Table 3 Continued

Case	Gender	Clinical diagnosis	Remarkable clinical features	CNV	Position	WGA-4500[b]	FISH[b]	Base position and size of the identified CNV[a]						Parental analysis	Protein-coding genes[c]	CNV assessment[d]	Corresponding or candidate gene(s)
								Start (max)	Start (min)	End (min)	End (max)	Size (min)	Size (max)				
27	F	MCA/MR		del	7q22.1q22.2	arr cgh 7q22.1q22.2 (RP11-10D8→ RP11-72J24)x1	ish del(7)(q22.1q22.2) (RP11-124G15+,RP11-188E1-,RP11-95P19-)	97 314 215	98 261 079	105 604 920	106 451 506	7 343 841	9 137 291		135	P	
28	F	MCA/MR	Epilepsy	del	12q13.13	arr cgh 12q13.13 (RP11-74I8→ RP11-624J6)x1	ish del(12)(q13.13) (RP11-624J6-)	50 987 232	51 016 427	51 956 291	52 180 088	939 864	1 192 856		44	P	
29	M	MCA/MR		dup	16q22.3	arr cgh 16q22.3 (RP11-90L19→ RP11-89K4)x3	ish dup(16)(q22.3) (RP11-115E3++, RP11-90L19++)	70 355 260	70 848 592	72 328 913	73 785 124	1 480 321	3 429 864		25	P	
30	M	RTS susp.		dup	16q24.1	arr cgh 16q24.1 (RP11-140K16→ RP11-442O1)x3	ish dup(16)(q24.1) (RP11-770B4++, RP11-140K16++)	82 699 729	82 797 548	83 749 375	84 123 857	951 827	1 424 128		16	P	
31	M	MCA/MR	Epilepsy	del	2q24.2q24.3	arr cgh 2q24.2 (RP11-89L13→ RP11-79L13)x1	ish del(2)(q24.2) (RP11-638N12-)	160 407 234	161 072 815	162 883 584	166 923 475	1 810 769	6 516 241		28	P	*TBR1*
32	M	MCA/MR		del	3p26.2	arr cgh 3p26.2 (RP11-32F23)x1	ish del(3)(p26.2) (RP11-32F23-)	3 943 353	4 016 797	4 198 468	4 329 970	181 671	386 617		2	P	*SUMF1*
33	M	MCA/MR	IgA deficiency	del	7q21.11	arr cgh 7q21.11 (RP11-22M18)x1	ish del(7)(q21.11) (RP11-115M2+, RP11-353O4-, RP11-22M18-)	83 597 839	83 601 541	84 549 609	84 788 160	948 068	1 190 321		3	P	*SEMA3A*
34	M	MCA/MR		dup	14q32.2	arr cgh 14q32.2 (RP11-128L1)x3	ish dup(14)(q32.2) (RP11-177F8++)	99 330 486	99 337 358	99 841 558	99 845 472	504 200	514 986		7	P	*EML1, YY1*
35	M	MCA/MR	Epilepsy	dup	16p13.3	arr cgh 16p13.3 (RP11-349I11)x3	ish dup(16)(p13.3) (RP11-349I11++)	4 851 459	5 678 447	5 906 909	6 165 923	228 462	1 314 464		9	P	*A2BP1*
36	M	MCA/MR		dup	Xp22.2p22.13	arr cgh Xp22.2p22.13 (RP11-2K15→ RP11-115I10)x3	not performed (X-tiling array)	16 874 735	16 952 121	17 596 600	17 638 351	644 479	763 616		2	P	
				dup	Xp21.3	arr cgh Xp21.3 (RP11-438J7)x3	not performed (X-tiling array)	28 704 076	28 704 076	28 868 075	28 868 075	163 999	163 999		1	P	*IL1RAPL1*
37	F	MCA/MR		del	1p34.3	arr cgh 1p34.3 (RP11-89N10→ RP11-416A14)x1	ish del(1)(p34.2) (RP11-195A8+, RP11-166F21-)dn	37 830 131	38 338 265	39 466 349	39 583 645	1 128 084	1 753 514	*dn*	7	P	
38	M	MCA/MR	Hyper IgE	dup	1q25.2	arr cgh 1q25.2 (RP11-177A2→ RP11-152A16)x3	ish dup(1)(1q25.2) (RP11-177A2++, RP11-152A16++)	177 088 480	177 196 858	177 535 659	177 859 828	338 801	771 348	*dn*	9	P	
39	M	MCA/MR		del	2p24.1p23.3	arr cgh 2p24.1p23.3 (RP11-80H16→ RP11-88F6)x1	ish del(2)(p23.3) (RP11-88F6-, RP11-373D23+)dn	20 037 821	23 094 244	26 815 794	28 414 457	3 721 550	8 376 636	*dn*	86	P	
40	F	MCA/MR	CHD	del	3p26.1p25.3	arr cgh 3p26.1p25.3 (RP11-128A5→ RP11-402P11)x1	ish del(3)(p26.1p25.3) (RP11-936E1-, RP11-402P11-, RP11-1079H21+) dn	8 190 557	8 497 949	9 930 973	10 026 217	1 433 024	1 835 660	*dn*	18	P	

Table 3 Continued

Case	Gender	Clinical diagnosis	Remarkable clinical features	CNV	Position	WGA-4500[b]	FISH[b]	Base position and size of the identified CNV[a]						Parental analysis	Protein-coding genes[c]	CNV assessment[d]	Corresponding or candidate gene(s)
								Start (max)	Start (min)	End (min)	End (max)	Size (min)	Size (max)				
41	M	MCA/MR		del	3p22.1p21.31	arr cgh 3p22.1p21.31 (RP11-241P3→ RP11-88B8)x1	ish del(3)(p22.1) (RP11-61H16+, RP11-241P3-, RP11-78O10+)dn	41 365 663	42 284 365	48 177 538	49 198 542	5 893 173	7 832 879	*dn*	123	P	
42	M	MCA/MR	Corneal opacity	del	3p14.3p14.2	arr cgh 3p14.3p14.2 (RP11-80H18→ RP11-79J9)x1	ish del(3)(p14.2) (RP11-79J19-, RP11-230A22+)mat	57 370 434	58 149 199	58 742 633	58 887 574	593 434	1 517 140	*mat*	11	B	
				del	8q21.11q21.13	arr cgh 8q21.11q21.13 (RP11-225J6→ RP11-214E11)x1	ish del(8) (q21.11q21.13) (RP11-225J6-, RP11-48B3+)dn	75 722 961	75 821 163	81 110 557	81 493 446	5 289 394	5 770 485	*dn*	12	P	
43	M	MCA/MR		del	3q26.31q26.33	arr cgh 3q26.31-q26.33 (RP11-292L5→ RP11-355N16)x1	ish del(3)(q26.32) (RP11-300L9+, RP11-105L6-)dn	175 650 310	176 531 688	180 613 203	181 653 281	4 081 515	6 002 971	*dn*	12	P	
44	M	MCA/MR	CHD	del	13q13.2q13.3	arr cgh 13q13.2 (RP11-269G10→ 90F5)x1	ish del(13)(q13.2) (RP11-142E9+, RP11-381E21-, RP11-98D3+)dn	33 451 136	33 895 560	34 813 379	34 909 905	917 819	1 458 769	*dn*	1	P	
				del	22q11.21	arr cgh 22q11.21 (RP11-155F20→ 54C2)x1	ish del(22)(q11.21) (RP11-155F20-, RP11-590C5-, RP11-54C2-)pat	19 310 307	19 310 307	19 590 642	19 590 642	280 335	280 335	*pat*	15	B	
45	F	aRS		del	18q21.2	arr cgh 18q21.2 (RP11-89B14)x1	ish del(18)(q21.2) (RP11-159D14+, RP11-186B13-, RP11-111C17-)dn	48 218 621	49 166 752	51 288 665	51 861 143	2 121 913	3 642 522	*dn*	9	P	
46	M	MCA/MR		dup	19p13.3	arr cgh 19p13.3 (RP11-49M3→ RP11-268O21)x3		1 095 485	2 418 857	3 499 581	4 460 252	1 080 724	3 364 767	*dn*	113	P	
47	F	MCA/MR	Autism	del	19p13.3	arr cgh 19p13.3 (RP11-30F17→ RP11-330I7)x1	ish del(19)(p13.3) (RP11-330I7-)dn	4 844 383	6 043 505	6 859 584	6 881 792	816 079	2 037 409	*dn*	23	P	
48	M	MCA/MR		del	Xp11.3	arr cgh Xp11.3 (RP11-151G3→ RP11-48J14)x0	ish del(X)(p11.3) (RP11-203D16-)mat	44 403 077	44 433 162	46 795 584	46 795 588	2 362 422	2 392 511	*mat*	18	P	
49	M	MCA/MR		dup	3p26.3	arr cgh 3p26.3 (RP11-63O1)x3	ish dup(3)(p26.3) (RP11-63O1++)pat	2 377 366	2 443 357	2 619 407	2 628 216	176 050	250 850	*pat*	1	B	
50	M	MCA/MR		dup	5p14.3	arr cgh 5p14.3 (RP11-91A5)x3	ish dup(5)(p14.3) (RP11-91A5++)pat	19 046 234	19 485 530	19 656 108	20 798 445	170 578	1 752 211	*pat*	1	B	
51	M	MCA/MR		dup	5q13.3	arr cgh 5q13.1 (RP11-40N8→ RP11-91C10)x3	ish dup(5)(q13.1) (RP11-105A11++)mat	66 417 271	66 481 371	67 501 700	67 838 977	1 020 329	1 421 706	*mat*	3	B	

