

# Isolation and Characterization of the Faciogenital Dysplasia (Aarskog–Scott Syndrome) Gene: A Putative Rho/Rac Guanine Nucleotide Exchange Factor

N. German Pasteris,\* Amy Cadle,\* Lindsay J. Logie,†  
Mary E. M. Porteous,† Charles E. Schwartz,‡  
Roger E. Stevenson,‡ Thomas W. Glover,\*§  
R. Sid Wilroy,|| and Jerome L. Gorski\*§

\*Department of Human Genetics

§Department of Pediatrics

University of Michigan

Ann Arbor, Michigan 48109-0688

†Department of Human Genetics

University of Edinburgh

Edinburgh EH4 2XU

Scotland

‡Greenwood Genetic Center

Greenwood, South Carolina 29646

||Department of Pediatrics

University of Tennessee Memphis Health Science Center

Memphis, Tennessee 38105

plasia (Melnick and Shields, 1976). Urogenital malformations include scrotal anomalies, hypospadias, and kidney hypoplasia (Gorlin et al., 1990). Based on these phenotypes, it has been suggested that FGDY be considered a multiple congenital anomaly syndrome involving at least two developmental field defects, the skeletal and urogenital systems (Gorlin et al., 1990).

Pedigree analyses of families segregating FGDY strongly suggested an X-linked recessive pattern of inheritance (Gorlin et al., 1990). Typically, the phenotype of obligate female heterozygotes was limited to relatively short stature and subtler craniofacial anomalies. Although the observation of male-to-male transmission of an FGDY-like syndrome suggested that FGDY may be genetically heterogeneous (Grier et al., 1983; van de Vooren et al., 1983), genetic linkage studies confirmed that, in most affected families, FGDY was X linked and mapped to the pericentric region of the X chromosome (Porteous et al., 1992; Stevenson et al., 1994). FGDY was also mapped to this region based on the observation of a mother and son who both displayed all of the major characteristics of FGDY in association with a reciprocal X;8 chromosome translocation (Bawle et al., 1984). Since the overwhelming majority of disease-associated reciprocal translocations have been found to have breakpoints within a candidate gene locus (Tommerup, 1993), it was reasonable to hypothesize that the FGDY-specific breakpoint directly disrupted the disease locus. Glover et al. (1993) recently sublocalized the FGDY X chromosome breakpoint to a region within Xp11.21 that was flanked by loci *ALAS2* and *DXS323*. Earlier radiation hybrid mapping experiments indicated that these loci were separated by approximately 350 kb (Gorski et al., 1992). Together, these results suggested that it would be possible to use a positional cloning strategy to isolate the FGDY gene.

We report here the isolation and characterization of a candidate gene, termed *FGD1*, for X-linked FGDY. The *FGD1* gene maps to the appropriate physical interval and is disrupted by the disease-specific t(X;8) breakpoint. An *FGD1* mutation detected by single strand conformation polymorphism (SSCP) analysis cosegregates with the disease in a second affected family. *FGD1* shows strong homology to Rho/Rac guanine nucleotide exchange factors (GEFs), proteins involved in growth regulation and signal transduction. These results suggest that molecular defects in this gene cause the developmental growth anomalies underlying FGDY.

## Results

### Isolation of Yeast Artificial Chromosomes Spanning an FGDY Breakpoint

Markers that detect the closest known loci flanking the FGDY X chromosomal breakpoint, *ALAS2* and *DXS323* (Figure 1A), were used to screen a human X chromosome-specific yeast artificial chromosome (YAC) library. Three clones, including *21G3* and *29D4*, were found to contain

## Summary

**Faciogenital dysplasia (FGDY), also known as Aarskog–Scott syndrome, is an X-linked developmental disorder characterized by disproportionately short stature and by facial, skeletal, and urogenital anomalies. Molecular genetic analyses mapped FGDY to chromosome Xp11.21. To clone this gene, YAC clones spanning an FGDY-specific translocation breakpoint were isolated. An isolated cDNA, *FGD1*, is disrupted by the breakpoint, and *FGD1* mutations cosegregate with the disease. *FGD1* codes for a 961 amino acid protein that has strong homology to Rho/Rac guanine nucleotide exchange factors (GEFs), contains a cysteine-rich zinc finger-like region, and, like the RasGEF mSos, contains two potential SH3-binding sites. These results provide compelling evidence that *FGD1* is responsible for FGDY and suggest that *FGD1* is a Rho/RacGEF involved in mammalian development.**

## Introduction

Faciogenital dysplasia (FGDY), which is also known as Aarskog–Scott syndrome, is a rare inherited multisystemic developmental disorder described in the early 1970s by Aarskog (1970) and Scott (1971). The disease phenotype consists of short stature and characteristic facial, skeletal, and urogenital anomalies. Impaired growth is a major manifestation of the disease, and affected males rarely exceed 160 cm in height. Stature is disproportionate and the distal extremities are most severely shortened (Duncan et al., 1977). Radiographic abnormalities include hypoplastic phalanges and retarded bone maturation (Grier et al., 1983). A variety of vertebral anomalies, including cervical spina bifida occulta, odontoid hypoplasia, and segmentation anomalies, occur in about half of affected males (Fryns, 1992). Facial features typically consist of widely spaced eyes, external ear anomalies, and maxillary hypo-

