SCIENTIFIC REPORTS

OPEN

Intestine specific regulation of pig cytidine-5′-monophospho-*N*-acetylneuraminic acid hydroxylase gene for *N*-glycolylneuraminic acid biosynthesis

Kwon-Ho Song[1], Choong-Hwan Kwak[1], Tae-Wook Chung[2], Sun-Hyung Ha[1], Jun-Young Park[1], Ki-Tae Ha [2], Seung-Hak Cho[3], Young-Choon Lee[4] & Cheorl-Ho Kim [1,5]

N-glycolylneuraminic acid (Neu5Gc), a generic form of sialic acid, is enzymatically synthesized by cytidine-5′-monophospho-*N*-acetylneuraminic acid hydroxylase (CMAH). Although expression of pig CMAH gene *pcmah* encoding CMAH has been reported to be regulated by pathogenic infection and developmental processes, little is known about the mechanisms underlying the regulation of *pcmah* gene expression. The objective of this study was to determine mechanism(s) involved in intestine specific regulation of *pcmah* gene by identifying several *cis*-acting elements and nuclear transcription factors that could directly interact with these *cis*-acting elements. We identified intestine specific promoter region (Pi) of *pcmah* gene located at upstream regions of the 5′flanking region of exon 1a and found that the promoter region is responsible for the transcriptional regulation of 5′*pcmah*-1. Based on reporter assays using serially constructed luciferase genes with each deleted promoter, we demonstrated that the Pi promoter activity was more active in intestinal IPI-2I cells than that in kidney PK15 cells, corresponding to both mRNA expression patterns in the two cell lines. In addition, we found that Sp1 transcription factor was necessary for basal activity of Pi promoter and that Ets-1 contributed to intestine-specific activity of Pi promoter. This study helps us understand transcriptional regulation of *pcmah* in the intestine of pig tissues. It also allows us to consider potential roles of Neu5Gc in interaction with environmental factors present in the intestinal tissue during pathogenic infection and developmental process.

Gastrointestinal tract lining functions as a barrier against hostile milieu such as foreign parasitic infection where epithelial cells are targets for infecting microbes and viruses[1]. Carbohydrates on the surface of intestinal epithelial cells have been implicated as the primary compounds that interact with microbes during the infection process. In particular, sialic acid (Sia) occupies the terminal position within glycan molecules and acts as the receptor for certain bacteria and viruses[2–5].

The 9-carbon structure of Sia can be found in Deuterostome lineaged animals. It is frequently modified. One of modified structures of Sia is *N*-glycolylneuraminic acid (Neu5Gc) which is one of the most common Sia derivatives[6,7]. Neu5Gc is synthesized by cytidine-5′-monophospho (CMP)-N-acetylneuraminic acid hydroxylase (CMAH) which catalytically converts sugar nucleotide CMP-Neu5Ac to CMP-Neu5Gc[8–10]. Neu5Gc is a major xeno-antigen that is expressed in pigs and apes, but not in human due to genetic loss of CMAH in human[8–10].

[1]Molecular and Cellular Glycobiology Unit, Department of Biological Science, Sungkyunkwan University, Seoburo 2066, Jangan-Gu, Suwon, Gyunggi-Do, 16419, Korea. [2]Division of Applied Medicine, School of Korean Medicine, Pusan National University, Yangsan City, Korea. [3]Korea National Institute of Health, Division of Bacterial Disease Research, 202, Osongsaengmyeong 2-ro, Osong-eup, Heungdeok-gu, Korea. [4]Department of Medicinal Biotechnology, Dong-A University, Busan, 49315, Korea. [5]Samsung Advanced Institute for Health Sciences & Technology (SAIHST), Samsung Medical Center, Seoul, 06351, Korea. Kwon-Ho Song, Choong-Hwan Kwak and Tae-Wook Chung contributed equally. Correspondence and requests for materials should be addressed to Y.-C.L. (email: yclee@dau.ac.kr) or C.-H.K. (email: chkimbio@skku.edu)

1

However, it can be introduced into the human body by Neu5Gc-containing dietary sources, mainly red meat intake, leading to the production of anti-Neu5Gc antibodies in human serum[10–12]. It has been suggested that binding of Neu5Gc to anti-Neu5Gc antibodies is associated with the occurrence of chronic inflammatory diseases such as cancer and atherosclerosis[11–15]. Neu5Gc is considered as a tumor marker in various human cancers including colon carcinoma, retinoblastoma, breast cancer, and melanoma[15–17]. Neu5Gc also plays a vital role as a non-Gal antigen in immune antigenic rejection during pig-to-human xenotransplantation[18–20]. In addition, Neu5Gc acts as a target receptor for pathogens such as *Escherichia coli* K99 and for bacterial toxins such as subtilase cytotoxin secreted by Shiga toxigenic *E. coli*[3,21].

Previous studies have shown that expression levels of Neu5Gc and CMAH gene are tissue-dependent[8,22]. In addition, the expression of CMAH gene is regulated by intracellular activation signals and disease status[23,24]. In mice, transcriptional mRNA expression of CMAH gene seems to be regulated by Gram-negative bacterial endotoxin and lipopolysaccharide (LPS)-induced mouse B-cell activation[25]. In the case of pigs, small intestinal synthesis of Neu5Gc in developing pigs is controlled by enzyme activity of hydroxylase and mRNA expression level of CMAH[24]. Although CMAH gene expression is basically regulated by development-derived factors and pathogenic infectious environments, the molecular mechanism(s) underlying the regulation of the CMAH gene, particularly in the viewpoint of directly control of Neu5Gc biosynthesis, is still unclear in any organism. In a recent report describing Sia levels in porcine milk during lactation, Neu5Ac has been found to be the major form of Sia (93–96%) while Neu5Gc (3–6%) is a minor form based on their contents irrespective of milk lactation stage[26]. However, Neu5Gc was unexpectedly the major Sia form in the porcine spleen (67–79%) and lung (36–49%) based on their contents. Skeletal muscle contained the lowest concentration of Neu5Gc among porcine tissues. In addition, all other organs have been found to contain higher levels of NeuGc than skeletal muscles during development[27].

In a previous report[22], our group has presented the complete pig CMAH gene *pcmah* with two 5′ alternative transcription variants (5′*pcmah*-1 and -2). Additionally, there are strong evidences for the existence of alternatively working promoters for *pcmah*[22]. Tissue specific expression pattern of *pcmah* suggests that its expression is regulated by alternative promoter utilization in a tissue-specific manner. Based on the dual existence of Neu5Ac and Neu5Gc in pig tissues, the biological role of alternative splicing of *pcmah* is likely to be important for their functions in endogenous and exogenous responses. Therefore, differences in Neu5Gc biosynthesis or CMAH enzyme activity in various pig tissues need to be clarified. Of the two alternative promoters of *pcmah*, promoter P2 has been demonstrated to be responsible for the expression of house-keeping gene. Hence, it was named promoter Ph[28].

The objective of this study was to explain mechanism(s) involved in intestine specific regulation of *pcmah* gene through identification of *cis*-acting elements and nuclear transcription factors. In the present study, we identified an intestine specific promoter region and named it Pi promoter of the 5′flanking region of exon 1a. We further found that this intestine specific promoter was responsible for cell specific expression pattern of 5′*pcmah*-1. We finally demonstrated that Sp1 and Ets-1 transcription factors were necessary for basal activity and intestine specific activity of the intestine specific promoter of *pcmah*, respectively. This study helps us understand transcriptional regulation of *pcmah* in different pig tissues. It may contribute to our understanding of differential expression of Neu5Gc in different tissues of nonhuman animals.

