Letters to the Editor

Phosphomannose Isomerase Deficiency: A Carbohydrate-Deficient Glycoprotein Syndrome with Hepatic-Intestinal Presentation

To the Editor:

Carbohydrate-deficient glycoprotein (CDG) syndromes are genetic diseases that are due to defects in the glycosylation of glycoproteins (Jaeken et al. 1993, 1997*b*). With one exception (Billette de Villemeur et al. 1995), all reported patients presented moderate to severe brain disease (Jaeken and Casaer 1997). The diagnosis is usually made by isoelectrofocusing of serum transferrin showing different patterns of cathodal shift due to the deficiency of the terminal sialic acid (Jaeken et al. 1984; Stibler and Jaeken 1990). The majority of patients show the so-called type I pattern characterized by an increase of di- and asialotransferrin. About 70% of patients with the type I pattern have a deficiency of phosphomannomutase (PMM) (type IA) (Van Schaftingen and Jaeken 1995; Jaeken et al. 1997*a*), the enzyme catalyzing the second step in the conversion of fructose 6-phosphate to GDP-mannose (fig. 1), and mutations in the PMM2 gene have been found in all patients with a PMM deficiency (Matthijs et al. 1997*a*, 1998).

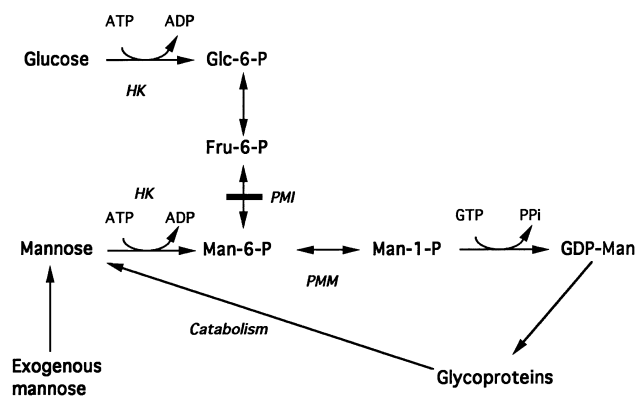Among the patients from which we analyzed the fi-

Figure 1 Scheme of mannose metabolism. Phosphomannose isomerase = PMI; phosphomannomutase = PMM; hexokinase = HK. The horizontal bar indicates the defect in phosphomannose isomerase.

Table 1

Enzyme Activities in Fibroblasts (in mU/mg protein) of Controls, Patients with Phosphomannomutase Deficiency, and Three Patients with Phosphomannose Isomerase Deficiency

	PMI Activities	PMM Activities
Controls	6.8 ± 2.4 (7)	3.8 ± .9 (8)
PMM deficient	8.0 ± 4.4 (13)	.5 ± .3 (13)
Patient A	.2, .3	6.3
Patient B	.2, .8	3.2
Patient C	.2, .2	3.5

NOTE.—Phosphomannomutase and protein were measured as described elsewhere (Van Schaftingen and Jaeken 1995; Jaeken et al. 1997*a*). Phosphomannose isomerase was assayed at 30°C in a reaction mixture (1 ml) containing 50 mM Hepes, pH 7.1, 5 mM MgCl₂, 25 mM KCl, 1 mM dithiothreitol, 0.6 mM NAD⁺, 0.5 mM mannose 6-phosphate, 2.5 U/ml glucose 6-phosphate dehydrogenase from *Leuconostoc mesenteroides,* and 10 µg/ml phosphoglucose isomerase with 10 µl of an extract containing 5–20 mg protein/ml. Control and PMM deficient measures are mean values ± SD. Where two data are given, the values were obtained on two different subcultures.

broblasts (Van Schaftingen and Jaeken 1995; Jaeken et al. 1997*a*; Matthijs et al. 1997*b*), ~30 did not show PMM deficiency. On the rationale that type I glycosylation pattern may be due to other enzymatic defects in the pathway leading to GDP-mannose, the activity of phosphomannose isomerase, which catalyzes its first step, was systematically analyzed in the samples from the latter patients. As shown in table 1, three of them showed a marked deficiency in this enzyme, whereas PMM activity was normal. Experiments in which extracts were mixed indicated that the deficiency was not due to the presence of an inhibitor. As reported elsewhere (Van Schaftingen and Jaeken 1995), PMI activity was normal in PMM-deficient patients. For patient C, the PMI deficiency was independently found at about the same time by other investigators (Niehues et al., 1998).This patient, who had a similar liver disease associated with protein-losing enteropathy (Freeze et al. 1997; Niehues et al., 1998) as patient A (see below) was not further investigated in the present study. The parents of patient B, but not of patient A, were available for testing; surprisingly, normal activities of PMI were found

in their leukocytes (mother 5.5 mU/mg protein; father 6.7 mU/mg protein; controls 5.4 $\pm$ 0.9 for $n = 6$).

On the basis of the available sequence of the PMI cDNA (Proudfoot et al. 1994), primers were designed for the amplification and analysis of PMI. Primers PMI-Biot (5′-Biotin-CGA GCA TGG CCG CTC CGC-3′) and PMI-R2 (5′-CCA GCT CTT CCA GGG TAT AC-3′) were used to amplify the PMI cDNA from total RNA isolated from fibroblasts according to established procedures (Sambrook et al. 1989; Matthijs et al. 1997*b*). The amplified fragment was sequenced using three fluorescently labeled primers: PMI-F1: 5′-FITC-CCA ATC AGG AAC TGA AAC TC-3′, PMI-F2: 5′-FITC-ACA GGG TTG GCA CTT CAA TG-3′, and PMI-F3: 5′-FITC-GAA TTT AGG GTG GCT GGC AG-3′. The cDNA fragment generated using primer PMI-R2 and PMI-F1 (identical to PMI-Biot but lacking the biotin) was cloned in pCR.2 by use of the TA-cloning kit (Invitrogen). The mutations were retrieved by sequencing with PMI-F1. To confirm the phase of the mutations, two clones representing the two alleles of the patients were entirely sequenced using the previously described primers and the M13-reverse and universal primers.

Patient A was found to be compound heterozygous for a C→T mutation at position 304, causing a Ser→Leu substitution at codon 102, and a T→C mutation at position 413, replacing Met at codon 138 by Thr (fig. 2). Both mutations concern positions that are conserved from *Saccharomyces cerevisiae* and *Candida albicans* to human PMI (Proudfoot et al. 1994). The corresponding amino acids are situated near the active site, as determined by X-ray crystallography of the *C. albicans* enzyme (Cleasby et al. 1996). The Glu residue adjacent to Met138 (Glu138 in *C. albicans* corresponding to Glu137 in human PMI) is one of the four residues involved in binding the catalytic Zn^{2+} ion. The Lys residue at position 100 in *C. albicans* (corresponding to Lys99 in human PMI) points inward into the active site, and the mutation at Ser102 may well affect this orientation. Thus, the mutations in the PMI gene are likely to explain the lack of activity of this enzyme in patient A.

Except for a polymorphism at codon 377 (GTA/GTG, both coding for Val), no sequence variation was detected in PCR-amplified cDNA from patient B. The presence of the polymorphism allowed us to select and fully sequence two clones derived from different alleles. Since there is no mutation, the low activity in fibroblasts of patient B could be due to a decrease in the transcription of the gene because of mutations either in the promoter or in a transcription factor required for the expression of the enzyme. The finding of a normal activity in leukocytes of the parents argues for this second possibility, if at least the same isozyme is expressed in fibroblasts and leukocytes. The fact that the supposed genetic defect in patient B is different from that in patient A probably

```
canal  95 LPFLFKVLSIEKVLSIQAH 113
sacce  93 LPFLFKVLSIEKVLSIQAH 111
emeni  91 LPFLFKVLSIRKALSIQAH 109
caeel  83 LSFLFKVLSVLGPLSIQIH 101
human  94 LPFLFKVLSVETPLSIQAH 112
              L        *  *

canal 129 NYPDDNHKPEMAIAVT 145
sacce 127 NYPDDNHKPEMAIAVT 143
emeni 125 NYPDDNHKPEMTIAIT 141
caeel 117 NYPDDNHKPEIAIALT 133
human 128 HYPDANHKPEMAIALT 144
                 *T
```

Figure 2 Sequence alignment showing the position of the mutations in patient A. The sequences of the enzymes from *Candida albicans* (canal), *Saccharomyces cerevisiae* (sacce), *Emericella nidulans* (emeni), *Caenorhabditis elegans* (caeel), and man (human) are shown. Numbering of the residues starts from the initiator methionine. Conserved residues are in bold. Three of the four residues liganding Zn^{2+} in the *C. albicans* enzyme are indicated by an asterisk (*). The mutated residues are shown below the alignment.

explains why the enzyme deficiency is not as pronounced in the former.

The clinical description of patient A has been briefly reported elsewhere (Billette de Villemeur et al. 1995). He was the third child of unrelated Lebanese parents. His siblings were healthy. He was admitted at the age of 5.5 mo for chronic diarrhea since the age of 3 mo and hypoglycemia with convulsions, coma, and apnea. There was no dysmorphy. Liver was 6 cm below the costal margin and firm. Stool culture showed *Pseudomonas aeruginosa*. Hypoglycemia tended to recur particularly after 4 h following feeding. Liver biopsy revealed fibrosis of the portal spaces and microvesicular steatosis. At the age of 10 mo, he was readmitted with generalized edema secondary to hypoalbuminemia. Weight was 6.7 kg ($< P_3 = 7.8$), length 65 cm ($< P_3 = 69$), and head circumference 43 cm ($< P_3 = 44.2$). Gross motor development was moderately retarded. Liver was unchanged. Liver function testing showed slightly increased bilirubin (50 mM) and glutamate-oxaloacetate transaminase (82 U/liter). IgG (2.6 g/liter) and IgA (0.32 g/liter) were decreased. Prothrombin time was 63%. Several clotting and anticlotting factors were decreased: factor XI (20%), antithrombin III (45%), protein C (29%) and total protein S (51%). There was intermittent slight proteinuria and moderate hyperaminoaciduria (particularly of glycine). Further extensive metabolic screening did not reveal the cause of the hypoglycemias. Serum transferrin isoelectrofocusing, however, showed the typical CDG syndrome type I pattern (decreased hexa-, penta- and tetrasialotransferrin, and increased di- and asialotransferrin) but phosphomannomutase activity in fibroblasts was normal. Radiological examination of the intestine suggested malabsorption, with fragmentation

and flocculation of the contrast substance. Sweat test was normal as well as intestinal disaccharidase activities. On endoscopy, fragility of the intestinal mucosa with easy bleeding was noted. Histology of duodenal biopsies showed partial villus atrophy with hypercellularity and only rare and discrete lymphangiectasias. Repeat liver biopsy at the age of one year showed a more pronounced picture of congenital hepatic fibrosis with stellate fibrosis of the portal spaces and proliferation of numerous biliary canaliculi without any inflammation and without notable alteration of the hepatocytes. Renal hyperechogenicity was found on ultrasound examination. Treatment consisted of Diazoxide, which normalized glycemia, regular intravenous administration of immunoglobulins and albumin, and nasogastric tube feeding with only a hydrolyzed, MCT-supplemented milk. He suffered from frequent bacterial as well as viral gastroenteritis. At the age of 26 months, weight was 9 kg ($< P_3 = 10.8$) and length 77.5 cm ($< P_3 = 82$). He had a large abdomen with pronounced collateral circulation, numerous disseminated angiomas, and unchanged hepatomegaly. Since oral feeding was still not supported, the nasogastric tube feeding was changed to tube feeding via a gastrostomy. Neurological examination was normal. He was last seen at the age of 2 9/12 years with persisting protein-losing enteropathy. However, psychomotor development was normal. He died at the age of 4 years from an unknown cause.

Patient B had a similar clinical picture consisting mainly of liver disease, which disappeared after the introduction of solid food at 3 mo. The case will be more extensively presented elsewhere.

PMM deficiency results in a multisystem disease with severe neurological involvement and dysmorphy, besides, usually mild, liver, intestinal, and other organ disease (Jaeken et al. 1993, 1997a, 1997b). PMI deficiency causes the same kind of sialotransferrin electrofocusing pattern as PMM deficiency, but its clinical presentation appears to be quite different in that the disease was enterohepatic in patients A and C and hepatic with mild neurological signs in patient B. Different explanations can be provided for the fact that the clinical manifestations of PMI deficiency are mainly hepatic and intestinal. A first one is the possible existence of different isozymes of PMI, as is the case for PMM (Matthijs et al. 1997a, 1997b; Pirard et al. 1997). The cDNA that has been identified and sequenced by Proudfoot et al. (1994) appears to be derived from a gene that is only poorly expressed in liver, as shown by northern blot analysis. Still, the fact that mutations in this gene lead to liver involvement in patient A indicates that this form of PMI is expressed in liver. Furthermore, all human expressed sequence tags found in data banks seem to be derived from this gene, except maybe one that could possibly encode an isozyme. However, we have been unable to PCR-amplify human liver and leukocyte cDNA with primers designed on the basis of this sequence.

A second explanation stems from the different position of PMM and PMI in the pathway leading to the formation of GDP-mannose. Both enzyme deficiencies are expected to cause a decrease in the biosynthesis of dolichol-pyrophosphate-oligosaccharide in the endoplasmic reticulum (Powell et al. 1994; Krasnewich et al. 1995), and both defects are indeed known to be lethal in yeast (Kepes and Schekman 1988). There is, however, a major difference between PMM and PMI in that the latter, but not the former, can be by-passed by the low-K_m hexokinases. These enzymes can indeed convert mannose to mannose 6-phosphate, displaying for this substrate a rather high affinity with K_m 5 μM in the case of the brain enzyme hexokinase I (Sols and Crane 1954). In contrast, glucokinase—also known as "hexokinase IV"—the major, if not the only, hexokinase present in hepatocytes (Reyes and Cardenas 1984) has a very low affinity for mannose with a K_m equal to 33 mM according to Weinhouse (1976) and is therefore rather inefficient to phosphorylate this substrate at the concentrations prevailing in plasma (50 μM in controls; Etchison and Freeze 1997). Blood mannose, which derives from glycoprotein degradation and alimentary mannose in unknown proportion, is therefore most likely sufficient to support N-glycosylation in tissues other than the liver. The exception appears also to apply to the intestinal mucosa, as evidenced by the marked enteropathy found in patients A and C (Niehues et al. 1988; Freeze et al. 1997). Remarkably, the specific activity of PMM is more than four times higher in the rat mucosa than in other tissues, which indicates that its GDP-mannose requirements are elevated (M. Pirard and E. Van Schaftingen, unpublished data). Furthermore, it is a tissue in which the glucose concentration is elevated, due to active absorption from the intestinal lumen and, during postabsorptive period, to equilibration with blood glucose through Glut2 (Thorens et al 1988). Glucose can therefore competitively inhibit the phosphorylation of mannose. Difference in tissue involvement of PMI deficiency could also possibly arise from difference in expression of the specific, low K_m mannose transporter (Panneerselvam and Freeze 1996), which appears to be involved in the preferential utilization of radiolabeled mannose over radiolabeled glucose in cultured fibroblasts (Panneerselvam et al. 1997).

A logical consequence of the fact that hexokinase phosphorylates mannose to mannose 6-phosphate is that PMI deficiency, unlike PMM deficiency (Jaeken and Casaer 1997; Marquardt et al. 1997; Mayatepek et al. 1997), should be treatable by administration of mannose supplements. This appears to be the case for patient C (Niehues et al. 1988; Freeze et al. 1997), in which we

have also found PMI deficiency. Unfortunately, patient A died before trials with mannose could be attempted. It would be interesting to know if the improvement of the symptomatology observed in patient B following introduction of solid food is due to the presence of mannose in the latter.

Because of its very different clinical presentation, this new type of CDG syndrome was probably not recognized as a glycosylation defect until now. In 1985, the same clinical picture as that of patient A and C was reported as a new syndrome in four infants from the northeastern part of Quebec. They died between the ages of 4 and 21 mo. They had also antithrombin III deficiency (Pelletier et al. 1985), a typical feature of CDG syndromes. We strongly suspect that these patients had a PMI deficiency. In conclusion, PMI deficiency is a newly recognized, lethal but potentially treatable CDG syndrome that should be searched for in unexplained enteropathy and/or liver disease.