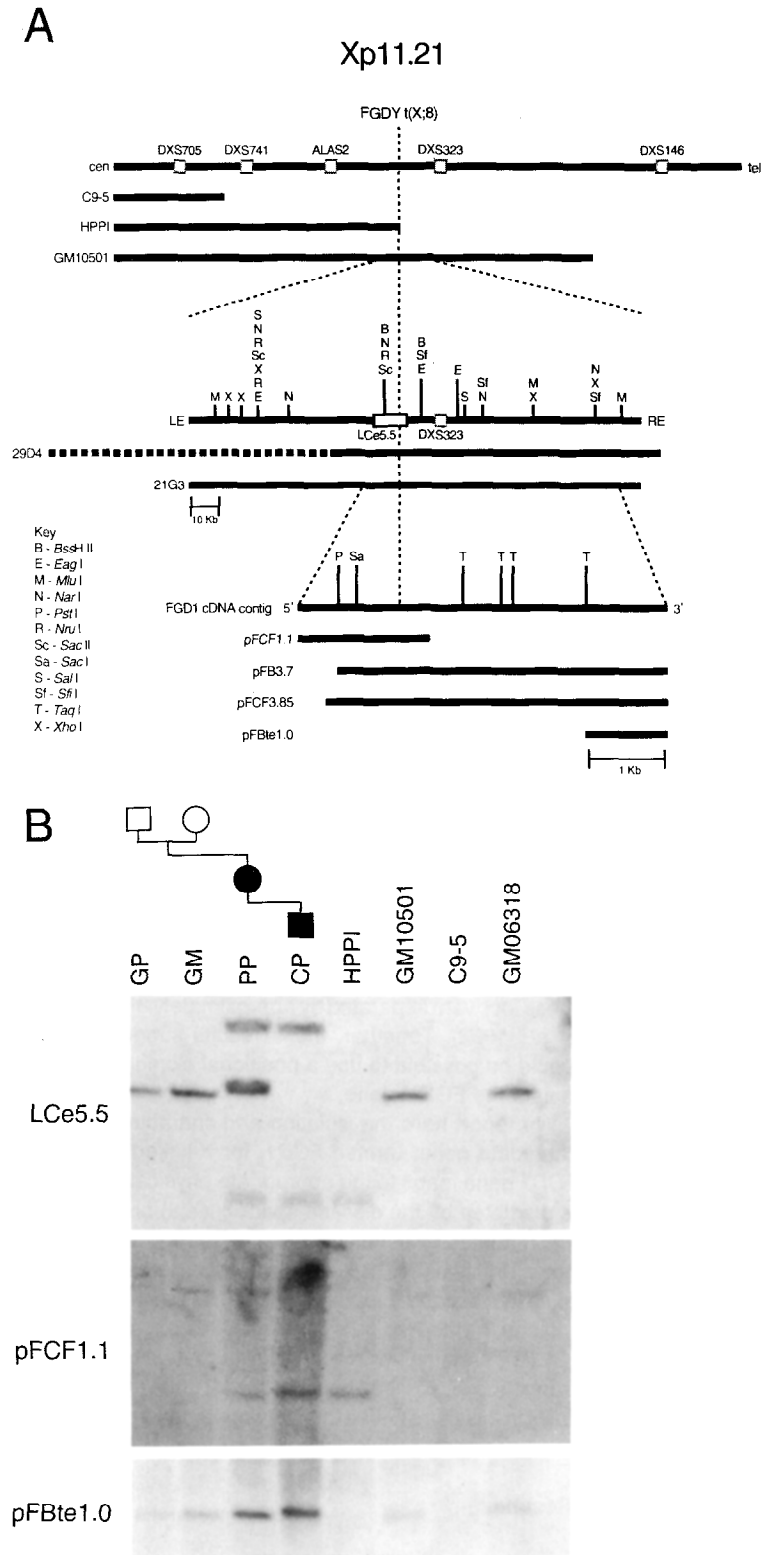


Figure 1. Molecular Map of the FGDY Region and the *FGD1* cDNA Contig

(A) A schematic representation of the FGDY region within band Xp11.21. Locus order was determined previously (Gorski et al., 1992); the FGDY X;8 translocation breakpoint was mapped to the region between loci *ALAS2* and *DXS323* (Glover et al., 1993). Bars indicate the relative X chromosomal content of somatic cell hybrid cell lines C9-5, HPP1, and GM10501 as previously reported (Lafreniere et al., 1991; Gorski et al., 1992); all bars extend to Xqter. A detailed composite long-range restriction map of the FGDY breakpoint region derived from YAC clones 21G3 and 29D4 is shown below; bars indicate relative clone content. LE and RE indicate the left and right ends of clone 21G3, respectively; the interrupted bar indicates a chimeric clone segment. A composite restriction map of the *FGD1* cDNA is shown below the YAC diagram; bars indicate individual cDNA clone content.

(B) Mapping of genomic and cDNA clones to a region spanning the FGDY *t(X;8)* breakpoint. Probe *LCe5.5* and cDNA clones pFCF1.1 and pFBte1.0 were sequentially hybridized to a filter containing *EcoRI*-digested hybrid and cell line DNA. Relative probe location is shown schematically in (A). In the pedigree, affected FGDY family members with the *t(X;8)* translocation are indicated with closed symbols (PP and CP), and unaffected members are indicated with open symbols (GP and GM).

distal marker *DXS323*. To map YAC inserts relative to the *t(X;8)* breakpoint, we recovered YAC insert end clones and mapped them by using a regional Xp11.21 somatic cell hybrid mapping panel. Although *DXS323* mapped dis-

tal to the *t(X;8)* breakpoint and was not contained in a hybrid cell line containing the FGDY-specific derivative X chromosome, hybrid HPP1 (Glover et al., 1993), the left end of clone 21G3 was present in HPP1 (data not shown).

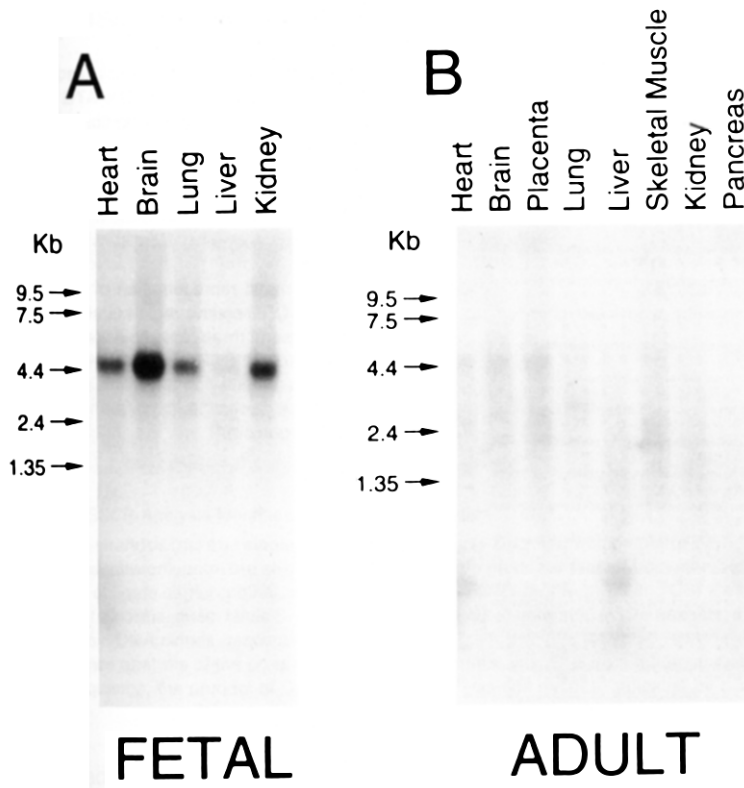


Figure 2. *FGD1* Is Widely Expressed in Human Tissues

The expression of *FGD1* mRNA in selected fetal (A) and adult (B) tissues was investigated by Northern blot analysis. The position of RNA size markers in kilobases is shown.

Fluorescent in situ hybridization studies showed that clone 21G3 was nonchimeric and hybridized to the rearranged X and 8 chromosomes and the normal X homolog (data not shown). Together, these results confirmed that clone 21G3 spanned the FGDY t(X;8) breakpoint. The constructed restriction maps of clones 21G3 and 29D4 were generally in agreement, suggesting that the two conforming YAC insert segments were colinear with genomic DNA (Figure 1A).

Long-range restriction analyses tentatively localized the FGDY t(X;8) breakpoint to a 10 kb *Bss*HII fragment (data not shown); this fragment was contained in clone 21G3 and was flanked by two potential CpG islands (Figure 1A). A  $\lambda$  phage subclone contig of clone 21G3 was constructed (Pasteris et al., 1993), and phage DNA inserts were examined to identify low copy (LC) and evolutionarily conserved sequences. One such fragment was identified from the 10 kb interval, a 5.5 kb *Eco*RI fragment, *LCe5.5* (Figure 1A). When hybridized to a filter containing DNA derived from FGDY t(X;8) family members, probe *LCe5.5* detected the normal 5.5 kb *Eco*RI fragment in the unaffected family members GP and GM and in affected member PP and detected two unique fragments, 8.0 and 3.0 kb in size, in affected members PP and CP (Figure 1B). Only the 3.0 kb fragment was detected in cell line HPPI. The 8.0 and 3.0 kb fragments were detected only in the DNA of the affected family members. The 5.5 kb fragment alone was observed in the DNA of 16 different unrelated individuals; similar results were obtained by using other restriction enzymes (data not shown). Together, these results indicated that probe *LCe5.5* spanned the FGDY translocation breakpoint and detected the junctional *Eco*RI fragments

of both the derivative X and the reciprocal derivative 8 chromosome. These results also showed that the 3.0 kb fragment was derived from the FGDY derivative X chromosome and that, at this level of resolution, no apparent deletion was associated with the translocation.

#### Isolation of Candidate cDNA Clones

Probe *LCe5.5* was used to screen human 16-week-old fetal brain (FB) and fetal craniofacial (FCF) cDNA libraries. We isolated eight independent FCF clones and two distinct FB clones from  $2 \times 10^6$  clones screened. Of the positive FCF clones, three were found to be identical by restriction mapping and were designated clone pFCF1.1. This clone was used to rescreen the FCF library and isolate 30 additional clones; these clones were analyzed to construct a cDNA contig (Figure 1A). To verify that the isolated cDNA clones spanned the FGDY breakpoint, cDNA subclones derived from the extreme portions of the constructed contig were hybridized to FGDY t(X;8) family member DNA. As shown in Figure 1B, pFCF1.1, a probe containing the 5' end of the cDNA, detected a pattern of 5.5 and 3.0 kb fragments that was superimposable to that detected by *LCe5.5*, including the 3.0 kb fragment detected in HPPI DNA. In contrast, pFBte1.0, a probe containing the 3' end of the cDNA, detected a 2.5 kb *Eco*RI fragment that was present in the DNA of FGDY t(X;8) family members and in hybrid cell lines GM10501 and GM06318, but absent in HPPI DNA. These results mapped pFCF1.1 proximal to (and pFBte1.0 distal to) the FGDY breakpoint and confirmed that the cDNA, termed *FGD1*, was directly disrupted by the FGDY t(X;8) breakpoint.