Results

Identification of intestine specific promoter (Pi) of *pcmah*.
Results obtained from promoter analysis clearly showed that pig CMAH gene *pcmah* had two different splicing forms (5′*pcmah*-1 and -2) of promoters as 5′alternative forms (Fig. 1a). It exhibited a tissue-specific pattern of expression. The 5′*pcmah*-1 splicing form was mainly expressed in the intestinal tissue while the 5′*pcmah*-2 form was ubiquitously expressed in most tissues[28]. In order to investigate whether tissue specific expression of 5′*pcmah*-1 was relevant to cell type derived from different tissues or organs, we analyzed mRNA levels of 5′*pcmah*-1 in pig kidney-derived PK15 cells, pig small intestine-derived IPI-2I cells, and pig aorta-derived MYP30 cells. 5′*pcmah*-1 was specifically expressed in IPI-2I cells, but not in PK15 or MYP30 cells (Fig. 1b). This result corresponds with a previous result that shows tissue specific expression pattern of 5′*pcmah-1*[28], suggesting that 5′*pcmah*-1 is an intestine specific splicing variant. In order to investigate the molecular mechanisms involved in 5′*pcmah*-1 regulation, we isolated the 5′flanking region of exon 1a of *pcmah* and named it promoter Pi (intestine specific promoter) (Fig. 1a,c). To analyze Pi promoter, three fragments (1600, 1100, and 700) were recombinantly inserted into pGL3 basic vector and then transfected into PK15, IPI-2I, and MYP30 cells, respectively. The overall activity of Pi promoter in IPI-2I cells was higher than that in PK15 cells or MYP30 cells (Fig. 1c). These results also corresponded with cell-specific expression patterns of these two 5′*pcmah*-1 and -2 transcripts (Fig. 1b), suggesting that the Pi promoter could confer intestine-specific activity for expression regulation of 5′*pcmah*-1.

Analysis of regulatory elements and characterization of Pi promoter.
Because Pi-700 fragment had the highest activity in IPI-2I cells, it was further analyzed for promoter characterization in IPI-2I cells. This promoter lacked TATA and CCAAT boxes. Instead, several polyA repeat sequences were present in Pi promoter (Fig. 2a). We then identified putative binding sites for CRE-BP, AP-4, HSF-2, Ets-1, NRF-2, YY1, and Sp1 in Pi promoter region of *pcmah* (Fig. 2a). In order to elucidate the functional role of putative transcription factor binding sites in Pi promoter activity, serially constructed 5′-deletion mutants Pi-542, Pi-260, and Pi-233, in addition to Pi-700, were generated and analyzed for their promoter activities in IPI-2I cells. The deletion of 282 and 27 nucleotides from Pi-542 and Pi-260, respectively, decreased Pi promoter activity, suggesting that these regions might contain positive regulatory elements required for Pi promoter activity (Fig. 2b). When these regions were assessed as transcription factor candidates, putative transcription factor elements including HSF2, Ets-1, NRF2, YY1, and Sp1 were found (Fig. 2b).
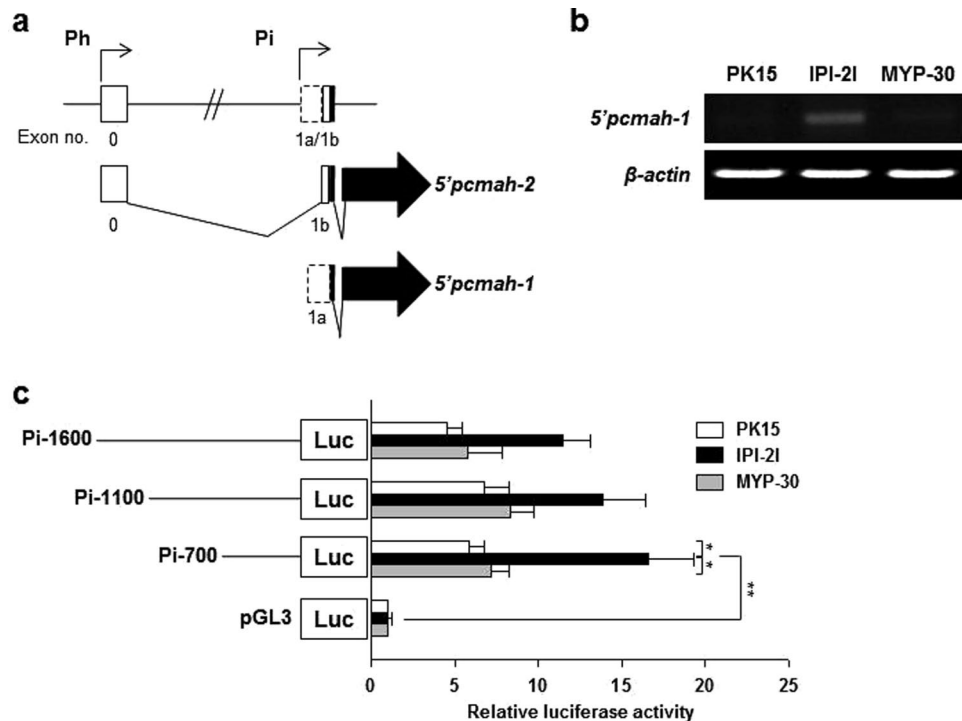
Figure 1. Identification of intestine specific promoter (Pi) of the *pcmah* gene. (**a**) Genomic structure of *pcmah*. Ubiquitous housekeeping promoter region of *pcmah* plays a role in the expression of 5′*pcmah*-2. Putative promoter region of the intestine specific splicing variant of *pcmah* (5′*pcmah*-1) is indicated by 'Pi (intestine specific promoter)'. Shaded boxes indicate the coding exons, while open boxes indicate untranslated exons. Two splicing variants of *pcmah* share a common ORF region (shaded arrow boxes). (**b**) The relative levels of 5′*pcmah*-1 expression were measured and analyzed by RT-PCR in three different cells of PK15, IPI-2I, and MYP30. β-Actin was used as a positive control. (**c**) The relative reporter activities of the luciferase reporter plasmids with the proximal region of exon 1a in PK15, IPI-2I, and MYP30 cells. DNA fragments containing various lengths of the Pi promoter region, which is the proximal region of exon 1a, were subcloned into a PGL3-basic vector (pGL3), and each generated plasmid was transiently transfected into the three different pig cells of PK15, IPI-2I, and MYP30. Two enzyme activities of luciferase and β-Gal in each of the transfected cells were assayed as described in the 'Materials and Methods' section. In each transfected cell, the luciferase enzyme activity was normalized with the measured β-Gal activity. The relative fold value was then calculated from the ratio of the normalized activity and the activity in the empty pGL3-basic vector-transfected cells. Statistic bars represent the mean ± SE of three independent determinations. Differences in the fold value of luciferase enzyme activities were statistically analyzed using the Student *t*-test: *$p < 0.05$; **$p < 0.01$.

Transcription factor Sp1 is crucial for basal activity of Pi promoter.