JAAK JAEKEN,[1] GERT MATTHIJS,[2] JEAN-MARIE SAUDUBRAY,[3] CARLO DIONISI-VICI,[4] ENRICO BERTINI,[5] PASCALE DE LONLAY,[3] HUGUES HENRI,[6] HUBERT CARCHON,[1] ELS SCHOLLEN,[2] AND EMILE VAN SCHAFTINGEN[7]

[1]Department of Pediatrics, and [2]Center for Human Genetics, University of Leuven, Leuven; [3]Clinique de Génétique Médicale, Hôpital des Enfants Malades, Paris; Departments of [4]Metabolism and [5]Neuropediatrics, "Bambino Gesù" Hospital, Rome; [6]Laboratoire Central de Chimie Clinique, Centre Hospitalier Universitaire Vaudois, Lausanne; and [7]Laboratory of Physiological Chemistry, International Institute of Cellular and Molecular Pathology, University of Louvain, Brussels

References

Billette de Villemeur T, Poggi-Travert F, Laurent J, Jaeken J, Saudubray JM (1995) Le syndrome d'hypoglycosylation des protéines: un nouveau groupe de maladies héréditaires d'expression multisystémique. In: Arthuis M, Beaufils F, Caille B, Dommergues JP, Fontaine JL, Griscelli C, Job JC, et al (eds) Journées Parisiennes de Pédiatrie 1995. Paris. Médecine-Sciences Flammarion, 119–124

Cleasby A, Wonacott A, Skarzynski T, Hubbard RE, Davies GJ, Proudfoot AE, Bernard AR, et al (1996) The X-ray crystal structure of phosphomannose isomerase from Candida albicans at 1.7 angstrom resolution. Nat Struct Biol 3: 470–479

Etchison JR, Freeze HH (1997) Enzymatic assay of D-mannose in serum. Clin Chem 43:533–538

Freeze H, Niehues R, Hasilik M, Marquardt T, Etchison J, Panneerselvam K, Harms K (1997) Initial results of mannose therapy in N-glycosylation disorders. Glycobiology 7:1020

Jaeken J, Artigas J, Barone R, Fiumara A, De Koning TJ, Poll-The BT, de Rijk-van Andel JF, et al (1997a) Phosphomannomutase deficiency is the main cause of carbohydrate-deficient glycoprotein syndrome with type I isoelectrofocusing pattern of serum sialotransferrins. J Inherit Metab Dis 20: 447–449

Jaeken J, Carchon H, Stibler H (1993) The carbohydrate-deficient glycoprotein syndromes: pre-Golgi and Golgi disorders? Glycobiology 3:423–428

Jaeken J, Casaer P (1997) Carbohydrate-deficient glycoconjugate (CDG) syndromes: a new chapter of neuropaediatrics. Eur J Paediatr Neurol 2/3:61–66

Jaeken J, Matthijs G, Barone R, Carchon H (1997b) Carbohydrate deficient glycoprotein (CDG) syndrome type I. J Med Genet 34:73–76

Jaeken J, van Eijk HG, van der Heul C, Corbeel L, Eeckels R, Eggermont E (1984) Sialic acid-deficient serum and cerebrospinal fluid transferrin in a newly recognized genetic syndrome. Clin Chim Acta 144:245–247

Kepes F, Schekman R (1988) The yeast SEC53 gene encodes phosphomannomutase. J Biol Chem 263:9155–9161

Krasnewich DM, Holt GD, Brantly M, Skovby F, Redwine J, Gahl WA (1995) Abnormal synthesis of dolichol-linked oligosaccharides in carbohydrate-deficient glycoprotein syndrome. Glycobiology 5:503–510

Marquardt T, Hasilik M, Niehues R, Herting M, Muntau A, Holzbach U, Hanefeld F, et al (1997) Mannose therapy in carbohydrate-deficient glycoprotein syndrome type 1 - first results of the German multicenter study. Amino Acids 12: 389

Matthijs G, Schollen E, Pardon E, Veiga-da-Cunha M, Jaeken J, Cassiman J-J, Van Schaftingen E (1997a) Mutations in PMM2, a phosphomannomutase gene on chromosome 16p13, in carbohydrate-deficient glycoprotein type I syndrome (Jaeken syndrome). Nat Genet 16:88–92

Matthijs G, Schollen E, Pirard M, Budarf ML, Van Schaftingen E, Cassiman J-J (1997b) PMM (PMM1), the human homologue of SEC53 or yeast phosphomannomutase, is localized on chromosome 22q13. Genomics 40:41–47

Matthijs G, Schollen E, Van Schaftingen E, Cassiman J-J, Jaeken J (1998) Lack of homozygotes for the most frequent disease allele in carbohydrate-deficient-glycoprotein syndrome type Ia. Am J Hum Genet 62:542–550

Mayatepek E, Schröder M, Kohlmüller D, Bieger WP, Nützenadel W (1997) Continuous mannose infusion in carbohydrate-deficient glycoprotein syndrome type I. Acta Paediatr 86:1138–1140

Niehues R, Hasilik M, Alton G, Körner C, Schiebe-Sukumar M, Koch HG, Zimmer K-P, et al (1998) Carbohydrate-deficient glycoprotein syndrome type Ib: phosphomannose isomerase deficiency and mannose therapy. J Clin Invest 101: 1414–1420

Panneerselvam K, Etchison JR, Freeze HH (1997) Human fibroblasts prefer mannose over glucose as a source of mannose for N-glycosylation. Evidence for the functional importance of transported mannose. J Biol Chem 272: 23123–23129

Panneerselvam K, Freeze HH (1996) Mannose enters mammalian cells using a specific transporter that is insensitive to glucose. J Biol Chem 271:9417–9421

Pelletier VA, Galeano N, Brochu P, Morin CL, Weber AM, Roy CC (1986) Secretory diarrhea with protein-losing enteropathy, enterocolitis cystica superficialis, intestinal lymphangiectasia, and congenital hepatic fibrosis: a new syndrome. J Pediatr 108:61–65

Pirard M, Collet JF, Matthijs G, Van Schaftingen E (1997) Comparison of PMM1 with the phosphomannomutases expressed in rat liver and in human cells. FEBS Lett 411: 251–254

Powell LD, Paneerselvam K, Vij R, Diaz S, Manzi A, Buist N, Freeze H, et al (1994) Carbohydrate-deficient glycoprotein syndrome: not an N- linked oligosaccharide processing defect, but an abnormality in lipid-linked oligosaccharide biosynthesis? J Clin Invest 94:1901–1909

Proudfoot AE, Turcatti G, Wells TN, Payton MA, Smith DJ (1994) Purification, cDNA cloning and heterologous expression of human phosphomannose isomerase. Eur J Biochem 219:415–423

Reyes A, Cardenas ML (1984) All hexokinase isoenzymes coexist in hepatocytes. Biochem J 221:303–309

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2d ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sols A, Crane RK (1954) Substrate specificity of brain hexokinase. J Biol Chem 210:581–595

Stibler H, Jaeken J (1990) Carbohydrate deficient serum transferrin in a new systemic hereditary syndrome. Arch Dis Child 65:107–111

Thorens B, Sarkar HK, Kaback HR, Lodish HF (1988) Cloning and functional expression in bacteria of a novel glucose transporter present in liver, intestine, kidney and β-pancreatic islet cells. Cell 55:281–290

Van Schaftingen E, Jaeken J (1995) Phosphomannomutase deficiency is a cause of carbohydrate-deficient glycoprotein syndrome type I. FEBS Letters 377:318–320

Weinhouse S (1976) Regulation of glucokinase in liver. Curr Top Cell Regul 11:1–50

Address for correspondence and reprints: Dr. J. Jaeken, University Hospital Gasthuisberg, Herestraat 49, B-3000 Leuven, Belgium. E-mail: Jo.Vencken@uz.kuleuven.ac.be

---

Temperature-Sensitive Phenotypes of Peroxisome-Assembly Processes Represent the Milder Forms of Human Peroxisome-Biogenesis Disorders

To the Editor:

Peroxisome-biogenesis disorders (PBDs) are lethal hereditary diseases caused by abnormalities in the assembly processes of peroxisomes (Moser et al. 1995). The peroxisome is a ubiquitous organelle involved in vital metabolic functions, such as oxidative processes involving H_2O_2, β-oxidation of fatty acids, and biosynthesis of plasmalogens (Van den Bosch et al. 1992). PBDs are characterized by multiple defects in these functions, as well as by the lack of morphologically normal peroxisomes. They are genetically classified into complementation groups (CGs), the number of which is ≥ 11 (Shimozawa et al. 1993; Moser et al. 1995; Poulos et al. 1995). Each CG contains significantly different clinical phenotypes—for example, Zellweger syndrome (ZS), neonatal adrenoleukodystrophy (NALD), and infantile Refsum disease (IRD). ZS patients have severe neurological abnormalities, dysmorphic features, hepatomegaly, and multiple renal cysts, and most die at age <6 mo. NALD patients have similar symptoms, but they survive considerably longer, dying during early childhood. In contrast, IRD patients do not exhibit significant abnormalities in the CNS, and they have the longest average life span among patients with PBDs (Lazarow and Moser 1995; Moser et al. 1995). Although the causal genes (*PEX*s) for several CGs have been cloned and the mutations have been identified at the molecular level (Shimozawa et al. 1992 ; Dodt et al. 1995; Wiemer et al. 1995; Fukuda et al. 1996; Yahraus et al. 1996; Chang et al. 1997; Okumoto and Fujiki 1997; Portsteffen et al. 1997; Reuber et al. 1997), it is unknown why such diverse clinical phenotypes occur in the same CGs although, in all CGs, the phenotypes are very similar. We report that milder forms of PBDs are characterized by temperature-sensitive (TS) phenotypes of peroxisome-assembly processes in the fibroblasts of patients.

In spite of the variations in the clinical features, the fibroblasts from patients of all three PBD phenotypes generally lack peroxisomes. Although the occurrence of a reduced number of peroxisomes occasionally has been noted in several PBD cell lines (Arias et al. 1985; Wiemer et al. 1991; Slawecki et al. 1995), no correlation with clinical features has been apparent. We assumed that limited types of leaky mutations in the *PEX* genes could be the causes of the milder forms of PBDs. As a possible parameter representing such leakiness, we examined temperature sensitivity. Fibroblasts from PBD patients with different CGs were incubated at 30°C and at 37°C and were subjected to immunofluorescence staining with anti-catalase antibody. After 72 h incubation at 30°C, punctate staining of catalase typical of peroxisomes was detected in the fibroblasts of all six patients with IRD and in three of five of those with NALD, belonging to four different CGs (fig. 1*b* and table 1), whereas no peroxisomes appeared in the same cells after incubation at 37°C (fig. 1*a*). Catalase and the 70-kD peroxisomal membrane protein (PMP70) were colocalized in these
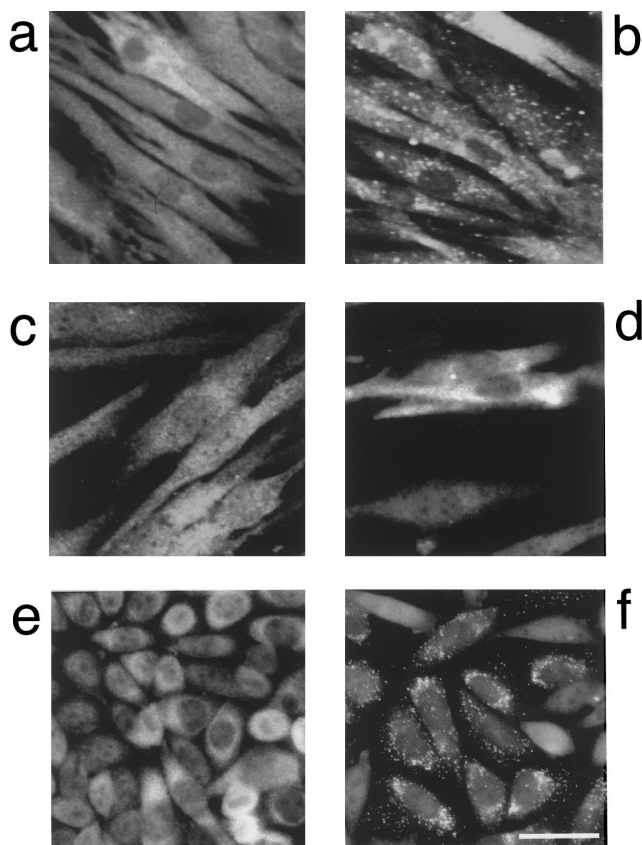
Figure 1 Immunofluorescence staining of peroxisomes in patients' fibroblasts and in Z65 mutant Chinese-hamster-ovary cells. Cells were cultured for 72 h at either 37°C (*a, c,* and *e*) or 30°C (*b, d,* and *f*) and were stained with either anti–human catalase antibody (*a–d*) or anti–rat catalase rabbit antibody (*e* and *f*). *a* and *b,* Fibroblasts of an IRD patient (F-05). *c* and *d,* Fibroblasts of a ZS patient (F-01). *e* and *f,* Z65 transformant with *Pex2*[E55K]. (Scale bar = 50 μm)

cells after incubation for 72 h at 30°C (data not shown), thereby confirming the identity of these catalase-positive granules as peroxisomes. In the fibroblasts of ZS patients with any CGs tested, no peroxisomes were found at either 30°C or 37°C incubation (fig. 1*c* and *d* and table 1). Fibroblasts of normal controls had punctate patterns of catalase at both 30°C and 37°C incubation (data not shown).

We confirmed the formation of functional peroxisomes in the TS cells at 30°C, on the basis of three biochemical criteria: the peroxisomal β-oxidation activity of very-long-chain fatty acids (Suzuki et al. 1991), the activity of peroxisomal dihydroxyacetonephosphate acyltransferase (Shimozawa et al. 1988) involved in the plasmalogen biosynthesis, and the proteolytic processing of peroxisomal acyl-CoA oxidase (Tsukamoto et al. 1995), which occurs within the peroxisomes. These activities were deficient in the ZS cells (F-01), at both 30°C and 37°C incubation, whereas in the TS cells (F-05),

these functions were markedly improved by incubation at 30°C but not by incubation at 37°C (data not shown).

To identify the mutation responsible for the TS phenotype, the *PEX2* gene (GenBank), the causal gene for CG-F (Shimozawa et al. 1992), was amplified by PCR from the genomic DNA of TS cells F-05. PCR reaction was performed by use of *PEX2* gene–specific antisense (complementary to the *PEX2* cDNA sequence of positions 967–991, 5′-ATA CTT AGG ATG ACT AAT ATT AAG-3′) and sense (an intron sequence of the *PEX2* gene, starting 48 bp upstream of the 5′ end of the coding exon, 5′-CAA GAT TGC AAC TCT TTG CTA ATG-3′) oligonucleotides. PCR conditions were as follows: initial heating for 1 min at 94°C, followed by 40 cycles of denaturation for 1 min at 94°C, annealing for 2 min at

Table 1

Temperature Sensitivity of Peroxisome Biogenesis in Fibroblasts of PBD Patients

CG AND PATIENT[a]	PHENOTYPE	PEROXISOME-POSITIVE CELLS INCUBATED AT[b] (%) 37°C	30°C	AGE AT DEATH OR LAST FOLLOW-UP[c]
A (8):				
A-06	ZS	0	10	4 mo
A-05	NALD	0	90	
A-08	NALD	0	80	3 years 1 mo
A-04[d]	IRD	0	60	
C (4):				
C-03[e]	ZS	0	0	8 mo
C-08	ZS	0	0	4 mo
E (1):				
E-14	ZS	0	0	4 mo
E-01[f]	NALD	5	5	2 years 9 mo
E-13	NALD	1	1	1 year 8 mo
E-05[d]	IRD	0	90	
E-24	IRD	0	60	1 year 7 mo[g]
E-25	IRD	0	60	10 years 7 mo[g]
E-26	IRD	0	50	6 years 1 mo[g]
F (10):				
F-01[h]	ZS	0	0	8 mo
F-04[i]	ZS	0	0	3 mo
F-05	IRD	0	70	
...[j] (6):				
6-01	NALD	5	80	4 years 6 mo

[a] The letter designation is that provided by Gifu University (Japan), and the number designation (in parentheses) is that provided by the Kennedy Krieger Institute.

[b] Data are averages of several view fields, at × 200.

[c] Data are for traceable cases only.

[d] Purchased from Coriell Cell Repositories (Camden, NJ); the cell line designations are −−−GM08771 (A-04) and GM08770 (E-05).

[e] Source: Fukuda et al. (1996).

[f] Source: Maeda et al. (1990).

[g] Parents were alive after the age at the last follow-up (i.e., age shown).

[h] Source: Shimozawa et al. (1992).

[i] Source: Shimozawa et al. (1998).

[j] No designation of CG was available from Gifu University.

55°C, and extension for 3 min at 72°C. Nucleotide-sequence comparison with the normal *PEX2* gene revealed that this patient was heterozygous for two point mutations. One was a G→A substitution at nucleotide position 163,, relative to the A residue of the initiation codon, causing an amino acid alteration (E55K). The other was a C→T substitution at nucleotide position 355, resulting in the change of the codon 119 to a stop codon, TGA (R119 Stop). These two mutations were also found in the *PEX2* cDNA obtained by reverse transcriptase–PCR of the mRNA from F-05. The latter nonsense mutation is identical to that reported by Shimozawa et al. in both ZS patient F-01 (Shimozawa et al. 1992) and another CG-F patient (Shimozawa et al. 1993) (both cases were homozygous for the mutation), and it previously had been established that this mutation is nonfunctional. Accordingly, we investigated the relationship between the E55K mutation and the TS phenotype, by gene transfection. The *PEX2*^{E55K} gene sequence subcloned in the expression vector pUcD2SRαMCS (Tsukamoto et al. 1995) was transfected to a *PEX2*-deficient Chinese-hamster-ovary cell mutant (Z65) (Tsukamoto et al. 1991, 1994), and stable transformants were produced. The transformants revealed a punctate distribution of catalase after 72 h incubation at 30°C, whereas no catalase-positive granules were observed for incubation at 37°C (fig. 1*e* and *f*); Z65 transfected with wild-type *PEX2* had catalase-positive granules at both 30°C and 37°C; and the cells transfected with the empty vector exhibited no peroxisomal staining at either 30°C or 37°C (data not shown). Thus, the TS phenotype of peroxisome biogenesis of the IRD fibroblasts (F-05) is caused by the E55K mutation of the *PEX2* gene.

The present results indicate that the peroxisome-assembly process is TS in the fibroblasts of patients with the mildest form of PBD (i.e., IRD), irrespective of CGs, and that this phenotype is directly linked to the specific genotype of the responsible *PEX* gene, at least in F-05. Such a TS phenotype was not observed for the most severe form of PBD (i.e., ZS), whereas only a subset of the cell lines were TS for NALD, the intermediate form of PBD. In this regard, it is interesting to note that the NALD patients with the TS phenotype (patients A-08 and 6-01) had longer life spans than did those with the non-TS phenotype (patients E-01 and E-13), even though the latter two patients exhibited slight leakiness at both temperatures. Thus, the TS phenotypes of peroxisome assembly in the cultured fibroblasts represent the mildness of the clinical symptoms of PBD. Patients with the TS phenotypes may be mosaic for peroxisome occurrence from cell to cell in the body at normal body temperatures. A mosaicism of peroxisomes was indeed reported in the liver of a PBD patient who had a relatively long life span (Giros et al. 1996). It is also possible that TS patients have partially functional peroxisomes.

In any case, the TS patients probably have higher gross peroxisomal activities than do the patients with non-TS leaky phenotypes (patients E-01 and E-13).

Temperature- or cold-sensitive phenotypes have been noted in a few genetic diseases. In epidermolysis bullosa simplex, the disturbance of the skin becomes worse at higher temperatures, whereas, in paramyotonia congenita, exposure to lower temperatures causes myotonia. In these instances, the symptoms are understood to be direct effects of the temperature- and cold-sensitive phenotypes of the corresponding gene products, keratin (Morley et al. 1995) and Na$^+$-channel protein (McClatchey et al. 1992), respectively. In maple syrup–urine disease, symptoms sometimes worsen when there is a high fever (Chuang and Shih 1995); however, the responsible mutation has not been identified. Thus, TS phenotypes directly linked to specific genotypes possibly occur in various genetic diseases. Among these, PBD cases are unique in that a complex cellular process, peroxisome assembly, becomes TS because of a single gene mutation, causing distinct clinical features.