A

```

5' - CAGGGCTCCTTCCGCGCCCGCCGCGAGACCCAGGCTGAAGTGGGGAGGAGGTGGAGCCGCGCCAGGCGCTCT
80 CCTGCCCCGCGGCTGGAGCTCTGGGACCTACCTGGGCTGAGGCGAGCCCGGACCCAGCCAGGACACGCGCCCGCCAGCCCGCTCTCT
182 TCCCTCCCTCTCTTCCCTCCCTCTCTCCGCGCCGCGCTGGGCTGGGGAGGAGAACCCCTCCCAACAGCCAGGAGAGCCAGATCATCTCC
284 ATTCGGGGAAGAGAGAGACCCCACTTGGATCCGACGACTTTCGCGCCGGGCGACGAAATGCTGCGGCTCCAGGAGTCCCGCCGAGCGCT
386 AGCCCTGATACCTCCGACAGCCGAGCCGATGAGAGGCGCTCCGAGCGGCTCCCGGCTCCGCGCCGCTCCGCGGCGGCGGCGGCGG
488 CGATCTCCGACAGCCGAGCGGCGGCGGCGGCTGCTGCTGCTACTGACCCGCTCCGACCGGCTCTCCGCGCGGCGGCGGCGGCGGCGG
590 TCCAGTGGGCTGCGCCGAGCTCCCGCCGCGCTGCGGCGGCGGCGGCGGCTCCCGCGCGGCTCCCGCCAGGAGCTGGAG
692 CCAAGCGGGAGCTCGGCGCCCGCCAGGCTCCGAGCTGCGGCGGCGGCGGCGGCGGCTTCGGAGCCCGAACCCGCGCCAGAA
      M D G H R A P Q G G R A F G A R T P F G H E
794 CCGCGCGCGCCGCTCCGCGGCTCTGGGAGCTTCGACCCGAGCTTCGGGCGGAGGGGCTGAGGTTCCGCTCTGGCGCCCA
P A G A A P P A C A D S D P G A S E P G L L A R R O S G S A L G G P
896 TGGGATCCCACTTGTGGGACCTTCGAGACCCAGCTGGGCGGCTTCGACGCGCCGCGGCTTTCGCTCCGCGGCGCACTTCAGCAGCACCACCGGCGCCT
L D P Q F V G P S D T S L G A A P G H R V L P C G F S P Q H R R A L
998 CGCTTCTTACCACTGGAGGCTCCGACGCTGGGCTGGGCTGACACAGGAAACCGATCTGCTGTAAGATTTGTCCTTACCCCTGGCCAAAGSCTA
R F S Y H L E G S Q P R P F G L H Q G N R I L V R S L S L D P G Q S L
1100 GAGCCTCTCCAGAGGTCCCGAGCGCTTCCGTCAGACCCAGGCTCCCGGCTGAAACCCCTAGCCAGGCTCTTACACCACTAAGCGGCGCAGCGGCGCG
E F P H P E C G P Q R L R S D P G P P T E T P S Q R P S P L K R A P G P
1202 AAGCAGAGGCTCCGCGGCTCCGCGGCTGCGGAGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
K P P Q V P P K P S Y L Q M P R M P P P L E P T I P P P S R P L P A D
1304 CACCAATCGGACCGGCTGGGCTCCGAGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
P R V Y G K G L A P R A E A S P S S A A V S S L I E K F E R E P V I V
1406 GCTGCGATAGACCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
A S D R P V P G P S P G P P E P V H L P Q P T S Q P P V P Q L P E G
1508 GAGCCCTGCGGCTCCGCGGCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
A E A S R C L F L L A P G P R D G E K V P N R D S G I D S I S S P S N
1610 AGCAGGAGACCTCTTCCGCTGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
S E E T C F V S D D G P P S H S L C P G P P A L A S V P V A L A D P
1712 CAGCCGCTGGCTCCGCGGCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
H R P G S Q E V D S D L E E E D D E E E E E E E K D R E I P V P L M E
1814 AGACAGAGCTCTGGGAGTGGAGTCCAGCAAGAGTCTTTCACATTCGCAATGAGCTCTGCAAACTGAGAGGCGCTACGTTTCAGGCTCCATCTCTG
R Q E S V E L T V Q Q K V P H I A N E L L Q T E K A Y V S R L H L L
1916 GATCAGGCTTCCGCGGCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
D Q V P C A R L L E E A R N R S P P A D V V H C I F S N I C S I Y
2018 TGCTCCAGCAGCTCCGCGGCTCCGCGGCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
C F H Q Q F L L F E L E K R M E R W D R Y P R J G D I L Q K L A P P
2120 CTCAGATATGTTGAGTCTGAGAACTTGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
L K M Y G E Y V K N F D R A V E L V N T W T E R S T Q F X V I H E
2222 GTGCGAGAGGAGGCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
V Q X E E A C N L T L Q H M L E P V Q R I P R Y E L L L K D Y L
2324 TTAAGCTCCGCGGCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
L K L P H G S P D S K D A Q K S L E L I A T A A E H S N A A I R K M
2426 GAGCGAATGAGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
E R R M H K L L X V Y E L L G G E E D I V S P T X E L I K E G H I L K
2528 CTTGACCAAGATGGGAGCCTCAAGACCGATCCCTACTATTCACAGCAGCGCTCTTACTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG
L S A K N G T T Q D R Y L I L F N D R L L Y C V P R L R L L G Q K F
2630 AGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
S V R A R I D V D G M E L K E S S N L N L P R T F L V S G K Q R S L
2732 GAGCTCCAGCAGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
E L Q A R T E E E K R D W V Q A I N S T L L K H E Q T L E T F K L L
2834 AACTCAACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
N S T N R E D E D T P P N S P N V D L G K R A P T P I R E K E V T M
2936 TGCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
C M H C Q R P F N S I T X R R H H C K A C G H V V C G K C S B P R A
3038 CGCTCTGATGACCAACCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
R L V Y D N N R S N R V C T D C Y V A L H G V P O S S P A C S Q H T
3140 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
P Q R R R S I L E K Q A S V A A E N S V I C S F L H Y M E K G G K G
3242 TGGCAGAGCAGCTTCTGCTGCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
M H K A M F V V P N R P L V L Y I Y G A P Q D V K A Q R S L P L I
3344 GCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
G P F V G P P E A G E R P D R R H V P K I T Q S H L S W Y F S P E T
3446 GAGAACTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
E E L Q R R W N A V L G Q R A G R C D T F C P G P T L S E R D R E M E
3548 GCACCGTGGCTTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
A P V A A L G A T A E P P E S P Q T R D K T
3550 CACTACTAGTCCGATGCTGATGGGCGGCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG
GTCTTGGTGTTTTAAATGCTCTTCCAGAAAACCGGCGGCTTTAAAAAATATATCCACAGTATGCAATTTTATTATCCCTCCCGGAGG
3752 GTGGAGCCGCTGCGGCTCCACTGAAATGAGTGTCTTCTCCCTATCCCTCAATACCTGAACTCCCATCCAGCGGCGGCTTCAATCCCGAATCTGG
3854 GTGGAGCCGCTGCGGCTCCACTGAAATGAGTGTCTTCTCCCTATCCCTCAATACCTGAACTCCCATCCAGCGGCGGCTTCAATCCCGAATCTGG
3956 CAGAGGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
4058 GTGGAGCCGCTGCGGCTCCACTGAAATGAGTGTCTTCTCCCTATCCCTCAATACCTGAACTCCCATCCAGCGGCGGCTTCAATCCCGAATCTGG
4160 GCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
4262 TACAG - 3'
    
```

B

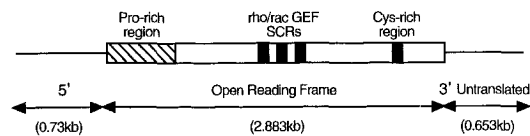


Figure 3. Nucleotide and Deduced Protein Sequence of FGD1

(A) DNA and deduced amino acid sequence of the gene encoding FGD1. The ATG codon representing the putative translation start site is underlined. An upstream and putatively non-functional ATG initiation codon and two in-frame TGA termination codons are boxed. Within the predicted FGD1 protein sequence, six potential N-linked glycosylation sites are underlined.