Previously, we have demonstrated that transcription factor Sp1 can specifically bind to its binding sites on housekeeping promoter (Ph) regions of *pcmah*. In addition, it positively regulates *pcmah* promoter activity in PK15 cells[28]. Interestingly, putative Sp1 binding sites overlapped by two Sp1 binding sites in Pi-260 fragment (Fig. 3a). We questioned whether Sp1 binding sites might be responsible for the basal activity of Pi promoter. In order to address this possibility, these putative Sp1 binding sites were experimentally altered by site-directed mutagenesis. Using altered clones, the effect of each mutation on promoter activity was examined in IPI-2I cells. In Pi promoter, luciferase activities of Pi-260 Sp1a Mut, Pi-260 Sp1b Mut, and Pi-260 Sp1a/b Mut (mutated in both Sp1a and Sp1b sites) were reduced by approximately 58% (9.2 to 3.9 folds), 62% (9.2 to 3.5 folds), and 84% (9 to 1.5 folds), respectively, compared to luciferase activity of wild type Pi-260 plasmid (Fig. 3b). Next, in order to confirm whether Sp1 could bind to Pi promoter *in vitro*, gel mobility shift assay was performed. Nuclear protein from IPI-2I cells was able to be bind to ³²-P labeled oligonucleotide probe. Pi-Sp1 contained two overlapping putative Sp1 binding sites of the Pi promoter region (Fig. 3c). The specificity of binding was confirmed under experimental condition of competition assay which used 100-fold molar excess of each unlabeled oligonucleotide probe and each unlabeled mutant probe that carried point mutation in the Sp1 consensus sequence of the region. As expected, results clearly showed that the unlabeled probe efficiently competed for the binding of the transcription factor to labeled oligonucleotide probe. In contrast, unlabeled mutant probe could not compete for the binding of the transcription factor (Fig. 3c). To further clarify whether the bound complex could be linked to transcription factor Sp1 protein, a super-shift assay was performed in the presence of anti-Sp1 antibody. Super-shift band was not detected in the assay. However, the binding intensity of the major Sp1 complexed form was dramatically decreased by treatment with anti-Sp1 specific antibody, indicating competitively specific binding by Sp1 (Fig. 3d). To demonstrate that transcription factor Sp1 could directly regulate Pi promoter activity of the gene, effects of mithramycin A, an Sp1
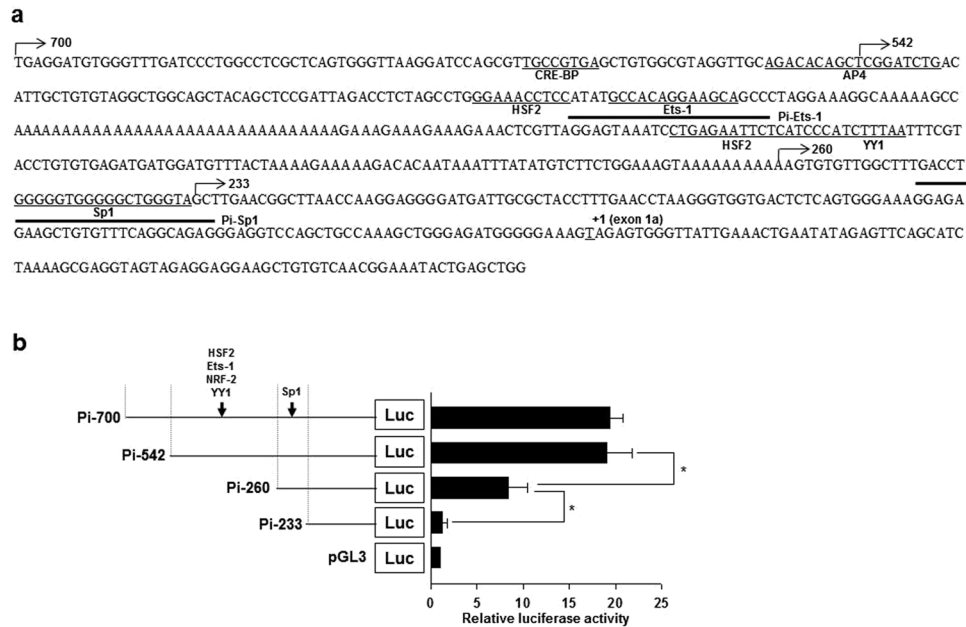
Figure 2. Detailed characterization and analysis of the Pi promoter region. (**a**) Analysis of nucleotide sequences of the Pi promoter region. For the transcription factors to interact, putative binding sites as DNA sequences are underlined. The end point of 5′deletion mutants is indicated by arrows. The position '+1' indicates the transcription start site of 5′pcmah-1. Thick underlines indicate the oligonucleotide DNA sequences for the EMSA experiment. (**b**) The 5′deletion analysis of the Pi promoter region of pcmah in IPI-2I cells. Transcription factors, which exist in the region and are indicated by black arrows, are represented. Statistic bars represent the mean ± SE obtained from three independent determinations. Differences in the fold value of luciferase enzyme activities were statistically analyzed using the Student t-test: $*p < 0.05$.

inhibitor, on Pi promoter was examined. As expected, results showed that treatment with increasing concentrations (25 nM to 100 nM) of mithramycin A decreased luciferase activity of Pi-260 in a dose-dependent manner (Fig. 3e). These results clearly suggest that Sp1 can bind to its putative *cis*-binding sites on the Pi promoter region of *pcmah* and that it is crucial for basic transcriptional regulation of Pi promoter activity.

Ets-1 is essential for intestine specific activity of Pi promoter.

In addition to Sp1, genetic deletion of the region containing putative binding sites for transcription factor HSF2, Ets-1, NRF2, and YY1 markedly decreased the luciferase activity of Pi promoter (Fig. 2b). Among them, Ets-1 transcription factors are widely expressed in developing and mature intestine[29]. For further characterization of possible intestine-specific *cis*-elements in Pi promoter, we investigated the role of putative Ets-1 binding sites in Pi promoter. Interestingly, mutation of Ets-1 binding site on Pi-542 significantly decreased the activity of Pi-542 in intestine IPI-2I cells, but not in kidney PK15 cells (Fig. 4a,b), suggesting that the putative Ets-1 binding site might act as a *cis*-acting element to confer intestine specific activity of the Pi promoter. To determine whether Ets-1 could directly interact with putative Ets-1 binding sites in the Pi promoter, electrophoretic mobility shift assay (EMSA) was performed. Nuclear protein isolated from IPI-2I cells was able to bind to [32]-P labeled Pi-Ets-1oligonucleotide probe (Fig. 4c). In addition, unlabeled probe was efficiently competitive for binding of Ets-1 to the labeled Pi-Ets-1 probe while no competition was observed for unlabeled mutant probe (Fig. 4c). A super-shift assay further verified the specificity of Ets-1 and DNA interaction (Fig. 4d). These results indicate that Ets-1 can specifically bind to its putative binding sites on the Pi promoter region of *pcmah*.

We reasoned that the dependency difference in Ets-1 binding site on IPI-2I and PK15 cells might be due to the spontaneous expression level of Ets-1. Thus, we compared protein levels of Ets-1 in IPI-2I and PK-15 cells. Indeed, Ets-1 protein level in IPI-2I cells was higher than that in PK15 cells (Fig. 4e). This suggests that Ets-1 is a *trans*-acting element that confers intestine-specificity of *pcmah* to IPI-2I cells. To further investigate the role of EtS-1 in *pcmah* expression, we transfected siRNAs targeting Ets-1 into IPI-2I cells. Notably, knock-down of Ets-1 decreased 5′*pcmah*-1 expression as well as luciferase activity of Pi-542 in IPI-2I cells (Fig. 4f,g). Furthermore, the amount of NeuGc was decreased in Ets-1 siRNA transfected IPI-2I cells as compared to that in NC siRNA-transfected cells (Fig. 4h). These results suggest that Ets-1 is essential for intestine specific expression of *pcmah*.

