blomstergren, A., A. Lundin, et al. (2004). "Comparative analysis of the complete cag pathogenicity island sequence in four Helicobacter pylori isolates." Gene **328**: 85.

http://www.sciencedirect.com/science/article/B6T39-4BK2JHH-5/2/ca78c51e54ec726d53a1a0feec077e42

The cytotoxin-associated gene (cag) pathogenicity island (PAI) is important for the virulence of Helicobacter pylori. In this study, we have compared the complete nucleotide sequence of the cag PAI in four clinical isolates. These isolates were selected from patients matched for age and sex from the same geographical region. The patients suffered from either gastric cancer (Ca52 and Ca73) or duodenal ulcer (Du23:2 and Du52:2). All four strains induced an interleukin (IL)-8 response in AGS cells and translocated CagA into host cells where the protein was tyrosine phosphorylated, and thus harboured a functional type IV secretion system encoded by the cag PAI. The cagA gene contains a variable region close to its 3' end. Different compositions of this region has been proposed to exert various degrees of morphological changes in cultured gastric epithelial cells, and there are indications that the structure of the repetitive region is connected to the severity of disease. One of the studied strains (Du23:2) possessed five Western-type repeat regions while the other three strains harboured one Western-type repeat. Strain Du23:2 also contained a major rearrangement or large insertion/duplication in the intergenic region between HP0546 and HP0547 (cagA). Sequence similar to that of genes HP0510 and HP0509 was found in the 5' end of this region. The 3' end was similar to the corresponding region of strain ATCC 43504, including a mini IS605 element and a duplication of the 3' end of the cag PAI. Finally, a novel gene was identified in the cag PAI in three of the sequenced strains at the position of HP0521. This gene, HP0521B, is present in approximately half of Swedish H. pylori isolates.

Bonello, J.-F., H.-G. Opsahl-Ferstad, et al. (2000). "Esr genes show different levels of expression in the same region of maize endosperm." Gene **246**(1-2): 219.

http://www.sciencedirect.com/science/article/B6T39-4019DM2-S/2/a7b9ca8d26cc05a862b73404b6ec2e7b

Bottner, M., M. Laaff, et al. (1999). "Characterization of the rat, mouse, and human genes of growth/differentiation factor-15/macrophage inhibiting cytokine-1 (GDF-15/MIC-1)." Gene **237**(1): 105.

http://www.sciencedirect.com/science/article/B6T39-3X944FT-D/2/282cce524b2964efa9703cfb2acec520

Brock, G. J. R., J. Charlton, et al. (1999). "Densely methylated sequences that are preferentially localized at telomere-proximal regions of human chromosomes." Gene **240**(2): 269.

http://www.sciencedirect.com/science/article/B6T39-3YXJ0JD-3/2/07b5d356cc11c08eadbdd48d9aa7901f

Bruvo, B., J. Pons, et al. (2003). "Evolution of low-copy number and major satellite DNA sequences coexisting in two Pimelia species-groups (Coleoptera)." Gene **312**: 85.

http://www.sciencedirect.com/science/article/B6T39-490H3SW-1/2/7c970c0ec47d30a5340758d19e3b808b

Satellite DNA sequence evolution has been studied in several insect species from the genus Pimelia (Tenebrionidae, Coleoptera). Low-copy number homologs of the previously characterized major satellite DNA from P. monticola (PMON) have been cloned and sequenced from six congeneric species belonging to two species groups: Ibero-Balearic and Moroccan. Sequence analysis of a sample of low-copy number repeats revealed two subfamilies, differing on average 17.5% due to randomly spread single point mutations. Each subfamily is specific for a group of taxa in congruence with their biogeography. Within each group, there is no significant species-specific clustering of the sequences. These results suggest that the two satellite subfamilies arose after the split of an ancestral lineage into the North African and Ibero-Balearic Pimelia species-groups, but before their subsequent radiation. Rate heterogeneity tests suggest that PMON sequences have evolved faster in the lineage leading to the Moroccan group. Comparison of sequence divergences between minor PMON and the previously characterized major PIM357 satellite obtained from the same taxa, points to similar evolutionary dynamics. Both sequences are evolving in parallel accumulating mutations in a gradual manner irrespectively of significant differences in abundance. These data show that copy number of the sequence families does not necessarily affect the sequence change dynamics of satellite repeats.

Brzozowska, A., H. Sundvold, et al. (2000). "Evolutionary conservation of the apolipoprotein E-C1-C2 gene cluster on bovine chromosome 18q24." Gene **241**(2): 241.

http://www.sciencedirect.com/science/article/B6T39-3Y6GVD3-6/2/d0f77a92fad7c8a5d6f7776f09531a73

Burris, P. A., Y. Zhang, et al. (1998). "The pore-forming and cytoplasmic domains of the neurogenic gene product, are conserved between Drosophila virilis and Drosophila melanogaster." Gene **206**(1): 69.

http://www.sciencedirect.com/science/article/B6T39-3S1PY3W-T/2/924351bce01047098f0599c4e8953c09

The Drosophila protein functions critically in the determination of neuroblasts in the embryonic ectoderm and many other cell types. is a member of the MIP family of transmembrane channel proteins. The conserved channel domain of is flanked by amino- and carboxy-terminal cytoplasmic domains of unique sequence, which comprise over two-thirds of the protein. To determine whether the cytoplasmic domains of are important for function, we have cloned and sequenced the bib gene of D. virilis and compared it with that of D. melanogaster. Here we report that the channel domain and both cytoplasmic domains are highly conserved between the two species. The conservation of the cytoplasmic domains indicates that they are critical to function. bib transcripts are found in similar patterns in both species, indicating that the developmental function(s) of have also been conserved.

Bustamante, V. H., J. Puente, et al. (1995). "Identification of Campylobacter jejuni and C. coli using the rpoB gene and a cryptic DNA fragment from C. jejuni." Gene **165**(1): 1.

http://www.sciencedirect.com/science/article/B6T39-3Y6HK4H-3K/2/ea6ecdb4579bf09b5995d128eb36bd65

Campylobacter jejuni (Cj) and C. coli (Cc) clinical isolates, obtained from three different sources, were characterized using two Cj DNA probes, CJ01 and CJ02. These probes were selected at random by virtue of their stability in Escherichia coli (Ec). CJ01 hybridized specifically with DNA from Cj reference strains, but not with DNA from Cc, C. lari (Cl) nor C. fetus (Cf) reference strains. Using clinical isolates characterized by genome-genome hybridization and biotype, CJ01 hybridized with DNA derived from all Cj strains. However, DNA from four out of ten Cc strains, from three different sources, also hybridized with CJ01, suggestive of this region being heterogeneous between clinical isolates of both species. The nucleotide sequence analysis of CJ01 reveals two incomplete open reading frames (ORFs) that did not show significant homology with any other known sequences. CJ02 hybridized specifically with DNA from Cj and Cc reference strains, but not with DNA from Cl and Cf reference strains. The specificity and sensitivity were maintained upon hybridization with DNA from 31 clinical isolates. CJ02 has an uninterrupted ORF whose deduced amino-acid sequence showed extensive homology with the central region of the Ec and Salmonella typhimurium (St) RNA polymerase [beta] subunits (52 and 66% similarity, respectively). The most conserved segments correspond to putative functional domains.

Cary, J. W., R. Brown, et al. (1995). "Cloning and characterization of a novel polygalacturonase-encoding gene from Aspergillus parasiticus." Gene **153**(1): 129.

http://www.sciencedirect.com/science/article/B6T39-3Y5MNMT-63/2/ce1e95fad93cd1f18487601178183c53

Pectinases produced by Aspergillus flavus and A. parasiticus are believed to play a significant role in the ability of these fungi to spread in cotton bolls and other crops. Utilizing a DNA probe, generated by PCR, of the Aspergillus niger pgaII gene, we have isolated a novel, constitutively expressed polygalacturonase (PG)-encoding gene (pecA) from an A. parasiticus cDNA library. DNA sequence analysis and the deduced amino acid (aa) sequence of pecA demonstrated significant identity at the nucleotide and aa levels with other PG of fungal origin. Northern blot analysis of RNA isolated from A. parasiticus grown on either glucose or pectin as the sole carbon source showed that pecA was expressed during growth in both media.

Chambers, C. and P. Russell (1993). "Sequence of the human lens [beta]B2-crystallin-encoding cDNA." Gene **133**(2): 295.

http://www.sciencedirect.com/science/article/B6T39-47PH8K0-24P/2/082da6877792c1ed9c58099c288b682a

This study reports the nucleotide (nt) sequence of the human [beta]B2-crystallin (h[beta]B2-Cry)-encoding cDNA (h[beta]B2-cry). The h[beta]B2-cry gene encodes a major structural protein in the lens of the vertebrate eye. Sequence information obtained from mouse genomic clones and the mouse [beta]B2-cry cDNA (m[beta]B2-cry) sequence was used to design oligodeoxyribonucleotide primers, corresponding to exon 1, transcription start points, and termination and polyadenylation sites, that were used in the polymerase chain reaction (PCR) to generate full-length cDNA clones from total human lens RNA. In addition, cDNA libraries were made with [lambda]gt11 from both human fetal and adult lens tissue RNAs, and screened with a m[beta]B2-cry cDNA clone. The full-length h[beta]B2-cry cDNA is 721 bp and contains an open reading frame of 612 nt. It codes for a 23-kDa protein of 205 amino acid residues. Comparison of the overall nt and deduced aa sequences shows a greater similarity of h[beta]B2-cry to bovine [beta]fB2-cry than to m[beta]B2-cry or rat [beta]B2-cry.

Chang, R., X. Xu, et al. (2002). "Molecular cloning, mapping and characterization of a novel mouse RING finger gene, Mrf1." Gene **291**(1-2): 241.

http://www.sciencedirect.com/science/article/B6T39-45NGR35-J/2/8af6e75a125e9bd779b97bd933776aa1

With a combined approach of database search, heterologous polymerase chain reaction (PCR), reverse transcription-PCR, rapid amplification of complementary DNA ends and genomic library screening, we have successfully cloned a mouse RING finger gene, mouse RING finger 1 (Mrf1). The Mrf1 gene has two exons of 63 and 2665 bp, respectively, and one intron of over 13 kb. An open reading frame was identified exclusively in exon 2, which encodes a putative protein of the RING-B box-coiled coil or the tripartite motif type of 403 amino acids. Mrf1 is moderately expressed in the spleen, brain and heart as a single 3.0 kb product and very highly expressed in the testis as two transcripts of 3.0 and 1.5 kb, respectively. The Mrf1 gene was mapped to mouse chromosome 3, between markers D3Mit70 and D3Mit277. Western blotting analysis indicated that an expected protein of approximately 44 kD was detected in the brain extracts of mouse, rat and human. The possible functions of Mrf1 are discussed in the contexts of protein-protein interactions, oncogenesis and ubiquitination.

Chave, K. J., T. J. Ryan, et al. (2003). "Identification of single nucleotide polymorphisms in the human [gamma]-glutamyl hydrolase gene and characterization of promoter polymorphisms." Gene **319**: 167.

http://www.sciencedirect.com/science/article/B6T39-49RCM38-7/2/78bfa84963bc6fe3a45fef0fdc4f3993

[gamma]-Glutamyl hydrolase (GGH) plays a central role in folate metabolism and antifolate action. Increased GGH activity has been found in rat hepatoma cells resistant to the cancer drug methotrexate (MTX). The aim of this study was to identify polymorphisms in the GGH gene that modulate GGH activity and that may affect methotrexate resistance. Exons of the human [gamma]-glutamyl hydrolase (hGGH) gene were amplified by polymerase chain reaction (PCR) from breast cancer tissue and leukemia cell lines. Single-stranded conformational polymorphism (SSCP) analysis was performed, and PCR products containing different patterns were cloned and sequenced. Six single nucleotide polymorphisms (SNPs) were identified, at bases -401C>T, -354G>T, -124T>G, +16T>C, +452C>T, and +1102A>G, relative to the A of the translation start codon being considered as +1. The SNP at +16, which changes codon -19 (relative to the start of the mature hGGH protein) in the endoplasmic reticulum targeting sequence of hGGH protein from cysteine to arginine, has previously been identified in this laboratory. The SNP at +452 changes the conserved hGGH protein codon 127 from threonine to isoleucine. The functions of SNPs in the promoter of the hGGH gene were studied by site-directed mutagenesis of a 516-bp region of the hGGH gene promoter in a luciferase reporter vector and transfection into HepG2 and MCF-7 cells. All of the promoter polymorphisms enhanced the production of luciferase compared to the wild-type hGGH gene promoter in HepG2 cells, and -401C>T and -124T>G enhanced luciferase expression in MCF-7 cells, suggesting that polymorphisms in the hGGH gene promoter may increase expression of hGGH protein.

Chavrier, P., K. Simons, et al. (1992). "The complexity of the Rab and Rho GTP-binding protein subfamilies revealed by a PCR cloning approach." Gene **112**(2): 261.

http://www.sciencedirect.com/science/article/B6T39-47T2THY-31/2/b02cc60d66a79c124ee1eaeb32c788ac

Partial sequences corresponding to eleven novel Rab proteins and one new Rho protein have been isolated using a PCR-based cloning approach. These results confirm that the overall diversity of the Rab and Rho protein subfamilies account for more than thirty different members in mammalian cells.

Chen, G.-Y., H. Muramatsu, et al. (2004). "ZEC, a zinc finger protein with novel binding specificity and transcription regulatory activity." Gene **340**(1): 71.

http://www.sciencedirect.com/science/article/B6T39-4CY0J8T-2/2/493423f5bb92a43c1efe3e0c805ffa36

A novel 114-kDa zinc finger protein, ZEC, has been found by cDNA cloning and characterized. ZEC was strongly expressed in the testis, liver and kidney, and also in embryonic stem cells. Epitope-tagged experiments indicated nuclear localization of ZEC. ZEC contained 18 C2H2 zinc fingers which were organized in two clusters. A ZEC binding DNA sequence, C/GA/TA/TGGTTGGTTGC, which we have designated the GT box, was identified by random oligonucleotide binding selection assay. The GT box did not contain binding sites for other previously characterized transcription factors and thus represented a potentially novel DNA target sequence. Electrophoretic mobility shift assay (EMSA) showed that both clusters of zinc fingers bound to the same DNA sequence. Site-directed mutagenesis revealed that the core sequence TTGGTT within the GT box was essential to ZEC binding, while DNA sequences outside of the core sequence enhanced this interaction. Furthermore, co-transfection assays demonstrated that ZEC could activate a reporter luciferase gene driven by this DNA sequence.

Chopra, A. K., A. R. Brasier, et al. (1994). "Improved synthesis of Salmonella typhimurium enterotoxin using gene fusion expression systems." Gene **144**(1): 81.

http://www.sciencedirect.com/science/article/B6T39-47P8KJR-1CP/2/840929109b1c535da59cf4acac4405f1

Salmonella enterotoxin (Stn) is a virulence factor in S. typhimurium strain Q1 that causes both fluid secretion in ligated intestinal loops of rabbits and elongation of Chinese hamster ovary (CHO) cells. High-level expression systems are needed to provide Stn in soluble form for detailed study of the biological activity of Stn. To maximize the synthesis and solubility of Stn, we systematically compared the production of native Stn synthesized with a T7 RNA polymerase/promoter system to that of two fusion proteins: glutathione S-transferase:Stn (Gst:Stn) and thioredoxin A:Stn (TrxA:Stn). The latter fusion protein expression systems resulted in a 64-fold increase in Gst:Stn and TrxA:Stn antigen concentration, as measured by specific anti-peptide antibodies in an enzyme-linked immunosorbent assay (ELISA). Most of the toxin derived using these vector systems was insoluble; however, the solubility of the TrxA:Stn antigen increased by at least 50-fold, with a concomitant increase in CHO cell elongation activity. In addition, stn gene expression was enhanced more than 50-fold by addition of 0.2-0.4 M NaCl to Luria-Bertani medium. The biological activity of Stn also was increased in the high-osmolarity medium. Consequently, the expression of stn may be regulated by DNA supercoiling.

Chuang, S.-E., V. Burland, et al. (1993). "Sequence analysis of four new heat-shock genes constituting

the hslTS/ibpAB and hslVU operons in Escherichia coli." Gene **134**(1): 1.

http://www.sciencedirect.com/science/article/B6T39-47PH7G4-1NJ/2/2751d14f42d7f4c978ae6066df43a836

Sequences of four new heat-shock (HS) genes of Escherichia coli organized into two operons were determined. The operon at 83 min specifies two proteins of 15.8 kDa (HslT) and 16.1 kDa (HslS), which are identical to IbpA and IbpB, respectively. Expression of mRNA from a [sigma]32-dependent promoter of the hslTS/ibpAB operon is stimulated 30-75-fold upon temperature upshift. The transcription start point (tsp) is located at a G, 96 bp upstream from the AUG start codon of hslT/ibpA. The deduced amino acid sequences of HslT/IbpA and HslS/IbpB are 48% identical to each other and were found to be remotely related to the chloroplast low-molecular-weight HS protein, which is highly conserved among plants. The second hs operon is much less actively stimulated by temperature upshift, although it has a hs promoter that perfectly matches the consensus of promoters recognized by [sigma]32. Located at 88.9 min, the hslVU operon specifies proteins of 19.1 kDa (HslV) and 49.6 kDa (HslU). Multiple tsp were found in this operon. HslV is remotely related to the eukaryotic proteasome proteins, and HslU is very similar to a Pasteurella haemolytica protein of unknown function. Both HslU and the P. haemolytica protein share a ATP/GTP-binding motif near their N-termini. The two operons described here are transcribed counterclockwise on the standard genetic map.

Cinquetti, R., F. Mazzotti, et al. (2003). "Influence of metal ions on gene expression of BALB 3T3 fibroblasts." Gene **318**: 83.

http://www.sciencedirect.com/science/article/B6T39-49M0W0P-2/2/5a70ace97faaaabadcffbd93b2f0c638

It is widely recognized that metal compounds may modify gene expression. In this context, we have searched for genes whose expression may be affected by cadmium and platinum ions within the context of a cell culture system. Cadmium is well known for its carcinogenic potential while platinum is destined to become more and more interesting because of its increasing use in the automotive industries. By applying differential display to cultures of mouse fibroblast, we have identified two transcripts (acute lymphoblastic leukemia-1, All-1, and a novel gene named metal-responsive gene, MERE-1) that were responsive to platinum and cadmium ions. Moreover, further experiments with a panel of metal compounds have shown that MERE-1 was strongly induced also by La(NO3)2 and Cr(NO3)3, and to a lesser extent, by Na2CrO4 and (NH4)2TeCl6.

Clark, T. G., T.-L. Lin, et al. (1999). "The gene for an abundant parasite coat protein predicts tandemly repetitive metal binding domains." Gene **229**(1-2): 91.

http://www.sciencedirect.com/science/article/B6T39-40TY7W0-C/2/feae569ea962be3a0920aa9447175799

Conte, I., M. Lestingi, et al. (2002). "Characterization of MPP4, a gene highly expressed in photoreceptor cells, and mutation analysis in retinitis pigmentosa." Gene **297**(1-2): 33.

http://www.sciencedirect.com/science/article/B6T39-46YJC6C-5/2/47e9476ab93c6aeea080c2b7a5ec8137

Membrane-associated guanylate kinase (MAGUK) proteins are cell-cell contact organizing molecules that mediate targeting, clustering and anchoring of proteins at synapses and other cell junctions. MAGUK proteins may contain multiple protein-protein interaction motifs including PDZ, SH3 and guanylate kinase (GuK) domains. In this study, we performed a detailed analysis of the expression pattern of MPP4, a recently described member of the MAGUK protein family. We confirmed that this gene is highly expressed in retina, and demonstrate that it is also present, at lower levels, in brain. We identified a new retina specific isoform encoding a predicted protein lacking 71 amino acids. This protein region contains a newly identified L27 domain, another module playing a role in protein-protein interaction. By RNA in situ hybridization, Mpp4 expression was found to be localized to photoreceptor cells in postnatal retina. The MPP4 gene is localized to chromosome 2, in band 2q31-33, where a locus for autosomal recessive retinitis pigmentosa (RP26) has been mapped. Mutation analysis of the entire open reading frame of the MPP4 gene in a RP26 family revealed no pathologic mutations. In addition, we did not identify mutations in a panel of 300 unrelated patients with retinitis pigmentosa.

Coutte, L., D. Monte, et al. (1999). "Genomic organization of the human e1af gene,a member of Ets transcription factors." Gene **240**(1): 201.

http://www.sciencedirect.com/science/article/B6T39-3YHWR69-R/2/e4898463d28e7a2e2a4b57f9a73d5114

Coy, J. F., S. Wiemann, et al. (2002). "Pore membrane and/or filament interacting like protein 1 (POMFIL1) is predominantly expressed in the nervous system and encodes different protein isoforms." Gene **290**(1-2): 73.

http://www.sciencedirect.com/science/article/B6T39-45M6JMW-2/2/f5056758ab3f88d22a1e1f2bea843a59

We have isolated and characterized a novel differentially spliced gene predominantly expressed in the nervous system, which encodes protein isoforms with significant homology to the [alpha]-actinin protein superfamily, the Caenorhabditis elegans UNC-53 protein and weak homology to the nuclear membrane protein POM121. Similar to POM121 the primary structures show a hydrophobic region that is likely to form one or more adjacent transmembrane segment(s). Indirect immunofluorescence with antibodies against a synthetic peptide gave staining of the nucleus. Target experiments with EGFP (enhanced green fluorescent protein)-fusion proteins confirmed the nuclear localization. Two further members of this gene family could be isolated. All three pore membrane and/or filament interacting like (POMFIL) genes are differentially expressed in neuronal tumor cell lines. In 40% of tested primary neuroblastomas expression of POMFIL1 is strongly reduced and after brain injury POMFIL1 protein expression is upregulated, indicating that POMFIL1 is involved in the process of neuron growth and regeneration, as well as in neural tumorigenesis.

Curach, N. C., V. S. J. Te'o, et al. (2004). "Isolation, characterization and expression of the hex1 gene from Trichoderma reesei." Gene **331**: 133.

http://www.sciencedirect.com/science/article/B6T39-4C1NGYS-1/2/7c9950446a64bb96e3821e2791824ca8

Polymers of the HEX1 protein produce Woronin bodies in filamentous fungi. We have isolated

and sequenced the hex1 gene and flanking regions from the industrially exploited fungus Trichoderma reesei. Multiple transcription start sites (TSS) and the 5' untranslated region (UTR) were identified by 5'RACE PCR. There are three hex1 transcript types, two of which originate from two TSSs at approximately -320 and -1335 from the start codon, which are separated by a 500-bp intron within the 5'UTR. The third transcript type results from alternative splicing of the intron within the coding sequence at the 3' end, which results in the inclusion or exclusion of an unconserved histidine-rich coding region. The three transcripts code for two forms of HEX1 protein. N-terminal sequencing of HEX1 separated by 2D gel electrophoresis confirms that there are two forms of HEX1 protein which are modified further by alternative cleavage of the N-terminus. The dominant form of HEX1 is coded by a cDNA with TSS at position -1335. Expression of hex1 on cellulase-inducing medium peaks strongly within 24 h of growth but the protein is expressed at a lower and more consistent level in medium containing glucose. This is the first investigation of expression of the hex1 gene encoding a protein unique to filamentous fungi.

Daibo, S., M. T. Kimura, et al. (2001). "Upregulation of genes belonging to the drosomycin family in diapausing adults of Drosophila triauraria." Gene **278**(1-2): 177.

http://www.sciencedirect.com/science/article/B6T39-44CMXX2-J/2/2b558dca9ac92b92bbc8315c39131193

Diapause-associated gene expression was studied in Drosophila triauraria using subtractive hybridization. Two genes that were shown to be upregulated in diapausing flies by Northern hybridization have similarity to genes encoding antifungal peptides of Drosophila melanogaster, members of the drosomycin family (drosomycin, CG10812, CG10813, CG10815 and CG11520). In addition, a signal peptide and Knot 1 domain are shared with them. The genes cloned from D. triauraria are tentatively named drosomycin-like. However, the similarities between drosomycin-like in D. triauraria and the members of the drosomycin family in D. melanogaster are quite lower than those between other homologous genes in these species. In addition, neighbor-joining analysis revealed that drosomycin-like in D. triauraria is not closely related to known members of the family in D. melanogaster. Thus, it is most plausible that drosomycin-like is not a D. triauraria counterpart of known members of the family, but a novel member belonging to the family. The drosomycin-like gene is expected to have a few copies, because at least two sequences having unique 3'-ends were obtained in RACE, and multiple bands were observed in Southern hybridization. However, these sequences from RACE had the same ORF. Probes for genes encoding additional antimicrobial peptides were used to evaluate expression during diapause. Like drosomycin-like, drosomycin was upregulated during diapause, but defensin and drosocin were not.

Davids, B. J., X.-J. Wu, et al. (1999). "Cloning of a [beta] integrin subunit cDNA from an embryonic cell line derived from the freshwater mollusc, Biomphalaria glabrata." Gene **228**(1-2): 213.

http://www.sciencedirect.com/science/article/B6T39-3W07P0S-S/2/426cc12ab2124c3b40629972335e69b9

Davidson, B. L., J. E. Brown, et al. (1993). "Synthesis of normal and variant human hypoxanthine-guanine phosphoribosyltransferase in Escherichia coli." Gene **123**(2): 271.

http://www.sciencedirect.com/science/article/B6T39-47PH7C3-

1MP/2/5ad87b0d143f82b166270294865c00ab

Naturally occurring mutations in hypoxanthine-guanine phosphoribosyltransferase (HPRT) have been identified by amino acid sequencing, cDNA cloning, and direct nucleotide sequencing of PCR-amplified transcripts. To determine the effect these mutations have on the catalytic properties of the molecule, knowledge of the three-dimensional structure of HPRT is required. A prerequisite for this, however, is the availability of a large amount of purified product for crystallization and x-ray diffraction analysis. For these reasons we have developed an effective means of producing high levels of human HPRT in Escherichia coli using the expression cassette PCR. By taking advantage of a T7 polymerase/promoter system, we have expressed both normal and variant human hprt sequences in E. coli. The proteins synthesized from these sequences are immunologically and enzymatically active, and are physically indistinguishable from the HPRT in B-lymphoblasts derived from normal and three HPRT-deficient subjects.

Davis-Taber, R., W. Choi, et al. (2000). "Molecular characterization of human SUR2-containing KATP channels." Gene **256**(1-2): 261.

http://www.sciencedirect.com/science/article/B6T39-41H3KVD-10/2/ab5ae7f0ae7d845e6e3f863905d4f198

De Antoni, A. and D. Gallwitz (2000). "A novel multi-purpose cassette for repeated integrative epitope tagging of genes in Saccharomyces cerevisiae." Gene **246**(1-2): 179.

http://www.sciencedirect.com/science/article/B6T39-4019DM2-M/2/94e988c05a529318d647406faf2ffc85

De Dominicis, A., F. Lotti, et al. (2000). "cDNA cloning and developmental expression of cellular nucleic acid-binding protein (CNBP) gene in Xenopus laevis." Gene **241**(1): 35.

http://www.sciencedirect.com/science/article/B6T39-3Y51FGV-5/2/69d5f5fa2e5decadb7534f42a06872e8

de la Rua-Domenech, R., M. Wiedmann, et al. (1996). "Equine motor neuron disease is not linked to Cu/Zn superoxide dismutase mutations: sequence analysis of the equine Cu/Zn superoxide dismutase cDNA." Gene **178**(1-2): 83.

http://www.sciencedirect.com/science/article/B6T39-3W2T6PW-G/2/a6272dc9fc5b64f10742a9951d4505f5

The cDNA encoding the equine copper/zinc superoxide dismutase (SOD1) was cloned from leukocyte total RNA from healthy horses and its nucleotide (nt) sequence was determined. We further sequenced the SODI gene from 16 horses diagnosed with equine motor neuron disease (EMND) and eight unrelated, clinically normal horses to determine if this disease, similar to amyotrophic lateral sclerosis (ALS) in humans, is linked to SODI mutations. The 465-bp SOD1 coding region in the horse encodes 153 amino acid (aa) residues. Equine SOD1 exhibited 81.8 and 79.9% sequence identity to the human homolog at the nt and aa levels, respectively, with only five distinct aa in the two loops that constitute the active site of the enzyme. None of the

human SODI mutations found in the familial form of ALS were detected in SODI of the 16 affected horses. Although DNA sequence analysis identified three potential polymorphisms in equine SOD1, these were silent and were found in both normal and EMND-afflicted horses. At this time, there is no conclusive evidence for EMND linkage to SOD1 mutations.

De Rossi, E., R. Leva, et al. (1995). "Cloning, sequencing and expression of the ilvBNC gene cluster from Streptomyces avermitilis." Gene **166**(1): 127.

http://www.sciencedirect.com/science/article/B6T39-3Y6HK1D-1T/2/ac867cfa6af185b131ecdfa26fe55ecd

The metabolism of the branched-chain amino acids (BCAA) isoleucine, leucine and valine is correlated to the production of polyketide antibiotics in many streptomycetes. Despite its significance, this biosynthetic pathway is poorly understood in Streptomyces. In order to develop a better understanding of Streptomyces BCCA biosynthesis, two genes, ilvBN and ilvC, encoding acetohydroxy acid synthase (AHS) and acetohydroxy acid isomeroreductase (IR), respectively, were cloned from Streptomyces avermitilis, a strain producing avermectins, potent antiparasitic compounds. The genes were isolated by applying a combination of PCR and genomic library screening. The deduced amino-acid sequences revealed significant homology to the AHS and IR proteins from other bacterial species. The ilvBN gene, expressed in Escherichia coli (Ec) by using the expression vector pGEX-4T-1, complemented the ilv- mutation of Ec PS1283. Ec transformants produced high levels of AHS, whose activity was feedback inhibited by valine.

DePriest, P. T. (1993). "Small subunit rDNA variation in a population of lichen fungi due to optional group-I introns." Gene **134**(1): 67.

http://www.sciencedirect.com/science/article/B6T39-47PH7G4-1NV/2/93dc9517c64ea3ac8529a29b225f5379

A natural population of the lichen-forming ascomycetous fungus, Cladonia chlorophaea, contained individuals with small subunit ribosomal DNA (SSU rDNA) of at least four different size classes and nine restriction-site patterns. The source of these differences was the variable occurrence of 200-400-nucleotide insertions, previously identified as small group-I introns, at five different positions within the SSU coding region. By specific amplification of the sequences flanking these five intron positions with the polymerase chain reaction (PCR), a minimum of nine types of rDNA repeats were defined that differ in number, position, restriction pattern and size of introns. The positions of the introns were verified by sequence analysis. The variable distribution of these introns suggests that they are currently mobile -- either by intron insertion, deletion or both -- within this species complex.

Deshpande, K. L. and J. R. Katze (2001). "Characterization of cDNA encoding the human tRNA-guanine transglycosylase (TGT) catalytic subunit." Gene **265**(1-2): 205.

http://www.sciencedirect.com/science/article/B6T39-42JHDKP-S/2/086ca9f6bca34f5c6ae01e3fea3432b3

Queuosine (Q) is a 7-deazaguanosine found in the first position of the anticodon of tRNAs that recognize NAU and NAC codons (Tyr, Asn, Asp and His). Eukaryotes synthesize Q by the base-for-base exchange of queuine (Q base) for guanine in the unmodified tRNA, a reaction catalyzed

by TGT. A search of the human EST database for sequences with significant homology to the well studied TGT from Escherichia coli identified several candidates for full-length (1.3-1.4 kb) cDNA clones. Three candidate cDNA clones, available from IMAGE Consortium, LLNL, (Lennon et al., 1996, Genomics 33, 151-152) were obtained: IMAGE Clone Id Nos. 611146, 1422928, and 72154. Here we report the complete sequences of these clones. IMAGE:72154 contains an ORF encoding a 44 kDa polypeptide with high homology to bacterial TGTs and was subcloned into the mammalian expression vector pMAMneo-Cat. When this construct was transfected into the TGT-negative cell line, GC3/c1 (Gunduz et al., 1992, Biochim. Biophys. Acta 1139, 229-238), it restored the ability of the cells to form Q-containing tRNA. This TGT cDNA sequence is encoded in human chromosome 19 clone CTC-539A10 (GenBank accession no. AC011475), enabling determination of the exon-intron boundaries for the TGT gene. The sequence of IMAGE:611146 is 5'-truncated by 76 bp compared to that from IMAGE:72154 and, except for two differences in the 3'-non-coding region, the remainder of the sequence is identical to that of IMAGE:72154. IMAGE:1422928 is a 1390 bp chimera: the 5'-portion, bp 1-708, is identical to a genomic DNA sequence from chromosome 15 (GenBank accession no. AC067805, bp 148976-149683); the 3'-end, bp 726-1390, is identical to the 3'-end of the TGT cDNA sequence from IMAGE:611146.

Dong, J., A. K. Kukula, et al. (2000). "Genomic organization and chromosome localization of the newly identified human heparanase gene." Gene **253**(2): 171.

http://www.sciencedirect.com/science/article/B6T39-40XNXHB-6/2/8556e5f24978dba2a7a111fe68d40d77

Dopf, J. and T. M. Horiagon (1996). "Deletion mapping of the Aequorea victoria green fluorescent protein." Gene **173**(1): 39.

http://www.sciencedirect.com/science/article/B6T39-3W0FCD6-7/2/f114b40e5310124ef56e99a1da8b7ee5

Aequorea victoria green fluorescent protein (GFP) is a promising fluorescent marker which is active in a diverse array of prokaryotic and eukaryotic organisms. A key feature underlying the versatility of GFP is its capacity to undergo heterocyclic chromophore formation by cyclization of a tripeptide present in its primary sequence and thereby acquiring fluorescent activity in a variety of intracellular environments. In order to define further the primary structure requirements for chromophore formation and fluorescence in GFP, a series of N- and C-terminal GFP deletion variant expression vectors were created using the polymerase chain reaction. Scanning spectrofluorometric analyses of crude soluble protein extracts derived from eleven GFP expression constructs revealed that amino acid (aa) residues 2-232, of a total of 238 aa in the native protein, were required for the characteristic emission and absorption spectra of native GFP. Heterocyclic chromophore formation was assayed by comparing the absorption spectrum of GFP deletion variants over the 300-500-nm range to the absorption spectra of full-length GFP and GFP deletion variants missing the chromophore substrate domain from the primary sequence. GFP deletion variants lacking fluorescent activity showed no evidence of heterocyclic ring structure formation when the soluble extracts of their bacterial expression hosts were studied at pH 7.9. These observations suggest that the primary structure requirements for the fluorescent activity of GFP are relatively extensive and are compatible with the view that much of the primary structure serves an autocatalytic function.

Dry, K., S. Kenwrick, et al. (2001). "The complete sequence of the human locus for NgCAM-related cell

adhesion molecule reveals a novel alternative exon in chick and man and conserved genomic organization for the L1 subfamily." Gene **273**(1): 115.

http://www.sciencedirect.com/science/article/B6T39-43KGNS7-F/2/a3f3a875b082b10f196f30cc59eb719c

NrCAM is one member of the L1 subfamily of cell surface recognition molecules implicated in nervous system development and function. Here we report the complete sequence of the human NRCAM locus. The gene comprises 34 exons and shows extensive conservation of exon/intron structure compared to L1, suggesting a common evolutionary ancestor. By human-chick sequence comparison we identified exons not previously found in mammalian NRCAM mRNAs. One of these encodes a premature stop codon that would give rise to an isoform of NRCAM lacking ankyrin-binding capacity. The availability of the complete sequence will allow an investigation of the potential role of these splice variants, and examination of the regulatory elements controlling NRCAM expression as well as the relationship of NRCAM to disorders involving 7q.

