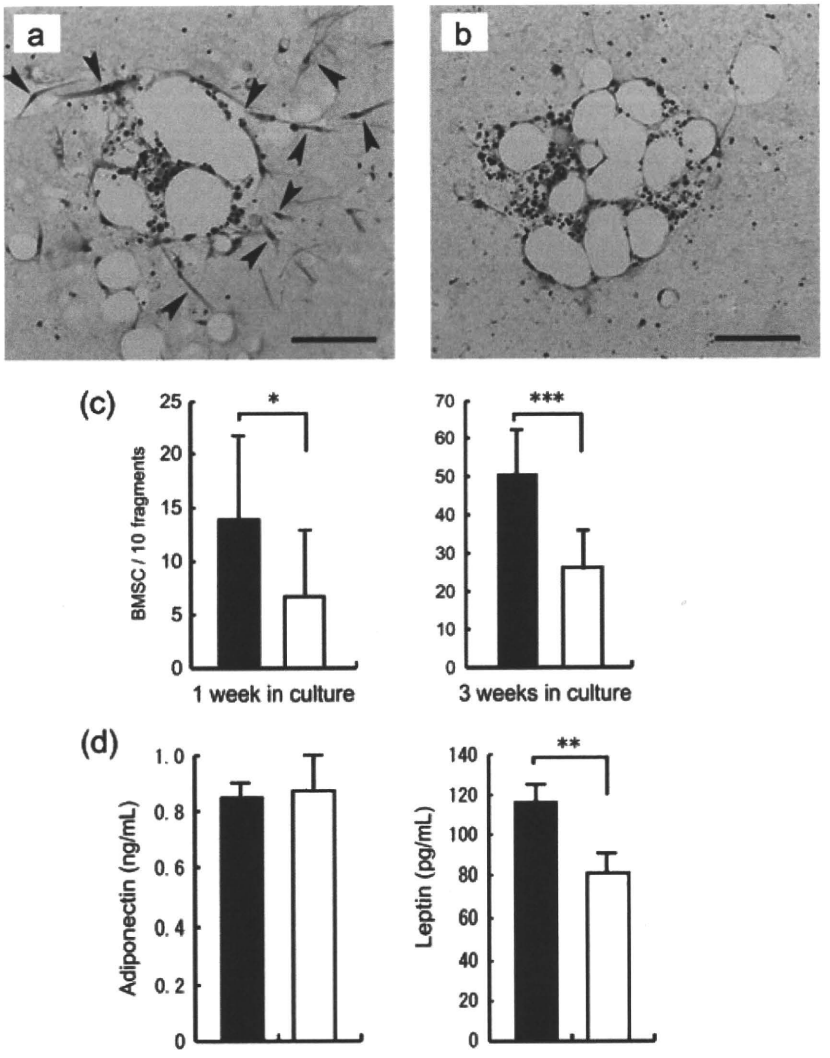

Figure 6 (**a–c**) Development of bone marrow stromal cells (BMSC) from bone marrow adipose tissue (BMAT) fragments and (**d**) production of adiponectin and leptin in culture with or without osteoblasts. (**a**) Numerous BMSC (arrowheads) appear around the BMAT fragments in the absence of osteoblasts at 3 weeks. (**b**) In contrast, development of BMSC from the fragments is inhibited in co-culture with osteoblasts. Bars, 100 μm. (**c**) Significant difference is seen between the number of BMSC in culture (□) with or (■) without osteoblasts (*$P < 0.05$ at 1 week; ***$P < 0.001$ at 3 weeks). (**d**) Adiponectin production in BMAT cultured (■) without osteoblasts is very low and (□) with osteoblasts is unaffected. Leptin is detected in cultures (■) without osteoblasts (control), and inhibited (□) with osteoblasts.

interaction between hematopoietic cells and adipocytes as a stromal cell type. To address an interaction between BMAT and hematopoietic cells in more detail, hematopoietic cell-rich marrow obtained from ilium or sternum would be more suitable as materials.

Our co-culture method creates the following active interaction between BMAT and osteoblasts: (i) osteoblasts inhibit the development of BMSC from BMAT, leptin secretion from BMAT and growth of leukocyte-linked hematopoietic cells; and (ii) BMAT in turn inhibits osteoblast growth and differentiation. These interaction-based phenomena took place actively in a paracrine manner, because the osteoblast layer was completely separated from the BMAT layer by the nitrocellulose membrane. Their reciprocal suppressive effects may be involved in the regulatory mechanisms of bone marrow homeostasis. In order to define the mechanisms underlying the osteoblast-BMAT interaction, further studies on the detection of soluble mediators are required.

In conclusion, our culture system established an active interaction between BMAT and osteoblasts. Thus, this method may prove a promising tool for the analysis of interactions between multiple cell types within BMAT, and for the study of skeletal bone marrow pathophysiology during osteoporosis and other hematopoietic disorders.

ACKNOWLEDGMENTS

This work was supported in part by Grants-in-Aid from the Japanese Ministry of Education, Culture, Sports, Science and Technology for Scientific Research nos. 18591871 and 20592023, and by personal grants from Koike Hospital, Sasebo Central Hospital and Yamada Clinic (to Professor Shuji Toda). We also thank Associate Professor M. Mawatari for helpful suggestions, and Mr H. Ideguchi, F. Mutoh, S. Nakahara and Mrs M. Nishida for their excellent technical assistance.

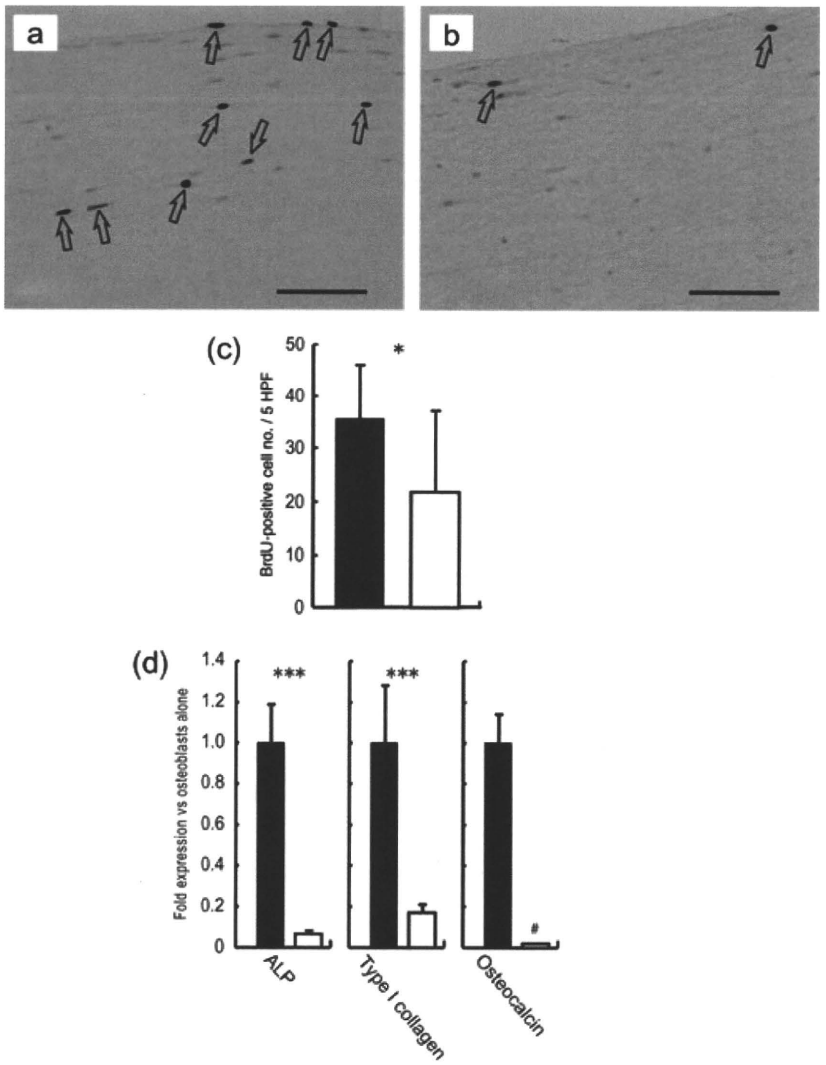

Figure 7 After 1 week in culture, the number of (**a**) bromodeoxyuridine (BrdU)-positive osteoblasts (arrows) without bone marrow adipose tissue (BMAT) fragments is higher than that (arrows) (**b**) with the fragments. Bars, 100 μm. (**c**) Significant difference is seen in the number of BrdU-positive osteoblasts between cultures (□) with and (■) without BMAT fragments (*P < 0.05). (**d**) Alkaline phosphatase (ALP), type I collagen and osteocalcin expression in osteoblasts (□) with or (■) without BMAT analyzed using real-time reverse transcription–polymerase chain reaction. BMAT inhibits ALP and type I collagen expression in osteoblasts (***P < 0.001). Osteoblasts without BMAT clearly express osteocalcin, whereas the cells with BMAT express it below the detectable range (#). HPF, high-power field.

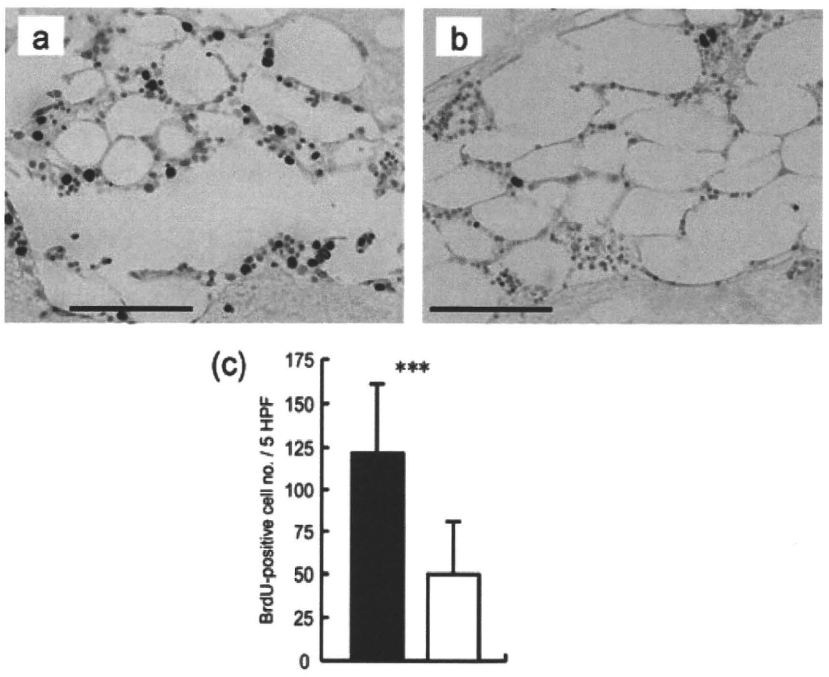

Figure 8 The number of bromodeoxyuridine (BrdU)-positive cells in culture (**a**) without osteoblasts (i.e. bone marrow adipose tissue (BMAT) alone) is higher than that (**b**) in culture with osteoblasts. Bars, 100 μm. (**c**) Significant difference is seen in the number of BrdU-positive cells between cultures (□) with and (■) without osteoblasts (***P < 0.001). These findings suggest that osteoblasts inhibit the growth of mononuclear cells within BMAT. HPF, high-power field.

REFERENCES

1 Gimble JM, Robinson CE, Wu X, Kelly KA. The function of adipocytes in the bone marrow stroma: An update. *Bone* 1996; **19**: 421–8.
2 Sonoda E, Aoki S, Uchihashi K *et al.* A new organotypic culture of adipose tissue fragments maintains viable mature adipocytes for a long term, together with development of immature adipocytes and mesenchymal stem cell-like cells. *Endocrinology* 2008; **149**: 4794–8.
3 Toda S, Uchihashi K, Aoki S *et al.* Adipose tissue-organotypic culture system as a promising model for studying adipose tissue biology and regeneration. *Organogenesis* 2009; **5**: 43–9.
4 Kitajima M, Shigematsu M, Ogawa K, Sugihara H, Hotokebuchi T. Effects of glucocorticoid on adipocyte size in human bone marrow. *Med Mol Morphol* 2007; **40**: 150–56.
5 Sugihara H, Funatsumaru S, Yonemitsu N, Miyabara S, Toda S, Hikichi Y. A simple culture method of fat cells from mature fat tissue fragments. *J Lipid Res* 1989; **30**: 1987–95.
6 Sugihara H, Yonemitsu N, Miyabara S, Yun K. Primary cultures of unilocular fat cells: Characteristics of growth *in vitro* and changes in differentiation properties. *Differentiation* 1986; **31**: 42–9.
7 Aoki S, Toda S, Sakemi T, Sugihara H. Coculture of endothelial cells and mature adipocytes actively promotes immature preadipocyte development *in vitro*. *Cell Struct Funct* 2003; **28**: 55–60.
8 Fawcett DW. *A Text Book of Histology*. New York: Chapman and Hall, 1994.
9 Braylan RC, Orfao A, Borowitz MJ, Davis BH. Optimal number of reagents required to evaluate hematolymphoid neoplasias: Results of an international consensus meeting. *Cytometry* 2001; **46**: 23–7.
10 Toda S, Matsumura S, Fujitani N, Nishimura T, Yonemitsu N, Sugihara H. Transforming growth factor-beta1 induces a mesenchyme-like cell shape without epithelial polarization in thyrocytes and inhibits thyroid folliculogenesis in collagen gel culture. *Endocrinology* 1997; **138**: 5561–75.
11 Aslan H, Zilberman Y, Kandel L *et al.* Osteogenic differentiation of noncultured immunoisolated bone marrow-derived CD105+ cells. *Stem Cells* 2006; **24**: 1728–37.
12 Zuk PA, Zhu M, Ashjian P *et al.* Human adipose tissue is a source of multipotent stem cells. *Mol Biol Cell* 2002; **13**: 4279–95.
13 Suzuki T, Tate G, Ikeda K, Mitsuya T. A novel multicolor immunofluorescence method using heat treatment. *Acta Med Okayama* 2005; **59**: 145–51.
14 Toda S, Sugihara H. Reconstruction of thyroid follicles from isolated porcine follicle cells in three-dimensional collagen gel culture. *Endocrinology* 1990; **126**: 2027–34.
15 Kras KM, Hausman DB, Martin RJ. Tumor necrosis factor-alpha stimulates cell proliferation in adipose tissue-derived stromal-vascular cell culture: Promotion of adipose tissue expansion by paracrine growth factors. *Obes Res* 2000; **8**: 186–93.
16 Canova N, Lincova D, Farghali H. Inconsistent role of nitric oxide on lipolysis in isolated rat adipocytes. *Physiological Res Acad Sci Bohemoslov* 2005; **54**: 387–93.
17 Sudo H, Kodama HA, Amagai Y, Yamamoto S, Kasai S. *In vitro* differentiation and calcification in a new clonal osteogenic cell line derived from newborn mouse calvaria. *J Cell Biol* 1983; **96**: 191–8.
18 Sugihara H, Yonemitsu N, Toda S, Miyabara S, Funatsumaru S, Matsumoto T. Unilocular fat cells in three-dimensional collagen gel matrix culture. *J Lipid Res* 1988; **29**: 691–7.
19 Gimble JM, Dorheim MA, Cheng Q *et al.* Adipogenesis in a murine bone marrow stromal cell line capable of supporting B lineage lymphocyte growth and proliferation: Biochemical and molecular characterization. *Eur J Immunol* 1990; **20**: 379–87.
20 Laharrague P, Larrouy D, Fontanilles AM *et al.* High expression of leptin by human bone marrow adipocytes in primary culture. *FASEB J* 1998; **12**: 747–52.
21 Wang B, Trayhurn P. Acute and prolonged effects of TNF-alpha on the expression and secretion of inflammation-related adipokines by human adipocytes differentiated in culture. *Pflugers Arch* 2006; **452**: 418–27.
22 Shimizu K, Sakai M, Ando M *et al.* Newly developed primary culture of rat visceral adipocytes and their *in vitro* characteristics. *Cell Biol Int* 2006; **30**: 381–8.
23 Thomas T, Gori F, Khosla S, Jensen MD, Burguera B, Riggs BL. Leptin acts on human marrow stromal cells to enhance differentiation to osteoblasts and to inhibit differentiation to adipocytes. *Endocrinology* 1999; **140**: 1630–38.
24 Umemoto Y, Tsuji K, Yang FC *et al.* Leptin stimulates the proliferation of murine myelocytic and primitive hematopoietic progenitor cells. *Blood* 1997; **90**: 3438–43.
25 Ricote M, Li AC, Willson TM, Kelly CJ, Glass CK. The peroxisome proliferator-activated receptor-gamma is a negative regulator of macrophage activation. *Nature* 1998; **391**: 79–82.
26 Takagi M. Cell processing engineering for ex-vivo expansion of hematopoietic cells. *J Biosci Bioeng* 2005; **99**: 189–96.
27 Dexter TM, Moore MA, Sheridan AP. Maintenance of hemopoietic stem cells and production of differentiated progeny in allogeneic and semiallogeneic bone marrow chimeras *in vitro*. *J Exp Med* 1977; **145**: 1612–16.

