THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 291, NO. 43, pp. 22594–22606, October 21, 2016

Epithelial Expression of Human ABO Blood Group Genes Is Dependent upon a Downstream Regulatory Element Functioning through an Epithelial Cell-specific Transcription Factor, Elf5*

Received for publication, April 5, 2016, and in revised form, August 30, 2016 Published, JBC Papers in Press, September 1, 2016, DOI 10.1074/jbc.M116.730655

Rie Sano[‡1], Tamiko Nakajima[‡†], Yoichiro Takahashi[‡], Rieko Kubo[‡], Momoko Kobayashi[‡], Keiko Takahashi[‡], Haruo Takeshita[§], Kenichi Ogasawara[¶], and Yoshihiko Kominato[‡]

From the [‡]Department of Legal Medicine, Gunma University Graduate School of Medicine, Maebashi, 371-8511 Japan, the [§]Department of Legal Medicine, Shimane University School of Medicine, Izumo, Japan, and the [¶]Japanese Red Cross Central Blood Institute, Tokyo, Japan

The human ABO blood group system is of great importance in blood transfusion and organ transplantation. The ABO system is composed of complex carbohydrate structures that are biosynthesized by A- and B-transferases encoded by the *ABO* gene. However, the mechanisms regulating *ABO* gene expression in epithelial cells remain obscure. On the basis of DNase I-hypersensitive sites in and around ABO in epithelial cells, we prepared reporter plasmid constructs including these sites. Subsequent luciferase assays and histone modifications indicated a novel positive regulatory element, designated the +22.6-kb site, downstream from *ABO*, and this was shown to enhance *ABO* promoter activity in an epithelial cell-specific manner. Expression of *ABO* and B-antigen was reduced in gastric cancer KATOIII cells by biallelic deletion of the +22.6-kb site using the CRISPR/Cas9 system. Electrophoretic mobility shift assay and chromatin immunoprecipitation assay demonstrated that the site bound to an epithelial cell-specific transcription factor, Elf5. Mutation of the Ets binding motifs to abrogate binding of this factor reduced the regulatory activity of the +22.6-kb site. Furthermore, *ELF5* knockdown with shRNA reduced both endogenous transcription from *ABO* and B-antigen expression in KATOIII cells. Thus, Elf5 appeared to be involved in the enhancer potential of the +22.6-kb site. These results support the contention that *ABO* expression is dependent upon a downstream positive regulatory element functioning through a tissue-restricted transcription factor, Elf5, in epithelial cells.

The human ABO blood group system is of great importance in blood transfusion and organ transplantation. The system comprises complex carbohydrate structures that are biosynthesized by the A- and B-transferases encoded by the *A* and *B* genes, respectively (1). The ABO genes consist of seven exons spanning more than 20 kb of genomic DNA, and two critical single-base substitutions in the last coding exon result in amino acid substitutions responsible for the difference in donor nucleotide sugar substrate specificity between A- and B-transferases (2). A single base deletion in exon 6 has been ascribed to a shift in the reading frame of codons and abolition of A-transferase activity in most *O* alleles. On the other hand, the distribution of the A- and B-antigens is cell type-specific; for example, the antigens are expressed on red blood cells and epithelial cells, as well as in salivary glands, although they are absent from the central nervous system, muscle, and connective tissue. Moreover, ABH antigens are known to be expressed during the maturation of erythroid and epithelial cells; for example, when erythroid cells differentiate *in vitro*, ABO is expressed at an undetectable level in the early phase, increases subsequently, and then decreases later (3, 4). In addition to the normal cell differentiation process, changes in ABH antigen expression have also been documented in abnormal processes such as tumorigenesis (1). Reduction or complete deletion of A/B antigen expression in primary lung, bladder, and colorectal carcinomas has been reported. This phenotypic change was well correlated with the invasive and metastatic potential of the tumors and with patient 5- or 10-year mortality rates (5, 6).

The DNA sequences in and around specific genes provide the code that dictates when, where, and at what level specific genes are transcribed. This code comprises three parts: the core promoter, the region proximal to the core promoter, and the more distant enhancer sequences. It has become obvious that enhancers usually work in groups (*i.e.* the locus control region and super enhancers), each being bound by several transcription factors (TFs),[2] forming a so-called enhanceosome. These enhanceosomes are nucleated by pioneer TFs early during differentiation and are subsequently replaced by other TFs that trigger polymerase II recruitment. Enhancers recruit the preinitiation complex and TFs and interact with each other through a multilooped structure (7, 8).

* This work was supported in part by a Grants-in-Aid 25460861 (to T. N.), 25860486 (to Y. T.), and 26293161 (to Y. K.) from the Japan Society for the Promotion of Science. The authors declare that they have no conflicts of interest with the contents of this article.

† This author is deceased.

[1] To whom correspondence should be addressed: Dept. of Legal Medicine, Gunma University Graduate School of Medicine, 3-39-22 Showa-machi, Maebashi, 371-8511 Japan. Tel.: 81-27-220-8031; Fax: 81-27-220-8035; E-mail: takagirie@gunma-u.ac.jp.

[2] The abbreviations used are: TF, transcription factor; qPCR, quantitative real time PCR; CGI, CpG island; HMM, hidden Markov model; HMEC, human mammary epithelial cell; NHLF, normal human lung fibroblast; EMT, epithelial-mesenchymal transition; TK, thymidine kinase; TMM, trimmed mean of *M* value; MFI, median fluorescence intensity; CTCF, CCCTC-binding factor.

The regulatory mechanisms underlying *ABO* expression have been studied using cultured cells and human genetic analysis. A proximal promoter has been found within the ABO CpG island (CGI) (9, 10). However, a cell type-specific, distal promoter has been demonstrated at the 5′ boundary of the CGI, although the amount of transcript from this promoter was very small (3). Because the ABO proximal promoter showed constitutive activity regardless of the cells examined in transfection experiments (3), it had been assumed that some cell-specific regulatory elements are involved in cell type-specific ABO expression. Recently, a candidate for erythroid cell-specific regulatory element, named the +5.8-kb site, has been proposed in the first intron of ABO (11). Human genetic analysis demonstrated a 5.8- or 3.0-kb deletion including this site in individuals with subgroup B_m, where B-antigen expression is barely detectable on erythrocytes, although the antigen is present in the saliva of secretor individuals (11, 12). Moreover, EMSA and ChIP assay have demonstrated that the transcription factors Runt-related transcription factor 1 (RUNX1), GATA-1 and GATA-2 are bound to the site through their recognition motifs, mutations of which were shown to reduce the transcriptional activity of the site (4, 11, 13, 14). Furthermore, natural deletion and mutation of their binding sites were involved in subgroups A_m and B_m (13–16). On the other hand, no epithelial cell-specific regulatory element has yet been characterized.

Among various cell-specific TFs, the Ets transcription factors play a crucial cell-specific role in transcriptional regulation of genes involved in a variety of developmental and cellular responses, including tumorigenesis and differentiation (17). The Ets family is quite large, comprising at least 26 unique members in mouse, all of which contain an evolutionarily conserved DNA-binding domain called the ETS domain. All DNA-binding ETS domains recognize a GGA(A/T) core sequence motif, although different Ets proteins exhibit a preference for different flanking sequences to bind differentially to specific DNA sites. Moreover, Ets proteins are expressed in a wide variety of tissues and organs and include some members that are expressed ubiquitously and others that display cell- and tissue-specific expression. For example, some Ets proteins such as Ets1 and Ets2 are widely expressed during development and differentiation of many tissues. On the other hand, other Ets proteins such as ESE-1/Elf3, ESE-2/Elf5, and ESE-3/EHF are expressed specifically in epithelial cells of various tissues and organs. ESE-1/Elf3 is broadly expressed in organs such as lung, stomach, kidney, colon, and skin and shows particularly high expression in the small intestine (18–20). Elf5 displays a more restricted pattern of expression in tissues that are rich in glandular or secretory epithelial cells, such as salivary gland, mammary gland, kidney, and stomach (21, 22). ESE-3/EHF is expressed constitutively in the bronchial and mucous gland epithelial cells of the lung in addition to prostate, pancreas, salivary gland, and trachea (23–25).

In the present study, we identified an epithelial cell-specific regulatory element downstream of *ABO* for which an epithelial cell-specific transcription factor, Elf5, was indicated to be involved in its enhancer potential. These findings support the contention that *ABO* expression is dependent upon a downstream positive regulatory element binding to Elf5 in epithelial cells.

Results

Identification of an Epithelial Cell-specific Regulatory Element Downstream of the ABO Gene—Publicly available data of Open Chromatin by DNaseI HS from ENCODE/OpenChrom (Duke University) ENCODE July 2012 Freeze (26) from the UCSC genome browser on human February 2009 (GRCh37/hg19) assembly showed genome-wide maps for DHSs in epithelial cells including pHTE, HPDE6-E6E7, and RWPE1 (Fig. 1), where there were many DHSs located around −0.1, +5.8, +22.6, and +36.0 kb from the ATG translation start site of exon 1 of *ABO*. Fig. 1 also shows DHSs in K562 cells representative of erythroid cells, as well as skeletal muscle cells, which are known not to express *ABO*, as controls. DHS −0.1 kb coincided with the proximal promoter within the *ABO* CpG island, whereas DHS +5.8 kb has been reported to correspond to an erythroid cell-specific regulatory element (11). Compared with those in K562 cells and skeletal muscle cells, DHS +22.6 kb seemed relatively specific to epithelial cells, whereas DHS +36.0 kb appeared without regard to cell type. Hereafter, we refer to the region comprising DHS around +22.6 or +36.0 kb as region +22.6 or +36.0, respectively.

Gastric cancer KATOIII cells are useful for investigating the mechanisms underlying *ABO* expression in epithelium, because B-antigens were expressed on the cell surface, and DNA methylation was undetectable from the 5′ end of the *ABO* CGI to the proximal promoter in cells of genotype *BB* (4, 9, 11). K562 cells were used as erythroid cells expressing *ABO*, whereas OUMS-36T-1 cells were chosen because of their absence of *ABO* expression (11). To examine the epithelial-specific enhancer potential, region +22.6 was PCR-amplified and subcloned upstream of the *ABO* proximal promoter sequence in the same orientation as that of the promoter in luciferase reporter plasmid +22.6/SN, and subsequent transient transfection experiments were carried out using KATOIII cells. Introduction of region +22.6 resulted in an ~3.4-fold increase of luciferase activity relative to SN, although no significant increase was observed in K562 cells and OUMS-36T-1 cells (Fig. 1). In sharp contrast, region +36.0 showed increase of luciferase activity independent of cell specificity. Those observations suggested that region +22.6 could contain a significant functional component specific to cells of epithelial lineage.

