TABLE. The Results of Clinical Assessments and the Correlation With Area of Abnormal FAF

Characteristics	Mean ± SD	ρ	*P* Value
Age, y	51.4 ± 17.4	−0.09	0.51
LogMAR VA	1.00 ± 0.57	0.23	0.08
GP I/4e scotoma size, cm^{2*}	18.8 ± 16.9	0.79	<0.001
Full-field ERG, μV			
Dark-adapted 0.01	44.0 ± 35.0	−0.63	<0.001
Dark-adapted 3.0 A wave	88.1 ± 59.0	−0.72	<0.001
Dark-adapted 3.0 B wave	127.2 ± 83.2	−0.66	<0.001
Light-adapted 3.0 B wave	43.1 ± 34.7	−0.44	<0.001
Light-adapted 3.0 flicker, 30 Hz	37.5 ± 30.0	−0.47	<0.001

ρ, correlation coefficient with the area of abnormal FAF; logMAR VA, logarithm of minimum angle of resolution visual acuity.

* Eight patients who were unable to detect the I/4e white test light were excluded.

mean value was used for analysis. The results obtained as pixel values were converted into a percentage of the elliptical area analyzed (abnormal FAF area/analyzed area).

Statistical Analyses

Statistical analyses were performed using statistical software (SPSS version 21.0; SPSS Science, Chicago, IL, USA). The results of descriptive analyses are reported as the mean ± standard deviation. Associations between clinical characteristics and the area of abnormal FAF were assessed with Spearman's rank correlation test. A *P* value of <0.05 was considered statistically significant.

RESULTS

Clinical Characteristics

We enrolled a total of 63 patients. Wide-field FAF was successfully performed in all cases. None of the patients complained of a deterioration in visual function or discomfort after the examination. After excluding two patients with poor-quality images and four patients who were found to share consanguinity with other participants, we evaluated 57 eyes of 57 patients (32 men and 25 women). The mean age was 51.4 ± 17.4 years (range, 12–82 years), and the mean logMAR score was 1.00 ± 0.57 units (range, 0–2 units). The study included 16 CD patients and 41 CRD patients. The inheritance pattern was autosomal dominant in 12 patients, autosomal recessive in 11 patients, X-linked in one patient, and sporadic in 33 patients. Thirty-two participants had previously submitted to causative mutation screening. The results showed *ABCA4* mutations in six patients (four CD patients and two CRD patients), *GUCY2D* mutations in two CD patients, and a *CRX* mutation in one CRD patient. There were no significant correlations between the area of abnormal FAF and age or logMAR score.

Correlation Between the Results of Visual Field and FAF Examinations

The scotoma size measurements obtained by Goldmann perimetry are presented in the Table. Eight patients could not recognize the I/4e white test light anywhere in the visual field. All of these were the patients with abnormal FAF throughout the fundus, who were excluded from the corresponding analysis. Even after excluding these patients, scotoma size correlated well with the area of abnormal FAF (Figs. 2, 3).

Correlation Between the Results of Electroretinography and FAF Examinations

The results of the full-field ERG examinations are presented as mean ± standard deviation in the Table. The area of abnormal FAF correlated well with ERG results under all conditions. The larger the area of abnormal FAF, the smaller the amplitude of ERG recordings (Fig. 3). The correlation was relatively strong

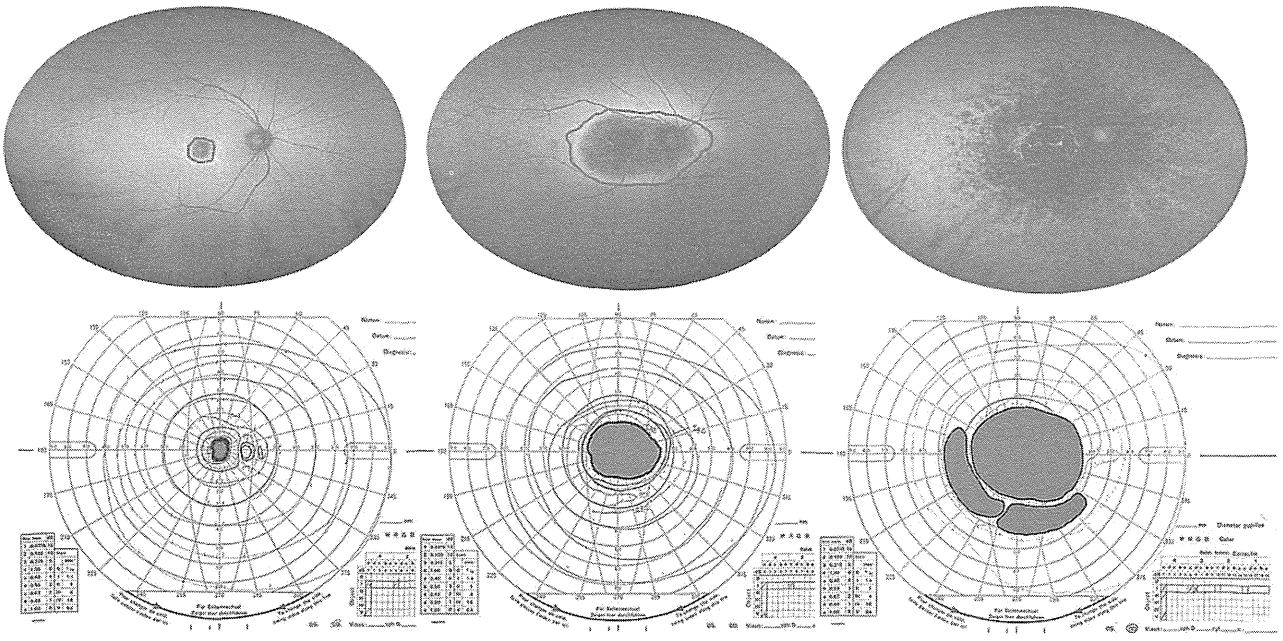

FIGURE 2. Representative images of wide-field fundus autofluorescence and Goldmann perimetry of eyes with cone or cone-rod dystrophy. Note that cases with larger areas of abnormal FAF showed larger scotoma areas defined by the I/4e white test light (area segmented in *gray, lower row*).

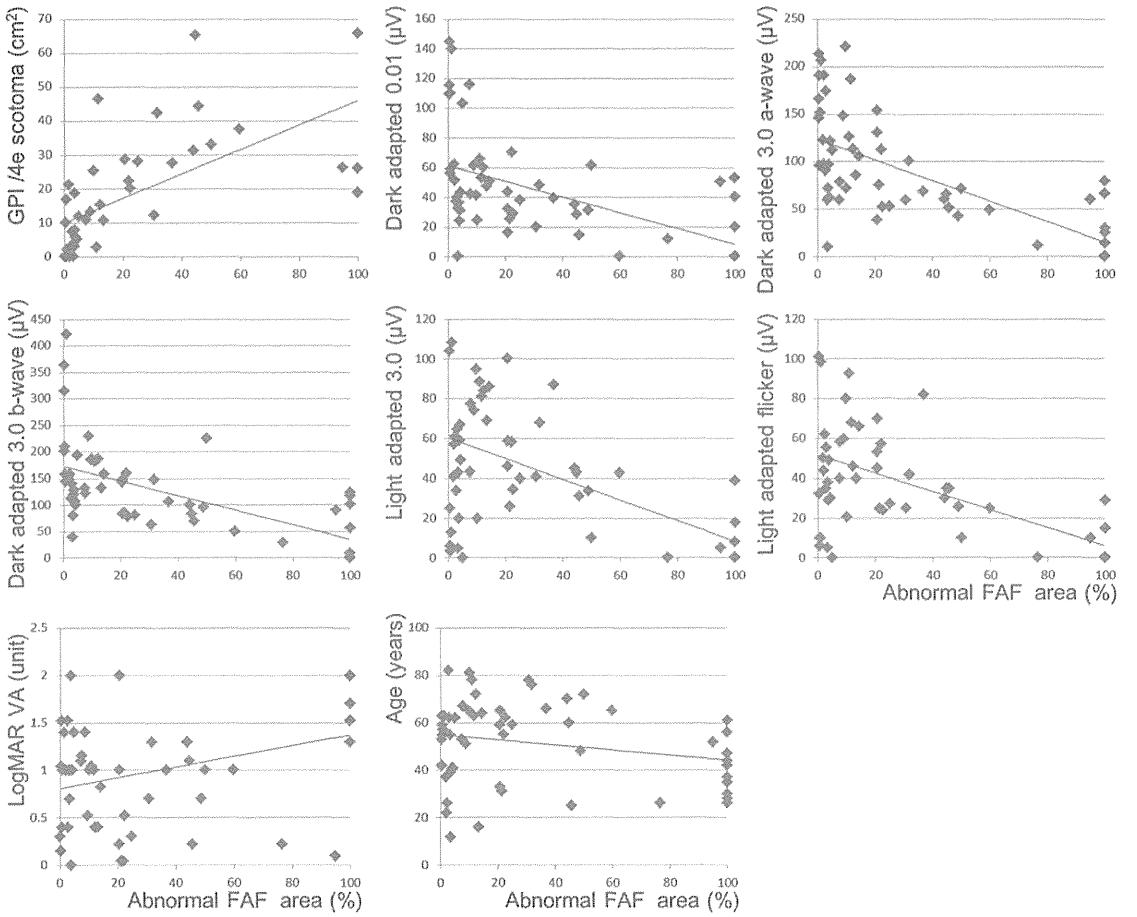

FIGURE 3. Scatter plots showing the relationship between the area of abnormal FAF and clinical characteristics. The area of abnormal FAF correlated well with visual function results on Goldman perimetry or ERG. The area of abnormal FAF showed no correlation with visual acuity or age.

in rod as well as combined responses and moderate for cone and flicker responses.

DISCUSSION

In the present study, we used wide-field FAF to evaluate patients with CD and CRD. The results showed that the extent of abnormal FAF correlates with the visual field and the results of ERG. This result demonstrates that wide-field FAF is clinically useful for predicting visual function in patients with CD and CRD.

The area of abnormal FAF was associated with scotoma size as measured with GP. In addition, the location and size of the scotoma seemed to correspond to the area of abnormal FAF as shown in Figure 2. The association between abnormal FAF and visual field defects was consistent with our previous report on retinitis pigmentosa.[13] The present findings confirm the relationship between abnormal FAF and visual field defects in patients with CD or CRD as well as rod-dominant retinal dystrophy.

The association between the area of abnormal FAF and retinal function was also confirmed by the results of full-field ERG. The amplitude of the rod, cone, or combined responses decreased as the area of abnormal FAF increased, which would be expected considering the close correlation between visual field defects and changes in ERG amplitude.[27] We consider that the use of wide-field FAF rather than conventional macular FAF is a reason for the strong correlation. Wide-field FAF can evaluate the peripheral retina; thus, the measurement correlated well with the results of GP or full-field ERG, which reflects function throughout the retina.

The association between cone function and the area of abnormal FAF was weaker than that of rod function and abnormal FAF area. This evidence suggests that FAF mainly reflects the function of rod photoreceptors. For example, the distribution of FAF roughly matches the distribution of rod photoreceptors.[28] In addition, the number of foveal cone-derived phagosomes in the RPE was one-third that of extrafoveal rod-derived phagosomes.[29] Larger areas of abnormal FAF represent more advanced stages or more severe phenotypes of the disease as manifest in rod function. Therefore, the association between larger areas of abnormal FAF and more pronounced cone dysfunction might reflect disease severity rather than cone cell loss itself.

There was no significant association between visual acuity and the area of abnormal FAF, as was expected from the nature of the examination. While the wide-field imaging device (Optos PLC) obtains a wide-field view of the retina, visual acuity only reflects foveal function. More specific examination tools such as static perimetry, microperimetry, contrast sensitivity measurement, and focal macular ERG would be more suitable for evaluating foveal function. Appropriate examinations should be employed to evaluate the area or the function of interest.

