Table 3. Novel loci identified in this study that are different from those in Table S1, for the segments overlapping in more than 1 unrelated individual and the common regions between the 2 siblings (cases h-1 and h-2).

Chromosome	Start	End	Samples	# Samples[a]	Length	Cytoband
1	146258078	148749860	h-1, h-2, a	3	2491783	1q21.1-q21.2
5	45437574	49631829	h-1, h-2, d	3	4194256	5p12-q11.1
5	117360252	120214932	h-1, h-2, f	3	2854681	5q23.1
5	120214932	122586267	h-1, h-2, f, g	4	2371336	5q23.1-23.2
7	57594442	62282881	h-1, h-2, b, f	4	4688440	7p11.2-q11.21
8	129121122	131617749	h-1, h-2, b	3	2496628	8q24.21-q24.22
8	132434559	139244531	h-1, h-2, b	3	6809973	8q24.22-24.23
10	37363792	37599485	h-1, h-2, e	3	235694	10p11.21
10	37599485	37874740	h-1, h-2, e, g	4	275256	10p11.21
10	37874740	42217616	h-1, h-2, c, e, g	5	4342877	10p11.21-q11.21
12	33982292	36255461	h-1, h-2, a, d	4	2273170	12p11.1-q11
13	35366458	43580724	h-1, h-2, g	3	8214267	13q13.3-14.11
16	28924029	29606107	h-1, h-2, c	3	682079	16p11.2
16	29606107	29657036	h-1, h-2, c, f	4	50930	16p11.2
16	29657036	29680943	h-1, h-2, c, d, f	5	23908	16p11.2
16	29680943	31277953	h-1, h-2, b, c, d, f	6	1597011	16p11.2
16	34467305	34647935	h-1, h-2, a, b, c, d, f, g	8	180631	16p11.1
16	34647935	45122807	h-1, h-2, a, c, d, f, g	7	10474873	16p11.1-q11.2
16	45122807	47094922	h-1, h-2, a, b, c, d, f, g	8	1972116	16q11.2-q12.1
17	29659797	32811528	h-1, h-2, a	3	3151732	17q12
19	37676724	40349191	h-1, h-2, a	3	2672468	19q13.11-13.12
21	19821557	20188026	h-1, h-2, g	3	366470	21q21.2

[a]Number of individuals (including h-1 and h-2) who shared the region; for example, 5 indicates that 3 other individuals shared the common region of the 2 siblings.
doi:10.1371/journal.pone.0020589.t003

Discussion

1. Samples

We recruited 9 offspring from first-cousin marriages (SCZ) and 92 from non-consanguineous marriages (non-CJ). As shown in Figure 2, our non-CJ dataset and publicly available HapMap3 JPT datasets showed a common cluster, except for the presence of 3 outliers that have been reported to be potentially from consanguineous families [14,15]. This concordance suggests that our experimental quality and data processing approaches were appropriate. In this study, we analyzed a limited number of samples; however, homozygosity mapping was a reasonable strategy to adopt because it requires relatively smaller number of samples than case-control studies. We did not use samples from the parents of patients in this study because these are not very informative in our strategy. On the other hand, affected and unaffected siblings in single families are strongly informative in homozygosity mapping, and we are continuously recruiting additional siblings for future study.

2. ROH analysis

Most of the previous homozygosity mapping studies were based on genotypes derived from microsatellites or simple tandem-repeat polymorphisms (STRP). The highly polymorphic nature of multi-allelic STRP markers is suitable to cover the whole genome with a fewer numbers of markers. However, recent DNA microarray technologies have enabled massive genome-wide SNP genotyping to be performed in a short time. The problem with homozygosity mapping based on SNPs is the accurate detection of regions with ROHs. As SNPs have a less informative biallelic nature, using the naïve definition of an ROH as just a contiguous homozygous region may skew the detection of ROHs because of the frequency of low minor allele SNPs, genotyping errors, and "no-call" SNPs.

The solution to this problem using the Affymetrix Human Genotyping 500K arrays and Illumina Infinium HumanHap300v2 arrays was the application of ROH detection bins sliding through each chromosome to filter out low SNP density bins and to allow the small number of heterozygous SNPs and no-call SNPs to be placed in a bin [6,17]. An alternative method to detect ROHs is to adopt an HMM. Partek GS software implements the HMM-based "LOH detection" algorithm. A similar algorithm is also implemented in the Affymetrix GeneChip Chromosome Copy Number Analysis Tool (CNAT), as described in the CNAT user guide [18]. The HMM-based algorithm of these tools takes not only the information of adjacent SNP genotypes but also the heterozygosity of SNPs as a reference baseline calculated from the genotyping results in the reference samples or the *a priori* default frequency. This method is expected to more accurately detect ROH regions that reflect actual recombination.

Selection of the reference population for the baseline data is crucial for the HMM-based detection of ROHs. If the reference population is carefully selected to match the background of the case population, the baseline generated from the observation of actual SNPs in the reference population can omit ROHs resulting from LD and regions with low SNP density, such as centromeres. However, if strict matching of the used population background is difficult, use of the fixed default heterozygous frequency, whose

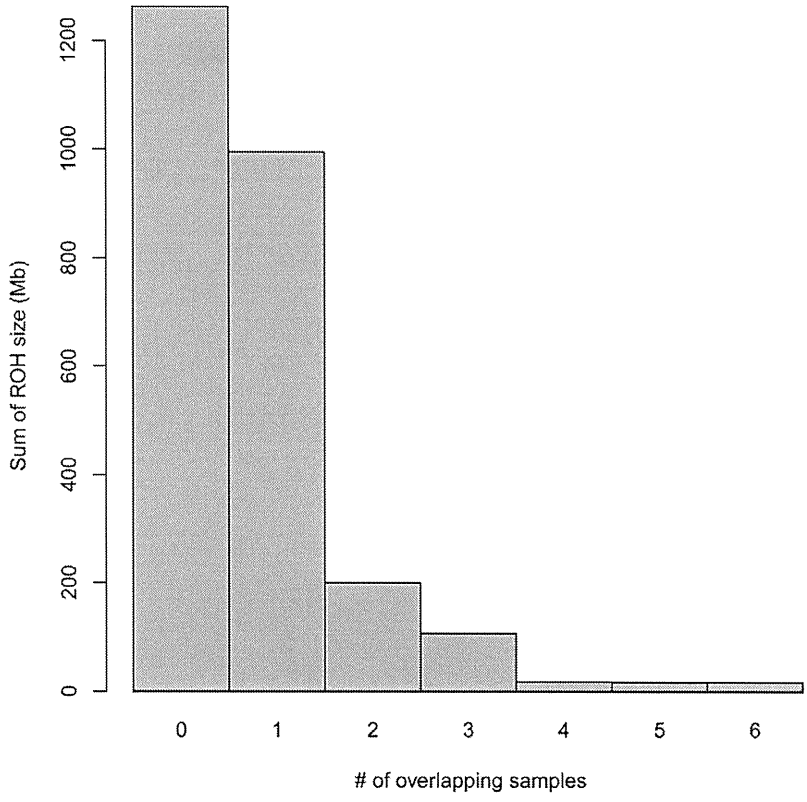

Figure 4. Sum of run of homozygosity (ROH) lengths and number of overlapping patients, excluding patient siblings. Y-axis indicates the sum of ROH lengths shared by a given number of patients. The zero column indicates the sum of ROHs not shared by any of the samples.
doi:10.1371/journal.pone.0020589.g004

default value is 0.3, still has the advantage of minimizing false-negatives in the detection of ROHs.

To determine the optimal length threshold of ROHs to extract autozygous segments from whole ROHs, we adopted OR = 3 for the analysis. This approach may work well when a large enough reference sample is available. When a reference population is not available, a threshold where the sum of the ROH length in the upper tail of its distribution is equal to the theoretical autozygous length of a genome, that is, 1/16 of a genome in the offspring of a first-cousin marriage, could be another option. In our SCZ dataset, the threshold using this approach was approximately 10.6 Mb.

Our results demonstrated obvious differences in the proportion of the length distribution of ROHs between the non-CJ and SCZ datasets. A recent report on European populations, including endogamy subpopulations, has shown that a higher proportion of individuals in endogamy subpopulations have ROHs longer than 1.5 Mb compared with other subpopulations [6]. Our scatter plot of the individual total number and size of ROHs (Figure 2) is not fully in agreement with this previous report, although the non-CJ dataset made a cluster and showed a positive correlation (Pearson product-moment correlation coefficient r = 0.773), and the SCZ dataset was scattered and showed a weak positive correlation (r = 0.432). This may be explained by the fact that the previous report excluded ROHs <500 kb to ignore ROHs that potentially resulted from LD and removed hemizygous deletions of ROHs. In this study, we did not adopt a strategy to filter ROHs by their size before the analyses because the discrimination of autozygous regions and LD simply by size is essentially impossible. Adopting a baseline file derived from a strictly matched population in the HMM-based detection of ROHs can be used instead. Additionally, differences in genotyping platforms with different SNP densities may affect the size distribution of ROHs. Although our data from the sparser Affymetrix Genotyping 10k SNP panel produced a similar bell curve-like ROH size distribution, the whole curve was shifted to the right (data not shown).

The size distribution of ROHs for a given population is affected by its inbreeding coefficient (F). Studies of consanguineous marriages in subpopulations from Japan during the 1980s compared the F values for Japan (F = 0.00134) to those in Kuwait (F = 0.0219), India (F = 0.02313), England (F = 0.00017), and the United States (F = 0.00003) [19,20]. These reports have also shown that despite the decrease in consanguineous marriages in Japan, local subpopulations have higher F-values. The same tendency has also been shown by a genealogical study that estimated inbreeding rates in large and semi-isolated populations on the basis of historical changes in population size [21]. Recently, the importance of studying endogamous populations has been stressed [22]; however, populations with intermediate F-values have advantages for our homozygosity mapping approach. This approach uses the differences in the size distribution of ROHs in a case population consisting of offspring from consanguineous marriages and a control population consisting of offspring from non-consanguineous marriages. A high F-value population may not have clear distribution differences between cases and controls. On the other hand, finding a sufficient number of cases in low F-value populations may not be easy. From this standpoint, an intermediate F-value population, such as the Japanese population,


A
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
B
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22

Figure 5. Overlapping autosomal runs of homozygosity. Each autosome is shown horizontally with the number of overlapping samples (upper) and chromosome ideograms (lower). Centromeres are shown by hatched boxes. A, overlapping segments shared among 1 (isolated) to 7 samples in a total of 9 patient samples. B, overlapping segments shared by 2 patient siblings (h-I and h-II) and an additional 1–4 patient samples.
doi:10.1371/journal.pone.0020589.g005

represents an interesting dataset for our homozygosity mapping approach, as was shown previously in the Costa Rican population [23,24].

We presented here some threshold lengths of ROHs to detect IBD regions from great-grandparents. As recombination will, of course, occur everywhere by chance, small autozygous regions could be overlooked with the threshold shown here. However, no systematic analyses have so far identified IBD regions in consanguineous marriages by whole-genome SNP typing. Our method shown here, which 1) detects longer ROHs in each individual and 2) aligns ROHs and identifies overlapping regions, will be helpful for autosomal recessive disorders and also for complex disorders resulting from rare variants. If collecting patients in geographically and historically isolated areas is possible, this homozygosity mapping approach is likely to be successful. Nonetheless, the effectiveness of homozygosity mapping for complex disorders remains controversial [4]. We believe that we can uncover new candidate loci through the application of whole-genome SNP typing to homozygosity mapping because of its high density genomic coverage and high-throughput ability.

3. Possible novel loci for schizophrenia

We identified several putative SCZ loci that are presented in Figure 5 and Tables S1 and 3. In our study, we assumed the 2 models outlined in the Methods section. Model I was designed to find shared causal loci among unrelated individuals. For the other model, we hypothesized that the siblings shared the same causal loci; thus, Model II was designed to find the common loci between the siblings and unrelated individuals.

For Model I (Table S1), from the analysis of 7 unrelated individuals, the loci included the 5q23.3–q31.1 region that was previously identified by linkage analysis in the Irish population [16]. Among the genes that mapped to 5q23.3–q31.1, *HINT1* [25,26] and *ACSL6* [27] were previously reported to be possibly associated with SCZ. In patients from consanguineous families we analyzed, homozygous genomic variations may be causative for the disease (Figure 5 and Tables S1 and 3). However, small sample size in our study may be a limiting factor to generalize such conclusions. In common diseases such as psychiatric disorders, including SCZ and bipolar disorders, especially in familial cases or in cases from relatively isolated areas, rare variants possibly contribute more than common variants to the disease phenotype [28–30].

On the basis of the rare variant-common disease hypothesis, it is appropriate that the genetic etiology between sibling cases and other unrelated cases may be various. In addition, our results suggested that multiple loci influenced the susceptibility to SCZ, as other reports have suggested [31].

We presented here the systematic analyses of the homozygosity mapping method using whole genome SNP typing, and we identified ROHs that potentially contain SCZ causative recessive regions that are shared among our samples. When we explain SCZ as a result of the homozygous state of rare variant mutations, the number of overlapping individuals may be challenging, as it is possible that each individual has a different variation. The heterogeneity of SCZ may explain the lack of overlap for our results with previously reported regions [32,33]; moreover, our methodology has a limitation for detecting causative genes that are included in shorter ROHs by chance. We have shown that the Affymetrix Genome-Wide Human SNP Array 5.0 or 6.0 could be applied to special cases including first-cousin marriages to identify genomic variations. Increasing number of samples obtained from patients from consanguineous families with SCZ is important to make our results more meaningful. Furthermore, we plan to analyze genetic variants in updated ROHs by the next-generation sequencing technologies.

