

some preparative chromatographic techniques and ^1H -nuclear magnetic resonance (NMR)¹ spectrometry [14–21]. They reported that milk oligosaccharides contain blood group-related antigens and that their relative abundances are characteristic among animal species. For example, bear milk contains oligosaccharides having an α -Gal epitope (Gal α 1–3Gal β 1–4GlcNAc-R), A blood antigen (GalNAc α 1–3[Fuc α 1–2]Gal-R), B blood antigen (Gal α 1–3[Fuc α 1–2]Gal-R), and Lewis^x antigen (Gal β 1–4[Fuc α 1–3]GlcNAc-R) [19]. Urashima and coworkers also reported that milk samples from bearded and hooded seals contain a large amount of neutral oligosaccharides, including lactose (Gal β 1–4Glc), 2'-fucosyllactose (Fuc α 1–2Gal β 1–4Glc), and lacto-N-fucopentaose (Fuc α 1–2Gal β 1–4GlcNAc β 1–3Gal β 1–4Glc) as major components [17,18]. In addition, both milk samples also contain branched oligosaccharides having lacto-N-neohexaose (LNnH, Gal β 1–4GlcNAc β 1–6[Gal β 1–4GlcNAc β 1–3]Gal β 1–4Glc) as a core, and most of them have one or two nonreducing α 1–2 linked Fuc. Although both milk samples contain sialylated oligosaccharides with high molecular masses, structural studies were not done because of the limited performance of the NMR method. It has also been revealed that milk samples in monotremes such as platypus and echidna contain Lewis^x and Lewis^y antigens (Fuc α 1–2Gal β 1–4[Fuc β 1–3]GlcNAc-R) [10,12,22]. In view of these species-specific structural features and distribution of diverse oligosaccharides in milk/colostrum of different animals, detailed structural studies are not only useful for understanding the underlying evolutionary significance but also promising for using these unique features for biomedical applications.

The dominant carbohydrate in mammalian milk is generally the disaccharide lactose, whereas the milk samples of phocid species, including hooded and bearded seals, contain a variety of oligosaccharides other than lactose [17,18]. The oligosaccharides in bearded and hooded seal milk contain lactose, lacto-N-neotetraose (LNnT, Gal β 1–4GlcNAc β 1–3Gal β 1–4Glc), and LNnH as core units. Furthermore, it is noteworthy that milk oligosaccharides from both species contain type II chain (Gal β 1–4GlcNAc-R) but not type I chain (Gal β 1–3GlcNAc-R). The presence of α 1–3 linked GlcNAc and type II chain suggests that seal mammary glands contain poly-N-acetyllactosamines, which are synthesized by β (1–4)galactosyltransferase as well as β (1–3)N-acetylglucosaminyltransferase. A search for higher oligosaccharides having poly-N-acetyllactosamine structure is an interesting target for understanding the regulation of biosynthesis because they are further modified to form functional oligosaccharides (e.g., sialyl Le^x) and/or branched structures.

NMR spectroscopy is the most important technique that provides sequence information including linkage and α/β -anomeric configurations. However, due to the complexity of extremely overlapping signals of the monosaccharide residues in similar environments, especially in the case of oligosaccharides having poly-Gal β 1–4GlcNAc (N-acetyllactosamine), it is often difficult to assign the branching pattern by only the NMR technique. In contrast, mass spectrometry (MS) has been an indispensable technique for structural analysis of oligosaccharides and useful for the analysis of higher oligosaccharides with high sensitivity. Finke and coworkers reported a method for the analysis of higher oligosaccharides (MW ~ 3000) by a combination of chromatographic

separation and matrix-assisted laser desorption/ionization–time-of-flight (MALDI–TOF) MS [23,24]. Recently, tandem mass spectrometry (MS/MS) with collision-induced dissociation (CID) has been applied to structural analysis of various oligosaccharides [25,26]. Chai and coworkers reported a method for the analysis of a complex mixture of isomeric neutral oligosaccharides in human urine and milk samples by nano-liquid chromatography–electrospray ionization–ion trap (LC–ESI–IT) mass spectrometer and identified three novel isomeric fucosylated lacto-N-hexaoses (LNHs) based on the studies using CID–MS/MS experiments [27,28]. However, the MS method often cannot differentiate isomeric branched or linear oligosaccharides such as LNnH and para-lacto-N-hexaose (pLNH, Gal β 1–4GlcNAc β 1–3Gal β 1–4GlcNAc β 1–3Gal β 1–4Glc). Thus, it is still difficult to characterize anomeric configurations, branching configurations, and epimeric forms. This information is often obtained by the analysis of the digestion products with specific exoglycosidases. The molecular mass obtained from MS analysis after digestion with well-defined exoglycosidases reveals the sequence of oligosaccharides and information on the branching pattern.

In this study, we isolated higher oligosaccharides (MW ~ 3800) from bearded and hooded seal milk samples and analyzed their structural characteristics by normal-phase (NP)–HPLC after desialylation with neuraminidase and also by MALDI–TOF MS. Furthermore, we confirmed the branching pattern of the oligosaccharides by a combination of sequential exoglycosidase digestions and MALDI–TOF MS and MALDI–quadrupole ion trap (QIT)–TOF MS.

Materials and methods

Materials

Milk samples from bearded seal (BS) and hooded seal (HS) were collected from a lactating female in Svalbard, Norway, and from animals on the drifting pack ice in the southern part of the Gulf of St. Lawrence, Canada, respectively. Both milk samples were stored at -20°C until use. α 1–2 Fucosidase derived from *Corynebacterium* sp. and α 1–3,4 fucosidase from *Streptomyces* sp. 142 were purchased from Takara Biochemicals (Kusatsu, Japan). α 2,3,6,8 Neuraminidase from *Arthrobacter ureafaciens* was kindly donated by Yasuhiro Ohta (Marukin Bio, Kyoto, Japan). β -Galactosidase and β -N-acetylhexosaminidase (both from jack beans) were obtained from Seikagaku Kogyo (Tokyo, Japan). All other reagents were analytical or HPLC grade.

Fractions containing acidic oligosaccharides from BS and HS milk samples

Samples of BS and HS milk (40 and 20 ml, respectively) were obtained after delipidation and protein precipitation according to the reported method [17,18]. Briefly, the milk samples were diluted with 4 volumes of distilled water and shaken vigorously with 4 volumes of chloroform/methanol (2:1, v/v). The chloroform layer and denatured protein were discarded. The methanol was removed from the upper layer by a rotary evaporator, and the resulting carbohydrate-containing solution was freeze-dried. Carbohydrate-containing fractions were fractionated on a Biogel P-2 column (2.5 \times 100 cm) previously equilibrated with water. An aliquot (0.5 ml) of each fraction was analyzed for hexose by phenol–sulfuric acid assay and for sialic acids by resorcinol assay [29]. Fractions eluted earlier were pooled and lyophilized to dryness (see Fig. 1 in Ref. [17] and Fig. 1 in Ref. [18]). The neutral oligosaccharides of both milk samples and a part of acidic oligosaccharides of bearded seal milk were already characterized in Urashima and coworkers' previous studies [17,18]. The fraction (1.7 mg), which was eluted

¹ Abbreviations used: NMR, nuclear magnetic resonance; LNnH, lacto-N-neohexaose; LNnT, lacto-N-neotetraose; MS, mass spectrometry; MALDI–TOF, matrix-assisted laser desorption/ionization–time-of-flight; MS/MS, tandem mass spectrometry; CID, collision-induced dissociation; LC–ESI–IT, liquid chromatography–electrospray ionization–ion trap; LNH, lacto-N-hexaose; pLNH, para-lacto-N-hexaose; NP, normal-phase; QIT, quadrupole ion trap; BS, bearded seal; HS, hooded seal; 2AA, 2-aminobenzoic acid; DHB, 2,5-dihydroxybenzoic acid; mw, molecular mass; LNnTD, lacto-N-neotetraodecaose; LNnD, lacto-N-neodecaose; LNnDD, lacto-N-neododecaose; LNnOD, lacto-N-neooctadecaose; GnT, β -N-acetylglucosaminyltransferase; iGnT, β -N-acetylglucosaminyltransferase; iGnT, β (1–6)N-acetylglucosaminyltransferase.

at void volumes, from the HS milk sample was used in the following preparations and characterization of each oligosaccharide.

