

Detecting tissue-specific alternative splicing and disease-associated aberrant splicing of the *PTCH* gene with exon junction microarrays

Kazuaki Nagao¹, Naoyuki Togawa², Katsunori Fujii³, Hideki Uchikawa^{1,3}, Yoichi Kohno³, Masao Yamada¹ and Toshiyuki Miyashita^{1,*}

This is the self-archive version of our report appeared in Hum. Mol. Genet., 14 (22), 3379-3388, 2005.

Our report is indexed in PubMed at http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=16203740&query_hl=2&itool=pubmed_docsum

The final version is available at the publisher's site, Oxford Journals, at <http://hmg.oxfordjournals.org/cgi/content/full/14/22/3379>

¹ Department of Genetics, National Research Institute for Child Health and Development, Tokyo 157-8535, JAPAN, ²Yokohama Research Laboratories, Mitsubishi Rayon Co., Ltd., Yokohama 230-0053, JAPAN, ³Department of Pediatrics, Graduate School of Medicine, Chiba University, Chiba 260-8670, JAPAN

* correspondence to Toshiyuki Miyashita
Department of Genetics, National Research Institute for Child Health and Development
2-10-1 Okura, Setagaya-ku, Tokyo 157-8535, JAPAN
Phone +81-3-3416-0181; Fax: +81-3-5494-7035; Email: tmiyashita@nch.go.jp

The nucleotide sequence data of human and mouse isoforms, +12b, has been deposited with the GenBank Library under Accession Nos. AB214500 and AB214501, respectively. Human isoforms, -3,4,5, +4' and -10, have been deposited with the GenBank Library under Accession Nos. AB233423, AB233424 and AB233422, respectively.

Abstract

Mutations in the human ortholog of *Drosophila patched* (*PTCH*) have been identified in patients with autosomal dominant nevoid basal cell carcinoma syndrome (NBCCS) characterized by minor developmental anomalies and an increased incidence of cancers such as medulloblastoma and basal cell carcinoma. We identified many isoforms of *PTCH* mRNA involving exons 1 to 5, exon 10 and a novel exon, 12b, generated by alternative splicing (AS), most of which have not been deposited in GenBank nor discussed earlier. In order to monitor splicing events of the *PTCH* gene, we designed oligonucleotide arrays on which exon probes and exon-exon junction probes as well as a couple of intron probes for the *PTCH* gene were placed in duplicate. Probe intensities were normalized based on the total expression of *PTCH* and probe sensitivity. Tissue-specific regulation of AS identified with the microarrays closely correlated with the results obtained by RT-PCR. Of note, the novel exon, exon 12b, was specifically expressed in the brain and heart, especially in the cerebellum. Additionally, using these microarrays, we were able to detect disease-associated aberrant splicings of the *PTCH* gene in two patients with NBCCS. In both cases, cryptic splice donor sites located either in an exon or in an intron were activated because of the partial disruption of the consensus sequence for the authentic splice donor sites due to point mutations. Taken together, oligonucleotide microarrays containing exon junction probes are demonstrated to be a powerful tool to investigate tissue-specific regulation of AS and aberrant splicing taking place in genetic disorders.

INTRODUCTION

Alternative splicing (AS) is one of the major mechanisms by which humans produce the complexity of the proteome. It has been estimated that greater than 55% of all genes and at least 74% of multi-exon genes are alternatively spliced in humans (1,2). Protein isoforms produced by AS can have antagonistic functions, such as anti-apoptotic Bcl-x_L vs. pro-apoptotic Bcl-x_S (3), or completely different amino acid compositions due to different reading frames, such as p16(INK4a) vs. p19(ARF) (4). In addition, AS is also implicated in pathophysiological processes and it has been estimated that at least 15% of point mutations that cause human genetic diseases affect splicing (5).

Nevoid basal cell carcinoma syndrome (NBCCS), also called Gorlin's syndrome, is an autosomal dominant neurocutaneous disorder characterized by large body size, developmental and skeletal abnormalities, radiation sensitivity, basal cell carcinoma (BCC), and an increased incidence of medulloblastoma (6). NBCCS is caused by inactivating mutations in the *Patched* (*PTCH*) gene (7,8). The human *PTCH* gene contains 23 exons spanning approximately 65 kb and is predicted to encode a protein of 1447 amino-acid residues containing 12 transmembrane-spanning domains and two large extracellular loops (7). Heterozygous loss of *PTCH* found in certain sporadic and familial cases of BCC indicates that *PTCH* is also a tumor suppressor gene (9,10). In vertebrates, a second *patched* gene (*PTCH2* in humans) was identified (11,12). To date, no mutations in *PTCH2* have been reported in NBCCS although a limited number of mutations were found in basal cell carcinoma and medulloblastoma (11).

Recently, we and others have identified that *PTCH* undergoes complex AS between multiple first exons and a second exon (13-16). In addition, we also have identified additional mRNA isoforms downstream of exon 2. Therefore, in order to survey tissue-specific AS and disease-associated aberrant splicing of *PTCH*, we have developed oligonucleotide microarrays designed for profiling AS. Current conventional microarray technologies are limited in their ability to distinguish and analyze mRNA isoforms. Several groups have reported a survey of human AS using exon junction microarrays that can circumvent this problem (2,17-20). In this study, considering the volume and complexity of the data produced by genome-wide studies and the cost-benefit ratio of these arrays, we have developed microarrays focused on the *PTCH* gene in which probes were designed for individual exons and splice junctions including junctions derived from rare AS. Using these arrays, we demonstrate a detailed evaluation of the data on tissue-specific regulation of AS. In addition, we describe the use of DNA microarrays to identify the aberrant splicing taking place in a genetic disorder.

RESULTS

Detection and validation of alternative splicing by oligonucleotide microarrays

To date, at least 5 alternatively used first exons have been reported in the *PTCH* gene (13-16). Recently, using RNA from various human tissues, we have identified additional mRNA isoforms generated by AS involving downstream exons (exon 2, 3, 4, 5, 10) as well as an alternative exon named 12b by RT-PCR and sequencing (Fig. 1A, B). The numbering of exons is according to Johnson et al. (7) with GenBank accession number U59464. Among these isoforms, the one skipping exon 10 has been briefly described previously (21) and the one skipping exon 4, 5 has been deposited in GenBank (Accession No. AB209495) but not

discussed in a prior publication. The rest of the isoforms are novel. Wicking et al. reported our exon 13 as exon 12b (22). However, since this is a constitutive exon, it is named as exon 13 in most publications and databases and the following exons are numbered by counting from 5' to 3'. No AS involving exon 13 or exons further downstream could be identified in any tissues examined. If alternative exons can be spliced independently, then the *PTCH* locus potentially encodes 483 isoforms. In order to monitor splicing events of the *PTCH* gene, we designed oligonucleotide arrays on which 17 exon probes and 23 exon-exon junction probes as well as a couple of intron probes were placed in duplicate (Fig. 1A). Before analyzing the data obtained from various tissues and NBCCS patients, we validated each probe using plasmids encoding various *PTCH* isoforms. For this purpose, each of the cDNA sequences encoding six *PTCH* isoforms was amplified by PCR followed by *in vitro* transcription and labeling with Cy-5 and then applied onto the microarrays. The intensity data prior to normalization represented as an intensity heat-map is shown in Fig. 1B. Each probe's intensity was then adjusted in two ways, first, for the mean probe intensity for each construct and then for each probe's sensitivity. Finally, mean probe intensities that should be positive according to the plasmid and probe sequences were adjusted to 1, and those that should be negative were adjusted to zero. As depicted in Fig. 1C, all constructs showed profiles expected from exon composition and could be clearly discriminated from one another. No intensities in-between were observed. However, when placing probes at exon-exon junctions, there is little choice as to the underlying nucleotide composition. Consequently, two junction probes, exon 9-11 and exon 12-13, were non-informative, e.g., saturated, constant intensity, and could lead to erroneous predictions, therefore excluded from the figure.