Journal of Human Genetics (2010) 55, 124–126

www.nature.com/jhg

SHORT COMMUNICATION

A type of familial cleft of the soft palate maps to 2p24.2–p24.1 or 2p21–p12

Masayoshi Tsuda[1,2,9], Takahiro Yamada[3,9], Tadashi Mikoya[4], Izumi Sogabe[5], Mitsuko Nakashima[1,2,6], Hisanori Minakami[3], Tatsuya Kishino[7], Akira Kinoshita[1], Norio Niikawa[8], Akiyoshi Hirano[2] and Koh-ichiro Yoshiura[1]

Cleft of the soft palate (CSP) and the hard palate are subtypes of cleft palate. Patients with either condition often have difficulty with speech and swallowing. Nonsyndromic, cleft palate isolated has been reported to be associated with several genes, but to our knowledge, there have been no detailed genetic investigations of CSP. We performed a genome-wide linkage analysis using a single-nucleotide polymorphism-based microarray platform and successively using microsatellite markers in a family in which six members, across three successive generations, had CSP. A maximum LOD score of 2.408 was obtained at 2p24.2–24.1 and 2p21–p12, assuming autosomal dominant inheritance. Our results suggest that either of these regions is responsible for this type of CSP.

Journal of Human Genetics (2010) **55,** 124–126; doi:10.1038/jhg.2009.131; published online 15 January 2010

Keywords: cleft of the soft palate; genome-wide linkage analysis; submucous cleft palate

INTRODUCTION

Orofacial cleft, one of the most common congenital malformations, is a heterogeneous group of complex traits. Orofacial cleft is classified into two main categories, cleft lip with or without cleft palate and cleft palate isolated (CPI). Both clefting phenotypes can appear to be related to some syndromes (syndromic orofacial cleft) or not be related to syndromes (nonsyndromic orofacial cleft). CPI is considered genetically distinct from cleft lip with or without cleft palate, on the basis of epidemiological evidence and the different developmental timing of lip and palate formation. Recent molecular genetic studies[1–5] have identified genes or loci that are responsible for CPI. However, fewer genes and/or loci-associated CPI have been reported in comparison with cleft lip with or without cleft palate.[6]

CPI is mostly classified into two subtypes morphologically: cleft of the hard palate (CHP) and cleft of the soft palate (CSP).[7] Submucous cleft palate (SMCP) is a small subgroup in the CPI. SMCP manifests with bifid uvula, separation of the muscle with an intact mucosa and a bony defect in the posterior edge of the hard palate.[8] Both CHP and CSP are caused by a failure of fusion of the palatal shelves, but little is known about the cause of the difference in their phenotypes. Christensen *et al.*[9] suggested that CHP and CSP might be etiologically distinct. Although patients with CSP have serious problems in speech and deglutition, as well as CHP, there have been no detailed genetic studies performed.

We recently encountered a Japanese family that included five CSP patients and one SMCP patient. The aim of this study was to identify the CSP/SMCP predisposing locus in this family using genome-wide single-nucleotide polymorphism (SNP)-based linkage analysis.

MATERIALS AND METHODS

Family and patients

A Japanese family included five patients (I-2, II-2, II-3, III-1 and III-2) with CSP and one patient (II-5) with SMCP across three generations (Figure 1). Two patients (II-2 and II-3) were monozygotic twins. The phenotypes of two patients (III-1 and II-5) were shown in Figure 2. All patients had no other symptoms such as mental retardation, and all family members were examined by one or two well-trained dentists.

The disease in the family was consistent with an autosomal dominant mode of inheritance. Blood samples were obtained with written informed consent from 15 cooperative family members (Figure 1). The study protocol was approved by the committee for ethical issues on the Human Genome and Gene Analysis of Nagasaki University.

[1]Department of Human Genetics, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan; [2]Department of Plastic and Reconstructive Surgery, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan; [3]Department of Obstetrics and Gynecology, Hokkaido University Graduate School of Medicine, Sapporo, Japan; [4]Center for Advanced Oral Medicine, Hokkaido University Hospital, Sapporo, Japan; [5]Department of Maxillofacial Surgery, Hokkaido University Graduate School of Medicine, Sapporo, Japan; [6]Laboratory of Molecular Medicine and Laboratory of Genome Technology of the Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan; [7]Division of Functional Genomics, Center for Frontier Life Sciences, Nagasaki University, Nagasaki, Japan and [8]Research Institute of Personalized Health Sciences, Health Sciences University of Hokkaido, Tobetsu, Japan
[9]These authors contributed equally to this work.
Correspondence: Dr K-i Yoshiura, Department of Human Genetics, Nagasaki University, Graduate School of Biomedical Sciences, 1-12-4, Sakamoto, Nagasaki 852-8523, Japan.
E-mail: kyoshi@nagasaki-u.ac.jp
Received 26 July 2009; revised 30 October 2009; accepted 13 November 2009; published online 15 January 2010

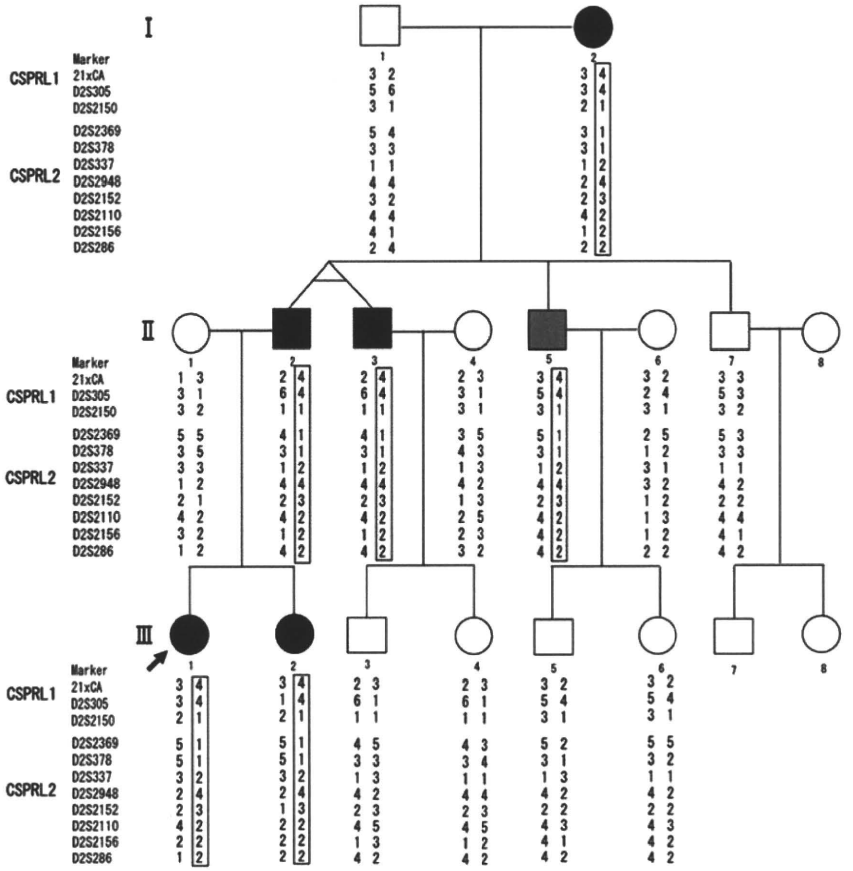

Figure 1 Family tree with haplotypes at 2p24.2–24.1 (CSPR1) and 2p21–p12 (CSPR2). Black closed, gray closed and open symbols indicate affected with cleft of the soft palate (CSP), affected with subcutaneous cleft palate (SMCP) and unaffected, respectively. An arrow indicates the proband. Genotypes of microsatellite markers defining the candidate intervals are shown below each individual. Boxed haplotype indicates possibly disease-associated haplotype.

SNP genotyping and linkage analysis

Genomic DNA was extracted from peripheral blood lymphocytes of the 15 members, using a QIAamp DNA Mini Kit (QIAGEN, Hilden, Germany). Their genotypes were determined using a GeneChip Human Mapping 10K 2.0 Xba Array (Affymetrix, Santa Clara, CA, USA). We used MERLIN software[10] to analyze compiled pedigree data sets. Mendelian errors were detected by PEDCHECK,[11] and SNPs with Mendelian error were not used in the data analysis. LOD scores were calculated under a parametric autosomal dominant model in which penetrance was set to 1.0 and disease allele frequency was 0.00001. As CSP and SMCP can be categorized together because of their similar anatomical features,[9] the patient with SMCP and the patients with CSP (II-5) were classified as 'affected' for linkage score calculations.

To confirm the result of linkage data using the GeneChip Human Mapping 10K 2.0 Xba Array, we performed a two-point linkage analysis using microsatellite markers by the method described elsewhere.[12] The two-point LOD score was calculated using MLINK program.[13]

RESULTS AND DISCUSSION

In the assay with the 10K-Array, the GeneChip call rates varied from 92.18 to 99.42% (with a mean of 97.54%). Two regions, 2p24.2–p24.1 (CSP region 1: CSPR1), a 4.5-Mb interval between rs1545497 and rs1872325, and 2p21–p12 (CSP region 2: CSPR2), a 34.5-Mb segment between rs940053 and rs310777, were CSP candidate loci with a maximum LOD score of 2.408 (Figure 3). The LOD scores of all other regions were below 1.000. Two-point LOD scores using microsatellite markers showed the same scores (2.408); therefore, the result of linkage analysis from SNP genotyping was reconfirmed (haplotype using microsatellite markers was shown in Figure 1). It is thus likely that a gene having a role in palatal fusion is located within either CSPR1 or CSPR2.

On the basis of our knowledge of oral palate development, we chose nine genes from the candidate CSP regions and performed mutation analysis. Of the nine candidate genes, three were from CSPR1: growth/differentiation factor 7 (*GDF7*), matrilin 3 (*MATN3*) and member B of the Ras homolog gene family (*RHOB*). The other six genes were from CSPR2: calmodulin 2 (*CALM2*), bone morphologic protein 10 (*BMP10*), sprouty-related EVH1 domain-containing protein 2 (*SPRED2*), transforming growth factor, alfa (*TGFA*), ventral anterior homeobox 2 (*VAX2*; 2p13.3) and stoned B-like factor/stonin 1 (*STON1*). Most of these genes are concerned with bone development, and with the transforming growth factor and mitogen-activated protein kinase signaling pathways, or are transcription factors related to homeobox genes. However, no pathogenic mutation was found within any of its exons or intron–exon boundaries of all nine genes.

To detect structural genomic alterations that may cause CSP within the candidate regions, we performed copy number analysis with the proband's DNA using the Genome-Wide Human SNP Array 5.0 (Affymetrix). Although several copy number alterations were detected (data were not shown), all were already registered as copy number variations on the UCSC Genome Browser (http://genome.ucsc.edu/) and none of them coincided with regions with positive LOD scores.

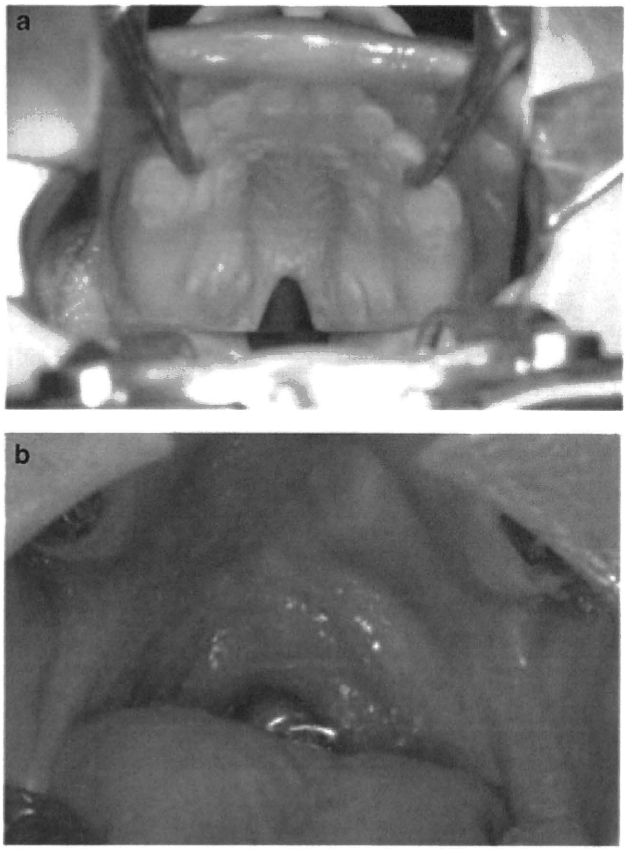

Figure 2 Views of palates. The palate of individual III-1 with CSP (**a**) showing a cleft limited to the soft palate, and that of individual II-5 with SMCP (**b**) showing a translucent zone in the soft palate resulting from a separation of the muscle.