The fraction from the BS milk sample was further separated by ion exchange chromatography, as indicated in the previous study [18]. The lyophilized material was dissolved in 50 mM Tris-HCl buffer (pH 8.7, 2.0 ml) and subjected to anion exchange chromatography on a DEAE Sephadex A-50 (1.5 × 35 cm). The unadsorbed oligosaccharide fractions were used for structural study of the oligosaccharides in the previous study [18]. The adsorbed oligosaccharides were eluted by linear gradient elution with changing NaCl concentrations from 0 to 0.25 M in the same buffer. Two fractions (BS1 and BS2) obtained by linear gradient elution were pooled and lyophilized to dryness. The lyophilized material was dissolved in water and passed through a Biogel P-2 column (2.5 × 100 cm). The fractions eluted at the void volume were pooled and lyophilized to dryness to yield a mixture of acidic oligosaccharides (2.0 and 2.3 mg of BS1 and BS2, respectively).

Fluorescent labeling of oligosaccharides with 2AA

Fluorescent labeling of oligosaccharides was performed according to the method reported previously [30,31]. Briefly, a solution (250 μ l) of 2-aminobenzoic acid (2AA) and NaBH₃CN, prepared by dissolution of both reagents (30 mg each) in methanol (1 ml) containing 4% CH₃COONa and 2% boric acid, was added to a mixture of oligosaccharides (100 μ g). The mixture was kept at 80 °C for 60 min. After cooling, water (250 μ l) was added, and the mixture was applied to a small column (1 × 50 cm) of Sephadex LH-20 previously equilibrated with 50% aqueous methanol. The earlier eluted fluorescent fractions that contained labeled oligosaccharides were collected and evaporated to dryness. The residue was dissolved in water (1 ml), and the solution was stored at -20 °C until analysis.

Preparation of asialo-oligosaccharides

A mixture of 2AA-labeled acidic oligosaccharides (~10 μ g) was dissolved in 20 mM acetate buffer (pH 5.0, 50 μ l), and neuraminidase (10 mU, 10 μ l) was added to the mixture. After incubation at 37 °C for 24 h, the reaction mixture was kept in the boiling water bath for 5 min. After centrifugation of the mixture at 10,000g for 10 min, a portion of the supernatant was used for the analysis.

α -Fucosidase digestion

A mixture of 2AA-labeled asialo-oligosaccharides (~2 μ g), as described above, was dissolved in 20 mM phosphate buffer (pH 7.5, 50 μ l) for α 1-2 fucosidase digestion or in 20 mM phosphate buffer (pH 6.0, 50 μ l) for α 1-3,4 fucosidase. α 1-2 Fucosidase (40 μ U, 2 μ l) or α 1-3,4 fucosidase (10 μ U, 10 μ l) was added to the mixture. After incubation at 37 °C for 24 h, the reaction mixture was kept in the boiling water bath for 5 min, and centrifuged at 10,000g for 10 min. The supernatant was diluted with water to adjust the volume of 200 μ l. A portion of each solution (5 μ l) was used for NP-HPLC analysis.

Sequential exoglycosidase digestion of oligosaccharides

Each oligosaccharide isolated by NP-HPLC was dissolved in 20 mM citrate buffer (pH 3.5, 8 μ l), and β -galactosidase (1 mU, 2 μ l) was added to the mixture. After incubation at 37 °C for 12 h, the reaction mixture was kept in the boiling water bath for 5 min. After centrifugation of the mixture, the supernatant was diluted with water (10 μ l). A portion of the solution (2 μ l) was analyzed by MALDI-TOF MS. Another portion (5 μ l) was mixed with 30 mM citrate buffer (pH 5.0, 5 μ l) containing β -N-acetylhexosaminidase (5 mU), and the reaction mixture was kept at 37 °C for

12 h. The supernatant was diluted with water (10 μ l), and then a portion of the solution (2 μ l) was also analyzed by MALDI-TOF MS.

Separation of the 2AA-labeled oligosaccharides

HPLC was performed with a Shimadzu apparatus equipped with two LC-6ADvp pumps and an FP-920 fluorescence detector (Waters). Separation was done with an Amide 80 column (TOSOH, 4.6 mm i.d. × 250 mm) using a linear gradient formed by 2% acetic acid in acetonitrile (solvent A) and 5% acetic acid in water containing 3% triethylamine (solvent B). The column was initially equilibrated and eluted with 70% solvent A for 2 min. After 2 min, solvent B was increased to 95% over 80 min and kept for further 20 min [32]. Fluorescence detection was performed at 410 nm by irradiating at 325 nm light.

MALDI-TOF MS

MALDI-TOF mass spectra were acquired with a Voyager DE-PRO mass spectrometer (PE Biosystems, Framingham, MA, USA). A nitrogen laser was used to irradiate samples, and an average shot of 50 times was taken. The instrument was operated in a linear mode at an accelerating voltage of 20 kV. An aqueous sample solution (2 μ l) was mixed with a matrix solution (2 μ l) of 1% 2,5-dihydroxybenzoic acid (DHB) in methanol/water (1:1). The mixture was applied to a polished stainless-steel target and then dried in atmosphere for a few hours.

MALDI-QIT-TOF MS

MALDI-QIT-TOF mass spectra were acquired on an AXIMA-QIT-TOF mass spectrometer (Shimadzu, Kyoto, Japan). A nitrogen laser was used to irradiate samples, and an average shot of 50 times was taken. Argon was used for CID. The instrument was operated in positive and reflectron mode. An aqueous sample solution (2 μ l) was mixed with a matrix solution (2 μ l) of 1% DHB in ethanol/water (1:1), and the mixture was applied to a polished stainless-steel target and dried in atmosphere for a few hours.

Results

Acidic oligosaccharides having high molecular masses in BS and HS milk samples

The method for the preparation of the oligosaccharide samples used in the current study was reported previously [17,18]. Two fractions (BS1 and BS2, 2.0 and 2.3 mg, respectively) from the BS milk sample (40 ml) and a fraction containing acidic oligosaccharides (HS, 1.7 mg) were used in the current study. Because the oligosaccharides from the HS and BS milk samples contained type II chain (Gal β 1-4GlcNAc-R) but not type I chain (Gal β 1-3GlcNAc-R), we add "neo" to all core oligosaccharide structures.

Oligosaccharides obtained from BS1 and BS2 were fluorescently labeled with 2AA and analyzed by MALDI-TOF MS. As shown in Fig. 1A, a large number of ion signals were observed at the range from m/z 1484.8 to m/z 3530.4. In BS1, two major molecular ions were observed at m/z 2362.3 and 2653.3, which were due to monofucosyl LNnH with one and two NeuAc residues (Nac1H4N2F1-2AA and Nac2H4N2F1-2AA). Ions at m/z 2151.0, 3027.9, and 3318.6 are 80 mass units larger than the m/z values of monofucosyl LNnH (theoretical molecular mass [mw] 2071.2), difucosyl lacto-N-neotetraose (LNnTD, theoretical mw 2946.7), and monosialyl difucosyl LNnTD (theoretical mw 3238.6), respectively. These data indicate that these oligosaccharides are substituted with one SO₃H group. In BS2, we observed two major ions at m/z 2337.2 and

