Contents lists available at ScienceDirect

Biochemistry and Biophysics Reports

Tandem repeats of the 5′ flanking region of human *MUC5AC* have a role as a novel enhancer in *MUC5AC* gene expression

Natsuko Kageyama-Yahara, Nobutake Yamamichi*, Yu Takahashi, Chihiro Takeuchi, Yuta Matsumoto, Yoshiki Sakaguchi, Kazuhiko Koike

Department of Gastroenterology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8655, Japan

ARTICLE INFO

ABSTRACT

MUC5AC is a well-known gastric differentiation marker, which has been frequently used for the classification of stomach cancer. However, the molecular mechanism of regulation of *MUC5AC* expression remains to be elucidated. In previous studies, we have shown that Gli regulated *MUC5AC* transcription through the Gli-binding motif in the 5′ region of *MUC5AC*. Gli played important roles, but independently was not sufficient for *MUC5AC* expression. In this study, we analyzed a 4010 bp fragment of the 5′-flanking promoter region of the human *MUC5AC* gene by luciferase assay, and found a novel distal enhancer region located between −1434 bp to −3000 bp upstream from the first ATG initiation codon. This region is composed of repetitive DNA sequences 5′-TCACTCAC-3′. The strength of enhancer activities depended on the length of the repetitive region. The tandem repeats are conserved among primates, but not in other mammals. The tandem repeat regions enhanced promoter activities not only of *MUC5AC* but also of other genes. The enhancer effect of the tandem repeat regions was maintained even when inverted. ChIP analysis revealed that H3K9me3 binds to the tandem repeat regions. Together, our results suggest that the tandem repeat region in the *MUC5AC* promoter has the potential to act as a strong enhancer, and H3K9me3 may contribute to histone modifications of this region.

1. Introduction

The gastric mucosal barrier protects the gastric mucosa from hydrochloric acid and various noxious agents. *MUC5AC*, a gel-forming mucin, is well-known as the main component of the gastric mucous layer [1,2]. *MUC5AC* is considered a very important prognostic marker of gastric cancer and is often used for clinical assessment [3]. We have previously reported that expression of *MUC5AC* in the stomach decreased in association with development of intestinal metaplasia [1] and that *MUC5AC* expression was related to the tumor stage: advanced gastric cancers presented reduced levels of *MUC5AC* compared with early gastric cancer [4]. Our aim is to elucidate molecular mechanisms of regulation of *MUC5AC* expression.

In our previous study, we found that Gli enhanced *MUC5AC* expression by binding to the highly conserved region containing Gli-binding (HCR-Gli) at −125/-111 bp in the promoter region of *MUC5AC*. Deletion mutant lacking HCR-Gli lost its promoter activity in luciferase assay, indicating that the Gli binding to the HCR-Gli of *MUC5AC* upstream is essential for *MUC5AC* expression. Furthermore, we found that the far distal upstream region from −1434 to −4010 bp

enhanced Gli-dependent *MUC5AC* expression [4].

In this study, we focused on the distal upstream region *MUC5AC*. The region from −1434 to −3000 bp contains high copies of tandem repeat sequences 5′-TCACTCAC-3′. The tandem repeats are conserved especially among primates. Interestingly, the tandem repeat region from −1434 to −3000 bp highly enhanced the promoter activity not only of *MUC5AC* but also of other genes, and it was effective even when inverted, indicating that the tandem repeats of the *MUC5AC* promoter region act as an enhancer. In addition, the degree of enhancement was different depending on cell lines, suggesting that cell-specific regulatory mechanisms might exist on the enhancer function of the tandem repeats. By ChIP analysis, we found that H3K9me3 bound to this tandem repeat region, suggesting that it is possible that H3K9me3 may contribute to the compaction of the chromatin in this region. Together, our results demonstrate that the tandem repeats of the *MUC5AC* promoter region are a potential enhancer of *MUC5AC* expression, and that H3K9me3 might be involved in the regulation of *MUC5AC* expression through its interactions with this tandem repeat region.

---

* Corresponding author.
  E-mail address: nyamamic-tky@umin.ac.jp (N. Yamamichi).

2. Materials and methods

2.1. Cell culture

Cancer cell lines were maintained in high-glucose DMEM with 10% fetal calf serum (Gibco/Invitrogen, Carlsbad, CA) at 37 °C in a humidified 5% CO_2 atmosphere. Human WI-38, T98G, A172, SH-10-TC, H-III-TC, MKN-7, GCIY, MKN-1, MKN-45, KE-39, KE-97 and ECC-10 cell lines were purchased from the RIKEN Bio Resource Center (Tsukuba, Japan); and AGS, SW480, and PANC-1 were purchased form the American Type Culture Collection (Manassas, VA).