Duchange, N., J. Pidoux, et al. (2000). "Alternative splicing in the human interleukin enhancer binding factor 3 (ILF3) gene." Gene **261**(2): 345.

http://www.sciencedirect.com/science/article/B6T39-4292HH9-N/2/638ae1fd312dda3916730fd27199e4b2

The Interleukin Enhancer Binding Factor 3 (ILF3) gene has been mapped to chromosome 19 in humans and to chromosome 9 in mice. Several reported double-stranded RNA binding proteins including NF90, ILF3, MPP4 and DRBP76 have been suggested to be isoforms of the ILF3 gene but this has not been clearly established. We isolated several ilf3 transcripts from a melanoma cDNA library and two corresponding genomic fragments, and report alternative splicing and polyadenylation site selection in the human ILF3 gene. We show the existence of an alternative splice site responsible for the sequence divergence in the 3' part of the transcripts. Another alternative splicing event at a site between the two double-stranded RNA binding motifs leads to the additional presence in some cases of a four amino acids NVKQ peptide. We also describe the utilization of three distinct polyadenylation signals and the generation of an ilf3 transcript with a long extended 3' UTR. The expression of the different transcripts was evaluated. We used a GenBank sequence for the part of chromosome 19 corresponding to the ILF3 gene to determine the exon-intron organization of the entire gene which spans 38 kb and is divided into 21 exons.

Duim, B., P. Ruiter, et al. (1997). "Sequence variation in the hpd gene of nonencapsulated Haemophilus influenzae isolated from patients with chronic bronchitis." Gene **191**(1): 57.

http://www.sciencedirect.com/science/article/B6T39-3RM6VMT-9/2/d1e1fe60a32d7fbf75a13fec1faec061

The molecular diversity of protein D of nonencapsulated Haemophilus influenzae strains isolated from persistently infected patients with chronic bronchitis was studied by sequencing the hpd gene of four independently obtained isolates. The nucleotide (nt) sequences of the hpd genes of two strains were identical. The other two hpd sequences showed nt substitutions which were mostly synonymous. As a consequence the deduced amino acid (aa) sequences differed from the consensus sequence only by a few aa. No changes in the hpd genes were observed among the four variants of the four strains persisting in chronic bronchitis patients for 9, 11, 8 and 3 months, respectively, although variation in their major outer membrane proteins P2 and P5 occurred. We

conclude that the hpd gene is conserved during chronic infections of nonencapsulated H. influenzae.

Dumonceaux, T. J., K. A. Bartholomew, et al. (1998). "Cloning and sequencing of a gene encoding cellobiose dehydrogenase from Trametes versicolor." Gene **210**(2): 211.

http://www.sciencedirect.com/science/article/B6T39-3SRBHWR-1B/2/216a04cc45ec209a4cbce0c735171fee

Cellobiose dehydrogenase (CDH) is an enzyme produced under lignocellulose-degrading conditions by Trametes versicolor strain 52J (Tv) and several other wood-degrading fungi, including Phanerochaete chrysosporium (Pc). In order to understand better the nature and properties of this enzyme, we isolated a genomic clone of Tv cdh using heterologous probes derived from the sequence of Pc cdh. DNA sequence analysis revealed that Tv cdh consists of 3091 bp of coding sequence interrupted by 14 introns. Southern blotting showed that the gene was present in a single copy in the strain of Tv analyzed. Tv cdh was shown by Northern blot analysis to be expressed as a single transcript under cellulolytic conditions. RT-PCR of poly(A)+ RNA isolated under cellulolytic conditions was used to generate a near full-length cDNA copy of the cdh mRNA. The deduced protein encoded by Tv cdh consists of 768 amino acids (aa), including a predicted 19 aa signal peptide. The protein had 73% identity to the corresponding protein from Pc, which is the only other CDH-encoding gene that has been cloned. Based upon its deduced primary structure and alignment to similar sequences, Tv CDH shares a general structural organization with Pc CDH and other hemoflavoenzymes. Amino acid residues H-109 and M-61 in the N-terminal heme domain are hypothesized to function in heme binding; the C-terminal flavin domain contained a consensus sequence for flavin binding between residues 217-222. Although the protein is known to bind to cellulose, no obvious homology to bacterial or fungal cellulose binding domains was observed. However, a strong homology was observed to a region of Pc CDH that is hypothesized to be involved in cellulose binding.

El Haj, A. J., S. L. Tamone, et al. (1997). "An ecdysteroid-responsive gene in a lobster - a potential crustacean member of the steroid hormone receptor superfamily." Gene **201**(1-2): 127.

http://www.sciencedirect.com/science/article/B6T39-3RD1R45-3T/2/6361257328d462a07f05df0c6aea5b87

The role of ecdysteroids in modulating exoskeletal growth during the moult cycle of Crustacea has been well described. However, little is known about the action of ecdysteroids at the level of gene transcription and regulation in Crustacea. This paper reports the cloning of an ecdysteroid responsive gene, HHR3, a potential Manduca sexta MHR3 homologue in the American lobster, Homarus americanus. Levels of HHR3 expression are up-regulated in response to in vivo injections of premoult concentrations (10-6 M) of 20-hydroxyecdysone in the epidermal and muscle tissue of the lobster after 6 h. Maximal mRNA levels are observed after 21 h before returning to basal levels. In muscle tissue, elevated levels of HHR3 mRNA follow a time course similar to elevated actin mRNA expression in response to hormonal injection. In contrast, in eyestalk tissue, the HHR3 levels decline up to 21 h post-injection before rising to basal levels after 48 h. Eyestalk, epidermal and leg muscle tissue was extracted over the moult cycle to determine the levels of expression. In muscle, HHR3 is high during the premoult period that corresponds to the period of the moult cycle when the ecdysteroid titre is high. In the epidermis, HHR3 levels are also high during the premoult with elevated levels maintained into the postmoult period. In the eyestalk, mRNA levels of HHR3 show an opposite pattern of expression with low levels during premoult and postmoult and high levels found during the intermoult period. Our results provide novel evidence for an ecdysteroid responsive gene in a crustacean that has many

similarities to MHR3 in Manduca and DHR3 in Drosophila melanogaster. This raises the question of whether a similar cascade of ecdysteroid responsive genes exist in other members of Arthropoda such as the Crustacea, as has been demonstrated in Drosophila. In addition, we provide further evidence for negative feedback regulation of ecdysteroids at the site of moult-inhibiting hormone (MIH) production in the lobster eyestalk.

Esaki, T., K. Roy, et al. (1998). "Cloning of mouse [gamma]-glutamyl hydrolase in the form of two cDNA variants with different 5' ends and encoding alternate leader peptide sequences." Gene **219**(1-2): 37.

http://www.sciencedirect.com/science/article/B6T39-3WBG1Y1-R/2/96924a772ebf72c7fb7a1a606ed4694a

Mouse-liver [gamma]-glutamyl hydrolase (GH) is a lysosomal endopeptidase with an acid pH optimum that is activated by sulfhydryl compounds and preferentially hydrolyzes the most proximal [gamma]-glutamyl linkage of longer chain polyglutamates of folates and their analogues. We describe the cloning of this mouse lysosomal cDNA enzyme from liver GH mRNA in the form of two cDNA variants (1.295 and 1.268 kb in length) differing 14-fold (Variant I versus Variant II) in relative frequency that exhibited 5'-end heterogeneity and encoded alternate leader peptides. The 5' UTR in these variants also differs in length by 27 nucleotides. Otherwise, the ORF and 3' UTR in each case are the same. These cDNAs encode a protein in which the deduced amino acid sequence shares 78.9 and 69.1% identity to rat and human GH sequences, respectively. Amino acid sequence comparisons among the three species identified three conserved Asn sites and two conserved Cys residues that may be sites of glycosylation and sulfhydryl compound activation, respectively. Variant I GH mRNA was more abundant than Variant II GH mRNA in all mouse tissues examined. Variant I GH mRNA levels were extremely high in salivary gland, moderately high in kidney, liver, lung, stomach and uterus, low in small intestine, brain and fetal liver and relatively rare in thymus, spleen and skeletal muscle. Abundance of GH mRNA among tumors varied from low to high, with no discernible correlation with their tissue of origin.

Falaschi, A., M. Giacca, et al. (1993). "Searching for replication origins in mammalian DNA." Gene **135**(1-2): 125.

http://www.sciencedirect.com/science/article/B6T39-47PH672-14V/2/30b6de2ba290dfe9e726e898585b609a

The attempts at identifying precise replication origins (ori) in mammalian DNA have been pursued mainly through physico-chemical and biochemical approaches, in view of the essential failure of the search for autonomously replicating sequences in cultured cells. These approaches involve the mapping of short stretches of nascent DNA, the identification of the regions where either leading or lagging strands switch polarity, or the localization of replication intermediates by two-dimensional gel electrophoresis. Due to the complexity of animal cell genomes, most of these studies have been performed on amplified domains and with the use of synchronization procedures. The results obtained have been controversial. In order to avoid the use of experimental procedures potentially affecting the physiological mechanism of DNA replication, we have developed a method for the localization of ori in single-copy loci in exponentially growing cells. This method entails the absolute quantification of the abundance of selected DNA fragments along a genomic region within samples of newly synthesized DNA by competitive polymerase chain reaction (PCR); the latter is immune to all the uncontrollable variables which severely affect the reproducibility of conventional PCR. The application of this method to SV40 ori-driven plasmid replication precisely identifies the known ori localization. Using the same approach, we have mapped an ori for bi-directional DNA replication in a 13.7-kb locus of human

chromosome 19 encoding lamin B2.

Figueiredo, C., W. G. V. Quint, et al. (2000). "Genetic organization and heterogeneity of the iceA locus of Helicobacter pylori." Gene **246**(1-2): 59.

http://www.sciencedirect.com/science/article/B6T39-4019DM2-6/2/3a60fce44bc86600d08aae94c7508ba3

Fischer, C., L. Bouneau, et al. (2004). "Global heterochromatic colocalization of transposable elements with minisatellites in the compact genome of the pufferfish Tetraodon nigroviridis." Gene **336**(2): 175.

http://www.sciencedirect.com/science/article/B6T39-4CNJD7P-1/2/f5e8fa800aab3b441ccb3a569bd47c14

Because of its unusual high degree of compaction and paucity of repetitive sequences, the genome of the smooth pufferfish Tetraodon nigroviridis is the subject of a well-advanced sequencing project. An astonishing diversity of transposable elements not found in the human and the mouse has been observed in the genome of T. nigroviridis. Due to the difficulty of assembling repeat-rich regions, the whole genome shotgun sequencing approach will probably fail to reveal the general organisation of this compact vertebrate genome. Therefore, in order to gain new insights into the global distribution pattern of repeated DNA in the genome of T. nigroviridis, we have reconstructed partial/complete repetitive sequences from data generated by the genome project and performed double-colour fluorescent in situ hybridization (FISH) analysis for representatives of three major categories of repeated sequences including two minisatellites (ms100 and ms104), two DNA transposons (Tol2 and Buffy1) and two non-long terminal repeat (LTR) retrotransposons (Rex3 and Babar). We show that DNA transposons and retroelements very frequently colocalize with minisatellites and mostly accumulate within heterochromatic regions. These results, which have not been reported so far for the fugu Takifugu rubripes, show that repeated elements are generally excluded from gene-rich regions in T. nigroviridis and underline the extreme degree of compartmentalization of this compact genome. The genome organization of the pufferfish is clearly different from that observed in humans, where repeated sequences make up an important fraction of euchromatic DNA, and is more similar to that observed in the fruit fly Drosophila melanogaster.

Fisher, K. L. and J. P. Woods (2000). "Determination of [beta]-glucosidase enzymatic function of the Histoplasma capsulatum H antigen using a native expression system." Gene **247**(1-2): 191.

http://www.sciencedirect.com/science/article/B6T39-402KBCD-P/2/b88fc71c56e3700a684412ce292ebf65

Franchi, M. and E. Gallori (2005). "A surface-mediated origin of the RNA world: biogenic activities of clay-adsorbed RNA molecules." Gene **346**: 205.

http://www.sciencedirect.com/science/article/B6T39-4FC449B-5/2/08e436b944deb163367b2c51b9c7613e

The involvement of clay surfaces in the origin of the first genetic molecules on Earth has long been suggested. However, the formation of these polymers was not sufficient by itself to initiate the evolutionary process leading to the appearance of life. These macromolecules had to persist in primeval habitats so that their biological potentiality could be expressed.In this study, we assess the possibility of development of the RNA world on a clay substrate by investigating the capacity of different RNA molecules adsorbed/bound on the clay minerals montmorillonite (M) and kaolinite (K) to persist in the presence of a degrading agent (RNase-A), to interact specifically with complementary RNA strands, and to transmit the information contained in their nucleotide sequences.The RNase-A degradation of clay-adsorbed 23S rRNA from Escherichia coli was significantly slower (75-80%) than that observed for free rRNA, and the complete digestion of nucleic acid in the presence of clay was obtained in 2 vs. 1 h.Clay-adsorbed Poly[A] homopolymer was able to recognize the complementary Poly[U] homopolymer present in the surrounding water solution and to establish a specific interaction (association) with it, possibly leading to the formation of double strands.Reverse transcription and amplification (RT-PCR) amplification of free and clay-adsorbed 16S indicated that the presence of clay particles partially reduced the efficiency and processivity of reverse transcriptase but did not inhibit its activity, demonstrating that clay-adsorbed RNA is still available for enzymatic replication.These findings indicate that primordial genetic molecules adsorbed on clay minerals would have been protected against degrading agents present in the environment and would have been in the right conditions to undergo evolutionary processes.

Freemantle, S. J., H. B. Portland, et al. (2002). "Characterization and tissue-specific expression of human GSK-3-binding proteins FRAT1 and FRAT2." Gene **291**(1-2): 17.

http://www.sciencedirect.com/science/article/B6T39-45NGR35-C/2/b72096117aed1a12615bbcece5bd460d

We have isolated the entire coding sequence of human FRAT2 (frequently rearranged in advanced T-cell lymphomas-2). It exhibits appreciable amino acid identity to FRAT1 (77%) which was initially isolated as frequently being overexpressed in a murine leukemia virus insertion model in murine tumors. FRAT proteins are thought to play a role in Wnt signaling. They can bind to glycogen synthase kinase-3 (GSK-3) and Dishevelled, two proteins involved in Wnt signal transduction. Both hFRAT1 and hFRAT2 are intronless genes localized to the same portion of chromosome 10q24.1 and separated by only 10.7 kb. In a broad range of human tissues FRAT1 and FRAT2 are readily detected and expressed in a near identical pattern. Both species are repressed when the human embryonal carcinoma cell line, NT2/D1, is induced to differentiate with all-trans retinoic acid (RA). This treatment had no appreciable effect on FRAT levels in two other RA-sensitive cell lines that were not of germ cell tumor origin. The overlapping expression patterns suggest these two genes share a regulatory region. Both FRAT genes exhibited three species of mRNA, which varied in representation between tissues. When transiently overexpressed in COS-1 cells, the FRAT proteins were detected in the cytosol and concentrated in the nucleus. Both hFRAT1 and hFRAT2 are implicated in the selective modulation of GSK-3 activity via the Wnt signaling pathway. This study provides a foundation from which to examine the role these proteins play in Wnt-dependent and -independent processes.

Fu, Y., N. Comella, et al. (1999). "Cloning of DLM-1, a novel gene that is up-regulated in activated macrophages, using RNA differential display." Gene **240**(1): 157.

http://www.sciencedirect.com/science/article/B6T39-3YHWR69-J/2/bbcccba853577491d2e5adc809583215

Fuchs, B., K. Zhang, et al. (2000). "Differential mRNA fingerprinting by preferential amplification of coding sequences." Gene **258**(1-2): 155.

http://www.sciencedirect.com/science/article/B6T39-41TMP6Y-K/2/2b1d2a9014507def73dfa0707d881124

Fuchs, B., K. Zhang, et al. (2001). "Identification of twenty-two candidate markers for human osteogenic sarcoma." Gene **278**(1-2): 245.

http://www.sciencedirect.com/science/article/B6T39-44767RX-1/2/b27d7c2a81018dba65327556033a0a76

Since osteogenic sarcoma (OGS) predominantly affects children, its etiology and progression may be determined more by genetic than environmental factors. A few genes have been associated with OGS, however, their value in the diagnosis and/or prognosis of the disease remains poor. Evidently, more markers need to be identified for improving management of patients with OGS. To identify potential genetic markers for OGS, we have extended preferential amplification of coding sequences (PACS) to screen multiple samples simultaneously. The extended method is termed multi-PACS. Multi-PACS was applied between a normal osteoblast and four OGS-derived cell lines to identify differentially expressed coding sequence tags (dCST) that identified 145 dCSTs. Subsequently, differential mRNA expression was validated for a chosen subset of 22 dCSTs. These chosen dCSTs include among others cyclins D and E, two cyclin dependent kinases, two other kinases, transcription factors E2F4, E2F5, and p130, a DNA repair gene, a gene for the signalosome subunit, and potential guanine nucleotide binding factors. We infer that these genes could be so easily identified because PACS preferentially identifies coding instead of non-coding sequences. We also infer that these genes identify signaling pathways pertinent to OGS. mRNA expression profile of these 22 genes/dCSTs generated distinct expression signature of the OGS-derived cell lines suggesting that further work on clinical samples with these dCSTs will yield valuable information for OGS. We conclude that these 22 genes/dCSTs are candidate markers for OGS.

Fulop, C., R. V. Kamath, et al. (1997). "Coding sequence, exon-intron structure and chromosomal localization of murine TNF-stimulated gene 6 that is specifically expressed by expanding cumulus cell-oocyte complexes." Gene **202**(1-2): 95.

http://www.sciencedirect.com/science/article/B6T39-3TVNRN6-G/2/623cda9603b5014f8c455b25d9355aa3

Tumor necrosis factor stimulated gene-6 (TSG-6) has been previously shown to be induced in vitro in several cell types by proinflammatory cytokines, and in vivo in pathological conditions such as rheumatoid arthritis. In this study, we report the complete coding sequence for the mouse TSG-6 protein, and the exon-intron structure and the chromosomal localization of the gene. We have identified a 1605 nt cDNA sequence from mouse cumulus cell-oocyte complexes (COCs) induced to expand in vivo. The sequence contains an open reading frame of 825 nt that codes for the 275 amino acid TSG-6 protein. The gene contains six exons separated by 1.1-5.8 kb introns and has been localized to the murine chromosome 2 by linkage analysis. Comparative reverse transcription-polymerase chain reaction studies have revealed that TSG-6 mRNA is specifically expressed after COC expansion induced in vivo, identifying the first non-pathological process in which TSG-6 may play an important role. Since TSG-6 binds to hyaluronan and interacts with inter-[alpha]-trypsin inhibitor (I[alpha]I), molecules that are essential for matrix formation by COCs, this protein may have a structural role in the matrix or may enhance the antiproteolytic

effect of I[alpha]I to protect the matrix from degradation.

Funakoshi, E., K.-y. Nakagawa, et al. (2005). "Molecular cloning and characterization of gene for Golgi-localized syntaphilin-related protein on human chromosome 8q23." Gene **344**: 259.

http://www.sciencedirect.com/science/article/B6T39-4F0GR5F-3/2/eac48deaffbd41945402580780a1d27b

Loci for several human genetic diseases including glaucoma have been mapped to q23 region on chromosome 8. We carried out homology search analysis of the genomic sequence of a bacterial artificial chromosome (BAC) clone, KB1590E11, on 8q23 region, and mapped a previously described cDNA, KIAA1472, to this BAC clone. In this study, we determined the complete genomic structure of the KIAA1472 gene and its expression in various tissues and cell lines. Four mRNA species (types 1a, 1b, 1c, and 2) were produced from this gene by use of alternative transcription start sites and alternative-splicing events. These mRNAs were expressed in various tissues, except for type 1a, which was found only in the brain. Further, type 1 mRNA could be translated into two protein isoforms with different N-terminal sequences; and type 2 mRNA, into another type of isoform. All three of these KIAA1472 gene products were localized in Golgi apparatus and contained a C-terminal hydrophobic segment characteristic of a transmembrane domain, thus indicating them to be Golgi membrane-bound proteins. Furthermore, these proteins were homologous to syntaphilin, a molecule involved in guiding vesicular transport. These results indicate that KIAA1472 gene products may play an important role in vesicular traffic in various tissues including the brain.

Futami, K., T. Komiya, et al. (2000). "Determination of heterogeneous transcription start points of two c-myc genes from the common carp (Cyprinus carpio)." Gene **245**(1): 43.

http://www.sciencedirect.com/science/article/B6T39-3YS34MW-5/2/9b74818782fd75c9df13fb653e1351c1

Futami, K., T. Komiya, et al. (2001). "Differential expression of max and two types of c-myc genes in a tetraploid fish, the common carp (Cyprinus carpio)." Gene **269**(1-2): 113.

http://www.sciencedirect.com/science/article/B6T39-433NSPG-D/2/b0404bd0eca08cdc5ddced60c60956e1

We cloned the full-length cDNA of max gene from the common carp (Cyprinus carpio). The cDNA clone of carp max consists of 1209 bp and contained an ATG-initiated ORF consisting of 156 aa. The carp MAX share 76.7-93.8% aa identity with those of human, mouse, rat, chicken, Xenopus and zebrafish, respectively. The 15 bp alternative splicing was observed in the loop region of helix-loop-helix and is not previously described in mammalian max sequences. Transcripts of max gene were observed in all of the tissues of carp investigated in this study. The highest expression was found in the ovary, and the transcripts in hepatopancreas and heart were low. Two carp c-myc genes (CAM1 and CAM2) showed differential expression pattern. The expression of max was concomitant with CAM2 expression, but not with CAM1. It has been reported that MYC/MAX heterodimer as a regulator of gene expression has been maintained throughout vertebrate evolution, and the expression of c-myc has been concomitant with max expression. In addition, according to phylogenetic analysis, CAM1 is evolving faster than CAM2 after gene duplication. Therefore, this result suggests that CAM1 may evolve to obtain a new

function different from c-myc.

Gamel, P. H. and J.-C. Piot (1992). "Characterization and properties of a novel plasmid vector for Bacillus thuringiensis displaying compatibility with host plasmids." Gene **120**(1): 17.

http://www.sciencedirect.com/science/article/B6T39-47P8JW4-14R/2/6bf2be7df3d24f39cb2e2edfc437b5cf

A novel plasmid vector, composed of a 1.7-kb Bacillus thuringiensis (B.t.) replicon, a multiple cloning site, and an erythromycin-resistance marker gene from Bacillus subtilis, was constructed for use in B.t. Unlike other vectors which have been reported to be acceptable for B.t., this new B.t. vector was stably maintained in the absence of Er and did not displace host plasmids, some of which carry crystal protein-encoding genes (cry genes). The compatibility of this B.t. vector with native plasmids is highly desirable when introducing new cry genes into a wild-type B.t. strain. When a cryIIIA gene of B.t. tenebrionis was cloned m this vector and introduced into B.t. kurstaki (kur) HD119, cryIIIA was highly expressed without affecting the level of expression of native cry genes. The stability of this vector and its compatibility with native B.t. plasmids were achieved by subcloning only nucleotide sequences required for the vector to replicate in B.t. The origin of replication was first cloned on a 9.6-kb BglII fragment from a 75-kb plasmid of B.t. kur HD73 and then localized to a 2.4-kb region within the 9.6-kb fragment. Sequencing of the 2.4-kb region revealed the presence of an open reading frame (ORF), encoding a putative 312-amino acid (aa) protein. The deduced aa sequence of the ORF showed no homology to any published aa sequences. Deletion analysis indicated that the B.t. vector required at least the ORF and up to 300 bp surrounding the ORF, in order to replicate.

Garcia, C. A., A. Ahmadian, et al. (2000). "Mutation detection by pyrosequencing: sequencing of exons 5-8 of the p53 tumor suppressor gene." Gene **253**(2): 249.

http://www.sciencedirect.com/science/article/B6T39-40XNXHB-H/2/51c323c80d584d639043dd2d0d0362de

Garcia, J. and J.-L. Castrillo (2005). "Identification of two novel human genes, DIPLA1 and DIPAS, expressed in placenta tissue." Gene **344**: 241.

http://www.sciencedirect.com/science/article/B6T39-4DW90R3-2/2/5d62df8ed8a21aba9c65b1b51177e2c6

Here we report the identification and expression analysis of two novel human genes--DIPLA1 (Differentially expressed in placenta 1) and DIPAS (DIPLA1 Antisense). These genes are located at chromosomal region 9q33.1, in opposite orientations, and are flanked by the pregnancy-associated plasma protein-A (PAPP-A) and astrotactin 2 (ASTN2) genes. The mRNA sequences of both genes contain several upstream AUGs (uAUG) and various potential open reading frames (ORFs). DIPLA1 mRNA is 1.8 kb long and contains a 285 nt ORF coding for a polypeptide designated as replicative senescence up-regulated (RSU) protein. Antisense DIPAS mRNA is 2.7 kb long and contains a 309 nt ORF coding for a protein with partial similitude to the [gamma] isoform variant of the human Ca2+/calmodulin (CaM)-dependent protein kinase II. Both genes are conserved in placental-species and are presumably transcribed from initiator (Inr) promoter elements located at opposite strands. In 20 human normal tissues tested, DIPLA1 mRNA expression was placenta-specific, whereas DIPAS mRNA expression was higher in placenta,

brain, kidney and testis. In addition, DIPAS mRNA hybridizes with the 3'UTR region from PAPP-A mRNA, which spans over 4 kb more than previously reported, forming a potential sense-antisense double stranded RNA (dsRNA) duplex. Our results are of interest for placenta gene expression regulation and for the identification of novel genes in the human genome.

Gaughan, D. J., S. Barbaux, et al. (2000). "The human and mouse methylenetetrahydrofolate reductase (MTHFR) genes: genomic organization, mRNA structure and linkage to the CLCN6 gene." Gene **257**(2): 279.

http://www.sciencedirect.com/science/article/B6T39-41N572W-D/2/25c21946afcb005fd86cd5c7a9341210

Gaur, A., L. F. Lemanski, et al. (1995). "Identification and expression of a homologue of the murine HoxA5 gene in the Mexican axolotl (ambystoma mexicanum)." Gene **162**(2): 249.

http://www.sciencedirect.com/science/article/B6T39-3Y6HGRC-F/2/07ce763d58b3fc99c5c551317853c388

An excellent model for studying heart development in vertebrates is the cardiac non-function lethal mutant (gene c) Mexican axolotl, Ambystoma mexicanum. In order to facilitate our analyses of the mutant system, we have undertaken a search for stage-specific molecular markers during embryonic development of the axolotl. We have concentrated on homeobox genes 'as suitable candidates for monitoring molecular changes during development. A 270-bp probe encoding a portion of the axolotl homeobox gene Ahox-1 was generated by PCR from a stage-18 axolotl embryonic cDNA library. 32P-labelled PCR-amplified Ahox-1 DNA was used as the probe for screening a [lambda]AM18 cDNA library using moderately stringent conditions. We isolated six clones and determined their partial nucleotide (nt) sequences. One of the clones, which has very high homology to human, mouse and rat Hox A5 (83 and 99% at the nt and amino-acid levels, respectively, in the homeodomain region), was analyzed further. RT-PCR analyses show that the level of expression of HoxA5 is very low at stage 11 of embryonic development (gastrula). The level of expression reaches maximum at stage 25 (tailbud) and then plateaus at stages 30 and 35 (heartbeat onset). Although the expression of Ahox-1 was also found to start at stage 11, it reaches a maximum level at stage 25 and declines at stage 35. We have also studied, using RT-PCR, the tissue-specific expression of HoxA5 and Ahox-1 in juvenile axolotl.

Ghosh, I., N. Raghavan, et al. (1995). "Nucleoside diphosphate kinase from the parasitic nematode Brugia malayi." Gene **164**(2): 261.

http://www.sciencedirect.com/science/article/B6T39-3Y6HK60-52/2/998cedab88b0d2cf538df71cbd7ca4d9

Using a reverse transcription-polymerase chain reaction (RT-PCR) procedure that exploited the presence of a conserved 22-nucleotide spliced leader (SL) sequence that is trans-spliced to the 5' end of nematode transcripts, a novel Brugia malayi (Bm) infective-stage SL cDNA expression library was constructed and characterized. The library was immunoscreened with rabbit anti-infective-stage antibodies (Ab) and an immunodominant clone, BmG4-7, was identified and characterized. BmG4-7 contained a full-length cDNA that had significant sequence similarity to nucleoside diphosphate kinase (NDK)-encoding sequences reported from a number of species, including Drosophila melanogaster and humans. BmNDK was found to be constitutively

transcribed during all stages of parasite development. An anti-BmNDK Ab was used to immunostain a Western blot of extracts from adult and larval parasites. The Ab specifically recognized a 17.5-kDa molecule in all of the parasite extracts. Molecular modeling of the BmNDK showed several regions surrounding the conserved catalytic site that may be important in the design of drugs specific for the disruption of NTP synthesis in filarial parasites.

Glowacki, G., R. Braren, et al. (2001). "Structure, chromosomal localization, and expression of the gene for mouse ecto-mono(ADP-ribosyl)transferase ART5." Gene **275**(2): 267.

http://www.sciencedirect.com/science/article/B6T39-44357HF-9/2/65f075b208ede1d2f2699ea599a99074

Mono(ADP-ribosyl)transferases regulate the function of target proteins by attaching ADP-ribose to specific amino acid residues in their target proteins. The purpose of this study was to determine the structure, chromosomal localization, and expression profile of the gene for mouse ecto-ADP-ribosyltransferase ART5. Southern blot analyses indicate that Art5 is a single copy gene which maps to mouse chromosome 7 at offset 49.6 cM in close proximity to the Art1, Art2a and Art2b genes. Northern blot and RT-PCR analyses demonstrate prominent expression of Art5 in testis, and lower levels in cardiac and skeletal muscle. Sequence analyses reveal that the Art5 gene encompasses six exons spanning 8 kb of genomic DNA. The 5' end of the Art5 gene overlaps with that of the Art1 gene. A single long exon encodes the predicted ART5 catalytic domain. Separate exons encode the N-terminal leader peptide and a hydrophilic C-terminal extension. Sequencing of RT-PCR products and ESTs identified six splice variants. The deduced amino acid sequence of ART5 shows 87% sequence identity to its orthologue from the human, and 37 and 32% identity to its murine paralogues ART1 and ART2. Unlike ART1 and ART2, ART5 lacks a glycosylphosphatidylinositol-anchor signal sequence and is predicted to be a secretory enzyme. This prediction was confirmed by transfecting an Art5 cDNA expression construct into Sf9 insect cells. The secreted epitope-tagged ART5 protein resembled rat ART2 in exhibiting potent NAD-glycohydrolase activity. This study provides important experimental tools to further elucidate the function of ART5.

Graham, L. A. and P. L. Davies (2002). "The odorant-binding proteins of Drosophila melanogaster: annotation and characterization of a divergent gene family." Gene **292**(1-2): 43.

http://www.sciencedirect.com/science/article/B6T39-45Y6MPT-6/2/c53f8d585da8fa2fba12be48eead9ba9

Insect odorant-binding proteins (OBPs) are thought to facilitate the delivery of hydrophobic odorants, such as sex pheromones or food odors, to receptors on sensory neurons. Increasingly, OBP family members are also being found in non-sensory tissues where they might carry other types of small hydrophobic molecules. They are identifiable by four or six conserved Cys residues and contain six [alpha]-helices which enclose a hydrophobic ligand-binding pocket. Through exhaustive BLAST searches we have increased the total number of OBPs identified in Drosophila melanogaster to 38, and have amplified the DNA complementary to RNA corresponding to 21 of these by reverse transcriptase polymerase chain reaction. Isoforms frequently share less than 30% amino acid identity and appear to have radically changed since the separation of the major insect orders. However, their sequences are consistent with known OBP structures. Most are located in clusters of between four and 14 genes and several were unusual in that they contained additions, deletions, or fusions. These hexa-helical insect OBPs are structurally unrelated to the functionally analogous lipocalin-like [beta]-barrel OBPs of vertebrates. As only two lipocalin-like proteins have been found in D. melanogaster, these helical proteins appear to be the dominant carrier of small hydrophobic molecules in insects.

Gravina, P., N. Campioni, et al. (2002). "Complementary DNA analysis, expression and subcellular localization of hnRNP E2 gene in Xenopus laevis." Gene **290**(1-2): 193.

http://www.sciencedirect.com/science/article/B6T39-45HWVR3-7/2/c6ade71ca6644518d4d2c973d3db9f7f

The cloning and sequencing of complementary DNAs corresponding to the two copies (a and b) of the Xenopus laevis gene for hnRNP E2 is presented. Comparison of the two sequences reveals that while they are somewhat divergent at the nucleotide level, they are very conserved at the amino acid level. The analysis also showed two transcripts of different length ([alpha] and [beta]), likely generated by alternative processing. There are indications that either gene copy can generate both type of transcripts. Northern blot analysis in oocytes and developing embryos showed that hnRNP E2 RNA is constantly present and that increases in amount at tadpole stage. A semiquantitative reverse transcriptase polymerase chain reaction analysis performed with RNA from developing embryos showed that long ([alpha]) transcript accumulation is constant during development, whereas the short one ([beta]) accumulation increases at later stages, thus determining the observed increase in total RNA. Nucleo-cytoplasm localization experiments indicated that in oocyte hnRNP E2 is exclusively cytoplasmic, whereas in somatic cells it is distributed in both compartments.Comparison of the amino acid sequence of the two X. laevis hnRNP E2 with the corresponding mammalian sequences shows a high homology along the molecule except for the region subjected to alternative splicing, which is completely different. Moreover, there are indications that the homologous of mammalian hnRNP E1 gene, very related to and derived from hnRNP E2 by retrotransposition, is not expressed or even not present in X. laevis, suggesting that mammalian hnRNP E1 gene may have originated after mammal/amphybia divergence.

Grell, M. N., P. Mouritzen, et al. (2003). "A Blumeria graminis gene family encoding proteins with a C-terminal variable region with homologues in pathogenic fungi." Gene **311**: 181.

http://www.sciencedirect.com/science/article/B6T39-48S38KG-C/2/60efe4712b96f8c211eb44b2cde5360a

In a study aimed at characterising, at the molecular level, the obligate biotrophic fungus Blumeria graminis f. sp. hordei (Bgh), we have identified a novel group of genes, the Egh16H genes, and shown that two of these are up-regulated during primary infection of barley leaves. The genes have partial homology to a previously characterised Bgh gene family, Egh16. Egh16 and Egh16H are subfamilies of a larger multigene family with presently about 15 members identified in Bgh. Egh16H has about ten members, and we show that five of these are expressed as highly conserved mRNAs that are predicted to encode proteins with a C-terminal variable region. Egh16H has high homology to sequences in Magnaporthe grisea and other plant pathogenic fungi, as well as sequences of both the insect pathogen Metarhizium anisopliae and the human pathogen Aspergillus fumigatus. No close homologues of Egh16H were found in the non-pathogenic fungi Neurospora crassa and Aspergillus nidulans. We predict that Egh16H plays a general role in the interaction between pathogenic fungi and their hosts. At present, the large number of gene family members with C-terminal variation appears to be unique for Bgh, and the Egh16/Egh16H gene family is to our knowledge the largest gene family so far characterised in this fungus.

Gunn, J. S. and D. C. Stein (1993). "Natural variation of the NgoII restriction-modification system of

Neisseria gonorrhoeae." Gene **132**(1): 15.

http://www.sciencedirect.com/science/article/B6T39-47PHB28-2KP/2/a19232fa5935bbd4b0da59cfa298ef36

The NgoII restriction-modification (R-M) system of Neisseria gonorrhoeae recognizes the sequence 5'-GGCC-3'. This system is encoded by two separate genes, dcmB for the methyltransferase (MTase) and dcrB for the restriction endonuclease (ENase). Three strains that vary in their NgoII phenotype were examined. Strain Pgh3-2 produced detectable levels of both enzymes, strain F62 lacked detectable levels of the dcrB gene product, and strain WR302 failed to produce either gene product. Strains that lacked either enzyme activity still possessed the genes that encode them. Transcriptional fusions of dcrB in strains F62 and Pgh3-2 indicate that this gene is transcribed at nearly identical levels in each strain. The DNA encoding the NgoII R-M system was cloned from the three strains, and the nucleotide sequence was determined. The dcrB genes of WR302 and F62 possess the same frameshift mutation (base position 1435) which would result in a truncated protein. The WR302 dcmB was found to have a point mutation that changed Arg288 (a residue that is conserved in all prokaryotic and phage cytosine MTases sequenced to date) to Trp.

