

# Blood group *ABO* gene–encoded A transferase catalyzes the biosynthesis of FORS1 antigen of FORS system upon Met69Thr/Ser substitution

Emili Cid,<sup>1,\*</sup> Miyako Yamamoto,<sup>1,\*</sup> and Fumiichiro Yamamoto<sup>1,2</sup>

<sup>1</sup>Laboratory of Immunohematology and Glycobiology, Josep Carreras Leukaemia Research Institute, and <sup>2</sup>Programa de Medicina Predictiva i Personalitzada del Càncer, Institut d'Investigació Germans Trias i Pujol, Campus Can Ruti, Badalona, Barcelona, Spain

## Key Points

- Met69Thr/Ser substitutions conferred blood group A glycosyltransferase with weak FS activity to produce FORS1 antigen of the FORS system.
- Met69Thr and LeuGlyGly266-268GlyGlyAla substitutions synergistically increase the acquired FS activity of A glycosyltransferase.

Blood group A/B glycosyltransferases (AT/BTs) and Forssman glycolipid synthase (FS) are encoded by the evolutionarily related *ABO* (A/B alleles) and *GBGT1* genes, respectively. AT/BT and FS catalyze the biosynthesis of A/B and Forssman (FORS1) oligosaccharide antigens that are responsible for the distinct blood group systems of ABO and FORS. Using genetic engineering, DNA transfection, and immunocytochemistry and immunocytometry, we have previously shown that the eukaryotic expression construct encoding human AT, whose LeuGlyGly tripeptide at codons 266 to 268 was replaced with FS-specific GlyGlyAla tripeptide, induced weak appearance of FORS1 antigen. Recently, we have shown that the human AT complementary DNA constructs deleting exons 3 or 4, but not exons 2 or 5, induced moderate expression of FORS1 antigen. The constructs containing both the GlyGlyAla substitution and the exon 3 or 4 deletion exhibited an increased FS activity. Here, we report another molecular mechanism in which an amino acid substitution at codon 69 from methionine to threonine or serine (Met69Thr/Ser) also modified enzymatic specificity and permitted FORS1 biosynthesis. Considering that codon 69 is the first amino acid of exon 5 and that the coinroduction of Met69Thr and GlyGlyAla substitutions also enhanced FS activity, the methionine substitutions may affect enzyme structure in a mode similar to the exon 3 or 4 deletion but distinct from the GlyGlyAla substitution.

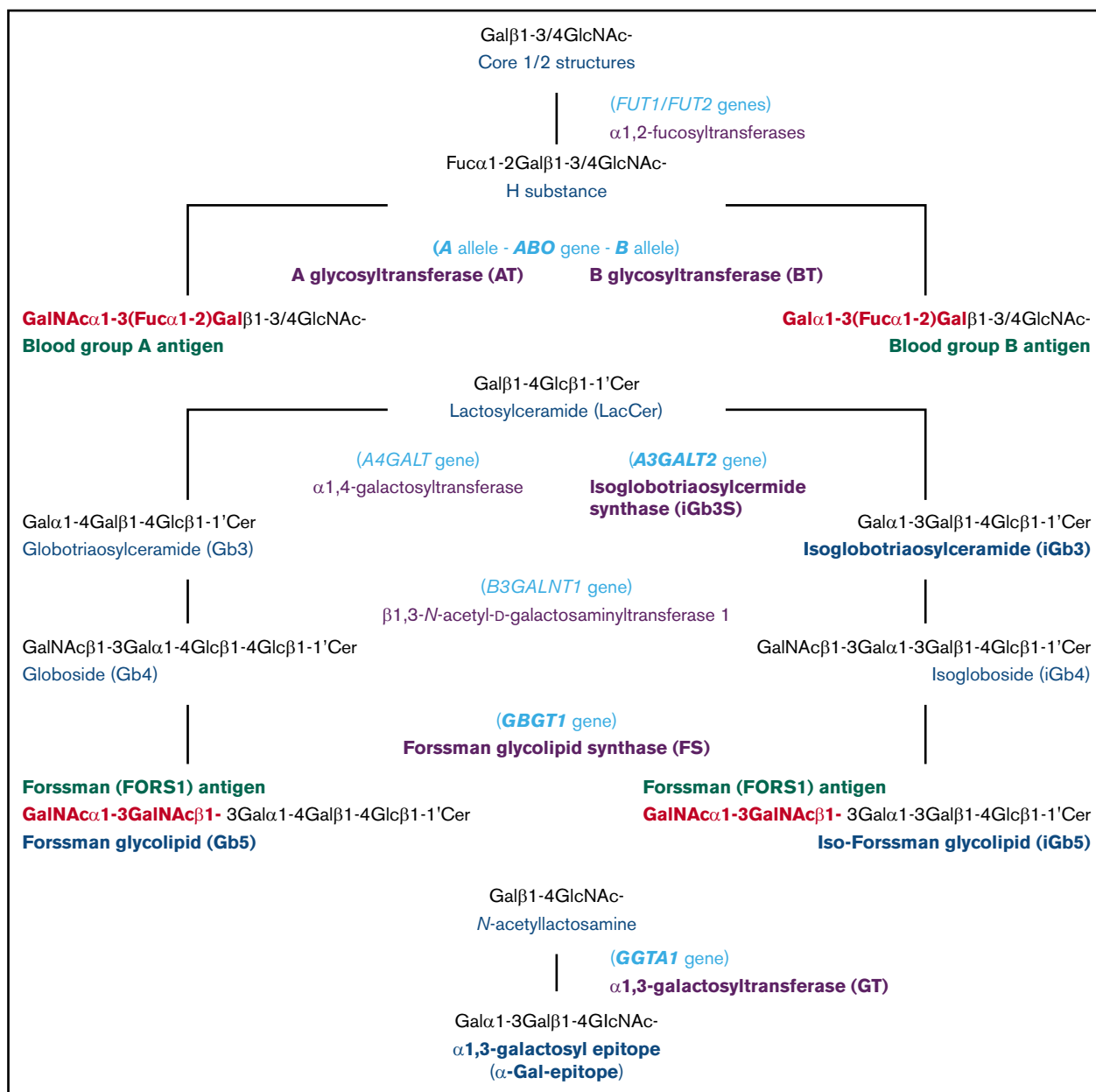
## Introduction

Human individuals can be classified into 4 blood groups, A, B, AB and O, depending on whether one, both, or neither of blood group A and B antigens are present on red blood cells (RBCs). The biochemistry and genetics knowledge underlying this ABO system has been well established. The A and B antigens are oligosaccharide antigens, whose chemical structures are GalNAc $\alpha$ 1-3(Fuc $\alpha$ 1-2)Gal and Gal $\alpha$ 1-3(Fuc $\alpha$ 1-2)Gal-, respectively. They are synthesized from the common precursor substrate H substance (Fuc $\alpha$ 1-2Gal-).<sup>1,2</sup> Their biosynthetic pathways and the structurally related oligosaccharide antigens are represented schematically in Figure 1. The functional A and B alleles at the *ABO* genetic locus encode A and B glycosyltransferases (AT and BT, respectively), which catalyze the last biosynthetic step to form A and B oligosaccharides. Group A individuals have 1 or 2 A alleles encoding AT that synthesize A antigen. Group B individuals have 1 or 2 B alleles encoding BT that synthesize B antigen. Group AB individuals have A and B alleles and express both AT and BT. On the other hand, group O individuals have two non-functional O alleles but no A or B allele. Without modification to A or B antigen, type O RBCs express H substance. In addition to A and B antigens, anti-A and anti-B antibodies are also important components of the ABO system, and the presence or absence of these antibodies is reciprocally specified by the ABO alleles. For instance, group A individuals possess anti-B antibodies,

Submitted 20 February 2018; accepted 21 May 2018. DOI 10.1182/bloodadvances.2018017293.

