Table 3 Comparison of the non-synonymous variant frequencies between the patients with CP and controls

Amino acid change	Genotype	All CP (%)	HGVD (%)	*P* value (vs. HGVD)				
				All CP	Alcoholic CP	Nonalcoholic CP	Idiopathic CP	Hereditary/ familial CP
p.R31C	CT	3/193 (1.6)	12/1102 (1.1)	0.48	>0.99	0.41	0.18	>0.99
p.R31H	GA	1/193 (0.5)	0	–	–	–	–	–
p.I125T	TC	3/193 (1.6)	5/1102 (0.5)	0.11	>0.99	0.057	0.15	0.13
p.K411E	AG	1/193 (0.5)	0	–	–	–	–	–
p.V470M	GA	97/193 (50.3)	573/1199 (47.8)	0.66	0.57	0.68	0.38	0.12
	AA	25/193 (13.0)	185/1199 (15.4)					
p.I556V	AG	10/193 (5.2)	78/1150 (6.8)	0.70	0.79	0.81	>0.99	0.45
	GG	0/193 (0)	3/1150 (0.3)					
p.E585X	GT	1/193 (0.5)	0	–	–	–	–	–
p.L957fs		1/193 (0.5)	0	–	–	–	–	–
p.L1156F	GT	14/193 (7.3)	45/1136 (4.0)	0.04	0.06	0.07	0.11	0.30
	TT	1/193 (0.5)	1/1136 (0.1)					
p.G1349S	GA	1/193 (0.5)	4/1094 (0.4)	0.56	0.19	>0.99	>0.99	>0.99
p.Q1352H	GC	20/193 (10.4)	57/1153 (4.9)	0.009	0.12	0.037	0.17	0.062
	CC	0/193 (0)	1/1153 (0.1)					
p.R1453W	CT	10/193 (5.2)	42/1144 (3.7)	0.32	0.25	0.49	0.45	>0.99

CP chronic pancreatitis, *HGVB* Human Genetic Variation Database

P values were determined versus HGVD by the Fisher's exact test

Table 4 Synonymous variants in the exons of the *CFTR* gene detected in this study

Exon	Synonymous variant	Amino acid change	dbSNP135	Genotype	Alcoholic CP (%)	Idiopathic CP (%)	Hereditary/ familial CP (%)
4	c.372C>T	p.G124=	–	CT	0/46 (0)	1/121 (0.8)	0/26 (0)
13	c.1731C>T	p.Y577=	rs55928397	CT	0/46 (0)	1/121 (0.8)	0/26 (0)
15	c.2562T>G	p.T854=	rs1042077	TG	20/46 (43.5)	69/121 (57.0)	12/26 (46.2)
				GG	6/46 (13.0)	18/121 (14.9)	0/26 (0)
23	c.3723C>A	p.G1241=	rs185065886	CA	1/46 (2.2)	0/121 (0)	0/26 (0)
25	c.3975A>G	p.R1325=	–	AG	0/46 (0)	1/121 (0.8)	0/26 (0)
27	c.4254G>A	p.E1418=	–	GA	0/46 (0)	1/121 (0.8)	0/26 (0)
27	c.4389G>A	p.Q1463=	rs1800136	GA	1/46 (2.2)	3/121 (2.5)	0/26 (0)

CP chronic pancreatitis

heterozygous form (Table 6). The nonsense variant c.1753G>T (p.E585X) was found in a patient with alcoholic CP. He was diagnosed as having alcoholic CP at 28 years old. The c.1231A>G (p.411E) variant was found in a 19-year-old male with idiopathic CP. He had suffered from pancreatitis attacks since 12 years old. ERCP showed multiple stones in the main pancreatic duct. He underwent extracorporeal shock wave lithotripsy for the treatment of pancreatic stones. The patient also had the c.3468G>T (p.L1156F) variant in a heterozygous form. None of these three patients had known pancreatitis susceptibility mutations in the *PRSS1*, *SPINK1*, *CTRC*, or *CPA1* genes (Table 6). All of the patients carrying the novel synonymous variants were idiopathic CP (Table 4).

The frequency of the c.4056G>C (p.Q1352H) variant was higher in all patients with CP than that in controls ($P = 0.009$; Table 3). Stratification based on the etiologies showed that the association was significant in patients with nonalcoholic CP (combination of cases with idiopathic, hereditary, and familial CP) ($P = 0.037$). The frequency of the c.3468G>T (p.L1156F) variant was also higher in patients with CP than that in controls ($P = 0.04$). There were no significant difference for any other non-synonymous or synonymous variants detected in the exons

Table 5 Comparison of the synonymous variant frequencies between the patients with CP and controls

Synonymous variant	Genotype	All CP (%)	HGVD (%)	*P* value (vs. HGVD)				
				All CP	Alcoholic CP	Nonalcoholic CP	Idiopathic CP	Hereditary/ familial CP
c.C372T	CT	1/193 (0.5)	0	–	–	–	–	–
c.1731C>T	CT	1/193 (0.5)	0	–	–	–	–	–
c.2562T>G	TG	101/193 (52.3)	528/1154 (45.8)	0.22	0.81	0.11	0.045	0.033
	GG	24/193 (12.4)	181/1154 (15.7)					
c.3723C>A	CA	1/193 (0.5)	3/671 (4.5)	>0.99	0.23	>0.99	>0.99	>0.99
c.3975A>G	AG	1/193 (0.5)	0	–	–	–	–	–
c.4254G>A	GA	1/193 (0.5)	0	–	–	–	–	–
c.4389G>A	GA	4/193 (2.1)	40/1112 (3.6)	0.48	>0.99	0.53	0.81	>0.99
	AA	0/193 (0)	1/1112 (0.1)					

CP chronic pancreatitis, *HGVD* Human Genetic Variation Database

P values were determined against HGVD by the Fisher's exact test

Table 6 Total *CFTR* sequencing results of patients carrying rare non-synonymous *CFTR* variants

Case#	Etiology	Age at onset	Rare variant	Additional non-synonymous variants	c.1210-34TG(9_13) c.1210-12T(5_9)	Mutation in other pancreatitis susceptibility genes[a]
A1	Idiopathic	34	p.R31C/-	p.R1453W/-	TG11/TG11, 7T/7T	–
A2	Idiopathic	8	p.R31C/-	–	TG11/TG12, 7T/7T	–
A3	Idiopathic	16	p.R31C/-	–	TG11/TG12, 7T/7T	–
A4	Idiopathic	10	p.R31H/-	–	TG11/TG12, 7T/7T	–
A5	Idiopathic	16	p.I125T/-	p.L1156F/-	TG11/TG12, 7T/7T	*CTRC* p.R29Q/-
A6	Idiopathic	2	p.I125T/-	–	TG11/TG12, 7T/7T	–
A7	Hereditary	28	p.I125T/-	p.R1453W/-	TG11/TG12, 7T/7T	–
A8	Idiopathic	19	p.K411E/-	p/L1156F/-	TG11/TG12, 7T/7T	–
A9	Alcoholic	28	p.E585X/-	p.I556V/-	TG11/TG11, 7T/7T	–
A10	Idiopathic	21	p.L957fs/-	p.Q1352H/-	TG11/TG12, 7T/7T	–
A11	Alcoholic	40	p.G1349S/-	–	TG11/TG11, 7T/7T	–

[a] Pancreatitis-associated mutations in the *PRSS1*, *SPINK1*, *CTRC*, and *CPA1* genes

between all patients with CP and controls (Tables 3, 5). The frequency of the c.2562T>G variant was different between the controls and the patients with idiopathic or hereditary/familial CP.

The 5T and, more rarely, 3T splicing variants of the intron 9 acceptor splice site [c.1210-12T(5_9)] are considered to be variants associated with CFTR-RD [16]. The 5T or 3T allele is a polymorphic variant with variable penetrance, causing less efficient exon 10 splicing and a lower *CFTR* transcript level [28]. The splicing efficiency of exon 10 is further affected by the length of the adjacent TG repeat [c.1210-34TG(9_13)]. The distribution of the c.1210-34TG(9_13) and c.1210-12T(5_9) variants is shown in Table 7. In our cohort, nine patients with CP had the 5T allele, all in a heterozygous form. Four patients (two alcoholic, one idiopathic, one hereditary) had the 5T-TG13. No patient had the haplotype TG10-7T-M470, which was reported to increase the risk of idiopathic CP [28].

It has been increasingly recognized that compound and trans-heterozygosity in the pancreatitis susceptibility genes are an overt risk factor for idiopathic CP [29–32]. Among the 193 patients with CP enrolled in this study, 29 patients had pancreatitis-associated mutations in the *PRSS1*, *SPINK1*, *CTRC*, and *CPA1* genes. Among these, nine patients had the non-synonymous *CFTR* variants, which are probably damaging based on the SIFT and/or the PolyPhen-2 prediction (Table 8).

Discussion

In this study, we performed comprehensive analysis of the variants in the *CFTR* gene by targeted NGS. To our knowledge, this is the first study to analyze pancreatitis susceptibility genes by targeted NGS. Comprehensive analysis by targeted NGS enabled us to identify novel and

Table 7 Distribution of the c.1210-34TG(9_13) and c.1210-12T(5_9) variants in patients with CP

c.1210-34TG(9_13), c.1210-12T(5_9)	All CP (%)	Alcoholic CP (%)	Idiopathic CP (%)	Hereditary/ familial CP (%)
TG10/TG11, 7T/9T	1/193 (0.5)	0/46 (0)	1/121 (0.8)	0/26 (0)
TG11/TG11, 7T/7T	46/193 (23.8)	15/46 (32.6)	23/121 (19.0)	8/26 (30.8)
TG11/TG11, 7T/9T	4/193 (2.1)	1/46 (2.2)	3/121 (2.5)	0/26 (0)
TG11/TG12, 5T/7T	5/193 (2.6)	0/46 (0)	4/121 (3.3)	1/26 (3.8)
TG11/TG12, 6T/7T	1/193 (0.5)	0/46 (0)	1/121 (0.8)	0/26 (0)
TG11/TG12, 7T/7T	124/193 (64.2)	27/46 (58.7)	81/121 (66.9)	16/26 (61.5)
TG11/TG13, 6T/7T	1/193 (0.5)	1/46 (2.2)	0/121 (0)	0/26 (0)
TG11/TG13, 7T/7T	1/193 (0.5)	0/46 (0)	1/121 (0.8)	0/26 (0)
TG12/TG12, 7T/7T	6/193 (3.1)	0/46 (0)	6/121 (5.0)	0/26 (0)
TG12/TG13, 5T/7T	4/193 (2.1)	2/46 (4.3)	1/121 (0.8)	1/26 (3.8)

CP chronic pancreatitis

rare variants in the *CFTR* gene. The c.1753G>T (p.E585X) variant is a nonsense variant, and the c.2869delC (p.L957fs) variant leads to a stop codon afterward at amino acid 967. These variants result in a heavily truncated protein missing nearly two-thirds (p.E585X) or more than one-third (p.L957fs) of its amino acids. Because we did not perform functional assays, we do not have direct evidence that these two variants cause loss of the CFTR expression and/or function. However, a general acknowledgment has been agreed that mutations of this type, called class I mutations, are associated with complete loss or near complete loss of the CFTR function (<3 % of wild-type CFTR function) [33, 34]. The pathogenic potential of another novel variant, c.1231A>G (p.K411E), is currently unknown, but the in silico analyses suggest that this variant is deleterious. Importantly, the clinical phenotype of this patient might be complicated by the presence of another variant, p.L1156F. Noone et al. [30] reported that pancreatitis risk was increased approximately 40-fold by having two *CFTR* mutations. This is also the case with the c.374T>C (p.I125T) variant. Two of the three patients carrying this variant had other non-synonymous variants (p.I556V and p.R1453W). This p.I125T variant was originally reported in Chinese patients with idiopathic bronchiectasis and considered to be associated with CFTR-RD [35].

There are considerable regional and ethnic variations in the spectrum of the *CFTR* mutations [15]. Approximately 70 % of individuals with CF in the Caucasian population are homozygous for the F508del mutation, and almost 90 % of the patients have at least one F508del allele [36]. This mutation is extremely rare in the Japanese population, accounting for the rare presentation of classical CF in this region (approximately 1/350,000 live births) [37]. It is not surprising that the CF-causing mutations are frequently found in Caucasians, but very rarely in East Asia. Audrézet et al. [29] reported from France that at least 20 % of the patients with idiopathic CP carried one of the most common *CFTR* mutations. Fujiki et al. [38] reported from Japan that none of the 20 common CF-causing mutations was found in 65 Japanese patients with CP (51 alcoholic and 14 idiopathic). Wang et al. [39] reported comprehensive screening of pancreatitis susceptibility genes including *CFTR* in 75 pediatric patients with idiopathic CP from China. They identified a novel 8-bp deletion in exon 4, but not the common CF-causing mutations. In this study, we found no common severe CF-causing mutations, in agreement with these previous studies from East Asia.