Supporting Information

Table S1 Novel loci identified in this study.
(DOC)

Acknowledgments

We thank Dr. Haruko Ichinose from Eijinkai Ariake Hoyouin Hospital and Drs. Takehito Sakai and Sumihisa Honda from Nagasaki University for their assistance. We also thank Dr. Pawel Stankiewicz from Baylor College of Medicine for a critical review of the manuscript.

Author Contributions

Conceived and designed the experiments: NK KI-Y HO. Performed the experiments: NK ST HM. Analyzed the data: SO AI TK NN KT. Contributed reagents/materials/analysis tools: SO AI TK NN KT. Wrote the paper: NK KI-Y HO.

References

1. Burmeister M, McInnis MG, Zöllner S (2008) Psychiatric genetics: progress amid controversy. Nat Rev Genet 9: 527–540.
2. Morrow EM, Yoo SY, Flavell SW, Kim TK, Lin Y, et al. (2008) Identifying autism loci and genes by tracing recent shared ancestry. Science 321: 218–223.
3. Bulayeva KB (2006) Overview of genetic-epidemiological studies in ethnically and demographically diverse isolates of Dagestan, Northern Caucasus, Russia. Croat Med J 47: 641–648.
4. Rudan I, Campbell H, Carothers AD, Hastie ND, Wright AF (2006) Contribution of consanguinuity to polygenic and multifactorial diseases. Nat Genet 38: 1224–1225.
5. Mansour H, Fathi W, Klei L, Wood J, Chowdari K, et al. (2010) Consanguinity and increased risk for schizophrenia in Egypt. Schizophr Res 120: 108–112.
6. McQuillan R, Leutenegger AL, Abdel-Rahman R, Franklin CS, Pericic M, et al. (2008) Runs of homozygosity in European populations. Am J Hum Genet 83: 359–372.
7. The International HapMap Consortium (2003) The International HapMap Project. Nature 426: 789–796.
8. Hong H, Su Z, Ge W, Shi L, Perkins R, et al. (2008) Assessing batch effects of genotype calling algorithm BRLMM for the Affymetrix GeneChip Human Mapping 500 K array set using 270 HapMap samples. BMC Bioinformatics 9: S17.
9. BRLMM-P: a Genotype Calling Method for the SNP 5.0 Array. Available: http://www.affymetrix.com/support/technical/whitepapers/brlmmp_whitepaper.pdf. Accessed 14 Mar 2011.
10. R Development Core Team () R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2006). Available: http://www.R-project.org/. Accessed 14 Mar 2011.
11. Wand MP, Johnes MC (1995) Kernel Smoothing. Chapman and Hall, London.
12. Korn JM, Kuruvilla FG, McCarroll SA, Wysoker A, Nemesh J, et al. (2008) Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. Nat Genet 40: 1253–1260.
13. The International HapMap Consortium (2005) A haplotype map of the human genome. Nature 437: 1299–1320.
14. Yang H, Chang L, Huggins RM, Chen C, Mullighan CG (2011) LOHAS: loss-of-heterozygosity analysis suite. Genet Epidemiol. In press.

15. USCS Genome Bioinformatics, Golden Path Statistics, NCBI Build 36.1 assembly, March 2006 (hg 18). Available: http://genome.ucsc.edu/goldenPath/stats.html#hg18. Accessed 14 Mar 2011.
16. Straub RE, MacLean CJ, Ma Y, Webb BT, Myakishev MV, et al. (2002) Genome-wide scans of three independent sets of 90 Irish multiplex schizophrenia families and follow-up of selected regions in all families provides evidence for multiple susceptibility genes. Mol Psychiatry 7: 542–559.
17. Lencz T, Lambert C, DeRosse P, Burdick KE, Morgan TV, et al. (2007) Runs of homozygosity reveal highly penetrant recessive loci in schizophrenia. Proc Natl Acad Sci U S A 104: 19942–19947.
18. Affymetrix GeneChip Chromosome Copy Number Analysis Tool (CNAT) Version 4.0 User Guide (2007) Affymetrix Inc.
19. Imaizumi Y (1986) A recent survey of consanguineous marriages in Japan. Clin Genet 30: 230–233.
20. Al-Awadi SA, Moussa MA, Naguib KK, Farag TI, Teebi AS, et al. (1985) Consanguinity among the Kuwaiti population. Clin Genet 27: 483–486.
21. Pattison JE (2004) A comparison of inbreeding rates in India, Japan, Europe and China. Homo 55: 113–128.
22. Editorial (2006) The germinating seed of Arab genomics. Nat Genet 38: 851.
23. McInnes LA, Service SK, Reus VI, Barnes G, Charlat O, et al. (2001) Fine-scale mapping of a locus for severe bipolar mood disorder on chromosome 18p11.3 in the Costa Rican population. Proc Natl Acad Sci U S A 98: 11485–11490.
24. Mathews CA, Reus VI, Bejarano J, Escamilla MA, Fournier E, et al. (2004) Genetic studies of neuropsychiatric disorders in Costa Rica: a model for the use of isolated populations. Psychiatr Genet 14: 13–23.
25. Chen Q, Wang X, O'Neill FA, Walsh D, Kendler KS, et al. (2008) Is the histidine triad nucleotide-binding protein 1 (HINT1) gene a candidate for schizophrenia? Schizophr Res 106: 200–207.
26. Chen X, Wang X, Hossain S, O'Neill FA, Walsh D, et al. (2006) Haplotypes spanning SPEC2, PDZ-GEF2 and ACSL6 genes are associated with schizophrenia. Hum Mol Genet 15: 3329–3342.
27. Luo XJ, Diao HB, Wang JK, Zhang H, Zhao ZM, et al. (2008) Association of haplotypes spanning PDZ-GEF2, LOC728637 and ACSL6 with schizophrenia in Han Chinese. J Med Genet 45: 818–826.
28. O'Donovan MC, Craddock NJ, Owen MJ (2009) Genetics of psychosis; insights from views across the genome. Hum Genet 126: 3–12.
29. Schork NJ, Murray SS, Frazer KA, Topol EJ (2009) Common vs. rare allele hypotheses for complex diseases. Curr Opin Genet Dev 19: 212–219.
30. Gorlov I, Gorlova O, Frazier M, Spitz M, Amos C (2011) Evolutionary evidence of the effect of rare variants on disease etiology. Clin Genet 79: 199–206.
31. Ioannidis JPA, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG (2001) Replication validity of genetic association studies. Nat Genet 29: 306–309.
32. Girard SL, Xiong L, Dion PA, Rouleau GA (2011) Where are the missing pieces of the schizophrenia genetics puzzle? Curr Opin Genet Dev. In press.
33. Glessner JT, Hakonarson H (2009) Common variants in polygenic schizophrenia. Genome Biol 10: 236.

The Breast Journal

LETTER TO THE EDITOR

Intracystic Papillary Carcinoma of Breast Harbors Significant Genomic Alteration Compared with Intracystic Papilloma: Genome-wide Copy Number and LOH Analysis Using High-Density Single-Nucleotide Polymorphism Microarrays

To the Editor:

Intracystic papillary breast tumors (ICPT) consist of benign papillomas, carcinomas in situ, and carcinomas with invasion, and they account for approximately 10% of benign breast tumors and less than 1% of malignant tumors, respectively (1,2). In breast lesions, indication for surgery is usually determined by pathological diagnosis together with radiologic findings, but differential, preoperative diagnosis of papillary carcinoma from papilloma is very difficult, even following needle biopsy (3) because of their nonspecific radiologic characteristics and their modest cytological and histologic appearance (4). To avoid excessive surgical intervention, another diagnostic procedure needs to be developed.

Cytogenetic studies of breast papillary tumors are limited, and cytogenetic differences between papillomas and papillary carcinomas are still controversial. Tsuda et al. (5,6) reported that papillary carcinomas have frequent changes in gene copy number and loss of heterozygosity (LOH), whereas papillomas did not show any gene copy number alteration or LOH at 16q and 1q. Boecker et al. (7) also reported that conventional comparative genomic hybridization (CGH) did not reveal any gene copy number change in papillomas. On the other hand, Lininger et al. (8) and Cristofano et al. (2) demonstrated that LOH at 16p or 16q was frequent in both papillomas and papillary carcinomas.

Address correspondence and reprint requests to: Koh-ichiro Yoshiura, MD, PhD, Department of Human Genetics, Nagasaki University Graduate School of Biomedical Science, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan, or e-mail: kyoshi@nagasaki-u.ac.jp.

DOI: 10.1111/j.1524-4741.2011.01110.x

The Breast Journal, Volume 17 Number 4, 2011 427–430

The purpose of this study was to determine the profile of genomic alterations in breast ICPT and to explore the possibility of using high-density oligonucleotide SNP arrays as the basis of a novel diagnostic method of ICPT. Ten formalin-fixed paraffin-embedded (FFPE) breast ICPT were obtained from the Department of Pathology, Nagasaki University Hospital. The samples included five benign papillomas (Pap), three papillary carcinomas in situ (PurePC), and two papillary carcinomas with invasion (PCinv). Pathological diagnosis was independently determined by two pathologists. Clinicopathological findings of these tumors are provided in Fig. 1 and Table 1.

Extracted DNA from each sample was processed following the manufacturer's protocol and hybridized on Affymetrix GeneChip Genome-Wide Human SNP Array 5.0® (Affymetrix, Santa Clara, CA, USA). The QC call rates, which is an index measuring the quality of a SNP microarray experiment, obtained from the FFPE samples were from 70.75% to 91.93%, with a mean of 80.72% (Table 1), which was comparable to the results from former cytogenetic studies using DNA extracted from FFPE samples(9–11).

Copy number change and LOH analyses (called here SNPaCGH) were conducted using the Partek Genomics Suite (PGS) version 6.3 (Partek, St. Louis, MI, USA). To estimate the total rate of a copy number changed region, each segment amplified or lost was summed and divided by 2,829 Mb, which is the total Mb in the genome, excluding heterochromatic, centromeric, and telomeric regions not covered by probes. Similarly, to estimate the total rate of genomic alteration, the sum of segments with copy number change and copy number neutral loss of heterozygosity (CNLOH) was divided by 2,829 Mb. To validate the

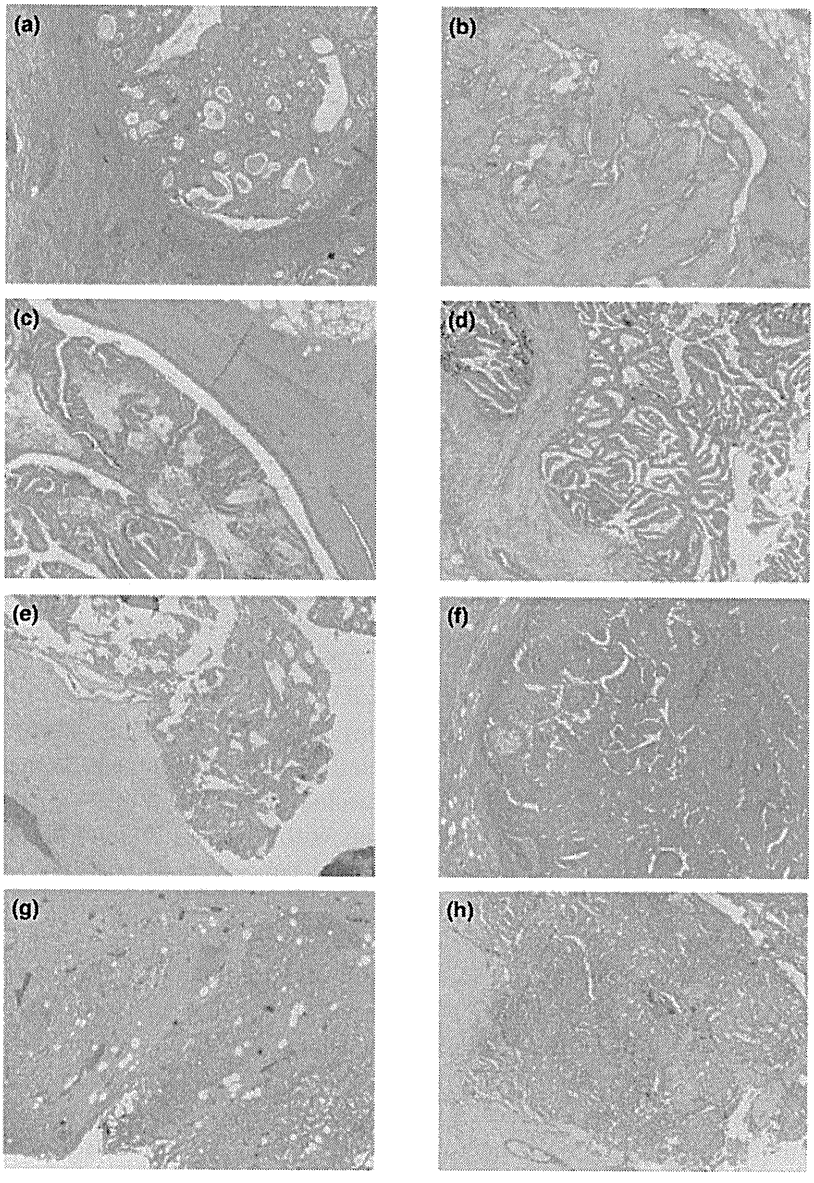

Figure 1. Hematoxylin-eosin stain in intracystic papillary tumors (Original magnification × 40). a–c, Intracystic papilloma (a: case 1, b: case 2, c: case 3). d–f, Intracystic papillary carcinoma in situ (d: case 6, e: case 7, f: case 8). g and h, Intracystic papillary carcinoma with invasion (g: case 9, h: case10).