\*E.C. and M.Y. contributed equally to this study.

© 2018 by The American Society of Hematology



**Figure 1. Schematic representation of biosynthesis of blood group A, B, FORS1, and related antigens.** The biosynthetic pathways of blood group A, B, FORS1, and related glycan antigens are schematically shown. The gene names, glycan names, and transferase names are shown in light blue, dark blue, and purple, respectively. Important genes/transferases/glycans are shown in bold. The immunodominant epitopes of A, B, and FORS1 are shown in red and bold.

whereas group B individuals possess anti-A antibodies. In other words, individuals possess antibodies against the A and/or B antigens they do not express (Landsteiner law). Consequentially, transfusions of ABO mismatched RBCs may induce potentially fatal, acute immune reactions in the recipient.

In 1990, we cloned the human A, B, and O allelic complementary DNAs (cDNAs) and determined the nucleotides characterizing those 3 major alleles. We also identified amino acids differentiating A and B transferases with different sugar specificities of GalNAc and galactose, respectively.<sup>3,4</sup> Furthermore, we correlated the nucleotide and deduced amino acid sequences with A and/or B antigen expression and experimentally demonstrated the central

dogma of ABO: A and B alleles encode A and B transferases, and these enzymes catalyze the biosyntheses of A and B oligosaccharides antigens, respectively. Human AT and BT are both 354-amino-acid proteins but differ by 4 amino acids (arginine [Arg], glycine [Gly], leucine [Leu], and Gly in AT and Gly and serine [Ser], methionine [Met], and alanine [Ala] in BT) at codons 176, 235, 266, and 268, respectively. We prepared 14 AT-BT chimeras that were different at those 4 positions, having either amino acid of AT or BT. We then examined their activity and sugar specificity by DNA transfection into human cancer cells of the uterus (HeLa cell line), which express cell surface H substance.<sup>5</sup> We found that the amino acids at codons 266 and 268 are critical in determining the GalNAc/galactose specificity,

whereas the amino acids at codons 176 and 235 exhibited no and slight effects, respectively. The importance of codons 266 and 268 has also been demonstrated from sugar specificity analyses of additional AT/BT derivatives of ours<sup>6,7</sup> and of others,<sup>8,9</sup> as well as the determination of 3-dimensional structures of AT/BT by X-ray crystallography<sup>10,11</sup> and the structural modeling of AT/BT.<sup>12</sup>

The International Society of Blood Transfusion recognizes a total of 36 RBC blood group systems, the ABO system being just one of these. It should be noted that many of the blood group antigens, especially oligosaccharide antigens, may also be expressed on cell types other than RBCs. Therefore, they may be more accurately referred to as histo-blood group antigens. In the case of A and B antigens, for instance, those antigens are also presented on epithelial cells of the gastrointestinal tract and endothelial cells lining blood capillaries, depending on the individual's ABO group. This is the reason why ABO matching is vital for not only RBC transfusion but also cell, tissue, and organ transplantations. Those antigens may also be expressed on proteins in secretion. The genetically well-defined and conserved A/B antigen expression has been widely used in forensic science. The ABO types of biological specimens, such as blood, skin, saliva, and seminal fluid, found at crime scenes are used as essential markers in identifying or distinguishing suspects in crimes.

In recent years, we have been interested in another blood group system: Forssman (FORS).<sup>13-16</sup> The Forssman antigen (FORS1) was initially recognized as an antigen in guinea pig and sheep, but not in rabbit.<sup>17</sup> Using antibodies against this antigen, species were categorized into FORS1 antigen-positive and antigen-negative species.<sup>18</sup> Mice (*Mus musculus*) belong to the former and humans (*Homo sapiens*) to the latter. However, rare individuals express FORS1 antigens on RBCs. Initially, those individuals were incorrectly categorized into a novel A subgroup, A<sub>pae</sub>. This was based on serological characterization of RBCs showing strong reaction with *Helix pomatia* lectin and weak reaction with polyclonal anti-A reagents but no reaction with monoclonal anti-A antibodies.<sup>19</sup> Glycolipid analysis later showed this to be a misnomer, because the antigen proved to be FORS1 antigen and not a blood group A antigen.<sup>20,21</sup> The immunodominant FORS1 epitope has the structure of GalNAc $\alpha$ 1-3GalNAc $\beta$ 1-, and it is carried on glycolipids such as pentasaccharide Forssman glycolipid (Gb5: GalNAc $\alpha$ 1-3GalNAc $\beta$ 1-3Gal $\alpha$ 1-4Gal $\beta$ 1-4Glc $\beta$ 1-1'Cer). This is in contrast to A and B antigens, which may be carried on either glycoproteins or glycolipids. FORS was recognized as a novel blood group system<sup>21</sup> because it fulfills the following criteria: (1) polymorphic FORS system is specified by functional and nonfunctional alleles at the genetic locus *GBGT1*, which is different from other blood group genes; (2) FORS1 is chemically distinct from other blood group antigens; and (3) there are individuals possessing anti-FORS1 antibodies.

The *GBGT1* gene is evolutionarily related to the *ABO* gene, and the functional *GBGT1* allele encodes Forssman glycolipid synthase (FS), which catalyzes the last biosynthetic step of Forssman glycolipid from its precursor globoside (Gb4: GalNAc $\beta$ 1-3Gal $\alpha$ 1-4Gal $\beta$ 1-4Glc $\beta$ 1-1'Cer) by transferring GalNAc.<sup>22</sup> While we were comparing the deduced amino acid sequences of FS proteins between FORS1 antigen-positive and antigen-negative species, we found many amino acids conserved within a group, but not between the 2 groups.<sup>13</sup> Therefore, we prepared structural chimeric expression constructs between functional mouse FS and nonfunctional human FS proteins,

performed DNA transfection, immunologically detected FORS1 antigen, and identified human *GBGT1* gene cDNA fragments annulling FS activity. We then prepared constructs containing potentially inactivating amino acid substitutions and finally showed that 2 substitutions, c.688G>A [p.Gly230Ser] and c.887A>G [p.Gln(glutamine)296Arg], are responsible for the FS inactivity of the human protein. We also showed that the reversion of either of the 2 substitutions conferred the human FS protein with weak FS activity and that the reversion of both endowed strong FS activity. The c.887G>A[p.Arg296Gln] was later discovered to be a dominant, activating mutation in the *GBGT1* gene of the A<sub>pae</sub> individuals.<sup>21</sup>

There have been additional advances in our FORS/*GBGT1*/FS/FORS1 studies. While analyzing species evolution of the *ABO* and *GBGT1* genes, we realized that the GlyGlyAla tripeptide sequence is conserved in the *GBGT1* gene-encoded FSs at codons corresponding to codons 266 to 268 of human AT/BT.<sup>7</sup> We also found that mouse *ABO* gene-encoded *cis*-AB transferase possesses the same tripeptide sequence.<sup>23</sup> Therefore, we prepared the expression construct, transfected DNA to appropriate cells, and examined whether the murine *cis*-AB transferase exhibits FS activity.<sup>14</sup> We found FORS1 antigen expression on some cells. Furthermore, we observed weak FORS1 expression when cells were transfected with the human AT construct containing the GlyGlyAla substitution. Because this substitution did not confer full FS activity, we searched for additional molecular mechanisms that might allow *ABO* gene-encoded AT to biosynthesize FORS1. Surprisingly, we discovered that the deletion of exons 3 or 4 of human AT cDNA, but not of exons 2 or 5, granted the encoded proteins with moderate FS activity.<sup>16</sup> This crosstalk between *ABO* and *GBGT1* genes may explain, at least partially, why some FORS1-negative individuals express FORS1 in malignant cells/tissues as alterations in messenger RNA (mRNA) splicing are frequent in cancer.<sup>24-26</sup> We also observed an enhanced FS activity by reintroducing the GlyGlyAla substitution with the deletion of exons 3 or 4.

