MYB and MYBL1 in adenoid cystic carcinoma: diversity in the mode of genomic rearrangement and transcripts

Yuki Togashi^{1,2,3} · Akito Dobashi¹ · Seiji Sakata¹ · Yukiko Sato² · Satoko Baba¹ · Akira Seto⁴ · Hiroki Mitani⁴ · Kazuyoshi Kawabata⁴ · Kengo Takeuchi^{1,2,3}

Received: 26 September 2017 / Revised: 28 November 2017 / Accepted: 3 December 2017 / Published online: 6 February 2018
© United States & Canadian Academy of Pathology 2018

Abstract

MYB-NFIB and *MYBL1-NFIB* have been reported in ~60% of adenoid cystic carcinoma cases, but driver alterations in the remaining ~40% of adenoid cystic carcinoma remain unclear. We examined 100 adenoid cystic carcinoma cases for *MYB* and *MYBL1* locus rearrangements by fluorescence *in situ* hybridization (FISH) with originally designed probe sets using formalin-fixed paraffin-embedded materials. Approximately one-third of samples were also analyzed by fusion transcript-specific RT-PCR and capture RNA sequencing. In the 27 cases with frozen materials, *MYB-NFIB* and *MYBL1-NFIB* fusion transcripts were detected in 9 (33%) and 6 cases (22%) by RT-PCR, respectively. Meanwhile, high expression of *MYB* (18 cases, 67%) or *MYBL1* (9 cases, 33%) was detected in all 27 cases in a mutually exclusive manner, regardless of its form (full-length, truncation, or fusion transcript). Interestingly, genomic rearrangements around the corresponding highly-expressed gene were observed in all 27 cases by FISH, suggesting a causative relationship between genomic rearrangements and gene expression. Among the 100 cases, including additional 73 cases, 97 harbored genomic rearrangements in the *MYB* (73 cases) or *MYBL1* locus (24 cases) including 10 cases with atypical FISH patterns undetectable through ordinary split FISH approaches: breakpoints far distant from *MYB* (5 cases) and a small *NFIB* locus insertion into the *MYB* (3 cases) or *MYBL1* locus (2 cases). In clinicopathological analyses, histological grade, primary tumor size, and lymph node metastasis were identified as prognostic factors, whereas *MYB/MYBL1* rearrangements were not, but were associated with histological grade. In the present study, *MYB* or *MYBL1* locus rearrangement was detected in nearly all adenoid cystic carcinoma cases, and therefore it would be a good diagnostic marker for adenoid cystic carcinoma. However, fusion transcript-specific RT-PCR for *MYB-NFIB* and *MYBL1-NFIB* and ordinary split FISH assays for *MYB* and *MYBL1* were less sensitive, and thus detection methods should be judiciously designed because of the diversity of rearrangement modes in adenoid cystic carcinoma.

Electronic supplementary material The online version of this article (<https://doi.org/10.1038/s41379-018-0008-8>) contains supplementary material, which is available to authorized users.

✉ Kengo Takeuchi
kentakuchi-tky@umin.net

- ¹ Pathology Project for Molecular Targets, The Cancer Institute, Japanese Foundation for Cancer Research, 3-8-31 Ariake, Koto, Tokyo 135-8550, Japan
- ² Division of Pathology, The Cancer Institute, Japanese Foundation for Cancer Research, 3-8-31 Ariake, Koto, Tokyo 135-8550, Japan
- ³ Department of JFCR Cancer Biology, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo, Tokyo 113-8510, Japan
- ⁴ Department of Head and Neck Oncology, The Cancer Institute Hospital, Japanese Foundation for Cancer Research, 3-8-31 Ariake, Koto, Tokyo 135-8550, Japan

Introduction

Adenoid cystic carcinoma is a rare salivary gland-type tumor that arises in exocrine glands including the salivary, lacrimal, mammary, and bronchial glands [1]. The median age at diagnosis is 54.6 years and the female ratio is ~60% [1]. Although locoregional recurrence and distant metastasis are common, adenoid cystic carcinoma is an indolent cancer (5-year overall survival: 77%, 10-year: 60%, and 15-year: 45%) [1].

Histopathologically, the most characteristic architecture of adenoid cystic carcinoma is the cribriform pattern showing tumor nests with gland-like structures filled with a basophilic substance. The tubular pattern is comprised of glands containing inner epithelial and outer myoepithelial layers. In the solid pattern, tumor nests consist of

monotonous proliferation of epithelial or myoepithelial tumor cells with occasional gland and gland-like structure formation. Depending on the case, these patterns are observed in a single tumor in different ratios. Other salivary-gland type tumors including basal cell adenoma, basal cell adenocarcinoma, pleomorphic adenoma, polymorphous low-grade adenocarcinoma, and epithelial–myoepithelial carcinoma may mimic these patterns of adenoid cystic carcinoma. Therefore, definitively diagnosing adenoid cystic carcinoma may be difficult, particularly when using small biopsy specimens.

In several salivary gland-type tumors, histopathological subtype-specific fusion genes have been identified, including *PLAG1* and *HMGA2* fusions in pleomorphic adenoma [2, 3], *CRTC1-MAML2* in mucoepidermoid carcinoma [4, 5], *ETV6-NTRK3* in secretory carcinoma [6], and *EWSR1-ATF1* in clear cell carcinoma [7]. In adenoid cystic carcinoma, *MYB-NFIB* [8–11] and *MYBL1-NFIB* [12–14] were detected in ~60% of cases [15]. However, driver alterations in the remaining 40% of adenoid cystic carcinoma remain unclear. In the present study, we aimed to clarify the significance of MYB family genes in the pathogenesis of adenoid cystic carcinoma using histopathology, molecular biology, and bioinformatics for 100 adenoid cystic carcinoma cases.

Materials and methods

Case selection and histological review

Three pathologists histopathologically reviewed the slides of 123 patients with adenoid cystic carcinoma diagnosed and treated in the Cancer Institute Hospital (Tokyo, Japan) between January 2005 and October 2016. Twenty-three cases were excluded because a consensus diagnosis of adenoid cystic carcinoma was not obtained in 14 cases and archival samples were insufficient in 9 cases (Supplementary Table 1). This study was approved by the Institutional Review Board of Japanese Foundation for Cancer Research.