Table 1. Characteristics of ten intracystic papillary lesions

	Clinicopathologic findings							Genetic findings		
Case	Diagnosis	Age	Size of cyst (mm)	MMG	US	FNAC	Receptor status	QC call rate (normal/tumor)	Ratio of CNC	Ratio of genomic alteration
1	Pap	43	80	Category 3	Category 3	Class 2	ER(+)	75.9%/82.9%	0.14%	0.24%
2	Pap	38	10	Category 1	Category 3	Class 3	NaN	83.4%/80.4%	0.66%	0.69%
3	Pap	49	25	Category 3	Category 3	Class 3	NaN	86.2%/86.5%	1.60%	1.60%
4	Pap	38	70	Category 3	Category 3	Class 2	NaN	89.9%/87.9%	0%	11.8%
5	Pap	49	75	Category 3	Category 3	Class 2	NaN	91.9%/89.8%	0%	0%
6	PurePC	61	31	Category 4	Category 3	Class 4	ER(+), HER2(1 +)	75.7%/76.2%	11.3%	24.1%
7	PurePC	58	49	Category 3	Category 4	Class 4	ER(+)	79.7%/70.8%	0.41%	8.83%
8	PurePC	43	16	Category 2	Category 4	Class 4	ER(+), HER2(1 +)	77.2%/79.9%	12.0%	13.2%
9	PCinv	60	96	NaN	Category 4	Class 1	ER(−), HER2(1 +)	71.6%/73.9%	16.6%	53.1%
10	PCinv	72	19	Category 4	Category 4	Class 5	ER(+), HER2(1 +)	82.0%/72.6%	16.0%	17.6%

Pap: intracystic papilloma, Pure PC: intracystic papillary carcinoma in situ, PC inv: intracystic papillary carcinoma with invasion, MMG: the mammographic features evaluated according to the Breast Imaging-Reporting and Data System (BI-RADS) of the American College of Radiology, US: the ultrasonographic features evaluated according to diagnostic guideline of the Japanese Association of Breast and Thyroid Sonology (JABTS), FNAC: the cytological features of fine needle aspiration cytology, ER: the status of estrogen receptor, HER2: the status of HER2/neu receptor, CNC: copy number change, genomic alteration: copy number change and copy neutral loss of heterozygosity, NaN: not analyzed.

copy number change identified by SNPaCGH, quantitative PCR assays were performed on a LightCycler® 480 Real-Time PCR System (Roche Diagnostics, Mannheim, Germany) at four selected loci, including independent genes (Table S1).

In SNPaCGH analysis, substantial divergence was observed between each ICPT subtype (Fig. 2). The mean rate of copy number change was 0.48% (from 0.0% to 1.60%), 7.89% (from 0.41% to 12.0%), and 16.3% (from 16.0% to 16.6%) in Pap, PC, and PCinv, respectively. The mean rate of genomic alteration (including copy number change and CNLOH) was 2.87% (from 0.00% to 11.8%), 15.4% (from 8.83% to 24.1%), and 35.3% (from 17.6% to 53.1%) in Pap, PC, and PCinv, respectively (Table 1). Malignant tumors (PurePC and PCinv) showed significantly more copy number changes and genomic alterations (copy number change and CNLOH) than benign tumors (Pap) (Wilcoxon's rank sum test, $p = 0.036$, 0.016, respectively) and these differences correlated with their malignant phenotype (Kruskal–Wallis' chi-squared test, $p = 0.046$, 0.043, respectively). The real time qPCR analysis to validate the copy number state in SNPaCGH demonstrated sufficient specificity, and thus all loci showing alteration in SNPaCGH were confirmed by real-time qPCR (Table S1). On the other hand, at 31 loci from the ten samples, where SNPaCGH showed the copy number state as disomy, ten loci were revealed to

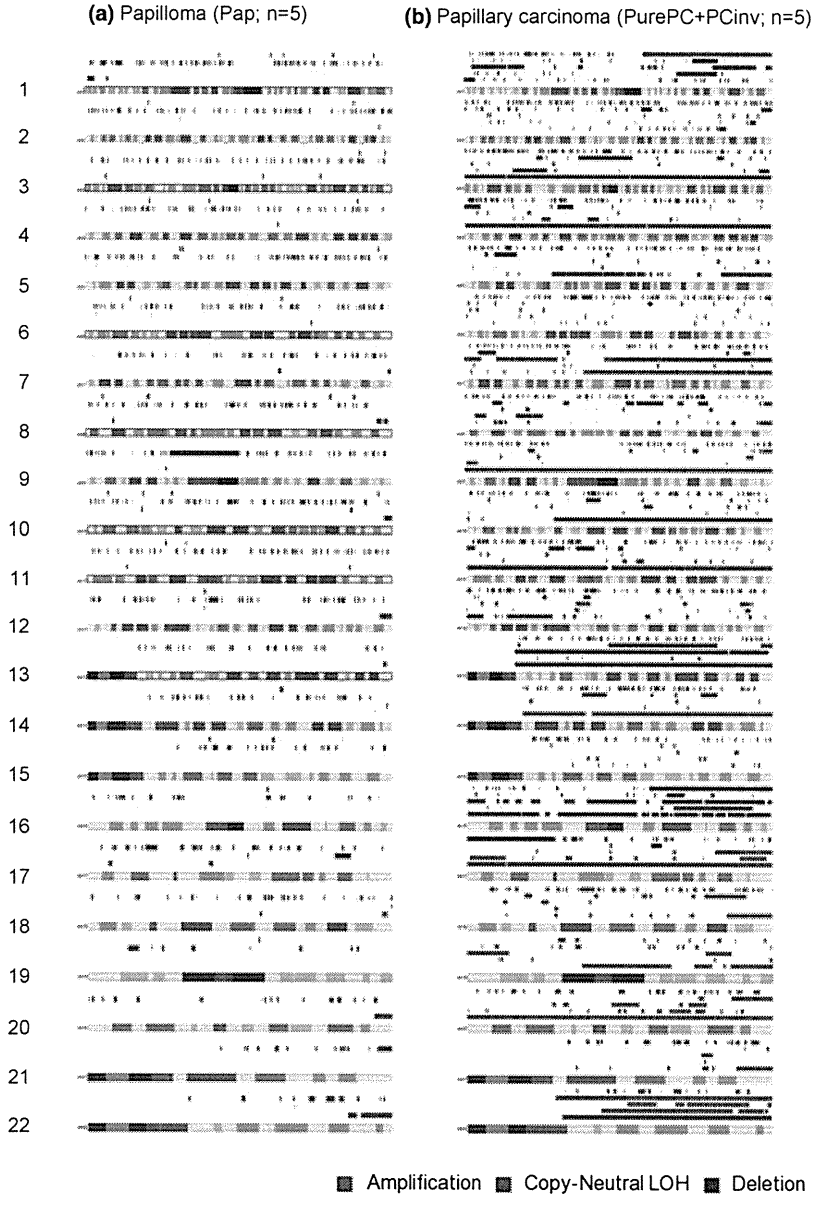

Figure 2. Graphic display of whole genomic alterations in papilloma (a) and papillary carcinoma (b). The color bar over each chromosome indicates copy number amplification (green color bars), copy-neutral LOH (blue color bars), and deletion (brown color bars) for each case. Papilloma includes five cases of Pap (a), and papillary carcinoma includes three cases of PurePC and two of PCinv (b).

have a copy number change by real-time qPCR. At regions determined to be disomy by SNPaCGH, 70.6% were determined using qPCR to be two copies. Previous studies have documented that the most common genomic alteration in papillary carcinoma is amplification on 1p, and deletion or LOH on 16q (2,5,6,8). Our study revealed that chromosomal regions at 3p21.31, 3p14.2, and 20q13.13 were commonly altered (≥4/5) among five carcinomas in addition to deletion or CNLOH on 16q (≥3/5) (Fig. 2, Table S2). The significance of 3p and 20q are currently unclear, but require further investigation.

In summary, we have elucidated significant differences in the molecular-cytogenetic profile between papilloma and papillary carcinoma. Thus, papillary carcinoma harbored significantly more genomic alterations than papilloma, even though papilloma had a number of genomic alterations, and the rate of genomic alteration correlated with pathological malignancy classification. These genome-wide findings could not be obtained by conventional cytogenetic study such a fluorescent in situ hybridization or conventional CGH. Our findings may aid clinical management of breast ICPT, and may provide insight into their carcinogenesis.

Masahiro Oikawa, MD*,†
Takeshi Nagayasu, MD, PhD†
Hiroshi Yano, MD, PhD†
Tomayoshi Hayashi, MD, PhD‡
Kuniko Abe, MD, PhD‡
Akira Kinoshita, PhD*
Koh-ichiroYoshiura, MD, PhD*
*Department of Human Genetics, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan
†Surgical Oncology, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan;
‡Department of Pathology, Nagasaki University Hospital, Nagasaki, Japan

REFERENCES

1. Fayanju OM, Ritter J, Gillanders WE, *et al.* Therapeutic management of intracystic papillary carcinoma of the breast: the roles of radiation and endocrine therapy. *Am J Surg* 2007;194:497–500.

2. Di Cristofano C, Mrad K, Zavaglia K, *et al.* Papillary lesions of the breast: a molecular progression? *Breast Cancer Res Treat* 2005;90:71–6.

3. Douglas-Jones AG, Verghese A. Diagnostic difficulty arising from displaced epithelium after core biopsy in intracystic papillary lesions of the breast. *J Clin Pathol* 2002;55:780–3.

4. Solorzano CC, Middleton LP, Hunt KK, *et al.* Treatment and outcome of patients with intracystic papillary carcinoma of the breast. *Am J Surg* 2002;184:364–8.

5. Tsuda H, Takarabe T, Susumu N, *et al.* Detection of numerical and structural alterations and fusion of chromosomes 16 and 1 in low-grade papillary breast carcinoma by fluorescence in situ hybridization. *Am J Pathol* 1997;151:1027–34.

6. Tsuda H, Takarabe T, Akashi-Tanaka S, Fukutomi T, Hirohashi S. Pattern of chromosome 16q loss differs between an atypical proliferative lesion and an intraductal or invasive ductal carcinoma occurring subsequently in the same area of the breast. *Mod Pathol* 2001;14:382–8.

7. Boecker W, Buerger H, Schmitz K, *et al.* Ductal epithelial proliferations of the breast: a biological continuum? Comparative genomic hybridization and high-molecular-weight cytokeratin expression patterns. *J Pathol* 2001;195:415–21.

8. Lininger RA, Park WS, Man YG, *et al.* LOH at 16p13 is a novel chromosomal alteration detected in benign and malignant microdissected papillary neoplasms of the breast. *Hum Pathol* 1998;29:1113–8.

9. Jacobs S, Thompson ER, Nannya Y, *et al.* Genome-wide, high-resolution detection of copy number, loss of heterozygosity, and genotypes from formalin-fixed, paraffin-embedded tumor tissue using microarrays. *Cancer Res* 2007;67:2544–51.

10. Lyons-Weiler M, Hagenkord J, Sciulli C, Dhir R, Monzon FA. Optimization of the Affymetrix GeneChip Mapping 10K 2.0 Assay for routine clinical use on formalin-fixed paraffin-embedded tissues. *Diagn Mol Pathol* 2008;17:3–13.

11. Nessling M, Richter K, Schwaenen C, *et al.* Candidate genes in breast cancer revealed by microarray-based comparative genomic hybridization of archived tissue. *Cancer Res* 2005;65:439–47.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table S1. The validation of Array-based comparative genomic hybridization by real time quantitative PCR.

Table S2. Regions of genomic alteration commonly shared in intracystic papillary carcinoma.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

RESEARCH ARTICLE

AMERICAN JOURNAL OF medical genetics PART A

Spectrum of *MLL2* (*ALR*) Mutations in 110 Cases of Kabuki Syndrome

Mark C. Hannibal,[1,2] Kati J. Buckingham,[1] Sarah B. Ng,[3] Jeffrey E. Ming,[4] Anita E. Beck,[1,2] Margaret J. McMillin,[2] Heidi I. Gildersleeve,[1] Abigail W. Bigham,[1] Holly K. Tabor,[1,2] Heather C. Mefford,[1,2] Joseph Cook,[1] Koh-ichiro Yoshiura,[5] Tadashi Matsumoto,[5] Naomichi Matsumoto,[6] Noriko Miyake,[6] Hidefumi Tonoki,[7] Kenji Naritomi,[8] Tadashi Kaname,[8] Toshiro Nagai,[9] Hirofumi Ohashi,[10] Kenji Kurosawa,[11] Jia-Woei Hou,[12] Tohru Ohta,[13] Deshung Liang,[14] Akira Sudo,[15] Colleen A. Morris,[16] Siddharth Banka,[17] Graeme C. Black,[17] Jill Clayton-Smith,[17] Deborah A. Nickerson,[3] Elaine H. Zackai,[4] Tamim H. Shaikh,[18] Dian Donnai,[17] Norio Niikawa,[13] Jay Shendure,[3] and Michael J. Bamshad[1,2,3]*

[1]Department of Pediatrics, University of Washington, Seattle, Washington

[2]Seattle Children's Hospital, Seattle, Washington

[3]Department of Genome Sciences, University of Washington, Seattle, Washington

[4]Department of Pediatrics, The Children's Hospital of Philadelphia, The University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania

[5]Department of Human Genetics, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan

[6]Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama, Japan

[7]Department of Pediatrics, Tenshi Hospital, Sapporo, Japan

[8]Department of Medical Genetics, University of the Ryukyus, Okinawa, Japan

[9]Department of Pediatrics, Dokkyo Medical University, Koshigaya Hospital, Saitama, Japan

[10]Division of Medical Genetics, Saitama Children's Medical Center, Saitama, Japan

[11]Division of Clinical Genetics, Kanagawa Children's Medical Center, Yokohama, Japan

[12]Department of Pediatrics, Chang Gung Children's Hospital, Taoyuan, Taiwan, Republic of China

[13]Research Institute of Personalized Health Sciences, Health Sciences University of Hokkaido, Hokkaido, Japan

[14]National Laboratory of Medical Genetics, Xiangya Hospital, Central South University, Republic of China

[15]Department of Pediatrics, Sapporo City General Hospital, Sapporo, Japan

[16]University of Nevada School of Medicine, Las Vegas, Nevada

[17]Department of Genetic Medicine, Manchester Academic Health Sciences Centre, University of Manchester, England

[18]Department of Pediatrics, University of Colorado, Denver, Colorado

Received 25 February 2011; Accepted 30 March 2011

Additional supporting information may be found in the online version of this article.