Previous studies focused on a ring of hyper-FAF around the degenerated retina. The finding was reported for patients with retinitis pigmentosa,[14,15] autoimmune retinopathy,[30] and age-related macular degeneration, respectively.[31,32] The increase in FAF adjacent to the atrophic area is considered to represent the presence of melanolipofuscin or changes in the metabolic activity of RPE cells.[7] This change sometimes precedes a visible change in appearance or retinal function[7] and is attracting attention. For example, in patients with retinitis pigmentosa, the size of the ring is associated with visual function,[9,10] and the radius of the ring constricts as the disease progresses.[14] As shown in the figures, the hyper-FAF ring was generally confirmed in cases whose decreased FAF area was confined to the area surrounding arcade vessels. Patients with decreased FAF that extends to the periphery will rarely, if ever, exhibit such a ring. One reason for the absence of the ring in advanced cases would be the distribution of lipofuscin: highest at approximately 10° from the fovea then decreasing toward the periphery.[33] The decreased background FAF may make it difficult to identify hyperautofluorescence in the periphery. Although we compared the clinical characteristics of patients with and without the ring, there was no significant difference. To evaluate the significance of the ring in CD and CRD, a longitudinal study is required.

Notably, this study included patients with CD and CRD. Although these diseases are differentiated clinically, they share major characteristics and there can be overlap between them.[34] For example, patients with CD can manifest rod dysfunction in the advanced stage of disease.[3,4] There is also overlap among the genes believed to cause these diseases.[2,3,35] Therefore, physicians must assess visual function in each patient without presumptions based upon the initial clinical diagnosis. The present results showed that the FAF pattern can roughly indicate the associated degree of retinal function regardless of a patient's clinical diagnosis. Accordingly, we might be able to evaluate patients with cone-dominant dystrophy to elaborate a spectrum of disease severity. Considering the difficulty in differentiating various manifestations of cone-dominant dystrophy, especially in advanced stages of disease, such an examination would facilitate patient treatment.

The present study has several limitations, including its cross-sectional study design and the relatively small number of patients, which was determined by the disease's prevalence. In addition, we had to exclude eight patients who did not respond to the I/4e isopter from the analysis. If these patients had been included, the difference between type 3 and type 1 or 2 would have been larger. Submitting each patient to mutation identification would have furthered our understanding.

Finally, we demonstrated the close correlation of wide-field FAF findings and visual function in CD and CRD. This type of noninvasive examination can be a practical indicator of the patient's visual field and retinal responses to light. Longitudinal studies will be necessary to further characterize the related decline in visual function. The findings would serve as a clinical guide when diagnosing, evaluating, or following patients with CD or CRD.

Acknowledgments

Supported in part by the Japan Ministry of Health, Labor and Welfare (No. 12103069).

Disclosure: **M. Oishi**, None; **A. Oishi**, None; **K. Ogino**, None; **Y. Makiyama**, None; **N. Gotoh**, None; **M. Kurimoto**, None; **N. Yoshimura**, Canon (F), Topcon (F), Nidek (F, C)

References

1. Hartong DT, Berson EL, Dryja TP. Retinitis pigmentosa. *Lancet*. 2006;368:1795-1809.
2. Berger W, Kloeckener-Gruissem B, Neidhardt J. The molecular basis of human retinal and vitreoretinal diseases. *Prog Retin Eye Res*. 2010;29:335-375.
3. Traboulsi EI. Cone dysfunction syndromes, cone dystrophies, and cone-rod degenerations. In: Traboulsi EI, ed. *Genetic Diseases of the Eye*. New York: Oxford University Press; 2012: 410-420.
4. Thiadens AA, Phan TM, Zekveld-Vroon RC, et al. Clinical course, genetic etiology, and visual outcome in cone and cone-rod dystrophy. *Ophthalmology*. 2012;119:819-826.
5. Delori FC, Dorey CK, Staurenghi G, Arend O, Goger DG, Weiter JJ. In vivo fluorescence of the ocular fundus exhibits retinal pigment epithelium lipofuscin characteristics. *Invest Ophthalmol Vis Sci*. 1995;36:718-729.
6. Holz FG, Fleckenstein M, Schmitz-Valckenberg S, Bird AC. Evaluation of fundus autofluorescence images. In: Holz FG, Schmitz-Valckenberg S, Spaide RF, Bird AC, eds. *Atlas of Fundus Autofluorescence Imaging*. Berlin: Springer-Verlag; 2007:71-76.
7. Schmitz-Valckenberg S, Holz FG, Bird AC, Spaide RF. Fundus autofluorescence imaging: review and perspectives. *Retina*. 2008;28:385-409.
8. Freund KB, Mrejen S, Jung J, Yannuzzi LA, Boon CJ. Increased fundus autofluorescence related to outer retinal disruption. *JAMA Ophthalmol*. 2013;131:1645-1649.
9. Robson AG, Egan CA, Luong VA, Bird AC, Holder GE, Fitzke FW. Comparison of fundus autofluorescence with photopic and scotopic fine-matrix mapping in patients with retinitis pigmentosa and normal visual acuity. *Invest Ophthalmol Vis Sci*. 2004;45:4119-4125.
10. Robson AG, Saihan Z, Jenkins SA, et al. Functional characterisation and serial imaging of abnormal fundus autofluorescence in patients with retinitis pigmentosa and normal visual acuity. *Br J Ophthalmol*. 2006;90:472-479.
11. Murakami T, Akimoto M, Ooto S, et al. Association between abnormal autofluorescence and photoreceptor disorganization in retinitis pigmentosa. *Am J Ophthalmol*. 2008;145:687-694.
12. Fleckenstein M, Charbel Issa P, Fuchs HA, et al. Discrete arcs of increased fundus autofluorescence in retinal dystrophies and functional correlate on microperimetry. *Eye (Lond)*. 2009;23: 567-575.
13. Oishi A, Ogino K, Makiyama Y, Nakagawa S, Kurimoto M, Yoshimura N. Wide-field fundus autofluorescence imaging of retinitis pigmentosa. *Ophthalmology*. 2013;120:1827-1834.
14. Lima LH, Burke T, Greenstein VC, et al. Progressive constriction of the hyperautofluorescent ring in retinitis pigmentosa. *Am J Ophthalmol*. 2012;153:718-727. e711-712.
15. Lima LH, Cella W, Greenstein VC, et al. Structural assessment of hyperautofluorescent ring in patients with retinitis pigmentosa. *Retina*. 2009;29:1025-1031.
16. Lois N, Halfyard AS, Bird AC, Holder GE, Fitzke FW. Fundus autofluorescence in Stargardt macular dystrophy-fundus flavimaculatus. *Am J Ophthalmol*. 2004;138:55-63.
17. Smith RT, Gomes NL, Barile G, Busuioc M, Lee N, Laine A. Lipofuscin and autofluorescence metrics in progressive STGD. *Invest Ophthalmol Vis Sci*. 2009;50:3907-3914.
18. Cukras CA, Wong WT, Caruso R, Cunningham D, Zein W, Sieving PA. Centrifugal expansion of fundus autofluorescence patterns in Stargardt disease over time. *Arch Ophthalmol*. 2012;130:171-179.
19. Kellner S, Kellner U, Weber BH, Fiebig B, Weinitz S, Ruether K. Lipofuscin- and melanin-related fundus autofluorescence in patients with ABCA4-associated retinal dystrophies. *Am J Ophthalmol*. 2009;147:895-902. 902 e891.
20. von Ruckmann A, Fitzke FW, Bird AC. In vivo fundus autofluorescence in macular dystrophies. *Arch Ophthalmol*. 1997;115:609-615.

21. Robson AG, Michaelides M, Luong VA, et al. Functional correlates of fundus autofluorescence abnormalities in patients with RPGR or RIMS1 mutations causing cone or cone rod dystrophy. *Br J Ophthalmol.* 2008;92:95–102.
22. Manivannan A, Plskova J, Farrow A, McKay S, Sharp PF, Forrester JV. Ultra-wide-field fluorescein angiography of the ocular fundus. *Am J Ophthalmol.* 2005;140:525–527.
23. Seidensticker F, Neubauer AS, Wasfy T, et al. Wide-field fundus autofluorescence corresponds to visual fields in chorioretinitis patients. *Clin Ophthalmol.* 2011;5:1667–1671.
24. Witmer MT, Cho M, Favarone G, Chan RV, D'Amico DJ, Kiss S. Ultra-wide-field autofluorescence imaging in non-traumatic rhegmatogenous retinal detachment. *Eye (Lond).* 2012;26:1209–1216.
25. Witmer MT, Kozbial A, Daniel S, Kiss S. Peripheral autofluorescence findings in age-related macular degeneration. *Acta Ophthalmol.* 2012;90:e428–433.
26. Marmor MF, Fulton AB, Holder GE, Miyake Y, Brigell M, Bach M. ISCEV Standard for full-field clinical electroretinography (2008 update). *Doc Ophthalmol.* 2009;118:69–77.
27. Zahid S, Jayasundera T, Rhoades W, et al. Clinical phenotypes and prognostic full-field electroretinographic findings in Stargardt disease. *Am J Ophthalmol.* 2013;155:465–473. e463.
28. Curcio CA, Sloan KR, Kalina RE, Hendrickson AE. Human photoreceptor topography. *J Comp Neurol.* 1990;292:497–523.
29. Anderson DH, Fisher SK, Erickson PA, Tabor GA. Rod and cone disc shedding in the rhesus monkey retina: a quantitative study. *Exp Eye Res.* 1980;30:559–574.
30. Lima LH, Greenberg JP, Greenstein VC, et al. Hyperautofluorescent ring in autoimmune retinopathy. *Retina.* 2012;32:1385–1394.
31. Bindewald A, Schmitz-Valckenberg S, Jorzik JJ, et al. Classification of abnormal fundus autofluorescence patterns in the junctional zone of geographic atrophy in patients with age related macular degeneration. *Br J Ophthalmol.* 2005;89:874–878.
32. Holz FG, Bindewald-Wittich A, Fleckenstein M, Dreyhaupt J, Scholl HP, Schmitz-Valckenberg S. Progression of geographic atrophy and impact of fundus autofluorescence patterns in age-related macular degeneration. *Am J Ophthalmol.* 2007;143:463–472.
33. Delori FC, Goger DG, Dorey CK. Age-related accumulation and spatial distribution of lipofuscin in RPE of normal subjects. *Invest Ophthalmol Vis Sci.* 2001;42:1855–1866.
34. van den Biesen PR, Deutman AF, Pinckers AJ. Evolution of benign concentric annular macular dystrophy. *Am J Ophthalmol.* 1985;100:73–78.
35. Garcia-Hoyos M, Auz-Alexandre CL, Almoguera B, et al. Mutation analysis at codon 838 of the Guanylate Cyclase 2D gene in Spanish families with autosomal dominant cone, cone-rod, and macular dystrophies. *Mol Vis.* 2011;17:1103–1109.