Table 3 Continued

Case	Gender	Clinical diagnosis	Remarkable clinical features	CNV	Position	WGA-4500[b]	FISH[b]	Base position and size of the identified CNV[a]						Parental analysis	Protein-coding genes[c]	CNV assessment[d]	Corresponding or candidate gene(s)
								Start (max)	Start (min)	End (min)	End (max)	Size (min)	Size (max)				
52	M	MCA/MR		dup	7p22.3	arr cgh 7p22.3 (RP11-23D23)x3	ish dup(7)(p22.3) (RP11-23D23++, RP11-1133D5+)mat	1	954 016	954 584	1 101 944	568	1 101 943	*mat*	12	B	
53	F	MCA/MR		dup	8p23.2	arr cgh 8p23.2 (RP11-79I19→ RP11-89I12)x3	ish dup(8)(p23.2) (RP11-89I19++, RP11-89I12++)pat	3 324 954	3 726 061	4 564 671	5 973 493	838 610	2 648 539	*pat*	1	B	
54	M	MCA/MR		dup	9q33.1	arr cgh 9q33.1 (RP11-150L1)x3	ish dup(9)(q33.1) (RP11-150L1++)pat	118 980 752	119 452 372	119 614 984	120 011 559	162 612	1 030 807	*pat*	2	B	
55	F	MCA/MR		dup	10q22.3	arr cgh 10q22.3 (RP11-79M9)x3	ish dup(10)(q22.3) (RP11-79M9++)mat	77 356 915	77 718 484	77 873 148	78 230 039	154 664	873 124	*mat*	1	B	
56	M	MCA/MR	ELBW, hepato-blastoma	dup	12q21.31	arr cgh 12q21.31 (RP11-91C4)x3	ish dup(12)(q21.31) (RP11-91C4++, RP11-142L2+)pat	80 924 954	82 678 148	82 830 190	85 768 388	152 042	4 843 434	*pat*	3	B	
57	M	GS		del	Xp11.23	arr cgh Xp11.23 (RP11-876B24) x0 mat	not performed (X-tiling array)	47 752 808	47 747 918	47 852 109	47 868 412	104 191	115 604	*mat*	3	B	
58	M	MCA/MR		dup	8q11.23	arr cgh 8q11.23 (RP11-221P7)x3	ish dup(8)(q11.23) (RP11-221P7++, RP11-26P22++)	53 665 974	53 717 675	54 235 229	54 576 654	517 554	910 680		3	VOUS	
59	F	MCA/MR	Micro-cephaly	dup	10q11.21	arr cgh 10q11.21 (RP11-178A10)x3	ish dup(10)(q11.21) (RP11-178A10++)	41 986 946	42 197 693	42 320 775	43 603 027	123 082	1 616 081		15	VOUS	
60	M	MCA/MR		dup	11p14.2p14.1	arr cgh 11p14.2p14.1 (RP11-1L12)x3	ish dup(11) (p14.2p14.1) (RP11-1L12++)	26 723 462	27 033 270	27 213 374	27 445 504	180 104	722 042		4	VOUS	
61	F	MCA/MR		dup	12p11.1	arr cgh 12p11.1 (RP11-88P4)x3	ish dup(12)(p11.1) (RP11-472A10++)	33 333 493	33 359 944	33 572 956	33 572 956	213 012	239 463		2	VOUS	
62	F	aRS		dup	12q21.31	arr cgh 12q21.31 (RP11-91I24→ RP11-91C4)x3	ish dup(12)(q21.31) (RP11-91C4++, RP11-142L2++)	79 949 648	82 172 368	83 968 319	85 768 388	1 795 951	5 818 740		12	VOUS	
63	F	MR	Congenital myopathy	dup	Xq12	arr cgh Xq12 (RP11-90P17→ RP11-383C12)x3	Not performed (X-tiling array)	66 212 661	66 216 353	66 921 699	66 948 538	705 346	735 877		1	VOUS	

Abbreviations: aRS, atyplical Rett syndrome; B, benign; CNV, copy-number variant; *dn*: *de novo* CNV observed in neither of the parents; ELBW, extremely low birth weight; FISH, fluorescence *in situ* hybridization; GS, Gillespie syndrome; *mat*: CNV identified also in mother; P, pathogenic; *pat*: CNV identified also in father; RTS, Rubinstein–Taybi syndrome; SMS, Smith–Magenis syndrome; VOUS, variant of uncertain clinical significance; ZLS, Zimmermann–Laband syndrome.
[a]The sizes were estimated by WGA-4500, X-array, FISH or Agilent Human Genome CGH microarray 244K.
[b]The notation systems is based on ISCN2005.[36]
[c]The number of protein-coding genes contained in the respective CNVs.
[d]The result of CNV assessment.

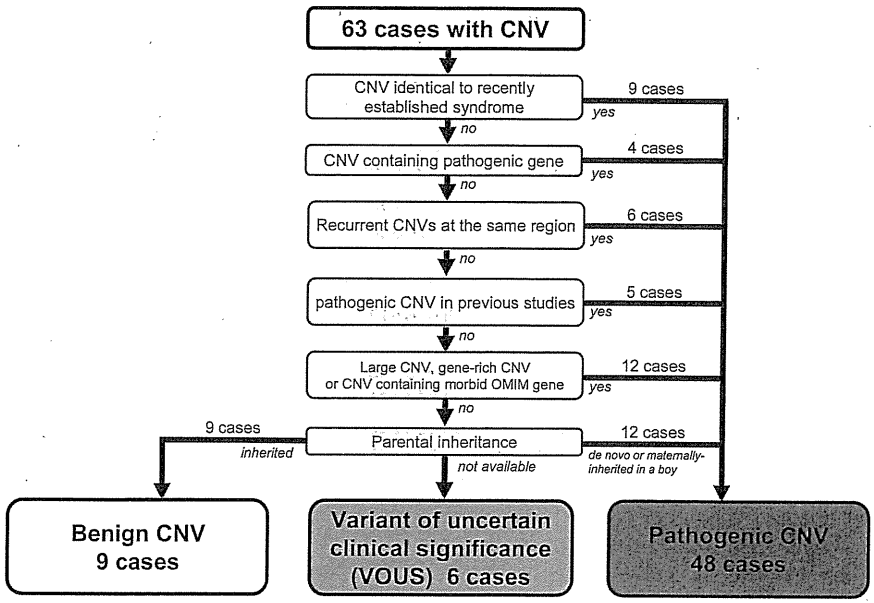

Figure 2 A flowchart of the assessment of CNVs detected in the second screening.

from several aspects. A CNV that contains abundant genes or is large (>3 Mb) has a high possibility to be pathogenic.[21] The CNVs in cases 25–30 probably correspond to such CNVs. Also, we judged a CNV containing a morbid OMIM gene as pathogenic:[21] *TBR1* (OMIM: *604616) in case 31,[56] *SUMF1* (OMIM: *607939) in case 32,[57,58] *SEMA3A* (OMIM: *603961) in case 33,[59] *EML1* (OMIM: *602033) and/or *YY1* (OMIM: *600013) in case 34,[60,61] *A2BP1* (OMIM: *605104) in case 35[62] and *IL1RAPL1* (OMIM: *300206) in case 36.[63] Several previous reports suggest that these genes are likely to be pathogenic, although at present no evidence of a direct association between these genes and phenotypes exists.

CNVs de novo or X maternally inherited. Among the remaining 27 cases, 12 cases had CNVs considered pathogenic as their CNVs were *de novo* (cases 37–47) or inherited del(X)(p11.3) from the mother (case 48). In the second screening we performed FISH for 36 CNVs of the 34 cases whose parental samples were available to confirm that 24 cases had *de novo* CNVs, which were probably pathogenic. A CNV in case 48, a boy with a nullizygous deletion at Xp11.3 inherited from his mother, was also probably relevant to his phenotype (Tables 3 and 4). Meanwhile, although case 57 was a boy with a deletion at Xp11.23 inherited from his mother, he was clinically diagnosed with Gillespie syndrome (OMIM: #206700) that was reported to show an autosomal dominant or recessive pattern,[64] thus we judged that the deletion was not relevant to his phenotype. As a result, cases 49–57 had only CNVs inherited from one of their parents which are likely to be unrelated to the phenotypes; that is, bCNV (Table 4).

As a result, we estimated that 48 cases among 349 analyzed (13.8%) had pCNV(s) in the second screening (Table 3; Figure 2). The CNVs of the remaining six cases, cases 58–63, were not associated with previously reported pathogenicity and their inheritance could not be evaluated, thus we estimated they were variants of uncertain clinical significance (VOUS).[38]

DISCUSSION

Because aCGH is a high-throughput technique to detect CNVs rapidly and comprehensively, this technique has been commonly used for analyses of patients with MCA and/or MR.[38,65–68] However, recent studies of human genomic variation have uncovered surprising properties of CNV, which covers 3.5–12% of the human genome even in healthy populations.[18–20,69] Thus analyses of patients with uncertain clinical phenotypes need to assess whether the CNV is pathogenic or unrelated to phenotypes.[21] However, such an assessment may diminish the rapidness or convenience of aCGH.

In this study, we evaluated whether our in-house GDA can work well as a diagnostic tool to detect CNVs responsible for well-established syndromes or those involved in subtelomeric aberrations in a clinical setting, and then explored candidate pCNVs in cases without any CNV in the first GDA screening. We recruited 536 cases that had been undiagnosed clinically and studied them in a two-stage screening using aCGH. In the first screening we detected CNVs in 54 cases (10.1%). Among them, 40 cases had CNV(s) at subtelomeric region(s) corresponding to the well-established syndromes or the already described disorders and the other 14 cases had CNVs in the regions corresponding to known disorders. Thus about three quarters of cases had genomic aberrations involved in subtelomeric regions. All the subtelomeric deletions and a part of the subtelomeric duplications corresponded to the disorders, indicating that especially subtelomeric deletions had more clinical significance compared to subtelomeric duplications, although the duplication might result in milder phenotypes and/or function as a modifier of phenotypes.[70] Moreover, parental analysis in three cases with two subtelomeric aberrations revealed that two of them were derived from the parental balanced translocations, indicating that such subtelomeric aberrations were potentially recurrent and parental analyses were worth performing. Recently several similar studies analyzed patients with MCA/MR or developmental delay using a targeted array for subtelomeric regions and/or known genomic disorders and detected clinically relevant CNVs in 4.4–17.1% of the patients.[28,65,70,71] Our detection rate in the first screening was equivalent to these reports. Although such detection rates depend on the type of microarray, patient selection criteria and/or number of subjects, these results suggest that at least 10% of cases with undiagnosed MCA/MR and a normal karyotype would be detectable by targeted array.

Table 4 Parental analysis of 34 cases in the second screening

Case	Gender	Clinical diagnosis	CNV		Size of CNV (bp)		Protein-coding genes	Parental analysis	Pathogenicity
			del/dup	Position	Min.	Max.			
1	M	MCA/MR	del	1p36.23p36.22	1 670 237	2 558 590	32	*de novo*	P
2	M	MCA/MR	del	1q41q42:11	5 001 798	6 481 439	35	*de novo*	P
7	M	MCA/MR	del	16p12.1p11.2	2 816 866	5 648 152	138	*de novo*	P
8	M	MCA/MR with CHD	del	16p11.2	951 773	4 258 984	134	*de novo*	P
10	M	MCA/MR	del	7p14.2p13	8 516 513	9 421 233	70	*de novo*	P
11	F	MCA/MR	del	14q22.1q22.3	2 746 662	3 089 980	18	*de novo*	P
12	M	MCA/MR	del	17q13.3	930 940	1 018 839	22	*de novo*	P
13	M	MCA/MR	del	Xp11.4p11.3	4 034 171	4 103 418	9	*de novo*	P
14	M	MCA/MR	del	6q12q14.1	14 194 290	16 071 847	56	*de novo*	P
18	M	MCA/MR	del	10q24.31q25.1	3 345 595	3 368 825	66	*de novo*	P
19	M	MCA/MR	del	10q24.32q25.1	2 077 638	2 093 622	41	*de novo*	P
21	M	MCA/MR	del	7p22.1	341 762	3 223 668	28	*de novo*	P
24	M	SMS susp.	del	19p13.2	1 719 919	3 304 902	23	*de novo*	P
37	F	MCA/MR	del	1p34.3	1 128 084	1 753 514	7	*de novo*	P
38	M	MCA/MR	dup	1q25.2	338 801	771 348	9	*de novo*	P
39	M	MCA/MR	del	2p24.1p23.3	3 721 550	8 376 636	86	*de novo*	P
40	F	MCA/MR	del	3p26.1p25.3	1 433 024	1 835 660	18	*de novo*	P
41	M	MCA/MR	del	3p22.1p21.31	5 893 173	7 832 879	123	*de novo*	P
42[a]	M	MCA/MR	del	8q21.11q21.13	5 289 394	5 770 485	12	*de novo*	P
42[a]	M	MCA/MR	del	3p14.3p14.2	593 434	1 517 140	11	Maternal	B
43	M	MCA/MR	del	3q26.31q26.33	4 081 515	6 002 971	12	*de novo*	P
44[b]	M	MCA/MR	del	13q13.2q13.3	917 819	1 458 769	1	*de novo*	P
44[b]	M	MCA/MR	del	22q11.21	917 819	1 458 769	15	Paternal	B
45	F	Rett syndrome	del	18q21.2	2 121 913	3 642 522	9	*de novo*	P
46	M	MCA/MR	dup	19p13.3	2 041 395	2 404 096	113	*de novo*	P
47	F	MCA/MR	del	19p13.3	816 079	2 037 409	23	*de novo*	P
48[c]	M	MCA/MR	del	Xp11.3	2 362 422	2 392 511	18	Maternal	P
49	M	MCA/MR	dup	3p26.3	176 050	250 850	1	Paternal	B
50	M	MCA/MR	dup	5p14.3	170 578	1 752 211	1	Paternal	B
51	M	MCA/MR	dup	5q13.3	1 020 329	1 421 706	3	Maternal	B
52	M	MCA/MR	dup	7p22.3	568	1 101 943	12	Maternal	B
53	F	MCA/MR	dup	8p23.2	838 610	2 648 539	1	Paternal	B
54	M	MCA/MR	dup	9q33.1	162 612	1 030 807	2	Paternal	B
55	F	MCA/MR	dup	10q22.3	154 664	873 124	1	Maternal	B
56	M	MCA/MR	dup	12q21.31	152 042	4 843 434	3	Paternal	B
57	M	Gillespie syndrome	del	Xp11.23	104 191	115 604	3	Maternal	B

Abbreviations: B, benign; CNV, copy-number variant; F, female; M, male; MCA/MR, multiple congenital anomalies and mental retardation; P, pathogenic.
[a]Two CNVs were detected in case 42.
[b]Two CNVs were detected in case 44.
[c]Nullizygous deletion inherited from his mother probably affected the phenotype.