Grant sponsor: National Institutes of Health/National Heart Lung and Blood Institute; Grant number: 5R01HL094976; Grant sponsor: National Institutes of Health/National Human Genome Research Institute; Grant numbers: 5R21HG004749, 1RC2HG005608, 5RO1HG004316, T32HG00035; Grant sponsor: National Institute of Heath/National Institute of Environmental Health Sciences; Grant number: HHSN273200800010C; Grant sponsor: National Institute of Neurological Disorders and Stroke; Grant number: RO1NS35102; Grant sponsor: NIHR Manchester Biomedical Research Centre; Grant sponsor: Ministry of Health, Labour and Welfare of Japan; Grant sponsor: Japan Science and Technology Agency; Grant sponsor: Society for the Promotion of Science; Grant sponsor: Life Sciences Discovery Fund; Grant numbers: 2065508, 0905001; Grant sponsor: Washington Research Foundation; Grant sponsor: National Institutes of Health/National Institute of Child Health and Human Development; Grant numbers: 1R01HD048895, 5K23HD057331.

Mark C. Hannibal, Kati J. Buckingham, and Sarah B. Ng contributed equally to this work.

*Correspondence to:
Michael J. Bamshad, M.D., Department of Pediatrics, University of Washington School of Medicine, Box 356320, 1959 NE Pacific Street, Seattle, WA 98195. E-mail: mbamshad@u.washington.edu

Published online 10 June 2011 in Wiley Online Library (wileyonlinelibrary.com).
DOI 10.1002/ajmg.a.34074

Kabuki syndrome is a rare, multiple malformation disorder characterized by a distinctive facial appearance, cardiac anomalies, skeletal abnormalities, and mild to moderate intellectual disability. Simplex cases make up the vast majority of the reported cases with Kabuki syndrome, but parent-to-child transmission in more than a half-dozen instances indicates that it is an autosomal dominant disorder. We recently reported that Kabuki syndrome is caused by mutations in *MLL2*, a gene that encodes a Trithorax-group histone methyltransferase, a protein important in the epigenetic control of active chromatin states. Here, we report on the screening of 110 families with Kabuki syndrome. *MLL2* mutations were found in 81/110 (74%) of families. In simplex cases for which DNA was available from both parents, 25 mutations were confirmed to be de novo, while a transmitted *MLL2* mutation was found in two of three familial cases. The majority of variants found to cause Kabuki syndrome were novel nonsense or frameshift mutations that are predicted to result in haploinsufficiency. The clinical characteristics of *MLL2* mutation-positive cases did not differ significantly from *MLL2* mutation-negative cases with the exception that renal anomalies were more common in *MLL2* mutation-positive cases. These results are important for understanding the phenotypic consequences of *MLL2* mutations for individuals and their families as well as for providing a basis for the identification of additional genes for Kabuki syndrome. © 2011 Wiley-Liss, Inc.

Key words: Kabuki syndrome; *MLL2*; *ALR*; Trithorax group histone methyltransferase

How to Cite this Article:
Hannibal MC, Buckingham KJ, Ng SB, Ming JE, Beck AE, McMillin MJ, Gildersleeve HI, Bigham AW, Tabor HK, Mefford HC, Cook J, Yoshiura K-i, Matsumoto T, Matsumoto N, Miyake N, Tonoki H, Naritomi K, Kaname T, Nagai T, Ohashi H, Kurosawa K, Hou J-W, Ohta T, Liang D, Sudo A, Morris CA, Banka S, Black GC, Clayton-Smith J, Nickerson DA, Zackai EH, Shaikh TH, Donnai D, Niikawa N, Shendure J, Bamshad MJ. 2011. Spectrum of *MLL2* (*ALR*) mutations in 110 cases of Kabuki syndrome.
Am J Med Genet Part A 155:1511–1516.

INTRODUCTION

Kabuki syndrome (OMIM#147920) is a rare, multiple malformation disorder characterized by a distinctive facial appearance, cardiac anomalies, skeletal abnormalities, and mild to moderate intellectual disability. It was originally described by Niikawa et al. [1981] and Kuroki et al. [1981] in 1981, and to date, about 400 cases have been reported worldwide [Niikawa et al., 1988; White et al., 2004; Adam and Hudgins, 2005]. The spectrum of abnormalities found in individuals with Kabuki syndrome is diverse, yet virtually all affected persons are reported to have similar facial features consisting of elongated palpebral fissures, eversion of the lateral third of the lower eyelids, and broad, arched eyebrows with lateral sparseness. Additionally, affected individuals commonly have severe feeding problems, failure to thrive in infancy, and height around or below the 3rd centile for age in about half of cases.

We recently reported that a majority of cases of Kabuki syndrome are caused by mutations in *mixed lineage leukemia 2* (*MLL2*; OMIM#602113), also known as either *MLL4* or *ALR* [Ng et al., 2010]. *MLL2* encodes a SET-domain-containing histone methyltransferase important in the epigenetic control of active chromatin states [FitzGerald and Diaz, 1999]. Exome sequencing revealed that 9 of 10 individuals had novel variants in *MLL2* that were predicted to be deleterious. A single individual had no mutation in the protein-coding exons of *MLL2*, though in retrospect, his phenotypic features are somewhat atypical of Kabuki syndrome. In a larger validation cohort screened by Sanger sequencing, we found *MLL2* mutations in approximately two-thirds of 43 Kabuki cases, suggesting that Kabuki syndrome is genetically heterogeneous.

Herein we report on the results of screening *MLL2* for mutations in 110 families with one or more individuals affected with Kabuki syndrome in order to: (1) characterize the spectrum of *MLL2* mutations that cause Kabuki syndrome; (2) determine whether *MLL2* genotype is predictive of phenotype; (3) assess whether the clinical characteristics of *MLL2* mutation-positive cases differ from *MLL2* mutation-negative cases; and (4) delineate the subset of Kabuki cases that are *MLL2* mutation-negative for further gene discovery studies.

MATERIALS AND METHODS

Subjects

Referral for inclusion into the study required a diagnosis of Kabuki syndrome made by a clinical geneticist. From these cases, phenotypic data were collected by review of medical records, phone interviews, and photographs. These data were collected from five different clinical genetics centers in three different countries and over a protracted period of time and forwarded for review to two of the authors (M.B. and M.H.). Data on certain phenotypic characteristics including stature, feeding difficulties, and failure to thrive was not uniformly collected or standardized. Therefore, we decided to be conservative in our analysis and use only phenotypic traits that could be represented by discrete variables (i.e., presence or absence) and for which data were available from at least 70% of cases. In addition, these clinical summaries were de-identified and therefore facial photographs were unavailable from most cases studied. Written consent was obtained for all participants who provided identifiable samples. The Institutional Review Boards of Seattle Children's Hospital and the University of Washington approved all studies. A summary of the clinical characteristics of 53 of these individuals diagnosed with Kabuki syndrome has been reported previously [Ng et al., 2010].

Mutation Analysis

Genomic DNA was extracted using standard protocols. Each of the 54 exons of *MLL2* was amplified using Taq DNA polymerase (Invitrogen, Carlsbad, CA) following manufacturer's recommendations and using primers previously reported [Ng et al., 2010]. PCR products were purified by treatment with exonuclease I (New England Biolabs, Inc., Beverly, MA) and shrimp alkaline phosphatase (USB Corp., Cleveland, OH), and products were sequenced using the dideoxy terminator method on an automated sequencer (ABI 3130xl). The electropherograms of both forward and reverse strands were manually reviewed using CodonCode Aligner (Dedham, MA). Primer sequences and conditions are listed in Supplementary Table I.

For *MLL2* mutation-negative samples, DNA was hybridized to commercially available whole-genome tiling arrays consisting of one million oligonucleotide probes with an average spacing of 2.6 kb throughout the genome (SurePrint G3 Human CGH Microarray 1 × 1 M, Agilent Technologies, Santa Clara, CA). Twenty-one probes on this array covered *MLL2* specifically. Data were analyzed using Genomics Workbench software according to manufacturer's instructions.

RESULTS

All 54 protein-coding exons and intron–exon boundaries of *MLL2* were screened by Sanger sequencing in a cohort of 110 kindreds with Kabuki syndrome. This cohort included 107 simplex cases (including a pair of monozygotic twins) and 3 familial (i.e., parent-offspring) cases putatively diagnosed with Kabuki syndrome. Seventy novel *MLL2* variants that were inferred to be disease-causing were identified in 81/110 (74%) kindreds (Fig. 1 and Supplementary Table II online). These 81 mutations included 37 nonsense mutations (32 different sites and five sites with recurrent mutations), 3 in-frame deletions or duplications (2 different sites and 1 site with a recurrent mutation), 22 frameshifts (22 different sites), 16 missense mutations (11 different sites and 4 sites with recurrent mutations), and 3 splice consensus site (or intron–exon boundary) mutations. None of these variants were found in dbSNP (build 132), the 1000 Genomes Project pilot data, or 190 chromosomes from individuals matched for geographical ancestry. In total, pathogenic variants were found at 70 sites. Additionally, there were 10 sites at which recurrent mutations were observed.

For 25 simplex cases in which we identified *MLL2* mutations, DNA was available from both unaffected parents, and in each case the mutation was confirmed to have arisen de novo (Supplementary Table II online). These included 14 nonsense, 5 frameshift, 3 missense, 2 splice site mutations, and 1 deletion. De novo events were confirmed at 6 of the 10 sites where recurrent mutations were noted. In addition to the 81 kindreds in which we identified causal *MLL2* mutations, we found two *MLL2* variants in each of three simplex cases. In each case, neither *MLL2* mutation could unambiguously

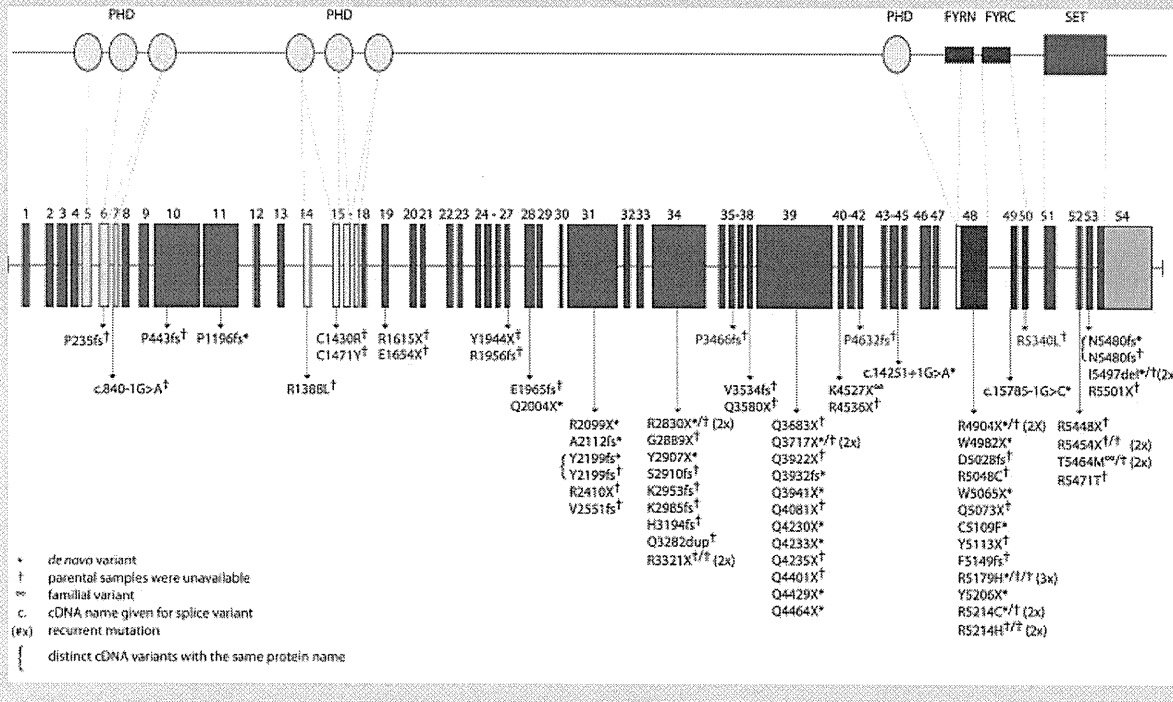

FIG. 1. Genomic structure and allelic spectrum of *MLL2* mutations that cause Kabuki syndrome. *MLL2* is composed of 54 exons that include untranslated regions (orange) and protein coding sequence (blue) including 7 PHD fingers (yellow), FYRN (green), FYRC (green), and a SET domain (red). Arrows indicate the locations of 81 mutations affecting 70 sites found in 110 families with Kabuki syndrome including: 37 nonsense, 22 frameshifts, 16 missense, 3 in-frame deletions/duplications, and 3 splice-site mutations. Asterisks indicate mutations that were confirmed to be de novo and crosses indicate cases for which parental DNA was unavailable. Figure adapted from Ng et al. [2010].

be defined as disease-causing (Supplementary Table II online). In one case, we found both a 21 bp in-frame insertion in exon 39 and a 1 bp insertion in exon 46 predicted to cause a frameshift. However, the unaffected mother also carried the 21 bp insertion suggesting that this is a rare polymorphism, and that the 1 bp deletion is the pathogenic mutation responsible for Kabuki syndrome.

