ORIGINAL PAPER

Identification of a recurrent transforming *UBR5–ZNF423* fusion gene in EBV-associated nasopharyngeal carcinoma

Grace TY Chung,[1,2#] Raymond WM Lung,[1,2#] Angela BY Hui,[3] Kevin YL Yip,[4] John KS Woo,[5] Chit Chow,[1] Carol YK Tong,[1] Sau-Dan Lee,[4] Jessie WF Yuen,[1] Samantha WM Lun,[1] Ken KY Tso,[4] Nathalie Wong,[1,2] Sai-Wah Tsao,[6] Timothy TC Yip,[7] Pierre Busson,[8] Hyungtae Kim,[9] Jeong-Sun Seo,[9,10] Brian O'Sullivan,[3,11] Fei-Fei Liu,[3,11] Ka-Fai To[1,2]* and Kwok-Wai Lo[1,2]*

[1] Department of Anatomical and Cellular Pathology, State Key Laboratory in Oncology in South China, Prince of Wales Hospital, The Chinese University of Hong Kong, SAR
[2] Li Ka Shing Institute of Health Science, The Chinese University of Hong Kong, SAR
[3] Ontario Cancer Institute, University Health Network, Toronto, ON, Canada
[4] Department of Computer Science and Engineering, The Chinese University of Hong Kong, SAR
[5] Department of Otorhinolaryngology, Head and Neck Surgery, The Chinese University of Hong Kong, SAR
[6] Department of Anatomy, University of Hong Kong, SAR
[7] Department of Clinical Oncology, Queen Elizabeth Hospital, Hong Kong SAR
[8] Laboratoire de Biologie des Tumeurs Humaines, Institut Gustave Roussy, Villejuif, France
[9] Macrogen Inc., Gasan-dong, Geumcheon-gu, Seoul, Republic of Korea
[10] Ilchun Genomic Medicine Institute, MRC, and Department of Biochemistry and Molecular Biology, Seoul National University College of Medicine, Republic of Korea
[11] Department of Radiation Oncology, Princess Margaret Cancer Center, ON, Canada

*Correspondence to: Kwok-Wai Lo, Department of Anatomical and Cellular Pathology, Prince of Wales Hospital, The Chinese University of Hong Kong, Shatin, NT, Hong Kong SAR. e-mail: kwlo@cuhk.edu.hk
Or Ka-Fai To, Department of Anatomical and Cellular Pathology, Prince of Wales Hospital, The Chinese University of Hong Kong, Shatin, NT, Hong Kong SAR. e-mail: kfto@cuhk.hk

#These authors contributed equally to this study.

Abstract

Nasopharyngeal carcinoma (NPC) is a distinct type of head and neck cancer which is prevalent in southern China, south-east Asia and northern Africa. The development and stepwise progression of NPC involves accumulation of multiple gross genetic changes during the clonal expansion of Epstein–Barr virus (EBV)-infected nasopharyngeal epithelial cell population. Here, using paired-end whole-transcriptome sequencing, we discovered a number of chimeric fusion transcripts in a panel of EBV-positive tumour lines. Among these transcripts, a novel fusion of ubiquitin protein ligase E3 component n-recognin 5 (*UBR5*) on 8q22.3 and zinc finger protein 423 (*ZNF423*) on 16q12.1, identified from the NPC cell line C666-1, was recurrently detected in 12/144 (8.3%) of primary tumours. The fusion gene contains exon 1 of *UBR5* and exons 7–9 of *ZNF423* and produces a 94 amino acid chimeric protein including the original C-terminal EBF binding domain (ZF29–30) of ZNF423. Notably, the growth of NPC cells with *UBR5–ZNF423* rearrangement is dependent on expression of this fusion protein. Knock-down of *UBR5–ZNF423* by fusion-specific siRNA significantly inhibited the cell proliferation and colony-forming ability of C666-1 cells. The transforming ability of *UBR5–ZNF423* fusion was also confirmed in NIH3T3 fibroblasts. Constitutive expression of *UBR5–ZNF423* in NIH3T3 fibroblasts significantly enhanced its anchorage-independent growth in soft agar and induced tumour formation in a nude mouse model. These findings suggest that expression of UBR5–ZNF423 protein might contribute to the transformation of a subset of NPCs, possibly by altering the activity of EBFs (early B cell factors). Identification of the oncogenic *UBR5–ZNF423* provides new potential opportunities for therapeutic intervention in NPC.
© 2013 The Authors. *Journal of Pathology* published by John Wiley & Sons Ltd on behalf of Pathological Society of Great Britain and Ireland.

Keywords: *UBR5–ZNF423* fusion; transcriptome sequencing; nasopharyngeal carcinoma; oncogene; gene rearrangement; Epstein–Barr virus

Introduction

Despite its rarity in most parts of the world, nasopharyngeal carcinoma (NPC) is one of the common cancers in southern China, south-east Asia and northern Africa. In endemic regions, it is consistently associated with Epstein–Barr virus (EBV) infection and appears as non-keratinizing carcinoma. Radiotherapy is an effective treatment for NPC patients with early disease, but therapeutic strategies for patients

presenting with metastatic disease or refractory cancer relapse remain less successful [1]. The limited knowledge on genetic lesions-driven initiation and progression of this cancer is a major barrier in advancing current therapeutic intervention. We have previously delineated multiple key genetic alterations, such as inactivation of *p16* and *RASSF1A* tumour suppressors and amplification of *LTBR*, that contribute to the tumourigenesis of NPC [2]. Our earlier cytogenetic and spectral karyotyping studies have demonstrated the prevalence of chromosomal translocations in EBV-positive NPC tumour lines [3–5]. These findings hint that gene rearrangements may also contribute to the genesis of this cancer. However, this hypothesis has not been proven, due to the technical limitations of conventional molecular and cytogenetic approaches. Advances in next-generation sequencing (NGS) technologies provide the means to systematically discover novel gene fusions caused by chromosomal translocations, inversions and deletions in cancer cells [6]. Using paired-end transcriptome sequencing, a number of recurrent gene rearrangements were successfully identified in various epithelial cancers, including prostate, lung, breast and colon cancers [7]. The new evidence demonstrates that chromosomal translocations or gene rearrangements are important for driving genetic change in solid cancers [7]). Here, we aimed to explore the involvement of gene rearrangements in EBV-associated NPC. We used pair-end whole-transcriptome sequencing to identify the fusion transcripts in a panel of EBV-positive tumours. A recurrent *UBR5–ZNF423* fusion gene detected in both NPC cell line and primary tumours was characterized. We demonstrated the oncogenic properties and transforming abilities of *UBR5–ZNF423* in NPC cells and NIH3T3 fibroblasts. This report provides compelling support for *UBR5–ZNF423* fusion as a crucial genetic change during tumourigenesis in a subset of NPCs.

