© 1998 Oxford University Press Human Molecular Genetics, 1998, Vol. 7, No. 3 407–414

A novel pseudoautosomal gene encoding a putative GTP-binding protein resides in the vicinity of the Xp/Yp telomere

Fernando Gianfrancesco[1], Teresa Esposito[1], Luisa Montanini[1], Alfredo Ciccodicola[2], Steven Mumm[3], Richard Mazzarella[4], Ercole Rao[5], Sabrina Giglio[6], Gudrun Rappold[5] and Antonino Forabosco[1,*]

[1]Department of Morphology and Legal Medicine Sciences, Histology, Embryology and Genetics Unit, University of Modena, 41100 Modena, Italy, [2]International Institute of Genetics and Biophysics (IIGB), CNR, 80125 Naples, Italy, [3]Division of Bone and Mineral Diseases and [4]Institute for Biomedical Computing and Center for Genetics in Medicine, Washington University School of Medicine, St Louis, MO 63110, USA, [5]Institute of Human Genetics, Heidelberg University, Im Neuenheimer Feld 328, 69120 Heidelberg, Germany and [6]Biologia Generale e Genetica Medica, Università di Pavia, Pavia, Italy

Received August 22, 1997; Revised and Accepted December 5, 1997 DDBJ/EMBL/GenBank accession no. Y14391

We report the cloning of a novel Xp/Yp pseudoautosomal gene called *PGPL*, and demonstrate that *PGPL*, like other pseudoautosomal genes, escapes X inactivation and has a functional homologue on the Y chromosome. This gene is expressed in all the tissues examined and is highly conserved across several species. The *PGPL* gene encodes a protein of 442 amino acids and shows the consensus sequences of a series of motifs of the GTP-binding protein domain. Using fluorescence *in situ* hybridization analysis on normal males and on patients with rearrangements in the pseudoautosomal region, the gene was localized within 500 kb of the telomere. Further refinement using a cosmid contig of the region places this novel gene within 80–110 kb of the telomere, making this the most telomeric gene on the short arms of the sex chromosomes.

INTRODUCTION

Mammalian sex chromosomes share two terminal regions of homologous DNA sequences, which pair and recombine during male meiosis. Genes in these regions can be exchanged between the X and Y chromosomes and are inherited as if autosomal (1–4). They are present in two doses in both males and females, and escape inactivation (with one exception noted below) of the X chromosome in females.

X chromosome inactivation results in the transcriptional inactivation of most genes on one of the pair of X chromosomes in females, thereby achieving dosage equivalence for X-linked genes between males and females (5). A growing number of human genes have been described that escape X inactivation, being expressed from both the active and inactive X chromosomes (6). Many of the genes that escape X inactivation are located in the pseudoautosomal regions (PARs) at the termini of the X and Y chromosomes (7–16), although an increasing number are being identified outside the PARs (17–30). Some of the interspersed X chromosome genes that escape X inactivation, like pseudoautosomal genes, have functional Y homologues, although these homologues are not identical to the X-linked genes at the DNA sequence level. Other genes also escape X inactivation, but they do not have a Y homologue (31). One of two genes located in the XqPAR (*IL9R*) escapes X inactivation and is expressed from the Y chromosome (16) while *SYBL1*, that has identical copies on X and Y, surprisingly does not escape X inactivation and, in addition, is inactive on the Y chromosome (15).

In our search for genes which escape X inactivation, we identified a novel pseudoautosomal gene that has a functional homologue on the Y chromosome. This gene is a putative GTP-binding protein which we have named *PGPL* (pseudoautosomal GTP-binding protein-like). Mapping information identifies this locus as the most telomeric gene in the Xp/Yp PAR.

RESULTS

Isolation and characterization of *PGPL*

Expressed sequence tag (EST) 917 was isolated from human skeletal muscle and mapped by silver-stained denaturing polyacrylamide gel electrophoresis (DPGE) on the X and Y chromosomes (32). By PCR-based screening (33) using the EST 917 primers, a cDNA of almost 1900 bp was recovered from a cDNA library derived from a human uninduced male teratocarcinoma cell line, NT2/D1 (34). The cDNA sequence was determined and is shown in Figure 1. It contains an interesting open