Another interesting observation in the first screening was that subtelomeric rearrangements frequently occurred even in patients with MCA/MR of uncertain whose karyotype had been diagnosed as normal. This result may be consistent with a property of subtelomeric regions whose rearrangements can be missed in conventional karyotyping,[72] and in fact other techniques involving subtelomeric FISH or MLPA also identified subtelomeric abnormalities in a number of patients with MCA and/or MR in previous reports.[70,73,74] Our result may support the availability of prompt screening of subtelomeric regions for cases with uncertain congenital disorders.

In the second screening we applied WGA-4500 to 349 cases to detect 66 candidate pCNVs in 63 cases (18.1%), and subsequently assessed the pathogenicity of these CNVs. The pCNVs included nine CNVs overlapping identical regions of recently recognized syndromes (cases 1–9; deletion at 1p36.23–p36.22, 1q41–q42.11, 1q43–q44, 2q23.1, 14q12, 15q26-qter and 16p11.2–p12.2, respectively), four CNVs containing disease-associated genes (cases 10–13; *GLI3*, *BMP4*, *YWHAE* and *CASK*, respectively), three pairs of CNVs of recurrent deletions (cases 14, 15: at 6q12–q14.1 and 6q14.1; case 16, 17: at 10p12.1–p11.23 and case 18, 19: at 10q24.31–q25.1 and 10q24.32–q25.1), five CNVs identical to pCNVs in previous studies (cases 20–24), six large and/or gene-rich CNVs (cases 25–30) and six CNVs containing a morbid OMIM gene (cases 31–36). For the remaining cases, we estimated the pathogenicity of the CNVs from a parental analysis (Table 4). We judged the 11 *de novo* CNVs (cases 37–47) and 1 CNV on chromosome Xp11.3 inherited from

the mother (case 48) as probably pathogenic. And nine inherited CNVs (cases 49–57) were probably benign. The clinical significance of CNVs in the other six cases, cases 58–63, remains uncertain (VOUS). As a result we estimated CNVs as pathogenic in 48 cases among 349 cases (13.8%) analyzed in the second screening. None of the pCNVs corresponded to loci of well-established syndromes. This may suggest that our two-stage screening achieved a good balance between rapid screening of known syndromes and investigation of CNV of uncertain pathogenicity.

Table 5 Summary of parental analyses

		Average size (bp)		*The average number of protein-coding genes*
		Min.	*Max.*	
Pathogenic CNVs[a]				
del	23	3 309 267	4 597 689	43
dup	2	1 190 098	1 587 722	61
Total	25	3 139 733	4 356 892	44
Benign CNVs[b]				
del	3	538 481	1 030 504	10
dup	8	334 432	1 740 327	3
Total	11	390 082	1 546 739	5

Abbreviation: CNV, copy-number variant.
[a]Twenty-four *de novo* CNVs and case 48.
[b]Eleven inherited CNVs other than case 48.

Among the cases with parental analyses, the 25 pCNVs had larger sizes and contained more protein-coding genes (average size, 3.1 Mb at minimum to 4.4 Mb at maximum; average number of genes, 44) as compared with the 11 inherited bCNVs that were probably unrelated to phenotypes (average size, 0.39 Mb at minimum to 1.5 Mb at maximum; average number of genes, 5) (Table 5). Although all of the 25 pCNVs except 2 were deletions, about three quarters (8 of 11 cases) of the inherited bCNVs were duplications (Table 5). These findings are consistent with previously reported features of pCNVs and bCNVs.[21,38]

We also compared our current study with recent aCGH studies meeting the following conditions: (1) a microarray targeted to whole genome was applied; (2) patients with MCA and/or MR of uncertain etiology, normal karyotype and the criteria for patients selection were clearly described; (3) pathogenicity of identified CNVs were assessed. On the basis of the above criteria, among studies reported in the past 5 years, we summarized 13 studies (Table 6).[10,14,15,17,54,55,75–81] Diagnostic yield of pCNVs in each study was 6.3–16.4%, and our current diagnostic yield of the second screening was 13.8%. Though cases with subtelomeric aberration detected in the first screening had been excluded, our diagnostic yield was comparable to those of the reported studies. It is not so important to make a simple comparison between diagnostic yields in different studies as they would depend on the conditions of each study, for example, sample size or array resolution,[38,82] however it seems interesting that the higher resolution of a microarray does not ensure an increase in the rate of detection of pCNVs. One recent study showed data that may explain the discrepancy between the resolution of microarray and diagnostic yield.[54,83] The authors analyzed 1001 patients with MCA and/or MR using one

Table 6 Previous studies of analyzing patients with MCA and/or MR using aCGH targeted to whole genome

	Applied array			*Patients*		*Pathogenic CNV*	
Author (year)	*Type*	*Number*[a]	*Distribution*[b]	*Number*	*Type of disorders*	*Number*	*%*
Schoumans *et al.*[75]	BAC	2600	1.0 Mb*	41	MCA and MR	4	9.8
de Vries *et al.*[76]	BAC	32 477	Tiling	100	MCA and/or MR	10	10.0
Rosenberg *et al.*[77]	BAC	3500	1.0 Mb*	81	MCA and MR	13	16.0
Krepischi-Santos *et al.*[78]	BAC	3500	1.0 Mb*	95	MCA and/or MR	15	15.8
Friedman *et al.*[14]	SNP	Affymetrix 100K	23.6 kb**	100	MR	11	11.0
Thuresson *et al.*[79]	BAC		1.0 Mb*	48	MCA and MR	3	6.3
Wagenstaller *et al.*[80]	SNP	Affymetrix 100K	23.6 kb**	67	MR	11	16.4
Fan *et al.*[55]	Oligo	Agilent 44K	24 kb–43 kb**	100[c]	MCA and MR, Autism	15[d]	15.0
Xiang *et al.*[15]	Oligo	Agilent 44K	24 kb–43 kb**	40[e]	MR, DD and autism	3	7.5
Pickering *et al.*[10]	BAC	2600	1 Mb*	354[f]	MCA and/or MR	36[g]	10.2
McMullan *et al.*[17]	SNP	Affymetrix 500K	2.5 kb–5.8 kb**	120	MCA and/or MR	18	15.0
Bruno *et al.*[81]	SNP	Affymetrix 250K	2.5 kb–5.8 kb**	117	MCA and/or MR	18	15.4
Buysse *et al.*[54]	BAC	3431	1 Mb*	298	MCA and/or MR	26	8.7
	Oligo	Agilent 44K	24 kb–43 kb**	703	MCA and/or MR	74	10.5
Our current study	BAC	4523	0.7 Mb	349	MCA and MR	48	13.8
Total				2613		305	11.7

Abbreviations: BAC, bacterial artificial chromosome; CNV, copy-number variant; DD, developmental delay; MCA, multiple congenital anomalies; MR, mental retardation; SNP, single nucleotide polymorphism.
[a]The number of clones or name of array is described.
[b]Each distribution referred to each article (*) or manual of each manufacturer (**).
[c]All cases were analyzed by both a targeted array and a genome-wide array.
[d]In five cases, CNVs were also identified by a targeted array.
[e]Ten cases with an abnormal karyotype were excluded.
[f]Only cases studied with an array throughout the genome are described. Ninety-eight cases were also analyzed by a targeted array.
[g]Seventeen cases with an abnormal karyotype were excluded.

of two types of microarray, BAC array and oligonucleotide array. The BAC array was applied for 298 patients to detect 58 CNVs in 47 patients, and among them 26 CNVs (8.7%) were determined to be causal (pathogenic). Conversely, the oligonucleotide arrays were applied for 703 patients to detect 1538 CNVs in 603 patients, and among them 74 CNVs (10.5%) were determined to be pathogenic. These results may lead to the following idea: a lower-resolution microarray detects a limited number of CNVs likely to be pathogenic, because such CNVs tend to be large, and a higher-resolution microarray detects an increasing number of bCNVs or VOUS.[38] Indeed, in studies using a high-resolution microarray, most of the CNVs detected were smaller than 500 kb but almost all pCNVs were relatively large.[54,81,83] Most of the small CNVs were judged not to be pathogenic, and the percentage of pCNVs stabilized at around 10%. This percentage may suggest a frequency of patients with MCA/MR caused by CNV affecting one or more genes, other than known syndromes and subtelomeric aberrations. The other patients may be affected by another cause undetectable by genomic microarray; for example a point mutation or microdeletion/duplication of a single gene, aberration of microRNA, aberration of methylation states, epigenetic aberration or partial uniparental disomy.

As recently hypothesized secondary insult, which is potentially another CNV, a mutation in a phenotypically related gene or an environmental event influencing the phenotype, may result in clinical manifestation.[84] Especially, in two-hit CNVs, two models have been hypothesized: (1) the additive model of two co-occurring CNVs affecting independent functional modules and (2) the epistatic model of two CNVs affecting the same functional module.[85] It also suggests difficulty in selecting an optimal platform in the clinical screening. Nevertheless, information on both pCNVs and bCNVs detected through studies using several types of microarrays is unambiguously significant because an accumulation of the CNVs will create a map of genotype–phenotype correlation that would determine the clinical significance of each CNV, illuminate gene function or establish a new syndrome.

ACKNOWLEDGEMENTS

We thank Ayako Takahashi and Rumi Mori for technical assistance. This study was supported by the Joint Usage/Research Program of Medical Research Institute, Tokyo Medical and Dental University. This work was also supported by grants-in-aid for Scientific Research on Priority Areas from the Ministry of Education, Culture, Sports, Science, and Technology, Japan; a grant from Core Research for Evolutional Science and Technology (CREST) of the Japan Science and Technology Corporation (JST); a grant from the New Energy and Industrial Technology Development Organization (NEDO); and in part by Grant-in-Aid for Scientific Research (B) (17390099, 20390301) of Japan Society for the Promotion of Science (JSPS); Health and Labour Sciences Research Grants for Research on information system of undiagnosed diseases (H21-nanchi-ippan-167) and Research on policy for intractable diseases (H22-nanchi-shitei-001) from the Ministry of Health, Labour and Welfare, Japan.

