Am. J. Hum. Genet. 65:370–386, 1999

Chromosome Breakage in the Prader-Willi and Angelman Syndromes Involves Recombination between Large, Transcribed Repeats at Proximal and Distal Breakpoints

James M. Amos-Landgraf,[1,*] Yonggang Ji,[1,*] Wayne Gottlieb,[2,†] Theresa Depinet,[1] Amy E. Wandstrat,[1,‡] Suzanne B. Cassidy,[1] Daniel J. Driscoll,[3] Peter K. Rogan,[4] Stuart Schwartz,[1] and Robert D. Nicholls[1]

[1]Department of Genetics, Case Western Reserve University School of Medicine, and Center for Human Genetics, University Hospitals of Cleveland, Cleveland; [2]Department of Neuroscience and [3]R.C. Philips Unit, Division of Genetics, Department of Pediatrics, University of Florida College of Medicine, Gainesville; [4]Section of Medical Genetics & Molecular Medicine, Children's Mercy Hospital, Kansas City, MO

Summary

Prader-Willi syndrome (PWS) and Angelman syndrome (AS) are distinct neurobehavioral disorders that most often arise from a 4-Mb deletion of chromosome 15q11-q13 during paternal or maternal gametogenesis, respectively. At a de novo frequency of ~.67–1/10,000 births, these deletions represent a common structural chromosome change in the human genome. To elucidate the mechanism underlying these events, we characterized the regions that contain two proximal breakpoint clusters and a distal cluster. Novel DNA sequences potentially associated with the breakpoints were positionally cloned from YACs within or near these regions. Analyses of rodent-human somatic-cell hybrids, YAC contigs, and FISH of normal or rearranged chromosomes 15 identified duplicated sequences (the *END* repeats) at or near the breakpoints. The *END*-repeat units are derived from large genomic duplications of a novel gene (*HERC2*), many copies of which are transcriptionally active in germline tissues. One of five PWS/AS patients analyzed to date has an identifiable, rearranged *HERC2* transcript derived from the deletion event. We postulate that the *END* repeats flanking 15q11-q13 mediate homologous recombination resulting in deletion. Furthermore, we propose that active transcription of these repeats in male and female germ cells may facilitate the homologous recombination process.

Received March 16, 1999; accepted for publication June 2, 1999; electronically published July 13, 1999.

Address for correspondence and reprints: Dr. Robert D. Nicholls, Department of Genetics, Case Western Reserve University, 10900 Euclid Ave., Cleveland, OH 44106-4955. E-mail: rxn19@po.cwru.edu

*These authors contributed equally to this work.

Present affiliations: †DeWitt Middle School, Ithaca, NY; ‡Center for Immunology, University of Texas Southwestern Medical Center, Dallas.

0002-9297/99/6502-0013$02.00

Introduction

Prader-Willi (PWS; MIM 176270) and Angelman (AS; MIM 105830) syndromes are clinically distinct neurobehavioral disorders, each occurring at a frequency of ~1/10,000-20,000 births (Clayton-Smith and Pembrey 1992; Cassidy 1997). These syndromes have related genetic origins; the nature of the disorder depends on the sex of the parent of origin of genetic abnormalities in the chromosome region 15q11-q13 (Nicholls et al. 1998). The genetic aberrations include 4-Mb 15q11-q13 deletions in 70%–80% of PWS and AS patients (Knoll et al. 1990; Kuwano et al. 1992; Zackowski et al. 1993; Christian et al. 1995), uniparental disomy (UPD) in 25% of PWS and 2% of AS patients (Nicholls et al. 1998), imprinting mutations in 2%–5% of PWS and AS sporadic or familial cases (Nicholls et al. 1998), and *UBE3A* gene mutations in ~5% of AS cases (Malzac et al. 1998). Since ~75% of PWS and AS patients have a deletion, of which ⩾95% have common deletion extents (Knoll et al. 1990; Kuwano et al. 1992; Christian et al. 1995; present report), the de novo frequency of 15q11-q13 deletions is ~.67–1/10,000 births, making this one of the most common recurring structural chromosomal changes in birth defects. Despite its frequent occurrence, the mechanism for a large common deletion in these syndromes is not yet understood at the molecular level.

The chromosome 15q11-q13 region is meiotically unstable, with an unusual variety of cytogenetic rearrangements, including the AS and PWS deletions, frequent inverted duplication (inv dup[15]) chromosomes (Robinson et al. 1993*a*; Huang et al. 1997; Wandstrat et al. 1998), rare duplications and triplications (Clayton-Smith et al. 1993*b*; Schinzel et al. 1994; Cassidy et al. 1996; Browne et al. 1997; Repetto et al. 1998), inversions (Clayton-Smith et al. 1993*a*), and balanced or unbalanced translocations (Butler 1990; Sun et al. 1996). The PWS/AS deletions have clustered breakpoints, with two proximal clusters and a single distal breakpoint re-

gion (Knoll et al. 1990; Kuwano et al. 1992; Christian et al. 1995). Inv dup(15) with breakpoints at 15q11.2 or 15q13 accounts for ~50% of observed small supernumerary marker chromosomes, representing an overall frequency of ~1/3,500–1/9,000 (Robinson et al. 1993*a*; Cheng et al. 1994; Webb et al. 1994; Crolla et al. 1995; Huang et al. 1997; Wandstrat et al. 1998). At least three regions in proximal 15q have been implicated as predisposing to these rearrangements, including one proximal and the distal PWS and AS breakpoints (Repetto et al. 1998), as well as a more distal breakpoint in large inv dup(15) chromosomes (Huang et al. 1997; Wandstrat et al. 1998).

There is evidence that region-specific, low-copy repeats occur within 15q11-q13. Such duplicated segments of the genome have been termed "duplicons," to distinguish them from repeats that are interspersed at low or high frequency throughout the genome (Eichler 1998). Buiting et al. (1992) isolated a microdissection clone (*D15F37;* MN7) from 15q11-q13 and presented data to suggest that four homologous loci were interspersed through 15q11-q13, both within and outside the critical deletion region, with related copies in chromosome 16p11.2. More recently, three of these *D15F37* sequences were localized at or near the proximal breakpoint cluster II and the distal deletion breakpoint region (Buiting et al. 1998). In the current work, positional cloning was used to isolate novel duplicons, termed the "*END* repeats," for which we localize three to five copies to each of the proximal and distal PWS/AS deletion breakpoint regions in 15q11-q13. The *END* repeats include the MN7 sequence and represent large genomic duplications of a novel unique gene, *HERC2* (Ji et al. 1999), with many actively transcribed copies. We suggest that the commonly occurring de novo deletions in chromosome 15q11-q13 involve misalignment and recombination within homologous *END*-repeat units, and we propose that active transcription in germ cells facilitates this process.

Material and Methods

Generation of Alu-Alu *and* Alu-*YAC End Microclones*

Alu-Alu and *Alu*-YAC–end PCR products were generated by use of various combinations of the following primers: *Alu* left (5′-GGAGGGCCCAAAGTGCTGGGATTACAGG-3′), *Alu* right (5′-GAGGATCCACTGCACTCCAGCCTGGG-3′), YAC left (5′-AAGTACTCTCGGTAGCCAAG-3′), and YAC right (5′-AGTCGAACGCCCGATCTCAA-3′) (Breukel et al. 1990). PCR was performed by 35 cycles of 1 min denaturation at 95°C, 1.5 min annealing at 65°C, and 5 min extension times at 72°C, with 0.5 units of *Taq* polymerase, 300 μM of each dNTP, 1 μM of each primer (0.1 μM of each *Alu* primer in YAC-*Alu* PCR), and 100 ng yeast DNA/50 μl reaction. PCR products were cloned into the pCR vector (Invitrogen).

Sequence Analysis

Sequencing of clone inserts was performed by use of Sequenase 2.0 (USB) and ^{35}S α-dATP (Amersham), and reactions were run on a 6% polyacrylamide gel. Sequences were analyzed by BLAST and MegAlign (DNASTAR).

PCR

Microclone PCRs were performed on TA-cloned inserts and YAC DNA (10 ng) by use of each of the following primer sets: 254RL2, RN202 (5′-CAAAGTCCTGGTTCAAATGC-3′) + RN203 (5′-CCGCGCAGAGCTGCCAC-3′), 158-bp fragment; 318RL3, RN191 (5′-CTTGTCTTCTACAGATAC-3′) + RN200 (5′-CAGCGATTTTGTTTATTTATATTCCC-3′), 127-bp fragment; 93RR2, RN188 (5′-GTATGCTGGAACCCTCAC-3′) + RN214 (5′-GAGTACTTGTTAGAGGTG-3′), 213-bp fragment; and λ11A1 CpG island, RN319 (5′-GCCAAGTCACAATGTCATCC-3′) + M13 reverse, 329-bp fragment amplified from a *Not*I-*Eco*RI subclone of λ11A1, including 247 bp of sequence 99% homologous to *HERC2* intron 1. Primers for polymorphic chromosome 15 markers were obtained from Genome Database and Research Genetics. PCR was performed either as specified elsewhere (Genome Database and Research Genetics) or in 50-μl reactions with 250 μM of each dNTP; 2.0 mM $MgCl_2$; 1 μM of each primer; and 1 U Ampli*Taq* (Perkin Elmer Cetus)—at 95°C for 5 min, 30 cycles of 94°C for 30 s, 55°C for 30 sec and 72°C for 1 min; and a 72°C 10 min final extension. Products were analyzed on 3% 3:1 NuSieve agarose gels (FMC Bioproducts).

Isolation of Genomic Clones and YACs

Bacteriophage λ clones were identified from a human male genomic library (Clontech). Phage were plated on 20 × 150-mm petri dishes at a density of 50,000 plaques/plate and were transferred to Hybond nylon membrane (Amersham). Membranes were hybridized by use of the 254RL2, 318RL3, and 93RR2 PCR products as probes, which were labeled by random priming (Boehringer Mannheim) with α[^{32}P]-dCTP (Amersham). *Eco*RI and *Eco*RI/*Not*I fragments were subcloned from two phage, λ6A1 and λ11A1. YACs 254B5, 318A1, and 93C9 were received from D. H. Ledbetter and A. C. Chinault of the Baylor College of Medicine, under the auspices of the Human Genome Project (Kuwano et al. 1992). Additional YACs were identified from the Whitehead database and were acquired through Research Ge-