The three pathologists independently graded the remaining 100 cases based on the proportion of solid component; grade 1 was assigned when the solid component was <30% of the tumor area; grade 2, 30%≤ solid component <50%; and grade 3, 50%≤ solid component. When an area of histological transformation was identified, the suffix t was added to the grade. Each pathologist's grade was defined as the consensus grade when it was agreed upon by 2 or 3 pathologists. Graded cases were grouped into low-grade (grades 1 and 2) and high-grade cases (grades 3, 1t, 2t, and 3t) for statistical analyses.

Fluorescence *in situ* hybridization (FISH)

Unstained formalin-fixed paraffin-embedded sections (4-μm-thick) were subjected to hybridization with bacterial artificial chromosome clone-derived probes for *MYB*, *MYBL1*, and *NFIB*. The clones used are shown in Supplementary Table 2 and genomic regions covered by the probes are illustrated in Supplementary Figure 1. For *MYB*, we developed two probe sets. One was designated as the wide-range *MYB* probe set covering 11.5 Mb upstream to 8.4 Mb downstream of *MYB* (Supplementary Figure 1A). The other was a conventionally designed probe set covering 386.7 kb upstream to 136.6 kb downstream of *MYB*, the region similar to those covered by previously reported and commercially available *MYB* probe sets (Supplementary Figure 1B) [11, 16–20]. The conventionally designed probe set was used for initial screening. Cases negative in both *MYB* and *MYBL1* split FISH assays were again examined for *MYB* locus rearrangement using the wide-range probe set. Probe sets for *MYBL1* and *NFIB* were designed to cover 681.1 kb and 631.8 kb, respectively (Supplementary Figure 1C, D). Hybridized slides were stained with DAPI and examined using a BX51 fluorescence microscope (Olympus, Tokyo, Japan).

Immunohistochemistry

MYB protein was detected on unstained 4-μm-thick formalin-fixed paraffin-embedded sections after deparaffinization with xylene using a rabbit monoclonal antibody to c-Myb (EP769Y) (Abcam, Cambridge, UK) at a dilution of 1:200.

Total RNA extraction and cDNA synthesis

For fresh frozen specimens, total RNA was extracted with RNeasy Mini Kit (QIAGEN, Hilden, Germany) and used to synthesize cDNA with SuperScript III Reverse Transcriptase (Thermo Fisher Scientific, Waltham, MA, USA) and random primers. For formalin-fixed paraffin-embedded specimens, regions of interest (tumor and normal regions) on 10-μm-thick sliced histopathological sections were manually dissected using disposable needles under a stereomicroscope and processed with RecoverAll Total Nucleic Acid Isolation (Thermo Fisher Scientific) to extract total RNA.

Multiplex RT-PCR for *MYB-NFIB* and *MYBL1-NFIB* fusions

To amplify *MYB-NFIB* and *MYBL1-NFIB* fusion cDNAs, multiplex reverse-transcription polymerase chain reaction (RT-PCR) analysis was conducted using AmpliTaq Gold

Table 1 Patient characteristics

	Head and neck (including orbital region) (<i>N</i> = 81)			All (<i>N</i> = 100)		
	Total	MYB-type (<i>N</i> = 64)	MYBL1-type (<i>N</i> = 17)	Total	MYB-type (<i>N</i> = 76)	MYBL1-type (<i>N</i> = 24)
Female/male	54/27	43/21	11/6	71/29	55/21	16/8
Age, median (range)	54 (15–80)	53 (15–78)	59 (34–80)	57 (15–85)	54 (15–85)	60 (34–80)
Primary tumor size (mm), median (range) (<i>N</i> = 74)	31 (8–79)	32 (8–79)	30 (13–50)	NA	NA	NA
Primary distant metastasis	9/81	7/64	2/17	9/98	7/74	2/24
Primary lymph node metastasis	5/81	5/64	0/17	7/99	6/75	1/24
Stage (UICC 8th) (%) (<i>N</i> = 69)						
I	6 (9)	4 (7)	2 (13)	NA	NA	NA
II	12 (17)	8 (15)	4 (27)	NA	NA	NA
III	28 (41)	22 (41)	6 (40)	NA	NA	NA
IVA/B/C	12/3/8 (33)	11/3/6 (37)	1/0/2 (20)	NA	NA	NA
Histological grade						
Low (1/2)	63 (53/10)	47 (38/9)	16 (15/1)	78 (65/13)	57 (45/12)	21 (20/1)
High (3/1t/2t/3t)	18 (12/4/ 1/1)	17 (11/4/1/ 1)	1 (1/0/0/0)	22 (15/5/ 1/1)	19 (12/5/1/ 1)	3 (3/0/0/0)

NA not applicable

DNA polymerase (Thermo Fisher Scientific) and the primers shown in Supplementary Table 3. The template cDNAs were incubated at 94 °C for 10 min and then subjected to 35 cycles of denaturation at 94 °C for 1 min, annealing at 64 °C for 1 min, and polymerization at 72 °C for 1 min. The PCR products were separated by agarose gel electrophoresis, followed by direct sequencing or sub-cloning, if needed.

3'-Rapid amplification of cDNA ends for *MYB* and *MYBL1*

For cases negative in the multiplex RT-PCR for *MYB-NFIB* and *MYBL1-NFIB* fusions, 3'-rapid amplification of cDNA ends was performed with one of the two forward primers designed for *MYB* and *MYBL1* (Supplementary Table 3) under previously reported conditions [21]. Amplified products were subjected to direct sequencing.

Capture RNA sequencing

TruSeq RNA Access Library Prep Kit (Illumina, San Diego, CA, USA), which includes biotinylated capture probes targeting coding RNA, was used for coding RNA sequencing with an input of 20 ng total RNA from fresh frozen specimens. TruSight RNA Pan-Cancer Panel (Illumina) was

used for RNA sequencing targeted to 1385 cancer-related genes with an input of 100 ng total RNA from formalin-fixed paraffin-embedded specimens. All samples were prepared according to the manufacturer's instructions.