1 Roeleveld, N., Zielhuis, G. A. & Gabreëls, F. The prevalence of mental retardation: a critical review of recent literature. *Dev. Med. Child Neurol.* **39**, 125–132 (1997).
2 Hunter, A. G. Outcome of the routine assessment of patients with mental retardation in a genetics clinic. *Am. J. Med. Genet.* **90**, 60–68 (2000).
3 Smith, D. W. & Bostian, K. E. Congenital anomalies associated with idiopathic mental retardation. *J. Pediatr.* **65**, 189–196 (1964).
4 Gustavson, K. H., Hagberg, B., Hagberg, G. & Sars, K. Severe mental retardation in a Swedish county. II. Etiologic and pathogenetic aspects of children born 1959–1970. *Neuropadiatrie* **8**, 293–304 (1977).
5 Fryns, J. P., Kleczkowska, A., Kubień, E. & Van den Berghe, H. Cytogenetic findings in moderate and severe mental retardation. A study of an institutionalized population of 1991 patients. *Acta. Paediatr. Scand. Suppl.* **313**, 1–23 (1984).
6 Gustavson, K. H., Holmgren, G. & Blomquist, H. K. Chromosomal aberrations in mildly mentally retarded children in a northern Swedish county. *Ups. J. Med. Sci. Suppl.* **44**, 165–168 (1987).
7 Schreppers-Tijdink, G. A., Curfs, L. M., Wiegers, A., Kleczkowska, A. & Fryns, J. P. A systematic cytogenetic study of a population of 1170 mentally retarded and/or behaviourly disturbed patients including fragile X-screening. The Hondsberg experience. *J. Genet Hum.* **36**, 425–446 (1988).
8 van Karnebeek, C. D., Koevoets, C., Sluijter, S., Bijlsma, E. K., Smeets, D. F., Redeker, E. J. *et al.* Prospective screening for subtelomeric rearrangements in children with mental retardation of unknown aetiology: the Amsterdam experience. *J. Med. Genet.* **39**, 546–553 (2002).
9 Vissers, L. E., de Vries, B. B., Osoegawa, K., Janssen, I. M., Feuth, T., Choy, C. O. *et al.* Array-based comparative genomic hybridization for the genomewide detection of submicroscopic chromosomal abnormalities. *Am. J. Hum. Genet.* **73**, 1261–1270 (2003).
10 Pickering, D. L., Eudy, J. D., Olney, A. H., Dave, B. J., Golden, D., Stevens, J. *et al.* Array-based comparative genomic hybridization analysis of 1176 consecutive clinical genetics investigations. *Genet. Med.* **10**, 262–266 (2008).
11 Bauters, M., Van Esch, H., Marynen, P. & Froyen, G. X chromosome array-CGH for the identification of novel X-linked mental retardation genes. *Eur. J. Med. Genet.* **48**, 263–275 (2005).
12 Hayashi, S., Honda, S., Minaguchi, M., Makita, Y., Okamoto, N., Kosaki, R. *et al.* Construction of a high-density and high-resolution human chromosome X array for comparative genomic hybridization analysis. *J. Hum. Genet.* **52**, 397–405 (2007).
13 Kok, K., Dijkhuizen, T., Swart, Y. E., Zorgdrager, H., van der Vlies, P., Fehrmann, R. *et al.* Application of a comprehensive subtelomere array in clinical diagnosis of mental retardation. *Eur. J. Med. Genet.* **48**, 250–262 (2005).
14 Friedman, J. M., Baross, A., Delaney, A. D., Ally, A., Arbour, L., Armstrong, L. *et al.* Oligonucleotide microarray analysis of genomic imbalance in children with mental retardation. *Am. J. Hum. Genet.* **79**, 500–513 (2006).
15 Xiang, B., Li, A., Valentin, D., Nowak, N. J., Zhao, H. & Li, P. Analytical and clinical validity of whole-genome oligonucleotide array comparative genomic hybridization for pediatric patients with mental retardation and developmental delay. *Am. J. Med. Genet.* **146A**, 1942–1954 (2008).
16 Shen, Y., Irons, M., Miller, D. T., Cheung, S. W., Lip, V., Sheng, X. *et al.* Development of a focused oligonucleotide-array comparative genomic hybridization chip for clinical diagnosis of genomic imbalance. *Clin. Chem.* **53**, 2051–2059 (2007).
17 McMullan, D. J., Bonin, M., Hehir-Kwa, J. Y., de Vries, B. B., Dufke, A., Rattenberry, E. *et al.* Molecular karyotyping of patients with unexplained mental retardation by SNP arrays: a multicenter study. *Hum. Mutat.* **30**, 1082–1092 (2009).
18 Iafrate, A. J., Feuk, L., Rivera, M. N., Listewnik, M. L., Donahoe, P. K., Qi, Y. *et al.* Detection of large-scale variation in the human genome. *Nat. Genet.* **36**, 949–951 (2004).
19 Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P. *et al.* Large-scale copy number polymorphism in the human genome. *Science.* **305**, 525–528 (2004).
20 Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Andrews, T. D. *et al.* Global variation in copy number in the human genome. *Nature* **444**, 444–454 (2006).
21 Lee, C., Iafrate, A. J. & Brothman, A. R. Copy number variations and clinical cytogenetic diagnosis of constitutional disorders. *Nat. Genet.* **39**, S48–S54 (2007).
22 Inazawa, J., Inoue, J. & Imoto, I. Comparative genomic hybridization (CGH)-arrays pave the way for identification of novel cancer-related genes. *Cancer Sci.* **95**, 559–563 (2004).
23 Hayashi, S., Kurosawa, K., Imoto, I., Mizutani, S. & Inazawa, J. Detection of cryptic chromosome aberrations in a patient with a balanced t(1;9)(p34.2;p24) by array-based comparative genomic hybridization. *Am. J. Med. Genet.* **139**, 32–36 (2005).
24 Shrimpton, A. E., Braddock, B. R., Thomson, L. L., Stein, C. K. & Hoo, J. J. Molecular delineation of deletions on 2q37.3 in three cases with an Albright hereditary osteodystrophy-like phenotype. *Clin. Genet.* **66**, 537–544 (2004).
25 Rauch, A. & Dörr, H. G. Chromosome 5q subtelomeric deletion syndrome. *Am. J. Med. Genet. C* **145C**, 372–376 (2007).
26 Horn, D., Tönnies, H., Neitzel, H., Wahl, D., Hinkel, G. K., von Moers, A. *et al.* Minimal clinical expression of the holoprosencephaly spectrum and of Currarino syndrome due to different cytogenetic rearrangements deleting the Sonic Hedgehog gene and the HLXB9 gene at 7q36.3. *Am. J. Med. Genet. A* **128A**, 85–92 (2004).
27 Tatton-Brown, K., Pilz, D. T., Orstavik, K. H., Patton, M., Barber, J. C., Collinson, M. N. *et al.* 15q overgrowth syndrome: a newly recognized phenotype associated with overgrowth, learning difficulties, characteristic facial appearance, renal anomalies and increased dosage of distal chromosome 15q. *Am. J. Med. Genet. A* **149A**, 147–154 (2009).
28 Lu, X., Shaw, C. A., Patel, A., Li, J., Cooper, M. L., Wells, W. R. *et al.* Clinical implementation of chromosomal microarray analysis: summary of 2513 postnatal cases. *PLoS One* **2**, e327 (2007).
29 Fernandez, T. V., García-González, I. J., Mason, C. E., Hernández-Zaragoza, G., Ledezma-Rodríguez, V. C., Anguiano-Alvarez, V. M. *et al.* Molecular characterization of a patient with 3p deletion syndrome and a review of the literature. *Am. J. Med. Genet. A* **146A**, 2746–2752 (2008).
30 Jones, K. L. *Smith's Recognizable Patterns of Human Malformation*, 6th edn. (Elsevier Saunders, Philadelphia, 2006).
31 Striano, P., Malacarne, M., Cavani, S., Pierluigi, M., Rinaldi, R., Cavaliere, M. L. *et al.* Clinical phenotype and molecular characterization of 6q terminal deletion syndrome: five new cases. *Am. J. Med. Genet. A* **140**, 1944–1949 (2006).