(B) A schematic representation of the *FGD1* gene. The ORF coding region is shown as a boxed segment; three closed blocks represent the Rho/RacGEF SCRs, the hatched area represents the proline-rich 5' region, and the 3' closed block represents the cysteine-rich zinc finger-like domain.

**Expression of the *FGD1* Gene**

Hybridization of the *FGD1* cDNA to a Northern blot containing poly(A)<sup>+</sup> RNA detected a 4.4 kb transcript in fetal heart, brain, lung, kidney, placenta, and (to a lesser extent) liver (Figure 2A). A similarly sized transcript was detected in adult heart, brain, lung, skeletal muscle, and (to a lesser extent) pancreas and liver (Figure 2B). Several other less-prominent transcripts that may represent alternatively spliced or cross-hybridizing transcripts were observed in placenta and adult brain, lung, and skeletal muscle.

***FGD1* cDNA Sequence Analysis**

Both strands of the composite *FGD1* cDNA were sequenced in at least two independent clones over the entire

length of the contig. In addition, 90% of the sequence was also confirmed by sequencing cloned genomic DNA (N. G. P. et al., unpublished data). The cDNA was 4266 bp in length and contained a 2883 nt open reading frame (ORF) that was predicted to encode a protein of 961 amino acids with a predicted mass of 107 kDa (Figure 3). The putative initiation codon of this ORF began at nucleotide 731, and the flanking sequences conformed to a translation initiation consensus sequence (Kozak, 1987). Preliminary data indicated that the first 936 bp of the cDNA was contained in a single exon that mapped proximal to the *FGDY* (X;8) breakpoint (data not shown), indicating that the predicted ORF was directly disrupted by the breakpoint.

The sequence 5' of the proposed initiating methionine

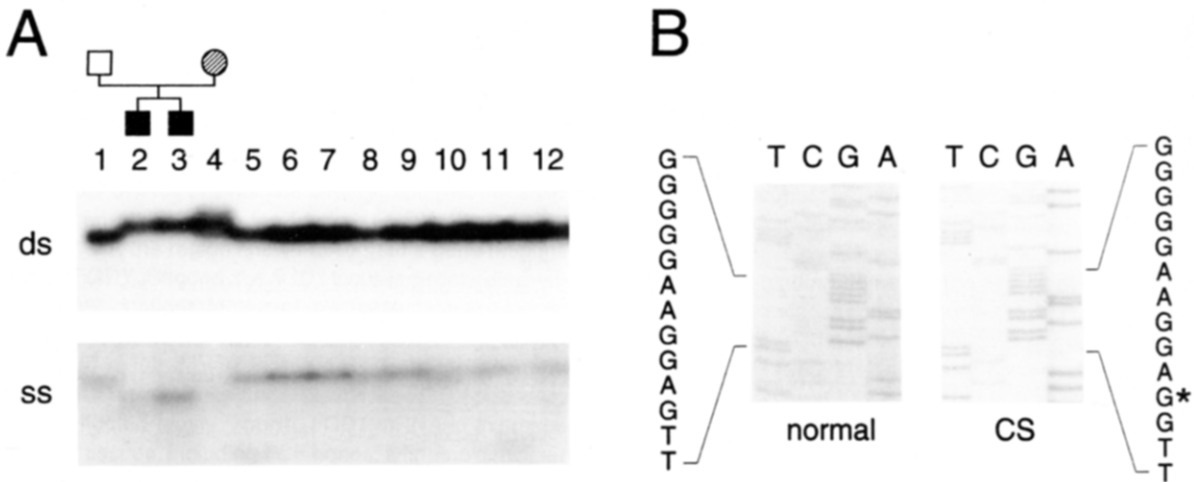


Figure 4. SSCP Analysis Identifies an Insertional Mutation

(A) Double-stranded (ds) and single-stranded (ss) products amplified with primer pair *E7-1* are shown; to be most informative, different exposures of the same autoradiogram are shown. Lanes 1–4 contain products derived from members of a family segregating FGDY; products of the normal father, the obligate carrier mother, and the two affected FGDY brothers (TS and CS) are in lanes 1, 4, 2, and 3, respectively. Lane 5 contains the products of a normal male; lanes 6–10 contain the products of unrelated FGDY patients; and lanes 11 and 12 contain products derived from cDNA and genomic DNA clones, respectively.

(B) Sequence analysis of the products amplified with primer pair *E7-2* from a normal male and the FGDY-affected male CS. Compared with the normal sequence, the product of CS had an additional guanine residue, indicated by a star, inserted at nucleotide 2122.

codon had stop codons in all three ORFs. An upstream and putatively nonfunctional initiation codon was found at nucleotide 490 (Figure 3A). Several observations suggested that this initiation codon was an unlikely translation start site: the potential ORF associated with this ATG codon was predicted to encode a low complexity 37 amino acid peptide, two in-frame termination codons were found between the first and second potential initiation codons, and this codon was predicted to be an unlikely site of translation initiation (Kozak, 1987).

#### Mutation Analysis

For the portions of the *FGD1* cDNA with known exon/intron boundaries, DNA from 18 unrelated FGDY males was analyzed by SSCP. Four different primer pairs were used collectively to amplify and analyze 526 bp (12.3%) of the cloned *FGD1* gene. All of the affected males exhibited the cardinal clinical features of FGDY. Although four of the affected males represented isolated cases, most patient families displayed X-linked inheritance and linkage to DNA markers within Xp11.21 (Porteous et al., 1992; Stevenson et al., 1994). In an FGDY family with two affected male siblings and an obligate carrier female, a similar band shift was seen in all three individuals for the *E7-1* primer pair (Figure 4A). The double-stranded product of each of the FGDY male siblings appeared to be slightly larger than the normal product. The DNA of the carrier female appeared to yield three products (a normally sized product, a mutant product, and a larger and presumably heteroduplex product), findings consistent with her known heterozygote status. The single-stranded products of all three individuals demonstrated anomalous gel mobility (Figure 4A). Compared with the normal product, sequence analysis showed that the aberrantly sized product contained an insertional

mutation with an additional guanine residue present at nucleotide 2122 (Figure 4B); the resulting frameshift mutation was predicted to cause premature translational termination at residue 469. In contrast, an analysis of 82 normal X chromosomes with the same primer set failed to yield an abnormal product.

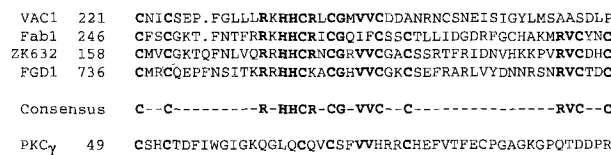
#### Predicted Amino Acid Sequence of FGD1

Comparison of the predicted FGD1 protein sequence with other sequences in the data bases indicated that the *FGD1* gene encoded a putative Rho/RacGEF. Rho/RacGEFs form a family of cytoplasmic proteins that activate the Ras-like family of Rho and Rac proteins by exchanging bound GDP for free GTP (Boguski and McCormick, 1993). Like the other Rho/RacGEF family members, the *FGD1*-derived amino acid sequence contained three structurally conserved regions (SCRs) spanning 180 amino acids (Figure 5A). Sequence alignment of the predicted FGD1 protein with other Rho/RacGEF proteins illustrated the remarkable similarity between the predicted FGD1 protein and other members of the GEF subfamily over the 180 residue region of functional significance. Among Rho/RacGEFs, the predicted FGD1 protein showed the greatest similarity ( $p = 1.4e-6$ ) to the *ect2* protein (Miki et al., 1993). However, the predicted FGD1 protein showed strong similarity to several other family members, including RasGRF ( $p = 1.3e-6$ ) (Shou et al., 1992), Dbl ( $p = 8.8e-5$ ) (Ron et al., 1991), and Bcr ( $p = 5.3e-5$ ) (Lifshitz et al., 1988). An examination of the sequence immediately 3' of the conserved Rho/RacGEF domain predicted that the FGD1 protein also contained a pleckstrin homology motif (data not shown), a feature common to other Rho/RacGEF family members (Musacchio et al., 1993).

