Human Molecular Genetics, 1999, Vol. 8, No. 6 **1025–1037**

Large genomic duplicons map to sites of instability in the Prader–Willi/Angelman syndrome chromosome region (15q11–q13)

Susan L. Christian, Judy A. Fantes, Stephanie K. Mewborn, Bing Huang[1] and David H. Ledbetter[+]

Department of Human Genetics, University of Chicago, 924 East 57th Street, Chicago, IL 60637, USA and [1]Genzyme Genetics, Longbeach, CA 90806, USA

Received December 24, 1998; Revised and Accepted March 17, 1999 DDBJ/EMBL/GenBank accession nos AF077842, AF077843, AF077845, AF077846, AF077848, AF077849, AF077851–4, AF129926–30

The most common etiology for Prader–Willi syndrome and Angelman syndrome is *de novo* interstitial deletion of chromosome 15q11–q13. Deletions and other recurrent rearrangements of this region involve four common 'hotspots' for breakage, termed breakpoints 1–4 (BP1–BP4). Construction of an ~4 Mb YAC contig of this region identified multiple sequence tagged sites (STSs) present at both BP2 and BP3, suggestive of a genomic duplication event. Interphase FISH studies demonstrated three to five copies on 15q11–q13, one copy on 16p11.1–p11.2 and one copy on 15q24 in normal controls, while analysis on two Class I deletion patients showed loss of approximately three signals at 15q11–q13 on one homolog. Multiple FISH signals were also observed at regions orthologous to both human chromosomes 15 and 16 in non-human primates, including Old World monkeys, suggesting that duplication of this region may have occurred ~20 million years ago. A BAC/PAC contig for the duplicated genomic segment (duplicon) demonstrated a size of ~400 kb. Surprisingly, the duplicon was found to contain at least seven different expressed sequence tags representing multiple genes/pseudogenes. Sequence comparison of STSs amplified from YAC clones uniquely mapped to BP2 or BP3 showed two different copies of the duplicon within BP3, while BP2 comprised a single copy. The orientation of BP2 and BP3 are inverted relative to each other, whereas the two copies within BP3 are in tandem. The presence of large duplicated segments on chromosome 15q11–q13 provides a mechanism for homologous unequal recombination events that may mediate the frequent rearrangements observed for this chromosome.

INTRODUCTION

Contiguous gene syndromes are caused by deletions or duplications of large regions of the human genome involving multiple unrelated genes physically contiguous on a chromosome. One of the most common sites of rearrangement is 15q11–q13, where multiple structural abnormalities, including deletions, supernumerary marker chromosomes, duplications, triplications and translocations, are observed. Interstitial deletions of proximal 15q account for ~70% of both Prader–Willi syndrome (PWS) and Angelman syndrome (AS) patients, with an estimated deletion frequency of ~1/10 000 live births (reviewed in ref. 1). A paternal deletion results in PWS, whereas a maternal deletion results in AS, due to genomic imprinting within this region (2,3). In addition to high frequency deletion events, chromosome 15 accounts for ~50% of all supernumerary marker chromosomes observed in man (4). The estimated frequency for chromosome 15-derived markers, commonly referred to as inv dup(15), is ~1/5000 live births (4). Less frequent rearrangements of this same region of chromosome 15 include duplications (5–9) and triplications (9–11).

Increasing evidence suggests the presence of four specific 'hotspots' or breakpoint clusters located in proximal 15q. Using restriction fragment length polymorphism (RFLP) analysis, Knoll *et al*. (12) first identified two classes of AS deletion (Class I and II) based on two proximal breakpoint clusters, flanking S18, with apparent uniformity in extent of deletion on the distal side. Kuwano *et al*. (13) used fluorescence *in situ* hybridization (FISH) analysis to identify a common proximal breakpoint within a single YAC clone (y254B5), in 10/10 PWS/AS deletion patients, and a recurring distal breakpoint within a single YAC (y93C9), in 8/9 deletion cases, consistent with the presence of 'hotspots' of rearrangement at these sites.

A first generation YAC contig of the chromosome 15q11–q13 region was constructed in 1993 which contained several gaps and a small number of polymorphic microsatellite markers useful for mapping the breakpoint regions (14). Additional microsatellite markers were developed from the proximal end of the contig and used to analyze patients known to contain deletions of this region by FISH (15). Microsatellite analysis of 53 PWS and 33 AS deletion patients showed 44% to be deleted for S542 (Class I), while 56% were not deleted (Class II) (15). The site of breakage for Class I patients is referred to as breakpoint (BP)1 and the site of breakage for Class II patients is

[+]To whom correspondence should be addressed. Tel: +1 773 834 0525; Fax: +1 773 834 0505; Email: dhl@genetics.uchicago.edu

referred to as BP2. Thus, both PWS and AS deletion cases show a remarkable consistency in deletion size and identify three hotspots for chromosomal breakage in the15q11–q13 region, hereafter referred to as BP1, BP2 and BP3.

Apparently consistent breakpoints are also associated with the inv dup(15) marker chromosome, producing two primary size classes which correlate with phenotype (16–19). Small inv dup(15) chromosomes do not contain the PWS/AS common deletion region and are usually associated with a normal phenotype. FISH and microsatellite analysis of 18 small inv dup(15) chromosomes showed that they involve rearrangement at the same two proximal breakpoint regions (BP1 and BP2) found in PWS/AS deletion patients (18).

The fourth hotspot for chromosome breakage, BP4, is located between S1031 and S1010 and is the breakpoint for most large inv dup(15) chromosomes (9), as well as in some cases of duplication and triplication (9; unpublished data). Patients with either three or four maternally derived copies of the PWS/AS critical region have an abnormal phenotype that includes developmental delay or mental retardation, seizures and autistic-like features (20,21). The breakpoints involved in formation of large inv dup(15) chromosomes have only recently begun to be characterized and appear to involve either BP3 or BP4 (9,19; unpublished data).

Maternal duplications including the PWS/AS region have also been observed in patients ascertained with typical autism, while individuals inheriting the same duplication of paternal origin appear to be normal (21). The distal breakpoint in interstitial duplications may involve either BP3 or BP4 (7–9). Interestingly, BP4 has also been implicated in patients with maternal triplications of proximal 15q where three copies of the BP1–BP4 region are present in the maternally derived chromosome, suggestive of a common mechanism of rearrangement (10,11).

The hypothesis that the genomic instability in 15q11–q13 might be due to repeated DNA elements in this region was first proposed by Donlon *et al*. (22). In this study, 10 chromosome 15-specific probes, later termed S9–S18 (23), were developed from flow-sorted large inv dup(15) chromosome-enriched phage libraries (22). Clones S9–S14 were mapped within the BP2–BP3 region, while S15–S18 were mapped outside it (22,23). Three unspecified clones were found to be present in multiple copies, which could potentially be involved in the breakpoint regions (22).

Further evidence of multiple copy sequence comes from data for a gene family identified within 15q11–q13. A clone from a proximal 15q microdissection library, MN7, identified a gene family with four to five copies on 15q11–q13 and at least one copy on 16p11.2 (24). Two cDNA clones were identified, one mapping to chromosome 15 and one to chromosome 16, and northern analysis indicated the presence of both a 15 and a 7 kb transcript on chromosome 15 (24). The 15 kb transcript corresponds to the *HERC2* (also known as the *ERY-1*) gene in human (25,26) and the *rjs* (27) or *herc2* (26) gene in mouse, while a nearly complete cDNA for the human 7 kb transcript (KIAA0393) was sequenced by Nagase *et al*. (28).

Recently, we developed a detailed ~4 Mb YAC contig providing the first complete, integrated physical map of the 15q11–q13 region (29). Eleven sequence tagged sites (STSs), including five expressed sequence tags (ESTs), were localized between S542 and S543, the site of BP2 (13,15,29). Unexpectedly, extension of this map demonstrated duplication of all 11 STSs between S542 and S543 to distal yeast artificial chromosomes (YACs) overlapping y93C9, the location of BP3, revealing a large duplicated segment at the sites of the common deletion breakpoints in PWS and AS. We now present detailed molecular and FISH analysis of the BP2 and BP3 duplicated segments associated with the common deletion breakpoints observed in PWS and AS.

RESULTS

YAC-based STS content mapping

The integrated YAC contig of the 15q11–q13 region provided the framework to begin mapping and cloning of the four breakpoint regions. Toward that goal, a minimum tiling path of YACs extending from the most proximal marker, NIB1540, to S144, located distal to BP4, was created (data not shown). Unexpectedly, several previously unmapped ESTs within this region, including A006B10, A008B26, SHGC15126 and SHGC17218, were found to be present in YACs at the sites of BP2, BP3 and some at BP4 (Fig. 1 and data not shown). Following this result, all STSs present in y931C4, the site of BP2, were tested against YACs at BP3 (e.g. y943D8, y962D11 and y893H9). Initially, 11 STSs were found to map to YACs at both BP2 and BP3 (Fig. 1).

BP2 was known to be localized between S542 and S543 (15); however, the distance between these markers was unknown. The identification of multiple ESTs (SGC32610, SHGC15126, SHGC17218, A006B10 and A008B26), potentially representing multiple genes/pseudogenes, within both the BP2 and BP3 regions suggested the presence of a large duplicated genomic segment. The other duplicated STSs included S15, S16 and S17, which could represent the multicopy probes mentioned in the early study of Donlon *et al*. (22), MN7, known to be present in multiple copies on 15q11–q13 (24–26), and two YAC ends, 368H3L and 166G7R. Recently, the term ‘duplicon’ has been used to describe duplication of genomic segments containing non-processed genes (30). Although the characterization of the multiple ESTs in this region is limited, we will refer to this duplicated genomic segment on 15q11–q13 as a duplicon.

The initial order of markers within the YAC contigs at each duplicon copy suggested an inverted orientation of BP2 relative to BP3 (Fig. 1). For example, S17 mapped to y254B5 and y166G7 at the distal end of BP2, but did not map to y764C6 at the distal end of BP3 (Fig. 1). To confirm the orientation of BP2 relative to BP3, S17 was used to identify bacterial artificial chromosome (BAC) clones GS124B5 and GS246N13 (Table 1). STS mapping of these BACs indicated that GS124B5 showed positive PCR results for 166G7R, S17 and S543, placing this clone at the distal boundary of BP2 (Fig. 1). BAC GS246N13 showed positive PCR results for S17 and S931 (exon 3 of the *P* gene), placing this clone at the proximal boundary of BP3, confirming the inverted orientation of BP2 relative to BP3 (Fig. 1).