GENETIC TESTING AND MOLECULAR BIOMARKERS
Volume 18, Number 11, 2014

Pp. 722–735
DOI: 10.1089/gtmb.2014.0109

The Use of Next-Generation Sequencing in Molecular Diagnosis of Neurofibromatosis Type 1: A Validation Study

Ryo Maruoka,[1,2] Toshiki Takenouchi,[1,3] Chiharu Torii,[1] Atsushi Shimizu,[4] Kumiko Misu,[1] Koichiro Higasa,[5] Fumihiko Matsuda,[5] Arihito Ota,[6] Katsumi Tanito,[6] Akira Kuramochi,[7] Yoshimi Arima,[8] Fujio Otsuka,[9] Yuichi Yoshida,[10] Keiji Moriyama,[2] Michihito Niimura,[6] Hideyuki Saya,[8] and Kenjiro Kosaki[1]

Aims: We assessed the validity of a next-generation sequencing protocol using in-solution hybridization-based enrichment to identify *NF1* mutations for the diagnosis of 86 patients with a prototypic genetic syndrome, neurofibromatosis type 1. In addition, other causative genes for classic genetic syndromes were set as the target genes for coverage analysis. *Results:* The protocol identified 30 nonsense, 19 frameshift, and 8 splice-site mutations, together with 10 nucleotide substitutions that were previously reported to be pathogenic. In the remaining 19 samples, 10 had single-exon or multiple-exon deletions detected by a multiplex ligation-dependent probe amplification method and 3 had missense mutations that were not observed in the normal Japanese SNP database and were predicted to be pathogenic. Coverage analysis of the genes other than the *NF1* gene included on the same diagnostic panel indicated that the mean coverage was 115-fold, a sufficient depth for mutation detection. *Conclusions:* The overall mutation detection rate using the currently reported method in 86 patients who met the clinical diagnostic criteria was 92.1% (70/76) when 10 patients with large deletions were excluded. The results validate the clinical utility of this next-generation sequencing-based method for the diagnosis of neurofibromatosis type 1. Comparable detection rates can be expected for other genetic syndromes, based on the results of the coverage analysis.

Introduction

GENETIC TESTING HAS HELPED clinicians to define the molecular pathology of diseases, especially when patients present with an atypical combination of phenotypic features. Our group developed a custom-designed mutation analysis panel using denaturing high-pressure liquid chromatography for the systematic screening of patients with classic genetic syndromes (Kosaki *et al.*, 2005). The system can be used to screen all the exons of the candidate gene quickly and has been helpful in confirming the clinical diagnosis, as published in a series of reports in this journal (Udaka *et al.*, 2005, 2006, 2007; Aramaki *et al.*, 2006; Samejima *et al.*, 2007; Hattori *et al.*, 2009). Nevertheless, the throughput of the system was not high enough to screen multiple candidate genes in a single testing.

The recent advent of a target sequencing panel with the next-generation sequencing technology has enabled many genes, regardless of size, to be analyzed in a systematic and comprehensive manner, as reviewed in this journal (Yan *et al.*, 2013). The strength of such a comprehensive approach is the ability to detect atypical presentations of classic syndromes, as illustrated by our recent reports on several patients with atypical presentations of mutations in the causative

[1]Center for Medical Genetics, Keio University School of Medicine, Tokyo, Japan.
[2]Section of Maxillofacial Orthognathics, Division of Maxillofacial/Neck Reconstruction, Department of Maxillofacial Reconstruction and Function, Graduate School, Tokyo Medical and Dental University, Tokyo, Japan.
[3]Department of Pediatrics, Keio University School of Medicine, Tokyo, Japan.
[4]Division of Biomedical Information Analysis, Iwate Tohoku Medical Megabank Organization, Iwate Medical University, Iwate, Japan.
[5]Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan.
[6]Department of Dermatology, Jikei University School of Medicine, Tokyo, Japan.
[7]Department of Dermatology, Saitama Medical University, Saitama, Japan.
[8]Division of Gene Regulation, Institute for Advanced Medical Research, Keio University School of Medicine, Tokyo, Japan.
[9]Department of Dermatology, Institute of Clinical Medicine, University of Tsukuba, Tsukuba, Japan.
[10]Division of Dermatology, Department of Medicine of Sensory and Motor Organs, Faculty of Medicine, Tottori University, Yonago, Japan.

genes of three classic genetic syndromes: the neonatal progeroid presentation of an *FBN1* mutation (Takenouchi *et al.*, 2013a), the Noonan-cafe au lait syndrome-like presentation of a *MAP2K2* mutation (Takenouchi *et al.*, 2013b), and Stickler syndrome-like presentation of *SOX9* mutation (Takenouchi *et al.*, 2014).

In this study, we assessed the analytical and clinical validity of the next-generation sequencing protocol with in-solution hybridization-based enrichment to identify disease-causing mutations in the diagnosis of a prototypic genetic syndrome, neurofibromatosis type 1, compared with direct capillary sequencing, which is the current gold standard methodology. The reason for the choice of the *NF1* gene, the causative gene for neurofibromatosis type 1, was twofold: (1) neurofibromatosis type 1 is a relatively common genetic condition with readily recognizable phenotypes: café-au-lait spots, cutaneous neurofibromas, axillary and inguinal freckling, and Lisch nodules (iris hamartomas) (Carey and Viskochil, 1999) and (2) the *NF1* gene comprised a total of 58 exons and is one of the largest genes in the human genome, making it a relatively difficult clinical target for direct capillary sequencing.

Materials and Methods

Patients

The current research protocol was approved by the institutional review board of Keio University and each participating center. Eighty-six patients with neurofibromatosis type 1 who met the NIH clinical diagnostic criteria (Neurofibromatosis Conference Statement, 1988) were recruited from multiple centers participating in the project. The NIH diagnostic criteria for neurofibromatosis type 1 defines an individual as neurofibromatosis type 1 when the person has two or more of the following features: six or more café-au-lait macules with a maximum diameter of over 5 mm in prepubertal individuals and with a maximum diameter of over 15 mm in postpubertal individuals; two or more neurofibromas of any type or 1 plexiform neurofibroma; freckling in the axillary or inguinal regions; optic glioma, two or more Lisch nodules; a distinctive osseous lesion, such as sphenoid dysplasia or tibial pseudarthrosis; and a first-degree relative (parent, sibling, or offspring) with neurofibromatosis type 1, as defined according to the above-mentioned criteria. After written consent was obtained at each participating center, the whole blood samples were sent to Keio University for genetic analysis.

Genomic DNA, sample preparation, targeted capturing, sequencing

Genomic DNA was extracted from peripheral blood according to standard procedures using the phenol–chloroform extraction method and checked for quality using Qubit (Life Technologies). The genomic DNA (3 μg) was fragmented into ~150 bp. In-solution hybridization-based enrichment was performed using the SureSelect Target Enrichment system (Agilent Technologies). The *NF1* gene (the canonical Refseq transcript NM_001042492.2) together with 108 causative genes for the more common classical congenital malformation syndromes selected from a standard textbook (Jones, 2005) was set as the target gene (Table 1). Genes that are responsible for a disease phenotype and involved in the RAS pathway (i.e., Rasopathy genes) (Aoki *et al.*, 2008) were included in the 108 genes set. A biotinylated RNA capture library was designed using the eArray system (Agilent Technologies) according to the manufacturer's protocol. The captured DNA was subjected to a 150-bp paired-end read sequencing on the MiSeq system (Illumina).

Bioinformatics pipeline

The sequence reads from the sequencer were exported as FASTQ format files and were analyzed using sets of open-source programs by means of the default parameters; the sequence reads were aligned to the human reference genome DNA sequence (hs37d5 assembly) using the Burrows–Wheeler Alignment (BWA) tool version 0.6.1 (Li and Durbin, 2009). The Genome Analysis Toolkit (GATK) package (McKenna *et al.*, 2010) was used to perform local realignment, base quality score recalibration, and SNP/indel calls. The called SNPs/indels were annotated using snpEff version 3.1 (Cingolani *et al.*, 2012), regarded as nonpathogenic, and excluded from further analysis when they were observed in the 1000 Genomes Project (www.1000genomes.org/) or in the Japanese SNP dataset of 1208 normal individuals (Japanese Genetic Variation Consortium, 2013). The variants and alignments were visually inspected using the Integrative Genomics Viewer version 2.1 (Thorvaldsdóttir *et al.*, 2013) and VarSifter version 1.5 (Teer *et al.*, 2012). Variants in the RAS pathway, including *PTPN11*, *KRAS*, *SOS1*, *RAF1*, *SHOC2*, *HRAS*, *BRAF*, *MAPK1*, *MAP2K1*, *MAP2K2*, *MAPK3*, *SPRED1*, and *RASA1*, were evaluated for pathogenicity. Other genes were not subject to further variant analysis to avoid potential issues with incidental findings. A statistical coverage analysis was performed as described below.

Coverage analysis

Information about enrichment performance and target coverage was obtained using the software NGSrich version 0.7.8 (Frommolt *et al.*, 2012). The following parameters were measured: information about the number of reads, mean coverage, fraction of the target region with a particular depth across the 109 genes, information on the number of genes that are poorly covered, and a summary table with exon-specific coverage information at the *NF1* locus.

Direct capillary sequencing for validation

When the next-generation sequencing protocol identified truncating mutations, including nonsense mutations, frameshift mutations, and mutations at the canonical splice sites, or missense mutations that had been previously reported as being pathogenic in the literature, the variants were validated with direct capillary sequencing. In the remaining samples, all the exons were analyzed using direct capillary sequencing (Richards *et al.*, 2008). For direct capillary sequencing, 56 pairs of polymerase chain reaction (PCR) primers were designed on flanking intronic and untranslated regions to encompass the coding regions of the 58 *NF1* exons and at least 30 bp of the intronic sequence surrounding each exon (Table 2). Three primers were designed newly using primer design software, Primer3 (Rozen and Skaletsky, 2000), and the remaining primers were described elsewhere (Purandare *et al.*,

TABLE 1. LIST OF THE 109 GENES

Gene	*Chromosome*	*Basepair position (GRCh37)*	*Disease*	*Gene*	*Chromosome*	*Basepair position (GRCh37)*	*Disease*
ACTA2	10	90,694,830–90,751,146	Multisystemic smooth muscle dysfunction syndrome	*MSX1*	4	4,861,391–4,865,662	Witkop syndrome
ACTC1	15	35,080,296–35,087,926	Atrial septal defect	*MYH7*	14	23,881,946–23,904,869	Scapuloperoneal syndrome, myopathic type
ACVRL1	12	52,300,656–52,317,144	Hereditary hemorrhagic telangiectasia	*MYH9*	22	36,677,322–36,784,106	Fechtner syndrome
BRAF	7	140,415,748–140,624,563	Cardiofaciocutaneous syndrome	*NF1*	17	29,421,944–29,704,694	Neurofibromatosis type 1
CBL	11	119,076,985–119,178,858	Noonan syndrome-like disorder	*NIPBL*	5	36,876,860–37,065,925	Cornelia de Lange syndrome
CDKL5	X	18,443,724–18,671,748	Angelman syndrome-like disorder	*NOTCH2*	1	120,454,175–120,639,879	Alagille syndrome
CHD7	8	61,591,320–61,780,586	CHARGE syndrome	*NRAS*	1	115,247,084–115,259,514	Noonan syndrome
COL11A1	1	103,342,022–103,574,051	Fibrochondrogenesis	*NRTN*	19	5,823,817–5,828,334	Hirschsprung disease
COL11A2	6	33,130,468–33,160,244	Stickler syndrome	*NSD1*	5	176,560,025–176,727,213	Sotos syndrome
COL1A1	17	48,261,456–48,279,002	Osteogenesis imperfecta	*OTX2*	14	57,267,424–57,277,193	Syndromic microphthalmia
COL1A2	7	94,023,872–94,060,543	Ehlers-Danlos syndrome	*PHOX2B*	4	41,746,098–41,750,986	Congenital central hypoventilation syndrome
COL2A1	12	48,366,747–48,398,284	Stickler syndrome	*PKHD1*	6	51,480,144–51,952,422	Polycystic kidney and hepatic disease
COL3A1	2	189,839,098–189,877,471	Ehlers-Danlos syndrome	*PLOD1*	1	11,994,723–12,035,598	Ehlers-Danlos syndrome
COL5A1	9	137,533,650–137,736,688	Ehlers-Danlos syndrome	*PSPN*	19	6,375,304–6,375,859	Hirschsprung's disease
COL5A2	2	189,896,640–190,044,667	Ehlers-Danlos syndrome	*PTCH1*	9	98,205,263–98,279,246	Basal cell nevus syndrome
COL9A1	6	70,925,742–71,012,785	Stickler syndrome	*PTPN11*	12	112,856,535–112,947,716	LEOPARD syndrome
COL9A2	1	40,766,161–40,782,938	Stickler syndrome	*RAD21*	8	117,858,172–117,887,104	Cornelia de Lange syndrome
COMP	19	18,893,582–18,902,113	Epiphyseal dysplasia	*RAF1*	3	12,625,099–12,705,699	LEOPARD syndrome
CREBBP	16	3,775,054–3,930,120	Rubinstein-Taybi syndrome	*RASA1*	5	86,564,069–86,687,742	Parkes Weber syndrome
CUL7	6	43,005,354–43,021,682	3-M syndrome	*RET*	10	43,572,516–43,625,798	MENII
DCC	18	49,866,541–51,062,272	Mirror movements	*RUNX2*	6	45,296,053–45,518,818	Cleidocranial dysplasia
DDX3X	X	41,192,560–41,209,526	Medulloblastoma	*SALL1*	16	51,169,885–51,185,182	Townes–Brocks syndrome
ECE1	1	21,543,739–21,672,033	Hirschsprung disease	*SALL4*	20	50,400,550–50,419,058	Duane-radial ray syndrome
EDN3	20	57,875,498–57,901,046	Central hypoventilation syndrome	*SCN1B*	19	35,521,554–35,531,352	Brugada syndrome
EDNRB	13	78,469,615–78,549,663	Waardenburg syndrome	*SHH*	7	155,595,557–155,604,966	Holoprosencephaly
EFNB1	X	68,048,839–68,062,006	Craniofrontonasal dysplasia	*SHOC2*	10	112,679,300–112,773,424	Noonan-like syndrome
ENG	9	130,577,290–130,617,051	Heredity hemorrhagic telangiectasia	*SIX3*	2	45,169,036–45,173,215	Holoprosencephaly