Apparent disease-causing variants were discovered in nearly half (i.e., 22/54) of all protein-coding exons of *MLL2* and in virtually every region known to encode a functional domain (Fig. 1). However, the distribution of variants appeared non-random as 13 and 12 novel variants were identified in exons 48 and 39, respectively. These sites accounted for 25, or more than one-third, of all the novel *MLL2* variants and 31/81 mutations that cause Kabuki syndrome in our cohort. Eleven of the 12 pathogenic variants in exon 39 were nonsense mutations and occurred in regions that encode long polyglutamine tracts.

Four of the families studied herein had two individuals affected with Kabuki syndrome. A pair of monozygous twins with a c.15195G>A nonsense mutation were concordant for mild developmental delay, congenital heart disease, preauricular pits, and palatal abnormalities, but discordant for hearing loss, and a central nervous system malformation. Concordance for mild developmental delay between an affected parent and child was observed in two families with *MLL2* mutations, one with a nonsense mutation, c.13579A>T, p.K4527X, and the other with a missense mutation, c.16391C>T, p.T5464M that was also found in a simplex case. No *MLL2* mutation was found in the remaining affected parent and child pair (Fig. 2).

To examine the relationship between genotype and phenotype, we first compared the frequency of developmental delay, congenital heart disease, cleft lip and/or palate, and structural renal defects between *MLL2* mutation-positive versus *MLL2* mutation-negative cases. No significant difference was observed between groups for three of these four phenotypes (Table Ia). However, renal anomalies were observed in 47% (31/66 cases) of *MLL2* mutation-positive cases compared to 14% (2/14 cases) of *MLL2* mutation-negative cases and this difference was statistically significant ($\chi^2 = 5.1$, df = 1, $P = 0.024$). In 35 cases in two clinical cohorts for whom more complete phenotypic data were available, short stature was observed in 54% (14/26) of *MLL2* mutation-positive cases compared to 33% (3/19 cases) of *MLL2* mutation-negative cases. We also divided the *MLL2* mutation-positive cases into those with nonsense and frameshift mutations and those with missense mutations and compared the frequency of developmental delay, congenital heart disease, cleft lip and/or palate, and structural renal defects between groups. No significant differences were observed between groups (Table Ib).

In 26 independent cases of Kabuki syndrome, including one parent-offspring pair, no *MLL2* mutation was identified. Both persons in the mother–child pair had facial characteristics consistent with Kabuki syndrome (Fig. 2), mild developmental delay, and no major malformations. The mother is of Cambodian ancestry and her daughter is of Cambodian and European American ancestry. In general, most of the *MLL2* mutation-negative Kabuki cases had facial characteristics (Fig. 3) similar to those of the *MLL2* mutation-positive Kabuki cases, and a similar pattern of major malformations (Table I) with the exception of fewer renal abnormalities.

TABLE I. Phenotypic Traits Grouped by *MLL2* Mutation Status (a) and Type (b)

Trait	*MLL2* +	*MLL2* −
Intellectual disability	74/74 (100%)	19/20 (95%)
Mild	51/74 (69%)	10/20 (50%)
Moderate	18/74 (24%)	4/20 (20%)
Severe	4/74 (5%)	3/20 (15%)
Cleft palate, CL/CP	29/72 (40%)	8/18 (44%)
Congenital heart defect	36/71 (51%)	8/19 (42%)
Renal abnormality	31/66 (47%)	2/14 (14%)

Trait	Truncating (N = 59)	Missense (N = 16)
Intellectual disability	54/54 (100%)	15/15 (100%)
Mild	36/54 (67%)	11/15 (73%)
Moderate	13/54 (24%)	4/15 (27%)
Severe	5/54 (9%)	0/15
Cleft palate, CL/CP	23/54 (43%)	3/14 (21%)
Congenital heart defect	30/54 (55%)	4/13 (30%)
Renal anomaly	9/44 (20%)	2/12 (17%)

We screened the *MLL2* mutation-negative cases by aCGH for large deletions or duplications that encompassed *MLL2*. Abnormalities were found in four cases. In one case, a 1.87 kb deletion of chromosome 5 (hg18, chr5:175,493,803–177,361,744) that included *NSD1* and had breakpoints in flanking segmental duplications identical to the microdeletion commonly found in Sotos syndrome, was found. This suggests that this individual has Sotos syndrome, not Kabuki syndrome [Kurotaki et al., 2002]. A second case had a novel 977-kb deletion of chromosome 19q13 (hg18, chr19:61,365,420–62,342,064) encompassing 20 genes. The majority of genes within the deleted region are zinc finger genes, some of which are known to be imprinted in both human and mouse. A third case had a complex translocation t(8;18)(q22;q21). Finally, a fourth case was found to have extra material for the entire chromosome 12. Average log 2 ratio across chromosome 12 was 0.49, most likely representing mosaic aneuploidy of chromosome 12. No aCGH abnormalities were observed in 21 cases and aCGH failed for one case.

DISCUSSION

We have expanded the spectrum of mutations in *MLL2* that cause Kabuki syndrome and explored the relationship between *MLL2* genotype and some of the major, objective phenotypic characteristics of Kabuki syndrome. The majority of variants found to cause Kabuki syndrome are either novel nonsense or frameshift mutations, and appear to arise de novo. While mutations that cause Kabuki syndrome are found throughout the *MLL2* gene, there appear to be at least two exons (39 and 48) in which mutations are identified with a considerably higher frequency. Mutations in these two exons account for nearly half of all mutations found in *MLL2*, while the length of these exons represents ~24% of the *MLL2* open reading frame (ORF). Furthermore, exon 48, the exon in which mutations are most common, comprises only ~7% of the

FIG. 2. Facial photographs of mother and daughter with Kabuki syndrome in whom no causative mutation in *MLL2* was identified. Both have mild developmental delay and no known major malformations.

FIG. 3. Facial photographs of four children diagnosed with Kabuki syndrome in whom no causative mutation in *MLL2* was found. The photograph in the upper left was reprinted from Ng et al. [2010].

MLL2 ORF. Exon 39 contains several regions that encode long polyglutamine tracts suggesting the presence of a mutational hotspot, although no such explanation is obvious for exon 48. A stepwise approach in which these regions are the first screened might be a reasonable approach to diagnostic testing. However, capture of all introns, exons, and nearby *MLL2* regulatory regions followed by next-generation sequencing would be more comprehensive and likely to be less costly over the long term.

Comparison of four of the objective clinical characteristics of *MLL2* mutation-negative versus *MLL2* mutation-positive cases allowed us to explore both the relationship between *MLL2* genotype and Kabuki phenotype and the phenotype of *MLL2* mutation-negative cases. Overall, the clinical characteristics of *MLL2* mutation-positive cases did not differ significantly from *MLL2* mutation-negative cases with the exception that renal anomalies were more common in *MLL2* mutation-positive cases. Similarly, we observed no significant phenotypic—including the severity of developmental delay—differences between individuals grouped by mutation type. However, the phenotypic data available to us for analysis was limited and, for many cases, we lacked specific information about each malformation present. Furthermore, the most typical phenotypic characteristic, the distinctive facial appearance, was not compared in detail between cases although it would be of interest to study facial images "blinded" to mutation status to investigate its power to predict genotype. Analysis of genotype–phenotype relationships using both a larger set of Kabuki cases, and with access to more comprehensive phenotypic information would be valuable.

No *MLL2* mutation could be identified in 26 of the cases referred to us with a diagnosis of Kabuki syndrome. In three of these cases, aCGH identified structural variants that could be of clinical significance although additional investigation is required. A fourth case had the classical deletion observed in individuals with Sotos syndrome, and in retrospect it appears that this case was included in the cohort erroneously. The 22 remaining cases, including 1 parent-offspring pair, represent individuals with fairly classic phenotypic features of Kabuki syndrome without a *MLL2* mutation. This observation suggests that Kabuki syndrome is genetically heterogeneous. To this end, in these 22 cases, we sequenced the protein-coding exons of *UTX*, a gene that encodes a protein that directly interacts with *MLL2* but no pathogenic changes were found (data not shown). Exome sequencing of a subset of these *MLL2* mutation-negative cases to identify other candidate genes for Kabuki syndrome is underway.

Whether Kabuki syndrome is the most appropriate diagnosis for the *MLL2* mutation-negative cases is unclear. Some of the *MLL2* mutation-negative cases appear to have a facial phenotype that differs somewhat from that of the *MLL2* mutation-positive cases. Whether these *MLL2* mutation-negative cases diagnosed by expert clinicians should be considered Kabuki syndrome, a variant thereof, or a separate disorder remains to be determined. Our opinion is that

there is simply not yet enough information to make an informed decision about this issue.

Most of the mutations in *MLL2* are predicted to result in haploinsufficiency. However, it is unclear by what mechanism(s) haploinsufficiency of *MLL2* could cause Kabuki syndrome. *MLL2* encodes a histone 3 lysine 4 (H3K4) methyltransferase, one of at least 10 proteins (genes for which have not to our knowledge yet been screened in Kabuki cases in which *MLL2* mutations were not found) that have been identified to specifically modify the lysine residue at the fourth amino acid position of the histone H3 protein [Kouzarides, 2007]. MLL2 has a SET domain near its C-terminus that is shared by yeast Set1, *Drosophila* Trithorax (TRX) and human MLL1 [FitzGerald and Diaz, 1999]. MLL2 appears to regulate gene transcription and chromatin structure in early development [Prasad et al., 1997]. In mice, loss of *MLL2* results in embryonic lethality before E10.5, and while $Mll2^{+/-}$ mice are viable, they are smaller than wild-type [Ng et al., 2010].

Kabuki syndrome is the most common of a small, but growing group of multiple malformation syndromes accompanied by developmental delay that are caused by mutations in genes that encode proteins involved in histone methylation [De Sario, 2009]. The most notable of these is CHARGE syndrome, which is one of the syndromes often considered in the differential diagnosis of children ultimately diagnosed with Kabuki syndrome. CHARGE syndrome is caused by mutations in *CHD7*, which encodes a chromodomain protein that recognizes the trimethylated H3K4 side chain [Vissers et al., 2004]. Other disorders caused by defects of histone methylation status include several intellectual disability syndromes, some of which are also characterized by malformations (e.g., cleft lip/palate) that overlap with those found in individuals with Kabuki syndrome.

Kabuki syndrome is one of the most common causes of heritable developmental delay. Discovery that mutations in *MLL2* are the most common cause of Kabuki syndrome highlights the role that disrupted regulation of histone methylation plays as a cause of human birth defects. Characterizing the spectrum of mutations in *MLL2* is a small but important first step toward understanding the mechanism(s) that underlies Kabuki syndrome.

ACKNOWLEDGMENTS

We thank the families for their participation and the Kabuki Syndrome Network for their support. Our work was supported in part by grants from the National Institutes of Health/National Heart Lung and Blood Institute (5R01HL094976 to D.A.N. and J.S.), the National Institutes of Health/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.), National Institute of Heath/National Institute of Environmental Health Sciences (HHSN273200800010C to D.N.), National Institute of Neurological Disorders and Stroke (RO1NS35102 to C.A.M.), NIHR Manchester Biomedical Research Centre (D. D.), Ministry of Health, Labour and Welfare (K.Y., N.M., T.O., and N.N.), Ministry of Health, Labour and Welfare of Japan (N.M), 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 National Institutes of Health/National Institute of Child Health and Human Development (1R01HD048895 to M.J.B. and 5K23HD057331 to A.E.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 National Institutes of Health/National Human Genome Research Institute (T32HG00035).

REFERENCES

Adam MP, Hudgins L. 2005. Kabuki syndrome: A review. Clin Genet 67: 209–219.

De Sario A. 2009. Clinical and molecular overview of inherited disorders resulting from epigenomic dysregulation. Eur J Med Genet 52:363–372.

FitzGerald KT, Diaz MO. 1999. ML L2: A new mammalian member of the trx/MLL family of genes. Genomics 59:187–192.

Kouzarides T. 2007. Chromatin modifications and their function. Cell 128:693–705.

Kuroki Y, Suzuki Y, Chyo H, Hata A, Matsui I. 1981. 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.