We found a significant association between the p.Q1352H variant and CP. This finding confirms the previous reports from Japan and Korea showing that this variant was over-presented in patients with CP compared to controls [38, 40]. Fujiki et al. [38] from Japan reported that the frequency of this variant was higher in patients with CP (8/65, 12.3 %) than in controls (6/162, 3.7 %). Lee et al. [40] reported from Korea that 14.3 % (4/28) of the patients with CP had this variant, whereas only 0.9 % (1/117) of the controls did. Glutamine at 1,352 is located in the second nucleotide-binding fold of CFTR, and its change to histidine (p.Q1352H) causes reductions in both the protein expression and channel activity of CFTR [40]. Similarly, we found that the p.L1156F variant was overexpressed in patients with CP. A functional study reported reduced Cl^-/HCO_3^- permeability in the presence of the p.L1156F variant [41].

Gene–gene interactions of known pancreatitis susceptibility genes, especially between the *CFTR* and *SPINK1* genes, have been increasingly recognized. Indeed, seven out of 25 patients carrying the *SPINK1* variant(s) had the *CFTR* p.Q1352H and/or p.L1156F variants. One patient was trans-heterozygous for the *CTRC* p.R29Q and *CFTR* p.I125T/p.L1156F variants. Noone et al. [30] reported that pancreatitis risk was increased approximately 40-fold by having two *CFTR* mutations, 20-fold by having the *SPINK1* p.N34S variant, and 900-fold by having both. Trans-heterozygosity of the *SPINK1* p.N34S with the

Table 8 Total *CFTR* sequencing results of patients with *SPINK1*, *PRSS1*, *CTRC*, or *CPA1* mutations

Case#	Etiology	*CFTR* variants[a]	c.1210-34TG(9_13) c.1210-12T(5_9)	*SPINK1*	*PRSS1*	*CTRC*	*CPA1*
B1	Familial	p.Q1352H/-	TG11/TG12, 7T/7T	p.N34S/p.N34S			
B2	Idiopathic	–	TG12/TG12, 7T/7T	p.N34S/p.N34S			
B3	Idiopathic	–	TG11/TG12, 7T/7T	p.N34S/p.N34S			
B4	Idiopathic	p.L1156F/-, p.Q1352H/-	TG11/TG11, 7T/7T	p.N34S/-			
B5	Idiopathic	p.Q1352H/-	TG11/TG12, 7T/7T	p.N34S/-			
B6	Idiopathic	p.Q1352H/-	TG11/TG12, 7T/7T	p.N34S/-			
B7	Idiopathic	–	TG11/TG12, 7T/7T	p.N34S/-			
B8	Idiopathic	–	TG11/TG12, 7T/7T	p.N34S/-			
B9	Idiopathic	–	TG11/TG12, 7T/7T	p.N34S/-			
B10	Idiopathic	–	TG11/TG12, 7T/7T	p.N34S/-			
B11	Idiopathic	–	TG11/TG12, 7T/7T	p.N34S/-			
B12	Idiopathic	–	TG11/TG12, 7T/7T	p.N34S/-			
B13	Idiopathic	–	TG11/TG12, 7T/7T	p.N34S/-			
B14	Alcoholic	–	TG12/TG13, 5T/7T	p.N34S/-			
B15	Idiopathic	–	TG11/TG12, 7T/7T	p.N34S/IVS3+2T>C			
B16	Idiopathic	p.R1453W/-	TG11/TG11, 7T/7T	p.N34S/IVS3+2T>C			
B17	Idiopathic	–	TG11/TG12, 7T/7T	IVS3+2T>C/IVS3+2T>C			
B18	Idiopathic	–	TG11/TG12, 7T/7T	IVS3+2T>C/IVS3+2T>C			
B19	Hereditary	p.I125T/-, p.L1156F/-	TG11/TG12, 5T/7T	IVS3+2T>C/-			
B20	Familial	p.L1156F/-	TG11/TG12, 7T/7T	IVS3+2T>C/-			
B21	Idiopathic	–	TG11/TG12, 7T/7T	IVS3+2T>C/-			
B22	Alcoholic	p.Q1352H/-	TG11/TG12, 7T/7T	IVS3+2T>C/-			
B23	Alcoholic	–	TG11/TG12, 7T/7T	IVS3+2T>C/-			
B24	Idiopathic	–	TG11/TG12, 7T/7T	p.P45S/-			
B25	Idiopathic	–	TG12/TG12, 7T/7T	IVS3+2T>C/-	p.R122H/-		
B26	Hereditary		TG11/TG12, 7T/7T		p.R122H/-		
B27	Idiopathic	p.I556V/-	TG11/TG12, 7T/7T		p.N29I/-		
B28	Idiopathic	p.I125T/-, p.L1156F/-	TG11/TG12, 7T/7T			p.R29Q/-	
B29	Idiopathic	–	TG11/TG12, 7T/7T				p.T368_Y369ins20/-

Nine patients had the non-synonymous *CFTR* variants, which are probably damaging based on the SIFT or the PolyPhen-2 prediction

The p.I556V variant appeared to be benign based on the SIFT or the PolyPhen-2 prediction

Case B28 is the same as A5 in Table 6

[a] We excluded the p.V470M variant from the list because of its similar frequencies in patients and controls

CFTR p.R75Q was reported to increase CP risk [31]. 6.5 % of the patients with idiopathic or hereditary CP carried variants in at least two pancreatitis susceptibility genes [32]. Whether the coinheritance of variants/mutations in pancreatitis susceptibility gene is a *bona fide* example of digenic inheritance or interaction between a disease-causing gene and a genetic modifier is unclear in most cases [42].

We used targeted sequence capture and high-throughput NGS to detect variants in the *CFTR* gene. Due to the large size (27 exons, 1,480 amino acids), traditional technologies, such as PCR and capillary sequencing, are time- and cost-consuming. A major advantage of the HaloPlex-targeted enrichment system is the convenient workflow, integrating both capture and library preparation. The protocol allows one person to prepare a set of finished libraries within two working days and requires no larger specialized instruments. Sequence capture eliminates the necessity of setting up hundreds of PCR, instead allowing for parallel

enrichment of target regions in a single experiment. A weakness of this method is that the detection of larger copy number variations would require different methods. We have designed the HaloPlex platform for more than 70 genes, including the known pancreatitis susceptibility genes such as *CFTR*, *PRSS1*, *SPINK1*, *CTRC*, and *CPA1*. This system has allowed us to perform rapid screening of the known susceptibility genes simultaneously and gives an overview of potentially pathogenic variants in patients with pancreatitis. In addition, our HaloPlex platform includes candidates of novel pancreatitis susceptibility genes such as pancreatic digestive enzymes, those highly expressed in the pancreas and those related to autophagy and endoplasmic reticulum stress. This system might contribute to the identification of novel pancreatitis susceptibility genes in the future.

Acknowledgments The authors are grateful to Ms. Yoko Tateda for the excellent technical assistance. This work was supported in part by the Pancreas Research Foundation of Japan (to E. Nakano), the HIROMI Medical Research Foundation (to A. Masamune), the Mother and Child Health Foundation (to A. Masamune), the Smoking Research Foundation (to A. Masamune), and by the Ministry of Health, Labour, and Welfare of Japan.

Conflict of interest None.

References

1. Steer ML, Waxman I, Freedman S. Chronic pancreatitis. *N Engl J Med.* 1995;332:1482–1490.
2. Witt H, Apte MV, Keim V, Wilson JS. Chronic pancreatitis: challenges and advances in pathogenesis, genetics, diagnosis, and therapy. *Gastroenterology.* 2007;132:1557–1573.
3. Whitcomb DC, Gorry MC, Preston RA, et al. Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene. *Nat Genet.* 1996;14:141–145.
4. Witt H, Luck W, Hennies HC, et al. Mutations in the gene encoding the serine protease inhibitor, Kazal type 1 are associated with chronic pancreatitis. *Nat Genet.* 2000;25:213–216.
5. Rosendahl J, Witt H, Szmola R, et al. Chymotrypsin C (CTRC) variants that diminish activity or secretion are associated with chronic pancreatitis. *Nat Genet.* 2008;40:78–82.
6. Witt H, Beer S, Rosendahl J, et al. Variants in CPA1 are strongly associated with early onset chronic pancreatitis. *Nat Genet.* 2013;45:1216–1220.
7. Masamune A. Genetics of pancreatitis: the 2014 update. *Tohoku J Exp Med.* 2014;232:69–77.
8. Kume K, Masamune A, Mizutamari H, et al. Mutations in the serine protease inhibitor Kazal Type 1 (SPINK1) gene in Japanese patients with pancreatitis. *Pancreatology.* 2005;5:354–360.
9. Masamune A, Nakano E, Kume K, Kakuta Y, Ariga H, Shimosegawa T. Identification of novel missense CTRC variants in Japanese patients with chronic pancreatitis. *Gut.* 2013;62:653–654.
10. Masamune A, Nakano E, Kume K, Takikawa T, Kakuta Y, Shimosegawa T. PRSS1 c.623G>C (p.G208A) variant is associated with pancreatitis in Japan. *Gut.* 2014;63:336.
11. Sharer N, Schwarz M, Malone G, et al. Mutations of the cystic fibrosis gene in patients with chronic pancreatitis. *N Engl J Med.* 1998;339:645–652.
12. Cohn JA, Friedman KJ, Noone PG, Knowles MR, Silverman LM, Jowell PS. Relation between mutations of the cystic fibrosis gene and idiopathic pancreatitis. *N Engl J Med.* 1998;339:653–658.
13. Bombieri C, Claustres M, De Boeck K, et al. Recommendations for the classification of diseases as CFTR-related disorders. *J Cyst Fibros.* 2011;10:S86–S102.
14. Riordan JR, Rommens JM, Kerem B, et al. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science.* 1989;245:1066–1073.
15. Ratjen F, Döring G. Cystic fibrosis. *Lancet.* 2003;361:681–689.
16. Dequeker E, Stuhrmann M, Morris MA, et al. Best practice guidelines for molecular genetic diagnosis of cystic fibrosis and CFTR-related disorders–updated European recommendations. *Eur J Hum Genet.* 2009;17:51–65.
17. Metzker ML. Sequencing technologies-the next generation. *Nat Rev Genet.* 2009;11:31–46.
18. Gilissen C, Hoischen A, Brunner HG, Veltman JA. Unlocking Mendelian disease using exome sequencing. *Genome Biol.* 2011;12:228.
19. Do R, Kathiresan S, Abecasis GR. Exome sequencing and complex disease: practical aspects of rare variant association studies. *Hum Mol Genet.* 2012;21:R1–R9.
20. Myllykangas S, Ji HP. Targeted deep resequencing of the human cancer genome using next-generation technologies. *Biotechnol Genet Eng Rev.* 2010;27:135–158.
21. Berglund EC, Lindqvist CM, Hayat S, et al. Accurate detection of subclonal single nucleotide variants in whole genome amplified and pooled cancer samples using HaloPlex target enrichment. *BMC Genom.* 2013;14:856.
22. Mertes F, Elsharawy A, Sauer S, et al. Targeted enrichment of genomic DNA regions for next-generation sequencing. *Brief Funct Genomics.* 2011;10:374–386.
23. Etemad B, Whitcomb DC. Chronic pancreatitis: diagnosis, classification, and new genetic developments. *Gastroenterology.* 2001;120:682–707.
24. Howes N, Lerch MM, Greenhalf W, et al. European Registry of Hereditary Pancreatitis and Pancreatic Cancer (EUROPAC). Clinical and genetic characteristics of hereditary pancreatitis in Europe. *Clin Gastroenterol Hepatol.* 2004;2:252–261.
25. Nishimori I, Kamakura M, Fujikawa-Adachi K, et al. Mutations in exons 2 and 3 of the cationic trypsinogen gene in Japanese families with hereditary pancreatitis. *Gut.* 1999;44:259–263.
26. Sunyaev S, Ramensky V, Koch I, Lathe W III, Kondrashov AS, Bork P. Prediction of deleterious human alleles. *Hum Mol Genet.* 2001;10:591–597.
27. Dorfman R, Nalpathamkalam T, Taylor C, et al. Do common in silico tools predict the clinical consequences of amino-acid substitutions in the CFTR gene? *Clin Genet.* 2010;77:464–473.
28. Steiner B, Rosendahl J, Witt H, et al. Common CFTR haplotypes and susceptibility to chronic pancreatitis and congenital bilateral absence of the vas deferens. *Hum Mutat.* 2011;32:912–920.
29. Audrézet MP, Chen JM, Le Maréchal C, et al. Determination of the relative contribution of three genes-the cystic fibrosis transmembrane conductance regulator gene, the cationic trypsinogen gene, and the pancreatic secretory trypsin inhibitor gene-to the etiology of idiopathic chronic pancreatitis. *Eur J Hum Genet.* 2002;10:100–106.
30. Noone PG, Zhou Z, Silverman LM, Jowell PS, Knowles MR, Cohn JA. Cystic fibrosis gene mutations and pancreatitis risk: relation to epithelial ion transport and trypsin inhibitor gene mutations. *Gastroenterology.* 2001;121:1310–1319.
31. Schneider A, Larusch J, Sun X, et al. Combined bicarbonate conductance-impairing variants in CFTR and SPINK1 variants are associated with chronic pancreatitis in patients without cystic fibrosis. *Gastroenterology.* 2011;140:162–171.