*To whom correspondence should be addressed. Tel: +39 59 424826; Fax: +39 59 424840; Email: forabosc@unimo.it

```
   1  GCGGGCCGCCGTACGCCCGGGGCTGCGGCTCTCCCGCGTGGGCCGCGGCCGCTCGGCTCCGCGGGCAGCCGCGCCGTCCTG    81
  82  CCCCGCGCGCGCGCTAGCCGCTGTCGGCCGCAGGAGCCCCGGGAATCTGGAGGGGCCGTGGGGCGGAGGGAGGGGCCTGCG   162
 163  GGCGGAAGGCGGACGAAAACAGAACGGAAGACGACAAGGAGGAGCCGGAAGATGCGGACGAGAACGCCGAGGAGGAGCTGC   243
   1                                                       M  R  T  R  T  P  R  R  S  C    10
 244  TGCGGGGAGAGCCTCTGCTGCCGGCGGGGACCCAAGCGCGTGTGTCTGGTTCACCCTGACGTCAAGTGGGGCCCAGGGAAA   324
  11   C  G  E  S  L  C  C  R  R  G  P  K  R  V  C  L  V  H  P  D  V  K  W  G  P  G  K    37
 325  TCGCAGATGACTCGAGCCGAGTGGCAGGTGGCGGAGGCCACAGCGCTGGTGCACACGCTGGACGGCTGGTCCGTGGTGCAC   405
  38   S  Q  M  T  R  A  E  W  Q  V  A  E  A  T  A  L  V  H  T  L  D  G  W  S  V  V  H    64
 406  ACAATGGTCGTGTCCACCAAAACGCCGGACAGGAAGCTCATCTTTGGCAAAGGGAACTTTGAGCACCTGACAGAAAAGATC   486
  65   T  M  V  V  S  T  K  T  P  D  R  K  L  I  F  G  K  G  N  F  E  H  L  T  E  K  I    91
 487  CGAGGGTCTCCAGACATCACGTGCGTCTTCCTGAACGTGGAGAGGATGGCTGCCCCGACCAAGAAAGAACTGGAAGCCGCC   567
  92   R  G  S  P  D  I  T  C  V  F  L  N  V  E  R  M  A  A  P  T  K  K  E  L  E  A  A   118
 568  TGGGGCGTGGAGGTGTTTGAGCGCTTCACGGTCGTCCTGCACATCTTCCGCTGTAACGCCCGCACGAAGGAGGCCCGGCTT   648
 119   W  G  V  E  V  F  E  R  F  T  V  V  L  H  I  F  R  C  N  A  R  T  K  E  A  R  L   145
 649  CAGGTGGCCCTGGCGGAGATGCCGCTGCACAGGTCGAACCTGAAAAGGGACGTCGCCCACCTGTACCGAGGAGTCGGCTCG   729
 146   Q  V  A  L  A  E  M  P  L  H  R  S  N  L  K  R  D  V  A  H  L  Y  R  G  V  G  S   172
 730  CGCTACATCATGGGGTCAGGAGAATCCTTCATGCAGCTGCAGCAGCGTCTCCTGAGAGAGAAGGAGGCCAAGATCAGGAAG   810
 173   R  Y  I  M  G  S  G  E  S  F  M  Q  L  Q  Q  R  L  L  R  E  K  E  A  K  I  R  K   199
 811  GCCTTGGACAGGCTTCGCAAGAAGAGGCACCTGCTCCGCCGGCAGCGGACGAGGCGGGAGTTCCCCGTGATCTCCGTGGTG   891
 200   A  L  D  R  L  R  K  K  R  H  L  L  R  R  Q  R  T  R  R  E  F  P  V  I  S  V  V   226
 892  GGGTACACCAACTGCGGAAAGACCACGCTGATCAAGGCACTGACGGGCGATGCCGCCATCCAGCCACGGGACCAGCTGTTT   972
 227   G  Y  T  N  C  G  K  T  T  L  I  K  A  L  T  G  D  A  A  I  Q  P  R  D  Q  L  F   253
 973  GCCACGCTGGACGTCACGGCCCACGCGGGCACGCTGCCCTCACGCATGACCGTCCTGTACGTGGACACCATCGGCTTCCTC  1053
 254   A  T  L  D  V  T  A  H  A  G  T  L  P  S  R  M  T  V  L  Y  V  D  T  I  G  F  L   280
1054  TCCCAGCTGCCGCACGGCCTCATCGAGTCCTTCTCCGCCACCCTGGAAGACGTGGCCCACTCGGATCTCATCTTGCACGTG  1134
 281   S  Q  L  P  H  G  L  I  E  S  F  S  A  T  L  E  D  V  A  H  S  D  L  I  L  H  V   307
1135  AGGGACGTCAGCCACCCCGAGGCGGAGCTCCAGAAATGCAGCGTTCTGTCCACGCTGCGTGGCCTGCAGCTGCCCGCCCCG  1215
 308   R  D  V  S  H  P  E  A  E  L  Q  K  C  S  V  L  S  T  L  R  G  L  Q  L  P  A  P   334
1216  CTCCTGGACTCCATGGTGGAGGTTCACAACAAGGTGGACCTCGTGCCCGGGTACAGCCCCACGGAACCGAACGTCGTGCCC  1296
 335   L  L  D  S  M  V  E  V  H  N  K  V  D  L  V  P  G  Y  S  P  T  E  P  N  V  V  P   361
1297  GTGTCTGCCCTGCGGGGCCACGGGCTCCAGGAGCTGAAAGCTGAGCTCGATGCGGCGGTTTTGAAGGCGACGGGGAGACAG  1377
 362   V  S  A  L  R  G  H  G  L  Q  E  L  K  A  E  L  D  A  A  V  L  K  A  T  G  R  Q   388
1378  ATCCTCACTCTCCGTGTGAGGCTCGCAGGGGCGCAGCTCAGCTGGCTGTATAAGGAGGCCACAGTTCAGGAGGTGGACGTG  1458
 389   I  L  T  L  R  V  R  L  A  G  A  Q  L  S  W  L  Y  K  E  A  T  V  Q  E  V  D  V   415
1459  ATCCCTGAGGACGGGGCGGCCGACGTGAGGGTCATCATCAGCAACTCAGCCTACGGCAAATTCCGGAAGCTCTTTCCAGGA  1539
 416   I  P  E  D  G  A  A  D  V  R  V  I  I  S  N  S  A  Y  G  K  F  R  K  L  F  P  G   442
1540  TGAACGGACGCCCACAGAGGCCTGCGGGGTGGGGGCATTGCTGCCTGGGGAGCTGAGGCGTTACCGCTGTGTTGGGGGCAG  1620
 443   *                                                                                 443
1621  CTTGGTGTCAGGTGCAGCAGGGTCCTCCTTGTCTGGTTCTGCACCCGTCTCGCTCCCAGCCATTTGCTGGGATGACCGTGC  1701
1702  AGGCCGGTGACACGGCCGCACCTGCCCCAAAGCGGGCCGCCCGAGCGTCCACTCCAAGCCTGAGCATCCACACAATTCCAG  1782
1783  TGGGCCCTCGGTGCCTGCTGTGAACTGCTTTCCCTCGGAATGTTTCCGTAACAGGACATTAAACCTTTGATTTTAAAAAAA  1863
1864  AAAAA                                                                            1868
```

Figure 1. The nucleotide sequence of the *PGPL* gene together with the conceptual translation of the open reading frame. GTP-binding protein-like domains are denoted by grey boxes. The primers used in the RT-PCR assay are indicated by solid lines.

reading frame (ORF) that starts at the ATG codon at nucleotide 214 and ends with the termination codon (TGA) at nucleotide 1542, encoding a putative protein of 442 amino acids (49 390 Da). The 3′ untranslated region (UTR) contains a putative polyadenylation signal (ATTAAA) at 11 nucleotides from the poly(A) tail. This putative protein contains four domains characteristic of a GTP-binding site at amino acids 227–234 (GYTNCGKT), amino acids 275–278 (DTIG), amino acids 344–347 (NKVD) and amino acids 363–366 (SALR).

In the review of Kjeldgaard *et al.* (35), the common structural core of the GTP-binding protein was described. The first domain, GXXXXGK(S/T), appears, with variations, in many nucleoside triphosphate-utilizing enzymes and represents the phosphate-binding loop. The second element, DXXG, is conserved in all GTPases and is implicated in conformational change between the GDP and GTP forms. The third element, NXXD, determines the specificity for the nucleotide guanine. The fourth element, (C/S)AXX, may be essential for the structure of the protein. Comparison of the deduced amino acid sequence of PGPL with those in the EMBL database revealed homology (34.5% identity and 43.5% similarity) with F46B6.4 (Z70780), a putative GTP-binding protein of *Caenorhabditis elegans* (36) (Fig. 6A) and lower homology with hflX (U14003), a GTP-binding protein of *Escherichia coli* required for high frequency lysogenization by bacteriophage λ (37). Analysis of the homologous eukaryotic proteins for cellular sorting signals shows that both the human and *C.elegans* sequences contain amphipathic α-helices at their N-termini characteristic of mitochondrial matrix proteins. The human protein is highly charged and contains a possible nuclear localization signal at amino acids 205–209 (RKKRH).