Our present results would raise several clinical implications. First, the severity of prognosis could be diagnosed, by examination of the temperature sensitivity of peroxisome assembly in the fibroblasts of newborn PBD patients. Second, precaution against fever may be necessary in the treatment TS PBD patients. Third, hypothermal therapy might be applicable to TS PBD patients. Such therapy might increase the frequency of cells having functional peroxisomes, thereby improving clinical symptoms.

ATSUSHI IMAMURA,[1] TOSHIRO TSUKAMOTO, [1]
NOBUYUKI SHIMOZAWA,[2] YASUYUKI SUZUKI,[2]
ZHONGHI ZHANG,[2] TSUNEO IMANAKA,[4] YUKIO FUJIKI,[5]
TADAO ORII,[3] NAOMI KONDO,[2] AND TAKASHI OSUMI[1]
[1]*Department of Life Science, Himeji Institute of Technology, Kamigori, Hyogo, Japan;* [2]*Department of Pediatrics, Gifu University School of Medicine, and* [3]*Faculty of Human Welfare, Chubu Gakuin University, Seki, Gifu, Japan;* [4]*Department of Microbiology and Molecular Pathology, Faculty of Pharmaceutical Sciences, Teikyo University, Sagamiko, Kanagawa, Japan; and* [5]*Department of Biology, Faculty of Science, Kyushu University, Fukuoka, Japan*

Electronic-Database Information

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

GenBank,   http://www.ncbi.nih.gov/Web/Search/index/html (for human peroxisome-assembly factor-1 [*hsPEX2;* accession number M86852])

References

Arias JA, Moser AB, Goldfischer SL (1985) Ultrastructural and cytochemical demonstration of peroxisomes in cultured fibroblasts from patients with peroxisomal deficiency disorders. J Cell Biol 100:1789–1792

Chang CC, Lee WH, Moser H, Valle D, Gould SJ (1997) Isolation of the human PEX12 gene, mutated in group 3 of the peroxisome biogenesis disorders. Nat Genet 15:385–388

Chuang DT, Shih VE (1995) Disorders of branched chain amino acid and keto acid metabolism. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) The metabolic and molecular bases of inherited disease. McGraw-Hill, New York, pp 1239–1278

Dodt G, Braverman N, Wong C, Moser A, Moser HW, Watkins P, Valle D, et al (1995) Mutations in the PTS1 receptor gene, PXR1, define complementation group 2 of the peroxisome biogenesis disorders. Nat Genet 9:115–125

Fukuda S, Shimozawa N, Suzuki Y, Zhang Z, Tomatsu S, Tsukamoto T, Hashiguchi N, et al (1996) Human peroxisome assembly factor-2 (PAF-2): a gene responsible for group C peroxisome biogenesis disorder in humans. Am J Hum Genet 59:1210–1220

Giros M, Roels F, Prats J, Ruiz M, Ribes A, Espeel M, Wanders RJ, et al (1996) Long survival in a case of peroxisomal biogenesis disorder with peroxisome mosaicism in the liver. Ann N Y Acad Sci 804:747–749

Lazarow PB, Moser HW (1995) Disorders of peroxisome biogenesis. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) The metabolic and molecular bases of inherited disease. McGraw-Hill, New York, pp 2287–2324

Maeda K, Terada A, Murakami Y, Terabe K, Oki S, Hattori M, Matsumoto N, et al (1990) Clinical and biochemical analysis of infant with neonatal adrenoleukodystrophy (in Japanese). Acta Paediatr Jpn 94:1869–1876

McClatchey AI, van den Bergh P, Pericak VM, Raskind W, Verellen C, McKenna YD, Rao K, et al (1992) Temperature-sensitive mutations in the III-IV cytoplasmic loop region of the skeletal muscle sodium channel gene in paramyotonia congenita. Cell 68:769–774

Morley SM, Dundas SR, James JL, Gupta T, Brown RA, Sexton CJ, Navsaria HA, et al (1995) Temperature sensitivity of the keratin cytoskeleton and delayed spreading of keratinocyte lines derived from EBS patients. J Cell Sci 108: 3463–3471

Moser AB, Rasmussen M, Naidu S, Watkins PA, McGuinness M, Hajra AK, Chen G, et al (1995) Phenotype of patients with peroxisomal disorders subdivided into sixteen complementation groups. J Pediatr 127:13–22

Okumoto K, Fujiki Y (1997) *Pex12* encodes an integral membrane protein of peroxisomes. Nat Genet 17:265–266

Portsteffen H, Beyer A, Becker E, Epplen C, Pawlak A, Kunau WH, Dodt G (1997) Human PEX1 is mutated in complementation group 1 of the peroxisome biogenesis disorders. Nat Genet 17: 449–452

Poulos A, Christodoulou J, Chow CW, Goldblatt J, Paton BC, Orii T, Suzuki Y, et al (1995) Peroxisomal assembly defects: clinical, pathologic, and biochemical findings in two patients in a newly identified complementation group. J Pediatr 127: 596–599

Reuber BE, Germain LE, Collins CS, Morrell JC, Ameritunga R, Moser HW, Valle D, et al (1997) Mutations in PEX1 are the most common cause of peroxisome biogenesis disorders. Nat Genet 17:445–448

Shimozawa N, Suzuki Y, Orii T, Moser A, Moser HW, Wanders RJ (1993) Standardization of complementation grouping of peroxisome-deficient disorders and the second Zellweger patient with peroxisomal assembly factor-1 (PAF-1) defect. Am J Hum Genet 52:843–844

Shimozawa N, Suzuki Y, Orii T, Yokota S, Hashimoto T (1988) Biochemical and morphologic aspects of peroxisomes in the human rectal mucosa: diagnosis of Zellweger syndrome simplified by rectal biopsy. Pediatr Res 24:723–727

Shimozawa N, Suzuki Y, Tomatsu S, Nakamura H, Kono T, Takada H, Tsukamoto T, et al (1998) A novel mutation, R125X in peroxisome assembly factor-1 responsible for Zellweger syndrome. Hum Mutat Suppl 1:S134–S136

Shimozawa N, Tsukamoto T, Suzuki Y, Orii T, Shirayoshi Y, Mori T, Fujiki Y (1992) A human gene responsible for Zellweger syndrome that affects peroxisome assembly. Science 255:1132–1134

Slawecki ML, Dodt G, Steinberg S, Moser AB, Moser HW, Gould SJ (1995) Identification of three distinct peroxisomal protein import defects in patients with peroxisome biogenesis disorders. J Cell Sci 108:1817–1829

Suzuki Y, Shimozawa N, Yajima S, Yamaguchi S, Orii T, Hashimoto T (1991) Effects of sodium 2-[5-(4-chlorophenyl)pentyl]-oxirane-2-carboxylate (POCA) on fatty acid oxidation in fibroblasts from patients with peroxisomal diseases. Biochem Pharmacol 41:453–456

Tsukamoto T, Miura S, Fujiki Y (1991) Restoration by a 35K membrane protein of peroxisome assembly in a peroxisome-deficient mammalian cell mutant. Nature 350:77–81

Tsukamoto T, Miura S, Nakai T, Yokota S, Shimozawa N, Suzuki Y, Orii T, et al (1995) Peroxisome assembly factor-2, a putative ATPase cloned by functional complementation on a peroxisome-deficient mammalian cell mutant. Nat Genet 11:395–401

Tsukamoto T, Shimozawa N, Fujiki Y (1994) Peroxisome assembly factor 1: nonsense mutation in a peroxisome-deficient Chinese hamster ovary cell mutant and deletion analysis. Mol Cell Biol 14:5458–5465

Van den Bosch H, Schutgens RBH, Wanders RJA, Tager JM (1992) Biochemistry of peroxisomes. Annu Rev Biochem 61: 157–197

Wiemer EA, Nuttley WM, Bertolaet BL, Li X, Francke U, Wheelock MJ, Anne UK, et al (1995) Human peroxisomal targeting signal-1 receptor restores peroxisomal protein im-

port in cells from patients with fatal peroxisomal disorders. J Cell Biol 130:51–65

Wiemer EA, Out M, Schelen A, Wanders RJ, Schutgens RB, van den Bosch H, Tager JM (1991) Phenotypic heterogeneity in cultured skin fibroblasts from patients with disorders of peroxisome biogenesis belonging to the same complementation group. Biochim Biophys Acta 1097:232–237

Yahraus T, Braverman N, Dodt G, Kalish JE, Morrell JC, Moser HW, Valle D, et al (1996) The peroxisome biogenesis disorder group 4 gene, PXAAA1, encodes a cytoplasmic ATPase required for stability of the PTS1 receptor. EMBO J 15:2914–2923

Address for correspondence and reprints: Dr. Takashi Osumi, Department of Life Science, Himeji Institute of Technology, Kamigori, Hyogo 678-1297, Japan. E-mail: osumi@sci.himeji-tech.ac.jp

A Breast Cancer Patient of Scottish Descent with Germ-Line Mutations in *BRCA1* and *BRCA2*

To the Editor:

Ramus et al. (1997) previously described an Ashkenazi Jewish patient found to have germ-line mutations in both breast and ovarian cancer–susceptibility genes, *BRCA1* and *BRCA2*. We report the first such example for the non-Jewish Caucasian population. The patient, who is indicated by an arrow in pedigree 232 (fig. 1), was of Scottish origin. She was diagnosed with breast cancer (grade 2 adenocarcinoma) at age 35 years. Simultaneous screening by protein truncation test of both *BRCA1* (exon 11) and *BRCA2* (exon 11) detected truncating
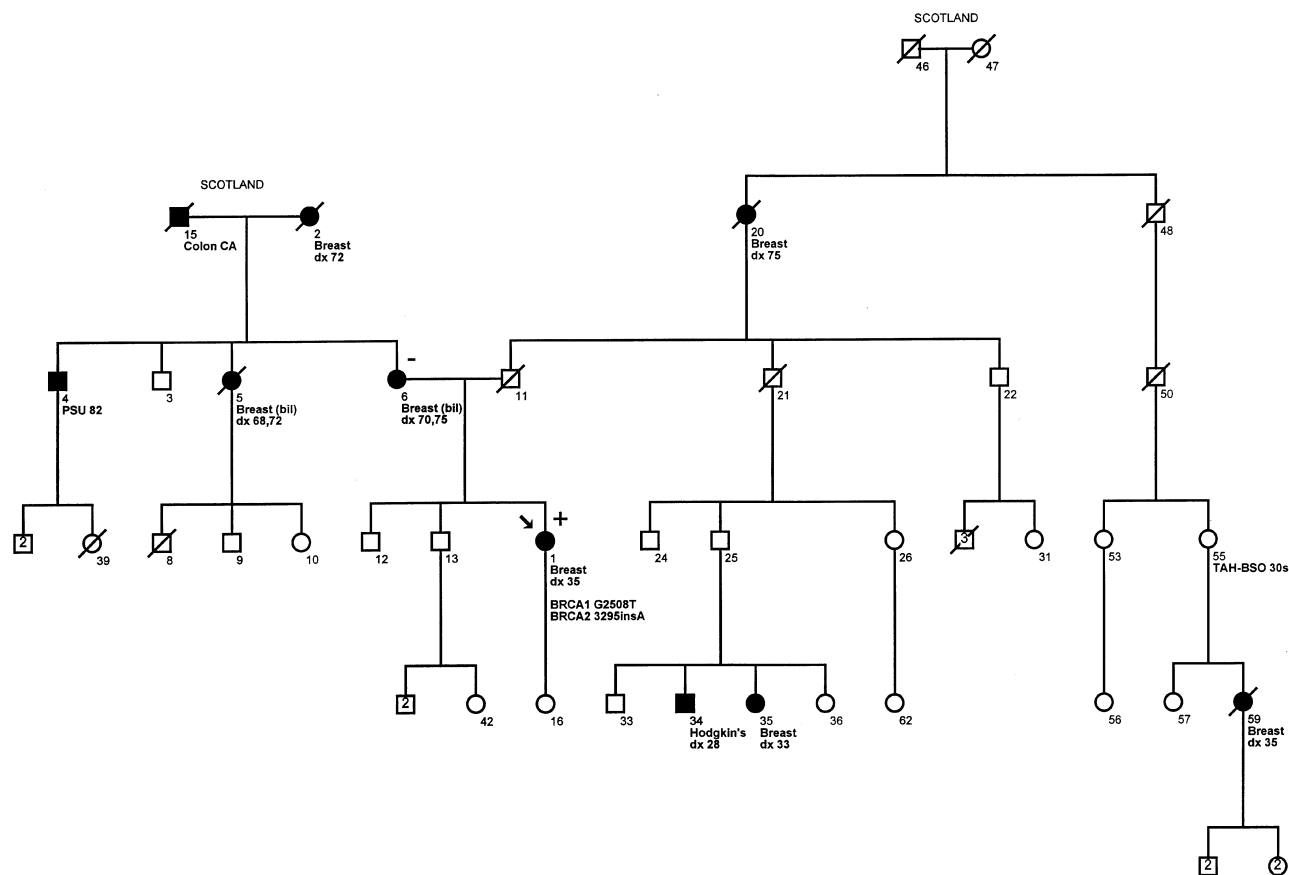

Figure 1   Pedigree of family 232. Blackened circles indicate affected women; blackened squares indicate affected men; and a diagonal slash indicates that the patient is deceased. Individual identification numbers appear directly below the symbols. Breast cancer is indicated with age at diagnosis ("dx"); bilateral breast cancer is indicated ("bil") with ages at diagnoses; "PSU" indicates primary site not known; and "TAH-BSO" indicates a complete abdominal hysterectomy, including bilateral oophorectomy. The plus sign (+) indicates the presence of the *BRCA1* G2508T mutation and the *BRCA2* 3295insA mutation, in the proband tested. The minus sign (−) indicates the absence of these two mutations in *BRCA1* and *BRCA2*.

mutations. Direct sequencing revealed *BRCA1* G2508T as a nonsense mutation resulting in Glu→stop codon and *BRCA2* 3295insA as an in-frame stop codon at position 1025. Mutation results were confirmed at a separate *BRCA1* and *BRCA2* testing facility in Toronto. These are novel mutations not previously reported in the Breast Cancer Information Core database.

This patient had a maternal and paternal history of breast cancer. The maternal side contained cases of post-menopausal breast cancer: her mother (patient 6) was diagnosed with bilateral breast cancer at ages 70 and 75 years, and her aunt (patient 5) was diagnosed at ages 68 and 72 years; her grandmother (patient 2) was diagnosed with breast cancer at age 72 years. The paternal side contained cases of premenopausal breast cancer: a cousin once removed (patient 35) was diagnosed with breast cancer at age 33 years, her grandmother (patient 20) was diagnosed at age 75 years, and a distant cousin (patient 59) was diagnosed at age 35 years. Interestingly, her mother did not have either mutation, suggesting that both *BRCA1* and *BRCA2* germ-line mutations originated from the father of the proband. This is consistent with the ages at onset of the women on the paternal side, indicating the presence of at least one of these mutations in the women diagnosed with premenopausal breast cancer. To date, no other family members are available for testing. For counseling of individuals identified as double heterozygotes for mutations in *BRCA1* and *BRCA2,* the risk of transmitting a breast cancer–susceptibility gene(s) to any offspring is 3/4.

The frequency of *BRCA1* and *BRCA2* mutations in the United Kingdom and Canada has been estimated at 1/850–1/500 individuals (Easton 1993; Ford et al. 1995); therefore, the likelihood of finding a double heterozygote in this population is between 1/700,000 and 1/250,000. It is estimated that, for individuals of Ashkenazi Jewish descent, the likelihood of being a carrier for one of three common *BRCA1* or *BRCA2* mutations is ⩾1/50 (Roa 1996; Tonin et al. 1996); therefore, the likelihood of finding a double or compound heterozygote is ~1/2,500 for Ashkenazi Jewish families unselected for cancer.

This individual is the first example to date of a double heterozygote for the high-penetrance breast cancer–susceptibility genes, *BRCA1* and *BRCA2,* outside of the Ashkenazi Jewish population (Ramus et al. 1997). This finding is predictably rare, with a maximum frequency of 1/250,000. Our patient was diagnosed with early-onset breast cancer at age 35 years, an age typical of other *BRCA1*/*BRCA2*–associated breast cancers. Her case does not suggest a more severe presentation or younger age at diagnosis for women found to harbor germ-line mutations in both genes. We cannot generalize that complete screening of *BRCA1* and *BRCA2* in families with a previously identified mutation in either gene

is necessary until other double heterozygotes are identified. However, if the family mutation is not found in women diagnosed with early-onset breast cancer or ovarian cancer, there is a basis for comprehensive screening of both genes.

ALEXANDER LIEDE, PAULINE REHAL, DANNY VESPRINI, ELAINE JACK, JOHN ABRAHAMSON, AND STEVEN A. NAROD
University of Toronto and Women's College Hospital and The Ontario Cancer Genetics Network, Toronto

Electronic-Database Information

URLs for data in this article are as follows:

Breast cancer information core database, http://www.nhgri.nih.gov/Intramural_research/Lab_transfer/Bic/

References

Ford D, Easton DF, Peto J (1995) Estimates of the gene frequency of BRCA1 and its contribution to breast and ovarian cancer incidence. Am J Hum Genet 57:1457–1462

Easton DF, Ford D, Peto J (1993) Inherited susceptibility to breast cancer. Cancer Surv 18:1–17

Ramus SJ, Friedman LS, Gayther SA, Ponder BAJ (1997) A breast/ovarian cancer patient with germline mutations in both *BRCA1* and *BRCA2*. Nat Genet 15:14–15

Roa BB, Boyd AA, Vocik K, Richards CS (1996) Ashkenazi Jewish population frequencies for common mutations in *BRCA1* and *BRCA2*. Nat Genet 14:185–187

Tonin P, Weber B, Offit K, Couch F, Rebbeck TR, Neuhausen S, Godwin AK, et al (1996) Frequency of recurrent *BRCA1* and *BRCA2* mutations in Ashkenazi Jewish breast cancer families. Nat Med 2:1179–1183

Evidence of Founder Mutations in Finnish BRCA1 and BRCA2 Families

To the Editor:

Approximately 4%–10% of breast and ovarian cancer is thought to be inherited (Newman et al. 1988). Most hereditary ovarian cancer cases and a significant portion

of breast cancer cases are due to mutations in the BRCA1 gene (Miki et al. 1994). In addition, BRCA2 mutations account for a large fraction of hereditary breast cancer but for only a small number of cases of hereditary ovarian cancer (Wooster et al. 1995). Other cancer phenotypes associated with mutations in these genes are prostate cancer and colon carcinoma, for BRCA1 (Ford et al. 1994), and prostate cancer and pancreatic carcinoma, for BRCA2 (Wooster et al. 1995). Most of the identified mutations appear to be highly penetrant, conferring a remarkably elevated lifetime cancer risk (Easton 1997). The proportion of breast/ovarian cancer families whose disease is attributable to BRCA1 or BRCA2 mutations varies, depending on the population studied. Although there seems to be a wide spectrum of different germ-line mutations in most countries, some geographically or ethnically isolated populations, such as Icelanders and Ashkenazi Jews, have shown clusterings of specific mutations (Szabo and King 1997; also see Breast Cancer Information Core).