2.2. Luciferase reporter assay

Luciferase reporter assays were carried out using Dual luciferase reporter assay system (Promega) as described previously [4]. Construction of pGL4.12-MUC5ACup plasmids, in which various lengths of 5′ upstream sequences of human *MUC5AC* were cloned into the pGL4.12, was described in previously [4]. To yield *LGR5* upstream plasmids, 2000 bp of *LGR5* promoter region was amplified from TIG-112 genome using the primers: 5′- CAAACTCGAGGGGTAGGAGAAGG GTGTGGG -3′ and 5′- TCATGGATCCGGTGCCCGAAGTAGGGGGCC -3′, treated with *Bam*HI and *Xho*I, and cloned into *Bgl*II-*Xho*I site of pGL4.12. The obtained plasmid was digested with *Xho*I and *Sph*I, subsequently treated with Klenow fragment (Takara) to create blunt end, and self-ligated. The resulting plasmid, which contains *LGR5* upstream 1155 bp was named pGL4.12-Lgr5up1155bp. For pTK4.12-MUC5ACup and pGL4.12-MUC5ACup-Lgr5up1155bp, *Pci*I (Klenow blunted)-*Spe*I fragment of pGL4.12-MUC5ACup3000bp (−3000 to −1425 bp upstream of *MUC5AC*) was cloned into *Kpn*I (T4-DNA-Polymerase (Takara) blunted)-*Spe*I sites of pTK4.12 and pGL4.12-LGR5up1155bp. For pGL4.12-inverted tandem repeats-MUC5ACup1433bp, *Pci*I (Klenow blunted)-*Nhe*I fragment of pGL4.12-MUC5ACup3000bp was cloned into *Kpn*I (T4-DNA-Polymerase blunted)-*Nhe*I site of pGL4.12-MUC5ACup1433bp.

2.3. Sequence data

Genomic sequences data of human *MUC5AC*-promoter region was from February 2009 human reference sequence (GRCh37/hg19). GenBank accession numbers of other species' *MUC5AC*-promoter region are: chimpanzee (NC_006478), rhusus monkey (NW_001100341), olive baboon (NC_018165), cattle (AC_000186), house mouse (NC_000073), horse (NC_009155), cat (NC_018732), dog (NC_006600).

2.4. RT-PCR

Total cellular RNA was prepared using the Isogen RNA isolation reagent (Wako Pure Chemical Industries, Osaka, Japan) as previously reported [5]. RT-PCR was performed via a Superscript One-Step reaction using the Platinum Taq (Invitrogen). The RT-PCR primer sequences used for *MUC5AC* and *GAPDH* are shown in Ref. [4].

2.5. Chromatin immunoprecipitation (ChIP) assay

ChIP analyses were performed using a Magna ChIP G (Millipore, Bedford, MA) according to the manufacture's instructions. After cross-linking using 1% formaldehyde, cells were sonicated (setting 5, Handy Sonic, model UR-20 P; Tomy Seiko, Co., Ltd., Tokyo, Japan) and confirmed that DNA fragments of ∼2000-bp size were mainly found by agarose gel electrophoresis. The sonicated sample was then subjected to immunoprecipitation using anti-H3K9me3 (Millipore, CMA308), anti-H3K27ac (Millipore, CMA309) antibodies and normal mouse IgG (Millipore). Immunoprecipitated DNA was analyzed by PCR using primers F1: 5′- TAACCCTGTCAGCCGCTCAGCCTTAAATGT -3′ and R1: 5′- AGAGCACTTCACATGTGGCAGGAGTGTGGG -3′ to amplify a fragment

of the *MUC5AC* promoter region.

2.6. PCR and sequencing analysis

1806 bp fragments in the tandem repeat regions of *MUC5AC* promoter were amplified using genome derived from various cells by using primers F1 and R1 used in ChIP assay. The 5′-upstream region from −4010 to −3194 bp (817 bp) were amplified using primers 5′- GAG CTCAGAAACAAGGCCCAGTGGGTTTTC -3′ and 5′- TTAAGGCTGAGCG GCTGACAGGGTTAGGGT -3′, the region from −1433 to −1 bp (1433 bp) were amplified using primers 5′- TGCCACATGTGAAGTGCTCTTTC TCTAGGC -3′ and 5′- TGTGTGGACGGCGGGGAAGAGTGCCCTGTC -3′. Amplified fragments were analyzed their length by agarose gel electrophoresis. 4010 bp of *MUC5AC* promoter regions were amplified from AGS, SH-10-TC, and KE-97 genome using primers MUC5ACup-*Sac*I-F2: 5′- TGCCCACAGAGCTCAGAAACAAGGCCCAGTGGG -3′ and MUC5ACup-R01re: 5′- CAACGGATCCTGTGTGGACGGCGGGGAAGAGT GCCCTGTC -3′, digested with *Sac*I and *Bam*HI, and cloned between the *Sac*I-*Bgl*II site of pGL4.12. The resulting plasmids were used for sequence analysis.