Discussion

During the course of evolution from their ancestors, mice and pigs shared a last common ancestor over 98 million years ago, with a large difference in the expression of a specific sialic acid called Neu5Gc. Neu5GC is widely expressed on most mammalian tissues, but not on human cells[30]. Apart from humans, recent studies have shown that several other groups of mammals have also lost the activity of CMAH. Mustelids (ferrets and
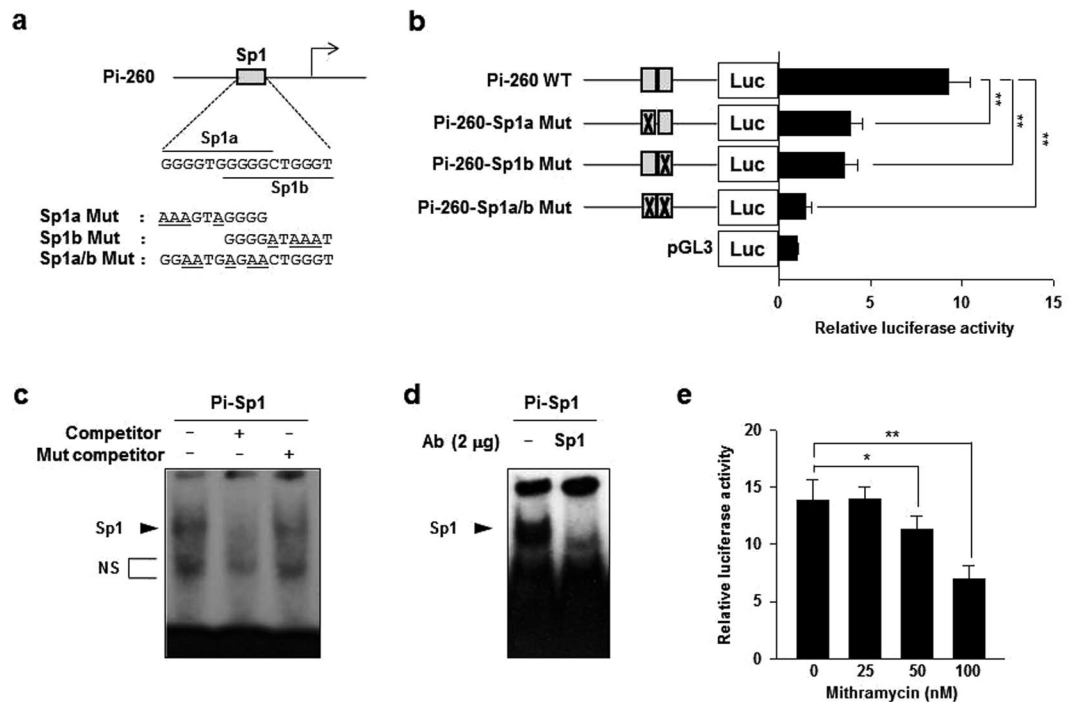
Figure 3. Transcription factor Sp1 is crucial for the basal activity of the Pi promoter. (**a**) Upper panel: nucleotide sequence of putative Sp1 binding sites (named here as Sp1a (upper lined) and Sp1b (under lined)) in Pi-175. Under panel: mutated bases in three Sp1 mutants are underlined. (**b**) Luciferase enzyme activities of three site-directed mutants and wild type Pi-175 in IPI-2I cells. Gray boxes and crossed gray boxes indicate the wild type or mutation of the putative Sp1 binding site(s), respectively. (**c**,**d**) Nuclear extracts of IPI-2I cells were incubated with ^{32}P labeled oligonucleotide DNA probes (Pi-Sp1) containing two overlapping Sp1 binding sites on the Pi promoter. The reaction mixture was then examined by EMSA as described in the 'Materials and Methods' section. (**c**) Competition binding assay was performed using a non-labeled probe DNA (competitor) or a non-labeled mutant DNA probe (Mut competitor) for the Sp1 binding site of the Pi promoter. (**d**) Super-shift assay using anti-Sp1 antibodies was then performed. The arrows indicate the specific binding complex for Sp1. 'NS' refers to non-specific bands. (**e**) The effect of Sp1 inhibitor on the P1 promoter activity. P1-175 plasmid was transiently transfected into IPI-2I cells and various concentrations (25, 50, and 100 nM) of mithramycin A were added to the cells for 20 h. Bars represent the mean $\pm$ SE obtained from three independent determinations. Differences in the fold value of luciferase enzyme activities were statistically analyzed using the Student t-test: *$p < 0.05$; **$p < 0.01$.

relatives) and pinnipeds (seals, sea lions and walruses) lack an active CMAH. Consequently, they do not produce Neu5Gc[31], although they belong to highly social species with very high chance of microbial infections. Ferret CMAH gene has been suggested to be deleted by an ancient mutation that is also shared by other members of the order Carnivora[31]. The exclusive expression of Neu5Ac in ferrets may contribute to their suitability as a model for human-adapted infectious diseases. Similarly, a CMAH mutation has also been known to occur in new world monkeys, differentiating them from old world monkey species which express Neu5Gc[32]. Compared to macaques (an old-world monkey species), marmosets (a new-world monkey species) have also been shown to be suitable for human-adapted transmission studies because of their small sizes requiring low doses for potential vaccine and drug studies and expression of symptoms similar to those seen in influenza virus-infected humans[33].