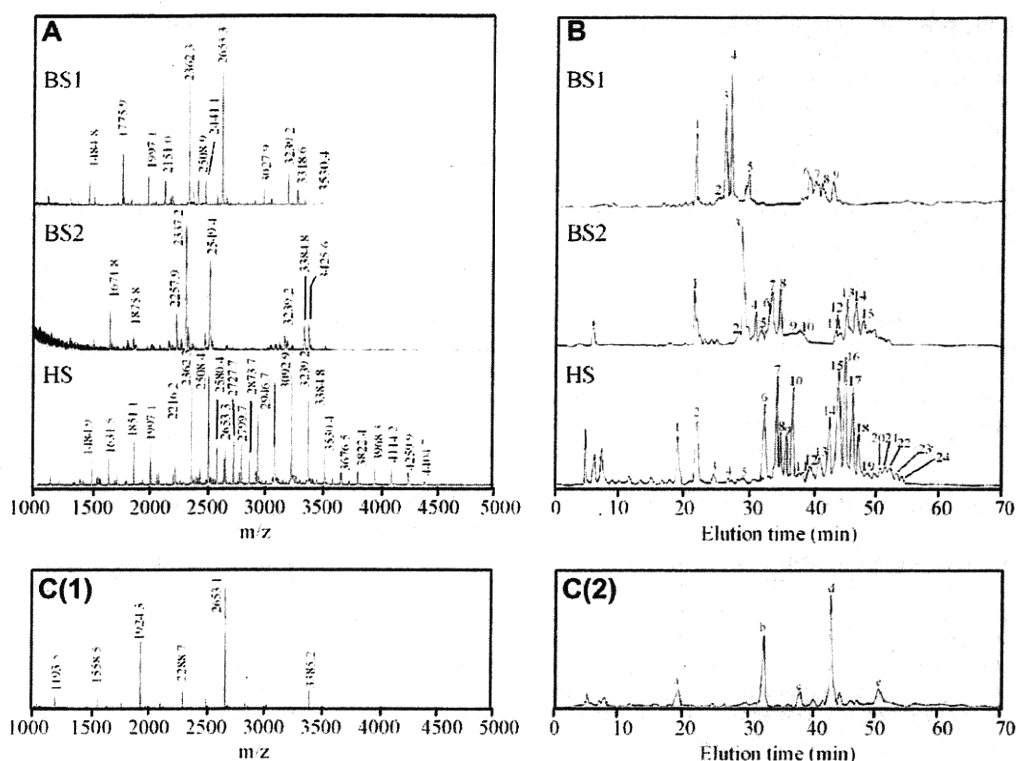


Fig. 1. MALDI-TOF MS and NP-HPLC analysis of higher oligosaccharides from bearded and hooded seal milk. (A) MALDI-TOF MS analysis of sialo-oligosaccharides from bearded and hooded seal milk. (B) NP-HPLC analysis of asialo-oligosaccharides from bearded and hooded seal milk. (C) MALDI-TOF MS and NP-HPLC analysis of defucosylated asialo-oligosaccharides from hooded seal milk. BS1, DEAE-adsorbed fraction 1 in BS milk oligosaccharides; BS2, DEAE-adsorbed fraction 2 in BS milk oligosaccharides; HS, higher oligosaccharide fraction in HS milk oligosaccharides; digestion product of HS with α 1-2 fucosidase. The monosaccharide compositions of asialo-oligosaccharides are summarized in Table 1.

2549.4, which were due to difucosyl decaose with one SO_3H group (H5N5F2- SO_3H -2AA) and monosialo-difucosyl neodecaose with one SO_3H group (Nac1H5N5F2- SO_3H -2AA), respectively. The molecular ion observed at m/z 3384.8 was due to monosialo LNnTD with three Fuc residues (Nac1H8N6F3-2AA).

HS showed characteristic ladder ions between m/z 1400 and m/z 4500. These ladder ions were classified into five groups based on the number of lactosamine (Gal β 1-4GlcNAc) units. The ions observed at m/z 1484.9 and 1631.5 have the composition of Nac1H4N2-2AA and Nac1H4N2F1-2AA, respectively, and are due to mono- and difucosyl LNnH with one NeuAc residue. Molecular ions at m/z 1851.1 and 1997.1 are due to the oligosaccharides having compositions of Nac1H5N3-2AA and Nac1H5N3F1-2AA, respectively. A series of the ions at m/z 2216.2, 2362.3, 2508.4, and 2653.3 were observed abundantly in HS and are due to the oligosaccharides having monosialo lacto-N-neodecaose (LNnD) core (Nac1H6N4-2AA) to which 0, 1, 2, and 3 Fuc residues are attached. Five signals from m/z 2946.4 to m/z 3530.4 were consistent with oligosaccharides of Nac1H8N6-2AA having 0 to 4 Fuc residues. In addition, we found characteristic glycans having extremely large molecular masses, as observed for the series of ions observed from m/z 3676.5 to m/z 4404.7. These oligosaccharides are considered to have the core of LNnTD to which 0 to 5 Fuc residues are attached.

To determine linkages of Fuc residues in HS oligosaccharides, we carried out specific fucosidase digestion. A mixture of HS asialo-oligosaccharides was digested with α 1-3/4 fucosidase and α 1-2 fucosidase, respectively, and the products were analyzed with MALDI-TOF MS. We found that α 1-3/4 fucosidase did not act on these oligosaccharides, whereas digestion with α 1-2 fucosidase (*Corynebacterium* sp.) caused disappearance of most ions, and six ions were observed at m/z 1193.5, 1558.5, 1924.3, 2288.7,

2653.1, and 3385.2 (Fig. 1C(1)). These ions are consistent with the theoretical m/z values of H4N2-2AA, H5N3-2AA, H6N4-2AA, H7N5-2AA, H8N6-2AA, and H10N8-2AA, respectively.

The oligosaccharides obtained from milk samples were also analyzed by HPLC using a TSK-Gel Amide-80 column after removing sialic acids with neuraminidase to improve resolution (Fig. 1B) [32]. We collected the major peaks and observed the molecular ions by MALDI-TOF MS. The results are summarized in Table 1.

BS1-1 obtained from the BS1 fraction was assigned as LNnH having two LacNAc units and core Gal β 1-4Glc unit from its molecular ion (m/z 1193.5) (Fig. 1B, top panel). Molecular ions (m/z 2071.8) of BS1-3 and BS1-4 are consistent with the theoretical m/z values of H6N4F1-2AA, suggesting the presence of isomers of monofucosyl LNnD. The peak observed at 30 min (BS1-5) gave two molecular ions (m/z 2802.6 and 2949.0). The molecular ion at m/z 2802.6 is consistent with the theoretical m/z value of H8N6F1, suggesting the structure of monofucosyl LNnTD. Likewise, the molecular ion at m/z 2949.0 was assigned as difucosyl LNnTD. Minor peaks (BS1-6 to BS1-9) were due to the oligosaccharides having a core structure of LNnD or LNnTD to which a sulfate group is attached (for confirmation of the structure, see the following section).

In the BS2 fraction, BS2-1 and BS2-3 are composed of H3N3F1-2AA and H5N5F2-2AA, respectively. Digestion of BS2-1 and BS2-3 with α 1-2 fucosidase caused the loss of one and two fucose residues, respectively. The defucosylated oligosaccharides gave molecular ions corresponding to the theoretical m/z values of H3N3-2AA and H5N5-2AA. From the monosaccharide compositions, these oligosaccharides are considered to be hexa- and decasaccharide, having LacNAc units at the reducing end (for confirmation of the structure, see the following section). Oligosaccharides observed be-

Table 1
List of asialo-oligosaccharides observed in bearded and hooded seal milk.

Peak number	Observed mass	Calculated mass	Composition
(a) BS1			
-1	1193.5	1194.1	H4N2-2AA
-2	2110.8	2111.9	H5N5F1-2AA
-3	2071.8	2071.2	H6N4F1-2AA
-4	2071.8	2071.2	H6N4F1-2AA
-5	2802.6	2801.7	H8N6F1-2AA
	2949.0	2947.9	H8N6F2-2AA
-6	2151.0	2151.2	H6N4F1-SO ₃ H-2AA
	3027.0	3027.9	H8N6F2-SO ₃ H-2AA
-7	2150.8	2151.2	H6N4F1-SO ₃ H-2AA
	3026.9	3027.9	H8N6F2-SO ₃ H-2AA
-8	2150.5	2151.2	H8N6F2-SO ₃ H-2AA
	3026.6	3027.9	H6N4F1-SO ₃ H-2AA
-9	2151.2	2151.2	H6N4F1-SO ₃ H-2AA
(b) BS2			
-1	1381.6	1381.3	H3N3F1-2AA
-2	2111.6	2111.9	H5N5F1-2AA
-3	2259.5	2258.2	H5N5F2-2AA
-4	2217.7	2217.7	H6N4F2-2AA
-5	2217.2	2217.7	H6N4F2-2AA
	2363.5	2363.7	H6N4F3-2AA
-6	2988.7	2988.0	H7N7F2-2AA
	3134.6	3134.1	H7N7F3-2AA
-7	3134.0	3134.1	H7N7F3-2AA
-8	2948.7	2947.9	H8N6F2-2AA
-9	3094.4	3093.5	H8N6F3-2AA
-10	3240.5	3239.2	H8N6F4-2AA
-11	2338.1	2338.2	H5N5F2-SO ₃ H-2AA
-12	2338.2	2338.2	H5N5F2-SO ₃ H-2AA
-13	2337.8	2338.2	H5N5F2-SO ₃ H-2AA
-14	2337.5	2338.2	H5N5F2-SO ₃ H-2AA
-15	3215.5	3215.1	H7N7F3-SO ₃ H-2AA
(c) HS			
-1	1193.2	1194.1	H4N2-2AA
-2	1339.3	1340.4	H4N2F1-2AA
-3	1558.6	1559.0	H5N3-2AA
-4	1704.9	1705.4	H5N3F1-2AA
-5	1761.9	1762.8	H5N4-2AA
-6	1923.9	1924.4	H6N4-2AA
-7	2068.7	2070.0	H6N4F1-2AA
-8	2069.5	2070.0	H6N4F1-2AA
-9	2215.2	2216.3	H6N4F2-2AA
-10	2215.5	2216.3	H6N4F2-2AA
-11	2288.8	2289.4	H7N5-2AA
	2363.5	2362.3	H6N4F3-2AA
-12	2435.6	2435.5	H7N5F1-2AA
	2492.7	2492.4	H7N6-2AA
-13	2580.9	2581.5	H7N5F2-2AA
	2637.0	2638.5	H7N6F1-2AA
-14	2652.7	2654.0	H8N6-2AA
-15	2798.6	2799.7	H8N6F1-2AA
-16	2944.5	2945.7	H8N6F2-2AA
-17	3089.9	3091.7	H8N6F3-2AA
-18	3236.4	3237.7	H8N6F4-2AA
-19	3236.7	3237.7	H8N6F4-2AA
-20	3382.8	3384.1	H10N8-2AA
-21	3528.8	3530.0	H10N8F1-2AA
-22	3674.6	3676.3	H10N8F2-2AA
-23	3819.8	3822.0	H10N8F3-2AA
-24	3820.0	3822.0	H10N8F3-2AA
(d) Defucosyl HS			
a	1193.5	1194.1	H4N2-2AA (lacto-N-neohexaose)
b	1924.2	1924.4	H6N4-2AA (lacto-N-neodecaose)
c	2288.4	2289.4	H7N5-2AA (lacto-N-neododecaose)
d	2653.1	2654.0	H8N6-2AA (lacto-N-neotetradecaose)
e	3382.5	3384.1	H10N8-2AA (lacto-N-neooctadecaose)