**A**



**B**



**C**

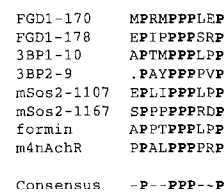


Figure 5. Comparative Analysis of Potential FGD1 Structural Domains

(A) Amino acid alignment between FGD1 and the SCR domains of Rho/RacGEF proteins. Amino acids comprising previously identified SCRs are boxed (Boguski and McCormick, 1993). Residues common to four or more proteins are indicated as consensus amino acids and shown in bold. Dbl and Bcr are human sequences, proto-Vav and ect2 are mouse sequences, and RasGRP and Cdc24 are from the rat and *S. cerevisiae*, respectively; corresponding GenBank accession numbers are P10911, L11316, Y00661, S23669, P28818, and P11433.

(B) Amino acid alignment between FGD1 and proteins containing cysteine-rich zinc finger-like domains. Residues common to three or more proteins are indicated as consensus amino acids and shown in bold. Vac1 and Fab1 are from *S. cerevisiae*, and ZK632.12 and PKC $\gamma$  are *C. elegans* and human sequences, respectively; corresponding GenBank accession numbers are P32609, P34756, P34657, and M13977.

(C) Amino acid alignment between FGD1 and proteins containing SH3-binding sites (Ren et al., 1993; Rozakis-Adcock et al., 1993). Residues common to all of the aligned amino acids are shown in bold and indicated as a consensus sequence. The m4nAChR sequence is from the rat (Bonner et al., 1987); 3BP1 and 3BP2 (Cicchetti et al., 1992), mSos2 (Bowtell et al., 1992), and formin (Woychik et al., 1990) are mouse sequences.

Comparison of the derived FGD1 protein sequence with other sequences in the data bases suggested that it contained two additional structural motifs (see Figure 3B). The 3' region of the sequence was predicted to encode a 50 amino acid cysteine-rich region that contained a zinc finger-like structural motif. Although this region showed the greatest similarity with two proteins of unknown function, the *Caenorhabditis elegans* protein ZK632.12 ( $p = 7.1 \times 10^{-12}$ ) and the *Saccharomyces cerevisiae* protein Fab1 ( $p = 1.9 \times 10^{-5}$ ), this region showed some degree of similarity to the zinc finger-like regions of both the Vac1 protein (Weisman and Wickner, 1992) and the diacylglycerol/phorbol diester-binding regulatory domain of protein kinase C $\gamma$  (PKC $\gamma$ ) (Quest et al., 1994). The sequence alignment of the predicted FGD1 protein with the other zinc finger-like proteins is shown in Figure 5B. This alignment illustrated that the predicted FGD1 zinc finger-like structural motif, C-X<sub>2</sub>-C-X<sub>18</sub>-C-X<sub>2</sub>-C-X<sub>4</sub>-C-X<sub>2</sub>-C-X<sub>17</sub>-C-X<sub>2</sub>-C (single letter code; X denoting variable residues), was identical to that found in the ZK632.12 and Fab1 protein sequences and similar to the PKC $\gamma$  motif. The remarkably strong conservation of this region among proteins derived from organisms as distantly related as humans, nematodes, and yeast inferred that this region was functionally significant and suggested

that this region contained a cysteine-rich zinc finger-like domain that was similar to, but distinct from, that contained in PKC $\gamma$ .

An analysis of the predicted FGD1 protein showed that the 5' region was remarkably proline-rich and that proline constituted 22% of the first 250 amino acid residues. Since proline-rich regions have been shown to contain Src homology 3 (SH3)-binding domains (Egan et al., 1993; Rozakis-Adcock et al., 1993), the 5' portion of the derived FGD1 protein sequence was compared with other sequences known to contain SH3-binding sites. As shown in Figure 5C, this comparison identified two partially overlapping segments of the predicted FGD1 sequence that exhibited strong similarity to the functionally significant regions of several proteins with demonstrated SH3 binding activity, including the Abl SH3-binding proteins 3BP1 and 3BP2 (Ren et al., 1993) and the RasGEF mSos2 (Egan et al., 1993; Rozakis-Adcock et al., 1993). The derived consensus sequence for the alignment was identical to that proposed by Ren et al. (1993), who showed that mutations within this consensus sequence abolished SH3 domain binding. Short peptides containing similar sequences have been shown to be necessary and sufficient for SH3 binding (Li et al., 1993; Rozakis-Adcock et al.,

1993). These results suggested that, like 3BP1, 3BP2, and mSos2, the FGD1 protein contained an SH3-binding domain.

## Discussion

A number of lines of evidence strongly suggest that *FGD1* is the gene responsible for X-linked FGDY. First, it maps to Xp11.21, the region known to contain a gene responsible for FGDY. Second, the *FGD1* locus is directly disrupted by a translocation breakpoint in a family in which the disease is segregating. Third, an insertional mutation, predicted to result in a severely abbreviated and nonfunctional FGD1 protein, segregates with the FGDY phenotype in an affected family. Fourth, *FGD1* mRNA is expressed in most tissues, including FCF bones, a finding consistent with the disease phenotype. Finally, the derived FGD1 amino acid sequence predicts a protein with Rho/RacGEF activity. In addition, this predicted sequence contains two other potential structural motifs: a cysteine-rich zinc finger-like structural domain that is related to the regulatory domain of PKC $\gamma$  and two putative SH3-binding domains. The remarkable similarity of *FGD1* to several proto-oncogenes involved in the regulation of growth-related signal transduction (i.e., *Dbl*, *Bcr*, *ect2*, and *Vav*), further suggests that the FGD1 protein is responsible for the altered pattern of embryonic development observed in FGDY.

Our assertion that FGD1 is a putative Rho/RacGEF is based upon the results of our comparative sequence analysis. The predicted FGD1 protein sequence contains a 180 amino acid structural motif that is significantly similar to other Rho/RacGEF family members, including Cdc24, Dbl, and *ect2*. Several lines of evidence indicate that members of this family do indeed represent Rho/RacGEFs. In yeast, CDC24 has been shown to regulate *CDC42Sc* (Bender and Pringle, 1989), whose human homolog, *CDC42Hs*, encodes a Rho-related GTP-binding protein involved in the regulation of cytoskeletal organization (Shinjo et al., 1990). Dbl has been shown specifically to stimulate guanine exchange activity for human CDC42Hs and RhoA (Hart et al., 1991, 1994). Although *ect2* has not been found to stimulate exchange activity for CDC42Mm, it has been shown specifically to bind to several Rho-related proteins, including Rac1, RhoA, and RhoC (Miki et al., 1993). Furthermore, the structurally conserved 180 amino acid motif identified in FGD1 represents the only common sequence contained in all Rho/RacGEF protein family members (Boguski and McCormick, 1993). It has recently been shown that the transforming and GEF activities of the Dbl protein are carried in a 238 amino acid peptide that contains this 180 residue domain (Hart et al., 1994). These results indicate that this domain is sufficient and specific for Rho/RacGEF activity and suggest that FGD1 is a family member. However, the recent demonstration that Vav, a putative Rho/RacGEF family member, stimulates the exchange of GTP for p21<sup>ras</sup> but not for Rho-related proteins (Gulbins et al., 1993), suggests that this motif may not be an absolute predictor of GEF specificity (Boguski and McCormick, 1993).