Previously, many reports have demonstrated that expression levels of Neu5Gc and CMAH are regulated during infection by intestinal parasites and during developmental process[2–5,34]. In rats, Neu5Gc level is down-regulated by a decrease in CMAH mRNA expression during infection with *Nippostrongylus brasiliensis*[1]. In pigs, ganglioside GM3 (Neu5Gc) or Neu5Gc content is maximal at birth and gradually decreased in adults. This may explain the susceptibility of newborn piglets to pig enteric pathogens such as *E. coli* K99[4,21,35,36]. These observations suggest that Neu5Gc production can be regulated at transcriptional level during infection and developmental processes. However, little is known about the mechanisms underlying the regulation of CMAH gene. Since intestine tissues are important for selective expression of this gene in response to pathogenic infection in pigs, intestine specific regulation of the promoter is in the best interest of pig Neu5Gc biosynthesis. In the present study, we elucidated the mechanism(s) involved in intestine specific regulation of the pig CMAH gene by identifying *cis*-acting elements and their interacting transcription factors crucial for *pcmah* intestinal promoter activity. Such mechanisms for *pcmah* transcriptional regulation may give us insight into the mechanism(s) of Neu5Gc regulation during infectious and developmental processes.
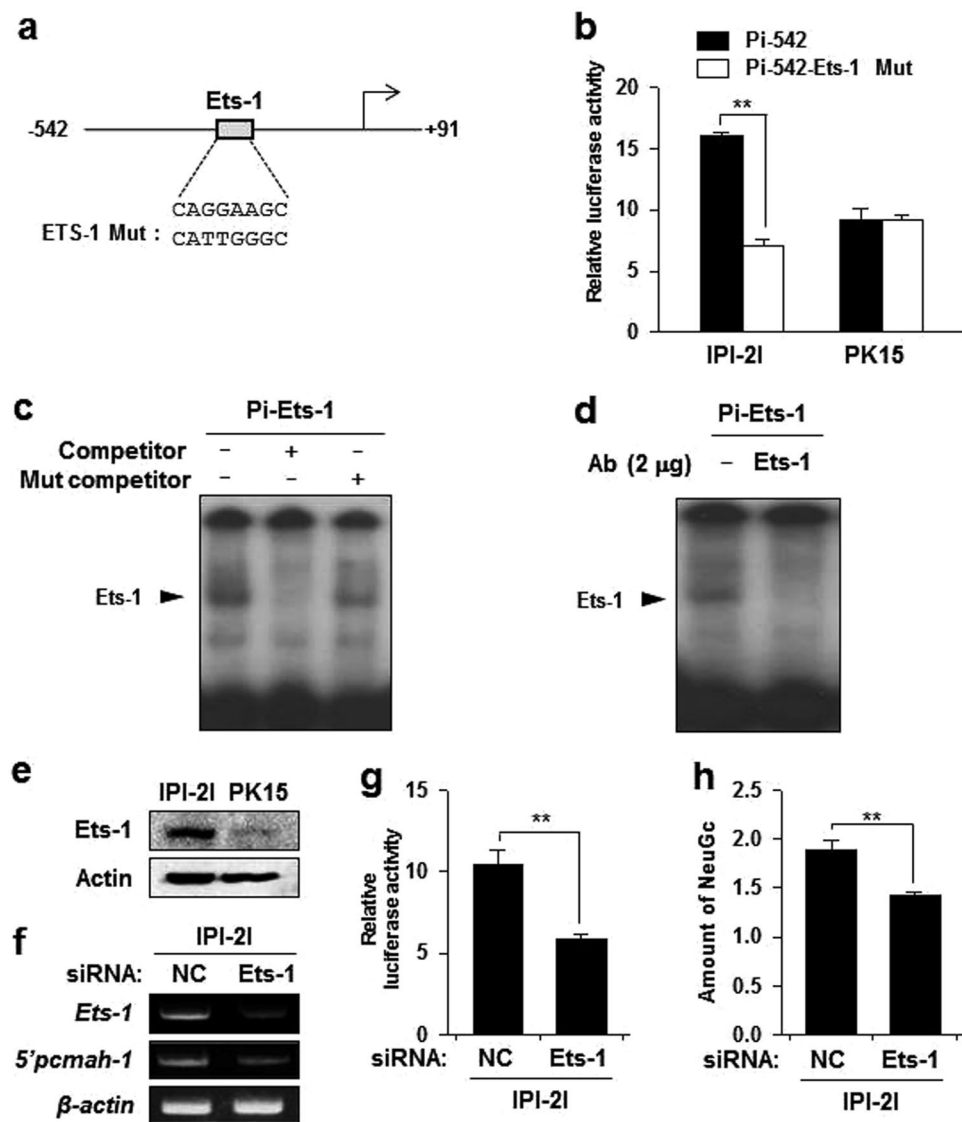
Figure 4. Ets-1 is essential for intestine specific activity of the Pi promoter. (**a**) Upper panel: nucleotide sequence of the putative Ets-1 binding site in Pi-457. Under panel; mutated bases in the Ets-1 mutant are underlined. (**b**) Luciferase activities of Ets-1 mutant and wild type of Pi-457 in IPI-2I and PK15 cells. (**c,d**) Nuclear extracts of IPI-2I cells were incubated with ^{32}P labeled oligonucleotide probes (Pi-Ets-1). The mixture was subsequently examined by EMSA as described in the 'Materials and Methods' section. (**c**) Competition binding assay was performed using a non-labeled probe (competitor) or a non-labeled mutant probe (Mut competitor) for Ets-1 binding site of the Pi promoter. (**d**) Super-shift assay using anti-Ets-1 antibodies was performed. The arrows indicate the specific binding complex for Ets-1. (**e**) Expression level of Ets-1 in IPI-2I and PK-15 cells was analyzed by immunoblotting assay. Actin was included as an internal loading control. (**f**) IPI-2I cells were transfected with negative control (NC) and siRNAs targeting Ets-1. The gene expression levels of Ets-1 and 5′pcmah-1 were analyzed by RT-PCR in IPI-2I cells. (**g**) Luciferase activities of Pi-457 in the siRNA transfected IPI-2I and PK15 cells. Bars represent the mean ± SE obtained from three independent determinations. (**h**) The level of NeuGc in Ets-1 siRNA-transfected IPI-2I cells was measured by ELISA assay. Differences in the fold value of optical density were statistically analyzed using the Student t-test: **$p < 0.01$.

In our previous studies[22,28], we have demonstrated that *pcmah* has two distinct 5′alternative splicing forms, namely 5′*pcmah*-1 and -2. We have also shown that *pcmah* expression has a tissue specific pattern such that 5′*pcmah*-1 is mainly expressed in the intestine while 5′*pcmah*-2 is rather ubiquitously expressed in most tissues[22,28]. The two alternative mRNA forms specifically expressed in the tissues are strictly corresponded with those observed in the PK15 and IPI-2I cells, which were derived from pig kidney tissue and small intestine tissue, respectively[22,28]. Tissue specific gene expression characterized by transcripts incorporating alternative 5′ ends is known to be under the regulation of multiple promoters[37–39]. In this study, we isolated the intestine-specific promoter (Pi) of *pcmah*. Based on luciferase reporter assay, we demonstrated that the Pi promoter, relative to the upstream region of 5′*pcmah*-1, was mainly active in IPI-2I cells (Fig. 2). These results corresponded with mRNA
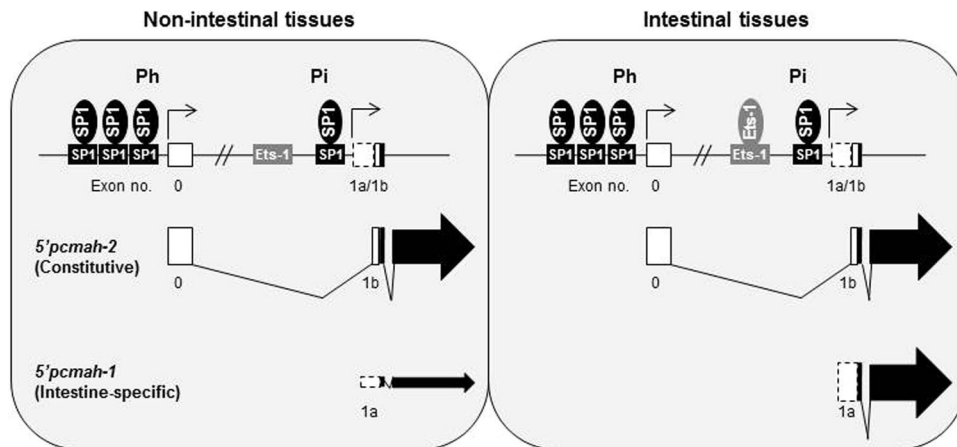
Figure 5. Schematic illustration of the intestine specific promoter (Pi) 5′*pcmah-1* and the house-keeping (constitutive) promoter (Ph) 5′*pcmah*-2 of the *pcmah* gene. The housekeeping promoter region of *pcmah*, starting at the 5′ flanking region including 3 independent SP1 of Exon 0, regulates the constitutive expression of 5′*pcmah*-2. Intestine specific splicing variant of *pcmah* (5′*pcmah*-1) is stringently controlled by the intestine specific promoter P1, starting at the 5′ flanking region, including Ets-1 and SP1 of Exon 1a.

expression patterns of 5′*pcmah*-1 shown in Fig. 1b, suggesting that these alternative promoters might confer differences in the regulation of *pcmah* transcription in pig tissues. However, PK15 cells have only basal activity of the Pi promoter and expression of *pcmah*-1 mRNA by promoter Pi was scarcely detected. Here, we used IPI-2I cells to further investigate elements related to Pi promoter activity. Based on the present results, transcription factor Sp1 can specifically bind to its putative binding sites on Pi promoter regions of *pcmah* and positively upregulate promoter activity in cells. Although it is clear that Sp1 is essential for the basal activity of Pi promoter, more complex regulatory element(s) are required for its full activity.