tween 30 and 38 min (BS2-4 to BS2-10) are considered to have multi-Fuc residues. BS2-4 and BS2-5 have the core structure of

the LNNH unit and contain Fuc α 1-2 residues at the nonreducing ends because these fucose residues were specifically released by digestion with α 1-2 fucosidase (data not shown). BS2-6 and BS2-7 showed molecular ions at m/z 2988.7 and 3134.6, respectively, which correspond to the compositions of H7N7F2-2AA and H7N7F3-2AA. Peaks BS2-8 to BS2-10 were LNNTD containing multiple Fuc α 1-2 residues. As a group of characteristic oligosaccharides in BS2, oligosaccharides having 80 mass units larger than BS2-3 were observed between 42 and 49 min. Oligosaccharides (BS2-11 to BS2-14) showed molecular ions at m/z 2338.2, indicating the composition of H5N5F2-SO₃H-2AA. These oligosaccharides are considered to be isomers having both Fuc and sulfate groups at different positions.

We found 24 oligosaccharide peaks in total in the HS milk sample. These oligosaccharides had LNNH, LNNd, lacto-N-neodecaose (LNNDD), LNNTD, and lacto-N-neooctadecaose (LNNOD) as core structures (Table 1, part c). HS-1 and HS-2 observed at 20.0 and 21.5 min, respectively, gave molecular ions at m/z 1193.2 and 1339.3, which correspond to H4N2-2AA and H4N2F1-2AA, respectively. HS-6, -7, -8, -9, and -10 showed molecular ions at m/z 1923.9, 2068.7, 2069.5, 2215.2, and 2215.5, respectively. The molecular ion of HS-6 is consistent with the theoretical mw of H6N4-2AA, suggesting the structure of LNNd. HS-7/8 (m/z 2068.7/2069.5) and HS-9/10 (m/z 2215.2/2215.5) showed larger molecular ions than those of HS-6 by one Fuc (146 mass units) and two Fuc (292 mass units), respectively. From these results, we concluded that these oligosaccharides had the core structure of LNNd to which different numbers of Fuc residues were attached (for confirmation of the structures, see below). The most abundant group of peaks (HS-14 to HS-19) commonly contains LNNTD (HS-14 at m/z 2652.7) as the core structure. HS-15, -16, -17, and -18/19 showed m/z values larger than LNNTD by one (146 mass units) to four (584 mass units) Fuc residues. These results indicate that HS-14 to HS-19 have LNNTD unit to which different numbers of Fuc residues are attached. The peaks (HS-20 to HS-24) having high molecular weights (m/z 3382.8 to 3820.0) were also observed between 50 and 54 min. These ladder peaks contained LNNOD (theoretical mw 3384.1) as the core structure to which one to four fucose residues are attached.

Urashima and coworkers reported that GlcNAc residues of LNNt and LNNH units in BS and HS oligosaccharides are not fucosylated. In contrast, most GlcNAc residues in bear milk oligosaccharides are fucosylated at OH-3 [17,18]. After digestion of asialo-oligosaccharides derived from HS with α 1-3,4 fucosidase from *Streptomyces* sp. 142 or α 1-2 fucosidase from *Corynebacterium* sp., the digestion products were analyzed by NP-HPLC. α 1-3,4 Fucosidase did not act on the HS oligosaccharides, indicating that the oligosaccharides were not substituted at OH-3/4 on GlcNAc residues with fucose residues (data not shown). In contrast, most peaks disappeared on digestion with α 1-2 fucosidase, and five peaks were observed at 19 min (peak a), 32 min (peak b), 38 min (peak c), 43 min (peak d), and 51 min (peak e) (Fig. 1C and Table 1, part d). These data indicated that all core oligosaccharides observed in Fig. 1C(2) were composed of one reducing terminal lactose and 2 to 8 LacNAc units. Peaks a and b showed molecular ions at m/z 1193.5 and 1924.2, which correspond to the molecular masses of LNNH and LNNd, respectively. Peak c showed a molecular ion at m/z 2288.4 of LNNDD. Peak d was the most abundant one in HS and showed the molecular ion of LNNTD at m/z 2653.1. In a similar manner, we confirmed that peak e was due to LNNOD.

Characterization of the branching pattern of BS oligosaccharides

The structures of dominant oligosaccharides in BS1 (BS1-1, -3, and -4 in Fig. 1B) were easily assigned. Digestion of BS1-1 with β -galactosidase from jack beans caused loss of two galactose res-

idues ($\Delta m/z$ 324). Further digestion with β -N-acetylhexosaminidase gave a molecular ion at m/z 461 corresponding to lactose with 2AA (data not shown). From the data, we concluded that BS1-1 was substituted with two Gal-GlcNAc residues at Gal OH-6 and Gal OH-3 of lactose. The BS1-3 and BS1-4 were digested with α 1-2 fucosidase to afford an ion at m/z 1923.5 corresponding to L_NnD. Digestion of the defucosylated oligosaccharide with β -galactosidase caused the loss of three galactose residues ($\Delta m/z$ 486), and the product showed a molecular ion at m/z 1435. From these results, we concluded that the core oligosaccharide of BS1-3 and BS1-4 is substituted with two LacNAc units on either branch of L_NnH (data not shown). Oligosaccharides observed between 38 and 42 min gave molecular ions H6N4F1-SO₃H-2AA and H8N6F1-SO₃H-2AA (m/z 2151.2 and 3026.6, respectively). Among these oligosaccharides, we obtained BS1-9 as nearly pure state (Fig. 2). Digestion of the BS1-9 with α 1-2 fucosidase caused the loss of one fucose residue and gave a molecular ion corresponding to L_NnD (m/z 2005.1) with a sulfate group. Serial digestions of the defucosylated oligosaccharide with β -galactosidase and β -N-acetylhexosaminidase caused the loss of two LacNAc units and gave a molecular ion corresponding to the composition of H4N2-SO₃H-2AA (m/z 1275.6). These results indicated that the defucosylated oligosaccharide has two nonsubstituted Gal residues at the nonreducing ends. Further digestion of the oligosaccharide with β -galactosidase gave a molecular ion, H3N2-SO₃H-2AA (m/z 1113.3). Urashima and coworkers reported that some oligosaccharides in BS milk were sulfated at the nonreducing terminal Gal OH-3 [18]. From this report and our observations, we concluded that the oligosaccharides from BS1-6 to BS1-9 were due to L_NnD and L_NnTD substituted with one sulfate group at the OH-3 position of the nonreducing terminal Gal residue.