Kurotaki N, Imaizumi K, Harada N, Masuno M, Kondoh T, Nagai T, Ohashi H, Naritomi K, Tsukahara M, Makita Y, Sugimoto T, Sonoda T, Hasegawa T, Chinen Y, Tomita Ha, Kinoshita HA, Mizuguchi A, Yoshiura T, Ki K, Ohta T, Kishino T, Fukushima Y, Niikawa N, Matsumoto N. 2002. Haploinsufficiency of NSD1 causes Sotos syndrome. Nat Genet 30:365–366.

Ng SB, Bigham AW, Buckingham KJ, Hannibal MC, McMillin MJ, Gildersleeve HI, Beck AE, Tabor HK, Cooper GM, Mefford HC, Lee C, Turner EH, Smith JD, Rieder MJ, Yoshiura K, Matsumoto N, Ohta T, Niikawa N, Nickerson DA, Bamshad MJ, Shendure J. 2010. Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. Nat Genet 42:790–793.

Niikawa N, Matsuura N, Fukushima Y, Ohsawa T, Kajii T. 1981. Kabuki make-up syndrome: A syndrome of mental retardation, unusual facies, large and protruding ears, and postnatal growth deficiency. J Pediatr 99:565–569.

Niikawa N, Kuroki Y, Kajii T, Matsuura N, Ishikiriyama S, Tonoki H, Ishikawa N, Yamada Y, Fujita M, Umemoto H, et al. 1988. Kabuki make-up (Niikawa-Kuroki) syndrome: A study of 62 patients. Am J Med Genet 31:565–589.

Prasad R, Zhadanov AB, Sedkov Y, Bullrich F, Druck T, Rallapalli R, Yano T, Alder H, Croce CM, Huebner K, Mazo A, Canaani E. 1997. 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.

Vissers LE, van Ravenswaaij CM, Admiraal R, Hurst JA, de Vries BB, Janssen IM, van der Vliet WA, Huys EH, de Jong PJ, Hamel BC, Schoenmakers EF, Brunner HG, Veltman JA, van Kessel AG. 2004. Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. Nat Genet 36:955–957.

White SM, Thompson EM, Kidd A, Savarirayan R, Turner A, Amor D, Delatycki MB, Fahey M, Baxendale A, White S, Haan E, Gibson K, Halliday JL, Bankier A. 2004. Growth, behavior, and clinical findings in 27 patients with Kabuki (Niikawa-Kuroki) syndrome. Am J Med Genet Part A 127A:118–127.

Proteasome assembly defect due to a proteasome subunit beta type 8 (PSMB8) mutation causes the autoinflammatory disorder, Nakajo-Nishimura syndrome

Kazuhiko Arima[a,1], Akira Kinoshita[b,1], Hiroyuki Mishima[b,1], Nobuo Kanazawa[c,1], Takeumi Kaneko[d], Tsunehiro Mizushima[e], Kunihiro Ichinose[a], Hideki Nakamura[a], Akira Tsujino[f], Atsushi Kawakami[a], Masahiro Matsunaka[c], Shimpei Kasagi[g], Seiji Kawano[g], Shunichi Kumagai[g], Koichiro Ohmura[h], Tsuneyo Mimori[h], Makito Hirano[i], Satoshi Ueno[i], Keiko Tanaka[j], Masami Tanaka[k], Itaru Toyoshima[l], Hirotoshi Sugino[m], Akio Yamakawa[n], Keiji Tanaka[o], Norio Niikawa[p], Fukumi Furukawa[c], Shigeo Murata[d], Katsumi Eguchi[a], Hiroaki Ida[a,q,2], and Koh-ichiro Yoshiura[b,2]

[a]Unit of Translational Medicine, Department of Immunology and Rheumatology, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki 852-8501, Japan; [b]Department of Human Genetics, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki 852-8523, Japan; [c]Department of Dermatology, Wakayama Medical University, Wakayama 641-0012, Japan; [d]Laboratory of Protein Metabolism, Graduate School of Pharmaceutical Sciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan; [e]Department of Life Science, Picobiology Institute, Graduate School of Life Science, University of Hyogo, Kamigori-cho, Ako-gun, Hyogo 678-1297, Japan; [f]Unit of Translational Medicine, Department of Neuroscience and Neurology, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki 852-8501; [g]Department of Clinical Pathology and Immunology, Kobe University Graduate School of Medicine, Kobe 650-0017, Japan; [h]Department of Rheumatology and Clinical Immunology, Graduate School of Medicine, Kyoto University, Kyoto 606-8507, Japan; [i]Department of Neurology, Nara Medical University, Kashihara, Nara 634-8522, Japan; [j]Department of Neurology, Kanazawa Medical University, Kahoku-gun, Ishikawa 920-0293, Japan; [k]Department of Neurology, Utano National Hospital, Ukyou-ku, Kyoto 616-8255, Japan; [l]Department of Neurology and Medical Education Center, Akita University School of Medicine, Akita 010-8543, Japan; [m]Sugino Pediatric Clinic, Asakita-ku, Hiroshima 731-0231, Japan; [n]Office of Strategic Management, Institute of Medical Science, The University of Tokyo, Minato-ku, Tokyo 108-8639, Japan; [o]Laboratory of Protein Metabolism, Tokyo Metropolitan Institute of Medical Science, Setagaya-ku, Tokyo 156-8506, Japan; [p]Research Institute of Personalized Health Sciences, Health Sciences University of Hokkaido, Ishikari-Tobetsu, Hokkaido 061-0293, Japan; and [q]Division of Respirology, Neurology, and Rheumatology, Department of Medicine, Kurume University School of Medicine, Kurume, Fukuoka 830-0011, Japan

Edited* by Daniel Kastner, National Institutes of Health, Bethesda, MD, and approved July 21, 2011 (received for review April 14, 2011)

Nakajo-Nishimura syndrome (NNS) is a disorder that segregates in an autosomal recessive fashion. Symptoms include periodic fever, skin rash, partial lipomuscular atrophy, and joint contracture. Here, we report a mutation in the human proteasome subunit beta type 8 gene (PSMB8) that encodes the immunoproteasome subunit β5i in patients with NNS. This G201V mutation disrupts the β-sheet structure, protrudes from the loop that interfaces with the β4 subunit, and is in close proximity to the catalytic threonine residue. The β5i mutant is not efficiently incorporated during immunoproteasome biogenesis, resulting in reduced proteasome activity and accumulation of ubiquitinated and oxidized proteins within cells expressing immunoproteasomes. As a result, the level of interleukin (IL)-6 and IFN-γ inducible protein (IP)-10 in patient sera is markedly increased. Nuclear phosphorylated p38 and the secretion of IL-6 are increased in patient cells both in vitro and in vivo, which may account for the inflammatory response and periodic fever observed in these patients. These results show that a mutation within a proteasome subunit is the direct cause of a human disease and suggest that decreased proteasome activity can cause inflammation.

Nakajo-Nishimura syndrome (NNS) (MIM256040, ORPHA-2615) is a distinct inflammatory and wasting disease. It was first reported by Nakajo in 1939, followed by Nishimura in 1950, and was called "secondary hypertrophic osteoperiostosis with pernio" (1, 2). More than 20 cases of this disease have been reported in various clinical fields, all from Japan (3–8). The disease was soon recognized as a new entity and was called "a syndrome with nodular erythema, elongated and thickened fingers, and emaciation" or "hereditary lipomuscular atrophy with joint contracture, skin eruptions and hyper-γ-globulinemia" on the basis of the common characteristic features (3, 4).

NNS usually begins in early infancy with a pernio-like rash. The patient develops periodic high fever, nodular erythema-like eruptions, and myositis. Lipomuscular atrophy and joint contractures gradually progress, mainly in the upper body, to form the characteristic thin facial appearance and elongated clubbed fingers. Inflammatory changes are marked and include constantly elevated erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP), hyper-γ-globulinemia, hepatosplenomegaly, basal ganglia calcification, and focal mononuclear cell infiltration with vasculopathy on histopathology. Autoantibodies are negative at the onset of NNS; although, in some cases, titers increase as the disease progresses.

Although NNS bears similarities to other autoimmune diseases, particularly dermatomyositis, it is only in recent years that its similarity to autoinflammatory periodic fever syndromes has been pointed out (5, 6). Oral steroids are effective in treating the inflammation, but not the wasting, and most patients die as a result of respiratory or cardiac failure. Despite the predicted segregation in an autosomal recessive fashion, the gene responsible has not been identified. Here, we describe a mutation in the human *PSMB8* that encodes the immunoproteasome subunit β5i in NNS patients.

Proteasomes collaborate with the ubiquitin system, which tags proteins with a polyubiquitin chain and marks them for degradation. The 26S proteasome is a multisubunit protease responsible for regulating proteolysis in eukaryotic cells in collaboration with the ubiquitin system. This ubiquitin–proteasome system is involved in various biological processes, including immune responses, DNA repair, cell cycle progression, transcription and protein quality control. It comprises a single catalytic 20S proteasome with 19S regulatory particles (RPs) attached to the ends (9–11). The 20S proteasome comprises 28 subunits arranged as a cylindrical particle containing four heteroheptameric rings: $\alpha_{1-7}\beta_{1-7}\beta_{1-7}\alpha_{1-7}$. Only three of the β subunits, β1, β2, and β5, are proteolytically active in the standard 20S pro-

Author contributions: N.K., A.Y., N.N., F.F., S.M., K.E., H.I., and K.-i.Y. designed research; K.A., A. Kinoshita, H.M., N.K., T.K., T. Mizushima, K.I., H.N., A.T., A. Kawakami, M.M., S. Kasagi, S. Kawano, S. Kumagai, K.O., T. Mimori, M.H., S.U., Keiko Tanaka, M.T., I.T., H.S., S.M., H.I., and K.-i.Y. performed research; K.A., A. Kinoshita, H.M., N.K., A.Y., S.M., H.I., and K.-i.Y. analyzed data; and N.K., Keiji Tanaka, F.F., S.M., H.I., and K.-i.Y. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

[1]K.A., A. Kinoshita, H.M., and N.K. contributed equally to this work.

[2]To whom correspondence may be addressed. E-mail: kyoshi@nagasaki-u.ac.jp or ida@med.kurume-u.ac.jp.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1106015108/-/DCSupplemental.

teasome. Each of the three β subunits preferentially cleaves an acidic, basic, or hydrophobic residue, activities often referred to as caspase-like, trypsin-like, or chymotrypsin-like, respectively.

In vertebrates, there are three additional IFN-γ–induced subunits: β1i, β2i, and β5i. These are preferentially incorporated into the 20S proteasome in place of the standard subunits to form the immunoproteasome in immune cells such as macrophages, T and B cells, and dendritic cells, whereas their expression is low in nonlymphoid peripheral tissues. This results in more efficient production of MHC class I epitopes (12). The present study analyzed the activity of proteasomes with a mutated β5i subunit, and the subsequent inflammatory signal transduction pathways in mutant cells. The results suggest that the *PSMB8* mutation evokes an inflammatory response in humans, and that the p38 pathway may play an important role in inflammation in NNS patients.

Recently, a different mutation in the *PSMB8* gene was reported in patients with a disease similar to, but distinct from, NNS: an autosomal recessive syndrome of joint contracture, muscular atrophy, microcytic anemia, and panniculitis-associated lipodystrophy (JMP) (13, 14). The mutation in JMP syndrome, T75M, causes a reduction in chymotrypsin-like activity only, without disrupting the activity of other peptidases (13). In contrast, the G201V mutation identified in NNS patients results in the loss of all peptidase activity because of assembly defects and reduced proteasome levels. Thus, the discovery of *PSMB8* mutations in these related diseases indicates the presence of a distinct class of proteasome-associated autoinflammatory disorders.

Results

Clinical Features of NNS Patients. National surveillance in Japan confirms that only around 10 NNS patients are alive today. Therefore, preserved fibroblasts from an autopsy case (patient 1) were provided for genetic analysis, following approval by the local ethical committee. Of the living cases, written informed consent to undertake genetic and molecular analyses was obtained from six patients. The clinical features of all seven cases are summarized in Table S1. Patients 1, 2, and 4 were born to consanguineous parents and their clinical features have been described previously (Fig. S1*A*) (6, 8). The other patients are sporadic cases collected for this study and were born in the limited area between south Osaka and Wakayama. A diagnosis of NNS is not difficult owing to the characteristic features, including the thin facial appearance and long clubbed fingers (Fig. S1*B*). The clinical course throughout childhood was variable: from no medical consultation in the case of patient 7, to administration of oral steroids since infancy in patients 3 and 6. Partial lipomuscular atrophy with long clubbed fingers plus a pernio-like, heliotrope-like, or nodular erythema-like skin rash were observed in all cases, and periodic fever and joint contractures in most but not all. Wheras hyperhidrosis was also observed in some cases, short stature and low IQ were seen only in patients 6 and 1, respectively. Indeed, patient 6 was treated with growth hormone, although growth retardation in this case may have been due, in part, to oral steroids. Chronic inflammation, indicated by elevated ESR and hyper-γ-globulinemia, were observed in all patients, and microcytic anemia, high serum creatine phosphokinase (CPK), hepatosplenomegaly, and basal ganglia calcification were present in most, but not all. Notably, various autoantibodies (with a mildly elevated titer of antinuclear antibodies) were detected in half of the patients. The most striking differences between NNS and JMP are the absence of fever in JMP syndrome and the absence of seizures in NNS (14) (Table S1).