In fact, serial 5′deletion analysis of Pi promoter region suggested that Ets-1 might additionally act as a positive regulatory element with Sp1 as evidenced by the fact that genetic deletion of DNA regions containing putative Ets-1 binding sites remarkably reduced reporter luciferase activity of Pi promoter. Mutation of Ets-1 *cis*-element or knockdown of the Ets-1 trans-element significantly disrupted P1 promoter activity in IPI-2I cells, but not in PK15 cells (Fig. 4b,g). Interestingly, it has been reported that Ets transcription factors are widely expressed in developing and mature intestines losely related to a certain inflammatory disease[29,40,41]. These results demonstrate that Ets-1 plays an important role in the activity of Pi promoter. Thus, it may contribute to the intestine-specific expression of *pcmah*.

In conclusion, as schematically illustrated in Fig. 5, two 5′alternative splicing forms of pig CMAH gene, 5′*pcmah*-1 and -2, are controlled by Pi (intestine specific) and Ph (house-keeping) promoters, respectively. Furthermore, each splicing variant is stringently regulated by different transcriptional *cis*- and *trans*-acting elements on each promoter region. Sp1 transcription factor binds to its specific binding sites on Pi and Ph promoter regions and positively upregulates the promoter activity in pig cells derived from tissues. On the other hand, Ets-1 transcription factor which is abundant in intestinal tissues is important for the activity of Pi promoter and expression of 5′*pcmah*-1, indicating that it may contribute to intestine-specific expression of *pcmah*. These findings provide enhanced insight into molecular regulation of CMAH gene expression at transcriptional level as well as new insights into the regulatory mechanisms of Neu5Gc biosynthesis during pathogenic infectious events and embryonic developmental stages.

Materials and Methods

Cell culture. Dulbecco′s modified Eagle's medium (DMEM; WelGENE, Daegu, Korea) was used for culturing PK15 pig kidney cells. IPI-2I cells derived from pig small intestinal tissue were cultured in DMEM containing 0.024 IU/ml insulin and 4 mM glutamine as additives. MYP30 pig endothelial cells were cultured in DMEM containing 0.1 mMol/L nonessential amino acids. All culture media were supplemented with 10% heat-inactivated fetal bovine serum (FBS; WelGENE, Korea), 100 unit/ml penicillin, and 100 μg/ml streptomycin. Cell cultures were routinely maintained at 37 °C with 5% CO_2 atmosphere.

RT-PCR. Total RNAs were isolated using commercially available TRIZOL reagent (Invitrogen, USA) and cDNA synthesis was performed using AccuPower® RT-PreMix kit (Bioneer, Daejon, Korea) according to the manufacturer's protocol. PCR was performed with EF-Taq polymerase (SolGent, Korea) using the following specific primers: 5′*pcmah*-1, 5′-GTCAACGGAAATACTGAGCTGGGT-3′ (forward) and 5′-TCGTCTTGACAGAAGCTTCCAGGA-3′ (reverse); Ets-1, 5′-AGAGTCCCGCTATACCTCGGA-3′ (forward) and 5′-AGGCTTGTCCTTGTTGAGGTC-3′ (reverse); β-*actin*, 5′-CACGCCATCCTGCG TCTGGA-3′ (forward) and 5′-TCTGCATCCTGTCGGCGATG-3′ (reverse). For the generalization of obtained data, equal amounts of mRNA were used. Expression levels of β-*actin* as internal control were analyzed and confirmed.

Primers for 5′ Deletion mutants[a]	
Pi-1600(Bgl II-F)	AC<u>AGATCT</u>TGAAGGACAGCTTTGGCTCT
Pi-1100(Bgl II-F)	AC<u>AGATCT</u>AAGGATAACCTGACCCCTTTGC
Pi-700(Bgl II-F)	AC<u>AGATCT</u>TGAGGATGTGGGTTTGATCC
Pi-542(Bgl II-F)	AC<u>AGATCT</u>TCGGATCTGACATTGCTG
Pi-260(Bgl II-F)	AC<u>AGATCT</u>AAGTGTGTTGGCTTTGACCTG
Pi-223(Bgl II-F)	AC<u>AGATCT</u>AGCTTGAACGGCTTAACCAAG
Pi(Hind III-R)	AC<u>AAGCTT</u>CAGTATTTCCGTTGACACAG
Primers for Direct mutagenesis[b]	
Pi-SP1a MutF	TTGACCTGaaaGTaGGGGCTGGGTAGCTTGAAC
Pi-SP1a MutR	TACCCAGCCCCtACtttCAGGTCAAAGCCAACA
Pi-SP1b MutF	GGTGGGGGaTaaaTAGCTTGAACGGCTTAACC
Pi-SP1b MutR	TCAAGCTAtttAtCCCCCACCCCCAGGTCAAAG
Pi-SP1ab MutF	TTGACCTGGGaaTGaGaaCTGGGTAGCTTGAACG
Pi-SP1ab MutR	CTACCCAGttCtCAttCCCAGGTCAAAGCCAAC
Pi-Ets-1 MutF	TATGCCACAttggGCAGCCCT
Pi-Ets-1 MutR	AGGGCTGCccaaTGTGGCATA
Probes for EMSA[b]	
Pi-SP1	ACCTGGGGGTGGGGGCTGGGTAGCTT
Pi-SP1 Mut	ACCTGGGGGTGGGGGaTaaaTAGCTT
Pi-Ets-1	TATGCCACAGGAAGCAGCCCT
Pi-Ets-1 Mut	TATGCCACAttggGCAGCCCT

Table 1. DNA sequences of primers and probes used in this study. [a]Restriction enzyme cleavage sites were underlined. [b]Lowercase letters indicate mutated nucleotides sequences.

DNA constructs. Fragments of different lengths in the 5′-flanking region of *pcmah* were amplified by PCR using pig BAC clone library which contained *pcmah* genome or cloned 5′-flanking DNAs as template with primer sets. Primer sequences are listed in Table 1. Obtained PCR products were digested using *Bgl* II and *Hind* III. Digested PCR products were then ligated with *Bgl* II and *Hind* III restriction sites of pGL3-Basic Vector as a vehicle (Promega, USA).

siRNA constructs. Synthetic small interfering RNA (siRNA) specific for Ets-1 were purchased from Genolution (Korea): Negative control (NC), 5′-CCUCGUGCCGUUCCAUCAGGUAGUU-3′ (sense), and 5′-CUACCUGAUGGAACGGCACGAGGUU-3′ (antisense); Ets-1, 5′-GGAAUUACUCACUGA UAAAUU-3′ (sense) and 5′-UUUAUCAGUGAGUAAUUCCUU-3′ (antisense). siRNA was delivered into 6-well plates at a dose of 100 pmol per well using commercially available WelFect-EXTM PLUS Transfection Reagent (WelGENE, Daegu, Korea).

Site directed mutagenesis. Site directed DNA mutagenesis was performed using a QuickChange XL Site-directed Mutagenesis kit (Agilent, USA) according to the manufacturer's instructions. Primer sequences are listed in Table 1. These primers were used to introduce desired DNA mutations to constructs of *pcmah* promoter. Multiple transcription factor binding sites were subjected to mutation from single mutation constructs by repeated mutagenesis with a different primer set. These generated mutation constructs were confirmed by DNA sequencing analysis.