Tissue-specific regulation of alternative splicing

In order to investigate tissue-specific regulation of AS using array data, we normalized probe intensities as described in Materials and Methods and in Fig. 4. If no AS takes place at the exons or exon-exon junctions for which probes are designed, then the normalized relative probe intensities are near 1.0 in all tissues. In contrast, probes involved in AS should give intensities with large standard deviations. As expected from mRNA isoforms we have identified so far, a marked variation in normalized probe intensities was observed in exon 2-5, 10 and 12b (Fig. 2A). The tissue-specific accuracy of these array data was confirmed by RT-PCR. For example, when we focused on exon 12b, in tissues in which elevated probe intensities for exon 12-12b, 12b and 12b-13 were observed (*i.e.* the brain and heart), RT-PCR products containing exon 12b were also evident (Fig. 2B, lanes 1 and 8). We next addressed the question of where in the brain exon 12b was highly expressed. The subsequent investigation demonstrated that exon 12b was particularly expressed in the cerebellum among various brain tissues (Fig. 2B, lane 2). Additionally, two independent methods, microarray analysis and RT-PCR, demonstrated a strong correlation (Fig. 2C), validating the method of array data normalization. Array data as to inclusion or exclusion of exon 10 was similarly compared with the RT-PCR results. As shown in Fig. 2D, the two methods again well correlated with each other. Isoforms skipping exon 10 were the most highly expressed in the thymus and the lung based on both methods.

As shown in Fig. 1A, AS involving exon 2 through 5 is relatively complicated and can not be explained simply by inclusion or exclusion of a single exon. When we focused on isoforms skipping exon 4 and 5, the data from the two methods still correlated (Fig. 2E). However, the degree of correlation represented by R^2 was weaker than that obtained with probes for exon 12b or 9-10, because AS skipping exon 4 and 5 is a rare event (up to 3% of the authentic isoform (shown by X-axis in Fig. 2E).

Next, we evaluated the microarray data regarding the usage of the alternative first exons. RT-PCR was performed using the same forward primers for each alternative exon 1 as those used for the microarray analysis. As shown in Fig. 2F and 2G, the two methods showed a good correlation regarding the usage of exon 1b. However, isoforms starting from exon 1a, 1d, and 1e failed to demonstrate significant correlations (data not shown). This is probably because the expression of isoforms starting from exon 1a and 1e is much lower than that starting from exon 1b (average expression of exon 1a and 1e is 3.4% and 13.8% of exon 1b, respectively, when pooled S/N ratios are compared). In addition, the PCR efficiency of the exon 1d primer is significantly lower than that of the exon 1b primer (37% of the exon 1b primer) (Supplementary Material, Figure S1).

Detection of aberrant splicing in NBCCS patients

We have been investigating *PTCH* mutations in NBCCS patients (23) and have detected mutations in 13 out of 17 cases analyzed so far. A list of all cases is presented in Table S1. Among 13 cases, G17 had a mutation, c.584G>A (As per Genbank entry NM_000264.2: the A of the ATG of the initiator Met codon is counted as nucleotide +1), on exon 3. This raised two possibilities that explain the effects on the coding for the PTCH protein. One is a missense mutation, p.R195K (as per Genbank entry NP_000255.1). However, since this point mutation was located at the 3'-end of exon 3, and potentially disrupts a splice donor site, we sought the second possibility that the mutation may affect splicing. The data of the microarray analysis showed a significant decreased intensity for the junctional probe exon 3-4, but otherwise the data looked normal indicating an abnormal splicing between exon 3 and exon 4 (Fig. 3A). The electrophoretogram of the RT-PCR product revealed an additional band which had a larger molecular weight indicating the presence of an aberrant splicing (Fig. 3B, left panel). Sequencing of the additional product demonstrated that abnormal splicing was indeed taking place in which a cryptic splice donor site located in intron 3 was activated, resulting in the insertion of a 37-bp intronic sequence between exon 3 and exon 4 causing a premature termination of the PTCH protein (Fig. 3C). We therefore concluded that it was the aberrant splicing rather than the missense mutation that caused the disease phenotype in this patient.

Another case, G9, did not contain mutations in any of the coding exons of *PTCH*. However, the array data demonstrated a markedly decreased intensity for the junctional probe exon 6-7 (Fig. 3A). The detection of the smaller RT-PCR product in G9 (Fig. 3B, right panel) prompted us to sequence this RT-PCR product, which again revealed the presence of an aberrant splicing. But in this case, a cryptic splice donor site located in exon 6 was activated generating the deletion of an exonic sequence of 87 bp (Fig. 3C). As a result, 29 amino acid residues located in the first extracellular loop important for Shh binding were deleted. Mutational analysis in intron 6 identified a point mutation, c.945+5G>T, that partially disrupts the consensus sequence for the splice donor site (Fig. 3C).

The third case, G8, had a mutation, c.1526G>A, on exon 11. This mutation presumably results in a missense mutation, p.G509D. However, a growing body of evidence indicates that some single base changes may be capable of switching regulation from positive to negative or vice versa by disrupting or creating exonic splicing enhancers or exonic splicing silencers that are just beginning to be understood (24). Therefore, we investigated such possibilities using microarrays. As shown in Fig. 3A, the splicing profile of G8 was similar to that of the control and no significant changes in probe intensity were observed. Consequently, we concluded that the mutation found in G8 was a *bona fide* missense mutation. Interestingly, substitutions of

the same amino acid residue have been reported in NBCCS (p.G509R and p.G509V) (25,26) and these mutations are predicted to disrupt the sterol sensing domain of the PTCH protein (27). It has also been observed that the mutation p.G509V results in dominant negative activity *in vivo* in *Drosophila* (28). Two out of 17 patients, G7 and G13, did not have a mutation in the region we have sequenced, nor have any abnormal splicing profile by microarray analysis (Table S1). They may have a mutation in non-coding region important for the transcription of *PTCH*, or have a mutation in a gene other than *PTCH*. Altogether, these results demonstrate that exon junction microarrays can be used for detecting the mutations which result in aberrant splicing.

DISCUSSION

In this paper, we described the use of oligonucleotide microarrays containing exon junction probes to investigate tissue-specific AS and to detect disease-associated aberrant splicing. Our microarrays have been demonstrated to be a particularly powerful tool to detect the exon skipping/inclusion type of AS, the most common type (38%) of AS (29), in which even a rare AS event (no more than 3%) can be quantitatively detected using microarrays. Theoretically, previously unrecorded mRNA isoforms can be predicted with this method. If unexpectedly high variation is detected in certain probes, then, AS is suggested in this region.

In this study, we found a novel alternative exon named exon 12b. High expression of exon 12b in the brain, particularly in the cerebellum, is intriguing, since proliferative effects of sonic hedgehog (Shh), the ligand of the PTCH protein, in the external granular layer of the cerebellum are well characterized (30) and patients with NBCCS are prone to develop the cerebellar tumor medulloblastoma (6). Since exon 12b has an in-frame stop codon, the mRNA isoform containing this exon encodes a truncated PTCH protein that is presumably non-functional. However, this protein may have a dominant-negative effect on the wild-type protein and such a possibility needs to be ruled out. Importantly, exon 12b was also found in mice and was also expressed in a brain- and heart-specific manner (data not shown). Recently, the splicing regulatory element, UGCAUG, was reported to be evolutionarily conserved in introns that flank brain-specific alternative exons (31). Such an element was indeed identified in the intron between exon 12b and 13 both in humans and in mice (data not shown). In contrast, skipping of exon 10 leads to a 52 amino acid in-frame deletion in the second and third transmembrane domain. The effect on the function of the PTCH protein is currently unknown and this PTCH isoform may be functional in some context.

Detection of AS using exon junction microarrays is limited in several ways. Firstly, since the detection requires probes that match specific exons and splice junctions, probe selection is tightly constrained, and some probes are non-informative due to cross-hybridization with somewhere else in human genes. In our study, two out of 42 probes, exon 9-11 and exon 12-13, gave constitutive high intensities and therefore were excluded from further study. In addition, detection is based on differential expression. Therefore, if two isoforms are present in the same proportion in every tissue, no prediction will result, because normalized probe intensities will behave similarly to the pool. Thirdly, precaution should be taken when interpreting the data on alternative usage of 5'-terminal exons, because not only expression levels but also other factors such as the sequences of forward primers constructed for 5'-terminal exons can have an effect on probe sensitivities.