Duplicon copy number in normal controls

To determine the copy number of the duplicon within chromosome 15, FISH was performed using P1-derived artificial chromo-

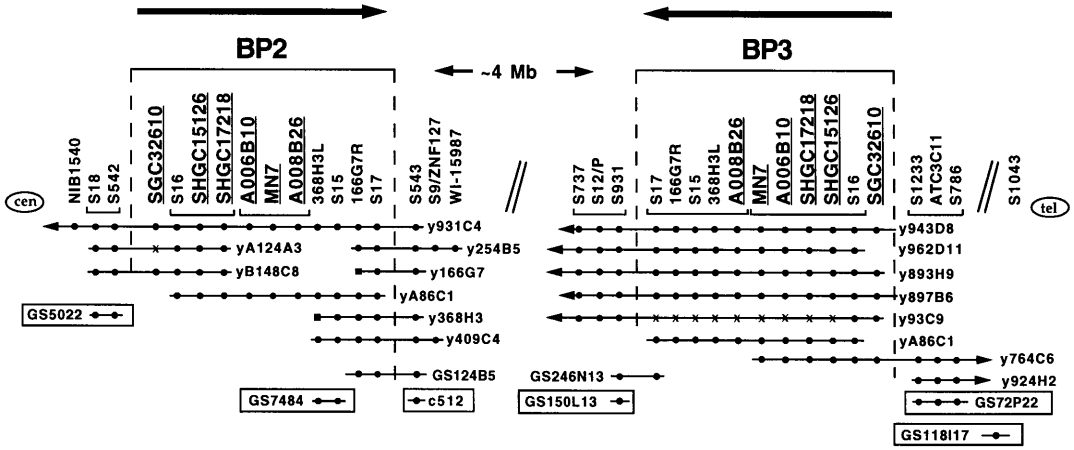

Figure 1. STS content map of BP2 and BP3 duplicons. The region is oriented from centromere (left) to telomere (right). The large arrows at the top of the figure indicate the orientation of the two duplicons relative to each other. STSs are listed vertically with ESTs/genes underlined. Below the STSs are horizontal lines which represent individual genomic clones. YAC clones begin with y, BAC/PAC clones with GS and the cosmid with c. A filled circle represents a positive PCR reaction for the STS on the clone indicated; an x represents a negative PCR reaction and a filled square represents an STS developed from a clone end. Rectangles enclose clones utilized in FISH experiments.

Table 1. BAC/PAC clones used in study

STS	Clone ID	Clone type	FISH signal	Location	Size (kb)
D15S18	GS5022	PAC	Single	15q11–q13	65
SGC32610	GS52I3	BAC	Multiple	15q11–q13	145
D15S16	GS82G5	BAC	Multiple	15q11–q13	110
D15S16	GS179A12	BAC	Multiple	15q11–q13	110
D15S16	GS160F18	BAC	Multiple	15q11–q13	65
SGC14317	GS179K9	BAC	Multiple	15q11–q13	170
A008B26	GS229K19	BAC	Multiple	15q11–q13	110
A008B26	GS98H3	BAC	Multiple	15q11–q13	150
A008B26	GS52G8	BAC	Multiple	15q11–q13	140
7484-SP6	GS253A21	BAC	Multiple	15q11–q13 + 16p11.1–p11.2 + 15q24	160
368H3L	GS7484	PAC	Multiple	15q11–q13 + 16p11.1–p11.2 + 15q24	85
368H3L	GS7483	PAC	Multiple	15q11–q13	65
S17	GS124B5	BAC	Multiple	15q11–q13 + 16p13 + 15q24	180
7483-SP6	GS119F13	BAC	Multiple	15q11–q13	180
S931	GS150L13	BAC	Single	15q11–q13	50
S17	GS246N13	BAC	Single	15q11–q13	115
S1233	GS72P22	BAC	Single	15q11–q13	190
S1043	GS118I17	BAC	Single	15q11–q13	160

some (PAC) GS7484, identified with STS 368H3L (Fig. 1 and Table 1). Hybridization of GS7484 to normal lymphocyte cultures demonstrated a large cluster of signals within the 15q11–q13 region on metaphase chromosomes (Fig. 2A). Weak signals on 15q24 and 16p11.1–p11.2 were also observed. In interphase nuclei there were three to five prominent signals in each chromosome 15 domain (Fig. 2A, inset). To confirm the copy number of GS7484 within 15q11–q13, hybridization was performed on G_0/G_1 synchronized nuclei from two normal fibroblast lines. In both samples the modal value was three signals with a range of one to seven signals in each chromosome 15 domain, consistent with the results observed in lymphocytes (data not shown).

Loss of duplicon copies in PWS/AS deletion patients

Previous microsatellite analysis using markers S542 and S543 localized the site of breakage observed in Class II PWS and AS deletions to the same site as the BP2 duplicon (15). To further assess the association of the duplicons to PWS/AS deletions, GS7484 was used as a FISH probe on two Class I and two Class II deletion patients. Co-hybridization was performed using clone GS118I17, a single copy probe isolated with S1043, a marker located distal to BP3 (Fig. 1 and Table 1). The signal from GS7484 is significantly reduced in intensity on the deleted chromosome 15 in two Class I deletion patients (Fig. 2B). In interphase nuclei, multiple signals (mode = 4) are seen on the normal chromosome, but only one faint