We have expanded the search for other molecular mechanisms. While we were studying the activity and subcellular localization of the *ABO* gene-encoded ATs with various N termini,<sup>27</sup> we prepared and examined a series of methionine (Met) to threonine (Thr) substitution constructs of human AT, including H\_ABO-A(Met1,20,26,43,53,69Thr), which contained the methionine-threonine substitutions at codons 1, 20, 26, 43, 53, and 69. Because those substitution constructs were available, we included them for the FS activity analysis. Unexpectedly, we observed weak FS activity with a couple of the constructs. Further studies have revealed that single amino acid substitutions at codon 69 from methionine to threonine or serine endow human AT with FS capability.

## Methods

### Preparation of amino acid substitution constructs by in vitro mutagenesis

Human AT and derivatives having GlyGlyAla or MetGlyAla tripeptide, mouse *cis*-AB transferase, and mouse FS constructs H\_ABO-A, H\_ABO-A(GlyGlyAla), H\_ABO-A(MetGlyAla), M\_ABO-AB, and M\_GBGT1, respectively, were prepared in the pSG5 eukaryotic expression plasmid vector as previously described.<sup>7,14,16</sup> Additional amino acid substitution constructs were prepared by in vitro mutagenesis, employing the 2-round primer-mediated polymerase chain reaction strategy using primers with nucleotide substitutions.<sup>7,16</sup>

**Table 1. FORS1 antigen expression after DNA transfection of a variety of expression constructs into COS1(B3GALNT1) cells analyzed by immunoperoxidase staining procedure**

Gene name	Tripeptide sequence	Other amino acid substitutions	Adjusted FORS1 antigen-positive cell %			
			Exp 1	Exp 2	Exp 3	FS activity
M_GBGT1	GlyGlyAla	—	100	100	100	+++++
M_ABO-AB	GlyGlyAla	—	68	29	66	++++
H_ABO-A	LeuGlyGly	—	0	0	0	—
		Met1Thr	0	0	0	—
		Met20Thr	0	0	0	—
		Met26Thr	0	0	0	—
		Met43Thr	0	0	0	—
		Met53Thr	0	0	0	—
		Met69Thr	1	1	7	++
		Met142Thr	0	0	0	—
		Met1,69Thr	0	0	2	+
		Met20,26Thr	0	0	0	—
		Met1,20,26Thr	0	0	0	—
		Met1,20,26,43,69Thr	2	0	2	+
		Met1,20,26,53,69Thr	3	2	2	++
		Met1,20,26,43,53,69Thr	2	0	4	+
		Met69Ser	1	0	5	+
	GlyGlyAla	—	2	0	3	+
		Met69Thr	56	119	100	+++++
	MetGlyAla	—	0	0	0	—
		Met69Thr	0	0	0	—

The experiments (Exp) were performed in triplicate. The original gene names are listed in the leftmost column. M\_GBGT1, M\_ABO-AB, and H\_ABO-A are the mouse *GBGT1* gene, mouse *ABO* gene, and the A allele of the human *ABO* gene, respectively. The tripeptide sequences at the codons corresponding to codons 266 to 268 of human AT/BT and the amino acid substitutions introduced are shown in the next 2 columns. The adjusted FORS1-positive cell percentages observed in the triplicate experiments (1-3) are shown in the next 3 columns. The numbers of cells positive with GFP, whose expression vector was cotransfected, were used to normalize FORS1 antigen positivity. The values are shown in percentage of the expression of the original mouse FS construct. The deduced FS activity is shown semiquantitatively in the rightmost column (—, no; +, very weak; ++, weak; +++, moderate; +++++, strong; ++++++, very strong activity).

## DNA transfection

HeLa(FUT2) cells were previously created from HeLa cells to achieve enhanced detection sensitivity of AT and BT activities by retrovirally transducing *FUT2* gene encoding  $\alpha$ 1,2-fucosyltransferase that catalyzes the last biosynthetic step of H substance, the acceptor substrate for AT and BT.<sup>27,28</sup> COS1(B3GALNT1) cells were similarly generated from African green monkey kidney COS1 cells by the modular expression of *B3GALNT1* gene-encoded  $\beta$ 1,3-*N*-acetyl-D-galactosaminyltransferase 1 to produce increased Gb4, the acceptor substrate for FS.<sup>14</sup> Following the manufacturer's protocols, Lipofectamine 2000 and 3000 reagents (Thermo Fisher Scientific) were used for DNA transfection for immunocytometry and immunocytochemistry, respectively. DNA from pEGFP plasmid vector expressing enhanced green fluorescent protein (GFP) or pLL3.7-mRFP vector expressing monomeric red fluorescent protein (mRFP) was co-transfected to estimate transfection efficiency.

## Immunological detection of FORS1 and A and B antigens

DNA-transfected COS1(B3GALNT1) cells were immunostained with clone FOM-1 rat immunoglobulin M (IgM) monoclonal antibody

against FORS1 (BMA Biomedicals). For immunocytochemistry, GFP-expressing cells were counted under fluorescence microscopy 2 days after DNA transfection. The following day cells were fixed on culture plates with 4% paraformaldehyde and thoroughly washed with phosphate-buffered saline. The cells were then treated first with the primary antibody and next with a biotinylated goat anti-rat IgG+IgM(H+L) secondary antibody (Jackson ImmunoResearch Laboratories). After washing, cells were incubated with the Vectastain ABC reagents (Vector Laboratories). For peroxidase-mediated color development, 3,3'-diaminobenzidine tetrahydrochloride from the same company was used. Positive cells were counted under the microscope and normalized using numbers of cells positively expressing cotransfected GFP. The percentage values were calculated using the mouse FS construct (M\_GBGT1) value of 100.