(*continued*)

TABLE 1. (CONTINUED)

Gene	*Chromosome*	*Basepair position (GRCh37)*	*Disease*	*Gene*	*Chromosome*	*Basepair position (GRCh37)*	*Disease*
EP300	22	41,488,613–41,576,080	Rubinstein-Taybi syndrome	*SIX6*	14	60,975,937–60,978,524	Microphthalmia with cataract
FBN1	15	48,700,502–48,937,984	Acromicric dysplasia	*SMC1A*	X	53,401,069–53,449,676	Cornelia de Lange syndrome
FBN2	5	127,593,600–127,873,734	Congenital contractural arachnodactyly	*SMC3*	10	112,327,448–112,364,391	Cornelia de Lange syndrome
FGFR1	8	38,268,655–38,326,351	Hypogonadotropic hypogonadism	*SOS1*	2	39,208,689–39,347,685	Noonan syndrome
FGFR2	10	123,237,843–123,357,971	Antley-Bixler syndrome	*SOX10*	22	38,368,318–38,380,555	PCWH syndrome
FGFR3	4	1,795,038–1,810,598	Achondroplasia	*SOX2*	3	181,429,711–181,432,223	Syndromic microphthalmia
GDNF	5	37,812,778–37,839,781	Central hypoventilation syndrome	*SPRED1*	15	38,544,924–38,649,449	Legius syndrome
GFRA1	10	117,816,435–118,033,125	Hirschsprung's disease	*SPRY2*	13	80,910,110–80,915,085	Holoprosencephaly
GFRA2	8	21,549,529–21,672,391	Hirschsprung's disease	*STAG1*	3	136,055,077–136,471,220	Cornelia de Lange syndrome
GLA	X	100,652,778–100,663,000	Fabry disease	*TAZ*	X	153,639,876–153,650,064	Barth syndrome
HRAS	11	532,241–535,560	Costello syndrome	*TBX22*	X	79,270,254–79,287,267	Abruzzo-Erickson syndrome
IHH	2	219,919,141–219,925,237	Acrocapitofemoral dysplasia	*TBX5*	12	114,791,734–114,846,246	Holt-Oram syndrome
IRF6	1	209,958,967–209,979,519	Van der Woude syndrome	*TCF4*	18	52,889,561–53,303,251	Pitt-Hopkins syndrome
JAG1	20	10,618,331–10,654,693	Alagille syndrome	*TCOF1*	5	149,737,201–149,779,870	Treacher Collins syndrome
KCNE1	21	35,790,909–35,884,572	Jervell and Lange-Nielsen syndrome	*TGFBR1*	9	101,867,411–101,916,473	Loeys-Dietz syndrome
KCNJ2	17	68,164,756–68,176,188	Andersen syndrome	*TGFBR2*	3	30,647,993–30,735,633	Loeys-Dietz syndrome
KCNQ1	11	2,466,220–2,870,339	Jervell and Lange-Nielsen syndrome	*TGIF1*	18	3,411,924–3,458,408	Holoprosencephaly
KIAA1279	10	70,748,476–70,776,738	Goldberg-Shprintzen megacolon syndrome	*TP63*	3	189,348,941–189,615,067	EEC syndrome
KIF26A	14	104,605,059–104,647,234	Megacolon	*TRAPPC10*	21	45,432,205–45,526,432	Holoprosencephaly
KRAS	12	25,358,179–25,403,869	Noonan syndrome	*TRIM37*	17	57,059,998–57,184,265	Mulibrey nanism
L1CAM	X	153,126,968–153,151,627	CRASH syndrome	*TSC1*	9	135,766,734–135,820,093	Tuberous sclerosis
LAMP2	X	119,560,002–119,603,203	Danon disease	*TSC2*	16	2,097,471–2,138,712	Tuberous sclerosis
MAP2K1	15	66,679,181–66,783,881	Cardiofaciocutaneous syndrome	*TWIST1*	7	19,039,314–19,157,294	Saethre Chotzen syndrome
MAP2K2	19	4,090,318–4,124,125	Cardiofaciocutaneous syndrome	*VHL*	3	10,183,318–10,195,353	Von Hippel-Lindau syndrome
MAPK1	22	22,113,945–22,221,969	Acromesomelic dysplasia	*VSX2*	14	74,706,174–74,729,440	Microphthalmia
MAPK3	16	30,125,425–30,134,629	Cardiac hypertrophy	*ZEB2*	2	145,141,941–145,277,957	Mowat-Wilson syndrome
MECP2	X	153,287,024–153,363,187	Rett syndrome	*ZIC2*	13	100,634,025–100,639,018	Holoprosencephaly
MID1	X	10,413,349–10,851,828	Opitz GBBB syndrome				

Table 2. List of Polymerase Chain Reaction Primers

Exon	Primer sequence (5′-3′)	Amplicon size	Reference	Exon	Primer sequence (5′-3′)	Amplicon size	Reference
1	CAGACCCTCTCCTTGCCTCTT GGATGGAGGGTCGGAGGCTG	439	Purandare *et al.* (1995)	29	ATATGGAGCAGGTATAATAAAC AAAACAGCGGTTCTATGTG	181	Bausch *et al.* (2007)
2	CGTCATGATTTTCAATGGCAAG GCTCACTGAATCTAAAACCCAGC	438	Bausch *et al.* (2007)	30	CGTTGCACTTGGCTTAATGTCTG CCATCAGCAGCTAGATCCTTCTTT	327	Bausch *et al.* (2007)
3	TTTCACTTTTCAGATGTGTGTTG TGGTCCACATCTGTACTTTG	245	Purandare *et al.* (1995)	31	TTTTCTGTGATTCATAGCC GATATTCTTAACAAACAGCA	400	This report
4	TTAAATCTAGGTGGTGTGT AAACTCATTTCTCTGGAG	517	Han *et al.* (2001)	32	CTTATACTCAATTCTCAACTCC GAATTTAAGATAGCTAGATTATC	226	Bausch *et al.* (2007)
5	GAGATACCACACCTGTCCCCTAA TTGACCCAGTGATTTTTTTCAGA	215	Bausch *et al.* (2007)	33	GACTTCATACAATAAATAATCTG TATTTGATTCAAACAGAGCAAC	195	Bausch *et al.* (2007)
6	TTTCCTAGCAGACAACTATCGA AGGATGCTAACAACAGCAAAT	308	Han *et al.* (2001)	34	CTCCATATTTGTAATCTTAGTTA GGAGAGTGTTCACTATCCC	298	Bausch *et al.* (2007)
7	GAAGGAAGTTAGAAGTTTGTG CACAAGTAGGCATTTAAAAGA	211	Bausch *et al.* (2007)	35	GTTACAAGTTAAAGAAATGTGTAG CTAACAAGTGGCCTGGTGGCAAAC	298	Purandare *et al.* (1995)
8	CATGTTTATCTTTTAAAAATGTTGCC ATAATGGAAATAATTTTGCCCTCC	301	Han *et al.* (2001)	36	TTTATTGTTTATCCAATTATAGACTT TCCTGTTAAGTCAACTGGGAAAAAC	296	Purandare *et al.* (1995)
9	CTGTTAATTTGCTATAATATTAGC CATAATACTTATGCTAGAAAATTC	328	Bausch *et al.* (2007)	37	TGAATCCAGACTTTGAAGAATTGTT CTAGGGAGGCCAGGATATAGTCTAGT	644	Bausch *et al.* (2007)
10	GTAATGTGTTGATGTTATTACATG GTCTTTTTGTTTATAAAGGATAACA	273	Bausch *et al.* (2007)	38	GGTTGGTTTCTGGAGCCTTTTAGA CAACAAACCCCAAATCAAACTGA	467	Bausch *et al.* (2007)
11	CTTTCTATTTGCTGTTCTTTTTGG CCTTTTTGAAAACCAAGAGTGCA	264	Bausch *et al.* (2007)	39	TTGGAACTATAAGGAAAAATACGTTT AGGGTTTTCTTTGAATTCTCTTAGA	321	Bausch *et al.* (2007)
12	ACGTAATTTTGTACTTTTTCTTCC CAATAGAAAGGAGGTGAGATTC	222	Purandare *et al.* (1995)	40	ATAATTGTTGATGTGATTTTCATTG AATTTTGAACCAGATGAAGAG	424	Han *et al.* (2001)
13	GCAAAAACGATTTTCATTGTTTTGT GCGTTTCAGCTAAACCCAATT	403	This report	41	TTGATTAGGCTGTTCCAATGAA CAAAACAAAAAACCTCCTGATGAT	298	Bausch *et al.* (2007)
14	ATTGAAGTTTCCTTTTTTTCCTTG GTATAGACATAAACATACCATTTC	275	Bausch *et al.* (2007)	42	GTGCTAAAACTTTGAGTCCCATGT ATAATCTATATTGATCAGGTGAAGTA	415	Bausch *et al.* (2007)
15	CCAAAAATGTTTGAGTGAGTCT ACCATAAAACCTTTGGAAGTG	256	Han *et al.* (2001)	43	GCAAGGAGCATTAATACAATGTATC CCATGCAAGTGTTTTTATTTAAGC	507	Bausch *et al.* (2007)

(*continued*)

TABLE 2. (CONTINUED)