3. Results

The 5′-upstream region from −1434 to −3000 bp in the *MUC5AC* promoter contains tandem repeat sequences and possesses an enhancer function to *MUC5AC* expression.

In our previous report, the 5′-upstream distal region from −1434 to −4010 bp in the *MUC5AC* promoter showed enhancing effects on *MUC5AC* expression in the gastrointestinal cell lines (SW480, SH-10-TC, and KE-39) [4]. We also examined the effects of the distal region on *MUC5AC* expression in gastric cancer (AGS), glioblastoma (T98G and A172), and non-cancerous lung fibroblast (WI-38) cell lines (Fig. 1A). The 5′-upstream region from −1434 to −3000 bp showed enhancing effects on *MUC5AC* expression in all these cell lines and the effect was most striking in AGS (Fig. 1B). As previously reported, we also confirmed that the HCR-Gli site at −125/-111 bp was essential for *MUC5AC* expression.

The region from −1434 to −3000 bp has characteristic features which contain tandem repeats of 5′-TCACTCAC-3′ (Fig. 2A). The repeat sequence is conserved among primates (human, chimpanzee, rhesus monkey, and olive baboon), but not in cattle, mouse, horse, cat, and dog (Fig. 2B). We analyzed *MUC5AC* promoter sequences of non-primate species using tandem repeat finder program [6] and found that there was no tandem repeat in *MUC5AC* promoter region of mouse, horse, cat and dog.

To verify the effect of the repetitive region as an enhancer, the −1425 to −3000 bp sequences were ligated to the upstream of herpes simplex virus-thymidine kinase (TK) and leucine rich repeat containing G protein-coupled receptor 5 (*LGR5*) promoters and their promoter activities were examined (Fig. 3). The 5′-upstream region from −1425 to −3000 bp enhanced not only the promoter activities of *MUC5AC* but also those of TK and *LGR5*. Because it has been reported that enhancer elements can be active in an orientation-independent manner [7], we constructed an inverted version of the −1425 to −3000 bp sequences. As shown in Fig. 3, the tandem repeats region showed enhancer effects even when inverted. The effect of enhancement on gene activation was highest in AGS among the gastrointestinal and the non-gastrointestinal cell lines used in this study.

3.1. H3K9me3 interacts to the tandem repeat sequences of MUC5AC promoter region

Since differential activity of the enhancer region was observed among cell lines, we next addressed epigenetic regulation on the enhancer. We searched the UCSC genome browser (http://genome.ucsc.edu) and the GRCh37/hg19 release of human genome and found that
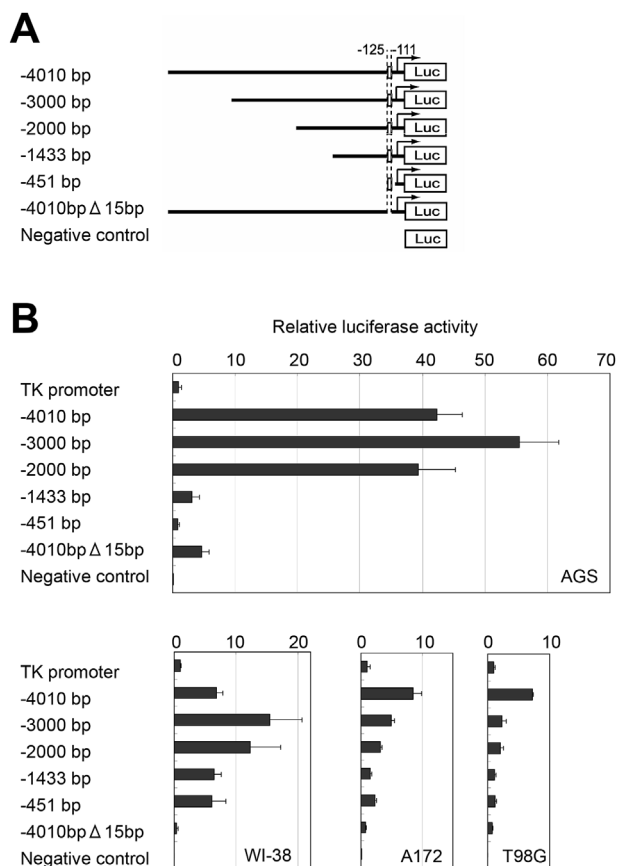
Fig. 1. Luciferase reporter analyses of a series of *MUC5AC* promoter constructs. (A) Representation of the various length of *MUC5AC* promoter constructs used in Luciferase reporter assay. (B) AGS, T98G, A172 and WI-38 cells were transfected with pTK4.12 (TK promoter), pGL4.12-MUC5ACup-4010bp, −3000bp, 2000bp, 1433bp, 451bp (−4010 bp, −3000 bp, −2000 bp, −1433 bp, −451 bp), and pGL4.12-MUC5ACup4010bpΔ15bp (a deletion mutant containing upstream 4010 bp but lacking the HCR-Gli, −4010bpΔ15bp), and examined the promoter activity at 24 h after transfection. Data represent the mean of luciferase activity (relative to TK-promoter activity) and error bars showed the standard deviation of the results from three independent experiments.