Table 1

Deletion Extent in PWS and AS Cases, by Microsatellite Mapping

	Proximal Marker[a]			Distal Marker[a]			
Patient	*D15S541*	*D15S1035*	*D15S543*	*D15S1002*	*D15S1048*	*D15S1019*	*D15S165*
PWS-H1	+? [b]	+	−PAT	−PAT	−PAT	+	+
PWS-S1	UI [b]	+	−PAT	UI	+	+	UI
PWS-S2	−PAT	ND	−PAT	−PAT	+	UI	+
PWS-F1	UI [b]	UI	−PAT	−PAT	UI	+	+
PWS-D1	+[b]	+	UI	ND	+	+	UI
PWS-B1	+	+?	UI	UI	+	UI	+
PWS-W1	−PAT	−PAT	UI	−PAT	+	+	+
PWS-M1	−PAT	ND	−PAT	UI	+	UI	UI
PWS-E1	+[b]	+	UI	UI	+	+	+
PWS-G1	UI	UI	UI	UI	UI	UI	+
PWS-K1	+[b]	+	UI	UI	+	UI	+
PWS-N1	UI	+?	UI	ND	ND	ND	ND
PWS-K2	+	+	UI	−PAT	+	+	+
PWS-T1	−PAT	UI	UI	−PAT	+	UI	UI
PWS-C1	−PAT	−PAT	−PAT	UI	+	+	+
AS-B1	+	UI	UI	−MAT	+	+	+
AS-F1	+	UI	UI	UI	+	+	UI
AS-H1	−MAT	−MAT	UI	−MAT	UI	UI	+
AS-K1	−MAT	−MAT	UI	UI	UI	+	UI
AS-M1	UI	UI	−MAT	UI	+	+	UI
AS140	−MAT	−MAT	−MAT	−MAT	+	+	+
AS145	+	+	UI	UI	+	UI	+
AS128	−MAT	−MAT	UI	−MAT	UI	+	+
AS144	−MAT	−MAT	UI	−MAT	UI	UI	+
AS160	UI	+	UI	−MAT	+	+	+
AS123	+	+	UI	−MAT	+	+	UI
AS154	UI	UI	UI	−MAT	+	UI	+
AS159	+	+	UI	UI	+	+	+
AS130	UI	+	−MAT	−MAT	+	+	+
AS184	+	UI	−MAT	UI	+	+	UI
AS138	UI	−MAT	UI	UI	+	+	+
AS165	+	+	−MAT	−MAT	+	UI	+
AS109	UI	UI	−MAT	−MAT	+	UI	+
AS151	UI	−MAT	UI	UI	UI	UI	+
AS135	+	+	UI	−MAT	−MAT	UI	+
AS131	UI	+	+	−MAT	+	UI	+
AS148	+	ND	−MAT	−MAT	+	+	UI
AS121	ND	ND	−MAT	−MAT	+	UI	UI

[a] + = intact; − = deleted; PAT = paternal; MAT = maternal; UI = uninformative; ND = not determined; ? = uncertain result.

[b] Intact (+) at *D15S542*, which is located only 1 kb from *D15S541* (see Christian et al. 1995).

netics. All YACs were grown in AHC media, and genomic DNA was isolated by standard methods.

Southern Blotting

Human, somatic-cell hybrid, or YAC DNA samples were digested with restriction enzymes and were electrophoresed in a .7% agarose gel with 1 × Tris-acetate EDTA buffer (40 mM Tris-HCl [pH 8.5], 40 mM sodium acetate, 2 mM EDTA) at 40 V for 16–20 h. DNA was transferred to nylon membranes (Genescreen; NEN) for hybridization by standard techniques. The "exon" and 1.1-kb *HERC2* cDNA probes were generated as described elsewhere (Ji et al. 1999), and other probes were generated as described above. Final wash conditions were high stringency (.1% SDS, .1% SSC) for 30–60 min at 56–65°C, depending on GpC and the repeat-element content of the probe. Preassociation (2 h) with an excess of sonicated human placental DNA (200 μg/ml) and pYAC4 DNA (5 μg/ml) was performed for probes containing repetitive sequences.

Cell Lines

Lymphoblastoid cell lines from patients were acquired from the Human Genetic Mutant Cell Repository of the National Institute of General Medical Sciences (Coriell): GM11404 (AS) and GM11385, GM11382, GM09189,

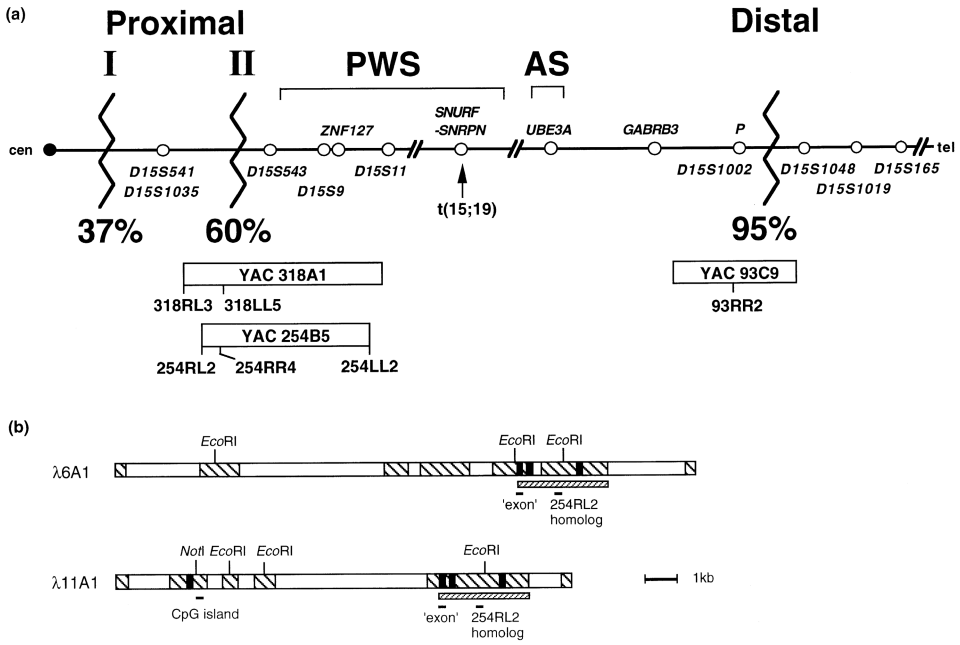

Figure 1 Deletion classes in chromosome 15q11-q13 in patients with PWS or AS, and isolation of microclones from the vicinity of the breakpoint regions. *A,* Map of chromosome 15q11-q13, showing the position of deletion breakpoints in PWS and AS (zigzag lines), along with the frequency of each deletion class determined in this study, genes and relevant markers in 15q11-q13 (*circles*), and a balanced translocation breakpoint within the *SNURF-SNRPN* gene. Below is shown the location of YACs and the microclones isolated from these YACs. *B,* Phage maps for λ6A1 and λ11A1, showing restriction sites, regions sequenced (*hatched*), and sequences homologous to *HERC2* exons (*blackened boxes*). Shown below are a region of 90% sequence similarity, over 2.6 kb, between the two λ clones (*hatched bar*), and the locations of sequences paralagous to the original 254RL2 microclone and other probes (CpG island and "exon") used in this study.

GM09133, GM09024, and GM04297 (PWS). Other Epstein-Barr virus–transformed lymphoblastoid cell lines (AS139 and PWS109) were established by standard techniques, or DNA was prepared directly from peripheral-blood leukocytes. Chromosome 15 somatic-cell–hybrid cell lines were as described elsewhere (Gabriel et al. 1998). Additional somatic-cell hybrids were generated from a fibroblast cell line established from a PWS patient with a balanced t(15;19)(q12;q13.41) (Sun et al. 1996). Fusions of the patient cell line with A9 mouse-fibroblast cell line were performed with polyethylene glycol by *HPRT* and *TK* selection, which selects for retention of the human X chromosome and against mouse-only cells (Willard and Holmes 1984), with random maintenance of other human chromosomes in different clones. Positive clones were identified by PCR with two chromosome 15–specific markers, *D15S11* and *D15S165*, which lie proximal and distal, respectively, to the breakpoint within the *SNURF-SNRPN* gene (Sun et al. 1996; Gray et al. 1999). Two clones were identified, one that contained the der(15), which is of paternal origin (2-3-4), and one that contained the normal maternal chromosome 15 (2-4-1). The nature of the chromosome 15 in both somatic-cell–hybrid lines was verified by FISH analysis using the *D15S11* and *PML* probes (see below), and by DNA-methylation analysis (Gabriel et al. 1998).

FISH

FISH was performed with bacteriophage λ and/or commercially available cosmid probes (Oncor) on unstained slides, by standard methods with minor modifications (Sullivan et al. 1996). At least 10 metaphases—and in most cases, 20 metaphases—were analyzed for the presence of the probe on both the normal chromosome and the rearranged chromosome. Digital images were captured with a Zeiss epifluorescence microscope equipped with a cooled CCD camera (Photometrics CH250) controlled by an Apple Macintosh computer. Gray-scale source images were captured separately with DAPI, fluorescein, and rhodamine filter sets, and were merged and pseudocolored with Gene Join software (Yale University). Detection of biotin-labeled probes was with fluorescein-labeled avidin or Texas-red avidin, and detection of digoxygenin-labeled probes was with fluorescein-labeled anti-digoxygenin or rhodamine-labeled anti-digoxygenin.

Northern Blot Analysis of mRNA Expression

A human multitissue northern blot filter, including ovary and testis, was hybridized by use of ExpressHyb solution under conditions recommended by the manufacturer (Clontech). RNA was extracted from lymphoblastoid cell lines with TRIzol (Gibco BRL). Ten micrograms of total RNA from each sample was run on a 0.7% agarose gel with 1 × MOPS buffer with 2.2 M formaldehyde. The gel was run for 5 h at 60 V, was stained with acridine orange, and was blotted in 20 × SSC as in Southern blots.

Results

Refinement of the PWS/AS Deletion Breakpoint Regions

To refine the locations of deletion endpoints in patients with PWS or AS, we typed eight polymorphic microsatellite markers mapped in the vicinity of these regions in 15 patients with the PWS deletion and 23 patients with the AS deletion and in their parents (table 1; fig. 1*a*). No differences in the deletion breakpoint locations were observed between maternally or paternally derived deletions in AS or PWS, respectively. In this series, 12 (37.5%) of 32 informative PWS/AS cases have deletions breaking between the four proximal markers and the centromere (class I deletions). Only one informative patient (AS131) was intact for *D15S543* (table 1), but, as in the majority of AS patients (Zackowski et al. 1993), this patient was deleted for *ZNF127/D15S9* (data not shown), and these two markers are located <275 kb apart (Christian et al. 1998; Jong et al. 1999). In contrast, 14 (93%) of 15 informative cases are deleted for *D15S543*, suggesting that the majority of the 23 uninformative cases are probably also deleted for this marker (similar results were found by Christian et al. [1995]). Seven cases showed deletion of *D15S543* but not of the group of three inseparable proximal markers (*D15S541/D15S542/D15S1035*), with 12 additional PWS and AS cases intact for at least one of the latter markers but uninformative at *D15S543* (one case deleted for *D15S543* was uninformative at proximal markers, and another was not tested). These data suggest that 19 (60%) of 32 PWS and AS deletions have class II deletions (fig. 1*a*).

At the distal end of the deletions (table 1 and fig. 1*a*), all 21 informative cases were deleted for *D15S1002* (none of the other 15 patients were heterozygous). In contrast, only 2 (7%) of 30 informative cases were deleted for *D15S1048*, and none were deleted for *D15S1019* (in 21 informative cases) or *D15S165* (in 26 informative cases). One of the patients deleted for *D15S1048*, PWS-H1, was heterozygous at *D15S1019*, indicating that the latter marker maps telomeric of *D15S1048*, whereas Buiting et al. (1998) placed these markers in the opposite orientation, on the basis of YAC mapping; this discrepancy likely reflects rearrangement in one of the DNA sources. In summary, our data suggest a single major PWS/AS breakpoint cluster in the distal region of 15q11-q13, between *D15S1002* and *D15S1048* (fig. 1*a*).