In this study, we monitored a relatively small number of AS events of one gene in a

relatively small set of samples. It should be noted, however, that this is the first report of microarrays used for the identification of aberrant splicings in a genetic disorder. In NBCCS, a considerable number of the patients do not have mutations within the coding region of *PTCH* (23, 32). In such patients, exon junction microarrays will be a valuable tool for detecting mutations affecting the splicing event. According to the *PTCH* mutation database (<http://www.cybergene.se/cgi-bin/w3-mysql/ptchbase/index.html>), at least 20 mutations have been reported to potentially result in abnormal splicing in *PTCH* and some of them have been proven experimentally (21,26). Thus, mutations having an effect on splicing events do not seem to be uncommon. Not only NBCCS, but an increasing number of genetic diseases are known to be caused by mutations that alter splicing in *cis* (at least 15% of point mutations) (5 and reviewed in 24). Some of these mutations weaken or activate *cis*-acting element such as intronic splicing enhancers or intronic splicing silencers that are sometimes located in intronic sequences distant from exons. In such cases, mutations are not identified by the sequencing of coding regions, warranting the use of exon junction microarrays like the ones described in this study to rapidly search the candidate regions to be sequenced.

Apart from *PTCH*, germline mutations in the genes encoding Shh signaling components such as *Shh* or *Gli3* are responsible for a variety of genetic disorders, most of which are accompanied by developmental anomalies in the central nervous system (33). In addition, somatic mutations of the genes involved in this signaling pathway, such as *PTCH2* or *Smoothed*, are associated with various sporadic cancers (33). Moreover, distinct roles of *PTCH2* splice variants in Shh signaling have recently been proposed (34). Therefore, oligonucleotide microarrays containing exon and exon junction probes widely covering these genes would be an attractive tool to study the pathogenesis of these disorders. In addition, these microarrays may also be useful to detect mutations in the dystrophin gene, since the removal of exon(s) is found in 60-65% of patients with Duchenne and Becker muscular dystrophies (35).

MATERIALS AND METHODS

Plasmids

The plasmid encoding myc-tagged full-length *PTCH* protein (exon 1b to 23) (pMyc-Ptc1) was kindly provided by Dr. Jeffrey Ming (15). The plasmids encoding other isoforms of *PTCH* were created by PCR-mediated mutagenesis as described previously (36) using pMyc-Ptc1 as a template. The details of construction are available on request.

Oligonucleotide microarray construction

All probes 34 to 76 bp in length were designed to have approximately the same annealing temperature (T_m). Splice junction probes between two exons were designed so that the T_m for the first exon sequence is the same as that for the second exon sequence (Supplementary Material, Table S2). The oligonucleotide microarray, GenoPalTM (Mitsubishi Rayon Co., Ltd.), was made in the following manner. Plastic hollow fibers were bundled in an orderly arrangement, and hardened with resin to form a block (Supplementary Material, Figure S2). Oligonucleotide capture probes were chemically bonded inside each hollow fiber with hydrophilic gel. The block was then sliced to make thin chips, each of which was set into a holder (<http://www.mrc.co.jp/genome/e/index.html> for details).

Preparation of labeled probe

Total RNA from a panel of human tissues was purchased from Ambion. Lymphoblastoid cell lines from NBCCS patients immortalized by Epstein-Barr virus were also used to obtain total RNA. All studies using patient samples were approved by the local ethic committee. The template for *in vitro* transcription was generated through a modified procedure using a SuperScript One-Step RT-PCR System with Platinum Taq (Invitrogen). In brief, 2.5 μ g of total RNA was reverse-transcribed with SuperScriptIII RT/Platinum Taq Mix and T7-exon15 reverse primer, 5'-TAATACGACTCACTATAGGGGTCATATTCTCTGGTTTCCCGAGGTACAATGTC-3', for 30min at 50°C. To evaluate the quality of RNA, T7-GAPDH reverse primer, 5'-TAATACGACTCACTATAGGGAGGAGGGGAGATTGAGTGTGGT-3' was also added in the reaction. The RNA was degraded with the addition of RNaseH (Invitrogen) for 15 min at 37°C. After the addition of forward primers specific for each first exon (5'-AGCGCCTGTTTACCCAGGAG-3' for exon 1a, 5'-GGACCGGGACTATCTGCACC-3' for exon 1b, 5'-AAATGCCGCGCCGGGGAGCAGCCT-3' for exon 1d and 5'-TTCTCGGCGGGGTCCAGTT-3' for exon 1e) and T7-exon15 reverse primer (final concentration 0.5 μ M each) and the activation of Platinum taq for 10min at 94°C, PCR was run for 30 cycles of denaturation at 94°C for 30sec, annealing at 56°C for 30sec, and extension at 72°C for 2min. Plasmid DNA encoding various isoforms of *PTCH* was also subjected to PCR using the exon 2 forward primer, 5'-GCTGAGAGCGAAGTTTCAGA-3', and T7-exon 15 reverse primer. PCR products generated from the reverse-transcribed cDNA and the plasmid DNA were purified with a PCR Purification kit (QIAGEN) according to the manufacturer's instructions. Then, a Cy-5-labeled probe was generated by using the PCR product as a template with a MEGAscript T7 kit (Ambion). The reaction contained a 1:1.5 mixture of uridine triphosphate (UTP) and Cy-5-UTP (Amersham Biosciences). The product was purified with an RNeasy mini kit (QIAGEN) following the manufacturer's instructions. One microgram of labeled probe was fragmented with RNA Fragmentation Reagents (Ambion) at 70°C for 3min.

Hybridization and detection

Hybridization was carried out in a final volume of 0.1 ml injected into a hybridization chamber at 65°C for 16–24 h in 0.5 \times SSC with 0.2% SDS. The microarray chip was then washed twice in 0.5 \times SSC with 0.2% SDS at 55°C for 20min and once in 0.5 \times SSC at 55°C for 10min, before being slowly cooled to room temperature. GenoPal™ was then scanned and the image was captured with a cooled CCD type Microarray Image Analyzer (Mitsubishi Rayon Co., Ltd.). Fluorescent intensity was analyzed with software developed by Mitsubishi Rayon Co., Ltd. Fluorescence throughout the three dimensional structure of each array feature can be efficiently captured due to the long focal depth of the optical system of the image analyzer (<http://www.mrc.co.jp/genome/e/index.html>).

Microarray data analysis

To analyze changes in the AS of a gene, changes in the total expression of a gene and differences in probe sensitivity should be separated and excluded. We used a simple and generalized pooling strategy presented by Le et al. (20) with modifications. The probe response P_{ij} (represented as an S/N ratio) for a specific probe i to a specific tissue sample j was normalized using the total expression of *PTCH* and probe sensitivity as described in Fig. 4.

ACKNOWLEDGEMENTS

We thank Kaori Takeuchi-Inoue and Mayu Yamazaki-Inoue for technical support, and Kayoko Saito for preparing the manuscript. This work was supported by Grants for Cancer Research and Child Health and Development from the Ministry of Health, Labour and Welfare; a Grant-in-Aid for Scientific Research and the Budget for Nuclear Research from the Ministry of Education, Culture, Sports, Science and Technology.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG Online.

(The supplemental materials presented to the publisher are included in this self-archived version at the bottom, after figures.)