H3K27Ac and H3K9me3 bind to the tandem repeats in the *MUC5AC* promoter region in PANC-1 (human pancreas carcinoma cell line) and U2OS (human bone osteosarcoma epithelial cell line), respectively (Supplement Fig. 1). Then we examined chromatin modification of active mark H3K27ac and repressive mark H3K9me3 by ChIP assays. As shown in Fig. 4A, the tandem repeats of the *MUC5AC* promoter region were co-immunoprecipitated with H3K9me3 in all cell lines we used, whereas existence of H3K27ac on the promoter was not observed.

3.2. Length variants are observed in tandem repeats of the MUC5AC promoter

In this ChIP assay, we used primer pairs which amplify 1806 bp including the tandem repeat region for PCR. From the results of this assay, we noticed that the length of PCR fragments amplified from the ChIP sample differed depending on the cell type. Amplified fragments from several cell-lines (SH-10-TC and SW480) were longer than those from a plasmid which contains 4010 bp of the *MUC5AC* promoter originally cloned from the TIG-112 (normal diploid fibroblast) genome. To verify the region responsible for the length differences, we amplified the promoter region of *MUC5AC* including the tandem repeats (−3219 to −1414 bp), and the regions directly upstream and downstream of the tandem repeats (−4010 to −3194 bp and −1433 to −1, respectively) of various cell lines (Fig. 4B). While the lengths of the regions

directly upstream and downstream of the tandem repeats were identical in all cell lines, there were length variants in the region including the tandem repeats in these cell lines. The length of this region was longest in SH-10-TC and shortest in KE-97 among the cells we used. Next we amplified the *MUC5AC* promoter region from AGS, SH-10-TC, and KE-97 genome by using primers originally designed to amplify −4010 bp of *MUC5AC* upstream of TIG-112, cloned into pGL4.12 plasmids, and performed DNA sequencing. Sequence analysis revealed that −2006 to −2713 bp of *MUC5AC* promoter region was deleted in the genome of KE-97. On the other hand, ~300-bp extension was observed in the tandem repeats of the *MUC5AC* promoter region in SH-10-TC cells (Fig. 4C). We could not read the entire sequences of the *MUC5AC* promoter in SH-10-TC, because of its length in the repetitive region and the sequencing technique limitation. The sequence of the tandem repeat region of the AGS cells was the same as reference genome hg19.

4. Discussion

MUC5AC, one of the most evident gastric marker genes, is often used for the clinical assessment of gastric cancer [15]. Our aim is to elucidate the molecular mechanism of regulation of *MUC5AC* expression in gastro-intestinal cells. In this paper, our results demonstrate that the 5′-upstream region from −1434 to −3000 bp in the *MUC5AC* promoter has the potential to act as a strong promoter. The promoter region is composed of a number of tandem repeats 5′-TCACTCAC-3′. Existence of H3K9me3 was detected by ChIP analysis. *MUC5AC* expression may be enhanced by the tandem repeat in its promoter region and the chromatin modification of the region may be regulated by H3K9me3.