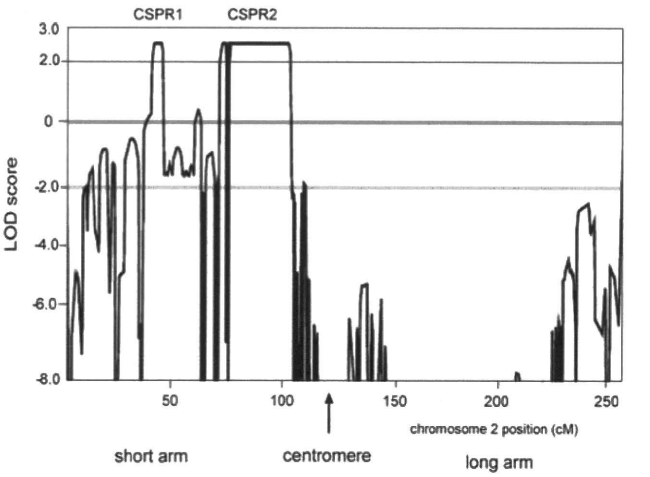

Figure 3 Multipoint LOD scores on chromosome 2. A 4.5-Mb (physical position, 18281893–22775527) interval from rs1545497 to rs1872325 corresponds to CSPR1, and a 34.5-Mb interval (45834656–80355227) from rs940053 to rs310777 corresponds to CSPR2.

In conclusion, this is the first report of a whole-genome linkage analysis scan for CSP. Although the LOD scores calculated are not high enough to assign the disease locus definitively, our data suggest that it lies at either 2p24.2–24.1 or 2p21–p12.

ACKNOWLEDGEMENTS

We are grateful to the members of the family for their participation in this research. We also thank Ms Miho Ooga and Ms Chisa Hayashida for their technical assistance. KY was supported partly by a Grant-in-Aid for Scientific Research from the Ministry of Health, Labour and Welfare, and partly by grants from the Takeda Scientific Foundation and the Naito Foundation.

1 FitzPatrick, D. R., Carr, I. M., McLaren, L., Leek, J. P., Wightman, P., Williamson, K. *et al.* Identification of SATB2 as the cleft palate gene on 2q32-q33. *Hum. Mol. Genet.* **12,** 2491–2501 (2003).
2 Braybrook, C., Doudney, K., Marçano, A. C., Arnason, A., Bjornsson, A., Patton, M. A. *et al.* The T-box transcription factor gene TBX22 is mutated in X-linked cleft palate and ankyloglossia. *Nat. Genet.* **29,** 179–183 (2001).
3 Koillinen, H., Lahermo, P., Rautio, J., Hukki, J., Peyrard-Janvid, M. & Kere, J. A. Genome-wide scan of non-syndromic cleft palate only (CPO) in Finnish multiplex families. *J. Med. Genet.* **42,** 177–184 (2005).
4 Shiang, R., Lidral, A. C., Ardinger, H. H., Buetow, K. H., Romitti, P. A. & Munger, R. G. Association of transforming growth-factor alpha gene polymorphisms with nonsyndromic cleft palate only (CPO). *Am. J. Hum. Genet.* **53,** 836–843 (1993).
5 Hwang, S. J., Beaty, T. H., Panny, S. R., Street, N. A., Joseph, J. M., Gordon, S. *et al.* Association study of transforming growth factor alpha (TGF alpha) TaqI polymorphism and oral clefts: indication of gene-environment interaction in a population-based sample of infants with birth defects. *Am. J. Epidemiol.* **141,** 629–636 (1995).
6 Carinci, F., Scapoli, L., Palmieri, A., Zollino, I. & Pezzetti, F. Human genetic factors in nonsyndromic cleft lip and palate: An update. *Int. J. Pediatr. Otorhinolaryngol.* **71,** 1509–1519 (2007).
7 Millard, D. R. in *Cleft Craft, the Evolution of its Surgery, Vol 1 The Unilateral Deformity* (eds Millard, D.R.) 41–55 (Little Brown, Boston, 1976).
8 Calnan, J. Submucous cleft palate. *Br. J. Plast. Surg.* **6,** 264–282 (1954).
9 Christensen, K. & Fogh-Anderson, P. Etiological subgroups in non-syndromic isolated cleft palate. A genetic-epidemiological study of 52 Danish birth cohorts. *Clin. Genet.* **46,** 329–335 (1994).
10 Abecasis, G. R., Cherny, S. S., Cookson, W. O. & Cardon, L. R. Merlin–rapid analysis of dense genetic maps using sparse gene flow trees. *Nat. Genet.* **30,** 97–101 (2002).
11 O'Connell, J. R. & Weeks, D. E. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am. J. Hum. Genet.* **63,** 259–266 (1998).
12 Nakashima, M., Nakano, M., Hirano, A., Kishino, T., Kondoh, S., Miwa, N. *et al.* Genome-wide linkage analysis and mutation analysis of hereditary congenital blepharoptosis in a Japanese family. *J. Hum. Genet.* **53,** 34–41 (2008).
13 Lathrop, G. M., Lalouel, J. M., Julier, C. & Ott, J. Strategies for multilocus linkage analysis in humans. *Proc. Natl Acad. Sci. USA* **81,** 3443–3446 (1984).

Journal of Human Genetics (2010) 55, 142–146

www.nature.com/jhg

ORIGINAL ARTICLE

Novel mutations in the *SIL1* gene in a Japanese pedigree with the Marinesco–Sjögren syndrome

Taichi Takahata[1,7], Koki Yamada[1,2,7], Yoshihisa Yamada[1], Shinji Ono[3,4], Akira Kinoshita[4], Tetsuo Matsuzaka[5,6], Koh-ichiro Yoshiura[4] and Takashi Kitaoka[1]

Marinesco–Sjögren syndrome (MSS) is a rare autosomal recessive disorder. Mutation in the *SIL1* gene accounts for the majority of MSS cases. However, some individuals with typical MSS without *SIL1* mutations have been reported. In this study, we identified two novel mutations in a Japanese pedigree with MSS, one of which was an intragenic deletion not detected using the PCR-direct sequencing protocol. This family consisted of three affected siblings, an unaffected sibling and unaffected parents. We found a homozygous 5-bp deletion, del598–602(GAAGA), in exon 6 of all affected siblings by PCR. Thus, we expected that both parents would be heterozygous for the mutation. As expected, the father was heterozygous, whereas the mother demonstrated no mutations. We then carried out array comparative genomic hybridization and quantitative PCR analyses, and identified an approximately 58 kb deletion in exon 6 in the patients and mother. As a result, the mother was hemizygous for a 58-kb deletion. The affected siblings contained two mutations, a 5-bp and a 58-kb deletion, resulting in *SIL1* gene dysfunction. It is possible that some reported cases of MSS without base alterations in the *SIL1* gene are caused by deletions rather than locus heterogeneity.
Journal of Human Genetics (2010) **55,** 142–146; doi:10.1038/jhg.2009.141; published online 29 January 2010

Keywords: array CGH; Marinesco–Sjögren syndrome; quantitative PCR; *SIL1* gene

INTRODUCTION

The Marinesco–Sjögren syndrome (MSS, OMIM 248800) is a rare, autosomal recessive disorder characterized by congenital cataracts, cerebellar ataxia, myopathy and mental retardation. Skeletal abnormalities including short stature, dysarthria, nystagmus and hypergonadotropic hypogonadism are also occasionally observed.

MSS was first described by Marinesco *et al.* in four Rumanian siblings in 1931.[1,2] Sjögren later reported 14 similar cases in six Swedish families and suggested an underlying autosomal recessive pattern of inheritance in 1950.[1,3] In 2003, Lagier-Tourenne *et al.*[4] identified a locus for MSS on chromosome 5q31 using homozygosity mapping in two consanguineous families of Turkish and Norwegian origin. Two years later, two groups independently identified several mutations in the *SIL1* gene located at chromosome 5q31 in MSS. In addition, Senderek *et al.*[5] identified nine different mutations in eight MSS families, and Anttonen *et al.*[6] found four different mutations in eight MSS families. Further novel mutations in the *SIL1* gene in MSS were subsequently identified.[7–9] Although mutations in the *SIL1* gene account for the majority of MSS cases, Senderrek *et al.*[5] reported four individuals with typical MSS lacking *SIL1* mutations. These reports suggest genetic heterogeneity in MSS. Here, we report novel mutations, including a deletion that is difficult to detect using conventional PCR for sequence analysis, in the *SIL1* gene in a Japanese family that included three individuals with MSS.

MATERIALS AND METHODS

Family

A Japanese family with MSS was investigated in this study. The clinical features of the affected individuals are summarized in Table 1. Each individual was diagnosed on the basis of the clinical features of MSS. The family consisted of three affected siblings, an unaffected sibling and their parents (Figure 1a). Parents were not consanguineous and all affected individuals were born after normal pregnancies.

The ophthalmological clinical features were as follows: The proband (II-1 in Figure 1a), an affected 14-year-old daughter, demonstrated slight bilateral cataract at 3 years of age, and received an operation for cataract at 4 years of age. Her best-corrected visual acuity after the operation was 1.0. An affected 10-year-old son (II-3) demonstrated slight bilateral posterior subcapsular cataracts at age 1 year and 6 months, and underwent an operation for cataract at 3 years of age. His best-corrected visual acuity after the operation was 1.2. An affected 8-year-old daughter (II-4) demonstrated bilateral total cataract at 4

[1]Department of Ophthalmology and Visual Sciences, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan; [2]Department of Ophthalmology, Sasebo City General Hospital, Sasebo, Japan; [3]Department of Psychiatry, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan; [4]Department of Human Genetics, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan; [5]Department of Pediatrics, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan and [6]Department of Pediatrics, Nagasaki Children's Medical and Welfare Center, Isahaya, Japan
[7]These authors contributed equally to this work.
Correspondence: Dr T Takahata, Department of Ophthalmology and Visual Sciences, Nagasaki University, 1-7-1 Sakamoto, Nagasaki 852-8501, Japan.
E-mail: takahata@nagasaki-u.ac.jp
Received 3 August 2009; revised 15 December 2009; accepted 16 December 2009; published online 29 January 2010

Table 1 Clinical features of the affected individuals

	II-1	*II-3*	*II-4*
Sex	F	M	F
Age	14	10	8
Bilateral cataract	+	+	+
Nystagmus	−	−	−
Strabismus	+	+	+
Skeletal deformities	+	+	+
Ataxia	+	+	+
Hypotonia	+	+	+
Spasticity	+	+	+
Mental retardation	+	+	+
Elevated serum CK	+	+	+
Myopathic EMG	+	NA	NA
Myopathic biopsy	+	NA	NA

Abbreviations: −, absent; +, present; CK, creatine kinase; EMG, electromyography; F, female; M, male; NA, not available.

years of age, and received an operation for cataract at the same age. Her best-corrected visual acuity after the operation was 0.6. Further ophthalmological examination of the three affected children revealed no abnormalities.

All samples from the family were collected after obtaining written informed consent, and the study protocol was preapproved by the Committee for the Ethical Issues on Human Genome and Gene Analysis in Nagasaki University. Genomic DNA was extracted directly from blood using the QIAamp DNA Blood mini kit (Qiagen, Hilden, Germany).

Mutation analysis

To identify mutations in the *SIL1* gene, PCR products were subjected to the direct sequencing protocol. Information regarding primer sequences was kindly provided by Dr Senderek (Department of Human Genetics, Aachen University of Technology) and Dr Anttonen (Folkhälsan Institute of Genetics and Neuroscience Center and Department of Medical Genetics, University of Helsinki). PCR was performed in a 30-μl reaction mixture containing 30 ng genomic DNA, 0.5 μM each of forward and reverse primers, 200 μM each of dNTP in 1× ExTaq buffer (Takara Bio, Shiga, Japan) and 0.75 U ExTaq (Takara Bio). PCR was performed in an iCycler thermal cycler (Bio-Rad, Hercules, CA, USA) and the PCR conditions were as follows: Taq activation step at 95 °C for 4 min, followed by 35 cycles at 95 °C for 30 s for denaturation, 58 °C for 30 s for annealing, 72 °C for 30 s for extension and finally one step at 72 °C for 10 min to ensure complete extension.

The PCR products were treated with ExoSAP-IT (USB, Cleveland, OH, USA) and directly sequenced using the BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems, Foster City, CA, USA). Samples were run on an ABI 3130-xl automated sequencer (Applied Biosystems) and electropherograms were aligned using ATGC software version 5.0 (Genetyx, Tokyo, Japan). Mutations were inspected visually.

Microsatellite analysis

Microsatellite analyses were carried out using the ABI PRISM linkage Mapping Set-MD10 (panel 8) and included eight markers on chromosome 5. PCR was performed in a 15-μl reaction mixture under the same conditions for the mutation analysis, with the exception that 55 °C for 30 s was used for annealing.

After mixing with GeneScan 400HD ROX size standard (Applied Biosystems) in deionized formamide, amplicons were separated on the ABI 3130-xl automated sequencer. Genotyping data were analyzed using GeneMapper 4.0 software (Applied Biosystems).

Array CGH

A DNA sample from patient II-3 (Figure 1a) was subjected to the high-density oligonucleotide-based array comparative genomic hybridization (CGH) assay. For this assay, we manufactured a custom-designed microarray targeted to a 300-kb genome region, including the *SIL1* gene, on 5q31.2 (Chr5:138 281 500–138 580 000 [NCBI Build 36.1, hg18]). We used the Agilent website (http://earray.chem.agilent.com/earray/) to design our custom array CGH. This array contained 2685 probes that were 60-mer in length (Agilent Technologies, Santa Clara, CA, USA). Experiments were performed according to the manufacturer's instructions. Briefly, patient and reference genomic DNA samples (1 μg per sample) were fluorescently labeled with Cy3 (patient) and Cy5 (reference) using the Agilent Genomic DNA Labeling Kit (Agilent Technologies). Labeled patient and reference DNA was then combined, denatured and preannealed with Cot-1 DNA (Invitrogen, Carlsbad, USA, USA) and blocking reagent (Agilent Technologies). The labeled samples were then hybridized to the arrays for 40 h in a rotating oven (Agilent Technologies) at 65 °C and 20 r.p.m. After hybridization and washing, the arrays were scanned at a 5-μm resolution with an Agilent G2565C scanner. The resulting images were analyzed using Feature Extraction Software 10.5.1.1 (Agilent Technologies).