Identification of Duplicated Sequences in the 15q11.2 and 15q13 Regions

YACs that were described (Kuwano et al. 1992) as potentially spanning the deletion breakpoints by FISH analysis—254B5 and 318A1 (proximal) and 93C9 (distal)—were utilized for the isolation of *Alu-Alu* and *Alu*-YAC–end PCR products. Of 53 isolated microclones, 24 appeared to be independent, on the basis of insert size, DNA sequence, and/or competitive-hybridization results. Four microclones originating from each of the three YACs were chosen at random and were sequenced. The size of each of the 12 clones was 190–1200 nucleotides. Two clones contained only tandem *Alu* sequences; several contained L1 sequences, but the majority contained unique sequences according to BLAST searches of databases (table 2). Both YAC ends of 254B5 were isolated (254RL2 and 254LL2; a 254B5 YAC right end is also listed in the work of Christian et al. [1998]), as was the YAC right end of 318A1 (318RL3).

Each microclone containing a putative unique sequence was then subjected to qualitative-dosage Southern hybridization analysis under competitive-hybridization conditions, to block *Alu*-repetitive and other repetitive sequences (fig. 2). As expected on the basis of prior polymorphism analyses (Zackowski et al. 1993),

Table 2

Characteristics of *Alu-Alu* and *Alu*-YAC Microclones

Name	Ends[a]	Size (bp)	GenBank Accession Number	Other Features
254RL2	Yr, Al	250	AF140512	Ll
254RR4	Ar, Yr(i)	650	…	…
254LL2	Al, Yl	1065	AF140513	Ll
254RR2	Ar, Ar	650	…	Ll
318RL3	Yr, Al	370	AF140514	Ll
318LL5	A1, A1	270	…	…
318RR5	Ar, Ar	500	…	Ll
318LL1	Yl(i), Al	190	…	A only
93RR2	Ar, Ar	400	AF140515	Ll
93LL5	Al, Al	250	…	…
93RL3	Al, Al	800	…	Ll
93RR6	Ar, Yr(i)	270	…	…
λ6A1 "254RL2"	…	180	AF140516	93% to 254RL2
λ11A1 "254RL2"	…	185	AF140518	96% to 254RL2
λ6A1 "exon"	…	228	AF140517	91% to *HERC2*
λ11A1 "exon"	…	178	AF140518	88% to *HERC2*
λ11A1 "CpG-island"	…	247	AF140519	99% to *HERC2*

[a]A = Alu; i = illegitimate PCR priming; l = left; r = right; Y = YAC; … = not applicable.

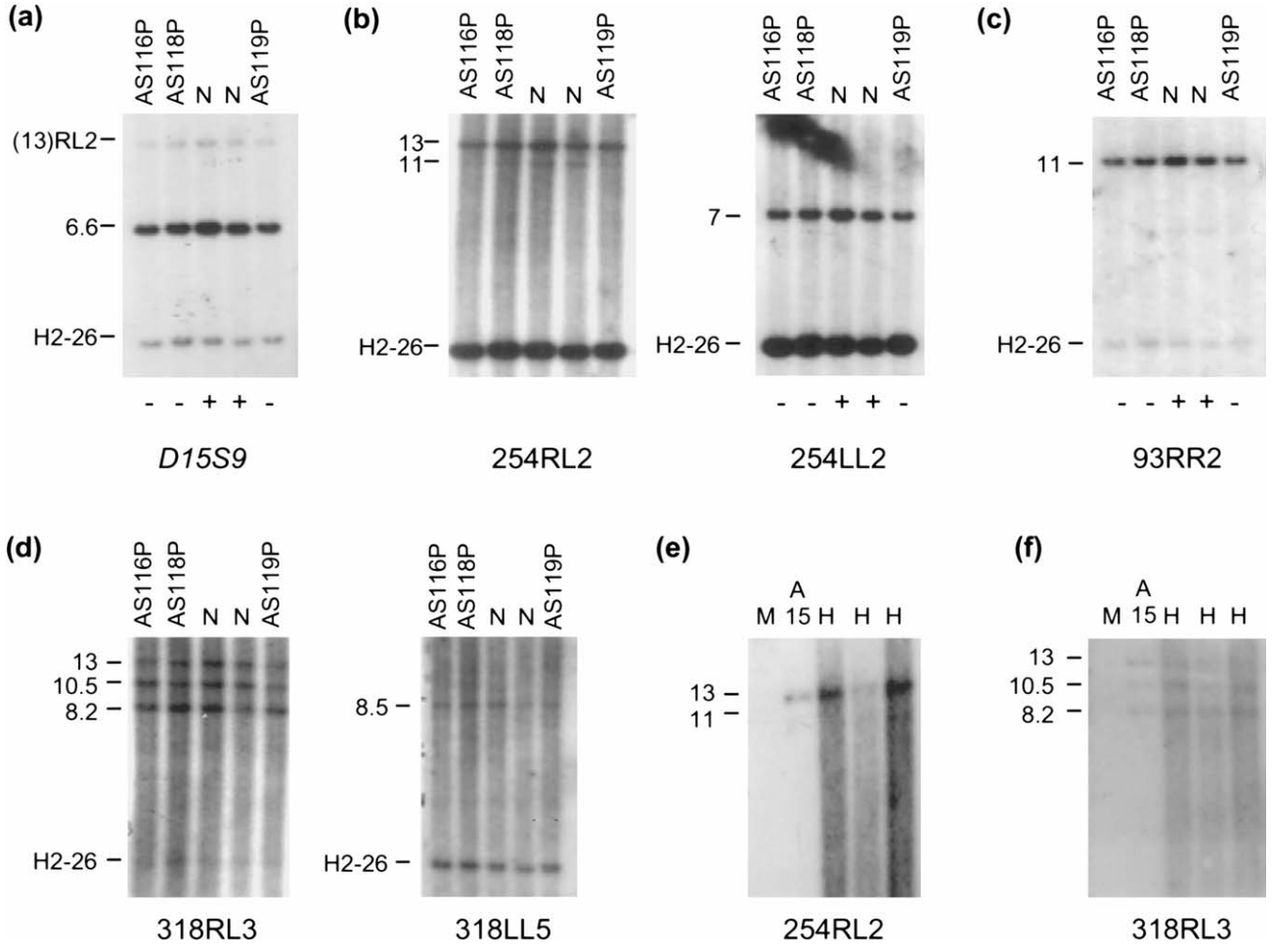

Figure 2 Identification of chromosome 15–specific duplicated sequences. Southern blots of *Hin*dIII-digested DNA from AS-deletion or normal (lanes N) DNA were probed with either (*A*) the control *D15S9* (p34) probe (standard conditions) or (*B–F*) with various *Alu-Alu* and *Alu*-YAC–end microclones as probes under competitive hybridization conditions. In *A–D*, filters are cohybridized with the control probe H2-26 from chromosome 13 (Tantravahi et al. 1989). The 6.6-kb p34 probe is located 0.5 kb centromeric to the *ZNF127*-gene 5′ end (fig. 1*A*; Jong et al. 1999). Qualitative-dosage interpretations for the strongest band(s) detected by the respective microclones are shown (+ = intact; − = deleted in the AS lanes). In *E* and *F*, a somatic-cell hybrid retaining a single human chromosome 15 is shown (lanes A15), along with mouse (lanes M) and normal human (lanes H) DNAs. The p34 probe in *A* was rehybridized to the 254RL2 blot, with residual signal present from the 13-kb 254RL2 band ([13]RL2).

the control *D15S9* probe was deleted in the three AS-deletion samples (fig. 2*a*). Six microclones yielded one to three specific *Hin*dIII bands in genomic DNA by this analysis, although the qualitative patterns differed (fig. 2*b–d*). Single bands produced by probe 254LL2 (fig. 2*b*) and by 93RR2 (fig. 2*c*) each showed results similar to those for *D15S9*, suggesting that they are localized within the AS deletion (fig. 1*a*), which was the expected result for the 254LL2 probe at the 254B5 YAC left end (Jong et al. 1999). In contrast, probes 254RL2, 318RL3, and 318LL5 (fig. 2*b* and 2*d*) show evidence of multiple bands of differing intensity, suggesting that these probes may detect duplicated sequences, which precluded interpretation of dosage. Probes 254RL2 and 318RL3 were therefore hybridized to Southern blots containing DNA from a rodent-human somatic-cell hybrid containing a single human chromosome 15, and each band was localized to chromosome 15 (fig. 2*e* and 2*f*).

These results suggested that several of the microclones, isolated from YACs at or close to the major PWS/AS proximal breakpoint region, contained chromosome-specific duplicated sequences. To generate reagents to further investigate this possibility by cytological techniques, we generated unique STS (sequence-tagged site) probes from several microclones, and we isolated clones positive for 254RL2 from a human genomic λ phage library. Of 21 initial positives, 2—λ11A1 and λ6A1—were chosen for detailed analysis. Partial sequence analysis indicated that these phage contained sequences closely related—but not identical—to each other and to

the original 254RL2 microclone (fig. 1*b* and table 2), confirming their origin from duplicated sequences. Both phage probes were independently used in FISH to normal metaphase chromosomes, which identified two distinct hybridization signals at 15q11-q13 (author's unpublished data). Two-color FISH with λ11A1 and a *D15S11* cosmid, the latter located in the 15q11-q13 PWS region proximal to *SNURF-SNRPN* (Fig. 1*a*), indicated that the two λ11A1 signals flanked the control probe (fig. 3). These data suggest the presence of a duplicated sequence located within proximal and distal 15q11-q13.

Further FISH experiments were performed with phage λ11A1 on metaphase spreads from a patient with the PWS deletion, a normal individual, a patient with a direct duplication of chromosome 15q11-q13 (dup [15]), and a patient with a triplication of chromosome 15q11-q13 (trip [15]) (Cassidy et al. 1996). The latter two abnormal chromosomes were identified by cytogenetic and FISH studies, with *D15S11* and *GABRB3* probes (data not shown). The results for λ11A1 confirm the presence of a duplicated sequence at proximal and distal 15q11-q13, with one, two, three, or four sets of signals for chromosomes containing zero, one, two, or three copies of 15q11-q13, respectively (fig. 4). The intensity of FISH signals produced by λ11A1 suggested that each copy of the "duplicated" region may actually represent several copies of homologous sequences. The isolation of these sequences from YAC clones located at or near the PWS and AS breakpoints in 15q11-q13, coupled with the FISH results, suggests that duplicated sequences are located at or near the *end*points of the PWS/AS deletion in 15q11-q13; therefore, we refer here to the duplicated sequences as the "*END* repeats."

Figure 3 FISH analysis, which localizes the *END* repeats to 15q11.2 and 15q13, flanking the *D15S11* locus. Two-color FISH on B29670 (PWS-deletion) cells, with λ11A1 (yellow) and *D15S11* cosmid (red) probes. The red signal at 15q22 represents the *PML* control probe. Only the normal chromosome 15 is shown, with the λ11A1 probe detecting a duplicated locus that flanks the *D15S11* signal.