Gustafson, C. E., R. A. Alm, et al. (1993). "Effect of heat denaturation of target DNA on the PCR amplification." Gene **123**(2): 241.

http://www.sciencedirect.com/science/article/B6T39-47PH7C3-1MF/2/12c98cfceb38fef156747e8d095dfa21

The polymerase chain reaction (PCR) and amplification of specific regions of DNA in vitro is a widely used and powerful technique, and the optimization of conditions used to maximize PCR product yield has received much attention. We have shown that lengthy denaturation times of template DNA ranging from 1 to 7 min at pH 7.0-8.0, that are often employed prior to the start of a PCR reaction, result in marked degradation of the template. This can result in a significant reduction in the yield of PCR products larger than 500 bp, by up to 99%. This effect was demonstrated for both complex genomic template DNA, and also for a 2691-bp linear piece of template DNA using both a rapid hot-air thermocycler and a conventional block thermocycler. This decrease in product yield is likely due to the increased degradation of the template or target DNA as a result of pre-amplification denaturation (PAD). We therefore recommend that when amplifying larger pieces of DNA, the template DNA should not be exposed to PAD prior to a PCR reaction, irrespective of the starting pH of the template solution.

Hanas, J. S., J. R. Hocker, et al. (2002). "cDNA cloning, DNA binding, and evolution of mammalian transcription factor IIIA." Gene **282**(1-2): 43.

http://www.sciencedirect.com/science/article/B6T39-44KWSWR-1/2/dd8a2f2ea3cb776b8732c17a27f17664

cDNA for rat transcription factor IIIA (TFIIIA) was cloned by degenerate PCR and rapid amplification of cDNA ends. This cDNA coded for a protein with nine Cys2His2 zinc fingers and a non-finger C-terminal tail; 63% amino acid (aa) sequence identity was observed with the Xenopus TFIIIA zinc finger region. Recombinant rat protein containing only the nine fingers afforded DNase I protection of the identical nucleotides protected by Xenopus laevis native TFIIIA on the Xenopus 5S RNA gene internal control region. A putative mouse TFIIIA clone was identified in an expressed sequence tag database by sequence similarity to rat TFIIIA. Recombinant nine-finger

protein from this clone afforded DNase I protection of the Xenopus 5S rRNA gene like the native frog protein as did a recombinant nine-finger form of a putative human TFIIIA clone. These DNA binding results demonstrate that these clones code for the respective mammalian TFIIIAs. Rodent and human TFIIIAs share about 87% aa sequence identity in their zinc finger regions and have evolved to about the same extent as X. laevis and Xenopus borealis TFIIIAs. A monoclonal antibody against human p53 tumor suppressor bound to rat and mouse TFIIIA but not to human TFIIIA in Western blots. The N-terminal regions of rodent and human TFIIIA do not contain the oocyte-specific initiating Met and accompanying conserved residues found in fish and amphibian TFIIIAs. In their non-finger C-terminal tails, mammalian and amphibian TFIIIAs share a conserved transcription activation domain as well as conserved nuclear localization and nuclear export signals.

Hanneman, W. H., K. J. Schimenti, et al. (1997). "Molecular analysis of gene conversion in spermatids from transgenic mice." Gene **200**(1-2): 185.

http://www.sciencedirect.com/science/article/B6T39-3Y9GDS5-R/2/2d3ff905043c104a2ebc76fee75c0096

Investigations into the mechanisms and properties of gene conversion in mammals are greatly restricted by the inability to recover all the products of a meiosis. Additionally, the study of this process has been hampered by the lack of visible markers to detect gene conversion, especially when the events are rare. In previous work, we developed a transgenic system for detection and quantitation of gene conversion events in the germline of mice (Murti, J.R., Bumbulis, M., Schimenti, J.C., 1992. High frequency germline gene conversion in transgenic mice. Mol. Cell. Biol. 12, 2545-2552) that could be exploited as an assay for recombinogenic chemicals (Murti, J.R, Schimenti, K.J., Schimenti, J.C., 1994. A recombination-based transgenic mouse system for genotoxicity testing. Mutat. Res. 307, 583-595). A specific intrachromosomal gene conversion event between two complementarily defective lacZ genes resulted in the production of [beta]-galactosidase in spermatids, enabling a measurement of conversion frequency. Here, we report that the anticancer drug, cisplatin, increased gene conversion in meiotic stage cells in these transgenic mice. Furthermore, a method was developed for direct molecular analysis of transgene conversion events in single or pooled lacZ-positive spermatids. The ability to identify gametes that have undergone a rare gene conversion event, followed by molecular amplification of the recombinant gene, should make it possible to investigate the mechanisms of genetic recombination in mammals in greater detail than previously possible.

Hardiman, G., S. Albright, et al. (1996). "The mouse Wnt-10B gene isolated from helper T cells is widely expressed and a possible oncogene in BR6 mouse mammary tumorigenesis." Gene **172**(2): 199.

http://www.sciencedirect.com/science/article/B6T39-3W496YH-4/2/8de1a79cf220242b82fb17bad21d7f69

From libraries made from activated mouse T lymphocytes, we have isolated cDNAs encoding Wnt-10B, a new member of the Wnt family of developmental control genes. This protein appears to be the mammalian orthologue of Wnt-10B, first identified in several non-mammalian vertebrates and recently in mouse. The mRNA expression pattern of mouse Wnt-10B indicates that it is induced following activation of helper T cells, but is also expressed in a variety of other tissues and cells of fetal or adult origin. 93 bp at the 5' end of the cDNA clone are identical to sequences previously reported as 3' flanking genomic DNA adjacent to a mouse mammary tumor virus (MMTV) provirus in the MMTV-induced BR6 mammary tumor, W26. Sequence analysis of tumor-derived genomic DNA confirms that the entire Wnt-10B gene is immediately adjacent to the provirus, suggesting that MMTV integration drives transcription of Wnt-10B, possibly contributing

to the oncogenic process. Consistent with this idea is the detection of hybrid MMTV-Wnt-IOB transcripts in BR6 tumor cells. T cells which produce abundant Wnt-10B mRNA were also found to produce protein.

Hart, P. E., G. M. Poynter, et al. (2001). "Characterization of the X-linked murine centrin Cetn2 gene." Gene **264**(2): 205.

http://www.sciencedirect.com/science/article/B6T39-42HNT5R-6/2/b00088b26a192da2307d1650660fd112

A multi-gene family (Cetn1, Cetn2, and Cetn3) encodes the calcium-binding protein, centrin, in the mouse. This work characterizes the Cetn2 gene. Structurally, Cetn2 consists of five exons and four introns, and contains a classical TATA-less promoter. Cetn2 has two alternate transcription start sites, and a single length 3' untranslated region. Fluorescence in situ hybridization demonstrates that Cetn2 is an X-linked gene whose alleles replicate asynchronously during S-phase. Cetn2 encodes a 172 amino acid protein, with a predicted molecular mass of 19,795 Da (pI=4.71), that contains all of the defining characteristics of centrin. Northern blot analysis indicates that Cetn2 is ubiquitously expressed in the tissues of adult mice. RT-PCR shows that Cetn2 and Cetn3, but not Cetn1, are expressed in NIH 3T3 cells. Immunofluorescence microscopy demonstrates that mouse centrin 2 protein localizes to the region immediately surrounding the centrioles in the centrosome of NIH 3T3 cells.

Hartweck, L. M., D. J. Llewellyn, et al. (1997). "The Arabidopsis thaliana genome has multiple divergent forms of phosphoinositol-specific phospholipase C." Gene **202**(1-2): 151.

http://www.sciencedirect.com/science/article/B6T39-3TVNRN6-S/2/bc7dcf03057c33cfb65b4fd197d72934

Highly degenerate primers to conserved regions of the eukaryotic phosphoinositol-specific phospholipase C (PLC) were used to amplify fragments of plant PLCs from Arabidopsis thaliana genomic DNA. Eight completely different fragment sequences that showed high homology to PLCs of both animals and plants were isolated. The variation between these putative PLCs was high and suggests that, like animals, plants have multiple isoforms of PLC. Using one of the PCR clones, we isolated a corresponding full-length Arabidopsis PLC gene (ATHATPLC1G), and sequence analysis indicated that it was most like a delta-type PLC. This gene is 2.5 kb and contains seven introns, all but one of which has intron/exon border sequences that conform to the Arabidopsis consensus. The structural complexity of the gene is relatively simple compared to mammalian [beta]-type PLCs that can be 15 kb long with up to 30 introns. The plant gene is a single copy and was mapped to four Arabidopsis YACs, one located on chromosome 2. The promoter region contained two TATA-like elements at -43 and -185 and other putative regulatory elements that suggest that this PLC is hormonally regulated. This is the first plant PLC gene and the first delta type-PLC gene from a higher organism to be sequenced.

Hazuka, C. D., S.-C. Hsu, et al. (1997). "Characterization of a cDNA encoding a subunit of the rat brain rsec6/8 complex." Gene **187**(1): 67.

http://www.sciencedirect.com/science/article/B6T39-3VWFYVS-9/2/72c748ab32c1eaeea0fb4467f574ea4a

Many proteins required for neurotransmission are homologous to proteins involved in the Golgi-to-plasma membrane stage of the yeast secretory pathway. A novel 17S complex composed of eight proteins including rsec6 and rsec8, the rat homologues of the yeast secretory proteins, Sec6p and Sec8p, has been identified in rat brain cytosol. Sec6p and Sec8p are components of a complex of at least seven proteins which are essential for secretion in yeast. While the complementary DNAs (cDNA) encoding rsec6 and rsec8 have been cloned [Ting et al. (1995) Proc. Natl. Acad. Sci. USA 92, 9613-9617], the other six components of the 17S complex remain undescribed. Using the peptide sequence obtained from p71, one of the subunits of the rat brain 17S complex, we isolated a full-length cDNA from a rat brain library. This cDNA is predicted to encode a hydrophilic protein of 82 kDa, similar in size to that observed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) for the endogenous rat brain rsec6/8 complex 71 kDa component. p71 contains domains of peptide sequence which display significant homology to regions of the tail domain of yeast type II myosin. Northern blot analysis of rat tissues indicates that messenger RNA transcripts of 3.1 and 4.4 kb encoding this protein are expressed broadly across several rat tissues in a pattern similar to that of rsec6 and rsec8 mRNA expression. A possible role for p71 as a point of interaction for proteins of the cytoskeleton and proteins involved in secretion is discussed.

Henry, T., B. Kliewer, et al. (1996). "Isolation and characterization of a bovine gene encoding phenol sulfotransferase." Gene **174**(2): 221.

http://www.sciencedirect.com/science/article/B6T39-3W257W9-6/2/2942ef4407db2cd8b82dfaa54fc1ef9b

Phenol sulfotransferases (PST) esterify hydroxylated aromatic compounds with sulfate, and therefore play a role in the metabolism of xenobiotics. In this investigation, a bovine PST gene of 2372 bp was amplified from genomic DNA. Sequence overlap with the cognate cDNA revealed seven exons, with all introns containing GT/AG splicing donor/acceptor sites. This product was subcloned into an SV40-based expression vector and transfected into WI-26 human lung fibroblasts. Immunoblot analyses revealed production of the anticipated 32-kDa protein, and the active enzyme displayed steady-state kinetic properties consistent with the enzyme characterized in bovine lung (apparent Km=6.5 [mu]M for 2-naphthol).

Herlitze, S. and M. Koenen (1990). "A general and rapid mutagenesis method using polymerase chain reaction." Gene **91**(1): 143.

http://www.sciencedirect.com/science/article/B6T39-47T307S-1F/2/ae3de030d0e5e807b96cd6b7eaa02879

The constructions of deletions, insertions and point mutations in DNA sequences is a powerful approach to analysing the function and structure of genes and their products. Here, we present a fast and efficient method using the polymerase chain reaction to introduce mutations into cDNAs coding for the [alpha]-, [gamma]- and [epsiv]-subunit of the rat muscle acetylcholine receptor. Two flanking primers and one mutant oligo, in conjunction with supercoiled plasmid DNA and a fragment of the target DNA are sufficient to introduce the mutation by two PCR amplifications. Our method permits directing the location of mutations anywhere in the target gene with a very low misincorporation rate, as no substitution could be delected within 9600 bp. The utility of this approach is demonstrated by the rapid introduction and analysis of eleven mutations into three different cDNAs. Any kind of mutation can be introduced with an efficiency of at least 50%.

Hester, S. D., G. B. Benavides, et al. (2002). "Normal gene expression in male F344 rat nasal transitional and respiratory epithelium." Gene **285**(1-2): 301.

http://www.sciencedirect.com/science/article/B6T39-454TC74-7/2/94ee95d56e840ea093728c3194787805

The nasal epithelium is an important target site for chemically-induced toxicity and carcinogenicity in rodents. Gene expression profiles were determined in order to provide normal baseline data for nasal transitional/respiratory epithelium from healthy rats. Cells lining the rat nasal passages were collected and gene expression analysis was performed using Clontech(TM) cDNA Rat Atlas 1.2 arrays (1185 genes). The percentages of genes within specific average expression ranges were 4.2% at 45,000-1000, 14.8% at 1000-200, 25.0% at 200-68, and 56.0% below 68. Nine out of a subset of ten genes were confirmed for relative signal intensity using quantitative real-time RT-PCR. The most highly expressed genes included those involved in phase I (e.g. cytochrome P450s) and phase II (e.g. glutathione S-transferases) xenobiotic metabolism, bioenergetics (e.g. cytochrome oxidase), osmotic balance (e.g. Na+/K+ ATPase) and epithelial ionic homeostasis (e.g. ion channels). Such baseline data will contribute to further understanding the normal physiology of these cells and facilitate the interpretation of responses by the nasal epithelial cells to xenobiotic treatment or disease.

Hey, P. J., R. C. J. Twells, et al. (1998). "Cloning of a novel member of the low-density lipoprotein receptor family." Gene **216**(1): 103.

http://www.sciencedirect.com/science/article/B6T39-3TDHM52-D/2/ee75ea90215dabda335b2fb837468f43

A gene encoding a novel transmembrane protein was identified by DNA sequence analysis within the insulin-dependent diabetes mellitus (IDDM) locus IDDM4 on chromosome 11q13. Based on its chromosomal position, this gene is a candidate for conferring susceptibility to diabetes. The gene, termed low-density lipoprotein receptor related protein 5 (LRP5), encodes a protein of 1615 amino acids that contains conserved modules which are characteristic of the low-density lipoprotein (LDL) receptor family. These modules include a putative signal peptide for protein export, four epidermal growth factor (EGF) repeats with associated spacer domains, three LDL-receptor (LDLR) repeats, a single transmembrane spanning domain, and a cytoplasmic domain. The encoded protein has a unique organization of EGF and LDLR repeats; therefore, LRP5 likely represents a new category of the LDLR family. Both human and mouse LRP5 cDNAs have been isolated and the encoded mature proteins are 95% identical, indicating a high degree of evolutionary conservation.

Heykants, L., E. Schollen, et al. (2001). "Identification and localization of two mouse phosphomannomutase genes, Pmm1 and Pmm2." Gene **270**(1-2): 53.

http://www.sciencedirect.com/science/article/B6T39-436W3KY-5/2/9b4ff33d180255b243e5a6941acc85a9

Phosphomannomutases catalyze the reversible conversion of mannose 6-phosphate to mannose 1-phosphate. In humans, two different isozymes have recently been identified, PMM1 and PMM2. We have previously shown that mutations in the PMM2 gene cause the most frequent type of the congenital disorders of glycosylation, CDG-Ia. Here, we present data on the two mouse orthologous genes, Pmm1 and Pmm2. The chromosomal localization of the two mouse genes has been determined. We also present the gene structure and the exon-intron organization of

Pmm1 and Pmm2. Pmm1 maps to mouse chromosome 15, Pmm2 to chromosome 16. These chromosomal regions are syntenic with regions on human chromosomes 22 and 16, respectively. The Pmm1 gene is composed of eight exons and spans approximately 9.5 kb. The genomic structure is extremely well conserved between the human and mouse gene. The Pmm2 gene consists of eight exons and spans a larger genomic region ([ap]20 kb). An alignment of the human and mouse protein sequences confirms the conservation among this family of phosphomannomutases. The two mouse genes are expressed in many tissues, but the expression pattern is slightly different between Pmm1 and Pmm2. The most striking difference is the high expression of Pmm1 in brain tissue, whereas Pmm2 is only weakly expressed in this tissue.

Hirochika, H. and H. Otsuki (1995). "Extrachromosomal circular forms of the tobacco retrotransposon Ttol." Gene **165**(2): 229.

http://www.sciencedirect.com/science/article/B6T39-3Y6HK2X-2P/2/8ecbcaf23d0aff623c9a195285015fe9

Extrachromosomal DNA forms of Drosophila retrotransposons (RTn) and retroviruses have been extensively analyzed. However, no such analysis with plant RTn has been reported. Here, we report the analysis of extrachromosomal forms of the tobacco RTn Ttol. Ttol is one of a few active RTn of plants and has been shown to be activated in tissue culture. Extrachromosomal circular DNA forms of Ttol, with one or two long terminal repeats (LTR), were found in cultured cells. Sequence analysis of the sites of circularization through joining two LTR showed that the junction between the LTR contains small deletions and/or insertions. The insertions are heterogeneous and do not show any homology to the Ttol sequence. Similar insertions have been detected in the extrachromosomal circular forms of the copia element of Drosophila and suggested to be the result of excision of genomic copia. The structural features of the junctions found in Ttol suggest that the insertions are produced by a mechanism other than excision. The potential mechanism of production of the extrachromosomal circular forms of Ttol is discussed.

Hirose, T., D. A. O'Brien, et al. (1995). "RTR: a new member of the nuclear receptor superfamily that is highly expressed in murine testis." Gene **152**(2): 247.

http://www.sciencedirect.com/science/article/B6T39-3Y5MNR1-7T/2/e8fc930799a698d93a793fada920dc22

We have identified and cloned a novel member of the nuclear receptor superfamily from murine testis, referred to as retinoid receptor-related testis-associated receptor or RTR. Degenerate PCR primers homologous to two conserved regions of the DNA-binding domain of members of this superfamily were employed to identify this gene. The amino-acid sequence of RTR is most closely related to that of the mouse RXRs with an overall identity of 32-34%; the highest similarity (61%) is observed in the DNA-binding domain. Northern blot analysis using RNA from multiple tissues showed that RTR is predominantly expressed in the testis. Northern blot analysis using RNA from different testicular cell types showed that RTR mRNA is not expressed in early germ cells or Sertoli cells but is most abundant in round spermatids. Our observations suggest that this putative transcription factor plays a role in the regulation of gene expression particularly during the post-meiotic phase of spermatogenesis.

Hiroshi, T., F. Kazuhiro, et al. (1994). "Gene sequence, purification and characterization of N-acetyl-

[beta]-glucosaminidase from a marine bacterium, Alteromonas sp. strain O-7." Gene **146**(1): 111.

http://www.sciencedirect.com/science/article/B6T39-47PNXG6-HG/2/00af5a9919995f88086ab24554a03b26

The gene (cht60) encoding N-acetyl-[beta]-glucosaminidase (Cht; EC 3.2.1.30) from the marine bacterium Alteromonas sp. strain O-7 was cloned into pUC18 in Escherichia coli JM109. The nucleotide (nt) sequence of cht60 was determined. A 1797-bp open reading frame encoded a polypeptide of 598 amino acids (aa) (Mr 64535). The aa sequence of the cloned enzyme (Cht60) deduced from the nt sequence showed no significant sequence homologies with available aa sequences from databases. Cht60 was purified from the periplasmic fraction of E. coli cells carrying pCHT982. The enzyme was most active towards ide(PNP-[beta]-GlcNAc) and diacetylchitobiose. The optimum pH and temperature of the enzyme were pH 7.5 and 37[deg]C, respectively. The N-terminal 11 aa residues of Cht60 were sequenced, and the location of the signal peptide cleavage site was clarified.

Holler, T. P., S. K. Foltin, et al. (1993). "HIV1 integrase expressed in Escherichia coli from a synthetic gene." Gene **136**(1-2): 323.

http://www.sciencedirect.com/science/article/B6T39-47P8H5B-K5/2/20735e327a7047aa633e3881addd2b36

Human immunodeficiency virus type 1 (HIV1) integrase is cleaved from the gag-pol precursor by the HIV1 protease. The resulting 32-kDa protein is used by the infecting virus to insert a linear, double-stranded DNA copy of its genome, prepared by reverse transcription of viral RNA, into the host cell's chromosomal DNA. In order to achieve high levels of expression, to minimize an internal initiation problem and to facilitate mutagenesis, we have designed and synthesized a gene encoding the integrase from the infectious molecular clone, pNL4-3. Codon usage was optimized for expression in Escherichia coli and unique restriction sites were incorporated throughout the gene. A 905-bp cassette containing a ribosome-binding site, a start codon and the integrase-coding sequence, sandwiched between EcoRI and HindIII sites, was synthesized by overlap extension of nine long synthetic oligodeoxyribonucleotides [90-120 nucleotides (nt)] and subsequent amplification using two primers (28-30 nt). The cassette was subcloned into the vector pKK223-3 for expression under control of a tac promoter. The protein produced from this highly expressed gene has the expected N-terminal sequence and molecular mass, and displays the DNA processing, DNA joining and disintegration activities expected from recombinant integrase. These studies have demonstrated the utility of codon optimization, and lay the groundwork for structure-function studies of HIV1 integrase.

Hong, C.-S., J.-H. Ji, et al. (2001). "Molecular cloning and characterization of mouse cardiac triadin isoforms." Gene **278**(1-2): 193.

http://www.sciencedirect.com/science/article/B6T39-44CMXX2-M/2/1647d958db41fedbb828a682ef299a47

Triadin is a ryanodine receptor and calsequestrin binding protein located in junctional sarcoplasmic reticulum of striated muscles. In the present study, mouse cardiac triadin cDNAs have been identified by cDNA library screening and RT-PCR. The deduced aa sequences show that the three isoforms consist of 277, 293 and 305 aa giving rise to the molecular weights of approximately 31,414, 33,066, and 34,328, respectively. The isoforms have identical 262 aa N-terminal sequences, whereas they have distinct C-terminal sequences. Northern blot analysis

using a cDNA probe representing the N-terminal common region of triadin revealed that the mouse triadins were present both in heart and skeletal muscles. The estimated sizes of the transcripts were approximately 1.3, 4.3 and 5 kb in heart and 5, 5.5 and 7 kb in skeletal muscle. Endo H treatment and Western blot analysis of isolated mouse cardiac sarcoplasmic reticulum and in vitro translation products indicate that there are three distinct mouse cardiac triadin isoforms having molecular weights of 35, 35.5 and 40 kDa. We termed those three isoforms as mouse cardiac triadin 1, mouse cardiac triadin 2 and mouse cardiac triadin 3.

Horiuchi, H., T. Inoue, et al. (2001). "Characterization and expression of three forms of cDNA encoding chicken platelet-derived growth factor-A chain." Gene **272**(1-2): 181.

http://www.sciencedirect.com/science/article/B6T39-43HT1F4-R/2/4fcf1c7903b32bbd0d6535e5597cd03f

Platelet-derived growth factor (PDGF) affects cell proliferation and differentiation during mammalian embryogenesis. In a number of avian species, PDGF-[alpha] receptors and PDGF-A chain (PDGF-A) are present during chicken limb and lens development. However, little is understood about the chicken PDGF-A gene. The present study identified short form type 1 (S1), long form (L) and short form type 2 (S2) cDNA clones encoding chicken PDGF-A chain (PDGF-A). These clones were isolated from a chicken hepatoma cell line (LMH) mRNA by reverse transcriptase-polymerase chain reaction (RT-PCR) and cDNA library cloning. Genomic sequencing and Southern blotting revealed that these forms were generated by alternative splicing. The mRNAs of S1 and L contained two transcription start sites on one exon. At the amino acid level, the mature protein encoded by the L clone showed 90 and 85% homology with the processed coding regions of the long form of human and Xenopus PDGF-A, respectively. The putative mature peptides of all forms of chicken PDGF-A encompassed the eight cysteine residues conserved in all known forms of PDGF. We examined the expression of the three forms in chicken tissues and cells using RT-PCR. Expression of these forms varied among tissues and cells. Levels of PDGF mRNAs were very low in chicken thrombocytes, which are analogous to mammalian platelets. However, the level of PDGF-A chain mRNA expression in chicken thrombocytes peaked 4 h after exposure to type 1 collagen or thrombin, and then decreased gradually with continued incubation. These results suggest that chicken PDGF in thrombocytes plays an important role in the vascular system and in healing damaged tissue.

Hosking, B. M., J. R. Wyeth, et al. (2001). "Cloning and functional analysis of the Sry-related HMG box gene, Sox18." Gene **262**(1-2): 239.

http://www.sciencedirect.com/science/article/B6T39-429XV3N-10/2/36a423e4eae972461bfa934540965af0

The Sox gene family (ry like HMG b gene) is characterised by a conserved DNA sequence encoding a domain of approximately 80 amino acids which is responsible for sequence specific DNA binding. We initially published the identification and partial cDNA sequence of murine Sox18, a new member of this gene family, isolated from a cardiac cDNA library. This sequence allowed us to classify Sox18 into the F sub-group of Sox proteins, along with Sox7 and Sox17. Recently, we demonstrated that mutations in the Sox18 activation domain underlie cardiovascular and hair follicle defects in the mouse mutation, ragged (Ra) (Pennisi et al., 2000. Mutations in Sox18 underlie cardiovascular and hair follicle defecs in ragged mice. Nat. Genet. 24, 434-437). Ra homozygotes lack vibrissae and coat hairs, have generalised oedema and an accumulation of chyle in the peritoneum. Here we have investigated the genomic sequences encoding Sox18. Screening of a mouse genomic phage library identified four overlapping clones, we sequenced a 3.25 kb XbaI fragment that defined the entire coding region and approximately 1.5 kb of 5'

flanking sequences. This identified (i) an additional 91 amino acids upstream of the previously designated methionine start codon in the original cDNA, and (ii) an intron encoded within the HMG box/DNA binding domain in exactly the same position as that found in the Sox5, -13 and -17 genes. The Sox18 gene encodes a protein of 468 aa. We present evidence that suggests HAF-2, the human HMG-box activating factor -2 protein, is the orthologue of murine Sox18. HAF-2 has been implicated in the regulation of the Human IgH enhancer in a B cell context. Random mutagenesis coupled with GAL4 hybrid analysis in the activation domain between amino acids 252 and 346, of Sox18, implicated the phosphorylation motif, SARS, and the region between amino acid residues 313 and 346 as critical components of Sox18 mediated transactivation. Finally, we examined the expression of Sox18 in multiple adult mouse tissues using RT-PCR. Low-moderate expression was observed in spleen, stomach, kidney, intestine, skeletal muscle and heart. Very abundant expression was detected in lung tissue.

Hsu, S.-C., F. Kirschenbaum, et al. (2002). "Structural and functional characterization of the upstream regulatory region of the human gene encoding prostate apoptosis response factor-4." Gene **295**(1): 109.

http://www.sciencedirect.com/science/article/B6T39-46RMWSK-H/2/544b1255ada28e48773587e82e250826

Prostate apoptosis response factor-4 (Par-4) is critical to cell growth and apoptosis. Induction of Par-4 expression has been shown to be required for apoptosis in a diversity of cellular systems, including neurons. Neuronal populations in individuals with degenerative disorders show elevated levels of Par-4 protein in advance of cellular and functional loss. To understand the regulation of par-4 expression, we isolated and characterized 5.7 kb of the human par-4 promoter. We demonstrated that the isolated promoter was functional. Similar to the endogenous par-4 gene, par-4 expression could be induced upon apoptotic insult with thapsigargin following introduction of the promoter DNA into human A375 cells. Also, increased levels of the atypical protein kinase C, [zeta]PKC, was shown to negatively regulate expression from the ectopic par-4 promoter. A 550 bp sequence immediately upstream to the 5'-untranslated region of the gene was found to be responsible for par-4 promoter induction to apoptosis by thapsigargin.

Huang, W., J. Escribano, et al. (1999). "Identification, expression and chromosome localization of a human gene encoding a novel protein with similarity to the pilB family of transcriptional factors (pilin) and to bacterial peptide methionine sulfoxide reductases." Gene **233**(1-2): 233.

http://www.sciencedirect.com/science/article/B6T39-3WRB6HH-X/2/216a9211a6125ee983486ced995899ba

Huo, L. and R. C. Scarpulla (1999). "Multiple 5'-untranslated exons in the nuclear respiratory factor 1 gene span 47 kb and contribute to transcript heterogeneity and translational efficiency." Gene **233**(1-2): 213.

http://www.sciencedirect.com/science/article/B6T39-3WRB6HH-V/2/f0d2ce151e389e8efc12d94f3f861670

Imai, K., A. H. Sarker, et al. (1998). "Genomic structure and sequence of a human homologue

(NTHL1/NTH1) of Escherichia coli endonuclease III with those of the adjacent parts of TSC2 and SLC9A3R2 genes." Gene **222**(2): 287.

http://www.sciencedirect.com/science/article/B6T39-3V7RVWT-K/2/d19ac92747c56f78ba5d2b75ebcb0ee5

Genomic cloning and sequencing of a human homologue (the gene name, endonuclease III-like 1; gene symbol, NTHL1 or NTH1) for Escherichia coli endonuclease III, that is involved in pyrimidine base excision repair, were performed. The sequence covered the entire NTHL1 gene consisting of six exons and five introns spanning 8 kb with 5' flanking (8 kb) and 3' flanking (3.8 kb) regions. Southern blot analysis suggested that the NTHL1 gene exists as a single copy in a haploid genome. The sequenced 5' flanking region lacks typical TATA and CAAT boxes, but contains a CpG island having putative binding sites for several transcription factors such as Ets1 and Sp1. The NTHL1 gene lies immediately adjacent to the tuberous sclerosis 2 (TSC2) gene on chromosome 16p13.3 in a 5'-to-5' orientation. Transcription initiation sites of both NTHL1 and TSC2 genes were suggested to be multiple by 5' RACE experiments. The northern hybridization experiment suggested that both genes are expressed in all tissues, but at different levels. Downstream of the NTHL1 gene, the gene for the regulatory factor 2 (SLC9A3R2/E3KARP; also called OCTS2, TKA-1 and SIP-1) of the solute carrier family 9 (sodium/hydrogen exchanger), isoform A3, lies in a 3'-to-3' orientation. This paper demonstrates for the first time the spatial relationship of these three genes (TSC2, NTHL1 and SLC9A3R2) at the nucleotide level, and the presence of multiple transcription initiation sites of the NTHL1 and TSC2 genes.

Inagaki, Y., Y. Johzuka-Hisatomi, et al. (1999). "Genomic organization of the genes encoding dihydroflavonol 4-reductase for flower pigmentation in the Japanese and common morning glories." Gene **226**(2): 181.

http://www.sciencedirect.com/science/article/B6T39-3VNR70Y-6/2/c98a02021cb701f8f3b1758b20375d49

Genomic DNA segments (approximately 17 kb) containing three DFR genes in the Japanese and common morning glories were sequenced. The three DFR genes in both plants were found to be arranged in a tandem array, and all of them comprised six exons with identical intron positions. Their DFR-B genes carrying longer introns than the DFR-A and DFR-C genes were expressed extensively in the young buds of pigmented flowers, and the transcription starting site for the DFR-B mRNA of the Japanese morning glory was determined. The DFR-B gene of the common morning glory was expressed considerably in stems, moderately in sepals and leaves, whereas the DFR-A and DFR-C genes of the same plant were expressed scarcely but significantly in the young flower buds and stems. Several novel mobile element-like sequences of around 200 bp were found in the genomic DFR regions. A phylogenetic tree indicated that each DFR gene in the Japanese morning glory is most closely related to the corresponding DFR gene in the common morning glory, and that the DFR-B gene is the most diversified gene among the three DFR genes. These structural and functional features of the DFR genes and their evolutionary implications are discussed.

Ishikawa, S., I. Kobayashi, et al. (2001). "Interaction of MCC2, a novel homologue of MCC tumor suppressor, with PDZ-domain Protein AIE-75." Gene **267**(1): 101.

http://www.sciencedirect.com/science/article/B6T39-42SPKX3-C/2/676a45487d97879ca85824b1db3ca3fd

AIE-75 is a protein identified as an autoantigen in patients with autoimmune enteropathy and as a colon cancer-related antigen. It has recently been assigned to be a causative gene for Usher type 1C congenital syndromic hearing loss. The novel protein has three PSD-95/Dlg/ZO-1 (PDZ) protein-protein interaction domains and is therefore implicated to function as a molecular anchor or sorter. We have identified a novel protein that binds to AIE-75 by yeast two-hybrid screening. The protein has a high homology to the tumor suppressor MCC (mutated in colon cancer; or MCC1 hereafter) and was named MCC2. MCC2 protein binds the first PDZ domain of AIE-75 with its C-terminal amino acids -DTFL. Since the MCC1 does not bind to AIE-75 and the MCC2 displays different expression patterns in various organs compared to MCC1, they appear to play distinct roles in cells. The MCC2 gene is located on chromosome 19p13 in the vicinity of APCL gene, while MCC1 maps near to APC tumor suppressor gene. Because of negative expression of MCC2 in a panel of cancer cell-lines compared to the corresponding normal tissues, we suggest that further study is necessary to investigate a possible role of MCC2 as a tumor suppressor.

Iwamoto, K., Y.-T. Huang, et al. (2000). "Genomic organization and alternative transcripts of the human PQBP-1 gene." Gene **259**(1-2): 69.

http://www.sciencedirect.com/science/article/B6T39-423RGKS-9/2/4c9781ba4ae9b856e250e3d4d765816e

PQBP-1 has been identified as a protein that binds to huntingtin, androgen receptor and transcription factor Brain-2 through their homopolymeric glutamine repeats. We here report the genomic organization of the human PQBP-1 gene and its multiple alternative transcripts. The coding region of PQBP-1 comprises six exons and five introns, and four types of alternative transcript, designated PQBP-1a to PQBP-1d, were found in addition to the PQBP-1 transcript reported originally. All of the PQBP-1 transcripts retain the WW domain in the N-terminal region, a potent transactivator domain. On the other hand, there is a wide variation in their C-terminal regions. Importantly, PQBP-1a and PQBP-1d lack the domain responsible for the interaction with homopolymeric glutamine repeats and a nuclear localization signal.

Iwata, H., T. Kato, et al. (2000). "Triparental origin of Damask roses." Gene **259**(1-2): 53.

http://www.sciencedirect.com/science/article/B6T39-423RGKS-7/2/900620355bb7191343618b4f30781654

Damask roses are one group of old rose varieties and a key material in old European rose improvement in the 19th century. To clarify the origin of Damask roses, we selected four varieties as the oldest Damask varieties and examined the relationship between the Damask varieties and their putative ancestors at the molecular level. Randomly amplified polymorphic DNA analysis of the Damask varieties proved that they had an identical profile, indicating they were established from a common ancestor. They have never been allowed to reproduce sexually; their reproduction depends entirely on vegetative propagation. We identified three Rosa species, R. moschata, R. gallica and R. fedschenkoana, as parental species of the original hybridization that contributed to forming the four oldest Damask varieties by sequencing the internal transcribed spacer of ribosomal DNA. We also found that all the four oldest Damask varieties had chloroplasts derived only from R. moschata, as judged from psbA-trnH spacer sequences. This triparental origin of the four oldest Damask varieties can explain some morphological characteristics of the four oldest Damask varieties, like fruit shape, leaf color and the 'Moss' character.