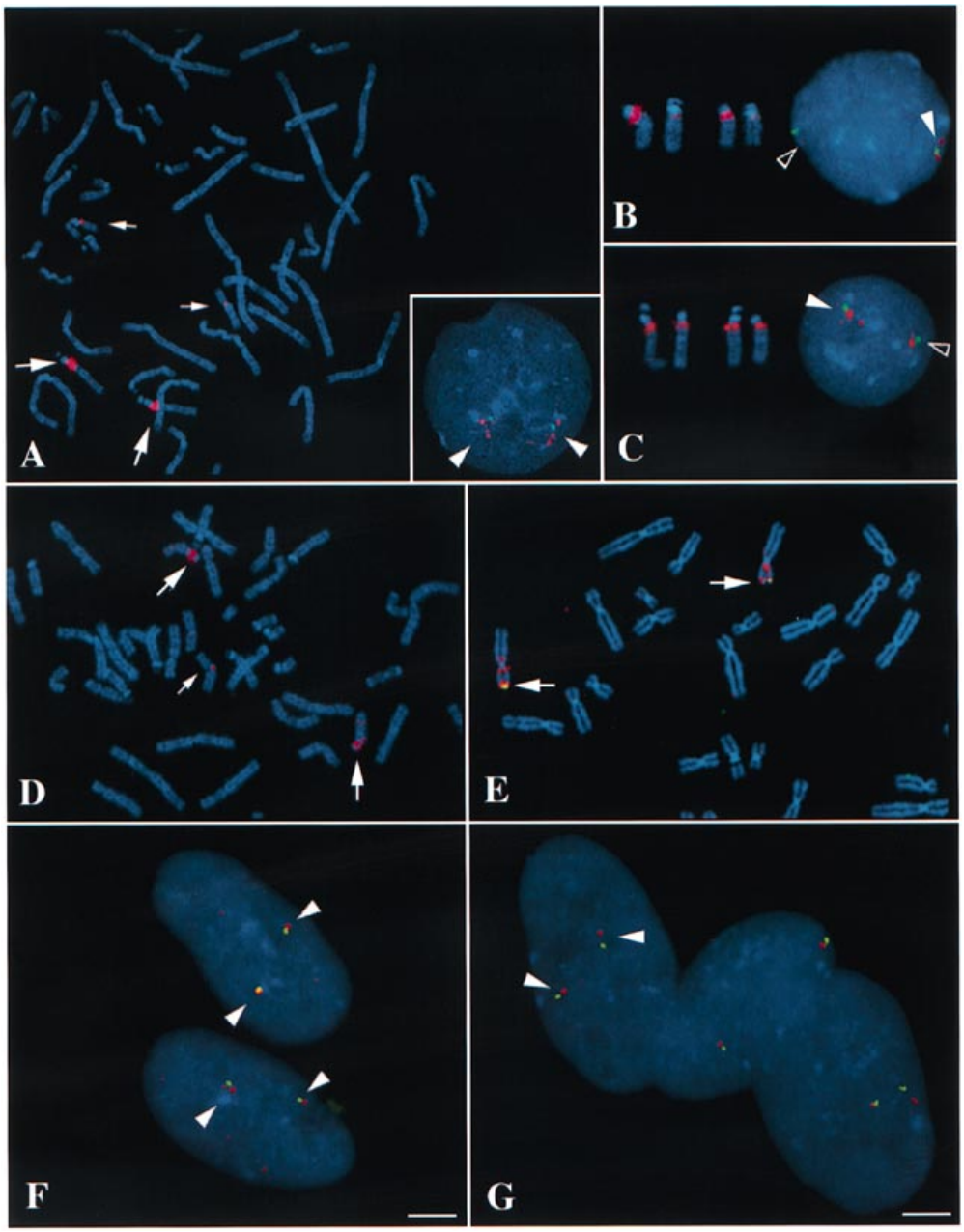

Figure 2. (**A**) FISH analysis on normal metaphase and interphase cells indicating duplicated sequences. GS7484 is labeled with digoxigenin and detected with rhodamine anti-digoxigenin (red), and pMC15 (D15Z3), a chromosome 15 α satellite probe, is labeled with biotin and detected with avidin–FITC (green). In metaphase chromosomes, GS7484 shows a large cluster of red signals at the 15q11–q13 region (large arrows) with a small signal at 16p11.1–p11.2 (small arrows). Very weak hybridization is also observed on 15q24. In interphase nuclei (inset) approximately three to five discreet red signals are observed (arrowheads). (**B**) FISH analysis on a Class I AS deletion patient. GS7484 is labeled in red and BAC GS118I17, located distal to BP3, is labeled in green. For clarity, the green signals from GS118I17 are only shown in interphase nuclei. In metaphase, GS7484 shows one homolog (left) with a large cluster of red signals and markedly reduced signal intensity in the second homolog (right). In interphase nuclei, GS7484 shows multiple red signals for one chromosome 15 (filled arrowhead) with only a single signal on the deleted homolog (open arrowhead). (**C**) FISH analysis on a Class II PWS deletion patient. The probes were labeled as in (b). GS7484 shows a significant cluster of red signals on both chromosomes 15; however, the homolog with prominent satellites (left) shows a stronger signal than that on the right. In interphase nuclei, GS7484 shows multiple red hybridization signals on one chromosome 15 (filled arrowhead) and a reduced number of red signals on the deleted homolog (open arrowhead). (**D**) Evolutionary conservation of duplicons in other primates. Hybridization of GS7484 (red) to metaphase chromosomes from chimpanzee (*P.troglodytes*) shows multiple red signals in the pericentromeric region of PTR16 (orthologous to human 15q11–q13) (large arrows) and weaker signals on distal PTR16 (orthologous to distal human 15q). An additional weak signal is observed near the centromere of PTR17 (orthologous to HSA16p11.2) (small arrow). (**E**) Hybridization of GS7484 (red) to metaphase chromosomes from crab-eating macaque (*M.fascicularis*) indicates multiple red signals on MFA7 (orthologous to HSA15). The single yellow signal at the telomere of MFA7 (arrows) indicates co-hybridization of PAC GS158H23 (green), a single copy probe containing UBE3A, with a red signal from GS7484. Therefore, this telomeric region is orthologous to HSA15q11–q13. Two additional red signals are observed on either side of the centromere of MFA7, the orthologous regions to distal 15q. (**F**) Interphase FISH distance measurements of BP2 using unique flanking clones. Cosmid 512 (red) and GS5022 (green) were hybridized to serum-starved G_0/G_1 fibroblast nuclei (arrowheads). Cosmid 512 also cross-hybridizes with 16p13.1, producing two additional signals in each interphase nucleus. From interphase distance measurements, the genomic distance is estimated to be ~250 kb between these two probes. The bar represents 5 μm. (**G**) Interphase FISH distance measurements of BP3 using unique flanking clones. GS150L13 (red) and GS72P22 (green) were hybridized to serum-starved G_0/G_1 fibroblast nuclei (arrowheads). From interphase distance measurements, the genomic distance is estimated to be ~500 kb between these two probes. The bar represents 5 μm.

Table 2. New chromosome 15q11–q13 STSs

Name	Primer sequences	Size (bp)	Genbank accession no.
GS52I3-T7	F: GGTCCTCTATTAATCTGCAGGTCC	210	AF077848
	R: CTCAGCACCCAGGGTTTCCAATG		
GS52I3-SP6	F: CACTTAATCTTCAATCTCTGTACT	171	AF077849
	R: CTAACCGCTTTGGGAGGCCGAG		
GS124B5-SP6	F: TCCTAGGAAACCCCCAAATA	147	AF077843
	R: TGCAATAAACATCTAGGTGCAGA		
GS124B5-T7	F: CCCAGGTCAACTTCTGCAAC	162	AF077842
	R: TTTCATAAACAGCAGCAATTTCA		
GS229K19-SP6	F: TGGTGGCTTCCTGATTCTTC	288	AF077845
	R: TACAGGGCAACCCTTTTCTG		
GS229K19-T7	F: TGGAAATGGCTTGGAAAAAC	145	AF077846
	R: TGATGAACAGTCAAGGCACA		
GS7483-T7	F: GTGAAGGTGAAATTTAAATACC	194	AF077852
	R: CAAACTCCTCACCTCAAGTAATCC		
GS7483-SP6	F: CCAAGATTGGGACAGTGAGACAGG	228	AF077851
	R: CAATTGCAATCCGCTGTTTCTG		
GS7484-T7	F: CCACCTGGGGGTCCACATGCAG	242	AF077854
	R: GCAGGGGTGGGCTCTAGGAGCG		
GS7484-SP6	F: CAACGACAAGACTGCTGCCAGAG	197	AF077853
	R: GTAAAGGATCGGGGTAGCCCCAG		

F, forward; R, reverse.

signal is present on the deleted chromosome consistent with loss of most GS7484 signals on the deleted homolog (Fig. 2B). In contrast, several signals from GS7484 remain on the deleted chromosome in Class II deletion patients (Fig. 2C), consistent with the presence of at least one additional copy of GS7484 remaining in a Class II deletion that is absent in Class I. Additionally, GS118I17 showed co-localization with GS7484 on metaphase chromosomes (data not shown), but a distinct signal on interphase nuclei (Fig. 2B and C), consistent with a close, but non-overlapping, localization of GS118I17 relative to GS7484.

Duplication events present in non-human primates

To begin to assess the evolutionary origin of the genomic duplications, FISH analysis using GS7484 was performed on metaphase preparations from two other great ape species [the chimpanzee, *Pan troglodytes* (PTR), and *Gorilla gorilla* (GGO)] and an Old World monkey [the crab-eating macaque, *Macaca fascicularis* (MFA)]. Multiple signals were seen on the chromosomal regions orthologous to human chromosomes 15 and 16 (HSA15 and HSA16) in chimpanzee (Fig. 2D), gorilla (data not shown) and macaque (Fig. 2E).

In the chimpanzee most of the GS7484 signals are near the pericentric region of PTR16, orthologous to HSA15q11–q13 (Fig. 2D). Additionally, weak cross-hybridization was seen at distal PTR16q and at the centromere of PTR18, orthologous to HSA16p11 (31,32), consistent with the signals observed in human (Fig. 2D). One distinguishing feature between human and *P.troglodytes* is a pericentric inversion on PTR16 (31–33), which places most of the signal from GS7484 at PTR16p11–p12. It would be interesting to speculate that this inversion event was facilitated by these duplicons subsequent to the divergence of chimpanzee and human.

In the gorilla the orthologous chromosomes to HSA15 and HSA16, GGO15 and GGO17, respectively, are known to have a similar banding and chromosomal painting pattern to human (31,32). FISH analysis in the gorilla using GS7484 showed an identical pattern of hybridization to that observed in human (data not shown).

The orthologous regions between human and crab-eating macaque have been well characterized, although some differences are present (34). For example, the orthologous chromosome to HSA15 in the macaque, MFA7, has been shown by chromosome painting to have homology with both HSA14 and HSA15 (34). FISH using GS7484 and GS158H23, a control probe containing UBE3A (35), shows strong hybridization to the MFA7 telomere, orthologous to HSA15q11–q13, as well as two sites corresponding to distal HSA15q (Fig. 2E). Weaker signals were also observed on MFA20, orthologous to HSA16p11, plus several other sites. The presence of multiple copies of GS7484 at the orthologous regions in *M.fascicularis* for HSA15q11–q13 is suggestive that these genomic duplication events may have occurred ~20–25 million years ago, prior to the divergence of Old World monkeys (Cercopithecidae) from great apes (Hominoidea) (36).

Size estimate of the duplicon by interphase FISH analysis

Although multiple STSs had been mapped to both BP2 and BP3, the sizes of the duplicons could not be estimated from the YAC contig data (Fig. 1). Therefore, initial size estimates of both BP2 and BP3 were obtained by interphase distance measurements between probes flanking each breakpoint region. For the BP2 region, GS5022 and cosmid 512 (c512) (18), both anchored to STSs outside the duplicated region (Fig. 1), were hybridized to G_0/G_1 synchronized fibroblast nuclei and the dis-

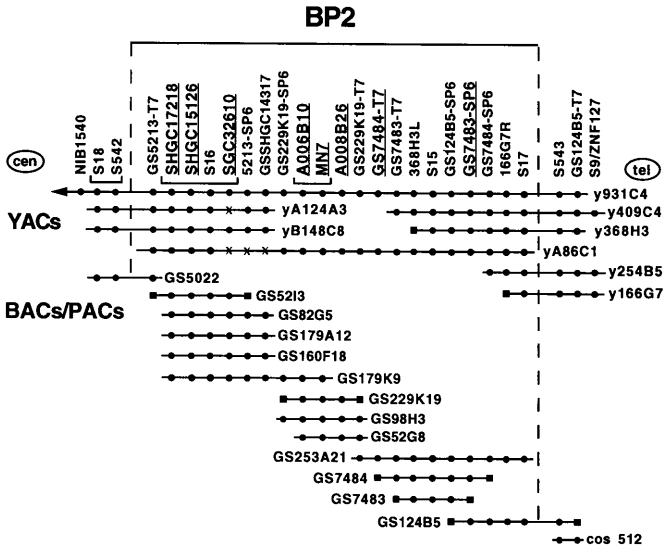

Figure 3. STS content map of a 'composite contig' of BAC and PAC clones across BP2. Below the STSs are horizontal lines representing individual genomic clones (YACs/BACs/PAC and cosmids). A filled circle represents a positive PCR reaction for the STS on the clone indicated; an x represents a negative PCR reaction and a small filled square represents an STS developed from a clone end. Underlined STSs represent ESTs/genes present in the duplicon.

tance between probe signals was measured (Fig. 2F). Although c512 cross-hybridizes weakly with 16p13.1, the chromosome 15 signals were easily distinguished (Fig. 2F). Based on these measurements, BP2 was estimated to be ~250 kb in size.

To provide a maximum size estimate of BP3, GS150L13, anchored on the proximal side to S931, and GS72P22, located distal to BP3 (Table 1 and Fig. 1), were used as the flanking probes. Each BAC gave a single hybridization signal on normal metaphase chromosomes and analysis of GS150L13 on PWS/AS deletion patients confirmed a location proximal to BP3 (Table 1 and data not shown). The two BACs appear to co-localize on metaphase chromosomes (data not shown) but in interphase nuclei are clearly separated (Fig. 2G). From these measurements, the maximum size of BP3 was estimated to be ~500 kb. These interphase distance measurements suggested that the duplicon at BP3 may be twice as large as the duplicon at BP2. However, the presence of a gap between GS72P22 and the distal end of BP3 could not be excluded.

Construction and characterization of a BAC/PAC contig across BP2

Although the YAC contigs were anchored for both BP2 and BP3, the order of markers within the regions relative to each other could not be determined unambiguously (Fig. 1). Therefore, a BAC/PAC contig was constructed for BP2 to fine map the duplicons. Initial screenings of a 2–3× BAC library yielded significant over-representation of positive clones for all duplicated STSs analyzed, consistent with multiple copies of these regions being present within the human genome. STS content mapping using PCR was performed to identify overlapping clones and BAC ends were sequenced to create new STSs to complete a 'composite contig' (Fig. 3 and Table 2). As seen by the presence of multiple copies of GS7484 by FISH (Fig. 2A), any of these individual clones could originate from BP2, BP3 or other copies of the duplicon (Fig. 3 and Table 2). Efforts to assign clones uniquely to BP2 or BP3 on the basis of nucleotide divergence are described below. A complete list of new BACs, PACs and STSs identified are presented in Tables 1 and 2. The proximal boundary for BP2 is contained within PAC GS5022, identified using S18 (Fig. 3), and the distal boundary is contained within BAC GS124B5, identified using S17 (Fig. 3). The presence of multiple BAC and PAC end STSs anchored to YACs at both ends of BP2 provides a high level of confidence in the overall ordering of STSs within this region.