RT-PCR for *MPZL1-NFIB*, *MPDZ-MYB*, *TRAM1-MYBL1*, *MYBL1-RAD51B*, and *RAD51B-AHI1*

To confirm the fusion transcripts detected by coding RNA sequencing, RT-PCR was performed using Platinum Taq DNA Polymerase High Fidelity (Thermo Fisher Scientific) and the specific primer sets shown in Supplementary Table 3.

Sequencing analyses of rearrangements of *MYB* with downstream regions of *NFIB*

For the three cases in which fusion transcripts of *MYB* exon 14 to downstream regions of *NFIB* were detected by capture RNA sequencing (ACC005, ACC008, and ACC089), genomic PCR was performed with the primers shown in Supplementary Table 3 using LA Taq (TaKaRa Bio, Shiga, Japan). The PCR mixture was incubated at 94 °C for 1 min and then subjected to 30 cycles of denaturation at 98 °C for 10 s, annealing, and polymerization at 68 °C for 5 min. The PCR products were separated by agarose gel electrophoresis, followed by direct sequencing.

Table 2 MYB/MYBL1 transcripts and FISH results of the 27 cases available for frozen materials and cases with atypical and negative FISH results

Case no.	Primary site	Frozen materials	RT-PCR, capture RNA sequencing, and 3'-RACE		FISH				
			Transcripts identified		MYB split (conventional)	MYB split (wide range)	MYB-NFIB fusion	MYBL1 split	MYBL1-NFIB fusion
ACC121	Buccal mucosa	Available	MYB exon 9 - NFIB exon 12	Out of frame	+	+	+	-	ND
ACC036	Lung	Available	MYB exon 9 - NFIB exon 11; MYB exon 9 - NFIB exon 12	In-frame; Out of frame	+	+	+	-	ND
ACC111	Root of tongue	Available	MYB exon 9 - NFIB exon 11	In-frame	+	+	+	-	ND
ACC021	Submandibular gland	Available	MYB exon 12 - NFIB exon 12	Out of frame	+	+	+	-	ND
ACC010	Hard palate	Available	MYB exon 14 - NFIB exon 12	Out of frame	+	+	+	-	ND
ACC016	Maxillary sinus	Available	MYB exon 14 - NFIB exon 12	Out of frame	+	+	+	-	ND
ACC033	Vulva	Available	MYB exon 14 - NFIB exon 12	Out of frame	+	+	+	-	ND
ACC035	Bartholin gland	Available	MYB exon 14 - NFIB exon 12	Out of frame	+	+	+	-	ND
ACC063	Parotid gland	Available	MYB exon 14 - NFIB exon 11	In-frame	+	+	+	-	ND
ACC098	Parotid gland	Available	Full-length MYB; MPZL1 exon 4 - NFIB exon 4	NA; Out of frame	+	+	+	-	ND
ACC005	Tongue	Available	MYB exon 14 - 207 kb downstream of NFIB	NA	+	+	+	-	ND
ACC089	Parotid gland	Available	MYB exon 14 - 3.7 kb downstream of NFIB	NA	+	+	+	-	ND
ACC110	Submandibular gland	Available	Full-length MYB	NA	+	+	+	-	ND
ACC120	Floor of the oral cavity	Available	Full-length MYB	NA	+	+	+	-	ND
ACC122	Parotid gland	Available	Full-length MYB	NA	+	+	+	-	ND
ACC056	Parotid gland	Available	Full-length MYB	NA	-	+	+	-	ND
ACC024	External ear canal	Available	Full-length MYB	NA	-	+	-	-	-
ACC102	Nasolacrimal duct	Available	Full-length MYB	NA	-	+	-	-	ND
ACC093	Submandibular gland	Available	NE		-	+	-	-	-
ACC112	Lung	Available	NE		-	+	-	-	-
ACC101	Lacrimal gland	Available	Full-length MYB; NFIB exon 4 - LINC01235 exon 2	NA; NA	-	-	+	-	-
ACC067	Sublingual gland	Available	MYBL1 exon 8 - NFIB exon 12	Out of frame	-	-	ND	+	+
ACC019	Lung	Available	MYBL1 exon 8 - NFIB exon 11	In-frame	-	-	ND	+	+
ACC109	Submandibular gland	Available	MYBL1 exon 8 - NFIB exon 11	In-frame	-	-	ND	+	+
ACC004	Maxillary sinus	Available	MYBL1 exon 8 - NFIB exon 11; MYBL1 exon 9 - NFIB exon 11; MYBL1 exon 9 - NFIB exon 12	In-frame; Out of frame	-	-	ND	+	+

Table 2 (continued)

Case no.	Primary site	Frozen materials	RT-PCR, capture RNA sequencing, and 3'-RACE	FISH			
				Transcripts identified	MYB split (conventional)	MYB split (wide range)	MYB-NFIB fusion
ACC103	Parotid gland	Available	MYBL1 exon 14 - NFIB exon 4; MYBL1 exon 15 - NFIB exon 4; MYBL1 125th nucleotide of exon 16 - NFIB 26th nucleotide of exon 4	Out of frame; Out of frame; In-frame	—	—	ND
ACC051	Lung	Available	MYBL1 exon 9 truncation	NA	—	—	ND
ACC064	Oropharynx (root of tongue)	Available	TRAM1 exon 1 - MYBL1 exon 11; MYBL1 exon 10 - RAD51B intron 9	In-frame; NA	—	—	ND
ACC003	Submandibular gland	Available	MYBL1 exon 9 truncation	NA	—	—	ND
ACC008	Submandibular gland	Available	MYB exon 14 - 2.9 kb downstream of NFIB; MPDZ exon 24 - MYB exon 15	NA; Out of frame	—	—	+
ACC062	Sublingual gland	Available	MYB exon 8 - NFIB exon 12	Out of frame	—	—	+
ACC037	Trachea	Available	MYBL1 exon 13 - NFIB exon 9; MYBL1 exon 13 - NFIB exon 11; MYBL1 exon 14 - NFIB exon 9; MYBL1 exon 14 - NFIB exon 12; MYBL1 exon 15 - NFIB exon 12	In-frame; In-frame; In-frame; Out of frame; Out of frame	—	—	ND
ACC055	Lacrimal sac		NE	—	—	—	—
ACC020	External ear canal		Full-length MYB	NA	—	—	—
ACC114	External ear canal		Full-length MYB	NA	—	—	—
ACC119	External ear canal		Full-length MYB	NA	—	—	—