However, no similar elevation of luciferase activity was demonstrated when KATOIII cells were transfected with luciferase reporter plasmid A-D/SN in which region +22.6 had been subcloned upstream of the promoter sequence in an orientation opposite to that of the promoter (Fig. 2). To identify any important element involved in transcription, subsequent transfection into KATOIII cells was performed using reporter plasmids containing four subregions of region +22.6, because the orientation-dependent activity could be ascribable to a mixture of positive and negative regulatory elements within the region. The results indicated that a cis-acting element critical for reporter expression resided within subregion C between +22,563 and +22,781. Approximately 9-fold higher luciferase

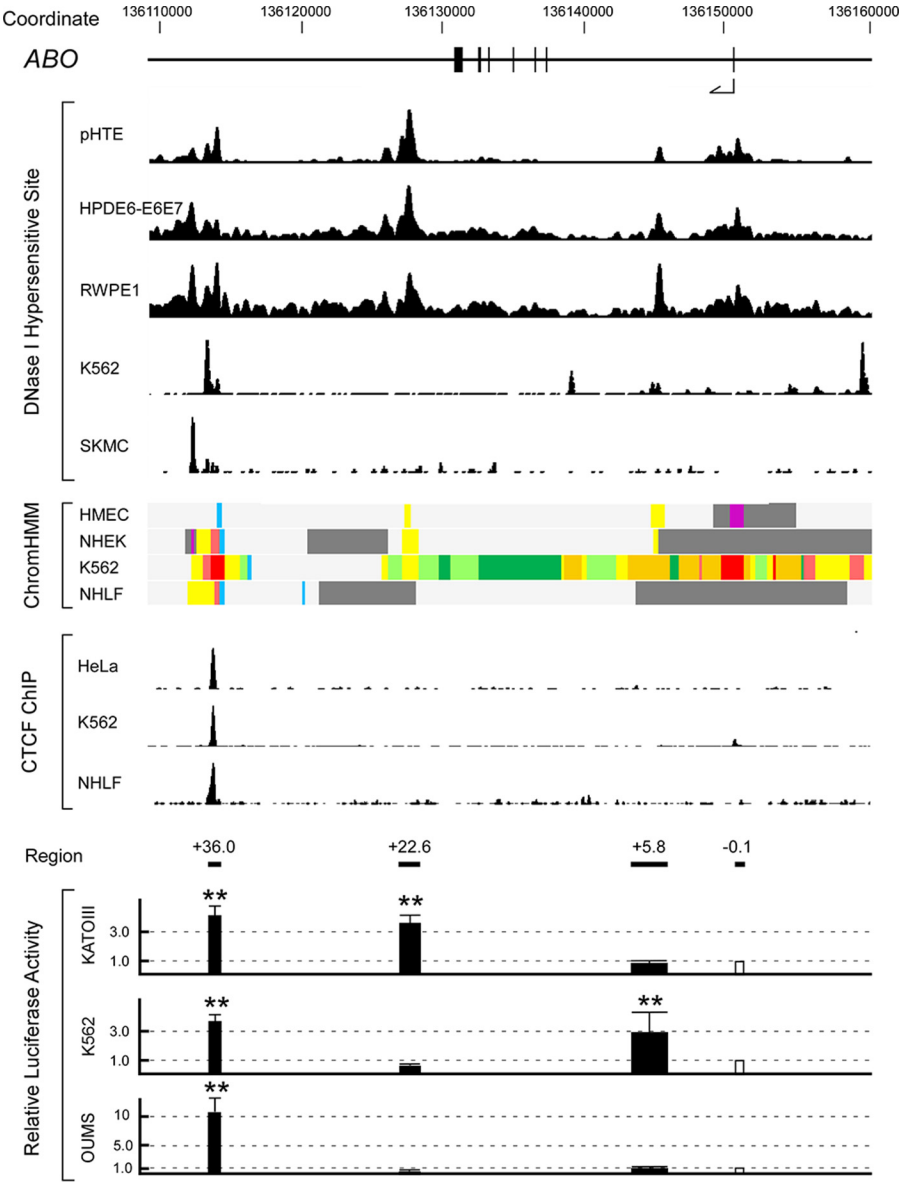

FIGURE 1. **A map of the 50-kb region of genomic DNA from upstream to downstream of the human *ABO* gene.** The *top diagram* indicates *ABO* gene exons 1–7 as represented by *vertical lines* with coordinates in hg19. The *second diagram* from the *top* indicates DHS signal tracks, which were constructed using from the UCSC genome browser among various cells. *pHTE*, primary tracheal epithelial cells; *HPDE6-E6E7*, pancreatic duct cells immortalized with the E6E7 gene of human papillomavirus; *RWPE1*, prostate epithelial cells; *K562*, erythroleukemia cells; *SKMC*, skeletal muscle cells. The *third diagram* from the *top* indicates chromatin state segmentation by the HMM from ENCODE/Broad. Each of the tracks from the various cell lines are shown: *HMEC*, human mammary epithelial cells; *NHEK*, normal human epidermal keratinocytes; *K562*, erythroleukemia cells; *NHLF*, normal human lung fibroblasts. The colored segmentations are represented as follows: *bright red*, active promoter; *light red*, weak promoter; *purple*, poised promoter; *orange*, strong enhancer; *yellow*, weak/poised enhancer; *blue*, insulator; *dark green*, transcriptional elongation; *light green*, weakly transcribed; *gray*, polycomb repressed; *light gray*, low signal or repetitive/copy number variation. The *fourth diagram* shows CTCF binding sites of HeLa cells, K562 cells, and NHLFs from CTCF binding sites by ChIP-seq from ENCODE/University of Washington. The *fifth diagram* indicates locations of DNA fragments that were subcloned upstream of the *ABO* promoter in reporter plasmids +5.8/SN, +22.6/SN, and +36.0/SN. The positions of regions +22.6 and +36.0 are denoted in Table 1. Transient transfection into KATOIII, K562, and OUMS-36T-1 cells was performed using firefly luciferase reporter plasmids and the *Renilla* luciferase reporter vector. The *diagram* at the *bottom* shows the relative activities of individual reporter plasmids, when the activity of reporter plasmid SN containing the promoter was assigned an arbitrary value of 1.0, indicated by an *open box*. All data represent means from more than three independent experiments, and the standard deviations are also shown. The significance of increases in reporter activities was determined by Student's *t* test at a significance level of $p < 0.01$ (**).

activity was observed when subregion C was subcloned upstream of the *ABO* promoter in a direction opposite to that of the promoter in plasmid C/SN, whereas introduction of the sequence in the promoterless vector C/Basic yielded luciferase activity similar to that for SN. In contrast, ~10-fold higher luciferase activity was observed when the subregion was inserted in the same direction as that of the promoter in plasmid rC/SN, whereas introduction of the sequence in the promoterless vector rC/Basic led to 4-fold higher luciferase activity relative to SN. These observations indicated that subregion C might have potential for enhancement of promoter activity, as well as initiation of transcription.

Furthermore, subregion C was proved to activate the *ABO* promoter independently of the distance from it; introduction of the sequence downstream of *luciferase* resulted in an ~2-fold increase of luciferase activity relative to SN (SN/C; Fig. 3). Next,

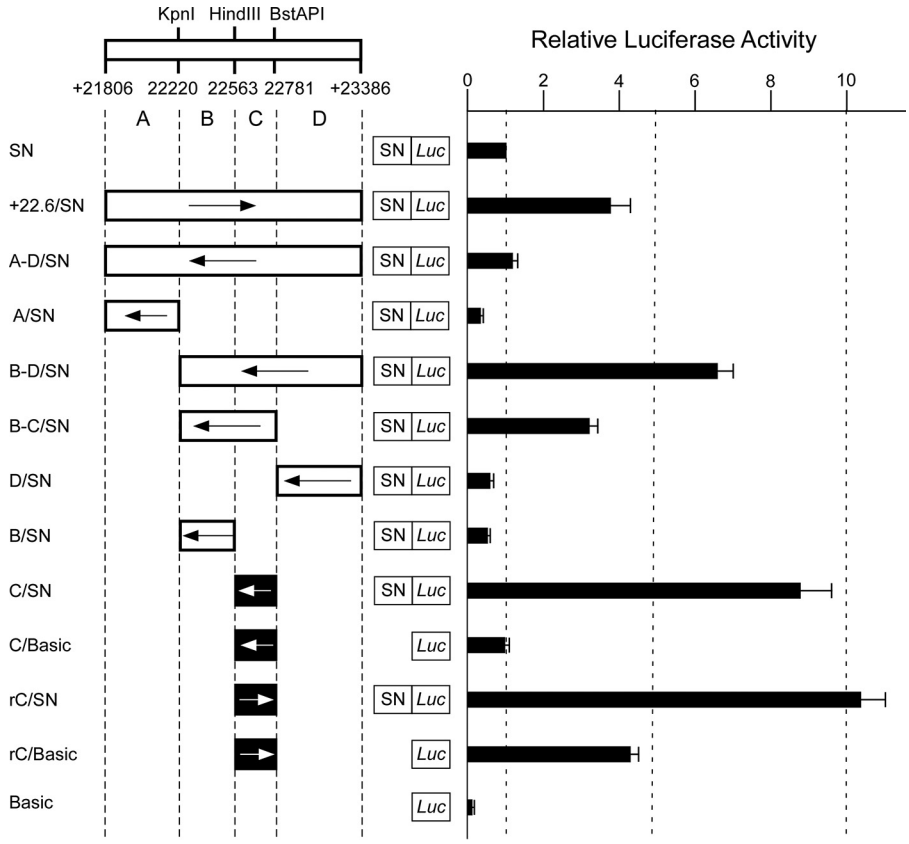

FIGURE 2. **Summary of the relative luciferase activities of reporter constructs containing various parts of DNA region +22.6 in *ABO*.** Transient transfection into KATOIII cells was performed using reporter plasmids in which various parts of region +22.6 were subcloned upstream of the *ABO* promoter sequence. Below the map of restriction enzyme sites with positions relative to the translation start site of exon 1, region +22.6 was subdivided into *regions A–D*. Construct names are indicated to the *left* of the *square*, and the locations of the fragments that were inserted upstream of the *ABO* promoter sequence are shown; the fragments were inserted in the same direction to that of the promoter in constructs +22.6/SN, rC/SN, and rC/Basic; in the other constructs, the fragments were inserted in a direction opposite to that of the promoter. *Arrows* to the *right* represent the inserts oriented in the same direction to that of the promoter. The +22,563 to +22,781 sequence (subregion C or the +22.6-kb site) is indicated by a *solid box* with an *arrow*. Basic represents the promoter-less pGL3-basic vector. Each construct as depicted on the *left* was used for transient transfection, and the luciferase activity obtained was normalized as shown in the right panel. To facilitate comparison of the corresponding reporter activity of each construct, the activity of reporter plasmid SN was assigned an arbitrary value of 1.0. The results are expressed as the relative activities observed. All data represent means from more than three independent experiments, and the standard deviations are also shown.

subregion C was verified to function independently of the promoter; insertion of the sequence downstream of *luciferase*, which was driven by the *SV40* or *TK* promoter, led to ~2-fold higher luciferase activity in comparison with SV or TK (SV/C, TK/C; Fig. 3). Thus, subregion C appeared to be involved in positive transcriptional regulation. When reporter construct SN/C was transfected into K562 and OUMS-36T-1 cells, luciferase activity was not increased relative to SN. These findings indicated that subregion C enhances *ABO* promoter activity in an epithelial cell-specific manner. Hereafter, we refer to subregion C as the +22.6-kb site.