Exon	*Primer sequence (5′-3′)*	*Amplicon size*	*Reference*	*Exon*	*Primer sequence (5′-3′)*	*Amplicon size*	*Reference*
16	AAACCTTACAAGAAAAACTAAGCT ATTACCATTCCAAATATTCTTCCA	303	Purandare *et al.* (1995)	44–45	GGTAACAGGTCACTTAATGACATCA GACCTCAAATTTAAACGTCTTTTAGA	512	Bausch *et al.* (2007)
17	CTCTTGGTTGTCAGTGCTTC CAGAAAACAAACAGAGCACAT	261	Han *et al.* (2001)	46	CATTCCGAGATTCAGTTTAGGAG AAGTAACATTCAACACTGATACCC	236	Abernathy *et al.* (1997)
18	CCCAAGTTGCAAATATATGTC GTGCTTTGAGGCAGACTGAG	336	Bausch *et al.* (2007)	47	TCCCCAAAAGAGAAAACATGG AGCAACAAGAAAAGATGGAAGAGT	334	Bausch *et al.* (2007)
19	TGAAGCATTTGCTCTGCTCT GTTTCAAACTTGATGTATATTAAA	347	Bausch *et al.* (2007)	48	CTACTGTGTGAACCTCATCAACC GTAAGACATAAGGGCTAACTTACTTC	284	Abernathy *et al.* (1997)
20	ACTTGGCTGTAGCTGATTGA ACTTTACTGAGCGACTCTTGAA	247	Han *et al.* (2001)	49	TCAGGGAAGAAGACCTCAGCAGATGC TGAACTTTCTGCTCTGCCACGCAACC	328	Abernathy *et al.* (1997)
21	GGAAGAAATGTTGGATAAAGCA AAACAAGTCACTCTATTCATAGA	579	Bausch *et al.* (2007)	50	GTGCACATTTAACAGGTACTAT CTTCCTAGGCCATCTCTAGAT	373	Han *et al.* (2001)
22	TATCTGTATGCTTATTTGGCTCTA GTGCAGTAAAGAATGGCCAG	385	Bausch *et al.* (2007)	51	CTTGGAAGGAGCAAACGATGGTTG CAAAAACTTTGCTACACTGACATGG	356	Abernathy *et al.* (1997)
23	AGAAGTTGTGTACGTTCTTTTCT CTCCTTTCTACCAATAACCGC	367	Purandare *et al.* (1995)	52	GCTCCAGGGATGTATTAGAGCTTT TGACTTTCATGTACTCTCCCACCT	325	Bausch *et al.* (2007)
24	TTGTTCCCTTCTGGCTTTTAT ATCTCAAAAGTTTAAATACACA	365	This report	53–54	TGAAGTGATTATCCAGGTGTTTGA AAAGACAGGCACGAAGGTGA	506	Bausch *et al.* (2007)
25	TGAGGGGAAGTGAAAGAACT GGCTTTATTTGCTTTTTGCT	235	Han *et al.* (2001)	55	AATTTTGGCACATTATTCTGGG AGCAAGTTCATCAACCATCCTT	290	Bausch *et al.* (2007)
26	CCACCCTGGCTGATTATCG TAATTTTTGCTTCTCTTACATGC	402	Purandare *et al.* (1995)	56	CTGTTACAATTAAAAGATACCTTGC TGTGTGTTCTTAAAGCAGGCATAC	185	Abernathy *et al.* (1997)
27	TGGTCTCATGCACTCCATA CATCTTTCTTCTGGCTCTGA	474	Han *et al.* (2001)	57	TTTTGGCTTCAGATGGGGATTTAC AAGGGAATTCCTAATGTTGGTGTC	351	Abernathy *et al.* (1997)
28	TGCTACTCTTTAGCTTCCTAC CCTTAAAAGAAGACAATCAGCC	331	Purandare *et al.* (1995)	58	AAGCGACACATGACTGCAATG TGGCTTTCATCACTGGCCA	571	Bausch *et al.* (2007)

1995; Abernathy *et al.*, 1997; Han *et al.*, 2001; Bausch *et al.*, 2007). The 3′ end of the primers were designed so as not to match the genomic sequences of any of the highly homologous pseudogene sequences to avoid mispriming to the pseudogenes. Direct capillary sequencing was performed using the ABI BigDye version 1.1 Terminator Cycle Kit (Life Technologies) and the ABI Prism 3500 Capillary Array Sequencer (Life Technologies). The sequence data were analyzed using Mutation Surveyor version 4.0.6 (Softgenetics) and Sequencher version 5.0 (Gene Codes Corp.).

Multiplex ligation-dependent probe amplification

When the next-generation sequencing protocol did not identify truncating mutations, canonical splice-site mutations, or other point mutations previously reported as pathological missense change or splicing defect, the remaining samples were screened for single/multiple exon deletions or duplications using a multiplex ligation-dependent probe amplification method (De Luca *et al.*, 2007) (SALSA P081/082-B2 NF1 MLPA assay kit; MRC-Holland) concurrently with the direct capillary sequencing of all the exons, as stated above.

Analysis algorithm of the variants

Missense variants that have not been reported as pathogenic in the literature and were not observed in the 1208 normal Japanese exome data were evaluated for potential pathogenicity using five bioinformatics programs, including SIFT (Kumar *et al.*, 2009), Polyphen2 (Adzhubei *et al.*, 2010), LRT (Chun and Fay, 2009), MutationTaster (Schwarz *et al.*, 2010), and PhyloP (Siepel *et al.*, 2009). When four of the five programs predicted the results as pathogenic ("damaging" with SIFT, "probably damaging" with PolyPhen2, "deleterious" with LRT, "disease causing" with MutationTaster, or "conserved" with PhyloP), we interpreted the clinical significance of the missense mutation as being putatively pathogenic.

Results

Performance of sequence capturing

In the custom-designed mutation analysis panel for the screening of classic genetic syndromes, the number of bases for targeted capturing was 459,952 bp over 1888 regions of the 109 target genes, including *NF1*. An average of 207,203 reads per sample were mapped and aligned uniquely to the targeted bases of the 109 genes among the 86 samples.

As far as the *NF1* locus was concerned, all the exons were highly covered with a coverage of 190.7x per sample. Overall, 99.3% of the regions were covered at least with a coverage of 5x and 98.8% of the regions were covered at least with a coverage of 30x. The mean coverage of all the exons in the 86 samples indicated that all the exons, but exon 1, were appropriate for base calling by next-generation sequencing (Table 3). Because of the poor coverage, exon 1 was sequenced using the direct capillary sequencing in all 86 samples, none of which had any variants.

The mean coverage over the entire targeted regions per sample was 131.0x, and most of the regions were well covered (Table 4). Overall, 97.1% of the regions were covered at least 5x coverage, and 84.4% of the regions were covered at least 30x coverage. Some exons of *NF1* and other regions were less well covered than others. Exon 15 and exon 50 of *NF1*, together with the *COMP* gene and the *PHOX2B* gene, had relatively low coverages of 86.9x, 79.1x, 55.3x, and 19.2x, respectively.

TABLE 3. MEAN COVERAGE OF *NF1* EXONS AMONG 86 PATIENTS

Exon	*Coverage (x)*	*Exon*	*Coverage (x)*
1	1.7	30	239.7
2	220.2	31	175.9
3	168.8	32	157.0
4	169.5	33	124.6
5	145.0	34	216.0
6	170.9	35	152.1
7	164.8	36	189.3
8	144.0	37	284.7
9	182.7	38	261.5
10	174.1	39	230.9
11	179.2	40	217.3
12	194.9	41	206.8
13	120.0	42	276.9
14	141.2	43	195.7
15	86.9	44	181.1
16	152.7	45	166.3
17	212.6	46	156.4
18	251.3	47	185.7
19	127.1	48	159.4
20	215.4	49	241.5
21	175.2	50	79.1
22	191.4	51	174.3
23	103.1	52	238.4
24	194.0	53	235.9
25	96.6	54	217.5
26	212.1	55	136.8
27	209.6	56	320.0
28	238.7	57	220.5
29	208.5	58	122.6

NF1 has seven highly homologous pseudogene sequences located in chromosomes other than chromosome 17 (2q12-q13, 12q11, 14p11-q11, 15q11.2, 18p11.2, 21p11-q11, and 22p11-q11), on which *NF1* resides (Upaddhyaya, 2008). We scrutinized the mapped reads among 10 arbitrarily selected patients; all the pseudogene sequences were mapped to their orthologous locations in the genome rather than the *NF1* locus on chromosome 17.

Coverage of the 108 genes other than the *NF1* gene was evaluated in all 86 samples. The mean coverage of all 108 genes on the same diagnostic panel indicated that the mean coverage ranged from 19.2x to 254.1x, with mean of 114.5x (Table 4).

Mutation detection

The next-generation sequencing protocol described above led to the identification of pathological *NF1* mutations in 70 of the 86 patients who met the NIH diagnostic criteria. The clinical information is listed in Table 5. All the 70 patients harbored mutations in a heterozygous state: 30 nonsense mutations, 19 frameshift mutations, 8 canonical splice-site mutations, and 6 point mutations that were previously reported and have been shown to lead to aberrant splicing

TABLE 4. SUMMARY OF THE COVERAGE OF 109 GENES

Gene	*Coverage (x)*	*Gene*	*Coverage (x)*
ACTA2	103.7	*MSX1*	49.4
ACTC1	111.4	*MYH7*	103.5
ACVRL1	60.4	*MYH9*	97.5
BRAF	160.0	*NF1*	190.7
CBL	192.3	*NIPBL*	175.9
CDKL5	146.1	*NOTCH2*	153.4
CHD7	150.6	*NRAS*	254.1
COL11A1	160.5	*NRTN*	45.8
COL11A2	66.8	*NSD1*	160.1
COL1A1	47.2	*OTX2*	115.1
COL1A2	127.0	*PHOX2B*	19.2
COL2A1	76.2	*PKHD1*	173.6
COL3A1	123.1	*PLOD1*	68.3
COL5A1	52.0	*PSPN*	66.5
COL5A2	159.2	*PTCH1*	111.0
COL9A1	147.4	*PTPN11*	152.6
COL9A2	52.4	*RAD21*	198.5
COMP	55.3	*RAF1*	154.9
CREBBP	50.1	*RASA1*	171.7
CUL7	68.8	*RET*	97.4
DCC	188.4	*RUNX2*	144.5
DDX3X	118.1	*SALL1*	91.7
ECE1	80.6	*SALL4*	93.8
EDN3	64.6	*SCN1B*	69.3
EDNRB	178.9	*SHH*	50.3
EFNB1	47.8	*SHOC2*	195.5
ENG	36.4	*SIX3*	80.0
EP300	191.0	*SIX6*	67.6
FBN1	177.2	*SMC1A*	134.7
FBN2	171.0	*SMC3*	157.2
FGFR1	102.7	*SOS1*	180.5
FGFR2	157.5	*SOX10*	45.1
FGFR3	34.8	*SOX2-OT*	89.0
GDNF	200.5	*SPRED1*	137.0
GFRA1	103.1	*SPRY2*	141.7
GFRA2	49.9	*STAG1*	193.3
GLA	121.1	*TAZ*	45.1
HRAS	44.4	*TBX22*	117.7
IHH	73.4	*TBX5*	124.2
IRF6	128.5	*TCF4*	170.8
JAG1	147.5	*TCOF1*	68.4
KCNE1	88.4	*TGFBR1*	190.0
KCNJ2	226.4	*TGFBR2*	89.6
KCNQ1	80.5	*TGIF1*	77.1
KIAA1279	186.5	*TP63*	182.5
KIF26A	33.7	*TRAPPC10*	139.7
KRAS	214.4	*TRIM37*	85.4
L1CAM	42.7	*TSC1*	157.8
LAMP2	128.2	*TSC2*	49.4
MAP2K1	151.4	*TWIST1*	47.9
MAP2K2	35.6	*VHL*	84.5
MAPK1	168.5	*VSX2*	29.7
MAPK3	87.1	*ZEB2*	218.9
MECP2	80.4	*ZIC2*	72.9
MID1	126.4		

according to reverse transcription (RT)-PCR studies, together with seven nonsynonymous substitutions (Table 5). Among the seven nonsynonymous substitutions, four were previously reported to be pathogenic based on functional assays or the inheritance pattern within the families (Li *et al.*, 1992; Fahsold *et al.*, 2000; Lee *et al.*, 2006).