The results of our comparative sequence analysis also

indicate that, like other Rho/RacGEFs, the predicted FGD1 sequence contains two additional structural domains that may either modify or regulate its activity or do both: a putative SH3-binding domain and a cysteine-rich zinc finger-like domain. SH3 domains have been shown specifically to bind to short (9–10 amino acid) proline-rich structurally conserved motifs contained in a variety of proteins, including the RasGEF Sos (Li et al., 1993; Ren et al., 1993; Rozakis-Adcock et al., 1993). Like the putative FGD1 SH3-binding domains, the Sos-binding motifs are quite similar to those contained in 3BP1, 3BP2, and formin, domains with demonstrated Abl SH3 domain specificity (Egan et al., 1993; Li et al., 1993; Rozakis-Adcock et al., 1993). Recently, it has been shown that a protein composed entirely of an SH2 domain flanked by two SH3 domains, Grb2 (homologous to Sem-5 in *C. elegans* and Drk in *Drosophila*), selectively binds the proline-rich motifs of the Sos protein to form a link in a signal transduction pathway that functionally ties tyrosine kinase receptors to Ras (Buday and Downward, 1993; Egan et al., 1993; Li et al., 1993; Rozakis-Adcock et al., 1993). Among the identified Ras and Rho/RacGEF family members, Sos is unique in containing an SH3- or Grb2-binding domain (Boguski and McCormick, 1993). The identification of a putative proline-rich SH3-binding domain in the predicted FGD1 protein implies that, like Sos, the location, activity, or both of the FGD1 protein may be modified by Abl- and Grb2-like proteins.

Like the product of the proto-oncogene Rho/RacGEF member Vav (Coppola et al., 1991), the predicted FGD1 sequence also contains a putative cysteine-rich zinc finger-like domain that is related to but distinct from the diacylglycerol and phorbol diester-binding regulatory domain of PKC $\gamma$  (Quest et al., 1994). The zinc finger-like domain of the Vav protein has been shown to be functionally significant; mutations of the conserved cysteine residues were found to abolish the transforming activity of the proteins (Coppola et al., 1991). Similar putative regulatory domains have been identified in a variety of Ras-associated proteins, including the product of the *raf* proto-oncogene, the Rho/Rac GTPase-activating protein n-chimerin, and diacylglycerol kinase (Liscovitch and Cantley, 1994). The observation that the predicted FGD1 sequence contains a putative zinc finger-like domain suggests that the FGD1 protein may interact with lipid second messenger molecules. Therefore, it is possible that the FGD1 protein may interact with the components of multiple signal transduction pathways.

Since both of the identified FGDY mutations predict the total absence of FGD1 protein in affected individuals, our findings predict that the constitutional loss of FGD1 Rho/RacGEF activity in humans results in the clinical FGDY phenotype of disproportionately short stature and craniofacial, skeletal, and urogenital anomalies. However, our results do not suggest a specific mechanism as to how the absence of the FGD1 protein would result in the observed developmental anomalies. In *S. cerevisiae*, *Schizosaccharomyces pombe*, *Drosophila melanogaster*, and *C. elegans*, the genetic loss of GEF function has effects similar to a loss of the associated Ras protein (see

Boguski and McCormick, 1993). In mammals, Rho protein family members consist of at least eight distinct proteins (RhoA, RhoB, RhoC, Rac1, Rac2, TC10, CDC42Hs1, and CDC42Hs2) (Downward, 1992; Boguski and McCormick, 1993). Rho-like proteins have been shown to be involved in the regulation of the architecture of the actin cytoskeleton (Ridley and Hall, 1992; Ridley et al., 1992), yeast bud site assembly and cell polarity (Drubin, 1991), and the NADPH oxidase system in phagocytes (Abo et al., 1991). Therefore, the study of FGD1 may provide important information regarding the roles that Rho- and Rac-like proteins play in signal transduction and mammalian development. Furthermore, if, as we expect, FGD1 interacts with other signal transduction components, genes that encode FGD1-associated proteins may be responsible for other similar inherited developmental disorders. Such anticipated interactions may provide an explanation for the previous reports of FGDY genetic heterogeneity.

Several Ras-associated proteins have been found to be responsible for human genetic diseases. Neurofibromin (NF1), the human protein defective in neurofibromatosis type 1, contains a RasGAP domain homologous to the catalytic domains of p120-GAP, IRA, and IRA2 (Buchberg et al., 1990). As recently reported by the European Chromosome 16 Tuberous Sclerosis Consortium (1993), tuberin, the gene responsible for the form of tuberous sclerosis mapped to chromosome 16 (*TSC2*), has some homology to the GAP3 protein. The gene responsible for choroderemia, a retinal degeneration syndrome, is similar to a Rab geranylgeranyl transferase (Seabra et al., 1993). However, because FGD1 is a member of the Rho/RacGEF family directly implicated in causing an inherited human disease, the study of FGD1, in the context of its normal function and regulation, will provide important information regarding the roles that Ras-like GEFs play in signal transduction and mammalian development. The demonstrated strong homology of other Ras-associated family members implies that it may be possible and valuable to study homologs of the FGD1 protein in other animal systems. Furthermore, studies of the features of FGD1 that resemble those of the transforming Rho/RacGEF family members may aid in the understanding of oncogene regulation and the events leading to transformation. Indeed, it will be interesting to determine whether FGD1 is involved in human neoplastic disease.

#### Experimental Procedures

##### Somatic Cell Hybrids and Lymphoblastoid Cell Lines

Epstein-Barr virus-transformed lymphoblastoid cell lines were derived from a female patient with FGDY and an X;8 translocation (PP), her similarly affected son (CP), and her normal parents (GM and GP) (Bawle et al., 1984). HPPI is a human-hamster hybrid cell line that contains the translocated X chromosome t(X;8)(p11.21;11.21) derived from PP (Glover et al., 1993). C9-5 contains the translocated X chromosome t(X;9)(p11.21;q34.3) derived from a female with incontinentia pigmenti type 1 (Gorski et al., 1989, 1992). GM10501 contains the translocation chromosome t(X;17)(17qter-17q11.2::Xp11.21-Xqter) (Lafreniere et al., 1991). GM06318 contains only a structurally intact human X chromosome (Coriell Medical Institute). This panel of hybrid cell lines has been previously described (Gorski et al., 1992).

##### YAC Clone Analyses

YAC clones were isolated from a human Xpter-Xq27.3-specific library by colony hybridization (Lee et al., 1992). DNA markers that detect *ALAS2* and *DXS323* have been described (Gorski et al., 1992). High molecular weight yeast DNA was prepared and analyzed as described (Gorski et al., 1992). Restriction maps were obtained by hybridizing YAC vector-specific and insert-specific probes to blots of fully and partially digested PFGE-fractionated YAC DNA. YAC insert end clones were recovered by inverse PCR (Arveiler and Porteous, 1991), confirmed by restriction analysis, and mapped to Xp11.21 by using fragments as probe to perform somatic cell hybrid and YAC DNA hybridizations.

##### Standard and PFGE Hybridization Analysis

DNA was isolated, digested, transferred to Hybond-N nylon membranes (Amersham), and hybridized as described (Gorski et al., 1992). Hybridizations were performed for 16–24 hr at 65°C, and membranes were washed to a final stringency of 0.1 × SSC, 0.1% SDS at 65°C for at least 10 min (Gorski et al., 1992). Plasmid DNA was prepared by standard techniques; prior to use as probe, DNA fragments were recovered from agarose gel by electroelution (Sambrook et al., 1989) or by adhesion to a silica matrix (BIO 101).

##### cDNA Isolation and Characterization

$\lambda$ ZAPII human 16-week-old FB and FCF cDNA libraries were screened by standard methods (Sambrook et al., 1989). Bluescript KS(+) plasmid clones containing cDNA inserts were isolated by *in vivo* excision as per recommendations of the supplier (Stratagene). Fetal and adult tissue poly(A)<sup>+</sup> RNA Northern blots were obtained from Clontech and hybridizations were performed as described (Sambrook et al., 1989). Double-stranded cDNA and genomic subclones were sequenced by using Sequenase version 2.0 as per the recommendations of the manufacturer (United States Biochemicals). Oligonucleotides for sequencing and PCR amplification were synthesized by standard phosphoramidite chemistry on an Applied Biosystems model 394 DNA synthesizer.