32 Lindstrand, A., Malmgren, H., Verri, A., Benetti, E., Eriksson, M., Nordgren, A. *et al.* Molecular and clinical characterization of patients with overlapping 10p deletions. *Am. J. Med. Genet. A* **152A,** 1233–1243 (2010).
33 Elbracht, M., Roos, A., Schönherr, N., Busse, S., Damen, R., Zerres, K. *et al.* Pure distal trisomy 2q: a rare chromosomal abnormality with recognizable phenotype. *Am. J. Med. Genet. A* **149A,** 2547–2550 (2009).
34 Lukusa, T. & Fryns, J. P. Pure *de novo* 17q25.3 micro duplication characterized by micro array CGH in a dysmorphic infant with growth retardation, developmental delay and distal arthrogryposis. *Genet. Couns.* **21,** 25–34 (2010).
35 Fukami, M., Kirsch, S., Schiller, S., Richter, A., Benes, V., Franco, B. *et al.* A member of a gene family on Xp22.3, VCX-A, is deleted in patients with X-linked nonspecific mental retardation. *Am. J. Hum. Genet.* **67,** 563–573 (2000).
36 Shaffer, L. G. & Tommerup, N. *An International System for Human Cytogenetic Nomenclature (2005)* (Karger, Basel, 2005).
37 Koolen, D. A., Pfundt, R., de Leeuw, N., Hehir-Kwa, J. Y., Nillesen, W. M., Neefs, I. *et al.* Genomic microarrays in mental retardation: a practical workflow for diagnostic applications. *Hum. Mutat.* **30,** 283–292 (2009).
38 Miller, D. T., Adam, M. P., Aradhya, S., Biesecker, L. G., Brothman, A. R., Carter, N. P. *et al.* Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am. J. Hum. Genet.* **86,** 749–764 (2010).
39 Shaffer, L. G., Theisen, A., Bejjani, B. A., Ballif, B. C., Aylsworth, A. S., Lim, C. *et al.* The discovery of microdeletion syndromes in the post-genomic era: review of the methodology and characterization of a new 1q41q42 microdeletion syndrome. *Genet. Med.* **9,** 607–616 (2007).
40 van Bon, B. W., Koolen, D. A., Borgatti, R., Magee, A., Garcia-Minaur, S., Rooms, L. *et al.* Clinical and molecular characteristics of 1qter microdeletion syndrome: delineating a critical region for corpus callosum agenesis/hypogenesis. *J. Med. Genet.* **45,** 346–354 (2008).
41 van Bon, B. W., Koolen, D. A., Brueton, L., McMullan, D., Lichtenbelt, K. D., Adès, L. C. *et al.* The 2q23.1 microdeletion syndrome: clinical and behavioural phenotype. *Eur. J. Hum. Genet.* **18,** 163–170 (2010).
42 Mencarelli, M. A., Kleefstra, T., Katzaki, E., Papa, F. T., Cohen, M., Pfundt, R. *et al.* 14q12 microdeletion syndrome and congenital variant of Rett syndrome. *Eur. J. Med. Genet.* **52,** 148–152 (2009).
43 Rump, P., Dijkhuizen, T., Sikkema-Raddatz, B., Lemmink, H. H., Vos, Y. J., Verheij, J. B. *et al.* Drayer's syndrome of mental retardation, microcephaly, short stature and absent phalanges is caused by a recurrent deletion of chromosome 15(q26.2→qter). *Clin. Genet.* **74,** 455–462 (2008).
44 Ballif, B. C., Hornor, S. A., Jenkins, E., Madan-Khetarpal, S., Surti, U., Jackson, K. E. *et al.* Discovery of a previously unrecognized microdeletion syndrome of 16p11.2-p12.2. *Nat. Genet.* **39,** 1071–1073 (2007).
45 Shinawi, M., Liu, P., Kang, S. H., Shen, J., Belmont, J. W., Scott, D. A. *et al.* Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioral problems, dysmorphism, epilepsy, and abnormal head size. *J. Med. Genet.* **47,** 332–341 (2010).
46 Kang, S. H., Scheffer, A., Ou, Z., Li, J., Scaglia, F., Belmont, J. *et al.* Identification of proximal 1p36 deletions using array-CGH: a possible new syndrome. *Clin. Genet.* **72,** 329–338 (2007).
47 Johnston, J. J., Olivos-Glander, I., Killoran, C., Elson, E., Turner, J. T., Peters, K. F. *et al.* Molecular and clinical analyses of Greig cephalopolysyndactyly and Pallister–Hall syndromes: robust phenotype prediction from the type and position of GLI3 mutations. *Am. J. Hum. Genet.* **76,** 609–622 (2005).
48 Johnston, J. J., Olivos-Glander, I., Turner, J., Aleck, K., Bird, L. M., Mehta, L. *et al.* Clinical and molecular delineation of the Greig cephalopolysyndactyly contiguous gene deletion syndrome and its distinction from acrocallosal syndrome. *Am. J. Med. Genet. A* **123A,** 236–242 (2003).
49 Hayashi, S., Okamoto, N., Makita, Y., Hata, A., Imoto, I. & Inazawa, J. Heterozygous deletion at 14q22.1-q22.3 including the BMP4 gene in a patient with psychomotor retardation, congenital corneal opacity and feet polysyndactyly. *Am. J. Med. Genet. A* **146A,** 2905–2910 (2008).
50 Hayashi, S., Mizuno, S., Migita, O., Okuyama, T., Makita, Y., Hata, A. *et al.* The CASK gene harbored in a deletion detected by array-CGH as a potential candidate for a gene causative of X-linked dominant mental retardation. *Am. J. Med. Genet. A* **146A,** 2145–2151 (2008).
51 Toyo-oka, K., Shionoya, A., Gambello, M. J., Cardoso, C., Leventer, R., Ward, H. L. *et al.* 14-3-3epsilon is important for neuronal migration by binding to NUDEL: a molecular explanation for Miller–Dieker syndrome. *Nat. Genet.* **34,** 274–285 (2003).
52 Mignon-Ravix, C., Cacciagli, P., El-Waly, B., Moncla, A., Milh, M., Girard, N. *et al.* Deletion of YWHAE in a patient with periventricular heterotopias and marked corpus callosum hypoplasia. *J. Med. Genet.* **47,** 132–136 (2010).
53 Haldeman-Englert, C. R., Gai, X., Perin, J. C., Ciano, M., Halbach, S. S., Geiger, E. A. *et al.* A 3.1-Mb microdeletion of 3p21.31 associated with cortical blindness, cleft lip, CNS abnormalities, and developmental delay. *Eur. J. Med. Genet.* **52,** 265–268 (2009).
54 Buysse, K., Delle Chiaie, B., Van Coster, R., Loeys, B., De Paepe, A., Mortier, G. *et al.* Challenges for CNV interpretation in clinical molecular karyotyping: lessons learned from a 1001 sample experience. *Eur. J. Med. Genet.* **52,** 398–403 (2009).
55 Fan, Y. S., Jayakar, P., Zhu, H., Barbouth, D., Sacharow, S., Morales, A. *et al.* Detection of pathogenic gene copy number variations in patients with mental retardation by genomewide oligonucleotide array comparative genomic hybridization. *Hum. Mutat.* **28,** 1124–1132 (2007).
56 Hevner, R. F., Shi, L., Justice, N., Hsueh, Y., Sheng, M., Smiga, S. *et al.* Tbr1 regulates differentiation of the preplate and layer 6. *Neuron* **29,** 353–366 (2001).
57 Cosma, M. P., Pepe, S., Annunziata, I., Newbold, R. F., Grompe, M., Parenti, G. *et al.* The multiple sulfatase deficiency gene encodes an essential and limiting factor for the activity of sulfatases. *Cell* **113,** 445–456 (2003).
58 Dierks, T., Schmidt, B., Borissenko, L. V., Peng, J., Preusser, A., Mariappan, M. *et al.* Multiple sulfatase deficiency is caused by mutations in the gene encoding the human C(alpha)-formylglycine generating enzyme. *Cell* **113,** 435–444 (2003).
59 Behar, O., Golden, J. A., Mashimo, H., Schoen, F. J. & Fishman, M. C. Semaphorin III is needed for normal patterning and growth of nerves, bones and heart. *Nature* **383,** 525–528 (1996).
60 Eudy, J. D., Ma-Edmonds, M., Yao, S. F., Talmadge, C. B., Kelley, P. M., Weston, M. D. *et al.* Isolation of a novel human homologue of the gene coding for echinoderm microtubule-associated protein (EMAP) from the Usher syndrome type 1a locus at 14q32. *Genomics* **43,** 104–106 (1997).
61 He, Y. & Casaccia-Bonnefil, P. The Yin and Yang of YY1 in the nervous system. *J. Neurochem.* **106,** 1493–1502 (2008).
62 Martin, C. L., Duvall, J. A., Ilkin, Y., Simon, J. S., Arreaza, M. G., Wilkes, K. *et al.* Cytogenetic and molecular characterization of A2BP1/FOX1 as a candidate gene for autism. *Am. J. Med. Genet.* **144B,** 869–876 (2007).
63 Tabolacci, E., Pomponi, M. G., Pietrobono, R., Terracciano, A., Chiurazzi, P. & Neri, G. A truncating mutation in the IL1RAPL1 gene is responsible for X-linked mental retardation in the MRX21 family. *Am. J. Med. Genet.* **140,** 482–487 (2006).
64 Nelson, J., Flaherty, M. & Grattan-Smith, P. Gillespie syndrome: a report of two further cases. *Am. J. Med. Genet.* **71,** 134–138 (1997).
65 Shaffer, L. G. & Bejjani, B. A. Medical applications of array CGH and the transformation of clinical cytogenetics. *Cytogenet. Genome Res.* **115,** 303–309 (2006).
66 Shaffer, L. G., Bejjani, B. A., Torchia, B., Kirkpatrick, S., Coppinger, J. & Ballif, B. C. The identification of microdeletion syndromes and other chromosome abnormalities: cytogenetic methods of the past, new technologies for the future. *Am. J. Med. Genet. C Semin. Med. Genet.* **145C,** 335–345 (2007).
67 Bejjani, B. A. & Shaffer, L. G. Clinical utility of contemporary molecular cytogenetics. *Annu. Rev. Genomics Hum. Genet.* **9,** 71–86 (2008).
68 Edelmann, L. & Hirschhorn, K. Clinical utility of array CGH for the detection of chromosomal imbalances associated with mental retardation and multiple congenital anomalies. *Ann. NY Acad. Sci.* **1151,** 157–166 (2009).
69 de Ståhl, T. D., Sandgren, J., Piotrowski, A., Nord, H., Andersson, R., Menzel, U. *et al.* Profiling of copy number variations (CNVs) in healthy individuals from three ethnic groups using a human genome 32K BAC-clone-based array. *Hum. Mutat.* **29,** 398–408 (2008).
70 Shao, L., Shaw, C. A., Lu, X. Y., Sahoo, T., Bacino, C. A., Lalani, S. R. *et al.* Identification of chromosome abnormalities in subtelomeric regions by microarray analysis: a study of 5,380 cases. *Am. J. Med. Genet. A* **146A,** 2242–2251 (2008).
71 Lu, X., Phung, M. T., Shaw, C. A., Pham, K., Neil, S. E., Patel, A. *et al.* Genomic imbalances in neonates with birth defects: high detection rates by using chromosomal microarray analysis. *Pediatrics* **122,** 1310–1318 (2008).
72 Xu, J. & Chen, Z. Advances in molecular cytogenetics for the evaluation of mental retardation. *Am. J. Med. Genet C Semin. Med. Genet.* **117C,** 15–24 (2003).
73 Ravnan, J. B., Tepperberg, J. H., Papenhausen, P., Lamb, A. N., Hedrick, J., Eash, D. *et al.* Subtelomere FISH analysis of 11 688 cases: an evaluation of the frequency and pattern of subtelomere rearrangements in individuals with developmental disabilities. *J. Med. Genet.* **43,** 478–489 (2006).
74 Ahn, J. W., Ogilvie, C. M., Welch, A., Thomas, H., Madula, R., Hills, A. *et al.* Detection of subtelomere imbalance using MLPA: validation, development of an analysis protocol, and application in a diagnostic centre. *BMC Med. Genet.* **8,** 9 (2007).
75 Schoumans, J., Ruivenkamp, C., Holmberg, E., Kyllerman, M., Anderlid, B. M. & Nordenskjold, M. Detection of chromosomal imbalances in children with idiopathic mental retardation by array based comparative genomic hybridisation (array-CGH). *J. Med. Genet.* **42,** 699–705 (2005).
76 de Vries, B. B., Pfundt, R., Leisink, M., Koolen, D. A., Vissers, L. E., Janssen, I. M. *et al.* Diagnostic genome profiling in mental retardation. *Am. J. Hum. Genet.* **77,** 606–616 (2005).
77 Rosenberg, C., Knijnenburg, J., Bakker, E., Vianna-Morgante, A. M., Sloos, W., Otto, P. A. *et al.* Array-CGH detection of micro rearrangements in mentally retarded individuals: clinical significance of imbalances present both in affected children and normal parents. *J. Med. Genet.* **43,** 180–186 (2006).
78 Krepischi-Santos, A. C., Vianna-Morgante, A. M., Jehee, F. S., Passos-Bueno, M. R., Knijnenburg, J., Szuhai, K. *et al.* Whole-genome array-CGH screening in undiagnosed syndromic patients: old syndromes revisited and new alterations. *Cytogenet. Genome Res.* **115,** 254–261 (2006).
79 Thuresson, A. C., Bondeson, M. L., Edeby, C., Ellis, P., Langford, C., Dumanski, J. P. *et al.* Whole-genome array-CGH for detection of submicroscopic chromosomal imbalances in children with mental retardation. *Cytogenet. Genome Res.* **118,** 1–7 (2007).
80 Wagenstaller, J., Spranger, S., Lorenz-Depiereux, B., Kazmierczak, B., Nathrath, M., Wahl, D. *et al.* Copy-number variations measured by single-nucleotide-polymorphism oligonucleotide arrays in patients with mental retardation. *Am. J. Hum. Genet.* **81,** 768–779 (2007).