Three samples with missense mutations that have never been reported in the literature were predicted to be pathogenic based on the consensus predication from multiple bioinformatics programs. Five programs, including SIFT, Polyphen2, LRT, Mutation Taster, and PhyloP, predicted potential pathogenicity as follows: c.2183T>G (p.Val728Gly) mutation was predicted to be pathogenic by all five programs, and c.2540T>G (p.Leu847Arg) and c.6818A>T (p.Lys2273Met) mutations were predicted to be pathogenic by four of the five bioinformatics programs. None of the three missense mutations resided within the critical functional domain, GAP-related domain that regulates the RasGAP activity.

Comparison of the distributions of nonsense, splice-site variants, and missense mutations in the Japanese population versus the northern European population, as reported by Messiaen *et al.* (2000), Nemethova *et al.* (2013), Sabbagh *et al.* (2013), and Valero *et al.* (2011), revealed no statistically significant differences among the groups ($p=0.203$ using the Fisher exact test for countable data).

Together with these 3 samples, which were subject to bioinformatics programs, 16 samples without truncating mutations or missense mutations, previously reported to be pathogenic, were further sequenced using direct capillary sequencing methods. All the exons were sequenced, including exon 1, and no additional point mutations or small indels were detected. These 19 patients were further screened for relatively large deletions that would span an entire exon or multiple exons and thus escape from direct capillary sequencing. Among 10 patients, 5 were shown to have a whole *NF1* deletion, 2 had multiple-exon deletions, and 3 had single-exon deletions. These five patients with a whole *NF1* deletion were apparently homozygous for all the SNPs for the entire *NF1* region according to the next-generation sequencing analysis.

Overall, no appreciable genotype–phenotype correlation was detected in the present study (Table 5). Variants were detected in genes other than *NF1* when the same criteria used in the *NF1* analysis were applied to these genes (Table 5). None of these variants was classified as truncating mutations and none of them listed in the Human Genome Mutation Database (HGMD) (Cooper *et al.*, 1998). Such rare variants of unknown significance among the genes on the panel were found in at least two-thirds of the patients. Patients with variants in genes other than *NF1* did not necessarily exhibit a severe *NF1* phenotype.

Discussion

The present study demonstrated that next-generation sequencing with in-solution hybridization-based enrichment provides a high mutation detection rate comparable to that of conventional direct capillary sequencing methods for the molecular diagnosis of neurofibromatosis. The overall mutation detection rate using the currently reported method in 86 patients who met the clinical diagnostic criteria was 81.4% (70/86). Among the 16 samples in which mutations were not detected using next-generation sequencing, 10 samples were later shown to have large deletions using a different method, multiplex ligation-dependent probe amplification (MLPA). Because of their large sizes, the 10 large deletions would not have been detected using the direct capillary sequencing

TABLE 5. SUMMARY OF PATHOGENIC MUTATIONS DETECTED BY NEXT-GENERATION SEQUENCING

Exon	*Genomic mutation*	*Amino acid substitution*	*Type of mutation*	*Reference*	*Age*	*Familial*	*Symptoms*	*Variations of unknown significance in rasopathy genes*	*Number of mutations in other genes*
2	c.83_84insG	p.Asn29Glufs*9	Frameshift		68	Yes	P,N	RASA1 c.293C>T p.Ala98Val	2
3	c.264_265insA	p.Thr89Asnfs*18	Frameshift		44	Yes	P,B,N		1
5	c.491T>A	p.Leu164*	Nonsense		50	Yes	P,B,O,N		1
5	c.495-498delTGTT	p.Cys167Glnfs*10	Frameshift		41	No	P,N,L		1
5	c.499_500insG	p.Cys167Trpfs*7	Frameshift		27	No	P,B,N,L		1
5	c.574C>T	p.Arg192*	Nonsense		32	No	P,N,L		2
10	c.1105C>T	p.Gln369*	Nonsense		40	Yes	P,N,L		1
11	c.1241T>G	p.Leu414Arg	Missense[a]	Lee *et al.* (2006)	21	No	P,N,L		1
11	c.1246C>T	p.Arg416*	Nonsense		32	Yes	P,B,N		1
12	c.1381C>T	p.Arg461*	Nonsense		3	No	P	RASA1 c.669G>C p.Gln223His	1
12	c.1381C>T	p.Arg461*	Nonsense		67	Yes	P,B,N		1
12	c.1381C>T	p.Arg461*	Nonsense		41	Yes	P,B,N		0
13	c.1466A>G	p.Tyr489Cys	Missense[a]	Messiaen *et al.* (2000)	36	No	P,N		1
13	c.1466A>G	p.Tyr489Cys	Missense[a]	Messiaen *et al.* (2000)	63	Yes	P,B,N		0
13	c.1466A>G	p.Tyr489Cys	Missense[a]	Messiaen *et al.* (2000)	71	No	P,N,L		1
13	c.1527+1_+4delGTAA		Splicing		30	No	P,N,L		2
14	c.1541_1542delAG	p.Gln514Argfs*43	Frameshift		52	No	P,B,N		1
15	c.1721+3A>G		Splicing	Purandare *et al.* (1994)	40	Yes	P,B,N		0
16	c.1726C>T	p.Gln576*	Nonsense		36	No	P,N		0
16	c.1754_1757delACTA	p.Thr586Valfs*18	Frameshift		49	Yes	P,N		0
16	c.1765C<T	p.Gln589*	Nonsense		40	No	P,N		1
16	c.1832delT	p.Asn614Ilefs*17	Frameshift		80	No	P,N,L		3
17	c.1876_1877insT	p.Tyr628Leufs*6	Frameshift		79	Yes	P,B,N,L		2
17	c.1885G>A	p.Gly629Arg	Missense[a]	Gasparini *et al.* (1996)	57	Yes	P,N		2
18	c.2041C>T	p.Arg681*	Nonsense		23	No	P,N		1
18	c.2041C>T	p.Arg681*	Nonsense		35	Yes	P,B,N		1
18	c.2087G>A	p.Trp696*	Nonsense		58	Yes	P,B,N,L		0
18[b]	c.2183T>G	p.Val728Gly	Missense		67	Yes	P,N		0
21	c.2423delT	p.His809Thrfs*12	Frameshift		43	Yes	P,N		1
21	c.2540T>C	p.Leu847Pro	Missense[a]	Fahsold *et al.* (2000)	33	Yes	P,N,L		0
21	c.2540T>C	p.Leu847Pro	Missense[a]	Fahsold *et al.* (2000)	59	Yes	P,B,N,L		0

(*continued*)

TABLE 5. (CONTINUED)

Exon	*Genomic mutation*	*Amino acid substitution*	*Type of mutation*	*Reference*	*Age*	*Familial*	*Symptoms*	*Variations of unknown significance in rasopathy genes*	*Number of mutations in other genes*
21[b]	c.2540T > G	p.Leu847Arg	Missense		55	No	P,N		0
21	c.2446C > T	p.Arg816*	Nonsense		52	Yes	P,N,L		0
22	c.2851-5_-2delTTTA		Splicing		19	No	P,B,N,L		1
23	c.3048T > A	p.Cys1016*	Nonsense		50	Yes	P,B,N		0
24	c.3132C > A	p.Tyr1044*	Nonsense		12	Yes	P,O,N		0
25	c.3213_3214delAA	p.Ser1072Hisfs*16	Frameshift		29	No	P,N,L		2
27	c.3595_3596insGG	p.Thr1199Argfs*17	Frameshift		20	No	P,N,L		1
27	c.3615_3616delTG	p.Phe1205Leufs*12	Frameshift		37	Yes	P,B,N		2
27	c.3615_3616delTG	p.Phe1205Leufs*12	Frameshift		64	Yes	P,B,N,L		1
28	c.3709-2A > G		Splicing		44	No	P,B,N,L		0
28	c.3765_3766insCT	p.Leu1257Cysfs*10	Frameshift		29	No	P,B,N,L		2
28	c.3826C > T	p.Arg1276*	Nonsense		21	No	P,O,B,N,L		0
29	c.3888T > A	p.Tyr1296*	Nonsense		49	No	P,N,L		0
30	c.4084C > T	p.Arg1362*	Nonsense		27	No	P,N		1
32	c.4329delA	p.Lys1444Argfs*25	Frameshift		50	Yes	P,B,N,L		0
32	c.4330A > G	p.Lys1440Glu	Missense[a]	Li *et al.* (1992)	40	No	P,N,L		0
33	c.4430 + 1G > A		Splicing		49	Yes	P,B,N		2
34	c.4544delA	p.Gln1515Argfs*59	Frameshift		35	Yes	P,N		2
35	c.4716_4724 + 6 delTATGACTAGGTAAAG		Splicing		50	No	P,B,N,L		1
36	c.4743_4744delAG	p.Glu1582Argfs*39	Frameshift		36	No	P,B,N,L		2
36	c.4769T > G	p.Leu1590*	Nonsense		45	No	P,N		1
37	c.4873_4874insA	p.Tyr1625*	Nonsense		63	No	P,B,N		1
37	c.5198T > G	p.Leu1733*	Nonsense		40	No	P,B,N,L		1
38	c.5269-6_5276 delTTCCAGGTTGGTTC		Splicing		38	No	P,N,L		1
38	c.5269-1G > A		Splicing		39	Yes	P,B,N,L		0
38	c.5516_5517insC	p.Glu1841Profs*21	Frameshift		31	Yes	P,B,N		1
38	c.5609G > A	p.Arg1870Gln	Missense[a]	Ars *et al.* (2003)	69	Yes	P,B,N		0
40	c.5902C > T	p.Arg1968*	Nonsense		22	No	P,N		1
44	c.6675G > A	p.Trp2225*	Nonsense		54	No	P,O,B,N		3
45	c.6772C > T	p.Arg2258*	Nonsense		69	Yes	P,N		0
45	c.6772C > T	p.Arg2258*	Nonsense		52	Yes	P,B,N,L		1
45[b]	c.6818A > T	p.Lys2273Met	Missense		46	No	P,N		1

(*continued*)

TABLE 5. (CONTINUED)

Exon	*Genomic mutation*	*Amino acid substitution*	*Type of mutation*	*Reference*	*Age*	*Familial*	*Symptoms*	*Variations of unknown significance in rasopathy genes*	*Number of mutations in other genes*
46	c.6850_6853delACTT	p.Tyr2285Thrfs*5	Frameshift		42	Yes	P,N		1
46	c.6853_6854insA	p.Tyr2285*	Nonsense		21	No	P,N		0
46	c.6853_6854insA	p.Tyr2285*	Nonsense		28	No	P,N		0
46	c.6904C>T	p.Gln2302*	Nonsense		37	Yes	P,N,L		1
47	c.6950G>A	p.Trp2317*	Nonsense		25	No	P,B,N,L		0
50	c.7348C>T	p.Arg2450*	Nonsense		46	No	P,B,N,L		0
54	c.7970+1_+4delGTAA		Splicing		41	Yes	P,N,L		2
			ex1 to 58 deletion		13	No	P,N,L		3
			ex1 to 58 deletion		29	No	P,N		1
			ex1 to 58 deletion		68	No	P,N		1
			ex1 to 58 deletion		58	No	P,B,N,L		1
			ex1 to 58 deletion		34	No	P,B,N		1
			ex1 deletion		68	No	P,N,L		1
			ex3 to 4 deletion		59	No	P,N,L		0
			ex6 to 51 deletion		36	Yes	P,N,L		2
			ex8 deletion		28	Yes	P,N		0
			ex12 deletion		55	No	P,N		1
					37	No	P		0
					50	No	P,N		0
					45	Yes	P,N,L		2
					30	No	P,N		0
					34	Yes	P,B,N		1
					25	No	P		0

[a]Previously reported to cause aberrant splicing.
[b]Predicted to be pathogenic by bioinformatics programs.
Symptoms: P, pigment; O, optic nerve tumor; B, bone manifestation; N, neurofibroma; L, Lisch nodules; HGMD; Human Genome Mutation Database.

method, which is currently considered to be the gold standard. The mutation detection rate was 92.1% (70/76) when these 10 samples were excluded from the calculation of the detection rate.