Refined size estimate of the BP2 duplicon

With the completion of a BAC/PAC contig across BP2, an independent size estimate of the duplicon could be made based on the size of a minimal tiling path of BAC/PAC clones across the region (Table 1). The non-overlapping BACs GS179K9 (160 kb) and GS253A21 (160 kb) cover the BP2 region with three gaps (Fig. 3), providing a minimum size estimate of 320 kb. Additionally, the gap between GS179K9 and GS253A21 is contained within a completely sequenced PAC clone, pDJ778A2 (GenBank accession no. AC004583). The 85 kb size of the region between MN7 (located within GS179K9) and GS253A21-T7, together with the 320 kb estimate above, provides a minimal size estimate of the duplicon of ~400 kb. The interphase FISH estimate (~250 kb) is, therefore, likely to be an underestimate of the true size of the duplicon, possibly due to chromosome condensation properties near the pericentromeric region.

Table 3. EST/genes identified for BP2/BP3

Name	IMAGE no.	dbEST no.	Size (bp)	GenBank accession no.	BLAST results
ESTs					
SGC32610	241447	H90407	958	AF129926	Unknown transcript
SHGC-15126	511845	AA088777	1016	AF129929	MYLE
SHGC17218	346304	W74131	1369	AF129927	Unknown transcript
A006B10	321824	W33186	873	AF129930	KIAA0393, MN7
A008B26	120151	T95087	799	AF129928	Unknown transcript
PAC ends					
GS7484-T7				AA478019	BEM-1/BUD5 suppressor-like protein
GS7483-SP6				U18237	ATP-binding cassette protein

Table 4. STS sequence variation

Locus	GenBank accession no.	Dx nt	Location	Clones
S15	AF017567	G^{165},T^{176}		
S15(S1)		A^{165},C^{176}	BP2	y931C4, y409C4, y368H3
S15(S2)		A^{165},T^{176}	BP3A	y962D11, y893H9, yA86C1, GS7484, GS253A21
S15(S3)		G^{165},T^{176}	BP3B	y962D11, y893H9, yA86C1, GS7483
S16	AF017568	$A^{68},A^{74},C^{160},G^{162},C^{168},T^{253},C^{303}$		
S16(S1)		$G^{68},A^{74},T^{160},T^{162},G^{168},C^{253},C^{303}$	BP2	yA124A3, yB148C8, GS179K9,
		$(G^{68},A^{74},T^{160},T^{162},G^{168},C^{253},T^{303})$[a]		(y931C4)[a]
S16(S2)		$A^{68},A^{74},C^{160},G^{162},C^{168},T^{253},C^{303}$	BP3A	y962D11, y893H9, y93C9, yA86C1, yA162B8, GS160F18, GS119F13
S16(S3)		$G^{68},\_^{74},T^{160},T^{162},G^{168},C^{253},C^{303}$	BP3B	y764C6, GS52I3, GS82G5, GS179A12
S17	AF017569	G^{81}		
S17(S1)		T^{81}	BP2	y409C4, y368H3, y254B5, y166G7, GS124B5
S17(S2)		G^{81}	BP3A/3B	y962D11, y893H9, y93C9, yA86C1, GS253A21, GS246N13, GS119F13, y931C4
MN7	X69635	A^{147},T^{175},C^{195} [b]		
(D15F37)				
MN7(S3)[c]		G^{147},C^{175},C^{195}	BP2	y931C4, GS179K9
MN7(S1)[c]		A^{147},T^{175},T^{195}	BP3A	y962D11, y893H9, yA86C1
MN7(S2)[c]		A^{147},T^{175},C^{195}	BP3B	y943D8, y962D11, y893H9, yA86C1, yA162B8, y764C6, GS229K19, GS98H3, GS52G8

Dx nt., diagnostic nucleotide which corresponds to the base number in the GenBank sequence.
[a]Probable polymorphism of position 303 in y931C4.
[b]Base numbers from GenBank sequence; represents the reverse sequence in Buiting *et al*. (25).
[c]MN7 sequences are numbered according to Buiting *et al*. (25).

Sequence analysis of cDNAs present within the duplicon

Five ESTs were initially mapped to the duplicon (Fig. 3). To determine whether any of these ESTs corresponds to the same gene, partial cDNA clones (800–1400 bp), representing the 3'-ends of the transcripts, were identified using the Unigene database (37) and sequenced (Table 3). BLAST2 comparisons were then performed on all possible pairs of sequence to detect any overlaps. Each sequence was found to be unique, suggesting that five different genes/pseudogenes may be represented by these ESTs.

BLASTN searches were then performed to determine whether these partial cDNA sequences matched any known genes. The sequence of IMAGE clone 321824, containing EST A006B10, identified MN7 and the cDNA clone KIAA0393 (28), indicating that they represent different ESTs for the same gene (Table 3). However, *HERC2*, the 15.0 kb MN7-containing transcript, was not identified. A BLASTN search of the 6.7 kb KIAA0393 sequence did identify *HERC2* (26) and its mouse homolog named *rjs* (27) or *herc2* (26), indicating that clone 321824 was specific for the 3'-end of KIAA0393. Additionally, a BAC clone (CIT987-SKA-17E1; GenBank accession no. AC002041), sequenced from chromosome 16p11.2, was found to contain part of KIAA0393 (bases 59639–87 066), S15 (bases 17449–17773, 94% identity) and GS7483-T7 (bases 36259–36507, 93% identity). The presence of ~65 kb of genomic sequence at both the 15q11–q13 duplicon and the pericentromeric region of 16p11 is consistent with the FISH data indicating duplication of GS7484 to 16p11.1.

A BLASTN search of IMAGE clone 511845, containing the sequence for SHGC15126, identified an unpublished gene named MYLE (GenBank accession no. AF108145, bases 65–1053, 99% identity) and S16 (GenBank accession no. AF017568, bases 1–332, 94% identity). BLASTN searches for three clones representing SHGC17218, SGC32610 and A008B26 indicated that these three cDNAs represent different, previously unknown transcripts (Table 3).

High sequence homology to two additional genes was identified using the end sequences of PACs GS7483 and GS7484 (GS7484-T7 and GS7483-SP6; Table 3). GS7484-T7 showed 90% identity to an uncharacterized gene located within the λ immunoglobulin light chain locus on 22q11 [GenBank accession no. D88269, bases 5515–5644 (38)] with homology to an EST representing a BEM-1/BUD5 suppressor-like protein (GenBank accession no. AA478019). GS7483-SP6 showed 92% identity to a human ATP-binding cassette protein mRNA localized to chromosome 1q42 (GenBank accession no. U18237) (39). Further characterization of these ESTs/cDNAs will be required to determine which of these may represent functional genes. In addition, polymorphisms will need to be identified to discriminate possible expression of a functional gene from more than one duplicon copy.

Two copies of the duplicon present in BP3

The 'composite contig' created across BP2 (Fig. 3) did not allow unambiguous localization of the 11 BAC/PAC clones contained completely within the duplicon. Only clones GS5022 and GS124B5, on the proximal and distal boundaries of BP2, respectively, could be uniquely mapped (Fig. 3). To determine if sufficient sequence divergence existed between BP2 and BP3 to allow precise mapping of the other 11 clones, sequencing of four STSs (S15, S16,

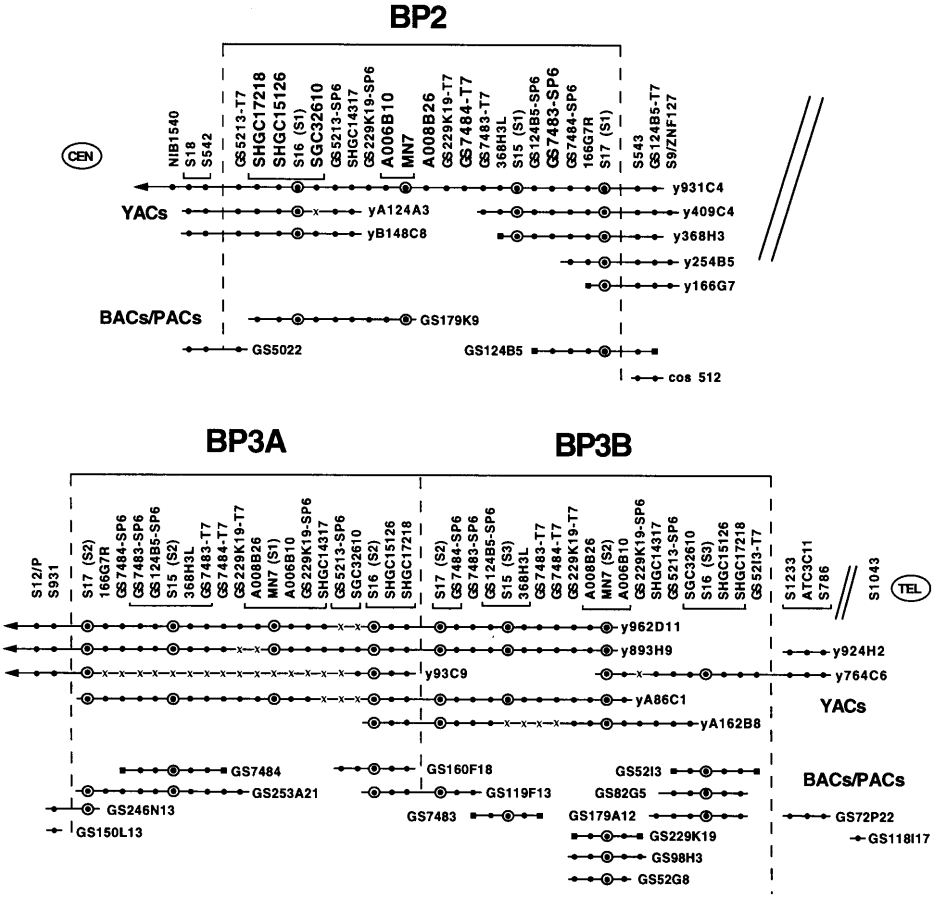

Figure 4. STS content map, including sequenced STSs, for BP2, BP3A and BP3B. Below the STSs are horizontal lines representing individual clones (YACs/BACs/PAC and cosmids). A filled circle represents a positive PCR reaction between the STS and the clone indicated and an x represents a negative PCR reaction. A small filled square represents an STS developed from a clone end and large open circles indicate STSs that have been sequenced. Note that only one BAC clone (GS179K9) from the 'composite contig' (Fig. 3) originates from the BP2 region, while multiple clones originate from BP3A and BP3B, including GS119F13, which spans the two copies within BP3. YAC y93C9 is the clone shown to contain BP3 by FISH (13) and yA86C1 and yA168B8 were identified using MN7 (24).