Quantitative PCR analysis

Real-time PCR was performed using a LightCycler 480 Instrument (Roche Applied Science, Penzberg, Germany) and SYTO13 dye (Molecular Probes, Eugene, OR, USA). Exons 2, 6 and 10 of the *SIL1* gene were selected as target exons for quantification. The *NSD1* gene was used as a reference gene (two copies in the reference DNA). Primers for the *SIL1* and *NSD1* genes were designed using Primer Express 1.5 (Applied Biosystems) and are listed in Table 2. Real-time PCR was performed in a 10-μl reaction mixture containing 10 ng genomic DNA, 0.5 μM each of forward and reverse primers, 200 μM each of dNTP, 1× ExTaq buffer, 0.2 μM SYTO13 dye and 0.5 U ExTaq.

Break-point determination

To determine the break point of deletion, we designed a deletion-specific amplification primer around the deletion break point detected by array CGH. The primer sequences were 5′-AGCGGATCAGTAAGGGTATT-3′ for SILint5-delF and 5′-CAGTGTCTGGAAGCACAAGC-3′ for SILint7delR. DNA from all family members and from 80 healthy individuals was subjected to PCR amplification using the same conditions as the mutation analysis, with the exception that 61 °C was used as the annealing temperature.

RESULTS

Mutation analysis

We sequenced all 10 exons of the *SIL1* gene in our MSS family members. Portions of the electropherograms are presented in Figure 1b. The electropherograms for the three affected siblings demonstrated that they were homozygous for a 5-bp deletion mutation, del598-602(GAAGA), in exon 6. No mutations were identified in the unaffected sibling. A heterozygous del598-602(GAAGA) mutation in exon 6 was also detected in the father. The mother's electropherogram did not show del598-602(GAAGA). del598-602(GAAGA) was not detected in any of the 80 healthy Japanese individuals.

Microsatellite analysis

All the eight microsatellite markers investigated showed the concordant inheritance pattern of the allele from both parents. This means that the parent–child relationship was confirmed, and that long-range uniparental disomy could be excluded. The inheritance pattern of the allele on chromosome 5 is summarized in Figure 1a.

Copy number analysis in the family members

We speculated that the patients and mother had a deletion in exon 6 of the *SIL1* gene on the basis of microsatellite and mutation analyses. We were able to identify the deletion within the *SIL1* genomic region using real-time PCR and array CGH (Figures 2 and 3a). The copy number state in the patients determined by real-time PCR was concordant with the results of array CGH (Figure 2).

Deletion break point

Using SILint5delF/SILint7delR primers we were able to amplify PCR products of the patients and mother, but not of the healthy

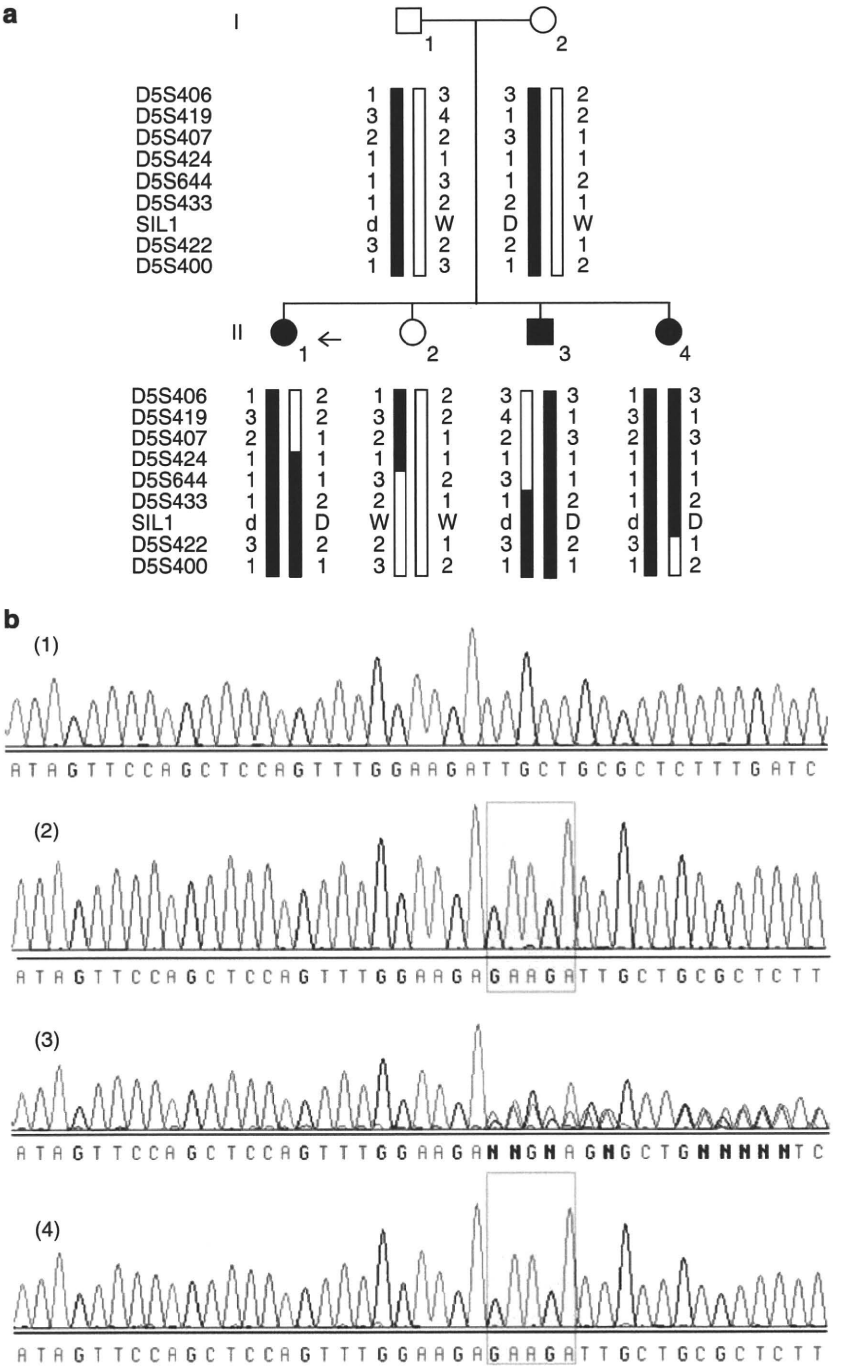

Figure 1 Microsatellite and mutation analysis. (**a**) Pedigree of the MSS family studied and the marker haplotypes. Closed symbols indicate individuals with MSS and open symbols represent clinically unaffected individuals. All investigated allele sizes of the four siblings corresponded to the allele sizes of either parent. The locus for the *SIL1* gene is between D5S433 and D45422, d: 5 bp deletion in exon 6, D: deletion including exon 6, W: wild type. Arrow indicates proband. (**b**) DNA sequence data in our MSS family. (1) Affected sibling (II-4 in panel a) with the homozygous del598-602(GAAGA) mutation, (2) unaffected sibling (II-2) without the mutation, (3) father (I-1) with the heterozygous del598-602(GAAGA) mutation and (4) mother without a mutation.

individuals. The deletion-specific product was subsequently processed for sequence analysis, confirming the 58 269-bp deletion [ch5:g.(138 339 032–138 397 300)del](NCBI Build 36.1, hg18) and the 4-bp insertion (Figure 3b). The telomeric break point within intron 5 was in the LINE/L1 repetitive sequence, whereas the centromeric break point within intron 7 was a unique sequence.

DISCUSSION

In this study, we identified novel mutations in the *SIL1* gene in a Japanese family that included three children with MSS. We sequenced all 10 exons of the *SIL1* gene, and identified a del598-602(GAAGA) mutation in exon 6 of the PCR products amplified from genomic DNA isolated from all three of the affected siblings. Thus, we expected

that the affected siblings would be homozygous for the mutation obtained from a parent, and that both parents would be heterozygous for the mutation. However, we found that only the father expressed the del598-602(GAAGA) mutation, whereas no mutations were identified in any of the 10 exons of the *SIL1* gene in the mother.

We next confirmed the parent–child relationship for each sibling using microsatellite markers on chromosome 5. The mutation and microsatellite analyses suggested that the mother may be hemizygous around exon 6. Quantitative PCR analyses in all family members indicated that the unaffected sibling and father expressed two copies of exon 6 in the *SLI1* gene, whereas the three affected siblings and mother expressed only one copy of exon 6. Therefore, we attempted to define the copy number state for the entire *SIL1* gene using array CGH to confirm the break point of deletion. As it is possible to speculate break points from the array CGH results, we were able to design primers to amplify the deletion-specific product using PCR. Using this method, we found a 58 269-bp deletion in the three affected siblings and mother. The character of break points was not specific, and did not indicate the recombination between the repetitive sequence or low copy repeats.

Table 2 Primer sequences for real-time PCR

Target sequence (*SIL1* exon 2)	
Forward primer	5′-CTCTTGTTGGATGGCTGGAC-3′
Reverse primer	5′-TGTGATTCCCATGTCGTCAC-3′
Target sequence (*SIL1* exon 6)	
Forward primer	5′-GGCAGATGTCTCCAACCAAT-3′
Reverse primer	5′-CTTGTTGATCAGCCGTACCA-3′
Target sequence (*SIL1* exon 10)	
Forward primer	5′-AGAGCTAGCCAGGTGTGAGC-3′
Reverse primer	5′-AGGAGGTGTACCTGGCGATA-3′
Reference sequence (*NSD1*)	
Forward primer	5′-ATGCTTTTTCAGCCCAAATG-3′
Reverse primer	5′-CTCCCTGCAGTACAGCATCA-3′

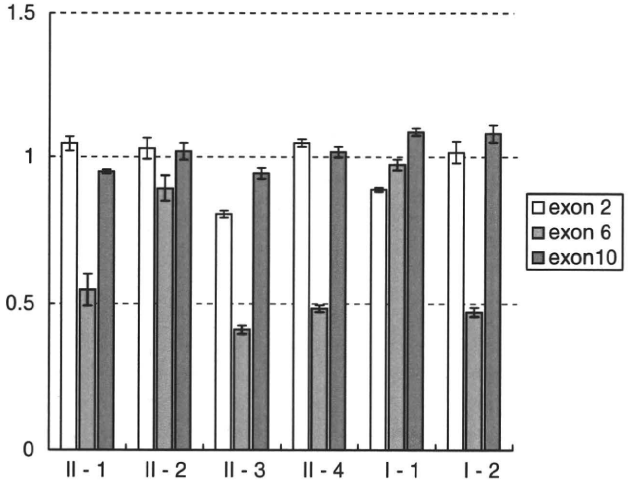

Figure 2 Copy number analysis. The *SIL1* to *NSD1* copy number ratio (*N*) for all family members. The sample with the deletion in *SIL1* is expected to yield *N*=0.5. The *N* values of exon 6 in affected siblings and the mother were 0.5472 (II-1), 0.414 (II-3), 0.483 (II-4) and 0.472 (I-1), respectively. The unaffected sibling was 0.897 (II-2) and the father (II-1) was 0.974. The fact that the *N* value of exons 2 and 10 for all family members was approximately 1.0 suggested that the deletion was not large enough to include the entire *SIL1* gene. The results are presented as mean ± s.d.

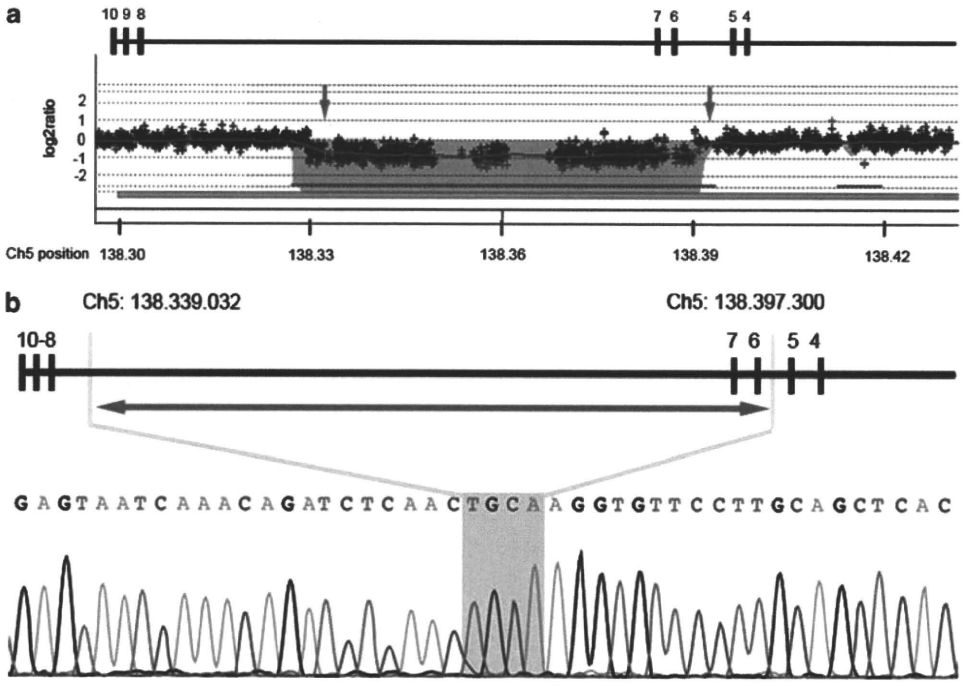

Figure 3 Break-point determination. (**a**) DNA Analytics view of the affected sibling (II-3) using the Agilent custom-designed array showing the approximately 58 kb deletion in the *SIL1* gene. Arrows indicate the break point. (**b**) The sequence results around the break point and the schematic drawing of the *SIL1* gene in the affected sibling (II-3) are shown. A 58 269 bp deletion at chr5: 138 339 032–138 397 300 (NCBI Build 36.1, hg18) and a 4b insertion (shaded region) were identified.

Table 3 Previously reported mutations in the *SIL1* gene of MSS patients

	Type	*Location*	*Nucleotide change*	*Amino-acid change*	*Origin*	P	*Ref.*
1	HM	Exon 3	212dupA	H71Qfs	France	1	6
2	HM	Exon 4	331C>T	R111X	Iran, Turkey, Italy	4	5,6,10
3	HM	Exon 6	506_509dupAAGA	D170fs	Finland, Norway	5	6
4	HM	Exon 6	645+1G>A	Skipping	Turkey	1	5
5	HM	Exon 9	936dupG	L313fs	Japan	2	8,9
6	HM	Exon 9	1029+1G>A	Skipping	Bosnia	1	5
7	HM	Exon 9	1030-9G>A	F345fs	Norway	2	9
8	HM	Exon 9	1249C>T	Q417X	Mali	1	5
9	HM	Exon 10	1312C>T	Q438X	Egypt	1	7
10	HM	Exon 10	1367T>A	L456X	Turkey	1	9
11	HM	Exon 10	1370T>C	L457Pro	Japan	1	9
12	CH	Exons 2, 4	178G>T 346delG	E60X G116fs	Vietnam	1	5
13	CH	Exon 6	506_509dupAAGA 645+2T>C	D170fs skipping	Sweden	1	6
14	CH	Exons 9, 10	947_948insT 1030-18G>A	L316fs M344fs	Germany	1	5
15	CH	Exons 9, 10	947_948insT 1366delT	L316fs 456fs	Russia	1	5

Abbreviations: CH, compound heterozygous; del, deletion; dup, duplication; fs, frameshift; HM, homozygous; ins, insertion; MSS, Marinesco–Sjögren syndrome; P, pedigree number; Ref., reference; X, stop.