Histone modifications were then examined to characterize the +22.6-kb site further. The ChIP assays comprised qPCR targets such as the *ABO* promoter, the +9.0-kb site, +22.6-kb site, and human *myoglobin* exon 2. The +9.0-kb site is located 9.0-kb downstream from the *ABO* translation start site. The +9.0-kb site and the *myoglobin* exon 2 were used as negative controls, because publicly available data of Chromatin State Segmentation by the hidden Markov model (HMM) from ENCODE/Broad from the UCSC genome browser suggested that the +9.0-kb site lacked any transcriptional regulatory element in human mammary epithelial cells (HMECs) and normal human lung fibroblasts (NHLFs) (Fig. 1) (27) and that the *myoglobin* exon 2 was not associated with any transcriptional regulatory element in HMEC, although it was a Polycomb-repressed region in NHLF, as also shown in the sessions that we have prepared (the tracked session can be found on the UCSC Genome Browser using the User Name: RieSANO, Session Name: myoglobin). Compared with OUMS-36T-1 cells not expressing *ABO*, ChIP assays of KATOIII cells showed that the +22.6-kb site was subjected to histone H3 monomethylation at lysine residue 4 (H3K4me1), whereas the *ABO* promoter was enriched for histone H3 trimethylation at lysine residue 4 (H3K4me3) (Fig. 4). In addition to covering the promoter sequences, the +22.6-kb site was also enriched for histone H3 acetylation at lysine residue 27 (H3K27ac), which was used as another indicator of a positive regulatory region. In contrast, neither sequence was enriched for histone H3 trimethylation at lysine residue 27 (H3K27me3), which was used as an indicator of a negative regulatory region. On the other hand, cross-linked chromatin derived from OUMS-36T-1 cells indicated that the *ABO* promoter, the +9.0-kb site, the +22.6-kb site, and the *myoglobin* exon 2 were enriched for H3K27me3 and that these regions were not enriched for the other modifications. Thus,

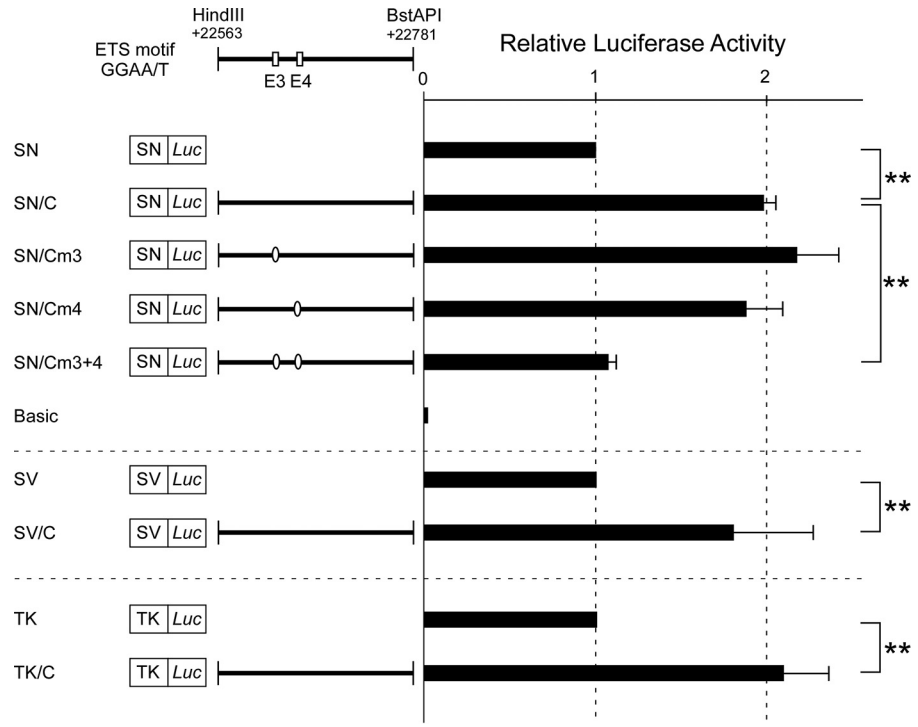

FIGURE 3. **Mutation of the Ets transcription factor binding motifs reduces the regulatory potential of the +22.6-kb site.** Subregion C or the +22.6-kb site was subcloned downstream of *luciferase* which was driven by the *ABO*, *SV40*, and *TK* promoters in constructs SN/C, SV/C, and TK/C, respectively. The site was inserted downstream of *luciferase* in the same orientation to each promoter in the constructs. We also prepared mutant construct SN/Cm3, SN/Cm4, or SN/Cm3 + 4 carrying GGAA to CCAA substitutions in the putative Ets factor binding site E3, E4, or both, respectively. These plasmids were transiently transfected into KATOIII cells. *Boxes* indicate the locations of sites E3 and E4. The *open circles* represent mutated sequences of the sites. To facilitate comparison of the corresponding reporter activity of each construct, the activity of reporter plasmid SN, SV, or TK was assigned an arbitrary value of 1.0. The results are expressed as the relative activities observed. All data represent means from more than three independent experiments, and the standard deviations are also shown. The significance of differences was determined by Student's *t* test at a significance level of $p < 0.01$ (**).

these observations suggested that the +22.6-kb site is a positive regulatory element.

Reduction of ABO and B-antigen Expression by Biallelic Deletion of the +22.6-kb site in KATOIII Cells—To examine whether the +22.6-kb site was involved in regulation of transcription from *ABO*, the CRISPR/Cas9 system with a pair of genome-editing plasmids 5′Enhancer and 3′Enhancer was used to create genomic deletions of the +22.6-kb site in KATOIII cells of genotype *BB*. After transfection and cloning, three clones with biallelic deletion were verified by PCR36 amplification and subsequent cloning and sequencing. The sequencing revealed genomic deletion of the sequences +22,579/+22,740 and +22,579/+22,741 in clone A4, the sequences +22,579/+22,740 and +22,579/+22,742 in clone B3, and the sequences +22,580/+22,740 and +22,580/+22,742 in clone B4. As a control, the CRISPR/Cas9 system was also used to create genomic deletion from the promoter proximal to exon 1 in KATOIII cells, yielding clone 1E9 with genomic deletion of the sequences −116/+12 and −116/+15. Quantitative real time PCR demonstrated that this biallelic deletion of the +22.6-kb site resulted in loss of 55–68% of the transcript amount relative to the wild-type sequence (Fig. 5*A*). RNA-seq using the wild-type, A4, B3, and B4 cells yielded results similar to those obtained by quantitative PCR (Fig. 5, *B* and *C*). Subsequent flow cytometric analysis demonstrated a moderate decrease of B-antigen expression on clones A4, B3, and B4 (Fig. 5D, *panels c–e*). Compared with the median fluorescence intensity of the wild-type cells, those of the deletion cells B3 and B4 were significantly reduced (Fig. 5*E*). In contrast, a large reduction of B-antigen expression was observed on 1E9 cells (Fig. 5, *D* and *E*), suggesting that transcription from the distal promoter might result in slight expression of B-antigen. Flow cytometric analysis might not be sufficiently sensitive for examining the effect of enhancer deletion. Because the glycosylation was stable and the reductions in the amounts of transcript were partial, deletion of the enhancer was expected to have a modest effect on the amount of B-antigen. On the basis of the overall data, we concluded that the +22.6-kb site plays an important role in transcriptional regulation of *ABO* expression in cells of epithelial lineage. However, biallelic deletion of the +22.6-kb site did not achieve complete loss of the expression, suggesting that another enhancer element may interact with the *ABO* promoter through a looped structure.

Involvement of Elf5 in Cell Type-specific Regulatory Activity of the +22.6-kb Site—Inspection of the nucleotide sequence between +22,563 and +22,781 revealed five putative binding sites for Ets transcription factors (Fig. 6*A*). RT-PCR showed that the *ELF3* and *ELF5* transcripts were expressed in KATOIII cells, but not in K562 and OUMS-36T-1 cells, and that *EHF* was detectable in those cells (data not shown). Considering its distribution in tissues rich in glandular or secretory epithelial cells, Elf5 could be involved in enhancer potential of the element. The relevant motifs for Elf5 were the sequences centered on positions +22,628 and +22,654, which were named sites E3 and E4, respectively. We examined whether the +22.6-kb site binds to the factor using EMSA with the labeled probe C-E4 involving site E4 (Fig. 6*A*). The oligonucleotide C-E4 probe produced a

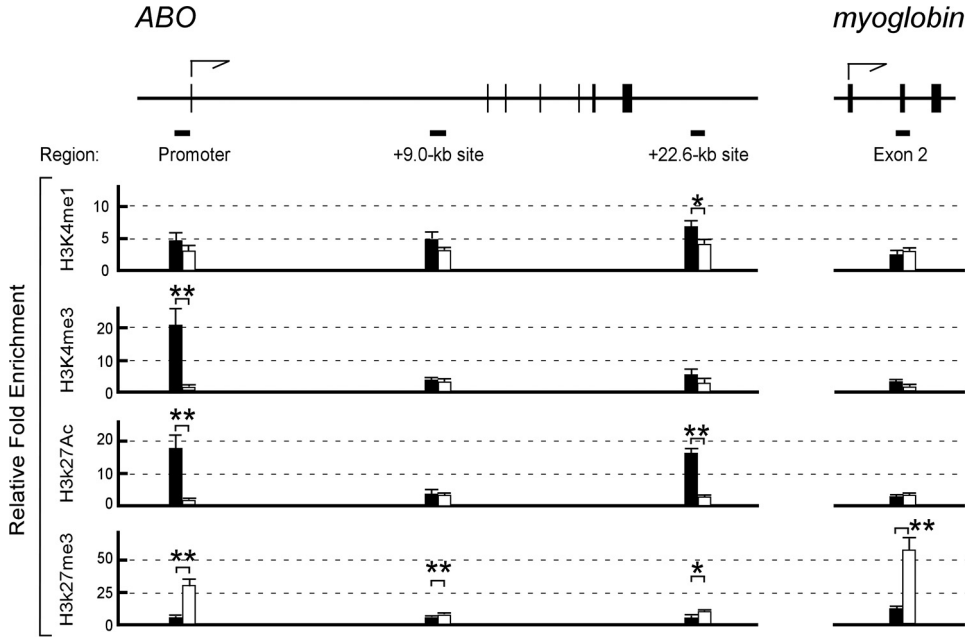

FIGURE 4. **Quantitative ChIP assay for evaluating histone modifications of the +22.6-kb site of *ABO*.** Cross-linked chromatin derived from KATOIII cells and OUMS-36T-1 cells was immunoprecipitated with anti-H3K4me1, anti-H3K4me3, anti-H3K27ac, and anti-H3K27me3 antibodies and control IgG. The precipitated DNA was subjected to qPCR with specific primers for the *ABO* promoter, the +9.0-kb site, the +22.6-kb site, and human *myoglobin* exon 2, calculating the percentage of the amount of precipitated DNA relative to the amount of input DNA for each target. The percentage for each target was divided by the percentage of the amount of DNA precipitated by control IgG to obtain the relative fold enrichment for target DNA precipitated by an individual antibody. The *top diagram* indicates the *ABO* gene exons 1–7 and *myoglobin* exons 1–3, which are denoted by *vertical lines*. *Horizontal bars* represent qPCR targets for the *ABO* promoter, the +9.0-kb site, the +22.6-kb site, and *myoglobin* exon 2. The diagrams under the *top diagram* indicate histone modifications such as H3K4me1, H3K4me3, H3K27ac, and H3K27me3 in KATOIII cells and OUMS-36T-1 cells. *Black bars* show the relative fold enrichment in each target region in KATOIII cells, and *open bars* denote that in OUMS-36-T1 cells. The *y* axis represents the relative fold enrichment of the target sequence immunoprecipitated by each individual antibody relative to the DNA precipitated by IgG, whereas the *x* axis indicates the locations of the target. The results are expressed as the relative enrichments observed. All data represent means from three independent experiments, and the standard deviations are also shown. The significance of differences in enrichment between KATOIII cells and OUMS-36-T1 cells in each region was determined by Student's *t* test at a significance level of $p < 0.05$ (*) or $p < 0.01$ (**).