Molecular Analysis of the END *Repeats*

The molecular nature and chromosomal location of the *END*-repeat elements was further investigated by Southern blot analysis of genomic DNA from somatic-cell hybrids, YACs, and patients with rearrangements of chromosome 15. Use of a 158-bp 254RL2 probe, free of *Alu* repeats, on human genomic DNA gave a pattern more complex than that given by competitive hybridization, consisting of a major 13-kb *Hin*dIII band and three less intense bands (fig. 5*a*, lanes 1 and 9, and fig. 6). All these bands are localized to chromosome 15, on the basis of analysis of chromosome 15 somatic-cell hybrids (fig. 5*a*, lanes 2–8)—including A15 and A15-1, which contain no other human chromosomes (McDaniel and Schultz 1992). Indeed, A15-1 contains only 15pter-q14, which, combined with the FISH data (see above), suggests that all copies of 254RL2 are sublocalized within 15q11-q13. We generated somatic-cell hybrids retaining either the der(15) (2-3-4) or the intact, maternally derived chromosome 15 (2-4-1), from a patient with a balanced translocation that breaks within the *SNURF-SNRPN* gene in central 15q11-q13 (see fig. 1*a*; Sun et al. 1996). Comparison of the 254RL2-probe hybridization patterns for these and other hybrids (fig. 5*a*) indicates that the 16.5-kb band (and perhaps the 11-kb band) is likely to be localized uniquely to proximal 15q11.2 in the der(15) (compare lane 3 with lane 2), the 15-kb band uniquely to 15q13 as not present in the der(15) hybrid (compare lane 3 and lanes 2 and 4-8), and the 13-kb band to both 15q11.2 and 15q13, since the latter is of reduced intensity in the der(15) hybrid. The intensity of the latter band in hybrids and human genomic DNA is also consistent with multiple copies of the 13-kb band, which was confirmed by YAC analysis (see below).

Analyses of YACs localized to proximal 15q11.2 and distal 15q13 by the 254RL2 probe (fig. 5*b*) and by a CpG-island probe from λ11A1 (figs. 1*b* and 5*c*) were complicated by several factors, including the relatively small size of YACs for low-copy repeat analysis and the instability of the *END* repeats within YACs, particularly at distal 15q13 (fig. 5*c*, lane 6 vs. lane 8; author's unpublished data; also see Buiting et al. 1998; Christian et al. 1998). Nevertheless, several important results emerge from these analyses. First, the 13-kb *Hin*dIII

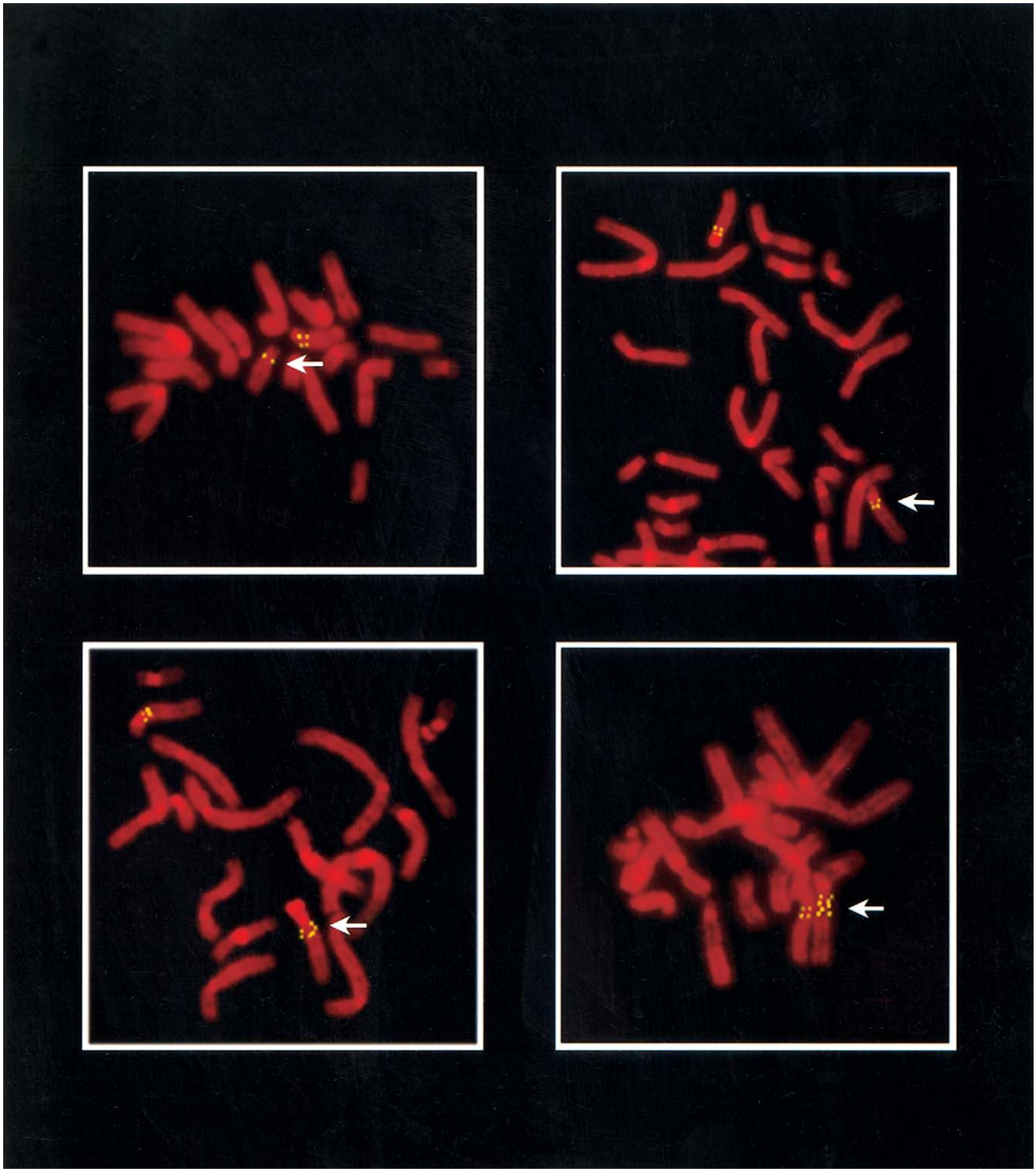

Figure 4 FISH analysis of the *END* repeats, for normal and rearranged chromosomes 15. In all panels, FISH is performed with the single λ11A1 probe. *Top left,* PWS-deletion cells (B29670). *Top right,* Normal cells (B29331). *Bottom left,* Intrachromosomal duplication (15) cells (R1000). *Bottom right,* Intrachromosomal triplication (15) cells (B27965). One to four sets of λ11A1 signals are seen on chromosomes containing zero to three copies of the 15q11-q13 region.

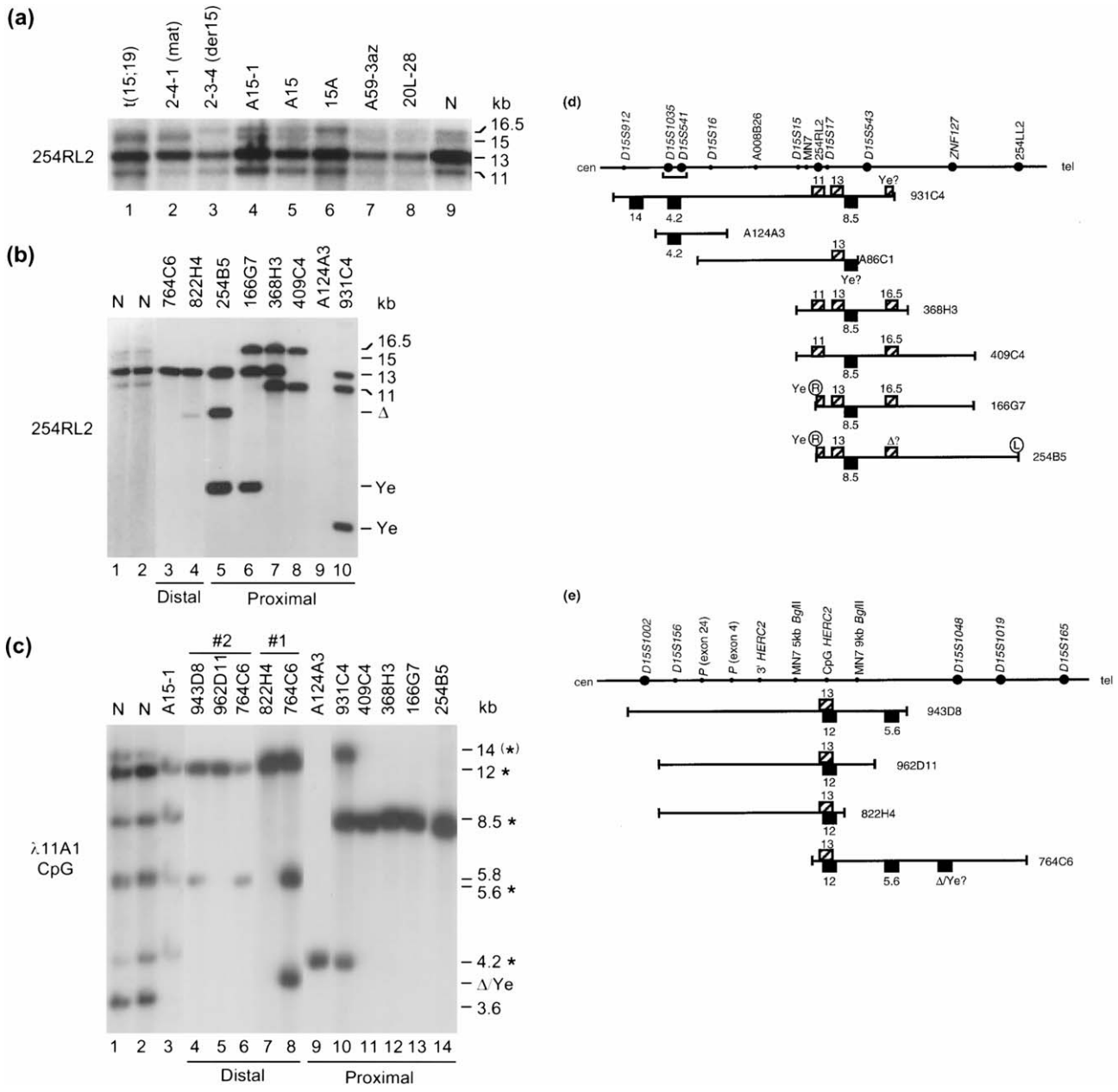

Figure 5 Somatic-cell hybrid and YAC mapping of *END* repeats at the proximal 15q11.2 and distal 15q13 deletion breakpoint regions. *A, Hin*dIII analysis of somatic-cell hybrids with the 254RL2 probe. Hybrids 2-4-1 (lane 2, maternal chromosome 15) and 2-3-4 (lane 3, der [15]) were established from a cell line from an individual with a balanced translocation breaking within the *SNURF-SNRPN* gene (lane 1; Sun et al. 1996). Other hybrids contain either a single chromosome 15 (lanes 5–8; Gabriel et al. 1998) or the 15pter-15q14 region (A15-1; lane 4). Hybrids A15 and A15-1 contain no other human chromosomes. Lane N, Normal cell line. *B,* Mapping of 254RL2-related sequences within YACs spanning the proximal and distal breakpoint regions. Novel bands, not present in normal genomic DNAs, represent the Ye fragments or a putative rearranged fragment (Δ). *C,* Mapping of λ11A1 CpG island–related sequences within YACs spanning the proximal and distal breakpoint regions. Abbreviations are as in panels *A* and *B*. * = Band present in chromosome 15pter-15q14 (hybrid A15-1; lane 3); (*) = polymorphic (+/− alleles) band present in proximal chromosome 15 YACs; #1 and #2 = two different YAC isolates. *D,* Marker and YAC map of the proximal 15q11.2 breakpoint region I–II, illustrating the location of 254RL2–related (*hatched boxes*) and λ11A1 CpG island–related (*blackened boxes*) sequences. Markers are indicated by large (present study) or small (location from Christian et al. 1998) blackened circles, and brackets indicate an unknown order with respect to the centromere (cen) and telomere (tel). Abbreviations are as for panel *B*. R = right; L = left; ? = most parsimonious position. *E,* Marker and YAC map of the distal 15q13 breakpoint region, illustrating the location of 254RL2- and λ11A1 CpG island-related sequences. Symbols and abbreviations as for panel *D*.