Northern analysis detected three RNA species in all human adult tissues tested. The most abundant transcript of 1900 bp was observed in all tissues analysed, with the highest levels of expression in muscle and heart (Fig. 2). Moreover, two less abundant transcripts (3.5 and 4.8 kb) were observed in all tissues.

To assess the transcriptional activity of *PGPL* on the X chromosome in various cell lines with either active or inactive X chromosomes, RNA was amplified by RT-PCR using a specific

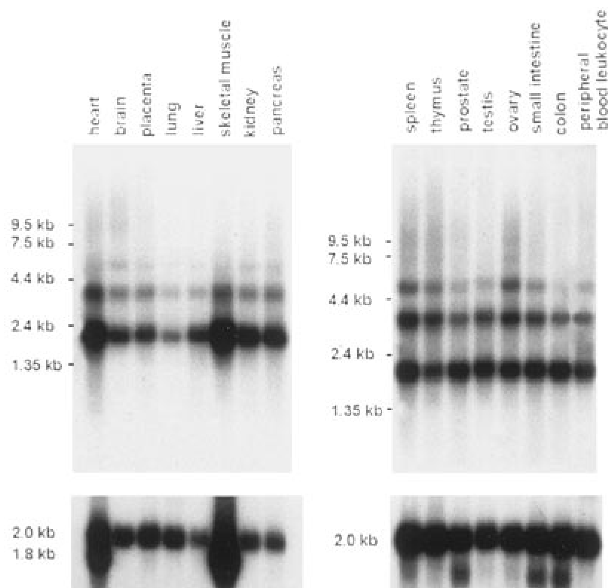

Figure 2. The expression of *PGPL* in human tissues by northern blot analysis. The bars on the left indicate the position of migration of RNA markers. Reprobing of the same filter with a β-actin probe is shown in the bottom panel.

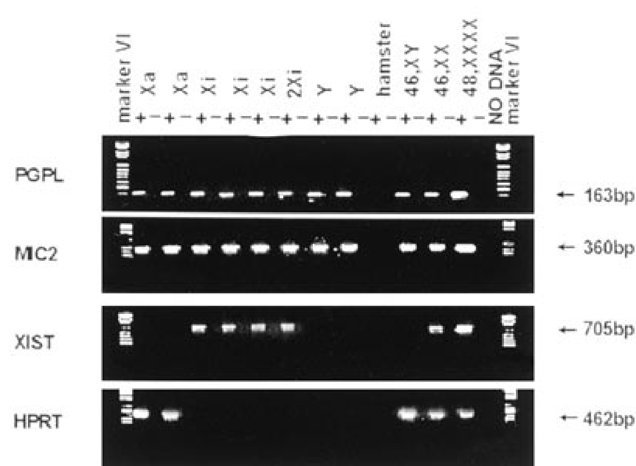

Figure 3. RT-PCR analysis of the expression of *PGPL* using the primer pair 917F/917R. Ethidium bromide-stained RT-PCR products were obtained after amplification with primers for *PGPL* as well as *MIC2*, *XIST* and *HPRT* control primers tested against RNA from somatic cell hybrids containing active X (Hy136c and GM06318B), inactive X (THX88, Hy70C4T3 and Y162.5E1T2), inactive XX (X8/6T2), Y (Hy853 and GM06317), a hamster cell line, male (46,XY) and female (46,XX) cell lines and a human lymphoblastoid cell line GM1416 (48,XXXX). Sizes are indicated on the right.

primer pair (see Materials and Methods and Fig. 1). RT-PCR amplification of *PGPL* was detected in both active and inactive X- as well as in the Y-containing hybrid cell lines, 4X cell line, male and female cell lines (Fig. 3). Because the primer pair (917F/917R) amplifies identical products from genomic DNA, RT-PCR assays were performed with and without reverse transcriptase to confirm that amplification was from RNA and not from contaminating DNA in the RNA preparations. Control genes showed the expected amplification products from the active X for the *HPRT* (hypoxanthine phosphoribosyl transferase) gene, a gene known to undergo X inactivation; from the inactive X for *XIST* (38) and from the active X, inactive X and Y for *MIC2* (39) (Fig. 3). Escape from X inactivation for *PGPL* was confirmed by RT-PCR assays with two other primer pairs (see Materials and Methods and Fig. 1) based on the cDNA sequence (data not shown).

Chromosomal localization of *PGPL*

The regional assignment of the *PGPL* gene on the X chromosome was inferred from the patterns of segregation of the amplification products in a somatic cell hybrid panel. This panel contained 28 human–hamster or human–mouse cell lines that subdivide the X chromosome into 31 intervals (40). Using this panel, we have mapped the gene in the cytogenetic band Xp22.33, telomeric to *DXS1233* (data not shown).

By Southern blot hybridization assay, the cDNA detected a 20 kb *Eco*RI fragment in male and female DNAs as well as in DNAs from hybrid cells containing either a human X chromosome or a human Y chromosome. The same fragment was detected in the hybrid EGMH49 that retains part of the human X chromosome between Xptel/Xq26.3 but not in the hybrid EGMH42 that contains the human X chromosome between *DXS1233*/Xqtel (data not shown). This result is in accordance with the PCR result.

Further mapping of the *PGPL* gene was achieved by fluorescence *in situ* hybridization (FISH) analysis. Using two cosmids (U55B10 and U63F1) positive for the *PGPL* sequence-tagged site (STS) (EST 917) as probes, FISH was performed on DNA from normal males, a patient with a deletion of the distal 1.5 Mb of the PAR region (YU; unpublished data) and a patient with an inversion breakpoint at 500 kb from the telomere [AT, (13)]. The results of this analysis show that *PGPL* is located in PAR1 at a maximum distance of 500 kb from the telomere. Finer analysis was performed by localizing the *PGPL* STS on a yeast artificial chromosome (YAC) and cosmid contig of the region (13,41). The STS was positive for cosmids LLNLN0434, U55B10 and U63F1, and places the gene ~80–110 kb from the telomere (13) (Fig. 4). This makes the *PGPL* gene the most telomeric gene on the short arms of the X and Y chromosomes.