major up-shifted band when the probe was mixed with the nuclear extract from KATOIII cells in the presence of a small amount of poly(dI-dC) (Fig. 6*B*), whereas the DNA-protein complex was not observed in mixtures containing ordinary amounts of poly(dI-dC) ranging from 0.1 to 1 μg in 15 μl of binding solution (data not shown). Formation of the up-shifted complexes, indicated by an *arrow*, was decreased by addition of competing unlabeled self-oligonucleotide or oligonucleotide Ccnd2 containing the Elf5 recognition motif in the *Cyclin D2* promoter (28), but not by addition of oligonucleotide mCcnd2 containing substitutions of GGAA with CCAA in the core sequence motif or oligonucleotide mC-E4, which was a mutated version of C-E4 comprising the same mutations. Furthermore, formation of the DNA-protein complex was reduced when the anti-Elf5 antibody was incubated with the nuclear extract. The specificity of the antibody was verified by Western blotting using the nuclear extract from KATOIII cells (Fig. 6*C*). In addition, formation of the DNA-protein complex was also reduced when an oligonucleotide C-E3 containing site E3 was incubated with the nuclear extract. These results suggested that Elf5 could be bound to the +22.6-kb site through sites E3 and E4. Next, ChIP assays were performed to evaluate the occupancy of Elf5 at the endogenous +22.6-kb site of *ABO* in KATOIII cells and K562 cells. Cross-linked chromatin was immunoprecipitated with the anti-Elf5 antibody and control IgG, and the precipitated DNA was subjected to ChIP-qPCR39 with specific primers for the +22.6-kb site (Fig. 6*D*). These results demonstrated that Elf5 occupied the endogenous +22.6-kb site of *ABO* in KATOIII cells, but not in K562 cells. Thus, Elf5 seemed to be bound to the +22.6-kb site in epithelial cells.

To investigate whether the function of the +22.6-kb site was dependent upon sites E3 and E4, we prepared reporter constructs SN/Cm3 and SN/Cm4 carrying the same mutations in sites E3 and E4, respectively, as those of oligonucleotide mC-E4 to abrogate Elf5 binding. Transient transfection into KATOIII cells demonstrated that neither mutation reduced any enhancer activity of the +22.6-kb site (Fig. 3). However, mutations of both sites led to a 92% reduction in construct SN/Cm3 + 4. Therefore, it was likely that the transcriptional activity of the +22.6-kb site might be dependent upon either site E3 or E4.

Next, to examine whether Elf5 affects the transcriptional potential of the +22.6-kb site, the *ELF5* or control shRNA expression plasmid was stably transfected into KATOIII cells. After limiting dilution of the cells stably transfected with the *ELF5* shRNA plasmid, clone 1 was selected because qPCR showed the most significant reduction of the *ELF5* transcripts within the cells compared with those obtained from the other four clones (Fig. 7*A*). This clone was used in the subsequent experiments. Quantitative PCR indicated that ~80% of the *ELF5* transcript was reduced in the cells transfected with *ELF5* shRNA relative to that of cells transfected with control shRNA. Western blotting demonstrated that Elf5 protein was decreased in the *ELF5* shRNA-transfected cells in comparison with the

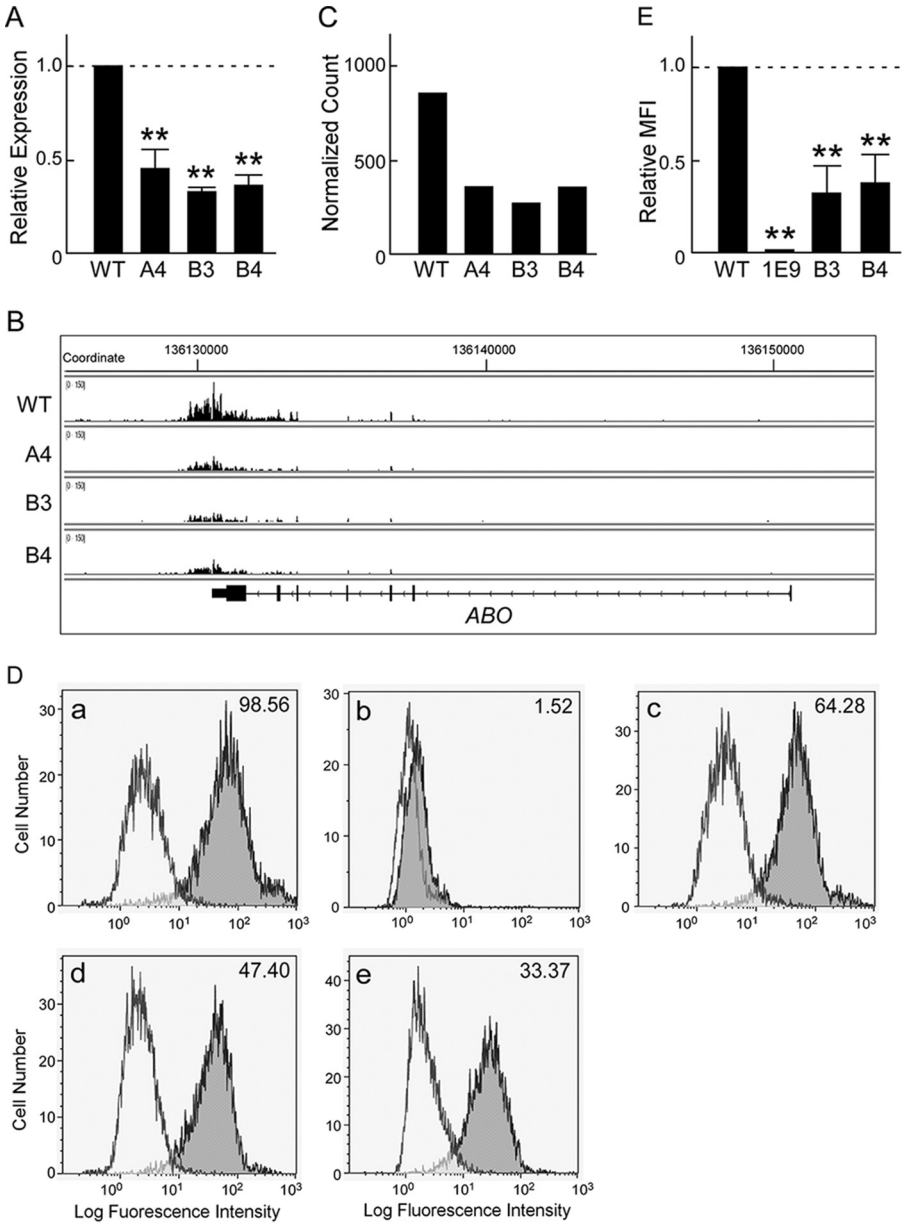

FIGURE 5. **Decrease of *ABO* and B-antigen expression on KATOIII cells harboring biallelic deletion of the +22.6-kb site.** *A*, quantitative analysis of the transcripts from the B-allele in the wild-type and mutant cells including clones A4, B3, and B4 using primers ABO+3 and ABO+98, which corresponded to exons 1 and 2, respectively. The ratio of the target transcripts was calculated by dividing by the copy number of *β-actin*. When the relative amount of the *B*-transcript in wild-type KATOIII cells was assigned an arbitrary value of 1.0, those of *B*-transcripts were calculated in the other clones. The results are indicated as the relative expressions observed. All data represent means from three independent experiments, and the standard deviations are also shown. The significance of differences was determined by Student's *t* test at a significance level of *p* value <0.01 (**). Because the region corresponding to primer ABO+3 was deleted in 1E9 cells, quantitative analysis was not undertaken for those cells. *B*, RNA-seq read alignment coverage of *ABO*. Read alignment coverages in the wild-type and mutant cells, including clones A4, B3, and B4, are displayed above *ABO* visualized with IGV version 2.3.47. *C*, TMM-normalized count of the *ABO* transcripts in the wild-type or mutant cells obtained from RNA-seq. The *y* axis represents the transcript count after normalization of TMM. *D*, flow cytometric analysis of B-antigens on the wild-type and mutant KATOIII cells. The wild-type and mutant cells including clones 1E9, A4, B3, and B4 were stained with mouse monoclonal anti-B IgG antibody and goat anti-mouse IgG conjugated to phycoerythrin. In the histograms, fluorescence is displayed on the *x* axis on a logarithmic scale and the number of cells on the *y* axis. Individual *panels a–e* shows representative histograms of cell surface B-antigen expression, as indicated in *gray*, of the wild-type, 1E9, A4, B3, or B4 cells, respectively. *Thin lines* display the negative control cells that were not stained with the anti-B antibody. Median fluorescence intensity (MFI) of the wild-type or mutant cells is shown in each panel. *E*, quantitative analysis of MFIs in the wild-type and mutant cells 1E9, B3, and B4. Relative MFIs were calculated when the MFI of the wild-type cells was assigned an arbitrary value of 1.0. The results are expressed as the relative MFIs observed. All data represent means from three independent experiments, and the standard deviations are also shown. The significance of differences was determined by Student's *t* test at a significance level of $p < 0.01$ (**). A4 cells were not available because of the problems with recovery from frozen stock.

control cells transfected with shRNA (Fig. 7*B*). When transient transfection experiments were performed using reporter plasmids SN/C and SN/Cm3 + 4, the difference of luciferase activity between these plasmids would indicate the enhancer potential that was dependent upon binding of Ets transcription factors to site E3 or E4. The difference was reduced in the knockdown cells relative to that in the control cells (Fig. 7*C*). Therefore, Elf5 protein appeared to be involved in the enhancer