For immunocytometry of DNA-transfected COS1(B3GALNT1) cells, cells were detached from culture plates with 5 mM EDTA in phosphate-buffered saline 1 day after DNA transfection, washed, and subjected to immunostaining, first with the primary antibody and then with Alexa Fluor 488 (AF488)-conjugated goat anti-rat IgM ( $\mu$ -chain specific) secondary antibody (Thermo Fisher Scientific). A BD LSRFortessa Analyzer (BD Biosciences) was used for

**Table 2. Expression of blood group antigens (FORS1, A, and B) after DNA transfection of the selected expression constructs into COS1 (B3GALNT1) and HeLa(FUT2) cells analyzed by FACS**

Gene name	Tripeptide sequence	Amino acid substitution	Cells/antigen					
			COS1(B3GALNT1)/FORS1 antigen		HeLa(FUT2)/A antigen		HeLa(FUT2)/B antigen	
			Pos cell %	MFI value	Pos cell %	MFI value	Pos cell %	MFI value
M_GBGT1	GlyGlyAla	—	9.3	2462	0.1	ND	0.1	ND
H_ABO-A	LeuGlyGly	—	0.4	ND	4.0	6361	0.2	ND
		Met43Thr	0.3	ND	9.6	5617	0.1	ND
		Met53Thr	0.4	ND	10.1	6506	0.0	ND
		Met69Thr	2.1	521	10.0	5558	0.0	ND
		Met69Ser	1.9	218	8.0	9169	0.1	ND
		Met142Thr	0.2	ND	7.6	9449	0.0	ND
	GlyGlyAla	—	0.6	79	5.2	6811	9.9	1307
		Met69Thr	3.5	497	6.8	6730	10.6	1244
	MetGlyAla	—	0.3	ND	0.5	ND	33.9	1493
		Met69Thr	0.2	ND	0.5	ND	42.0	1331
M_ABO-AB	GlyGlyAla	—	4.0	1158	4.0	14349	16.5	1696
Empty vector	—	—	0.3	ND	0.2	ND	0.2	ND

The original construct names, tripeptide sequences, and amino acid substitutions are shown in the 3 leftmost columns. The A, B, and FORS1 antigen-positive cell percentages and MFI values of the antigen-positive cell populations are shown in the 6 rightmost columns. The MFI values are shown for samples whose positive cell percentages surpassed the background level of 0.5%.

ND, not determined; Pos, positive.

fluorescence-activated cell sorting (FACS) analyses, employing the same FACS detection settings for a single set of experiments. Viable single cells were gated based on their forward and side scattering values. Fluorescence was measured at 575 nm to identify cotransfected mRFP expression and determine the transfection efficiency. FORS1 expression (AF488 fluorescence) was detected at 530 nm. FORS1-positive cell percentages and their B530-A median fluorescence intensity (MFI) values were determined. For immunocytometry of DNA-transfected HeLa(FUT2) cells, detached cells were incubated first with the mixture of anti-A murine monoclonal antibodies or the mixture of anti-B murine monoclonal antibodies (Ortho Clinical Diagnostics) and then with Alexa Fluor 647–conjugated goat anti-mouse IgM ( $\mu$  chain specific) secondary antibody (Thermo Fisher Scientific). The immunostained cells were FACS analyzed as described above with COS1(B3GALNT1) cells. The A or B antigen-positive cell percentages and MFI were determined by Alexa Fluor 647 fluorescence at 660 nm and GFP expression at 530 nm.

## Results

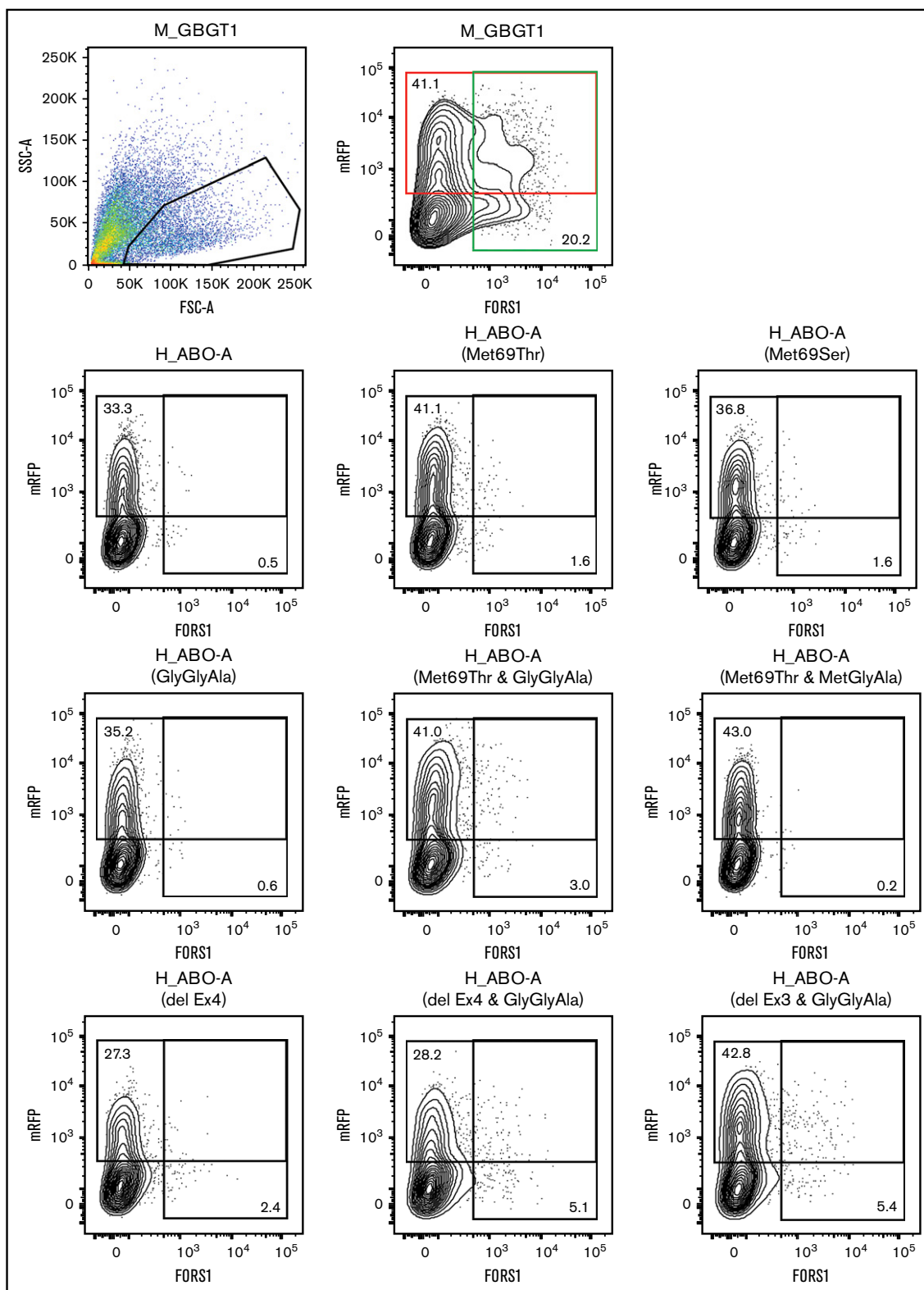
### Methionine-threonine substitution at codon 69, but not codon 1, 20, 26, 43, or 53, confers human AT with FS activity

In an attempt to search for additional molecular mechanisms that might confer human AT with FS activity, we analyzed a couple dozen AT derivatives, including H\_ABO-A(Met1,20,26,43,53,69Thr). We found that some COS1(B3GALNT1) cells transfected with H\_ABO-A(Met1,20,26,43,53,69Thr) expressed FORS1 antigen weakly. Because H\_ABO-A(Met1,20,26,43,53,69Thr) contained 6

methionine-threonine substitutions, we wondered which substitutions might be responsible for FS activity. Several intermediate constructs were prepared during the H\_ABO-A(Met1,20,26,43,53,69Thr) construction through sequential introduction of substitutions, which included H\_ABO-A(Met1,69Thr), H\_ABO-A(Met20,26Thr), H\_ABO-A(Met1,20,26Thr), H\_ABO-A(Met1,20,26,43,69Thr) and H\_ABO-A(Met1,20,26,53,69Thr). We also prepared the constructs afresh with a single substitution, H\_ABO-A(Met1Thr), H\_ABO-A(Met20Thr), H\_ABO-A(Met26Thr), H\_ABO-A(Met43Thr), H\_ABO-A(Met53Thr), H\_ABO-A(Met69Thr), and H\_ABO-A(Met142Thr) containing Met1Thr, Met20Thr, Met26Thr, Met43Thr, Met53Thr, Met69Thr, or Met142Thr, respectively. Codons 1, 20, and 26 are located in the cytoplasmic domain of human AT, whereas codons 43 and 53 are in the transmembrane (TM) region, and codons 69 and 142 in the lumen of Golgi apparatus.<sup>27</sup>

Together with positive controls of mouse *GBGT1* gene construct (M\_GBGT1) and mouse *ABO* gene construct (M\_ABO-AB) and a negative control of human AT construct (H\_ABO-A), we examined the FS activity of H\_ABO-A(Met1,20,26,43,53,69Thr) and 12 other methionine-threonine substitution constructs. Results are shown in Table 1.