ND not done, NE not evaluable, NA not applicable

Fig. 1 Breakpoints in *MYB*-*NFIB* and *MYBL1*-*NFIB* fusion transcripts. Breakpoints in *MYB* and *NFIB* (**a**) and *MYBL1* and *NFIB* (**b**) found to be fused to each other are shown in the same color. Arrows and triangles represent the breakpoints reported and those detected in the present study, respectively. Cases with asterisks showed alternative

splicing. The primer locations for multiplex RT-PCR are also shown. Note that all fusion-positive cases retained *MYB* or *MYBL1* DNA binding domain (DBD) and transactivation domain (TAD). NRD negative regulatory domain

Sequence data analysis

Quality control was performed on fastq files using FASTX-Toolkit version 0.0.13 (http://hannonlab.cshl.edu/fastx_toolkit/index.html), PRINSEQ version 0.20.4 [22], and cmpfastq_pe (http://compbio.brc.iop.kcl.ac.uk/software/cmpfastq_pe.php) software. Sequenced reads containing <80% of bases with a quality score of Q20 or higher were deleted. Bases at read ends with N sequences or whose quality scores were less than Q20 were removed. Among the remaining reads, those shorter than 20 base pairs (bp) and the resulting non-paired reads were deleted. The qualified reads were mapped to hg19 with STAR version 2.4.2a [23]. Fusion gene analyses were performed with Fusion-Catcher [24] and in-house programs. Expression analyses for *MYB* and *MYBL1* were performed using R version 3.3.2 [25] and HTSeq version 0.6.1 [26].

Statistics analysis

Fisher's exact test (two-tailed) was used to analyze the differences between *MYB* and *MYBL1* locus rearrangement-positive head and neck adenoid cystic carcinoma ($N=78$). For histological grade, Mann-Whitney U test was also used.

To assess the statistical significance of variables affecting overall survival, the log-rank test was employed for univariate analysis. Multivariate analysis was performed on prognostic factors with significant difference in univariate analysis (P value < 0.05), using the Cox proportional hazard model. Univariate analysis was carried out with GraphPad Prism® Version 6.05 (GraphPad, Inc., La Jolla, CA, USA) and multivariate analysis was performed using R version 3.4.1 and survival version 2.38.

Results

Histopathological grading

In the 100 cases with adenoid cystic carcinoma, there were 71 female and 29 male patients. The median age of onset was 56.8 years (range: 14.9–84.7). The sites of origin included major salivary glands ($N=40$), minor salivary glands ($N=37$), and other sites ($N=23$). The number of cases of head and neck (including four cases of orbital region) were 81 (Table 1; Supplementary Table 4). The 100 cases were histopathologically classified as follows: 65 cases as grade 1, 13 as grade 2, 15 as grade 3, five as grade

Fig. 2 Estimated genomic rearrangements in *MYB* split FISH-negative but *MYB-NFIB* fusion FISH-positive cases. **a** Transcripts of *MYB* exon 14-*NFIB* 3' downstream and *MPDZ* exon 24-*MYB* exon 15 were detected in ACC008, suggesting an ~0.9 Mb insertion of the *NFIB* locus into *MYB* intron 14. **b** In ACC101, a transcript of *NFIB* exon 4-

LINC01235 exon 2 was identified. Therefore, a part of the *NFIB* locus, which consisted of the *NFIB* sequence from intron 4, intergenic region (gray) between *NFIB* and *LINC01235*, and *LINC01235* intron 1, was likely to juxtapose the *MYB* locus. Not to scale

1t, one as grade 2t, and one as grade 3t, resulting in 78 low- and 22 high-grade cases. In 81 head-and-neck cases, 63 and 18 cases were graded as low and high, respectively (Table 1).

Expression of *MYB* and *MYBL1* transcripts in the 27 cases with available frozen materials

We developed a multiplex RT-PCR system that could theoretically detect all possible fusion variants coding for the *MYB-NFIB* or *MYBL1-NFIB* fusion protein. Among the 27 cases for which frozen materials were available, 9 (33%) and 6 (22%) cases were found to express *MYB-NFIB* and *MYBL1-NFIB* fusion transcripts, respectively. In the remaining 12 (44%) cases, neither transcript was expressed (Table 2, Fig. 1, and Supplementary Figure 2).

Next, we performed a conventional 3'-rapid amplification of cDNA ends assay and coding RNA sequencing using next-generation sequencing. Interestingly, all 27 cases were found to highly express either *MYB* (18 cases, 67%) or *MYBL1* (9 cases, 33%) in a mutually exclusive manner, and three additional *MYB* (ACC005, ACC008, and ACC089) and one additional *MYBL1* (ACC064) fusion transcripts were identified in the 12 multiplex RT-PCR-negative cases (Table 2). In the three additionally identified *MYB* transcripts, *MYB* exon 14 fused to 9p23 intergenic sequences, 207 kb (ACC005), 2.9 kb (ACC008), and 3.8 kb (ACC089) downstream of the 3' untranslated region (UTR) end of *NFIB* (Supplementary Figure 3). The genomic breakpoints in *MYB* intron 14 were at 4.4 kb (ACC005), 0.7 kb (ACC008), and 4.2 kb (ACC089) downstream from the end of *MYB* exon 14 (Supplementary Figure 3). In ACC008, an

Fig. 3 Clinicopathological and genetic landscape of 100 cases with adenoid cystic carcinoma. At the top of the figure, *MYB* or *MYBL1* expression level is shown using fragments per kilobase of exon per million mapped fragments (FPKM) (pink bar: *MYB*, green bar: *MYBL1*)