Luciferase assay. Cells were cultured after they were seeded onto 12-well culture plates one day prior to luciferase assay. Each cell was co-transfected with each luciferase reporter construct and β-galactosidase reporter plasmid using WelFect-EXTM PLUS Transfection Reagent (WelGENE, Korea). After 24 h of culturing, cells were harvested and subjected to cell lysis with 1 X Passive lysis buffer (Promega, USA). Enzyme activities of Luciferase and β-galactosidase were measured using luciferase and β-galactosidase activity assay system (Promega, USA). Luciferase enzyme activity was then normalized against β-galactosidase enzyme activity in cell lysate and calculated as an average of three sets of independently performed experiments.

EMSA. For EMSA, single-stranded oligonucleotides were commercially synthesized by IDT DNA (IDT, USA). DNA sequences of these synthesized probes are listed in Table 1. To generate double stranded DNA probes, complementary oligonucleotides were mixed together at a 1:1 molar ratio and subjected to heating at 95 °C for 2 min followed by cooling down to 25 °C for 45 min. To obtain nuclear protein of each cell, each nuclear extract was prepared and isolated as previously described[42]. EMSA experiment was performed using a gel shift assay system kit (Promega, USA) according to the supplier's recommended protocol and instructions. Briefly, 1.75 pmol/μl of double-stranded oligonucleotide probe was end-labeled with [γ-^{32}P]ATP (3000 Ci/mmol; BMS, USA) using T4 polynucleotide kinase (Promega, USA). For protein-DNA binding, reaction mixture contained 2 μg of nuclear extract and 2 μl of gel shift binding buffer [4% glycerol, 1 mM MgCl$_2$, 0.5 mM EDTA, 0.5 mM DTT, 50 mM NaCl, 10 mM Tris-HCl (pH 7.5), and 0.05 mg/ml poly (deoxyinosine-deoxycytosine)] with or without unlabeled wild type (wt) or mutant double-stranded oligonucleotide for binding competition. After preincubating at room

temperature for 10 min, the reaction mixture was further incubated with labeled probe for 20 min at room temperature. For antibody competition assay, the nuclear extract was incubated with 2 μg of anti-Sp1 (Santa Cruz, USA) or Ets-1 antibody (Santa Cruz, USA) for 8 h at room temperature or overnight at 4 °C in the presence of binding buffer prior to addition of the labeled probe. The incubated reaction mixture was then separated by electrophoresis gel using 4% nondenaturing polyacrylamide in 0.5 × Tris-borate EDTA buffer at 250 V for 30 min. The separated polyacrylamide gel was dried for detection by autoradiography.

Immunoblot Analysis.
IPI-2I and PK-15 cells were rinsed twice with PBS and lysed with RIPA lysis buffer containing 20 mM Tris/HCl pH 7.5, 150 mM NaCl, 1% Triton X-100, 2 mM EDTA, 10% glycerol, 0.1% SDS, and 0.5% sodium deoxycholate with protease inhibitor cocktail (1 mM Na_3VO_4, 20 μg/mL PMSF, 10 μg/μL leupeptin, and 50 mM NaF). Lysates were centrifuged at 13,000 rpm for 10 min. Concentrations of whole cell lysates were measured using Lowry assay (Bio-Rad, Hercules, CA, USA). Then 25 μg of total proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred onto nitrocellulose membranes using a Hoefer electrotransfer system (Amersham Biosciences, Amersham, UK). These transferred nitrocellulose membranes were incubated with anti-Ets1 and anti-β-actin antibodies (1:1000 dilution in 1% skim milk) at 4 °C overnight followed by incubation with a secondary antibody (1:5000 dilution in 1% skim milk) for 1 h at room temperature. Bands were visualized by enhanced chemiluminescence (Sigma-Aldrich) and analyzed using ChemiDoc (Davinchi-K, Seoul, Korea). Polyclonal rabbit anti Ets-1 and polyclonal mouse anti-actin antibodies were purchased from Santa Cruz Biotechnology (Santa Cruz, CA, USA). Secondary antibodies were horseradish peroxidase conjugated anti-mouse IgG and anti-rabbit IgG (Thermo Scientific Pierce, Rockford, IL, USA).

Enzyme linked immunosorbent assay (ELISA).
ELISA was performed using an Anti-Neu5Gc antibody Kit (BioLegend, San Diego, CA, USA), according to the supplier's recommended instruction. Whole cell lysates were coated onto ELISA plates overnight at 4 °C. Wells were then washed with PBS. After blocking, the plate was incubated with chicken polyclonal IgY anti-Neu5Gc antibody (1:1000) for three hours at ambient temperature. The plate was then incubated with *Horseradish* peroxidase conjugated anti-chicken antibody (1:5000) for 1 h at ambient temperature. Developing reagent, 3,3′,5,5′-tetramethylbenzidine, was then added to the plate and optical density value was measured at 450 nm.

References

1. Karlsson, N. G. *et al.* Identification of transient glycosylation alterations of sialylated mucin oligosaccharides during infection by the rat intestinal parasite *Nippostrongylus brasiliensis*. *Biochem. J.* **350**(Pt 3), 805–814 (2000).
2. Hanisch, F. G., Hacker, J. & Schroten, H. Specificity of S fimbriae on recombinant *Escherichia coli*: preferential binding to gangliosides expressing NeuGc alpha (2-3)Gal and NeuAc alpha (2-8)NeuAc. *Infect. Immun.* **61**, 2108–2115 (1993).
3. Byres, E. *et al.* Incorporation of a non-human glycan mediates human susceptibility to a bacterial toxin. *Nature* **456**, 648–652 (2008).
4. Jay, C. M., Bhaskaran, S., Rathore, K. S. & Waghela, S. D. Enterotoxigenic K99 + *Escherichia coli* attachment to host cell receptors inhibited by recombinant pili protein. *Vet. Microbiol.* **101**, 153–160 (2004).
5. Ito, T. *et al.* Recognition of N-glycolylneuraminic acid linked to galactose by thealpha2,3 linkage is associated with intestinal replication of influenza A virus in ducks. *J. Virol.* **74**, 9300–9305 (2000).
6. Schauer, R. *et al.* Functional and biosynthetic aspects of sialic acid diversity. *Indian J. Biochem. Biophys.* **34**, 131–141 (1997).
7. Chou, H. H. *et al.* A mutation in human CMP-sialic acid hydroxylase occurred after the Homo-Pan divergence. *Proc. Natl. Acad. Sci. USA* **95**, 11751–11756 (1998).
8. Kawano, T. *et al.* Molecular cloning of cytidine monophospho-N-acetylneuraminic acid hydroxylase. Regulation of species- and tissue-specific expression of N-glycolylneuraminic acid. *J. Biol. Chem.* **270**, 16458–16463 (1995).
9. Muchmore, E. A., Milewski, M., Varki, A. & Diaz, S. Biosynthesis of N-glycolyneuraminic acid. The primary site of hydroxylation of N-acetylneuraminic acid is the cytosolic sugar nucleotide pool. *J. Biol. Chem.* **264**, 20216–20223 (1989).
10. Suzuki, A. Genetic basis for the lack of N-glycolylneuraminic acid expression in human tissues and its implication to human evolution. *Proc. Natl. Acad. Sci. USA* **82**, 93–103 (2006).
11. Hedlund, M., Padler-Karavani, V., Varki, N. M. & Varki, A. Evidence for a human-specific mechanism for diet and antibody-mediated inflammation in carcinoma progression. *Proc. Natl. Acad. Sci. USA* **105**, 18936–18941 (2008).
12. Nguyen, D. H., Tangvoranuntakul, P. & Varki, A. Effects of natural human antibodies against a nonhuman sialic acid that metabolically incorporates into activated and malignant immune cells. *J. Immunol.* **175**, 228–236 (2005).
13. Taylor, R. H. *et al.* Novel mechanism for the generation of human xeno-autoantibodies against the nonhuman sialic acid n-glycolylneuraminic acid. *J. Exp. Med.* **207**, 1637–1646 (2010).
14. Samraj, A. N. *et al.* A red meat-derived glycan promotes inflammation and cancer progression. *Proc. Natl. Acad. Sci. USA* **112**, 542–547 (2015).
15. Samraj, A. N., Laubli, H., Varki, N. & Varki, A. Involvement of a non-human sialic acid in human cancer. *Front. Oncol.* **4**, 33, https://doi.org/10.3389/fonc.2014.00033 (2014).
16. Higashi, H. *et al.* Tumor-associated expression of glycosphingolipid Hanganutziu-Deicher antigen in human cancers. *Gan* **75**, 1025–1029 (1984).
17. Hirabayashi, Y., Higashi, H., Kato, S., Taniguchi, M. & Matsumoto, M. Occurrence of tumor-associated ganglioside antigens with Hanganutziu-Deicher antigenic activity on human melanomas. *Gan* **78**, 614–620 (1987).
18. Miwa, Y. *et al.* Are N-glycolylneuraminic acid (Hanganutziu-Deicher) antigens important in pig-to-human xenotransplantation? *Xenotransplantation* **11**, 247–253 (2004).
19. Burlak, C. *et al.* Reduced binding of human antibodies to cells from GGTA1/CMAH KO pigs. *Am. J. Transplant.* **14**, 1895–1900 (2014).
20. Kwon, D. N. *et al.* Production of biallelic CMP-Neu5Ac hydroxylase knock-out pigs. *Sci. Rep.* **3**, 1981, https://doi.org/10.1038/srep01981 (2013).
21. Kyogashima, M., Ginsburg, V. & Krivan, H. C. *Escherichia coli* K99 binds to N-glycolylsialoparagloboside and N-glycolyl-GM3 found in piglet small intestine. *Arch. Biochem. Biophys.* **270**, 391–397 (1989).
22. Song, K. H. *et al.* Cloning and functional characterization of pig CMP-N-acetylneuraminic acid hydroxylase for the synthesis of N-glycolylneuraminic acid as the xenoantigenic determinant in pig-human xenotransplantation. *Biochem. J.* **427**, 179–188 (2010).
23. Shaw, L. & Schauer, R. Detection of CMP-N-acetylneuraminic acid hydroxylase activity in fractionated mouse liver. *Biochem. J.* **263**, 355–363 (1989).