Population history explains much of the enrichment of certain disease-related gene defects in Finland. The concept of Finnish disease heritage currently includes >30 diseases that are more prevalent in Finland than in other populations (de la Chapelle 1993). The majority of these diseases are rare autosomal recessive disorders, but founder mutations can also be seen in the more common diseases, such as hereditary nonpolyposis colon cancer (HNPCC) (Nyström-Lahti et al. 1995). Therefore, we were interested to see whether we could find any mutational clustering in the BRCA1 and BRCA2 genes. Also, the phenotypic characteristics of the mutation carriers were determined.

Eighty-eight families from the Oulu University Hospital area (population 720,000), each with either at least two cases of breast and/or ovarian cancer in first-degree relatives or other characteristics of inherited susceptibility (e.g., bilateral/multiple primary tumors or age at diagnosis <40 years) were analyzed. Expression of breast and ovarian cancer in these families is shown in table 1. Pedigree data were obtained from the probands by means of a questionnaire and an interview. Permission for genetic testing was requested during genetic counseling. The study was approved by the Ethical Committee of the Oulu University Medical Faculty.

The protein-encoding and splice site–junction regions of BRCA1 and BRCA2 were examined mainly by conformation-sensitive gel electrophoresis, the mutation-detection sensitivity of which is estimated to be in the range of 60%–95% (Körkkö et al. 1998). Exon 11 of BRCA1 and exons 10 and 11 of BRCA2 were screened by use of protein-truncation test (Håkansson et al. 1997). For sequencing, we used Cyclist Exo⁻*Pfu* DNA Sequencing Kit (Stratagene). The primer sets and PCR conditions have been described elsewhere (Castilla et al. 1994;

Table 1

Classification of 88 Cancer Families According to Occurrence of Breast and Ovarian Cancer in 1st-Degree Relatives

No. of Cases of Breast Cancer	No. of Families in Which No. of Cases of Ovarian Cancer =				
	0	1	2	3	4
0		1	2	1	1
1	13	9	2		
2	27	5	1		
3	14	1			
4	8				
5	1	1	1		

Friedman et al. 1994, 1997; Couch et al. 1996). The four microsatellite markers used to determine the disease-linked BRCA1 haplotypes were D17S846, D17S855, D17S1322, and D17S1323 (Genome Database).

Altogether, five distinct mutations were observed—two in BRCA1 and three in BRCA2 (table 2). The proportion of disease-related BRCA1 and BRCA2 mutations in the population studied was 12.5% (11/88). The mutation prevalence of true high-risk families was 17% (6/36), and that for moderate-risk families was 10% (5/52), suggesting that a strong family history is not the only parameter that can predict mutations; early disease onset, tumor bilaterality, and multiple primary tumors should also be taken into account. Cancer phenotypes are shown in table 2. BRCA1 mutations were observed in 7% (6/88) of the studied kindreds: the 3745delT mutation leads to a premature protein-translation termination at codon 1209, and the A→G substitution at 4216-2nt generates an aberrant splice-acceptor recognition site. Both of these mutations were present in three families. The proportion of BRCA2 mutations was 6% (5/88): the 999del5 mutation leads to translation termination at codon 273, and the 6503delTT mutation, which also is a frameshift mutation, generates a stop codon 18 bp farther downstream. Each of these two mutations was found in one family. The A→G 9346-2nt splice-acceptor site mutation is predicted to result in the skipping of exon 24. Three families were found to carry this mutation. In addition to the disease-related mutations, other DNA alterations in the coding regions were seen (table 3). Furthermore, several polymorphisms were identified in the noncoding regions of BRCA1 (in introns 6, 9, 17, 18, and 21) and BRCA2 (in exon 2 and introns 8, 24, and 25).

Interestingly, four of the mutations identified in our study were found to be the same as those in another Finnish study, by Vehmanen et al. (1997*a*, 1997*b*). Combined data show the presence of the BRCA1 exon 11 3745delT mutation in six families, the BRCA1 exon 12 4216-2ntA→G mutation in five families, the BRCA2

exon 9 999del5 mutation in eight families, and the BRCA2 exon 24 9346-2ntA→G mutation in eight families, providing evidence of at least four different founder mutations in Finland. At present, the BRCA1 exon 12 and BRCA2 exon 24 splice-site mutations appear to be unique to the Finnish population, but the BRCA1 3745delT and BRCA2 999del5 mutations have also been observed in Sweden (Zelada-Hedman et al. 1997) and Iceland (Johannesdottir et al. 1996), respectively. The ancestors of all three families with the BRCA1 3745delT mutation have been traced back to the late 19th century and the same rural area south of Oulu. Families 014 and 016 share the same disease-linked haplotype for four BRCA1 markers, indicating a common origin of the mutation (data not shown). Interestingly, both families have relatives living in Stockholm, where the Swedish family with the BRCA1 3745delT mutations resides. Although extensive migration from Finland to Sweden has occurred for many centuries, there is no indication that the Swedish family is related to the Finnish families in our study. It is quite peculiar that the BRCA2 999del5 mutation, which occurs frequently in Iceland, appears to be a founder mutation in Finland also but, as yet, has not been reported in other Nordic countries. BRCA2 6503delTT is the only mutation with no present recurrence in the Finnish population. However, several other families with this mutation have been identified elsewhere (see Breast Cancer Information Core), and it is possible that this site represents a mutational hot spot. The observed phenotype of one breast cancer and two ovarian cancers in family 028 corresponds well to what was expected, since the mutation is located within the

Table 3

Polymorphisms and Unknown and Silent Variants Occurring in Exonic Regions

Gene and Exon	Alteration	Effect	Type (No. of Cases)[a]
BRCA1:			
9	C710T	Cys→Cys	S (2)
11	C3832T	Pro→Leu	UV (1)
13	T4427C	Ser→Ser	P (41)
15	G4654T	Ser→Ile	UV (1)
16	A4956G	Ser→Gly	P (27)
16	T5002C	Met→Thr	UV (3)
16	T5074C	Met→Thr	UV (1)
16	G5075A	Met→Ile	P (3)
BRCA2:			
25	T9639G	Thr→Thr	S (1)
27	A10462G	Ile→Val	UV (3)

[a] S = silent variant; UV = unknown variant; and P = polymorphism.

ovarian cancer cluster region. This family was negative for a polymorphic stop codon, lys3326ter, which has been detected in some individuals carrying the 6503delTT mutation (Gayther et al. 1997). Haplotype and genealogical studies to determine the age and possible common ancestors of the mutations that are now observed—and to elucidate their relationship to the same mutations observed in other populations—have been initiated.

The population history of Finland differs, in many respects, from that of most countries. The coastal areas were populated mainly during the 15th century, and it was not until the 17th century that the vast inland regions were gradually inhabited by a relatively small

Table 2

Mutations and Cancers in BRCA1 and BRCA2 Families

Family	Gene	Exon	Mutation	Breast/Ovarian Cancer(s) in 1st- and 2d-Degree Relatives (Age [years] at Diagnosis)[a]	Other Cancers (No. of Cases)[a]
014	BRCA1	11	3745delT	Ov (46, 47, 50, 70[b])	Kid (1)
016	BRCA1	11	3745delT	Br (47, u)	Col (1), End (1), Lip (2), Liv (1)
062	BRCA1	11	3745delT	Ov (44, 50)	...
015	BRCA1	12	4216-2ntA→G	Br (28, 51, u), Ov (u)	Bt (1), Liv (1)
017	BRCA1	12	4216-2ntA→G	Br (29,[c] 36, 42, 52, u), Ov (u), Br+Ov (52)	Bt (1), Leu (1), Liv (1), Sto (1), Tes (1), Ton (1)
113	BRCA1	12	4216-2ntA→G	Br (62, u)	Col (1), Cx (1), Lu (1), Sto (1)
034	BRCA2	9	999delTCAAA	Bil Br (68/68[d])	Bo (1), Col (1), End (1), Lu (1), Sar (1)
028	BRCA2	11	6503delTT	Br (77), Ov (64, u)	End (2), Lu (1)
002	BRCA2	24	9346-2ntA→G	Br (45, 47, 48, u), Bil Br (62/64)	Csu (3), Liv (3), Pro (1), Sto (1), Thy (1)
005	BRCA2	24	9346-2ntA→G	Br (37, 38, 46, 68)	Csu (1), Leu (1)
087	BRCA2	24	9346-2ntA→G	Br (39)	Lu (1), Mel (1), Pan (1), Sto (2)

[a] Bas = basalioma; Bil Br = bilateral breast cancer; Bo = bone cancer; Br = breast cancer; Bt = brain cancer; Col = colon cancer; Csu = cancer site unknown; Cx = cervical cancer; End = endometrial cancer; Kid = kidney cancer; Leu = leukemia; Lip = lip cancer; Liv = liver cancer; Lu = lung cancer; Mel = melanoma; Ov = ovarian cancer; Pan = pancreatic cancer; Pro = prostate cancer; Sar = sarcoma; Sto = stomach cancer; Tes = testicular cancer; Thy = thyroid cancer; Ton = cancer of the tongue; and u = age unknown. Mutation carrier status is not defined for all individuals.

[b] Endometrial cancer also present.

[c] Liver cancer also present.

[d] Basalioma (at age 69 years) and colon cancer (at age 72 years) also present.

number of individuals. The presence of several regionally occurring BRCA1 and BRCA2 founder mutations could reflect this development, a situation similar to that for HNPCC. For instance, the four BRCA1 and BRCA2 Finnish founder mutations are responsible for 91% of the mutations in the families studied in Oulu but for only 45% of the mutations observed in the Helsinki study. None of the other 12 mutations identified in the study of individuals from southern Finland (Vehmanen et al. 1997*a*, 1997*b*) was seen in our material.

The clustering of mutations provides significant diagnostic advantages. In HNPCC, for example, two mutations have been found to account for a majority (63%) of the Finnish cases. It has therefore become important to design simple PCR-based tests for rapid mutation detection (Nyström-Lahti et al. 1995). Since the BRCA1 and BRCA2 founder mutations are currently seen in 55% (27/49) of the Finnish mutation-positive kindreds, it makes meaningful the development of similar diagnostic tests for breast and ovarian cancer.

Acknowledgments

Pia Huusko,[1] Kati Pääkkönen,[1] Virpi Launonen,[1] Minna Pöyhönen,[1] Guillermo Blanco,[2] Antti Kauppila,[3] Ulla Puistola,[3] Heikki Kiviniemi,[4] Marika Kujala,[1] Jaakko Leisti,[1] and Robert Winqvist[1, 2]

Departments of [1]Clinical Genetics, [2]Oncology and Radiotherapy, [3]Obstetrics and Gynecology, and [4]Surgery, Oulu University Hospital, University of Oulu, Oulu, Finland

Electronic-Database Information

URLs for data in this article are as follows:

Breast Cancer Information Core, http://www.nhgri.nih.gov/Intramural_research/Lab_transfer/Bic/
Genome Database, http://gdbwww.gdb.org/

References

Castilla LH, Couch FJ, Erdos MR, Hoskins KF, Calzone K, Garber JE, Boyd J, et al (1994) Mutations in the BRCA1 gene in families with early-onset breast and ovarian cancer. Nat Genet 8:387–391

Couch FJ, Farid LM, DeShano ML, Tavtigian SV, Calzone K, Campeau L, Peng Y, et al (1996) BRCA2 germline mutations in male breast cancer cases and breast cancer families. Nat Genet 13:123–125

de la Chapelle A (1993) Disease gene mapping in isolated human populations: the example of Finland. J Med Genet 30:857–865

Easton D (1997) Breast cancer genes—what are the real risks? Nat Genet 16:210–211

Ford D, Easton DF, Bishop DT, Narod SA, Goldgar DE, Breast Linkage Consortium (1994) Risks of cancer in BRCA1-mutation carriers. Lancet 343:692–695

Friedman LS, Gayther SA, Kurosaki T, Gordon D, Noble B, Casey G, Ponder BAJ, et al (1997) Mutation analysis of BRCA1 and BRCA2 in a male breast cancer population. Am J Hum Genet 60:313–319

Friedman LS, Ostermeyer EA, Szabo CI, Dowd P, Lynch ED, Rowell SE, King MC (1994) Confirmation of BRCA1 by analysis of germline mutations linked to breast and ovarian cancer in ten families. Nat Genet 8:399–404

Gayther SA, Mangion J, Russell P, Seal S, Barfoot R, Ponder BAJ, Stratton MR, et al (1997) Variation of risks of breast and ovarian cancer associated with different germline mutations of the BRCA2 gene. Nat Genet 15:103–105

Håkansson S, Johannsson O, Johansson U, Sellberg G, Loman N, Gerdes A-M, Holmberg E, et al (1997) Moderate frequency of BRCA1 and BRCA2 germ-line mutations in Scandinavian familial breast cancer. Am J Hum Genet 60:1068–1078

Johannesdottir G, Gudmundsson J, Bergthorsson JT, Arason A, Agnarsson BA, Eiriksdottir G, Johannsson OT, et al (1996) High prevalence of the 999del5 mutation in Icelandic breast and ovarian cancer patients. Cancer Res 56:3663–3665

Körkkö J, Annunen S, Pihlajamaa T, Prockop DJ, Ala-Kokko L (1998) Conformation sensitive gel electrophoresis for simple and accurate detection of mutations: comparison with denaturing gradient gel electrophoresis and nucleotide sequencing. Proc Natl Acad Sci USA 95:1681–1685

Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, et al (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 266:66–71

Newman B, Austin MA, Lee M, King MC (1988) Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families. Proc Natl Acad Sci USA 85:3044–3048

Nyström-Lahti M, Kristo P, Nicolaides NC, Chang SY, Aaltonen LA, Moisio AL, Järvinen HJ, et al (1995) Founding mutations and Alu-mediated recombination in hereditary colon cancer. Nat Med 1:1203–1206

Szabo CI, King M-C (1997) Population genetics of BRCA1 and BRCA2. Am J Hum Genet 60:1013–1020

Vehmanen P, Friedman LS, Eerola H, McClure M, Ward B, Sarantaus L, Kainu T, et al (1997*a*) Low proportion of BRCA1 and BRCA2 mutations in Finnish breast cancer families: evidence for additional susceptibility genes. Hum Mol Genet 6:2309–2315

Vehmanen P, Friedman LS, Eerola H, Sarantaus L, Pyrhönen S, Ponder BAJ, Muhonen T, et al (1997*b*) A low proportion of BRCA2 mutations in Finnish breast cancer families. Am J Hum Genet 60:1050–1058

Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N, et al (1995) Identification of the breast cancer susceptibility gene BRCA2. Nature 378:789–792

Zelada-Hedman M, Wasteson Arver B, Claro A, Chen J, Wer-
    elius B, Kok H, Sandelin K, et al (1997) A screening for
    BRCA1 mutations in breast and breast-ovarian cancer fam-
    ilies from the Stockholm region. Cancer Res 57:2474–2477

Address for correspondence and reprints: Dr. Robert Winqvist or Dr. Pia
Huusko, Department of Clinical Genetics, Oulu University Hospital
Kajaanintie 50, FIN-90220 Oulu, Finland. E-mail: robert.winqvist@oulu.fi

—————

A Triplet Repeat on 17q Accounts for Most Expansions Detected by the Repeat-Expansion–Detection Technique

To the Editor:

Eight hereditary neurogenerative disorders have been identified that result from expansions of CAG trinucleotide repeats (Gilles et al. 1997). Thus, there has been great incentive to develop techniques to efficiently screen for repeat sequences in specific patient populations. The repeat-expansion detection (RED) is a widely used technique that screens for trinucleotide expansions without a requirement of prior knowledge of the disease locus (Schalling et al. 1993; Zander et al. 1997). The method uses genomic human DNA as a template, specific-repeat oligonucleotide primers, and a thermostable ligase to generate oligomers of the primer. However, there are limitations to the technique. Non–disease-related expansions occur frequently in the population, which complicate the interpretation of disease-association studies. For a repeat to be detected by the RED, it must stand out in size, and smaller expansions causing disease can be missed. Also, once an expanded repeat is found, there is no information about chromosomal localization.

Other techniques for identifying trinucleotide expansions have recently been developed (Sanpei et al. 1996; Koob et al. 1998). The direct identification of repeat expansion and cloning technique (DIRECT) was designed to enable the localization and cloning of expanded-repeat regions (Sanpei et al. 1996). By means of the DIRECT technique, a novel, long, and unstable CAG/CTG trinucleotide repeat (Dir I) was identified and localized to chromosome 17q (Ikeuchi et al. 1998). This repeat is highly polymorphic, ranging in size from 10 to 92 repeat copies (30–276 bp) in normal individuals (Ikeuchi et al. 1998). Independently, a second group identified this expanded repeat by cloning the gene fragment from RED positive DNAs (Nakamoto et al. 1997).