To examine whether similar sequences to the tandem repeat sequences of the *MUC5AC* promoter exist in the genome, data base searches were performed. BLAST search using the tandem repeat sequences identified some similar sequences which exist in human genome, e.g. 35547 bp at 5′ side of neurogenic locus notch homolog protein 1 isoform X1, 714304 bp at 5′ side of maestro heat-like repeat family member 5. Because all of these are more than 20 kDa upstream of 5′ end of the genes, the possibility that these sequences have an impact on the downstream genes is slim.

Our ChIP analysis found that the tandem repeat region was a target of repressive mark H3K9me3, indicating that enhancer activity of the region is epigenetically regulated. H3K9me3 is known to be a marker of constitutive heterochromatin [8]. Because heterochromatin is the condensed and transcriptionally inactive state of chromatin, our results suggest that the tandem repeat region of *MUC5AC* upstream might be densely packed and be in a transcriptionally repressed state.

At present it remains unclear whether the tandem repeat region of *MUC5AC* upstream can be opened and physiologically active in cells. Tandem repeats are abundant in the human genome and have been historically ignored as nonfunctional junk DNA. However, recent accumulated evidence suggests that tandem repeats and their variations affect gene expressions and hence phenotypic variability [9]. The results of our luciferase assays clearly demonstrated that the tandem repeat region of the *MUC5AC* promoter dramatically enhanced the downstream-gene expressions. In addition, it should be noted that the degree of enhancement depends on cell type and was most obvious in AGS, a typical human gastric carcinoma cell line. Not only the number of repeat sequences, but also epigenetic regulation, such as cell-type specific transcription factors, might also play an important role in transcriptional activation. We also noticed that the length of the tandem repeat regions of the *MUC5AC* promoter differed depending on cell type (Fig. 4B–C) and that the strength of promoter activities was changed in accordance with the length of tandem repeats (Fig. 1). Interestingly, recent evidence suggests that genetic variations of the tandem repeat regions of the *MUC5AC* promoter are associated with the susceptibility and progression of gastric cancer [10]. Together, the length and regulation of the tandem repeat regions might be important

A

B
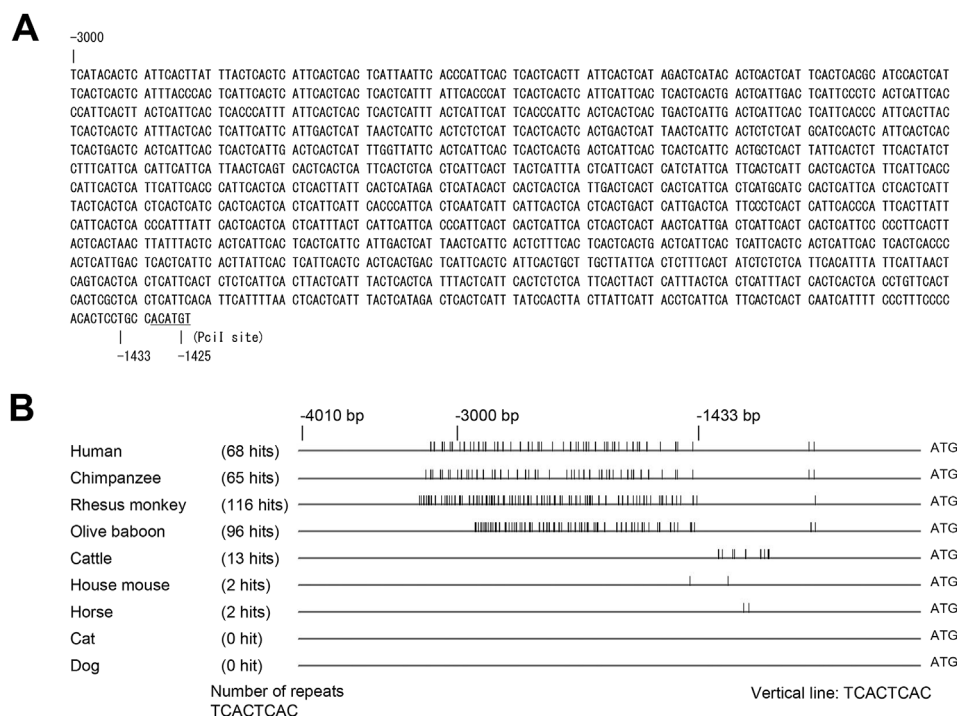
Fig. 2. Tandem repeat sequences of the *MUC5AC* promoter region. (A) Sequences of −3000 to −1424 bp upstream of the ATG start codon of human *MUC5AC*. *PciI* sites are underlined. (B) Schematic view of −4010 bp upstream of *MUC5AC* of 9 mammal species. TCACTCAC sequences were searched using pDRAW32 (AcaClone Software) and indicated in vertical lines. Number of TCACTCAC repeats was also shown.

for *MUC5AC* expression in living cells. Further examinations are required to explore pioneer factors, which force open chromatin on the repeat region.