MSS is a rare, autosomal recessive disorder. After the two initial groups independently identified several mutations in the *SIL1* gene in 2005,[5,6] only a few mutations in the *SIL1* gene have been reported since.[7–10] Karim *et al.*[7] located a novel mutation in an Egyptian family in 2006, and Eriguchi *et al.*[8] identified a novel mutation in three unrelated Japanese patients in 2008. All mutations in the *SIL1* gene reported previously to be associated with MSS are presented in Table 3. The mutation we found was located in exon 6, which encodes the BiP-interacting domain.[5] Zhao *et al.*[11] have reported that the SIL1 protein associates with the BiP chaperone to aid unfolded proteins in folding normally, and to help in the release of folded proteins. Thus, the loss of SIL1 protein function results in BiP recycling and the accumulation of unfolded proteins in the endoplasmic reticulum.[11–13]

Senderek *et al.*[5] were unable to identify any *SIL1* gene mutations in four individuals with typical MSS. These reports suggested genetic heterogeneity in MSS or that individuals exhibiting MSS may contain mutations that are difficult to detect. For example, compound heterozygous deletions that include different exons or intronic base changes affect the splicing process. In general, when gene mutations in a single gene defect syndrome are detected, it is essential to consider that deletion may not be detected using the PCR-direct sequencing protocol. Our results suggested that deletion assay, quantitative PCR, array CGH or multiple ligation-mediated PCR amplification should be performed to detect deletions of exons in MSS patients. It remains possible that some reported cases without base alterations in the *SIL1* gene are caused by small deletions rather than locus heterogeneity.

ACKNOWLEDGEMENTS

K Yamada was supported partly by a Grants-in-Aid for Scientific Research Category, no. 18791284 from the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT). K Yoshiura was supported partly by a Grant-in-Aid for Scientific Research from the Ministry of Health, Labour and Welfare, and partly by grants from the Takeda Scientific Foundation and the Naito Foundation. We are greatly indebted to all the participants of this research. We also thank Ms M Ooga and C Hayashida for their excellent technical assistance.

1 Andersen, B. Marinesco-Sjogren syndrome: spinocerebellar ataxia, congenital cataract, somatic and mental retardation. *Dev Med Child Neurol.* **47,** 249–257 (1965).
2 Marinesco, G., Draganesco, S. & Vasiliu, D. Novelle maladie familiale caracterisee par Une cataracte congenitale et un arret du development somato-neuro-phychique. *Lencephale.* **26,** 97–109 (1931).
3 Sjögren, T. Hereditary congenital spinocerebellar ataxia accompanied by congenital cataracts and oligophrenia. *Confin. Neurol.* **10,** 293–308 (1950).
4 Lagier-Tourenne, C., Tranebjaerg, L., Chaigne, D., Gribaa, M., Dollfus, H., Silvestri, G. *et al.* Homozygosity mapping of Marinesco-Sjogren syndrome to 5q31. *Eur. J. Hum. Genet.* **11,** 770–778 (2003).
5 Senderek, J., Krieger, M., Stendel, C., North, K., Muntoni, F., Quijano-Roy, S. *et al.* Mutations in SIL1 cause Marinesco-Sjogren syndrome, a cerebellar ataxia with cataract and myopathy. *Nat. Genet.* **37,** 1312–1314 (2005).
6 Anttonen, A. K., Mahjneh, I., Hämäläinen, R. H., Lagier-Tourenne, C., Kopra, O., Waris, L. *et al.* The gene disrupted in Marinesco-Sjogren syndrome encodes SIL1, an HSPA5 cochaperone. *Nat. Genet.* **37,** 1309–1311 (2005).
7 Karim, M. A., Parsian, A J., Cleves, M. A., Bracey, J., Elsayed, M. S., Elsobky, E. *et al.* A novel mutation in BAP/SIL1 gene causes Marinesco-Sjögren syndrome in an extended pedigree. *Clin. Genet.* **70,** 420–423 (2006).
8 Eriguchi, M., Mizuta, H., Kurohara, K., Fujitake, J. & Kuroda, Y. Identification of a new homozygous frameshift insertion mutation in the SIL1gene in 3 Japanese patients with Marinesco-Sjögren syndrome. *J. Neurol. Sci.* **270,** 197–200 (2008).
9 Anttonen, A. K., Siintola, E., Tranebjaerg, L., Iwata, N. K., Bijlsma, E. K., Meguro, H. *et al.* Novel SIL1 mutations and exclusion offunctional candidate genes in Marinesco-Sjögren syndrome. *Eur. J. Hum. Genet.* **16,** 961–969 (2008).
10 Annesi, G., Aguglia, U., Tarantino, P., Annesi, F., De Marco, E. V., Civitelli, D. *et al.* SIL1 and SARA2 mutations in Marinesco-Sjögren and chylomicron retention diseases. *Clin. Genet.* **71,** 288–289 (2007).
11 Zhao, L., Longo-Guess, C., Harris, B. S., Lee, J. W. & Ackerman, S.L. Protein accumulation and neurodegeneration in the woozy mutant mouse is caused by disruption of SIL1, a cochaperone of BiP. *Nat Genet.* **37,** 974–979 (2005).
12 Weitzmann, A., Volkmer, J. & Zimmermann, J. The nucleotide exchange factor activity of Grp170 may explain the non-lethal phenotype of loss of Sil1 function in man and mouse. *FEBS Lett.* **580,** 5237–5240 (2006).
13 Weitzmann, A., Baldes, C., Dudek, J. & Zimmermann, R. The heat shock protein 70 molecular chaperone network in the pancreatic endoplasmic reticulum. *FEBS J.* **274,** 5175–5187 (2007).

Human Reproduction, Vol.25, No.4 pp. 1076–1080, 2010
Advanced Access publication on February 6, 2010 doi:10.1093/humrep/deq006

human reproduction

CASE REPORT *Reproductive genetics*

A case of Kallmann syndrome carrying a missense mutation in alternatively spliced exon 8A encoding the immunoglobulin-like domain IIIb of fibroblast growth factor receptor 1

Kiyonori Miura[1,*], Shoko Miura[1], Koh-ichiro Yoshiura[2], Stephanie Seminara[3], Daisuke Hamaguchi[1], Norio Niikawa[4], and Hideaki Masuzaki[1]

[1]Department of Obstetrics and Gynecology, Nagasaki University Graduate School of Biomedical Sciences, 1-7-1 Sakamoto, Nagasaki, Japan [2]Department of Human Genetics, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan [3]Reproductive Endocrine Unit, Massachusetts General Hospital, Boston, MA 02114, USA [4]Research Institute of Personalized Health Sciences, Health Sciences University of Hokkaido, Hokkaido, Japan

[*]Correspondence address. Tel: +81-95-819-7363; Fax: +81-95-819-7365; E-mail: kiyonori@nagasaki-u.ac.jp

Submitted on August 24, 2009; resubmitted on December 17, 2009; accepted on January 4 2010

ABSTRACT: Fibroblast growth factor receptor 1 (*FGFR1*) is one of the causative genes for Kallmann syndrome (KS), which is characterized by isolated hypogonadotropic hypogonadism with anosmia/hyposmia. The third immunoglobulin-like domain (D3) of *FGFR1* has the isoforms *FGFR1-IIIb* and *FGFR1-IIIc*, which are generated by alternative splicing of exons 8A and 8B, respectively. To date, the only mutations to have been identified in D3 of *FGFR1* are in exon 8B. We performed mutation analysis of *FGFR1* in a 23-year-old female patient with KS and found a missense mutation (c.1072C>T) in exon 8A of *FGFR1*. The c.1072C>T mutation was not detected in her family members or in 220 normal Japanese and 100 Caucasian female controls. No mutation in other KS genes, KS 1, prokineticin-2, prokineticin receptor-2 and *FGF-8* was detected in the affected patient or in her family members. Therefore, this is the first case of KS carrying a *de novo* missense mutation in *FGFR1* exon 8A, suggesting that isoform *FGFR1-IIIb*, as well as isoform *FGFR1-IIIc*, plays a crucial role in the pathogenesis of KS.

Key words: Kallmann syndrome / *FGFR1b mutation* / fibroblast growth factor receptor 1 isoform expression

Introduction

Kallmann syndrome (KS), which is characterized by isolated hypogonadotropic hypogonadism (IHH) and anosmia/hyposmia, is a clinically and genetically heterogeneous disorder. To date, five causative genes for KS have been reported: KS 1 (*KAL1*, GenBank accession M97 252), prokineticin-2 (*PROK2*, GenBank accession NM 021935), prokineticin receptor-2 (*PROKR2*, GenBank accession NM 144773), fibroblast growth factor-8 (*FGF-8*, GenBank accession NM 033163) and fibroblast growth factor receptor 1 (*FGFR1*, GenBank accession NM 023110.2).

Although sporadic cases of KS are more frequent, families with KS have been reported with X-linked recessive or autosomal dominant or recessive modes of inheritance. Mutations in *KAL1* have been found in familial cases with X-linked recessive inheritance (Franco *et al.*, 1991; Legouis *et al.*, 1991). Mutations in *PROK2* were detected in the heterozygous state, whereas *PROKR2* mutations were found in the heterozygous, homozygous or compound heterozygous state (Dodé *et al.*, 2006). *PROKR2/PROK2* mutations with true pathogenic potential were found only in the homozygous state (Abreu *et al.*, 2008), and any dominant-negative effect of *PROKR2* mutations was ruled out (Monnier *et al.*, 2009). Mutations in *FGFR1* or *FGF8* underlie an autosomal dominant form with incomplete penetrance. Therefore, KS families harbouring heterozygous *FGFR1* or *FGF8* mutations display variable olfactory phenotypes (Dodé *et al.*, 2003; Falardeau *et al.*, 2008), and a few cases with heterozygous *FGFR1* mutations show a normosmic IHH (Pitteloud *et al.*, 2006a). The *FGFR1* gene, which is located on chromosome 8p12, comprises 18 exons (Ruta *et al.*, 1989), and various mutations, including missense and protein truncation mutations, have been reported (Trarbach *et al.*, 2007). The third immunoglobulin-like domain (D3) of *FGFR1* has the isoforms *FGFR1-IIIb* and *FGFR1-IIIc*, which are generated by alternative splicing

of exons 8A and 8B, respectively (Johnson *et al.*, 1991). To date, mutations in D3 of *FGFR1* have only been identified in exon 8B, which encodes immunoglobulin domain IIIc, suggesting that isoform FGFR1-IIIc plays a crucial role in the pathogenesis of KS (Pitteloud *et al.*, 2006b; Trarbach *et al.*, 2006; Dodé *et al.*, 2007).

Here, we report for the first time a KS case carrying a *de novo* missense mutation in the alternatively spliced exon 8A of *FGFR1-IIIb*.

Materials and Methods

Patient and family

Patient (Subject II-2) was a 23-year-old Japanese woman. When she was 18 years old, she was treated at Nagasaki University Hospital because of primary amenorrhea with anosmia. Her height was 159.2 cm and her weight was 72.0 kg. Her serum levels of luteinizing hormone (LH), follicle-stimulating hormone (FSH) and estradiol (E2) were less than 0.5 and 1.5m IU/ml and 10 pg/ml, respectively. Her LH frequent sampling study (sampling performed every 15 min) showed a low-amplitude pattern of LH pulsation (Fig. 1). Her brain magnetic resonance imaging (MRI) examination was negative for tumors and showed no anatomical abnormalities of the hypothalamic–pituitary region and olfactory bulbs. A scratch-and-sniff test (UPSIT, Sensonics, Haddon Hts, NJ, USA) (Doty *et al.*, 1985), which determines ability to smell, indicated anosmia. She was diagnosed as having KS and received hormone replacement therapy for 5 years. Her mother (Subject I-1) was normosmic and had normal puberty and regular menstrual cycles. Her father (Subject I-2), elder brother (Subject II-1) and younger brother (Subject II-3) were also normosmic and had normal puberty (Fig. 2).

Molecular analysis

DNA extraction

Whole blood samples were obtained from the KS patient and from her mother, father, elder and younger brothers. All samples were collected after obtaining written informed consent and the study protocol was approved by the Institutional Review Board of Nagasaki University. Genomic DNA from lymphocytes was extracted using a QIAamp DNA blood mini kit (Qiagen, Düsseldorf, Germany), according to the manufacturer's instructions.

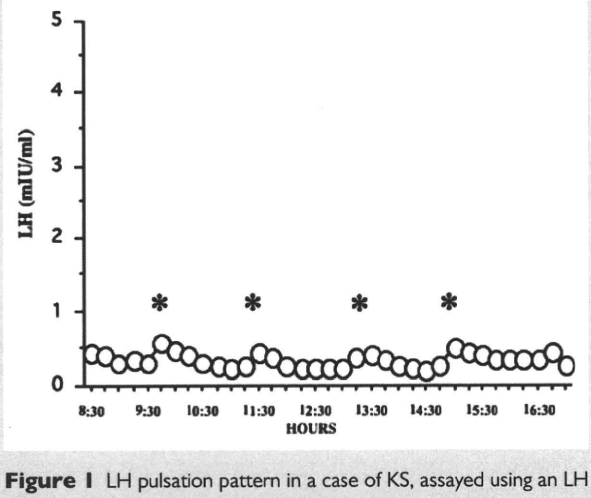

Figure 1 LH pulsation pattern in a case of KS, assayed using an LH frequent sampling study. LH frequent sampling was performed every 15 min. *Low-amplitude pattern of LH pulse.

Sequence analysis

FGFR1 consists of 18 coding exons. Intragenic mutations were investigated by PCR amplification and sequence analysis using 14 pairs of primers, as previously described (Dodé *et al.*, 2003; Sato *et al.*, 2004). Genomic DNA was PCR amplified using conditions of 95°C for 12 min followed by 95°C for 30 s, 59°C for 30 s and 72°C for 60 s for 35 cycles and a final cycle of 72°C for 10 min. PCR products were analyzed by agarose gel electrophoresis, purified with ExoSAP-IT and subjected to sequencing reactions. Sequencing reactions were performed using the BigDye terminator v.3.1 kit and analyzed with an ABI PRISM 3100 Genetic Analyzer™ (Applied Biosystems). The KS patient carrying a mutation in *FGFR1* and her family members were also screened for mutations in the other genes known to be involved in KS [*KAL1*, *PROK2*, *PROKR2* and *FGF8*]. Whether the mutation leads to a change in the protein structure and function was predicted bioinformatically using the ExPASy proteomics server (http://au.expasy.org/) and PolyPhen (http://genetics.bwh.harvard.edu/pph/).