DNA samples, obtained with informed consent from psychiatric patients with childhood onset of disease, were studied specifically to detect increased numbers of triplet repeats by use of the RED technique (Burgess et al. 1998). Since children and adolescents with psychiatric disorders appear to have functional brain abnormalities (McKenna et al. 1994), these patients are a valuable resource for such studies. Although expansions of trinucleotide repeats have been associated with several disorders affecting the brain and nervous system, their involvement in the etiology of psychiatric disorders has not been clearly demonstrated (Lindblad et al. 1995; Morris et al. 1995; O'Donovan et al. 1995). Since expanded repeats can be associated with genetic anticipation, patients with an early onset of disease are especially good candidates to evaluate.

The RED technique was used to identify trinucleotide expansions in 227 individuals, including 36 patients diagnosed with childhood-onset schizophrenia (COS); 21 diagnosed with atypical psychosis, termed by us as "multidimensionally impaired" (MDI) (Kumra et al. 1998); 46 patients with attention-deficit hyperactivity disorder (ADHD); 51 screened controls; and 73 relatives of probands. Patients were diagnosed according to standard *Diagnostic and Statistical Manual of Mental Disorders* definitions with standardized interviews as described elsewhere (Gordon et al. 1994; Castellanos et al. 1996). Diagnostic criteria for the MDI group has been discussed elsewhere (Kumra et al. 1998). The RED analysis was performed with a CTG_{10} oligonucleotide in the RED reaction, producing a repeat-size representation at 30-nucleotide intervals (Lindblad et al. 1996; Zander et al. 1997). The same samples were then analyzed for expansions of the polymorphic CAG/CTG Dir I trinucleotide repeat on chromosome 17q, by use of PCR conditions as described elsewhere (Ikeuchi et al. 1998; see fig. 1).

RED expansions of $\geqslant 180$ nucleotides were detected in a total of 99 (44%) of the 227 individuals screened, with the distribution of RED scores shown in table 1. When diagnoses were evaluated separately, RED scores of $\geqslant 180$ nucleotides were seen in 41% of COS patients ($n = 36$), 43% of MDI patients ($n = 21$), 43% of ADHD patients ($n = 24$), and 29% of the controls ($n = 51$) (fig. 2A).

In analyzing the CAG/CTG repeat on chromosome 17q, we scored the allele in each individual with the largest repeat size. A total of 81 (36%) of the 227 individuals screened had a chromosome 17q Dir I repeat expansion of >150 bp. Interestingly, 80 of the 81 individuals with a repeat size of >150 bp on chromosome 17q had RED scores of $\geqslant 180$ nucleotides. Thus, the RED technique appeared to detect this expansion reliably. There was also a strong correlation between the size of the Dir I expansion and the size of the expansion detected by RED (table 1).

Dir I repeats in excess of 50 copies (150 bp) were

observed in 33% of 36 COS patients ($n = 12$), 35% of 46 ADHD patients ($n = 16$), 33% of 21 MDI patients ($n = 7$), and 27% of the 51 controls ($n = 14$). On the basis of these frequencies and P values (COS, $P < .568$; ADHD, $P < .428$; MDI, $P < .607$), it does not appear that expansions at this locus are specifically associated with any of these disease phenotypes.

Thus, expansions of Dir I appear to account for the majority of elevated RED scores observed in our population. However, in each diagnostic group and in the controls there were ~20% of individuals with increased triplet repeats detected by the RED who did not have Dir I expansions (fig. 2*A*). These individuals were screened for expansions of another known heritable expanding CTG repeat in SEF2-1, a gene encoding a transcriptional factor protein found on the human chromosome 18q21.1 (Breschel et al. 1997). A 1.6-kb clone (termed "CTG 18.1"), which consists of portions of SEF2-1, contains an intronic $(CTG)_{24}$ repeat, which is highly polymorphic but not associated with an obvious abnormal phenotype.

In the majority of patients with elevated RED scores

Table 1

Correlation of Dir I Expansion Size with RED Score in Individuals with RED scores of $\geqslant 180$ Nucleotides

DIR I REPEAT SIZE (BP)	RED SCORE (NUCLEOTIDES)					Total
	180	210	240	270	300	
$\leqslant 150$	2	7	5	4	1	19
151–180	20	1	1	0	0	22
181–210	14	23	3	0	1	41
211–240	0	6	10	0	0	16
>240	0	0	0	1	0	1
Total	36	37	19	5	2	99

without an expansion of Dir I, examination of the CTG18.1 locus revealed an expanded allele of $\geqslant 150$ bp. Patients with the expanded CTG18.1 allele generally had very high RED values (many were $\geqslant 240$ nucleotides). Removing these samples greatly improved the correlation of RED score with Dir I expansion size (figs. 2*B* and 2*C*), changing the adjusted R^2 value from 0.500 to 0.740.
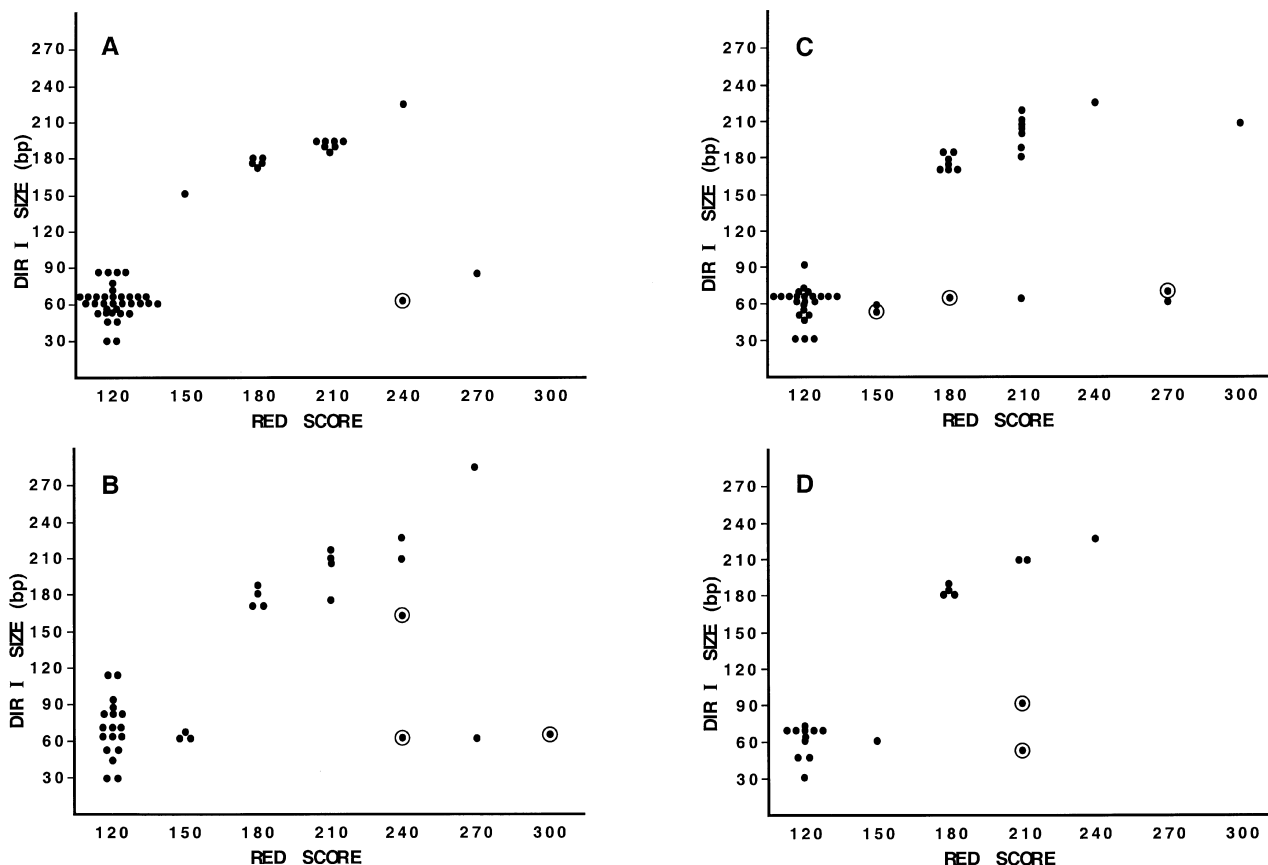

Figure 1 RED score (nucleotides) vs. Dir I size (bp): *A*, normal controls, *B*, patients with COS, *C*, patients with ADHD, and *D*, patients with MDI. Circled individuals were found to have expansions of CTG 18.1.
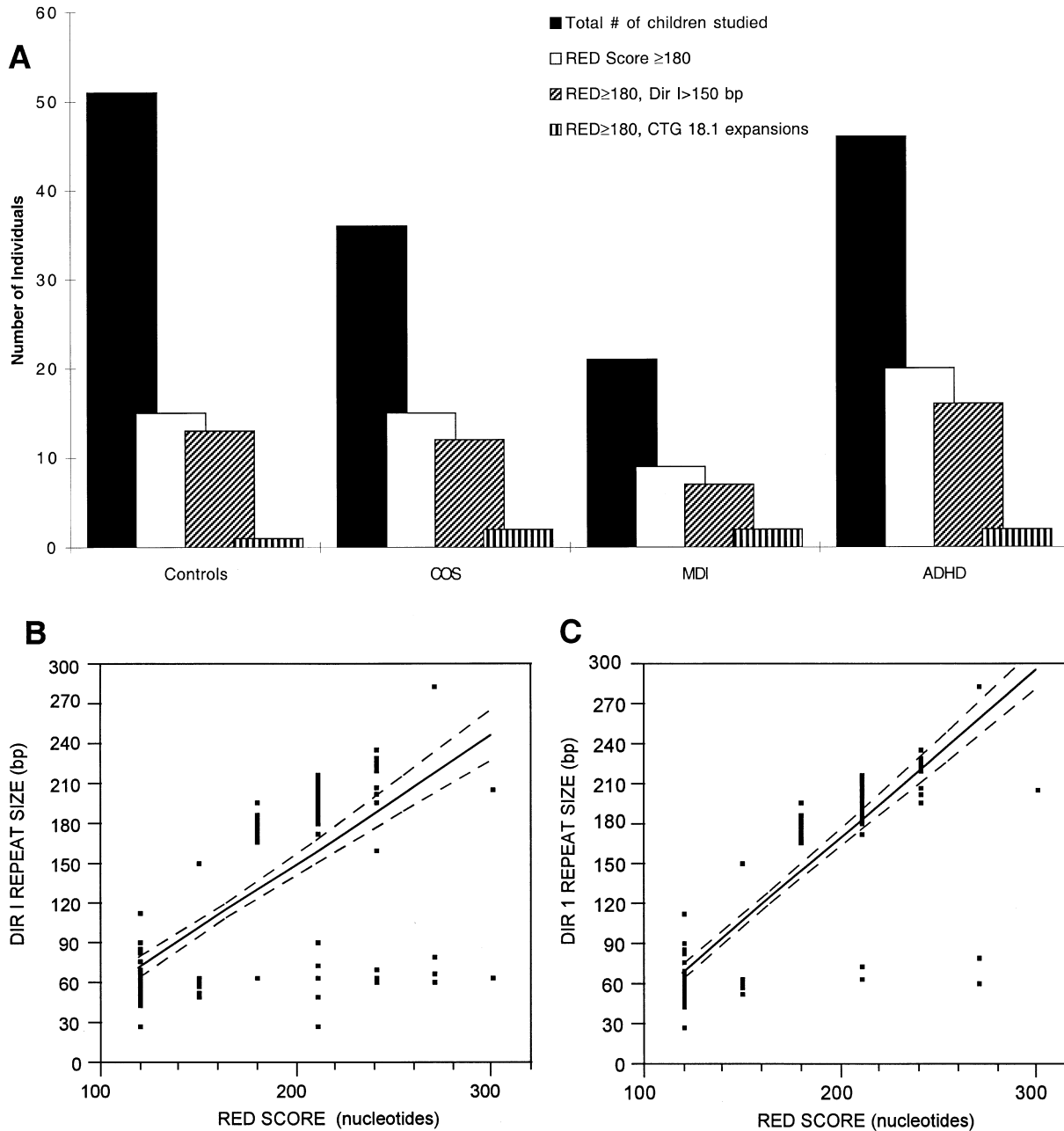
Figure 2   *A,* Distribution of elevated RED scores by diagnosis. *B,* Linear fit of RED scores plotted against Dir I repeat size: observations = 227, adjusted R^2 = 0.500, root mean square error = 15.62, analysis of variance $P \leqslant .001$. *C,* Linear fit when 15 individuals with expansions at CTG18.1 are removed: observations = 212, adjusted R^2 = 0.740, root mean square error = 11.42, analysis of variance $P \leqslant .001$.

This study demonstrates that up to 94% of trinucleotide-repeat expansions detected by RED can be accounted for by PCR analysis of two specific loci, Dir I and CTG18.1. The etiology of the remaining expansions requires further study. The combination of RED analysis with PCR of the Dir I and CTG18.1 loci may improve the ability to identify other disorders associated with repeat expansions. Clearly, although the RED technique reliably detects large trinucleotide repeats in the genome, many of the large repeats identified are not those related to illness.

ELLEN SIDRANSKY,[1] CATHERINE BURGESS,[3,*]
TAKESHI IKEUCHI,[4,*] KERSTIN LINDBLAD,[3]
ROBERT T. LONG,[1] ROBERT A. PHILIBERT,[1]
JUDITH RAPOPORT,[2] MARTIN SCHALLING,[3]
SHOJI TSUJI,[4] AND EDWARD I. GINNS[1]

[1]Clinical Neuroscience Branch and [2]Child Psychiatry
Branch, National Institute of Mental Health, National
Institutes of Health, Bethesda; [3]Department of Molecular
Medicine, Karolinska Hospital, Stockholm; and [4]Department
of Neurology, Brain Research Institute, Niigata University,
Niigata, Japan

References

Breschel TS, McInnis MG, Margolis RL, Sirugo G, Corneliussen B, Simpson SG, McMahon FJ, et al (1997) A novel, heritable, expanding CTG repeat in an intron of the SEF2-1 gene on chromosome 18q21.1. Hum Mol Genet 6: 1855–1863

Burgess CE, Lindblad K, Sidransky E, Yuan Q-P, Breschel T, Ross CA, McInnis M et al (1998) Large CAG/CTG repeat expansions are associated with childhood-onset schizophrenia. Mol Psychiatry (in press)

Castellanos FX, Giedd JN, Marsh WL, Hamburger SD, Vaituzis AC, Dickstein DP, Sarfatti SE, et al (1996) Quantitative brain magnetic resonance imaging in attention-deficit/hyperactivity disorder. Arch Gen Psychiatry 53:607–616

Gilles D, Abbas N, Stevanin G, Dürr A, Yvert G, Cancel G, Weber C, et al (1997) Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. Nat Genet 17:65–70

Gordon CT, Frazier JA, McKenna K, Giedd J, Zametkin A, Zahn T, Hommer D, et al (1994) Childhood onset schizophrenia: a NIMH study in progress. Schizophr Bull 20: 697–712

Ikeuchi T, Sanpei K, Takano H, Sasaki H, Tashiro K, Cancel G, Brice A, et al (1998) A novel long and unstable CAG/CTG trinucleotide repeat on chromosome 17q. Genomics (in press)

Koob MD, Benzow KA, Bird TD, Day JW, Moseley ML, Ranum LPW (1998) Rapid cloning of expanded trinucleotide repeat sequences from genomic DNA. Nat Genet 18: 72-75

Kumra S, Jacobsen LK, Lenane M, Zahn TP, Wiggs E, Alaghband-Rad J, Castellanos FX (1998) Multidimensionally impaired disorder: is it a variant of very early-onset schizophrenia? J Am Acad Child Adolesc Psychiatry 37:91–99

Lindblad K, Nylander P-O, Debruyn A, Sovrey D, Fander C, Engström C, Holmgren G, et al (1995) Detection of expanded CAG repeats in bipolar affective disorder using the repeat expansion detection (RED) method. Neurobiol Dis 2:55–62

Lindblad K, Savontaus M-L, Stevanin G, Holmberg M, Digre K, Zander C, Ehrsson H, et al (1996) CAG repeat expansions in spinocerebellar ataxia type 7. Genome Res 6: 965–971

McKenna K, Gordon CT, Lenane M, Kaysen D, Fahley K, Rapoport JL (1994) Childhood-onset schizophrenia: timely neurobiological research. J Am Acad Child Adolesc Psychiatry 33:771–781

Morris AG, Gaitonde E, McKenna PJ, Mollon JD, Hunt DM (1995) Repeat expansions and schizophrenia: association with disease in females and with early age-at-onset. Hum Mol Genet 4:1957–1961

Nakamoto M, Takebayashi H, Kawaguchi Y, Narumiya S, Taniwaki M, Nakamura Y, Ishikawa Y, et al (1997) A CAG/CTG expansion in the normal population. Nat Genet 17: 385–386

O'Donovan MC, Guy C, Craddock N, Murphy KC, Cardno AG, Jones LA, Owen MJ, et al (1995) Schizophrenia and bipolar disorder are associated with expanded CAG/CTG repeats. Nat Genet 10:380–381

Sanpei K, Takano H, Igarashi S, Sato T, Oyake M, Sasaki H, Wakisaka A, et al (1996) Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and cloning technique, DIRECT. Nat Genet 14: 277–284

Schalling M, Hudson TJ, Buetow KH, Housman DE (1993) Direct detection of novel expanded trinucleotide repeat in the human genome. Nat Genet 4:135–139

Zander C, Burgess CE, Johansen J, Lindblad K, Hudson T, Schalling M (1997) Repeat expansion detection. In: Taylor GR (ed) Laboratory methods for detection of mutations and polymorphisms in DNA. CRC Press, New York, pp 263–271

Paternal Isodisomy of Chromosome 7 Associated with Complete Situs Inversus and Immotile Cilia

To the Editor:

Uniparental disomy (UPD) refers to the inheritance of two homologous chromosomes from one parent, in a diploid individual. Heterodisomy is the inheritance of both parental homologues, whereas isodisomy implies the inheritance of two copies of a single parental homologue (Engel 1980). Uniparental inheritance of a human autosome in a cytogenetically normal individual was first recognized by Spence et al. (1988), on the basis
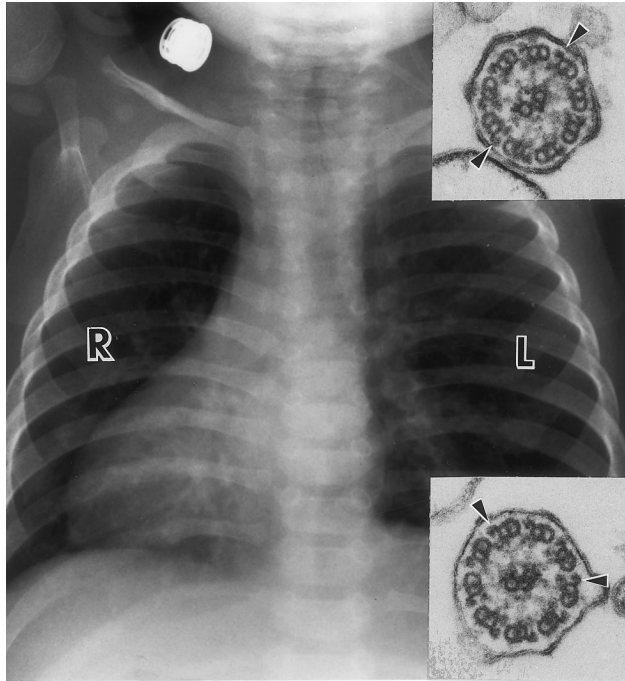
Figure 1   Complete situs inversus. A chest X-ray demonstrates the laterality defect manifested as dextrocardia and visceral situs inversus. The ultrastructural appearance of bronchial cilia is shown in the insets. Note normal axonemal structure with typical 9 + 2 doublet microtubular configuration. Arrowheads indicate normal inner and outer dynein arms (transmission electron microscopy).

of the failure to detect a paternal allele in the course of genotyping a patient with cystic fibrosis (CF; MIM 219700).