Evolutionary conservation of *PGPL*

Southern 'zoo' blot analysis, using the *PGPL* cDNA as a probe, was employed to study whether there are *PGPL*-homologous sequences present in other species. The cDNA clone hybridized to genomic DNA from a number of species including human, monkey, rat, mouse, dog, cow and rabbit (Fig. 5). Such conservation of the *PGPL* gene argues strongly in favour of an important biological role.

In an EST database search to identify ESTs of mouse related to the *PGPL* gene, we found six ESTs of mouse (see Materials and Methods) which showed significant homology to the *PGPL* gene. These ESTs were assembled by 'The EST Assembly Machine' program (http://gcg.tigem.it/cgi-bin/uniestass.pl) which gave a putative murine cDNA homologous to *PGPL*. An assembled fragment of 773 bp was obtained, and comparison of the deduced

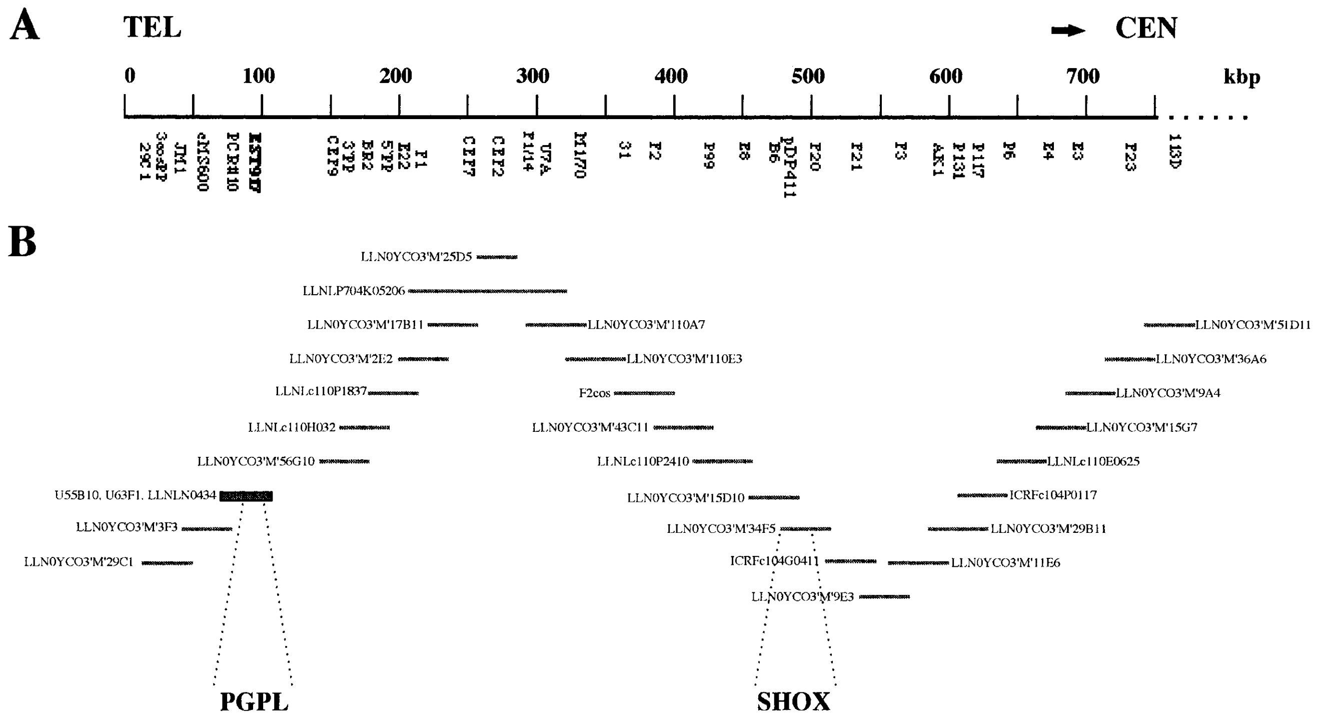

Figure 4. Relationship between the chromosomal marker map, cosmid contig and the genes in the 700 kb region bounded by the Xp–Yp telomere. (**A**) Markers are indicated vertically below the solid bar representing genomic DNA, with respect to their distance to the telomere. The telomere is indicated by zero on the left side. (**B**) Cosmids are positioned below (13). *PGPL* resides on cosmids U55B10, U63F1 and LLNLN0434 shown by a thick line. A further gene, *SHOX*, has also been mapped on this chromosomal interval.

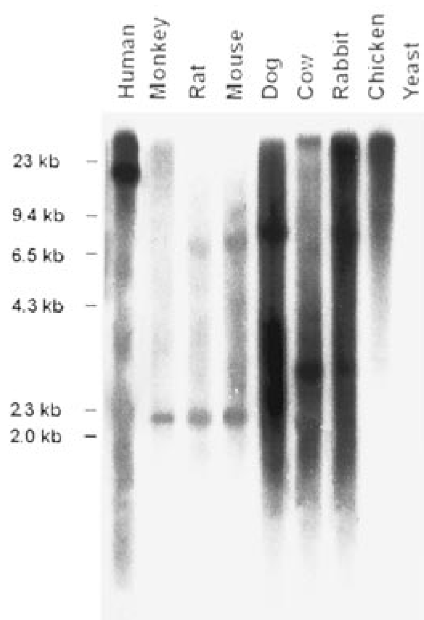

Figure 5. Evolutionary conservation of the *PGPL* gene. The bars on the left indicate the position of migration of the DNA markers.

amino acid sequence with those in the EMBL database revealed the same homologies as that of the *PGPL* human gene (F46B6.4 and hflX) and supports the hypothesis that the fragment obtained is part of the putative *pgpl* gene. In fact, comparison of the consensus nucleotide sequence of *pgpl* with *PGPL* revealed 71% homology (data not shown), while comparison of the deduced amino acid sequence of pgpl with PGPL revealed high homology (67% identity and 87% similarity) (Fig. 6B). Therefore, a more precise comparison with the full-length *pgpl* cDNA remains to be determined, together with its chromosomal localization and its expression pattern.

DISCUSSION

We have isolated a novel pseudoautosomal gene called *PGPL* which is the most telomeric gene on the short arms of the sex chromosomes. This gene shows the consensus sequences which are known to constitute a fingerprint of GTP-binding protein domains. GTPases are conserved molecular switches, built according to a common structural design. Turned on by binding GTP and off by hydrolysing GTP to GDP, the switch mechanism is remarkably versatile, enabling different GTPases to sort and amplify transmembrane signals, direct the synthesis and translocation of proteins, guide vesicular traffic through the cytoplasm, and control proliferation and differentiation of animal cells.