Among the 10 samples with large deletions, 5 patients with a whole *NF1* deletion could have been suspected of having a whole gene deletion, in that these patients were apparently homozygous for all the SNPs for the entire *NF1* region according to the next-generation sequencing data. The remaining five patients with a partial deletion of the *NF1* gene, as documented using MLPA, would not have been reliably inferred to have such a deletion based on the relatively short runs of homozygosity.

Recent reports on comprehensive *NF1* screening using the direct capillary sequencing method revealed that the detection rate was 89.5–96.3% when cases with large deletions detectable only by using MLPA were excluded [93.4%: Valero *et al.* (2011), 89.5%: Nemethova *et al.* (2013), 96.3%: Sabbagh *et al.* (2013)]. Hence, the performance of the presently reported protocol was comparable with that of the direct capillary sequencing methods.

The present protocol uses genomic DNA as the starting material, unlike other protocols using puromycin-tested Epstein-Barr virus cell lines as the starting material for RT-PCR (Messiaen *et al.*, 2000). Apparently, the use of genomic DNA is much easier in clinical settings. Yet, genetic testing based on genomic DNA, including the previously reported protocol, cannot predict potential splicing defects caused by point mutations. The use of RNA would be more sensitive to splicing abnormalities, if any, because of the possibility of mutations located deep in the intron or aberrant splicing defects caused by point mutations within coding sequences that were not evaluated in the presently reported protocol. However, such deep intronic mutations or splicing defects may be relatively rare, given the high overall detection rate of 92.1% in the present study.

The mean coverage of the entire target regions per sample was 131.0x. This coverage figure was considered to be sufficient for the detection of heterozygous base changes. Furthermore, the observation that rare variants in some genes on the panel were found in at least two-thirds of the patients supports the notion that the diagnostic performance of the panel for other genes is as robust as it is for *NF1*. Thus, our results regarding the validity of next-generation sequencing for the molecular diagnosis of the *NF1* gene, in comparison with direct capillary sequencing, can be extrapolated to the molecular diagnosis of other classic malformation syndromes.

Nevertheless, exon-to-exon variations in the coverage figures should be carefully evaluated. The extremely low coverage of the *NF1* exon1 can be ascribed to its extremely high GC content of 77.5%, in that a GC content of 60% or higher is associated with a sharp decrease in the read depth (Chilamakuri *et al.*, 2014). Similarly, a relatively low coverage of the *COMP* gene of 55.3x may be associated with a GC content of 63.4%. Exon 15 and exon 50 of *NF1*, together with the *PHOX2B* gene, had relatively low coverages of 86.9x, 79.1x, and 19.2x, respectively. The underlying cause of such variations is currently unexplained in that the GC contents of these regions were 32.2%, 39.4%, and 54.5%, respectively.

We estimated that the cost for consumables would be about USD 400 for direct capillary sequencing of the *NF1* gene, excluding labor costs. The estimated cost for consumables for the NGS panel analysis would be comparable. Hence, if we were to screen for the single *NF1* gene, the cost–benefit of next-generation sequencing may not be advantageous. However, if we were to screen for genes associated with conditions to be differentiated from neurofibromatosis using direct capillary sequencing, the consumable cost would be multiplied, whereas the cost for the screening of extra genes using next-generation sequencing would remain fixed. Indeed, the molecular diagnosis of Legius syndrome and Noonan syndrome would be helpful for the clinical management and outcome predictions of patients with café-au-lait spots, since patients with these conditions are unlikely to develop neurofibromas or other hamartomatous complications.

The availability of a mutation analysis panel, like the one presented herein, plays a critical role in differentiating the underlying genetic cause of patients whose diagnosis is uncertain from a clinical standpoint (Takenouchi *et al.*, 2013a, 2013b). The use of a whole-exome panel would be advantageous because of its comprehensiveness. However, apart from the higher cost of a whole-exome analysis, a panel approach enables a higher sensitivity (Chin *et al.*, 2013) because the average coverage, and thus the sensitivity, is higher using a panel approach (close to 100%) compared with a whole-exome approach (85%–95%).

Acknowledgments

All the authors would like to express their sincere appreciation to Mr. Yuji Sugie for his special support and all the patients and their families who were enrolled in this study. This work was partly supported by Research on Applying Health Technology and Research on Rare and Intractable Diseases from the Ministry of Health, Labour and Welfare, Japan.

Author Disclosure Statement

The authors declare that they have no competing interests.

References

Abernathy CR, Rasmussen SA, Stalker HJ, *et al.* (1997) *NF1* mutation analysis using a combined heteroduplex/SSCP approach. Hum Mutat 9:548–554.

Adzhubei IA, Schmidt S, Peshkin L, *et al.* (2010) A method and server for predicting damaging missense mutations. Nat Methods 7:248–249.

Aoki Y, Niihori T, Narumi Y, *et al.* (2008) The RAS/MAPK syndromes: novel roles of the RAS pathway in human genetic disorders. Hum Mutat 29:992–1006.

Aramaki M, Udaka T, Torii C, *et al.* (2006) Screening for CHARGE syndrome mutations in the *CHD7* gene using denaturing high-performance liquid chromatography. Genet Test Mol Biomarkers 10:244–251.

Ars E, Kruyer H, Morell M, *et al.* (2003) Recurrent mutations in the *NF1* gene are common among neurofibromatosis type 1 patients. J Med Genet 40:e82.

Bausch B, Borozdin W, Mautner VF, *et al.* (2007) Germline *NF1* mutational spectra and loss-of-heterozygosity analyses in patients with pheochromocytoma and neurofibromatosis type 1. J Clin Endocrinol Metab 92:2784–2792.

Carey JC, Viskochil DH (1999) Neurofibromatosis type 1: A model condition for the study of the molecular basis of variable expressivity in human disorders. Am J Med Genet 89:7–13.

Chilamakuri CS, Lorenz S, Madoui MA, *et al.* (2014) Performance comparison of four exome capture systems for deep sequencing. BMC Genomics 15:449.

Chin E, Zhang V, Wang J, *et al.* (2013) Frequently asked questions about the clinical utility of next-generation sequencing in molecular diagnosis of human genetic diseases. In: Wong L-JC (ed) Next Generation Sequencing: Translation to Clinical Diagnostics. Springer Science+Business Media, New York, pp 287–299.

Chun S, Fay JC (2009) Identification of deleterious mutations within three human genomes. Genome Res 19:1553–1561.

Cingolani P, Platts A, Wang le L, *et al.* (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w^{1118}; iso-2; iso-3. Fly 6:80–92.

Cooper DN, Ball EV, Krawczak M (1998) The human gene mutation database. Nucleic Acids Res 26:285–287.

De Luca A, Bottillo I, Dasdia MC, *et al.* (2007) Deletions of *NF1* gene and exons detected by multiplex ligation-dependent probe amplification. J Med Genet 44:800–808.

Fahsold R, Hoffmeyer S, Mischung C, *et al.* (2000) Minor lesion mutational spectrum of the entire *NF1* gene does not explain its high mutability but points to a functional domain upstream of the GAP-related domain. Am J Hum Genet 66: 790–818.

Frommolt P, Abdallah AT, Altmüller J, *et al.* (2012) Assessing the enrichment performance in targeted resequencing experiments. Hum Mutat 33:635–641.

Gasparini P, D'Agruma L, Pio de Cillis G, *et al.* (1996) Scanning the first part of the neurofibromatosis type 1 gene by RNA-SSCP: identification of three novel mutations and of two new polymorphisms. Hum Genet 97:492–495.

Han SS, Cooper DN, Upadhyaya MN (2001) Evaluation of denaturing high performance liquid chromatography (DHPLC) for the mutational analysis of the neurofibromatosis type 1 (NF1) gene. Hum Genet 109:487–497.

Hattori M, Torii C, Yagihashi T, *et al.* (2009) Diagnosis of Russell-Silver syndrome by the combined bisulfite restriction analysis—denaturing high-performance liquid chromatography assay. Genet Test Mol Biomarkers 13:623–630.

Japanese Genetic Variation Consortium (2013) Human genetic variation browser. Available at www.genome.med.kyoto-u.ac.jp/SnpDB (accessed March 18 2014).

Jones K (2005) Smith's Recognizable Patterns of Human Malformation. Saunders, Philadelphia.

Kosaki K, Udaka T, Okuyama T (2005) DHPLC in clinical molecular diagnostic services. Mol Genet Metab 86:117–123.

Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 4:1073–1081.

Lee MJ, Su YN, You HL, *et al.* (2006) Identification of forty-five novel and twenty-three known *NF1* mutations in Chinese patients with neurofibromatosis type 1. Hum Mutat 27:832.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler Transform. Bioinformatics 25:1754–1760.

Li Y, Bollag G, Clark R, *et al.* (1992) Somatic mutations in the neurofibromatosis 1 gene in human tumors. Cell 69:275–281.

McKenna A, Hanna M, Banks E, *et al.* (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20: 1297–1303.

Messiaen LM, Callens T, Mortier G, *et al.* (2000) Exhaustive mutation analysis of the *NF1* gene allows identification of 95% of mutations and reveals a high frequency of unusual splicing defects. Hum Mutat 15:541–555.

Nemethova M, Bolcekova A, Ilencikova D, *et al.* (2013) Thirty-nine novel neurofibromatosis 1 (NF1) gene mutations identified in Slovak patients. Ann Hum Genet 77:364–379.

Neurofibromatosis Conference Statement (1988) National Institutes of Health Consensus Development Conference. Arch Neurol 45:575–578.

Purandare SM, Huntsman Breidenbach H, *et al.* (1995) Identification of neurofibromatosis 1 (NF1) homologous loci by direct sequencing, fluorescence *in situ* hybridization, and PCR amplification of somatic cell hybrids. Genomics 30:476–485.

Purandare SM, Lanyon WG, Connor JM (1994) Characterisation of inherited and sporadic mutations in neurofibromatosis type-1. Hum Mol Genet 3:1109–1115.

Richards CS, Bale S, Bellissimo DB, *et al.* (2008) ACMG recommendations for standards for interpretation and reporting of sequence variations: revisions 2007. Genet Med 10: 294–300.

Rozen S, Skaletsky H (2000) Primer3 on the www for general users and for biologist programmers. Methods Mol Biol 132:365–386.

Sabbagh A, Pasmant E, Imbard A, *et al.* (2013) NF1 molecular characterization and neurofibromatosis type I genotype-phenotype correlation: the French experience. Hum Mutat 34: 1510–1518.

Samejima H, Torii C, Kosaki R, *et al.* (2007) Screening for Alagille syndrome mutations in the *JAG1* and *NOTCH2* genes using denaturing high-performance liquid chromatography. Genet Test Mol Biomarkers 11:216–227.

Schwarz JM, Rödelsperger C, Schuelke M, *et al.* (2010) MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods 7:575–576.

Siepel A, Pollard KS, Haussler D (2009) New methods for detecting lineage-specific selection. Proceedings of the 10th International Conference on Research in Computational. Mol Biol 3909:190–205.

Takenouchi T, Hida M, Sakamoto Y, *et al.* (2013a) Severe congenital lipodystrophy and a progeroid appearance: mutation in the penultimate exon of *FBN1* causing a recognizable phenotype. Am J Med Genet A 161A:3057–3062.

Takenouchi T, Matsuzaki Y, Torii C, *et al.* (2014) *SOX9* dimerization domain mutation mimicking type 2 collagen disorder phenotype. Eur J Med Genet 57:298–301.