REFERENCES

1. Kan, Z., Rouchka, E. C., Gish, W. R., and States, D. J. (2001) Gene structure prediction and alternative splicing analysis using genomically aligned ESTs. *Genome Res.*, **11**, 889-900.
2. Johnson, J. M., Castle, J., Garrett-Engle, P., Kan, Z., Loerch, P. M., Armour, C. D., Santos, R., Schadt, E. E., Stoughton, R., and Shoemaker, D. D. (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science*, **302**, 2141-2144.
3. Boise, L. H., Gonzalez-Garcia, M., Postema, C. E., Ding, L., Lindsten, T., Turka, L. A., Mao, X., Nunez, G., and Thompson, C. B. (1993) bcl-x, a bcl-2-related gene that functions as a dominant regulator of apoptotic cell death. *Cell*, **74**, 597-608.
4. Quelle, D. E., Zindy, F., Ashmun, R. A., and Sherr, C. J. (1995) Alternative reading frames of the INK4a tumor suppressor gene encode two unrelated proteins capable of inducing cell cycle arrest. *Cell*, **83**, 993-1000.
5. Krawczak, M., Reiss, J., and Cooper, D. N. (1992) The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Hum. Genet.*, **90**, 41-54.
6. Gorlin, R. J. (1987) Nevoid basal-cell carcinoma syndrome. *Medicine*, **66**, 98-113.
7. Johnson, R. L., Rothman, A. L., Xie, J., Goodrich, L. V., Bare, J. W., Bonifas, J. M., Quinn, A. G., Myers, R. M., Cox, D. R., Epstein, E. H., Jr., et al. (1996) Human homolog of *patched*, a candidate gene for the basal cell nevus syndrome. *Science*, **272**, 1668-1671.
8. Hahn, H., Wicking, C., Zaphiropoulos, P. G., Gailani, M. R., Shanley, S., Chidambaram, A., Vorechovsky, I., Holmberg, E., Uden, A. B., Gillies, S., et al. (1996) Mutations of the human homolog of *Drosophila patched* in the nevoid basal cell carcinoma syndrome. *Cell*, **85**, 841-851.
9. Gailani, M. R., Stahle-Backdahl, M., Leffell, D. J., Glynn, M., Zaphiropoulos, P. G., Pressman, C., Uden, A. B., Dean, M., Brash, D. E., Bale, A. E., et al. (1996) The role of the human homologue of *Drosophila patched* in sporadic basal cell carcinomas. *Nat. Genet.*, **14**, 78-81.
10. Uden, A. B., Holmberg, E., Lundh-Rozell, B., Stahle-Backdahl, M., Zaphiropoulos, P. G., Toftgård, R., and Vorechovsky, I. (1996) Mutations in the human homologue of *Drosophila patched* (*PTCH*) in basal cell carcinomas and the Gorlin syndrome: different *in vivo* mechanisms of *PTCH* inactivation. *Cancer Res.*, **56**, 4562-4565.
11. Smyth, I., Narang, M. A., Evans, T., Heimann, C., Nakamura, Y., Chenevix-Trench, G., Pietsch, T., Wicking, C., and Wainwright, B. J. (1999) Isolation and characterization of human *patched 2* (*PTCH2*), a putative tumour suppressor gene in basal cell carcinoma and medulloblastoma on chromosome 1p32. *Hum. Mol. Genet.*, **8**, 291-297.
12. Zaphiropoulos, P. G., Uden, A. B., Rahnama, F., Hollingsworth, R. E., and Toftgård, R. (1999) *PTCH2*, a novel human patched gene, undergoing alternative splicing and up-regulated in basal cell carcinomas. *Cancer Res.*, **59**, 787-792.
13. Kogerman, P., Krause, D., Rahnama, F., Kogerman, L., Uden, A. B., Zaphiropoulos, P. G., and Toftgård, R. (2002) Alternative first exons of *PTCH1* are differentially regulated *in vivo* and may confer different functions to the *PTCH1* protein. *Oncogene*, **21**, 6007-6016.
14. Ågren, M., Kogerman, P., Kleman, M. I., Wessling, M., and Toftgård, R. (2004) Expression of the *PTCH1* tumor suppressor gene is regulated by alternative promoters and a single functional Gli-binding site. *Gene*, **330**, 101-114.

15. Nagao, K., Toyoda, M., Takeuchi-Inoue, K., Fujii, K., Yamada, M., and Miyashita, T. (2005) Identification and characterization of multiple isoforms of a murine and human tumor suppressor, *Patched*, having distinct first exons. *Genomics*, **85**, 462-471.
16. Shimokawa, T., Rahnema, F., and Zaphiropoulos, P. G. (2004) A novel first exon of the *Patched1* gene is upregulated by Hedgehog signaling resulting in a protein with pathway inhibitory functions. *FEBS Lett.*, **578**, 157-162.
17. Castle, J., Garrett-Engle, P., Armour, C. D., Duenwald, S. J., Loerch, P. M., Meyer, M.R., Schadt, E. E., Stoughton, R., Parrish, M. L., Shoemaker, D. D., et al. (2003) Optimization of oligonucleotide arrays and RNA amplification protocols for analysis of transcript structure and alternative splicing. *Genome Biol.*, **4**, R66.
18. Wang, H., Hubbell, E., Hu, J. S., Mei, G., Cline, M., Lu, G., Clark, T., Siani-Rose, M. A., Ares, M., Kulp, D. C., et al. (2003) Gene structure-based splice variant deconvolution using a microarray platform. *Bioinformatics*, **19** (Suppl 1), i315-i322.
19. Yeakley, J. M., Fan, J. B., Doucet, D., Luo, L., Wickham, E., Ye, Z., Chee, M. S., and Fu, X. D. (2002) Profiling alternative splicing on fiber-optic arrays. *Nat. Biotechnol.*, **20**, 353-358.
20. Le, K., Mitsouras, K., Roy, M., Wang, Q., Xu, Q., Nelson, S. F., and Lee, C. (2004) Detecting tissue-specific regulation of alternative splicing as a qualitative change in microarray data. *Nucleic Acids Res.*, **32**, e180.
21. Smyth, I., Wicking, C., Wainwright, B., and Chenevix-Trench, G. (1998) The effects of splice site mutations in patients with naevoid basal cell carcinoma syndrome. *Hum. Genet.*, **102**, 598-601.
22. Wicking, C., Gillies, S., Smyth, I., Shanley, S., Fowles, L., Ratcliffe, J., Wainwright, B., and Chenevix-Trench, G. (1997) De novo mutations of the *Patched* gene in nevoid basal cell carcinoma syndrome help to define the clinical phenotype. *Am. J. Med. Genet.*, **73**,304-307.
23. Fujii, K., Kohno, Y., Sugita, K., Nakamura, M., Moroi, Y., Urabe, K., Furue, M., Yamada, M., and Miyashita, T. (2003) Mutations in the human homologue of *Drosophila patched* in Japanese nevoid basal cell carcinoma syndrome patients. *Hum. Mutat.*, **21**,451-452.
24. Garcia-Blanco, M. A., Baraniak, A. P., and Lasda, E. L. (2004) Alternative splicing in disease and therapy. *Nat. Biotechnol.*, **22**, 535-546
25. Chidambaram, A., Goldstein, A. M., Gailani, M. R., Gerrard, B., Bale, S. J., DiGiovanna, J. J., Bale, A. E., and Dean, M. (1996) Mutations in the human homologue of the *Drosophila patched* gene in Caucasian and African-American nevoid basal cell carcinoma syndrome patients. *Cancer Res.*, **56**, 4599-4601.
26. Pastorino, L., Cusano, R., Nasti, S., Faravelli, F., Forzano, F., Baldo, C., Barile, M., Gliori, S., Muggianu, M., Ghigliotti, G., et al. (2005) Molecular characterization of Italian nevoid basal cell carcinoma syndrome patients. *Hum. Mutat.*, **25**, 322-323.
27. Kuwabara, P.E., Labouesse, M. (2002) The sterol-sensing domain: multiple families, a unique role? *Trends Genet.*, **18**, 193-201.
28. Hime, G. R., Lada, H., Fietz, M. J., Gillies, S., Passmore, A., Wicking, C., and Wainwright, B. J. (2004) Functional analysis in *Drosophila* indicates that the NBCCS/PTCH1 mutation G509V results in activation of smoothed through a dominant-negative mechanism. *Dev. Dyn.*, **229**, 780-790.
29. Sugnet, C. W., Kent, W. J., Ares, M., Jr., and Haussler, D. (2004) Transcriptome and genome conservation of alternative splicing events in humans and mice. *Pac. Symp. Biocomput.*, 66-77.
30. Dahmane, N., Altaba, A. (1999) Sonic hedgehog regulates the growth and patterning of the cerebellum. *Development*, **126**, 3089-3100.

31. Minovitsky, S., Gee, S. L., Schokrpur, S., Dubchak, I., and Conboy, J. G. (2005) The splicing regulatory element, UGCAUG, is phylogenetically and spatially conserved in introns that flank tissue-specific alternative exons. *Nucleic Acids Res.*, **33**, 714-724.
32. Wicking, C., Shanley, S., Smyth, I., Gillies, S., Negus, K., Graham, S., Suthers, G., Haites, N., Edwards, M., Wainwright, B., et al. (1997) Most germ-line mutations in the nevoid basal cell carcinoma syndrome lead to a premature termination of the PATCHED protein, and no genotype-phenotype correlations are evident. *Am. J. Hum. Genet.*, **60**, 21-26.
33. Cohen, M. M., Jr. (2003) The hedgehog signaling network. *Am. J. Med. Genet.*, **123A**, 5-28.
34. Rahnama, F., Toftgård, R., and Zaphiropoulos, P. G. (2004) Distinct roles of PTCH2 splice variants in Hedgehog signalling. *Biochem.J.*, **378**, 325-334.
35. Muntoni, F., Torelli, S., and Ferlini, A. (2003) Dystrophin and mutations: one gene, several proteins, multiple phenotypes. *Lancet Neurol.*, **2**, 731-740.
36. Imai, Y., Matsushima, Y., Sugimura, T., and Terada, M. (1991) A simple and rapid method for generating a deletion by PCR. *Nucleic Acids Res.*, **19**, 2785.