Confirmation of the alternatively spliced exon

Isolation of a full-length murine *Fgfr1-IIIb* showed that Fgfr1-IIIb was a transmembrane receptor (Beer *et al.*, 2000). Although the mRNA encoding exon IIIb has been found in human (Johnson *et al.*, 1991), the presence of sequences encoding the intracellular domain has not yet been demonstrated. Therefore, to determine the splice site of exon 8A and to detect *FGFR1-IIIb* mRNA encoding the intracellular domain, we performed RT–PCR using specific primers to amplify the splice isoform containing exon 8A. Kal23 is designed to span exons 7 and 8A for specific annealing to the *FGFR1-IIIb* isoform, which is spliced from exon 7 to exon 8A (Fig. 3A). Kal5 is designed within exon 8B for specific annealing to the *FGFR1-IIIc* isoform, which is spliced from exon 7 to exon 8B. Kal2 and Kal6 are designed within exons 7 and 9, respectively, for annealing to the D3 isoforms of *FGFR1*. Primer sequences were as follows: kal2: 5′-GACAGAAGGTCGGTTATGTC-3′, Kal23: 5′-CAGATCTTGAAGC ATTCGGG-3′, Kal5: 5′-GGTGGTATTAACTCCAGCAG-3′ and Kal6: 5′-GTACAGGGGCGAGGTCATCA-3′. The BD multiple tissue complementary DNA (cDNA, MTC) panels Human I and Human II (BD Biosciences Clontech, Mountain View, CA, USA) were used to detect the expression of each isoform of *FGFR1*. PCR amplification was performed on cDNAs as follows: 94°C, 30 s; 62°C, 30 s; 72°C, 1 min; 40 cycles. PCR products were analyzed by agarose gel electrophoresis and sequenced using then ABI PRISM 3100 Genetic Analyzer™.

Results

Sequence analysis of the entire coding region of *FGFR1*, including exon–intron boundary regions, showed that the KS patient had a mutation (c.1072C>T) in exon 8A of *FGFR1-IIIb*, while the other family members did not (Fig. 2). However, the full-length *FGFR1* mRNA that includes exon 8A is not deposited in the full-length cDNA database (GenBank accession no. NM 023110.2). RT–PCR analysis indicated that most transcripts containing exon 8A were spliced to exon 8B in all adult tissues except bone marrow (data not shown). We wished to demonstrate the existence of an alternative transcript, exon 8A which was spliced to exon 9 encoding the transmembrane helix; therefore, RT–PCR products from human fetal brain were cloned and sequenced. In 1 of 27 clones exon 8A was spliced to exon 9 (designated here '*FGFR1-IIIb*', GenBank accession FJ809917, see Fig. 3B), while in the other clones exon 8A was spliced to exon 8B (designated here

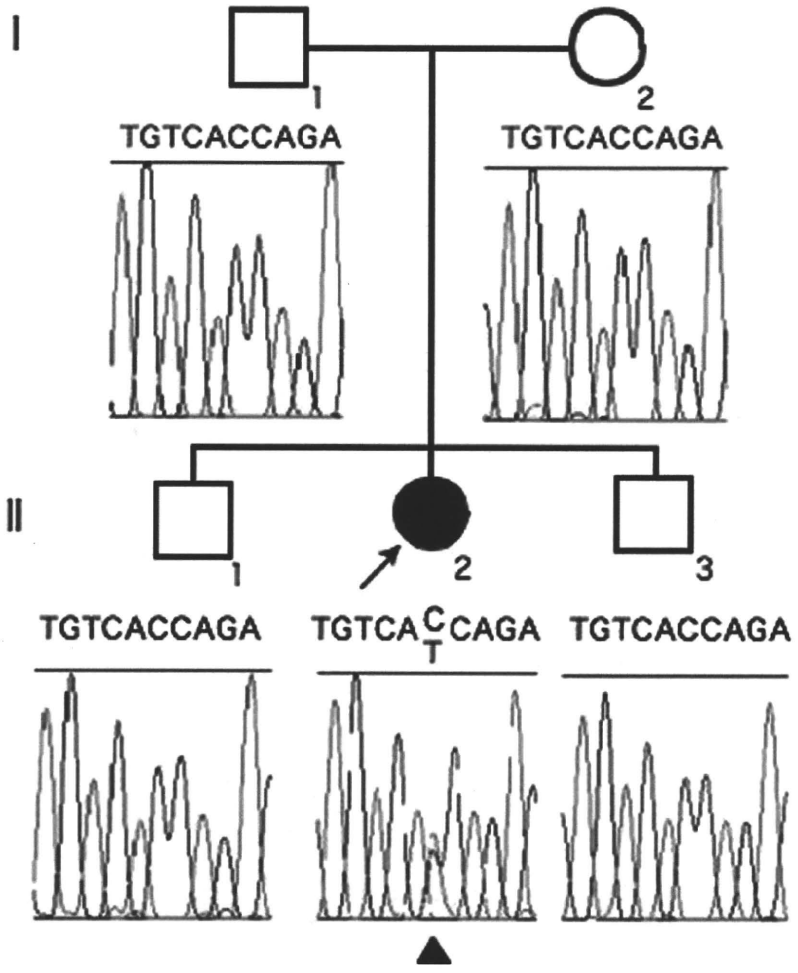

Figure 2 Pedigree of patient's family and the results of fibroblast growth factor receptor 1 (*FGFR1*) sequence analysis. II-2 is a 23-year-old woman with KS. The patient had a mutation in *FGFR1* (C>T) but the other family members did not. The mutation was *de novo* because parentage was assured. Arrowhead under the electropherogram indicates the mutation site.

'*FGFR1-secr*', GenBank accession FJ809916, see Fig. 3B). The exact acceptor and donor sites of exon 8A in '*FGFR1-IIIb*' mRNA, which produces a membrane-bound FGFR1-containing D3, were determined by sequence analysis of splice isoforms, '*FGFR1-IIIb*' and '*FGFR1-secr*', (Fig. 3). As most full-length *FGFR1* cDNAs in the database were transcripts containing exon 7–exon 8B–exon 9 (designated here '*FGFR1-IIIc*', GenBank accession NM 023110.2, see Fig. 2B) without exon 8A, '*FGFR1-IIIc*' is likely to be the most abundantly expressed human isoform. The EST, CA488712.1, was the only isoform in the EST database corresponding to '*FGFR1-IIIb*'. Although both '*FGFR1-IIIb*' and '*FGFR1-IIIc*' encode membrane-bound *FGFR1*, '*FGFR1-secr*' encodes a secreted form of *FGFR1* because of a sequence frameshift and a termination codon in exon 9.

The exons 8A and 8B of the human FGFR1 isoforms shared the amino acid sequence at 354–357; WLTV. However, exon 8A ends with six extra amino acids at 358–363, TRPVAK, whereas exon 8B ends with only two, LE. These sequences are identical in the mouse Fgfr1 isoforms (Beer *et al.*, 2000). The mutation in exon 8A of *FGFR1-IIIb* (GenBank accession no. FJ809917 bankit1193625) is c.1072C>T at the cDNA level and p.T358I at the amino acid level. Bioinformatic analysis shows the mutated amino acid residue to be conserved between human and mouse and to be located in D3 of *FGFR1-IIIb*, which is a critical region for FGF ligand binding. However, the mutation was not predicted to produce a change in the human protein structure. The c.1072C>T mutation was not detected in 220 normal Japanese women or in 100 normal Caucasian women. The patient had no mutation in any of the other four KS genes.

Discussion

Mouse *Fgfr1-IIIb* has a low level of expression in a wide variety of adult tissues, but a high level of expression in skin and brain, indicating the existence of specific splicing factors in skin and brain that recognize the relatively weak *Fgfr1-IIIb* splice site (Beer *et al.*, 2000). Consistent with the expression pattern of mouse *Fgfr1-IIIb*, we could isolate human *FGFR1-IIIb* from a fetal brain cDNA library but not from adult

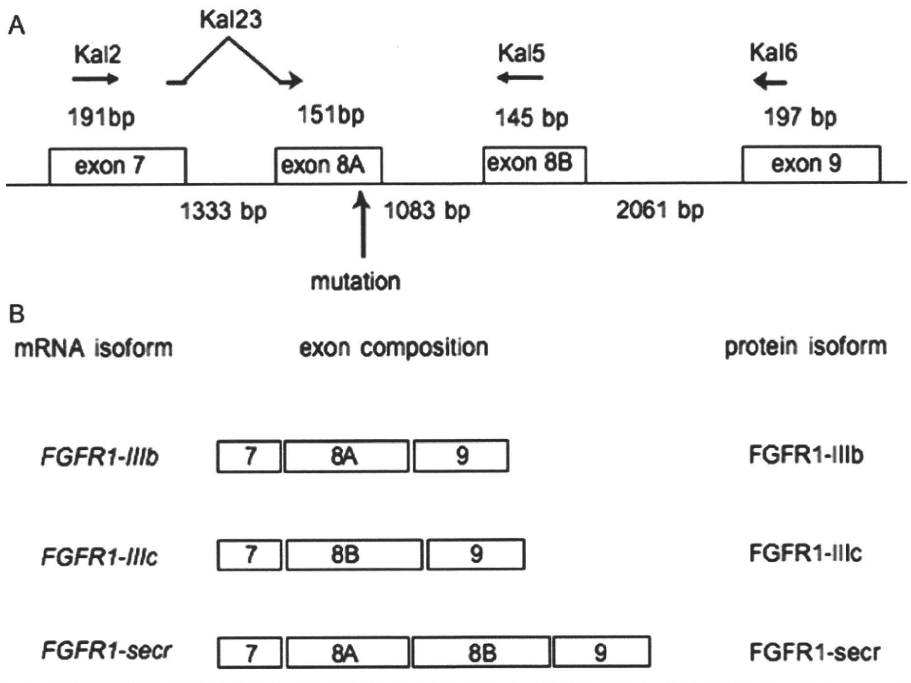

Figure 3 Genomic organization of *FGFR1* around exon 8A. (A) The numbers with bp indicate exon length (over the line) and intron length (under the line). The mutation is located near the end of exon 8A. Horizontal arrows indicate the locations of primers used to perform RT–PCR amplification of the isoform containing exons 7 and 8A. Primer Kal23 is designed to span exons 7 and 8A for specific annealing to the *FGFR1-IIIb* isoform, which is spliced from exon 7 to exon 8A (Fig. 3A). Kal5 is designed within exon 8B for specific annealing to the *FGFR1-IIIc* isoform, which is spliced from exon 7 to exon 8B. Kal2 and Kal6 are designed within exons 7 and 9, respectively, for annealing to the D3 isoforms of *FGFR1*. Vertical arrow indicates mutation site. (B) Composition of mRNA isoforms and of putative protein structures. FGFR1-IIIb: membrane-bound form of FGFR1 with immunoglobulin-like domain IIIb encoded by exon 8A, FGFR1-IIIc: membrane-bound form of FGFR1 with immunoglobulin-like domain IIIc encoded by exon 8B, FGFR1-secr: a secreted form of FGFR1.

tissues. Most full-length *FGFR1* cDNAs in the database represent *FGFR1-IIIc*. *FGFR1-IIIc* is expressed at high levels, but *FGFR1-IIIb* is expressed at very low levels (Johnson *et al.*, 1991). We can amplify only a tiny amount of *FGFR1-IIIb* that has exon 8A spliced to exon 9 by RT–PCR using the primers Kal23 and Kal6. Most of the sequenced RT–PCR products corresponded to *FGFR1-secr*, suggesting that the expression level of the three isoforms is '*FGFR1-IIIc*' ≫ '*FGFR1-IIIsecr*' ≫ '*FGFR1-IIIb*' (Fig. 3B).

Several studies suggested that mutations in exon 8B of isoform *FGFR1-IIIc* are implicated in the pathogenesis of KS (Pitteloud *et al.*, 2006b; Trarbach *et al.*, 2006; Dodé *et al.*, 2007). Mice homozygous for alleles with a stop codon in exon IIIc displayed phenotypes resembling those of embryos homozygous for null alleles, while mice carrying an in-frame stop codon in exon IIIb were viable and fertile (Partanen *et al.*, 1998). Therefore, *Fgfr1-IIIc* is the dominant isoform that carries out the majority of the biological functions of the *Fgfr1* gene, whereas *Fgfr1-IIIb* plays a minor and to some extent redundant role (Partanen *et al.*, 1998). A receptor-binding analysis revealed no difference in the binding specificity between the endogenous *Fgfr1-IIIb* and an artificially created *Fgfr1-IIIb*, which had two different amino acids in the 3′-end of the unique IIIb exon (Beer *et al.*, 2000), suggesting that the carboxyl terminus of D3 may not overtly influence binding specificity. However, being expressed at low levels does not imply that '*FGFR1-IIIb*' has an unimportant role. The mutation we found, p.T358I, was located in exon 8A of *FGFR1-IIIb*. Therefore, p.T358I may affect ligand binding and cause the KS phenotype, although how this mutation affects the loss of function of *FGFR1-IIIb* is unknown. The spatio-temporal expression of any gene involved in development is key; therefore, the expression and functional involvement of *FGFR1-IIIb* may be important in the early embryonic brain, in particular during GnRH neuronal development. KS missense mutations in *FGFR1* are distributed in the first, second and third immunoglobulin-like domains (D1–D3), in the tyrosine kinase domain and also in the intracellular domain (Dodé *et al.*, 2003; Sato *et al.*, 2004; Albuisson *et al.*, 2005; Pitteloud *et al.*, 2006a; Trarbach *et al.*, 2006; Dodé *et al.*, 2007); therefore, the membrane-bound form of *FGFR1* is probably important for the KS phenotype. The membrane-bound form of *FGFR1-IIIb* could, therefore, be the critical isoform for the KS phenotype.

We present here, for the first time, a case of KS carrying a missense mutation in exon 8A of *FGFR1*, suggesting that the minor isoform '*FGFR1-IIIb*' as well as the major isoform '*FGFR1-IIIc*' has a crucial role in the pathogenesis of KS. Therefore, immunoglobulin-like domain IIIb may have an essential role in GnRH neuronal migration, which is initiated from the nasal placode and runs towards the forebrain following the olfactory sensory neuron axonal connection with the developing olfactory bulb. Further experiments are needed to show that the mutation in

exon 8A causes KS, such as expression of the mutated isoform in transfected cells to analyze receptor stability and signaling efficiency. Although *FGFR1* containing immunoglobulin-like domain IIIb has not been analyzed intensively, our mutation report should encourage researchers to analyze immunoglobulin-like domain IIIb function and the spatio-temporal expression of exon 8A in fetal brain development.

Acknowledgements

We thank Ms Yasuko Noguchi and Miho Ooga for their technical assistance and we thank Drs Nelly Pitteloud and William Crowley for their valuable contribution.