Jin, J., C. B. Poole, et al. (1999). "Alternative splicing creates sex-specific transcripts and truncated forms of the furin protease in the parasite Dirofilaria immitis." Gene **237**(1): 161.

http://www.sciencedirect.com/science/article/B6T39-3X944FT-M/2/f2c81fa4c8a0ae4c2e6c56c3abaa8c9b

Jirholt, P., M. Ohlin, et al. (1998). "Exploiting sequence space: shuffling in vivo formed complementarity determining regions into a master framework." Gene **215**(2): 471.

http://www.sciencedirect.com/science/article/B6T39-3VGR2GV-2W/2/15d9a583ef41189a9aef8d751839cf58

A novel approach in molecular design is presented, where in vivo formed complementarity determining regions (CDR) from antibody genes were shuffled into a specific framework region. A synthetic gene library of soluble VH-fragments was created and the complexity of the library was determined by sequencing. The synthetic genes were diverse and contained random combinations of CDR from different germlines. All CDR were randomised in one step and by using in vivo formed CDR, the length, sequence and combination were varied simultaneously.

Jung, Y.-K., J.-H. Jeong, et al. (2004). "Gene expression profile of human chondrocyte HCS-2/8 cell line by EST sequencing analysis." Gene **330**: 85.

http://www.sciencedirect.com/science/article/B6T39-4BVNYC4-1/2/66dd933ca8e2309db15a47102c977576

Large-scale single-pass sequencing of randomly selected cDNA clones from cell type specific libraries has proven to be a powerful approach for the discovery of novel gene functions, identification of novel gene family members, and definition of gene expression profiles. HCS-2/8 chondrocyte has been used as a cell culture model to study chondrocyte differentiation. Here we performed 3350 single-pass sequencing reactions obtained from the 5' ends of cDNAs from HCS-2/8 cells. To define the expression profiles of HCS-2/8 chondrocytes, we analyzed the identity of these representative cDNA sequences using database searches (BLAST). The sequences represent 1927 unique genes with known function (i.e., unigene clusters), 38 transcripts that are similar to genes with known function, 739 expressed genes with unknown function (i.e., expressed sequence tags), and 18 cDNAs which have not previously been sequenced. Interestingly, many transcripts were expressed from chromosome 12 compared with total genes, while the fewer numbers of cDNAs were derived from genes on chromosomes 14, 18 and Y. The chondrocytic phenotype of HCS-2/8 cells is reflected by abundant expression of genes related to cell structure and motility and the 20 most frequently expressed unigenes reflect a chondrocyte-related gene expression signature. Thus, our data establish a representative set of more than 2000 genes expressed in a chondrocytic cell line. This finding provides a framework for understanding cell growth and differentiation of chondrocytes and their metabolic function in the formation and remodeling of cartilage.

Kakar, S. S. (1999). "Molecular cloning, genomic organization, and identification of the promoter for the human pituitary tumor transforming gene (PTTG)." Gene **240**(2): 317.

http://www.sciencedirect.com/science/article/B6T39-3YXJ0JD-8/2/e4dc301b1ddca13ec0279a2df43a250d

Kapoor, M., L. Zhang, et al. (1995). "Synthesis and characterization of an infectious dengue virus type-2 RNA genome (New Guinea C strain)." Gene **162**(2): 175.

http://www.sciencedirect.com/science/article/B6T39-3Y6HGRC-1/2/cdbebc4cdccd27d8767969f8254bda22

Dengue virus type 2 (DEN-2), a member of the Flaviviridae family, has a positive-strand RNA genome, 10 723 nucleotides (nt) in length and encoding a single polyprotein precursor consisting of 3391 amino acids (aa). In order to construct a full-length cDNA clone, the viral genome was cloned into 5' (nt 1-2203 under the control of the T7 promoter (pT7)) and 3' (nt 2203-10723) constructs. A full-length DEN-2 cDNA under pT7 control was assembled in vitro after excising the two cDNA inserts from the 5' and 3' constructs, and joining them with T4 DNA ligase. The RNA produced by in vitro transcription of the cDNA using T7 RNA polymerase was infectious, as shown by transfection of permissive BHK-21 and Vero cells, and propagation of the virus particles released into the culture media. The virus particles stably maintained the conservative mutation introduced into the 5' construct, and the cells infected with the infectious RNA-derived virus synthesized virus-specific DEN-2 antigens, as shown by immunofluorescence and immunoprecipitations. The full-length infectious clone for DEN-2 should be useful for the study of molecular mechanisms involved in viral RNA replication and virus assembly.

Karnovsky, A. M., L. F. Gotow, et al. (2003). "A cluster of novel serotonin receptor 3-like genes on human chromosome 3." Gene **319**: 137.

http://www.sciencedirect.com/science/article/B6T39-49RCM38-4/2/5a6b733afc86e587f63982ffc21eaf2a

The ligand-gated ion channel family includes receptors for serotonin (5-hydroxytryptamine, 5-HT), acetylcholine, GABA, and glutamate. Drugs targeting subtypes of these receptors have proven useful for the treatment of various neuropsychiatric and neurological disorders. To identify new ligand-gated ion channels as potential therapeutic targets, drafts of human genome sequence were interrogated. Portions of four novel genes homologous to 5-HT3A and 5-HT3B receptors were identified within human sequence databases. We named the genes 5-HT3C1-5-HT3C4. Radiation hybrid (RH) mapping localized these genes to chromosome 3q27-28. All four genes shared similar intron-exon organizations and predicted protein secondary structure with 5-HT3A and 5-HT3B. Orthologous genes were detected by Southern blotting in several species including dog, cow, and chicken, but not in rodents, suggesting that these novel genes are not present in rodents or are very poorly conserved. Two of the novel genes are predicted to be pseudogenes, but two other genes are transcribed and spliced to form appropriate open reading frames. The 5-HT3C1 transcript is expressed almost exclusively in small intestine and colon, suggesting a possible role in the serotonin-responsiveness of the gut.

Katoh, T., B. Munkhbat, et al. (2005). "Genetic features of Mongolian ethnic groups revealed by Y-chromosomal analysis." Gene **346**: 63.

http://www.sciencedirect.com/science/article/B6T39-4FC449B-3/2/6aed8791ffeed54dda5b46ba8e0e8b00

About 20 ethnic groups reside in Mongolia. On the basis of genetic and anthropological studies, it is believed that Mongolians have played a pivotal role in the peopling of Central and East Asia.

However, the genetic relationships among these ethnic groups have remained obscure, as have their detailed relationships with adjacent populations. We analyzed 16 binary and 17 STR polymorphisms of human Y chromosome in 669 individuals from nine populations, including four indigenous ethnic groups in Mongolia (Khalkh, Uriankhai, Zakhchin, and Khoton). Among these four Mongolian populations, the Khalkh, Uriankhai, and Zakhchin populations showed relatively close genetic affinities to each other and to Siberian populations, while the Khoton population showed a closer relationship to Central Asian populations than to even the other Mongolian populations. These findings suggest that the major Mongolian ethnic groups have a close genetic affinity to populations in northern East Asia, although the genetic link between Mongolia and Central Asia is not negligible.

Kazuko, I., M. Kenji, et al. (1993). "Isolation of a yeast essential gene, COF1, that encodes a homologue of mammalian cofilin, a low-Mr actin-binding and depolymerizing protein." Gene **124**(1): 115.

http://www.sciencedirect.com/science/article/B6T39-47P8K4K-17V/2/c24c16c56ad7ee95859f03255da3b386

We have cloned a Saccharomyces cerevisiae gene (COF1) encoding a low-Mr actin-binding protein of 143 amino acid (aa) residues (yeast cofilin; Cof); its aa sequence is 35% identical to porcine Cof. The yeast recombinant Cof produced in Escherichia coli exhibited in vitro activities on actin filaments similar to those of mammalian and avian Cof. Gene disruption and tetrad analysis showed that gene COF1 is essential for yeast cell growth. Expression of the cDNA of porcine Cof or destrin (Des), the latter a Cof-related protein, complemented the cof 1 null allele in yeast cells.

Kersulyte, D., N. S. Akopyants, et al. (1998). "Novel sequence organization and insertion specificity of IS605 and IS606: chimaeric transposable elements of Helicobacter pylori." Gene **223**(1-2): 175.

http://www.sciencedirect.com/science/article/B6T39-3VCK52C-P/2/6be2bf9b4ded1b7ef99b1a5a1dde8e6c

IS605, an insertion sequence (IS) that is unusual in containing homologs of genes for the single putative transposases of two other unrelated IS elements (IS200 and IS1341), was found in nearly one-third of a set of 238 independent isolates of the gastric pathogen Helicobacter pylori. Hybridization and PCR tests indicated that any strain carrying one of these ORFs also carried the other, which implies that both ORFs are in the same unit of transposition. The IS605 ends and target sites for insertion were identified by sequencing eight preexisting insertions in strain NCTC11638, corresponding empty sites in other strains, and new transpositions in E. coli of an IS605 derivative marked with a selectable chloramphenicol-resistance gene. These tests showed that IS605 is also unusual in: (1) having unique, not inverted repeat, ends; (2) not duplicating (or deleting) target sequences during transposition; and (3) inserting with its left (IS200-homolog) end next to 5'-TTTAA or 5'-TTTAAC. IS605 was implicated in at least two genome rearrangements in strain NCTC11638. A second member of the IS605 family, called IS606 (25% amino acid identity to IS605 in inferred proteins) was found in one-third of 38 H. pylori strains tested, many of which did not carry IS605. The features of these two chimaeric IS elements are discussed in terms of possible transposition mechanisms, IS element evolution, and effects of IS elements on genome organization and evolution in the microbes that they inhabit.

Kersulyte, D., B. Rajendra Krishnan, et al. (1992). "Nonrandom orientation of transposon Tn5supF

insertions in phage lambda." Gene **114**(1): 91.

http://www.sciencedirect.com/science/article/B6T39-47F6XNG-G/2/6aa5e9b5ddfeb8fdb136e9776ab71755

Transposition of mini-transposon Tn5supF to phage [lambda] can be selected in two ways: (i) by plaque formation on a dnaB amber strain of Escherichia coli, which requires expression of the transposon-borne suppressor tRNA gene (supF) during lytic phage growth, or (ii) by lysogenization of a strain with amber mutations in tet and amp resistance genes, and selection of TcRApR (Sup+) transductant colonies. Tn5supF insertions in several [lambda] clones were isolated and mapped using a polymerase chain reaction (PCR) amplification method. Among insertions selected during lytic growth, more than 90% were oriented such that supF could be transcribed from an upstream [lambda] promoter. In contrast, half of those selected by transduction were in each orientation. These results indicate that Tn5supF insertion occurs with equal frequency in each orientation. However, Tn5supF insertion phages in which transcription from the [lambda] and supF promoters would collide tend to be lost when supF is selected during lytic growth. The tendency to recover Tn5supF insertions in only one orientation is useful in a transposon- and crossover-PCR-based method for preparing templates for DNA sequencing.

Keryanov, S. and K. L. Gardner (2002). "Physical mapping and characterization of the human Na,K-ATPase isoform, ATP1A4." Gene **292**(1-2): 151.

http://www.sciencedirect.com/science/article/B6T39-45SRKFF-2/2/6e3303edb1f8950e8b9dfd274bcf6928

Four isoforms of the catalytic [alpha] subunit of the Na,K-ATPase have been previously identified. We characterized and mapped a genomic copy of the human ATP1A4 isoform between D1S2707 and WI-9524, telomeric to a nearby isoform ATP1A2, and within a candidate region at 1q23 for familial hemiplegic migraine (FHM). Human ATP1A4 gene shares 84% identity with the mouse Atp1a4 gene, and both consist of 22 exons and 21 introns. The predicted polypeptide is 1029 amino acids and shares 82 and 79.8% identity, respectively, with human ATP1A2 and ATP1A1. ATP1A4 is larger than other isoforms and most divergent at the N-terminus. ATP1A4 and ATP1A2 are paralogous genes with the same number and organization of putative H-transmembrane domains, conserved exon-intron boundaries, and are found approximately 8.5 kb apart. Expression analysis of the ATP1A4 gene revealed a new major ~7.5 kb transcript in human skeletal muscle, with expression also shown in mouse muscle. Predictive analysis of promoter regions identified muscle specific regulatory elements for ATP1A4 and Atp1a4. Mutation analysis among eight affected individuals from a single large, highly penetrant FHM family was negative in ATP1A4 and ATP1A2 although multiple polymorphisms were identified.

Kim, S.-H., C. C. Titlow, et al. (2000). "An approach for preventing recombination-deletion of the 40-50 anti-digoxin antibody VH gene from the phage display vector pComb3." Gene **241**(1): 19.

http://www.sciencedirect.com/science/article/B6T39-3Y51FGV-3/2/372d8019a96489a578f3492a403b6228

Kim, S., I. Sohn, et al. (2004). "Hepatic gene expression profiles in a long-term high-fat diet-induced obesity mouse model." Gene **340**(1): 99.

http://www.sciencedirect.com/science/article/B6T39-4CXV04K-2/2/00931e3103427753b7c5542438c5e33d

To understand the molecular mechanisms underlying alterations in the pathophysiologic status of dietary obesity, we examined hepatic genes differentially expressed in a long-term high-fat intake-induced obesity mouse model. C57BL/6J male mice were fed with two kinds of diets for 12 weeks; a low-fat diet (LFD), a high-fat diet (HFD; n=8), and the expression levels of ~10,000 transcripts in liver tissues from the two groups were assessed using cDNA microarray analysis. Twelve-week feeding with the HFD resulted in significant increase in body weight, visceral fat accumulation and circulating cholesterol concentration, compared with the LFD group. The cDNA microarray analysis revealed marked differences in the expressions of 97 hepatic genes. These genes were categorized into seven groups:1. metabolism;2. defense, stress, and inflammation responses;3. signal transduction, apoptosis, and cell cycle;4. transcription regulation;5. protein synthesis and modification;6. transport; and7. cellular adhesion, cytoskeleton and trafficking.The expression of genes involved in fatty acid catabolism and ketone body synthesis, such as acyl-CoA oxidase1 (Acox1) and HMG-CoA lyase (Hmgcl), was significantly increased, and expression of genes involved in lipogenesis and cholesterol synthesis, such as acetyl-CoA synthetase2 (Acs2), fatty acid synthase (Fasn), and squalene epoxidase (Sqle), was drastically decreased in the HFD group. Interestingly, the genes implicated in defense and stress responses, such as glutathione S-transferases (GSTs) and heat shock proteins (Hsps), were also highly represented in the HFD group. Besides, a number of previously unappreciated regulatory molecules were changed by the HFD. These results revealed a transcriptional adaptation to long-term HFD and provided interesting information about the molecules involved in the development and maintenance of the obesity phenotype in vivo.

Kimura, E., K. Hidaka, et al. (2004). "Serine-arginine-rich nuclear protein Luc7l regulates myogenesis in mice." Gene **341**: 41.

http://www.sciencedirect.com/science/article/B6T39-4DBKH0V-1/2/9192c5e90193e4791beacd51a89949ed

Using a gene trap technique, we identified a murine homologue of the yeast LUC7-like gene (Luc7l), which is a serine-arginine-rich protein (SR protein) that localizes in the nucleus through its arginine-serine-rich domain (RS domain) at the C-terminus and shows a speckled distribution pattern. Although its transcripts are widely expressed in embryos and adults, they are rarely detected in adult skeletal muscle, and Luc7l expression was found to be negatively regulated during the course of development of limb skeletal muscle, as well as during in vitro differentiation of the myoblast cell lines Sol8 and C2C12. We also demonstrated that forced expression of Luc7l protein inhibited myogenesis in vitro. Based on our results, Luc7l is thought to play an important role in the regulation of muscle differentiation.

Kjellman, C., H.-O. Sjogren, et al. (1995). "The Y chromosome: a graveyard for endogenous retroviruses." Gene **161**(2): 163.

http://www.sciencedirect.com/science/article/B6T39-3Y6HGSR-16/2/cafcf0cd97cb9dc1ca90ad26c65f6a6d

We have isolated 20 different human endogenous retroviruses (ERV) related to ERV3, Hsrirt and Humer 4-1. Phylogeny and the presence of these ERV among different primates were determined by computer and Southern blot analyses. Preferential localization of ERV to the human, chimpanzee and orangutan Y chromosomes among the low-copy-number ERV is demonstrated.

The reason for this accumulation of ERV on the strongly heterochromatic Y chromosome is probably mediated by (i) the absence of recombination of the Y chromosome that makes it more difficult for sequences to be lost, and (ii) integration of retroviruses in heterochromatic regions is less harmful to the organism. If ERV located on the Y chromosome are transcribed and translated to peptides, such peptides could be potential HY-antigens.

Kloek, A. P., D. R. Sherman, et al. (1993). "Novel gene structure and evolutionary context of Caenorhabditis elegans globin." Gene **129**(2): 215.

http://www.sciencedirect.com/science/article/B6T39-47PH60Y-114/2/4e8271277becb51267f3c7abaab68b88

Animal and plant globin-encoding genes (Glo) contain two introns in strictly conserved positions. Plant Glo genes possess an additional, centrally located intron. We have determined the cDNA sequence and gene structure of a putative Glo gene from the free-living nematode, Caenorhabditis elegans. The gene encodes a one-domain globin with a single intron, corresponding to the central intron of plant Glo genes. The two introns common to virtually all animal and plant Glo genes are missing. Comparison with the related organisms Trichostrongylus colubriformis, Ascaris suum and Pseudoterranova decipiens, provides evidence of gene duplication, intron loss, and functional divergence within the Glo genes of the nematode phylum. It is now apparent that differential intron loss during evolution has generated Glo genes with a panoply of exon/intron permutations.

Knobloch, J., R. Winnen, et al. (2002). "A novel Syk-family tyrosine kinase from Schistosoma mansoni which is preferentially transcribed in reproductive organs." Gene **294**(1-2): 87.

http://www.sciencedirect.com/science/article/B6T39-46RN5P2-B/2/c47f1db1c914f5bb131ee6640fe4d356

The complete coding deoxyribonucleic acid for a novel tyrosine kinase (TK) of the human parasite Schistosoma mansoni has been cloned and characterized. The molecule was designated TK4. The sequence predicts a translation product of about 140 kDa containing two Src homology 2 domains and a tyrosine kinase domain. Data base analyses indicate that TK4 belongs to the Syk family of TKs which has not been identified in schistosomes or other Acoelomata yet. The presence of a member of the Syk family in this phylum supports previous findings demonstrating that TK subclasses were established early in evolution. Although Northern blot and reverse transcription polymerase chain reaction analyses show transcription of TK4 in larval stages and adult schistosomes of both genders, TK4 is more abundantly transcribed in males. In situ hybridization data demonstrate the gender-independent occurrence of TK4 transcripts in parenchymatic cells. Significant signals were detected in the oocytes of the female and in the spermatocytes of the male suggesting that TK4, among other functions, may play a role in germ cell development. This is an unexpected finding considering that Syk-family TKs of invertebrates and vertebrates described so far are not involved in the differentiation of the gonads.

Kobayashi, K., M. Ouchida, et al. (2002). "Reduced expression of the REIC/Dkk-3 gene by promoter-hypermethylation in human tumor cells." Gene **282**(1-2): 151.

http://www.sciencedirect.com/science/article/B6T39-44KWSWR-B/2/71fa98e0754bff4c39f300d659837a29

The human REIC gene is a recently found mortalization-related gene and a candidate tumor suppressor gene expression of which is largely attenuated in many immortalized and tumor-derived cell lines (Biochem. Biophys. Res. Commun. 268 (2000) 20-24). To gain insight into the mechanisms of the down-regulation, we investigated the genomic structure and promoter activity of the human REIC gene. The gene, identical with the DKK-3 gene, resides on chromosome 11p15.1, consists of nine exons, and has two promoters. Methylation in the main promoter region was detected in 11 out of 21 cell lines tested (52%) derived from a variety of human tumors, in which the expression of the REIC gene was decreased. In ten of these 11 cell lines the minor promoter was also methylated. Similarly, the REIC gene expression was decreased in 14 of 24 fresh non-small cell lung cancer specimens (58%) compared to that in corresponding non-cancerous tissue, though allelic loss and tumor-specific mutation were rare. Of these 14 tumors, at least five tumors exhibited heavy methylation of the REIC promoter region. These results indicate that the down-regulation of the REIC gene expression is ascribed to the aberrant promoter hyper-methylation at least in a subset of human tumors. The expression was restored upon treatment of SQ5 cells with 5-aza-deoxycytidine, confirming DNA methylation as the mode of downregulation. A notable single nucleotide polymorphism in the coding region (cSNP) with an amino acid substitution of glycine (GGG) to arginine (AGG) was found at codon 335 of the REIC gene. However, the distribution of the cSNP showed no significant difference between lung cancer patients and healthy population.

Kondo, H., K. Morinaga, et al. (2005). "Characterization of the pufferfish Takifugu rubripes apolipoprotein multigene family." Gene **346**: 257.

http://www.sciencedirect.com/science/article/B6T39-4FD9SYC-1/2/64481ef8ed745d9ba8ceac0728d55b8a

We have characterized the apolipoprotein multigene family of the pufferfish Takifugu rubripes. The pufferfish mainly contains 28-kDa, 27-kDa, and 14-kDa apolipoproteins in its plasma and was designated apo-28 kDa, apo-27 kDa, and apo-14 kDa, respectively. N-terminal amino acid sequencing revealed that pufferfish apo-28 kDa and apo-27 kDa have an identical amino acid sequence except an additional propeptide in the former; and both are homologues of apoA-I from other animals. The sequence of pufferfish apo-14 kDa is homologous to that of eel apo-14 kDa previously reported, both being apparently specific to fish. In silico screening, using the publicly available Fugu genome database confirmed the pufferfish apoA-I and apo-14 kDa genes. The database further contained the genes encoding four types of apoA-IV, one apoC-II and two types of apoE. Thus, pufferfish contains nine genes encoding apolipoprotein multigene family. Two apoA-IV and one apoE genes were tandemly arrayed and located on one scaffold. Thus two sets of these genes formed two gene clusters. The apoC-II and apo-14 kDa genes are also located on a single scaffold. apoA-I and apo-14 kDa gene transcripts were mainly expressed in liver and less abundantly in brain. The transcripts of the former gene were also observed in intestine. In contrast, the transcripts encoding four apoA-IVs, one apoC-II, and two apoEs were mainly expressed in intestine. These structural details of pufferfish apolipoproteins and tissue distribution of their gene transcripts provide a novel evidence for better understanding of evolutionary relationships of apolipoprotein multigene family.

Kono, T., T. Sakai, et al. (2003). "Molecular cloning and expression analysis of a novel caspase recruitment domain protein (CARD) in common carp Cyprinus carpio L." Gene **309**(1): 57.

http://www.sciencedirect.com/science/article/B6T39-48B5V01-3/2/5e3005ed81a62f19d2ce7420e7659cef

A novel caspase recruitment domain protein (CARD) was isolated from common carp Cyprinus

carpio L. by expressed sequence tag analysis. This gene consist of a 2016 bp open reading frame and untranslated regions, which is putatively translated to a protein of 535 amino acid residues. The gene harbors domains (CARD and Coiled-coil domain), which are conserved in proteins of CARD family. The CARD domain have carp was similar to human CARD9 with 72.4% identity. Expression analysis revealed that CARD gene of carp (carp-CARD) expressed in normal tissues of head kidney, spleen, liver, heart and brain. Here we demonstrated that the expression of carp-CARD increased by cortisol treatment in all the tissues and had a high and long lasting expression in cortisol treated spleen.

Korth, M. J., C. N. Lyons, et al. (1996). "Cloning, expression, and cellular localization of the oncogenic 58-kDa inhibitor of the RNA-activated human and mouse protein kinase." Gene **170**(2): 181.

http://www.sciencedirect.com/science/article/B6T39-3VXNK10-4/2/a2752673af9335c6220d196b875e45a8

The 58-kDa inhibitor of the interferon-induced double-stranded RNA-activated protein kinase (PKR) is a cellular protein that is activated during influenza virus infection to down-regulate the activity of PKR. This study was initiated to further our understanding of the inhibitor which, when overproduced, has the capacity to malignantly transform cells. We report here the isolation and characterization of cDNA clones encoding the inhibitor, designated p58, from human HeLa and mouse NIH 3T3 cells. The human and mouse p58 cDNAs were 6.5 and 1.6 kb in length, respectively. Surprisingly, the deduced amino acid sequences of the human and mouse p58 were 96% identical, indicating a remarkably high degree of conservation between species. An examination of p58 mRNA expression in human tissues revealed a 6.5-kb transcript in all tissues examined, with a particularly high level of expression present in the pancreas and liver, and also in certain leukemic cell lines. Similarly, p58 expression was detected in all mouse tissues examined, with the highest level of expression found in liver. In contrast to human tissues, three p58 transcripts of approximately 1.7, 3.3 and 5.4 kb were observed in mouse tissues, suggesting that p58 expression may be regulated differently in human and mouse cells. Western blot analysis of subcellular fractions and indirect immunofluorescence analysis of intact cells revealed that p58 was found predominantly in the cytoplasm, consistent with its function as an inhibitor of PKR, which is also a predominantly cytoplasmic protein.

Kratzschmar, J., B. Haendler, et al. (1993). "Bovine urokinase-type plasminogen activator and its receptor: Cloning and induction by retinoic acid." Gene **125**(2): 177.

http://www.sciencedirect.com/science/article/B6T39-47PNYB9-XP/2/9c58f127c8a155d1c2a0c52245cec83a

Full-length cDNAs encoding bovine urokinase-type plasminogen activator (u-PA) and urokinase receptor (u-PAR) were cloned from an aortic endothelial cell cDNA library using PCR-amplified cDNA fragments as probes. Bovine u-PA amino acid identity ranges from 79.5 to 70.9% when compared to its pig, human, baboon and mouse analogues, while bovine u-PAR is 61.8 and 59.6% identical to its human and mouse counterparts, respectively. All Cys residues previously found in mature u-PA and u-PAR from these different species are also conserved in the bovine molecules. Bovine u-PA and its cell-surface receptor display one and six potential sites of N-linked glycosylation, respectively. Northern blot hybridization demonstrated a moderate induction of u-PA and u-PAR mRNA in bovine aortic endothelial cells after treatment with 10 nM and 1 [mu]M retinoic acid for 8 hours.

Kricker, M. C. and K. R. Tindall (1989). "Direct sequencing of bacteriophage T4 DNA with a thermostable DNA polymerase." Gene **85**(1): 199.

http://www.sciencedirect.com/science/article/B6T39-47PH5MW-W0/2/edf8c1fa996a861cdd05dfda94a39b5f

We present a simple and convenient protocol for the direct sequencing of bacteriophage T4 genomic DNA. The method utilizes the thermostable DNA polymerase from Thermus aquaticus (Taq) and 32P-end-labeled oligodeoxyribonucleotide primers to produce extension products that allow the analysis of at least 200 nucleotides (nt) on a single sequencing gel. Single-nt changes in the template were easily detectable following an overnight exposure of the autoradiograms. Comparison of sequences from fully modified T4 DNA containing glucosylated hydroxymethyldeoxycytosine or from templates containing cytosine showed little difference in sequence clarity. These techniques considerably simplify the molecular analysis of T-even bacteriophages and should be compatible with automated sequencing methods which employ 5'-end-labeled primers.

Kudia, B. and A. Nicolas (1992). "A multisite integrative cassette for the yeast Saccharomyces cerevisiae." Gene **119**(1): 49.

http://www.sciencedirect.com/science/article/B6T39-47P8FWJ-5X/2/0f1f62a2b93e5c4c617d7bb121c14596

We have developed a cassette for the integration of cloned DNA sequences at multiple sites in the Saccharomyces cerevisiae genome, taking advantage of the naturally repeated sigma sequences. This cassette contains one engineered sigma element which allows the targeting of an embedded gene at different genomic sigma elements by gene replacement. Two yeast genes, ARG4 and URA3, were thus integrated in the absence of any bacterial sequences, individually or sequentially on twelve chromosomes. Consequently, these studies led to the genetical tagging of individual members of the sigma family.

Kullik, I., R. Jenni, et al. (1998). "Sequence of the putative alanine racemase operon in Staphylococcus aureus: insertional interruption of this operon reduces -alanine substitution of lipoteichoic acid and autolysis." Gene **219**(1-2): 9.

http://www.sciencedirect.com/science/article/B6T39-3WBG1Y1-M/2/f66b3a566d41b5c158a6c91e6b68104a

A gene cluster comprising the alanine racemase gene alr was identified 5' to the sigB operon in Staphylococcus aureus. It is flanked upstream by four ORFs of which one shows similarity to the dpj gene of Escherichia coli, and downstream by two ORFs of which the last shows similarity to the E. coli pemK gene. Preliminary data suggest that the seven ORFs orf1-orf2-orf3-dpj-alr-orf6-pemK may form an operon. Disruption of the proposed operon by insertional mutagenesis leads to a drastic loss in the -alanine (-Ala) substitution of lipoteichoic acid and to delayed autolysis, without affecting the -Ala substitution of the wall teichoic acid.

Kurdyukov, S. G., Y. B. Lebedev, et al. (2001). "Full-sized HERV-K (HML-2) human endogenous retroviral LTR sequences on human chromosome 21: map locations and evolutionary history." Gene **273**(1): 51.

http://www.sciencedirect.com/science/article/B6T39-43KGNS7-6/2/a4e2bef4c7f8156d927fcd256a40bde4

One of the evolutionary mechanisms for acquisition of novel functional sequences can be domestication of exogenous retroviruses that have been integrated into the germ line. The whole genome mapping of such elements in various species could reveal differences in positions of the retroviral integration and suggest possible roles of these differences in speciation. Here, we describe the number, locations and sequence features of the human endogenous retrovirus HERV-K (HML-2) long terminal repeat (LTR) sequences on human chromosome 21. We show that their distribution along the chromosome is not only non-random but also roughly correlated with the gene density. Amplification of orthologous LTR sites from a number of primate genomes produced patterns of presence and absence for each LTR sequence and allowed determination of the phylogenetic ages and evolutionary order of appearance of individual LTRs. The identity level and phylogenetic age of the LTRs did not correlate with their map locations. Thus, despite the non-random distribution of LTRs, they have apparently been inserted randomly into the chromosome relative to each other. As evidenced in previous studies of chromosomes 19 and 22, this is a characteristic of HERV-K integration.

Kurotaki, N., N. Harada, et al. (2001). "Molecular characterization of NSD1, a human homologue of the mouse Nsd1 gene." Gene **279**(2): 197.

http://www.sciencedirect.com/science/article/B6T39-44CNR71-1/2/29d28ef57b9185f13b5a9652ffe454a8

NR-binding SET-domain-containing protein (NSD1) is a mouse nuclear protein containing su(var)3-9, enhancer-of-zeste, trithorax (SET), proline-tryptophan-tryptophan-proline (PWWP) and plant homeodomain protein (PHD)-finger domains (Huang et al., EMBO J. 17 (1998) 3398). This protein also has two other distinct nuclear receptor (NR)-interaction domains, called NID-L and NID+L, and acts as both a NR corepressor and a coactivator by interacting directly with the ligand-binding domain of several NRs. Thus, NSD1 is a bifunctional, transcriptional, intermediary factor. We isolated the human homologue (NSD1) of the mouse NSD1 gene (Nsd1), mapped it to human chromosome 5q35, and characterized its genomic structure. NSD1 consists of at least 23 exons. Its cDNA is 8552 bp long, has an 8088 bp open reading frame, contains at least six functional domains (SET, PWWP-I, PWWP-II, PHD-I, PHD-II, and PHD-III) and ten putative nuclear localization signals, and encodes 2696 amino acids. NSD1 shows 86% identity with the mouse Nsd1 at the nucleotide level, and 83% at the amino acid level. NSD1 is expressed in the fetal/adult brain, kidney, skeletal muscle, spleen, and the thymus, and faintly in the lung. Two different transcripts (9.0 and 10.0 kb) were consistently observed in various fetal and adult tissues examined. These findings favor the character of NSD1 as a nucleus-localized, basic transcriptional factor and also a bifunctional transcriptional regulator, such as that of the mouse Nsd1. It remains to be investigated whether mutations of NSD1 lead to a specific phenotype in man.

Lareu, M. V., M. del Carmen Pestoni, et al. (1996). "Sequence variation of a hypervariable short tandem repeat at the D12S391 locus." Gene **182**(1-2): 151.

http://www.sciencedirect.com/science/article/B6T39-3W2XNW0-20/2/831205fde5b35982452049c8787d61d0

A short tandem repeat (STR) in the D12S391 locus was sequenced in more than 40 individuals. Twenty different alleles were found and these could be grouped into 12 allelic classes in

accordance with the total number of repeats. This is a compound STR consisting of blocks of (AGAT) and (AGAC) repeats with basic sequence structure (AGAT)8-17(AGAC)6-10(AGAT)0-1. Whereas smaller alleles (15-18) have variation limited to the (AGAT) unit, in larger alleles the complexity is greater with variation in the number of tandem arrays in the two motifs (AGAT) and (AGAC). Population data showed that this is a highly polymorphic STR with a heterozygosity of more than 0.9. This fact, together with its simple structure, makes this STR an interesting DNA polymorphism for forensic and genetic purposes.

Larochelle, S. and B. Suter (1995). "The Drosophila melanogaster homolog of the mammalian MAPK-activated protein kinase-2 (MAPKAPK-2) lacks a proline-rich N terminus." Gene **163**(2): 209.

http://www.sciencedirect.com/science/article/B6T39-3Y6HKBK-8K/2/5885fc6228d0f1ddd8f60f9465c432e7

Recently, a mammalian kinase cascade was discovered that is triggered by stress and heat shock, and leads to the stimulation of mitogen-activated protein kinase (MAPK)-activated protein kinase-2 (MAPKAPK-2). Surprisingly, this process turns out to be independent of the classical MAPK. The stress-induced activation of MAPKAPK-2, in turn, results in the phosphorylation of small heat-shock proteins (Hsp). We have isolated a Drosophila melanogaster (Dm) cDNA encoding a polypeptide that has extensive sequence similarity to the mammalian MAPKAPK-2. As in mammalian MAPKAPK-2, the Dm MAPKAPK-2 possesses a MAPK phosphorylation site and a nuclear targeting sequence located C-terminal to the catalytic domain. However, in contrast to its mammalian counterpart, it lacks the Pro-rich N-terminal region proposed to form Src-homology domain 3 (SH3) binding domains. A 2.4-kb MAPKAPK-2 message is expressed throughout development, while two shorter transcripts of 2.3 and 1.8 kb appear to be specifically expressed in the germline. The 1.8-kb transcript results from the usage of an atypical germline-specific polyadenylation signal (AATATA) located early within the 3' untranslated region. Dm MAPKAPK-2 is located at cytological position 5D in the Dm genome.

Laten, H. M. and R. O. Morris (1993). "SIRE-1, a long interspersed repetitive DNA element from soybean with weak sequence similarity to retrotransposons: initial characterization and partial sequence." Gene **134**(2): 153.

http://www.sciencedirect.com/science/article/B6T39-47PH62P-11T/2/84ec6f9b6ed73a3e012492bad6f1576e

Gm776 is a 776-bp subregion of a member of an interspersed family of relatively homogeneous repetitive DNA elements from soybean (Glycine max). The fragment was originally amplified from soybean DNA by the polymerase chain reaction using a single 22-nucleotide primer, and consequently terminates in an inverted repeat. The elements defined by Gm776 are at least 10.6 kb in length and constitute a family of 500-800 members per haploid genome. The family has been designated SIRE-1 (soybean interspersed repetitive element 1). Overlapping regions of Gm776 exhibit suggestive DNA sequence similarity to Tal and Tyl, copia-like retrotransposons from Arabidopsis thaliana and Saccharomyces cerevisiae, respectively. However, there are no similarities at the amino acid level, and the regions of similarity are not functionally related.