Now we are searching for interacting proteins to the tandem repeat region of *MUC5AC* promoter using DNA affinity precipitation assay, and have identified hnRNPL as a binding protein to this region (our unpublished result). HnRNPL is known as a multifunctional RNA-binding protein containing four RNA-recognition motifs (RRMs) and to be ubiquitously expressed in all tissues [11]. Genome-wide RNA binding sites of hnRNPL were monitored by crosslinking-

immunoprecipitation (iCLIP) analysis, and sequence motif analysis revealed significant enrichment of CA-repeat and CA-rich motifs [12,13]. Although hnRNPL is known to preferentially bind to CA repeats and CA-rich sequences of RNA, little is known about its interaction to DNA. Precise experiments to address the physiological roles of hnRNPL on the tandem repeat region of the *MUC5AC* promoter is now underway.

Interestingly, the tandem repeat region of *MUC5AC* is conserved among primates, but not in other mammals such as mouse and cat. Stomach size and shape vary widely among vertebrate species to accommodate dietary variations [14] and secreted enzymes are also
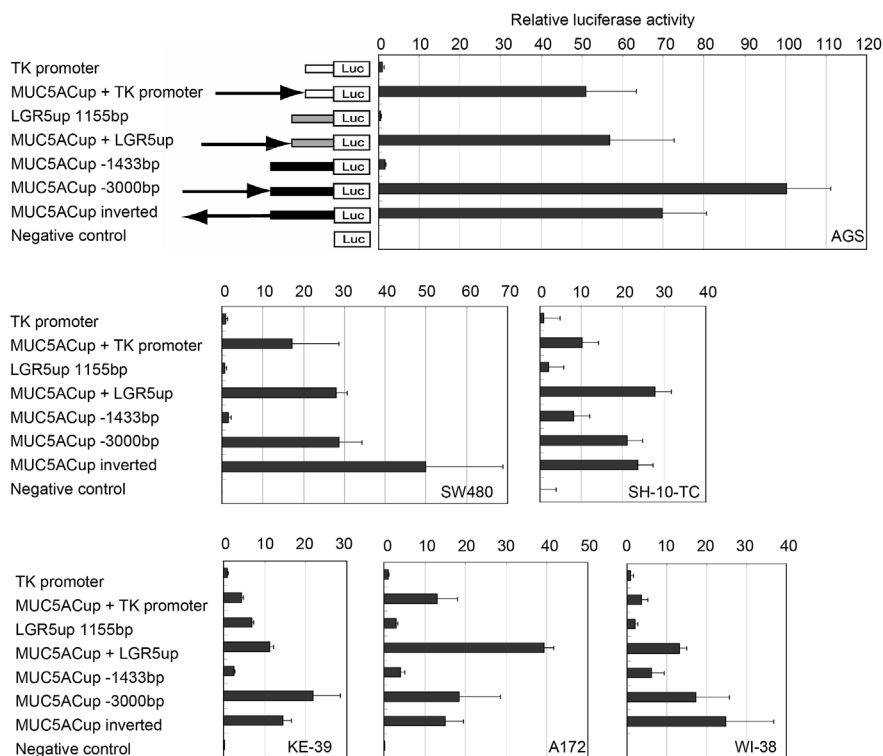
Fig. 3. Effect of tandem repeat sequences of the *MUC5AC* promoter region as an enhancer element. Representation of the various promoter constructs used in luciferase reporter assay is shown. Fragments of −3000 to −1425 bp upstream of *MUC5AC* were inserted in the upstream of TK and *LGR5* promoters (MUC5ACup + TK promoter, MUC5ACup + LGR5up). −3000 to −1425 bp upstream of *MUC5AC* were also inverted and ligated to the −1433 bp promoter region of *MUC5AC* (MUC5ACup inverted). AGS, SW480, SH-10-TC, KE-39, A172, WI-38 cells were transfected with pTK4.12 (TK promoter), pTK4.12-MUC5ACup (MUC5ACup + TK promoter), pGL4.12-LGR5up1155bp (LGR5up 1155bp), pGL4.12-MUC5ACup-LGR5up1155bp (MUC5ACup + LGR5up), pGL4.12-MUC5ACup-1433bp and −3000bp (MUC5ACup −1433bp, −3000bp), pGL4.12-inverted tandem repeats-MUC5ACup1433bp (MUC5ACup inverted), and pGL4.12 (negative control), and luciferase activities were measured at 24 h after transfection. Data represent the mean of luciferase activity (relative to TK-promoter activity) and error bars showed the standard deviation of the results from three independent experiments.