   Patient CC was born at 34 wk gestation to a 20-year-old Hispanic mother after an uncomplicated pregnancy. His birth weight was 2,540 g and the newborn physical exam was reported as normal. The family history was unremarkable, and there was no known consanguinity. He developed respiratory distress at 10 d of age, and a chest X-ray and echocardiogram demonstrated dextrocardia with a structurally normal heart (fig. 1). An abdominal ultrasound demonstrated visceral situs inversus with a structurally normal spleen. At 6 mo of age, the child was growth retarded with signs of significant respiratory disease. A tracheal aspirate grew *Pseudomonas aerugenosa* and *Staphylococcus aureus,* and a sweat chloride test confirmed the clinical diagnosis of CF. Because of the presence of situs inversus and severe lung disease, a bronchial biopsy was performed to evaluate ciliary function. The respiratory mucosal brushings were obtained by the endoscopist, transported in isotonic saline, and immediately examined by the pathologist. The endoscopist was instructed to avoid mucosal areas with acute and/or chronic inflammatory changes, such as erosions, ulcerations, or hyperemia, and to sample only

those areas that appear clinically normal. Microscopic examination for ciliary motility was performed by use of dark-field microscopy with the cytologic material diluted with saline (Rutland and Cole 1980, 1981). An exhaustive search with numerous wet preparation slides was performed. No normal ciliary motion was observed, despite the lack of an inflammatory infiltrate. Electron microscopy of the sample demonstrated structurally normal cilia (fig. 1, *inset*).

   A peripheral lymphocyte karyotype was normal. DNA testing showed the patient to be homozygous for the ΔF508 CFTR (*c*ystic *f*ibrosis *t*ransmembrane-conductance *r*egulator) mutation, with his mother lacking any detectable alteration. To further define the mode of inheritance, DNA was extracted from lymphoblasts of the patient and his mother; a sample from the father was not available for study. PCR amplification was carried

Table 1

DNA Markers and Results of Genotyping

		ALLELES		
CHROMOSOME AND MARKER[a]	LOCATION	Mother	Patient	HETEROZYGOSITY
7:				
D7S531[b]	7pter-p15	2,3	1,1	.77
D7S481	7pter-p15	1,2	2,2	.84
D7S507[b]	7pter-p15	1,2	3,3	.89
D7S488[b]	7p21-p15	1,2	3,3	.83
D7S673[b]	7	1,2	3,3	.86
D7S526[b]	7p21-p15	1,2	3,3	.71
D7S485[b]	7p15	2,2	1,1	.78
Elastin	7q11.23	1,2	1,1	.63
D7S849	7p11-q21	1,2	2,2	.57
D7S440	7q	1,1	1,1	.73
D7S492	7q15-q22	1,2	1,1	.76
D7S501	7q31	1,1	1,1	.81
D7S471[b]	7	1,3	2,2	.83
D7S650	7q	1,1	1,1	.85
D7S640[b]	7q	2,3	1,1	.85
D7S495	7q31-q35	1,2	1,1	.81
D7S636	7	2,3	1,1	.90
D7S483[b]	7q31-qter	1,2	3,3	.82
D7S550[b]	7q31-qter	2,3	1,1	.83
D7S489	7p15-q22	1,2	1,1	.38
D7S467	7q35-q36	1,1	1,1	.93
Other:				
D1S249		1,2	2,3	.87
D9S195		1,1	1,2	.74
D9S67		2,2	1,2	.64
D10S187		1,2	1,1	.83
D10S211		1,2	2,3	.83
D14S67		2,3	1,2	.87
D14S68		2,2	1,2	.88
D16S406		1,1	1,1	.83
D17S806		1,3	1,2	.90
D21S1255		1,2	1,2	.80

  [a] Twenty one polymorphic loci were examined for evidence of UPD for chromosome 7, and 10 markers from other chromosomes were used to examine maternity.
  [b] Demonstrated uniparental isodisomy.

out under standard conditions for the di- and trinucleotide repeat loci listed in table 1 (Gyapay et al. 1994), by use of a Perkin-Elmer 480 device and polyacrylamide electrophoresis. A total of 21 loci along the length of chromosome 7 were used to demonstrate paternal isodisomy for the chromosome. The patient was homozygous at all 21 loci, with 10 of the markers being fully informative for paternal inheritance. To establish that maternity is correct, 10 highly polymorphic markers (6–16 alleles for each) for autosomal loci other than chromosome 7 were tested. The patient shared at least one allele at each locus with his mother, indicating a high likelihood (>99%) of maternity.

To address the possibility of a second independent etiology for situs abnormalities, the patient's DNA was examined for deletions within the critical region of the X-linked heterotaxy locus in Xq26.2 (MIM 306955; Casey et al. 1993), and the coding region of a recently identified gene that when altered leads to heterotaxy (Gebbia et al. 1997) was sequenced. Males with X-linked heterotaxy typically have complex heart anomalies in addition to visceral situs ambiguous (Casey et al. 1993), a feature lacking in this patient. Sporadic cases of heterotaxy have been described with submicroscopic deletions within this critical region (Ferrero et al. 1997). A total of nine sequence-tagged-site markers distributed throughout the critical region were used to search for submicroscopic deletions, and none were identified (data not shown). Furthermore, the protein-encoding exons from the gene now known to be responsible for X-linked heterotaxy (ZIC3; Gebbia et al. 1997) was sequenced from the patient, and no mutation was identified. In addition, the differential display technique (Liang and Pardee 1992) was used to compare transcripts present in lymphoblasts from this patient, a patient with maternal isodisomy of chromosome 7 (Spence et al. 1988), and from control cell lines. This analysis failed to identify any differentially expressed transcripts among these cell lines (except for immunoglobulin light-chain transcripts), possibly because of the cell type used as the source of RNA.

Several possible mechanisms for UPD have been proposed, including gamete complementation, postzygotic monosomy duplication, reduction to disomy of a trisomic conception, and somatic crossing over (Spence et al. 1988). The two more likely explanations in this case would be (1) monosomy duplication associated with a nullisomic maternal gamete or (2) a paternal meiosis II nondisjunction resulting in a trisomic conceptus, with subsequent reduction to disomy through loss of the maternal chromosome 7 (trisomic rescue). UPD is increasingly recognized as a mechanism for the non-Mendelian transmission of a recessive disorder (reviewed by Ledbetter and Engel 1995). In addition to CF, examples include methylmalonic aciduria (MIM 251000; Abra-

mowicz et al. 1994), rod monochromacy (MIM 216900; Pentao et al. 1992), Bloom syndrome (MIM 210900; Woodage et al. 1994), lipoprotein lipase deficiency (MIM 238600; Benlian et al. 1996), Duchenne muscular dystrophy in a female (MIM 310200; Quan et al. 1997), Herlitz junctional epidermolysis bullosa (MIM 226700; Pulkkinen et al. 1997), and congenital chloride diarrhea (CLD; MIM 214700; Hoglund et al. 1994). As with several cases of maternal disomy of chromosome 7, the patient was initially identified because of CF. Maternal isodisomy for chromosome 7 or 7p is consistently associated with short stature (Spence et al. 1988; Voss et al. 1989; Spotila et al. 1992; Eggarding et al. 1994; Langlois et al. 1995), and in one patient this has been attributed to growth-hormone deficiency (Hubbard et al. 1980). Similarly, both hetero- and isodisomy of chromosome 7 has been reported in association with dwarfism, either unspecified dwarfism or with clinical features of Russell-Silver syndrome (MIM 180860; Kotzot et al. 1995; Preece et al. 1997). The finding of maternal heterodisomy implies that genomic imprinting may be the underlying cause of the dwarfism. Since it is known that the region of conserved synteny in the mouse undergoes imprinting (Cattanach and Kirk 1985), it is reasonable to predict that there is one or more loci on chromosome 7 involved in the growth-hormone axis that undergoes imprinting during development. Potential genes located on chromosome 7 include, among others, insulin-like growth-factor binding protein 1 (IGF-BP1), insulin-like growth-factor binding protein 3 (IGF-BP3), or the growth hormone–releasing hormone receptor (GHRHR). In contrast, paternal isodisomy for the entire chromosome has been reported in only a single patient, who came to medical attention because of CLD, a disorder previously mapped to chromosome 7 (Hoglund et al. 1994). The patient was reported to have normal stature and, apart from CLD, has no clinically detectable abnormalities except for mild high-frequency hearing loss.

A parsimonious interpretation of this case would suggest that the child inherited two recessive disorders from his father: CF, and immotile cilia associated with complete situs inversus. It is, of course, possible that the two disorders are unrelated in this patient; for example, he may coincidentally have autosomal dominant or recessive heterotaxy, especially in light of the unavailability of the father for a detailed clinical evaluation. X-linked heterotaxy is less likely given the failure to identify any sequence alterations in the causative gene or any small deletions in the adjoining regions.

Immotile-cilia syndrome (MIM 242650) is likely a genetically heterogeneous disorder that typically comes to medical attention because of sinopulmonary infections (rhinosinusitus, bronchitis, and bronchiectasis) and male infertility. In the face of situs inversus, it is referred

to eponymically as "Kartagener syndrome" (MIM 244400), with an estimated prevalence of 1/20,000–1/60,000 individuals (Afzelius and Mossberg 1995). Although, by use of electron microscopy, abnormalities of ciliary structures are a frequent finding, there are a number of cases of Kartagener syndrome with structurally normal cilia that are dysmotile or immotile (Herzon and Murphy 1980; Pedersen 1983), and a recent report indicates that subtle structural deficiencies may be much more common than has been previously recognized (Teknos et al. 1997). In examining ciliary function, care must be taken to process the sample properly, and acute or chronic inflammation from a variety of sources may lead to false-positive results, especially in the case of a CF patient. In light of the fact that only half of the patients with immotile-cilia syndrome have situs inversus, electron microscopic evaluation of ciliary structure in those CF patients with unusually severe respiratory disease may be warranted. Interestingly, one other well-established case of CF in association with situs inversus that was interpreted to be Kartagener syndrome has been reported (Burnell and Robertson 1974). In contrast, Liechti-Gallati and Kraemer (1995) failed to detect any CF mutations in a cohort of patients with immotile-cilia syndrome.

It is noteworthy that laterality defects have been associated with complex chromosome rearrangements involving chromosome 7. One patient has been described (Koiffman et al. 1993) with an insertion of chromosome 8 material into 7q22 (46,XY,ins[7;8][q22;q12q24]), whereas the offspring of a mother with ectrodactyly and an apparently balanced reciprocal translocation, t(2;7)(q21.1;q22.1), was found to have situs ambiguus and complex heart disease (analysis of a karyotype from the infant was not performed) (Genuardi et al. 1993). Although genes such as *nodal* and *Vg1* have been implicated in the lateralization of vertebrate embryos (Collignon et al. 1996; Hyatt et al. 1996), a variety of other genes may be involved in laterality and ciliary defects, including other extracellular signalling molecules, transcription factors, and cytoskeletal components. In this regard, the gene for actin-binding protein 2 (FLN2), previously localized to chromosome 7, may represent a candidate gene. It is proposed that chromosome 7 should be examined in a large collection of families with Kartagener syndrome.

Yanzhen Pan,[1] Christopher D. McCaskill,[1]
Katherine H. Thompson,[1] John Hicks,[2] Brett Casey,[2]
Lisa G. Shaffer,[1] and William J. Craigen[1,3]
Departments of [1]*Molecular and Human Genetics*,
[2]*Pathology, and* [3]*Pediatrics, Baylor College of Medicine*,
Houston

Electronic-Database Information

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

Online Mendelian inheritance in man (OMIM), http://www.ncbi.nlm.gov/htbin-post/Omim (for cystic fibrosis [MIM 219700], X-linked heterotaxy [MIM 306955], methylmalonic aciduria [MIM 251000], rod monochromacy [MIM 216900], Bloom syndrome [MIM 210900], lipoprotein lipase deficiency [MIM 238600], Duchenne muscular dystrophy [MIM 310200], Herlitz junctional epidermolysis bullosa [MIM 226700], congenital chloride diarrhea [MIM 2147000], Russell-Silver syndrome [MIM 180860], immotile-cilia syndrome [MIM 242650], and Kartagener syndrome [MIM 244400])

References

Abramowicz MJ, Andrien M, Dupont E, Dorchy H, Parma J, Duprez L, Ledley FD, et al (1994) Isodisomy of chromosome 6 in a newborn with methylmalonic acidemia and agenesis of pancreatic beta cells causing diabetes mellitus. J Clin Invest 94:418–421

Afzelius B, Mossberg B (1995) Immotile-cilia syndrome (primary ciliary dyskinesia), including Kartagener syndrome. In: Scriver C, Beaudet AL, Sly W, Valle D (eds) The metabolic and molecular bases of inherited disease, 7th ed. McGraw-Hill, New York, pp 3943–3954

Benlian P, Foubert L, Gagné E, Bernard L, De Gennes JL, Langlois S, Robinson W, et al (1996) Complete paternal isodisomy for chromosome 8 unmasked by lipoprotein lipase deficiency. Am J Hum Genet 59:431–436

Burnell RH, Robertson EF (1974) Cystic fibrosis in a patient with Kartagener syndrome. Am J Dis Child 127:746–747

Casey B, Devoto M, Jones K, Ballabio A (1993) Mapping a gene for familial situs abnormalities to human chromosome Xq24-q27.1. Nat Genet 5:403–407

Cattanach B, Kirk M (1985) Differential activity of maternally and paternally derived chromosome regions in mice. Nature 315:496–498

Collignon J, Varlet I, Robertson EJ (1996) Relationship between asymmetric *nodal* expression and the direction of embryonic turning. Nature 381:155–158

Eggerding FA, Schonberg SA, Chehab FF, Norton ME, Cox VA, Epstein CJ (1994) Uniparental isodisomy for paternal 7p and maternal 7q in a child with growth retardation. Am J Hum Genet 55:253–265

Engel E (1980) A new genetic concept: uniparental disomy and its potential effect, isodisomy. Am J Med Genet 6:137–143

Ferrero GB, Gebbia M, Pilia G, Witte D, Peier A, Hopkin RJ, Craigen WJ, et al (1997) A submicroscopic deletion in Xq26 associated with familial situs ambiguus. Am J Hum Genet 61:395–401

Gebbia M, Ferrero GB, Pilia G, Bassi MT, Aylsworth AS, Bird LM, Penman-Splitt M, et al (1997) X-linked situs abnormalities result from mutations in ZIC3. Nat Genet 17:305–308

Genuardi M, Pomponi M, Sammito V, Bellussi A, Zollino M, Neri G (1993) Split hand/split foot anomaly in a family segregating a balanced translocation with breakpoint on 7q22.1. Am J Med Genet 47:823–831

Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, Marc S, et al (1994) The 1993–94 Généthon human genetic linkage map. Nat Genet 7:246–339

Herzon F, Murphy S (1980) Normal ciliary ultrastructure in children with Kartagener's syndrome. Ann Otol Rhinol Laryngol 89:81–83

Höglund P, Holmberg C, de la Chapelle A, Kere J (1994) Paternal isodisomy for chromosome 7 is compatible with normal growth and development in a patient with congenital chloride diarrhea. Am J Hum Genet 55:747–752

Hubbard V, Davis P, DiSant'Agnese PA, Gordon P, Schwartz R (1980) Associated growth hormone deficiency and cystic fibrosis: a report of two cases. Am J Dis Child 134:317–319

Hyatt BA, Lohr JL, Yost HJ (1996) Initiation of vertebrate left-right axis formation by maternal vg1. Nature 384:62–65

Koiffmann CP, Wajntal A, de Souza DH, Gonzalez CH, Coates MV (1993) Human situs determination and constitution 46,XY,ins (7;8)(q22;q12q24). Am J Med Genet 47:568–569

Kotzot D, Schmitt S, Bernasconi F, Robinson W, Lurie I, Ilyina H, Mehes K, et al (1995) Uniparental disomy 7 in Silver-Russell syndrome and primordial growth retardation. Hum Mol Genet 4:583–587

Langlois S, Yong S, Wilson R, Kwong L, Kalousek D (1995) Prenatal and postnatal growth failure associated with maternal heterodisomy for chromosome 7. J Med Genet 32:871–875

Ledbetter D, Engel E (1995) Uniparental disomy in humans: development of an imprinting map and its implications for prenatal diagnosis. Hum Mol Genet 4:1757–1764

Liang P, Pardee AB (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257:967–971

Liechti-Gallati S, Kraemer R (1995) Cystic fibrosis mutations and immotile cilia syndrome. Clin Genet 47:328–329

Pedersen M (1983) Specific types of abnormal ciliary motility in Kartagener's syndrome and analogous respiratory disorders. Eur J Respir Dis Suppl 64:78–90