The comparison of the human protein against its *C.elegans* and *E.coli* homologues suggests that they have the same or very similar functions since the vital regions of the GTP-binding domain is conserved and the regions of homology are almost at equal distances from one another. All of these proteins appear to

A

```
PGPL         MRTRTPRRSCCGESLCCRRGPKRVCLVHPDVKWGPGK-SQMTRAEWQVAEATA   52
F46B6.4   SRQTVLSFRNLSIGSSEVAAPSAFANDRWSVLVVHPKVRWGSGSASVLKQADRQLEEAVA   66

PGPL      LVHTLDGWSVVHTMVVSTK-TPDRKLIFGKGNFEHLTEKIRGSPDITCVFLNVERMAAPT  111
F46B6.4   LVDNLPNMNAVDSLIMPVDYNTKRKAVWASGNLEKLIAR-REAARATALMVNVDALSPSQ  125

PGPL      KKELEAAWGVEVFERFTVVLHIFRCNARTKEARLQVALAEMP-----LHRSNLKR-----  161
F46B6.4   QQELYRIFEVPIFDRYNIVLATFKQFAKTEEARIQIAIAEIPYIKHRIHALSSKRLHSRP  185

PGPL      DVAHLYRGVGSRYIMGSGESFMQLQQRLLREKEAKIRKAL-DRLRKKRHLLRRQRTRREF  220
F46B6.4   DILH----IDSHYSDIDGD-----LNEILRKREQDLRKELKDVTRKNVGQLGVRNSSDA-  235

PGPL      PVISVVGYTNCGKTTLIKALTGDAAIQPRDQLFATLDVTAHAGTLPSRMTVLYVDTIGFL  280
F46B6.4   -VVAVVGYTNSGKTSLVKKLTGAASLTPKDQLFATLDTTRHLAKLPSGRSAVFTDTIGFL  294

PGPL      SQLPHGLIESFSATLEDVAHSDLILHVRDVSHPEAELQKCSVLSTLRGLQLPAPLLDS-M  339
F46B6.4   SDLPMHLIAAFEATLAHVKSADVIIHLRDISNPDWKAQEEDVLATLKSIGVTDYVLNERI  354

PGPL      VEVHNKVDLVPGYSPTEPN-VVPVSALRGHGLQELKAELDAAVLKATGRQILTLRVRLAG  398
F46B6.4   ISVDNKIDKESAFPTSESNNSVRISCKTGDGMHELIDVINDKVTMVTKCKTIRLRLDARS  414

PGPL      AQLSWLYK-EATVQEVDVIPEDGAADV-RVIISNSAYGKFRKLFPG*               443
F46B6.4   PVIEWLYHNELVVIEPTI---DGNYLIFDVVMNESEIGRFRKKFAHLKKKNSQSVSL     468
```

B

```
Human     QLQQRLLREKEAKIRKALDRLRKKRHLLRRQRTRREFPVISVVGYTNCGKTTLIKALTGD  243
mouse     ELRARALRDRELRLRRVLERLRDKRRLMRKERVRREFPVVSVVGYTNCGKTTLIQALTGE

Human     AAIQPRDQLFATLDVTAHAGTLPSRMTVLYVDTIGFLSQLPHGLIESFSATLEDVAHSDL  303
mouse     ASLQPRDQPFATLDVTVHAGLLPSRLRILYVDTIGFLSQLPHSLIHAFSATLEDVAYSDV

Human     ILHVRDVSHPEAELQKCSVLSTLRGLQLPAPLLDSMVEVHNKVDLVPGYSPTEPNVVPVS  363
mouse     LVHVTDVSHPDAELQKATVLSTLRGLHLPPALLESALEVHSKVDLVPGYTPPCSGALAVS

Human     ALRGHGLQELKAELDAAVLKATGRQILTLRVRLAGAQLSWLYKEATVQEVDVIPEDGAAD  423
mouse     AISGRGLDELKAALEASVLRSTGRQVLTLCVRLGGPQLGWLYKEAVVQQVQELPEGDAAH

Human     VRVIISNSAYGKFRKLFPG*                                          443
Mouse     VTVVITQASYGRFRKLFPI*
```

Figure 6. Comparison of the amino acid sequences. Amino acids are given in the single letter code. The vertical lines represent identical amino acids and the colons represent conservative amino acid changes. (**A**) Comparison between F46B6.4 of *C.elegans* and the predicted PGPL peptide. (**B**) Comparison between the predicted PGPL peptide and the putative homologous murine pgpl.

be in their own GTP family and do not appear to be part of the Ras or any other known GTP-binding family. Analysis of the PGPL sequence shows that it contains an N-terminal sequence which can function as a mitochondrial matrix targeting sequence. Functional studies in *E.coli* suggest that the prokaryotic homologue may form a subunit of a protein complex that may act as a switch controlling the activity of a protease. This GTP-coupled proteolysis system appears to affect the stability of specifically targeted proteins (37).

Seven human pseudoautosomal genes in the Xp/Yp PAR have been described to date. *MIC2* was the first pseudoautosomal gene described in man, and is a ubiquitously expressed housekeeping

gene that encodes a cell surface antigen (7). *XE7* appears to be ubiquitously expressed, and alternative splicing results in two very hydrophilic protein isoforms. Presently there is no clue to its biological function (8). The product of *ASMT* has been suggested as a candidate for psychiatric disorders because of its tissue-specific expression in brain and retina and previous association and linkage studies of pseudoautosomal markers in schizophrenia patients (9). *ANT3* represents a highly conserved gene from the ADP/ATP translocase family and thus plays a fundamental role in the energy metabolism of the eukaryotic cell (10). *IL3RA* and *CSF2RA* map close and their products share the same β subunit whereas the α subunits are distinct (11,12). Recently, two groups independently have isolated two identical genes, *SHOX* and *PHOG*, that were suggested to cause growth failure in idiopathic short stature and Turner syndrome (13,14).

The high degree of conservation across several species of the *PGPL* gene implies an important function for the locus in humans. Moreover, its expression pattern as well as the important biological role carried out by GTP-binding proteins suggest that the gene may not have a specific phenotypic effect in a particular tissue or cell type but may be crucially involved in the vitality of an organism.

MATERIALS AND METHODS

Southern blots

Southern blot hybridization was carried out in 5× SSPE/5× Denhardt's/0.5% SDS at 65°C; the filter was washed once in 2× SSC (300 mM NaCl, 30 mM sodium citrate, pH 7.5), 1% SDS at 65°C and twice in 0.2× SSC, 0.2% SDS at 65°C. The membrane was exposed to XAR-5 film at –80°C for 24 h. Zoo-blot (Clontech) was pre-hybridized and hybridized in 5× SSPE/5× Denhardt's/0.5% SDS at 50°C; the membrane was washed once in 2× SSC, 1% SDS at 50°C and once in 0.5× SSC, 0.2% SDS at 50°C. The membrane was exposed to XAR-5 film at –80°C for 24 h.