32. Rosendahl J, Landt O, Bernadova J, et al. CFTR, SPINK1, CTRC and PRSS1 variants in chronic pancreatitis: is the role of mutated CFTR overestimated? *Gut*. 2013;62:582–592.
33. Welsh MJ, Smith AE. Molecular mechanisms of CFTR chloride channel dysfunction in cystic fibrosis. *Cell*. 1993;73:1251–1254.
34. Castellani C, Cuppens H, Macek M Jr, et al. Consensus on the use and interpretation of cystic fibrosis mutation analysis in clinical practice. *J Cyst Fibros*. 2008;7:179–196.
35. Ngiam NS, Chong SS, Shek LP, et al. Cystic fibrosis transmembrane conductance regulator (CFTR) gene mutations in Asians with chronic pulmonary disease: a pilot study. *J Cyst Fibros*. 2006;5:159–164.
36. Watson MS, Cutting GR, Desnick RJ, et al. Cystic fibrosis population carrier screening: 2004 revision of American College of Medical Genetics mutation panel. *Genet Med*. 2004;6:387–391.
37. Yamashiro Y, Shimizu T, Oguchi S, Shioya T, Nagata S, Ohtsuka Y. The estimated incidence of cystic fibrosis in Japan. *J Pediatr Gastroenterol Nutr*. 1997;24:544–547.
38. Fujiki K, Ishiguro H, Ko SB, et al. Genetic evidence for CFTR dysfunction in Japanese: background for chronic pancreatitis. *J Med Genet*. 2004;41:e55.
39. Wang W, Sun XT, Weng XL, et al. Comprehensive screening for PRSS1, SPINK1, CFTR, CTRC and CLDN2 gene mutations in Chinese paediatric patients with idiopathic chronic pancreatitis: a cohort study. *BMJ Open*. 2013;3:e003150.
40. Lee JH, Choi JH, Namkung W, et al. A haplotype-based molecular analysis of CFTR mutations associated with respiratory and pancreatic diseases. *Hum Mol Genet*. 2003;12:2321–2332.
41. Ko S, Zeng W, Fujiki K, et al. Functional characterization of L1156F CFTR: a newly identified mutation in Japanese patients with chronic pancreatitis. *J Physiol Sci*. 2006;56:S71. (abstract).
42. Chen JM, Férec C. Chronic pancreatitis: genetics and pathogenesis. *Annu Rev Genomics Hum Genet*. 2009;10:63–87.

CLINICAL REPORT

AMERICAN JOURNAL OF medical genetics PART A

A Novel Heterozygous *MAP2K1* Mutation in a Patient with Noonan Syndrome with Multiple Lentigines

Eriko Nishi,[1,2] Seiji Mizuno,[3] Yuka Nanjo,[4] Tetsuya Niihori,[4] Yoshimitsu Fukushima,[2] Yoichi Matsubara,[4,5] Yoko Aoki,[4] and Tomoki Kosho[1,2]*

[1]Division of Medical Genetics, Nagano Children's Hospital, Azumino, Japan

[2]Department of Medical Genetics, Shinshu University School of Medicine, Matsumoto, Japan

[3]Department of Pediatrics, Central Hospital, Aichi Human Service Center, Kasugai, Japan

[4]Department of Medical Genetics, Tohoku University School of Medicine, Sendai, Japan

[5]National Research Institute for Child Health and Development, Tokyo

Manuscript Received: 31 March 2014; Manuscript Accepted: 1 October 2014

Noonan syndrome with multiple lentigines (NSML), formerly referred to as LEOPARD syndrome, is a rare autosomal-dominant condition, characterized by multiple lentigines, electrocardiographic conduction abnormalities, ocular hypertelorism, pulmonary stenosis, abnormal genitalia, growth retardation, and sensorineural deafness. To date, *PTPN11*, *RAF1*, and *BRAF* have been reported to be causal for NSML. We report on a 13-year-old Japanese boy, who was diagnosed with NSML. He was found to have a novel heterozygous missense variant (c.305A > G; p.E102G) in *MAP2K1*, a gene mostly causal for cardio-facio-cutaneous syndrome (CFCS). He manifested fetal macrosomia, and showed hypotonia and poor sucking in the neonatal period. He had mild developmental delay, and multiple lentigines appearing at approximately age 3 years, as well as flexion deformity of knees bilaterally, subtle facial characteristics including ocular hypertelorism, sensorineural hearing loss, and precocious puberty. He lacked congenital heart defects or hypertrophic cardiomyopathy, frequently observed in patients with NSML, mostly caused by *PTPN11* mutations. He also lacked congenital heart defects, characteristic facial features, or intellectual disability, frequently observed in those with CFCS caused by *MAP2K1* or *MAP2K2* mutations. This may be the first patient clinically diagnosed with NSML, caused by a mutation in *MAP2K1*.

Key words: Noonan syndrome with multiple lentigines (NSML); *MAP2K1*; cardio-facio-cutaneous syndrome (CFCS)

How to Cite this Article:
Nishi E, Mizuno S, Nanjo Y, Niihori T, Fukushima Y, Matsubara Y, Aoki Y, Kosho T. 2015. A novel heterozygous *MAP2K1* mutation in a patient with Noonan syndrome with multiple lentigines.
Am J Med Genet Part A 167A:407–411.

INTRODUCTION

Noonan syndrome with multiple lentigines (NSML), formerly referred to as LEOPARD syndrome, is a rare autosomal-dominant multiple congenital anomaly condition, characterized by multiple lentigines, electrocardiographic (ECG) abnormalities, ocular hypertelorism, pulmonary stenosis, genital abnormalities, growth retardation, and sensorineural deafness [Sarkozy et al., 2008; Gelb and Tartaglia, 2010; Martínez-Quintana and Rodríguez-González, 2012]. The diagnosis of NSML is made on clinical grounds by observation of specific features. Standard diagnostic criteria for NSML, proposed by Voron et al. [1976]; included multiple lentigines and two other cardinal features.

Together with Noonan syndrome (NS), Costello syndrome, cardio-facio-cutaneous syndrome (CFCS), and neurofibromatosis type 1, NSML is classified as RASopathy, a disorder affecting the RAS-MAPK signal transduction pathway [Aoki and Matsubara, 2013]. NSML is genetically heterogeneous and three causative genes have been identified, accounting for approximately 95% of affected individuals [Martínez-Quintana and Rodríguez-González, 2012]. Approximately 85% of patients with NSML have heterozygous

Conflict of interest: none.
Grant sponsor: Ministry of Health, Labour and Welfare, Japan; Grant sponsor: Nagano Children's Hospital Research Foundation.
*Correspondence to:
Tomoki Kosho, M.D., Department of Medical Genetics, Shinshu University School of Medicine, 3-1-1 Asahi, Matsumoto 390-8621, Japan.
E-mail: ktomoki@shinshu-u.ac.jp
Article first published online in Wiley Online Library (wileyonlinelibrary.com): 25 November 2014
DOI 10.1002/ajmg.a.36842

missense mutations in the protein-tyrosine phosphatase, non-receptor type 11 (*PTPN11*) gene (OMIM#151100). To date, 11 different *PTPN11* mutations, all localized in the protein-tyrosine phosphatase (PTP) domain, have been reported in NSML, two of which (p.T279C and p.T468M) constitute approximately 65% of the cases [Martínez-Quintana and Rodríguez-González, 2012]. Two unrelated patients with NSML were found to have heterozygous missense mutations in the v-Raf-1 murine leukemia viral oncogene homolog 1 (*RAF1*) gene (p.L613V and p.S257L) (OMIM#611554) [Pandit et al., 2007]. The p.L613V mutation increases kinase activity and enhances downstream ERK activation [Pandit et al., 2007]. Two unrelated patients with NSML had heterozygous missense mutations in the v-Raf murine sarcoma viral oncogene homolog B1 (*BRAF*) gene (p.T241P and p.L245F) (OMIM#613707) [Koudova et al., 2009; Sarkozy et al., 2009].

Mitogen-activated protein kinase 1 (MAP2K1) and MAP2K2 are dual-specificity protein kinases, which function as effectors of the serine/threonine kinase *RAF* family members by phosphorylating

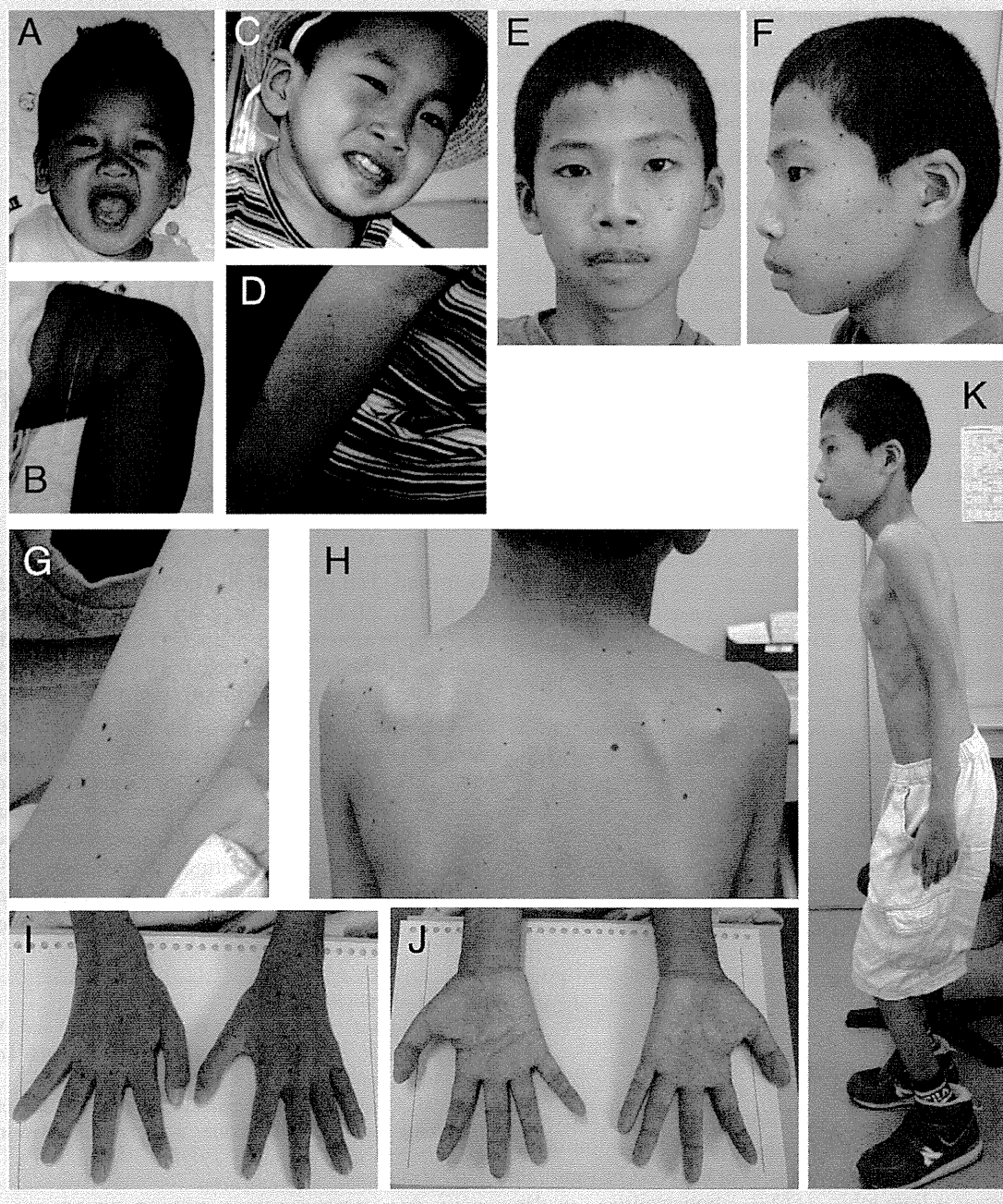

FIG. 1. Clinical photographs of the patient at the age 7 months (A, B), at 2 5/12 years (C, D), and at 11 years (E–K).

and activating ERK proteins. A heterozygous missense mutation in *MAP2K1* is known to be causal for CFCS or NS [Allanson and Roberts, 2011; Rauen, 2012]. To date, all published *MAP2K1* mutations occurred in exons 2, 3, and 6.