Funding

K.M. was supported, in part, by Seeds (No.15-B09) from the Japan Science and Technology Agency (JST), by grants from the Naito Foundation, by Grant-in-Aid for Young Scientists (B) (no. 21791567) from the Ministry of Education, Sports, Culture, Science and Technology of Japan, by a Grant for Child Health and Development (20C-1) from the Ministry of Health, Labor and Welfare, and by Grant-in-Aid for Scientific Research from Nagasaki University, Japan.

References

Abreu AP, Trarbach EB, de Castro M, Frade Costa EM, Versiani B, Matias Baptista MT, Garmes HM, Mendonca BB, Latronico AC. Loss-of-function mutations in the genes encoding prokineticin-2 or prokineticin receptor-2 cause autosomal recessive Kallmann syndrome. *J Clin Endocrinol Metab* 2008;**93**:4113–4118.

Albuisson J, Pêcheux C, Carel JC, Lacombe D, Leheup B, Lapuzina P, Bouchard P, Legius E, Matthijs G, Wasniewska M *et al.* Kallmann syndrome: 14 novel mutations in KAL1 and FGFR1 (KAL2). *Hum Mutat* 2005;**25**:98–99.

Beer HD, Vindevoghel L, Gait MJ, Revest JM, Duan DR, Mason I, Dickson C, Werner S. Fibroblast growth factor (FGF) receptor 1-IIIb is a naturally occurring functional receptor for FGFs that is preferentially expressed in the skin and the brain. *J Biol Chem* 2000; **275**:16091–16097.

Dodé C, Levilliers J, Dupont JM, De Paepe A, Le Dû N, Soussi-Yanicostas N, Coimbra RS, Delmaghani S, Compain-Nouaille S, Baverel F *et al.* Loss-of-function mutation in *FGFR1* cause autosomal dominant Kallmann syndrome. *Nature Genet* 2003;**33**:463–465.

Dodé C, Teixeira L, Levilliers J, Fouveaut C, Bouchard P, Kottler ML, Lespinasse J, Lienhardt-Roussie A, Mathieu M, Moerman A *et al.* Kallmann syndrome: mutations in the genes encoding prokineticin-2 prokineticin receptor-2. *PLoS Genet* 2006;**2**:e175.

Dodé C, Fouveaut C, Mortier G, Janssens S, Bertherat J, Mahoudeau J, Kottler ML, Chabrolle C, Gancel A, François I *et al.* Novel FGFR1 sequence variants in Kallmann syndrome, and genetic evidence that the FGFR1c isoform is required in olfactory bulb and palate morphogenesis. *Hum Mutat* 2007;**28**:97–98.

Doty RL, Applebaum S, Zusho H, Settle RG. Sex differences in odor identification ability: a cross-cultural analysis. *Neuropsychologia* 1985; **23**:667–672.

Falardeau J, Chung WC, Beenken A, Raivio T, Plummer L, Sidis Y, Jacobson-Dickman EE, Eliseenkova AV, Ma J, Dwyer A *et al.* Decreased FGF8 signaling causes deficiency of gonadotropin-releasing hormone in humans and mice. *J Clin Invest* 2008;**118**:2822–2831.

Franco B, Guioli S, Pragliola A, Incerti B, Bardoni B, Tonlorenzi R, Carrozzo R, Maestrini E, Pieretti M, Taillon-Miller P *et al.* A gene deleted in Kallmann's syndrome shares homology with neural cell adhesion and axonal path-finding molecules. *Nature* 1991;**353**:529–536.

Johnson DE, Lu J, Chen H, Werner A, Williams LT. The human fibroblast growth factor receptor genes: a common structure arrangement underlies the mechanisms for generating receptor forms that differ in their third immunoglobulin domain. *Mol Cell Biol* 1991;**11**:4627–4634.

Legouis R, Hardelin JP, Levilliers J, Claverie JM, Compain S, Wunderle V, Millasseau P, Le Paslier D, Cohen D, Caterina D *et al.* The candidate gene for the X-linked Kallmann syndrome encodes a protein related to adhesion molecules. *Cell* 1991;**67**:423–435.

Monnier C, Dodé C, Fabre L, Teixeira L, Labesse G, Pin JP, Hardelin JP, Rondard P. PROKR2 missense mutations associated with Kallmann syndrome impair receptor signalling activity. *Hum Mol Genet* 2009; **18**:75–81.

Partanen J, Schwartz L, Rossant J. Opposite phenotypes of hypomorphic and Y766 phosphorylation site mutations reveal a function for Fgfr1 in anteroposterior patterning of mouse embryos. *Genes Dev* 1998; **12**:2332–2344.

Pitteloud N, Acierno JS Jr, Meysing A, Eliseenkova AV, Ma J, Ibrahimi OA, Metzger DL, Hayes FJ, Dwyer AA, Hughes VA *et al.* Mutations in fibroblast growth factor receptor 1 cause both Kallmann syndrome and normosmic idiopathic hypogonadotropic hypogonadism. *Proc Natl Acad Sci USA* 2006a;**103**:6281–6286.

Pitteloud N, Meysing A, Quinton R, Acierno JS Jr, Dwyer AA, Plummer L, Fliers E, Boepple P, Hayes F, Seminara S *et al.* Mutations in fibroblast growth factor receptor 1 cause Kallmann syndrome with a wide spectrum of reproductive phenotypes. *Mol Cell Endocrinol* 2006b; **254–255**:60–69.

Ruta M, Burgess W, Givol D, Epstein J, Neiger N, Kaplow J, Crumley G, Dionne C, Jaye M, Schlessinger J. Receptor for acidic fibroblast growth factor is related to the tyrosine kinase encoded by the fms-like gene (FLG). *Proc Natl Acad Sci USA* 1989;**86**:8722–8726.

Sato N, Katsumata N, Kagami M, Hasegawa T, Hori N, Kawakita S, Minowada S, Shimotsuka A, Shishiba Y, Yokozawa M *et al.* Clinical assessment and mutation analysis of Kallmann syndrome 1 (KAL1) and fibroblast growth factor receptor 1 (FGFR1, or KAL2) in five families and 18 sporadic patients. *J Clin Endorinol Matab* 2004; **89**:1079–1088.

Trarbach EB, Costa EM, Versiani B, de Castro M, Baptista MT, Garmes HM, de Mendonca BB, Latronico AC. Novel fibroblast growth factor receptor 1 mutations in patients with congenital hypogonadotropic hypogonadism with and without anosmia. *J Clin Endocrinol Metab* 2006;**91**:4006–4012.

Trarbach EB, Silveira LG, Latronico AC. Genetic insights into human isolated gonadotropin deficiency. *Pituitary* 2007;**10**:381–391.

nature genetics

Exome sequencing identifies *MLL2* mutations as a cause of Kabuki syndrome

Sarah B Ng[1,7], Abigail W Bigham[2,7], Kati J Buckingham[2], Mark C Hannibal[2,3], Margaret J McMillin[2], Heidi I Gildersleeve[2], Anita E Beck[2,3], Holly K Tabor[2,3], Gregory M Cooper[1], Heather C Mefford[2], Choli Lee[1], Emily H Turner[1], Joshua D Smith[1], Mark J Rieder[1], Koh-ichiro Yoshiura[4], Naomichi Matsumoto[5], Tohru Ohta[6], Norio Niikawa[6], Deborah A Nickerson[1], Michael J Bamshad[1–3] & Jay Shendure[1]

We demonstrate the successful application of exome sequencing[1–3] to discover a gene for an autosomal dominant disorder, Kabuki syndrome (OMIM%147920). We subjected the exomes of ten unrelated probands to massively parallel sequencing. After filtering against existing SNP databases, there was no compelling candidate gene containing previously unknown variants in all affected individuals. Less stringent filtering criteria allowed for the presence of modest genetic heterogeneity or missing data but also identified multiple candidate genes. However, genotypic and phenotypic stratification highlighted *MLL2*, which encodes a Trithorax-group histone methyltransferase[4]: seven probands had newly identified nonsense or frameshift mutations in this gene. Follow-up Sanger sequencing detected *MLL2* mutations in two of the three remaining individuals with Kabuki syndrome (cases) and in 26 of 43 additional cases. In families where parental DNA was available, the mutation was confirmed to be *de novo* ($n = 12$) or transmitted ($n = 2$) in concordance with phenotype. Our results strongly suggest that mutations in *MLL2* are a major cause of Kabuki syndrome.

Kabuki syndrome is a rare, multiple malformation disorder characterized by a distinctive facial appearance (**Supplementary Fig. 1**), cardiac anomalies, skeletal abnormalities, immunological defects and mild to moderate mental retardation. Originally described in 1981 (refs. 5,6), Kabuki syndrome has an estimated incidence of 1 in 32,000 (ref. 7), and approximately 400 cases have been reported worldwide. The vast majority of reported cases have been sporadic, but parent-to-child transmission in more than a half dozen instances[8] suggests that Kabuki syndrome is an autosomal dominant disorder. The relatively low number of cases, the lack of multiplex families and the phenotypic variability of Kabuki syndrome have made the identification of the gene(s) underlying this disorder intractable to conventional approaches of gene discovery, despite aggressive efforts.

We sequenced the exomes of ten unrelated individuals with Kabuki syndrome: seven of European ancestry, two of Hispanic ancestry and one of mixed European and Haitian ancestry (**Supplementary Fig. 1** and **Supplementary Table 1**). Enrichment was performed by hybridization of shotgun fragment libraries to custom microarrays followed by massively parallel sequencing[1–3]. On average, 6.3 gigabases of sequence were generated per sample to achieve 40× coverage of the mappable, targeted exome (31 Mb). As with our previous studies, we focused our analyses here primarily on nonsynonymous variants, splice acceptor and donor site mutations and coding indels, anticipating that synonymous variants were far less likely to be pathogenic. We also predicted that variants underlying Kabuki syndrome are rare, and therefore likely to be previously unidentified. We defined variants as previously unidentified if they were absent from all datasets used for comparison, including dbSNP129, the 1000 Genomes Project, exome data from 16 individuals previously reported by us[2,3] and 10 exomes sequenced as part of the Environmental Genome Project (EGP).

Under a dominant model in which each case was required to have at least one previously unidentified nonsynonymous variant, splice acceptor and donor site mutation or coding indel variant in the same gene, only a single candidate gene (*MUC16*) was shared across all ten exomes (**Table 1** and **Supplementary Table 2**). However, we considered *MUC16* as a likely false positive due to its extremely large size (14,507 amino acids). Potential explanations for our failure to find a compelling candidate gene in which newly identified variants were seen in all affected individuals included: (i) Kabuki syndrome is genetically heterogeneous and therefore not all affected individuals will have mutations in the same gene; (ii) we failed to identify all mutations in the targeted exome; and (iii) some or all causative mutations were outside of the targeted exome, for example, in noncoding regions or unannotated genes. To allow for a modest degree of genetic heterogeneity and/or missing data, we conducted a less stringent analysis by looking for candidate genes shared among subsets of affected individuals. Specifically, we searched

[1]Department of Genome Sciences, University of Washington, Seattle, Washington, USA. [2]Department of Pediatrics, University of Washington, Seattle, Washington, USA. [3]Seattle Children's Hospital, Seattle, Washington, USA. [4]Department of Human Genetics, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan. [5]Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama, Japan. [6]Research Institute of Personalized Health Sciences, Health Sciences University of Hokkaido, Hokkaido, Japan. [7]These authors contributed equally to this work. Correspondence should be addressed to J.S. (shendure@u.washington.edu) or M.J.B. (mbamshad@u.washington.edu).

Received 28 April; accepted 21 July; published online 15 August 2010; doi:10.1038/ng.646

Table 1 Number of genes common to any subset of *x* affected individuals.

Subset analysis (any *x* of 10)	1	2	3	4	5	6	7	8	9	10
NS/SS/I	12,042	8,722	7,084	6,049	5,289	4,581	3,940	3,244	2,486	1,459
Not in dbSNP129 or 1000 Genomes	7,419	2,697	1,057	488	288	192	128	88	60	34
Not in control exomes	7,827	2,865	1,025	399	184	90	50	22	7	2
Not in either	6,935	2,227	701	242	104	44	16	6	3	1
Is loss-of-function (nonsense or frameshift indel)	753	49	7	3	2	2	1	0	0	0

The number of genes with at least one nonsynonymous variant (NS), splice-site acceptor or donor variants (SS) or coding indel (I) are listed under various filters. Variants were filtered by presence in dbSNP or 1000 Genomes (not in dbSNP129 or 1000 genomes) and control exomes (not in control exomes) or both (not in either); control exomes refer to those from 8 Hapmap[3], 4 FSS[3], 4 Miller[2] and 10 EGP samples. The number of genes found using the union of the intersection of *x* individuals is given.

for subsets of *x* out of 10 exomes having ≥1 previously unidentified variant in the same gene, with $x = 1$ to $x = 10$. For $x = 9$, $x = 8$ and $x = 7$, previously unidentified variants were shared in 3 genes, 6 genes and 16 genes, respectively (**Table 1**). However, there was no obvious way to rank these candidate genes.

We speculated that genotypic and/or phenotypic stratification would facilitate the prioritization of candidate genes identified by subset analysis. Specifically, we assigned a categorical rank to each individual with Kabuki syndrome based on a subjective assessment of the presence of, or similarity to, the canonical facial characteristics of Kabuki syndrome (**Supplementary Fig. 1**) and the presence of developmental delay and/or major birth defects (**Supplementary Table 1**). The highest-ranked individual was one of a pair of monozygotic twins with Kabuki syndrome. We then categorized the functional impact (that is, nonsense versus nonsynonymous substitution, splice-site disruption and frameshift compared to in-frame indel) of each newly identified variant in candidate genes shared by each subset of two or more ranked cases. Manual review of these data highlighted distinct, previously unidentified nonsense variants in *MLL2* in each of the four highest-ranked cases. After sequential analysis of phenotype-ranked cases with a loss-of-function filter, *MLL2* was the only candidate gene remaining after addition of the second individual (**Table 2**). We found no such variant in *MLL2* in the individual with Kabuki syndrome ranked fifth; hence, the number of candidate genes dropped to zero after the individual ranked fourth in the set (**Table 2**). However, we found a 4-bp deletion in the individual ranked sixth, and we found nonsense variants in the individuals ranked seventh and ninth. Thus, exome sequencing identified a nonsense substitution or frameshift indel in *MLL2* in seven of the ten individuals with Kabuki syndrome analyzed here.

Retrospectively, we applied a loss-of-function filter to the subset analysis of exome data (**Table 1**), and at $x = 7$, found *MLL2* to be the only candidate gene. We also developed a *post hoc* ranking of candidate genes based on the functional impact of the variants present (variant score) and the rank of the cases in which each variant was observed (case score). When this was applied to the exome data as a combined metric, *MLL2* emerged as the top candidate gene (**Supplementary Fig. 2**).