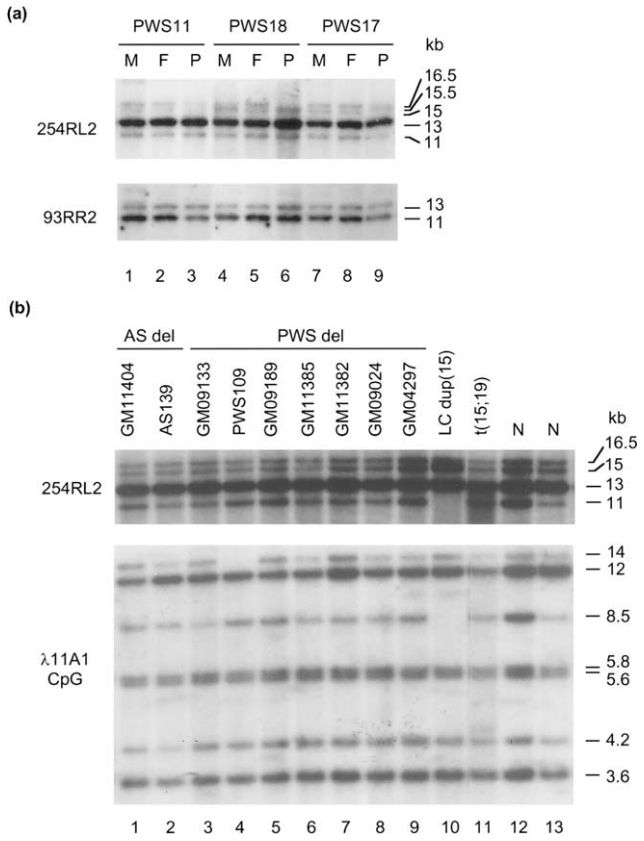

Figure 6 Southern analyses using *END*-repeat probes for patients with chromosome 15q11-q13 rearrangements. *A,* Use of repeat-free probes for 254RL2 (*top*) and 93RR2 (*bottom*), on Southern blots of *Hin*dIII-digested DNAs from PWS-deletion and parental (M = mother; F = father; P = proband) samples. On the basis of IR39 (*D15S18*) polymorphism analysis, PWS11P is heterozygous and hence is a class II deletion, whereas, on the basis of dosage studies, PWS17P is a class I deletion and PWS18P is a class II deletion (Mascari et al. 1992). *B,* Use of the 254RL2 (*top*) and λ11A1 CpG (*bottom*) probes on Southern blots of *Hin*dIII-digested DNAs from AS- and PWS-deletion (del) cell lines, a patient with a duplication of 15q11-q13 (dup [15]), a patient with a balanced translocation t(15;19), and two normal controls (N: PWS19M [lane 12] and PWS23F [lane 13]). Of the nine patients with AS or PWS who are shown in panel (*b*), on the basis of heterozygosity at *D15S541* and/or *D15S1035*, all had class II deletions, with the exception of cell line AS139, which is either class I or II (one allele at both markers); GM09024 was not analyzed.

band detected by 254RL2 is present in both proximal (15q11.2) and distal (15q13) YACs (fig. 5*b*, lanes 3 and 4 vs. lanes 5–7 and 10), confirming interpretations of hybrid data (see above). Second, four proximal YACs each contain at least three copies of the 254RL2 sequence, on the basis of *Hin*dIII restriction-fragment-length variants (RFLVs) between different copies of the *END*-repeat sequence detected by 254RL2 (fig. 5*b*, lanes 5–7 and 10). Some of these fragments are not found in human genomic DNA and, hence, likely represent YAC ends (Ye), or rearrangements; the Ye fragment identified as the strongest-intensity band in YAC 254B5 (fig. 5*b*, lane 5) was predicted, since 254RL2 was derived from the right end of this YAC (fig. 1*a* and table 2). The same end fragment was identified in YAC 166G7 (fig. 5*b*, lane 6), which was confirmed by PCR using two different YAC-end and 254RL2-specific primers (data not shown). Combining the 254RL2 YAC data with STS data for overlap of these YACs from other studies (Christian et al. 1998) allowed the generation of the most parsimonious map in which a total of at least three copies of 254RL2-related sequences are spread throughout the proximal (15q11.2) breakpoint region II (fig. 5*d*). Nevertheless, because of the significantly increased relative intensity of the 13-kb 254RL2 *Hin*dIII band in genomic DNA compared with YACs, we cannot rule out the presence of additional copies of related sequences more proximal in 15q11.

The λ11A1 CpG-island probe may detect one endpoint of the duplicated sequences, since we detect the highest frequency of *Hin*dIII RFLVs for different *END*-repeat copies by using this probe, with at least seven fragments detected (figs. 5*c* and 6). Alternatively, rearrangements subsequent to duplication may have generated this diversity. It should be noted that the probe is only 247 bp and does not include a *Hin*dIII site (nor do two related genomic copies; see below) and that similar results are obtained with other enzymes (data not shown). By analysis of hybrid A15-1, at least four of these CpG-island bands map to chromosome 15pter-q14 (fig. 5*c*, lane 3), although YAC studies (lane 10) and other hybrids (A9+15, 2-4-1 and 2-3-4; data not shown) also map the 14-kb band to proximal 15q11.2; the absence of this fragment in A15-1 likely represents a null polymorphism (see below). The 3.6-kb genomic band detected with this probe but not present in either A15-1 or other chromosome 15 hybrids (data not shown) derives from chromosome 16p11.2, since a bacterial artificial chromosome (GenBank AC002041) derived from this locus contains a 3.6-kb *Hin*dIII fragment with a sequence 88% identical to that of the CpG island present in λ11A1.

All five chromosome 15–specific fragments detected by the CpG-island probe can be mapped to proximal 15q11.2 YACs or distal 15q13 YACs (fig. 5*c–e*). Three different but related loci map within YAC 931C4 (fig. 5*c*, lane 10); on the basis of overlap with other YACs, the 8.5-kb band maps distally and is shared with four other YACs (fig. 5*c*, lanes 11–14, and fig. 5*d*), the 4.2-kb band is shared with YAC A124A3 (fig. 5*c*, lane 9, and fig. 5*d*), and the 14-kb band maps most centromeric (under the assumption that other YACs do not contain the null allele at this locus). Interestingly, in addition to these three bands (4.2, 8.5, and 14 kb), hybrid 2-3-4 contains both a 12-kb CpG-island band that appears to be multicopy in genomic DNA (figs. 5*c* and 6), and a

9-kb band novel to this cell line, indicating that additional copies of related sequences may be found proximal to the translocation breakpoint (fig. 1*a* and data not shown; on the basis of the absence of a 3.6-kb band, this hybrid does not appear to have chromosome 16).

Similarly, YAC 764C6 from 15q13 contains three related loci, the 12-kb and 5.6-kb genomic bands and a putative rearranged (Δ; or YAC-end) fragment (Fig. 5c, lane 8). Since YACs 822H4, 962D11, and 943D8 contain the 12-kb CpG-island band (fig. 5*c*, lanes 4, 5, and 7) and the *HERC2* gene (see below), it is likely that the 12-kb band corresponds to the promoter of this gene (fig. 5*e*). The 5.6-kb band maps more distal (fig. 5*c*, lanes 4, 6, and 8 and fig. 5*e*). When these data are combined with other YAC-overlap data (Buiting et al. 1998), the most parsimonious map suggests that three copies of the CpG island–related sequence occur in 15q13 (fig. 5*e*). We predict that the 764C6 Δ or YAC-end band corresponds to a second copy of the 12-kb fragment, since the 12-kb band shows increased intensity in genomic DNA but intensity equal to that of the 5.8-kb and Δ or YAC-end bands in 764C6 (fig. 5*c*).

Since the *END*-repeat probes detect duplicons localized at or close to the putative PWS- and AS-deletion breakpoints, we analyzed, by Southern hybridization with the 254RL2 (fig. 6*a* and *b*), 93RR2 (fig. 6*a*) and CpG-island (fig. 6*b*) probes, a series of probands with 15q11-q13 rearrangements. Parents were also analyzed for several PWS deletions (fig. 6*a*). Although no rearrangements were identified (fig. 6*a* and *b*), qualitative-dosage analyses allowed interpretations in families with PWS (fig. 6*a*). All 254RL2 and 93RR2 bands show reduced intensity in PWS17P (fig. 6*a*; lane 9 has loading equal to that of 17F in lane 8 and more than that of 17M in lane 7), consistent with a class I deletion. PWS11P, a class II deletion, shows reduced intensity and hence a deletion for only the 11-kb 93RR2 band and the 15+16.5-kb 254RL2 band, whereas PWS18P, also a class II deletion, does not show deletion of any band, for either of these two probes, and appears to show an increase in intensity of the 13-kb 254RL2 *Hin*dIII fragment (fig. 6*a*). Allelic polymorphism was also noted for four fragments, with two of the probes: (i) the 16.5-kb 254RL2 band is replaced by a 15.5-kb allele in some individuals (fig. 6*a*, lanes 4 and 6; data not shown); (ii) the 14-kb CpG-island band in one patient with PWS appears to contain a null allele (fig. 6*b*, lane 4), as found above for the normal chromosome 15 segment in hybrid A15-1; and (iii) the 254RL2 11-kb band and the 8.5-kb CpG-island band appear to represent a linked null allele in one patient with a dup (15) (fig. 6*b*, lane 10; YAC data also link these two fragments; fig. 5*d*). Since these null alleles occur in patients who have one normal chromosome 15, the absence of these bands cannot reflect solely the de novo rearrangement and must at least represent a null polymorphism on the normal chromosome 15.

A Novel Gene Family Associated with the END *Repeat*

Sequence analysis of two genomic λ phage clones and the 254RL2 microclone identified three paralogous sequences (see above). Database searches identified a flanking "exon" sequence in λ6A1 that was paralogous to an expressed sequence tag, EST05046, which subsequently led to the cloning of a paralogous full-length 15.5-kb transcript for the human *HERC2* gene (Ji et al. 1999). The λ11A1 clone also contains an "exon" paralogous to that in λ6A1 and *HERC2* (table 2) but is truncated, by 46 bp, at the 5′ end of this exon. The λ11A1 clone has a genomic deletion 5′ of this, with respect to *HERC2* (data not shown), but does contain a CpG island and two other exons paralogous to those in *HERC2* (fig. 1*b*; these λ11A1 exons are identical to EST05046). These data indicate that the *END* repeats, containing the 254RL2 and the CpG-island sequences, are derived from evolutionary duplications of the *HERC2* gene (Ji et al. 1999).