81 Bruno, D. L., Ganesamoorthy, D., Schoumans, J., Bankier, A., Coman, D., Delatycki, M. *et al.* Detection of cryptic pathogenic copy number variations and constitutional loss of heterozygosity using high resolution SNP microarray analysis in 117 patients referred for cytogenetic analysis and impact on clinical practice. *J. Med. Genet.* **46**, 123–131 (2009).
82 Sagoo, G. S., Butterworth, A. S., Sanderson, S., Shaw-Smith, C., Higgins, J. P. & Burton, H. Array CGH in patients with learning disability (mental retardation) and congenital anomalies: updated systematic review and meta-analysis of 19 studies and 13,926 subjects. *Genet. Med.* **11**, 139–146 (2009).
83 Wincent, J., Anderlid, B. M., Lagerberg, M., Nordenskjöld, M. & Schoumans, J. High-resolution molecular karyotyping in patients with developmental delay and/or multiple congenital anomalies in a clinical setting. *Clin. Genet.* (e-pub ahead of print 8 May 2010).
84 Girirajan, S., Rosenfeld, J. A., Cooper, G. M., Antonacci, F., Siswara, P., Itsara, A. *et al.* A recurrent 16p12.1 microdeletion supports model for severe developmental delay. *Nat. Genet.* **42**, 203–209 (2010).
85 Veltman, J. A. & Brunner, H. G. Understanding variable expressivity in microdeletion syndromes. *Nat. Genet.* **42**, 192–193 (2010).

Journal of Human Genetics (2011) 56, 156–160

www.nature.com/jhg

ORIGINAL ARTICLE

Breakpoint determination of X;autosome balanced translocations in four patients with premature ovarian failure

Akira Nishimura-Tadaki[1], Takahito Wada[2], Gul Bano[3], Karen Gough[4], Janet Warner[4], Tomoki Kosho[5], Noriko Ando[6], Haruka Hamanoue[1,7], Hideya Sakakibara[7], Gen Nishimura[8], Yoshinori Tsurusaki[1], Hiroshi Doi[1], Noriko Miyake[1], Keiko Wakui[5], Hirotomo Saitsu[1], Yoshimitsu Fukushima[5], Fumiki Hirahara[7] and Naomichi Matsumoto[1]

Premature ovarian failure (POF) is a disorder characterized by amenorrhea and elevated serum gonadotropins before 40 years of age. As X chromosomal abnormalities are often recognized in POF patients, defects of X-linked gene may contribute to POF. Four cases of POF with t(X;autosome) were genetically analyzed. All the translocation breakpoints were determined at the nucleotide level. Interestingly, *COL4A6* at Xq22.3 encoding collagen type IV alpha 6 was disrupted by the translocation in one case, but in the remaining three cases, breakpoints did not involve any X-linked genes. According to the breakpoint sequences, two translocations had microhomology of a few nucleotides and the other two showed insertion of 3–8 nucleotides with unknown origin, suggesting that non-homologous end-joining is related to the formation of all the translocations.
Journal of Human Genetics (2011) **56**, 156–160; doi:10.1038/jhg.2010.155; published online 9 December 2010

Keywords: *COL4A6*; critical region; non-homologous end-joining; premature ovarian failure; X;autosome translocation

INTRODUCTION

Premature ovarian failure (POF) is a disorder characterized by amenorrhea and elevated serum gonadotropins before 40 years of age. The risk of this disorder or natural menopause before 40 years is approximately 1% of women.[1] Heterogeneous etiology should be involved in POF, such as environmental, autoimmune and genetic factors. X chromosomal abnormalities (partial monosomies and X;autosome-balanced translocations) are often observed in POF patients. These rearrangements cluster at Xq13–q26 called the critical region (for POF).[2,3] The critical region is separated into two: critical region 1 at Xq13–q21 and critical region 2 at Xq23–q26.[2,4] It was suggested that several X-linked loci expressing on both X chromosomes, which were required in a higher dosage for normal ovarian function, were involved in POF.[5] Furthermore, genetic factors for POF may be more complex as X;autosome translocations often disrupt no genes; therefore, other factors, such as position effects on autosomal genes, are proposed.[6] We had an opportunity to analyze four cases of POF each having t(X;autosome). Precise determination of translocation breakpoints in these patients may reveal direct evidence of POF-related genes and mechanisms of the formation of chromosomal translocations. Breakpoint sequences will be presented and discussed in relation to genes and formation process.

MATERIALS AND METHODS

Patients and genomic DNA preparation

A total of four POF patients with t(X;autosome) were recruited to this study. Case 1 had secondary amenorrhea and the other three (cases 2, 3 and 4) presented with primary amenorrhea. Cases 1, 3 and 4 are Japanese and case 2 is Caucasian. Case 2 was reported previously.[7] Chromosome analysis revealed 46,X,t(X;4)(q21.3;p15.2) in case 1, 46,X,t(X;2)(q22;p13) in case 2, 46,X,t(X;4)(q22.1;q12) in case 3 and 46,X,t(X;14)(q24;q32.1) in case 4. All translocations occurred *de novo*. In addition, 11 other POF patients were collected to check candidate gene abnormality. After informed consent was obtained, genomic DNA was prepared from peripheral blood leukocytes using QuickGene-610L (Fujifilm, Tokyo, Japan). Institutional review board approved the research protocol.

Fluorescence *in situ* hybridization

Metaphase chromosomes were prepared from peripheral blood lymphocytes of POF cases. Bacterial artificial chromosome DNA was labeled with fluorescein isothiocyanate- or Cy3-11-dUTP by Vysis Nick Translation kit (Vysis, Downers

[1]Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama, Japan; [2]Division of Neurology, Clinical Research Institute, Kanagawa Children's Medical Center, Yokohama, Japan; [3]Department of Cellular and Molecular Medicine, St George's University of London, London, UK; [4]Mater Pathology, Mater Health Services, Brisbane, Queensland, Australia; [5]Department of Medical Genetics, Shinshu University School of Medicine, Matsumoto, Japan; [6]Division of Obstetrics and Gynecology, Yokohama Municipal Citizens Hospital, Yokohama, Japan; [7]Department of Obstetrics and Gynecology and Reproductive Science, Yokohama City University Graduate School of Medicine, Yokohama, Japan and [8]Department of Radiology, Tokyo Metropolitan Kiyose Children's Hospital, Tokyo, Japan
Correspondence: Professor N Matsumoto, Department of Human Genetics, Yokohama City University Graduate School of Medicine, 3-9 Fukuura, Kanazawa-ku, Yokohama 236-0004, Japan.
E-mail: naomat@yokohama-cu.ac.jp
Received 17 September 2010; revised 10 November 2010; accepted 12 November 2010; published online 9 December 2010

Grove, IL, USA), and denatured at 70 °C for 10 min. Probe–hybridization mixtures (15 µl) were applied to chromosomes, incubated at 37 °C for 16–72 h, and then washed and mounted in antifade solution (Vector, Burlingame, CA, USA) containing 4′,6-diamidino-2-phenylindole. Photographs were taken on an AxioCam MR CCD fitted to Axioplan2 fluorescence microscope (Carl Zeiss, Oberkochen, Germany).

Southern blot and inverse PCR

Genomic DNA was digested with restriction enzymes: *Eco*RI and *Hin*dIII for case 1 and her parents, *Sac*I and *Eco*RI for case 2 and a normal female control, *Bam*HI and *Eco*RI for case 3 and her parents and *Nde*I and *Bgl*II for case 4 and her mother. Probes were made by PCR and labeled using DIG synthesis kit (Roche Applied Science, Basel, Switzerland). Hybridization, wash and detection were performed according to the manufacturer's protocol. Inverse PCR was performed using self-ligated DNA after digestion with *Eco*RI (cases 1 and 3), *Sac*I (case 2) and *Bgl*II (case 4). All the breakpoints were determined by sequencing inverse PCR products. Information of primers used is available on request.

Mutation analysis

Genomic DNA was obtained from peripheral blood leukocytes by standard methods and used for mutational screening. Protein coding exons of *COL4A6* (exons 1–45), insulin-like growth factor binding protein 7 (*IGFBP7*) (exons 1–5) and *C14orf159* (exons 4–16) were screened by high-resolution melt analysis using LightCycler 480 system II (Roche Applied Science, Tokyo, Japan), except for exon 1 of *IGFBP7*, which were analyzed by direct sequencing. PCR mixture contained 20 ng genomic DNA, 1× ExTaq buffer, 0.2 mM each dNTPs, 0.3 µM each primer, 0.25 µl SYTO9 (Invitrogen, Carlsbad, CA, USA) and 0.25 U ExTaq HS (Takara, Ohtsu, Japan). PCR was initially denatured at 94 °C for 2 min and cycled 45 times for 10 s at 94 °C, 15 s at 60 °C and 15 s at 72 °C, and then finalized at 72 °C for 1 min. High-resolution melt analysis was then performed. For exon 1 of *IGFBP7*, PCR mixture contained 20 ng genomic DNA, 1× GC buffer II, 0.4 mM each dNTPs, 1 µM each primers, 2% dimethylsulfoxide and 0.04 U LaTaq HS (Takara), and then PCR was initially denatured at 94 °C for 2 min and cycled 35 times at 94 °C for 20 s, at 60 °C for 20 s, at 72 °C for 1 min, and then finalized at 72 °C for 2 min. If a sample showed an aberrant melting curve pattern, the PCR product was purified using ExoSAP-IT (USB, Cleveland, OH, USA) and sequenced by a standard method using BigDye terminator ver.3 (Applied Biosystems, Foster City, CA, USA) on the ABI PRISM 3100 Genetic analyzer (Applied Biosystems). Sequences were compared with reference sequences using SeqScape version 2.7 (Applied Biosystems).

X-inactivation assay

Human androgen receptor assay and FRAXA locus methylation assay were performed as described previously,[8,9] with a slight modification. In brief, genomic DNA of patients, their parents and a female control was digested with two methylation-sensitive enzymes, *Hpa*II and *Hha*I. Subsequently, PCR was performed using digested and undigested DNA with human androgen receptor assay primers (FAM-labeled ARf: 5′-TCCAGAATCTGTTCCAGAG CGTGC-3′; ARr: 5′-CTCTACGATGGGCTTGGGGAGAAC-3′)[10] and FRAXA primers (FAM-labeled FRM1f: 5′-AGCCCCGCACTTCCACCACCAGCTCCT CCA-3′; FMR1r: 5′-GCTCAGCTCCGTTTCGGTTTCACTTCCGGT-3′), electrophoresed on ABI PRISM 3100 Genetic analyzer and analyzed with GeneMapper™ Software version 3.5 (Applied Biosystems).