Legends to Figures

Figure 1. Detection and validation of the known isoforms of *PTCH*. (A) Exon structure of the *PTCH* gene. Positions of the exon probes and the exon junction probes are indicated. (B) Probe intensities obtained with Cy-5-labeled RNA from constructs encoding various isoforms of *PTCH*. The labeling reaction was performed with T7 RNA polymerase by using the PCR product as a template obtained from the constructs. Each matrix point shows the S/N ratio of one exon or exon junction probe in one construct. Probes are ordered horizontally, 5' to 3', together with some probes for infrequently spliced junctions such as exon 2-6. Exon 3-4' is an exon junction probe between exon 3 and the alternative splice acceptor site located in intron 3. Probe sequences for 5'-most exons were not included in some of the constructs and were excluded from this figure. Hybridization samples form the vertical axis of each matrix. Exon compositions of the expression plasmids are briefly depicted at the bottom. (C) Normalized probe intensities. The mean positive intensity and the mean negative intensity are adjusted to 1 and 0, respectively.

Figure 2. Detection of tissue-specific regulation of AS. (A) AS profiles in various tissues. Probes were ordered horizontally as in Fig. 1B. Standard deviations for each probe are indicated in parentheses. (B) RT-PCR results from a primer pair hybridizing to exon 11 and exon 14. AS events corresponding to each RT-PCR product are depicted at the right. (C) Comparison of the data obtained by the two methods. The Y-axis represents normalized relative probe intensity shown in (A). RT-PCR products were applied onto Agilent Bioanalyzer and the percentage of the isoform containing exon 12b is presented by the X-axis. (D) Two data obtained by microarray and RT-PCR were compared as in (C). A primer pair was constructed on exon 8 and exon 13. X-axis represents the percentage of the isoform skipping exon 10. (E) The data obtained by microarray analysis and RT-PCR were compared as in (C). A primer pair was constructed on exon 1b and exon 6. The X-axis represents the percentage of the isoform skipping exon 4 and 5. (F, G) The data obtained by microarray analysis (probe 1b in (F) and junction probe 1b-2 in (G)) and RT-PCR were compared as in (C). A primer pair was constructed on exon 1b and exon 2. The X-axis represents the estimated molarity of the isoform starting from exon 1b.

Figure 3. Detection of aberrant splicing in NBCCS patients. (A) Normalized probe intensities obtained from NBCCS patients, G8, G9 and G17. Data was analyzed as described in Fig. 2A. (B) RT-PCR analysis of AS in the *PTCH* gene. Exon 3 (G17) and exon 6 (G9) and flanking exons were amplified by RT-PCR and subjected to agarose gel electrophoresis. The positions of extra bands not observed in a control healthy individual (Cont) are indicated by arrows. (C) Schematic representation of abnormal splicing identified in G17 and G9. The positions of the point mutations are indicated by asterisks.

Figure 4. Matrix representation. The relationship between probes and tissues can be represented by an m -by- n matrix P . The total number of probes is m , and n is the total number of tissues. Let M be an n -by- n diagonal matrix where \bar{P}_j represents the mean P in tissue j . The probe intensities normalized by total *PTCH* expression in each tissue can then be expressed as PM . Let N be an m -by- m diagonal matrix where \bar{p}_i represents the mean normalized P in probe i . The matrix NPM therefore represents the probe intensities normalized by both factors described above and $P_{ij}/\bar{p}_i\bar{P}_j$ is used in Fig. 2 and 3.

Legends to Supplemental Figures

Figure S1. Real-time PCR using cDNA for *PTCH* isoforms starting from exon 1b or 1d as a template and a pair of primers constructed for exon 1b or 1d (forward) and exon 2 (reverse). PCR was performed in triplicate. Blue and green lines depict representative amplification curves of exon 1b-2 and exon 1d-2, respectively. Both reactions were started using the equimolar cDNAs. Data was obtained with ABI PRISM 7700 (Applied Biosystems).

Figure S2. Schematic figures of oligonucleotide microarray construction. See materials and methods section for details.

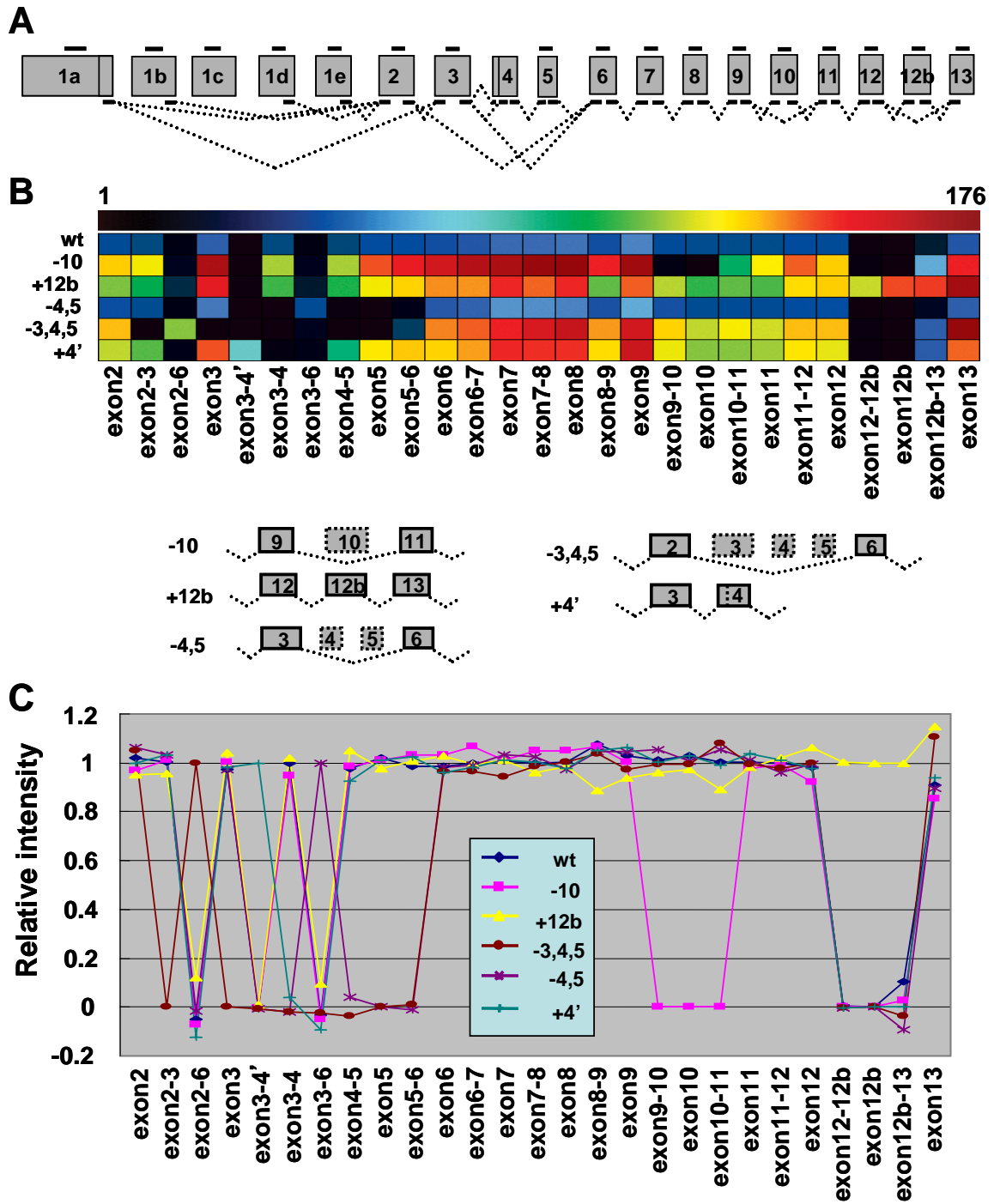


Fig. 1 Nagao et al.