We observed cell-surface FORS1 antigen expression on some cells transfected with human ATs with Met1,69Thr, Met1,20,26,43,69Thr, Met1,20,26,53,69Thr, or Met1,20,26,43,53,69Thr. As the Met69Thr substitution was shared among those constructs, this substitution was presumed to be responsible for FS activity. The positive result obtained with H\_ABO-A(Met69Thr), having a single substitution at codon 69, confirmed this prediction. None of the other single methionine-to-threonine substitutions at codons 1, 20, 26, 43, or 53 induced FORS1 antigen expression, suggesting that they were ineffective, although they did not inhibit Met69Thr-induced FS activity of human AT.



**Figure 2. Cytometry analysis of COS1(B3GALNT1) cells transfected with M\_GBGT1, H\_ABO-A, and 8 representative human A transferase derivative constructs.** Cells were cotransfected with DNA from selected expression constructs together with a vector expressing mRFP. Cells were then detached and stained with anti-FORS1 antibody followed by AF488-conjugated secondary antibody. The first 2 panels correspond to M\_GBGT1 transfection. On the left, the selection of single viable cell population by forward (FSC-A) and side (SSC-A) scattering amplitudes is depicted. Events are plotted in pseudocolor, with red representing higher frequency and dark blue the lowest. This selection was maintained for all the other samples. The right panel shows a contour plot of the mRFP fluorescence intensity, used as transfection control,

## Methionine-serine substitution at codon 69 (Met69Ser) also confers human AT with weak FS activity, and cointroduction of Met69Thr and GlyGlyAla substitutions enable strong FS activity

In order to examine whether the threonine residue at codon 69 is essential, we prepared H\_ABO-A(Met69Ser) with serine rather than threonine at codon 69. In addition, we prepared H\_ABO-A (Met69Thr&GlyGlyAla) and H\_ABO-A(Met69Thr&MetGlyAla) with both Met69Thr and FS-specific LeuGlyGly266-268GlyGlyAla or BT-specific LeuGlyGly266-268MetGlyAla substitutions. DNA from those constructs was also transfected, and FS activity was examined (Table 1). Human AT with Met69Ser also exhibited weak FS activity. The combination of Met69Thr with GlyGlyAla enhanced FS activity, suggesting that the molecular mechanisms to confer FS activity are different and synergistic between those 2 modifications. On the other hand, the combination with MetGlyAla eliminated FS activity, indicating that the negative effect of MetGlyAla is dominant over the FS activity of human AT induced by Met69Thr.

## FACS immunocytometry confirmed the immunocytochemistry results

Immunocytochemistry presented semiquantitative results due to cell detachment during immunostaining and difficult microscopical observation of weakly stained cells. To better quantify FS activity, we performed FACS analyses, which permitted more accurate measurements without enzyme-mediated signal amplification. We repeated transfection experiments using several important constructs. We selected human AT and their derivative constructs containing Met43Thr, Met53Thr, Met69Thr, Met69Ser, or Met142Thr substitution and GlyGlyAla or MetGlyAla tripeptide alone or in combination with Met69Thr. They were examined together with M\_GBGT1 and M\_ABO-AB as positive controls, and H\_ABO-A and the empty vector as negative controls. DNA from those 13 constructs was separately transfected into COS1(B3GALNT1) cells, together with DNA from pLL3.7-mRFP vector. After transfection, cells were detached and incubated with anti-FORS1 monoclonal antibody followed by AF488-conjugated secondary antibody. The cells were then analyzed by FACS. Results were compared among constructs and are shown in Table 2. In addition to M\_GBGT1 and M\_ABO-AB, FORS1 expression above background levels was observed for H\_ABO-A(GlyGlyAla), H\_ABO-A (Met69Thr), H\_ABO-A(Met69Ser), and H\_ABO-A(Met69Thr&GlyGlyAla). The FACS results presented a similar tendency in positive cell percentages with the immunocytochemistry results.

## Met69Thr/Ser substitutions confer FS activity without affecting AT and BT activity

In addition to COS1(B3GALNT1) cells, the same set of constructs were transfected to HeLa(FUT2) cells, and AT and BT activity was examined by FACS. A antigen-positive and B antigen-positive HeLa(FUT2) cell percentages and MFI values are also shown in

Table 2. A antigen expression was observed in HeLa(FUT2) cells transfected with M\_ABO-AB and H\_ABO-A constructs with LeuGlyGly or GlyGlyAla irrespective of the presence or absence of Met69Thr/Ser. B antigen expression was observed in cells transfected with M\_ABO-AB and H\_ABO-A constructs with GlyGlyAla or MetGlyAla irrespective of Met69Thr. Some fluctuations were observed, but Met69Thr/Ser did not greatly affect the expression of A and/or B antigens.

## Met69Thr/Ser substitutions may affect human AT in a manner similar to exon 3 or 4 deletion of human AT mRNA

We performed additional FACS analyses of COS1(B3GALNT1) cells transfected with selected constructs. In addition to H\_ABO-A (Met69Thr), H\_ABO-A(Met69Ser), H\_ABO-A(GlyGlyAla), H\_ABO-A(Met69Thr&GlyGlyAla), and H\_ABO-A(Met69Thr&MetGlyAla), the exon deletion constructs were also included for comparison.