Structures of dominant oligosaccharides in BS2 (BS2-1 and BS2-3) were confirmed in a similar manner. Digestion of BS2-1 with α 1-2 fucosidase caused the loss of one fucose residue (Fig. 3A). Further digestion with β -galactosidase gave a molecular ion (m/z 911.4) corresponding to H3N3-2AA. Finally, digestion with β -N-acetylhexosaminidase gave a molecular ion at m/z 505.1 corresponding to H1N1-2AA. Accordingly, we concluded that the core disaccharide at the reducing end in BS2-1 was Gal-GlcNAc and that BS2-1 was a hexasaccharide substituted with two LacNAc units at OH-6 and OH-3 of Gal residue of the terminal Gal-GlcNAc. Oligosaccharide BS2-3 was also digested with α 1-2 fucosidase to give a glycan showing the molecular ion at m/z 1966.5 corresponding to H5N5-2AA (Fig. 3B). Digestion of the core oligosaccharide with β -galactosidase caused the loss of three galactose residues ($\Delta m/z$ 486). Further digestion with β -N-acetylhexosaminidase gave a molecular ion (m/z 869.9) corresponding to H2N2-2AA. Thus, we concluded that one of the branched units on BS2-1 was further substituted with two LacNAc residues. The characteristic oligosaccharides from BS2-11 to BS2-14 showed molecular ions at m/z 2238.1, which are consistent with the composition of H5N5F2-SO₃H-2AA. These oligosaccharides were digested with α 1-2 fucosidase to produce a signal at m/z 2046.3 corresponding to H5N5-SO₃H-2AA. Further digestion of the defucosylated oligosaccharide with β -galactosidase caused the loss of two Gal residues. These observations indicated that the oligosaccharides from BS2-11 to BS2-14 have two LacNAc branches substituted with α 1-2 Fuc residue (data not shown).

Characterization of the branching pattern of HS oligosaccharides

Digestion of the core oligosaccharide (A, peak a in Fig. 1C(2)) with β -galactosidase caused the loss of two galactose residues

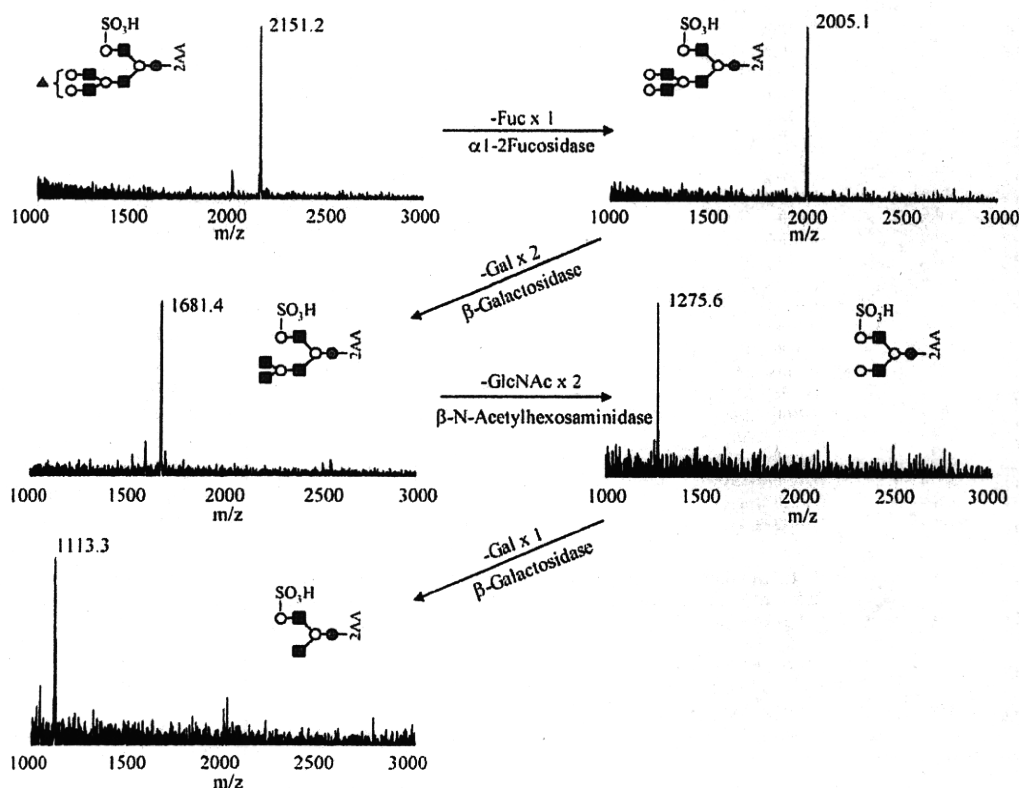


Fig. 2. Stepwise exoglycosidase digestion of characteristic oligosaccharide BS1-9 observed in Fig. 1B. Conditions for the enzymatic reaction with exoglycosidases are shown in Materials and methods. Symbols: open circles; Gal; filled circles, Glc; filled squares, GlcNAc; filled triangles, Fuc. Linkage positions are assigned tentatively.

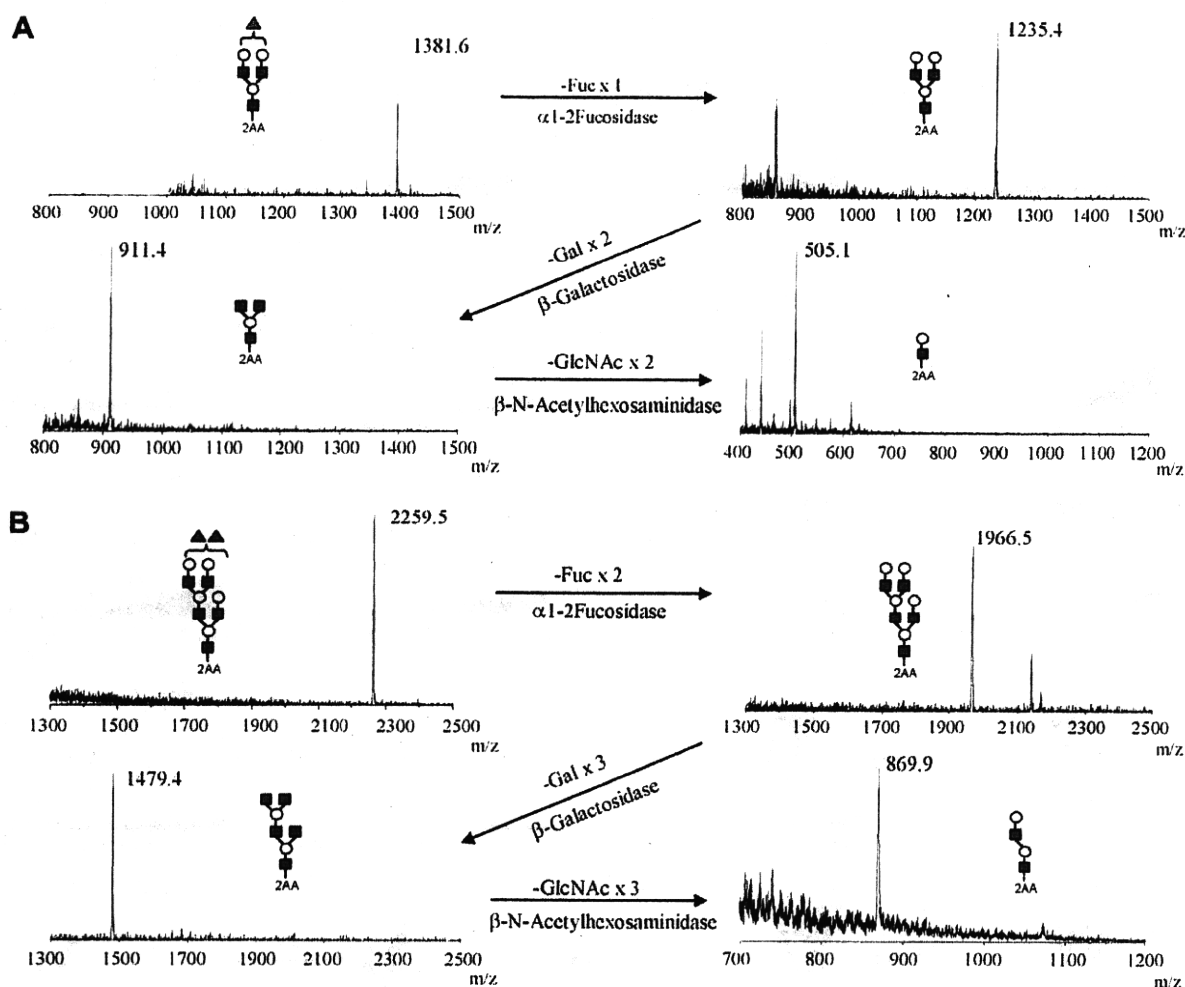


Fig. 3. Stepwise exoglycosidase digestion of characteristic oligosaccharides BS2-1 (A) and BS2-3 (B) observed in Fig. 1B. Symbols: open circles, Gal; filled circles, Glc; filled squares, GlcNAc; filled triangles, Fuc. Linkage positions are assigned tentatively.