Genetic Mapping and Mutation Searches. We examined genomic DNA samples from five patients (patients 1–5) and three unaffected siblings of patient 4 using an Affymetrix GeneChip Human Mapping 500K array set (Nsp I and Sty I arrays), and the BRLMM genotyping algorithm. Because the runs of homozygosity (ROHs) shared by all patients were expected to be candidate regions containing the gene responsible for the disease, we identified a region spanning 1.1 Mb on chromosome 6p21.31–32 [from 32,798,004–33,903,106; National Center for Biotechnology Information (NCBI) build 36.1] as the sole candidate region responsible for NNS (Fig. 1*A*). We directly sequenced 436 coding exons in the 44 genes within this candidate region, including the splicing sites. A single nonsynonymous variation (not registered in the dbSNP database) was identified in exon 5 of *PSMB8* (NM_148919 in the NCBI database), designated *LMP7* or *RING10*, which encodes the LMP7 protein (β5i subunit) of the immunoproteasome. This mutation was a guanine to thymine transversion at nucleotide position 602 (c.602G > T) (Fig. 1*B*). Haplotype analysis indicated that the G201V mutation was probably introduced into the Japanese population by a single founder, as the haplotype around this mutation was identified in all patients (Fig. S1*C*). Gly201, which is a highly conserved residue in the β5i subunit (Fig. 1*C*) and among mature proteasome subunits in vertebrates (Fig. S1*D*), is substituted by Val (G201V) (Fig. 1*B*).

Impaired Immunoproteasome Assembly and Peptidase Activity. In silico modeling of the mutant β5i ($\beta5i^{G201V}$) subunit was used to infer the conformational impact of this mutation because the assembly of the proteasome is a highly orchestrated and complex process (9, 15). The β5i subunit is cleaved between amino acid residues Gly72 and Thr73 to yield the active form (16), in which the catalytic center is generated by Thr73, Asp89, Arg91, and Lys105. The mutated residue at position 201 was located at the edge of the S8 β-sheet of β5i and was close to its catalytic threonine residue Thr73 (Fig. 2*A*). The G201V substitution caused conformational changes not only in Thr73 but also in Lys105 within the catalytic center (Fig. S2). The mutation resulted in further conformational changes in the S8–H3 loop located at the

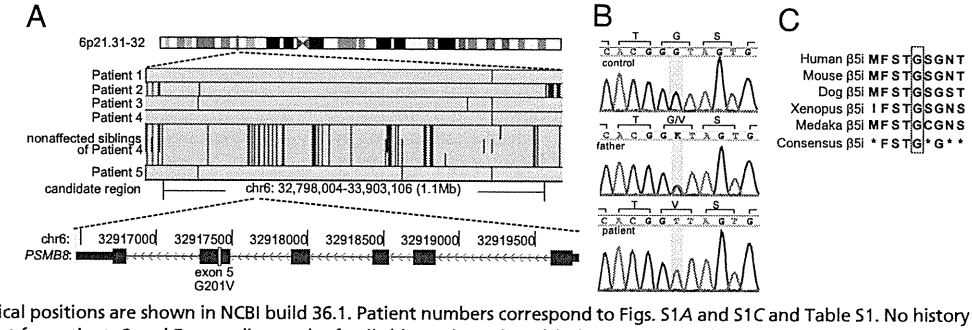

Fig. 1. SNP microarray-based homozygosity mapping and mutation search. (*A*) Homozygosity mapping for NNS patients and nonaffected siblings. ROH regions were detected using a hidden Markov model-based algorithm. The sole candidate region identified within 6p21.31–32 is shown. Green vertical lines indicate heterozygous SNPs and the background gray area indicates a region without heterozygous SNP calls. To be conservative, we did not regard isolated single heterozygous calls as delimiting ROH regions. The physical positions are shown in NCBI build 36.1. Patient numbers correspond to Figs. S1*A* and S1*C* and Table S1. No history of consanguineous marriage was apparent for patients 3 and 5, according to the family history interview. (*B*) Chromatograms for a control, a patient's father, and a patient. A mutation in *PSMB8* exon 5 identified in NNS patients by sequencing is highlighted in yellow. (*C*) Amino acid comparisons with other species. The glycine at the mutation site (red box) is highly conserved among vertebrates.

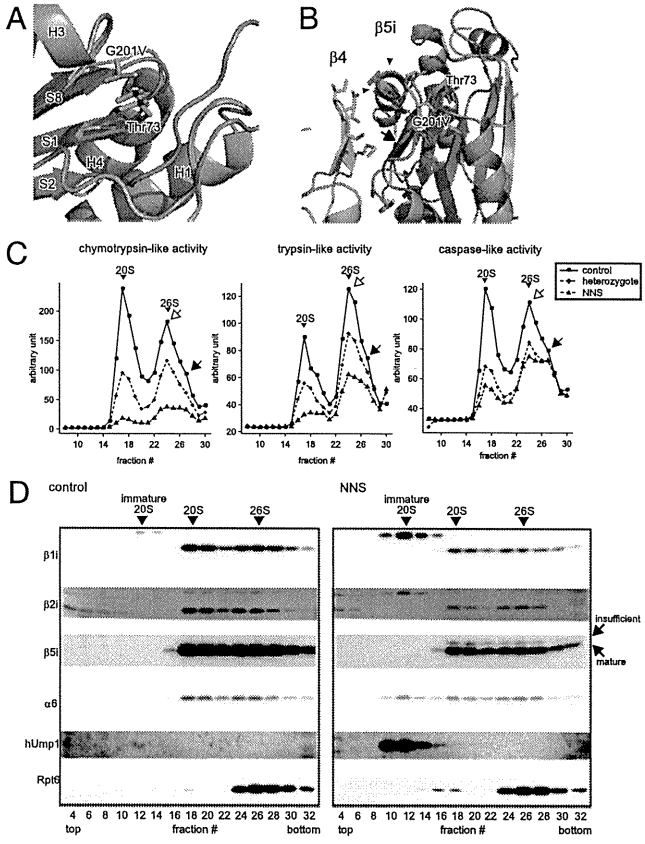

Fig. 2. G201V mutation in β5i reduces proteasome activity in immunoproteasome-expressing cells. (*A*) Close-up view of the mutation site (G201V) within β5i. Structural models of G201V β5i (orange) and wild-type β5i (green) were created from the β5-subunit structure [Protein Data Bank (PDB) ID code 1IRU]. The secondary structure elements for β5i are labeled. Val201 and Thr73 are shown in the stick model. Thr73 is a catalytic residue of β5i. (*B*) A ribbon diagram of the β4–β5i complex. The arrow shows the difference in the β-sheet between β5i (green) and $\beta5i^{G201V}$ (orange). Arrowheads show the protruding S8–H3 loop of $\beta5i^{G201V}$. (*C*) Peptidase activity of LCLs. Extracts were fractionated by glycerol gradient centrifugation (8–32% glycerol from fraction 1–32). Arrowheads indicate the peak positions of the 20S and 26S proteasomes (open arrows, single-capped 26S; closed arrows, double-capped 26S). (*D*) Western blot analysis of fractionated total LCL extracts. Western blot analysis of proteasome subunits from fractions 1–32 fractionated in *C*. The sedimenting positions of the immature 20S, 20S, and 26S proteasomes are indicated by arrowheads. The mature and incompletely cleaved $\beta5i^{G201V}$ subunits are indicated by arrows. The mature β5i subunit is cleaved within a C-terminal polypeptide between Gly72 and Thr73. The insufficiently cleaved β5i subunit is probably cleaved at a site toward the N terminus site, yielding a fragment with a higher molecular weight. The same amount of protein was subjected to glycerol gradient ultracentrifugation. The level of proteasome is reduced in NNS patients. Control, LCL extract from healthy control; NNS, LCL extract from patient with NNS.

interface between β4 and β5i, which affected the surface contact of β5i with the adjacent β4 subunit (Fig. 2*B*). These results suggest that the G201V mutation affects both β5i catalytic activity and assembly of the 20S proteasome.

According to Sijts and Kloetzel (17), the β1 subunit has a caspase-like function, the β2 subunit has trypsin-like activity, and the β5 subunit has chymotrypsin-like activity. Although it has not been clearly confirmed which of the immunoproteasome subunits possess which peptidase activity, it is generally thought that β5i has chymotrypsin-like activity. We next examined the influence of the β5i mutation on proteasome peptidase activity. Extracts from immortalized lymphoblastoid cell lines (LCLs) that constitutively expressed the immunoproteasome, rather than the standard proteasome, were obtained from an NNS patient, his heterozygous parent, and a healthy control, and were separated by glycerol gradient centrifugation. The fractions were then assayed for chymotrypsin-like, trypsin-like, and caspase-like activity mediated by the 20S/26S proteasomes. The results showed that not only was chymotrypsin-like activity markedly decreased in NNS cells, but the other two enzyme-like activities were also decreased (Fig. 2*C*).

Reduced Proteasome Levels. To gain further insight into the molecular mechanisms affecting peptidase activity in the mutant cells, the glycerol density gradient fractions were subjected to Western blot analysis (Fig. 2*D*). Assembly of the mammalian 20S proteasome begins with the formation of the α-ring in conjunction with a dedicated assembly chaperone, PAC1-4. The β-ring is then formed on the α-ring with the aid of another chaperone, hUmp1, resulting in the formation of half-sized immature proteasomes. The immature proteasomes then dimerize to form the 20S proteasome accompanied by cleavage of β-subunit propeptides and the degradation of hUmp1 (9). Our most noteworthy finding was the accumulation of immature 20S proteasome precursors in NNS cells before incorporation of β5i and dimerization, as indicated by the presence of the proforms of β1i and β2i, α6 and hUmp1, and the absence of β5i (Fig. 2*D*, fractions 10–14) (18). Computer modeling suggests that this assembly defect could be due to the fact that β5i, β4, and β6 line up next to each other and that the interaction between mutant $\beta5i^{G201V}$ and β4 may be disturbed (Fig. 2*B*). The reduction in peptidase activity was unlikely due to differences in the ability of 20S to associate with 19S RP, because single-capped and double-capped 26S proteasomes were detected in the glycerol fractions from an NNS patient and control LCLs (Fig. 2*C*). The assembly defect caused a reduction in the number of 20S and 26S proteasomes in NNS cells (Fig. 2*D*), which accounts for the observed decrease in activity of all three peptidases. Another intriguing observation was that a portion of the $\beta5i^{G201V}$ subunit incorporated into the mature proteasome appeared as a slower migrating band, suggesting the presence of an insufficiently cleaved form of $\beta5i^{G201V}$ (Fig. 2*D*) (16). This may have contributed to the markedly reduced chymotrypsin-like activity seen in NNS cells compared with the other two peptidase activities.

Decreased Proteolytic Activity and Accumulation of Ubiquitinated and Oxidized Proteins. To examine proteolytic activity in vitro, the ornithine decarboxylase (ODC) degradation assay was performed (19). Proteolytic activity was significantly decreased in mutant proteasomes (Fig. 3*A*). As a consequence of the altered proteasome levels and incomplete cleavage of the subunits, proteolytic activity decreased and ubiquitinated proteins accumulated in LCLs (Fig. 3*B*) and fibroblasts from NNS patients (Fig. 3*C*). In particular, there was an obvious accumulation of K48 polyubiquitinated proteins in fibroblasts (Fig. 3*C*).

Because the immunoproteasome is important for degrading oxidized proteins and defective ribosomal products (20), we examined whether such proteins accumulated in NNS cells. We found that the level of oxidized proteins increased in cultured NNS fibroblasts and after stimulation with IFN-γ (Fig. 3*D*). Taken together, these results show that the G201V substitution within β5i severely impairs assembly of the immunoproteasome, leading to decreased proteasome levels and activity in β5i-expressing cells.

We then examined whether the defect in proteasome activity was apparent in situ in NNS patients. We stained skin biopsy sections obtained from an NNS patient and used sections from a monocytic fasciitis patient as a control. CD68 is a marker for monocyte/macrophages, a cell type known to predominantly express the immunoproteasome rather than the standard proteasome (21). Inflammatory responses characterized by the infiltration of numerous $CD68^+$ cells into the skin were observed in both NNS and fasciitis samples. However, the $CD68^+$ cells in the NNS sec-

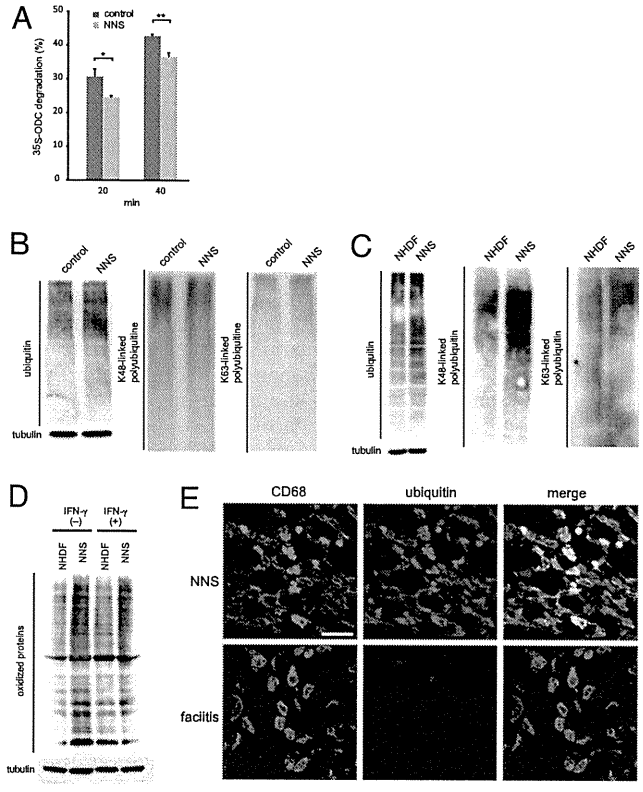

Fig. 3. Decrease of proteolytic activity and accumulation of polyubiquitinated and oxidized proteins in NNS cells. (*A*) In vitro proteolytic activity of the mutant proteasome. Degradation of recombinant ^{35}S-labeled ODC was expressed as % total ODC as described previously (11). Error bars indicated the SD of the mean ($n = 3$). $*P < 0.05$, $**P < 0.01$. (*B* and *C*) Accumulation of ubiquitinated proteins in LCLs (*B*) and fibroblasts (*C*). Western blot analysis of ubiquitinated proteins using an antiubiquitin antibody (*Left*), an anti-K48 polyubiquitinated protein antibody (*Middle*), and an anti-K63 polyubiquitinated protein antibody (*Right*). Tubulin was used as a loading control (*Lower*). NHDF, adult normal human dermal fibroblasts. (*D*) Levels of oxidized proteins determined by Oxyblot. NHDF and NNS fibroblasts were stimulated with or without 100 units of IFN-γ for 24 h. Tubulin was used as a loading control. (*E*) Immunofluorescence staining of CD68 and ubiquitinated proteins. Staining for CD68 (green) and ubiquitinated proteins (red) in skin sections from an NNS patient and a fasciitis patient. NNS ubiquitin signals showed a 4.7-fold increase with ImageJ (http://rsb.info.nih.gov/ij/) compared with fasciitis signals. (Scale bar, 10 μm.)