Leal, A., S. Endele, et al. (2003). "A novel myosin heavy chain gene in human chromosome 19q13.3." Gene **312**: 165.

http://www.sciencedirect.com/science/article/B6T39-492090F-6/2/5685e47e18d303cd5e8bd15c39c01f32

A human myosin heavy chain gene was identified in chromosome 19q13 by computational sequence analysis, RT-PCR and DNA sequencing of the cDNA. The complete cDNA has a length of 6786 bp and comprises 41 exons (40 coding) included in 108 kb of genomic sequence. Alternative splicing variants were also identified. The gene is expressed in a multitude of tissues, but mainly in small intestine, colon and skeletal muscle. The putative protein (228 kDa) carries the common myosin domains and presents high homology with the non-muscle myosin heavy chains (MYH9 and MYH10) as well as the smooth muscle myosin heavy chain MYH11. Nevertheless, phylogenetic analysis indicated that these homologous proteins are more related among themselves than to MYH14, suggesting that possibly this myosin heavy chain should be classified in a new myosin-subfamily.

Lebedev, Y. B., O. S. Belonovitch, et al. (2000). "Differences in HERV-K LTR insertions in orthologous loci of humans and great apes." Gene **247**(1-2): 265.

http://www.sciencedirect.com/science/article/B6T39-402KBCD-Y/2/c15a78cdd4f2999d3a8f374f8c8f7d69

Lee, J.-S., M. Miya, et al. (2001). "The complete DNA sequence of the mitochondrial genome of the self-fertilizing fish Rivulus marmoratus (Cyprinodontiformes, Rivulidae) and the first description of duplication of a control region in fish." Gene **280**(1-2): 1.

http://www.sciencedirect.com/science/article/B6T39-44GM1PG-6/2/300631cf6406fd0b492d2d89d0dedbc0

We isolated Rivulus marmoratus mitochondrial DNA by long-polymerase chain reaction with conserved primers, and sequenced it with 36 sets of internal conserved primers, which were designed from the extensive sequence similarities of mitochondrial DNA from several fish species. The R. marmoratus mitochondrial DNA has 17,329 bp with a conserved structural organization compared to those of other fish. Rivulus marmoratus mitochondrial DNA also has two nearly identical control regions. The basic characteristics of the R. marmoratus mitochondrial genome are discussed.

Leonard, P. M., C. M. Adema, et al. (2001). "Structure of two FREP genes that combine IgSF and fibrinogen domains, with comments on diversity of the FREP gene family in the snail Biomphalaria glabrata." Gene **269**(1-2): 155.

http://www.sciencedirect.com/science/article/B6T39-433NSPG-J/2/000e0154282ad06cfe975923f6288f24

Upon exposure to infection with digenetic trematodes such as Echinostoma paraensei, the freshwater snail Biomphalaria glabrata produces increased quantities of hemolymph lectins, some of which are unique polypeptides containing both immunoglobulin superfamily (IgSF) and fibrinogen domains. These unusual lectins have been termed fibrinogen-related proteins (FREPs), and recognize and precipitate digenean antigens. We here report 11 distinct FREP-encoding sequences from B. glabrata, and provide the complete genomic sequence for two of the most frequently recovered FREPs. The unique juxtaposition of IgSF and fibrinogen domains,

previously known only from incomplete cDNAs, is confirmed. Sequences corresponding to known peptides derived from FREPs from hemolymph were found in one of these genes. Both genes contain four exons, the first encodes a putative signal peptide, the second and third a portion of an IgSF-type loop, and the fourth a fibrinogen domain. Cysteines, postulated to form an intrachain loop, are present in the IgSF domain and are separated from one another by 78 or 79 residues. The IgSF sequences most closely resemble V (variable)-type Ig domains, based on canonical and hydrophobic residues and predicted secondary structure. Some minor differences in genomic fragments isolated for each of the two sequences were noted and may represent allelic variants. The results may be of relevance in understanding the role of B. glabrata in transmission of Schistosoma mansoni, a digenean parasite that infects nearly 100 million people in the tropics.

Li, B., T. Chang, et al. (2000). "Identification of mRNAs expressed in tumor-infiltrating lymphocytes by a strategy for rapid and high throughput screening." Gene **255**(2): 273.

http://www.sciencedirect.com/science/article/B6T39-41C2P1P-G/2/7cb9860a96f6ca7e581e7cbe0ef64dd3

Li, D., O. Gonzalez, et al. (2000). "Human protein tyrosine phosphatase-like gene: expression profile, genomic structure, and mutation analysis in families with ARVD." Gene **256**(1-2): 237.

http://www.sciencedirect.com/science/article/B6T39-41H3KVD-W/2/fab67b7301ec46ec5b0bed1c0e6ae831

Lim, K. Y., C.-S. Hong, et al. (2000). "cDNA cloning and characterization of human cardiac junctin." Gene **255**(1): 35.

http://www.sciencedirect.com/science/article/B6T39-413KWNH-16/2/4f272766b6ebf6dc8a8f55cab5a7ea5d

Lindas, A.-C. and B. Tomkinson (2005). "Identification and characterization of the promoter for the gene encoding human tripeptidyl-peptidase II." Gene **345**(2): 249.

http://www.sciencedirect.com/science/article/B6T39-4F60WVW-3/2/78b10ee55a8f1a9e22a184b6287ade32

Tripeptidyl-peptidase II (TPP II) is a ubiquitously expressed exopeptidase. The expression of this enzyme is increased, e.g. in some tumor cells, but the regulation of the expression of the gene has not been investigated previously. The gene encoding human TPP II (TPP2) is 82 kb and consists of 30 exons. An 8 kb NcoI fragment covering the 5'-flanking region of the TPP2 gene, including the initiation codon, was cloned into a luciferase-containing reporter vector. Human embryonic kidney cells (HEK-293 cells) and murine fibroblasts (NIH3T3 cells) were transiently transfected with the construct. Through sequential deletions and analysis of short PCR-fragments, the promoter could be localized to a 215 bp fragment upstream of the initiation codon. This region is GC-rich, lacks a TATA-box and contains two inverted CCAAT-boxes and a GC-box. Electrophoretic mobility shift assays showed that nuclear proteins bind to the promoter fragment. The 85 bp 5'-end of the promoter fragment is essential for transcriptional activation. Out of this a 44 bp fragment suffices to compete with binding of nuclear proteins to the 215 bp fragment.

Supershift assays demonstrated that the CCAAT-binding factor (CBF; NF-Y) is involved in the formation of a complex with the 215 bp fragment. Although Sp1 binds to the promoter fragment in vitro, it was found to bind to the 3'-end of the 215 bp fragment which is not essential for transcription. The potential role of Sp1 in transcription of TPP2 therefore remains to be established.

Lindgren, R.-M., J. Zhao, et al. (2005). "Molecular cloning and characterization of two novel truncated isoforms of human Na+/Ca2+ exchanger 3, expressed in fetal brain." Gene **348**: 143.

http://www.sciencedirect.com/science/article/B6T39-4FNW4N5-4/2/249ee2c343f2c9679b4683352edaec74

The human gene encoding the Na+/Ca2+ exchanger family member 3 (NCX3) undergoes extensive alternative splicing, with four variants previously identified. In this study, we report two novel alternative transcripts encoding two N-terminally truncated NCX3 proteins specifically expressed in human fetal brain. The identified transcripts, designated NCX3-tN.1 and NCX3-tN.2, are approximately 2.8 kb and 2.9 kb, respectively. The open reading frames (ORFs) are predicted to encode separately a 284 and a 298 amino acid (aa) polypeptide. Sequence analysis and bioinformatics reveal that NCX3-tN.1 and NCX3-tN.2 are the result of alternative splicing of the NCX3 gene. They have their own potential start codons and unique 5' untranslated regions (UTRs) that are different from those of the known NCX3 variants. The variants include a part of intron 2 of the original gene organization as their first exon (exon "a") at the 5' end of the novel transcripts. NCX3-tN.2 consists of six exons including exon "a" and exons 4, 6, 7, 8 and 9 of NCX3, while NCX3-tN.1 lacks exon 4, but is otherwise similar to NCX3-tN.2. Expression studies show that both variants can be translated into protein and NCX3-tN.1 seems more efficiently translated. Based on their structural features, NCX3-tN.1 and NCX3-tN.2 proteins are potentially involved in regulation of Na+/Ca2+ homeostasis.

Lindqvist, A., P. Rouet, et al. (1999). "The alpha1-microglobulin/bikunin gene: characterization in mouse and evolution." Gene **234**(2): 329.

http://www.sciencedirect.com/science/article/B6T39-3WWDH80-J/2/b147081af8a90edab0e4c713bbb51e76

Liu, C.-Y., H. Arar, et al. (2000). "Identification of a 3.2 kb 5'-flanking region of the murine keratocan gene that directs [beta]-galactosidase expression in the adult corneal stroma of transgenic mice." Gene **250**(1-2): 85.

http://www.sciencedirect.com/science/article/B6T39-40GHS13-9/2/56f0728222cae7932c13d3972e359baa

Liu, M., R. Reimschuessel, et al. (2002). "Molecular cloning of the fish interferon stimulated gene, 15 kDa (ISG15) orthologue: a ubiquitin-like gene induced by nephrotoxic damage." Gene **298**(2): 129.

http://www.sciencedirect.com/science/article/B6T39-473HYTB-3/2/f4d969bbc5a261dc4aabd60ef7758d84

In mammals, the response to nephrotoxicant-induced renal injury is limited to repair of the proximal tubule by surviving epithelial cells. In contrast, bony fish are capable of both repair, and de novo production of nephrons in response to renal damage. Importantly, toxicant-induced nephron neogenesis in goldfish (Carassius auratus) parallels nephron development in the mammalian embryo, providing a vertebrate model for kidney development. We utilized this model system to identify genes induced by the renal toxicant, gentamicin, that may function in nephron neogenesis. A novel ubiquitin-like (UBL) gene, 40.1, was identified by differential display analysis of control and gentamicin-treated goldfish kidney. 40.1 was induced dramatically 3-7 days following a sublethal dose of gentamicin, and returned to basal level by 14 days post-treatment. The induction of 40.1 coincided with early renal injury in the proximal tubules of gentamicin-injected fish; however, expression was not restricted to the kidney, suggesting that 40.1 induction may be a more general response to cell injury. Sequence analysis revealed that 40.1 contains tandem UBL domains, and shares homology with ISG15, a 15 kD interferon-(IFN) stimulated UBL found in mammals. Analysis of the genome database for the pufferfish, Fugu rubrides, identified a goldfish ISG15 (gfISG15) homologue with an IFN-stimulated response element in the promoter region, providing further evidence that gfISG15 is the true teleost ISG15 orthologue. Zebrafish and catfish ISG15 genes were subsequently identified by sequence analysis. Consistent with its predicted function as a UBL, gfISG15 formed conjugates with cellular proteins in vitro and in transient transfections. Similar to the induction of mammalian ISG15 by microbial challenge, gfISG15 was induced in the spleen of mycobacteria-infected fish. These studies identified the first teleost ISG15 orthologue. The induction of gfISG15 as an early genetic event in response to a renal toxicant, and its conserved, stress-associated, expression in higher vertebrates suggests that ISG15 is an important component of the host response to diverse stress stimuli.

Liu, N., X. Wang, et al. (2000). "Developmentally regulated expression of two transcripts for heme oxygenase-2 with a first exon unique to rat testis: control by corticosterone of the oxygenase protein expression." Gene **241**(1): 175.

http://www.sciencedirect.com/science/article/B6T39-3Y51FGV-R/2/bdbe94f819ad11e163b71fafdf0e90a6

Liu, Q., C. Bai, et al. (1998). "Uncoupling protein-3: a muscle-specific gene upregulated by leptin in ob/ob mice." Gene **207**(1): 1.

http://www.sciencedirect.com/science/article/B6T39-3S1PY3W-1/2/232163a05a1599ebf9b5146790802ac4

We identified and partially characterized another member of the uncoupling protein termed UCP3. Human and mouse UCP3 protein sequences are 86% identical to each other, and 73% and 59% identical to UCP2 and UCP1, respectively. Expression of human UCP3 in yeast resulted in a drastic decrease of mitochondria membrane potential. Northern analysis showed that UCP3 was highly expressed in skeletal muscle in human, rat, and mouse. Mapping of UCP3 placed it to the same chromosomal region of UCP2 in both human and mouse, a region that is linked to obesity and hyperinsulinemia. Furthermore, adenovirus-mediated leptin expression in obese ob/ob mice led to increased expression of UCP3 in skeletal muscle. The data indicate that UCP3 encodes a muscle-specific uncoupling protein that may play an important role in the regulation of energy expenditure and development of obesity.

Lopez, C. C., I. Kamnert, et al. (1999). "Interspersed DNA element restricted to a specific group of

telomeres in the dipteran Chironomus pallidivittatus." Gene **233**(1-2): 249.

http://www.sciencedirect.com/science/article/B6T39-3WRB6HH-10/2/6e2d8a2025793ff1f679ac8783b23d1b

Luckhart, S. and R. Rosenberg (1999). "Gene structure and polymorphism of an invertebrate nitric oxide synthase gene." Gene **232**(1): 25.

http://www.sciencedirect.com/science/article/B6T39-3XJKJSF-3/2/8fa8eda0bd1ca3f05588fad161edfd8f

Lutz, G. J., S. Razzaghi, et al. (2000). "Cloning and characterization of the S1 domain of four myosin isoforms from functionally divergent fiber types in adult Rana pipiens skeletal muscle." Gene **250**(1-2): 97.

http://www.sciencedirect.com/science/article/B6T39-40GHS13-B/2/44b8524cbf70698ede7b9772aac2ff16

Lycan, D. E., K. A. Stafford, et al. (1996). "A new Saccharomyces cerevisiae ankyrin repeat-encoding gene required for a normal rate of cell proliferation." Gene **171**(1): 33.

http://www.sciencedirect.com/science/article/B6T39-3W4972K-12/2/00cfd0c7a37958d628731e72eff2516c

We have characterized a new ankyrin (ANK) repeat-containing Saccharomyces cerevisiae gene, YAR1, located between the HSP82 and SUI3 genes on chromosome XVI. YAR1 encodes a 200-amino-acid (aa) protein with two ANK repeat motifs and an acidic C terminus rich in PEST-like sequences. The Yarl ANK repeats are most similar to the conserved ANK repeats in the yeast cell cycle transcription factor, Swi6. We show that YAR1 is transcribed as an 800-nucleotide (nt) poly(A)+mRNA from a promoter lacking a consensus TATA sequence. YAR1 is transcribed in both haploid and diploid cells, and in haploid cells arrested in G1 with a-factor or in S phase with hydroxyurea. YAR1 shares an intergenic region with HSP82, and while HSP82 transcription is induced 15-fold by heat shock, transcription of YAR1 is transiently repressed by heat shock. We show that YAR1 is not an essential gene, but that haploid cells bearing a yarl deletion grow significantly more slowly than do isogenic wild-type cells, especially at low temperature.

Maeda, A., S. Hongo, et al. (2004). "Genomic organization, expression, and comparative analysis of noncoding region of the rat Ndrg4 gene." Gene **324**: 149.

http://www.sciencedirect.com/science/article/B6T39-4B290CR-3/2/20702f2fe1252ebf69987794e2895864

Rat Ndrg4 is a member of the NDRG gene family and has been suggested to relate to brain development. The structure of the rat Ndrg4 gene was studied to understand the mechanism for the expression of multiple forms of Ndrg4 protein, which were revealed in the brain. Subcloning and DNA sequencing analysis of a bacterial artificial chromosome (BAC) clone, together with analysis of a transcriptional start site by a cap-site hunting, indicated that the Ndrg4 gene spans

about 39 kilobases (kb) and consists of 19 exons, in which the first and second exons were first found in rat. An alternative promoter usage at different transcriptional start sites may produce three types of messages, Ndrg4-A, Ndrg4-B, and Ndrg4-C, and there is a variant that lacks exon 18 for each type of transcript. Thereby, Ndrg4-A1, Ndrg4-A2, Ndrg4-B1, Ndrg4-B2, Ndrg4-C1, and Ndrg4-C2 were identified to be expressed. These six variants might explain the heterogeneity of the Ndrg4 protein in the brain. The variants without exon 18 were revealed in the embryonic and early postnatal brains while those with exon 18 were detected in the maturing and adult brains. Radiation hybrid mapping suggests that the rat Ndrg4 gene is located on chromosome 19 at 90.6 centirays (cR) from the top. Comparison of the noncoding sequence of the rat Ndrg4 gene to those of the orthologous mouse and human genes suggests that the AP-1 binding site is a candidate regulatory element.

Magnusson, K. P., M. Sandstrom, et al. (2000). "p53 splice acceptor site mutation and increased HsRAD51 protein expression in Bloom's syndrome GM1492 fibroblasts." Gene **246**(1-2): 247.

http://www.sciencedirect.com/science/article/B6T39-4019DM2-W/2/9a0bacd60f934affbfc1a88c1d30adce

Majumder, K. (1992). "Ligation-free gene synthesis by PCR: synthesis and mutagenesis at multiple loci of a chimeric gene encoding OmpA signal peptide and hirudin." Gene **110**(1): 89.

http://www.sciencedirect.com/science/article/B6T39-47T30FY-3V/2/0874dfd68554961086824dd2115c7e6c

A unique kination and ligation-free method that allows de novo synthesis ofa gene through a novel application of polymerase chain reaction (PCR) involving stepwise elongation of sequence (SES) is described. SES-PCR is simple and efficient. Optimal utilization of nucleotides, ability to use only partially purified oligodeoxyribonucleotides, and elimination of kination and ligation of intermediates make SES-PCR-mediated gene synthesis more conomical in terms of time, labour and money. Site-directed mutagenesis and/or gene fusion by SES-PCR is not limited by the prior availability of the gene(s) in question. The potentials of this novel method in gene synthesis, mutagenesis at multiple loci of DNA and gene fusion have been demonstrated using a chimeric gene encoding fusion between OmpA signal peptide and hirudin, as an example. The SES-PCR product was cloned and sequencing of positive clones demonstrated the presence of genes with expected sequence and bearing only the desired mutations. A nearly 100% efficiency of mutation was easily achieved by the design of the method.

Majumder, K., S. Choudhury, et al. (1994). "Recombinant enrichment by exploitation of restriction sites with interrupted palindromes: design, synthesis and incorporation of zero-background linkers in cloning and expression vectors." Gene **151**(1-2): 147.

http://www.sciencedirect.com/science/article/B6T39-47P8201-2C/2/aac666a38b9f1e91a9a83f59f3bff81e

A universal approach for improving the efficiency of cloning through a selective enrichment of recombinants has been developed. This was achieved by using novel polylinkers or multiple cloning sites (MCS) termed zero-background linkers (ZBL). These MCS have short-cutter site (s) incorporated in the wobble portion of the recognition sequence of the long-cutter (s) in such a fashion that cloning at the short-cutter site disrupts the continuity of the long-cutter-specific

sequence. Consequently, digestion of the ligation mixture with the long-cutter, prior to transformation, essentially eliminates the insert negative clones from transformants. The usefulness of such background removal has been illustrated through the design and synthesis of a model ZBL molecule and then its incorporation into various popular cloning and expression vectors. We demonstrate that replacement of the residing MCS of a vector with ZBL leads to enhanced cloning efficiency, as evidenced by the marked increase in the ratio of recombinants to non-recombinants. Such vector improvement is nearly universal as zero backgrounding is an add-on feature and can be adapted to practically any vector system, without having to alter the essential features of the parent vectors.

Malecova, B., J. Ramser, et al. (2003). "Honeybee (Apis mellifera L.) mrjp gene family: computational analysis of putative promoters and genomic structure of mrjp1, the gene coding for the most abundant protein of larval food." Gene **303**: 165.

http://www.sciencedirect.com/science/article/B6T39-47PGBTF-3/2/781a6464a63c3a984cf35a90fd58f3fe

Mrjp1 gene belongs to the honeybee mrjp gene family encoding the major royal jelly proteins (MRJPs), secreted by nurse bees into the royal jelly. In this study, we have isolated the genomic clone containing the entire mrjp1 gene and determined its sequence. The mrjp1 gene sequence spans over 3038 bp and contains six exons separated by five introns. Seven mismatches between the mrjp1 gene sequence and two previously independently published cDNA sequences were found, but these differences do not lead to any change in the deduced amino acid sequence of MRJP1. With the aid of inverse polymerase chain reaction we obtained sequences flanking the 5' ends of other mrjp genes (mrjp2, mrjp3, mrjp4 and mrjp5). Putative promoters were predicted upstream of all mrjp genes (including mrjp1). The predicted promoters contain the TATA motif (TATATATT), highly conserved both in sequence and position. Ultraspiracle (USP) transcription factor (TF) binding sites in putative promoter regions and clusters of dead ringer TF binding sites upstream of these promoters were predicted computationally. We propose that USP, as a juvenile hormone (JH) binding TF, might possibly act as a mediator of mrjp expression in response to JH. Mrjp1's genomic locus is predicted to encode an antisense transcript, partially overlapping with five mrjp1 exons and entirely overlapping with the putative promoter and predicted transcriptional start point of mrjp1. This finding may shed light on the mechanisms of regulation of mrjps expression. Southern blot analysis of genomic DNA revealed that all so far known members of mrjp gene family (mrjp1, mrjp2, mrjp3, mrjp4 and mrjp5) are present as single-copy genes per haploid honeybee genome. Although MRJPs and the yellow protein of Drosophila melanogaster share a certain degree of similarity in aa sequence and although it has been shown that they share a common evolutionary origin, neither structural similarities in the gene organization, nor significant similarities between intron sequences of mrjp1 gene and fourteen yellow-like genes of D. melanogaster were found.

Marshall, B., G. Isidro, et al. (1996). "Insertion of a short Alu sequence into the hMSH2 gene following a double cross over next to sequences with chi homology." Gene **174**(1): 175.

http://www.sciencedirect.com/science/article/B6T39-3W2T6C3-V/2/02941c925004da7b13d2b92513a03103

Alu repeat sequences and other multiple copy repetitive elements are present throughout the human genome and are active in promoting recombination. It is believed that reverse transcription of transcribed Alu repeats followed by chromosomal integration has been responsible for the wide dispersion and high copy number of these sequences. During studies on the hMSH2 gene we have used RT-PCR to amplify from peripheral blood lymphocytes a cDNA

species in which 553 base pairs of hMSH2 cDNA have been deleted to be replaced by a short 36 base pair Alu sequence as a result of a genomic insertion/deletion event. The 36 base pair Alu insert is homologous to a 26 base pair Alu sequence previously implicated in the promotion of recombination and contains the GCTGG motif which is part of the prokaryotic chi sequence. A second chi-like sequence is also located within the deleted hMSH2 region. Both chi-like sequences are located within 4 bp of the two 4-bp regions of cross over containing the insertion/deletion breakpoints. This suggests that a double recombination event has occurred, providing direct evidence for the recombinogenic activity of this Alu element. Furthermore, it suggests that chi-like sequences may define recombination hotspots as in prokaryotes.

Marson, A. L., D. E. K. Tarr, et al. (2001). "Macrophage migration inhibitory factor (mif) transcription is significantly elevated in Caenorhabditis elegans dauer larvae." Gene **278**(1-2): 53.

http://www.sciencedirect.com/science/article/B6T39-44CMXX2-5/2/8140b9eb29f317eb2dbdb400b4e9b7f0

Macrophage migration inhibitory factor (MIF) from vertebrate species is a molecule that exerts a wide-range of effects in inflammatory responses, cell activation and cell differentiation. Several species of parasitic nematodes have been shown to express genes encoding orthologues of the mammalian MIF that appear to play a key role in immune evasion by modifying the activity of host cells. In addition, MIF accumulates in nematode somatic cells where its role has not yet been defined. In order to identify the role that MIF plays in the cell biology of nematodes, we have characterized the members of the mif gene family in the free-living species Caenorhabditis elegans. Unlike the single mif gene found in humans and mice, C. elegans expresses four distinct mif genes: Ce-mif-1, Ce-mif-2, Ce-mif-3 and Ce-mif-4. The Ce-MIF proteins are between 15-30% identical to each other, 34-38% identical to the MIFs from the parasitic nematode Brugia malayi, and 22-35% identical to mammalian MIFs. The transcription of Ce-mif-2 and Ce-mif-3, but not Ce-mif-1, was upregulated >100-fold compared to L2 levels when the worms entered the dauer stage. The transcription levels of Ce-mif-2 and Ce-mif-3 fell to near baseline a few hours after exit from dauer. Ce-MIF/GFP transgenic animals and immunostaining were used to demonstrate that the main sites of MIF production are in the hypodermis, body wall muscles and in the nuclei of developing embryos. The results suggest a role for C. elegans MIF in cellular maintenance during periods of adverse conditions that lead to developmental arrest.

Martins, A. S., L. J. Greene, et al. (1995). "The cDNA encoding canine dihydrolipoamide dehydrogenase contains multiple termination signals." Gene **161**(2): 253.

http://www.sciencedirect.com/science/article/B6T39-3Y6HGSR-1T/2/b52e0df550cace03d9ba2d198e7dc259

A 2288-bp cDNA sequence encoding dihydrolipoamide dehydrogenase (DLDH; dihydrolipoamide: NAD+ oxidoreductase; EC 1.8.1.4) was obtained by isolating a 1762-bp cDNA clone from a canine skeletal muscle library in the vector, [lambda]UNIZAP, combined with PCR amplification of the 5' end of the mRNA. The DLDH cDNA sequence contains a 49-bp G+C-rich 5'-untranslated region (UTR), followed by 1527 bp of coding region, and 695 bp of 3'-UTR preceding a 17-bp poly(A) tail. The single open reading frame encodes a precursor DLDH of 509 amino acids (aa) that begins with a 35-aa leader sequence. The 3'-UTR includes six possible polyadenylation signals (three AATAAA, one TATAAA and two AATGAA) and one potential stem-loop region extending from bp 1969-1991. Alignment studies of the canine and human DLDH demonstrate homology within the coding region of 98% at the aa level and 94% at the nt level. Northern blot analysis using the cDNA clone as probe showed wide tissue distribution of the mRNA, with differences in the level of expression among tissues and possible utilization of different

polyadenylation sites.

Marzullo, L., A. Tosco, et al. (2004). "Identification of dietary copper- and iron-regulated genes in rat intestine." Gene **338**(2): 225.

http://www.sciencedirect.com/science/article/B6T39-4CX728C-3/2/add293f73a2a1008a853798270f22543

Copper and iron act at different levels on gene expression. Due to their chemical reactivity, both metals could play a role in the regulation of the protein machinery involved in their metabolism, and/or of the metabolic function they are involved in. Experimental and clinical evidences raise also the hypothesis of the existence of genes commonly regulated by both metals. Purpose of this work was to find genes modulated by copper and iron in the rat intestine. A panel of 24 animals was randomly divided into three nutritional treatments including a control, a copper-deficient and an iron-deficient diet. The positive regulation of iron responsive element (IRE)-DMT1 gene was found, with different extent, in both experimental groups. A differential display reverse transcription (DDRT)-polymerase chain reaction (PCR) analysis carried out on the rat intestinal mRNAs demonstrated the differential expression of five cDNA fragments. Among these, the Cytochrome c oxidase (COX) subunit II mitochondrial gene resulted to be regulated by both metals, the Serum and Glucocorticoids-regulated Kinase (SGK) gene mainly by iron, and an Ebnerin-like 2 kb mRNA dramatically down-regulated by copper. Two residual clones showed low identity scores with sequences present in data bank. Finally, we observed that both iron and copper are able to modulate the expression of the three characterized genes in some tissues, other than intestine.

Matsumoto, M., M. Kamohara, et al. (2000). "The novel G-protein coupled receptor SALPR shares sequence similarity with somatostatin and angiotensin receptors." Gene **248**(1-2): 183.

http://www.sciencedirect.com/science/article/B6T39-4067BC2-N/2/55044220843402c52f26d0ad37696538

Matsuzaka, Y., K. Okamoto, et al. (2004). "Identification, expression analysis and polymorphism of a novel RLTPR gene encoding a RGD motif, tropomodulin domain and proline/leucine-rich regions." Gene **343**(2): 291.

http://www.sciencedirect.com/science/article/B6T39-4DM2G9T-1/2/c5598acc45ee1ba43a89fd50ce53a23a

We describe the isolation and characterization of a full-length cDNA encoded by a gene that was significantly down-regulated in the affected skin of patients with psoriasis vulgaris. The cDNA was isolated from a keratinocyte cDNA library and its sequence was found to correspond to a hypothetical locus recorded in GenBank with the accession number LOC146206. The nucleotide sequence of the full-length cDNA was found to have an open reading frame of 1365 amino acids and to span approximately 12 kb of genomic DNA with 39 exons on chromosome 16q22. The deduced amino acid sequence contains four distinct structural regions, an RGD motif, a leucine-rich repeat (LRR) region, a tropomodulin domain, and a proline-rich domain. The gene was consequently designated as RLTPR (RGD, leucine-rich repeat, tropomodulin and proline-rich containing protein). The RLTPR hypothetical protein has a functional domain organization similar to Acan125, a myosin-binding protein expressed by Acanthamoeba castellanni. RT-PCR with

RLTPR PCR primers amplified products from cDNAs prepared from all of the 30 different tissues that we examined including thymus, spleen, colon, skin, skin keratinocytes, skin fibroblasts and fetal skin. During the course of screening the human keratinocyte cDNA library, some alternative splicing was also detected in three regions of the RLTPR gene. In addition, sequence analysis of the RLTPR genes from eight psoriasis patients and eight healthy controls revealed a number of synonymous and nonsynonymous SNPs that may be useful markers for future disease association studies.

McDermott, J. B., C. A. Peterson, et al. (1992). "Structure and lens expression of the gene encoding chicken [beta]A3/A1-crystallin." Gene **117**(2): 193.

http://www.sciencedirect.com/science/article/B6T39-47FW95F-5/2/fb752857a0e9c384ef0e0549915c32c7

The [beta]A1- and [beta]A3-crystallins are major polypeptides in the lenses of vertebrates. We present evidence that a single [beta]A3/A1 gene encodes these two proteins in the chicken. The [beta]A3/A1 gene has been sequenced and its functional promoter identified in transfection experiments. The chicken [beta]A3/A1 gene has the same structure as the human orthologue: six exons with standard splice sites and two alternative start codons from which the protein products are apparently translated. Northern analysis revealed an abundant 0.9-kb transcript in the lenses of 1-2-day-old chickens and no detectable transcripts in the rest of the eye, brain, heart, kidney, liver or skeletal muscle. The 5' -flanking sequence of the chicken [beta]A3/A1 gene is very similar to that of the human and mouse genes, suggesting conservation of important putative regulatory sequences in addition to the TATA box. A thymidine-rich element (bp -218 to -163) and a potential AP-1-binding site (bp -264 to -258) 3re present within the chicken 5' -flanking region. A DNA fragment from -382 to +22 of the chicken [beta]A3/A1 gene is sufficient to promote expression of the bacterial cat gene in transfected chicken primary lens epithelial cells, but not in transfected dermal fibroblasts. Moreover, the sequence from positions -382 to -143 of the chicken [beta]A3/A1 promoter appears to be critical for proper transcription and expression in the transfected lens cells.

Miyamoto, Y., J. Kim, et al. (2002). "Genomic organization and chromosomal localization of the importin [alpha]1 gene in the mouse." Gene **288**(1-2): 49.

http://www.sciencedirect.com/science/article/B6T39-45G023J-2/2/1043b932d8d188d4561db82644846e02

Importin [alpha]1 (also referred to as NPI1 or importin [alpha]S1) gene encodes a member of the cytosolic receptor protein superfamily that recognizes classical monopartite and bipartite basic type nuclear localization signals and mediates nuclear protein import via an importin [beta]-dependent pathway. Here we report on the organization of the importin [alpha]1 locus in the mouse genome. The gene is approximately 40 kb in length from the translation initiation codon to the poly(A) additional site. The translated region of the gene is comprised of 13 coding exons and the exon-intron boundaries conform to the GT/AG rule. Importin [alpha]1 was mapped to a middle region of mouse chromosome 16 by fluorescence in situ hybridization analysis. Moreover, it was found by reverse transcriptase polymerase chain reaction analysis that importin [alpha]1 is widely expressed in various tissues in adult mice and at various stages during embyogenesis. This study is the first example that provides detailed genomic information on nuclear transport factors such as importins and exportins and provides a basis for further studies such as the generation of mutants in mice for purposes of investigating the role of importin [alpha]1 in development and differentiation.

Moir, D. T., T. E. Dorman, et al. (1993). "Rapid identification of overlapping YACs in the MEN2 region of human chromosome 10 by hybridization with Alu element-mediated PCR products." Gene **136**(1-2): 177.

http://www.sciencedirect.com/science/article/B6T39-47P8H5B-J6/2/07cf28f6b0780c517b1fe78b4b18bd69

An overlapping set of 21 yeast artificial chromosomes (YACs) spanning the RET proto-oncogene [Takahashi et al., Oncogene 3 (1988) 571-578] and D10S102 markers on human chromosome 10 was isolated in a series of hybridization-based chromosomal walks in a YAC library. Genetic linkage analyses implicate this chromosomal region as the location of the gene (MEN2A) responsible for multiple endocrine neoplasia type 2A. Four YACs carrying a RETsequence-tagged site (STS) and two YACs carrying a D10S102 STS were used to initiate chromosome walks. These were based on hybridization of Alu element-mediated polymerase chain reaction (Alu-PCR) products from YACs to dot blots of Alu-PCR products from complex pools of YAC clones. The hybridization anchor content of YACs identified in the walks was confirmed by probing blots of Alu-PCR products from individual YACs and by comparing Alu-PCR fingerprints of each YAC. Ten hybridization-based Alu-PCR anchors and three STS anchors were ordered within eleven intervals created by the 21 overlapping YACs. The order of anchors requiring the fewest gaps in the YACs is consistent with the walking results and establishes the STS anchor order as D10S102-D10S94-RET. The overlapping set of YACs represents about 1.55 Mb of the human genome according to restriction mapping of four representative YACs in the contig. These results demonstrate the power of Alu-PCR hybridization for chromosomal walking and provide a rich source of overlappping YACs which can be used to identify candidate MEN2A genes.

Monroe, R. S. and B. E. Huber (1994). "The major form of the murine asialoglycoprotein receptor: cDNA sequence and expression in liver, testis and epididymis." Gene **148**(2): 237.

http://www.sciencedirect.com/science/article/B6T39-47PNWHM-1V/2/99525dc060e99a1e955c07b7e478d7ef

Northern blot analysis of poly(A)+ RNAs isolated from mouse liver or mouse testis (Te)/epididymis (Ep) reveals that both tissues express 1.5- and 7.5-kb transcripts which have extensive homology to the major form of the rat asialoglycoprotein receptor (ASGP-R). In situ hybridization studies have localized the expression of this ASGP-R-like transcript to late-stage sperm from Te and Ep of several different strains of mice. Swiss Webster mice express this ASGP-R-like transcript in late-stage spermatids at the time of release into the seminiferous tubule and in Ep sperm, while Balb/C, NIH Swiss and C57B1/6 mice express this ASGP-R-like transcript predominantly in Ep sperm. cDNAs containing the entire coding region for this ASGP-R-like transcript have been cloned from mouse liver and mouse Te/Ep. These cDNAs are 100% identical in the coding region and 3'-untranslated region (UTR), but diner in the 5'-UTR. The gene encoding these cDNAs is called MHL-1, designating the major form of the mouse ASGP-R. The deduced amino acid (aa) sequence of MHL-1 shares 88% homology to the rat hepatic (He) lectin form 1 (RHL-1) and 78% homology to the human asialoglycoprotein receptor form I (HI). The three sites for N-linked glycosylation in the RHL-1 sequence are all conserved in the deduced MHL-1 sequence. Taken collectively, these data describe the cloning and sequencing of the MHL-1 cDNA and illustrate its deduced aa homology to RHL-1 and H1. Most importantly, these data show that mouse late-stage spermatids and Ep sperm express the authentic MHL-1 gene, suggesting this receptor may have an important role in spermatogenesis, in addition to its He function.