Takenouchi T, Shimizu A, Torii C, *et al.* (2013b) Multiple Café au Lait spots in familial patients with *MAP2K2* mutation. Am J Med Genet A 164A:392–396.

Teer JK, Green ED, Mullikin JC, *et al.* (2012) VarSifter: visualizing and analyzing exome-scale sequence variation data on a desktop computer. Bioinformatics 28:599–600.

Thorvaldsdóttir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14:178–192.

Udaka T, Imoto I, Aizu Y, *et al.* (2007) Multiplex PCR/liquid chromatography assay for screening of subtelomeric rearrangements. Genet Test Mol Biomarkers 11:241–248.

Udaka T, Kurosawa K, Izumi K, *et al.* (2006) Screening for partial deletions in the *CREBBP* gene in Rubinstein-Taybi syndrome patients using multiplex PCR/liquid chromatography. Genet Test Mol Biomarkers 10:265–271.

Udaka T, Torii C, Takahashi D, *et al.* (2005) Comprehensive screening of the thiopurine methyltransferase polymorphisms by denaturing high-performance liquid chromatography. Genet Test Mol Biomarkers 9:85–92.

Upaddhyaya M (2008) *NF1* gene structure and *NF1* genotype/phenotype correlations. In: Kaufmann D (ed) Neurofibromatoses. Karger, Basel, pp 46–62.

Valero MC, Martín Y, Hernández-Imaz E, *et al.* (2011) A highly sensitive genetic protocol to detect NF1 mutations. J Mol Diagn 13:113–122.

Yan D, Tekin M, Blanton SH, *et al.* (2013) Next-generation sequencing in genetic hearing loss. Genet Test Mol Biomarkers 17:581–587.

Address correspondence to:
Kenjiro Kosaki, MD, FACMG
Center for Medical Genetics
Keio University School of Medicine
35 Shinanomachi, Shinjuku-ku
Tokyo 160-8582
Japan

E-mail: kkosaki@z3.keio.jp

Genomics Data 2 (2014) 55–59

Contents lists available at ScienceDirect

Genomics Data

journal homepage: http://www.journals.elsevier.com/genomics-data/

Data in Brief

A definitive haplotype map of structural variations determined by microarray analysis of duplicated haploid genomes

Tomoko Tahira [a,*], Koji Yahara [b], Yoji Kukita [a,c], Koichiro Higasa [a,d], Kiyoko Kato [e], Norio Wake [e], Kenshi Hayashi [a,*]

[a] Medical Institute of Bioregulation, Kyushu University, Fukuoka, Japan
[b] Biostatistics Center, Kurume University, Kurume, Japan
[c] Research Institute, Osaka Medical Center for Cancer and Cardiovascular Diseases, Osaka, Japan
[d] Center for Genomic Medicine, Kyoto University, Kyoto, Japan
[e] Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

ARTICLE INFO

Article history:
Received 11 April 2014
Accepted 11 April 2014
Available online 24 April 2014

Keywords:
Complete hydatidiform moles
Definitive haplotypes
Single nucleotide polymorphism
Copy Number Variation
LD-bin

ABSTRACT

Complete hydatidiform moles (CHMs) are tissues carrying duplicated haploid genomes derived from single sperms, and detecting copy number variations (CNVs) in CHMs is assumed to be sensitive and straightforward methods. We genotyped 108 CHM genomes using *Affymetrix SNP 6.0* (GEO#: GSE18642) and *Illumina 1 M-duo* (GEO#: GSE54948). After quality control, we obtained 84 definitive haplotype consisting of 1.7 million SNPs and 2339 CNV regions. The results are presented in the database of our web site (http://orca.gen.kyushu-u.ac.jp/cgi-bin/gbrowse/humanBuild37D4_1/).

Specifications	
Organism/cell line/tissue	Homo sapiens/complete hydatidiform moles (CHMs)
Sex	Duplicated haploids whose genomes are from single sperms harboring X
Sequencer or array type	*Affymetrix SNP 6.0* and *Illumina 1 M-duo*
Data format	*Affymetrix* Raw data: CEL files, normalized data: SOFT, MINIML and TXT *Illumina* Raw data: GSE54948_ signal_intensities.txt.gz, normalized data: SOFT, MINIML, TXT and GSE54948_ matrix_processed.txt.gz
Experimental factors	Single nucleotide polymorphism (SNP), copy number variation (CNV), LD-bin, CNV segments, CNV regions, definitive haplotypes
Experimental features	Whole genome SNP/CNV haplotyping of 84 duplicated haploid samples
Consent	All patients (donors) gave their written informed consent before study entry.
Sample source location	Japan

* Corresponding authors at: Division of Genome Analysis, Research Center for Genetic Information, Medical Institute of Bioregulation, Kyushu University, Fukuoka 812–8582, Japan. Tel.: +81 92 642 6171.
E-mail addresses: tomo.tahira@gmail.com (T. Tahira), hayashi.kenshi@gmail.com (K. Hayashi).

Direct link to deposited data

http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE18642
http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE54948

Experimental design, materials and methods

Samples

Complete hydatidiform mole tissues dissected from patients and the blood sample of one patient served as sources of DNAs for array hybridization experiments as described previously [1]. The informed consent was obtained from each donor. This study was approved by the Institutional Review Board (Ethical Committee of Kyushu University).

SNP genotyping

The raw data files of *Affymetrix SNP 6.0* arrays (CEL files) and sample attribute files of 94 CHM samples and one blood sample that has passed quality control in the previous study [1] were reanalyzed by *Birdseed v2* of *Geotyping Console 4. 1. 1. 834* (*GTC 4.1*), together with CEL files and sample attribute files of 45 *HapMap-JPT* samples (obtained from *Affymetrix*). The locations of markers in genome coordinate of *GRCh37* were according to *GenomeWideSNP_6.na32* that was obtained from

http://dx.doi.org/10.1016/j.gdata.2014.04.006

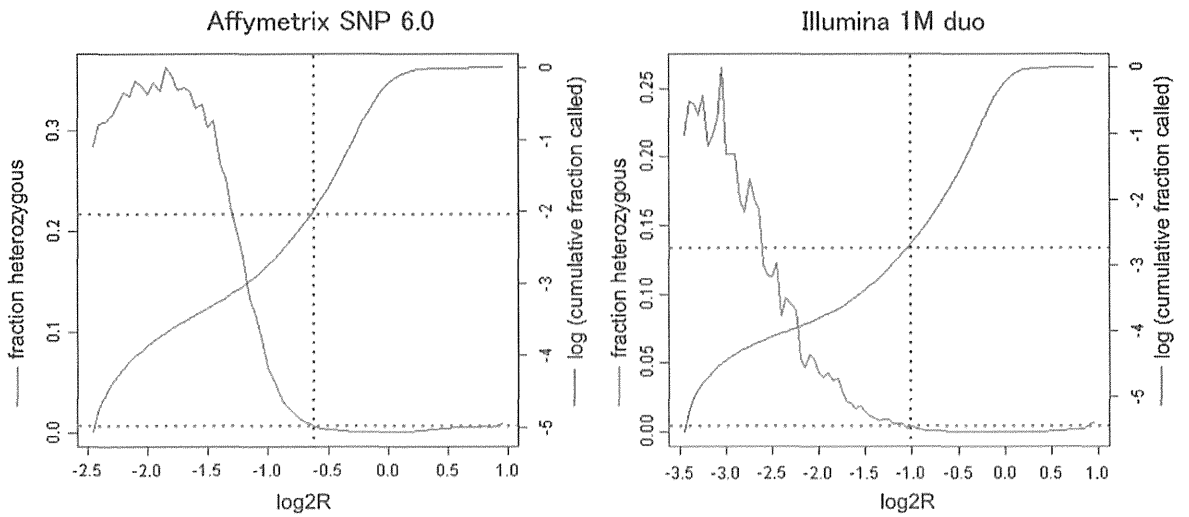

Fig. 1. Increased heterozygosity of calls at a low signal intensity. The genotype calls at the relative signal intensity where heterozygosity was approximately 1% (horizontal red dotted lines) or greater were regarded to contain significant fraction of unreliable calls. Blue horizontal lines indicate the fraction of cumulative calls at the reliability thresholds.

Affymetrix. A total of 905,025 SNP genotypes (excluding chromosome Y and mitochondria) were obtained, at an initial average call rate for the 94 CHMs of 99.2%.

Array hybridization experiments using *Illumina 1 M-duo* was performed for 98 CHM samples that included the 94 samples and one blood samples mentioned above by previously described procedures [1]. The genotypes were called using *GenTrain 2.0* cluster algorithm of *Genome Studio 2011.1*, *Illumina*. *Human1M-Duov3_H.egt* (based on *GRCh37*) was used as the manifest file and *Human1M-Duov3_H.bpm* as the cluster file. The initial average call rate was 99.5%.

Copy number analysis

The CEL files of *Affymetrix* arrays were subjected to *Copy Number/LOH analysis* module of *GTC 4.1* without regional GC correction. The 94 CHM samples, one blood sample mentioned above and four male samples from *HapMap JPT* (*NA18940*, *NA18943*, *NA18944* and *NA18945*) served as references to obtain "Log2Ratio" (abbreviated as log2R in this paper) data. Then, the data of markers on chromosome Y and mitochondria were excluded and the remaining data were exported as *CNCHP.txt*. The "log R Ratio" (abbreviated as logRR in this paper) data of *Illumina* arrays were calculated by *Genome Studio 2011.1* using the cluster file (*Human1M-Duov3_H.bpm*) as a reference.

Results and discussion

SNP genotyping of haploid samples

CHM genomes are supposed to be genome-widely homozygous. However, the genotypes obtained by the two systems revealed small fractions (0.27% of *Affymetrix* call and 0.01% of *Illumina* call) of heterozygous calls. The dramatic increase of heterozygous calls for the markers at lower relative signal intensities (log2R of *Affymetrix* arrays and logRR of *Illumina* arrays) indicated that the calls were falsely made for the markers at (homozygously) deleted regions where no genotypes should be called, although some of them might be ascribed to the markers in divergent paralogous regions (Fig. 1). These findings provided us an additional quality control measure of SNP genotype calling, that

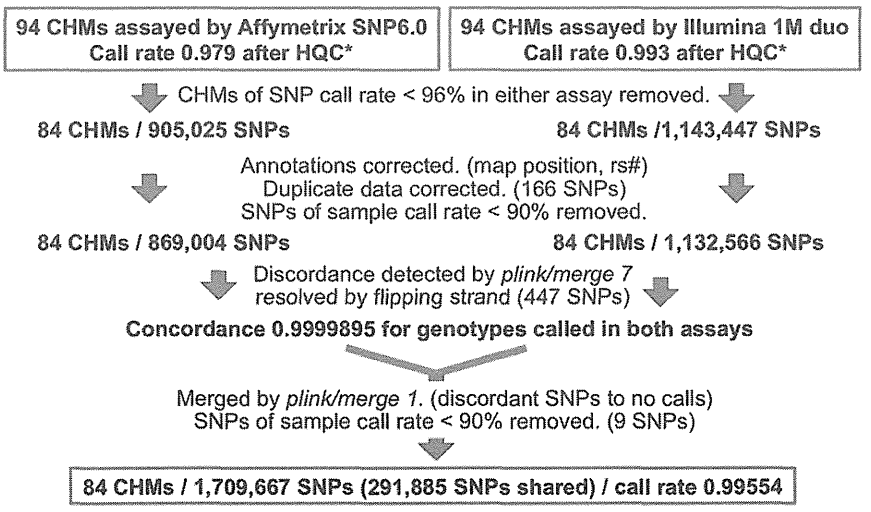

Fig. 2. Overview of SNP genotyping and its quality control. *HQC: haploid quality control, that is, heterozygous calls and weak signal calls were forced to no calls. See text for detail.