Materials and methods

Cell lines, xenografts and primary tumours

Six EBV-positive xenografts (xeno-666, xeno-2117, xeno-1915, xeno-99186, C15 and C17) and a cell line (C666-1) established in our laboratories were included in the study [3–5]. The primary tumour samples include 42 frozen and 102 formalin-fixed, paraffin-embedded specimens retrieved from the tissue bank of the Department of Anatomical and Cellular Pathology at the Prince of Wales Hospital in Hong Kong and the Ontario Cancer Institute in Canada, respectively. The study protocol was approved by the respective clinical research ethics committees and Institutional Review Boards.

Paired-end transcriptome sequencing and identification of fusion genes

Total RNA was extracted from the tumour lines and its quality was assessed using an Agilent Bioanalyser.

cDNA libraries were prepared and sequenced (100 nt paired-end) on the Illumina Hi-seq2000, as previously described, to a depth of 50–80 million paired-end reads per sample [8]. To identify fusion genes from transcriptome sequencing, the data were analysed by the computational pipeline called deFuse, which uses clusters of discordant paired-end alignments to inform a split-read alignment analysis for finding fusion boundaries [8]. The UCSC *H. sapiens* reference genome (build hg19) was used for alignments.

RT–PCR and direct DNA sequencing

For reverse transcription–polymerase chain reaction (RT–PCR) analysis, total RNA was extracted from frozen specimens and microdissected paraffin-embedded tissue, using Trizol (Invitrogen) and Recoverall Total Nucleic Acid Isolation Kit for FFPE samples (Ambion), respectively. To confirm the expression of potential fusion transcripts in NPC tumour lines, RT–PCR was performed. The fusion-specific primers were designed within the margins of the paired-end read sequences and are listed in Table S1 (see supplementary material). The amplified PCR products were isolated from the gel, purified and subjected to direct DNA sequencing to confirm the sequences and fusion breakpoints. DNA sequencing was carried out using a BigDye 3.1 Cycle Sequencing Kit (Applied Biosystems) and analysed using an ABI 3130X Genetic Analyzer (Applied Biosystems).

Quantitative RT–PCR (qRT–PCR)

qRT–PCR was performed using the Taqman gene expression assay and the ABI 7500 Fast Real-time PCR system (Applied Biosystems), following the manufacturer's protocol. All reactions were performed in triplicate. Expression levels of each target relative to the housekeeping gene β-*actin* were determined on the basis of the comparative threshold cycle C_T method ($2^{-\Delta\Delta C_T}$). The primer and probe sequences used in these experiments are listed in Table S1 (see supplementary material).

Fluorescence *in situ* hybridization (FISH) analysis

To detect the translocation involving *UBR5* and *ZNF423* in the tumour samples, interphase FISH analysis was performed using either break-apart or co-localizing probe strategies. The ZNF423 break-apart probe was composed of two bacterial artificial chromosome (BAC) clones, PR11-48I18 (green) and RP11-426M23 (red) which locate at the 5′ and 3′ regions of the *ZNF423* gene at 16q22.3, respectively. The co-localizing probes for detecting the fusions includes the BAC clones RP11-12K18 (green) at the 5′ region of *UBR5* and RP11-426M23 (red) at the 3′ region of *ZNF423*. All BAC clones were purchased from Invitrogen (Carlsbad, CA, USA). The probe DNA was prepared using the Qiagen Plasmid Maxi Kit (Qiagen)

after colony purification. DNA was labelled by nick translation with spectrum green–dUTP or spectrum orange–dUTP (Vysis, Abbott Molecular, IL, USA). For each sample, at least 50 nuclei were evaluated. Fusion signals were defined as a single yellow overlapping signal or a red and green signal < two signal diameters apart. The break-apart signals were those separated by ≥ two signal diameters or only single red signal of the 3′ region of *ZNF423*. Loss of green signal in cases with t(8;16)(q22;q12) may be due to the deletion of chromosome 16q12-ter, including the 5′-regions of *ZNF423* in NPCs.

Cloning of *UBR5–ZNF423* fusion construct

A 359 bp DNA sequence containing the coding region of the *UBR5–ZNF423* fusion gene was amplified from C666-1 cDNA, using the *UBR5*-F primer, 5′-AA GCTTGGAAAGCACCATGACGTCCATC (NM_015 902.5:446-467), and the *ZNF423*-R primer, TCTAGA TCACTGTGCGTGCTGGCTC (NM_015069.2:4134-4153). The PCR product was cloned into pcDNA3.1 expression vector via *Hin*dIII and *Xba*I sites. The fusion gene sequence was validated by direct sequencing. The fusion construct was transfected into NP69 or NIH3T3 cells, using Lipofectamine™ LTX reagent (Invitrogen) according to the manufacturer's instructions. Stable NIH3T3 cells were obtained by selecting the transfectant in the culture medium containing 400 μg/ml G418 (Invitrogen) for 6 weeks.

Western blotting

The expression of *UBR5–ZNF423* was detected by western blotting, as described [9]. The anti-*ZNF423* antibody (ab94451; Abcam, MA, USA) specific for the C-terminal sequences (amino acids 1235–1284) was used. The blot was reprobed with monoclonal antibody against β-actin (Sigma, St. Louis, MO, USA) as control.

siRNA transfection

To knock down the expression of *UBR5–ZNF423*, two independent fusion-specific siRNA duplexes (S1 and S2; see supplementary material, Table S1) were transfected into C666-1 cells, using LipofectAMINE 2000 (Invitrogen, Carlsbad, CA, USA) as described [9]. Non-specific control siRNA and reagent control were included in the experiments.

Cell proliferation and colony-formation assay

Cell proliferation and anchorage-dependent growth of siRNA-transfected C666-1 cells was determined by performing WST-1 and colony-formation assays as previously described [9,10]. All the experiments were carried out in triplicate.

Anchorage-independent growth and *in vivo* tumourigenicity assays

The soft agar assay for anchorage-independent growth was carried out using 5×10^5 stable NIH3T3 cells containing *UBR5–ZNF423* or empty vector in 4 ml medium supplemented with 0.35% agarose and layered on a 5 ml base of 0.7% agarose [11]. Experiments were carried out in triplicate. After 40 days, cells were stained with 0.8 mM *p*-iodonitrotetrazolium violet (Sigma-Aldrich). The *in vivo* tumourigenicity assay was performed as described previously [11]. Stable NIH3T3 cells (1×10^6) containing *UBR5–ZNF423* or empty vector were injected subcutaneously into four 5 week-old male Balb/c nude mice. All experimental procedures were approved by the Animal Ethics Committee of the Chinese University of Hong Kong.