Monroe, R. S. and B. E. Huber (1994). "Characterization of the "hepatic" asialoglycoprotein receptor in rat late-stage spermatids and epididymal sperm." Gene **148**(2): 261.

http://www.sciencedirect.com/science/article/B6T39-47PNWHM-1Y/2/367a1c2ef3e8695bc37fe6e8e12fcb50

Northern blot analysis of rat testicular (Te) poly(A)+ RNA reveals that a transcript homologous to the major form of the asialoglycoprotein receptor (ASGP-R), designated RHL-1, is expressed as early as one week postnatally and that steady-state levels are approx. 8-times higher in the Te of an 8-week-old rat (sexually mature) as compared to an 84-week-old rat (aged). Partial cDNAs encoding RHL-1 and the minor form of the ASGP-R, designated RHL-2/3, have been cloned from two rat Te/epididymal (Ep) cDNA libraries and rat Te poly(A)+RNA. Sequence analysis of the Te/Ep RHL-1 cDNA and the Te/Ep RHL-2/3 cDNA indicates that these cDNAs are identical to the forms expressed in rat liver. Western blot analysis demonstrates the presence of a 49-kDa Te/Ep RHL-1-related protein band and a 54-kDa Te/Ep RHL-2/3-related protein band in both rat Te membrane fractions (MF) and rat Ep sperm MF. The RHL-1-related protein has been localized to late-stage Te spermatids at the time of release from the seminiferous tubules and to Ep sperm in the region of the sperm tail, referred to as the middle piece. Taken collectively, these data indicate that the authentic RHL-1 and RHL-2/3 genes of the ASGP-R are expressed in late-stage spermatids; however, the Te/Ep RHL-1-related protein differs in size from the hepatic RHL-1 polypeptide, possibly indicating a specific function of the RHL-1-related protein in spermatogenesis.

Moretti, P., P. Simmons, et al. (1994). "Identification of homeobox genes expressed in human haemopoietic progenitor cells." Gene **144**(2): 213.

http://www.sciencedirect.com/science/article/B6T39-47P8GYD-FM/2/1b548339eab005cdea629d629e6191e5

Homeodomain (HD)-containing proteins have been shown to regulate cellular commitment and differentiation in fungal, invertebrate and vertebrate systems. Bone marrow cells synthesizing the CD34 antigen are a complex mix of early, stem and progenitor cells at various stages of commitment to the many haemopoietic lineages. Here, we report the cloning and sequencing of 31 homeobox (HB) sequences, identified using degenerate oligodeoxyribonucleotide primers, in a polymerase chain reaction with cDNA derived from a purified CD34+ population of human haemopoietic cells. Of these sequences, 16 correspond to previously identified genes, and 13 are located within the HOXA, B and C clusters. Ten of the clones most likely represent human homologues of genes identified previously in other species. Five of the clones reported here represent novel HD sequences. The identification of five new genes using a subclass-specific 5' primer, designed from the engrailed and Xanfl sequences, suggests that there still remain several uncharacterised HB genes in the human genome. Haemopoietic cells purified on the basis of CD34 antigen synthesis are a rich source of regulatory genes consistent with their ability to differentiate into diverse haemopoietic cell types.

Muftuoglu, M., R. Selzer, et al. (2002). "Phenotypic consequences of mutations in the conserved motifs of the putative helicase domain of the human Cockayne Syndrome Group B gene." Gene **283**(1-2): 27.

http://www.sciencedirect.com/science/article/B6T39-44R2SP4-3/2/f41fd55d2a89d84fb49541c8732453d1

Cockayne syndrome (CS) is a human genetic disorder characterized by several neurological and developmental abnormalities. Two genetic complementation groups, CS-A and CS-B, have been identified. The CSB protein belongs to helicase superfamily 2, and to the SWI/SNF family of proteins. The CSB protein is implicated in transcription-coupled repair (TCR), basal transcription and chromatin remodeling. In addition, CS cells undergo UV-induced apoptosis at much lower doses than normal cells. However, the molecular function of the CSB protein in these biological pathways has remained unclear. Evidence indicates that the integrity of the Walker A and B boxes (motifs I and II) are important for CSB function, but the functional significance of the helicase motifs Ia, III-IV has not been previously examined. In this study, single amino acid changes in highly conserved residues of helicase motifs Ia, III, V, VI and a second putative nucleotide-binding motif (NTB) of the CSB protein were generated by site-directed mutagenesis to analyze the genetic function of the CSB protein in survival, RNA synthesis recovery and apoptosis after UV treatment. The survival analysis of these CS-B mutant cell lines was also performed after treatment with the chemical carcinogen, 4-nitroquinoline-1-oxide (4-NQO). The lesions induced by UV light, cyclobutane pyrimidine dimers, are known to be repaired by TCR whereas the lesions induced by 4-NQO are repaired by global genome repair. The results of this study demonstrate that the point mutations in highly conserved residues of helicase motifs Ia, III, V and VI abolished the genetic function of the CSB protein in survival, RNA synthesis recovery and apoptosis after UV treatment. Similarly, the same mutants failed to complement the sensitivity toward 4-NQO. Thus, the integrity of these helicase motifs is important for the biological function of the CSB protein. On the contrary, a point mutation in a C-terminal, second, NTB motif of the CSB protein showed full complementation in the ability to repair damage induced by UV light or 4-NQO, suggesting that this motif is not important for the CSB repair function.

Mukai, H., M. Miyahara, et al. (1995). "Identification of Schizosaccharomyces pombe gene pskl+, encoding a novel putative serine/threonine protein kinase, whose mutation conferred resistance to phenylarsine oxide." Gene **166**(1): 155.

http://www.sciencedirect.com/science/article/B6T39-3Y6HK1D-20/2/a3bd23a133990f637cd309f5825f218c

We have identified a novel putative protein kinase-encoding gene from Schizosaccharomyces pombe (Sp), designated pskl+, by using a highly conserved amino acid (aa) sequence motif to design amplification of DNA fragments using PCR. The putative translation product of pskl + contains 436 aa, with a molecular mass of 49 317 Da. A single pskl + was identified by genomic Southern blot analysis, and the genomic mapping indicated that psk1 + was localized in Sp chromosome III. Growth of wild-type Sp cells was inhibited by 0.5 [mu]M phenylarsine oxide, a protein tyrosine phosphatase inhibitor, but pskl- cells were relatively resistant to this drug.

Muramoto, Y., A. Watanabe, et al. (1999). "Enhanced expression of a nuclease gene in leaves of barley plants under salt stress." Gene **234**(2): 315.

http://www.sciencedirect.com/science/article/B6T39-3WWDH80-G/2/a97488f4cfcbdab97c56f3c3fc65aed0

Murphy, K. C., K. G. Campellone, et al. (2000). "PCR-mediated gene replacement in Escherichia coli." Gene **246**(1-2): 321.

http://www.sciencedirect.com/science/article/B6T39-4019DM2-15/2/0b555343955a22bbfee23c08ff4e6285

Nakano, M., K.-i. Yoshiura, et al. (1998). "Identification, characterization and mapping of the human ZIS (zinc-finger, splicing) gene." Gene **225**(1-2): 59.

http://www.sciencedirect.com/science/article/B6T39-3VH7JJP-8/2/37c3839af653c9101d894c3c41b8ba2c

From a human fetal brain cDNA library, we isolated two transcripts (ZIS-1 and ZIS-2) corresponding to the human ZIS gene, an ortholog of the rat Zis (zinc finger, splicing). A comparison of base sequences of the cDNA and its corresponding genomic DNA (a P1-derived artificial chromosome clone) revealed that both transcripts have an ORF of 1011 bp and encodes 337 amino acids, but ZIS-1 has 10 exons and ZIS-2 contains 11 exons. Although both transcripts share the first nine exons, exon 10 of ZIS-2 is lacking in ZIS-1, and instead, exon 11 (10th exon) of ZIS-1 is larger in size, leading to the longer 3'-UTR. Thus, the two transcripts result from differential splicing. A Northern blot analysis on various adult and fetal tissues revealed that 5.2- and 3.2-kb transcripts were ubiquitously expressed, and 3.9- and 1.9-kb transcripts were highly expressed in the fetal brain and kidney, respectively. There were several other transcripts that may be alternatively processed forms of the human ZIS. Considering the ZIS gene size, the 3.2-kb transcripts most likely corresponds to ZIS-1 and may act as a major transcript of ZIS. The human ZIS has a high homology to the rat Zis for the coding DNA sequence with 91% identity and for the amino acid sequence with 87% identity. ZIS and Zis contain the same numbers of exons and introns. Both genes have unusually long 3'-UTR, and their encoding proteins contain similar components, i.e. a zinc finger domain, a nuclear localization signal, an Asp-Glu region, and a Ser-Arg-rich region. Furthermore, the expression patterns of the two genes in tissues are similar each other. Thus, the human ZIS may act as a transcriptional factor to regulate transcription and/or splicing, as does the rat Zis.

Nikali, K., J. Saharinen, et al. (2002). "cDNA cloning, expression profile and genomic structure of a novel human transcript on chromosome 10q24, and its analyses as a candidate gene for infantile onset spinocerebellar ataxia." Gene **299**(1-2): 111.

http://www.sciencedirect.com/science/article/B6T39-4778G3N-5/2/ec91d70d64a3279056370aa6a42f480f

In our search for the disease gene underlying autosomally recessively inherited infantile onset spinocerebellar ataxia (IOSCA), we identified an expressed sequence tag cluster representing a previously uncharacterized transcript in the restricted genomic sequence covering the IOSCA locus on chromosome 10q24, and for mutation analyses in IOSCA patients isolated the corresponding novel human cDNA, C10orf6. Multiple tissue cDNA and Northern analyses showed that this gene is ubiquitously expressed, with expression levels highest in the skeletal muscle and less abundant in the brain, liver, and heart than in other tissues examined. C10orf6 consists of 20 exons forming a 7.3 kb cDNA which is capable of encoding a 1173 amino acid polypeptide and possesses orthologues in other mammals. Sequencing of RT and genomic PCR products of the gene revealed no alterations in IOSCA patients when compared to control subjects, and neither could differences be detected in expression levels between patient and control brain RNA samples, thus excluding mutation(s) in this novel gene as causative for IOSCA. However, this study facilitates future investigations on both the role of C10orf6 gene product in human cells as well as its possible involvement in the pathogenesis of other hereditary diseases mapped to chromosome 10q24.

Odeberg, J., A. Ahmadian, et al. (1999). "Context-dependent Taq-polymerase-mediated nucleotide alterations, as revealed by direct sequencing of the ZNF189 gene: implications for mutation detection." Gene **235**(1-2): 103.

http://www.sciencedirect.com/science/article/B6T39-3YRWC1C-D/2/9e7602d47c08b7e042c2f915d7bace53

Okamoto, K., Y. Matsuzaka, et al. (2003). "Identification of NAD+-dependent isocitrate dehydrogenase 3 [gamma]-like (IDH3GL) gene and its genetic polymorphisms." Gene **323**: 141.

http://www.sciencedirect.com/science/article/B6T39-4B1X6XH-4/2/580c5cbeb2921d80ef5808e2ff686607

We have identified a novel human gene designated as IDH3GL (isocitrate dehydrogenase 3 [gamma]-like) that is expressed specifically in human testis. The gene corresponds in sequence to an EST (expressed sequence tag) A1476435 that was first detected by differential expression analysis using a microarray assay. The full-length cDNA sequence (1037 bp) was isolated from the human testis 5'-3'-RACE cDNA libraries and found to have 83% nucleotide sequence identity with part of the IDH3G (isocitrate dehydrogenase 3 [gamma]). The IDH3GL gene consists of 3 exons spanning approximately 220 kb within the region of the NELL1 gene on chromosome 11p15.1. Sequence analysis of the IDH3GL cDNA revealed the presence of a premature stop codon at nucleotide positions 337-339 that results in a truncated peptide with 112 amino acids. This stop codon is conserved in various human ethnic populations and in the chimpanzee (Pan troglodytes). In order to assess the functional status of IDH3GL, especially in relation to the presence of the putative premature stop codon, single nucleotide polymorphisms (SNPs) were screened in the upstream, coding and non-coding regions of the IDH3GL gene in a Japanese population. As a result, a total of 10 SNPs were identified, seven were novel and one of them was a non-synonymous amino acid substitution from Leu to Val. We conclude that the IDH3GL gene sequence is a splice variant of the NELL1 gene and that it probably evolved from a transposed pseudogene of the IDH3 gene.

Olkkonen, V. M., J. R. Peterson, et al. (1994). "Isolation of a mouse cDNA encoding Rab23, a small novel GTPase expressed predominantly in the brain." Gene **138**(1-2): 207.

http://www.sciencedirect.com/science/article/B6T39-47PNXC3-G8/2/09086d5155052a0ebe3025e08ab614c8

The full-length cDNA encoding Rab23, a novel Ras-related small GTPase, was isolated using the sequence of a previously described [Chavrier et al., Gene 112 (1992) 261-264] short cDNA fragment and the rapid amplification of cDNA ends (RACE) PCR techniques. The deduced amino acid sequence was not very closely related to any previously described small GTPase, but was within the Rab subfamily. A Northern analysis revealed that the rab23 mRNA is predominantly expressed in the brain, which places the protein, together with Rab3a and Rab 15, in the group of small GTPases characteristic of the nervous system.

Ong, E., W. B. R. Pollock, et al. (1997). "Cloning and sequence analysis of two laccase complementary DNAs from the ligninolytic basidiomycete Trametes versicolor." Gene **196**(1-2): 113.

http://www.sciencedirect.com/science/article/B6T39-3R3GCDP-H/2/b4a0b2c4d55d4d611e34e8ebf93af5f9

Laccases are oxidoreductase enzymes involved in the oxidation of various phenolic compounds. They may play a role in the biodegradation of lignin and in the dechlorination of chlorophenols. The cDNAs encoding laccase LccI and a putative laccase LccIV and the gene for LccI from the white-rot basidiomycete Trametes versicolor were cloned, sequenced and characterized. The genomic DNA of lccI consists of 2128 bp, with the coding region interrupted by 10 introns; the cDNA consists of a 1560 bp open reading frame (ORF). The cDNA of the putative lccIV gene consists of a 1581 bp ORF, with a 794 bp 5' untranslated region. The size of the major transcript for both lccI and lccIV is approximately 2.3 kb. Transcription of lccIV was induced by 2,5-dimethylaniline, whereas the opposite effect was observed for lccI. Laccases I and IV contain highly conserved histidinyl and cysteinyl residues, believed to be involved in binding copper, and share extensive sequence similarity with other laccases produced by both ligninolytic and non-ligninolytic fungi.

Padas, P. M., K. S. Wilson, et al. (1992). "The DNA-binding protein HU from mesophilic and thermophilic Bacilli: gene cloning, overproduction and purification." Gene **117**(1): 39.

http://www.sciencedirect.com/science/article/B6T39-47T30VW-7W/2/31469f53a3499def50e1b83f232fc7bd

The major histone-like bacterial protein (HU)-encoding genes (hup) from five different Bacilli have been cloned, sequenced and overexpressed in Escherichia coli. The five Bacilli selected are closely related, but have different optimum growth temperatures: > 70[deg]C for Bacillus caldolyticus and B. aldotenax; 60-65[deg]C for B. stearothermophilus (Bst); 37[deg]C for B. subtilis and 30[deg]C for B. globigii. The deduced amino acid (aa) sequences from the three thermophiles are identical. Those from the two mesophiles are also identical and differ from those of the thermophiles at eleven aa positions. The mesophilic proteins have an extra two aa at the C terminus. Cells harbouring plasmids containing the hup genes can produce HU. An efficient purification scheme using cation-exchange chromatography and fast protein liquid chromatography is presented. This gives approx. 30-40 mg of more than 95% pure Bst HU per litre of E. coli culture.

Parthasarathy, L., R. Parthasarathy, et al. (1997). "Molecular characterization of coding and untranslated regions of rat cortex lithium-sensitive myo-inositol monophosphatase cDNA." Gene **191**(1): 81.

http://www.sciencedirect.com/science/article/B6T39-3RM6VMT-D/2/fb459cfd837979b702e89ff938688440

Lithium sensitive myo-inositol monophosphatase (IMPase) is a pivotal enzyme which controls the levels of brain inositol within the inositol-based signaling system. Its capacity to release free myo-inositol from inositol monophosphates generated from receptor-linked and de novo pathways is crucial to the maintenance of appropriate amounts of intracellular myo-inositol, which is essential for both inositol-based cell signaling and cell volume control. We present here the full length cDNA encompassing the coding and untranslated regions (5'- and 3'-UTRs) of rat brain IMPase. This cDNA was derived from rat cortex mRNA by the RT-PCR technique. Analysis of this cDNA revealed several interesting features which include a short 5'-untranslated region (5'-UTR) of 68 nucleotides followed by coding region of approximately 0.8 kb and a long 3'-untranslated region (3'-UTR) of 1.2 kb. Both 5'-rapid amplification of cDNA ends (5'-RACE) and 3'-RACE techniques were carried out to isolate both UTRs and double stranded sequencing was carried out to its

entirety (~2.1 kb) by 'gene walking' using several oligonucleotide primers. All nucleotides were sequenced unambiguously using the sense and antisense strands of DNA. PCR analysis for the coding region and the deduced amino acid sequence demonstrated a DNA fragment of 831 bp and 277 amino acids, respectively, which are strikingly similar to human hippocampal IMPase. The 5'-UTR demonstrated distinct CpG doublets, characteristic of 'housekeeping' genes. The sequence around the initiator methionine, AAGATGG, conforms well to the Kozak consensus sequence for mammalian protein biosynthesis and the 3'-UTR demonstrated three canonical and one unusual poly-adenylation signals followed by a 31 base poly(A) tail. The presence of a CCTGTG in the 3'-UTR (putative carbohydrate response element) links IMPase mRNA to brain carbohydrate metabolic pathways. Computer analyses demonstrated several unique features of this mRNA, including the potential formation of hairpin loops which might be important in its intracellular regulation and turn-over. In summary, this lithium-sensitive brain IMPase mRNA has the following characteristics: a 5'-CpG-rich short untranslated segment, a highly conserved coding region, and a long 3'-untranslated region with several polyadenylation signals.(c) 1997 Elsevier Science B.V. All rights reserved.

Pasteris, N. G., K.-i. Nagata, et al. (2000). "Isolation, characterization, and mapping of the mouse Fgd3 gene, a new Faciogenital Dysplasia (FGD1; Aarskog Syndrome) gene homologue." Gene **242**(1-2): 237.

http://www.sciencedirect.com/science/article/B6T39-3YGDG07-X/2/bbc0e9069af1beee45deb7c18def9f91

Patel, D. H., C. R. Crawford, et al. (2000). "Cloning, genomic organization and chromosomal localization of the gene encoding the murine sodium-dependent, purine- selective, concentrative nucleoside transporter (CNT2)." Gene **242**(1-2): 51.

http://www.sciencedirect.com/science/article/B6T39-3YGDG07-5/2/5fed520f76f84c48a8db0cdddc74ee84

Pawlik, K. M., C.-W. Sun, et al. (1995). "End joining of genomic DNA and transgene DNA in fertilized mouse eggs." Gene **165**(2): 173.

http://www.sciencedirect.com/science/article/B6T39-3Y6HK2X-2C/2/0df61a0bf275e7d8cdcfef56b995866b

A linear 5.2-kb HS2/[beta]-globin construct with an upstream KpnI terminus (4-nucleotide (nt) 3' protruding single strand, PSS) and a downstream SalI terminus (4-nt 5' PSS) was microinjected into fertilized mouse eggs. The injected DNA fragments integrated into the mouse genome primarily as a head-to-tail tandem array. Chromosome/transgene junctions were obtained from seven of eight transgenic animals. All of the junctions occurred in the proximity of a transgene KpnI end; a maximum loss of 8 nt from the transgene terminus was observed. Two of these junctions completely preserved the 4-nt KpnI 3' PSS. Transgene/transgene junctions from two animals were analyzed. SalI/KpnI junctions that completely preserved both the SalI 5' PSS and the KpnI 3' PSS were found in each animal. These are the first examples of complete nt preservation at junctions formed between a 5' PSS terminus and a 3' PSS terminus in transgenic mice. The data are consistent with the fill-in model of Thode et al. [Cell 60 (1990) 921-928] in which alignment proteins juxtapose 5' PSS and 3' PSS termini; DNA polymerase then utilizes the recessed 3'-OH of the 5' PSS terminus as a primer to synthesize DNA across the gap. This

mechanism results in the formation of junctions with no loss of sequence. The results described in the present paper suggest that this mechanism may be involved in the formation of junctions in transgenic mice.

Persic, L., M. Righi, et al. (1997). "Targeting vectors for intracellular immunisation." Gene **187**(1): 1.

http://www.sciencedirect.com/science/article/B6T39-3VWFYVS-1/2/127b969cc4feaa515561eab22b8bb0d5

We define intracellular immunization as the inhibition or inactivation of the function of a molecule by the ectopic intracellular expression of antibody binding domains which recognise the molecule. Such recombinant antibodies can be directed to different compartments of eukaryotic cells by means of previously defined targeting signals, thus permiting the study of any molecule in any cellular compartment for which an antibody is available. For this purpose, we have created a set of vectors based on the VHExpress vector described [Persic, L., Roberts, A., Wilton, J., Cattaneo, A., Bradbury, A. and Hoogenboom, H.R. (1997) An integrated vector system for the eukaryotic expression of antibodies or their fragments after selection from phage display libraries. Gene 187, 000-000], which has been modified to express scFvs (single chain fragments) linked to specific targeting signals. These permit the localisation of scFvs to different intracellular compartments: the endoplasmic reticulum (scFvE-er), the nucleus (scFvE-nuclear), the mitochondria (scFvE-mit), the cytoplasm (scFvE-cyto), and as secreted proteins (scFvE-sec). The function of these vectors has been assessed by the immunofluorescence of COS cells transiently transfected with constructs containing the [alpha]D11 scFv.

Persic, L., A. Roberts, et al. (1997). "An integrated vector system for the eukaryotic expression of antibodies or their fragments after selection from phage display libraries." Gene **187**(1): 9.

http://www.sciencedirect.com/science/article/B6T39-3VWFYVS-2/2/39702616ab8bf2596a4ce6d7ea318779

Phage display is now an established method to select antibody fragments specific for a wide range of diverse antigens. In particular, isolation of human monoclonal antibodies has become a reality and for most purposes bacterial expression of the selected recombinant antibody fragments is sufficient. However, there are some cases where the expression of complete human immunoglobulin in mammalian cells is, if not essential, at least desirable. For this reason we have designed and constructed a set of mammalian expression vectors which permit facile and rapid cloning of antibody genes for both transient and stable expression in mammalian cells. Immunoglobulin genes may be cloned into these expression vectors as V regions or as Fabs for expression as either complete antibodies or as Fab fragments, using restriction sites which are rare in human V genes. All the important elements in the vectors - promoter, leader sequence, constant domains and selectable markers - are flanked by unique restriction sites, allowing simple substitution of elements. The vectors have been evaluated using the variable regions from the neutralizing anti-nerve growth factor (NGF) antibody, [alpha]D11, and the V regions from 2E10, a scFv selected from a scFv phagemid library.

Peterson, M. J. and J. F. Morris (2000). "Human myeloid zinc finger gene MZF produces multiple transcripts and encodes a SCAN box protein." Gene **254**(1-2): 105.

http://www.sciencedirect.com/science/article/B6T39-413KWGN-

D/2/15393c94992253c476f751f44ca1a6ec

Pinto, J. P., M. C. P. Ohresser, et al. (2001). "Cloning of the bone Gla protein gene from the teleost fish Sparus aurata. Evidence for overall conservation in gene organization and bone-specific expression from fish to man." Gene **270**(1-2): 77.

http://www.sciencedirect.com/science/article/B6T39-436W3KY-8/2/3e6733c85ad1f6e040c7756226f890a6

Bone Gla protein (BGP, Osteocalcin) is a bone-specific vitamin K-dependent protein which has been intensively studied in mammals. Although BGP is the most abundant non-collagenous protein of bone, its mode of action at the molecular level remains unclear. From an evolutionary point of view, the appearance of BGP seems to parallel the appearance of hydroxyapatite-containing bone structures since it has never been found in elasmobranchs, whose skeleton is composed of calcified cartilage. Accordingly, recent work indicates that, in mammalian bone, BGP is required for adequate maturation of the hydroxyapatite crystal. Taken together, these data suggest that teleost fishes, presumably the first vertebrates to develop a BGP-containing skeleton, may be a useful model to further investigate BGP function. In addition, fish offer several advantages over mammalian models, due to a large progeny, external embryonic development and transparency of larvae. In the present work, the BGP cDNA and gene were cloned from a teleost fish, Sparus aurata, and its tissue distribution, pattern of developmental expression and evolutionary pathways analyzed. The molecular organization of the Sparus BGP (spBGP) gene is similar to mammalian BGP genes, and its expression throughout development follows the onset of calcification. The spBGP gene encodes a pre-propeptide of 97 amino acid residues, expressed only in bone and showing extensive homology to its mammalian homologs. Phylogenetic analysis of the available BGP sequences supports the hypothesis that all BGPs have a single origin and share a common ancestor with a related vitamin K-dependent protein (Matrix Gla protein).

Pradet-Balade, B., C. Salmon, et al. (1998). "Heterogeneity of eel thyrotropin [beta] mRNAs is due to a minisatellite in the 3' untranslated region of the gene." Gene **215**(2): 251.

http://www.sciencedirect.com/science/article/B6T39-3VGR2GV-22/2/20ffc9019471865e649ab5be3c8dea37

The aim of this study was to determine the causes of the high heterogeneity, in the number and the length, of the thyrotropin (TSH) [beta] mRNA in the European eel. Northern blot analysis showed that removal of the poly(A) tail did not affect this heterogeneity. PCR amplification on reverse-transcribed pituitary RNAs (RT) showed the main source of heterogeneity to be a highly variable region in the 3' untranslated region (UTR). PCR amplification of the 3' UTR from RTs and genomic DNAs demonstrated that the high variability reflected polymorphism within the eel TSH [beta] gene. Isolation and sequencing of 3' UTR amplification fragments showed that the variable region comprised more or less exact repetitions of a 26-42-bp fragment. The number of repetitions varied from one allele to another. This variable region could be characterized as a minisatellite. In conclusion, instability of a minisatellite in the 3' UTR of the TSH [beta] gene generated the multiple and widely differing TSH [beta] mRNAs.

Prentki, P., A. Binda, et al. (1991). "Plasmid vectors for selecting IS1-promoted deletions in cloned DNA: sequence analysis of the omega interposon." Gene **103**(1): 17.

http://www.sciencedirect.com/science/article/B6T39-47PNWW7-68/2/08947e64fd5417c7800e3b8aa0639aa8

We have constructed two plasmid vectors which allow selection for in vivo deletions within cloned DNA fragments. The plasmids are derivatives of pBR322 which carry the Escherichia coli rpsL (strA) gene, known to confer a dominant streptomycin (Sm)-sensitivity phenotype to the host cell, and a copy of the IS1 transposable element. Sm-resistant strains that harbor these plasmids display sensitivity to Sm. Spontaneous IS1-promoted deletions across the rpsL gene can be isolated simply by selection for Sm resistance. Hence, nested sets of deletions of a cloned DNA can be obtained and sequenced with an IS1-specific primer. Using this approach, we have determined the complete nucleotide sequence of the [Omega] interposon [Prentki and Krisch, Gene 29 (1984) 303-313].

Rafti, F., D. Scarvelis, et al. (1996). "A Drosophila melanogaster homologue of the human DEAD-box gene DDX1." Gene **171**(2): 225.

http://www.sciencedirect.com/science/article/B6T39-3W4975X-2B/2/aecfa70474c49db4ce65c2eabce50912

DEAD-box genes are found throughout evolution and encode RNA-binding proteins. Such proteins include eukaryotic initiation factor-4A, which is essential for protein translation, Vasa, which is essential for germ line development, and a number of nuclear and mitochondrial RNA splicing factors. Transcription of a human DEAD-box gene, DDX1, is elevated in two retinoblastoma cell lines as a result of amplification of the immediate chromosomal region surrounding it, suggesting an important role for this gene in control of cell growth and division. We have isolated a Drosophila melanogaster (Dm) homologue (Ddxl) of DDX1 which is strikingly similar to the human gene. The similarity (58.3% amino acid (aa) identity over 720 aa) extends beyond regions conserved in all DEAD-box proteins and covers the entire lengths of the proteins. The 2.7-kb Dm Ddxl RNA is expressed throughout development, but its levels are elevated in early embryos. Ddxl maps to polytene chromosome band 79D4 on the left arm of Dm chromosome 3.

Regnier, V., J. Novelli, et al. (2003). "Characterization of chicken CENP-A and comparative sequence analysis of vertebrate centromere-specific histone H3-like proteins." Gene **316**: 39.

http://www.sciencedirect.com/science/article/B6T39-49KH2XM-4/2/cf4e4a409f1266f698028bc88381d88a

Centromere protein A (CENP-A) is a centromere-specific histone H3 variant conserved amongst all eukaryotes. We have isolated the chicken gene for CENP-A (GgCENP-A). It encodes a 131-amino-acid polypeptide that possesses an average identity of 54% with human CENP-A, reaching 69% in the histone-fold domain. The gene spans 1.7 kb of genomic DNA and contains four exons that range in size from 78 to 186 bp. The exon/intron organisation of the chicken gene is conserved with its mammalian counterparts in the carboxy-terminal histone-fold domain (exons 2 to 4), consistent with the strong conservation of this domain at the amino acid level. Sequence analysis of the chicken CENP-A locus revealed that the gene is located within the class III genes of the major histocompatibility complex (MHC), and extended the previously defined limit of the compact chicken MHC complex. We compared the sequences of CENP-A from mammals, chicken and fishes and thereby identified conserved motifs in the otherwise variable amino-terminal tail that may be important for functional reasons. We also identified evolutionarily variable regions within the conserved histone-fold domain. We found that loop 1 between the first

and second [alpha]-helix is the region that diverged most widely. This finding is in agreement with evolutionary studies in Drosophila species, and suggests that this domain could play a role in species-specific centromere targeting of CENP-A. In addition, protein sequence comparison of several vertebrate species revealed that the RT-PCR strategy we have developed for isolating the chicken centromeric histone H3 variant gene should be applicable to the isolation of CENP-A from a wide range of vertebrates.

Remillard, S. P., E. Y. Lai, et al. (1995). "A calcineurin-B-encoding gene expressed during differentiation of the amoeboflagellate Naegleria gruberi contains two introns." Gene **154**(1): 39.

http://www.sciencedirect.com/science/article/B6T39-3Y5MNJ3-3B/2/9d3b3693d545136b8dec65ba491c6a4b

One of two similar genes in the unicellular eukaryote Naegleria gruberi is shown to encode calcineurin B (CnB), the regulatory subunit of calcium-calmodulin-regulated protein phosphatase 2B. Over a span of 156 amino acids, excluding divergent N-termini, the encoded sequence shows 62% identity with vertebrate CnB, and also shows sequence elements specific, among calcium-binding proteins, to CnB. In contrast, the sequence shows only 23% identity with N. gruberi flagellar calmodulin. CNB mRNA is readily detected in amoebae; its abundance increases fourfold during differentiation to flagellates, reaches a peak at 50-70 min, when flagella are forming, and then declines. A genomic clone matches an expressed cDNA, except that it is interrupted by two phase I introns. The position of one intron, which separates the divergent N-terminal domain from the four calcium-binding domains (EF hands), is shared with a yeast CNB gene; the other is located in the central helix between the two pairs of calcium-binding loops; features that support an ancient origin. These introns, the first found in protein-coding genes of Naegleria, are flanked by characteristic splice junction sequences. N. gruberi CnB also shares similarities with recoveries. The finding in a protist of a CNB gene that contains two introns separating functional domains, shares similarities to recoveries and shows increased expression during differentiation is provocative. If the phylogeny of major groups derived from ribosomal RNA is accepted, Naegleria is among the earliest branching eukaryotes known to contain canonical pre-mRNA introns.

Retallack, D. M. and J. P. Woods (1999). "Transcript splicing is essential for functional Histoplasma capsulatum URA5 expression." Gene **230**(2): 181.

http://www.sciencedirect.com/science/article/B6T39-3W846W4-7/2/67c76ba02ae52358b5ba2a46221eaf1c

Rie, T., B. Anderson, et al. (2003). "Rapid and efficient transposon mutagenesis of Bartonella henselae by transposome technology." Gene **313**: 103.

http://www.sciencedirect.com/science/article/B6T39-493HSN6-4/2/6c6d9a80b7df176c7359dcb41abc4b93

Molecular genetics are difficult to perform in Bartonella henselae, the causative agent of cat scratch disease and the vasculoproliferative disorders bacillary angiomatosis and bacillary peliosis. To elucidate the underlying bacterial pathogenic mechanisms, genetic manipulation of B. henselae is the method of choice. We describe how to perform transposon mutagenesis in B. henselae using transposome technology. B. henselae mutants revealed by this technique showed

random transpositional insertion into the chromosome. In contrast to transposon mutagenesis by conjugational transfer, transposome technology allows transposon mutagenesis of early passaged Bartonella spp. with approximately 100-fold higher efficiency. The results show that transposome technique is a rapid, efficient and simple method to generate transposon mutants of B. henselae.

Riggio, M., R. Scudiero, et al. (2000). "Sex- and tissue-specific expression of aspartic proteinases in Danio rerio (zebrafish)." Gene **260**(1-2): 67.

http://www.sciencedirect.com/science/article/B6T39-41YG91G-7/2/657506567e76329e435e7ed170b22345

Roberts, L. R., L. A. Nichols, et al. (1995). "cDNA and amino-acid sequences and organization of the gene encoding the B[beta] subunit of fibrinogen from Xenopus laevis." Gene **160**(2): 223.

http://www.sciencedirect.com/science/article/B6T39-3Y6HGWH-3S/2/5d2fcb1334ca20b4043bbb80df09631e

Fibrinogen, the major blood-clotting protein, is made up of three chains, A[alpha], B[beta] and [gamma], which are synthesized and secreted by the liver. In this communication, we describe the complete cDNA sequence, deduced amino acid (aa) sequence and organization of the gene encoding the B[beta] subunit of fibrinogen from Xenopus laevis (Xl). The cDNA representing the predominant form of the B[beta] mRNA comprises 2390 nucleotides (nt), with an open reading frame of 1467 nt coding for a 488-aa protein. The percent identity between Xl B[beta] and that of other animals ranges from 50% for lamprey to 66% for human. The Xl B[beta] gene consists of nine exons, one more than found in the human gene. The exon/intron boundaries in the frog and human B[beta] genes are in exactly conserved positions, except for junctions in the highly variable fibrinopeptide-encoding regions. Three of the exon/intron boundaries in the Xl B[beta] gene are also analogous to ones in A[alpha] and [gamma] genes of other species, supporting the notion of a close evolutionary relationship between the genes for all three subunits. This analysis of B[beta] from an amphibian provides the first complete description of the arrangement of exons and introns in any fibrinogen subunit gene from a non mammal and gives insight into the most highly conserved aspects of fibrinogen protein structure and gene organization.

Robinson, C. A., A. Hayward-Lester, et al. (1997). "Quantification of alternatively spliced RUSH mRNA isoforms by QRT-PCR and IP-RP-HPLC analysis: a new approach to measuring regulated splicing efficiency." Gene **198**(1-2): 1.

http://www.sciencedirect.com/science/article/B6T39-3VBSRC2-1/2/87f887f85ef657a00fae07e3522b62f4

Quantitative reverse transcriptase-polymerase chain reaction (QRT-PCR) and the ion-pair reverse-phase (IP-RP)-HPLC product purification and detection system were developed to facilitate the isolation and proportional quantification of alternatively spliced RUSH mRNAs. RUSH isoforms result from alternative splicing of a 57-bp exon and encode SNF/SWI-related proteins that bind to the uteroglobin promoter. QRT-PCR was performed using total RNA, and a pair of primers designed to flank the 57-bp exon. When more than one splice variant was expressed, IP-RP-HPLC identified the specific homoduplex products, as well as the heteroduplexes formed as a consequence of partial sequence complementarity between the

products. Data analysis included the correct re-allocation of heteroduplex components to achieve accurate quantitation of changes in the relative levels of RUSH message isoforms. The preferential expression of the RUSH-1[alpha] isoform by all the tissues except estrous uterine endometrium and lactating mammary gland indicates RUSH pre-mRNAs are alternatively spliced in a tissue-specific manner. A 61-fold difference in the relative rate of RUSH pre-mRNA splicing is indicated by the difference in the ratios of RUSH mRNA isoforms from uterine endometrium and testis. Clearly, QRT-PCR and IP-RP-HPLC are powerful and versatile tools for the detection and quantitation of mRNA splice variants.