In this report, we present a patient clinically diagnosed with NSML, who had a de novo novel and heterozygous *MAP2K1* variant with probable pathogenicity.

CLINICAL REPORT

The patient, a 13-year-old Japanese boy, was the second child of a healthy 30-year-old mother and a healthy 35-year-old nonconsanguineous father. His two brothers were healthy. He was born by normal vaginal delivery at 41 weeks and 4 days of gestation after an uncomplicated pregnancy. His birth weight was 4,350 g (+3.2 SD), length was 51 cm (+1.0 SD), and OFC was 37 cm (+2.6 SD). He showed hypotonia and sucked poorly in the neonatal period. He raised his head at age 3 months, rolled over at 4 months, and sat unsupported at 7 months. He showed no distinctive facial features and only a few lentigines in infancy (Fig. 1A, B).

His growth was impaired with a weight of 8.25 kg (−2.1 SD), height of 76.9 cm (−1.6 SD), and OFC of 45.6 cm (−1.4 SD) at age 1 7/12 years. His weight was 11 kg (−2.5 SD), height was 90.0 cm (−2.4 SD), and OFC was 49 cm (−0.4 SD) at age 2 10/12 years. Lentigines increased on the face and the limbs (Fig. 1C, D). He walked unassisted at age 3 3/12 years, and spoke a two-word sentence at 3 years. His intellectual quotient was 60 at 4 years, and 82 at 7 years. He showed growth acceleration from age 8.5 years, accompanied by a change in voice, and was diagnosed as precocious puberty at 9 years with an advanced bone age of 11.5 years. At age 10 years, his weight was 22.1 kg (−1.5 SD), height was 130 cm (−1.2 SD), and OFC was 51.8 cm (−1.0 SD). He underwent surgical elongation of his hamstrings, which reduced the limitation of bilateral knee extension from −60° degrees to −20° degrees.

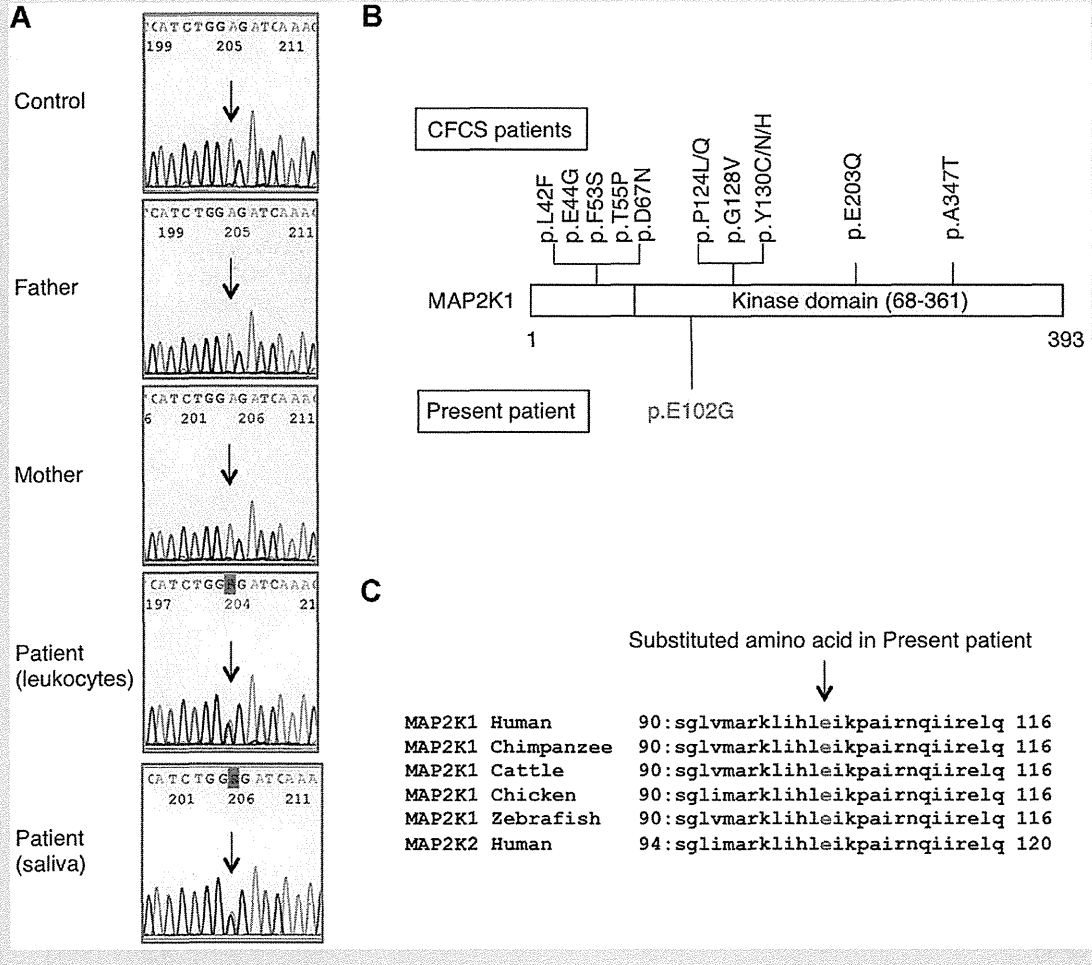

FIG. 2. A: Sanger sequencing of *MAP2K1*, showing an A→G substitution (c.305A > G, p.E102G) in exon 3, which was detected in the patient's DNA from leukocytes and saliva, but not detected in parental samples. B: MAP2K1 domain structure and location of residues altered in the present patient and previously reported patients with cardio-facio-cutaneous syndrome (CFCS). C: MAP2K1 amino acid alignment around the residue where the present amino acid change occurred. This residue is evolutionally conserved.

At age 11 years, his facial features included ocular hypertelorism, a long philtrum, thick upper and lower lip vermilions, and thickened ear helices (Fig. 1E, F). He had hyperextensible and dark skin with multiple lentigines all over the body, several café-au-lait spots, and fine wrinkles on the palms (Fig. 1G–J). He had a slender habitus with pectus carinatum, mild scoliosis, slender extremities, and limited extension of both elbows and knees (Fig. 1K). His weight was 23.0 kg (−1.8 SD) and height was 141 cm (−0.4 SD).

He had no abnormalities in the external genitalia. Resting or 24-hour ECG detected no conduction abnormalities. Echocardiography showed no congenital heart defects, pulmonary valve stenosis, or hypertrophic cardiomyopathy (HCM). Brain magnetic resonance imaging showed no structural abnormalities. He had bilateral mild sensorineural hearing loss with the threshold of 40 dB at approximately 2 kHz. G-banded chromosomes were normal.

MOLECULAR INVESTIGATION

Genomic DNA was isolated from the patient's leukocytes and saliva and his parents' leukocytes after appropriate informed consent. All coding exons and flanking introns in *PTPN11*, *KRAS*, *HRAS*, and *SOS1*, exons 6 and 11–16 in *BRAF*, exons 7, 14, and 17 in *RAF1*, exons 2 and 3 in *MAP2K1/2*, and exon 1 in *SHOC2* were amplified by polymerase chain reaction (PCR) with primers based on GenBank sequences. The primer sequences are available on request. PCR amplification was performed under standard condition using Taq DNA polymerase. After amplification, the PCR products were gel-purified and sequenced on the ABI 3500xL automated DNA sequencer (Applied Biosystems, Carlsbad, CA). A heterozygous missense variant (c.305A > G, p.E102G) was identified in exon 3 of *MAP2K1* in the patient's DNA extracted from his leukocytes and saliva. The variant was not detected in the parental samples (Fig. 2A). No mutation, other than c.305A > G in *MAP2K1*, was identified by the analysis using custom HaloPlex panel (Agilent Technologies, Santa Clara, CA) designed to identify mutations in exons and exon-intron boundaries of the following RASopathy-related genes: *PTPN11*, *HRAS*, *KRAS*, *NRAS*, *BRAF*, *RAF1*, *MAP2K1/2*, *SOS1*, *SHOC2*, *CBL*, *RIT1*, *NF1*, *SPRED1*, and *RRAS*.

DISCUSSION

The present patient had multiple lentigines, café-au-lait spots, ocular hypertelorism, growth impairment, sensorineural hearing loss, hypotonia, low average intelligence, and skeletal abnormalities.

TABLE I. Clinical Features of the Present Patient, Patients With Noonan Syndrome With Multiple Lentigines (NSML), and Patients With Cardio-Facio-Cutaneous Syndrome (CFCS) Caused by *MAP2K1* or *MAP2K2* Mutations

	Present patient	Patients with NSML [Gelb and Tartaglia, 2010]	Patients with CFCS caused by *MAP2K1* or *MAP2K2* mutations [Dentici et al., 2009]	
Causative gene	*MAP2K1*	*PTPN11* (90%) *RAF1* (n = 2) *BRAF* (n = 2)	*MAP2K1* (n = 41)	*MAP2K2* (n = 20)
Sex	Male	Male > Female	Male:Female = 9:14	
Nevi/lentigines	+	<100%	11/34 (32%)	
Café-au-lait spots	+	70–80%	5/30 (17%)	
Congenital heart defects	−	85%	25/39 (64%)	
HCM	−	70%	14/42 (33%)	
ECG abnormalities	−	23%	2/28 (7%)	
Pulmonary valve stenosis	−	25%	17/42 (40%)	
Polyhydramnios	−		20/32 (63%)	
Fetal macrosomia	+		13/25 (52%)	
Short stature	+	<50%	30/38 (79%)	
Macrocephaly	+		26/34 (76%)	
Hypertelorism	+		23/30 (77%)	
Thickened helix	+		27/30 (90%)	
Sparse hair	−		33/49 (67%)	
Sparse eyebrow	−		35/38 (92%)	
Palpebral ptosis	−		18/27 (67%)	
Flat nasal bridge	−		10/12 (83%)	
Joint limitation	+			
Failure to thrive	+		29/35 (83%)	
Intellectual disability	−	30%	43/46 (93%)	
Development delay	+		43/45 (96%)	
Hypotonia	+		40/45 (89%)	
Sensorineural hearing loss	+	<20%		
Seizures	−		16/44 (36%)	

CFCS, cardio-facio-cutaneous syndrome; ECG, electrocardiograph; HCM, hypertrophic cardiomyopathy , NSML, Noonan syndrome with multiple lentigines.

He lacked ECG conduction abnormalities, pulmonary stenosis, or abnormal genitalia. These findings were compatible with the standard diagnosis of NSML by Voron et al. [1976]. The variant c.305A > G, p.E102G was found de novo and not detected in db SNP Release 137 (http://www.ncbi.nlm.nih.gov/projects/SNP/), the Exome Sequencing Project (NHLBI-ESP) database (ESP6500SI-V2) (http://evs.gs.washington.edu/EVS/), the 1000 Genomes Project (1KGP) (http://www.1000genomes.org/), or the Human Gene Mutation Database (http://www.hgmd.cf.ac.uk/ac/index.php). In the COSMIC database, c.302_307delTGGAGA, resulting in an in-frame deletion (p.E102_I103delEI), has been identified in two samples with malignant melanoma and lung cancer (http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/). The glutamine residue at codon 102 is located in the kinase domain (residues 68–361) of *MAP2A1* (Fig. 2B) and is conserved in higher organisms (Fig. 2C). Polymorphism Phenotyping v2 (PolyPhen-2) (http://genetics.bwh.harvard.edu/pph2/) predicts the variant to be possibly damaging, with a score of 0.711. In view of this evidence, the variant p.E102G may be causal for various clinical features consistent with NSML in the patient. However, no functional characterization of the variant was available and Sorting Intolerant From Tolerant (SIFT) (http://sift.jcvi.org) predicts the variant to be tolerated, with a score of 0.09.

We reviewed clinical features of the present patient, previously reported patients with NSML caused by *PTPN11* mutations in most (including two caused by *RAF1* mutations and two caused by *BRAF* mutations), and patients with CFCS caused by *MAP2K1* or *MAP2K2* mutations (Table I) [Pandit et al., 2007; Dentici et al., 2009; Koudova et al., 2009; Sarkozy et al., 2009]. Patients with NSML frequently had congenital heart defects and/or HCM, and sometimes had pulmonary valve stenosis and/or ECG abnormalities [Wakabayashi et al., 2011; Martínez-Quintana and Rodríguez-González, 2012], none of which were found in the present patient. Both patients with NSML caused by *RAF1* mutations had HCM, additionally, one had pulmonary valve stenosis, and the other had a mitral valve anomaly [Pandit et al., 2007]. One of the two patients with NSML caused by *BRAF* mutations had tetralogy of Fallot and the other had mitral and aortic valve dysplasia [Koudova et al., 2009; Sarkozy et al., 2009]. Patients with CFCS caused by *MAP2K1* or *MAP2K2* mutations frequently had congenital heart defects, polyhydramnios, characteristic facial "coarseness" (sparse hair/eyebrows, palpebral ptosis, and flat nasal bridge), and intellectual disability [Dentici et al., 2009], which were not found in the present patient. They rarely or sometimes had nevi, café-au-lait spots, or sensorineural hearing loss [Dentici et al., 2009], which were found in the present patient. Fetal macrosomia, postnatal failure to thrive/growth impairment, macrocephaly, hypotonia, developmental delay, and facial features including hypertelorism and thickened helices were shared by the present patient and over half of the patients with CFCS caused by *MAP2K1* or *MAP2K2* mutations.