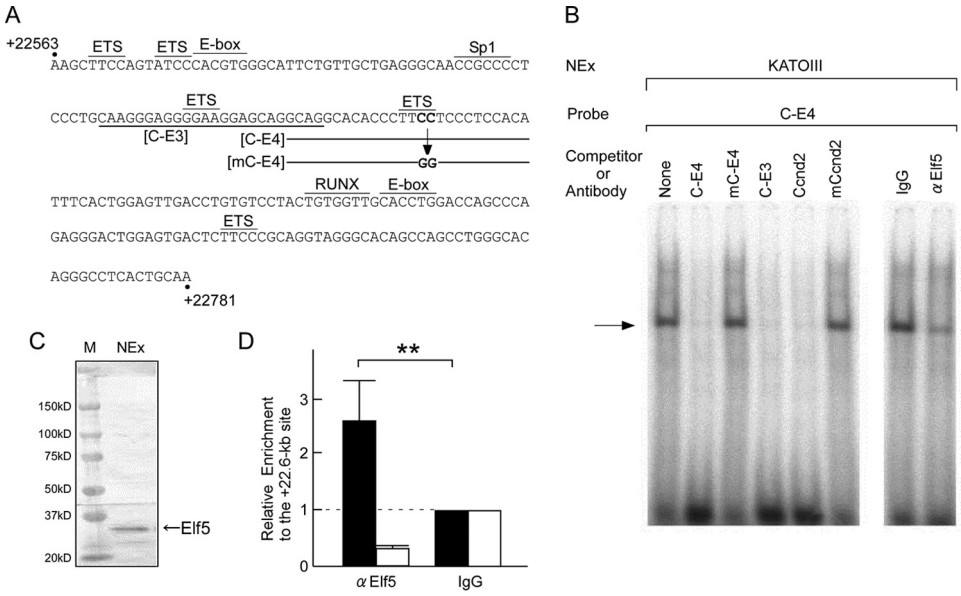

FIGURE 6. **The +22.6-kb site binds to transcription factor Elf5.** *A*, nucleotide sequences of the +22.6-kb site. The sequence from positions +22,563 to +22,781 is shown relative to the ATG translation start site of *ABO*. The sequence JN863720 was derived from the genomic DNA of K562 cells. The motifs for several relevant transcription factors and the E-box are indicated by *overbars*. The oligonucleotides C-E3, C-E4, and mC-E4 used in EMSA are shown under the sequences. The position and identity of the mutations in the GGAA/T core sequence motif are represented in oligonucleotide mC-E4. *B*, the +22.6-kb site binds to a nuclear factor through site E3 or E4. EMSAs were performed using the nuclear extract from KATOIII cells. DNA-protein interaction was investigated using radiolabeled probe C-E4 in the presence or absence of a 200-fold molar excess of competing unlabeled oligonucleotides such as C-E3, C-E4, and a mutated version of oligonucleotide C-E4 or mC-E4, as well as oligonucleotides Ccnd2 and mCcnd2, which included a wild and mutated type of Elf5 recognition motif in the *Cyclin D2* promoter, respectively. Oligonucleotide mCcnd2 comprised two-nucleotide substitutions in the GGAA core sequence motif with CCAA. The major shifted complex is indicated by an *arrow*. The nuclear extract prepared from KATOIII cells was preincubated with the anti-Elf5 antibody or control IgG prior to addition of the radiolabeled probe. *C*, validation of the Elf5 antibody specificity using the nuclear extract prepared from KATOIII cells. The specificity of the Elf5 antibody was evaluated by Western blotting using the nuclear extract prepared from KATOIII cells. The molecular weight of each protein was estimated using Precision Plus Protein™ Kaleidoscope™ (Bio-Rad). *D*, Elf5 interactions with the +22.6-kb site in KATOIII cells. Quantitative ChIP assays were carried out. The *y* axis represents relative enrichment of Elf5 compared with the IgG control. *Solid bars* show relative enrichment in KATOIII cells, and *open bars* show that in K562 cells. All data represent means from three independent experiments, and the standard deviations are also shown. The significance of enrichment was determined by Student's *t* test at a significance level of $p < 0.01$ (**).

potential of the +22.6-kb site through site E3 or E4. Also, quantitative PCR indicated that the amount of *ABO* transcript was decreased significantly in the knockdown cells in comparison with the control cells (Fig. 7*D*). Although *ELF5* shRNA expression resulted in the loss of one-third of the *ABO* transcripts in the knockdown cells relative to the control cells, this appeared to be compatible with the fact that half of the Elf5 protein was decreased in the knockdown cells and that approximately two-thirds of the *ABO* expression was repressed by biallelic deletion of the +22.6-kb site bound to Elf5 (Fig. 5*A*). Subsequent flow cytometric analysis indicated that the amount of B-antigen was moderately decreased in the knockdown cells relative to the control cells (Fig. 7*E*). Because the *ABO* gene encodes glycosyltransferase, which produces blood group antigens, it was unlikely that the extent of transcript reduction was concordant with that of antigen loss. Therefore, we concluded that *ABO* expression is regulated by the +22.6-kb site, which is dependent upon binding of Elf5.

Discussion

In the present study, we have identified an epithelial cell-specific positive regulatory element, named the +22.6-kb site, downstream of *ABO* employing transient transfection experiments with luciferase reporter plasmids that were prepared on the basis of DHSs within a 50-kb region of genomic DNA in and around *ABO* in epithelial cells. The +22.6-kb site appeared to enhance the *ABO* promoter activity in an epithelial cell-specific manner. In addition, B-antigen expression was reduced in KATOIII cells after biallelic deletion of the site using the genome editing technique. Furthermore, subsequent EMSAs, ChIP assays and knockdown experiments revealed that the enhancer potential of the site was dependent upon binding of an epithelial-specific transcription factor, Elf5. Therefore, it is likely that *ABO* expression is dependent upon the tissue-specific enhancer, which is regulated by binding of Elf5 in epithelial cells.

Recently, genome-wide approaches for the discovery of enhancers have become available. Regulatory elements are often characterized by the presence of DHSs, which can mark the position where TFs bind to DNA. Other chromatin features found at distant regulatory elements are histone signatures including modifications associated with activation such as H3K4me1, H3K9ac, and H3K27ac (29, 30). All these features can be used to identify putative functional elements, and these powerful strategies are now being widely applied. Publicly available genome-wide data on histone modifications in HMEC and normal human epidermal keratinocytes (NHEK) on Chromatin State Segmentation by the HMM from ENCODE/Broad *ENCODE Mar 2012 Freeze* (*June 2011 Analysis Pubs*) have suggested that the +22.6-kb site is a weak/poised enhancer (Fig. 1) (31). Based on data from CCCTC-binding factor (CTCF) binding sites by ChIP-seq in several cells, region +36.0 shows binding of the insulator protein CTCF (Fig. 1) and the cohesin sub-

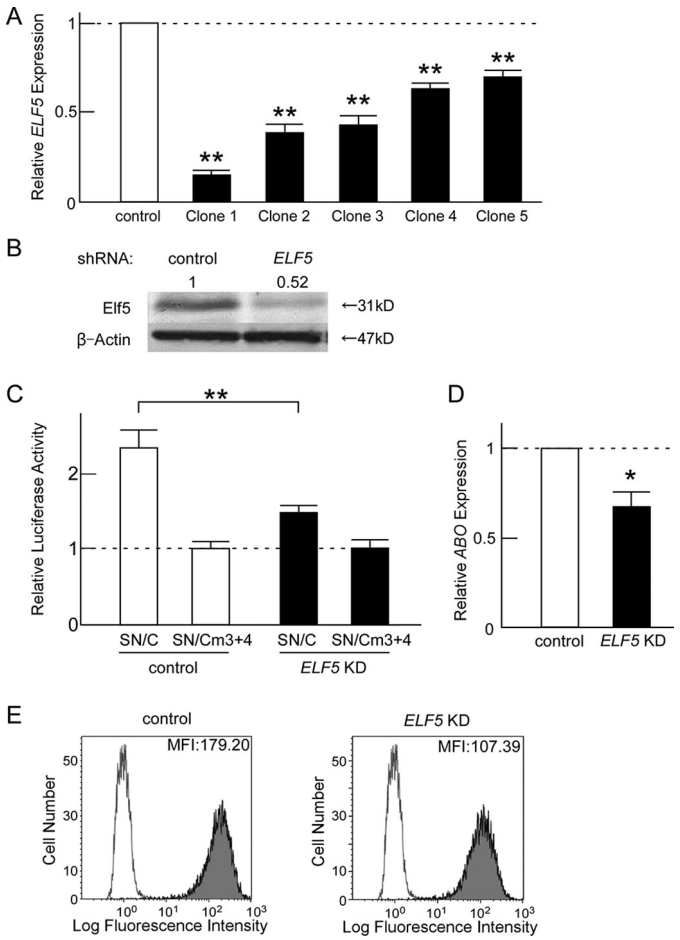

FIGURE 7. **ELF5 knockdown with shRNA reduces endogenous *ABO* transcription and B-antigen expression in KATOIII cells.** *A*, *ELF5* knockdown with shRNA reduces endogenous *ELF5* transcripts in the clones obtained. KATOIII cells were stably transfected with the *ELF5* or control shRNA expression plasmid. Based on subsequent antibiotic selection and limiting dilution, five stable clones transfected with the *ELF5* shRNA plasmid were obtained, followed by quantification of the *ELF5* and *β-actin* transcripts using qPCR. However, the cells transfected with the control plasmid were not cloned. Relative expression of the target *ELF5* transcripts was calculated relative to the copy number of *β-actin* transcripts. Subsequently, the relative expression of *ELF5* was calculated for the clones in the presence of *ELF5* shRNA, assigning that in the cells with control shRNA an arbitrary value of 1.0. The panel indicates the relative expression of *ELF5* in the control shRNA-transfected cells and clones 1–5. All the data represent the means of triplicates, and the standard deviations are also shown. The significance of differences was determined by Student's *t* test at a significance level of $p < 0.01$ (**). *Open* and *solid bars* indicate the relative expression in the control shRNA-transfected cells and the *ELF5* shRNA-transfected cells, respectively. *B*, reduction of Elf5 protein in the *ELF5* shRNA-transfected cells. Amounts of Elf5 and β-actin protein were evaluated by Western blotting using whole cell lysates prepared from the control and *ELF5* shRNA-transfected cells, followed by densitometry measurements. Molecular weight of each protein was estimated using Precision Plus Protein™ Kaleidoscope™ (Bio-Rad). The relative amount of Elf5 protein in the *ELF5* shRNA-transfected cells is indicated above the images when that of the control shRNA-transfected cells is assigned an arbitrary value of 1.0. *C*, decreased enhancer potential of the +22.6-kb site in the *ELF5* shRNA-transfected cells. Luciferase assays with reporter vectors SN/C and SN/Cm3 + 4 were performed using the control and knockdown cells. The average relative luciferase activity of SN/C is expressed when that of SN/Cm3 + 4 is assigned an arbitrary value of 1.0. All data represent means from three independent experiments, and the standard deviations are also shown. The significance of differences was determined by Student's *t* test at a significance level of $p < 0.01$ (**). The *left panel* indicates the relative luciferase activities of SN/C and SN/Cm3 + 4 in the control cells, whereas the *right panel* shows those in the knockdown cells. *D*, reduction of the *ABO* transcripts in KATOIII cells stably transfected with the *ELF5* shRNA expression plasmid. Real time PCR was performed to measure the quantity of the *ABO* and *β-actin* transcripts. Relative amounts of the target *ABO* transcripts were calculated by dividing by the copy number of *β-actin*. Relative expression of *ABO* was calculated in KATOIII cells in the presence of *ELF5* shRNA, when that in the cells with control shRNA was assigned an arbitrary value of 1.0. The results are indicated as the relative expressions observed. The data represent mean values from three independent experiments, and the standard deviation is also shown. The significance of differences was determined by Student's *t* test at a significance level of $p < 0.05$ (*). *Open* and *solid bars* indicate the relative expression in the control shRNA-transfected cells and the *ELF5* shRNA-transfected cells, respectively. *E*, reduction of B-antigen expression on the *ELF5* shRNA-transfected cells. Flow cytometric analysis of B-antigens was carried out on the control and knockdown cells. These cells were stained with mouse monoclonal anti-B IgG antibody and goat anti-mouse IgG conjugated to phycoerythrin. B-antigen expression is indicated in *gray*. MFI of the control cells and the knockdown cells is shown in each panel. For comparison, *thin lines* display B-antigen expression of the control 1E9, which was stained with the anti-B antibody and the goat anti-mouse antibody.

unit RAD21 as shown in the session that we have prepared (the tracked session can be found on the UCSC Genome Browser using the User Name: RieSANO, Session Name: ABO-donstream_Enhancer_ver1), suggesting that the region might be an insulator and involved in chromatin looping (32). Although the region showed more transcriptional activities in the luciferase assays of OUMS-36T-1 cells than in KATOIII cells and K562 cells (Fig. 1), further investigations are needed to characterize the region and reveal its involvement in transcriptional regulation.