($\Delta m/z$ 324) and gave a product ion at m/z 868.8 (Fig. 4A). The oligosaccharide at m/z 868.8 was further digested with β -N-acetylhexosaminidase, and a new ion corresponding to Gal-Glc-2AA was observed at m/z 462.4. From these observations, we concluded that the oligosaccharide (peak a) has the structure of Gal β 1-4GlcNAc β 1-6[Gal β 1-4GlcNAc β 1-3]Gal β 1-4Glc (LNnH). Digestion of the core oligosaccharide (B, peak b in Fig. 1C(2)) with β -galactosidase caused the loss of three galactose residues ($\Delta m/z$ 486), and the product showed a molecular ion at m/z 1437.5 (Fig. 4B). The oligosaccharide at m/z 1437.5 was further digested with β -N-acetylhexosaminidase to release three GlcNAc residues. These observations indicated that the core oligosaccharide (peak b) has a triantennary structure. The produced oligosaccharide corresponding to Gal-GlcNAc-Gal-Glc-2AA (m/z 826.7) was again digested with β -galactosidase to produce a peak at m/z 665.1 (GlcNAc-Gal-Glc-2AA). The structure was confirmed by comparison of the retention time with that of trisaccharide (GlcNAc β 1-3Gal β 1-4Glc-2AA) prepared by digestion of lacto-N-tetraose (Gal β 1-3GlcNAc β 1-3Gal β 1-4Glc) with β -galactosidase using HPLC on an ODS column and capillary electrophoresis (data not shown). These observations indicated that the core oligosaccharide (peak b) has two LacNAc units on the α 1-3 LacNAc branch of LNnH. Digestion of the core oligosaccharide (C, peak c in Fig. 1C(2)) with β -galactosidase caused the loss of three Gal residues ($\Delta m/z$ 486), and the

product showed a molecular ion at m/z 1801.9. The oligosaccharide at m/z 1801.9 was further digested with β -N-acetylhexosaminidase to produce a molecular ion at m/z 1193.8. The oligosaccharide has the structure of H6N4-2AA. These results indicated that peak c has a triantennary structure. The oligosaccharide (m/z 1193.8) was again digested with β -galactosidase to produce a peak at m/z 869.1 by the loss of two galactose residues. The course of digestion by a combination of exoglycosidases revealed that peak c was a dodecasaccharide having three LacNAc residues at nonreducing ends, and we concluded that the oligosaccharide of peak c has two LacNAc units and one LacNAc unit on both branches of LNnH. Digestion of the core oligosaccharide (D, peak d in Fig. 1C(2)) with β -galactosidase caused the loss of four galactose residues ($\Delta m/z$ 648), and the product showed a molecular ion at m/z 2004.5. The product was further digested with β -N-acetylhexosaminidase to produce an ion at m/z 1193.8 corresponding to LNnH. These observations indicated that peak d has a tetraantennary structure. The produced oligosaccharide corresponding to LNnH was again digested with β -galactosidase to produce a molecular ion at m/z 869.1. From these results, we concluded that peak d was a tetradecasaccharide having four LacNAc residues at the nonreducing ends and that both branches of LNnH were substituted with two LacNAc units. The oligosaccharide (E, peak e in Fig. 1C(2)) having the largest molecular mass (m/z 3382.5) present in HS milk caused

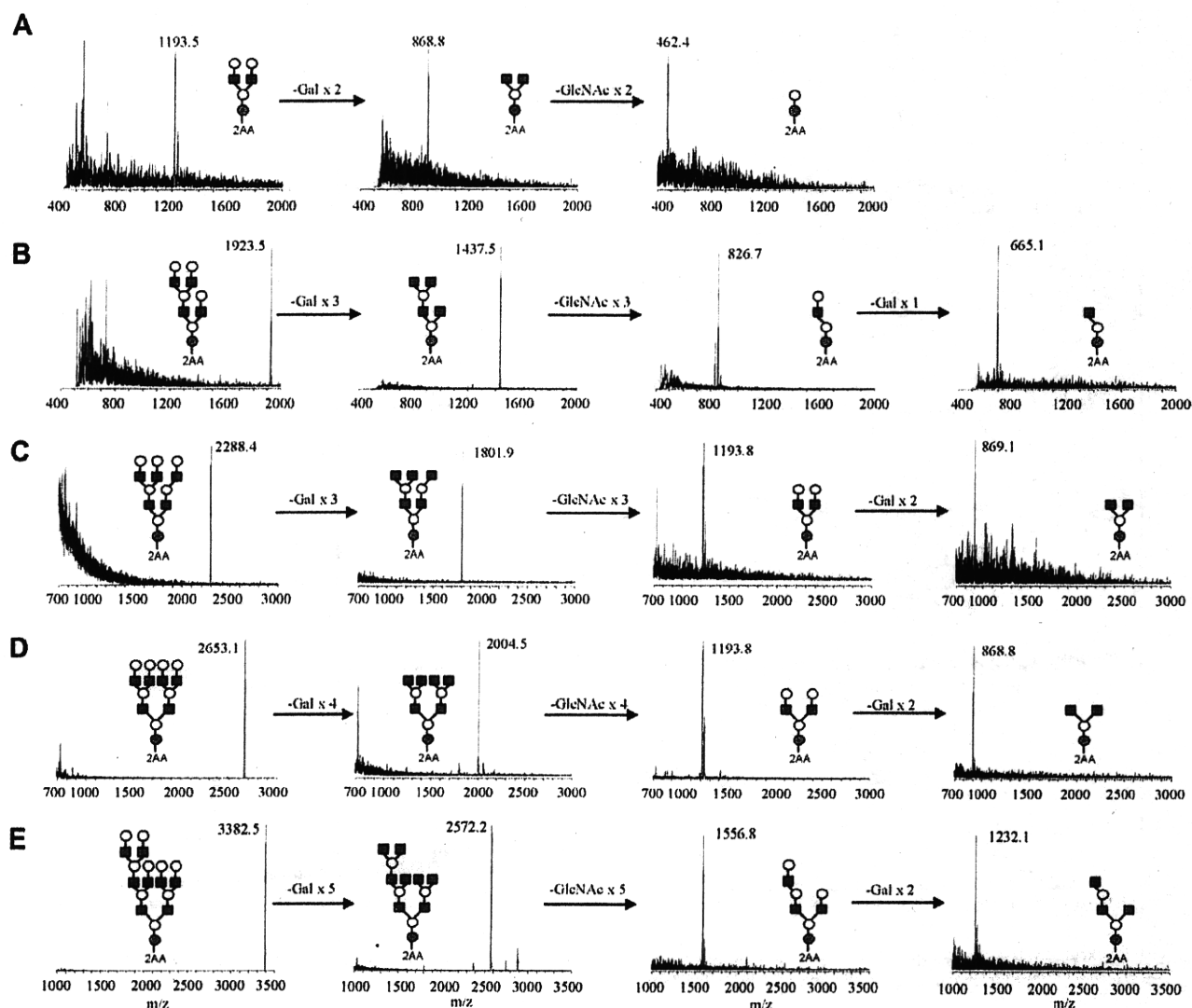


Fig. 4. Stepwise exoglycosidase digestion of core oligosaccharides derived from HS observed in Fig. 1B: (A) lacto-*N*-neohexaose (LNnH); (B) lacto-*N*-neodecaose (LNnD); (C) lacto-*N*-neododecaose (LNnDD); (D) lacto-*N*-neotetradecaose (LNnTD); (E) lacto-*N*-neooctaodecaose (LNnOD). Symbols: open circles, Gal; filled circles, Glc; filled squares, GlcNAc.

the loss of five galactose residues ($\Delta m/z$ 810) by digestion with β -galactosidase, and the product showed a molecular ion at m/z 2572.2. The product was further digested with β -*N*-acetylhexosaminidase to produce a molecular ion at m/z 1556.8. The oligosaccharide was again digested with β -galactosidase to produce a peak at m/z 1232.1. From these results, we concluded that the oligosaccharide derived from peak e was an octadecasaccharide, as shown in Fig. 4E.