In conclusion, the present patient may be the first to fit the standard clinical diagnostic criteria for NSML by Voron et al. [1976]; associated with a *MAP2K1* mutation. He lacked congenital heart defects or HCM, frequently observed in those with NSML, mostly caused by *PTPN11* mutations. He had fetal macrosomia, postnatal failure to thrive/growth impairment, macrocephaly, hypotonia, developmental delay, and hypertelorism but lacked congenital heart defect, characteristic facial features, or intellectual disability; which are frequently observed features in CFCS caused by *MAP2K1* or *MAP2K2* mutations. These observations could offer new insight into the phenotypic spectrum of RASopathies.

ACKNOWLEDGMENTS

We thank the patient and his parents for participating in this study. This work was supported by Research on Intractable Diseases from Ministry of Health, Labour and Welfare, Japan (Y.M., K.T.) and Nagano Children's Hospital Research Foundation (E.N.).

REFERENCES

Allanson JE, Roberts AE. 2011. Noonan syndrome. In: Pagon RA, Adam MP, Bird TD, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington; 1993–2014. Available from: http://www.ncbi.nlm.nih.gov/books/NBK1116/. Accessed March 16, 2014. .

Aoki Y, Matsubara Y. 2013. Ras/MAPK syndromes and childhood hematooncological diseases. Int J Hematology 97:30–36.

Dentici ML, Sarkozy A, Pantaleoni F, Carta C, Lepri F, Ferese R, Cordeddu V, Martinelli S, Briuglia S, Digilio MC, Zampino G, Tartaglia M, Dallapiccola B. 2009. Spectrum of MEK1 and MEK2 gene mutations in cardio-facio-cutaneous syndrome and genetype-phenotype correlations. Eur J Hum Genet 17:733–740.

Gelb BD, Tartaglia M. 2010. LEOPARD syndrome. In: Pagon RA, Adam MP, Bird TD, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington; 1993–2014. Available from: http://www.ncbi.nlm.nih.gov/books/NBK1116/. Accessed March 16, 2014. .

Koudova M, Seemanova E, Zenker M. 2009. Novel BRAF mutation in a patient with LEOPARD syndrome and normal intelligence. Eur J Med Genet. 52:337–340.

Martínez-Quintana E, Rodríguez-González F. 2012. LEOPARD syndrome: Clinical features and gene mutations. Mol Syndromol 3:145–157.

Pandit B, Sarkozy A, Pennacchio LA, Carta C, Oishi K, Martinelli S, Pogna EA, Schackwitz W, Ustaszewska A, Landstrom A, Bos JM, Ommen SR, Esposito G, Lepri F, Faul C, Mundel P, López Siguero JP, Tenconi R, Selicorni A, Rossi C, Mazzanti L, Torrente I, Marino B, Digilio MC, Zampino G, Ackerman MJ, Dallapiccola B, Tartaglia M, Gelb BD. 2007. Gain-of-function RAF1 mutations cause Noonan and LEOPARD syndromes with hypertrophic cardiomyopathy. Nat Genet 39:1007–1012.

Rauen KA. 2012. Cardiofaciocutaneous syndrome. In: Pagon RA, Adam MP, Bird TD, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington; 1993–2014. Available from: http://www.ncbi.nlm.nih.gov/books/NBK1116/. Accessed March 16, 2014. .

Sarkozy A, Digilio MC, Dallapiccola B. 2008. Leopard syndrome. Orphanet J Rare Dis 3:13.

Sarkozy A, Carta C, Moretti S, Zampino G, Digilio MC, Pantaleoni F, Scioletti AP, Esposito G, Cordeddu V, Lepri F, Petrangeli V, Dentici ML, Mancini GM, Selicorni A, Rossi C, Mazzanti L, Marino B, Ferrero GB, Silengo MC, Memo L, Stanzial F, Faravelli F, Stuppia L, Puxeddu E, Gelb BD, Dallapiccola B, Tartaglia M. 2009. Germline BRAF mutations in Noonan, LEOPARD, and cardiofaciocutaneous syndromes: Molecular diversity and associated phenotypic spectrum. Hum Mutat 30:695–702.

Voron DA, Hatfield HH, Kalkhoff RK. 1976. Multiple lentigines syndrome. Case report and review of the literature. Am J Med 60:447–456.

Wakabayashi Y, Yamazaki K, Narumi Y, Fuseya S, Horigome M, Wakui K, Fukushima Y, Matsubara Y, Aoki Y, Kosho T. 2011. Implantable cardioverter defibrillator for progressive hypertrophic cardiomyopathy in a patient with LEOPARD syndrome and a novel PTPN11 mutation Gln510His. Am J Med Genet A 155A:2529–2533.

Hum. Reprod. Advance Access published January 20, 2015
Human Reproduction, Vol.0, No.0 pp. 1–8, 2015
doi:10.1093/humrep/deu364

human reproduction ORIGINAL ARTICLE *Andrology*

Molecular basis of non-syndromic hypospadias: systematic mutation screening and genome-wide copy-number analysis of 62 patients

M. Kon[1,2], E. Suzuki[1], V.C. Dung[3], Y. Hasegawa[4], T. Mitsui[2], K. Muroya[5], K. Ueoka[6], N. Igarashi[7], K. Nagasaki[8], Y. Oto[9], T. Hamajima[10], K. Yoshino[11], M. Igarashi[1], Y. Kato-Fukui[1], K. Nakabayashi[12], K. Hayashi[12], K. Hata[12], Y. Matsubara[13], K. Moriya[2], T. Ogata[1,14], K. Nonomura[2], and M. Fukami[1,*]

[1]Department of Molecular Endocrinology, National Research Institute for Child Health and Development, Tokyo 157-8535, Japan [2]Department of Renal and Genitourinary Surgery, Hokkaido University Graduate School of Medicine, Sapporo 060-8638, Japan [3]Department of Endocrinology, The Vietnam National Hospital of Pediatrics, Hanoi, Vietnam [4]Department of Endocrinology and Metabolism, Tokyo Metropolitan Children's Medical Center, Tokyo 183-8561, Japan [5]Department of Endocrinology and Metabolism, Kanagawa Children's Medical Center, Yokohama 232-0066, Japan [6]Division of Urology, National Center for Child Health and Development, Tokyo 157-8535, Japan [7]Department of Pediatrics, Toyama Prefectural Central Hospital, Toyama 930-0975, Japan [8]Department of Pediatrics, Niigata University School of Medicine, Niigata 951-8510, Japan [9]Department of Pediatrics, Dokkyo Medical University Koshigaya Hospital, Koshigaya 343-0845, Japan [10]Division of Endocrinology and Metabolism, Aichi Children's Health and Medical Center, Obu 474-8710, Japan [11]Department of Urology, Aichi Children's Health and Medical Center, Obu 474-8710, Japan [12]Maternal-Fetal Biology, National Research Institute for Child Health and Development, Tokyo 157-8535, Japan [13]National Research Institute for Child Health and Development, Tokyo 157-8535, Japan [14]Department of Pediatrics, Hamamatsu University School of Medicine, Hamamatsu 431-3192, Japan

*Correspondence address. E-mail: fukami-m@ncchd.go.jp

Submitted on October 1, 2014; resubmitted on December 16, 2014; accepted on December 30, 2014

STUDY QUESTION: What percentage of cases with non-syndromic hypospadias can be ascribed to mutations in known causative/candidate/susceptibility genes or submicroscopic copy-number variations (CNVs) in the genome?

SUMMARY ANSWER: Monogenic and digenic mutations in known causative genes and cryptic CNVs account for >10% of cases with non-syndromic hypospadias. While known susceptibility polymorphisms appear to play a minor role in the development of this condition, further studies are required to validate this observation.

WHAT IS KNOWN ALREADY: Fifteen causative, three candidate, and 14 susceptible genes, and a few submicroscopic CNVs have been implicated in non-syndromic hypospadias.

STUDY DESIGN, SIZE, DURATION: Systematic mutation screening and genome-wide copy-number analysis of 62 patients.

PARTICIPANTS/MATERIALS, SETTING, METHODS: The study group consisted of 57 Japanese and five Vietnamese patients with non-syndromic hypospadias. Systematic mutation screening was performed for 25 known causative/candidate/susceptibility genes using a next-generation sequencer. Functional consequences of nucleotide alterations were assessed by *in silico* assays. The frequencies of polymorphisms in the patient group were compared with those in the male general population. CNVs were analyzed by array-based comparative genomic hybridization and characterized by fluorescence *in situ* hybridization.

MAIN RESULTS AND THE ROLE OF CHANCE: Seven of 62 patients with anterior or posterior hypospadias carried putative pathogenic mutations, such as hemizygous mutations in *AR*, a heterozygous mutation in *BNC2*, and homozygous mutations in *SRD5A2* and *HSD3B2*. Two of the seven patients had mutations in multiple genes. We did not find any rare polymorphisms that were abundant specifically in the patient group. One patient carried mosaic dicentric Y chromosome.

LIMITATIONS, REASONS FOR CAUTION: The patient group consisted solely of Japanese and Vietnamese individuals and clinical and hormonal information of the patients remained rather fragmentary. In addition, mutation analysis focused on protein-altering substitutions.

WIDER IMPLICATIONS OF THE FINDINGS: Our data provide evidence that pathogenic mutations can underlie both mild and severe hypospadias and that *HSD3B2* mutations cause non-syndromic hypospadias as a sole clinical manifestation. Most importantly, this is the first report documenting possible oligogenicity of non-syndromic hypospadias.

STUDY FUNDING/COMPETING INTERESTS: This study was funded by the Grant-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology; by the Grant-in-Aid from the Japan Society for the Promotion of Science; by the Grants from the Ministry of Health, Labour and Welfare, from the National Center for Child Health and Development and from the Takeda Foundation. The authors have no competing interests to disclose.

TRIAL REGISTRATION NUMBER: Not applicable.

Key words: copy-number / hypospadias / mutation / polymorphism / susceptibility

Introduction

Hypospadias is a relatively common form of 46,XY disorders of sex development (DSD) observed in ~4–40 per 10 000 live births (Kurahashi *et al.*, 2004; Nassar *et al.*, 2007; Blaschko *et al.*, 2012). Hypospadias occurs either as an isolated anomaly or as a component of congenital malformation syndromes (Wu *et al.*, 2002; Kurahashi *et al.*, 2004). Although non-syndromic hypospadias is a multifactorial disorder induced by both genetic and environmental factors, this condition can also take place as a result of single gene mutations (Kurahashi *et al.*, 2004; Wang *et al.*, 2004; Chen *et al.*, 2007; Köhler *et al.*, 2009). Previous studies revealed familial aggregation of non-syndromic hypospadias (Schnack *et al.*, 2008; van Rooij et al., 2013). In most cases, familial hypospadias is equally transmitted from the paternal and maternal sides of the family and shows similar recurrence risks between the brothers and sons of patients, indicating a significant role of single gene mutations in the development of the disease (Schnack *et al.*, 2008).

In 2012, van der Zanden *et al.* (2012) reviewed 162 prior studies and listed 15 causative genes and three candidate genes for this condition. They also introduced 49 polymorphisms in 13 genes associated with disease risk, together with one susceptibility gene *CYP1A1* whose risk allele is yet to be determined. To date, however, there is no single report of systematic mutation analysis of the causative/candidate/susceptible genes. Likewise, while a small number of submicroscopic copy-number variations (CNVs) have been identified in patients with non-syndromic hypospadias (Tannour-Louet *et al.*, 2010), genome-wide copy-number analysis has been performed only in exceptional cases. Thus, the contribution of single gene mutations and submicroscopic CNVs to the etiology of non-syndromic hypospadias remains unknown.

The aim of this study was to clarify the frequency and type of genetic defects in patients with non-syndromic hypospadias. This study consisted of systematic mutation screening using next-generation sequencing (NGS) technology and cytogenetic analyses using comparative genomic hybridization (CGH) and fluorescence *in situ* hybridization (FISH).