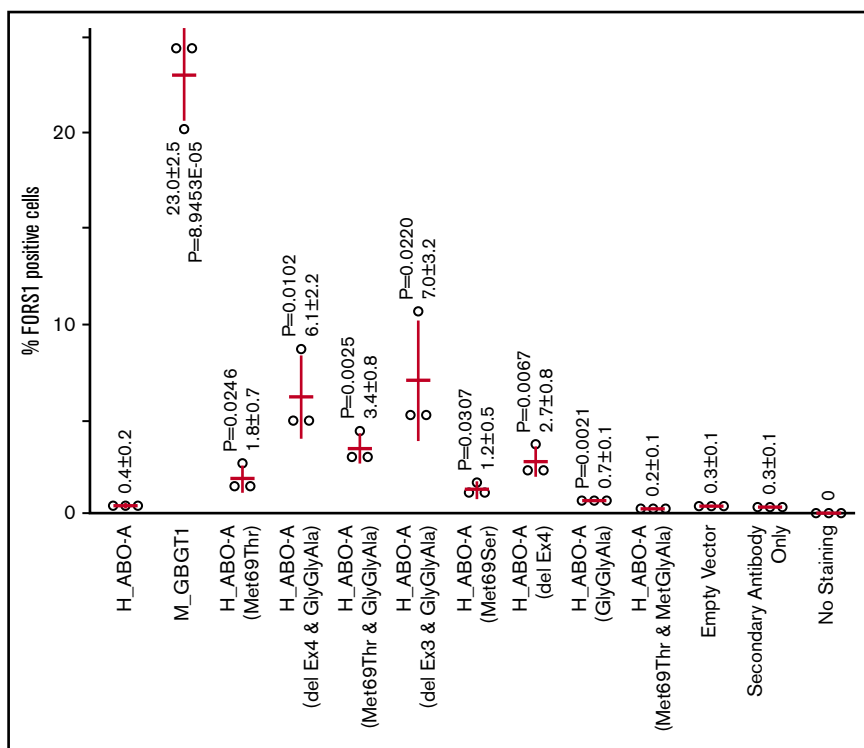
Representative cytometry results are shown in Figure 2. The results clearly show FORS1 appearance after the substitution of the methionine residue with threonine or serine at codon 69 of human AT. We also observe that the double mutants containing Met69Thr and LeuGlyGly266-268GlyGlyAla exhibited larger number of FORS1-positive cells than the single-substitution constructs containing either Met69Thr or LeuGlyGly266-268GlyGlyAla.

FORS1 antigen-positive cell percentages were determined from the cytometry experiments, which were repeated in triplicate. Results are shown in a scatterplot in Figure 3. It was observed that FS activity takes place in human AT with Met69Thr/Ser. FS activity was enhanced by adding the GlyGlyAla substitution, as has been reported of the combinations between exon 3 or 4 deletions and GlyGlyAla substitution. On the other hand, H\_ABO-A(Met69Thr&MetGlyAla) did not induce FORS1 antigen expression. Rather, MetGlyAla substitution nullified the FS activity of human AT induced by Met69Thr.

## Discussion

In living cells and organisms, a variety of enzymes catalyze biological reactions in a specific and efficient manner at ambient pH and temperature. This wide enzymatic repertoire has resulted from vertically transmitted gene duplications, exon shuffling, and divergence. Enzyme-encoding genes may also have been transmitted horizontally from other organisms. Catalytic activities and substrate specificities may have been modified from existing enzymes or created anew from nonfunctional proteins by mutations. In addition, alterations in intracellular localization and protein conformation may have influenced enzymatic activity due to reduced or enhanced substrate availability or utility. Linkages to different promoters, enhancers, or insulators may have triggered gene expression changes.

**Figure 2. (continued)** on the y-axis, and the fluorescence corresponding to FORS1 on the x-axis. Both axes are on logarithmic scales. The upper red rectangle represents mRFP-positive cells, and the percentage of mRFP-positive cells is also indicated at the upper left corner. The rightmost green rectangle corresponds to AF488 fluorescence, and the percentage of FORS1-positive cells is indicated at the lower right corner. The other panels show the graph for the other constructs, using the same gate, second row: H\_ABO-A, H\_ABO-A(Met69Thr), and H\_ABO-A(Met69Ser); third row: H\_ABO-A(GlyGlyAla), H\_ABO-A(Met69Thr&GlyGlyAla), and H\_ABO-A(Met69Thr&MetGlyAla); and fourth row: H\_ABO-A(delEx4), H\_ABO-A(delEx4&GlyGlyAla), and H\_ABO-A(delEx3&GlyGlyAla).



**Figure 3. Expression of FORS1 antigen after DNA transfection of the selected expression constructs into COS1(B3GALNT1) cells analyzed by FACS cytometry.**

After DNA transfection of COS1(B3GALNT1) with a variety of expression constructs and a vector expressing mRFP, cells were immunostained using anti-FORS1 antibody, followed by FACS analysis. The percentages of FORS1-positive cells for each construct and for all 3 independent experiments are depicted as open circles. Their mean  $\pm$  standard deviation is written in numbers and represented on the graph as red bars (horizontal for the mean and vertical for the standard). The construct names are shown under the x-axis. Two-tailed, homoscedastic Student *t* tests comparing each set of data to the empty vector set were performed, and  $P < .05$  is shown.

The evolutionarily related *ABO* and *GBGT1* genes belong to the  $\alpha$ 1,3-Gal(NAc) transferase gene family. The other members include *A3GALT2* gene encoding isoglobotriaosylceramide synthase (iGb3S) involved in the biosynthesis of isoglobotriaosylceramide (iGb3: Gal $\alpha$ 1-3Gal $\beta$ 1-4Glc $\beta$ 1-1'Cer) and *GGTA1* gene encoding  $\alpha$ 1,3-galactosyltransferase (GT) involved in the biosynthesis of  $\alpha$ 1,3-galactosyl epitope ( $\alpha$ -Gal-epitope: Gal $\alpha$ 1-3Gal $\beta$ 1-4GlcNAc).<sup>29</sup> In 2014, we illustrated the  $\alpha$ 1,3-Gal(NAc) transferase gene evolution.<sup>7</sup> The ancestral gene underwent repeated gene duplications, followed by separations into different gene lineages. Gene absence in some species indicates that they are dispensable for species survival. In addition to species variations, polymorphisms may also exist. Therefore, those genes are also dispensable for individual survival. However, interspecies variations and intraspecies polymorphisms may have affected survival, considering that cell-surface glycans play a pivotal role as receptors for bacteria, viruses, and toxins.<sup>30,31</sup>

While studying the activity and subcellular localization of ATs with various N-terminal deletions, we prepared H\_ABO-A (Met1,20,26,43,53,69Thr).<sup>27</sup> This construct encoded AT protein, whose methionine residues in the N-terminal cytoplasmic, TM, and stem regions were all replaced by threonines. We hypothesized that truncated enzymes translated from internal methionines propelled A antigen synthesis in those alleles lacking a proper translation initiation codon.<sup>32</sup> However, the encoded protein exhibited AT activity and induced A antigen expression on HeLa(FUT2) cells after DNA transfection. The N-terminal amino acid sequencing of the encoded protein revealed that the threonine residue at codon 1 was actually the translation initiation codon, demonstrating that codons other than AUG may start protein synthesis. The current study

showed that this construct also exhibited weak FS activity and that the Met69Thr substitution was responsible for the FS activity. Together with our previous finding that the GlyGlyAla substitution at codons 266 to 268 of human AT and the deletion of exon 3 or 4 of AT mRNAs also endowed FS activity,<sup>14,16</sup> it is evident that multiple molecular mechanisms exist allowing the crosstalk between ABO and FORS systems. The fact that FS activity was enhanced by the cointroduction of GlyGlyAla tripeptide and Met69Thr substitutions suggests 2 different synergistic mechanisms. The TM region of human AT/BT spans over codons 33 to 54, which coincides well with codons 33 to 52 contained in exon 3. Codons 53 to 68, present in exon 4, and codon 69, the first amino acid of exon 5, are located in the stem region of the enzymes between the TM region and catalytic domain. Accordingly, the molecular mechanisms may be similar between the deletion of exons 3 or 4 and the Met69Thr/Ser substitution, potentially affecting intra-Golgi localization and/or protein folding.