We have demonstrated elsewhere (Ji et al. 1999) that a subset of the *HERC2*-derived duplications include the *D15F37* (MN7) family of 6–7 kb transcripts. Further analysis of genomic sequence for the 5′ half of the *HERC2* gene (GenBank accession number AC004583), a duplicated but highly deleted chromosome 16p11.2 copy (GenBank accession number AC002041), and of a clone containing a duplicated segment corresponding to *HERC2* cDNA coordinates 9576–12291 (GenBank accession number AC004460) identified sequences homologous to (i) 254RL2 at the expected location in a *HERC2* intron (at 2932–2933) and (ii) 318RL3 at a position just upstream of cDNA coordinate 9576, presumably in a *HERC2*-related intron. Therefore, most if not all of the duplicated sequences identified in 15q11.2 and 15q13, in both this study and our previous study (Ji et al. 1999), appear to originate from the *HERC2*-derived *END* repeats.

A 1.1-kb cDNA probe from the ancestral *HERC2* gene, which is repeated in multiple copies at proximal and distal 15q11-q13 but is not part of the *D15F37* transcripts (Ji et al. 1999; data not shown) is ubiquitously expressed as a 15.5-kb transcript (Ji et al. 1999). The 15.5-kb *HERC2* transcript is detected at high levels in male and female germline tissues (fig. 7*a*); similarly, the MN7 probe detects a very high level of diffuse transcripts in the 6–7 kb region in testis and ovary, in addition to weakly detecting the *HERC2* transcript (fig. 7*b*). Therefore, large parts of the *END* repeats are actively transcribed in germline tissues.

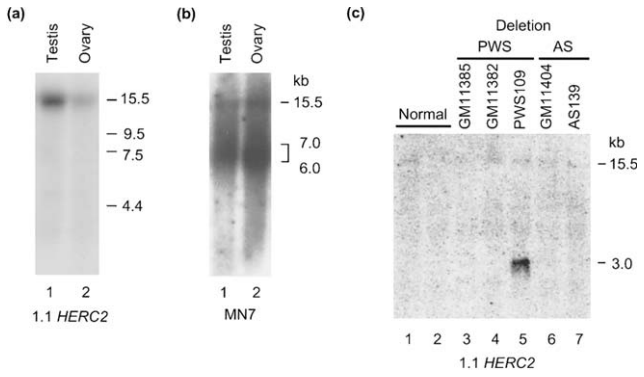

Figure 7 Expression analyses of the *END* repeats in germline tissues, and identification of a putative junction transcript in a patient with PWS. Northern blot analysis of polyA(+) mRNA by the 1.1 *HERC2* probe (*A*) or MN7 probe (*B*), showing abundant expression of *HERC2* (15.5 kb) and duplicated transcripts (6–7 kb), respectively, in testis and ovary. *C*, Northern blot analysis of total RNA for patients with PWS or AS, and for normal controls, with a *HERC2* probe. A faint 15.5-kb *HERC2* transcript from the normal allele is seen in all lanes, whereas a novel 3.0-kb transcript is also seen in PWS109, which has a class II deletion. For deletion class designations, see figure 6*b*.

Preliminary Screen for Deletion Breakpoints in Patients with either the PWS Deletion or the AS Deletion

Because of the complexity of size and copy number of the *END* repeats, probably coupled with the variable but highly methylated nature of proximal 15q11-q12 sequences analyzed by pulsed-field gel electrophoresis (PFGE; Nicholls et al. 1989), standard PFGE methods for analysis of PWS and AS breakpoints have, to date, been unsuccessful. Furthermore, because of the high sequence similarity between *END*-repeat units at proximal and distal 15q11-q13, crossovers may not be detected by standard Southern blot analysis, since the breakpoint restriction fragments may be identical in size to the "parental" recombining fragments. However, on the basis of the large transcripts produced by the *HERC2*-duplicated loci within the *END* repeats, we reasoned that some deletion breakpoints within repeat units located in the same orientation proximally (15q11.2) and distally (15q13) might be detected as abnormally sized transcripts, by northern blot analysis.

Therefore, total RNA was prepared from three PWS- and two AS-deletion cell lines and from two normal control cell lines, and northern blots were probed with a 5′ *HERC2*-specific probe, to maximize detection of altered transcripts. The wild-type 15.5-kb *HERC2* transcript was detected at low levels by this analysis, in each normal cell line and from the normal chromosome present in patients with a deletion (fig. 7*c*). No other transcript has been detected on northern blots from the *HERC2* gene, in normal human or mouse tissues at various stages of development (Ji et al. 1999; Walkowicz et al. 1999). In contrast, in one (PWS109) of five deletion samples, an abundant steady-state level of an abnormal 3.0-kb transcript was detected (fig. 7*c*, lane 5). This represents a putative truncated or fusion transcript generated from the recombinant chromosome in this patient.

Discussion

We have identified duplicated sequences (the *END* repeat) at the proximal and distal regions at which patients with PWS or AS show clustered deletion breakpoints. Analyses of the proximal extent of deletion in a large series of patients with PWS or AS confirmed that ~60% have class II and ~40% have class I deletions (Knoll et al. 1990; Christian et al. 1995). Our studies have also refined the distal breakpoint cluster and indicate that ~95% of PWS and AS deletions break in a small region between *D15S1002* and *D15S1048*. The latter region is even smaller, since the *P* gene (Spritz et al. 1997) and 3′ *HERC2* (Ji et al. 1999)—two loci that are telomeric of *D15S1002*—are also deleted in the majority of patients with PWS or AS. Indeed, the *END* repeats have derived evolutionarily by genomic duplications from the ancestral large and unique gene, *HERC2*, located at the distal PWS/AS breakpoint (Ji et al. 1999). On the basis of the presence of at least three to five *END*-repeat units at both 15q11.2 and 15q13, many of which are transcriptionally active, we suggest a simple model (fig. 8) in which these sequences are directly involved in the misalignment of chromosome 15 chromatids during meiosis. This allows homologous recombination between proximal and distal *END*-repeat units to generate the common deletion classes observed in PWS and AS, and a similar event may also be associated with some instances of duplications. A rearrangement between inverted *END*-repeat units can also provide a simple ex-

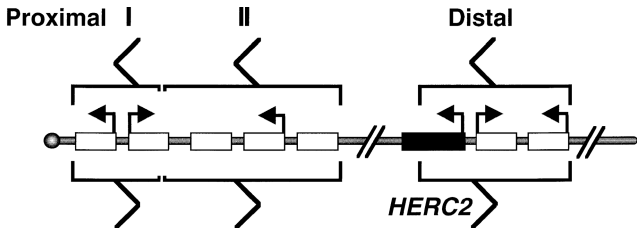

Figure 8 Model for the arrangement of transcriptionally active, paralagous *END*-repeat units within chromosome 15q11-q13. Eight *END*-repeat units, each derived from a genomic duplication of a *HERC2*- or *HERC2*-related locus, have confirmed locations at either end of 15q11-q13. On the basis of hybrid 2-3-4 data, two additional *END*-repeat units map proximal of copies shown here. Active transcription is shown by arrows, although the orientation of *END*-repeat units is arbitrary, except for *HERC2* (Ji et al. 1999). The presence of large, duplicated structures at either end of 15q11-q13 suggests a simple model for chromatid misalignment and recombination.

planation for inversions in this region. Alternatively, even if some recombination breakpoints map outside the actual repeat units, these large blocks of homologous sequence might stabilize interacting chromosome 15 homologues. Finally, both intra- and interchromosomal 15q11-q13 deletions have been observed in PWS and AS (Carrozzo et al. 1997; Robinson et al. 1998), and it is possible that a large chromatin loop of 3–4 Mb (Sachs et al. 1995) brings about the initial *END*-repeat interaction, at least in the intrachromosomal events.

The structure of the *END* repeats is clearly complex. From data presented here and elsewhere (see Buiting et al. 1998; Ji et al. 1999), we estimate that ⩾8 (fig. 8)—and, at most, 10-12—copies of this sequence are located within 15q11 and 15q13. At least two additional copies of highly homologous sequences occur within 16p11.2 (Buiting et al. 1992, 1998; Ji et al. 1999). Furthermore, each of the *END*-repeat units is large, spanning 50–200 kb of sequences related to the ancestral *HERC2* gene, although different subsets of duplicated sequence derived from *HERC2*-related sequence are present in each copy (Ji et al. 1999; present study). At present, it is unknown whether other genes are located within the duplicated *END*-repeat segments or whether unique-sequence "islands" are dispersed within or adjacent to duplicated segments. The homology within small parts of *HERC2*-related sequences of the various *END*-repeat units sequenced to date is 90%–99% (Ji et al. 1999; present study). Nevertheless, the complexity of the *END* repeats flanking 15q11-q13, as well as the similar complexity of repeat structures flanking regions prone to rearrangement in 17p11.2 in Smith-Magenis syndrome (SMS) (MIM 182290; Chen et al. 1997), in 7q11.23 in Williams syndrome (WS) (MIM 194050; Osborne et al. 1997; Pérez Jurado et al. 1998), in 22q11.2 in DiGeorge/velocardiofacial syndrome (DG/VCFS) (MIM 188400; Baumer et al. 1998; Edelmann et al. 1999), and in other regions in genomic diseases (Purandare and Patel 1997; Eichler 1998; Lupski 1998), indicates that the sequence analysis of such regions will be exceedingly difficult and will not be accomplished by standard sequencing methods employed by the Human Genome Project. Therefore, we strongly endorse Eichler's proposal (1998) that these regions receive particularly intense scrutiny, in order to fully understand the role that they play in recurrent chromosomal structural rearrangements in disease.

Since multiple copies of the *END* repeats located at each of the proximal and distal breakpoint regions are identical in restriction-fragment length, homologous recombination would generate an identical fragment length that would make detection of recombination events difficult. Furthermore, PFGE analysis has, to date, not proved useful in analysis of breakpoints associated with this chromosomal region. Nevertheless, in one of five PWS- and AS-deletion samples, we detected, using a 5′ *HERC2* probe and a novel RNA expression assay, a putative truncated or fusion transcript generated from the recombinant chromosome in a patient with PWS (PWS109). This represents the first molecular identification of a putative breakpoint junction in patients with either the PWS or the AS common deletion. Since this probe does not detect the family of 6–7-kb MN7 transcripts (see fig. 7*b*) also derived from duplicated *HERC2*-related sequences (Ji et al. 1999), the use of additional probes detecting these family members should increase the percentage of breakage events detected by this novel RNA-expression assay. Nevertheless, since this method depends on the detection of stable abnormal RNA transcripts, truncated transcripts that are unstable (Maquat 1996) may not be detectable.