Materials and Methods

Patients

A total of 57 Japanese and 5 Vietnamese patients with hypospadias participated in the study (Table I). All patients were referred to our clinics because of hypospadias. Patients with additional clinical features except for cryptorchidism and micropenis and those with cytogenetically detectable chromosomal abnormalities were excluded from this study. The 62 patients had no family history of 46,XY DSD. One of the 62 patients (case 18) was born to consanguineous parents. Hospital records of genital features at birth were obtained for 49 patients. Eleven patients manifested relatively mild hypospadias with the urethral opening at the anterior portion of the penis, while 14 and 24 patients presented with moderate (middle) and severe (posterior) hypospadias, respectively. Cryptorchidism and micropenis were observed in 5 and 11 patients, respectively.

Ethical approval

This study was approved by the Institutional Review Board Committee at the National Center for Child Health and Development and performed after obtaining written informed consent from the parents of patients.

Identification of nucleotide substitutions

Sequence analysis was carried out for 25 known causative/candidate/susceptible genes for non-syndromic hypospadias, i.e. *AR*, *ATF3*, *BMP4*, *BMP7*, *BNC2*, *CTGF*, *CYP1A1*, *CYR61*, *DGKK*, *EGF*, *ESR1*, *ESR2*, *FGF8*, *FGFR2*, *GSTM1*, *GSTT1*, *HOXA4*, *HOXB6*, *HSD3B2*, *HSD17B3*, *MAMLD1*, *MID1*, *NR5A1* (alias *SF1*), *SRD5A2*, and *WT1* (van der Zanden *et al.*, 2012). The coding regions of these genes were amplified from genomic DNA using the Haloplex Target Enrichment System (Design ID 02185-1348467147) (Agilent Technologies, Palo Alto, CA, USA), and were sequenced as 150 bp paired-end reads on a MiSeq sequencer (Illumina, San Diego, CA, USA). The average read depth of each amplicon was 115.0. Subsequently, nucleotide alterations in the samples were called by the Surecall system (Agilent Technologies) and SAMtools 0.1.17 software (http://samtools.sourceforge.net, 12 January 2015, date last accessed) (Li *et al.*, 2009). In the present study, we focused on non-synonymous substitutions in the coding regions and nucleotide changes at splice sites. Substitutions detected by NGS were confirmed by Sanger direct sequencing. The primers utilized in the present study are available upon request.

Characterization of nucleotide substitutions

Functional consequences of nucleotide alterations were predicted by *in silico* analyses. Single nucleotide polymorphisms (SNPs) with allele frequencies of >1.0% in the general population (dbSNP, http://www.ncbi.nlm.nih.gov/, 12 January 2015, date last accessed), except for those that have been reported as risk alleles (van der Zanden *et al.*, 2012), were excluded from further analyses. The effects of missense substitutions on protein function were predicted using Polyphen2 (http://genetics.bwh.harvard.edu/pph2/, 12 January 2015, date last accessed) (Adzhubei *et al.*, 2010), and those of intronic substitutions on splicing were assessed using Genome Project

Table I Nucleotide alterations identified in the present study.

Case[a]	Ethnic origin	Putative pathogenic mutation	Putative risk variant	Probable benign change	Copy-number alteration	Position of urethral opening[b]	Cryptorchidism	Micropenis
1	J	***AR* (p.S176R)**				Anterior	No	No
2	J	***AR* (p.A403V)**				No data	No data	No data
3	J	***AR* (p.R841S)**	*HSD17B3* (p.G289S)			Posterior	No	Yes
4	J	***AR* (delins[c])** *HOXB6* (p.S2N)	***MAMLD1* (p.N662S)**			No data	No data	No data
5	J	*BNC2* (p.M801R)				Posterior	No	No
6	V	***SRD5A2* (p.R227Q)**[d]	*HSD17B3* (p.G289S)			Posterior	No	Yes
7	V	***HSD3B2* (p.A10T)**	*SRD5A2* (p.R227Q)[d]			Posterior	Yes (right)	Yes
8	J		*HSD17B3* (p.G289S)	*CYP1A1* (p.T173R)	Y chromosome[e]	Posterior	No	No
9	J		***MAMLD1* (p.N662S)**			Anterior	No	No
10	J		*CYP1A1* (p.Q75P)			Middle	No data	No data
11	J		*CYP1A1* (p.A62P)			Middle	No	No
12	J		*BMP7* (p.T170M)			Middle	No	No
13	V		*HSD17B3* (p.G289S)			No data	No	No
14	J		*HSD17B3* (p.G289S)			Posterior	No	No
15	J		***HSD17B3* (p.G289S)**			Posterior	Yes	No
16	J		*HSD17B3* (p.G289S)			Posterior	No	No
17	J		*HSD17B3* (p.G289S)			Posterior	Yes (right)	Yes
18	J		*HSD17B3* (p.G289S)			Posterior	No	No
19	J		*HSD17B3* (p.G289S)			Middle	Yes (right)	Yes
20	J		*HSD17B3* (p.G289S)			Middle	No data	No data
21	J		***HSD17B3* (p.G289S)**			Middle	No	Yes
22	J		*HSD17B3* (p.G289S)			Anterior	No	No
23	J		***HSD17B3* (p.G289S)**			No data	No data	No data
24	J		*HSD17B3* (p.G289S)			No data	No data	No data
25	J		*HSD17B3* (p.G289S)			No data	No data	No data
26	J		*HSD17B3* (p.G289S) ***MAMLD1* (p.N662S)**			Middle	No	No
27	J		*HSD17B3* (p.G289S)	*BNC2* (p.M539V)		No data	No data	No data
28	J		*HSD17B3* (p.G289S)	*BNC2* (p.P614S)		No data	No data	No data
29	J		***MAMLD1* (p.N662S)**	*EGF* (p.S16R)		Posterior	No	No
30	J		*HSD17B3* (p.G289S)	*FGFR2* (p.M97V)		Anterior	No data	No data
31	J		*HSD17B3* (p.G289S)	*EGF* (p.S16R)		Middle	No	No
32	J		***MAMLD1* (p.N662S)**	*HSD3B2* (p.S284I) *EGF* (p.S16R)		Posterior	No	No

Continued

Table I *Continued*

Case[a]	Ethnic origin	Putative pathogenic mutation	Putative risk variant	Probable benign change	Copy-number alteration	Position of urethral opening[b]	Cryptorchidism	Micropenis
33	J		*HSD17B3* (p.G289S)	*HSD3B2* (p.R362W)		Anterior	No data	No data
34	J			*NR5A1* (g.IVS2-5G>A)		Posterior	No data	No data
35	J			*HOXB6* (p.P40S)		Posterior	No	Yes
36	J			***MAMLD1* (p.N675K)**		Posterior	No data	No data
37	J			*ESR2* (p.G67S)		Posterior	No	No
38	J			*EGF* (p.S16R) *BNC2* (p.I974V)		Middle	No data	No data

J, Japanese; V, Vietnamese.
Homozygous or hemizygous mutations/variations are boldfaced, and heterozygous substitutions are lightfaced.
[a]Cases 39–62 carried no nucleotide alterations in the target genes.
[b]Detailed clinical information was obtained only from 49 of the 62 patients.
[c]c.1995delTGAAGGCTATGAATGTCinsCAGAA; p.666delEGYECQinsRK.
[d]Homozygosity and heterozygosity of this mutation were described as a pathogenic defect and a disease-susceptible alteration, respectively.
[e]Copy-number gain of the region from Ypter to Yq11.223 and copy-number loss of the remaining Y chromosomal region.

Data (http://www.fruitfly.org/seq_tools/splice.html, 12 January 2015, date last accessed) (Reese *et al.*, 1997). Nucleotide deletions and insertions in the coding regions were assessed as 'probably damaging'.

Nucleotide alterations were classified into the following three groups: (i) putative pathogenic mutations: mutations that have been associated with 46,XY DSD or hitherto unreported nucleotide changes in causative genes that were assessed as 'probably damaging' or 'possibly damaging' by *in silico* analyses; (ii) putative risk variants: previously reported risk SNPs or novel substitutions in susceptibility genes, or rare SNPs in causative genes that were assessed as 'probably damaging' or 'possibly damaging'; and (iii) probable benign changes: nucleotide substitutions in causative/susceptible/candidate genes that were assessed as 'benign'. To determine the possible association between the SNPs (putative risk variants and probable benign changes) and disease risk, we compared allele frequencies in the patient group with those in the male general population. In the SNP analysis, we focused on Japanese patients, for whom the allele frequencies in the general population were available in the public database (dbSNP, http://www.ncbi.nlm.nih.gov/, 12 January 2015, date last accessed).

Statistical analysis

The statistical significance of the comparison of allele frequency in the patient group and the general population was evaluated using χ^2 and Fisher's exact probability tests.

Copy-number analyses

CNVs in the genome were screened by CGH using a catalog human array (8 × 60 k format, catalog number G4450A, Agilent Technologies), according to the manufacturers' instructions. In this study, we focused on copy-number alterations affecting genomic intervals larger than 1.5 Mb, which have a higher probability of being associated with disease phenotypes (Cooper *et al.*, 2011). We referred to the Database of Genomic Variants (http://projects.tcag.ca/variation/, 12 January 2015, date last accessed) to exclude known benign variants. Genomic structures of CNVs were characterized by FISH analysis.

Results

Identification and characterization of nucleotide substitutions

Eight putative pathogenic mutations were identified in seven patients (Table I and Fig. 1). The eight mutations consisted of three hemizygous missense mutations and one hemizygous deletion/insertion in *AR*, one heterozygous missense mutation in *HOXB6*, one heterozygous missense mutation in *BNC2*, and apparent homozygous mutations in *SRD5A2* and *HSD3B2*. Of these, the *AR* mutation in case 3 and the *SRD5A2* mutation in case 6 were previously identified in patients with 46,XY DSD (Melo *et al.*, 2003 in which the p.R841S mutation in *AR* was described as p.R840S; Sasaki *et al.*, 2003; van der Zanden *et al.*, 2012), while the other mutations were first identified in the present study.

Putative risk variants were identified in 30 patients (Table I and Supplementary Table SI). These variants included three known risk alleles for hypospadias and/or micropenis: rs2066476 in *HSD17B3*, rs2073043 in *MAMLD1* and rs9332964 in *SRD5A2* (Sasaki *et al.*, 2003; Fukami *et al.*, 2008; Sata *et al.*, 2010; Kalfa *et al.*, 2011; van der Zanden *et al.*, 2012). The SNPs in *HSD17B3* and *MAMLD1* were identified in the Japanese patient group and the male general population at similar frequencies. We also identified a rare SNP in the causative gene *CYP1A1* which was shared by the Japanese patients and the male

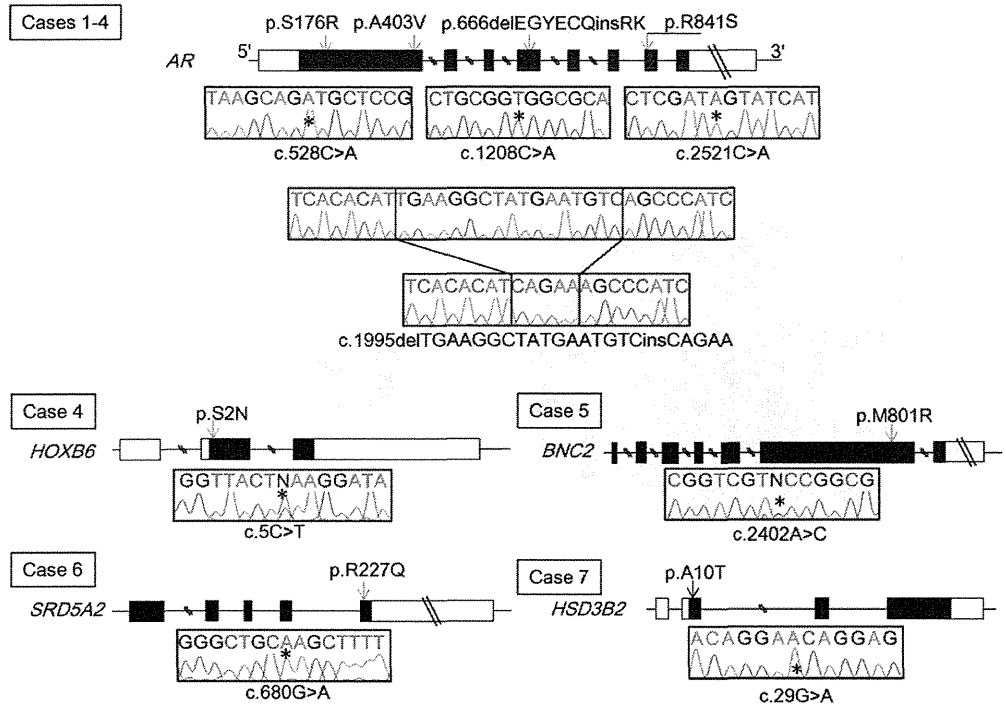

Figure 1 Putative pathogenic mutations identified in the present study. Genomic positions and chromatograms of the nucleotide substitutions are shown. Asterisks indicate the mutated nucleotides.

general population at a similar frequency, together with a SNP in *BMP4* whose frequency in the general population is unknown.