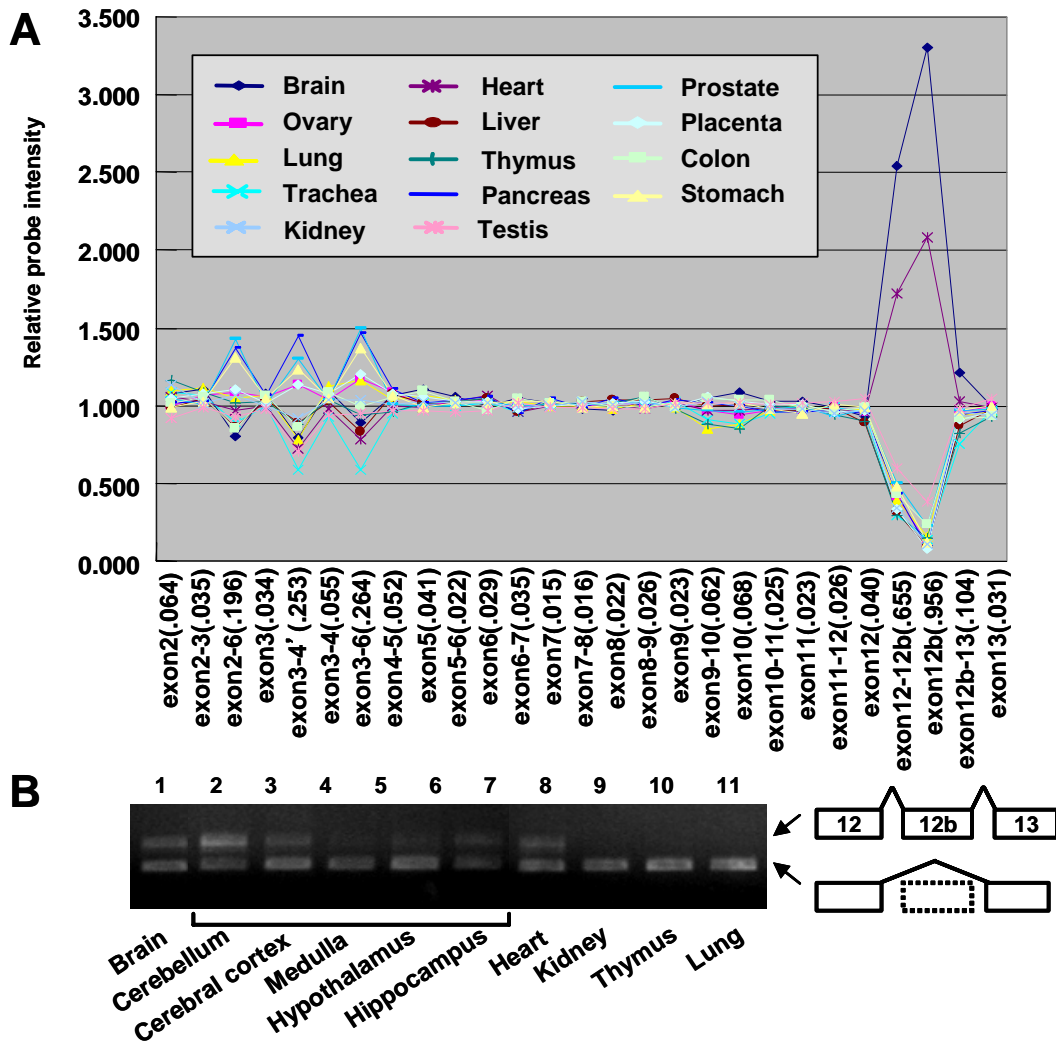


Fig. 2A,B Nagao et al.

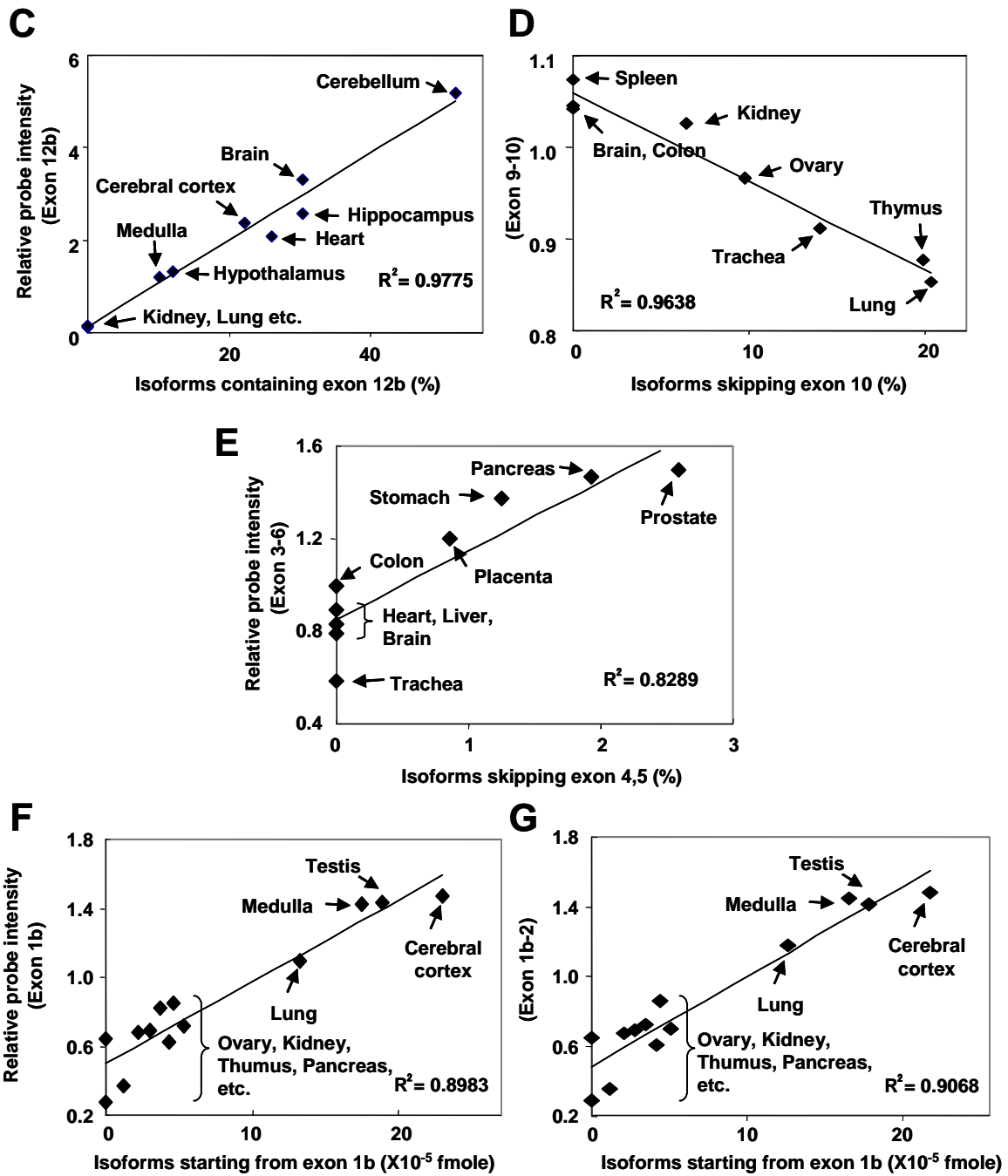


Fig. 2C-G Nagao et al.