Results

Identification of novel *UBR5–ZNF423* fusion transcripts in NPC

In this study, we comprehensively searched for gene rearrangements in EBV-associated NPCs by paired-end whole-transcriptome sequencing. A large number of potential fusion transcripts were identified from the transcriptome sequencing results of six EBV-positive NPC tumour lines, using the deFuse gene fusion discovery algorithm. To discover the functional chimeric genes and avoid false-positive nominations, candidate fusion transcripts containing coding regions and with > 0.85 prediction probability were prioritized. Sixteen candidate non-adjacent fusions were selected and subjected to validation (Table 1). As shown in Figure 1a, b (see also supplementary material, Figure S1), the fusion transcripts were confirmed in NPC tumour lines by RT–PCR and direct sequencing of PCR products. Among the top-ranked fusion transcripts, we focused on a fusion between exon 1 of *UBR5* on chromosome 8q22.3 and exon 7 of *ZNF423* on chromosome 16q12.1 in the EBV-positive cell line C666-1 (Figure 1a, b). *ZNF423* is a frequent target of retroviral integration in murine B cell lymphomas and aberrant expression of *ZNF423* induces blast crisis of chronic myelogenous leukaemia [12,13]. Mutations of *ZNF423* cause dysregulated DNA damage response signalling and contribute to the pathogenesis of nephronophthisis-related ciliopathies (NPHP-RC) [14]. These data indicate that rearrangement of the *ZNF423* gene may contribute to the tumourigenesis of NPC. The inter-chromosomal translocation t(8;16)(q22;q12) in C666-1 was delineated in our spectral karyotyping (SKY) study (Figure 1b; see also supplementary material, Figure S2) [5]. The fusion of the *UBR5* and *ZNF423* genes was also confirmed by FISH analysis, using both break-apart and fusion probes (Figure 1c). To define the genomic breakpoint, we reviewed the whole-genome sequencing data of C666-1, which were generated from paired 100-base reads with average of

Table 1. Predicted chimeric fusion transcripts from six EBV-positive NPC tumour lines

Sample	Splitr count	Span count	Gene 1	Gene 2	Ref Seq 1	Ref Seq 2	Chr 1	Strand 1	Chr 2	Strand 2	Breakpoint homology	Deletion	Eversion
C666-1	8	6	CTAGE5	PSMA3	ENSG00000150527	ENSG00000100567	14	+	14	+	2	Y	N
C666-1	62	68	UBR5	ZNF423	ENSG00000104517	ENSG00000102935	8	−	16	−	1	N	N
C666-1	28	35	RARS	MAD1L1	ENSG00000113643	ENSG00000002822	5	+	7	−	4	N	N
X2117	56	92	NCAPD2	PRPF3	ENSG00000010292	ENSG00000117360	12	+	1	+	2	N	N
X2117	160	104	RANBP9	CDKAL1	ENSG00000010017	ENSG00000145996	6	−	6	+	3	N	N
X2117	24	5	ADRBK1	IGHMBP2	ENSG00000173020	ENSG00000132740	11	+	11	+	4	Y	N
X1915	1	6	IGSF3	GGT1	ENSG00000143061	ENSG00000100031	1	−	22	+	7	N	N
X99186	17	13	ANAPC1	RGPD3	ENSG00000153107	ENSG00000153165	2	−	2	−	2	Y	N
C15	11	8	RCC1	UBE2D2	ENSG00000180198	ENSG00000131508	1	+	5	−	5	N	N
C17	169	25	GTF2I	CLIP2	ENSG00000077809	ENSG00000106665	7	+	7	+	13	N	Y
C17	104	64	KIAA1967	SORBS3	ENSG00000158941	ENSG00000120896	8	+	8	+	4	N	Y
C17	8	11	KIAA1217	FCN1	ENSG00000120549	ENSG00000120549	10	+	9	−	15	N	N
C17	2	7	VCL	ABCC8	ENSG00000035403	ENSG00000006071	10	+	11	−	8	N	N
C17	91	138	RAB9A	EGFL6	ENSG00000123595	ENSG00000198759	X	+	X	+	1	N	Y
C17	9	9	ENSA	AADACL2	ENSG00000143420	ENSG00000197953	1	−	3	+	17	N	N
C17	11	10	GTF2H3	DDX55	ENSG00000111358	ENSG00000111364	12	+	12	+	0	N	Y

×60 coverage (unpublished data). The genomic breakpoint of this fusion gene was found in intron 1 of *UBR5* (nt#103379335 on chromosome 8) and intron 6 of *ZNF423* (nt#49650741 on chromosome 16). The genomic junction of *UBR5* and *ZNF423* shows microhomology within the three-base (CTA) region. The fusion DNA sequence was confirmed by genomic PCR analysis and Sanger sequencing (Figure 1b).

UBR5–ZNF423 fusion transcripts in primary NPC

To address the prevalence of *UBR5–ZNF423* fusion transcripts in NPC, 42 primary tumours from patients in Hong Kong were examined in our preliminary study. We detected the recurrent *UBR5–ZNF423* fusion transcripts in four of these cases by RT–PCR analysis (Figure 2a). The fusion transcripts were confirmed by direct DNA sequencing of PCR products and FISH analysis (Figure 2b, c). We did not detect *UBR5–ZNF423* in all of the 22 normal nasopharyngeal epithelial samples (see supplementary material, Figure S3). A cohort of 102 paraffin-embedded primary tumours from Toronto was then recruited for confirming the prevalence of *UBR5–ZNF423* fusion and determining its clinicopathological significance. By RT–PCR analysis and direct DNA sequencing, the gene rearrangement was confirmed in 8/102 tumour specimens of NPC patients (see supplementary material, Figure S2). In total, the recurrent fusion gene was detected in 12/144 (8.3%) primary tumours. This is the first time a recurrent gene rearrangement has been identified in EBV-associated NPC. As shown in Table 2, the *UBR5–ZNF423* fusion occurred in patients with advanced disease.

Characterization of full-length *UBR5–ZNF423* fusion gene

By 5'- and 3'- Rapid Amplification of cDna Ends (RACE) and PCR amplification, we revealed that the full-length *UBR5–ZNF423* fusion gene includes the 5'-UTR and exon 1 of *UBR5* and exons 7–9 of *ZNF423*

Table 2. Correlation between *UBR5–ZNF423* fusion gene expression and clinicopathological features in 102 patients with non-keratinizing NPC

Variables	No. of patients	Fusion gene		p*
		Present	Absent	
Age (years)				
≤ 50	44	2	42	0.46
> 50	58	6	52	
Gender				
Male	70	4	66	0.25
Female	32	4	28	
Metastasis				
Absent (M0)	98	7	91	0.28
Present (M1)	4	1	3	
Clinical stage				
Early (stages 1 and 2)	30	0	30	0.10
Late (stages 3 and 4)	72	8	64	

*As determined by Fischer's exact test.

and is 1031 bp in length (GenBank Accession No. KC155256; see supplementary material, Figure S4). Exon-specific gene expression analysis of transcriptome sequencing data and qRT–PCR analysis confirmed that the *ZNF423* exons after fusion breakpoint (exons 7–9) are highly expressed in C666-1 (Figure 3a, b). Expression levels of these three exons in C666-1 are significantly higher than those of other NPC tumour lines. In other nasopharyngeal epithelial cells, we detected only weak expression or absence of *ZNF423*. Exons 7–9 of *ZNF423* in C666-1 cells might be expressed exclusively from the fusion gene, rather than the natural *ZNF423*. On the other hand, constitutive expression of all exons of *UBR5* was detected in all NPC tumour lines (see supplementary material, Figure S5). The t(8;16)(q22;p12) translocation links the 3'-region of *ZNF423*, which contains intron 6, to 3'-UTR to intron 1 of *UBR5*. Thus, the over-expression of *ZNF423* exons 7–9 might be driven by the *UBR5* promoter, which is constitutively activated in NPC cells.