Structural determination of fucosylated deca-saccharides by MALDI-QIT-TOF MS

The core oligosaccharides in HS milk are substituted with a different number of fucose residues, as shown by characteristic ladder patterns (Fig. 1B). We purified monofucosylated LNnD (MFLNnD, HS-7, and HS-8 in Fig. 1B) and difucosylated LNnD (DFLNnD, HS-9, and HS-10 in Fig. 1B) and analyzed them using MALDI-QIT-TOF MS. Fig. 5 shows the MS/MS spectra using $[M + Na]^+$ observed at m/z 2093.1 for the purified MFLNnD (Fig. 5A and B). The Y ion at m/z 1947.5/1947.2 corresponding to the loss of 146 mass units (dHex-18 mass) from $[M + Na]^+$ indicates the presence of a Fuc residue.

The Y ions at m/z 1728.0/1728.1 and 1581.9/1582.0 are due to H5N3F1-2AA and H5N3-2AA, respectively. These fragment ions were commonly observed in HS-7 and HS-8. We also found the set of B ion series, $[H2N2]^+$ at m/z 753.5/754.5, $[H3N3]^+$ at m/z 1118.7/1118.8, $[H3N3F1]^+$ at m/z 1264.8/1264.9, and $[H4N3F1]^+$ at m/z 1791. Characteristic ions observed at m/z 1264.8 (Fig. 5A) and m/z 1118.8 (Fig. 5B) suggested the difference in the linkage positions of Fuc residues at the nonreducing Gal residues. The B ion at m/z 1264.8 ($B_{5\alpha}$) indicates that a Fuc residue is linked to the most outer LacNAc residue. The B ion at m/z 1118.8 corresponding to three LacNAc units suggests that one Fuc is attached to the 6-branch side of the reducing terminal lactose. Urashima and coworkers reported that small oligosaccharides in HS milk contained type II LacNAc (Gal β 1-4GlcNAc-R) but not type I LacNAc (Gal β 1-3GlcNAc-R) [14]. Thus, the oligosaccharides HS-7 and HS-8 are assigned to those as indicated in Fig. 5.

Fig. 6 shows the MS/MS spectra of the ions at m/z 2239.8 for the $[M + Na]^+$ of DFLNnD (HS-9 and HS-10 in Fig. 1B). The Y ions at m/z 2093.3 corresponding to a loss of 146 (dHex-18 mass) from $[M + Na]^+$ indicate the presence of Fuc residue. In a similar manner, in the case of MS/MS of MFLNnD (Fig. 5), the Y ions observed at m/z

1728.1 and 1582.1 correspond to [H5N3F1–2AA] and [H5N3–2AA], respectively. The Y ion at *m/z* 851.7 corresponds to the composition of [H3N1–2AA]⁺. These fragment ions of the Y ion series were commonly observed in HS-9 and HS-10. We also observed the set of B ion series, [H2N2]⁺ at *m/z* 753.5, [H2N2F1]⁺ at *m/z* 1118.5, [H3N3F1]⁺ at *m/z* 1264.5, and [H3N3F2]⁺ at *m/z* 1411.0. A characteristic ion at *m/z* 1411.0 observed in HS-9 (Fig. 6A), which corresponds to three LacNAc units having two Fuc residues, indicates that two Fuc residues are attached to the nonreducing Gal residues of both LacNAc branches linked to Galβ1–4GlcNAcβ1–3Galβ1–4Glc, as shown in Fig. 6A. Thus, the structure of HS-9 is assigned as shown in Fig. 6A. In contrast, a characteristic ion at *m/z* 1264.5 (B_{5x}) observed in HS-10 (Fig. 6B) indicates the attachment of only one Fuc to the most outer LacNAc residue. Therefore, the oligosaccharide (HS-10) is assigned to the structure as shown in Fig. 6B.

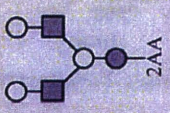
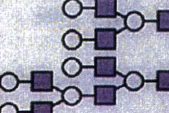
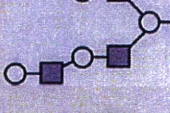
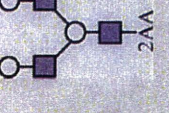
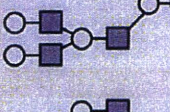
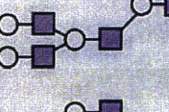
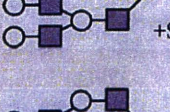
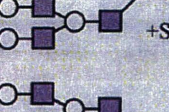
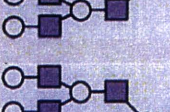
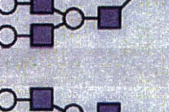
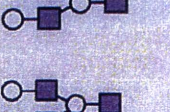
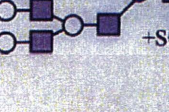
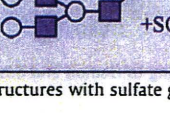
Among LNnD substituted with α1–2 Fuc residues, HS-7 and HS-10 were abundantly present (Fig. 1B) and both oligosaccharides have an α1–2 Fuc residue on either LacNAc on the α1–3 branch

of the LNnH core. These observations suggest that the modification of multibranched core oligosaccharides with α1–2 Fuc residues proceeds preferably at LacNAc residues of the elongated branches.

Discussion

We studied structural features of oligosaccharides from the milk samples of bearded and hooded seals by NP-HPLC and MALDI-TOF MS. The combination of sequential digestion of the oligosaccharides with exoglycosidases and MALDI-TOF MS was a useful technique for elucidation of the branching patterns and modification of oligosaccharides with fucose and/or sulfate group(s). Table 2 shows a list of asialo-oligosaccharides found in bearded and hooded seal milk. The oligosaccharides are categorized into nine core structures (A–I) based on the monosaccharide compositions. LNnD and LNnTD (C and E in Table 2) were observed as common core structures in both milk samples, but the two species showed quite different features. The most characteristic fea-

Table 2
Structural features of the oligosaccharides derived from bearded and hooded seal milk.

ID	Core structure	Number of Fuc residues					ID	Core structure	Number of Fuc residues				
		0	1	2	3	4			0	1	2	3	4
A		HS	HS	–	–	–	F		HS	HS	HS	HS	HS
B		HS	HS	–	–	–	G		–	BS	–	–	–
C		HS	HS BS	HS BS	HS BS	–	H		–	BS	BS	–	–
Cs		–	BS	–	–	–	HS		–	–	BS	–	–
D		HS	HS	HS	–	–	I		–	–	BS	BS	–
E		HS	HS BS	HS BS	HS BS	HS BS	Is		–	–	–	BS	–
Es		–	–	BS	–	–							

*Core structures with sulfate group at 3-OH position of nonreducing terminal Gal.

ture of oligosaccharides in both milk samples is that multi-branched oligosaccharides were present and linear oligosaccharides were not detected in the current study.

BS milk contained characteristic oligosaccharides having monosaccharide compositions of H3N3F1 (G in Table 2), H5N5F1 (H), and H7N7F3 (I), and these oligosaccharides were confirmed as Gal β 1–4GlcNAc β 1–3[Gal β 1–4GlcNAc β 1–6]Gal β 1–4GlcNAc and Gal β 1–4GlcNAc β 1–3[Gal β 1–4GlcNAc β 1–6]GlcNAc β 1–3[Gal β 1–4GlcNAc β 1–6]Gal β 1–4GlcNAc. Free oligosaccharides having LacNAc at the reducing end have been reported in bovine and caprine colostrum [33–35]. In mammary glands, lactose is synthesized by lactose synthase, a complex of β 4 galactosyltransferase I (β GalT-I) and α -lactalbumin [36]. β GalT-I is also involved in the synthesis of Gal β 1–4GlcNAc in the case of the absence of α -lactalbumin. However, α -lactalbumin in lactating mammary glands changes the preferred acceptor of β GalT-I from GlcNAc to Glc [37]. Interestingly, bovine colostrum contains oligosaccharides such as NeuAc α 2–6Gal β 1–4GlcNAc, Gal β 1–4(Fuc α 1–3)GlcNAc, and Gal β 1–3(Fuc α 1–4)GlcNAc, but their concentrations decrease dramatically to the trace level 7 days after parturition [38,39]. The presence of large oligosaccharides such as H, HS, I, and is in BS milk strongly suggests that biosynthesis starts from LacNAc as the core structure.