Probable benign changes were found in 13 patients (Table I and Supplementary Table SI). These substitutions included a rare SNP in *EGF* which was identified in the patient group and in the general population at similar frequency. We also detected SNPs in *ESR2* and *BNC2* that had unknown frequencies in the general population, together with a novel substitution in intron 2 of *NR5A1* (g.IVS2-5G>A) that was predicted to not affect splicing.

Copy-number analyses

One of the 62 patients (case 8) carried CNVs on the Y chromosome (Fig. 2A). These alterations consisted of copy-number gain of a ~23 Mb region from Ypter to Yq11.223 and copy-number loss of the remaining Y chromosomal region. The log2 signal ratios of most probes corresponding to the amplified and deleted regions were lower than +1.0 and higher than −2.0 respectively, indicating mosaicism of these CNVs. FISH analysis using a *SRY*-containing probe showed that case 8 had mosaic dicentric Y (Fig. 2B). CGH analysis for case 6 with an apparently homozygous *SRD5A2* mutation and case 7 with an apparently homozygous *HSD3B2* mutation excluded compound heterozygosity for a mutation and deletion (data not shown).

Clinical findings of patients with putative pathogenic defects

Putative pathogenic defects were associated with both anterior and posterior hypospadias (Table II). Endocrine evaluation of cases 1–8 remained fragmentary; blood hormone levels in cases 3 and 7 were within the normal range (Table II).

Discussion

Systematic mutation screening identified putative pathogenic mutations in 7 of 62 patients with non-syndromic hypospadias. These results, in conjunction with previous studies showing that ~30% of cases with severe hypospadias are ascribable to specific defects such as mutations in *AR* or *SRD5A2* (Albers *et al.*, 1997; Boehmer *et al.*, 2001), demonstrate the significant role of mutations in known causative genes in the etiology of non-syndromic hypospadias. Furthermore, our results support the previously proposed notion that genetic defects in *AR* account for a substantial percentage of cases with various types of 46,XY DSD (Albers *et al.*, 1997; Boehmer *et al.*, 2001; Audi *et al.*, 2010) and that mutations in *HSD3B2* can lead to non-syndromic hypospadias as a sole clinical manifestation, although *HSD3B2* plays an essential role in adrenal function (Boehmer *et al.*, 2001; Codner *et al.*, 2004; Audi *et al.*, 2010). Case 3 carried the p.R841S mutation in *AR*, which have been identified in patients with ambiguous genitalia (Melo *et al.*, 2003), suggesting the phenotypic diversity of missense mutations in *AR*. Notably, two of our patients had putative pathogenic mutations in multiple genes. Case 4 carried a hemizygous in-frame deletion/insertion in *AR* and a heterozygous missense substitution in *HOXB6*. Likewise, case 7 with a homozygous missense mutation in *HSD3B2* had an additional heterozygous missense mutation in *SRD5A2* that retains 3% of enzymatic activity (Makridakis *et al.*, 2000; Sasaki *et al.*, 2003). These data imply for the

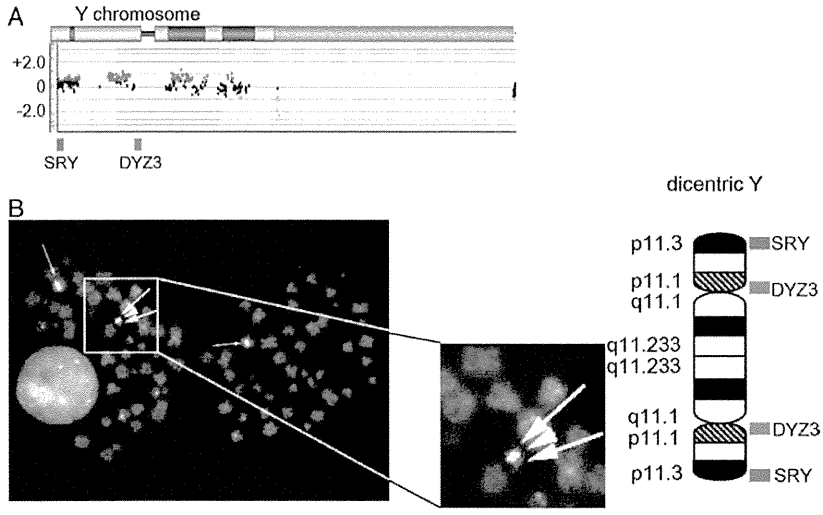

Figure 2 Copy-number alterations identified in case 8. Results of array-based comparative genomic hybridization (CGH) (**A**) and fluorescence *in situ* hybridization (FISH) analysis and schematic representation of the dicentric Y chromosome (**B**) are shown. The black, red and green dots in CGH denote signals indicative of the normal, increased (> +0.4) and decreased (< −0.8) copy-numbers, respectively. The arrowhead and thick arrows in FISH indicate a signal of DYZ3 (Y centromeric probe) and signals of *SRY*-containing probe (Yp11.3), respectively. The thin arrows in the left panel indicate signals of X centromeric probe.

Table II Molecular and clinical findings of patients with putative pathogenic abnormalities.

	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8
Affected gene/region	*AR*	*AR*	*AR*	*AR/HOXB6*	*BNC2*	*SRD5A2*	*HSD3B2*	Y chromosome[a]
Ethnic origin	Japanese	Japanese	Japanese	Japanese	Japanese	Vietnamese	Vietnamese	Japanese
Family history of DSD	No	No data	No	No data	No	No	No	No
Clinical features								
Hypospadias[b]	Anterior	No data	Posterior	No data	Posterior	Posterior	Posterior	Posterior
Cryptorchidism	No	No data	No	No data	No	No	Yes (right)	No
Micropenis	No	No data	Yes	No data	No	Yes	Yes	No
Other features	No	No data	No	No data	No	No	No	Borderline MR
Endocrine findings								
Age at examination	No data	No data	15 months	No data	No data	No data	3.5 years	No data
LH (IU/l)[c]	No data	No data	<0.2 (<0.2–0.3)	No data	No data	No data	No data	No data
FSH (IU/l)[c]	No data	No data	<1.0 (<1.0–1.5)	No data	No data	No data	No data	No data
Testosterone (nmol/l)[c]	No data	No data	0.17 (0.10–0.45)	No data	No data	No data	0.16 (0.10–0.45)	No data

DSD, disorders of sex development; MR, mental retardation; LH, luteinizing hormone; FSH, follicle stimulating hormone.
[a]Copy-number alterations on Y chromosome.
[b]Position of urethral opening.
[c]Hormone values in parentheses indicate the reference ranges of age- and sex-matched control individuals.

first time that non-syndromic hypospadias results from digenic mutations. On the other hand, we did not observe the accumulation of rare SNPs in the patient group. Our data suggest that previously reported susceptibility SNPs play no or only minor roles in the development of non-syndromic hypospadias in the Japanese population. However, we cannot exclude the possibility that oligogenicity of these SNPs increases the risk of the disease, because a small number of our patients carried these SNPs as biallelic or digenic substitutions. Considering the small number of participants of this study, further investigations are necessary to clarify the possible association between rare SNPs and the disease phenotype.

Genome-wide copy-number analysis identified cryptic CNVs only in one patient. Case 8 carried a copy-number gain of a ~23 Mb region

on Yp and Yq and copy-number loss of the remaining Y chromosomal region. FISH analysis revealed that case 8 had mosaic dicentric Y, which has been described in multiple patients with hypospadias (Drummond-Borg *et al.*, 1988; Kojima *et al.*, 2001). It has been proposed that dicentric Y results in hypospadias by mosaic loss of the rearranged Y chromosome or by aberrant expression of Y chromosomal genes (Drummond-Borg *et al.*, 1988; Kojima *et al.*, 2001). The lack of pathogenic CNVs in the remaining 61 cases suggests the rarity of cryptic CNVs as genetic causes of non-syndromic hypospadias.

In this study, putative pathogenic defects were identified predominantly in patients with severe (posterior) hypospadias, while an *AR* mutation was detected in case 1, who manifested mild (anterior) hypospadias without micropenis or cryptorchidism. In this regard, previous studies have shown that syndromic hypospadias often arises from known gene mutations or chromosomal rearrangements (van der Zanden *et al.*, 2012). These data imply that monogenic mutations can underlie various types of hypospadias, although they are more strongly associated with severe or syndromic hypospadias than with mild non-syndromic hypospadias. Since identification of pathogenic defects can help to predict disease outcomes and improves the accuracy of genetic counseling, genetic analyses should be considered in patients with hypospadias of various clinical severities.

It should be pointed out that the present study has some limitations. First, the patient group consisted of only Japanese and Vietnamese individuals. Since the prevalence of hypospadias varies among countries (Nassar *et al.*, 2007; Serrano *et al.*, 2013), there may be ethnicity-specific causes of hypospadias. For example, mutations in *ATF3*, which account for ~10% of cases in the USA (Kalfa *et al.*, 2008), were absent from our cohort. In contrast, the p.A10T mutation in *HSD3B2* and the p.R227Q mutation in *SRD5A2* were detected exclusively in Vietnamese patients in homozygous state. Thus, our results are not simply applicable to other ethnic groups. Second, the frequency of monogenic defects may be underestimated in this study, because we focused on protein-altering mutations in 25 genes. Mutations/variations in regulatory regions, defects in unexamined genes and epigenetic abnormalities may be hidden in our mutation-negative patients. Lastly, clinical information of our patients remained fragmentary. Although previous studies have revealed that several factors such as low birthweight, placental insufficiency and maternal hypertension are associated with the risk of hypospadias (Stoll *et al.*, 1990; Weidner *et al.*, 1999; Fredell *et al.*, 2002; Brouwers *et al.*, 2010), the contributions of such factors to the disease phenotype of our patients are yet to be studied. Moreover, since endocrine data were unavailable for most of our mutation-positive cases, further studies are needed to elucidate the hormonal characteristics of each monogenic disorder.

Conclusion

The present study indicates that mutations in known causative genes and submicroscopic CNVs account for >10% of cases with non-syndromic hypospadias. Pathogenic defects appear to underlie both severe and mild hypospadias. On the other hand, previously reported risk SNPs are unlikely to play a major role in the development of the disease; further studies are required to validate this observation. Most importantly, this is the first report documenting the possible oligogenicity of non-syndromic hypospadias.

Supplementary data

Supplementary data are available at http://humrep.oxfordjournals.org/.

Authors' roles

M.K., K.No., T.O., and M.F. designed the study. M.K., E.S., V.C.D., Y.H., T.M., K.Mu., K.U., N.I., K.Nag., Y.O., T.H., K.Y., M.I, Y.K.-F., K.Nak., K.Hay., K.Hat., Y.M., K.Mo., and T.O. contributed to the acquisition of data. M.K. and M.F. analyzed data and wrote the paper. All authors were involved in revising the paper and approved the final version of the manuscript for submission.

Funding

This study was funded by the Grant-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology; by the Grant-in-Aid from the Japan Society for the Promotion of Science; by the Grants from the Ministry of Health, Labour and Welfare, from the National Center for Child Health and Development and from the Takeda Foundation.

Conflict of interest

The authors declare that there is no conflict of interest.

References

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;**7**:248–249.

Albers N, Ulrichs C, Glüer S, Hiort O, Sinnecker GH, Mildenberger H, Brodehl J. Etiologic classification of severe hypospadias: implications for prognosis and management. *J Pediatr* 1997;**131**:386–392.

Audi L, Fernández-Cancio M, Carrascosa A, Andaluz P, Torán N, Piró C, Vilaró E, Vicens-Calvet E, Gussinyé M, Albisu MA *et al.* Novel (60%) and recurrent (40%) androgen receptor gene mutations in a series of 59 patients with a 46,XY disorder of sex development. *J Clin Endocrinol Metab* 2010;**95**:1876–1888.

Blaschko SD, Cunha GR, Baskin LS. Molecular mechanisms of external genitalia development. *Differentiation* 2012;**84**:261–268.

Boehmer AL, Nijman RJ, Lammers BA, de Coninck SJ, Van Hemel JO, Themmen AP, Mureau MA, de Jong FH, Brinkmann AO, Niermeijer MF *et al.* Etiological studies of severe or familial hypospadias. *J Urol* 2001; **165**:1246–1254.

Brouwers MM, van der Zanden LF, de Gier RP, Barten EJ, Zielhuis GA, Feitz WF, Roeleveld N. Hypospadias: risk factor patterns and different phenotypes. *BJU Int* 2010;**105**:254–262.

Chen T, Li Q, Xu J, Ding K, Wang Y, Wang W, Li S, Shen Y. Mutation screening of BMP4, BMP7, HOXA4 and HOXB6 genes in Chinese patients with hypospadias. *Eur J Hum Genet* 2007;**15**:23–28.