In addition to H\_ABO-A(Met69Thr&GlyGlyAla), we also analyzed the FS activity of H\_ABO-A(Met69Thr&MetGlyAla). No FORS1 was detected. The LeuGlyGly266-268MetGlyAla substitution of human AT converts the sugar specificity from GalNAc to galactose, producing B antigens in place of A antigens.<sup>5,6</sup> We could not immunologically detect such an epitope, because antibodies that recognize the Gal $\alpha$ 1-3GalNAc $\beta$ 1- structure are not available. Therefore, a possibility exists that AT with Met69Thr and MetGlyAla still utilizes globoside/isogloboside substrates but transfers galactose in place of GalNAc. We have observed that AT and BT activities were not very affected by the Met69Thr/Ser substitution, although weak FS activity was acquired (Table 2). However, the study did not analyze qualitative changes. Therefore, it is possible



that carrier molecule preference may have been altered by those amino substitutions.

In vitro mutagenesis has also been employed to characterize the substrate specificity of other glycosyltransferases that synthesize blood group antigens and otherwise biologically important sugars. For example, Dupuy et al found that a single amino acid at codon 111 in the hypervariable stem domain of vertebrate  $\alpha$ 1,3/1,4-fucosyltransferases determines H-type 1 or H-type 2 transfer specificity.<sup>33</sup> *FUT3*-encoded Lewis  $\alpha$ 1,3/1,4-fucosyltransferases and *FUT5*-encoded  $\alpha$ 1,3/1,4-fucosyltransferases, which can transfer fucose in an  $\alpha$ 1,4-linkage, have tryptophan (Trp) at codon 111, whereas it is Arg in other  $\alpha$ 1,3-fucosyltransferases. They showed that the Trp111Arg substitution of Lewis  $\alpha$ 1,3/1,4-fucosyltransferase changed the specificity from H-type 1 to H-type 2 acceptors. Kaczmarek et al<sup>34</sup> studied human *A4GALT* gene-encoded  $\alpha$ 1,4-galactosyltransferase. Codon 211 is Gln in the majority of individuals, whereas it is glutamic acid (Glu) in rare individuals carrying RBCs expressing terminal disaccharide of the NOR antigen. It was shown that consensus human  $\alpha$ 1,4-galactosyltransferase may synthesize the P<sup>k</sup> and P1 blood group antigens and that its Gln211Glu variant is able to synthesize all the 3 P1PK blood group antigens (P<sup>k</sup>, P1, and NOR). The specificity of  $\beta$ 1,4-galactosyltransferase that synthesize Gal $\beta$ 1-4GlcNAc- disaccharide unit of glycoconjugates was also studied by in vitro mutagenesis.<sup>35</sup> Although the normal enzyme transfers galactose and not GalNAc, the Tyr(tyrosine)289Leu substitution was found to change the donor nucleotide-sugar specificity and to transfer GalNAc to the acceptor GlcNAc. Our finding of the crosstalk between AT and FS is unique in the sense that the acceptor specificities were modified across the gene and the blood group boundary between *ABO* and *GBGT1* and *ABO* and *FORS*. However, it is obvious that the modifications in substrate specificities have been frequent in the evolution of 200-plus glycosyltransferases.

The *FORS1* antigen has been detected in human cancer.<sup>36-42</sup> We have recently proposed that the appearance of genetically inappropriate heterophile *FORS1* antigen in cancer cells/tissues may result from the exon 3 or 4 deletion of human *AT* mRNA,<sup>16</sup> because altered splicing is one of the hallmarks of cancer.<sup>24-26</sup> In contrast, Met69Thr/Ser may not be responsible for this *FORS1* antigen expression in human cancer. We have examined the SNP database, but no SNPs have been found corresponding to the Met69Thr/Ser substitutions in the human *ABO* gene. We also searched the Catalogue of Somatic Mutations in Cancer database for somatic

mutations resulting in the Met69Thr/Ser missense mutations, but none were found. However, it should be noted that other amino acid substitutions at codon 69 or neighboring codons of human *AT* might also confer the FS activity. Together with the conversion issue of sugar specificity discussed above, further studies will be needed to clarify this possibility.

## Acknowledgments

The authors thank Marco Antonio Fernández Sanmartín and Gerard Requena Fernández at the Cytometry Core Facility of the Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol for their help with FACS analyses, Alba Garcia and Albert Perez for their technical assistance, and Michael Maher for editing the manuscript. The work was performed in part at the now-defunct Institut de Medicina Predictiva i Personalitzada del Càncer and in part at the Josep Carreras Leukaemia Research Institute (IJC).

The study was supported by the Spanish Health Research Foundation (Instituto de Salud Carlos III) (PI11/00454), the fund from Agència de Gestió d'Ajuts Universitaris i de Recerca (2014 SGR 1269 and 2017 SGR 529), and institutional start-up funds from Institut de Medicina Predictiva i Personalitzada del Càncer and IJC—"La Caixa" Foundation (F.Y.). IJC and Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol are CERCA centers supported by CERCA Programme/Generalitat de Catalunya. IJC is also supported by the Josep Carreras Foundation.

## Authorship

Contribution: M.Y. and F.Y. conceived the experiments; M.Y., E.C., and F.Y. designed the experiments; M.Y. prepared expression constructs and performed DNA transfection and immunocytochemistry; E.C. performed DNA transfection and FACS immunocytometry; and F.Y. wrote the manuscript with contributions from the other authors.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: E.C., 0000-0002-5025-352X; M.Y., 0000-0001-9516-1402; F.Y., 0000-0001-9690-7034.

Correspondence: Fumiichiro Yamamoto, Josep Carreras Leukaemia Research Institute, Can Ruti Campus, Ctra de Can Ruti, Camí de les Escoles s/n, 08916 Badalona, Barcelona, Spain; e-mail: fyamamoto@carrerasresearch.org.

## References

1. Watkins WM. Biochemistry and Genetics of the ABO, Lewis, and P blood group systems. *Adv Hum Genet.* 1980;10:1-136, 379-385.
2. Watkins WM. The ABO blood group system: historical background. *Transfus Med.* 2001;11(4):243-265.
3. Yamamoto F, Marken J, Tsuji T, White T, Clausen H, Hakomori S. Cloning and characterization of DNA complementary to human UDP-GalNAc: Fuc $\alpha$ 1 $\rightarrow$ 2Gal  $\alpha$ 1 $\rightarrow$ 3GalNAc transferase (histo-blood group A transferase) mRNA. *J Biol Chem.* 1990;265(2):1146-1151.
4. Yamamoto F, Clausen H, White T, Marken J, Hakomori S. Molecular genetic basis of the histo-blood group ABO system. *Nature.* 1990;345(6272):229-233.
5. Yamamoto F, Hakomori S. Sugar-nucleotide donor specificity of histo-blood group A and B transferases is based on amino acid substitutions. *J Biol Chem.* 1990;265(31):19257-19262.
6. Yamamoto F, McNeill PD. Amino acid residue at codon 268 determines both activity and nucleotide-sugar donor substrate specificity of human histo-blood group A and B transferases. In vitro mutagenesis study. *J Biol Chem.* 1996;271(18):10515-10520.
7. Yamamoto F, Cid E, Yamamoto M, Saitou N, Bertranpetit J, Blancher A. An integrative evolution theory of histo-blood group *ABO* and related genes. *Sci Rep.* 2014;4(1):6601.