Romio, L., L. Musante, et al. (1999). "Characterization of a murine gene homologous to the bovine CaCC chloride channel." Gene **228**(1-2): 181.

http://www.sciencedirect.com/science/article/B6T39-3W07P0S-N/2/f56ba3e1f7f2517ded0fd115007b08a6

Roovers, E., M. E. Vincent, et al. (1995). "Characterization of a putative molluscan insulin-related peptide receptor." Gene **162**(2): 181.

http://www.sciencedirect.com/science/article/B6T39-3Y6HGRC-2/2/16ab33ff25ccaeeba200ffd7fb1fa74a

In the pond snail Lymnaea stagnalis (Ls), growth and associated processes are likely to be controlled by a family of molluscan insulin-related peptides (MIP). Here we report on the cloning of a cDNA encoding a putative receptor for these MIP. This cDNA was isolated from Ls via PCR with degenerate oligodeoxynucleotides corresponding to conserved parts of the tyrosine kinase domain of the human insulin receptor and its Drosophila homologue. Many of the typical insulin-receptor features, including a cysteine-rich domain, a single transmembrane domain and a tyrosine-kinase domain are conserved in the predicted, 1607-amino acid (aa) protein. Comparison of the aa sequence of the molluscan receptor to other insulin-receptor sequences revealed strong variations in the percentage of sequence identity for the different domains, ranging from 70% sequence identity in the tyrosine-kinase domain to virtually no sequence identity in the C-terminal sequence. Striking differences are the absence of a clear tetrabasic cleavage site, and the extremely long C-terminus of 308 aa that contains seven Tyr residues. Southern blot analyses at varying stringencies, extensive screening of cDNA- and genomic libraries, and PCR experiments indicate the presence of a single putative MIP receptor. This suggests that the four different MIP may exert their functional role in Ls by binding to the same receptor.

Rosin-Arbesfeld, R., D. Willbold, et al. (1998). "The Tat protein of equine infectious anemia virus (EIAV) activates cellular gene expression by read-through transcription." Gene **219**(1-2): 25.

http://www.sciencedirect.com/science/article/B6T39-3WBG1Y1-P/2/7b9ecb30633b0c11d299f3dc8f31ed90

The Tat protein of equine infectious anemia virus, EIAV, was shown to augment viral gene expression, presumably through interaction with the Tat responsive element, TAR. Recently, cell-free polyadenylation assays suggested that perturbation of the EIAV TAR secondary structure diminished polyadenylation efficiency. The present study indicates that the EIAV TAR regulates the efficiency of the 3'-end processing of viral RNA also in transfected cells. Moreover, our data

suggest that the provision of the EIAV Tat protein in trans potentiates read-through transcription through the 3' viral long terminal repeat (3' LTR), thus suggesting activation of downstream-located cellular genes.

Roux, D. T.-L., M. Senorale-Pose, et al. (1994). "Three novel SMR1-related cDNAs characterized in the submaxillary gland of mice show extensive evolutionary divergence in the protein coding region." Gene **142**(2): 175.

http://www.sciencedirect.com/science/article/B6T39-47P8HFC-MP/2/91447eb635fb4c397586a0c801830ef4

We have previously characterized an abundant male-specific mRNA from the submaxillary gland (SMG) of rats, encoding the SMR1 (androgen-regulated) protein, which has the structure of a prohormone and is processed by maturation enzymes to release a small peptide in the blood and saliva. We have now characterized three SMR1 -related cDNAs in the SMG of Balb/c mice. These cDNAs encode three novel proteins, designated MSG1, MSG2 and MSG3. They are 639, 662 and 471 nucleotides (nt) long, respectively, and the corresponding mRNAs appear to be expressed only in the SMG. The putative polypeptides they encode carry an N-terminal secretory peptide sequence and are, therefore, presumably secreted into saliva. Although closely related, the three mRNAs show striking differences: a particularly different expression pattern and an extremely high degree of variability observed in the central part of the molecules. The MSG1 and MSG3 cDNAs are identical, except for a 173-bp insert found only in MSG1. This insert contains three Pro-rich repeats (GPGIGRPPPPPP), reminiscent of the most abundant multigenic family of the SMG, the Pro-rich proteins (PRP). Although MSG1 shares several common features with PRP, it is structurally related to SMR1. The unusually high ratio of replacement/silent nt changes provides a basis to address complex aspects concerning the molecular events leading to the emergence of new proteins in the SMG.

Sakamoto, H., T. Araki, et al. (2000). "Expression of a subset of the Arabidopsis Cys2/His2-type zinc-finger protein gene family under water stress." Gene **248**(1-2): 23.

http://www.sciencedirect.com/science/article/B6T39-4067BC2-3/2/062edce8b6a17759534e9367074af72a

San Mateo, L. R., K. L. Toffer, et al. (1998). "The sodA gene of Haemophilus ducreyi encodes a hydrogen peroxide-inhibitable superoxide dismutase." Gene **207**(2): 251.

http://www.sciencedirect.com/science/article/B6T39-3SBNJXV-N/2/8989c92fe7fb3d3165190d37afa8910d

Haemophilus ducreyi is the etiologic agent of the sexually transmitted disease chancroid, an ulcerative condition implicated in increased HIV transmission. There is increasing evidence for the roles of oxidative stress proteins including superoxide dismutase enzymes in the survival and persistence of pathogenic organisms within the host. The sodA gene of Haemophilus ducreyi was isolated from a genomic plasmid library on the basis of its ability to rescue the hydrogen peroxide hypersensitivity of an Escherichia coli sodA-sodB- strain. The H. ducreyi SodA protein also complemented the aerobic growth defect of the E. coli sodA-sodB- strain in minimal medium. The deduced amino-acid sequence of the H. ducreyi sodA gene product is 74 and 70% identical to the Mn-SODs of Haemophilus influenzae and E. coli, respectively. However, unlike Mn-SODs, the H.

ducreyi SodA protein was inhibited by hydrogen peroxide in native gels stained for SOD activity.

Schroeder, A. A., A. M. Brown, et al. (1998). "Identification and cloning of a developmentally regulated Cryptosporidium parvum gene by differential mRNA display PCR." Gene **216**(2): 327.

http://www.sciencedirect.com/science/article/B6T39-3TJ452W-D/2/38f24a572fa315da5c5c4708959fd4b4

To identify Cryptosporidium parvum genes expressed during intracellular development, differential mRNA display was used to detect differences in gene expression between mock-infected and C. parvum-infected human epithelial cells. A reproducible band present only in C. parvum-infected cells, ddHC-23, was isolated and cloned. Southern blot analysis demonstrated that ddHC-23 represented a C. parvum gene. RT-PCR revealed that HC-23 mRNA levels decreased from 6 to 12 h post-infection (pi), were maximally expressed at 24 h pi, and returned to low levels at 48 and 72 h pi. Northern blot analysis determined that the approx. 3.6 kb transcript is expressed by sporozoites prior to invasion of epithelial cells. Screening of a C. parvum genomic library with ddHC-23 isolated a genomic subclone which contained a 2790 bp ORF, uninterrupted by introns. Sequence analysis indicated that the encoded protein, which displayed no similarity to any sequences in the public databases, contained a high proportion of polar amino acids, with the most abundant being Asp (17.3%), Ser (15.8%) and Gly (8.1%). Numerous potential sites for posttranslational modification were present including: casein kinase II and protein kinase C phosphorylation sites, N-myristolation sites and N-glycosylation sites. These findings demonstrate the usefulness of differential mRNA display for identifying developmentally regulated C. parvum genes within the background of genes expressed by the host cell.

Schwartz, F. and T. Ota (1997). "The 239AB gene on chromosome 22: a novel member of an ancient gene family." Gene **194**(1): 57.

http://www.sciencedirect.com/science/article/B6T39-3V4BS2H-2S/2/b88313bb6701f17d723995cd924dd82c

A novel family of genes expressed in human brain has recently been identified. Gene 239FB, transcribed extensively in fetal brain, was isolated from the chromosome 11p13 region associated with mental retardation component of the WAGR (Wilms tumor, aniridia, genitourinary anomalies, mental retardation) syndrome. This report presents a cDNA sequence and expression profile of a related gene, 239AB, isolated from adult brain library, that was mapped to chromosome 22. While similar in structure, the two genes differ in their expression pattern and may have different roles in central nervous system development and function. In contrast to the 239FB, which is expressed predominantly in fetal brain, the 239AB gene is transcribed in adult tissues. Both human genes encode novel proteins of unknown function that are highly conserved from Caenorhabditis elegans to birds and mammals. Phylogenetic analysis suggested that the two lineages of the ancient gene family represented by 239FB and 239AB have been in existence prior to the emergence of modern animals.

Sekizawa, Y., T. Kubo, et al. (1997). "Molecular cloning of cDNA for lysenin, a novel protein in the earthworm Eisenia foetida that causes contraction of rat vascular smooth muscle." Gene **191**(1): 97.

http://www.sciencedirect.com/science/article/B6T39-3RM6VMT-

G/2/6a30b0dd311ee7f8239b8769cc4b049a

Lysenin, which causes contraction of rat vascular smooth muscle, is a protein that was isolated from the earthworm Eisenia foetida. A cDNA encoding lysenin was isolated by use of a partial cDNA probe that had been generated by the PCR with a primer designed by reference to an internal peptide sequence of lysenin. This clone had an ORF encoding 297 amino acid residues. The amino acid sequence deduced from the cDNA revealed the absence of any significant homology to those of previously characterized vasoactive substances. The recombinant lysenin was produced in Escherichia coli. This protein and native lysenin isolated from the earthworm had similar contractive activities when tested on rat aorta. Northern blot analysis of the RNA from various tissues of the earthworm indicated that lysenin is produced by the coelomocytes.

Shahrestanifar, M., D. P. Saha, et al. (1994). "Cloning of a human cDNA encoding a putative nucleotide-binding protein related to Escherichia coli MinD." Gene **147**(2): 281.

http://www.sciencedirect.com/science/article/B6T39-47P8MDG-1NJ/2/1ff84c05be523f6fa696fe7d6617880c

A novel human cDNA encoding a putative nucleotide-binding protein (NBP) was obtained by screening a human SHSY5Y neuroblastoma library. The deduced protein contains 320 amino acids (aa) with a Mr of 34 540. NBP displays sequence similarity with the product of the minD gene from Escherichia coli. MinD is involved in the proper placement of the division septum, and has ATPase activity. NBP and MinD contain consensus nucleotide (nt)-binding domains. The NBP mRNA is approx. 1500 nt in length and is expressed in several human cell lines and in all rat tissues examined, with the highest levels in lung and testis.

Shamsher, M. and X. Montano (1996). "Analysis of intron 4 of the p53 gene in human cutaneous melanoma." Gene **176**(1-2): 259.

http://www.sciencedirect.com/science/article/B6T39-3W25BP2-33/2/87a95b6defb316f4e922108dee91ccbe

DNA sequencing of intron 4 of the p53 gene from seven cutaneous melanoma cell lines showed an absence of mutations. However, both control and melanoma cell lines sequences were different from the reference source obtained from GenBank databank (accession No. X54156). Base pairs 101 and 689 were determined to be T (instead of A) and C (instead of G). Also, an additional C was not detected at position 371. Comparative analysis with p53 DNA-binding sequences, a sequence recognized by a p53 intron 4-binding protein and consensus sequences recognized by transcription factors demonstrated that intron 4 contains putative sequences for NF-[kappa]B, SP1, AP1 and TFIID binding. Binding of transcription factors could be one of the mechanisms by which intron 4 modulates human p53 expression.

Shimizu, C., M. Kubo, et al. (1997). "Genomic organization of the mouse adrenocorticotropin receptor." Gene **188**(1): 17.

http://www.sciencedirect.com/science/article/B6T39-3RD1R45-3/2/fd4a00418e86919e37f9a6bf42e3d381

As a step toward understanding the transcriptional regulation of the adrenocorticotropin receptor

(ACTH-R) gene, we examined the full length cDNA sequence of the mouse ACTH-R by rapid amplification of cDNA ends, and the organization of the gene. Mouse ACTH-R mRNA consists of 374 bp in the 5'-untranslated region (UTR), 888 bp in the coding sequence, and 445 bp in the 3'-UTR, the 1707 bp being fairly compatible with the 1.8-kb adrenal mRNA detected by Northern analysis. The mouse ACTH-R gene consists of at least four exons; the first three exons encode 5'-UTR and the fourth exon encodes part of 5'-UTR, the entire coding region, and the whole of 3'-UTR. We also defined two mRNA species, one with and one without the 57-bp exon 2, produced by alternative splicing.

Shintani, S., M. Kobata, et al. (2002). "Identification and characterization of ameloblastin gene in a reptile." Gene **283**(1-2): 245.

http://www.sciencedirect.com/science/article/B6T39-44KPS90-3/2/f7e2f214a6e631c3594ad7024ae480f9

Ameloblastin (AMBN) is one of the enamel sheath proteins which presumably has a role in determining the prismatic structure of growing enamel crystals. There may therefore be a relationship between the molecular evolution of the AMBN gene and the development of enamel prismatic structures. To investigate whether such a relationship exists, it was necessary to identify the homologues of the AMBN gene in a reptile whose teeth lack an enamel prismatic structure. To this end, several clones containing AMBN cDNA were isolated from caiman jaws using the reverse transcription-polymerase chain reaction (RT-PCR) method. Sequence analysis of the AMBN cDNA revealed an open reading frame of 1221 bp encoding a 407-amino-acid protein. Translation of the caiman cDNA starts at the methionine corresponding to the second of two putative start codons conserved in mammalian AMBN genes. The N-terminal part of the caiman AMBN shows high amino acid sequence similarities to human, pig, cattle, rat and mouse AMBN sequences, as well as several other features that have been conserved throughout the evolution of reptiles and mammals. Unexpectedly, the nucleotide sequences of the 3' untranslated region (UTR) are also conserved, not only within mammalian genes but also between reptilian and mammalian genes. The caiman AMBN gene is a single-copy gene, transcribed only in the jaws, presumably in teeth.

Shintani, S., M. Kobata, et al. (2003). "Identification and characterization of ameloblastin gene in an amphibian, Xenopus laevis." Gene **318**: 125.

http://www.sciencedirect.com/science/article/B6T39-49M0W0P-4/2/9928ddc99ee56172baf8eeded30768eb

Ameloblastin (AMBN) is an enamel sheath protein that presumably has a role in determining the prismatic structure of growing enamel crystals. To investigate the relationship between the molecular evolution of the AMBN gene and development of enamel prismatic structures, it is considered to be of great significance in the identification of homologues of the AMBN genes in nonmammals whose teeth lack an enamel prismatic structure. Several clones containing AMBN cDNA were isolated from an African clawed toad tooth cDNA library by screening with a polymerase chain reaction (PCR) method. Sequence analysis of the clones revealed that they were derived from different genes (toad-A and toad-B), which were found to contain ORFs encoding 408- and 352-amino-acid proteins, respectively. The N-terminal part of the toad AMBN proteins and the phosphorylation motif for casein kinase II, as well as several features, were found to be highly conserved throughout the evolution of tetrapods. Exon-intron boundaries were shared by toad and caiman genes with the exception of exons 6, 7 and 10 while human and caiman genes shared them exclusive of exons 8 and 9 which have been found only in the human. As for exon 7, it was absent in both toad genes. Moreover, the AMBN genes were transcribed

only in the upper jaw, presumably in teeth. These results may provide useful information for investigation of the evolution of enamel.

Siever, D. A. and M. F. Verderame (1994). "Identification of a complete Cek7 receptor protein tyrosine kinase coding sequence and cDNAs of alternatively spliced transcripts." Gene **148**(2): 219.

http://www.sciencedirect.com/science/article/B6T39-47PNWHM-1S/2/2bfb5112ef4b12e7f09762a80e5d6a99

Receptor protein tyrosine kinases (RPTK) are critical components of signal transduction pathways in multicellular organisms. Identification of new RPTK constitutes an initial step in understanding the variety of signalling pathways in which these proteins participate. In this study, a cDNA containing a complete coding sequence for Cek7 (chicken RPTK) has been cloned from a chicken embryo expression library using anti-phosphotyrosine antibodies (Ab). Cek7 is a member of the EPH (human RPTK) subfamily of RPTK; this subfamily is characterized by extracellular domains containing an immunoglobulin-like motif, a Cys-rich region and two fibronectin type-III repeats. Analysis of additional cDNAs revealed that two positions of alternative splicing in primary transcripts may produce several isoforms of this RPTK; cDNAs corresponding to three isoforms of this receptor are reported. These isoforms are predicted to have altered extracellular ligand-binding domains and/or altered cytoplasmic juxtamembrane regions. The nucleotide sequence of cek7 cDNAs identified in this study diverges at the 3' end from the sequence found in a recently described partial cek7 cDNA [Sajjadi and Pasquale, Oncogene 8 (1993) 1807-1813]. Therefore, a third position of alternative splicing may produce Cek7 RPTK with divergent C-terminal tails. RNA blot analysis revealed expression of this receptor at highest levels in the central nervous system and eyes of 10-day-old chicken embryos.

Siigur, E., A. Aaspollu, et al. (2001). "Sequence diversity of Vipera lebetina snake venom gland serine proteinase homologs - result of alternative-splicing or genome alteration." Gene **263**(1-2): 199.

http://www.sciencedirect.com/science/article/B6T39-42D2CGD-R/2/110029785d424fc0aa5092cce91b1bf1

Four clones encoding homologous protein(ase)s were isolated from the Vipera lebetina (snake) venom gland cDNA library. One of them represented DNA encoding factor V activating enzyme (Siigur et al., 1999), the other is homologous to VLFVA but has two principal discrepancies in the translated protein sequence in comparison with snake venom serine proteinase structures: in the active site triad Ser195 is replaced by Asn195 and His57 by Arg57. The third and the fourth clone represent combinations of the first two clones. The possibilities of generation of such clones via trans-splicing of the primary gene transcript, by exon shuffling or by unequal crossing-over on the genome level are discussed.

Singh, K. P. and D. Roy (2001). "Identification of novel breast tumor-specific mutation(s) in the q11.2 region of chromosome 17 by RAPD/AP-PCR fingerprinting." Gene **269**(1-2): 33.

http://www.sciencedirect.com/science/article/B6T39-433NSPG-4/2/2c6e6b8d2b3a4724958c447a13d88770

Analysis of genetic instability in breast cancer tissues compared to uninvolved breast tissues from the same individuals by RAPD (random amplified polymorphic DNA)/AP-PCR (arbitrarily primed

PCR) fingerprinting using 30 arbitrary primers revealed 190 amplified DNA fragments. Presumably, each of these represents a gene locus in a different region of the genome of breast cancer tissues. Among these amplified DNA fragments, 65 (34.2%) exhibited presence and absence or reductions and enhancements in the intensity in breast cancer tissues compared to uninvolved breast tissues from the same individuals, and 11 amplified DNA fragments (5.7%) represented polymorphisms in the uninvolved human breast tissues. Reductions and enhancements in the intensity of some of the amplified fragments were observed indicating allelic gains or losses in the breast tumor genome compared to the matched uninvolved tissue genome. The presence or absence of some of the amplified DNA fragments were observed in this study indicating homozygous deletions or insertions in the breast tumor DNA compared to the matched uninvolved tissue DNA. Notably, an insertion of a 1270 bp amplified fragment was observed in 81% (17 of 21) of the tumor samples using the primer, OPC04. This amplified fragment resolved into two, 1200 and 1300 bp, single-stranded amplified fragments on the denaturing sequencing gel. This separation into single-stranded fragments suggests that the amplified fragment contains a conformation that is semistable. The 1270 bp amplified fragment localizes to the q11.2 region of chromosome 17. Sequence analysis of this fragment showed a significant DNA base sequence similarity (93%) with one of the breast tumor-specific human EST. The similarity with EST sequences and RT-PCR analysis showed that a part of this amplified fragment is from the coding region of the genome. Any one of the events observed in this study could play an important role in the development of breast cancer or could occur during the clonal expansion of the genetically unstable breast cells.

Smith, C. A., M. J. Smith, et al. (1999). "Gene expression during gonadogenesis in the chicken embryo." Gene **234**(2): 395.

http://www.sciencedirect.com/science/article/B6T39-3WWDH80-T/2/39d4781ba8971fc4f15cbbfe02668e24

Strazzullo, M., T. Parisi, et al. (1998). "Characterization and genomic mapping of chimeric ERV9 endogenous retroviruses-host gene transcripts." Gene **206**(1): 77.

http://www.sciencedirect.com/science/article/B6T39-3S1PY3W-V/2/3bc4901c55f9580ae5d26036a80fc963

ERV9 is a low repeated family of human endogenous retroviral elements, which has close to 50 members, in addition to at least 4000 solitary LTRs. Previous work has shown that randomly selected LTRs can promote transcription of reporter genes, raising the possibility that these sequences may affect the expression of adjacent cellular genes. We performed Northern blot experiments using sequences from ERV9-LTR, and we observed a different pattern of expression in several different hemopoietic tumor cell lines. It is possible that by the result of a somatic integration event, or by virtue of their original dispersal in the genome, ERV9-LTRs may specifically induce the expression of different cellular sequences in different cell lineages. Here, we describe the identification and analysis of four chimeric cDNA clones isolated from the T-lymphoma Peer cell line, having a structure consistent with transcription initiation from an ERV9-LTR. All the cDNA clones represent transcripts derived from unique cellular sequences. We also report the genomic localization of these cDNA clones.

Sumathy, K., K. V. Desai, et al. (1997). "Isolation of Transforming Growth Factor-[beta]2 cDNA from a fish, Cyprinus carpio by RT-PCR." Gene **191**(1): 103.

http://www.sciencedirect.com/science/article/B6T39-3RM6VMT-H/2/5324121cd662b85469db143156268435

Transforming Growth Factors-[beta] (TGF-[beta]s) have been described in many vertebrate species of amphibians, aves and mammals. In this report we demonstrate the presence of TGF-[beta]2 in pisces. TGF-[beta]2 has been cloned from a fish, Cyprinus carpio, by RT-PCR using degenerate oligonucleotide primers. Sequence analysis of the amplified product and alignment of the deduced amino acid sequence with the human TGF-[beta]2 amino acid sequence revealed 81% and 93% identity in the precursor and the mature regions, respectively. The northern blot analysis of fish heart RNA shows a major messenger RNA species of about 8.0 kb and two messages of very low abundance of about 5.0 kb and 4.0 kb. The identification of TGF-[beta]2 isoform in Pisces and it's high degree of homology with the mammalian isoform suggests that among all TGF-[beta] isoforms, TGF-[beta]2 is the most conserved during evolution.

Sumiyama, K., T. Kitano, et al. (2000). "Gene diversity of chimpanzee ABO blood group genes elucidated from exon 7 sequences." Gene **259**(1-2): 75.

http://www.sciencedirect.com/science/article/B6T39-423RGKS-B/2/87487bde5f3d6c615349686e781da05d

Human and non-human primate ABO blood group genes show relatively large numbers of nucleotide differences. In this study, we determined exon 7 sequences for 10 individuals of common chimpanzee and for four individuals of bonobo to estimate nucleotide diversities among them. Sequence data showed the existence of chimpanzee specific 9-base deletion in the beginning of the exon 7 coding region. From a phylogenetic network of exon 7 sequences of ABO blood group genes for human, common chimpanzee, bonobo and gorilla, effects of parallel substitutions and/or some kinds of convergent events are inferred in the chimpanzee lineage. We also estimated nucleotide diversities for common chimpanzee and bonobo ABO blood group genes, and these values were 0.4% and 0.2%, respectively. These values are higher than that of most human genes.

Suzuki, M., T. Yamasaki, et al. (2004). "Cloning and reporter analysis of human mitochondrial phosphoenolpyruvate carboxykinase gene promoter." Gene **338**(2): 157.

http://www.sciencedirect.com/science/article/B6T39-4CXMYKG-4/2/718e37ddff59952a28096762744bc541

Phosphoenolpyruvate carboxykinase (PEPCK) is one of the key regulatory enzymes in gluconeogenesis. In human liver, PEPCK is about equally distributed in both cytosol (PEPCK-1) and mitochondria (PEPCK-2). The human pepck2 gene and cDNA have been reported, but the cloning of the promoter region of the pepck2 gene has not been elucidated yet. We isolated and characterized human genomic P1-artificial chromosome (PAC) clones carrying the human pepck2 gene promoter. The oligocapping method revealed that the transcriptional start point (tsp) of the human pepck2 gene is located at 97 bp upstream of the first adenine residue of the translation start site. We also determined the nucleotide sequence to 1819 bp upstream of tsp. Sequence analysis of this region revealed that it contained several potential regulatory elements, including five GC boxes and three CCAAT boxes.Reporter analysis using transient transfection with firefly luciferase synthetic gene indicated 5' flanking region up to 822 bp, and 317 bp upstream of tsp had transcriptional activity. These results suggest that these regions of the human pepck2 gene play an important role for its expression.

Svensson, M. E., H. Johannesson, et al. (2000). "The LAMB1 gene from the clubmoss, Lycopodium annotinum, is a divergent MADS-box gene, expressed specifically in sporogenic structures." Gene **253**(1): 31.

http://www.sciencedirect.com/science/article/B6T39-40VT3F0-4/2/780b6a6239a3396b195cf769d37cb19a

Tang, L., Y. J. Yoon, et al. (1998). "Characterization of the enzymatic domains in the modular polyketide synthase involved in rifamycin B biosynthesis by Amycolatopsis mediterranei." Gene **216**(2): 255.

http://www.sciencedirect.com/science/article/B6T39-3TJ452W-4/2/e88f7245228884526fd3e575fbb3c4d7

Five clustered polyketide synthase (PKS) genes, rifA-rifE, involved in rifamycin (Rf) biosynthesis in Amycolatopsis mediterranei S699 have been cloned and sequenced (August, P.R. et al., 1998. Chem. Biol. 5, 69-79). The five multifunctional polypeptides constitute a type I modular PKS that contains ten modules, each responsible for a specific round of polyketide chain elongation. Sequence comparisons of the Rf PKS proteins with other prokaryotic modular PKSs elucidated the regions that have an important role in enzyme activity and specificity. The [beta]-ketoacyl:acyl carrier protein synthase (KS) domains show the highest degree of similarity between themselves (86-90%) and to other PKSs (78-85%) among all the constituent domains. Both malonyl-coenzyme A (MCoA) and methylmalonyl-coenzyme A (mMCoA) are substrates for chain elongation steps carried out by the Rf PKS. Since acyltransferase (AT) domains of modular PKSs can distinguish between these two substrates, comparison of the sequence of all ten AT domains of the Rf PKS with those found in the erythromycin (Er) (Donadio, S. and Katz, L., 1992. Gene 111, 51-60) and rapamycin (Rp) (Haydock, S. et al., 1995. FEBS Lett. 374, 246-248) PKSs revealed that the AT domains in module 2 of RifA and module 9 of RifE are specific for MCoA, whereas the other eight modules specify mMCoA. Dehydration of the [beta]-hydroxyacylthioester intermediates should occur during the reactions catalysed by module 4 of RifB and modules 9 and 10 of RifE, yet only the active site region of module 4 conforms closely to the dehydratase (DH) motifs in the Er and Rp PKSs. The DH domains of modules 9 and 10 diverge significantly from the consensus sequence defined by the Er and Rp PKSs, except for the active site His residues. Deletions in the DH active sites of module 1 in RifA and module 5 in RifB and in the N- and C-terminal regions of module 8 of RifD should inactivate these domains, and module 2 of RifA lacks a DH domain, all of which are consistent with the proposed biosynthesis of Rf. In contrast, module 6 of RifB and module 7 of RifC appear to contain intact DH domains even though DH activity is not apparently required in these modules. Module 2 of RifA lacks a [beta]-ketoacyl:acyl carrier protein reductase (KR) domain and the one in module 3 has an apparently inactive NADPH binding motif, similar to one found in the Er PKS, while the other eight KR domains of the Rf PKS should be functional. These observations are consistent with biosynthetic predictions. All the acyl carrier protein (ACP) domains, while clearly functional, nevertheless have active site signature sequences distinctive from those of the Er and Rp PKSs. Module 2 of RifA has only the core domains (KS, AT and ACP). The starter unit ligase (SUL) and ACP domains present in the N-terminus of RifA direct the selection and loading of the starter unit, 3-amino-5-hydroxybenzoic acid (AHBA), onto the PKS. AHBA is made by the products of several other genes in the Rf cluster through a variant of the shikimate pathway (August, P.R. et al., inter alia). RifF, produced by the gene immediately downstream of rifE, is thought to catalyse the intramolecular cyclization of the PKS product, thereby forming the ansamacrolide precursor of Rf B.

Tekki-Kessaris, N., J. V. Bonventre, et al. (1999). "Characterization of the mouse Kid1 gene and identification of a highly related gene, Kid2." Gene **240**(1): 13.

http://www.sciencedirect.com/science/article/B6T39-3YHWR69-2/2/40f600119a706601d2ae1e44060de742

Thomas, C. P., S. D. Auerbach, et al. (1999). "The structure of the rat amiloride-sensitive epithelial sodium channel gamma subunit gene and functional analysis of its promoter." Gene **228**(1-2): 111.

http://www.sciencedirect.com/science/article/B6T39-3W07P0S-D/2/d1b4b1fbe6dd3e15683fe47128c38ac1

Thorpe, K. L., P. Gorman, et al. (1997). "Chromosomal localization, gene structure and transcription pattern of the ORFX gene, a homologue of the MHC-linked RING3 gene." Gene **200**(1-2): 177.

http://www.sciencedirect.com/science/article/B6T39-3Y9GDS5-P/2/f99635e80796efe42d3ee4fee9c56689

We have mapped the human ORFX gene to chromosome 9q34 and determined its complete gene structure. Comparison with RING3, the human MHC-linked homologue on 6p21.3, shows the two gene structures to be highly conserved but with an approximate threefold expansion in the ORFX introns. RING3 and ORFX are found to be ubiquitously expressed in human adult and foetal tissues. Evidence suggests that the two genes may have arisen from an ancient duplication in a common ancestral chromosome.

Timms, K. M., M. A. Ansari-Lari, et al. (1998). "The genomic organization of Isopeptidase T-3 (ISOT-3), a new member of the ubiquitin specific protease family (UBP)." Gene **217**(1-2): 101.

http://www.sciencedirect.com/science/article/B6T39-3V7WV9H-D/2/f633eebf286da8bd447e14314bdd71ba

A novel Isopeptidase T gene (ISOT-3) has been identified on human chromosome 3q26.2-q26.3. This gene shows 67.3% nucleotide identity and 54.8% amino acid identity to human Isopeptidase T-1 (ISOT-1). Northern blot analysis has shown that ISOT-3 is highly expressed in ovary and testes, with low-level expression in six other tissues tested. In contrast, ISOT-1 is expressed at high levels in the brain, and there is no detectable expression in ovary. The exonic organization of these two genes is highly conserved with only one variant intron position. Intron 15 in ISOT-3 is absent in ISOT-1, but there is an alternate splice site at the same location. Although the exon-intron structure has been conserved between the two genes, ISOT-3 has significantly larger intronic regions, and the overall size of this gene is at least 90 kb compared to 15 kb for ISOT-1. These data suggest that both ISOT-1 and ISOT-3 have descended from a common ancestor. In addition, the low overall sequence identity and different expression patterns may reflect differences in substrate specificity.

Tolner, B., K. Roy, et al. (1998). "Structural analysis of the human RFC-1 gene encoding a folate transporter reveals multiple promoters and alternatively spliced transcripts with 5' end

heterogeneity." Gene **211**(2): 331.

http://www.sciencedirect.com/science/article/B6T39-3YJB847-M/2/79221ac0df8a97510d6fb0c5afa17f7d

The organization and structure of the human RFC-1 gene encoding a folate transporter were determined. The RFC-1 gene spans 22.5 kb and was found to be distributed in eight exons, including five primary exons and three alternatives of exon 1. Most splice junctions conform to consensus sequences for such junctions. The human RFC-1 gene differs from the mouse and hamster genes both in terms of the total number of exons and in regard to alternatives of exon 1 which encode 5' end heterogeneity. Previously described cDNA variants (GenBank/EMBL accession no. U19720) are now shown to incorporate one of two alternatives (exons 1a and 1b) to exon 1 and exons 2-6 as a result of RNA splicing. Another variant also described may not be full length in that it incorporates a probable alternative (exon 1c) to exon 1 along with exon 2 and a truncated exon 3. A relatively GC- rich region of the genome 5' of the alternatives to exon 1 appears to be distinctly promoter like and incorporates a number of putative cis-acting elements, including multiple SP1 sites, involved in the regulation of transcription. Primer extension analysis of this upstream region in two human cell types revealed a similar pattern of multiple transcription start sites (tsp) proximal to the 5' end of exon 1. However, there was a greater number of potential tsp within the region immediately upstream of exon 1b than within the regions upstream of exons 1a and 1c. The existence of true alternatives to exon 1 in this gene incorporating different 5' ends indicates that its transcription is under the control of multiple promoters. The identity of two such promoters was obtained by functional deletion analysis, showing that expression of a luciferase reporter gene was directed separately by discrete stretches of nucleotide sequence proximal to exon 1a (promoter 1) or exon 1b (promoter 2) in transient transfection experiments. Promoter 1 appeared to have a three-fold lower basal activity than promoter 2, but was enhanced up to nine-fold in fusion constructs containing an SV40 enhancer element. Also, promoter 2 partly consists of a highly GC-rich direct repeat element containing at least three putative SP-1 and 3 putative MZF1 sites. Finally, the activity of these promoters relative to each other was consistent with the results of primer extension analysis showing a greater multiple and usage of tsp within promoter 2 (exon 1b) than within promoter 1 (exons 1a and 1c), suggesting that the variant incorporating exon 1b was the most abundant.

Toramoto, T., D. Ikeda, et al. (2004). "Multiple gene organization of pufferfish Fugu rubripes tropomyosin isoforms and tissue distribution of their transcripts." Gene **331**: 41.

http://www.sciencedirect.com/science/article/B6T39-4BYR95B-1/2/1005179671831b74eab27d37b3e26951

The Japanese pufferfish, torafugu (Fugu rubripes), has a haploid genome of about 400 Mb in size, which has been sequenced to approximately 90% coverage. Here we identified six Fugu tropomyosin (TPM) gene sequences by using the BLASTN program and the sequence of the white croaker TPM1 gene in our collection against the draft assembly of the Fugu genomic sequence database. TPM2, TPM3 and TPM4 genes were identified together with a set of two potentially duplicated genes of TPM1 (TPM1-1 and TPM1-2) as described in our previous report and TPM4 (TPM4-1 and TPM4-2) newly found in this study. The expression patterns of these Fugu TPM genes were determined by reverse transcription polymerase chain reaction (RT-PCR). A phylogenetic tree was constructed using the deduced amino acid sequences, which were encoded by the exons common to all vertebrate TPM genes. This indicated that the Fugu TPM1 and TPM4 genes had resulted from a gene duplication in the fish evolutionary lineage.

Tozaki, T., K.-i. Hirota, et al. (2005). "Prospects for whole genome linkage disequilibrium mapping in

thoroughbreds." Gene **346**: 127.

http://www.sciencedirect.com/science/article/B6T39-4F9N6YJ-1/2/14b62b360ea2e592adc5f7efafb4d6db

Linkage disequilibrium (LD) mapping is often used in searches for genes governing economically significant traits and diseases. The D' coefficient is a commonly used measure of the extent of LD between all possible pairs of alleles at two markers. This study aimed to test the utility of the D' coefficient for LD mapping of a trait in a thoroughbred population. Microsatellite genotype data and grey coat colour as a trait model in a thoroughbred population were used to assess the extent of LD. We demonstrated that LD mapping was a reasonable approach for initial genome-wide scans in a thoroughbred population. Significant LD was demonstrated at approximately 7 cM, implying that roughly 430 appropriately spaced microsatellites were needed for systematic whole-genome LD mapping in this model. LD mapping methods using D' in a thoroughbred population were useful for identifying the chromosomal regions for diseases and economic trait loci (ETL). It was suggested that a thoroughbred population represented a population particularly suitable for LD mapping.