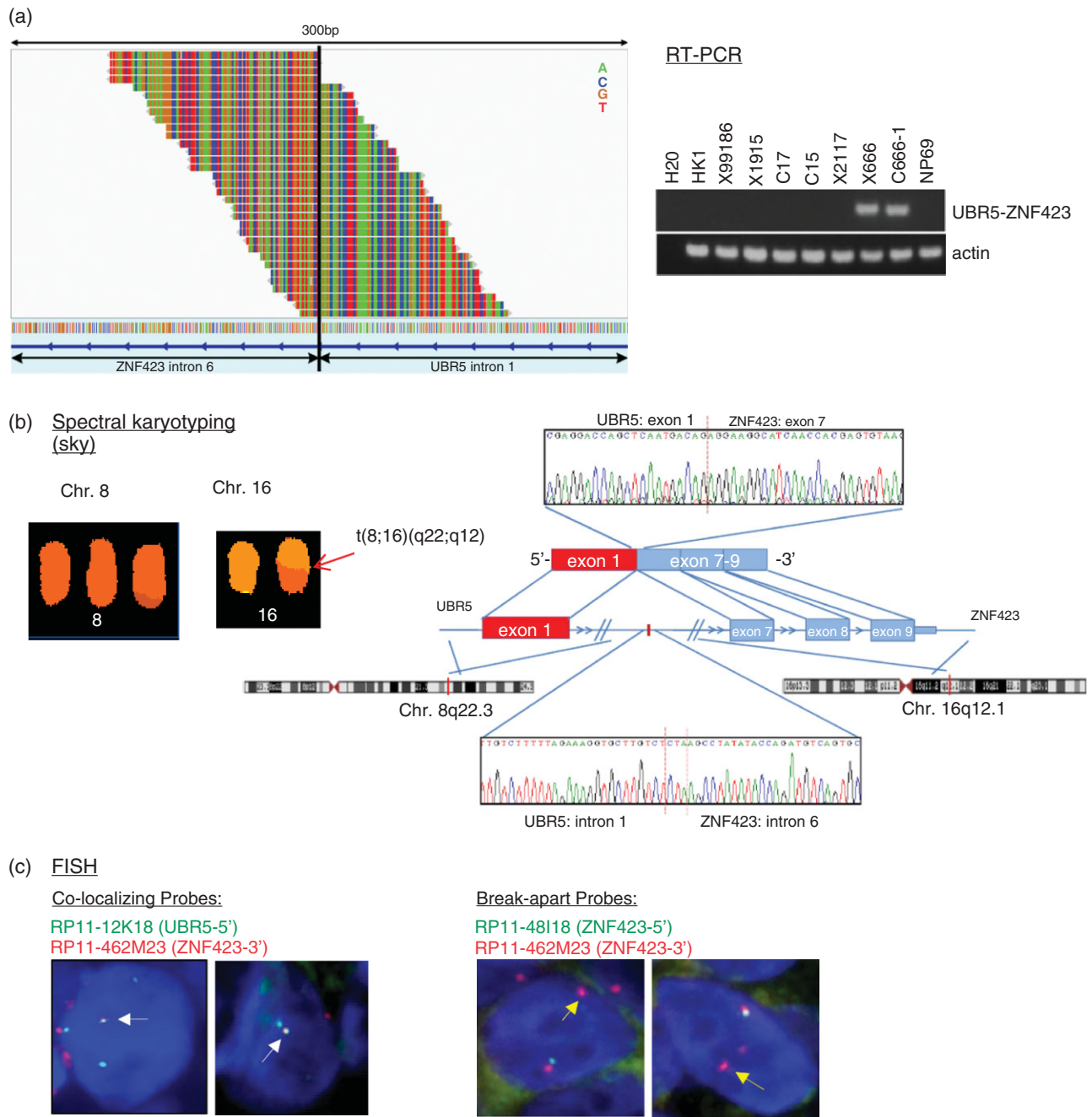The predicted in-frame protein product of *UBR5–ZNF423* fusion contains 94 amino acids, with a

Figure 1. Discovery of the *UBR5–ZNF423* gene fusion in EBV-positive NPC. (a) By whole transcriptome sequencing and deFuse analysis, an *UBR5–ZNF423* fusion transcript was identified in the NPC cell line C666-1. The spanning reads spanning the junction breakpoint within the *UBR5–ZNF423* gene fusion are shown on the left. A spanning read is a read, one of whose end-sequences is aligned across the junction of the predicted fusion transcript. The *UBR5–ZNF423* fusion transcripts in C666-1 and xeno-666 were confirmed by RT–PCR. The C666-1 cell line was derived from a NPC xenograft, xeno-666. (b) Direct sequencing confirmed that the chimeric transcripts contained the fusion of *UBR5* exon 1 and *ZNF423* exon7. Fusion junctions with respective exon numbers comprising the chimeric transcripts are indicated. The genomic fusion of *UBR5* intron 1 on chromosome 8 and *ZNF423* intron 6 on chromosome 16q was detected by direct DNA sequencing. Red bar indicates the 3 bp (CTA) microhomology region of the junction. Spectral karyotyping (SKY) analysis also showed the presence of a derivative chromosome, t(8; 16)(q22;q12), in C666-1. (c) The *UBR5–ZNF423* fusion in C666-1 cells was validated by FISH analysis, using both break-apart and fusion probes. White arrows, fusion signal when co-localizing probes were used; yellow arrows, distinct red signal when break-apart probes were used.

predicted molecular mass of 10.8 kDa (Figure 3c). The fusion protein mainly harbours the original C-terminal EBF-binding domain (ZF28-30) of ZNF423. Only two amino acids at the N-terminal of the fusion protein are encoded from *UBR5* sequences. *ZNF423* encodes a nuclear protein which contains a