##### Mutation Analysis

Portions of the *FGD1* cDNA for which exon/intron boundaries were known were PCR amplified and analyzed by SSCP (Orita et al., 1989). Primer pairs were used as follows: *E7-1*, GGACCGCTATCCACGCATTGGAGA and TCAAAGTTCTTCACATACTCACC, which amplify cDNA sequences from bases 2072 to 2154; *E4-1*, ATACCTGCTTCCATGTGTGTCC and GAAACAGGCAGCGGGAGGCCTCA, which amplify cDNA sequences from bases 1388 to 1529; *E4-2*, TGAGGCTCCCGTGCCTGTTTC and GGCAACAGGCACACTAGCCAGGG, which amplify cDNA sequences from bases 1507 to 1699; *E4-3*, CCCTGGCTAGTGTGCTGTTGCC and CTCTCGGTGAACACAGGTCAGT, which amplify cDNA sequences from bases 1677 to 1830. Exon amplifications were performed in 50  $\mu$ l volumes containing 50 mM KCl, 10 mM Tris (pH 8.0), 10 mg/ml BSA, 1.5 mM MgCl<sub>2</sub>, 200  $\mu$ M each of dATP, dGTP, and dTTP, 20  $\mu$ M dCTP, 1  $\mu$ Ci of [ $\alpha$ -<sup>32</sup>P]dCTP, 100 ng of genomic DNA, 0.5  $\mu$ M primers, and 2.5 U of AmpliTaq (Perkin-Elmer Cetus). Amplifications were performed in a DNA Thermal Cycler 480 (Perkin-Elmer Cetus) with 35 cycles of 1 min of denaturation at 94°C, 1 min of annealing at 55°C–59°C, and 2–4 min of extension at 72°C.

For SSCP analysis, 2  $\mu$ l of a PCR was mixed with an equal volume of stop solution (United States Biochemicals), boiled for 5 min, cooled on ice for 2 min, and loaded on a 6% polyacrylamide gel containing 5% glycerol and 0.5 × TBE buffer. Samples were electrophoresed at both 4°C for 4–6 hr at 20 W and overnight at room temperature at 1–5 W of constant power. PCR products were sequenced as described (Casanova et al., 1990) with slight modifications. Amplification was performed by substituting primer pair *E7-2* (GGACCGCTATCCACGCATTGGAGA and ATGCGCTGCACAGGCTCCAGCAT), which amplifies cDNA sequences from bases 2072 to 2289, and 180  $\mu$ M dCTP for the 1  $\mu$ Ci of [ $\alpha$ -<sup>32</sup>P]dCTP. Product was purified by adhesion to silica matrix (BIO 101) and resuspended in 21  $\mu$ l of water; 7  $\mu$ l of the product (approximately 200 ng) was sequenced as above by using primer CTGCACCTCATGGATGATGACTT (bases 2227–2205), added in 20 M excess. Prior to sequencing, the reaction was boiled for 2 min and



snap frozen in dry ice and ethanol. The reaction was performed for 45 s at room temperature and terminated for 5 min at 37°C.

#### Analysis of Derived Amino Acid Sequences

Homology searches were carried out at the National Center for Biotechnology Information by using the BLAST network service (Altschul et al., 1994). Protein alignments and consensus determinations were performed by using the PILEUP program of the Genetics Computer Group software package (Devereux et al., 1984), with minimal manual modification.

#### Acknowledgments

Correspondence and requests for reprints should be addressed to J. L. G. We wish to thank the members of the Aarskog Syndrome Support Group and Michael Schneider for their contributions of time and patient materials; Robert Nussbaum for the human YAC library; Jeff Murray for the FCF cDNA library; David Kurnit for the brain cDNA library; Richard Jove, Michael Long, and Diane Robins for critically reviewing the manuscript; and Diane Voss for assistance in preparing this manuscript. The work was supported in part by the General Clinical Research Center at the University of Michigan and was funded by National Center for Research Resources grant RR-00042 and by March of Dimes-Birth Defects Foundation basic science grant 1-93-0326 and National Institutes of Health grant NS-30771 to J. L. G.

Received August 9, 1994; revised September 1, 1994.

#### References

Aarskog, D. (1970). A familial syndrome of short stature associated with facial dysplasia and genital anomalies. *J. Pediatr.* **77**, 856-861.

Abo, A., Pick, E., Hall, A., Totty, N., Teahan, C., and Segal, A. (1991). Activation of the NADPH oxidase involves the small GTP-binding protein p21<sup>rac1</sup>. *Nature* **353**, 668-670.

Altschul, S. F., Boguski, M. S., Gish, W., and Wootton, J. C. (1994). Issues in searching molecular sequence databases. *Nature Genet.* **6**, 119-129.

Arveiler, B., and Porteous, D. J. (1991). Amplification of end fragments of YAC recombinants by inverse-polymerase chain reaction. *Technique* **3**, 24-28.

Bawle, E., Tyrkus, M., Lipman, S., and Bozomowski, D. (1984). Aarskog syndrome: full male and female expression associated with an X-autosome translocation. *Am. J. Med. Genet.* **17**, 595-602.

Bender, A., and Pringle, J. R. (1989). Multicopy suppression of the *cdc24* budding defect in yeast by *CDC42* and three newly identified genes including the *ras*-related gene *RSR1*. *Proc. Natl. Acad. Sci. USA* **86**, 9976-9980.

Boguski, M. S., and McCormick, F. (1993). Proteins regulating Ras and its relatives. *Nature* **366**, 643-653.

Bonner, T. I., Buckley, N. J., Young, A. C., and Brann, M. R. (1987). Identification of a family of muscarinic acetylcholine receptor genes. *Science* **237**, 527-531.

Bowtell, D., Fu, P., Simon, M., and Senior, P. (1992). Identification of murine homologues of the *Drosophila son of sevenless* gene: potential activators of ras. *Proc. Natl. Acad. Sci. USA* **89**, 6511-6515.

Buchberg, A. M., Cleveland, L. S., Jenkins, N. A., and Copeland, N. G. (1990). Sequence homology shared by neurofibromatosis type-1 gene and *IRA* and *IRA2* negative regulators of the Ras cyclic AMP pathway. *Nature* **347**, 291-294.

Buday, L., and Downward, J. (1993). Epidermal growth factor regulates p21<sup>ras</sup> through the formation of a complex of receptor, Grb2 adapter protein, and Sos nucleotide exchange factor. *Cell* **73**, 611-620.

Casanova, J. L., Pannetier, C., Jaulin, C., and Kourilsky, P. (1990). Optimal conditions for directly sequencing double stranded PCR products with sequenase. *Nucl. Acids Res.* **18**, 4028.

Cicchetti, P., Mayer, B. J., Thiel, G., and Baltimore, D. (1992). Identification of a protein that binds to the SH3 region of Abl and is similar

to Bcr and GAP-rho. *Science* **257**, 803-806.

Coppola, J., Bryant, S., Koda, T., Conway, D., and Barbacid, M. (1991). Mechanism of activation of the *Vav* protooncogene. *Cell Growth Differ.* **2**, 95-105.

Devereux, J., Haeblerli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucl. Acids Res.* **12**, 387-395.

Downward, J. (1992). Rac and Rho in tune. *Nature* **359**, 273-274.

Drubin, D. G. (1991). Development of cell polarity in budding yeast. *Cell* **65**, 1093-1096.

Duncan, P. A., Klein, R. M., Wilmot, P. L., and Shapiro, L. R. (1977). Additional features of the Aarskog syndrome. *J. Pediatr.* **91**, 769-770.

Egan, S. E., Giddings, B. W., Brooks, M. W., Buday, L., Sizeland, A. M., and Weinberg, R. A. (1993). Association of Sos Ras exchange protein with Grb2 is implicated in tyrosine kinase signal transduction and transformation. *Nature* **363**, 45-51.

The European Chromosome 16 Tuberous Sclerosis Consortium (1993). Identification and characterization of the tuberous sclerosis gene on chromosome 16. *Cell* **75**, 1305-1315.

Fryns, J. P. (1992). Aarskog syndrome: the changing phenotype with age. *Am. J. Med. Genet.* **43**, 420-427.

Glover, T. W., Verga, V., Rafael, J., Gorski, J. L., Bawle, E., and Higinis, J. V. (1993). Translocation breakpoint in Aarskog syndrome maps to Xp11.21 between *ALAS2* and *DXS323*. *Hum. Mol. Genet.* **10**, 1717-1718.