The present results suggest that the enhancer potential of the +22.6-kb site is dependent upon binding of Elf5, which has been reported to be involved in regulation of epithelium-specific genes. For example, it modulates the differentiation of mammary gland and alveoli (33, 34). On the other hand, loss of *ELF5* in the mammary gland results in epithelial-mesenchymal transition (EMT) through derepression of *SNAIL2* (35). EMT, which was first recognized by Hay, is a morphological change that allows the escape of epithelial cells from the structural constraints imposed by tissue architecture (36). In epithelial cancer cells, the EMT program has been implicated as the crucial mechanism for acquisition of malignant phenotypes (37). Interestingly, reduction or complete deletion of A/B antigen expression is reportedly correlated with progression of carcinomas such as lung carcinomas (5, 6). Such poor prognosis is also suggested to be caused by high mobility of malignant cells that have lost these antigens (38). As to the mechanisms underlying this antigen reduction, Orlow *et al.* have reported cases of bladder cancer in which allelic loss appeared to lead to loss of A/B antigens (39). We and others have also demonstrated that hypermethylation of the *ABO* promoter could be responsible for the absence of ABO transcript and A-antigen in gastric and colon cancer cell lines (40, 41). Using clinical samples of oral squamous cell cancer, Gao *et al.* showed that loss of A/B antigens was responsible for molecular events such as loss of the A/B allele or *ABO* promoter hypermethylation in two-thirds of tissue samples they examined (42). Thus, an additional mechanism for loss of A/B antigens other than allelic loss or promoter hypermethylation remains to be clarified in one-third of such cases. Consistently, the present observations suggest that loss or reduction of Elf5 could trigger the EMT process and down-regulate the enhancer potential of the +22.6-kb site of *ABO* in malignant cells, followed by invasion and metastasis of the cells with A/B antigen reduction, thus leading to poor outcome. This possibility will be investigated using clinical samples in the next study.

TABLE 1
Sequences of oligonucleotides used for PCR

Oligonucleotide[a]	5′-Primer (5′ → 3′)	Application	Target
ABO+21806[b]	CGTGTTGCTGGCTTTTCCAG	PCR36	Region + 22.6
ABO+23386[b]	GGTCACCTGCCCAACTGATG	PCR36	Region + 22.6
ABO+35008[b]	CACCGTGTTAGCCAGGATGGTTTCAGTC	PCR37	Region + 36.0
ABO+37985[b]	AACAGAGGCAAGCCGTTCCTGTCTC	PCR37	Region + 36.0
ABO−25	GGAGGCCGAGACCAGACGC	ChIP-qPCR38	Promoter
ABO+44	GGTCACCTGCCCAACTGATG	ChIP-qPCR38	Promoter
ABO+22581	GTGGGCATTCTGTTGCTGAG	ChIP-qPCR39	+22.6-kb site
ABO+22666	AATGTGGAGGGAGGAAGGGT	ChIP-qPCR39	+22.6-kb site
ABO+8997	GGTGGCGCCATCTCAACTCAC	ChIP-qPCR40	+9.0-kb site
ABO+9093	GGTGGGTGCCTGTAGTCCCA	ChIP-qPCR40	+9.0-kb site

[a] Oligonucleotide names involve numerals showing 5′ positions relative to the ATG translation start site of exon 1 in *ABO* on genomic DNA of NCBI reference sequence NT_035014.4.
[b] Oligonucleotide primers were used for PCR amplification of genomic DNA to prepare reporter plasmid constructs +22.6/SN and +36.0/SN. Region +36.0 between +35848 and +36819 was prepared by SacII and BamHI digestion of the PCR37 product.

At present, the physiological role of the human ABO blood group system remains elusive, whereas pathogenic implications have been revealed such as host-pathogen interactions and differential susceptibility to *Helicobacter pylori* and Noroviruses among individuals with different glycosylation profiles (43, 44). In addition, genome-wide association studies have indicated associations between the ABO locus and cardiovascular diseases such as venous thromboembolism and coronary artery disease (44). Thus, the biological aspects of ABO blood groups have been clarified with delineation of transcriptional regulation, providing a basis for development of treatment that might reduce the risks for those diseases and ABO-incompatible organ transplantation.

Experimental Procedures

Cells—Human gastric cancer cell line KATOIII (JCRB0611), human erythroleukemia cell line K562 (JCRB0019), and human embryo fibroblast cell line OUMS-36T-1 (JCRB1006.1) were cultured as described previously (3, 11). OUMS-36T-1 is a normal human embryo fibroblast cell line transfected with the human telomerase reverse transcriptase gene.

Plasmids—Luciferase reporter plasmids SN and +5.8/SN were described previously (11). The *ABO* proximal promoter located between −150 and −2 relative to the translation start site was subcloned upstream of the luciferase gene in reporter SN, whereas region +5.8 located between +4602 and +7196 was inserted upstream of the promoter in the same orientation to the promoter in reporter +5.8/SN. DNA regions +22.6 and +36.0 were PCR-amplified using the genomic DNA obtained from K562 cells. Sequences of the primers used and conditions for PCR36 and 37 are shown in Tables 1 and 2, respectively. The PCR products were subsequently cloned into the pUC118 vector using a Mighty cloning reagent set (Blunt End) (TaKaRa, Shiga, Japan). The nucleotide sequences of the amplified fragments were determined with a BigDye® Terminator v1.1 cycle sequencing kit (Thermo Fisher Scientific) with both M13 forward and reverse primers, specific primers for the target, and the primers used for PCR. The sequencing run was performed on a Genetic Analyzer (Thermo Fisher Scientific). DNA region +22.6 was subcloned at the SacI and MluI sites just upstream of the *ABO* proximal promoter sequence in the same orientation to the promoter in reporter plasmid +22.6/SN. Also, both sites were used to insert the subregions of region +22.6, which were prepared by restriction enzyme digestion, upstream of the promoter sequence in an orientation opposite to that of the promoter in constructs A/SN, B/SN, C/SN, D/SN, A-D/SN, B-D/SN, B-C/SN, and C/SN. Subregion C was inserted upstream of the promoter sequence in the same orientation as that of the promoter in reporter plasmid rC/SN and inserted into the BamHI and SalI sites downstream of *luciferase* to generate construct SN/C. Subregion C was also inserted upstream of *luciferase* in an orientation opposite to or the same as that of *luciferase* in the promoter-less constructs C/Basic and rC/Basic, respectively. Subregion C was also subcloned downstream of *luciferase*, which was driven by the SV40 promoter and the thymidine kinase (TK) promoter in constructs SV/C and TK/C, respectively. Mutation of the Ets motif in subregion C was generated by overlapping PCR mutagenesis in constructs SN/Cm3, SN/Cm4, and SN/Cm3 + 4. Region +36.0 was prepared by SacII and BamHI digestion of the PCR37 product, and the region was also inserted upstream of the *ABO* proximal promoter in the same orientation as that of the promoter in reporter plasmid +36.0/SN.

The directions of the inserts for all of the constructs used in this study were verified by detailed restriction enzyme mapping and DNA sequence analysis as described above. For all constructs derived by overlapping PCR mutagenesis, sequencing was performed over the entire region of the amplified sequences as described above. Plasmid DNA was purified using a HiSpeed® plasmid maxi kit (Qiagen).

Transfection and Luciferase Assay—Transient transfection of K562 cells or OUMS-36T-1 cells was performed as reported previously (11). Transient transfection of KATOIII cells was carried out using Lipofectamine LTX reagent (Thermo Fisher Scientific) with 1 μg of reporter plasmid and 0.001 μg of pRL-SV40 *Renilla* reporter in accordance with the manufacturer's instructions. The human *ELF5* expression plasmid was purchased from Sino Biological (Beijing, China). After collecting the cells, cell lysis and luciferase assays were performed using the Dual Luciferase reporter assay system (Promega, Madison, WI) to measure the activities of firefly and *Renilla* luciferases. Variations in transfection efficiency were normalized to the activities of *Renilla* luciferase expressed from the cotransfected pRL-SV40 *Renilla* luciferase reporter.

ChIP Assay—ChIP assay was performed using a HighCell# ChIP Kit (Diagenode, Liège, Belgium) in accordance with the manufacturer's instructions. The antibodies employed were

TABLE 2
PCR conditions for qPCR and genomic DNA amplification

PCR*	Initial denaturation	Cycles;	Denaturation	Annealing	Extension	Incubation
PCR36	94ºC for 3 min	35 cycles;	98ºC for 10 sec	60ºC for 15 sec	68ºC for 2 min	68ºC for 7 min.
PCR37	94ºC for 3 min	35 cycles;	98ºC for 10 sec	60ºC for 15 sec	68ºC for 3 min	68ºC for 7 min.
ChIP-qPCR38	95ºC for 3 min	35 cycles;	95ºC for 3 sec	66ºC for 30 sec		
ChIP-qPCR38	95ºC for 3 min	35 cycles;	98ºC for 3 sec	60ºC for 30 sec		
ChIP-qPCR40	95ºC for 3min	35 cycles;	98ºC for 3 sec	60ºC for 30 sec		

[a] PCR36 and 37 by the three-step PCR method, and ChIP-PCR38 – 40 were performed by the two-step PCR method.

anti-H3K4me1, anti-H3K4me3, anti-H3K27me3 (Diagenode), anti-H3K27ac antibodies (ab4729; Abcam, Cambridge, UK), and anti-Elf5 antibody (C-18; Santa Cruz Biotechnology, Dallas, TX). The anti-Elf5 is a goat polyclonal antibody raised against a peptide mapped at the C terminus of Elf-5 of human origin. ChIP experiments were analyzed using qPCR, named ChIP-qPCR38, 39, and 40, employing StepOne and SYBR Select (Thermo Fisher Scientific). The qPCR was carried out in 20 μl of reaction mixture containing 10 μl of SYBR Select Master Mix (2×), 4 pmol each primer, and 2 μl of the ChIP sample. The primers for the *ABO* proximal promoter, the +22.6-kb site and the +9.0-kb site in ChIP-qPCR38, 39, and 40, respectively, are shown in Table 1, and the PCR conditions are shown in Table 2. The primers for human *myoglobin* exon 2 were purchased from Diagenode, and the qPCR conditions employed were described previously (4).