DNA-binding domain and 30 Kruppel-like C2H2 zinc fingers (Figure 3c) [18–20]. As a multifunctional transcriptional regulator, it contributes to the regulation of different signalling pathways (eg NOTCH, BMP, RA, EBF) through its distinct sets of zinc fingers [13–16]. However, the translocation leads to the
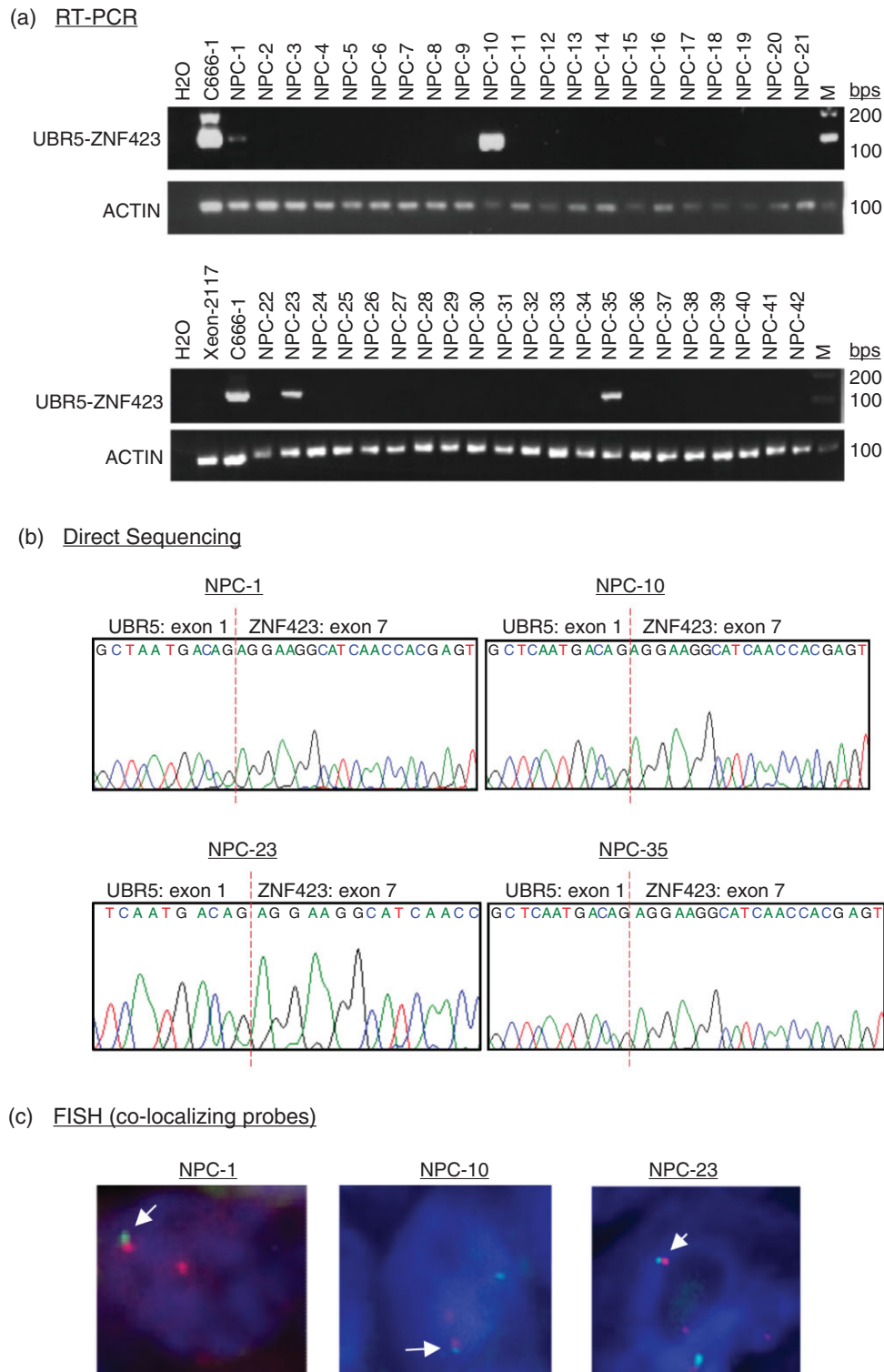
(a) RT-PCR

(b) Direct Sequencing

(c) FISH (co-localizing probes)

Figure 2. Recurrent *UBR5–ZNF423* fusion transcripts in primary NPCs. (a) By RT–PCR, *UBR5–ZNF423* fusion transcripts were detected in 4/42 primary tumours from Hong Kong NPC patients. (b) The PCR products of *UBR5–ZNF423* fusion transcripts in primary NPCs were validated by DNA sequencing. (c) Primary NPC cases with *UBR5–ZNF423* fusion were validated by FISH analysis, using break-apart and fusion probes; the fusion signals are indicated by white arrows.

expression of a truncated ZNF423 protein containing C-terminal EBF-binding domain (ZF28-30) only. To identify expression of the UBR5–ZNF423 fusion protein, western blotting was performed with anti-body specific for the C-terminal region of ZNF423 in C666-1 cells. Despite the absence of full-length ZNF423 protein, a putative UBR5–ZNF423 fusion protein, which was represented by a band at approx-imately 10.8 kDa, was detected (Figure 3d). The result indicated that only the C-terminal EBF-binding domain (ZF28-30) of ZNF423 expresses in these cells.
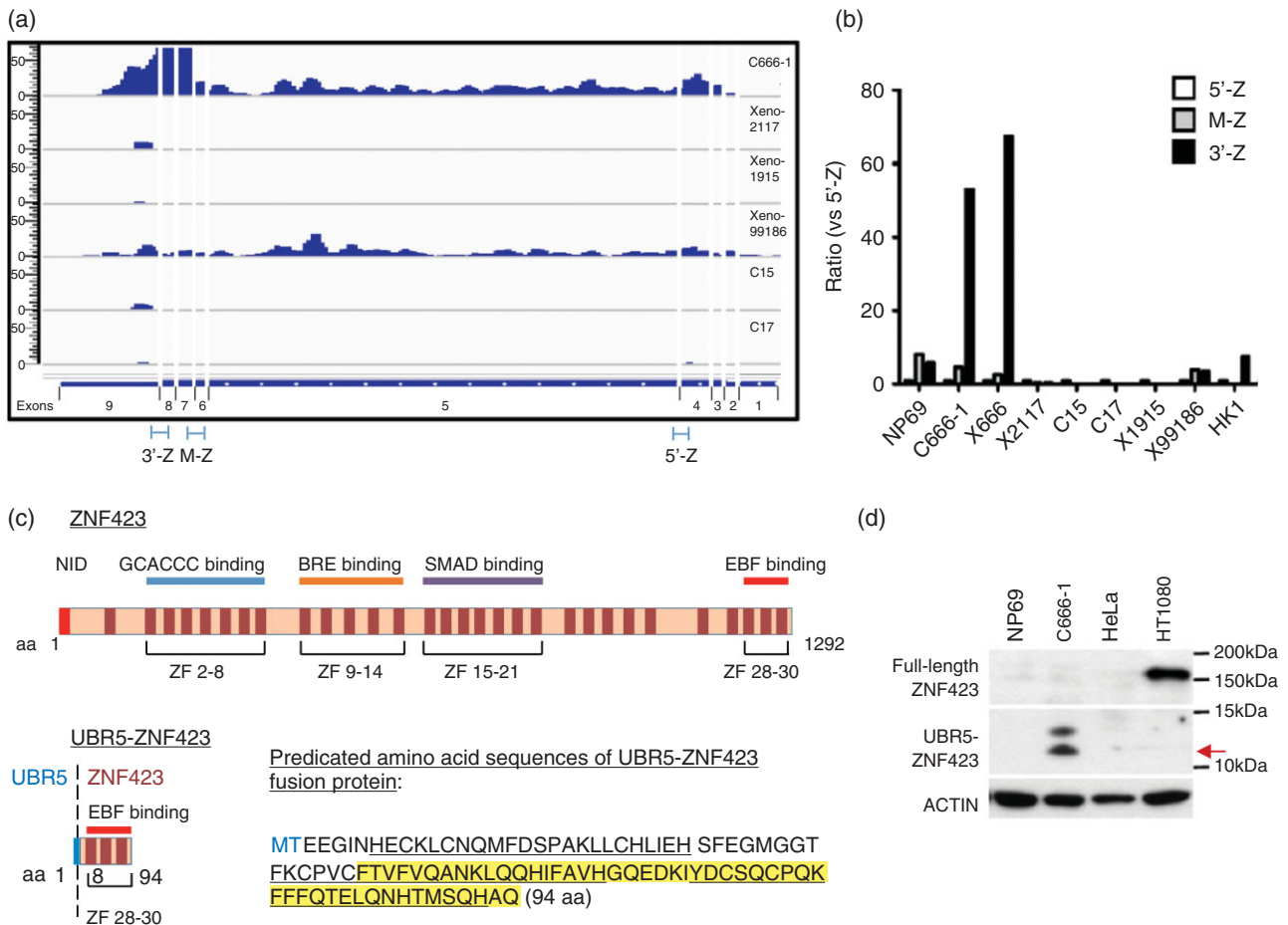
Figure 3. Expression of *UBR5–ZNF423* fusion transcripts and protein in C666-1. Exon-specific gene expression analysis of whole-transcriptome sequencing (a) and quantitative RT–PCR (b) revealed the over-expression of exons 7–9 of *ZNF423* in C666-1. NP69 (immortalized normal nasopharyngeal epithelial cells) and HK1 (EBV-negative well differentiated NPC cell line) were recruited as additional references. 5'-Z (Hs01046870_m1, Applied Biosystems), M-Z (Hs00391820_m1, Applied Biosystems), and 3'-Z (Hs00323880_m1, Applied Biosystems) indicate the regions in ZNF423 assessed by quantitative RT–PCR assay. (c) Predicted amino acid sequences and domain of UBR5–ZNF423 fusion protein: for the predicted amino acid sequences, the amino acids derived from UBR5 sequences are highlighted in blue; the zinc finger sequences are underlined. (d) A chimeric UBR5–ZNF423 protein of approximately 10.8 kDa (red arrow) was detected in C666-1 cells by western blotting. Full-length ZNF423 protein is not expressed in either C666-1 or the immortalized nasopharyngeal epithelial cells NP69. HT1080 and HeLa are positive and negative controls, respectively, for wild-type ZNF423 expression. The amino acid sequences of the UBR5–ZNF423 fusion protein specific for the anti-ZNF423 antibody (ab4451; Abcam) is highlighted in yellow.