8. Seto NO, Palcic MM, Compston CA, Li H, Bundle DR, Narang SA. Sequential interchange of four amino acids from blood group B to blood group A glycosyltransferase boosts catalytic activity and progressively modifies substrate recognition in human recombinant enzymes. *J Biol Chem.* 1997; 272(22):14133-14138.
9. Seto NO, Compston CA, Evans SV, Bundle DR, Narang SA, Palcic MM. Donor substrate specificity of recombinant human blood group A, B and hybrid A/B glycosyltransferases expressed in *Escherichia coli*. *Eur J Biochem.* 1999;259(3):770-775.
10. Patenaude SI, Seto NO, Borisova SN, et al. The structural basis for specificity in human ABO(H) blood group biosynthesis. *Nat Struct Biol.* 2002;9(9): 685-690.
11. Alfaro JA, Zheng RB, Persson M, et al. ABO(H) blood group A and B glycosyltransferases recognize substrate via specific conformational changes. *J Biol Chem.* 2008;283(15):10097-10108.
12. Heissigerova H, Breton C, Moravcova J, Imberty A. Molecular modeling of glycosyltransferases involved in the biosynthesis of blood group A, blood group B, Forssman, and iGb3 antigens and their interaction with substrates. *Glycobiology.* 2003;13(5):377-386.
13. Yamamoto M, Cid E, Yamamoto F. Molecular genetic basis of the human Forssman glycolipid antigen negativity. *Sci Rep.* 2012;2(1):975.
14. Yamamoto M, Cid E, Yamamoto F. Crosstalk between ABO and Forssman (FORS) blood group systems: FORS1 antigen synthesis by ABO gene-encoded glycosyltransferases. *Sci Rep.* 2017;7:41632.
15. Yamamoto F. Evolutionary divergence of the ABO and GBT1 genes specifying the ABO and FORS blood group systems through chromosomal rearrangements. *Sci Rep.* 2017;7(1):9375.
16. Yamamoto M, Cid E, Yamamoto F. ABO blood group A transferases catalyze the biosynthesis of FORS blood group FORS1 antigen upon deletion of exon 3 or 4. *Blood Adv.* 2017;1(27):2756-2766.
17. Forssman J. Die Herstellung hochwertiger spezifischer Schafhämonsine ohne Verwendung von Schafblut. Ein Beitrag zur Lehre von heterologer Antikörperbildung. *Biochem Z.* 1911;37:78-115.
18. Leduc EH, Tanaka N. A study of the cellular distribution of Forssman antigen in various species. *J Immunol.* 1956;77(3):198-212.
19. Stamps R, Sokol RJ, Leach M, Herron R, Smith G. A new variant of blood group A. A<sub>pae</sub>. *Transfusion.* 1987;27(4):315-318.
20. Svensson L, Bindila L, Angström J, et al. The structural basis of blood group A-related glycolipids in an A<sub>3</sub> red cell phenotype and a potential explanation to a serological phenomenon. *Glycobiology.* 2011;21(2):162-174.
21. Svensson L, Hult AK, Stamps R, et al. Forssman expression on human erythrocytes: biochemical and genetic evidence of a new histo-blood group system. *Blood.* 2013;121(8):1459-1468.
22. Siddiqui B, Hakomori S. A revised structure for the Forssman glycolipid hapten. *J Biol Chem.* 1971;246(18):5766-5769.
23. Yamamoto M, Lin XH, Kominato Y, et al. Murine equivalent of the human histo-blood group ABO gene is a cis-AB gene and encodes a glycosyltransferase with both A and B transferase activity. *J Biol Chem.* 2001;276(17):13701-13708.
24. Venables JP, Klinck R, Bramard A, et al. Identification of alternative splicing markers for breast cancer. *Cancer Res.* 2008;68(22):9525-9531.
25. He C, Zhou F, Zuo Z, Cheng H, Zhou R. A global view of cancer-specific transcript variants by subtractive transcriptome-wide analysis. *PLoS One.* 2009;4(3):e4732.
26. Shapiro IM, Cheng AW, Flytzanis NC, et al. An EMT-driven alternative splicing program occurs in human breast cancer and modulates cellular phenotype. *PLoS Genet.* 2011;7(8):e1002218.
27. Cid E, Yamamoto M, Yamamoto F. Non-AUG start codons responsible for ABO weak blood group alleles on initiation mutant backgrounds. *Sci Rep.* 2017;7:41720.
28. Cid E, Yamamoto M, Buschbeck M, Yamamoto F. Murine cell glycolipids customization by modular expression of glycosyltransferases. *PLoS One.* 2013; 8(6):e64728.
29. Turcot-Dubois AL, Le Moullac-Vaidye B, Despiau S, et al. Long-term evolution of the CAZY glycosyltransferase 6 (ABO) gene family from fishes to mammals: a birth-and-death evolution model. *Glycobiology.* 2007;17(5):516-528.
30. Garratty G. Blood group antigens as tumor markers, parasitic/bacterial/viral receptors, and their association with immunologically important proteins. *Immunol Invest.* 1995;24(1-2):213-232.
31. Yamamoto F, Cid E, Yamamoto M, Blancher A. ABO research in the modern era of genomics. *Transfus Med Rev.* 2012;26(2):103-118.
32. Seltsam A, Das Gupta C, Bade-Doeding C, Blasczyk R. A weak blood group A phenotype caused by a translation-initiator mutation in the ABO gene. *Transfusion.* 2006;46(3):434-440.
33. Dupuy F, Petit JM, Mollicone R, Oriol R, Julien R, Maftah A. A single amino acid in the hypervariable stem domain of vertebrate alpha1,3/1,4-fucosyltransferases determines the type 1/type 2 transfer. Characterization of acceptor substrate specificity of the lewis enzyme by site-directed mutagenesis. *J Biol Chem.* 1999;274(18):12257-12262.
34. Kaczmarek R, Duk M, Szymczak K, et al. Human Gb3/CD77 synthase reveals specificity toward two or four different acceptors depending on amino acid at position 211, creating P(k), P1 and NOR blood group antigens. *Biochem Biophys Res Commun.* 2016;470(1):168-174.
35. Qasba PK, Ramakrishnan B, Boeggeman E. Structure and function of beta -1,4-galactosyltransferase. *Curr Drug Targets.* 2008;9(4):292-309.
36. Kawanami J. The appearance of Forssman hapten in human tumor. *J Biochem.* 1972;72(3):783-785.
37. Hakomori S, Wang SM, Young WW Jr. Isoantigenic expression of Forssman glycolipid in human gastric and colonic mucosa: its possible identity with "A-like antigen" in human cancer. *Proc Natl Acad Sci USA.* 1977;74(7):3023-3027.
38. Yoda Y, Ishibashi T, Makita A. Isolation, characterization, and biosynthesis of Forssman antigen in human lung and lung carcinoma. *J Biochem.* 1980; 88(6):1887-1890.

39. Mori T, Fujii G, Kawamura A Jr, Yasuda T, Naito Y, Tsumita T. Forssman antibody levels in sera of cancer patients. *Immunol Commun.* 1982;11(3):217-225.
40. Mori E, Mori T, Sanai Y, Nagai Y. Radioimmuno-thin-layer chromatographic detection of Forssman antigen in human carcinoma cell lines. *Biochem Biophys Res Commun.* 1982;108(3):926-932.
41. Uemura K, Hattori H, Ono K, Ogata H, Taketomi T. Expression of Forssman glycolipid and blood group-related antigens A, Le(x), and Le(y) in human gastric cancer and in fetal tissues. *Jpn J Exp Med.* 1989;59(6):239-249.
42. Ono K, Hattori H, Uemura K, Nakayama J, Ota H, Katsuyama T. Expression of Forssman antigen in human large intestine. *J Histochem Cytochem.* 1994;42(5):659-665.