FISH mapping

Biotinylated DNA of cosmids U55B10 and U63F1 was hybridized to metaphase chromosomes from stimulated lymphocytes of patients AT (13) and YU (unpublished data) under conditions as described previously (42). The hybridized probe was detected via avidin-conjugated fluorescein isothiocyanate (FITC).

cDNA and sequence analysis

The cDNA library used in this work was from a human uninduced male teratocarcinoma cell line, NT2/D1 (34). The cDNA was subcloned into vector pGEM-4Z (Promega-Biotech), and analysed by dye-terminator cycle sequencing on an Applied Biosystem 373A automated sequencer. The cDNA sequence was deposited in the EMBL database with the accession no. Y14391. Sequence databases were searched using the BLAST sequence alignment program (43).

ESTs

The ESTs for the putative *pgpl* gene are: mu22h01.r1 (AA197541), mv45g12.r1 (AA198285), mu40e06.r1 (AA209886), ve36b07.r1 (AA423167), ve91a02.r1 (AA432941) and ma24a09.r1 (W50953). These ESTs were assembled by the 'EST Assemble Machine' program (http://gcg.tigem.it/cgi-bin/uniestass.pl).

Northern blot analysis

Two human multiple-tissue northern blots (Clontech) were hybridized with the full-length human cDNA, and with a β-actin probe to verify the relative normalization of mRNA amounts. The northern blots were pre-hybridized, hybridized and washed by an ExpressHyb™ Hybridisation Solution protocol (Clontech) (44). The filters were exposed for 16 h for the PGPL probe and 2 h for the β-actin probe.

Cell lines

The panel of somatic cell hybrid lines used in RT-PCR assays comprised two hybrids retaining the active human X chromosome, Hy136c (45) and GM06318B (NIGMS, Camden, NJ); three hybrids retaining an inactive human X chromosome, THX88 (46), Hy70C4T3 (15) and Y.162.5E1T2 (from M. Rocchi) [the hybrid Y.162.5E1T2 retains an inactive human X chromosome, a portion of chromosome 5 (5pter–5cen) and a fragment of chromosome 12 (12q24.3–qter)]; a hybrid retaining two inactive human X chromosomes, X8/6T2 (from M. Rocchi); and two hybrids retaining the human Y chromosome, GM06317 (NIGMS, Camden, NJ) and Hy853 (47,48). The human lymphoblastoid cell line GM1416 (48,XXXX) (NIGMS, Camden, NJ) and normal male and female cell lines were used as positive controls. The hamster cell line YH21 (45) was used as negative control.

The somatic cell hybrid lines used in the Southern blot assay comprised two hybrids: EGMH42 and EGMH49 retain part of the human X chromosome between *DXS1233*/Xqtel and Xptel/Xq26.3 respectively (40).

Analysis of expression from active and inactive X chromosomes

RNAs from cell lines were obtained by extraction in guanidinium thiocyanate followed by centrifugation in caesium chloride solution (49). Ten µg of total RNA, 1× reaction buffer (40 mM Tris–HCl pH 7.9; 10 mM NaCl; 6 mM $MgCl_2$), 16.5 U of RNasin (Promega) and 7 U of RQ1DNase (Promega) were incubated in 50 µl of total volume at 37°C for 30 min. After incubation, the RNA was purified on Strataclean resin (Stratagene). About 1 µg of total RNA was reverse-transcribed in a 50 µl reaction mixture containing 1× RT buffer (20 mM Tris–HCl pH 8.4, 50 mM KCl, 2.5 mM $MgCl_2$), 10 mM dithiothreitol (DTT), 0.5 mM dNTP, 0.2 µg of random hexamers (Boehringer Mannheim) and 200 U of SuperScript reverse transcriptase (Gibco BRL). After a 60 min incubation at 37°C, 1 µg of DNase-free RNase was added and incubated for 10 min at 37°C. The cDNA formed was extracted twice with phenol:chloroform:iso-amylalcohol (25:24:1). The cDNA was then precipitated overnight with 1/10 vol. of 3 M sodium acetate, pH 5.2, and 2.5 vol. of ethanol. Approximately 1 µg of cDNA was recovered. RT-PCR experiments were carried out using 100 ng of RNA and 5 ng of cDNA as the template in a 10 µl PCR reaction containing 1× TNK 100 buffer (50), 0.2 mM dNTP, 0.35 U of AmpliTaq polymerase (Boehringer Mannheim) and 0.5 µM each of the primer sequences derived from the EST. Using a DNA Thermal Cycler MJR (M.J. Research Inc.), we carried out 35 cycles of amplification using a step programme: 1 min at 94°C; 2 min at

60°C; 2 min at 72°C. The *PGPL* primers used in RT-PCR assay, are: 917F, 5′-GTCTCGCTCCCAGCCATTTGCTGGGATGA-C-3′; 917R, 5′-GGAAACATTCCGAGGGAAAGCAGTTCA-CAG-3′; PGPLa, 5′-CGGGGAGAGCCTCTGCTGCC-3′; PG-PLb, 5′-CTTTCTTGGTCGGGGCAGCC-3′; PGPLe, 5′-GGT-ACACCAACTGCGGAAAG-3′; PGPLf, 5′-TGTCTCCCCGT-CGCCTTCAA-3′.

ACKNOWLEDGEMENTS

We thank Professors D. Schlessinger and O. Zuffardi for helpful suggestions; Professor M. Rocchi and TIGEM core for providing hybrid cell lines; Dr C. Sala for CEPH YAC library screening; Dr C. Migliaccio and A. Terracciano for sequence analysis; and Jonathan Jones and Marie-Christine Meffert for technical assistance. S.M. was supported by a fellowship from The Shriners Hospital for Children. This work was supported by Telethon Italy, grant no. E546 to A.C., a Deutsche Forschungsgemeinschaft grant to G.R. and MURST 40% to A.F.