Tribioli, C. and T. Lufkin (1997). "Molecular cloning, chromosomal mapping and developmental expression of BAPX1, a novel human homeobox-containing gene homologous to Drosophila bagpipe." Gene **203**(2): 225.

http://www.sciencedirect.com/science/article/B6T39-3RM6VMT-19/2/3a77499018895f4fd7ed7e1c5f00d666

We describe here the cloning of the human BAPX1 gene, a homologue of the Drosophila bagpipe gene which has 87% aa identity within the homeodomain relative to the fly gene. We recently have identified the murine bagpipe homolog. The predicted aa sequence of the human gene has 85% overall identity to the murine gene, with 100% identity in the homeodomain. In mouse, this gene maps to the proximal portion of chromosome 5. We show that the human gene maps to 4p16.1, the human region syntenic with mouse chromosome 5. Expression of BAPX1 was evaluated during human embryonic development by RT-PCR analysis and by RNA in situ hybridization. RT-PCR analysis showed that BAPX1 is expressed in embryo tissues, particularly the limb, and at a lower level in an embryonic lung cell line. RNA in situ hybridization revealed that BAPX1 is predominantly expressed in mesenchymal condensations of the fetal limb and axial skeleton, and in lateral plate mesoderm giving rise to visceral muscle. The expression pattern of BAPX1 combined with the chromosomal localization to 4p16.1, where several human genetic diseases involving dysmorphology of the skeleton have been assigned, raises the potential of it being a candidate gene for one of these disorders.

Trueman, L. J., A. Richardson, et al. (1996). "Molecular cloning of higher plant homologues of the high-affinity nitrate transporters of Chlamydomonas reinhardtii and Aspergillus nidulans." Gene **175**(1-2): 223.

http://www.sciencedirect.com/science/article/B6T39-3W25BFT-18/2/3e44a525c22b98655965cf93a9652f17

The crnA nitrate transporter from Aspergillus nidulans was identified as belonging to the major facilitator superfamily (MFS) of membrane transporters. Degenerate oligonucleotides corresponding to the crnA sequences at the locations of two conserved sequence motifs were designed and used in the polymerase chain reaction (PCR) to amplify related sequences from

barley root poly(A)+ RNA. A 130 bp cDNA fragment with sequence similarities to crnA was amplified and used as a probe to screen a barley root cDNA library. Two full-length clones (pBCH1 and pBCH2) were isolated. The nt sequences of pBCH1 and pBCH2 are closely related (80% identical) and potentially encode hydrophobic polypeptides of 54.7 and 55.0 kDa respectively, with twelve predicted transmembrane domains. The encoded polypeptides are 41-43% identical to the A. nidulans CRNA protein and 56-57% identical to NAR-3, a high-affinity nitrate transporter from the eukaryotic alga Chlamydomonas reinhardtii. Phylogenetic analysis indicated that crnA, nar-3 and the barley homologues belong to a new family within the MFS, a family that also includes narK, the gene for a nitrite efflux pump in Escherichia coli. In northern blots, BCH1 hybridised to a mRNA species of 1.9 kb which is rapidly induced in barley roots by NO3-, but not by NH4+, and genomic Southern blots indicated that there may be seven to ten BCH1-related genes in the barley genome.

van den Berg, L., S. Imholz, et al. (2004). "Isolation and characterization of the canine serotonin receptor 1B gene (htr1B)." Gene **326**: 131.

http://www.sciencedirect.com/science/article/B6T39-4B8BX05-3/2/f474301cbf1df8838a88593e2fa85723

The serotonin receptor 1B gene (htr1B) has been suggested to be implicated in mental disorders in both humans and other species. We have isolated a canine bacterial artificial chromosome (BAC) clone containing htr1B, revealed the coding and surrounding DNA sequence of canine htr1B and designed primer sets for genomic sequencing of the gene. A mutation scan in 10 dogs revealed five single nucleotide polymorphisms in the htr1B coding sequence. By random sequencing of subclones of the BAC a polymorphic microsatellite repeat was found. We found evidence for at least four extended haplotypes in six dogs of the same breed. The chromosomal localization of the gene was confirmed by fluorescence in situ hybridisation and radiation hybrid mapping. This work provides a starting point for mutation scans and association studies on dogs with behavioural problems.

Varadaraj, K. and D. M. Skinner (1994). "Denaturants or cosolvents improve the specificity of PCR amplification of a G + C-rich DNA using genetically engineered DNA polymerases." Gene **140**(1): 1.

http://www.sciencedirect.com/science/article/B6T39-47PNX4N-9Y/2/56f7893627e7850db854e4a61aeea93e

We describe conditions that improve the specificity of amplification of a G + C-rich (57% G + C) DNA by PCR. Under standard conditions a 368-bp segment of the approx. 2.1-kb repeat unit of a satellite DNA that accounts for approx. 3% of the genome of the Bermuda land crab, Gecarcinus lateralis, was not amplified specifically. To establish optimal conditions for amplification of the segment of the G + C-rich satellite, we used two genetically engineered enzymes, AmpliTaq DNA polymerase and AmpliTaq DNA polymerase. Stoffel fragment (SF), and a number of denaturants or co-solvents. In the absence of denaturants or co-solvents, amplified products of both enzymes contained non-specific bands upon gel electrophoresis. Addition of certain denaturants or co-solvents to PCR mixtures resulted in the production of the single specific band of the expected size. Reagents that improved specificity of the amplified product were formamide, glycerol, DMSO, Tween-20 and NP-40; on the other hand, urea, ethanol and 1-methyl-2-pyrrolidone (NMP) inhibited amplification. Of the two enzymes, SF was more specific and efficient. The products of AmpliTaq DNA polymerase included one or more extra bands, even in the presence of denaturants or co-solvents, except for glycerol or DMSO.

Vazza, G., S. Picelli, et al. (2003). "Identification and characterization of C3orf6, a new conserved human gene mapping to chromosome 3q28." Gene **314**: 113.

http://www.sciencedirect.com/science/article/B6T39-49D256V-3/2/8c77a4bbc7ac0b349c1ca68d18e328ab

This study reports the characterization of a novel human gene, chromosome 3 open reading frame 6 (C3orf6), mapped to chromosome 3q28, within the critical region of hereditary spastic paraplegia SPG14 locus. Based on computational "spliced" EST alignment and RT-PCR, two C3orf6 transcript variants were identified. The longer C3orf6 transcript contains a 1449-nt ORF, encoding a protein of 482 aa, while the shorter variant contains a 921-nt ORF, encoding for a protein of 306 aa. C3orf6 gene is organised on 12 exons and the shorter transcript comes from an alternative splicing event skipping exon 6. The two mRNA are differentially expressed in brain and in several other human tissues with a predominant level for the shorter transcript. By database analysis, EST assembling and RT-PCR, we identified the transcripts of mouse and rat C3orf6 orthologous genes. The involvement of C3orf6 in the spastic paraplegia was investigated by sequencing all coding exons and flanking sequences in the SPG14 family, excluding the presence of causative mutations.

Ventura, M., M. Boniotto, et al. (2001). "Characterization of a highly repeated DNA sequence family in five species of the genus Eulemur." Gene **275**(2): 305.

http://www.sciencedirect.com/science/article/B6T39-44357HF-F/2/67f92ffe7ca6abe10b4c4f8bda4b6d7b

The karyotypes of Eulemur species exhibit a high degree of variation, as a consequence of the Robertsonian fusion and/or centromere fission. Centromeric and pericentromeric heterochromatin of eulemurs is constituted by highly repeated DNA sequences (including some telomeric TTAGGG repeats) which have so far been investigated and used for the study of the systematic relationships of the different species of the genus Eulemur. In our study, we have cloned a set of repetitive pericentromeric sequences of five Eulemur species: E. fulvus fulvus (EFU), E. mongoz (EMO), E. macaco (EMA), E. rubriventer (ERU), and E. coronatus (ECO). We have characterized these clones by sequence comparison and by comparative fluorescence in situ hybridization analysis in EMA and EFU. Our results showed a high degree of sequence similarity among Eulemur species, indicating a strong conservation, within the five species, of these pericentromeric highly repeated DNA sequences.

Vigetti, D., C. Monetti, et al. (2002). "Genomic organization and chromosome localization of the murine and human allantoicase gene." Gene **289**(1-2): 13.

http://www.sciencedirect.com/science/article/B6T39-45G023J-6/2/84f779cb1c21538f742b9b18201ba162

Allantoicase is one of the enzymes involved in uricolysis. The enzymes of this catabolic pathway (i.e. allantoinase, allantoicase, ureidoglycolate lyase and urease) were lost during vertebrate evolution and the causes for this loss are still unclear. In mammals, as well as in birds and reptiles, the activity of allantoicase is absent; notwithstanding, we recently cloned human and mouse cDNA sequences with high similarity with previously characterized allantoicases. In the present paper, we report the genomic organization of the allantoicase gene in mouse and in man.

Both genes are constituted by 11 exons that appear to be very conserved; introns are more variable in length while maintain the same phase but for intron 4. We have also detected a second transcript of the human allantoicase gene in which exon 1 is absent. Moreover, the mouse gene maps in chromosome 12 at 13.0 cM from the centromere.

Wadskov-Hansen, S. L. L., J. Martinussen, et al. (2000). "The pyrH gene of Lactococcus lactis subsp. cremoris encoding UMP kinase is transcribed as part of an operon including the frr1 gene encoding ribosomal recycling factor 1." Gene **241**(1): 157.

http://www.sciencedirect.com/science/article/B6T39-3Y51FGV-N/2/b3d34d6e35c9de768ba4f85673ff879d

Walston, J., A. Lowe, et al. (1997). "The [beta]3-adrenergic receptor in the obesity and diabetes prone rhesus monkey is very similar to human and contains arginine at codon 64." Gene **188**(2): 207.

http://www.sciencedirect.com/science/article/B6T39-3RD1R45-14/2/6f22ac00538fcb89501d848b52f6f272

The [beta]3-adrenergic receptor (ADR[beta]3) is a seven-membrane spanning, G-protein linked receptor expressed in brown adipose tissue in rodents, and visceral adipose tissue in humans. Stimulation of the receptor by norepinephrine leads to lipolysis and thermogenesis. In rodent models of obesity and diabetes, administration of [beta]3-agonists results in weight loss and improved glucose tolerance. Studies indicate that the pharmacological properties of the ADR[beta]3 differ markedly between rodents and humans, making generalizations of rodent studies to humans difficult. We hypothesized that the obesity and diabetes prone rhesus monkey (Macaca mulatta) would provide an excellent animal model to study the role of the ADR[beta]3 in the development of obesity and diabetes as well as for assessment of the therapeutic efficacy of [beta]3-agonists. We sequenced the entire coding region of the rhesus ADR[beta]3 gene. Like humans, the rhesus ADR[beta]3 has two exons. There is 89% amino acid (aa) identity between human and rhesus compared to 82% aa identity between human and mouse. A single base deletion results in divergence of the intracellular carboxy terminus accounting for 26 of the 45 aa changes and 10 additional aa. Of the 15 rhesus monkeys studied, all were homozygous for Arg64. In humans, Arg64 (rather than Trp) is associated with increased body mass index, insulin resistance, and an earlier onset of type II diabetes mellitus. We conclude that the rhesus ADR[beta]3 is more similar to the human ADR[beta]3 than to the rodent ADR[beta]3 suggesting that this primate model may be more appropriate for physiologic and therapeutic studies of the ADR[beta]3 axis, and that Arg64 may influence susceptibility in this species to obesity, insulin resistance, and type II diabetes.

Wang, W., Q. Zhang, et al. (1999). "Molecular characterization and mapping of canine cGMP-phosphodiesterase delta subunit (PDE6D)." Gene **236**(2): 325.

http://www.sciencedirect.com/science/article/B6T39-3X64DJG-F/2/fc5d0c0ca155ffa50af07566a1877286

Wasenius, V.-M., J. Merilainen, et al. (1993). "Sequence of a chicken cDNA encoding a GRB2 protein." Gene **134**(2): 299.

http://www.sciencedirect.com/science/article/B6T39-47PH62P-12K/2/3d87c627798d2b2e4fb4fd69e15fe815

The nucleotide (nt) sequence of a chicken cDNA encoding a protein homologous to the human GRB2 (growth factor receptor-binding protein) was determined. Remarkably high identities were found on the nt (88%) and deduced amino acid sequence (96%) levels.

Watanabe, Y., T. Tenzen, et al. (2000). "Replication timing of the human X-inactivation center (XIC) region: correlation with chromosome bands." Gene **252**(1-2): 163.

http://www.sciencedirect.com/science/article/B6T39-40RTKY5-J/2/fa00a703bd07a7e0694314e0c02ae606

Westendorf, J. J., R. Mernaugh, et al. (1999). "Identification and characterization of a protein containing formin homology (FH1/FH2) domains." Gene **232**(2): 173.

http://www.sciencedirect.com/science/article/B6T39-3WHKRH2-4/2/e0ab8879760adecf2c3d753c9ed349fa

Western, P. S., J. L. Harry, et al. (2000). "Temperature-dependent sex determination in the American alligator: expression of SF1, WT1 and DAX1 during gonadogenesis." Gene **241**(2): 223.

http://www.sciencedirect.com/science/article/B6T39-3Y6GVD3-4/2/b5447e9973c906d1e148dc69429cc8a1

Wieland, K.-P., B. Wieland, et al. (1995). "A promoter-screening plasmid and xylose-inducible, glucose-repressible expression vectors for Staphylococcus carnosus." Gene **158**(1): 91.

http://www.sciencedirect.com/science/article/B6T39-40PRVR9-K/2/efa84ad50d9e743c6a1e23f71a1809d4

We describe a series of plasmid vectors for DNA cloning in staphylococci. pPSII is a promoter probe plasmid containing a promoterless lipase (Lip)-encoding gene (lip). Insertion of a promoter-bearing DNA fragment at the single BamHI site turns on lip expression. Lip activity can be easily determined to estimate the strength of the inserted promoter. pPSII served also as a basis for the construction of vectors which allow xylose-inducible gene expression in Staphylococcus carnosus (Sc). Using plasmid pCX15, we studied xylose-inducible lip expression in Sc. The lip expression is under transcriptional control of the repressor, XylR. The xylR gene, the XylR target sequence and the xylA promoter/ operator sequence with the cis-acting catabolite-responsive element (cre) are derived from the xyl operon of S. xylosus. The single BamHI site in front of the lip ribosome-binding site (RBS) also makes it possible to put other promoterless genes under transcriptional control of XylR. To facilitate the controlled expression of genes which merely start with the start codon and have no RBS, or to insert genes with their own RBS, pCX26 and pCX26[Delta]lip were constructed. The influence of xylose and glucose on lip expression was studied both in a batch culture and in a fermentor under controlled pH conditions. With pCX15, the presence of xylose led to a 40-fold increase in extracelluar Lip activity, while the presence of glucose caused a repression of lip expression. The results suggest that the xylA promoter is subject to two different

regulatory mechanisms, one of which involves the repression of the xylA promoter by XylR in the absence of xylose, and the other involves a glucose-mediated catabolite repression which dominates over the xylose induction.

Xie, G.-x., E. Ito, et al. (1999). "The promoter region of human prepro-nociceptin gene and its regulation by cyclic AMP and steroid hormones." Gene **238**(2): 427.

http://www.sciencedirect.com/science/article/B6T39-3YDNWHX-1C/2/05bfae0ecc2324628e12ec134cd06db1

Xu, Z., D. M. Jablons, et al. (2001). "Expression sequence tag-specific full-length cDNA cloning: actin cDNAs." Gene **263**(1-2): 265.

http://www.sciencedirect.com/science/article/B6T39-42D2CGD-11/2/1760ad3dac7a96ad14b0db7820c8b505

Current strategies for cDNA cloning are based on construction of cDNA libraries and colony screening. The process of obtaining a full-length cDNA clone can be highly time and labor intensive. Using the human actin gene as a model target cDNA, we have developed an RNA-capture method for rapid cloning of full-length cDNAs. The approach involves the capture of mRNA with expressed sequence tag (EST)-derived, biotin labeled antisense 'capture' primers and streptavidin-coated magnetic beads. Full-length cDNA is then synthesized from purified EST-specific mRNA and cloned directly into plasmid vectors. The results of using [beta]-actin-based capture primers on cytoplasmic RNA were the isolation of both [beta]- and [gamma]-actin cDNA clones. Of the 16 actin-specific cDNA clones analyzed, 15 (93%) were full-length. This approach for cloning full-length cDNAs from available ESTs or partial cDNA sequences will facilitate a more rapid and efficient characterization of gene structure and function.

Xue, H., D. O'Neill, et al. (1999). "A novel mouse gene, HemT, encoding an hematopoietic cell-specific transcript." Gene **231**(1-2): 49.

http://www.sciencedirect.com/science/article/B6T39-3WD5BHJ-6/2/e4132f6ed4c5c15896f75affee5a5450

Yamada, S., S. Murakami, et al. (2001). "Expression profile of active genes in human periodontal ligament and isolation of PLAP-1, a novel SLRP family gene." Gene **275**(2): 279.

http://www.sciencedirect.com/science/article/B6T39-44357HF-B/2/cf98395a21a92e0d1301b7c97b9be969

Periodontal ligament (PDL) is one of the most important tissues in maintaining the homeostasis of tooth and tooth-supporting tissue, periodontium. In this study, we investigated the expression profile of active genes in the human PDL obtained by collecting sequences with a 3'-directed cDNA library, which faithfully represents the composition of the mRNA population. We succeeded in obtaining a total of 1752 cDNA sequences by sequencing randomly selected clones and identified a total of 1318 different species as gene signatures (GS) by their sequence identity, 344 of which were known genes in the GenBank, and 974 of which were new genes. The resulting

expression profile showed that collagen type I and type III were the most abundant genes and that osteogenesis-related proteins, such as SPARC/osteonectin and osteoblast specific factor 2, were highly expressed. By comparing the expression profile of PDL with 44 profiles similarly obtained with unrelated human cell/tissue, nine novel genes, which are probably expressed specifically in PDL, were discovered. Among them, we cloned a full-length cDNA of GS5096, which is frequently expressed in freshly-isolated periodontal tissue. We found that it encodes a novel protein, which is a new member of the class I small leucine-rich repeat proteoglycan family, and designated it PLAP-1 (periodontal ligament associated protein-1). PLAP-1 mRNA expression was confirmed in in vitro-maintained PDL cells and was enhanced during the course of the cytodifferentiation of the PDL cells into mineralized tissue-forming cells such as osteoblasts and cementoblasts. These findings suggest the involvement of PLAP-1 in the mineralized matrix formation in PDL tissues.

Yamaki, A., J. Tochigi, et al. (2001). "Molecular mechanisms of human single-minded 2 (SIM2) gene expression: identification of a promoter site in the SIM2 genomic sequence." Gene **270**(1-2): 265.

http://www.sciencedirect.com/science/article/B6T39-436W3KY-10/2/3a8462abb92cf2315d8430181be55f8d

We previously postulated that the single-minded 2 (SIM2) gene identified on the human chromosome 21q22.2 is a good candidate gene for the pathogenesis of mental retardation in Down syndrome because its mouse homolog exhibits preferential expression in the mouse diencephalon during early embryogenesis. We analyzed the genomic sequence of the entire SIM2 gene which consists of 11 exons and spans over 50 kb. As a step toward understanding the molecular mechanisms of SIM2 gene expression, we have analyzed the human SIM2 gene expression in nine established human cell lines. Three transcripts of 3.6, 4.4, and 6.0 kb were detected in the glioblastoma cell line, T98G, neuroblastoma cell line, TGW, and transformed embryonic kidney cell line, 293. The RACE analysis using SIM2-expressing human cell line T98G provided evidence for the transcription start site at ~1.2 kb upstream of the translation initiation site. The transfection assay using various deletion constructs with reporter gene suggested the presence of a presumptive promoter region. Transient transfection assay in T98G cell line revealed a significant promoter activity located in the 60 bp sequence between nt -1385 and -1325 upstream region of the translation initiation site. This 60 bp sequence contains cis-elements for c-myb, E47 and E2F transcription factors. Moreover, the gel retardation assay using oligo-DNA of various cis-element sequences indicated the presence of protein factor(s) which bind to the cis-element for c-myb. These results suggested that binding of a protein transcription factor(s) such as c-myb or that alike regulates transcription of the SIM2 gene by binding to a small upstream region.

Yang, S. and S. F. Parmley (1997). "Toxoplasma gondii expresses two distinct lactate dehydrogenase homologous genes during its life cycle in intermediate hosts." Gene **184**(1): 1.

http://www.sciencedirect.com/science/article/B6T39-3RD1R45-7F/2/4cfff88aa996bd976d5387764c1d11c4

Two Toxoplasma gondii genes were characterized that are differentially expressed during the parasite's life cycle. The genes named LDH1 and LDH2, respectively, encode polypeptides similar to the enzyme lactate dehydrogenase (LDH; -lactate:NAD+ oxidoreductase, EC 1.1.1.27) from a variety of organisms. They show 64.0% nucleotide identity in the coding region and both have an intron at the same relative position. The deduced amino acid sequences of LDH1 and LDH2 share 71.1% identity. LDH1 and LDH2 are most similar to an LDH of Plasmodium falciparum (46.5% and 48.5% amino acid identities, respectively). The mRNA of LDH2 was only

detected in the bradyzoite stage, while the mRNA of LDH1 was detected in both the bradyzoite and tachyzoite stages. However, by isoelectric focusing and immunoblot analysis, only one LDH isoform was found to be expressed in each stage. Furthermore, the expression of a reporter gene carrying chloramphenicol acetyltransferase (CAT) coding sequence and the putative LDH2 promoter sequence was significantly up-regulated by growing parasites in tissue culture in media with alkaline pH (pH 8.2, a condition known to induce the expression of bradyzoite-specific antigens), while the expression of a CAT reporter construct carrying the putative LDH1 promoter sequence was down-regulated by similar treatment. These results indicate that LDH expression is developmentally regulated in T. gondii and suggest a possible correlation between stage conversion and alteration in carbohydrate or energy metabolism in this parasite.

Yin, X., T. O'Hare, et al. (2003). "Identification and cloning of genes encoding viomycin biosynthesis from Streptomyces vinaceus and evidence for involvement of a rare oxygenase." Gene **312**: 215.

http://www.sciencedirect.com/science/article/B6T39-491J42W-3/2/5a93f6358ad740c284f7c891696f0633

The tuberactinomycins are a family of basic cyclic peptides that exhibit potent antitubercular activity. These peptides are characterized by the presence of an amino acid with a 6-membered cyclic guanidine side chain (capreomycidine) and two or more 2,3-diaminopropionate residues. Viomycin (tuberactinomycin B) is a well-studied member of the family, was once prescribed for the treatment of tuberculosis, and has been shown to block translocation during protein biosynthesis. The gene cluster encoding viomycin biosynthesis was identified and cloned from Streptomyces vinaceus. The cluster was identified by screening genomic libraries with the viomycin phosphotransferase self-resistance gene (vph) and non-ribosomal peptide synthetase (NRPS) gene probes amplified from S. vinaceus genomic DNA. The viomycin cluster was localized to ca. 120 kb of contiguous DNA defined by four overlapping cosmid inserts. Each cosmid hybridized with one or more peptide synthetase gene probes and two also hybridized with vph. Confirmation that the cluster encoded viomycin biosynthesis was obtained from the disruption of two NRPS adenylation domains. Partial sequence analysis revealed an ORF (svox) predicted to encode a rare non-heme iron, [alpha]-ketoglutarate dependent oxygenase proposed to function in the oxidative cyclization of arginine to the capreomycidine residue. Insertional disruption of svox resulted in complete loss of viomycin production, confirming its involvement in the pathway.

Yoshida, S., M. Taniguchi, et al. (1998). "Sequence analysis and expression of human neuropsin cDNA and gene." Gene **213**(1-2): 9.

http://www.sciencedirect.com/science/article/B6T39-3VGR20J-2/2/608448a61d65ac3bc5636ef48e60c62b

Neuropsin is a serine protease which is thought to function in a variety of tissues including the brain and skin. This protease has been shown to have important roles in neural plasticity in mice. Here we have cloned a cDNA and analyzed the gene for human neuropsin by polymerase chain reaction-based strategies. The cDNA had 72% identity to mouse neuropsin. The deduced amino acid sequence showed 72% identity to mouse neuropsin. Key amino acid residues for the enzyme activity and all cysteine residues were conserved between human and mouse neuropsin. The gene for human neuropsin had six exons and five introns, and the gene organization is similar to trypsin-type serine proteases. The mRNA was expressed in primary cultures of keratinocytes.

Yousef, G. M., A. Scorilas, et al. (2000). "The KLK7 (PRSS6) gene, encoding for the stratum corneum chymotryptic enzyme is a new member of the human kallikrein gene family -- genomic characterization, mapping, tissue expression and hormonal regulation." Gene **254**(1-2): 119.

http://www.sciencedirect.com/science/article/B6T39-413KWGN-F/2/5be6ee63cfdc65bcbb49213b9bd6801f

Yowe, D. L. and B. N. Ames (1998). "Quantitation of age-related mitochondrial DNA deletions in rat tissues shows that their pattern of accumulation differs from that of humans." Gene **209**(1-2): 23.

http://www.sciencedirect.com/science/article/B6T39-3SRBHWR-3/2/7b053fad8ea937b3192cfca2af4094fe

Three age-related mtDNA deletions were identified, and the competitive polymerase chain reaction (PCR) was used to quantitate their levels in different Fisher 344 rat tissues. Deletions that removed 4834; 13273; or 13415 nt of the mitochondrial genome were shown to be associated with 16 (mtDNA4834), nine (mtDNA13273), or five (mtDNA13415) nt direct repeats, respectively. The mtDNA4834 deletion was detected in an age-related manner in all tissues screened; the mtDNA13415 deletion was detected in old heart, and in both young and old brain; and the mtDNA13273 deletion was only detected in old brain tissues. The mtDNA4834 deletion was found to be at its highest level in the liver (1.88 x 10-2%), followed by the brain (0.22 x 10-2%) and kidney (0.40 x 10-2%) of old animals. Much lower levels were observed in old heart (0.07 x 10-2%) and lung (0.04 x 10-2%). This distribution of mtDNA deletions in old rat tissues is in contrast to work done in humans where age-related deletions are present at the highest levels in post-mitotic tissues with much lower levels in more mitotic tissues. An inverse relationship was observed between the level of mtDNA deletions and the size of the deleted region, since the mtDNA13415 deletion was present at about a 100-fold lower level (0.53 x 10-5%) than the smaller mtDNA4834 deletion in old heart tissue.

Yuka, U., I. Harue, et al. (1991). "Nonradioactive labeling with chemically modified cytosine tails by the polymerase chain reaction." Gene **108**(1): 103.

http://www.sciencedirect.com/science/article/B6T39-47PNY80-X0/2/928a9526f380b837008e74dcd9636888

We developed a modified nonradioactive method for the detection of DNA. This method makes use of the polymerase chain reaction for preparation of probes; that is, a DNA fragment inserted in the polylinker region of an M13 or pUC vector is amplified with primers that have a modified cytosine tail at the 5' terminus (C-tailed primers). By this method, large amounts of labeled probes can be obtained easily. After hybridization, modified cytosine tails can be detected immunologically. DNA labeled by this method could be used in plaque hybridization. We could detect 0.05 pg of dot-blotted labeled DNA in 30 min with an enzyme-catalyzed chemiluminescence reaction.

Zaiss, D. and J. M. Belote (1997). "Molecular cloning of the Drosophila melanogaster gene [alpha]5_dm encoding a 20S proteasome [alpha]-type subunit." Gene **201**(1-2): 99.

http://www.sciencedirect.com/science/article/B6T39-3RD1R45-3N/2/be41d1df758d73615b71df6684703597

Proteasomes are large, multisubunit particles that act as the proteolytic machinery for most regulated intracellular protein breakdown in eukaryotic cells. The core proteinase of this complex, known as the 20S proteasome, is a hollow barrel-shaped structure made up of four stacked rings of seven subunits each, with the outer two rings each being made up of seven distinct [alpha]-type subunits, and the two inner rings composed of seven different [beta]-type subunits. Here we present the cloning, sequencing and genetic mapping of a Drosophila melanogaster gene, [alpha]5_dm, encoding one of the proteasome [alpha] subunits. This gene, which is homologous to the yeast PUP2 and the human Zeta genes, maps to chromosome 2 at position 54B3-5. The map positions of the previously cloned proteasome genes Pros25 and Pros29 were also determined, and found to lie at positions 87B and 57B, respectively. A search for other D. melanogaster [alpha]5_dm-like genes encoding potential isoforms of this subunit failed to identify any closely related genes.

Zeng, G., L. Gao, et al. (2005). "Expression of the mouse WNK1 gene in correlation with ganglioside GD3 and functional analysis of the mouse WNK1 promoter." Gene **344**: 233.

http://www.sciencedirect.com/science/article/B6T39-4DWHJ4Y-2/2/3cd09ad7c7a2fac603835f367e07745b

WNK1 is one of WNK (With No K=Lysine) protein kinases which comprise a newly described subfamily. Our studies showed that expression of the mouse WNK1 gene was dramatically suppressed in a tumor cell line when its phenotype was altered by suppression of the GD3-synthase gene expression. The mouse WNK1 gene was expressed at a high level at early stage of embryonic brain and its expression decreased as brain developed, similar to the expression pattern of the GD3-synthase gene. To study transcriptional regulation, we cloned a 5'-flanking 1239-bp fragment of the mouse WNK1 gene. This fragment contains a number of potential consensus binding sites for transcription factors, including Sp1, AP2, CCAAT, Est-1, Oct-1, CNBP, and NFkB, but lacks a TATA box. Primer extension identified multiple putative transcriptional initiation sites, including several sites downstream of the ATG codon. Activities of the promoter fragments were assessed in mouse breast Sa/R-MT cells by transient transfection and the results showed that the promoter elements between -700 and -977 is required for maintaining a high level of promoter activity of the TATA-less mouse WNK1 gene.

Zhang, J., W.-L. Liu, et al. (2002). "Identification and characterization of a novel member of olfactomedin-related protein family, hGC-1, expressed during myeloid lineage development." Gene **283**(1-2): 83.

http://www.sciencedirect.com/science/article/B6T39-44JJHYF-2/2/5c19e2831f383a621af717b029ed7eaf

We have cloned a novel hematopoietic granulocyte colony-stimulating factor (G-CSF)-induced olfactomedin-related glycoprotein, termed hGC-1 (human G-CSF-stimulated clone-1). mRNA differential display was used in conjunction with a modified two-phase liquid culture system. Cultures were enriched for early precursors of erythroid, myeloid, and megakaryocytic lineages, which were isolated after induction with erythropoietin, G-CSF, and thrombopoietin, respectively. RNA from the enriched cells was subjected to differential display analysis to identify lineage-specific expressed genes. One clone specifically induced by G-CSF, hGC-1, was characterized. The 2861 bp cDNA clone of hGC-1 contained an open reading frame of 1530 nucleotides,

translating into a protein of 510 amino acids with a signal peptide and six N-linked glycosylation motifs. The protein sequence of hGC-1 showed it to be a glycoprotein of the olfactomedin family, which includes olfactomedin, TIGR, Noelin-2 and latrophilin-1. Olfactomedin-like genes show characteristic tissue-restricted patterns of expression; the specific tissues expressing these genes differ among the family members. hGC-1 was strongly expressed in the prostate, small intestine, and colon, moderately expressed in the bone marrow and stomach, and not detectable in other tissues. In vitro translation and ex vivo expression showed hGC-1 to be an N-linked glycoprotein. The hGC-1 gene locus mapped to chromosome 13q14.3. Together, our findings indicate that hGC-1 is primarily expressed as an extracellular olfactomedin-related glycoprotein during normal myeloid-specific lineage differentiation, suggesting the possibility of a matrix-related function for hGC-1 in differentiation.

Zhang, K., N. Dion, et al. (2002). "The human homolog of yeast SEP1 is a novel candidate tumor suppressor gene in osteogenic sarcoma." Gene **298**(2): 121.

http://www.sciencedirect.com/science/article/B6T39-473HYTB-B/2/d217cb862551b806b0def9c5b2520c5e

The hSEP1 gene is the human homolog of yeast SEP1. Yeast SEP1 is a multifunctional gene that regulates a variety of nuclear and cytoplasmic functions including homologous recombination, meiosis, telomere maintenance, RNA metabolism and microtubule assembly. The function of hSEP1 is not known. We show loss or reduced expression of hSEP1 messenger RNA (mRNA) in three of four primary osteogenic sarcoma (OGS)-derived cell lines and in eight of nine OGS biopsy specimen. In addition, we find a heterozygous missence mutation (Valine1484>Alanine) at a conserved amino acid in the primary OGS-derived cell line U2OS. Importantly, we identified a homozygous missense mutation involving a CG-dinucleotide leading to a change in a conserved amino acid, aspartic acid1137 >asparagine, in the primary OGS-derived cell line, TE85. hSEP1 mRNA expression was nearly undetectable in TE85 and low in U2OS cell lines. None of these mutations were identified in 20 normal samples consisting of bone, cartilage and fibroblast. The hSEP1 gene is located in chromosome 3 at 3q25-26.1 between markers D3S1309 and D3S1569. An adjacent locus defined by the polymorphic markers D3S1212 and D3S1245 has previously been reported to undergo loss of heterozygosity (LOH) at a >70% frequency in OGS and claimed to harbor an important tumor suppressor gene in osteosarcoma. The homozygous mutation in the hSEP1 mRNA in TE85 cell line suggest that this gene itself is subject to LOH. Taken together, these results suggest that hSEP1 acts as a tumor suppressor gene in OGS.

Zhang, X. and W. S. McIntire (1996). "Cloning and sequencing of a copper-containing, topa quinone-containing monoamine oxidase from human placenta." Gene **179**(2): 279.

http://www.sciencedirect.com/science/article/B6T39-3W2698D-F/2/68ca013408f16ff937df2f60f668a962

A 4040-bp cDNA was cloned from a human placenta library by screening with a polymerase chain reaction-amplified fragment. The fragment was generated from the library using primers corresponding to conserved sequences encompassing the topa quinone (TPQ) cofactor sites of the copper-containing proteins, bovine serum amine oxidase (BSAO) and human kidney diamine oxidase (DAO). The cloned cDNA contains a coding sequence from positions 161 to 2449. Between bases 2901 and 2974, in a very long 1591-bp 3'-untranslated region, there is a G/A-rich region in the minus strand, which contains a (AGG)5 tandem repeat. The human placenta cDNA sequence and its translated amino acid sequence are 84% and 81% identical to the corresponding BSAO sequences, while the identities for the placenta sequences and those for

human kidney DAO are 60% and 41%, respectively. The TPQ consensus nucleotide and protein sequences are identical for the placenta enzyme and BSAO, but the corresponding sequences for human kidney DAO are nonidentical. Three His residues that have been identified as Cu(II) ligands in other amine oxidases are conserved in the human placenta amine oxidase protein sequence. It was concluded that the placenta cDNA open-reading frame codes for a copper-containing, TPQ-containing monoamine oxidase. A putative 19-amino acid signal peptide was identified for human placenta amine oxidase. The resulting mature protein would be composed of 744 amino acids, and would have a Mr of 82 525. Comparison of the human placenta amine oxidase with DNA sequences found in GenBank suggests that the gene for this enzyme is located in the q21 region of human chromosome 17, near the BRCA1 gene.

Zolezzi, F. and S. Linn (2000). "Studies of the murine DDB1 and DDB2 genes." Gene **245**(1): 151.

http://www.sciencedirect.com/science/article/B6T39-3YS34MW-K/2/9d5d43554a86d69f17c475e7e64c99b5