~0.9 Mb fragment of chromosome 9 was considered inserted in *MYB* intron 14 because the following two fusion transcripts were identified: *MYB* exon 14 fused to 3.6 kb downstream of the *NFIB* 3'UTR and *MPDZ* exon 24 fused to *MYB* exon 15 (Fig. 2a; Supplementary Figure 4A). In ACC064, *MYBL1* exon 10-*RAD51B* intron 9 and *TRAM1* exon 1-*MYBL1* exon 11 fusion transcripts were expressed and confirmed by specific RT-PCR (Supplementary Figure 4B). *MYBL1* and *RAD51B* were fused to each other in the opposite direction in this case (Supplementary Figure 5). The genomic breakpoint of *MYBL1* was in intron 10, 87 bp downstream of exon 10, which was fused to intron 9 of *RAD51B*, 49.5 kb downstream of exon 9 (Supplementary Figure 5). As *MYBL1*, *TRAM1*, and *RAD51B* are located at 8q13.1, 8q13.3, and 14q24.1, the chromosomal rearrangement involving *MYBL1* in ACC064 was considered complex.

In the remaining eight cases, full-length *MYB* (ACC024, ACC098, ACC102, ACC110, ACC120, and ACC122) and truncated *MYBL1* (ACC003 and ACC051) were highly

expressed. In a case with the full-length *MYB* transcript (ACC098), an *MPZL1* (1q24.2) exon 4-*NFIB* exon 4 fusion transcript was detected (Table 2; Supplementary Figure 6). In confirmatory RT-PCR, a transcript with an insertion of 69bases from 6q25.3 between *MPZL1* exon 4 and *NFIB* exon 4 was also identified (Supplementary Figure 4C; Supplementary Figure 6).

Taken together, the following results were obtained in the analyses of *MYB* and *MYBL1* transcripts in the 27 cases. Either *MYB* or *MYBL1* transcripts were expressed in all cases in a mutually exclusive manner, regardless of its form (full-length, truncation, or fusion). In the fusion forms, both in-frame and out-of-frame fusions were identified. In all *MYB* and *MYBL1* transcripts, the DNA-binding domain and trans-activation domain were retained (Table 2; Fig. 1).

FISH for *MYB* and *MYBL1* loci in the 27 cases

We subsequently performed FISH assays using formalin-fixed paraffin-embedded specimens to determine whether

Fig. 4 Overall survival curves in the 78 *MYB/MYBL1* locus rearrangement-positive cases whose primary sites were the head and neck (including orbital region). Cases with clinicopathological factors:

a >30 mm primary tumor size ($N=38$) ($P=0.0028$), **b** primary lymph node metastasis ($N=5$) ($P<0.0001$), and **c** high histological grade ($N=18$) ($P=0.0022$)

genomic rearrangements occurred around *MYB* and *MYBL1* in these 27 cases. In adenoid cystic carcinoma, recent studies reported the fusion transcripts of *NFIB* and genes around *MYB* (6q23.3), such as *NKAIN2* (6q22.31) and *AIG1* (6q24.2) [13]. These data prompted us to design probes that covered the ~20 Mb region around *MYB* (Supplementary Figure 1A), which was much broader than those covered by previously reported and commercially available FISH probes (Supplementary Figure 1B) [11, 16–20]. Among the 27 cases, 17 and 8 were positive in *MYB* and *MYBL1* split FISH assays, respectively (Table 2; Supplementary Figure 7). One of the remaining two cases (ACC008) was negative for *MYB* split signals, but was positive in the *MYB-NFIB* fusion FISH assay. The other case (ACC037) was negative for *MYBL1* split signals, but was positive in the *MYBL1-NFIB* fusion FISH assay.

Collectively, genomic rearrangements around the corresponding highly-expressed gene were observed in all 27 examined adenoid cystic carcinoma cases (18 cases for *MYB* and 9 cases for *MYBL1*), and the *NFIB* locus (9p23-22.3) was involved in 16 of 18 *MYB* and 7 of 9 *MYBL1* locus rearrangements (Table 2). In addition, it was shown that *MYB* and *MYBL1* were the target genes for over-expression in *MYB* and *MYBL1* locus rearrangements, but not the other adjacent genes around *MYB* (*ALDH8A1*, *HBS1L*, *AHI1*, and *LINC00271*) and *MYBL1* (*TRIM55*,

ADHFE1, *C8orf46*, and *C8orf44-SGK3*), respectively (Supplementary Table 5; Supplementary Figure 8).

Validation of *MYB* and *MYBL1* rearrangements using 73 additional cases

On the basis of the fact that all of the examined 27 adenoid cystic carcinoma cases harbored genomic rearrangements in the *MYB* or *MYBL1* locus, we further performed FISH analyses using formalin-fixed paraffin-embedded specimens of 73 additional adenoid cystic carcinoma cases. *MYB* split signals and *MYBL1* split signals were identified in 52 and 15 cases, respectively. In ACC062 and ACC101, *MYB-NFIB* fusion signals were positive, although *MYB* split signals were negative as in ACC008. In ACC055, *MYBL1-NFIB* fusion signals were positive, although *MYBL1* split signals were negative as in ACC037 (Fig. 3).

Among the 100 cases examined, 97 showed either of the following signal patterns: both *MYB* split and *MYB-NFIB* fusion signals in 57, only *MYB* split signals in 12, only *MYB-NFIB* fusion signals in 3, both *MYBL1* split and *MYBL1-NFIB* fusion signals in 18, only *MYBL1* split signals in 5, and only *MYBL1-NFIB* fusion signals in 2. The remaining three cases (ACC020, ACC114, and ACC119) were negative for any of these split/fusion FISH signals (Fig. 3).

Further analyses of genomic rearrangements in cases with atypical FISH patterns

Thirteen cases with atypical FISH patterns were further analyzed by capture RNA sequencing (Table 2). In the five cases negative for *MYB* split with the conventional probe set but positive with the wide-range set (ACC024, ACC056, ACC093, ACC102, and ACC112), full-length *MYB* expression was detected except for ACC093 and ACC112, which were not evaluable due to the low quality of RNA.