Oncogenic properties and transforming activity of *UBR5–ZNF423*

To assess the effects of *UBR5–ZNF423* on the growth of NPC cells, its expression in C666-1 cells was knocked down by siRNAs specific for the fusion transcripts (Figure 4a). As shown in Figure 4b, c, knock-down of *UBR5–ZNF423* significantly inhibited cell proliferation and colony-forming ability. Furthermore, we established stable *UBR5–ZNF423*-transfected NIH3T3 fibroblast cells which expressed the fusion protein at levels comparable to those present in C666-1 (Figure 5a). Constitutive expression of *UBR5–ZNF423* significantly enhanced the anchorage-independent growth of NIH3T3 cells in soft agar (Figure 5b). In the nude mouse model, large tumours were consistently detected in the sites implanted with NIH3T3 cells expressing *UBR5–ZNF423* over a time course of 5 weeks (Figure 5c). The results indicate

that *UBR5–ZNF423* is able to induce tumourigenic transformation of NIH3T3 cells. Our findings strongly suggest that *UBR5–ZNF423* is a novel oncogenic fusion which plays a role as a driver of genetic change in the genesis of a subset of NPCs.

Discussion

In this study, we discovered a novel *UBR5–ZNF423* transforming fusion gene in 8.3% of NPCs by whole-transcriptome sequencing. Our finding provides the first evidence to support the important role of gene rearrangement in NPC. As shown in Figure 3, fusion of the *UBR5* and *ZNF423* genes results in the *de novo* expression of a truncated ZNF423 protein. *ZNF423* is the human homologue of *Zfp423*, which was originally identified as a binding partner and negative
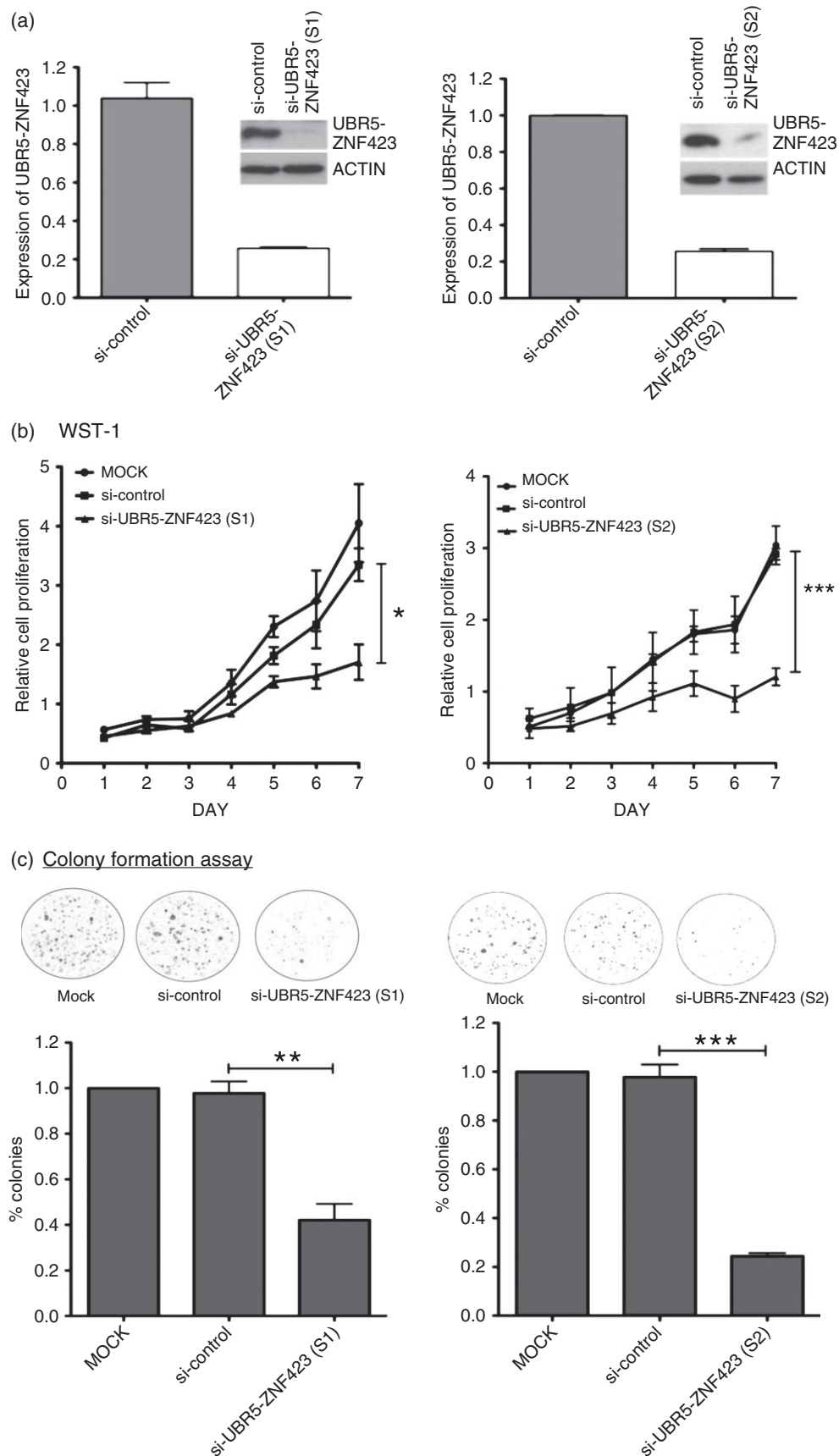
Figure 4. Knock-down of *UBR5–ZNF423* inhibits cell proliferation and colony formation of NPC cells. (a) Expression of *UBR5–ZNF423* was knocked down by the fusion-specific siRNA. The suppression of *UBR5–ZNF423* in C666-1 was confirmed by quantitative RT–PCR and western blotting. (b) WST-1 assay demonstrated that cell proliferation was significantly reduced in C666-1 treated with siRNAs (S1 and S2) targeting *UBR5–ZNF423* fusion. (c) Knock-down of *UBR5–ZNF423* by siRNA significantly inhibited the colony-forming ability of C666-1 cells.
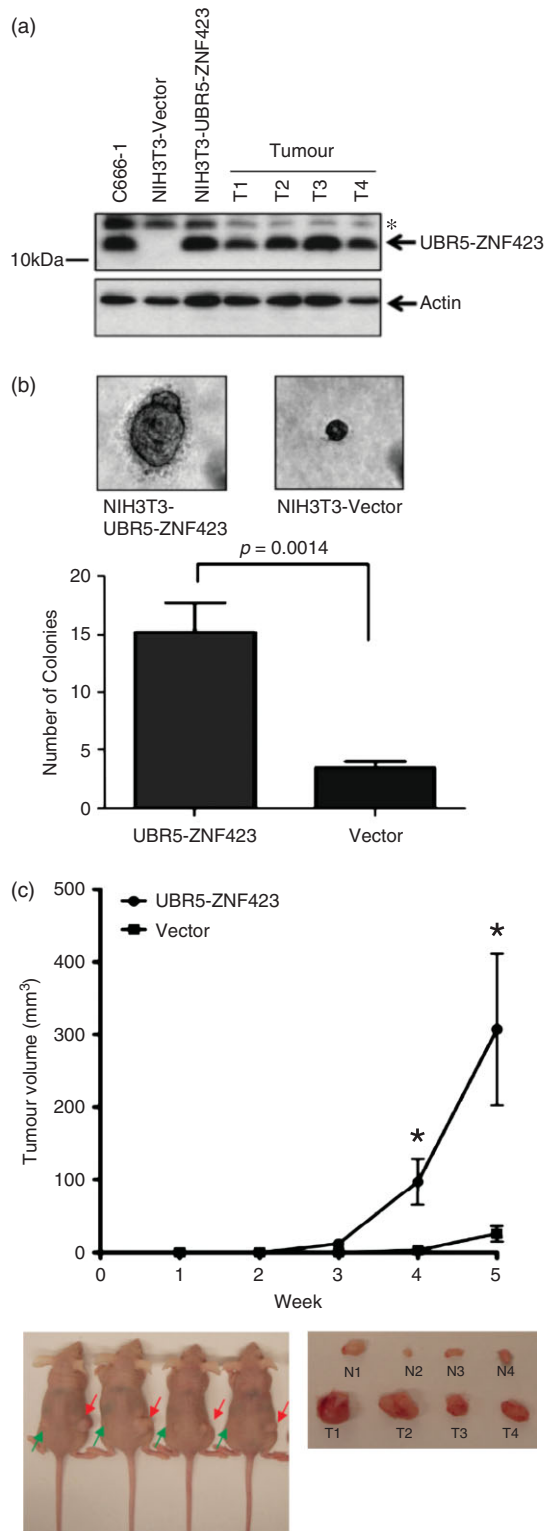
Figure 5. Transforming activity of *UBR5–ZNF423*. (a) By western blotting, the expression of UBR5–ZNF423 fusion protein in the stable *UBR5–ZNF423*-transfected NIH3T3 cells and tumours dissected from the xenografts (T1–T4) was detected; *non-specific bands. (b) Stable expression of *UBR5–ZNF423* induces the anchorage-independent growth of NIH3T3 cells. Significant increase in number and size of colonies in the stable *UBR5–ZNF423*-expressing cells was demonstrated by soft agar assay. (c) *In vivo* tumourgenic assay in nude mice showed that tumours formed in the sites implanted with NIH3T3 cells expressing *UBR5–ZNF423* (T1–T4, red arrows) were consistently larger than those implanted with vector controls (N1–N4, green arrows).