RESULTS

Breakpoint sequences

Using fluorescence *in situ* hybridization analysis of metaphase chromosomes, we could identify Bacterial artificial chromosome clones spanning translocation breakpoints in each patient: RP11-636H11 at Xq22.3 (case 1), RP11-815E21 at Xq22.3 (case 2), RP11-589G9 at 4q12 (case 3) and RP11-904N19 at Xq24 (case 4). Southern blot analysis could identify aberrant bands in all the patients (Figure 1) and subsequent inverse PCR successfully cloned all breakpoints in the four cases. Breakpoint sequences are shown in Figure 2. Junction sequencing of der(X) and der(4) in case 1 revealed a 4192-bp deletion of chromosome X (UCSC genome browser coordinates March 2006 version: chr. X: 107 322 866–107 327 057 bp) and 7082-bp deletion (chr. 4: 11 846 359–11 853 387 bp) of chromosome 4, respectively. In addition, five unknown nucleotides were recognized at the der(X)

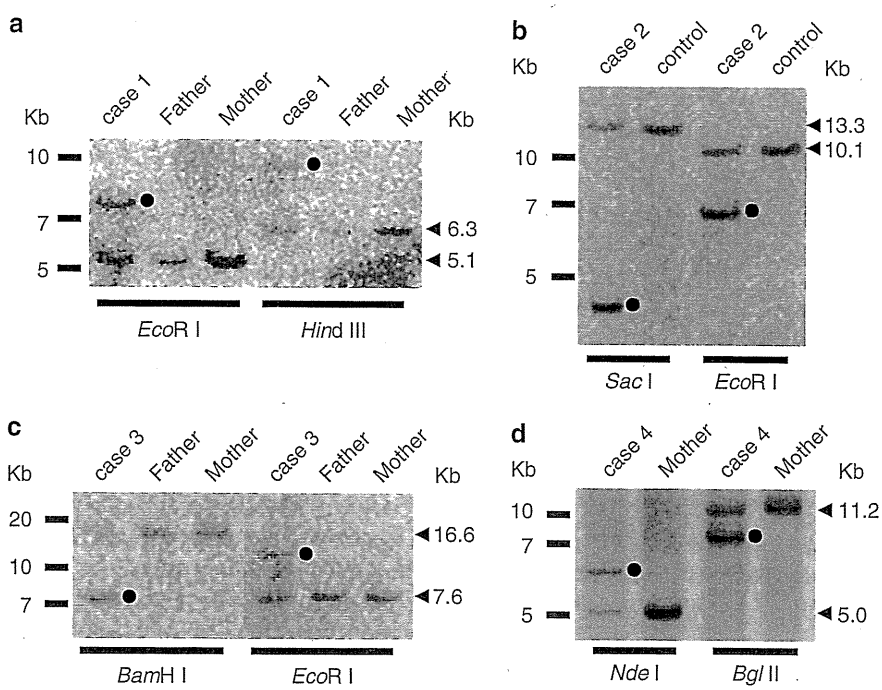

Figure 1 Southern blot analysis of four cases. (**a**) case 1, (**b**) case 2, (**c**) case 3 and (**d**) case 4. Aberrant bands are indicated with dots.

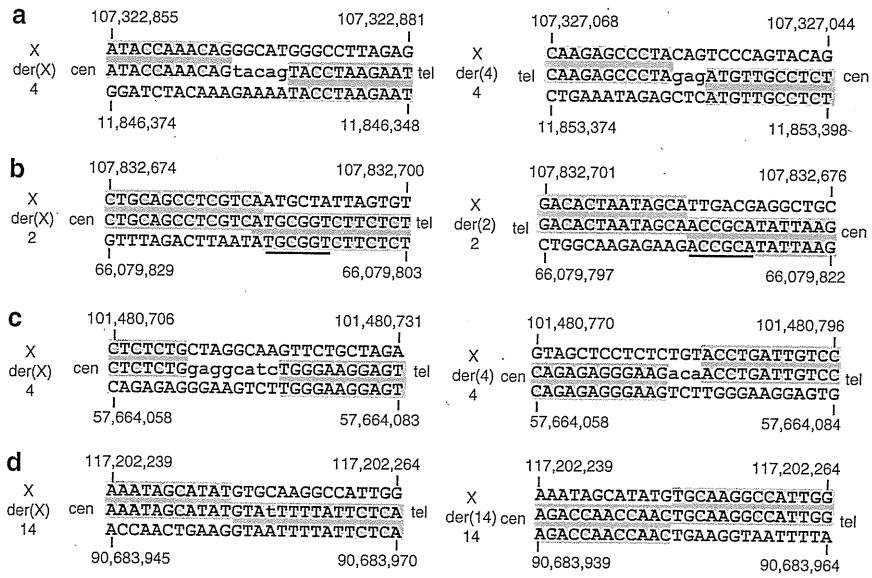

Figure 2 Breakpoint sequences of t(X;autosome) in four cases. (**a**) case 1, (**b**) case 2, (**c**) case 3 and (**d**) case 4. Top, middle and bottom sequences indicate those of normal X, derivative and normal autosomal chromosomes. Upper and lower cases indicate nucleotides of known and unknown origin, respectively. Matched sequences are with gray shadow. Underline indicates duplicated sequence. Numbers are based on the nucleotide position of the UCSC genome browser coordinates March 2006 version.

junction as well as three unknown nucleotides at der(4) (Figure 2a). Sequences of der(X) and der(2) in case 2 indicated six nucleotides (chr. 2: 66 079 810–66 079 816 bp) were duplicated (Figure 2b). In case 3, 71 nucleotides (chr. X: 101 480 713–101 480 782 bp) were deleted, and unknown eight nucleotides were inserted in der(X), and unknown three nucleotides were also recognized in der(4) (Figure 2c). In case 4, a nucleotide in chromosome X (chr. X: 117 202 250 bp) and five nucleotides in chromosome 14 (chr. 14: 90 683 951–90 683 955 bp) were missing (Figure 2d). The locations of X-chromosome breakpoints are shown in Figure 3. Translocation breakpoints disrupted *COL4A6* at Xq22.3 in case 1 (Figures 3a and b, Table 1), *IGFBP7* at 4q12 in case 3 (Table 1) and *C14orf159* at 14q32.12 in case 4 (Table 1). Other breakpoints did not involve any functional genes. Adjacent genes to breakpoints (less than 100 kb away) are *COL4A5* (5 kb away at Xq22.3) and *IRS4* (30 kb away at Xq22.3) in case 2, *NXF2* (12 or 21 kb away at Xq22.1) in case 3 and *KLHL13* (67 kb away at Xq24) in case 4 (Table 1). *COL4A6*, encoding collagen type IVα6, was the only disrupted X-linked gene in our POF patients.

X-inactivation assay

Human androgen receptor assay in cases 2 and 3 and FRAXA assay in cases 1 and 4 clearly indicated skewed X inactivation in all cases (100% in case 1, 94% in case 2, 98% in case 3 and 100% in case 4) and random patterns in their mothers available for this study (20–80%). Eleven other POF patients also showed random inactivation patterns (30–70%).

Mutation search

Considering accumulation of X-chromosome structural abnormalities in POF, X-chromosomal genes disrupted by rearrangements would be the primary target of this study. As *COL4A6* was completely disrupted in case 1 (Figure 3b), we started analyzing *COL4A6* as a candidate in 11 other POF patients. We found one heterozygous missense change, c.1460G>T (p.Gly487Val) (Figure 3c). This mutation was not recognized in 247 ethnically matched female controls (494 alleles). Web-based SIFT (http://sift.jcvi.org/) and PolyPhen (http://genetics.bwh.harvard.edu/pph/) did not indicate harmful effects of the amino-acid change on protein function: 0.26 by SIFT (predictable functional damage is <0.05) and 'benign' by PolyPhen, but the amino acid was evolutionally conserved (Figure 3d). The Gly487 was located between the (Gly–*X*–*Y*)*n* repeats. Parental origin of the change could not be confirmed as parental samples were unavailable. As *IGFB7* at 4q12 and *C14orf159* at 14q32.12 were also disrupted, both genes were analyzed in the 11 POF patients, but no mutation was found.

DISCUSSION

In this study, we could successfully determine the translocation breakpoints at nucleotide level in all the four cases analyzed. *COL4A6* at Xq22.3 in case 1, *IGFBP7* at 4q12 in case 3 and *C14orf159* at 14q32.12 in case 4 were disrupted. No genes were disrupted in case 2. Importantly, *COL4A6* was the only X-linked gene that was our primary target as a causative gene for POF. One missense change with benign nature outside the functional repeats was found in another POF patient who showed random X inactivation (35%).

In case 1, based on the skewing of X inactivation, der(X) should be active and normal X should be inactive. Thus, *COL4A6* is predicted to be functionally null in case 1 as the active allele is disrupted by the translocation. Collagen type IV is an essential component of basement membrane, consisting of six distinct α-chains (α1–α6) encoded by *COL4A1* to *COL4A6*. These six genes are located in three pairs with head-to-head orientation, *COL4A1*/*COL4A2* on chromosome 13, *COL4A3*/*COL4A4* on chromosome 2 and *COL4A5*/*COL4A6* on chromosome X. The chains interact and assemble with specificity to form three distinct patterns: α1α1α2, α3α4α5 and α5α5α6.[11] The α5- and α6-chains are found in the basement membrane of skin, smooth muscle and kidney.[12] Two transcripts of *COL4A6* are known, isoforms A and B (Figure 3b). The protein structure of collagen type IV contains an amino-terminal collagenous domain (also called 7S domain), a triple-helical region (Gly–*X*–*Y*) and a carboxyl-terminal

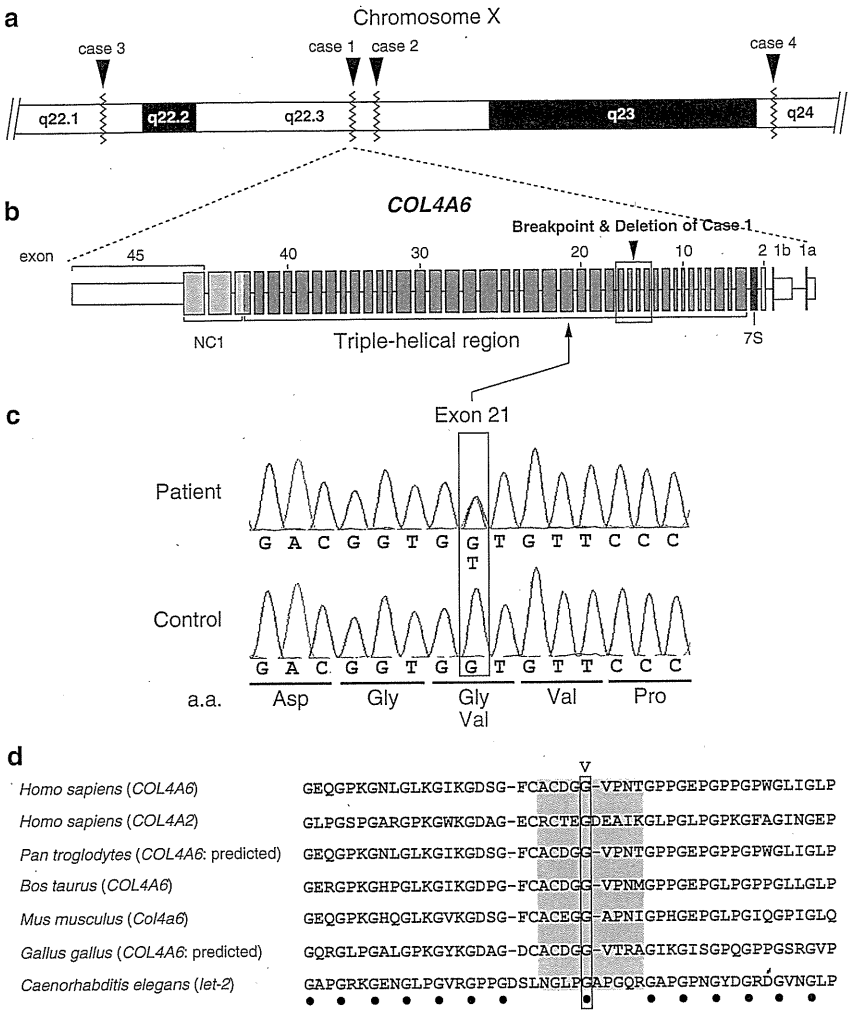

Figure 3 Location of the X-chromosome breakpoints and the *COL4A6* gene. (**a**) Breakpoint locations (zigzag lines) of four cases around Xq22.1–q24. (**b**) Schema of the *COL4A6* gene. Boxes are exons with numbering. White, dark gray, gray and light gray boxes indicate UTRs, 7S domain, triple-helical region and non-collagenous (NC1) domain, respectively. Breakpoint of the translocation with associated deletion is shown above boxes. (**c**) Heterozygous missense mutation, c.1460G>T (p.Gly487Val) at exon 21, is shown in the upper panel and wild-type sequence is shown in the lower panel. a.a.: amino acid. (**d**) Evolutionary conservation of the Gly487. CLUSTALW (http://align.genome.jp/) was used for this analysis. Dots show perfect conservation. Gray box is a space between the Gly–*X*–*Y* repeats.