S17 and MN7) was performed on YACs anchored to either BP2 or BP3. All four STSs showed sequence divergence ranging from one to six nucleotides in the BP2 and BP3 YACs (Table 4). The presence of two different sequences within BP3 YACs for three STSs (S15, S16 and MN7) suggested the presence of two different copies of the duplicon at BP3. In previous studies, two different sequences had been identified for MN7 in yA86C1 and y962D11 (24,25); however, the distance between them had not been determined. Additionally, the unique sequences of S16 present in YACs y962D11 and y764C6 indicated that these YACs did not overlap each other, providing further evidence for the presence of two copies of the duplicon within BP3 (Figs 1 and 4).

To fine map the BACs to specific BP regions, at least one of the four STSs (S15, S16, S17 or MN7) was sequenced for all BAC/PACs (Table 4). This analysis showed that only clone GS179K9 originated from BP2, whereas the other 10 clones mapped to two different sites within the BP3 region (Fig. 4 and Table 4). Although several YACs showed the presence of two copies of S15 and MN7 (y962D11, y893H9 and A86C1), the BAC/PAC clones showed only one copy of any individual STS, indicating separation of the duplicated STSs by a distance greater than the size of a single BAC clone. Sequencing of S16 and S17 in GS119F13, a clone which did not map to BP2, indicated a location which crosses the boundary between two copies of the duplicon (Fig. 4). The presence of two sequences for S15 and MN7 in BP3 YACs with single sequences in individual BAC clones confirms the presence of two copies of the duplicon within BP3, which we have termed BP3A and BP3B. Linkage of the two BP3 duplicons in a head-to-tail manner in GS119F13 indicates a direct (tandem) orientation for BP3A and BP3B.

DISCUSSION

Until recently, relatively few data were available on the molecular mechanisms of recurring constitutional chromosome rearrangements in humans. Detailed physical maps of human chromosomes have led to recognition of the importance of genome organization or architecture in the mechanisms of some human genetic diseases (40). The presence of low copy number, chromosome-specific repeated DNA sequences provides an opportunity for homologous recombination events leading to chromosomal rearrangements which have been referred to as 'genomic disorders' (40), to distinguish them from the more common mutational mechanisms associated with most Mendelian genetic diseases.

There are now several examples of recurring chromosomal abnormalities, including deletions, duplications and inversions, mediated by multiple copies of a DNA element with high sequence homology. Most often, misalignment of these homologous sequences within a chromosomal region leads to unequal meiotic exchange. The products of these homologous recombination events will be dependent on: (i) the orientation of the DNA elements; (ii) the type of DNA strand exchange; and (iii) whether the rearrangements are inter- or intrachromosomal. The DNA elements that mediate these rearrangements may be copies of related genes or pseudogenes (e.g. globin), common interspersed repeats (e.g. Alu or LINE elements) or other chromosome-specific low copy repeats (40–42).

The now classic paradigm for chromosome-specific repeats mediating unequal homologous recombination is Charcot–Marie–Tooth syndrome type 1A (CMT1A) and hereditary neuropathy with liability to pressure palsies (HNPP) (43). CMT1A and HNPP result from reciprocal duplication and deletion events, respectively, involving a 1.5 Mb region on 17p11.2 flanked by two copies of a 24 kb repeat sequence (CMT1A-REP) arranged in a direct orientation (43,44). Interchromosomal misalignment and crossing-over within the repeats during meiosis I produces these reciprocal duplication and deletion events. The breakpoints have been fine mapped to a 557 bp region within the CMT1A-REPs containing a 456 bp region of perfect sequence identity. This is thought to represent a minimum efficient processing segment, required for homologous recombination and is located adjacent to a *mariner* transposon-like element (45). A similar example of chromosome-specific low copy number repeats predisposing to chromosomal rearrangement is that of X-linked ichthyosis due to steroid sulfatase (SS) deficiency. Approximately 90% of these patients have a 1.9 Mb deletion on Xp22.3, mediated by VNTR-like repeats flanking the deletion interval and resulting in complete deletion of the *SS* gene (46,47).

In contrast, crossing-over of duplicated DNA sequences that are in an inverted orientation mediates intrachromosomal inversions. There are three such examples of duplicated sequences on distal Xq which mediate inversions and are associated with genetic disease. In hemophilia A, ~45% of mutations are due to an inversion event involving two copies of 'gene A'. One copy of this gene is located within the factor VIII gene, while two additional copies are located ~500 kb distal on Xq (48). A foldback pairing mechanism for the telomeric portion of Xq has been proposed as an intermediate structure which is followed by a cross-over event between two 'gene A' copies (48). This creates a paracentric inversion, which disrupts the factor VIII gene by separating exons 1–22 from exons 23–26 (48). Similarly, the gene deficient in Hunter syndrome, iduronate 2-sulfatase (*IDS*), is disrupted in ~13% of patients due to a paracentric inversion event involving an *IDS* pseudogene (*IDS-2*), located 20 kb distal in an inverted orientation (49). The *emerin* gene, responsible for Emery–Dreifuss muscular dystrophy, is flanked by two copies of an 11.3 kb repeat in an inverted orientation (50). This has led to a common inversion polymorphism, in which ~33% of normal females are heterozygous for a paracentric inversion. More rarely, complete deletion of the *emerin* gene has been identified due to a more complicated rearrangement in this region (50,51).

Intrachromosomal crossing-over of the inverted duplicons on chromosome 15 would predict a paracentric inversion event similar to that observed on Xq28. Although such inversions have not been reported, they would be difficult to detect by standard G-banded chromosome analysis and a systematic search for an inverted order of DNA markers by FISH may be necessary to address this issue. In contrast, interchromosomal crossing-over of these inverted duplicons would predict formation of an inv dup(15) plus an acentric fragment, so alternative DNA strand exchange events are needed to mediate the interstitial deletions observed in 15q11–q13.

One model for an intrachromosomal deletion event between inverted duplicons would require a foldback or stem–loop intermediate structure, with excision or deletion of the intervening 'looped-out' DNA segment. The best characterized example of this type of rearrangement is V(D)J recombination observed with the immunoglobulin heavy and light chain genes (52). Additionally, several examples of deletions in human have been proposed as being mediated by such stem–loop structure intermediates. In autosomal recessive juvenile nephronophthisis (NPH), a homozygous deletion of an ~250 kb region on chromosome 2q13 is the most frequent mutation mechanism observed. The deleted segment is flanked by two copies of an ~100 kb duplicated region arranged in an inverted orientation (53). Overall, 80% of familial cases and 65% of sporadic cases of NPH showed the presence of a homozygous deletion of this region (53). In familial hypercholesterolemia, deletions of the LDL receptor have been reported in which the rearrangement was mediated by Alu elements arranged in an inverted orientation (54).

Support for intrachromosomal homologous recombination mediating some deletions in PWS/AS comes from two studies. In one study, the grandparental origin of the deletion in seven PWS families was determined using polymorphic markers flanking the deletion (55). Recombination between the grandparental alleles flanking the site of deletion was observed in five cases (suggestive of interchromosomal events), but two cases showed no recombination (consistent with an intrachromosomal event) (55). In a similar study, analysis of the grandparental origin in nine PWS/AS deletion patients demonstrated five cases with recombination between grandparental alleles and two cases with no recombination (9). Together, four cases are consistent with an intrachromosomal deletion event, while 10 cases are suggestive of an interchromosomal exchange. However, in these latter 10 cases, one cannot distinguish an interchromosomal deletion associated with recombination from an intrachromosomal deletion after recombination.

An interchromosomal deletion mechanism would also predict formation of the reciprocal duplication product, as seen with CMT1A/HNPP (43). Two cases of interstitial duplications on 15q11–q13 have been observed which involve BP3, consistent with this prediction (7). However, interstitial duplications have also been identified which involve BP4 (9), indicating multiple mechanisms producing these rearrangements. The high frequency of PWS/AS deletions [~1/10 000 (2,3)] and the rarity of reported cases of interstitial duplication 15 (5–9) suggests that intrachromosomal mechanisms may be a common cause of deletions of 15q11–q13. However, lack of ascertainment of duplication 15 cases may also contribute to the significantly smaller number of cases reported.

Currently, seven genes/pseudogenes, have been identified which are present in both the proximal (BP2) and distal (BP3A/BP3B) regions. Three cDNAs (IMAGE 120151, 241447 and 346304) represent unknown transcripts, whereas two others (IMAGE 321824 and 511845) are highly similar to the MYLE

and KIAA0393 mRNA sequences, respectively. Two PAC ends show high sequence identity to a BEM-1/BUD5 suppressor-like protein (GS7484-T7) and an ATP-binding cassette protein (GS7483-SP6), located on other chromosomes, and may represent pseudogenes in the 15q11–q13 duplicons.

One gene has been identified within the duplicon termed *HERC2* (26) or *ERY-1* (25). A006B10 is contained within the sequence for KIAA0393, the 7.0 kb transcript containing MN7 (28), but not the *HERC2* gene (26), indicating that the 7.0 and 15.0 kb MN7-containing transcripts may represent alternatively spliced products of the same gene. The mouse homolog for *HERC2*, termed *rjs* (27) or *herc2* (26), has recently been identified and maps to a single location on mouse chromosome 7 orthologous to BP3A (28), suggesting that BP3A is the ancestral duplicon in human. The cDNA sequence data together with the BLASTN analysis of all STSs demonstrates the presence of multiple genes/pseudogenes in these duplicons with their functional status yet to be determined.