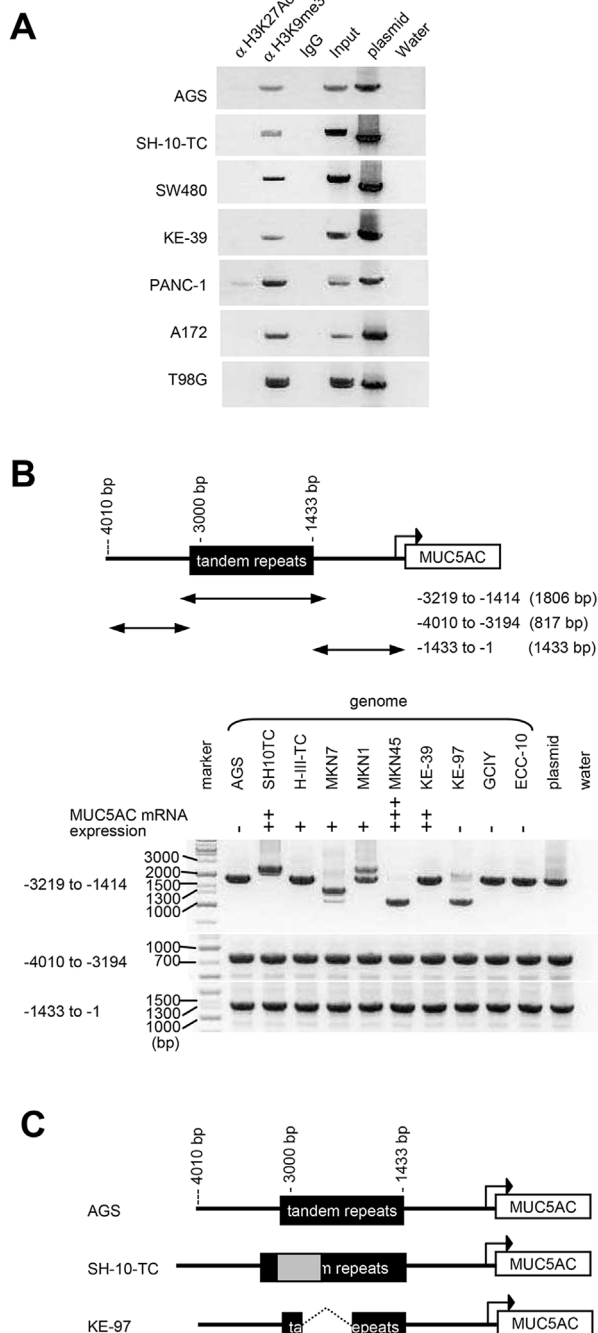
Fig. 4. ChIP analysis and length variants of tandem repeat sequences in the *MUC5AC* promoter region. (A) ChIP analysis of the tandem repeats of the *MUC5AC* promoter region. Anti-H3K27Ac, and anti-H3K9me3 antibodies and non-immunized mouse IgG were used for immunoprecipitation. pGL4.12-MUC5ACup4010bp (plasmid) was used as a control for PCR. (B) Tandem repeats of the *MUC5AC* promoter region (−3219 to −1414 bp; 1806bp) and the regions from −4010 to −3194 bp (817 bp) and from −1433 to −1 bp (1433 bp) were amplified from genomic DNA of various cell lines. mRNA expression levels of *MUC5AC* (the results of our previous studies [4]) were indicated as - no expression, + faint, ++ medium, +++ high. (C) Schematic view of representative variants of the *MUC5AC* promoter region. 4010 bp of the *MUC5AC* promoter region of AGS, SH-10-TC, and KE-97 genome were cloned and their DNA sequences were analyzed. Inserted (in SH-10-TC) and deleted (in KE-97) regions are indicated by the gray column and thin dotted line, respectively.

differed to adapt to different types and amount of food [15]. Because the high copies of tandem repeats in *MUC5AC* promoter is specific to primates, the system of the explosive *MUC5AC* expression might be necessary only in primates. An evolutionary viewpoint might be needed to elucidate primate-specific mechanism of *MUC5AC* expression.

Conflicts of interest

None of the authors have any conflicts of interest associated with this study.

Acknowledgements

Transparency document

Transparency document related to this article can be found online at https://doi.org/10.1016/j.bbrep.2019.100632

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.bbrep.2019.100632.