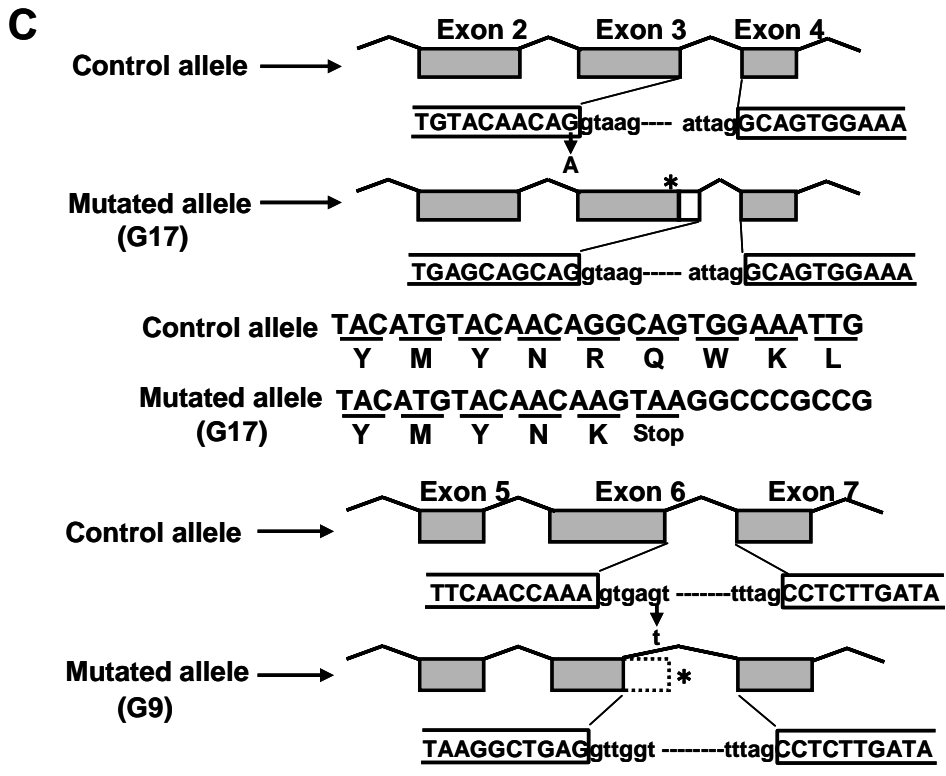
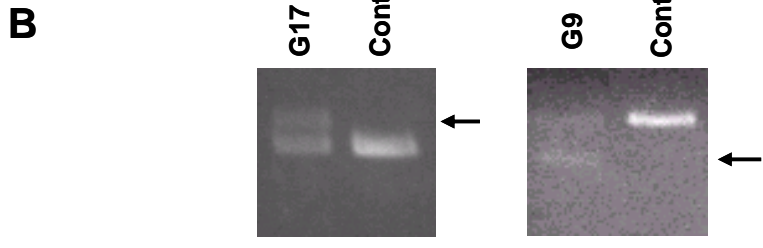
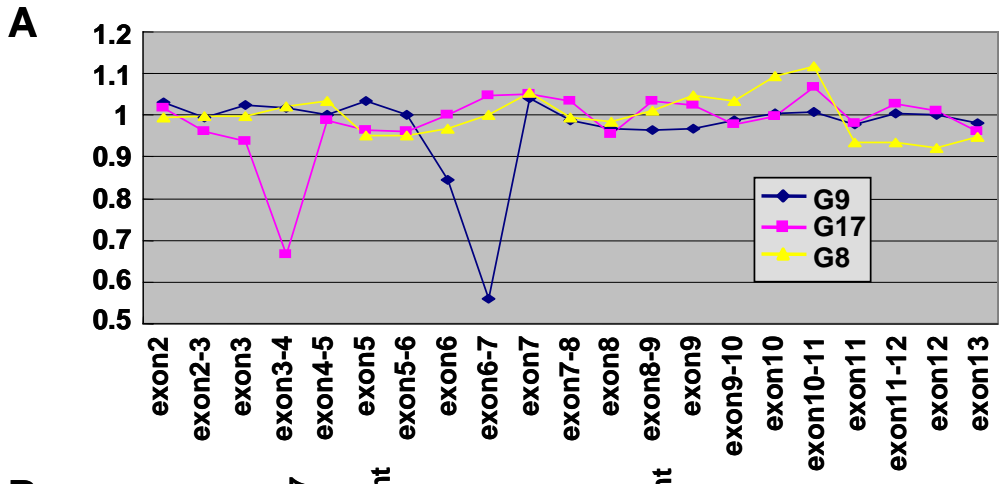


Fig. 3 Nagao et al.

$$P = \begin{pmatrix} P_{11} & \cdots & \cdots & \cdots & P_{1n} \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ P_{m1} & \cdots & \cdots & \cdots & P_{mn} \end{pmatrix} \quad M = \begin{pmatrix} 1/\bar{P}_1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1/\bar{P}_n \end{pmatrix} \quad \bar{P}_j = (\sum_i P_{ij})/m$$

$$PM = \begin{pmatrix} P_{11}/\bar{P}_1 & P_{12}/\bar{P}_2 & \cdots & \cdots & P_{1n}/\bar{P}_n \\ P_{21}/\bar{P}_1 & P_{22}/\bar{P}_2 & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ P_{m1}/\bar{P}_1 & \cdots & \cdots & \cdots & P_{mn}/\bar{P}_n \end{pmatrix}$$

$$N = \begin{pmatrix} 1/\bar{p}_1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1/\bar{p}_m \end{pmatrix} \quad \bar{p}_i = (\sum_j P_{ij}/\bar{P}_j)/n$$

$$NPM = \begin{pmatrix} P_{11}/\bar{p}_1\bar{P}_1 & P_{12}/\bar{p}_1\bar{P}_2 & \cdots & \cdots & P_{1n}/\bar{p}_1\bar{P}_n \\ P_{21}/\bar{p}_2\bar{P}_1 & & & & \vdots \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ P_{m1}/\bar{p}_m\bar{P}_1 & \cdots & \cdots & \cdots & P_{mn}/\bar{p}_m\bar{P}_n \end{pmatrix}$$

Fig. 4 Nagao et al.