Recent physical mapping data for several microdeletion syndromes observed in human has demonstrated the involvement of large duplicated genomic segments at the common deletion breakpoints (40). This includes Smith–Magenis syndrome (SMS) at 17p11.2, Williams syndrome (WS) at 7q11.2 and DiGeorge (DGS)/velo-cardio-facial syndrome (VCFS) at 22q11.2. The best characterized of these is the ~5 Mb deletion associated with SMS. Located centromeric to the CMT1A/HNPP deletion/duplication region, the SMS deletion region is flanked by duplicons termed SMS-REPs (56). Partial cosmid contigs have been developed across the SMS-REPs, estimated to be >200 kb in size (40,56), but the orientation of the duplicons is currently unknown (56). Four genes, *SRP*, *TRE*, *KER* and *CLP*, are located within both the proximal and distal SMS-REPs and analysis of SMS deletion patients using the *CLP* cDNA probe identified a 1.2 Mb novel junction fragment in 29 of 31 patients analyzed (56).

Williams syndrome is associated with deletions of 7q11, including the elastin and *LIMK1* genes, in >90% of patients (57–59). Two duplicated genes, *PMS2L* and *GTF2I*, are both present in the regions flanking the WS deletion region, suggesting that duplicons may be involved in these rearrangements (60–62). A novel >3 Mb junction fragment was detected in WS deletion patients using a cDNA probe for IB291, located within the duplicated flanking regions (62).

Two deletions of either 3.0 or 1.5 Mb on 22q11.2 are associated with DGS/VCFS (63). The proximal breakpoint region is the same in both deletions; however, two different distal breakpoints are observed. Preliminary characterization of the large deletion has identified ~300 kb duplicons arranged in tandem at the sites of rearrangement and containing at least five genes/pseudogenes (64). Interestingly, the supernumerary marker 22 chromosomes, observed in the cat eye syndrome, involve the same breakpoint regions observed in the DGS/VCFS deletions (65), possibly indicating a similar mechanism of formation as the small and large supernumerary marker chromosomes derived from 15q11–q13 (17).

The fact that these four common microdeletion syndromes are all located in relatively close proximity to the pericentromeric region of the chromosome raises interesting questions regarding the mechanism of duplication of these large genomic segments. Evidence is emerging that the pericentromeric regions, located adjacent to the α satellite arrays of the centromeres, contain paralogous copies of chromosomal regions duplicated and translocated from other locations within the human genome (30). The pericentromeric region of chromosome 15 is known to contain partial copies of the immunoglobulin heavy chain V and D segments (66,67), *NF1* pseudogenes (68–70) and *GABRA5* pseudogenes (71). Additionally, while a single copy of MN7 is present in mouse, located in a region syntenic with the BP3A site, multiple copies are present in human near the pericentromeric regions on both 15q11.2 and 16p11.2 (24). The microsatellite D15S543, located just distal to BP2, is also present on 16p13, indicating a second duplication event between chromosome 16 and 15q11.2. These data suggest that duplication of large genomic regions, near pericentromeric regions, can create genomic instability, which may predispose to further chromosomal rearrangements.

MATERIALS AND METHODS

Genomic clones

YACs used in this study are derived from the CEPH Mark I, CEPH Mark II or St Louis libraries and were previously mapped by STS typing by Christian *et al.* (29). YAC yA162B8 was identified using STS MN7 (24). YACs were acquired from either the Baylor College of Medicine Genome Center (Houston, TX), the National Human Genome Research Institute (NIH, Bethesda, MD), Research Genetics (Huntsville, AL) or CEPH (Paris, France).

PAC GS5022 was identified by screening a total human P1 library using STS D15S18 (Genome Systems, St Louis, MO). This clone was previously published by Huang *et al*. (18) as clone 770c6. PACs GS7483 and GS7484 were isolated from the same P1 library using STS 368H3L. Cosmid 512 (c512) was identified using STS S543 as previously described (18). All BACs were identified by PCR screening of a total human BAC library (Genome Systems), using the STSs listed in Table 1. DNA from genomic clones was isolated using an AutoGen 740 (Integrated Separations Systems, Natick, MA). Sizing of BAC/PAC clones was performed by digesting 1 µg DNA with *Not*I for 3 h and separating on a 1.0% SeaKem GTG agarose gel (FMC Bioproducts, Rockland, ME) using the Bio-Rad Chef Mapper apparatus (Bio-Rad, Hercules, CA) with an auto-algorithm mode ranging from 5 to 300 kb at 14°C. The gels were stained with ethidium bromide, visualized on a UVP GDS8000 gel documentation system (UVP, Upland, CA) and sizes determined by extrapolation with known size markers.

Chromosome 15 STSs

New STSs were developed by sequencing the ends of BACs GS52I3, GS124B5 and GS229K19 and PACs GS7483 and GS7484 (Table 2). Briefly, BAC DNA was isolated using the AutoGen 740 and purified using Microcon 100 microconcentrators according to the manufacturer's protocol (Amicon, Beverly, MA). One microgram of BAC DNA and 40 pmol of primers were used for sequencing with the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction kit and analyzed using the ABI 377 automated DNA sequencer (Applied Biosystems, Foster City, CA). STS primers were designed using the Primer3 program from the Whitehead Institute, Massachusetts Institute of Technology (MIT; http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi). The complete sequences for the BAC ends have been submitted

to GenBank with the accession numbers listed in Table 2. All other STSs have been described previously (29).

STS content mapping using PCR

Ten microliter reactions containing 1.0 µl Perkin Elmer 10× buffer I, 200 µM dNTP mix, 0.5 µM primers, 0.5 U Amplitaq Gold (Perkin Elmer, Foster City, CA) and 5–25 ng template DNA were analyzed for each STS against all YACs, BACs and PACs mapping to these regions. PCR was performed and analyzed as previously described (29).

FISH

Chromosome preparations were made from peripheral blood or lymphoblastoid cell lines from normal controls and PWS/AS deletion patients using standard methods. In addition, metaphase preparations were made from lymphoblastoid cell lines from a common chimpanzee (*P.troglodytes*) and gorilla (*G.gorilla*), as well as a fibroblast cell line from a crab-eating macaque (*M.fascicularis*) (GM03446; Coriell Institute for Medical Research, Camden, NJ). Probe labeling, DNA hybridization and antibody detection were carried out using methods described previously (72). FISH slides were analyzed using a Zeiss Axiophot microscope with filters to detect DAPI, FITC and rhodamine separately as well as a triple bandpass filter (no. 83000; Chroma Technology, Brattleboro, VT) to detect signals simultaneously. Images were collected and merged using a cooled CCD camera (KAF 1400; Photometrics, Tucson, AZ) and IP Lab Spectrum (Signal Analytics, Vienna, VA) or Quips mFISH software (Vysis, Downer's Grove, IL). At least 10 metaphase cells were analyzed for each probe to verify probe location and to detect any cross-hybridization to other chromosomes. D15Z3 (pMC15), an α satellite probe for chromosome 15 (73), was added to the hybridization solution to aid in identification. If a probe appeared to give multiple signals on chromosome 15, additional interphase nuclei were examined.

Interphase FISH distance measurements

For genomic distances in the range 50 kb–2 Mb, the physical distance between a pair of probes in interphase (i.e. interphase distance or ID) is related to the genomic distance (74). Probes labeled with either biotin or digoxigenin were hybridized in pairs to G_0/G_1 synchronized fibroblast nuclei and the distance between probe signals was measured using the measure length command contained in the IP Lab Spectrum v.3 software package (Signal Analytics). The measure length command was calibrated using images from a Zeiss stage micrometer and at least 60 interphase distances were measured for each pair of probes. Genomic distance was estimated from a calibration curve generated in our laboratory using cosmids from 4p16.3, which were previously used to generate the relationship between the mean interphase distance squared and genomic distance (75).

Sequencing of genomic clones

PCR reactions using four STSs (S15, S16, S17 and MN7) were used to amplify DNA from YAC, BAC and PAC clones. The reactions were performed in 100 µl reactions by scaling up the conditions indicated above. The PCR products were purified using Microcon 30 or Microcon 100 filters according to the manufacturer's protocol (Amicon). The samples were resuspended to a final volume of 100 µl, and 5 µl were used for visualization on a 1.0% agarose gel to check for the presence of a single band. PCR products were quantitated and then sequenced using the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction kit and the ABI 377 automated DNA sequencer (Applied Biosystems).

Sequencing of cDNA clones

Clones corresponding to ESTs SHGC17218 (IMAGE 346304), SHGC15126 (IMAGE 511845), SGC32610 (IMAGE 241447), A006B10 (IMAGE 321824) and A008B26 (IMAGE 120151) were acquired from the American Type Culture Collection (Rockville, MD) or Research Genetics (Huntsville, AL). DNA was isolated from 3 ml cultures of cDNA clones 241447 and 321824 using the AutoGen 740 in a total volume of 50 µl. The DNA was purified and concentrated using Microcon 100 filters according to the manufacturer's protocol (Amicon). Aliquots of 1.5 µl of DNA concentrate and 40 pmol of T3 or T7 primers were used for sequencing each clone end using the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction kit and the ABI 377 automated DNA sequencer. Sequence analysis using Sequencher 3.1 (Gene Codes, Ann Arbor, MI) indicated no overlap in the sequence, so new primers were designed for the 3'-end of the sequence using the Primer3 program from MIT. A second round of sequencing was performed and Sequencher 3.1 analysis confirmed completion of the single pass sequencing. Single pass sequencing was performed on clones 120151, 346304 and 511845 by a commercial source (SeqWright, Houston, TX).

ABBREVIATIONS

AS, Angelman syndrome; BAC, bacterial artificial chromosome; BP, breakpoint; DGS, DiGeorge syndrome; EST, expressed sequence tag; FISH, fluorescence *in situ* hybridization; IDS, iduronate 2-sulfatase; NPH, nephronophthisis; PAC, P1-derived artificial chromosome; PWS, Prader–Willi syndrome; RFLP, restriction fragment length polymorphism; SMS, Smith–Magenis syndrome; SS, steroid sulfatase; STS, sequence tagged site; VCFS, velo-cardio-facial syndrome; WS, Williams syndrome; YAC, yeast artificial chromosome.

ACKNOWLEDGEMENTS

We would like to thank Julie Kuc for expert technical assistance. The 4p16.3 cosmids were generously provided by B. Trask and M. McDonald. This work was supported in part by NIH grant HD36111.