regulator of *Ebf1* (early B cell factor) in the rat [14]. The protein functions as a DNA-binding transcription factor by using distinct zinc fingers in different signalling pathways (eg NOTCH, BMP, RA, EBF) and is essential for B cell and olfactory nervous system development. In this study, we found that only the EBF-binding domain (ZF28-30) of ZNF423 is retained in the UBR5–ZNF423 fusion protein. By co-immunoprecipitation, UBR5–ZNF423 protein was shown to bind to EBF3, which is consistently expressed in NPC tumours (see supplementary material, Figure S6). The fusion protein may facilitate its transforming activity through binding to EBF3. A number of studies indicate that the EBFs (eg EBF1, EBF3) act as tumour suppressors and inactivation of EBFs contributes to the development of both haematological and solid cancers [17–19]. Translocation of *EBF1* and its target gene, *PAX5*, was frequently found in paediatric acute lymphoblastic leukaemia (ALL) [18]. Inactivation of EBF3 by promoter methylation was shown in tumour cell lines of various tissues, including colon, breast and brain. However, EBF3 might play a pro-oncogenic role in some human cancers. A recent study has shown that over-expression of EBF3 in medulloblastoma cancer stem cells CSCs greatly enhances its tumourigenic ability [20]. Notably, EBF is also required for EBNA2 regulation of the promoter of *LMP1*, an EBV oncoprotein [21]. The constitutive expression of the EBF-binding domain encoded by *UBR5–ZNF423* may disrupt the EBF-mediated transcriptional regulation of EBV and cellular target genes. Currently, study on the function of EBFs, interaction between the fusion protein and EBFs and their roles in EBV-associated NPC are ongoing in our laboratories. In conclusion, identification of the *UBR5–ZNF423* fusion gene might provide new opportunities for the development of novel biomarkers and therapeutic strategies for NPC patients. Detection of chimeric *UBR5–ZNF423* fusion transcripts in patients' peripheral blood and plasma will be useful for the prediction of remission disease in the subgroup of NPC patients. Furthermore, screening of inhibitors or small molecules that interfere with *UBR5–ZNF423* binding activity may enable the development of NPC-targeting therapies.