MYB split was negative but *MYB-NFIB* fusion was positive in three cases (ACC008, ACC062, and ACC101). In ACC008, the discrepant FISH results support our above-mentioned estimation that a ~900 kb fragment of chromosome 9p23-p22.3 (3.6 kb downstream from *NFIB* 3'UTR to *MPDZ* exon 24) was probably inserted in *MYB* intron 14 (Fig. 2a), as the insertion size is far below the resolution limit of the split FISH assay on formalin-fixed paraffin-embedded slices. In ACC062, an *MYB* exon 8-*NFIB* exon 12 fusion transcript was identified, indicating that a small fragment containing *NFIB* exon 12 and after was inserted in *MYB* intron 8 through a similar mechanism as in ACC008. In ACC101, full-length *MYB* was highly expressed and an *NFIB* exon 4-*LINC01235* exon 2 fusion transcript was identified. *LINC01235* is located 650 kb downstream from *NFIB*, and therefore the region from *NFIB* intron 4 to *LINC01235* intron 1 may be juxtaposed to *MYB* (Fig. 2b).

MYBL1 split was negative while *MYBL1-NFIB* fusion was positive in two cases (ACC037 and ACC055). A *MYBL1* exon 14-*NFIB* exon 9 fusion transcript was expressed in ACC037. Therefore, as in the three cases (ACC008, ACC062, and ACC101), a small fragment containing *NFIB* exon 9 and after was likely inserted into *MYBL1* intron 14. Capture RNA sequencing was not successful in ACC055 for unknown reasons.

In the three cases negative in all FISH analyses (ACC020, ACC114, and ACC119), the expression of full-length *MYB* mRNA and MYB protein was detected (Supplementary Figure 9). Any causative genomic rearrangements for the expression of *MYB* were not observed in the present FISH assays or capture RNA sequencing. Interestingly, all three FISH-negative cases were of external ear canal origin, and the remaining external ear canal case was also atypical in FISH (ACC024), which was negative for *MYB* split with the conventional probe set but positive with the wide-range set.

Secondary *MYBL1* locus rearrangement as a passenger event

In the 97 cases showing rearrangements, those in the *MYB* and *MYBL1* loci were fundamentally mutually exclusive (Fig. 3), except for one case (ACC018). The sample from

the primary lesion (lacrimal gland) was positive in *MYB* split FISH and negative in *MYB-NFIB* fusion FISH and *MYBL1* split FISH. However, *MYBL1* split signals were also detected in samples from relapsed lesions. We examined all four lesions (primary and relapsed at 2, 4, and 6 years after) by capture RNA sequencing with TruSight RNA Pan-Cancer Panel using RNA extracted from formalin-fixed paraffin-embedded tissues. Full-length *MYB* mRNA expression was constantly detected, while *MYBL1* expression remained very low (Supplementary Figure 10). Therefore, the *MYBL1* locus rearrangement in this *MYB*-rearranged case was likely a secondary and passenger event. A fusion transcript of *RAD51B* exon 7 (14q24.1) fused to *AH11* exon 23 (114 kb downstream from *MYB*) was detected, which was confirmed by RT-PCR (Supplementary Figure 4D), supporting the FISH results (positive for *MYB* split and negative for *MYB-NFIB* fusion).

Clinicopathological analysis of 78 head and neck cases with confirmed *MYB*/*MYBL1* locus rearrangement

Prognostic analyses were performed on 78 head and neck cases (including four cases in the ocular regions) with confirmed *MYB* or *MYBL1* rearrangement. Three variables, primary tumor size (30 mm<, $P = 0.0028$), primary lymph node metastasis (positive, $P < 0.0001$), and histological grade (high, $P = 0.0022$) were significantly associated with poorer overall survival in univariate analysis (Supplementary Table 6; Fig. 4). In multivariate analysis of these three variables, primary tumor size ($P = 0.003$), primary lymph node metastasis ($P = 0.002$), and histological grade ($P < 0.001$) were independent prognostic factors.

Any clinical features tested including overall survival did not significantly differ between the cases with *MYB* and *MYBL1* alterations in the present cohort. However, histological grade was generally higher in cases with *MYB* alterations ($P = 0.10$, Fisher's exact test; $P = 0.023$, Mann-Whitney U test) (Supplementary Table 7). Given that histological grade was found to be an independent prognostic factor, a large-scale study with detailed examination of *MYB*/*MYBL1* alterations would reveal the prognostic significance of the alteration.

Discussion

In the present study, we found that 97 of the 100 adenoid cystic carcinoma cases were associated with genomic rearrangement either in the *MYB* locus (73 cases) or in the *MYBL1* locus (24 cases). In addition, all 27 cases examined expressed *MYB* (18/27, 67%) or *MYBL1* (9/27, 33%) transcripts in a mutually exclusive manner. There were various

forms of the transcripts including full-length, truncated, and fusion forms. In fusion forms, both in-frame and out-of-frame fusion transcripts were detected. In all transcripts, the DNA-binding domain and trans-activation domain were retained. Taken together, these observations suggest that the fundamental alteration in the development of adenoid cystic carcinoma is dysregulated overexpression of the two domains in *MYB* or *MYBL1* led by translocation of super-enhancers (mostly those of *NFIB*) in the proximity of *MYB* or *MYBL1*.

Given its high frequency, *MYB/MYBL1* locus rearrangement can be a sensitive diagnostic marker of adenoid cystic carcinoma. However, the diversity of the rearrangements was also revealed and thus the method of detection should be designed judiciously. Although the frequency of *MYB-NFIB* fusion transcript in adenoid cystic carcinoma in the initial report was 100% (11/11) [8], the frequency obtained by RT-PCR was relatively low considering subsequent reports (51.1%, 162/317) [8–10, 12, 16, 18, 19, 27–31]. Aside from technical issues (primer design, formalin-fixed paraffin-embedded vs. frozen materials, etc.), this was explained by the presence of other types of *MYB* expression including full-length *MYB*, truncated *MYB*, and *MYB* fusion transcript with other partner genes (*PDCD1LG2* at 9p24.1, *EFR3A* at 8q24.22, *TGFBR3* at 1p22.1, and *RAD51B* at 14q24.1) [28, 32]. In addition, the *MYBL1-NFIB* fusion was reported in 2015. The reported frequency of the fusion varies depending on methodology used (10.9%, 16/147, transcript-based; 11.9%, 16/135, *MYBL1-NFIB* fusion FISH), and other types of *MYBL1* expression including truncated *MYBL1* and *MYBL1* fusion with other partner genes have been reported [12–14, 33]. Although *MYB/MYBL1* alterations are highly characteristic to adenoid cystic carcinoma, this diversity in *MYB/MYBL1* transcripts dampens the efficacy of RT-PCR-based approaches for confirmatory diagnosis of adenoid cystic carcinoma.