REFERENCES

1. Ledbetter, D.H. and Ballabio, A. (1995) Molecular cytogenetics of contiguous gene syndromes: mechanisms and consequences of gene dosage imbalance. In Scriver, C.R., Beaudet, A.L., Sly, W.S. and Valle, D. (eds) *The Metabolic and Molecular Bases of Inherited Disease*. McGraw-Hill, New York, NY, Vol. 1, pp. 811–839.
2. Nicholls, R.D., Saitoh, S. and Horsthemke, B. (1998) Imprinting in Prader–Willi and Angelman syndromes. *Trends Genet.*, **14**, 194–200.
3. Jiang, Y., Tsai, T.F., Bressler, J. and Beaudet, A.L. (1998) Imprinting in Angelman and Prader–Willi syndromes. *Curr. Opin. Genet. Dev.*, **8**, 334–342.
4. Webb, T. (1994) Inv dup(15) supernumerary marker chromosomes. *J. Med. Genet.*, **31**, 585–591.

5. Pettigrew, A.L., Gollin, S.M., Greenberg, F., Riccardi, V.M. and Ledbetter, D.H. (1987) Duplication of proximal 15q as a cause of Prader–Willi syndrome. *Am. J. Med. Genet.*, **28**, 791–802.
6. Clayton-Smith, J., Webb, T., Cheng, S.J., Pembrey, M.E. and Malcolm, S. (1993) 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.
7. Repetto, G.M., White, L.M., Bader, P.J., Johnson, D. and Knoll, J.H. (1998) Interstitial duplications of chromosome region 15q11q13: clinical and molecular characterization. *Am. J. Med. Genet.*, **79**, 82–89.
8. Browne, C.E., Dennis, N.R., Maher, E., Long, F.L., Nicholson, J.C., Sillibourne, J. and Barber, J.C. (1997) Inherited interstitial duplications of proximal 15q: genotype–phenotype correlations. *Am. J. Hum. Genet.*, **61**, 1342–1352.
9. Robinson, W.P., Dutly, F., Nicholls, R.D., Bernasconi, F., Penaherrera, M., Michaelis, R.C., Abeliovich, D. and Schinzel, A.A. (1998) The mechanisms involved in formation of deletions and duplications of 15q11–q13. *J. Med. Genet.*, **35**, 130–136.
10. Schinzel, A.A., Brecevic, L., Bernasconi, F., Binkert, F., Berthet, F., Wuilloud, A. and Robinson, W.P. (1994) Intrachromosomal triplication of 15q11–q13. *J. Med. Genet.*, **31**, 798–803.
11. Long, F.L., Duckett, D.P., Billam, L.J., Williams, D.K. and Crolla, J.A. (1998) Triplication of 15q11–q13 with inv dup(15) in a female with developmental delay. *J. Med. Genet.*, **35**, 425–428.
12. Knoll, J.H., Nicholls, R.D., Magenis, R.E., Glatt, K., Graham, J.M.Jr, Kaplan, L. and Lalande, M. (1990) Angelman syndrome: three molecular classes identified with chromosome 15q11q13-specific DNA markers. *Am. J. Hum. Genet.*, **47**, 149–155.
13. Kuwano, A., Mutirangura, A., Dittrich, B., Buiting, K., Horsthemke, B., Saitoh, S., Niikawa, N., Ledbetter, S.A., Greenberg, F., Chinault, A.C. and Ledbetter, D.H. (1992) Molecular dissection of the Prader–Willi/Angelman syndrome region (15q11–13) by YAC cloning and FISH analysis. *Hum. Mol. Genet.*, **1**, 417–425.
14. Mutirangura, A., Jayakumar, A., Sutcliffe, J.S., Nakao, M., McKinney, M.J., Buiting, K., Horsthemke, B., Beaudet, A.L., Chinault, A.C. and Ledbetter, D.H. (1993) A complete YAC contig of the Prader–Willi/Angelman chromosome region (15q11–q13) and refined localization of the SNRPN gene. *Genomics*, **18**, 546–552.
15. Christian, S.L., Robinson, W.P., Huang, B., Mutirangura, A., Line, M.R., Nakao, M., Surti, U., Chakravarti, A. and Ledbetter, D.H. (1995) Molecular characterization of two proximal deletion breakpoint regions in both Prader–Willi and Angelman syndrome patients. *Am. J. Hum. Genet.*, **57**, 40–48.
16. Leana-Cox, J., Jenkins, L., Palmer, C.G., Plattner, R., Sheppard, L., Flejter, W.L., Zackowski, J., Tsien, F. and Schwartz, S. (1994) Molecular cytogenetic analysis of inv dup(15) chromosomes, using probes specific for the Prader–Willi/Angelman syndrome region: clinical implications. *Am. J. Hum. Genet.*, **54**, 748–756.
17. Crolla, J.A., Harvey, J.F., Sitch, F.L. and Dennis, N.R. (1995) Supernumerary marker 15 chromosomes: a clinical, molecular and FISH approach to diagnosis and prognosis. *Hum. Genet.*, **95**, 161–170.
18. Huang, B., Crolla, J.A., Christian, S.L., Wolf-Ledbetter, M.E., Macha, M.E., Papenhausen, P.N. and Ledbetter, D.H. (1997) Refined molecular characterization of the breakpoints in small inv dup(15) chromosomes. *Hum. Genet.*, **99**, 7–11.
19. Wandstrat, A.E., Leana-Cox, J., Jenkins, L. and 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.
20. Flejter, W.L., Bennett-Baker, P.E., Ghaziuddin, M., McDonald, M., Sheldon, S. and Gorski, J.L. (1996) Cytogenetic and molecular analysis of inv dup(15) chromosomes observed in two patients with autistic disorder and mental retardation. *Am. J. Med. Genet.*, **61**, 182–187.
21. Cook, E.H.Jr, Lindgren, V., Leventhal, B.L., Courchesne, R., Lincoln, A., Shulman, C., Lord, C. and Courchesne, E. (1997) Autism or atypical autism in maternally but not paternally derived proximal 15q duplication. *Am. J. Hum. Genet.*, **60**, 928–934.
22. Donlon, T.A., Lalande, M., Wyman, A., Bruns, G. and Latt, S.A. (1986) Isolation of molecular probes associated with the chromosome 15 instability in the Prader–Willi syndrome. *Proc. Natl Acad. Sci. USA*, **83**, 4408–4412.
23. Tantravahi, U., Nicholls, R.D., Stroh, H., Ringer, S., Neve, R.L., Kaplan, L., Wharton, R., Wurster-Hill, D., Graham, J.M.Jr, Cantu, E.S., Frias, J.L., Kousseff, B.G. and Latt, S.A. (1989) Quantitative calibration and use of DNA probes for investigating chromosome abnormalities in the Prader–Willi syndrome. *Am. J. Med. Genet.*, **33**, 78–87.
24. Buiting, K., Greger, V., Brownstein, B.H., Mohr, R.M., Voiculescu, I., Winterpacht, A., Zabel, B. and Horsthemke, B. (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.
25. Buiting, K., Gross, S., Ji, Y., Senger, G., Nicholls, R.D. and 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.
26. Ji, Y., Walkowicz, M.J., Buiting, K., Johnson, D.K., Tarvin, R.E., Rinchik, E.M., Horsthemke, B., Stubbs, L. and Nicholls, R.D. (1999) The ancestral gene for transcribed, low-copy repeats in the Prader–Willi/Angelman region encodes a large protein implicated in protein trafficking, which is deficient in mice with neuromuscular and spermiogenic abnormalities. *Hum. Mol. Genet.*, **3**, 533–542.
27. Lehman, A.L., Nakatsu, Y., Ching, A., Bronson, R.T., Oakey, R.J., Keiper-Hrynko, N., Finger, J.N., Durham-Pierre, D., Horton, D.B., Newton, J.M., Lyon, M.F. and Brilliant, M.H. (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.
28. Nagase, T., Ishikawa, K., Nakajima, D., Ohira, M., Seki, N., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N. and Ohara, O. (1997) Prediction of the coding sequences of unidentified human genes VII. The complete sequences of 100 new cDNA clones from brain which can code for large proteins *in vitro*. *DNA Res.*, **4**, 141–150.
29. Christian, S.L., Bhatt, N.K., Martin, S.A., Sutcliffe, J.S., Kubota, T., Huang, B., Mutirangura, A., Chinault, A.C., Beaudet, A.L. and Ledbetter, D.H. (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.
30. Eichler, E.E. (1998) Masquerading repeats: paralogous pitfalls of the human genome. *Genome Res.*, **8**, 758–762.
31. Yunis, J.J. and Prakash, O. (1982) The origin of man: a chromosomal pictorial legacy. *Science*, **215**, 1525–1530.
32. Jauch, A., Wienberg, J., Stanyon, R., Arnold, N., Tofanelli, S., Ishida, T. and Cremer, T. (1992) Reconstruction of genomic rearrangements in great apes and gibbons by chromosome painting. *Proc. Natl Acad. Sci. USA*, **89**, 8611–8615.
33. Luke, S. and Verma, R.A. (1995) The genomic sequence for Prader–Willi/Angelman syndromes' loci of human is apparently conserved in the great apes. *J. Mol. Evol.*, **41**, 250–252.
34. Wienberg, J., Stanyon, R., Jauch, A. and Cremer, T. (1992) Homologies in human and *Macaca fuscata* chromosomes revealed by *in situ* suppression hybridization with human chromosome specific DNA libraries. *Chromosoma*, **101**, 265–270.
35. Sutcliffe, J.S., Jiang, Y.H., Galijaard, R.J., Matsuura, T., Fang, P., Kubota, T., Christian, S.L., Bressler, J., Cattanach, B., Ledbetter, D.H. and Beaudet, A.L. (1997) The E6-Ap ubiquitin-protein ligase (*UBE3A*) gene is localized within a narrowed Angelman syndrome critical region. *Genome Res.*, **7**, 368–377.
36. Kumar, S. and Hedges, B.S. (1998) A molecular timescale for vertebrate evolution. *Nature*, **392**, 917–920.
37. Schuler, G.D., Boguski, M.S., Stewart, E.A., Stein, L.D., Gyapay, G., Rice, K., White, R.E., Rodriguez-Tome, P., Aggarwal, A., Bajorek, E., Bentolila, S., Birren, B.B., Butler, A., Castle, A.B., Chiannilkulchai, N., Chu, A., Clee, C., Cowles, S., Day, P.J., Dibling, T., Drouot, N., Dunham, I., Duprat, S., East, C., Hudson, T.J. *et al.* (1996) A gene map of the human genome. *Science*, **274**, 540–546.
38. Kawasaki, K., Minoshima, S., Nakato, E., Shibuya, K., Shintani, A., Schmeits, J.L., Wang, J. and Shimizu, N. (1997) One-megabase sequence analysis of the human immunoglobulin lambda gene locus. *Genome Res.*, **7**, 250–261.
39. Allikmets, R., Gerrard, B., Glavac, D., Ravnik-Glavac, M., Jenkins, N.A., Gilbert, D.J., Copeland, N.G., Modi, W. and Dean, M. (1995) Characterization and mapping of three new mammalian ATP-binding transporter genes from an EST database. *Mamm. Genome*, **6**, 114–117.
40. Lupski, J. (1998) Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet.*, **14**, 417–422.
41. Cooper, D., Krawczak, M. and Antonarakis, S. (1995) The nature and mechanisms of human gene mutation. In Scriver, C.R., Beaudet, A.L., Sly, W.S. and Valle, D. (eds), *The Metabolic and Molecular Bases of Inherited Disease.* McGraw-Hill, New York, NY, Vol. 1, pp. 259–291.
42. Purandare, S.M. and Patel, P.I. (1997) Recombination hot spots and human disease. *Genome Res.*, **7**, 773–786.
43. Lupski, J.R. (1998) Charcot–Marie–Tooth disease: lessons in genetic mechanisms. *Mol. Med.*, **4**, 3–11.