Genome-editing Plasmids, Transfection, and Screening Clones—Single-guide RNA-specifying sequences were chosen to minimize the likelihood of off target cleavage based on publicly available on-line tools (45). Each single-guide RNA specified sequences corresponding to the *ABO* proximal promoter or the +22.6-kb site with respect to genomic DNA of NCBI reference sequence NT_035014.4. Genome-editing plasmids 5′ promoter, 3′ promoter, 5′ enhancer, and 3′ enhancer were produced using oligonucleotide pairs Pro−101, Pro+11, +22.6-kb+22,578, and +22.6-kb+22,741, respectively (Table 3). 100 pmol of each oligonucleotide corresponding to the upper or lower strand of the targeted sequence was mixed respectively in a 20-μl final volume containing 40 mM Tris-HCl, 50 mM NaCl, and 20 mM $MgCl_2$. Annealed DNA fragments were subcloned into the BbsI site of the pX330-U6-Chimeric BB-CBH-hSpCas9 vector (42230; Addgene, Cambridge, MA) using DNA ligation kit Mighty Mix (TaKaRa) in accordance with the manufacturer's instructions. For all constructs, the inserted sequences were confirmed by subsequent sequencing using a BigDye® Terminator v1.1 cycle sequencing kit with specific primers for the pX-330 plasmid.

Despite their complex karyotype, KATOIII cells exhibited karyotype stability, and two copies were present for chromosome 9 when biallelic deletion of the promoter or the +22.6-kb site was verified (data not shown). 100,000 cells were cotransfected with 2.5 μg of the genome-editing plasmids 5′ promoter and 3′ promoter, as well as the plasmids 5′ enhancer and 3′ enhancer using Lipofectamine LTX reagent. Two days after the transfection, 30 cells/96-well plate were cultured to isolate single cell-derived clones. Biallelic deletion clones were defined when PCR amplification detected the deletion band whose sequence was subsequently verified by cloning into the pUC118 vector and nucleotide determination.

Flow Cytometric Analysis—Flow cytometric analysis was performed on a Cytomics FC 500 flow cytometer (Beckman Coulter, Brea, CA) according to the method described previously (4).

RT-PCR and qPCR—RNA purification and cDNA preparation were performed as reported previously (4, 11). Quantification of the *ABO* and human *β-actin* transcripts was performed using StepOne and SYBR Select (4). RT-PCR of the *ELF3*, *ELF5*, and *EHF* transcripts was performed using the RT^2 qPCR primer assay (Qiagen) in accordance with the manufacturer's protocol. Quantification of the *ELF5* transcript was also performed using the same kit in accordance with the manufacturer's protocol.

RNA-seq—RNA-seq analyses were performed by DNA ChIP Research Inc. The details are as follows. Total RNA obtained from each sample was subjected to sequencing library construction using the NEBNext Ultra Directional RNA library prep kit for Illumina (New England Biolabs, MA) in accordance with the manufacturer's protocol. The quality of the libraries was assessed with an Agilent 2200 TapeStation High Sensitivity D1000 (Agilent Technologies, Santa Clara, CA). The pooled libraries of the samples were sequenced using the Illumina HiSeq system in 51-bp single-end reads.

Sequencing adaptors, low quality reads, and bases were trimmed with Trimmomatic-0.32 tool (46). The sequence reads were aligned to the human reference genome (GRCh37/hg19) using Tophat 2.0.13 (bowtie2–2.2.3) (47), which can adequately align reads onto the location including splice sites in the genome sequence. Files of the gene model annotations and known transcripts were downloaded from the Illumina iGenomes web site, being necessary for the whole transcriptome alignment with Tophat.

The aligned reads were subjected to downstream analyses using StrandNGS 2.6 software (Agilent Technologies). The read counts allocated for each gene and transcript (UCSC version 2013.12.31) were quantified using a trimmed mean of *M* value (TMM) method (48).

Preparation of Nuclear Extracts and EMSA—A nuclear extract from KATOIII cells and probe were prepared as reported previously (9). The EMSA was carried out according to the method described previously (11), except for use of 0.05 μg of poly(dI-dC) in 15 μl of binding solution. The double-stranded oligonucleotides C-E3, C-E4, mC-E4, Ccnd2, and mCcnd2 were obtained by annealing two chemically synthe-

TABLE 3
Sequences of oligonucleotides used for construction of genome-editing plasmids and EMSA

Oligonucleotide pair	Sequences[a]
Pro−101	5′-CACCGTGTTCGGCCTCGGGAAGTCG-3′ and 5′-AAACCGACTTCCCGAGGCCGAACAC-3′
Pro+11	5′-CACCGGCCAGCGTCCGCAACACCT-3′ and 5′-AAACAGGTGTTGCGGACGCTGGCC-3′
+22.6-kb+22578	5′-CACCGAAGCTTCCAGTATCCCACGT-3′ and 5′-AAACACGTGGGATACTGGAAGCTTC-3′
+22.6-kb+22741	5′-CACCGAGTGACTCTTCCCGCAGGTA-3′ and 5′-AAACTACCTGCGGGAAGAGTCACTC-3′
C-E3	5′-CAAGGGAGGGGAAGGAGCAGGCAG-3′ and 5′-CTGCCTGCTCCTTCCCCTCCCTTG-3′
C-E4	5′-GCAGGCACACCCTTCCTCCCTCCACA-3′ and 5′-TGTGGAGGGAGGAAGGGTGTGCCTGC-3′
mC-E4	5′-GCAGGCACACCCTTGGTCCCTCCACA-3′ and 5′-TGTGGAGGGACCAAGGGTGTGCCTGC-3′
Ccnd2	5′-ACGCTAGAGGAGGGGAGGAAAGGGGAGGAGGAA-3′ and 5′-AGGTTCCTCCTCCCCCTTTCCTCCCCTCCTCTAG-3′
mCcnd2	5′-ACGCTAGAGGAGGGGACCAAAGGGGAGGAGGAA-3′ and 5′-AGGTTCCTCCTCCCCCTTTGGTCCCCTCCTCTAG-3′

[a] The sequences of each oligonucleotide are shown. The double-stranded oligonucleotides were obtained by annealing two chemically synthesized strands. The sequences underlined in oligonucleotide pairs Pro−101, Pro+11, +22.6-kb+22578, and +22.6-kb+22741 represent the sequences corresponding to targeted genome. The sequences underlined in oligonucleotides mC-E4 and mCcnd2 indicate the sites of the specific mutations.

sized strands (Table 3). Two microliters of the anti-Elf5 antibody or control IgG was added to the nuclear extracts overnight at 4 °C before addition of the radiolabeled probe. The DNA-protein complex was quantified with a BAS-1800 image analyzer (FujiFilm, Tokyo, Japan).

Transfection of shRNA—The shRNA expression plasmids for human *ELF5* and control plasmid, MISSION® shRNA plasmid DNA (NM_001422) and MISSION® TRC2 pLKO.5-puro non-mammalian shRNA control plasmid DNA, respectively, were purchased from Sigma-Aldrich. The shRNA expression plasmid (2.5 μg) was transfected into KATOIII cells (2.5×10^5 cells/transfection) using Lipofectamine 3000 (Thermo Fisher Scientific) in accordance with the manufacturer's protocol. Two days after transfection, the cells transfected with the *ELF5* shRNA expression plasmid were subjected to limiting dilution in growth medium containing 2 μg/ml puromycin (Thermo Fisher Scientific) to isolate single cell-derived stable clones. However, the cells transfected with the control plasmid were not cloned. 1 or 2 weeks after transfection, colonies surviving in the medium containing puromycin were harvested. We routinely pooled 50–200 colonies, and aliquots were kept in liquid nitrogen for further investigation. We propagated the transfectants in the presence of puromycin for the minimum amount of time necessary to accumulate sufficient cells for quantification of the *ELF5*, *ABO*, and human *β-actin* transcripts, quantification of Elf5 and β-actin proteins, and B-antigen expression. A whole cell lysate prepared from the transfectants using a total protein extraction kit (TaKaRa) was used for Western blot analysis with the anti-Elf5 antibody, followed by densitometry measurements with a LAS-3000 and MultiGauge v3.0 (FujiFilm). B-antigen expression was analyzed by flow cytometry.

Statistical Analyses—All data are expressed as mean values and error bars representing standard deviation from at least three independent experiments. Data analyses for the two groups were performed using Student's *t* test (*, $p < 0.05$; **, $p < 0.01$)

Author Contributions—R. S. conceived, designed, coordinated, performed research, analyzed data, and wrote the paper; T. N., R. K., M. K., K. T., and K. O. performed research; Y. T. designed genome-editing plasmids; and H. T. and Y. K. wrote the paper.

Acknowledgments—We thank Tomoko Miwa for preparation of this manuscript. We express our profound grief and pay our last respects to T. N. for enthusiasm in ABO research.