REFERENCES

1. Cooke, H.J., Brown, W.R.A. and Rappold, G.A. (1985) Hypervariable telomeric sequences from the human sex chromosomes are pseudoautosomal. *Nature*, **317**, 687–692.
2. Simmler, M.-C., Rouyer, F., Vergnaud, G., Nystro-Lahti, M., Ngo, K.Y., de la Chapelle, A. and Weissenbach, J. (1985) Pseudoautosomal DNA sequences in the pairing region of the human sex chromosomes. *Nature*, **317**, 692–697.
3. Rappold, G.A. (1993) The pseudoautosomal regions of the human sex chromosomes. *Hum. Genet*., **92**, 315–324.
4. Freije, D., Helms, C., Watson, M.S. and Donis-Keller, H. (1992) Identification of a second pseudoautosomal region near the Xq and Yq telomeres. *Science*, **258**, 1784–1787.
5. Lyon, M.F. (1961) Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature*, **190**, 372–373.
6. Brown, C.J. and Willard, H.F. (1993) Molecular and genetic studies of human X chromosome inactivation. *Adv. Dev. Biol*., **2**, 37–72.
7. Goodfellow, P., Pym, B., Mohandas, T. and Shapiro, L.J. (1984) The cell surface antigen locus, MIC2X, escapes X-inactivation. *Am. J. Hum. Genet*., **36**, 777–782.
8. Ellison, J.W., Ramos, C., Yen, P.H. and Shapiro, L.J. (1992) Structure and expression of the human pseudoautosomal gene XE7. *Hum. Mol. Genet*., **1**, 691–696.
9. Yi, H., Donohue, S.J., Klein, D.C. and McBride, O.W. (1993) Localization of the hydroxyindole-*O*-methyltransferase gene to the pseudoautosomal region: implications for mapping of psychiatric disorders. *Hum. Mol. Genet*, **2**, 127–131.
10. Schiebel, K., Weiss, B., Wohrle, D. and Rappold, G. (1993) A human pseudoautosomal gene, ADP/ATP translocase, escapes X-inactivation whereas a homologue in Xq is subject to X-inactivation. *Nature Genet*., **3**, 82–87.
11. Gough, N.M., Gearing, D.P., Nicola, N.A., Baker, E., Pritchard, M., Callen, D.F. and Sutherland, G.R. (1990) Localization of the human GM-CSF receptor gene to X–Y pseudoautosomal region. *Nature*, **345**, 734–736.
12. Milatovich, A., Kitamura, T., Miyajima, A. and Francke, U. (1993) Gene for the α-subunit of the human interleukin-3 receptor (IL3RA) localized to the X–Y pseudoautosomal region. *Am. J. Hum. Genet*., **53**, 1146–1153.
13. Rao, E., Weiss, B., Fukami, M., Rump, A., Niesler, B., Mertz, A., Muroya, K., Binder, G., Kirsch, S., Winkelmann, M, Nordsiek, G, Heinrich, U., Breuning, M.H., Ranke, M.B., Rosenthal, A., Ogata,T. and Rappold, G.A. (1997) Pseudoautosomal deletions encompassing a novel homeobox gene cause growth failure in idiopathic short stature and Turner syndrome. *Nature Genet*., **16**, 54–63.
14. Ellison, J.W., Wardak, Z., Young, M.F., Robey, P.G., Laig-Webster, M. and Chiong, W. (1997) PHOG, a candidate gene for involvement in the short stature of Turner syndrome. *Hum. Mol. Genet*., **6**, 1341–1347.
15. D'Esposito, M., Ciccodicola, A., Gianfrancesco, F., Esposito, T., Flagiello, L., Mazzarella, R., Schlessinger, D. and D'Urso, M. (1996) A synaptobrevin-like gene in the Xq28 pseudoautosomal region undergoes X inactivation. *Nature Genet*., **13**, 227–229.
16. Vermeesch, J.R., Petit, P., Kermouini, A., Renauld, J.-C., Van Den Berghe, H. and Marynen, P. (1997) The IL-9 receptor gene, located in the Xq/Yq pseudoautosomal region, has an autosomal origin, escapes X inactivation and is expressed from the Y. *Hum. Mol. Genet*., **6**, 1–8.
17. Franco, B., Meroni, G., Parenti, G., Levilliers, J., Bernard, L., Gebbia, M., Cox, L., Maroteaux, P., Sheffield, L., Rappold, G.A., Andria, G., Petit, C. and Ballabio, A. (1995) A cluster of sulfatase genes on Xp22.3: mutations in chondrodysplasia punctata (CDPX) and implications for warfarin embryopathy. *Cell*, **81**, 15–25.
18. Yen, P.H., Ellison, J., Salido, E.C., Mohandas, T. and Shapiro, L. (1992) Isolation of a new gene from the distal short arm of the human X chromosome that escapes X-inactivation. *Hum. Mol. Genet*., **1**, 47–52.
19. Yen, P.H., Allen, C., Marsh, B., Mohandas, T., Wang, N., Taggart, R.T. and Shapiro, L.J. (1987) Cloning and expression of steroid sulfatase cDNA and the frequent occurrence of deletions in STS deficiency: implication for X–Y interchange. *Cell*, **49**, 443–454.
20. Franco, B., Guioli, S., Pragliola, A., Incerti, B., Bardoni, B., Tonlorenzi, R., Carrozzo, R., Maestrini, E., Pieretti, M., Taillon-Miller, P., Brown, C.J., Willard, H.F., Lawrence, C., Persico, M.G., Camerino, G. and Ballabio, A. (1991) A gene deleted in Kallmann's syndrome shares homology with neural cell adhesion and axonal path-finding molecules. *Nature*, **353**, 529–536.
21. Ellison, J., Passage, M., Yu, L.C., Yen, P., Mohandas, T.K. and Shapiro, L. (1992) Directed isolation of human genes that escape X inactivation. *Somat. Cell Mol. Genet*., **18**, 259–268.
22. Schneider-Gadicke, A., Beer-Romero, P., Brown, L. G., Nussbaum, R. and Page, D.C. (1989) ZFX has a gene structure similar to ZFY, the putative sex determinant, and escapes X inactivation. *Cell*, **57**, 1247–1258.
23. Jones, M.H., Furlong, R.A., Burkin, H., Chalmers, I.J., Brown, G.M., Khwaja, O. and Affara, N.A. (1996) The *Drosophila* developmental gene fat facets has a human homologue in Xp11.4 which escapes X-inactivation and has related sequences on Yq11.2. *Hum. Mol. Genet*., **5**, 1695–1701.
24. Brown, C.J. and Willard, H.F. (1990) Localization of a gene that escapes inactivation to the X chromosome proximal short arm: implications for X inactivation. *Am. J. Hum. Genet*., **46**, 273–279.
25. Carrel, L., Clemson, C.M., Dunn, J.M., Miller, A.P., Hunt, P.A., Lawrence, J.B. and Willard, H.F. (1996) X inactivation analysis and DNA methylation studies of the ubiquitin activating enzyme E1 and PCTAIRE-1 genes in human and mouse. *Hum. Mol. Genet*., **5**, 391–401.
26. Miller, A.P., Gustashaw, K., Wolff, D.J., Rider, S.H., Monaco, A.P., Eble, B., Schlessinger, D., Gorski, J.L., van Ommen, G.-J., Weissenbach, J. and Willard, H.F. (1995) Three genes that escape X chromosome inactivation are clustered within a 6 Mb YAC contig and STS map in Xp11.21–11.22. *Hum. Mol. Genet*., **4**, 731–739.
27. Wu, J., Ellison, J., Salido, E., Yen, P., Mohandas, T. and Shapiro, L.J. (1994) Isolation and characterization of XE169, a novel human gene that escapes X-inactivation. *Hum. Mol. Genet*., **3**, 153–160.
28. Brown, C.J., Miller, A.P., Carrel, L., Rupert, J.L., Davies, K.E. and Willard, H.F. (1995) The DXS423E gene in Xp11.21 escapes X chromosome inactivation. *Hum. Mol. Genet*., **4**, 251–255.
29. Fisher, E.M.C., Beer-Romero, P., Brown, L.G., Ridley, A., McNeil, J.A., Lawrence, J.B., Willard, H.F., Bieber, F.R. and Page, D.C. (1990) Homologous ribosomal protein genes on the human X and Y chromosomes: escape from X inactivation and implications for Turner syndrome. *Cell*, **63**, 1205–1218.
30. Esposito, T., Gianfrancesco, F., Ciccodicola, A., D'Esposito, M., Nagaraja, R., Mazzarella, R., D'Urso, M. and Forabosco, A. (1997) Escape from X-inactivation of two new genes associated with DXS6974E and DXS7020E. *Genomics*, **43**, 183–190.
31. Disteche, C.M. (1995). Escape from X inactivation in human and mouse. *Trends Genet.*, **11**, 17–22.
32. Muraro, T., Stephan, D., Tiso, N., Zimbello, R., Danieli, G.A., Hoffman, E.H., Valle, G. and Lanfranchi, G. (1997) Chromosome assignment of 115 expressed sequence tags (ESTs) from human skeletal muscle. *Cytogenet. Cell Genet*., **76**, 144–152.
33. D'Esposito, M., Mazzarella, R., Pengue, G., Jones, C., D'Urso, M. and Schlessinger, D. (1994) PCR-based immortalization and screening of hierarchical pools of cDNAs. *Nucleic Acids Res.*, **22**, 4806–4809.
34. Skowronski, J., Fanning, T.G. and Singer, M.F. (1988) Unit-length Line-1 transcripts in human teratocarcinoma cells. *Mol. Cell. Biol*., **8**, 1385–1397.