Gorlin, R. J., Cohen, M. M., and Levin, L. S. (1990). *Syndromes of the Head and Neck, Third Edition* (New York: Oxford University Press).

Gorski, J. L., Stein, C. K., and Glover, T. W. (1989). A somatic cell hybrid panel to facilitate identification of DNA sequences in the vicinity of the incontinentia pigmenti locus (*IP1*). *Cytogenet. Cell Genet.* **52**, 90-92.

Gorski, J. L., Boehnke, M., Reyner, E. L., and Burright, E. N. (1992). A radiation hybrid map of the proximal short arm of the human X chromosome spanning incontinentia pigmenti 1 (*IP1*) translocation breakpoints. *Genomics* **14**, 657-665.

Grier, R. E., Farrington, F. H., Kondig, R., and Mamunes, P. (1983). Autosomal dominant inheritance of the Aarskog syndrome. *Am. J. Med. Genet.* **15**, 39-46.

Gulbins, E., Coggeshall, K. M., Baier, G., Katzav, S., Burn, P., and Altman, A. (1993). Tyrosine kinase-stimulated guanine nucleotide exchange activity of Vav in T cell activation. *Science* **260**, 822-825.

Hart, M. J., Eva, A., Evans, T., Aaronson, S. A., and Cerione, R. A. (1991). Catalysis of guanine nucleotide exchange on the CDC42Hs protein by the *dbl* oncogene product. *Nature* **354**, 311-314.

Hart, M. J., Eva, A., Zangrill, D., Aaronson, S. A., Evans, T., Cerione, R. A., and Zheng, Y. (1994). Cellular transformation and guanine nucleotide exchange activity are catalyzed by a common domain on the *dbl* oncogene product. *J. Biol. Chem.* **269**, 62-65.

Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucl. Acids Res.* **15**, 8125-8148.

Lafreniere, R. G., Brown, C. J., Powers, V. E., Carrel, L., Davies, K. E., Barker, D. F., and Willard, H. F. (1991). Physical mapping of 60 DNA markers in the p21.1-q21.3 region of the human X chromosome. *Genomics* **11**, 352-363.

Lee, J. T., Murgia, A., Sosnoski, D. M., Olivos, I. M., and Nussbaum, R. L. (1992). Construction and characterization of a yeast artificial chromosome library for Xpter-Xq27.3: a systematic determination of cloning rate and X-chromosome representation. *Genomics* **12**, 526-533.

Li, N., Batzer, A., Daly, R., Yajnik, V., Skolnik, E., Chardin, P., Bar-Sagi, D., Margolls, B., and Schlessinger, J. (1993). Guanine-nucleotide-releasing factor hSos1 binds to Grb2 and links receptor tyrosine kinases to Ras signaling. *Nature* **363**, 85-88.

Lifshitz, B., Fainstein, E., Marcelle, C., Shtivelman, E., Amson, R., Gale, R. P., and Canaani, E. (1988). *Bcr* genes and transcripts. *Oncogene* **2**, 113-117.

Liscovitch, M., and Cantley, L. C. (1994). Lipid second messengers. *Cell* 77, 329–334.

Melnick, M., and Shields, E. D. (1976). Aarskog syndrome: new oral-facial findings. *Clin. Genet.* 9, 20–24.

Miki, T., Smith, C. L., Long, J. E., Eva, A., and Fleming, T. P. (1993). Oncogene *ect2* is related to regulators of small GTP-binding proteins. *Nature* 362, 462–465.

Musacchio, A., Gibson, T., Rice, P., Thompson, J., and Saraste, M. (1993). The PH domain: a common piece in the structural patchwork of signalling proteins. *Trends Biochem. Sci.* 18, 343–348.

Orita, M., Suzuki, Y., Sekiya, T., and Hayashi, K. (1989). Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* 5, 874–879.

Pasteris, N. G., Bialecki, M. D., and Gorski, J. L. (1993). YAC subclone contig assembly by serial interspersed repetitive sequence (IRS)-PCR product hybridizations. *Nucl. Acids Res.* 21, 5275–5276.

Porteous, M. E. M., Curtis, A., Lindsay, S., Williams, O., Goudie, D., Kamakari, S., and Bhattacharya, S. S. (1992). The gene for Aarskog syndrome is located between *DXS255* and *DXS566* (Xp11.2-Xq13). *Genomics* 14, 298–301.

Quest, A. F. G., Bardes, E. S. G., and Bell, R. M. (1994). A phorbol ester binding domain of protein kinase C $\gamma$ : deletion analysis of the Cys2 domain defines a minimal 43-amino acid peptide. *J. Biol. Chem.* 269, 2961–2970.

Ren, R., Mayer, B. J., Cicchetti, P., and Baltimore, D. (1993). Identification of a ten-amino acid proline-rich SH3 binding site. *Science* 259, 1157–1161.

Ridley, A. J., and Hall, A. (1992). The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors. *Cell* 70, 389–399.

Ridley, A. J., Paterson, H. F., Johnston, C. L., Diekmann, D., and Hall, A. (1992). The small GTP-binding protein rac regulates growth factor-induced membrane ruffling. *Cell* 70, 401–410.

Ron, D., Zannini, M., Lewis, M., Wickner, R. B., Hunt, L. T., Graziani, G., Tronick, S. R., Aaronson, S. A., and Eva, A. (1991). A region of proto-*dbl* essential for its transforming activity shows sequence similarity to a yeast cell cycle gene, *CDC24*, and the human breakpoint cluster gene, *bcr*. *New Biol.* 3, 372–379.

Rozakis-Adcock, M., Fernley, R., Wade, J., Pawson, T., and Bowtell, D. (1993). The SH2 and SH3 domains of mammalian Grb2 couple the EGF receptor to the Ras activator mSos1. *Nature* 363, 83–85.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual, Second Edition* (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Scott, C. I., Jr. (1971). Unusual facies, joint hypermobility, genital anomaly and short stature: a new dysmorphic syndrome. In *The Clinical Delineation of Birth Defects, Volume 10: The Endocrine System* (Baltimore: Williams and Wilkins), pp. 240–246.

Seabra, M. C., Brown, M. S., and Goldstein, J. L. (1993). Retinal degeneration in choroideremia: deficiency of Rab geranylgeranyl transferase. *Science* 259, 377–381.

Shinjo, K., Koland, J. G., Hart, M. J., Narasimhan, V., Johnson, D. I., Evans, T., and Cerione, R. A. (1990). Molecular cloning of the gene for the human placental GTP-binding protein Gp (G25K): identification of this GTP-binding protein as the human homolog of the yeast cell-division-cycle protein CDC42. *Proc. Natl. Acad. Sci. USA* 87, 9853–9857.

Shou, C., Farusworth, C. L., Neel, B. G., and Feig, L. A. (1992). Molecular cloning of cDNAs encoding a guanine nucleotide releasing factor for Ras p21. *Nature* 358, 351–354.

Stevenson, R. E., May, M., Arena, J. F., Millar, E. A., Scott, C. I., Jr., Schroer, R. J., Simensen, R. J., Lubs, H. A., and Schwartz, C. E. (1994). Aarskog-Scott syndrome: confirmation of linkage to the pericentric region of the X chromosome. *Am. J. Med. Genet.*, in press.

Tommerup, N. (1993). Mendelian cytogenetics: chromosome rearrangements associated with Mendelian disorders. *J. Med. Genet.* 30, 713–727.

van de Vooren, M. J., Niermeijer, M. F., and Hoogeboom, A. J. M. (1983). The Aarskog syndrome in a large family, suggestive for autosomal

dominant inheritance. *Clin. Genet.* 24, 439–445.

Weisman, L. S., and Wickner, W. (1992). Molecular characterization of *VAC1*, a gene required for vacuole inheritance and vacuole protein sorting. *J. Biol. Chem.* 267, 618–623.

Woychik, R. P., Maas, R. L., Zeller, R., Vogt, T. F., and Leder, P. (1990). Formins: proteins deduced from the alternative transcripts of the limb deformity gene. *Nature* 346, 850–853.

#### GenBank Accession Number

The accession number for the sequence reported in this paper is U11690.