Codner E, Okuma C, Iñiguez G, Boric MA, Avila A, Johnson MC, Cassorla FG. Molecular study of the 3 beta-hydroxysteroid dehydrogenase gene type II in patients with hypospadias. *J Clin Endocrinol Metab* 2004;**89**:957–964.

Cooper GM, Coe BP, Girirajan S, Rosenfeld JA, Vu TH, Baker C, Williams C, Stalker H, Hamid R, Hannig V *et al.* A copy number variation morbidity map of developmental delay. *Nat Genet* 2011;**43**:838–846.

Drummond-Borg M, Pagon RA, Bradley CM, Nordlund J, Salk D. Nonfluorescent dicentric Y in males with hypospadias. *J Pediatr* 1988; **113**:469–473.

Fredell L, Kockum I, Hansson E, Holmner S, Lundquist L, Lackgren G, Pedersen J, Stenberg A, Westbacke G, Nordenskjöld A. Heredity of hypospadias and the significance of low birth weight. *J Urol* 2002;**167**:1423–1427.

Fukami M, Wada Y, Okada M, Kato F, Katsumata N, Baba T, Morohashi K, Laporte J, Kitagawa M, Ogata T. Mastermind-like domain-containing 1 (MAMLD1 or CXorf6) transactivates the Hes3 promoter, augments testosterone production, and contains the SF1 target sequence. *J Biol Chem* 2008;**283**:5525–5532.

Kalfa N, Liu B, Klein O, Wang MH, Cao M, Baskin LS. Genomic variants of ATF3 in patients with hypospadias. *J Urol* 2008;**180**:2183–2188.

Kalfa N, Cassorla F, Audran F, Oulad Abdennabi I, Philibert P, Béroud C, Guys JM, Reynaud R, Alessandrini P, Wagner K *et al.* Polymorphisms of MAMLD1 gene in hypospadias. *J Pediatr Urol* 2011;**7**:585–591.

Köhler B, Lin L, Mazen I, Cetindag C, Biebermann H, Akkurt I, Rossi R, Hiort O, Gruters A, Achermann JC. The spectrum of phenotypes associated with mutations in steroidogenic factor 1 (SF-1, NR5A1, Ad4BP) includes severe penoscrotal hypospadias in 46,XY males without adrenal insufficiency. *Eur J Endocrinol* 2009;**161**:237–242.

Kojima Y, Hayashi Y, Yanai Y, Tozawa K, Sasaki S, Kohri K. Molecular analysis of hypospadias in a boy with dicentric Y chromosome. *J Urol* 2001; **165**:1244–1245.

Kurahashi N, Murakumo M, Kakizaki H, Nonomura K, Koyanagi T, Kasai S, Sata F, Kishi R. The estimated prevalence of hypospadias in Hokkaido, Japan. *J Epidemiol* 2004;**14**:73–77.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009;**25**:2078–2079.

Makridakis NM, di Salle E, Reichardt JK. Biochemical and pharmacogenetic dissection of human steroid 5 alpha-reductase type II. *Pharmacogenetics* 2000;**10**:407–413.

Melo KF, Mendonca BB, Billerbeck AE, Costa EM, Inácio M, Silva FA, Leal AM, Latronico AC, Arnhold IJ. Clinical, hormonal, behavioral, and genetic characteristics of androgen insensitivity syndrome in a Brazilian cohort: five novel mutations in the androgen receptor gene. *J Clin Endocrinol Metab* 2003;**88**:3241–3250.

Nassar N, Bower C, Barker A. Increasing prevalence of hypospadias in Western Australia, 1980–2000. *Arch Dis Child* 2007;**92**:580–584.

Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection in Genie. *J Comput Biol* 1997;**4**:311–323.

Sasaki G, Ogata T, Ishii T, Kosaki K, Sato S, Homma K, Takahashi T, Hasegawa T, Matsuo N. Micropenis and the 5alpha-reductase-2 (SRD5A2) gene: mutation and V89L polymorphism analysis in 81 Japanese patients. *J Clin Endocrinol Metab* 2003;**88**:3431–3436.

Sata F, Kurahashi N, Ban S, Moriya K, Tanaka KD, Ishizuka M, Nakao H, Yahata Y, Imai H, Kakizaki H *et al.* Genetic polymorphisms of 17 beta-hydroxysteroid dehydrogenase 3 and the risk of hypospadias. *J Sex Med* 2010;**7**:2729–2738.

Schnack TH, Zdravkovic S, Myrup C, Westergaard T, Christensen K, Wohlfahrt J, Melbye M. Familial aggregation of hypospadias: a cohort study. *Am J Epidemiol* 2008;**167**:251–256.

Serrano T, Chevrier C, Multigner L, Cordier S, Jégou B. International geographic correlation study of the prevalence of disorders of male reproductive health. *Hum Reprod* 2013;**28**:1974–1986.

Stoll C, Alembik Y, Roth MP, Dott B. Genetic and environmental factors in hypospadias. *J Med Genet* 1990;**27**:559–563.

Tannour-Louet M, Han S, Corbett ST, Louet JF, Yatsenko S, Meyers L, Shaw CA, Kang SH, Cheung SW, Lamb DJ. Identification of *de novo* copy number variants associated with human disorders of sexual development. *PLoS One* 2010;**5**:e15392.

van der Zanden LF, van Rooij IA, Feitz WF, Franke B, Knoers NV, Roeleveld N. Aetiology of hypospadias: a systematic review of genes and environment. *Hum Reprod Update* 2012;**18**:260–283.

van Rooij IA, van der Zanden LF, Brouwers MM, Knoers NV, Feitz WF, Roeleveld N. Risk factors for different phenotypes of hypospadias: results from a Dutch case-control study. *BJU Int* 2013;**112**:121–128.

Wang Y, Li Q, Xu J, Liu Q, Wang W, Lin Y, Ma F, Chen T, Li S, Shen Y. Mutation analysis of five candidate genes in Chinese patients with hypospadias. *Eur J Hum Genet* 2004;**12**:706–712.

Weidner IS, Møller H, Jensen TK, Skakkebaek NE. Risk factors for cryptorchidism and hypospadias. *J Urol* 1999;**161**:1606–1609.

Wu WH, Chuang JH, Ting YC, Lee SY, Hsieh CS. Developmental anomalies and disabilities associated with hypospadias. *J Urol* 2002; **168**:229–232.

© 2014. Published by The Company of Biologists Ltd | Development (2014) 141, 91-100 doi:10.1242/dev.103168

RESEARCH ARTICLE

STEM CELLS AND REGENERATION

CCAAT/enhancer binding protein-mediated regulation of TGFβ receptor 2 expression determines the hepatoblast fate decision

Kazuo Takayama[1,2,3], Kenji Kawabata[4], Yasuhito Nagamoto[1,2], Mitsuru Inamura[1], Kazuo Ohashi[5], Hiroko Okuno[1], Tomoko Yamaguchi[4], Katsuhisa Tashiro[4], Fuminori Sakurai[1], Takao Hayakawa[6], Teruo Okano[5], Miho Kusada Furue[7,8] and Hiroyuki Mizuguchi[1,2,3,9,*]

ABSTRACT

Human embryonic stem cells (hESCs) and their derivatives are expected to be used in drug discovery, regenerative medicine and the study of human embryogenesis. Because hepatocyte differentiation from hESCs has the potential to recapitulate human liver development *in vivo*, we employed this differentiation method to investigate the molecular mechanisms underlying human hepatocyte differentiation. A previous study has shown that a gradient of transforming growth factor beta (TGFβ) signaling is required to segregate hepatocyte and cholangiocyte lineages from hepatoblasts. Although CCAAT/enhancer binding proteins (c/EBPs) are known to be important transcription factors in liver development, the relationship between TGFβ signaling and c/EBP-mediated transcriptional regulation in the hepatoblast fate decision is not well known. To clarify this relationship, we examined whether c/EBPs could determine the hepatoblast fate decision via regulation of TGFβ receptor 2 (TGFBR2) expression in the hepatoblast-like cells differentiated from hESCs. We found that *TGFBR2* promoter activity was negatively regulated by c/EBPα and positively regulated by c/EBPβ. Moreover, c/EBPα overexpression could promote hepatocyte differentiation by suppressing TGFBR2 expression, whereas c/EBPβ overexpression could promote cholangiocyte differentiation by enhancing TGFBR2 expression. Our findings demonstrated that c/EBPα and c/EBPβ determine the lineage commitment of hepatoblasts by negatively and positively regulating the expression of a common target gene, *TGFBR2*, respectively.

KEY WORDS: Hepatoblasts, c/EBP, CEBP, Human ESCs

INTRODUCTION

Many animal models, such as chick, *Xenopus*, zebrafish and mouse, have been used to investigate the molecular mechanisms of liver development. Because many functions of the key molecules in liver development are conserved in these species, studies on liver development in these animals can be highly informative with respect that in humans. However, some functions of important molecules in liver development might differ between human and other species. Although analysis using genetically modified mice has been successfully performed, it is not of course possible to perform genetic experiments to elucidate molecular mechanisms of liver development in human. Pluripotent stem cells, such as human embryonic stem cells (hESCs), are expected to overcome some of these problems in the study of human embryogenesis, including liver development, because the gene expression profiles of this model are similar to those in normal liver development (Agarwal et al., 2008; DeLaForest et al., 2011).

During liver development, hepatoblasts differentiate into hepatocytes and cholangiocytes. A previous study has shown that a high concentration of transforming growth factor beta (TGFβ) could give rise to cholangiocyte differentiation from hepatoblasts (Clotman et al., 2005). To transmit the TGFβ signaling, TGFβ receptor 2 (TGFBR2) has to be stimulated by TGFβ1, TGFβ2 or TGFβ3 (Kitisin et al., 2007). TGFβ binding to the extracellular domain of TGFBR2 induces a conformational change, resulting in the phosphorylation and activation of TGFBR1. TGFBR1 phosphorylates SMAD2 or SMAD3, which binds to SMAD4, and then the SMAD complexes move into the nucleus and function as transcription factors to express various kinds of differentiation-related genes (Kitisin et al., 2007). Although the function of TGFBR2 in regeneration of the adult liver has been thoroughly examined (Oe et al., 2004), the function of TGFBR2 in the hepatoblast fate decision has not been elucidated.

CCAAT/enhancer binding protein (c/EBP) transcription factors play decisive roles in the differentiation of various cell types, including hepatocytes (Tomizawa et al., 1998; Yamasaki et al., 2006). The analysis of *c/EBPα* (*Cebpa*) knockout mice has shown that many abnormal pseudoglandular structures, which co-express antigens specific for both hepatocytes and cholangiocytes, are present in the liver parenchyma (Tomizawa et al., 1998). These data demonstrated that c/EBPα plays an important role in hepatocyte differentiation. It is also known that the suppression of c/EBPα expression in periportal hepatoblasts stimulates cholangiocyte differentiation (Yamasaki et al., 2006). Although the function of c/EBPα in liver development is well known, the relationship between TGFβ signaling and c/EBPα-mediated transcriptional regulation in the hepatoblast fate decision is poorly understood. c/EBPβ is also known to be an important factor for liver function (Chen et al., 2000), although the function of c/EBPβ in the cell fate decision of hepatoblasts is not well known. c/EBPα and c/EBPβ bind to the same DNA binding site. However, the promoter activity of hepatocyte-specific genes, such as those encoding hepatocyte nuclear factor 6 (HNF6, also known as ONECUT1) and UGT2B1,

[1]Laboratory of Biochemistry and Molecular Biology, Graduate School of Pharmaceutical Sciences, Osaka University, Osaka 565-0871, Japan. [2]Laboratory of Hepatocyte Differentiation, National Institute of Biomedical Innovation, Osaka 567-0085, Japan. [3]iPS Cell-based Research Project on Hepatic Toxicity and Metabolism, Graduate School of Pharmaceutical Sciences, Osaka University, Osaka 565-0871, Japan. [4]Laboratory of Stem Cell Regulation, National Institute of Biomedical Innovation, Osaka 567-0085, Japan. [5]Institute of Advanced Biomedical Engineering and Science, Tokyo Women's Medical University, Tokyo 162-8666, Japan. [6]Pharmaceutical Research and Technology Institute, Kinki University, Osaka 577-8502, Japan. [7]Laboratory of Embryonic Stem Cell Cultures, Department of Disease Bioresources Research, National Institute of Biomedical Innovation, Osaka 567-0085, Japan. [8]Department of Embryonic Stem Cell Research, Field of Stem Cell Research, Institute for Frontier Medical Sciences, Kyoto University, Kyoto 606-8507, Japan. [9]The Center for Advanced Medical Engineering and Informatics, Osaka University, Osaka 565-0871, Japan.

*Author for correspondence (mizuguch@phs.osaka-u.ac.jp)

Received 27 August 2013; Accepted 3 October 2013