35. Kjeldgaard, M., Nyborg, J. and Clark, B.F.C. (1996) The GTP binding motif: variations on a theme. *FASEB J*., **10**, 1347–1368.
36. Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A., Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M., Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laister, N., Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B., O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A., Saunders, D., Shownkeen, R., Smaldon, N., Smith, A., Sonnhammer, E., Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M., Vaughan, K., Waterston, R., Watson, A., Weinstock, L., Wilkinson-Sproat, J. and Wohldman P. (1994) 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C.elegans*. *Nature*, **368**, 32–38.
37. Noble, J.A., Innis, M.A., Koonin, E.V., Rudd, K.E., Banuett, F. and Herskowitz, I. (1993) The *Escherichia coli* hflA locus encodes a putative GTP-binding protein and two membrane proteins, one of which contains a protease-like domain. *Proc. Natl Acad. Sci. USA*, **90**, 10866–10870.
38. Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M., Tonlorenzi, S. and Willard, H.F. (1991) A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature*, **349**, 38–44.
39. Brown, C.J., Flenniken, A.M., Williams, B.R. and Willard, H.F. (1990) X chromosome inactivation of the human TIMP gene. *Nucleic Acids Res*., **18**, 4191–4195.
40. Gianfrancesco, F., Esposito, T., Ruini, L., Houlgatte, R., Nagaraja,R., D'Esposito, M., Rocchi, M., Auffray, C., Schlessinger, D., D'Urso, M. and Forabosco, A. (1997) Mapping of 59 EST gene markers in 31 intervals spanning the human X chromosome. *Gene*, **187**, 174–184 .
41. Ried, K., Mertz, A., Nagaraja, R., Trusgnich, M., Riley, J.H., Anand, R., Lehrach, H., Page, D., Ellison, J.W. and Rappold, G. (1995) Characterization of a YAC contig spanning the pseudoautosomal region. *Genomics*, **29**, 787–792.
42. Lichter, P. and Cremer, T. (1992) *Human Cytogenetics: A Practical Approach*. IRL Press at Oxford University Press, Oxford.
43. Altschul, S.F., Gish, W., Miller, W., Meyers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol*., **215**, 403–410.
44. Yang, T.T. and Kain, S.R. (1995) Fast hybridization solution for the detection of immobilized nucleic acids. *Biotechniques*, **18**, 498–503.
45. Antonacci, R., Marzella, R., Finelli, P., Lonoce, A., Forabosco, A., Archidiacono, N. and Rocchi, M. (1995) A panel of subchromosomal painting libraries representing over 300 regions of the human genome. *Cytogenet. Cell Genet*., **68**, 25–32.
46. Aguinik, A.I., Mitchell, M.J., Mattei, M.-G., Borsani, G., Avner, P.A., Lerner, J.L. and Bishop, C.E. (1994) A novel X gene with a widely transcribed Y-linked homologue escapes X-inactivation in mouse and human. *Hum. Mol. Genet*., **3**, 879–884.
47. Burk, R.D., Ma, P. and Smith, K.D. (1985) Characterization and evolution of a single-copy sequence from the human Y chromosome. *Mol. Cell Biol*., **5**, 576–581.
48. Wieacker, P., Davies, K.E., Cooke, H.J., Pearson, P.L., Williamson, R., Bhattacharya, S., Zimmer, J. and Ropers, H.H. (1984) Toward a complete linkage map of human X chromosome: regional assignment of 16 cloned single-copy DNA sequences employing a panel of somatic cell hybrids. *Am. J. Hum. Genet*., **36**, 265–276.
49. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
50. Blanchard, M.M., Taillon-Miller, P. and Nowotny, V. (1993) PCR buffer optimization with uniform temperature regimen to facilitate automation. *PCR Methods Appl*., **2**, 234–240.