Table 1 Genes within a 100-kb distance from translocation breakpoints

Case	*Chromosome X*	*Autosome*
1	*COL4A6*	None
	[q22.3]	[4p15.33]
2	*COL4A5* (5 kb)	None
	IRS4 (30 kb)	[2p14]
	[q22.3]	
3	*NXF2*[a] (12 or 21 kb)	*IGFBP7*
	[q22.1]	[4q12]
4	*KLHL13* (67 kb)	*C14orf159*
	[q24]	[14q32.12]

Round and square brackets indicate a distance from a breakpoint and chromosomal location, respectively.
Underlined genes are disrupted by breakpoints.
[a]*NXF2* is mapped to two adjacent segmental duplications.

non-collagenous (NC1) domain (Figure 3b).[13] We found a missense change, c.1460G>T (p.Gly487Val), in exon 21 in another POF patient (Figure 3c). Although this change is not found in 247 Japanese controls, its benign nature is suspected based on the web-based programs, the location outside the functional repeats and random X inactivation leading to the production of normal α6-chain. Parental samples were unfortunately unavailable to test the origin of the nucleotide change.

COL4A6 abnormality is known to be related to Alport syndrome with diffuse leiomyomas (AL-DS). *COL4A6* deletions in AL-DS are limited to exons 1, 1′ and 2 always together with *COL4A5* deletion in diverse extent.[14,15] In this paper, we first describe a POF patient (case 1) with disruption of only *COL4A6* not involving *COL4A5*. The inactivated normal X chromosome as well as the der(X) with disrupted *COL4A6* should lead to functionally null status in the patient. Extracellular matrix proteins (including COL4A6) have been shown to alter Leydig cell steroidogenesis *in vivo*, implying that Leydig cell steroidogenic

activity and matrix environment are interdependent.[16] Therefore, *COL4A6* depletion in ovarian extracellular matrix may alter normal steroidogenesis even in the ovary and have been possibly the cause of POF, especially in case 1. So far, there have been at least eight POF genes registered in OMIM: *FMR1* at Xq27.3 (POF1, OMIM no. 311360), *DIAPH2* at Xq22 (POF2A, OMIM no. 300511), *POF1B* at Xq21 (POF2B, OMIM no. 300604), *FOXL2* at 3q23 (POF3, OMIM no. 608996), *BMP15* at Xp11.2 (POF4, OMIM no. 300510), *NOBOX* at 7q35 (POF5, OMIM no. 611548), *FIGLA* at 2p12 (POF6, OMIM no. 612310) and *NR5A1* at 9q33 (POF7, OMIM no. 312964). Furthermore, *XPNPEP2* at Xq25,[17] *DACH2* at Xq21.2[18] and *CHM* (Xq21.2)[19] have also been described as being disrupted by translocations. *COL4A6* may possibly be an additional X-linked gene related to POF.

Two autosomal genes were disrupted: a gene encoding *IGFBP7* at 4q12 and *C14orf159* on 14q32.12. *IGFBP7* (also known as *IGFBP-rP1* or *MAC25*) is a secreted 31-kDa protein, belonging to the IGFBP superfamily. *IGFBP7* is involved in proliferation, senescence and apoptosis. Recently, it is reported that *IGFBP7* loss has a functional role in thyroid carcinogenesis.[20] *C14orf159* is a hypothetical protein with unknown function. Both disrupted genes were relatively expressed in ovary based on the GeneCards database (http://www.genecards.org/). We could not find any sequence aberrations in either gene among other POF patients. Further analysis might be necessary in relation to POF.

According to the precise breakpoint locations in all the cases reported here, *COL4A5* and *IRS4* (case 2), *NXF2* (case 3) and *KLHL13* (case 4) were localized near to breakpoints (within less than a 100-kb distance). All the adjacent genes except for *KLHL13* are shown to be expressed in human ovary in the GeneCards database. Interestingly, it was suggested that IRS4 protein expression was decreased in theca cells of polycystic ovary syndrome[21] and IGFBP7 protein suppressed estrogen production in granulose cells.[22] Reduced expression of these genes owing to the position effects by translocations could affect to normal ovarian function.

On the basis of the breakpoint sequences, two translocations (in cases 2 and 4) had microhomology (defined as the presence of the same short sequence of bases) of a few nucleotides and the other two (in cases 1 and 3) showed insertion of 3–8 nucleotides of unknown origin, suggesting that non-homologous end-joining is related to the formation of all the translocations in our patients.[23]

In conclusion, we could determine four t(X;autosome) breakpoints at the nucleotide level. We found that only one X-linked gene, *COL4A6*, was disrupted, resulting in functionally null status. All the four translocations are formed by non-homologous end-joining.

ACKNOWLEDGEMENTS

This work was supported by grants from the Ministry of Education, Culture, Sports, Science and Technology (NM), the Japan Science and Technology Agency (NM), the Ministry of Health, Labour and Welfare, Japan (NM) and Japan Society for the Promotion of Science (JSPS) (NM and AN). AN is a JSPS fellow.

1 Coulam, C. B., Adamson, S. C. & Annegers, J. F. Incidence of premature ovarian failure. *Obstet. Gynecol.* **67,** 604–606 (1986).
2 Therman, E., Laxova, R. & Susman, B. The critical region on the human Xq. *Hum. Genet.* **85,** 455–461 (1990).
3 Schlessinger, D., Herrera, L., Crisponi, L., Mumm, S., Percesepe, A., Pellegrini, M. *et al.* Genes and translocations involved in POF. *Am. J. Med. Genet.* **111,** 328–333 (2002).
4 Rizzolio, F., Bione, S., Sala, C., Goegan, M., Gentile, M., Gregato, G. *et al.* Chromosomal rearrangements in Xq and premature ovarian failure: mapping of 25 new cases and review of the literature. *Hum. Reprod.* **21,** 1477–1483 (2006).
5 Sala, C., Arrigo, G., Torri, G., Martinazzi, F., Riva, P., Larizza, L. *et al.* Eleven X chromosome breakpoints associated with premature ovarian failure (POF) map to a 15-Mb YAC contig spanning Xq21. *Genomics* **40,** 123–131 (1997).
6 Rizzolio, F., Sala, C., Alboresi, S., Bione, S., Gilli, S., Goegan, M. *et al.* Epigenetic control of the critical region for premature ovarian failure on autosomal genes translocated to the X chromosome: a hypothesis. *Hum. Genet.* **121,** 441–450 (2007).
7 Bano, G., Mansour, S. & Nussey, S. The association of primary hyperparathyroidism and primary ovarian failure: a *de novo* t(X; 2) (q22p13) reciprocal translocation. *Eur. J. Endocrinol.* **158,** 261–263 (2008).
8 Allen, R. C., Zoghbi, H. Y., Moseley, A. B., Rosenblatt, H. M. & Belmont, J. W. Methylation of *Hpa*II and *Hha*I sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *Am J Hum. Genet.* **51,** 1229–1239 (1992).
9 Carrel, L. & Willard, H. F. An assay for X inactivation based on differential methylation at the fragile X locus, FMR1. *Am. J. Med. Genet.* **64,** 27–30 (1996).
10 Kubota, T., Nonoyama, S., Tonoki, H., Masuno, M., Imaizumi, K., Kojima, M. *et al.* A new assay for the analysis of X-chromosome inactivation based on methylation-specific PCR. *Hum. Genet.* **104,** 49–55 (1999).
11 Khoshnoodi, J., Pedchenko, V. & Hudson, B. G. Mammalian collagen IV. *Microsc. Res. Technol.* **71,** 357–370 (2008).
12 Borza, D. B., Bondar, O., Ninomiya, Y., Sado, Y., Naito, I., Todd, P. *et al.* The NC1 domain of collagen IV encodes a novel network composed of the alpha 1, alpha 2, alpha 5, and alpha 6 chains in smooth muscle basement membranes. *J. Biol. Chem.* **276,** 28532–28540 (2001).
13 Zhou, J., Ding, M., Zhao, Z. & Reeders, S. T. Complete primary structure of the sixth chain of human basement membrane collagen, alpha 6(IV). Isolation of the cDNAs for alpha 6(IV) and comparison with five other type IV collagen chains. *J. Biol. Chem.* **269,** 13193–13199 (1994).
14 Sugimoto, K., Yanagida, H., Yagi, K., Kuwajima, H., Okada, M. & Takemura, T. A Japanese family with Alport syndrome associated with esophageal leiomyomatosis: genetic analysis of COL4A5 to COL4A6 and immunostaining for type IV collagen subtypes. *Clin. Nephrol.* **64,** 144–150 (2005).
15 Hertz, J. M., Juncker, I. & Marcussen, N. MLPA and cDNA analysis improves COL4A5 mutation detection in X-linked Alport syndrome. *Clin. Genet.* **74,** 522–530 (2008).
16 Mazaud Guittot, S., Verot, A., Odet, F., Chauvin, M. A. & le Magueresse-Battistoni, B. A comprehensive survey of the laminins and collagens type IV expressed in mouse Leydig cells and their regulation by LH/hCG. *Reproduction* **135,** 479–488 (2008).
17 Prueitt, R. L., Ross, J. L. & Zinn, A. R. Physical mapping of nine Xq translocation breakpoints and identification of XPNPEP2 as a premature ovarian failure candidate gene. *Cytogenet. Cell Genet.* **89,** 44–50 (2000).
18 Prueitt, R. L., Chen, H., Barnes, R. I. & Zinn, A. R. Most X;autosome translocations associated with premature ovarian failure do not interrupt X-linked genes. *Cytogenet. Genome Res.* **97,** 32–38 (2002).
19 Lorda-Sanchez, I. J., Ibanez, A. J., Sanz, R. J., Trujillo, M. J., Anabitarte, M. E., Querejeta, M. E. *et al.* Choroideremia, sensorineural deafness, and primary ovarian failure in a woman with a balanced X-4 translocation. *Ophthalmic Genet.* **21,** 185–189 (2000).
20 Vizioli, M. G., Sensi, M., Miranda, C., Cleris, L., Formelli, F., Anania, M. C. *et al.* IGFBP7: an oncosuppressor gene in thyroid carcinogenesis. *Oncogene* **29,** 3835–3844 (2010).
21 Yen, H. W., Jakimiuk, A. J., Munir, I. & Magoffin, D. A. Selective alterations in insulin receptor substrates-1, -2 and -4 in theca but not granulosa cells from polycystic ovaries. *Mol. Hum. Reprod.* **10,** 473–479 (2004).
22 Tamura, K., Matsushita, M., Endo, A., Kutsukake, M. & Kogo, H. Effect of insulin-like growth factor-binding protein 7 on steroidogenesis in granulosa cells derived from equine chorionic gonadotropin-primed immature rat ovaries. *Biol. Reprod.* **77,** 485–491 (2007).
23 Gu, W., Zhang, F. & Lupski, J. R. Mechanisms for human genomic rearrangements. *Pathogenetics* **1,** 4 (2008).