44. Pentao, L., Wise, C.A., Chinault, A.C., Patel, P.I. and Lupski, J.R. (1992) Charcot–Marie–Tooth type 1A duplication appears to arise from recombination at repeat sequences flanking the 1.5 Mb monomer unit. *Nature Genet.*, **2**, 292–300.
45. Reiter, L.T., Hastings, P.J., Nelis, E., De Jonghe, P., Van Broeckhoven, D. and Lupski, J.R. (1998) Human meiotic recombination products revealed by sequencing a hotspot for homologous strand exchange in multiple HNPP deletion patients. *Am. J. Hum. Genet.*, **62**, 1023–1033.
46. Ballabio, A., Bardoni, B., Guioli, S., Basler, E. and Camerino, G. (1990) Two families of low-copy-number repeats are interspersed on Xp22.3: implications for the high frequency of deletions in this region. *Genomics*, **8**, 263–270.
47. Yen, P.H., Li, X.M., Tsai, S.P., Johnson, C., Mohandas, T. and Shapiro, L.J. (1990) Frequent deletions of the human X chromosome distal short arm result from recombination between low copy repetitive elements. *Cell*, **61**, 603–610.
48. Lakich, D., Kazazian, H.H.Jr, Antonarakis, S.E. and Gitschier, J. (1993) Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. *Nature Genet.*, **5**, 236–241.
49. Bondeson, M.L., Dahl, N., Malmgren, H., Kleijer, W.J., Tonnesen, T., Carlberg, B.M. and Pettersson, U. (1995) Inversion of the IDS gene resulting from recombination with IDS-related sequences is a common cause of the Hunter syndrome. *Hum. Mol. Genet.*, **4**, 615–621.
50. Small, K., Iber, J. and Warren, S.T. (1997) Emerin deletion reveals a common X-chromosome inversion mediated by inverted repeats. *Nature Genet.*, **16**, 96–99.
51. Small, K. and Warren, S.T. (1998) Emerin deletions occurring on both Xq28 inversion backgrounds. *Hum. Mol. Genet.*, **7**, 135–139.
52. Lewis, S.M. and Wu, G.E. (1997) The origin of V(D)J recombination. *Cell*, **88**, 159–162.
53. Konrad, M., Saunier, S., Heidet, L., Silbermann, F., Benessy, F., Calado, J., Le Paslier, D., Broyer, M., Gubler, M.C. and Antignac, C. (1996) Large homozygous deletions of the 2q13 region are a major cause of juvenile nephronophthisis. *Hum. Mol. Genet.*, **5**, 367–371.
54. Lehrman, M.A., Schneider, W.J., Sudhof, T.C., Brown, M.S., Goldstein, J.L. and Russell, D.W. (1985) Mutation in LDL receptor: Alu–Alu recombination deletes exons encoding transmembrane and cytoplasmic domains. *Science*, **227**, 140–146.
55. Carrozzo, R., Rossi, E., Christian, S.L., Kittikamron, K., Livieri, C., Corrias, A., Pucci, L., Fois, A., Simi, P., Bosio, L., Beccaria, L., Zuffardi, O. and Ledbetter, D.H. (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.
56. Chen, K.S., Manian, P., Koeuth, T., Potocki, L., Zhao, Q., Chinault, A.C., Lee, C.C. and Lupski, J.R. (1997) Homologous recombination of a flanking repeat gene cluster is a mechanism for a common contiguous gene deletion syndrome. *Nature Genet.*, **17**, 154–163.
57. Lowery, M.C., Morris, C.A., Ewart, A., Brothman, L.J., Zhu, X.L., Leonard, C.O., Carey, J.C., Keating, M. and Brothman, A.R. (1995) Strong correlation of elastin deletions, detected by FISH, with Williams syndrome: evaluation of 235 patients. *Am. J. Hum. Genet.*, **57**, 49–53.
58. Nickerson, E., Greenberg, F., Keating, M.T., McCaskill, C. and Shaffer, L.G. (1995) Deletions of the elastin gene at 7q11.23 occur in approximately 90% of patients with Williams syndrome. *Am. J. Hum. Genet.*, **56**, 1156–1161.
59. Budarf, M.L. and Emanuel, B.S. (1997) Progress in the autosomal segmental aneusomy syndromes (SASs): single or multi-locus disorders? *Hum. Mol. Genet.*, **6**, 1657–1665.
60. Osbourne, L.R., Herbrick, J.A., Greavette, T., Heng, H.H., Tsui, L.C. and Scherer, S.W. (1997) PMS2-related genes flank the rearrangement breakpoints associated with Williams syndrome and other diseases on human chromosome 7. *Genomics*, **45**, 402–406.
61. Perez Jurado, L.A., Peoples, R., Kaplan, P., Hamel, B.C. and Francke, U. (1996) Molecular definition of the chromosome 7 deletion in Williams syndrome and parent-of-origin effects on growth. *Am. J. Hum. Genet.*, **59**, 781–792.
62. Perez Jurado, L.A., Wang, Y.K., Peoples, R., Coloma, A., Cruces, J. and 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.
63. Morrow, B., Goldberg, R., Carlson, C., Das Gupta, R., Sirotkin, H., Collins, J., Dunham, I., O'Donnell, H., Scambler, P., Shprintzen, R. and Kucherlapati, R. (1995) Molecular definition of the 22q11 deletion in velo-cardio-facial syndrome. *Am. J. Hum. Genet.*, **56**, 1391–1403.
64. Edelmann, L., Pandita, R.K., McCain, N., Goldberg, R., Shprintzen, R. and Morrow, B. (1998) The VCFS common deletion occurs within two tandem gene clusters on 22q11. *Am. J. Hum. Genet.*, **63** (suppl.), A54.
65. McTaggert, K.E., Budarf, M.L., Driscoll, D.A., Emanuel, B.S., Ferreira, P. and McDermid, H.E. (1998) Cat eye syndrome chromosome breakpoint clustering: identification of two intervals also associated with 22q11 deletion syndrome breakpoints. *Cytogenet. Cell Genet.*, **81**, 222–228.
66. Tomlinson, I.M., Cook, G.P., Carter, N.P., Elaswarapu, R., Smith, S., Walter, G., Buluwela, L., Rabbitts, T.H. and Winter, G. (1994) Human immunoglobulin VH and D segments on chromosomes 15q11.2 and 16p11.2. *Hum. Mol. Genet.*, **3**, 853–860.
67. Nagaoka, H., Ozawa, K., Matsuda, F., Hayashida, H., Matsumura, R., Haino, M., Shin, E.K., Fukita, Y., Imai, T., Anand, R., Yokoyama, K., Eki, T., Soeda, E. and Honjo, T. (1994) Recent translocation of variable and diversity segments of the human immunoglobulin heavy chain from chromosome 14 to chromosomes 15 and 16. *Genomics*, **22**, 189–197.
68. Purandare, S.M., Huntsman Breidenbach, H., Li, Y., Zhu, X.L., Sawada, S., Neil, S.M., Brothman, A., White, R., Cawthon, R. and Viskochil, D. (1995) Identification of neurofibromatosis 1 (NF1) homologous loci by direct sequencing, fluorescence *in situ* hybridization and PCR amplification of somatic cell hybrids. *Genomics*, **30**, 476–485.
69. Regnier, V., Meddeb, M., Lecointre, G., Richard, F., Duverger, A., Nguyen, V.C., Dutrillaux, B., Bernheim, A. and Danglot, G. (1997) Emergence and scattering of multiple neurofibromatosis (NF1)-related sequences during hominoid evolution suggest a process of pericentromeric interchromosomal transposition. *Hum. Mol. Genet.*, **6**, 9–16.
70. Kehrer-Sawatzki, H., Schwickardt, T., Assum, G., Rocchi, M. and Krone, W. (1997) A third neurofibromatosis type 1 (NF1) pseudogene at chromosome 15q11.2. *Hum. Genet.*, **100**, 595–600.
71. Ritchie, R.J., Mattei, M.-G. and Lalande, M. (1998) A large polymorphic repeat in the pericentromeric region of human chromosome 15q contains three partial gene duplications. *Hum. Mol. Genet.*, **7**, 1253–1260.
72. Chong, S.S., Pack, S.D., Roschke, A.V., Tanigami, A., Carrozzo, R., Smith, A.C., Dobyns, W.B. and Ledbetter, D.H. (1997) A revision of the lissencephaly and Miller–Dieker syndrome critical regions in chromosome 17p13.3. *Hum. Mol. Genet.*, **6**, 147–155.
73. Finelli, P., Antonacci, R., Marzella, R., Lonoce, A., Archidiacono, N. and Rocchi, M. (1996) Structural organization of multiple alphoid subsets coexisting on human chromosomes 1, 4, 5, 7, 9, 15, 18 and 19. *Genomics*, **38**, 325–330.
74. Trask, B.J., Allen, S., Massa, H., Fertitta, A., Sachs, R., van den Engh, G. and Wu, M. (1993) Studies of metaphase and interphase chromosomes using fluorescence *in situ* hybridization. *Cold Spring Harbor Symp. Quant. Biol.*, **58**, 767–775.
75. van den Engh, G., Sachs, R. and Trask, B.J. (1992) Estimating genomic distance from DNA sequence location in cell nuclei by a random walk model. *Science*, **257**, 1410–1412.