In the present study, we detected 10 cases (10%) whose *MYB/MYBL1* locus rearrangements could not be detected using the conventionally designed split FISH probe set (Table 2). In five cases, genomic rearrangements in the *MYB* locus turned out to be visible with the wide-range probe set. The probe set covered the ~20 Mb region around *MYB* (Supplementary Figure 1A), which was extremely broad compared with those covered by previously reported and commercially available FISH probes (Supplementary Figure 1B) [11, 16–20]. Thus, the breakpoints in the *MYB* locus are far distant from *MYB* (~10 Mb) in these cases. In the other five cases, a small insertion (<1 Mb) that could not be recognized by split FISH assay was considered to occur within *MYB* (ACC008 and ACC062), close to *MYB* (ACC101), within *MYBL1* (ACC037), and within or close to *MYBL1* (ACC055). The super-enhancer elements in

NFIB, *TGFBR3*, or *RAD51B* loci translocated to the proximity of *MYB* (regardless of whether it was upstream or downstream) caused overexpression of *MYB* protein, which binds the elements and sustains *MYB* expression in a positive feedback manner [32]. On the basis of previous results and those of the present study, *MYB* can be upregulated even if a genomic region containing the super-enhancer elements is translocated “far distant” from *MYB* (but the elements should physically interact with *MYB* promoter region) and/or if the genomic region is too small to be observed by conventional karyotyping and FISH assays.

FISH assay is a powerful tool for detecting genomic rearrangements. Particularly, split FISH assay is suitable for screening for rearrangement of genes of interest, as it is available regardless of the rearranged partners. For example, the split FISH assay for *ALK* has been widely used as a companion diagnostic for *ALK* inhibitors in *ALK*-positive lung cancers, in which *ALK* is rearranged with various potential fusion partner genes including *EML4* [34], *KIF5B* [35], *KLC1* [21], and others. However, the presence of above-mentioned atypical patterns of *MYB/MYBL1* locus rearrangement in adenoid cystic carcinoma may limit FISH as a confirmatory diagnostic assay. Using conventionally designed split FISH probe sets, the frequency of *MYB/MYBL1* locus rearrangement in our cohort was 87% (87/100), which is comparable to that in a study in which both *MYB* and *MYBL1* were examined by FISH (88%, 29/33) [33]. Therefore, to maintain the sensitivity of FISH for detecting *MYB/MYBL1* locus rearrangement in adenoid cystic carcinoma, combined analysis using the split FISH assay with a wide-range probe set and fusion FISH assay using probes for known fusion partners (at least *NFIB*) is necessary. For example, one of the practical ways is following; 3-color FISH using a mixture of a conventionally designed *MYB* probe set (e.g., set A in Supplementary Table 2) and *NFIB* (e.g., set C) is performed in a case. If the result is negative, a mixture of a *MYBL1* probe set (e.g., set B) and *NFIB* (e.g., set C) is performed. If the result is negative, FISH using wide-range *MYB* probe set is performed. The first two or all three FISH tests can be done at the same time.

Among the 14 excluded cases without a consensus diagnosis of adenoid cystic carcinoma, 11 were evaluable and negative for *MYB/MYBL1* rearrangement (data not shown). Capture RNA sequencing was performed in 4 cases. *MYBL1* expression was not detected in the 4 cases, but the expression of full-length *MYB* transcript was detected in 3 cases. This means that the expression-based assays that cannot distinguish *MYB* forms (full length, fusion, and truncation), such as anti-*MYB* immunohistochemistry, are less useful in differentiating adenoid cystic carcinoma from its mimics in diagnostic settings.