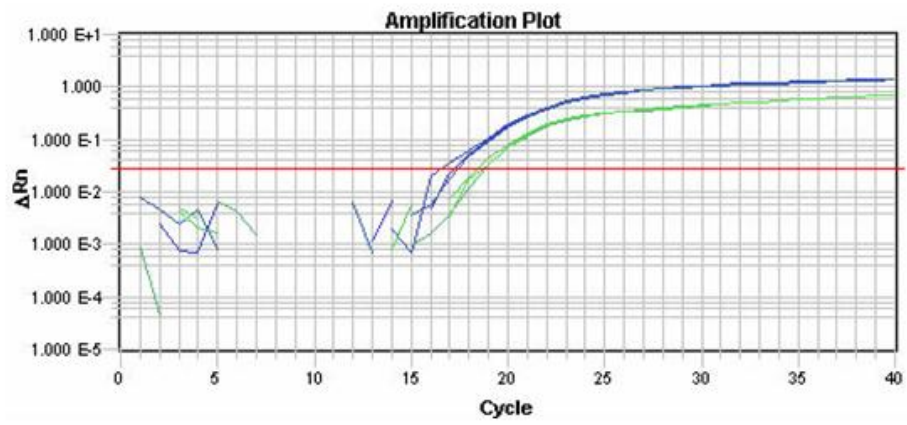


Figure S1

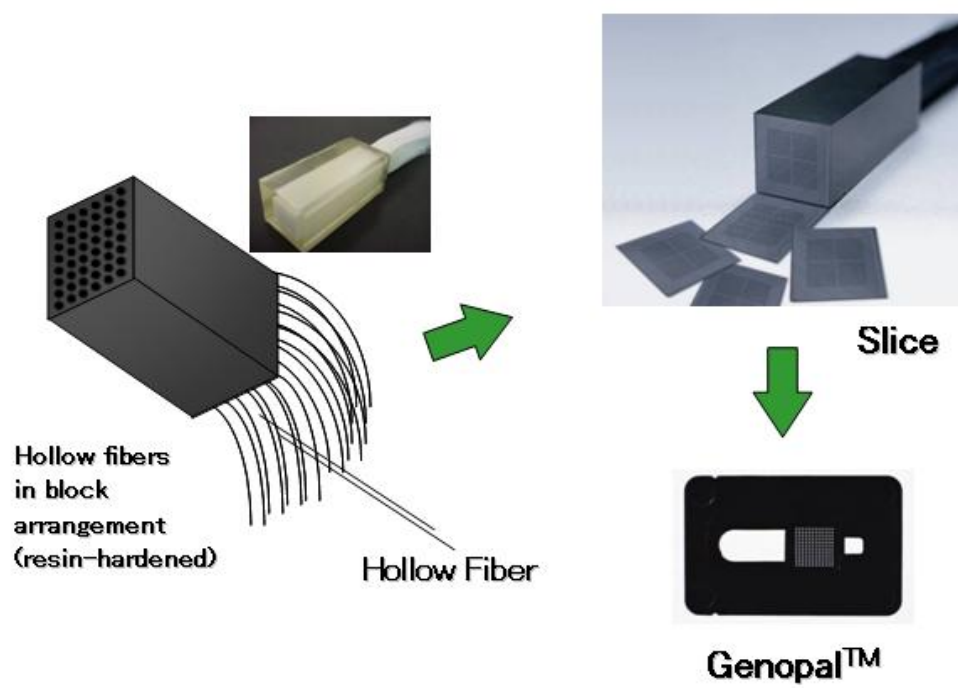


Figure S2

Supplemental Table 1 A list of NBCCS patients

Patient	PTCH mutation	Effect on Coding	Microarray Analysis	Comment
G15	c.459delT	Frameshift, truncation	RNA not available	
G17	c.584G>A	Splice variant	Abnormal (Fig. 3A)	
G1	c.912delC	Frameshift, truncation	ND	
G2	c.912delC	Frameshift, truncation	ND	G1's brother
G9	c.945+5G>T	Splice variant	Abnormal (Fig. 3A)	
G4	c.1259insT	Frameshift, truncation	ND	
G14	c.1412_1415dupTGGC	Frameshift, truncation	ND	
G8	c.1526G>A	p.G509D	Normal (Fig. 3A)	
G3	c.2011delC	Frameshift, truncation	ND	
G21	c.2724_2725insT	Frameshift, truncation	ND	
G11	c.3130_3131dupGC	Frameshift, truncation	RNA not available	
G12	c.3130_3131dupGC	Frameshift, truncation	RNA not available	G11's mother
G20	c.3364_3365delAT	Frameshift, truncation	Normal	
G6	Not detected		RNA not available	
G7	Not detected		Normal	
G13	Not detected		Normal	
G19	Not detected		RNA not available	

Nucleotide numbering is based on NM_000264.2: the A of the ATG of the initiator Met codon is counted as nucleotide +1

Amino Acid numbering is based on NP_000255.1

ND: not done

Supplemental Table 2 . Oligonucleotide probes used for the microarray construction

	probe name	Sequence
Junction	exon1b-exon2	GCTCTGGAGCAGATTTCCAAGGGGAAGGCTACTGGCCG
Junction	exon2-exon3	GAGGAGCTGTGGGTGGAAGTTGGAGGACGAGTAAGTCGTGAATT
Junction	exon3-exon4	CCGTGTCCATGTATACATGTACAACAGGCAGTGGAAATTGGAACATTTGTGTTACAA
Junction	exon4-exon5	AGCTTATCACAGAAACAGGTTACATGGATCAGATAATAGAATATCTTTACCCTTGTTTGATTATTACACCTTTGGA
Junction	exon5-exon6	AAATTACAGTCTGGGACAGCATACCTCCTAGGTAAACCTCCTTTGCGGTGGACAAA
Junction	exon6-exon7	CACAGCCCCAACAAAAATTCAACCAAACCTCTTGATATGGCCCTTGTTTTGAATG
Junction	exon7-exon8	ACAGCACTGGAAAACCTCGTCAGCGCCCATGCCCTGCAGA
Junction	exon8-exon9	CTGGCAGAGGACATATGTGGAGGTGGTTCATCAGAGTGTTCGCACA
Junction	exon9-exon10	TGGCCAGCGGCTACTTACTCATGCTCGCCTATGCCTGTCTAACCAT
Junction	exon10-11	GAATTTCCCTTTAACGCTGCAACAACCTCAGGTTTTGCCATTTCTCGCTCTTGGTGTTG
Junction	exon11-12	GAAACAGGACAGAATAAAAAGAATCCCTTTTGAGGACAGGACCGGGGAGTGC
Junction	exon12-13	CGGGCGTTCTCCCTCCAGGCAGCGGTAGTAGTGGTGTTC AATTTT
Junction	exon1a-exon2	GTTGACGGCCGGCTATGGGGAAGGCTACTGGCCG
Junction	exon1d-exon2	CCCTCTCCGATCACAAATATTGTTCGGGAAGGCTACTGGCCG
Junction	exon2-exon6	TGGAGGAGCTGTGGGTGGAAGAGGTAAACCTCCTTTGCGGTGGA
Junction	exon3-exon6	CCGTGTCCATGTATACATGTACAACAGAGGTAAACCTCCTTTGCGGTGGA
Junction	exon1a-exon3	GGGTTGACGGCCGGCTATGTTGGAGGACGAGTAAGTCGTGAATTA AATT
Junction	exon9-exon11	GGCCAGCGGCTACTTACTCATGGTTTTGCCATTTCTCGCTCTTGGTGTT
Junction	exon12-exon12b	GGGCGTTCTCCCTCCAGTACATACTGATGGCTCATCGAGG
Junction	exon3-exon4'	CCGTGTCCATGTATACATGTACAACAGTGAGAAATTTTTGTCTCTGCTTTTCATTTTATTAG
Junction	exon1e-exon2	CCTCCCTTCTCAGCTCTGGGGGAAGGCTACTGGCCG
Junction	exon12b-exon13	TCCAGTGCGCTAAATTTCTCCAAGGCAGCGGTAGTAGTGGTGTTC AATT
Intron	Intron2-3	GTCTCATGACTAGAGTTCATCTATTGTGTATCCTATGGCAGGTAGTCAGATAACAGATAAAA
Intron	Intron12b-13	TTTGAGGGGATGGGGCATTTCCTATGAGTACAATAACTAGAATTGATTCACAGTTGCAGTAT
Exon	exon1d	TGAGCCCTTCCCTTTGCACCTCGGCTGTTTTTTACGTTTAACCAGAAAGGAAGGGAGA
Exon	exon1b	GCGACGCCGCCTTCGCTCTGGAGCAGATTTCCAAG
Exon	exon1a	TTTCTTGAGTTTATTGTAAAGGGGTAAAGTTTTTCGGATCCGTCACGTGACCCTGACAGGT
Exon	exon2	AAGTTCCTTGGTTGTGGCCCTCCTCATATTTGGGGCCTTCGCGGTGGGATTA AAGCAGCGAA
Exon	exon3	AATCCTCAACTCATGATACAGACCCCTAAAGAAGAAGGTGCTAATGTCCTGACCACAGAAG
Exon	exon4	GCAGTGGAAATTGGAACATTTGTGTTACAAATCAGGAGAGCTTATCACAGAAACAGGTTACATGGATCAG
Exon	exon5	ATCTTTACCCTTGTTTGATTATTACACCTTTGGACTGCTTCTGGGAAGGGGCGAAATTACAGTCTG
Exon	exon6	AAATAAACTATCAAGTGGACAGCTGGGAGGAAATGCTGAATAAGGCTGAGGTTGGTCATGGTTACATGGA
Exon	exon7	TCTTGATATGGCCCTTGTTTTGAATGGTGGATGTCATGGCTTATCCAGAAAGTATATGCACTGGCA
Exon	exon8	ATGTTCCAGTTAATGACTCCCAAGCAAATGTACGAGCACTTCAAGGGGTACGAGTATGTCT
Exon	exon9	TGGTTCATCAGAGTGTTCGCACAGAACTCCACTCAAAGGTGCTTTTCTTCACCACCACGA

Exon	exon10	TGGCTGCAGGACTGGGCCTGTGCTCATTGATCGGAATTTCTTTAACGCTGCAACAACCTCA
Exon	exon11	TGGTGTGGATGATGTTTTTCTTCTGGCCCACGCCTTCAGTGAAACAGGACAGAATAAAAAGAAT
Exon	exon12	ATCAGCAATGTCACAGCCTTCTTCATGGCTGCGTTAATCCCAATTCCCCTCTGCGGGCGTTCT
Exon	exon12b	TTAATGATACCCTGTGGTGTGGTGGTCTGAAAAGCTACATGAGGTTTCCTTACGAGGA
Exon	exon13	CAATTCTCAGCATGGATTTATATCGACGCGAGGACAGGAGACTGGATATTTTCTGCTG
control	BC013835	CCAGGGCTTACCTGTACACTGACTTGAGACCAGTTGAATAAAAAGTGCACACCT
	Homo sapiens actin, beta, mRNA	
control	NM_002046	TGGCATTGCCCTCAACGACCACTTTGTCAAGCTCATTTCTGGTATGACAACGAATTT
	Homo sapiens glyceraldehyde-3-phosphate dehydrogenase (GAPD), mRNA	