24. Malykh, Y. N. *et al*. Regulation of N-glycolylneuraminic acid biosynthesis in developing pig small intestine. *Biochem. J.* **370**, 601–607 (2003).
25. Naito, Y. *et al*. Germinal center marker GL7 probes activation-dependent repression of N-glycolylneuraminic acid, a sialic acid species involved in the negative modulation of B-cell activation. *Mol. Cell. Biol.* **27**, 3008–3022 (2007).
26. Jahan, M., Wynn, P. C. & Wang, B. Molecular characterization of the level of sialic acids N-acetylneuraminic acid, N-glycolylneuraminic acid, and ketodeoxynonulosonic acid in porcine milk during lactation. *J. Dairy Sci.* **99**, 8431–8442 (2016).
27. Ji, S. *et al*. Developmental changes in the level of free and conjugated sialic acids, Neu5Ac, Neu5Gc and KDN in different organs of pig: a LC-MS/MS quantitative analyses. *Glycoconj. J.* **34**, 21–30 (2017).
28. Song, K. H. *et al*. Housekeeping promoter 5′pcmah-2 of pig CMP-N-acetylneuraminic acid hydroxylase gene for NeuGc expression. *Glycoconj. J.* **33**, 779–788 (2016).
29. Oliver, J. R., Kushwah, R. & Hu, J. Multiple roles of the epithelium-specific ETS transcription factor, ESE-1, in development and disease. *Lab.Invest.* **92**, 320–330 (2012).
30. Varki, A. Colloquium paper: uniquely human evolution of sialic acid genetics and biology. *Proc. Natl. Acad. Sci. USA* **107**(Suppl 2), 8939–8946 (2010).
31. Ng, P. S. *et al*. Ferrets exclusively synthesize Neu5Ac and express naturally humanized influenza A virus receptors. *Nat. Commun.* **5**, 5750, https://doi.org/10.1038/ncomms6750 (2014).
32. Springer, S. A., Diaz, S. L. & Gagneux, P. Parallel evolution of a self-signal: humans and new world monkeys independently lost the cell surface sugar Neu5Gc. *Immunogenetics* **66**, 671–674 (2014).
33. Moncla, L. H. *et al*. A novel nonhuman primate model for influenza transmission. *PLoS One* **8**, e78750, https://doi.org/10.1371/journal.pone.0078750 (2013).
34. Muchmore, E. A., Varki, N. M., Fukuda, M. & Varki, A. Developmental regulation of sialic acid modifications in rat and human colon. *FASEB J.* **1**, 229–235 (1987).
35. Teneberg, S., Willemsen, P., de Graaf, F. K. & Karlsson, K. A. Receptor-active glycolipids of epithelial cells of the small intestine of young and adult pigs in relation to susceptibility to infection with *Escherichia coli* K99. *FEBS Lett.* **263**, 10–14 (1990).
36. Yuyama, Y., Yoshimatsu, K., Ono, E., Saito, M. & Naiki, M. Postnatal change of pig intestinal ganglioside bound by *Escherichia coli* with K99 fimbriae. *J. Biochem.* **113**, 488–492 (1993).
37. Nakamuta, M. *et al*. Alternative mRNA splicing and differential promoter utilization determine tissue-specific expression of the apolipoprotein B mRNA-editing protein (Apobec1) gene in mice. Structure and evolution of Apobec1 and related nucleoside/nucleotide deaminases. *J. Biol. Chem.* **270**, 13042–13056 (1995).
38. Rajagopalan, S. *et al*. Six mRNAs with different 5′ ends are encoded by a single gamma-glutamyltransferase gene in mouse. *Proc. Natl. Acad. Sci. USA* **90**, 6179–6183 (1993).
39. Tolner, B., Roy, K. & Sirotnak, F. M. Structural analysis of the human RFC-1 gene encoding a folate transporter reveals multiple promoters and alternatively spliced transcripts with 5′ end heterogeneity. *Gene* **211**, 331–341 (1998).
40. Al-Sadi, R., Guo, S., Ye, D. & Ma, T. Y. TNF-alpha modulation of intestinal epithelial tight junction barrier is regulated by ERK1/2 activation of Elk-1. *Am. J. Pathol.* **183**, 1871–1884 (2013).
41. Al-Sadi, R., Ye, D., Said, H. M. & Ma, T. Y. Cellular and molecular mechanism of interleukin-1beta modulation of Caco-2 intestinal epithelial tight junction barrier. *J. Cell. Mol. Med.* **15**, 970–982 (2011).
42. Chung, T. W., Lee, Y. C., Ko, J. H., Kim, C. H. & Hepatitis, B. Virus X protein modulates the expression of PTEN by inhibiting the function of p53, a transcriptional activator in liver cells. *Cancer Res.* **63**, 3453–3458 (2003).

Acknowledgements

Author Contributions

C.H.K. (Kim) conceived the experiments. K.H.S., C.H.K. and T.W.C. conducted the experiments, Y.C.L. and K.T.H. helped in study design and data interpretation. H.S.H., J.Y.P. and S.H.C. have discussed and carefully checked the Discussion and Introduction Section. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-40522-9.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.