Preliminary evidence has led to a suggestion that there may be an association between paternal occupational hydrocarbon exposure and PWS (Strakowski and Butler 1987; Cassidy et al. 1989; Åkefeldt et al. 1995), although large epidemiological studies have not been performed and the possibility has not been investigated in AS. However, it has been shown that low x-ray doses enhance recombination of a genomic duplication in mice (Schiestl et al. 1994). Therefore, it will be important to determine whether environmental factors can enhance the frequency of recombination between *END*-repeat units or similar types of region-specific repetitive sequences in the genome (Eichler 1998; Lupski 1998). Recognition of the molecular and environmental risk factors, if any, that are associated with chromosomal rearrangements mediated by duplicated sequences may allow both prevention of preconceptional exposure and a significant reduction in the frequency of associated genomic disorders.

The physical distances between microsatellite markers that flank the breakpoint regions and *END* repeats—*D15S543* and *D15S541*/*D15S1035* in the proximal region and *D15S1002* and *D15S1048*/*D15S1019* in the distal region—have been estimated by YAC contig information as ~1 Mb and 1.5 Mb, respectively. The genetic distance between the two sets of proximal markers is 2.7 cM (Christian et al. 1995), and that between the distal two sets of markers is 6 cM (Robinson et al. 1993*b*). The large genetic distances and relatively reduced physical distances suggest the possibility of an increase in recombination between these markers, which would colocalize with the *END* repeats and PWS/AS breakpoint regions. Nevertheless, more-exact data will be necessary to determine whether the apparent PWS/AS deletion–breakpoint hot spots at proximal and distal 15q11-q13 are also hot spots for normal meiotic recombination.

Although de novo chromosome rearrangements between duplicons in Charcot-Marie-Tooth syndrome type

1A (MIM 118220) and spinal muscular atrophy (MIM 253300) occur predominantly in spermatogenesis (Palau et al. 1993; Blair et al. 1996; Bort et al. 1997; Wirth et al. 1997), those in WS, SMS, DG/VCS, and PWS and AS occur with equal frequency during male and female meiosis (Dutly and Schinzel 1996; Juyal et al. 1996; Urbán et al. 1996; Carrozzo et al. 1997; Baumer et al. 1998; Robinson et al. 1998). We suggest that one factor that enhances the likelihood of 15q11-q13 recombination in both maternal and paternal germ cells is active transcription of the *END* repeats in germ cells. Similarly, in yeast, hot spots of meiotic and mitotic recombination occur at active promoter regions (Voelkel-Meiman et al. 1987; Nicolas et al. 1989; Wu and Lichten 1994; Klein et al. 1996), and transcription is essential for rearrangement of immunoglobulin genes (Stavnezer 1996; Bachl et al. 1998) and homologous recombination in mammalian cell culture (Nickoloff 1992). Consistent with this hypothesis, the *HERC2* sequence and some related sequences that comprise the *END* repeats have been shown here to include the CpG-island promoter (see also Ji et al. 1999) and to be transcribed at high levels in germline tissues, and there is evidence from mouse that the *Herc2* gene is active in germ cells. As further discussed below, mutations in *Herc2* are responsible for the juvenile development and fertility 2 (*jdf2*) syndrome (Rinchik et al. 1995; Ji et al. 1999; Walkowicz et al. 1999). In the male, at least, these defects are integral to the germ cell and phenotypically are indistinguishable from meiotic mutants (Rinchik et al. 1995), which suggests that the *Herc2*-gene product may be required during meiosis. Therefore, it is possible that the *END* repeats are transcribed in meiotic germ cells, which is consistent with our hypothesis that active transcription promotes chromosome 15q11-q13 recombination to generate deletions.

Since at least five *END*-repeat units map within 15q11 (the proximal PWS/AS breakpoint) and at least three in 15q13 (the distal PWS/AS breakpoint), it is likely that recombination also occurs between clustered duplicons at each of these two locations. Indeed, our identification of several null alleles for RFLVs within these repeats is consistent with this hypothesis. Any recombination within the functional *HERC2* gene in 15q13 (Ji et al. 1999) will lead to a null allele, and it is known that recessive mutations in the orthologous gene in mice lead to the severe juvenile-lethal *jdf2* syndrome (Lehman et al. 1998; Ji et al. 1999; Walkowicz et al. 1999). In addition, all patients studied (10/10) with either the PWS or the AS deletion are hemizygous at *HERC2* (Ji et al. 1999). Therefore, a recessive human disease equivalent to mouse *jdf2* may occur at a reasonably high frequency, as a consequence of a high mutation rate in *HERC2*. Given the likely propensity for recombination either between *END*-repeat units within proximal or distal 15q11-q13 or between proximal and distal 15q11-q13 as seen in PWS and AS, it will be important to determine both the exact number and orientation of copies of the *END*-repeat units at each location and the extent of polymorphic variation in these elements, within different individuals and population groups. This should be possible with techniques such as fiber-FISH (Heiskanen et al. 1995). Since these 15q11-q13 rearrangements in disease and polymorphic variation reflect the evolutionary process, further analysis of the evolutionary origin of the *END* repeats by successive duplication events during primate evolution (see Ji et al. 1999) will allow a better understanding of the molecular evolution, structure, and plasticity of these sequences during gametogenesis.

Acknowledgments

We thank Drs. Y. Sun and C. Palmer, for the t(15;19) cell line; Dr. R. Schultz, for the A15 and A15-1 hybrids; Dr. W. P. Robinson, for advice on microsatellites; Drs. D. H. Ledbetter and A. C. Chinault, for the first series of YAC clones; Drs. B. Horsthemke and K. Buiting, for the MN7 probe, YACs, and discussion; J. Warren, for assistance with some sequencing; and Drs. E. E. Eichler and H. F. Willard, for critical discussion. This work was supported by clinical research grants from the March of Dimes Birth Defects Foundation (to R.D.N.), by the Pew Scholars Program in the Biomedical Sciences (support to R.D.N.), and by donations from anonymous Prader-Willi Syndrome Association members (to R.D.N. and S.B.C.). J.M.A.-L. is supported by National Institutes of Health Training Grant GM08613.

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

BLAST search engine, http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST/nph-blast

Genbank, http://www.ncbi.nlm.nih.gov/Web/Genbank/ (see table 2 for *HERC2* genomic sequence [AC004583] and *HERC2*-related genomic sequence [AC002041 and AC004460])

Genome Database, http://gdbwww.gdb.org/ (for markers for microsatellite studies)

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim (for AS [105830], PWS [176270], CMT1A [118220], SMS [182290], DG/VCF [188400], WS [194050], and SMA [253300])

Whitehead Institute Database, http://www-genome.wi.mit.edu (for YACs)

References

Åkefeldt A, Anvret M, Grandell U, Nordlinder R, Gillberg C (1995) Parental exposure to hydrocarbons in Prader-Willi syndrome. Dev Med Child Neurol 37:1101–1109

Bachl J, Olsson C, Chitkara N, Wabl M (1998) The Ig mutator is dependent on the presence, position, and orientation of the large intron enhancer. Proc Natl Acad Sci USA 95: 2396–2399

Baumer A, Dutly F, Balmer D, Riegel M, Tukel T, Krajewska-Walaseck M, Schinzel AA (1998) High level of unequal meiotic crossovers at the origin of the 22q11.2 and 7q11.23 deletions. Hum Mol Genet 7:887–894

Blair IP, Nash J, Gordon MJ, Nicholson GA (1996) Prevalence and origin of de novo duplications in Charcot-Marie-Tooth disease 1A: first report of a de novo duplication with a maternal origin. Am J Hum Genet 58:472–476

Bort S, Martínez F, Palau F (1997) Prevalence and parental origin of de novo 1.5-Mb duplication in Charcot-Marie-Tooth disease type 1A. Am J Hum Genet 60:230–233

Breukel C, Wijnen J, Tops C, Klift H v/d, Dauwerse H, Khan PM (1990) Vector-Alu PCR: a rapid step in mapping cosmids and YACs. Nucleic Acids Res 18:3097

Browne CE, Dennis NR, Maher E, Long FL, Nicholson JC, Sillibourne J, Barber JCK (1997) Inherited interstitial duplications of proximal 15q: genotype-phenotype correlations. Am J Hum Genet 61:1342–1352

Buiting K, Greger V, Brownstein BH, Mohr RM, Voiculescu I, Winterpacht A, Zabel B, et al (1992) A putative gene family in 15q11-13 and 16p11.2: possible implications for Prader-Willi and Angelman syndromes. Proc Natl Acad Sci USA 89:5457–5461

Buiting K, Groß S, Ji Y, Senger G, Nicholls RD, Horsthemke B (1998) Expressed copies of the MN7 (*D15F37*) gene family map close to the common deletion breakpoints in the Prader-Willi/Angelman syndromes. Cytogenet Cell Genet 81:247–253

Butler MG (1990) Prader-Willi syndrome—current understanding of cause and diagnosis. Am J Med Genet 35: 319–332

Carrozzo R, Rossi E, Christian SL, Kittikamron K, Livieri C, Corrias A, Pucci L, et al (1997) Inter- and intrachromosomal rearrangements are both involved in the origin of 15q11-q13 deletions in Prader-Willi syndrome. Am J Hum Genet 61:228–231

Cassidy SB (1997) Prader-Willi Syndrome. J Med Genet 34: 917–923

Cassidy SB, Conroy J, Becker L, Schwartz S (1996) Paternal triplication of 15q11-q13 in a hypotonic, developmentally delayed child without Prader-Willi or Angelman syndrome. Am J Med Genet 62:206–207

Cassidy SB, Gainey AJ, Butler MG (1989) Occupational hydrocarbon exposure among fathers of Prader-Willi syndrome patients with and without deletion of 15q. Am J Hum Genet 44:806–810

Chen KS, Manian P, Koeuth T, Potocki L, Zhao Q, Chinault AC, Lee CC, et al (1997) Homologous recombination of a flanking repeat gene cluster is a mechanism for a common contiguous gene deletion syndrome. Nat Genet 17:154–163

Cheng S-D, Spinner NB, Zackai EH, Knoll JHM (1994) Cytogenetic and molecular characterization of inverted duplication chromosomes 15 from 11 patients. Am J Hum Genet 55:753–759

Christian SL, Bhatt NK, Martin SA, Sutcliffe JS, Kubota T, Huang B, Mutirangura A, et al (1998) Integrated YAC contig map of the Prader-Willi/Angelman region on chromosome 15q11-q13 with average STS spacing of 35 kb. Genome Res 8:146–157

Christian SL, Robinson WP, Huang B, Mutirangura A, Line MR, Nakao M, Surti U, et al (1995) Molecular characterization of two proximal deletion breakpoint regions in both Prader-Willi and Angelman syndrome patients. Am J Hum Genet 57:40–48

Clayton-Smith J, Driscoll DJ, Waters MF, Webb T, Andrews T, Malcolm S, Pembrey ME, et al (1993*a*) Difference in methylation patterns within the *D15S9* region of chromosome 15q11-q13 in first cousins with Angelman syndrome and Prader-Willi syndrome. Am J Med Genet 47:683–686

Clayton-Smith J, Pembrey ME (1992) Angelman syndrome. J Med Genet 29:412–415

Clayton-Smith J, Webb T, Cheng XJ, Pembrey ME, Malcolm S (1993*b*) Duplication of chromosome 15 in the region 15q11-13 in a patient with developmental delay and ataxia with similarities to Angelman syndrome. J Med Genet 30: 529–531