Pentao L, Lewis RA, Ledbetter DH, Patel PI, Lupski JR (1992) Maternal uniparental isodisomy of chromosome 14: association with autosomal recessive rod monochromacy. Am J Hum Genet 50:690–699

Preece MA, Price SM, Davies V, Clough L, Stanier P, Trembath RC, Moore GE (1997) Maternal uniparental disomy 7 in Silver-Russell syndrome. J Med Genet 34:6–9

Pulkkinen L, Bullrich F, Czarnecki P, Weiss L, Uitto J (1997) Maternal uniparental disomy of chromosome 1 with reduction to homozygosity of the LAMB3 locus in a patient with Herlitz junctional epidermolysis bullosa. Am J Hum Genet 61:611–619

Quan F, Janas J, Toth-Fejel S, Johnson DB, Wolford JK, Popovich BW (1997) Uniparental disomy of the entire X chromosome in a female with Duchenne muscular dystrophy. Am J Hum Genet 60:160–165

Rutland J, Cole PJ (1980) Non-invasive sampling of nasal cilia for measurement of beat frequency and study of ultrastructure. Lancet 2:564–565

—— (1981) Nasal mucociliary clearance and beat frequency in cystic fibrosis compared with sinusitis and bronchiectasis. Thorax 36:654–658

Spence JE, Perciaccante RG, Grieg GM, Willard HF, Ledbetter DH, Hejtmancik JF, Pollack MS, et al (1988) Uniparental disomy as a mechanism for human genetic disease. Am J Hum Genet 42:217–226

Spotila LD, Sereda L, Prockop DJ (1992) Partial isodisomy for maternal chromosome 7 and short stature in an individual with a mutation at the COL1A2 locus. Am J Hum Genet 51:1396–1405

Teknos TN, Metson R, Chasse T, Balercia G, Dickersin GR (1997) New developments in the diagnosis of Kartagener's syndrome. Otolaryngol Head Neck Surg 116:68–74

Voss R, Ben-Simon E, Avital A, Godfrey S, Zlotogora J, Dagan J, Tikochinski Y, et al (1989) Isodisomy of chromosome 7 in a patient with cystic fibrosis: could uniparental disomy be common in humans? Am J Hum Genet 45:373–380

Woodage T, Prasad M, Dixon JW, Selby RE, Romain DR, Columbano-Green LM, Graham D, et al (1994) Bloom syndrome and maternal uniparental disomy for chromosome 15. Am J Hum Genet 55:74–80

Address for correspondence and reprints: Dr. William J. Craigen, S821, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030. E-mail: wcraigen@bcm.tmc.edu

Familial Skewed X Inactivation and X-Linked Mutations: Unbalanced X Inactivation is a Powerful Means to Ascertain X-Linked Genes That Affect Cell Proliferation

To the Editor:

We read with great interest the article by Pegoraro et al. (1997) in which they report familial skewed X inactivation associated with a submicroscopic deletion at Xq28. This deletion, which spans ~800 kb near the F8C locus, is transmitted in the expected X-linked fashion through females, but no males have the deletion. Although we agree that the skewed X-inactivation pattern in this family is attributable to the region around or included in the deletion, we disagree with their inter-

pretation of these results and with some assertions that may lead to confusion about the underlying basis for skewed X inactivation in this family.

The authors suggest that the simplest interpretation of their data is that this disorder behaves like an X dominant with lethality in the male fetus. All the evidence supports this hypothesis: The Xq28 deletion is lethal when present in hemizygous males and confers a proliferative disadvantage to cells of heterozygous females in whom the deleted chromosome is active, which explains why these females are clinically normal.

However, Pegoraro et al. (1997) considered the hypothesis of an X-dominant disorder with male lethality inconsistent with their data on the basis that the relevant locus cannot be a genetic or cell lethal, because, if so, affected males would die soon after fertilization. In fact, the time at which a male embryo carrying the deletion would become inviable, or at which the population of mutant cells in the female would be eliminated, depends on the nature of the mutant gene product—when it is needed, in which cells it is expressed, and how detrimental (or advantageous) it is in the presence of wild-type cells in the mosaic population. Clearly, a male with this small deletion might not be aborted before the pregnancy is recognizable; some males with macroscopic deletions of the X chromosome can survive until late gestation or term (Rosenberg et al. 1987; Cremers et al. 1988).

This family resembles those in which selection has been shown to underlie X inactivation. The familial skewing seen with mutations at the incontinentia pigmenti (IP) locus, in cytogenetically normal females, is also associated with recurrent abortion (Migeon et al. 1989; Parrish et al. 1996). However, IP heterozygotes do not eliminate their mutant population of cells as completely as do the heterozygotes in this family, and hence some IP heterozygotes may be clinically abnormal. Large differences in the growth of normal and abnormal cells can result in early loss of disfavored cells, whereas more subtle differences account for progressive loss during the lifetime of the heterozygote (Migeon et al. 1981, 1988). The familial skewed X inactivation associated with adrenoleukodystrophy was also mapped to Xq28, by tight linkage with the G6PD locus (Migeon et al. 1981), and was shown to be attributable to a proliferative advantage of the mutant gene.

Pegoraro et al. (1997) suggest two alternative novel hypotheses to explain their findings: (1) the deleted region in Xq28 is directly involved in the primary mechanisms of X inactivation; or (2) the region is directly responsible for skewing. In either case, the observed skewing is attributed to a gene in Xq28 that determines the choice of X chromosome to inactivate or the ability of an X to inactivate. As Pegoraro et al. (1997) note, there is no evidence in either mouse or human to support the existence of a such a gene within the deletion at Xq28. In addition, there is uncited evidence (discussed below) that argues against the existence of such a hypothetical locus in Xq28.

There may be multiple *trans*-acting factors involved in the initiation and maintenance of X inactivation. Observations of triploid embryos suggest that these are likely to reside on autosomes (discussed by Jacobs and Migeon 1989). There is no convincing evidence that factors responsible for *cis* inactivation lie outside the X chromosome–inactivation center (XIC) at Xq13.2. Studies of chromosomal deletions and translocations leading to deletions indicate that this segment is the site of the XIC (Therman et al. 1979; Mattei et al. 1981), and all X-linked genes shown to have an effect on choice (Xce) or on inactivation (XIST) lie within this region. Evidence that other regions of the X are unlikely to be relevant in the choice of X to inactivate comes from recent studies of the murine XIC region transfected into male embryonic stem cells, which indicate that the XIC region contains all the genetic information needed for *cis* inactivation of an X chromosome, including choice of inactive X (Lee et al. 1996; Lee and Jaenisch 1997).

None of the reported families with skewed patterns of X inactivation has been shown, convincingly, to involve the mechanisms of X inactivation. Most of them can be explained by mutations or cytogenetic deletions that adversely affect the proliferation of one of the cell populations produced by random X inactivation (reviewed in Belmont 1996). In light of this, it seems that relatively few cases of familial unbalanced X inactivation will be due to mutations causing primary nonrandom inactivation. Although mutations affecting the process of X inactivation are real possibilities, one expects that those severe enough to interfere with X inactivation would be embryonic lethals, because of the functional X disomy. Mutations in the inactivation machinery that are associated with normal phenotypes and two intact X chromosomes are most likely to be the ones that influence only the choice of X to inactivate and therefore would be at the XIST or an Xce-like locus or in other constituents of the XIC. However, even linkage to the XIC or with specific XIST alleles (Plenge et al. 1997) is not sufficient evidence that skewing is in fact attributable to these genes, because it could be due to a hitchhiker effect, resulting from mutations in a close neighbor with potential for cell selection.

We believe that familial skewing, which maps to the X chromosome but is unlinked to the XIC, most likely results from mutations that do not interfere with the inactivation machinery but that affect the proliferation of the cell populations produced following normal random X inactivation. Females who manifest X-recessive diseases provide a powerful means to detect unbalanced X inactivation (Migeon 1993). The corollary is that

skewed patterns of X inactivation, familial or not, provide a powerful means of ascertaining mutations that influence cell proliferation.

BARBARA R. MIGEON AND CAMILLE HAISLEY-ROYSTER
*Department of Pediatrics and
Center for Medical Genetics
The Johns Hopkins University School of Medicine
Baltimore*

References

Belmont JW (1996) Genetic control of X inactivation and processes leading to X-inactivation skewing. Am J Hum Genet 58:1101–1108

Cremers FPM, van de Pol TJR, Wieringa B, Hofker MH, Pearson PL, Pfeiffer RA, Mikkelsen M, et al (1988) Molecular analysis of male-viable deletions and duplications allows ordering of 52 DNA probes on proximal Xq. Am J Hum Genet 43:452–461

Jacobs PA, Migeon BR (1989) Studies of X-chromosome inactivation in trisomies. Cytogenet Cell Genet 50:75–77

Lee JT, Jaenisch R (1997) Long-range *cis* effects of ectopic X-inactivation centres on a mouse autosome. Nature 386: 275–279

Lee JT, Strauss WM, Dausman JA, Jaenisch R (1996) A 450 kb transgene displays properties of the mammalian X-inactivation center. Cell 86:83–94

Mattei MG, Mattei JF, Vidal I, Giraud F (1981) Structural anomalies of the X chromosome and inactivation center. Hum Genet 56:401–408

Migeon BR (1993) The postulated X-inactivation center at Xq27 is most reasonably explained by ascertainment bias: heterozygous expression of recessive mutations is a powerful means of detecting unbalanced X inactivation. Am J Hum Genet 52:431–432

Migeon BR, Axelman J, Jan de Beur S, Valle D, Mitchell GA, Rosenbaum KN (1989) Selection against lethal alleles in females heterozygous for incontinentia pigmenti. Am J Hum Genet 44:100–106

Migeon BR, Moser HW, Moser AB, Sprenkle JA, Sillence D, Norum RA (1981) Adrenoleukodystrophy: evidence for X-linkage, inactivation and selection favoring the mutant allele in heterozygous cells. Proc Natl Acad Sci USA 78: 5066–5070

Parrish JE, Scheuerle AE, Lewis RA, Levy ML, Nelson DL (1996) Selection against mutant alleles in blood leukocytes is a consistent feature in incontinentia pigmenti type 2. Hum Mol Genet 5:1777–1783

Pegoraro E, Whitaker J, Mowery-Rushton P, Surti U, Lanasa M, Hoffman EP (1997) Familial skewed X inactivation: a molecular trait associated with high spontaneous-abortion rate maps to Xq28. Am J Hum Genet 61:160–170

Plenge RM, Hendrich BD, Schwartz C, Arena F, Naumova A, Sapienza C, Winter R, et al (1997) A promoter mutation in the XIST gene in two unrelated families with skewed X-chromosome inactivation. Nat Genet 17:353–356

Rosenberg T, Niebuhr E, Yang H-M, Parving A, Schwartz M (1987) Choroideremia, congenital deafness and mental retardation in a family with an X chromosomal deletion. Ophthalmic Paediatr Genet 8:139–143

Therman E, Sarto GE, Palmer CG, Kallio H, Denniston C (1979) Position of the human X inactivation center on Xq. Hum Genet 50:59–64

Reply to Migeon and Haisley-Royster

To the Editor:

We thank Drs. Migeon and Haisley-Royster (1998 [in this issue]) for their interest in our research. We are, however, a bit puzzled by their letter to the editor, since they write that they disagree with the interpretation of our results yet then restate what was already written in our previously published article (Pegoraro et al. 1997).

The 50-member pedigree that we reported showed an X-linked dominant disorder with male lethality. There is no question of this fact, because we found a deletion mutation of Xq28 associated with skewed X inactivation and recurrent pregnancy loss (LOD = 6.92). The deletion included the factor VIII gene, yet there were no males from 50 females with factor VIII deficiency, again clearly proving that this family had an X-linked dominant disorder with male lethality.

Drs. Migeon and Haisley-Royster appear to wish to address two issues: (1) interpretation of the likely mechanisms that would cause X-inactivation skewing in the females in this family; and (2) transcriptional timing of the deleted gene or gene products in Xq28 and the observed effect on miscarriage detection. There is very little to disagree with in Drs. Migeon and Haisley-Royster's interpretation of our results; they suggest that a growth disadvantage is probably playing a role, which is precisely what we stated in our discussion. We, too, feel that growth disadvantage is the most likely mechanism causing skewing of X-chromosome inactivation. However, in the absence of characterization of the causative genes in Xq28, it seems unreasonable to dismiss the possibility that the gene(s) may actually be involved in the process of X inactivation. This is the least likely mechanism, but it does not seem to warrant exclusion from discussion.

The timing of transcription of the gene products in Xq28 undoubtedly affects when the miscarriage occurs. In fact, the issue of timing is central to the inferred ge-

netic mechanism. A cell-lethal trait expressed very early in embryonic development would be undetectable or perhaps would cause a "biochemical pregnancy." Activation later in embryonic life would still cause male lethality but would be less likely to cause complete skewing of X inactivation in multiple tissues in the heterozygous female. In view of this delicate balance in timing, we feel that the genes in question are most likely to be transcribed early in fetal development and to impart a growth disadvantage rather than being cell lethal. The size of the deletion mutation, however, is less important to when the miscarriage occurs: size is simply being used as a surrogate to the assumed importance of the deletion region and gene(s) contained in that region. In the end, this is all an exercise in mental gymnastics, since the characterization of the causative gene(s) will enlighten us all as to the true mechanism.

ERIC HOFFMAN, ELENA PEGORARO, AND
MARK LANASA
Department of Molecular Genetics and Biochemistry
University of Pittsburgh School of Medicine
Pittsburgh

References

Migeon BR, Haisley-Royster C (1998) Familial skewed X inactivation and X-linked mutations: unbalanced X inactivation is a powerful means to ascertain X-linked genes that affect cell proliferation. Am J Hum Genet 62:1555–1557 (in this issue)

Pegoraro E, Whitaker J, Mowery-Rushton P, Surti U, Lanasa M, Hoffman EP (1997) Familial skewed X inactivation: a molecular trait associated with high spontaneous-abortion rate maps to Xq28. Am J Hum Genet 61:160–170

Address for correspondence and reprints: Dr. Eric Hoffman, Department of Molecular Genetics and Biochemistry, University of Pittsburgh School of Medicine, E1240 Biomedical Science Tower, Pittsburgh, PA 15261. E-mail: eric@hoffman.mgen.pitt.edu

Difficulties in the Estimation of Ethnic Affiliation

To the Editor:

Although I disagree with their results, I am indebted to Shriver et al. (1997) for reawakening my attention to the interesting but tricky subject of the inference of ethnic origin by DNA typing.

They have taken the novel and daunting approach of culling through a vast catalogue of candidate DNA loci to find those which are particularly discriminating. They list a battery of 10 loci, obtained mostly through such a search, which they claim will be effective in determining whether an unknown stain is of African American (AA) or European American (EA) origin. Specifically, they predict that only "0.01% [of individuals will] show log likelihoods <3.0" favoring one origin over the other (Shriver et al. 1997, p. 962). If a prior probability of 50% is assumed for each alternative, this implies the posterior ability to make a correct guess at least 999 times in 1,000. Categorizing Americans as black or white by interviewing them probably does not achieve such a high level of reproducibility, so it seems natural to review with care the basis for such a claim.

I am concerned that the claim rests on serious flaws in statistical methodology. My reanalysis shows that the estimates of efficacy for race determination are significantly overstated because of bias in the algorithm for prediction of likelihood ratios. This is true even for the handful of loci from the literature the authors say that they were able to verify as useful. As for the majority of the recommended loci—those discovered by surveying the catalogue—there is an additional bias that is probably even more serious. I shall discuss a computer simulation that shows that the apparent good performance of the culled loci may be completely illusionary, explainable as mere sampling variation.

These concerns can be conveniently discussed and illustrated in the context of D7S657, the most highly rated of the loci found by the statistical survey. Figure 2 of the Shriver et al. article reveals enough information to allow a check of the calculations for this locus, calculations that assert a typical likelihood ratio of $r = 19$ ($\log_{10} r = 1.276$). I will argue that that number is inflated both by algorithmic errors and by sampling bias. A more realistic likelihood ratio estimation algorithm will reduce the value from 19 to ~8 ($\log_{10} r = 0.9$), and consideration of sampling bias will show that a value of 2.5 ($\log_{10} r = 0.4$) or even less is plausible and consistent with the reported results.

Let $a_1, a_2, \ldots$ and $b_1, b_2, \ldots$ be the allele frequencies at some locus in populations A and B, for alleles 1, 2, $\ldots$, respectively. Then for an allele whose true origin is A and for allele frequencies that are known,

$$\log_{10} r_{\mathrm{AB}} = \sum a_i \log_{10}(a_i/b_i) \qquad (1)$$

is the expected value of the logarithm of the likelihood ratio that the origin is the reference population A rather than the target population B. The formulas in the article by Shriver et al. are equivalent, except that their notation refers to genotypes rather than to single alleles (which explains why their formula has factors of "$\frac{1}{2}$," whereas mine does not), and they formulate a statistic that is

symmetric with respect to A and B. Consequently, they define the "log-likelihood level," which in my notation is $\log_{10} r$ where $r = r_{AB} r_{BA}$. For r itself, the geometric expected value of the likelihood ratio, I use the phrase "typical" likelihood ratio.

In practice, it is necessary to deal with sample frequencies rather than with true frequencies. Therein lies the rub.

An obvious problem with (1) is possible division by zero when some allele is not observed in the target sample. The best strategy in this situation is probably the one that Shriver et al. (1997, p. 958) chose: "an allele not found in a sample is assumed to be the next allele to be observed"; that is, assume a minimum frequency of $b_i = 1/(2n + 1)$, where n is the number of people in the target sample. Apart from this proviso, the authors make computation (1) with sample frequencies $\hat{a}_i$, $\hat{b}_i$, as if they were population frequencies (the "hat" [^] crowning a variable indicates and emphasizes that it is a "statistic" computed from a population sample, and thus a mere estimate of the hatless "parameter"). If the only purpose is a rough comparison of loci, this approach could be accepted uncritically. However, since the likelihood ratio statistic is to be interpreted as what it is—as predicting the likelihood ratio performance that can be expected in practice—then it must be an unbiased estimate. For example, it must satisfy the validation criterion (Brenner 1997) that, if one compares two identical populations, then the likelihood ratio statistic should not imply that people will be more likely to come from one than from the other. Imagine collecting two n-person samples S and T from the same population A. The sample frequencies will, of course, by random chance, differ somewhat, so a likelihood ratio analysis of the samples, even a correct one, will sometimes suggest that A can be distinguished from A. However, when an average is taken over all samples S and T, there should be no expected distinction. The average log likelihood ratio should be zero. As an experiment, I posited a population with allele frequencies approximately like the D7S657 AA sample frequencies. Testing the Shriver et al. formulas on repeated computer samples S and T from this population, I found that $\log_{10} \hat{r}$ averaged ∼0.11 ($\hat{r} = 1.3$) for distinguishing the simulated AA population from itself. In other words, there was a bias of ∼30%.