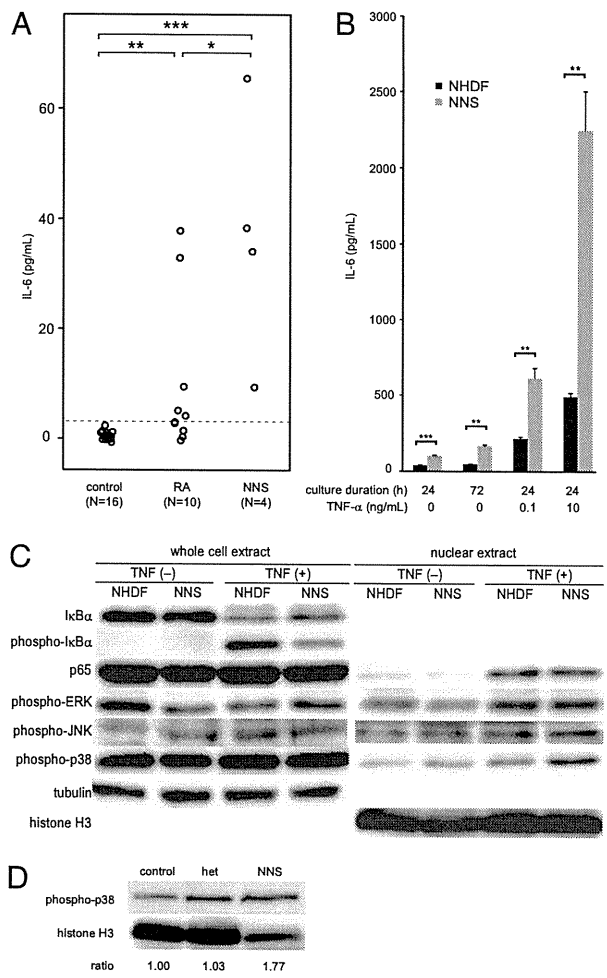

Fig. 4. Analyses of the level of IL-6 in NNS and the signal transduction system related to cytokine production. (*A*) IL-6 concentrations in sera from healthy controls, patients with NNS, and patients with rheumatoid arthritis. IL-6 levels in sera were determined by ELISA. (*B*) IL-6 production by cultured fibroblasts. The concentrations of IL-6 in conditioned media were determined by ELISA (in triplicate). (*C*) Western blot analysis for NF-κB and MAPK. Whole cell extracts and nuclear extracts were immunoblotted using antibodies against IκBα, p-IκBα, p65, p-ERK, p-JNK, and p-p38. (*D*) Western blot analysis of p-p38 in peripheral blood lymphocytes. Nuclear extracts from the peripheral blood lymphocytes of a healthy control, a heterozygous family member, and a NNS patient were blotted and visualized with anti–p-p38. Error bars indicate SD of the mean. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ [Mann-Whitney *u* test (*A*) and two-tailed Welch's *t* test (*B*)]. Signal intensities were quantified using ImageJ and expressed as fold changes relative to a healthy control normalized to histone H3 (*D*).

tions were strongly positive for ubiquitin, whereas ubiquitin was only faintly detectable in the fasciitis sections (Fig. 3*E*).

Increased IL-6 and IP-10 Levels in NNS Patient Sera and Signal Transduction in NNS Fibroblasts. We next screened NNS patient sera for inflammatory cytokines using a multiplex bead-based ELISA on a suspension array. The results showed a significant increase in the levels of interleukin (IL)-6, IFN-γ–inducible protein (IP)-10, granulocyte colony stimulating factor, and monocyte chemoattractant protein-1 (Fig. S3*A*). IL-6 was of particular interest because it is a pleiotropic cytokine with a wide range of biological activities, and it plays a key role of immune regulation, hematopoiesis, oncogenesis, and inflammation (22–24). Increased IL-6 levels in NNS sera were confirmed using a standard ELISA (Fig. 4*A*). IL-6 production was significantly higher in NNS patient fibroblasts than in healthy control fibroblasts both in the presence and absence of TNF-α (Fig. 4*B*). The serum concentration of IP-10 was also higher than that in healthy controls (Fig. S3 *A* and *B*). We measured the level of IP-10 in conditional media from cultured fibroblasts using an ELISA, but found no significant difference under the conventional culture condition, although NNS cells tended to overproduce IP-10 after stimulation with 10 ng/mL TNF-α (Fig. S3*C*).

We next investigated the various signal transduction pathways that could be responsible for IL-6 overproduction by NNS fibroblasts. Nuclear factor (NF)-κB and AP-1 are the two major transcription factors that induce proinflammatory cytokines, including IL-6 (25, 26). We used an EMSA to detect activated NF-κB in cells treated with TNF-α; however, no differences in the amount of the p65/p50 heterodimer were observed in nuclear extracts from NNS fibroblasts and healthy control fibroblasts (Fig. S4 *A* and *B*). Consistent with this result, IκBα degradation and nuclear translocation of NF-κB were not enhanced in NNS fibroblasts (Fig. 4*C*). Although activation of NF-κB is largely dependent on the ubiquitin–proteasome system, these results

suggest that decreased proteasome activity does not have much influence on the regulation of NF-κB signaling in NNS cells.

We next measured the molecules that activate AP-1, including JNK1/2/3, ERK1/2, and p38, by Western blot analysis (27, 28). The amount of phosphorylated p38 (p-p38) in the nuclear extracts from NNS fibroblasts was increased (Fig. 4*C*), irrespective of TNF-α stimulation; however, there was no obvious difference in the levels of JNK1/2/3 and ERK1/2 (Fig. 4*C*). We also observed increased levels of p-p38 in the nuclear extracts from NNS peripheral blood lymphocytes (Fig. 4*D*). The build-up of oxidized proteins and/or reactive oxygen species (ROS) within NNS fibroblasts may be one of the mechanisms responsible for the accumulation of p-p38 (29, 30).

Discussion

We have identified a point mutation in the gene encoding the immunoproteasome subunit β5i as the cause of NNS. This mutation interferes with the assembly of the 20S proteasome in cells expressing immunoproteasomes. The mutation is described as c.602G > T, and results in a Gly201 to Val (G201V) (NM_148919) substitution in the immunoproteasome β5i subunit. Although a heterozygous carrier showed reduced proteasome peptidase activity, carriers had no clinical symptoms. Thus, the NNS phenotype may be due to a reduction in total proteasome enzymatic activity below the threshold necessary for maintaining cellular homeostasis in homozygous individuals.

The *PSMB8* mutation, c.224C > T (Thr75Met), occurs in patients with JMP syndrome (13). Mutant β5i in JMP patients results in a clear reduction in chymotrypsin-like activity only, with no disruption of other peptidase activities. However, the G201V mutation we identified in NNS patients causes losses of all peptidase activity owing to assembly defects and reduced proteasome levels. The T75M mutation is probably rapidly incorporated to the proteasome complex during biogenesis and is specific for chymotrypsin-like activity. The differences between the JMP syndrome and NNS phenotypes, including cytokine production by various cells during inflammatory or noninflammatory states, need to be clarified because these differences could result from a reduction in chymotrypsin-like activity in JMP syndrome or from reductions in chymotrypsin-, trypsin-, and caspase-like activity in NNS. One of the main differences between NNS and JMP syndrome is the level of IFN-γ. IFN-γ levels are increased in JMP patients, but are within the normal range in NNS patients (Fig. S3*A*). The basis for this difference is unclear. It is possible that IFN-γ levels may not increase when all three peptidase activities are inhibited.

We also found increased IP-10 levels in patient sera using ELISA on suspension arrays. There were no significant differences in IP-10 levels between nonstimulated NNS fibroblasts and control cells, although NNS fibroblasts tended to overproduce IP-10 after stimulation with TNF-α (Fig. S3 *B* and *C*). This may reflect the proinflammatory state in NNS cells, or an increased sensitivity to cytokines (31). Because IP-10 is categorized as an inflammatory chemokine produced by various types of cells, it may play an important role in leukocyte homing to inflamed tissues and in perpetuating inflammation in various autoimmune diseases such as rheumatoid arthritis, systemic lupus erythematosus, systemic sclerosis, and multiple sclerosis (32). Thus, IP-10 may enhance inflammation in NNS patients and be associated with the autoantibody production that is occasionally observed.

A single base deletion in the 5′-UTR of hUmp1 causes keratosis linearis with ichthyosis congenita and sclerosing keratoderma (KLICK) syndrome, which is characterized by palmoplantar keratoderma (33) related to proteasome activity. This mutation results in changes in hUmp1 levels and alterations in the epidermal distribution of hUmp1 and proteasomal subunits. It is unclear how the proteasome functions in KLICK syndrome, although it is clear that disturbances in proteasome function cause clinical phenotypes in humans.

Studies in animal models indicate that cells deficient in various immunoproteasome subunits show poor CD8 responses when challenged with epitopes (34, 35) and may display alterations in the T-cell receptor (TCR) repertoire (36). In particular, β5i-deficient mice show increased susceptibility to pathogens, most likely due to the reduced efficiency of antigen presentation by β5i-deficient cells (12). Actually, in NNS patients, unresponsiveness to an intradermally applied purified protein derivative of *Mycobacterium tuberculosis* has been reported; however, there are no documented changes in susceptibility to pathogens, and no abnormalities in the number of any particular T-cell subset have been observed, apart from reduced NK activity (4). Conversely, there are no reports that β5i-deficient mice show the type of systemic inflammation observed in NNS patients.

In general, gene-deficient mice are very useful tools for analyzing the functions of target genes; however, the $\beta5i^{G201V}$ mutation shows a type of "enzymatic dominant-negative interference," which abrogates not only chymotrypsin-like activity (due to the mutation) but also the activity of the entire proteasome (due to defective assembly). Thus, it is not surprising that the phenotype seen in NNS is different from that seen in *Psmb8* knockout mice (12, 20) or after treatment with PR-957 inhibitors (37). Thus, analysis of patients with NNS and JMP syndrome and mice knocked in with these mutations would provide new insights into the function of the immunoproteasome in vivo.

Finally, we observed increased levels of p-p38 in nuclear extracts from NNS peripheral blood lymphocytes (Fig. 4*D*), although it remains unknown precisely how attenuation of proteasome activity causes accumulation of p-p38 in the nucleus. The

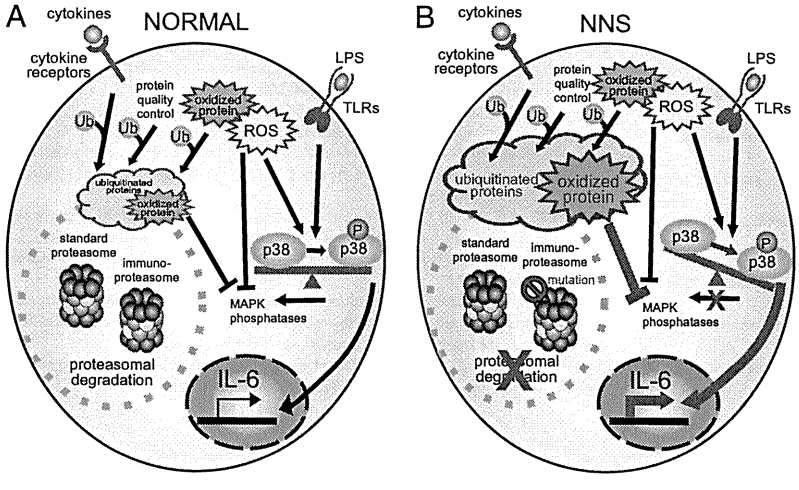

Fig. 5. Schematic model showing induction of inflammation in NNS patients with the *PSMB8* mutation. Our data are based on the scheme proposed by Bulua et al. (40). (*A*) In a normal cell, ubiquitinated or oxidized proteins generated by various stressors, including cytokines, are cleared by proteasomes. (*B*) The ubiquitinated and oxidized proteins accumulate in a cell with the *PSMB8* mutation (NNS cell). ROS and/or oxidized proteins may cause phosphorylation of p-38 to predominate over the nonphosphorylated form by inhibiting MAPK phosphatase or by activating MAPK.