In parallel with these analyses, we applied genomic evolutionary rate profiling (GERP)[9] to the exome data. GERP uses mammalian genome alignments to define a rejected substitution score for each variant regardless of functional class. We have previously shown that the quantitative ranking of candidate genes by the rejected substitution scores of their variants can facilitate the exome-based analysis of Mendelian disorders[10]. Following subset analysis with GERP-based ranking, *MLL2* remained on the candidate list up to $x = 8$, ranking third in a list of 11 candidate genes at this threshold (**Table 3** and **Supplementary Fig. 3**). Notably, the additional *MLL2* variant contributing to this analysis (such that *MLL2* was still considered at $x = 8$) was a synonymous substitution with a rejected substitution score of 0.368 in the individual ranked fifth.

We sought to confirm all newly identified variants in *MLL2*, particularly because loss-of-function variants identified through massively parallel sequencing have a high prior probability of being false positives. All seven loss-of-function variants in *MLL2* were validated by Sanger sequencing. We further analyzed the three cases in which we did not initially find a loss-of-function variant in *MLL2*, first by array comparative genomic hybridization (aCGH) to determine any gross structural changes and then by Sanger sequencing of all exons of *MLL2* in case of false negatives by exome sequencing. Because an average of 96% of the coding bases in *MLL2* were called at sufficient quality and coverage for single nucleotide variant detection, we anticipated that any missed variants were more likely to be indels because of the higher coverage required for confident indel detection in short-read sequence data. Indeed, although aCGH did not find any structural variants in the region, Sanger sequencing did identify frameshift indels in two of these three cases (specifically, the cases ranked eighth and tenth).

Ultimately, loss-of-function mutations in *MLL2* were identified in nine out of ten cases in the discovery cohort (**Fig. 1**), making this gene a compelling candidate for Kabuki syndrome. For validation, we screened all 54 exons of *MLL2* in 43 additional cases by Sanger sequencing. Previously unidentified nonsynonymous, nonsense or frameshift mutations in *MLL2* were found in 26 of these 43 cases (**Fig. 1** and **Supplementary Table 3**). In total, through either exome sequencing or targeted sequencing of *MLL2*, 33 distinct *MLL2* mutations were identified in 35 of 53 families (66%) with Kabuki syndrome (**Fig. 1** and **Supplementary Table 3**). In each of 12 cases for which DNA from both parents was available, the *MLL2* variant was found to have occurred *de novo*. Three mutations were found in two individuals each. One of these three mutations was confirmed to have arisen *de novo* in one of the cases, indicating that some mutations in individuals with Kabuki syndrome are recurrent. In addition, *MLL2* mutations (resulting in p.4527K>X and p.5464T>M) were also identified in each of two families in which Kabuki syndrome was transmitted from parent to child.

Table 2 Number of genes common in sequential analysis of phenotypically ranked individuals

Sequential analysis	1	+2	+3	+4	+5	+6	+7	+8	+9	+10
NS/SS/I	5,282	3,850	3,250	2,354	2,028	1,899	1,772	1,686	1,600	1,459
Not in dbSNP129 or 1000 Genomes	687	214	145	84	63	54	42	40	39	34
Not in control exomes	675	134	50	26	13	13	8	5	4	2
Not in either	467	89	34	18	9	8	4	4	3	1
Is loss-of-function (nonsense/frameshift indel)	25	1	1	1	0	0	0	0	0	0

Variants were filtered as in **Table 1**. Exomes were added sequentially to the analysis by ranked phenotype; for example, column "+3" shows the number of genes at the intersection of the three top ranked cases (**Supplementary Fig. 1**). The gene with at least one NS/SS/I in all individuals is *MUC16*, which is very likely to be a false positive due to its extreme length (14,507 amino acids).

Table 3 Analysis of exome variants using genomic evolutionary rate profiling

GERP score analysis (at least *x* of 10)	1	2	3	4	5	6	7	8	9	10
Variant RS score > 0	7,176	2,360	754	269	106	39	20	11	3	1
MLL2 rank	3,732	1,232	399	136	47	14	6	3	NA	NA

The number of genes with at least a single previously unidentified variant with a rejected substitution score[10] > 0 in at least *x* individuals is given. A gene rank is assigned based on the average GERP score[9] over all newly identified variants with rejected substitution score > 0 in all affected individuals.

None of the additional *MLL2* mutations was found in 190 control chromosomes from individuals of matched geographical ancestry.

Our results strongly suggest that mutations in *MLL2* are a major cause of Kabuki syndrome. *MLL2* encodes a large 5,262-residue protein that is part of the SET family of proteins, of which Trithorax, the *Drosophila* homolog of MLL, is the best characterized[11]. The SET domain of MLL2 confers strong histone 3 lysine 4 methyltransferase activity and is important in the epigenetic control of active chromatin states[12]. In mice, loss of *Mll2* on a mixed 129Sv/C57BL/6 background slows growth, increases apoptosis and retards development, leading to early embryonic lethality due in part to misregulation of homeobox gene expression[13]. However, no morphological defects have been reported in $Mll2^{+/-}$ mice[13].

Most of the *MLL2* variants identified in individuals with Kabuki syndrome are predicted to truncate the polypeptide chain before translation of the SET domain. Though it is not certain whether Kabuki syndrome results from haploinsufficiency or from a gain of function at *MLL2*, haploinsufficiency seems to be the more likely mechanism. Deletion of chromosome 12q12–q13.2, which encompasses *MLL2*, has been reported in a child with characteristics of Noonan syndrome[14]. However, we re-analyzed this case using oligo aCGH (including 21 probes that cover *MLL2*) and found the distal breakpoint to be located ~700 kb proximal to *MLL2* (data not shown). Also, all of the pathogenic missense variants identified here are located in regions of *MLL2* that encode C-terminal domains. This suggests that missense variants elsewhere in *MLL2* may be better tolerated or, alternatively, may be embryonically lethal.

For the 18 of 53 cases for which no previously unidentified protein-altering variant was found, it is possible that noncoding or other missed mutations in *MLL2* are responsible for this disorder. Alternatively, Kabuki syndrome could be genetically heterogeneous, and further analysis of these cases by exome sequencing may elucidate additional genes for Kabuki syndrome and potentially explain some of the phenotypic heterogeneity seen in this disorder. Notably, 9 of 10 individuals in the discovery cohort (90%), but only 26 of 43 individuals in the replication cohort (60%), were ultimately found to have mutations in *MLL2*. It is therefore possible that the careful selection of canonical Kabuki cases for the discovery cohort enriched for a shared genetic basis. This underscores the importance of access to deeply phenotyped and well-characterized cases.

In summary, we applied exome sequencing of a small number of unrelated individuals with Kabuki syndrome to discover that mutations in *MLL2* underlie this disorder. As predicted in previous analyses[2,3], allowing for even a small degree of genetic heterogeneity or missing data substantially confounds exome analysis by increasing the number of candidate genes consistent with the model of inheritance. To facilitate the prioritization of genes under such criteria, we stratified data by ranked phenotypes and found that *MLL2* was prominent in the higher ranked cases. However, nine of the ten individuals with Kabuki syndrome in the discovery cohort were ultimately found to have *MLL2* mutations, such that stratification by phenotype was of less importance than originally appeared to have been the case. Nonetheless, the sequential analysis of ranked cases may have reduced the probability of confounding due to genetic heterogeneity. All of the *MLL2* mutations found in the discovery set via exome sequencing were loss-of-function variants. As a result, *MLL2* ranked highly among candidate genes assessed by predicted functional impact. Such a pattern will likely occur for some, but not all, Mendelian phenotypes subjected to this approach. We anticipate that the further development of strategies to stratify data at both the genotypic and phenotypic level will be critical for exome and whole-genome sequencing to reach their full potential as tools for discovery of genes underlying Mendelian and complex diseases.

URLs. RefSeq 36.3, ftp://ftp.ncbi.nlm.nih.gov/genomes/MapView/Homo_sapiens/sequence/BUILD.36.3/updates/seq_gene.md.gz; Phaster, http://www.phrap.org; SeattleSeq Annotation, http://gvs.gs.washington.edu/SeattleSeqAnnotation/; 1000 Genomes Project, http://www.1000genomes.org/page.php/; dbGaP accession, http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000295.v1.p1.

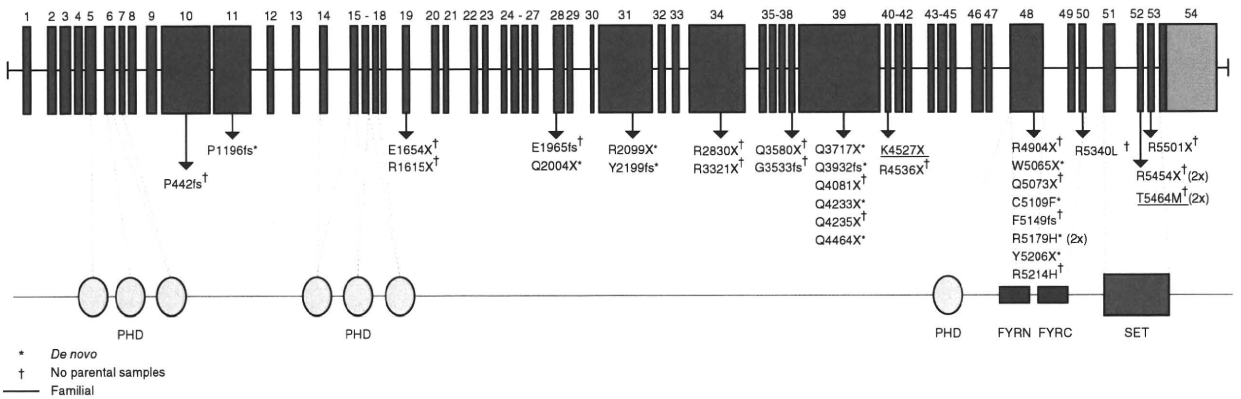

Figure 1 Genomic structure and allelic spectrum of *MLL2* mutations that cause Kabuki syndrome. *MLL2* is composed of 54 exons that encode untranslated regions (orange) and protein coding sequence (blue) including 7 PHD fingers (yellow), FYRN (green), FYRC (green) and a SET domain (red). Arrows indicate the locations of 32 different mutations found in 53 families with Kabuki syndrome including 20 nonsense mutations, 7 indels and 5 amino acid substitutions. Asterisks indicate mutations that were confirmed to be *de novo* and crosses indicate cases for which parental DNA was unavailable. The two underlined mutations were transmitted each within a family, from an affected parent to an affected child.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Accession codes. Exome data for the discovery cohort is available via the NCBI dbGaP repository under accession number phs000295.v1.p1.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank the families for their participation and the Kabuki Syndrome Network for their support. We thank J. Allanson, J. Carey and M. Golabi for referral of cases and M. Emond for helpful discussion. We thank the 1000 Genomes Project for early data release that proved useful for filtering out common variants. Our work was supported in part by grants from the US National Institutes of Health (NIH)–National Heart, Lung, and Blood Institute (5R01HL094976 to D.A.N. and J.S.), the NIH–National Human Genome Research Institute (5R21HG004749 to J.S., 1RC2HG005608 to M.J.B., D.A.N. and J.S.; and 5RO1HG004316 to H.K.T.), NIH–National Institute of Environmental Health Sciences (HHSN273200800010C to D.N. and M.J.R.), Ministry of Health, Labour and Welfare (K.Y., N.M., T.O. and N.N.), Japan Science and Technology Agency (N.M.), Society for the Promotion of Science (N.M.), the Life Sciences Discovery Fund (2065508 and 0905001), the Washington Research Foundation and the NIH–National Institute of Child Health and Human Development (1R01HD048895 to M.J.B.). S.B.N. is supported by the Agency for Science, Technology and Research, Singapore. A.W.B. is supported by a training fellowship from the NIH–National Human Genome Research Institute (T32HG00035).

AUTHOR CONTRIBUTIONS

The project was conceived and the experiments were planned by M.J.B., D.A.N. and J.S. The review of phenotypes and the sample collection were performed by M.J.B., M.C.H., M.J.M., K.Y., N.M., T.O. and N.N. Experiments were performed by S.B.N., K.J.B., A.E.B., C.L., H.C.M., J.D.S., M.J.R., E.H.T. and H.I.G. Ethical consultation was provided by H.K.T. Data analysis was performed by A.W.B., M.J.B., K.J.B., G.M.C., S.B.N. and J.S. The manuscript was written by M.J.B., S.B.N. and J.S. All aspects of the study were supervised by M.J.B. and J.S.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Choi, M. *et al.* Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc. Natl. Acad. Sci. USA* **106**, 19096–19101 (2009).
2. Ng, S.B. *et al.* Exome sequencing identifies the cause of a Mendelian disorder. *Nat. Genet.* **42**, 30–35 (2010).
3. Ng, S.B. *et al.* Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* **461**, 272–276 (2009).
4. FitzGerald, K.T. & Diaz, M.O. MLL2: A new mammalian member of the *trx/MLL* family of genes. *Genomics* **59**, 187–192 (1999).
5. Niikawa, N., Matsuura, N., Fukushima, Y., Ohsawa, T. & Kajii, T. Kabuki make-up syndrome: a syndrome of mental retardation, unusual facies, large and protruding ears, and postnatal growth deficiency. *J. Pediatr.* **99**, 565–569 (1981).
6. Kuroki, Y., Suzuki, Y., Chyo, H., Hata, A. & Matsui, I. A new malformation syndrome of long palpebral fissures, large ears, depressed nasal tip, and skeletal anomalies associated with postnatal dwarfism and mental retardation. *J. Pediatr.* **99**, 570–573 (1981).
7. Niikawa, N. *et al.* Kabuki make-up (Niikawa-Kuroki) syndrome: a study of 62 patients. *Am. J. Med. Genet.* **31**, 565–589 (1988).
8. Courtens, W., Rassart, A., Stene, J.J. & Vamos, E. Further evidence for autosomal dominant inheritance and ectodermal abnormalities in Kabuki syndrome. *Am. J. Med. Genet.* **93**, 244–249 (2000).
9. Cooper, G.M. *et al.* Distribution and intensity of constraint in mammalian genomic sequence. *Genome Res.* **15**, 901–913 (2005).
10. Cooper, G.M. *et al.* Single-nucleotide evolutionary constraint scores highlight disease-causing mutations. *Nat. Methods* **7**, 250–251 (2010).
11. Prasad, R. *et al.* Structure and expression pattern of human *ALR*, a novel gene with strong homology to *ALL-1* involved in acute leukemia and to *Drosophila* trithorax. *Oncogene* **15**, 549–560 (1997).
12. Issaeva, I. *et al.* Knockdown of ALR (MLL2) reveals ALR target genes and leads to alterations in cell adhesion and growth. *Mol. Cell. Biol.* **27**, 1889–1903 (2007).
13. Glaser, S. *et al.* Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. *Development* **133**, 1423–1432 (2006).
14. Tonoki, H., Saitoh, S. & Kobayashi, K. Patient with del(12)(q12q13.12) manifesting abnormalities compatible with Noonan syndrome. *Am. J. Med. Genet.* **75**, 416–418 (1998).