Acknowledgements We thank Mr. Motoyoshi Iwakoshi, Ms. Keiko Shiozawa, and Ms. Tomoyo Kakita for their technical assistance and Ms. Sayuri Sengoku for her management assistance in the research project. We also thank Dr. Hiroyuki Yonekawa, Dr. Hirofumi Fukushima, Dr. Toru Sasaki, Dr. Wataru Shinbashi, Dr. Kazutoshi Toda, Dr. Aya Ebina, and Dr. Hideki Tsuji for providing clinical specimens and Dr. Futoshi Akiyama and Dr. Yutaka Takazawa for giving advice on histological review. Computations were partially performed on the NIG supercomputer at ROIS National Institute of Genetics. This study was supported in part by the Japan Society for the Promotion of Science KAKENHI grant number 16K21591 (to Yuki Togashi).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Lloyd S, Yu JB, Wilson LD, et al. Determinants and patterns of survival in adenoid cystic carcinoma of the head and neck, including an analysis of adjuvant radiation therapy. *Am J Clin Oncol*. 2011;34:76–81.
- Kas K, Voz ML, Roijer E, et al. Promoter swapping between the genes for a novel zinc finger protein and beta-catenin in pleomorphic adenomas with t(3;8)(p21; q12) translocations. *Nat Genet*. 1997;15:170–4.
- Geurts JM, Schoenmakers EF, Roijer E, et al. Expression of reciprocal hybrid transcripts of HMGIC and FHIT in a pleomorphic adenoma of the parotid gland. *Cancer Res*. 1997;57:13–7.
- Nordkvist A, Gustafsson H, Juberg-Ode M, et al. Recurrent rearrangements of 11q14-22 in mucoepidermoid carcinoma. *Cancer Genet Cytogenet*. 1994;74:77–83.
- Winnes M, Molne L, Suurkula M, et al. Frequent fusion of the CRTC1 and MAML2 genes in clear cell variants of cutaneous hidradenomas. *Genes Chromosomes Cancer*. 2007;46:559–63.
- Skalova A, Vanecek T, Sima R, et al. Mammary analogue secretory carcinoma of salivary glands, containing the ETV6-NTRK3 fusion gene: a hitherto undescribed salivary gland tumor entity. *Am J Surg Pathol*. 2010;34:599–608.
- Antonescu CR, Katabi N, Zhang L, et al. EWSR1-ATF1 fusion is a novel and consistent finding in hyalinizing clear-cell carcinoma of salivary gland. *Genes Chromosomes Cancer*. 2011;50:559–70.
- Persson M, Andren Y, Mark J, et al. Recurrent fusion of MYB and NFIB transcription factor genes in carcinomas of the breast and head and neck. *Proc Natl Acad Sci USA*. 2009;106:18740–4.
- Mitani Y, Li J, Rao PH, et al. Comprehensive analysis of the MYB-NFIB gene fusion in salivary adenoid cystic carcinoma: Incidence, variability, and clinicopathologic significance. *Clin Cancer Res*. 2010;16:4722–31.
- Brill LB 2nd, Kanner WA, Fehr A, et al. Analysis of MYB expression and MYB-NFIB gene fusions in adenoid cystic carcinoma and other salivary neoplasms. *Mod Pathol*. 2011;24:1169–76.
- West RB, Kong C, Clarke N, et al. MYB expression and translocation in adenoid cystic carcinomas and other salivary gland tumors with clinicopathologic correlation. *Am J Surg Pathol*. 2011;35:92–9.
- Brayer KJ, Frerich CA, Kang H, et al. Recurrent fusions in MYB and MYBL1 define a common, transcription factor-driven oncogenic pathway in salivary gland adenoid cystic carcinoma. *Cancer Discov*. 2016;6:176–87.
- Mitani Y, Liu B, Rao PH, et al. Novel MYBL1 gene rearrangements with recurrent MYBL1-NFIB Fusions in salivary adenoid cystic carcinomas lacking t(6;9) translocations. *Clin Cancer Res*. 2016;22:725–33.
- Rettig EM, Talbot CC Jr., Sausen M, et al. Whole-genome sequencing of salivary gland adenoid cystic carcinoma. *Cancer Prev Res*. 2016;9:265–74.
- Wysocki PT, Izumchenko E, Meir J, et al. Adenoid cystic carcinoma: emerging role of translocations and gene fusions. *Oncotarget*. 2016;7:66239–54.
- von Holstein SL, Fehr A, Persson M, et al. Adenoid cystic carcinoma of the lacrimal gland: MYB gene activation, genomic imbalances, and clinical characteristics. *Ophthalmology*. 2013;120:2130–8.
- Zhang J, Wu G, Miller CP, et al. Whole-genome sequencing identifies genetic alterations in pediatric low-grade gliomas. *Nat Genet*. 2013;45:602–12.
- Martelotto LG, De Filippo MR, Ng CK, et al. Genomic landscape of adenoid cystic carcinoma of the breast. *J Pathol*. 2015;237:179–89.
- North JP, McCalmont TH, Fehr A, et al. Detection of MYB alterations and other immunohistochemical markers in primary cutaneous adenoid cystic carcinoma. *Am J Surg Pathol*. 2015;39:1347–56.
- Rettig EM, Tan M, Ling S, et al. MYB rearrangement and clinicopathologic characteristics in head and neck adenoid cystic carcinoma. *Laryngoscope*. 2015;125:E292–9.
- Togashi Y, Soda M, Sakata S, et al. KLC1-ALK: a novel fusion in lung cancer identified using a formalin-fixed paraffin-embedded tissue only. *PLoS ONE*. 2012;7:e31323.
- Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics*. 2011;27:863–4.
- Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29:15–21.
- Daniel Nicorici MS, Henrik Edgren, Sara Kangaspeska, et al. FusionCatcher—a tool for finding somatic fusion genes in paired-end RNA-sequencing data. *bioRxiv*. 2014.
- RC Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2016).
- Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31:166–9.
- Fehr A, Kovacs A, Loning T, et al. The MYB-NFIB gene fusion—a novel genetic link between adenoid cystic carcinoma and dermal cylindroma. *J Pathol*. 2011;224:322–7.
- Mitani Y, Rao PH, Futreal PA, et al. Novel chromosomal rearrangements and break points at the t(6;9) in salivary adenoid cystic carcinoma: association with MYB-NFIB chimeric fusion, MYB expression, and clinical outcome. *Clin Cancer Res*. 2011;17:7003–14.
- Persson M, Andren Y, Moskaluk CA, et al. Clinically significant copy number alterations and complex rearrangements of MYB and NFIB in head and neck adenoid cystic carcinoma. *Genes Chromosomes Cancer*. 2012;51:805–17.
- Wetterskog D, Lopez-Garcia MA, Lambros MB, et al. Adenoid cystic carcinomas constitute a genomically distinct subgroup of triple-negative and basal-like breast cancers. *J Pathol*. 2012;226:84–96.
- D'Alfonso TM, Mosquera JM, MacDonald TY, et al. MYB-NFIB gene fusion in adenoid cystic carcinoma of the breast with special

- focus paid to the solid variant with basaloid features. *Hum Pathol.* 2014;45:2270–80.
32. Drier Y, Cotton MJ, Williamson KE, et al. An oncogenic MYB feedback loop drives alternate cell fates in adenoid cystic carcinoma. *Nat Genet.* 2016;48:265–72.
 33. Fujii K, Murase T, Beppu S, et al. MYB, MYBL1, MYBL2 and NFIB gene alterations and MYC overexpression in salivary gland adenoid cystic carcinoma. *Histopathology.* 2017;71:823–34.
 34. Soda M, Choi YL, Enomoto M, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature.* 2007;448:561–6.
 35. Takeuchi K, Choi YL, Togashi Y, et al. KIF5B-ALK, a novel fusion oncokinase identified by an immunohistochemistry-based diagnostic system for ALK-positive lung cancer. *Clin Cancer Res.* 2009;15:3143–9.