Acknowledgements

Author contributions

KWL, KFT and GTYC designed the study; GTYC, RWML, ABYH, CC, CYKT, JWFY, SWML and NW carried out experiments; KYLY, SDL, KKYT, HK and JSS conducted the massive parallel genome sequencing and bioinformatics analysis; JKSW, SWT, TTCY, PB, BO, FFL and KFT provided the NPC tumour models, primary tumour specimens and clinical data; and KWL, KFT, GTYC and RWML were involved in data analysis and writing the paper. All authors had final approval of the submitted manuscript.

Abbreviations

EBV, Esptein–Barr virus; FISH, fluorescence *in situ* hybridization; NPC, nasopharyngeal carcinoma; qRT–PCR, quantitative reverse transcription–polymerase chain reaction.

References

1. Lo KW, To KF, Huang DP. Focus on nasopharyngeal carcinoma. *Cancer Cell* 2004; **5:** 423–428.
2. Lo KW, Chung GT, To KF. Deciphering the molecular genetic basis of NPC through molecular, cytogenetic, and epigenetic approaches. *Semin Cancer Biol* 2012; **22:** 79–86.
3. Huang DP, Ho JH, Chan WK, *et al.* Cytogenetics of undifferentiated nasopharyngeal carcinoma xenografts from southern Chinese. *Int J Cancer* 1989; **43:** 936–939.
4. Bernheim A, Rousselet G, Massaad L, *et al.* Cytogenetic studies in three xenografted nasopharyngeal carcinomas. *Cancer Genet Cytogenet* 1993; **66:** 11–15.
5. Wong N, Hui AB, Fan B, *et al.* Molecular cytogenetic characterization of nasopharyngeal carcinoma cell lines and xenografts by comparative genomic hybridization and spectral karyotyping. *Cancer Genet Cytogenet* 2003; **140:** 124–132.
6. Maher CA, Kumar-Sinha C, Cao X, *et al.* Transcriptome sequencing to detect gene fusions in cancer. *Nature* 2009; **458:** 97–101.
7. Edwards PA. Fusion genes and chromosome translocations in the common epithelial cancers. *J Pathol* 2010; **220:** 244–254.
8. Ju YS, Lee WC, Shin JY, *et al.* A transforming *KIF5B* and *RET* gene fusion in lung adenocarcinoma revealed from whole-genome and transcriptome sequencing. *Genome Res* 2012; **22:** 436–445.
9. Man CH, Lun SW, Hui JW, *et al.* Inhibition of NOTCH3 signalling significantly enhances sensitivity to cisplatin in EBV-associated nasopharyngeal carcinoma. *J Pathol* 2012; **226:** 471–481.
10. Kwong J. Chow LSN, Wong AYH, *et al.* Epigenetic inactivation of deleted in lung and esophageal cancer 1 (*DLEC1*) gene in nasopharyngeal carcinoma. *Genes Chromosomes Cancer* 2007; **46:** 171–180.
11. Chow LS, Lo KW, Kwong J, *et al. RASSF1A* is a target tumor suppressor from 3p21.3 in nasopharyngeal carcinoma. *Int J Cancer* 2004; **109:** 839–847.
12. Warming S, Suzuki T, Yamaguchi TP, *et al.* Early B-cell factor-associated zinc-finger gene is a frequent target of retroviral integration in murine B-cell lymphomas. *Oncogene* 2004; **23:** 2727–2731.
13. Miyazaki K, Yamasaki N, Oda H, *et al.* Enhanced expression of p210BCR/ABL and aberrant expression of Zfp423/ZNF423 induce blast crisis of chronic myelogenous leukemia. *Blood* 2009; **113:** 4702–4710.
14. Chaki M, Airik R, Ghosh AK, *et al.* Exome capture reveals ZNF423 and CEP164 mutations, linking renal ciliopathies to DNA damage response signaling. *Cell* 2012; **150:** 533–548.
15. Tsai RY, Reed RR. Cloning and functional characterization of Roaz, a zinc finger protein that interacts with O/E-1 to regulate gene expression: implications for olfactory neuronal development. *J Neurosci* 1997; **17:** 4159–4169.
16. Tsai RY, Reed RR. Identification of DNA recognition sequences and protein interaction domains of the multiple Zn-finger protein Roaz. *Mol Cell Biol* 1998; **18:** 6447–6456.
17. Liao D. Emerging roles of the EBF family of transcription factors in tumor suppression. *Mol Cancer Res* 2009; **712:** 1893–1901.
18. Mullighan CG, Goorha S, Radtke I, *et al.* (2007) Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* 446; **7137:** 758–764.
19. Zhao LY, Niu Y, Santiago A, *et al.* (2006) An EBF3-mediated transcriptional program that induces cell cycle arrest and apoptosis. *Cancer Res* 2006; **66:** 9445–9452.
20. Corno D, Pala M, Cominelli M, *et al.* Gene signatures associated with mouse postnatal hindbrain neural stem cells and medulloblastoma cancer stem cells identify novel molecular mediators and predict human medulloblastoma molecular classification. *Cancer Discov* 2012; **2:** 554–568.
21. Zhao B, Zou J, Wang H, *et al.* Epstein–Barr virus exploits intrinsic B-lymphocyte transcription programs to achieve immortal cell growth. *Proc Natl Acad Sci USA* 2011; **108:** 14902–14907.

SUPPLEMENTARY MATERIAL ON THE INTERNET

The following supplementary material may be found in the online version of this article:

Table S1. Sequence of primers and siRNAs.

Figure S1. Validation of chimeric fusion transcripts in NPC tumour lines by RT–PCR and direct DNA sequencing.

Figure S2. Spectral karyotyping of NPC cell line C666-1.

Figure S3. Detection of *UBR5–ZNF423* fusion transcripts in normal nasopharyngeal tissues and an independent panel of primary NPCs from PMH, Toronto.

Figure S4. Sequences of the 1031 bp full-length *UBR5–ZNF423* fusion gene determined by 5′- and 3′-RACE and PCR amplification.

Figure S5. Constitutive expression of all exons of UBR5 was detected in all NPC tumour lines by exon-specific gene expression analysis of whole-transcriptome sequencing.

Figure S6. Interaction between *UBR5–ZNF423* and EBFs.