Crolla JA, Harvey JF, Sitch FL, Dennis NR (1995) Supernumerary marker 15 chromosomes: a clinical, molecular and FISH approach to diagnosis and prognosis. Hum Genet 95: 161–170

Dutly F, Schinzel A (1996) Unequal interchromosomal rearrangements may result in elastin gene deletions causing the Williams-Beuren syndrome. Hum Mol Genet 5:1893–1898

Edelmann L, Pandita RK, Morrow BE (1999) Low-copy repeats mediate the common 3-Mb deletion in patients with velo-cardio-facial syndrome. Am J Hum Genet 64: 1076–1086

Eichler EE (1998) Masquerading repeats: paralogous pitfalls of the human genome. Genome Res 8:758–762

Gabriel JM, Higgins MJ, Gebuhr TC, Shows TB, Saitoh S, Nicholls RD (1998) A model system to study genomic imprinting of human genes. Proc Natl Acad Sci USA 95: 14857–14862

Gray TA, Saitoh S, Nicholls RD (1999) An imprinted, mammalian bicistronic transcript encodes two independent proteins. Proc Natl Acad Sci USA 96:5616–5621

Heiskanen M, Hellsten E, Kallioniemi OP, Makela TP, Alitalo K, Peltonen L, Palotie A (1995) Visual mapping by fiber-FISH. Genomics 30:31–36

Huang B, Crolla JA, Christian SL, Wolf-Ledbetter ME, Macha ME, Papenhausen PN, Ledbetter DH (1997) Refined molecular characterization of the breakpoints in small inv dup(15) chromosomes. Hum Genet 99:11–17

Ji Y, Walkowicz MJ, Buiting K, Johnson DK, Tarvin R, Rinchik EM, Horsthemke B, et al (1999) The ancestral gene for transcribed, low-copy repeats in the Prader-Willi/Angelman region encodes a large protein implicated in protein trafficking that is deficient in mice with neuromuscular and spermiogenic abnormalities. Hum Mol Genet 8:533–542

Jong MTC, Gray TA, Ji Y, Glenn CC, Saitoh S, Driscoll DJ, Nicholls RD (1999) A novel imprinted gene, encoding a RING zinc-finger protein, and overlapping antisense transcript in the Prader-Willi syndrome critical region. Hum Mol Genet 8:783–793

Juyal RC, Figuera LE, Hauge X, Elsea SH, Lupski JR, Greenberg F, Baldini A, et al (1996) Molecular analyses of 17p11.2

deletions in 62 Smith-Magenis syndrome patients. Am J Hum Genet 58:998–1007

Klein S, Zenvirth D, Sherman A, Ried K, Rappold G, Simchen G (1996) Double-strand breaks on YACs during yeast meiosis may reflect meiotic recombination in the human genome. Nat Genet 13:481–484

Knoll JHM, Nicholls RD, Magenis RE, Glatt K, Graham JM Jr, Kaplan L, Lalande M (1990) Angelman syndrome: three molecular classes identified with chromosome 15q11q13-specific DNA markers. Am J Hum Genet 47:149–154

Kuwano A, Mutirangura A, Dittrich B, Buiting K, Horsthemke B, Saitoh S, Niikawa N, et al (1992) Molecular dissection of the Prader-Willi/Angelman syndrome region (15q11-13) by YAC cloning and FISH analysis. Hum Mol Genet 1: 417–425

Lehman AL, Nakatsu Y, Ching A, Bronson RT, Oakey RJ, Keipo-Hrynko N, Finger JN, et al (1998) A very large protein with diverse functional motifs is deficient in *rjs* (runty, jerky, sterile) mice. Proc Natl Acad Sci USA 95:9436–9441

Lupski JR (1998) Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. Trends Genet 14:417–422

Malzac P, Webber H, Moncla A, Graham JM Jr, Kukolich M, Williams C, Pagon RA, et al (1998) Mutation analysis of UBE3A in Angelman syndrome patients. Am J Hum Genet 62:1353–1360

Maquat LE (1996) Defects in RNA splicing and the consequence of shortened translational reading frames. Am J Hum Genet 59:279–286

Mascari MJ, Gottlieb W, Rogan P, Butler MG, Armour J, Jeffreys A, Waller D, et al (1992) The frequency of uniparental disomy in Prader-Willi syndrome: implications for molecular diagnosis. N Engl J Med 326:1599–1607

McDaniel LD, Schultz RA (1992) Elevated sister chromatid exchange phenotype of Bloom syndrome cells is complemented by human chromosome 15. Proc Natl Acad Sci USA 89:7968–7972

Nicholls RD, Knoll JHM, Butler MG, Karam S, Lalande M (1989) Genetic imprinting suggested by maternal heterodisomy in non-deletion Prader-Willi syndrome. Nature 342: 281–285

Nicholls RD, Saitoh S, Horsthemke B (1998) Imprinting in Prader-Willi and Angelman syndromes. Trends Genet 14: 194–200

Nickoloff JA (1992) Transcription enhances intrachromosomal homologous recombination in mammalian cells. Mol Cell Biol 12:5311–5318

Nicolas A, Treco D, Schultes NP, Szostak JW (1989) An initiation site for meiotic gene conversion in the yeast *Saccharomyces cerevisiae*. Nature 338:35–39

Osborne LR, Herbrick J, Greavette T, Heng HHQ, Tsui L, Scherer SW (1997) PMS2- related genes flank the rearrangement breakpoints associated with Williams syndrome and other diseases on human chromosome 7. Genomics 45: 402–406

Palau F, Lofgren A, De Jonghe P, Bort S, Nelis E, Sevilla T, Martin J, et al (1993) Origin of *de novo* duplication in Charcot-Marie-Tooth disease type 1A: unequal nonsister chromatid exchange during spermatogenesis. Hum Mol Genet 2:2031–2035

Pérez Jurado LA, Wang Y, Peoples R, Coloma A, Cruces J, Francke U (1998) A duplicated gene in the breakpoint regions of the 7q11.23 Williams-Beuren syndrome deletion encodes the initiator binding protein TFII-I and BAP-135, a phosphorylation target of BTK. Hum Mol Genet 7: 325–334

Purandare SM, Patel PI (1997) Recombination hot spots and human disease. Genome Res 7:773–786

Repetto GM, White LM, Bader PJ, Johnson D, Knoll JHM (1998) Interstitial duplications of chromosome region 15q11q13: Clinical and molecular characterization. Am J Med Genet 9:82–89

Rinchik EM, Carpenter DA, Handel MA (1995) Pleiotropy in microdeletion syndromes: neurologic and spermatogenic abnormalities in mice homozygous for the p^{6H} deletion are likely due to dysfunction of a single gene. Proc Natl Acad Sci USA 92:6394–6398

Robinson WP, Binkert F, Gine R, Vazques C, Miller W, Rosenkranz W, Schinzel A (1993*a*) Clinical and molecular analysis of five inv dup(15) patients. Eur J Hum Genet 1:37–50

Robinson WP, Dutly F, Nicholls RD, Bernasconi F, Penaherrera M, Michaelis RC, Abeliovich D, et al (1998) The mechanisms involved in formation of deletions and duplications of 15q11-q13. J Med Genet 35:130–136

Robinson WP, Spiegel R, Schinzel AA (1993*b*) Deletion breakpoints associated with the Prader-Willi and Angelman syndromes (15q11-q13) are not sites of high homologous recombination. Hum Genet 91:181–184

Sachs RK, Van Den Engh G, Trask B, Yokota H, Hearst JE (1995) A random-walk/giant-loop model for interphase chromosomes. Proc Natl Acad Sci USA 92:2710–2714

Schiestl RH, Khogali F, Carls N (1994) Reversion of the mouse pink-eyed unstable mutation induced by low doses of X-rays. Science 266:1573–1576

Schinzel AA, Brecevic L, Bernasconi F, Binkert F, Berthet F, Wuilloud A, Robinson WP (1994) Intrachromosomal triplication of 15q11-q13. J Med Genet 31:798–803

Spritz RA, Bailin T, Nicholls RD, Lee S-T, Park S-K, Mascari MJ, Butler MG (1997) Hypopigmentation in the Prader-Willi syndrome correlates with *P* gene deletion but not with haplotype of the hemizygous *P* allele. Am J Med Genet 71: 57–62

Stavnezer J (1996) Immunoglobulin class switching. Curr Opin Immunol 8:199–205

Strakowski SM, Butler MG (1987) Paternal hydrocarbon exposure in Prader-Willi syndrome. Lancet 2:1458

Sullivan BA, Jenkins LS, Karson EM, Leana-Cox J, Schwartz S (1996) Evidence for structural heterogeneity from molecular cytogenetic analysis of dicentric Robertsonian translocations. Am J Hum Genet 59:167–175

Sun Y, Nicholls RD, Butler MG, Saitoh S, Hainline BE, Palmer CG (1996) Breakage in the *SNRPN* locus in a balanced 46,XY,t(15;19) Prader-Willi syndrome patient. Hum Mol Genet 5:517–524

Tantravahi U, Nicholls RD, Stroh HS, Ringer S, Neve RL, Kaplan L, Wharton R, et al (1989) Quantitative calibration and use of DNA probes for investigating chromosomal abnormalities in the Prader-Willi syndrome. Am J Med Genet 33:78–87

Urbán Z, Helms C, Fekete G, Csiszár K, Bonnet D, Munnich

A, Donis-Keller H, et al (1996) 7q11.23 deletions in Williams syndrome arise as a consequence of unequal meiotic crossover. Am J Hum Genet 59:958–962

Voelkel-Meiman K, Keil RL, Roeder GS (1987) Recombination-stimulating sequences in yeast ribosomal DNA correspond to sequences regulating transcription by RNA polymerase I. Cell 48:1071–1079

Walkowicz M, Ji Y, Ren X, Horsthemke B, Russell LB, Johnson DK, Rinchik EM, et al. Molecular characterization of radiation- and chemically-induced mutations associated with neuromuscular tremors, runting, juvenile lethality, and sperm defects in *jdf2* mice. Mamm Genome (in press)

Wandstrat AE, Leana-Cox J, Jenkins L, Schwartz S (1998) Molecular cytogenetic evidence for a common breakpoint in the largest inverted duplications of chromosome 15. Am J Hum Genet 62:925–936

Webb T (1994) Inv dup(15) supernumerary marker chromosomes. J Med Genet 31:585–594

Willard HF, Holmes MT (1984) A sensitive and dependable assay for distinguishing hamster and human X-linked steroid sulfatase activity in somatic cell hybrids. Hum Genet 66:272–275

Wirth B, Schmidt T, Hahnen E, Rudnick-Schöneborn S, Krawczak M, Müller-Myhsok B, Schonling J, et al (1997) De novo rearrangements found in 2% of index patients with spinal muscular atrophy: mutational mechanisms, parental origin, mutation rate, and implications for genetic counseling. Am J Hum Genet 61:1102–1111

Wu T, Lichten M (1994) Meiosis-induced double-strand break sites determined by yeast chromatin structure. Science 263: 515–518

Zackowski JL, Nicholls RD, Gray BA, Bent-Williams A, Gottlieb W, Harris PJ, Waters MF, et al (1993) Cytogenetic and molecular analysis in Angelman syndrome. Am J Med Genet 46:7–11