References

[1] M. Konno-Shimizu, N. Yamamichi, K.-i. Inada, N. Kageyama-Yahara, K. Shiogama, Y. Takahashi, I. Asada-Hirayama, M. Yamamichi-Nishina, C. Nakayama, S. Ono, S. Kodashima, M. Fujishiro, Y. Tsutsumi, M. Ichinose, K. Koike, Cathepsin E is a marker of gastric differentiation and signet-ring cell carcinoma of stomach: a novel suggestion on gastric tumorigenesis, PLoS One 8 (2013) e56766.

[2] C.A. Reis, L. David, P.A. Nielsen, H. Clausen, K. Mirgorodskaya, P. Roepstorff, M. Sobrinho-Simoes, Immunohistochemical study of MUC5AC expression in human gastric carcinomas using a novel monoclonal antibody, Int. J. Cancer 74 (1997) 112–121.

[3] K. Wakatsuki, Y. Yamada, M. Narikiyo, M. Ueno, T. Takayama, H. Tamaki, K. Miki, S. Matsumoto, K. Enomoto, T. Yokotani, Y. Nakajima, Clinicopathological and prognostic significance of mucin phenotype in gastric cancer, J. Surg. Oncol. 98 (2008) 124–129.

[4] N. Kageyama-Yahara, N. Yamamichi, Y. Takahashi, C. Nakayama, K. Shiogama, K.-i. Inada, M. Konno-Shimizu, S. Kodashima, M. Fujishiro, Y. Tsutsumi, M. Ichinose, K. Koike, Gli regulates MUC5AC transcription in human gastrointestinal cells, PLoS One 9 (2014) e106106.

[5] N. Yamamichi, M. Yamamichi-Nishina, T. Mizutani, H. Watanabe, S. Minoguchi, N. Kobayashi, S. Kimura, T. Ito, N. Yahagi, M. Ichinose, M. Omata, H. Iba, The Brm gene suppressed at the post-transcriptional level in various human cell lines is inducible by transient HDAC inhibitor treatment, which exhibits antioncogenic potential, Oncogene 24 (2005) 5471–5481.

[6] G. Benson, Tandem repeats finder: a program to analyze DNA sequences, Nucleic Acids Res. 27 (1999) 573–580.

[7] G.E. Zentner, P.C. Scacheri, The chromatin fingerprint of gene enhancer elements, J. Biol. Chem. 287 (2012) 30888–30896.

[8] B. Lehnertz, Y. Ueda, A.A.H.A. Derijck, U. Braunschweig, L. Perez-Burgos, S. Kubicek, T. Chen, E. Li, T. Jenuwein, A.H.F.M. Peters, Suv39h-Mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin, Curr. Biol. 13 (2003) 1192–1200.

[9] R. Gemayel, M.D. Vinces, M. Legendre, K.J. Verstrepen, Variable tandem repeats accelerate evolution of coding and regulatory sequences, Annu. Rev. Genet. 44 (2010) 445–477.

[10] C. Wang, J. Wang, Y. Liu, X. Guo, C. Zhang, MUC5AC upstream complex repetitive region length polymorphisms are associated with susceptibility and clinical stage of gastric cancer, PLoS One 9 (2014) e98327.

[11] S. Piñol-Roma, M.S. Swanson, J.G. Gall, G. Dreyfuss, A novel heterogeneous nuclear RNP protein with a unique distribution on nascent transcripts, J. Cell Biol. 109 (1989) 2575–2587.

[12] O. Rossbach, L.-H. Hung, E. Khrameeva, S. Schreiner, J. König, T. Curk, B. Zupan, J. Ule, M.S. Gelfand, A. Bindereif, Crosslinking-immunoprecipitation (iCLIP) analysis reveals global regulatory roles of hnRNP L, RNA Biol. 11 (2014) 146–155.

[13] J. Hui, L.-H. Hung, M. Heiner, S. Schreiner, N. Neumüller, G. Reither, S.A. Haas, A. Bindereif, Intronic CA-repeat and CA-rich elements: a new class of regulators of mammalian alternative splicing, EMBO J. 24 (2005) 1988–1998.

[14] T.-H. Kim, R.A. Shivdasani, Stomach development, stem cells and disease, Development 143 (2016) 554.

[15] J.B. Furness, J.J. Cottrell, D.M. Bravo, COMPARATIVE GUT PHYSIOLOGY SYMPOSIUM: comparative physiology of digestion 1, J. Anim. Sci. 93 (2015) 485–491.