A procedure that does seem to survive the validation test is the one used by Erikson and Svensmark (1994). They added one to the target sample count for every allele—not just for the unseen ones as do Shriver et al. When this method is applied to the D7S657 data, the likelihood ratio drops to $\hat{r} = 8$ ($\log_{10} \hat{r} = 0.9$).

The reader may wonder why the discrepancy between 19 and 8 is much larger than the 30% accounted for by the above validation test. It turns out that most of the discrepancy has a less subtle explanation. Those values

that can be checked in Shriver et al.'s (1997) table 1 are consistent with misinterpreting n as the sample size in alleles rather than in people, suggesting a solecism in their computer implementation (a bug). Correcting the arithmetic would give $\log_{10} \hat{r} = 1.09$ (so $\hat{r} = 12$).

The most powerful locus in the Shriver et al. article is FY-null (Duffy blank), for which they give $\log_{10} \hat{r} = 1.858$ ($\hat{r} = 72$) for distinguishing AA and EA. The above bug is not an issue here since there were no zero-count alleles; nor is sampling bias an issue since FY-null is among the loci from the literature rather than from the statistical survey. Nonetheless, substituting the Erikson and Svensmark procedure reduces $\hat{r}$ to 39 ($\log_{10} \hat{r} = 1.59$), and that is probably a fairer guess of the efficacy of this locus based on the data given. It is still a very discriminating locus. In ∼96% of the cases in which an unknown stain donor is African American, this locus alone will answer the question of ethnic origin. But a difficult question about allelic association: In estimating how many of the remaining 4% of such cases will be resolved by other loci, is it correct to use the overall AA allele frequencies, which after all come mostly from people who have FY-null?

It may be of interest to compare the r values discussed above with values for typical forensic loci, not intentionally selected for their ethnic-discrimination potential. Data from the Office of the Chief Medical Examiner of New York City (personal communication) comparing AA and EA samples ($n = 118$ and $n = 107$) in the tetrameric loci F13A, TH01, FES/FPS, and VWA, give values $\hat{r} = 2.56$, 1.56, 1.43, and 1.17 ($\log_{10} \hat{r} \leqslant 0.4$). With somewhat more care and difficulty, the same sort of evaluation can be made for RFLP loci. Six of them, in common use by United States law-enforcement and paternity laboratories for identification, average $r > 2$ ($\log_{10} r > 0.3$) per locus (Brenner 1997). Incidentally, whereas Shriver et al. (1997, p. 957) say in their paper that "most [DNA markers] offer little power to distinguish ethnicity," a handful of independent markers with $r = 2$ provide quite useful power (Brenner 1997), certainly better than "the best racial estimates [, which] are achieved" by bone and skull measurements, giving 80%–90% correct categorization as cited in Shriver et al.'s (1997, p. 958) paper. Further, with a computationally more sophisticated approach (Evett et al. 1992; Brenner 1997), the RFLP efficacy increases to about $r = 3$ ($\log_{10} r = 0.5$) per locus. Combining a handful of loci of such power is sufficient to decide most cases with confidence, but the predicted distribution is such that 5%–10% of cases remain elusively ambiguous.

So compared with the standard repertoire of forensic loci, the claims of Shriver et al., even trimmed back a few orders of magnitude by the arguments that I have made above, would still be impressive. By my rough

estimates, the top 10 loci would still rigidly categorize Americans as black or white, to an implausible extent.

My remaining concern, the most vexing and possibly the most telling, is sampling bias. Of the 20 high-performing loci in Shriver et al.'s table 1 for discrimination between AA and EA, 17 were obtained by canvassing >1,000 loci. To be more precise, what they canvassed is >1,000 *pairs of samples* and sometimes rather small samples (e.g., $n = 21$ people). This suggests the possibility that most of the "high performers" are really ordinary performers with an atypically lucky sample.

How much can be explained by luck depends on the sampling distribution of the likelihood level statistic, $\hat{r}$, which I have investigated with a Monte Carlo computer experiment. Each experiment begins with 1,000 simulated loci whose AA and EA allele frequencies are assigned according to one or another of the New York data mentioned above, so $1.17 \leqslant r \leqslant 2.56$ ($0.08 \leqslant \log_{10}r \leqslant 0.4$) for each simulated locus. For each of these loci, a 21-person sample and a 22-person sample (mimicking the D7S657 sample sizes) are randomly selected according to the assumed frequencies, and the statistic $\hat{r}$ is computed from the two samples. To simplify the comparison with Shriver et al.'s table 1, I used the same (albeit incorrect, as per above discussion) formulas as were used for that table.

The 17 largest $\hat{r}$ values from such a 1,000-locus experiment are similar to the values for the 17 canvassed loci (out of 20 total) in the AA/EA column of table 1. The largest value is sometimes a little larger, sometimes a little smaller, than $\hat{r} = 19$ ($\log_{10}\hat{r} = 1.276$) of D7S657. The 17th largest simulated $\hat{r} \approx 5$ ($\log_{10}\hat{r} \approx 0.7$)—easily comparable to $\hat{r} = 3$ ($\log_{10}\hat{r} = 0.498$) in table 1. One might say that what the computer experiment screens is not nature but sampling variation. It lists loci with merely ordinary ethnic-discrimination power, but with extraordinary statistics. From among 1,000 loci, one could similarly find a set of 10 loci that differentiate the 9-year-old children from the 10-year-olds in the local playground. In the phrase of one of the referees of this letter, the process has the potential to create the appearance of signal where there is only noise.

Is the sieving procedure of Shriver et al. any different from the computer experiment? The bias problem would be mitigated if their sample sizes were mostly larger, or if some loci were screened twice. This may have been done to some extent; the description in the Shriver et al. paper is not explicit. Also, there is of course a tendency for the better loci to achieve a better score. But as I have shown, there is a strong countervailing tendency that the list of top scores will be dominated by scores that are particularly biased. Therefore, I do not believe that their conclusion—namely, that they have found "a set of genetic markers that would allow the confident determi-

nation of ethnicity" (Shriver et al. 1997, p. 962)—is likely to be correct.

CHARLES H. BRENNER

Electronic-Database Information

Brenner CH (1997) http://www.ccnet.com/~cbrenner/race.htm

References

Brenner CH (1997) Probable race of a stain donor. In: Proceedings from the Seventh International Symposium on Human Identification 1996. Promega, Madison, pp 48–52

Erikson B, Svensmark O (1994) DNA polymorphism in Greenland. Int J Legal Med 106:254–257

Evett IW, Pinchin R, Buffery C (1992) An investigation of the feasibility of inferring ethnic origin from DNA profiles. J Forensic Sci Soc 32:301–306

Shriver MD, Smith MW, Jin L, Marcini A, Akey JM, Deka R, Ferrell RE (1997) Ethnic-affiliation estimation by use of population-specific DNA markers. Am J Hum Genet 60: 957–964

Address for correspondence and reprints: Dr. Charles H. Brenner, 2486 Hilgard Avenue, Berkeley, CA 94709. E-mail: cbrenner@ccnet.com

Reply to Brenner

To the Editor:

In response to the letter by Dr. Brenner (1998 [in this issue]), there are a number of issues open for discussion with regard to both our previously published article (Shriver et al. 1997) and, more generally, methods for estimation of biological ancestry. Dr. Brenner has identified some specific concerns with regard to our methods and results, which we address below. However, we remain confident of the main conclusions of our study: (1) the reliable estimation of ethnic affiliation by use of population-specific alleles (PSAs) is possible; and (2) many of the loci we identified will be useful markers for this effort.

We have examined the computer program that was used to calculate average single-locus log-likelihood levels and have found that Dr. Brenner is correct in his determination that alleles that were not observed were assigned a frequency of $1/(4n + 1)$, instead of $1/(2n + 1)$, where n is the number of individuals in the sample. The effect of this error was to inflate the average single-locus and multilocus log-likelihood estimates, to a small degree. Since the same program was used to screen all the allele-frequency data sets, it is reasonable to conclude

that the 40 loci with the highest log-likelihood levels, which we presented in tables 1 and 2 of our article (Shriver et al. 1997), are still good candidates for high performers among the loci tested.

Dr. Brenner is correct to recognize that our method for determining average single-locus log-likelihood ratios (LLRs) and multilocus ethnic-affiliation estimates is appropriate only when accurate allele-frequency data are available. We expect that, in the determination of biological ancestry, care will be taken to determine with precision the allele frequencies of potential contributing populations. If accurate allele frequencies are available (e.g., $n > 200$ individuals), no adjustment of the formula we presented will be needed. In cases for which frequency data are available only from small samples, the addition of one to the total allele count for each allele is a reasonable adjustment.

Dr. Brenner concludes that the differences in allele frequency that we observed between loci were largely due to bias resulting from small sample size. He bases this conclusion on a computer simulation in which he evidently resampled 1,000 × from frequency data on four short tandem-repeat identity markers. He then compared his results with the data in table 1 of our article (Shriver et al. 1997). We have two concerns with this approach. First, the 17 microsatellite PSAs that we presented in table 1 were culled from ~350 loci (1,000 loci/population combinations were tested in the work that we reported). Second, the range of variation in the frequency differential used in Dr. Brenner's model was very limited and, with only four loci (LLR of .08–.4), could not have reflected naturally observed levels of variation in the allele-frequency differential. We are well aware of the bias resulting from small sample sizes, which is why we presented a list of 20 loci in table 1 and not just the best 10. In fact, we stated, "It should be noted that the markers on this list need to be typed in larger samples from different parts of the country, both to have more accurate allele-frequency estimates and to identify the most efficient set for EAE [ethnic-affiliation estimation]" (Shriver et al. 1997, p. 963). Recently, we typed nine dimorphic autosomal PSAs in large samples from >20 ethnographically defined populations, including 12 African-American population samples, and indeed found these markers to be useful for the estimation of ethnic affiliation and admixture (Parra et al. 1997; E. J. Parra, A. Marcini, L. Jin, J. Akey, M. Batzer, R. Cooper, T. Forrester, et al., unpublished data). Overall and in view of Dr. Brenner's concerns, we still feel that this is a viable approach for the estimation of the biological ancestry of a person and that we have provided an important list of putative PSAs for this purpose.

Finally, in responding to Dr. Brenner's comments, we would like to suggest an alternative phrase that more accurately describes what is being estimated by means of the markers and methods that we, Dr. Brenner, and others have described. Ethnicity is a term that directly refers to the culture of a person or people and that encompasses their language, traditions, and national identity. Ethnicity is often related to biological ancestry but not always. In the United States, awkward terms that combine both ethnicity and biological ancestry are sometimes used—for example, "non-Hispanic whites," "black Hispanics," and "non-Hispanic blacks." Modern populations are highly complex, and the classification of genetic differences among individuals and populations is a potentially sensitive issue. We therefore propose and intend to use the term "estimation of biological ancestry," rather than "ethnic-affiliation estimation," to describe the methods that we have presented.

MARK D. SHRIVER,[1] MICHAEL W. SMITH,[2] AND LI JIN[3]
[1]Department of Human Genetics, Allegheny University of the Health Sciences, Pittsburgh; [2]National Cancer Institute, Frederick Cancer Research and Development Center, Frederick, MD; and [3]Human Genetics Centers, University of Texas, Houston

References

Brenner CH (1998) Difficulties in the estimation of ethnic affiliation. Am J Hum Genet 62:1558–1560 (in this issue)
Parra E, Marcini A, Akey J, Ferrell RE, Shriver MD (1997) A systematic study of African-American admixture using population-specific alleles. Am J Hum Genet Suppl 61:A17
Shriver MD, Smith MW, Jin L, Marcini A, Akey JM, Deka R, Ferrell RE (1997) Ethnic-affiliation estimation by use of population-specific DNA markers. Am J Hum Genet 60: 957–964

Address for correspondence and reprints: Dr. Mark D. Shriver, Department of Human Genetics, Allegheny University of the Health Sciences, 3290 William Pitt Way, Building B4, Room 125, Pittsburgh, PA 15212-4772. E-mail: mshriver@phg.auhs.edu

Discriminating between True and False-Positive Peaks in a Genomewide Linkage Scan, by Use of the Peak Length

To the Editor:
A standard method to map disease-susceptibility loci consists of collecting n affected sib pairs and their parents, genotyping them for a dense set of genetic markers, and counting, at each marker locus t, the number, X_t, of parental alleles shared identical by descent (IBD). According to current statistical practice (e.g., see Feingold

et al. 1993; Lander and Kruglyak 1995), only the height of the peak (i.e., $\max_t X_t$) is used to decide whether there is sufficient evidence in favor of linkage. Recently, Terwilliger et al. (1997) published in the *Journal* a paper in which they challenged the claim that "there is no way to distinguish between small peaks that represent weak true positives and peaks of the same height arising from random fluctuations" (Lander and Kruglyak 1995, p. 244). By relying on some deeper results of the theory of stochastic processes, Terwilliger et al. (1997) showed that true positive peaks are expected to be longer than false-positive peaks. The purpose of this letter is to explain and support their result by presenting an extremely simplified situation in which elementary argumentation is sufficient to obtain the same conclusion.

Let M_1 and M_2 denote two different but linked marker loci separated by a genetic distance of $\theta \in (0, \frac{1}{2})$. The data consist of a single affected sib pair and only one parent. Both markers are assumed to be completely polymorphic. Thus, it is possible to decide whether, from this parent, the sib pair has inherited the same allele ($X_i = 1$) or has not inherited the same allele ($X_i = 0$), at marker locus i ($i = 1,2$). Let $p = P(X_1 = 1)$ denote the probability that the sib pair share an allele IBD at M_1. If M_1 is unlinked to the disease, then $p = \frac{1}{2}$; if M_1 is a disease locus, then $p > \frac{1}{2}$. In both cases, the conditional probabilities for the IBD score at the second marker locus, given the IBD score at the first marker locus, depend only on the genetic distance θ between M_1 and M_2 —that is, $P(X_2 = 1 \mid X_1 = 1) = P(X_2 = 0 \mid X_1 = 0) = \theta^2 + (1-\theta)^2 =: \Psi$ and $P(X_2 = 1 \mid X_1 = 0) = P(X_2 = 0 \mid X_1 = 1) = 1 - \Psi$. The joint probability distribution of (X_1, X_2) is given in table 1.

Now the terms "peak" and "length of a peak" have to be defined. A peak occurs at marker locus i ($i \in \{1,2\}$) if $X_i = 1$. Given that there is a peak at marker locus i, the length of this peak is either 2 (if $X_{3-i} = 1$) or 1 (if $X_{3-i} = 0$). The following three conclusions are evident from table 1:

1. Given a peak at M_1, the length of this peak is 2, with probability Ψ. Thus, this probability does not depend on p.

2. However, given a peak at M_2, the length of this peak is 2, with probability $\Psi p / \{\Psi p + [(1-\Psi)(1-p)]\}$. Since $\Psi < 1$, this is a strictly increasing function in p, for $p \in [\frac{1}{2}, 1]$.

3. Given a peak "somewhere in the genome" (i.e., at M_1 and/or M_2), the length of this peak is 2, with probability $\Psi p / \{1 - [\Psi(1-p)]\}$. Since the value for this expression is strictly increasing with p, this shows that true peaks ($p > \frac{1}{2}$) are expected to be longer than false-positive peaks ($p = \frac{1}{2}$).

Table 1

Joint Probability Distribution of (X_1, X_2)

l	m	$P(X_1 = l, X_2 = m)$
1	1	$\Psi \cdot p$
1	0	$(1 - \Psi) \cdot p$
0	1	$(1 - \Psi) \cdot (1 - p)$
0	0	$\Psi \cdot (1 - p)$

This example can be extended to a consideration of statistical testing. Let α denote an arbitrary but fixed value, with $\alpha \leqslant \Psi/2$. Since $P_{p=\frac{1}{2}}(\max_{t \in \{1,2\}} X_t = 1) = 1 - (\Psi/2)$, a (randomized) level α test based on the test statistic $\max_{t \in \{1,2\}} X_t$ is obtained by rejection of the null hypothesis of no linkage, with probability $\gamma_1 := \alpha/[1 - (\Psi/2)]$. The power of this test is $\alpha \cdot \{1 - [\Psi(1 - p)]\}/[1 - (\Psi/2)]$. Alternatively, a test based on the length of a peak can be constructed in the following way: for $p = \frac{1}{2}$, the probability is $\Psi/2$ that a peak of length 2 occurs. Thus, a randomized level α test based on the length of a peak is obtained by rejection of the null hypothesis, with probability $\gamma_2 := 2\alpha/\Psi$. The power of this test is $\alpha \cdot 2 \cdot p$. Since $2p > \{1 - [\Psi(1-p)]\}/[1 - (\Psi/2)]$ for $p > \frac{1}{2}$, the second test, which is based on the length of a peak, is more powerful than the test based solely on the height of a peak. For a more realistic and relevant situation than the one considered in the present letter, it has to be determined how both height and length of a peak can be combined in a test for linkage. However, the observation described by Terwilliger et al. (1997) may prove very useful to increase the power for detection of disease-susceptibility loci by genomewide linkage scans.

MICHAEL KNAPP

Institute for Medical Statistics
University of Bonn
Bonn

References

Feingold E, Brown PO, Siegmund D (1993) Gaussian models for genetic linkage analysis using complete high-resolution maps of identity by descent. Am J Hum Genet 53:234–251

Lander ES, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat Genet 11:241–247

Terwilliger JD, Shannon WD, Lathrop GM, Nolan JP, Goldin LR, Chase GA, Weeks DE (1997) True and false positive peaks in genomewide scans: application of length-biased sampling to linkage mapping. Am J Hum Genet 61:430–438