References

1. Daniels, G. (2013) *Human Blood Groups*, 3rd Ed., pp. 21–25, Wiley-Blackwell, West Sussex, UK
2. Yamamoto, F. (2000) Molecular genetics of ABO. *Vox Sang.* **78,** 91–103
3. Kominato, Y., Hata, Y., Takizawa, H., Matsumoto, K., Yasui, K., Tsukada, J., and Yamamoto, F. (2002) Alternative promoter identified between a hypermethylated upstream region of repetitive elements and a CpG island in human ABO histo-blood group genes. *J. Biol. Chem.* **277,** 37936–37948
4. Sano, R., Nogawa, M., Nakajima, T., Takahashi, Y., Takahashi, K., Kubo, R., Kominato, Y., Yokohama, A., Tsukada, J., Yamao, H., Kishida, T., Ogasawara, K., and Uchikawa, M. (2015) Blood group B gene is barely expressed in *in vitro* erythroid culture of B_m-derived $CD34^+$ cells without an erythroid cell-specific regulatory element. *Vox Sang.* **108,** 302–309
5. Lee, J. S., Ro, J. Y., Sahin, A. A., Hong, W. K., Brown, B. W., Mountain, C. F., and Hittelman, W. N. (1991) Expression of blood-group antigen A: a favorable prognostic factor in non-small-cell lung cancer. *N. Engl. J. Med.* **324,** 1084–1090
6. Matsumoto, H., Muramatsu, H., Shimotakahara, T., Yanagi, M., Nishijima, H., Mitani, N., Baba, K., Muramatsu, T., and Shimazu, H. (1993) Correlation of expression of ABH blood group carbohydrate antigens with metastatic potential in human lung carcinomas. *Cancer* **72,** 75–81
7. Vernimmen, D. (2014) Uncovering enhancer functions using the α–globin locus. *PLoS Genet.* **10,** e1004668
8. Levine, M., Cattoglio, C., and Tjian, R. (2014) Looping Back to leap forward: transcription enters a new era. *Cell* **157,** 13–25
9. Kominato, Y., Tsuchiya, T., Hata, N., Takizawa, H., and Yamamoto, F. (1997) Transcription of human ABO histo-blood group genes is dependent upon binding of transcription factor CBF/NF-Y to minisatellite sequence. *J. Biol. Chem.* **272,** 25890–25898
10. Hata, Y., Kominato, Y., Yamamoto, F. I., and Takizawa, H. (2002) Characterization of the human ABO gene promoter in erythroid cell lineage. *Vox Sang.* **82,** 39–46
11. Sano, R., Nakajima, T., Takahashi, K., Kubo, R., Kominato, Y., Tsukada, J., Takeshita, H., Yasuda, T., Ito, K., Maruhashi, T., Yokohama, A., Isa, K., Ogasawara, K., and Uchikawa, M. (2012) Expression of ABO blood-group genes is dependent upon an erythroid cell-specific regulatory element, which is deleted in persons with the B_m phenotype. *Blood* **119,** 5301–5310
12. Sano, R., Kuboya, E., Nakajima, T., Takahashi, Y., Takahashi, K., Kubo, R., Kominato, Y., Takeshita, H., Yamao, H., Kishida, T., Isa, K., Ogasawara, K., and Uchikawa, M. (2015) A 3.0-kb deletion including an erythroid cell-specific regulatory element in intron 1 of the ABO blood group gene in an individual with the B_m phenotype. *Vox Sang.* **108,** 310–313
13. Nakajima, T., Sano, R., Takahashi, Y., Kubo, R., Takahashi, K., Kominato, Y., Tsukada, J., Takeshita, H., Yasuda, T., Uchikawa, M., Isa, K., and Ogasawara, K. (2013) Mutation of the GATA site in the erythroid cell-specific regulatory element of the ABO gene in a B_m subgroup individual. *Transfusion* **53,** 2917–2927
14. Takahashi, Y., Isa, K., Sano, R., Nakajima, T., Kubo, R., Takahashi, K., Kominato, Y., Tsuneyama, H., Ogasawara, K., and Uchikawa, M. (2014) Deletion of the RUNX1 binding site in the erythroid cell-specific regulatory element of the ABO gene in two individuals with the A_m phenotype. *Vox Sang.* **106,** 165–175

15. Thuresson, B., Hellberg, A., and Olsson, M. L. (2013) Disruption of a binding site for the erythroid transcription factor GATA-1 in intron 1 of the ABO gene in A_m and B_m subgroup samples. *Vox Sang.* **105,** 43
16. Oda, A., Isa, K., Ogasawara, K., Kameyama, K., Okuda, K., Hirashima, M., Ishii, H., Kimura, K., Matsukura, H., Hirayama, F., and Kawa, K. (2015) A novel mutation of the GATA site in the erythroid cell-specific regulatory element of the ABO gene in a blood donor with the A_mB phenotype. *Vox Sang.* **108,** 425–427
17. Sharrocks, A. D. (2001) The ETS-domain transcription factor family. *Nat. Rev. Mol. Cell Biol.* **2,** 827–837
18. Tymms, M. J., Ng, A. Y., Thomas, R. S., Schutte, B. C., Zhou, J., Eyre, H. J., Sutherland, G. R., Seth, A., Rosenberg, M., Papas, T., Debouck, C., and Kola, I. (1997) A novel epithelial-expressed ETS gene, ELF3: human and murine cDNA sequences, murine genomic organization, human mapping to 1q32.2 and expression in tissues and cancer. *Oncogene* **15,** 2449–2462
19. Andreoli, J. M., Jang, S. I., Chung, E., Coticchia, C. M., Steinert, P. M., and Markova, N. G. (1997) The expression of a novel, epithelium-specific ets transcription factor is restricted to the most differentiated layers in the epidermis. *Nucleic Acids Res.* **25,** 4287–4295
20. Ng, A. Y., Waring, P., Ristevski, S., Wang, C., Wilson, T., Pritchard, M., Hertzog, P., and Kola, I. (2002) Inactivation of the transcription factor Elf3 in mice results in dysmorphogenesis and altered differentiation of intestinal epithelium. *Gastroenterology* **122,** 1455–1466
21. Zhou, J., Ng, A. Y., Tymms, M. J., Jermiin, L. S., Seth, A. K., Thomas, R. S., and Kola, I. (1998) A novel transcription factor, ELF5, belongs to the ELF subfamily of ETS genes and maps to human chromosome 11p13–15, a region subject to LOH and rearrangement in human carcinoma cell lines. *Oncogene* **17,** 2719–2732
22. Lapinskas, E. J., Palmer, J., Ricardo, S., Hertzog, P. J., Hammacher, A., and Pritchard, M. A. (2004) A major site of expression of the ets transcription factor Elf5 is epithelia of exocrine glands. *Histochem. Cell Biol.* **122,** 521–526
23. Bochert, M. A., Kleinbaum, L. A., Sun, L. Y., and Burton, F. H. (1998) Molecular cloning and expression of Ehf, a new member of the ets transcription factor/oncoprotein gene family. *Biochem. Biophys. Res. Commun.* **246,** 176–181
24. Silverman, E. S., Baron, R. M., Palmer, L. J., Le, L., Hallock, A., Subramaniam, V., Riese, R. J., McKenna, M. D., Gu, X., Libermann, T. A., Tugores, A., Haley, K. J., Shore, S., Drazen, J. M., and Weiss, S. T. (2002) Constitutive and cytokine-induced expression of the ETS transcription factor ESE-3 in the lung. *Am. J. Respir. Cell Mol. Biol.* **27,** 697–704
25. Kas, K., Finger, E., Grall, F., Gu, X., Akbarali, Y., Boltax, J., Weiss, A., Oettgen, P., Kapeller, R., and Libermann, T. A. (2000) ESE-3, a novel member of an epithelium-specific ets transcription factor subfamily, demonstrates different target gene specificity from ESE-1. *J. Biol. Chem.* **275,** 2986–2998
26. ENCODE Project Consortium. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* **489,** 57–74
27. Ernst, J., Kheradpour, P., Mikkelsen, T. S., Shoresh, N., Ward, L. D., Epstein, C. B., Zhang, X., Wang, L., Issner, R., Coyne, M., Ku, M., Durham, T., Kellis, M., and Bernstein, B. E. (2011) Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473,** 43–49
28. Escamilla-Hernandez, R., Chakrabarti, R., Romano, R. A., Smalley, K., Zhu, Q., Lai, W., Halfon, M. S., Buck, M. J., and Sinha, S. (2010) Genome-wide search identifies *Ccnd2* as a direct transcriptional target of Elf5 in mouse mammary gland. *BMC Mol. Biol.* **11,** 68–68
29. Zhu, Y., Sun, L., Chen, Z., Whitaker, J. W., Wang, T., and Wang, W. (2013) Predicting enhancer transcription and activity from chromatin modifications. *Nucleic Acids Res.* **41,** 10032–10043
30. Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., Hanna, J., Lodato, M. A., Frampton, G. M., Sharp, P. A., Boyer, L. A., Young, R. A., and Jaenisch, R. (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 21931–21936
31. Ernst, J., and Kellis, M. (2010) Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat. Biotechnol.* **28,** 817–825
32. Rao, S. S., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., Sanborn, A. L., Machol, I., Omer, A. D., Lander, E. S., and Aiden, E. L. (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159,** 1665–1680
33. Zhou, J., Chehab, R., Tkalcevic, J., Naylor, M. J., Harris, J., Wilson, T. J., Tsao, S., Tellis, I., Zavarsek, S., Xu, D., Lapinskas, E. J., Visvader, J., Lindeman, G. J., Thomas, R., Ormandy, C. J., *et al.* (2005) Elf5 is essential for early embryogenesis and mammary gland development during pregnancy and lactation. *EMBO J.* **24,** 635–644
34. Choi, Y. S., Chakrabarti, R., Escamilla-Hernandez, R., and Sinha, S. (2009) Elf5 conditional knockout mice reveal its role as a master regulator in mammary alveolar development: Failure of Stat5 activation and functional differentiation in the absence of Elf5. *Dev. Biol.* **329,** 227–241
35. Chakrabarti, R., Hwang, J., Andres Blanco, M., Wei, Y., Lukačišin, M., Romano, R. A., Smalley, K., Liu, S., Yang, Q., Ibrahim, T., Mercatali, L., Amadori, D., Haffty, B. G., Sinha, S., and Kang, Y. (2012) Elf5 inhibits the epithelial-mesenchymal transition in mammary gland development and breast cancer metastasis by transcriptionally repressing Snail2. *Nat. Cell Biol.* **14,** 1212–1222
36. Hay, E. D. (1995) An overview of epithelio-mesenchymal transformation. *Acta Anat.* **154,** 8–20
37. Thiery, J. P. (2002) Epithelial-mesenchymal transitions in tumour progression. *Nat. Rev. Cancer* **2,** 442–454
38. Ichikawa, D., Handa, K., Withers, D. A., and Hakomori, S. (1997) Histoblood group A/B versus H status of human carcinoma cells as correlated with haptotactic cell motility: approach with A and B gene transfection. *Cancer Res.* **57,** 3092–3096
39. Orlow, I., Lacombe, L., Pellicer, I., Rabbani, F., Delgado, R., Zhang, Z. F., Szijan, I., and Cordón-Cardó, C. (1998) Genotypic and phenotypic characterization of the histoblood group ABO(H) in primary bladder tumors. *Int. J. Cancer* **75,** 819–824
40. Kominato, Y., Hata, Y., Takizawa, H., Tsuchiya, T., Tsukada, J., and Yamamoto, F. (1999) Expression of human histo-blood group ABO genes is dependent upon DNA methylation of the promoter region. *J. Biol. Chem.* **274,** 37240–37250
41. Iwamoto, S., Withers, D. A., Handa, K., and Hakomori, S. (1999) Deletion of A-antigen in a human cancer cell line is associated with reduced promoter activity of CBF/NF-Y binding region, and possibly with enhanced DNA methylation of A transferase promoter. *Glycoconj. J.* **16,** 659–666
42. Gao, S., Worm, J., Guldberg, P., Eiberg, H., Krogdahl, A., Liu, C. J., Reibel, J., and Dabelsteen, E. (2004) Genetic and epigenetic alterations of the blood group ABO gene in oral squamous cell carcinoma. *Int. J. Cancer* **109,** 230–237
43. Yamamoto, F., Cid, E., Yamamoto, M., and Blancher, A. (2012) ABO research in the modern era of genomics. *Transfus. Med. Rev.* **26,** 103–118
44. Zhou, S., and Welsby, I. (2014) Is ABO blood group truly a risk factor for thrombosis and adverse outcomes? *World J. Cardiol.* **6,** 985–992
45. Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., and Zhang, F. (2013) Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8,** 2281–2308
46. Bolger, A. M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina Sequence Data. *Bioinformatics* **30,** 2114–2120
47. Langmead, B., and Salzberg, S. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9,** 357–359
48. Robinson, M. D., and Oshlack, A. (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11,** R25