HS milk contained varieties of oligosaccharides having multi-branched core structures (i.e., cores E and F in Table 2). In addition, most oligosaccharides were substituted with different numbers of α 1–2 Fuc residues at the nonreducing terminal Gal residues. All oligosaccharides in HS milk have LNnH (Gal β 1–4GlcNAc β 1–3[Gal β 1–4GlcNAc β 1–6]Gal β 1–4GlcNAc) as a common core. They are preferentially elongated at the Gal β 1–4GlcNAc β 1–3 branch of the LNnH core unit. For example, the core structure having monosaccharide compositions of H6N4 (C in Table 2) has two LacNAc residues on the Gal β 1–4GlcNAc β 1–3 branch of LNnH (see Fig. 4B), in contrast to the branching of lacto-N-decaose in human milk [40]. Among the β -N-acetylglucosaminyltransferase (GnT) family, β 3-N-acetylglucosaminyltransferase (iGnT), which is a key enzyme for the elongation of LacNAc sequence, prefers type II chain (Gal β 1–4Glc/GlcNAc) [41–43]. Urashima and coworkers analyzed small oligosaccharides in HS milk by ^1H NMR spectroscopy and revealed that major oligosaccharides contained only type II chains [14]. The addition of Gal β 1–4 residue to terminal GlcNAc provides the preferable acceptor for iGnT enzyme. In contrast, HS milk contains multiantennary oligosaccharides (E in Table 2), which have two LacNAc residues on both branches of LNnH, indicating that the Gal residue on the β 1–6 branch of LNnH was substituted followed by modification of the β 1–3 branch of LNnH. iGnT was considered to be less efficient to the longer LacNAc repeats as the acceptors [44]. Furthermore, the efficiency of iGnT may be decreased by the presence of the Gal β 1–4GlcNAc β 1–6 branch to the Gal residue of Gal β 1–4GlcNAc β 1–3Gal β 1–4GlcNAc. In general, the sequence of Gal β 1–4GlcNAc β 1–3Gal β 1–4Glc/GlcNAc provides the preferred acceptor for β (1–6)N-acetylglucosaminyltransferase (IGnT), which is thought to be a key enzyme for the branching of oligosaccharides [45–47]. It is likely that human milk oligosaccharides preferentially is elongated at the Gal β 1–4GlcNAc β 1–6 branch of LNnH, and the Gal β 1–3GlcNAc β 1–3 branch of LNnH does not receive further modification with LacNAc [6]. The presence of multi-branched oligosaccharides in seal milk suggests that enzyme activities of IGnT and β 4GalTs in seal mammary glands are higher than those in other eutherian mammals.

In this article, we have focused on characterization of the branching pattern of oligosaccharides of high molecular masses in seal milk samples by means of the combination of MALDI-TOF MS and sequential exoglycosidase digestion. Branching is one of the major structural features of carbohydrates, and a relatively simple set of monosaccharides can form various branching configurations. Techniques based on MS/MS were used for the structural

characterization of oligosaccharides. Special emphasis was made so that the combined use of MALDI-TOF MS and sequential exoglycosidase digestion gave unambiguous structural details of multi-branched oligosaccharides, including linkage positions and anomeric configurations.

Acknowledgment

We thank M.O. Hammill of the Department of Fisheries and Oceans, Canada, for his help with collecting the hooded seal milk.

References

- [1] T. Feizi, Demonstration by monoclonal antibodies that carbohydrate structures of glycoproteins and glycolipids are onco-developmental antigens, *Nature* 314 (1985) 53–57.
- [2] T. Osanai, T. Feizi, W. Chai, A.M. Lawson, M.L. Gustavsson, K. Sudo, M. Araki, K. Araki, C.T. Yuen, Two families of murine carbohydrate ligands for E-selectin, *Biochem. Biophys. Res. Commun.* 218 (1996) 610–615.
- [3] T. Feizi, Carbohydrate-mediated recognition systems in innate immunity, *Immunol. Rev.* 173 (2000) 79–88.
- [4] M. Podbielska, S.A. Fredriksson, B. Nilsson, E. Lisowska, H. Krotkiewski, ABH blood group antigens in O-glycans of human glycoprotein A, *Arch. Biochem. Biophys.* 429 (2004) 145–153.
- [5] A. Kobata, K. Yamashita, Y. Tachibana, Oligosaccharides from human milk, *Methods Enzymol.* 50 (1978) 216–220.
- [6] D.S. Newburgh, S.H. Neubauer, *Handbook of Milk Composition*, Academic Press, San Diego, 1995.
- [7] H. Kogelberg, V.E. Piskarev, Y. Zhang, A.M. Lawson, W. Chai, Determination by electrospray mass spectrometry and ^1H -NMR spectroscopy of primary structures of variously fucosylated neutral oligosaccharides based on the iso-lacto-N-octaoase core, *Eur. J. Biochem.* 271 (2004) 1172–1186.
- [8] T. Urashima, S. Asakuma, M. Messer, *Comprehensive Glycoscience*, Elsevier, Amsterdam, 2007.
- [9] R. Jenness, E.A. Regehr, R.E. Sloan, Comparative biochemical studies of milk: II. Dialyzable carbohydrates, *Comp. Biochem. Physiol.* 13 (1964) 339–352.
- [10] M. Messer, Identification of N-acetyl-4-O-acetylneuraminyl-lactose in echidna milk, *Biochem. J.* 139 (1974) 415–420.
- [11] M. Messer, B. Green, Milk carbohydrates of marsupials: II. Quantitative and qualitative changes in milk carbohydrates during lactation in the tammar wallaby (*Macropus eugenii*), *Austral. J. Biol. Sci.* 32 (1979) 519–531.
- [12] J.P. Kamerling, L. Dorland, H. van Halbeek, J.F. Vliegthart, M. Messer, R. Schauer, Structural studies of 4-O-acetyl- α -N-acetylneuraminyl-(2–3)-lactose, the main oligosaccharide in echidna milk, *Carbohydr. Res.* 100 (1982) 331–340.
- [13] T. Urashima, T. Saito, Y. Tsuji, Y. Taneda, T. Takasawa, M. Messer, Chemical characterization of sialyl oligosaccharides isolated from tammar wallaby (*Macropus eugenii*) milk, *Biochim. Biophys. Acta* 1200 (1994) 64–72.
- [14] T. Urashima, T. Saito, T. Nakamura, M. Messer, Oligosaccharides of milk and colostrum in non-human mammals, *Glycoconj. J.* 18 (2001) 357–371.
- [15] T. Urashima, M. Yamamoto, T. Nakamura, I. Arai, T. Saito, M. Namiki, K. Yamaoka, K. Kawahara, Chemical characterization of the oligosaccharides in a sample of milk of a white-nosed coati, *Nasua narica* (Procyonidae: Carnivora), *Comp. Biochem. Physiol. A* 123 (1999) 187–193.
- [16] T. Nakamura, T. Urashima, T. Mizukami, M. Fukushima, I. Arai, T. Senshu, K. Imazu, T. Nakao, T. Saito, Z. Ye, H. Zuo, K. Wu, Composition and oligosaccharides of a milk sample of the giant panda, *Ailuropoda melanoleuca*, *Comp. Biochem. Physiol. B* 135 (2003) 439–448.
- [17] T. Urashima, M. Arita, M. Yoshida, T. Nakamura, I. Arai, T. Saito, J.P. Arnould, K.M. Kovacs, C. Lydersen, Chemical characterisation of the oligosaccharides in hooded seal (*Cystophora cristata*) and Australian fur seal (*Arctocephalus pusillus doriferus*) milk, *Comp. Biochem. Physiol. B* 128 (2001) 307–323.
- [18] T. Urashima, T. Nakamura, D. Nakagawa, M. Noda, I. Arai, T. Saito, C. Lydersen, K.M. Kovacs, Characterization of oligosaccharides in milk of bearded seal (*Erignathus barbatus*), *Comp. Biochem. Physiol. B* 138 (2004) 1–18.
- [19] T. Urashima, T. Nakamura, K. Teramoto, I. Arai, T. Saito, T. Komatsu, T. Tsubota, Chemical characterization of sialyl oligosaccharides in milk of the Japanese black bear, *Ursus thibetanus japonicus*, *Comp. Biochem. Physiol. B* 139 (2004) 587–595.
- [20] T. Urashima, T. Nakamura, K. Yamaguchi, J. Munakata, I. Arai, T. Saito, C. Lydersen, K.M. Kovacs, Chemical characterization of the oligosaccharides in milk of high Arctic harbour seal (*Phoca vitulina vitulina*), *Comp. Biochem. Physiol. A* 135 (2003) 549–563.
- [21] T. Urashima, H. Sato, J. Munakata, T. Nakamura, I. Arai, T. Saito, M. Tetsuka, Y. Fukui, H. Ishikawa, C. Lydersen, K.M. Kovacs, Chemical characterization of the oligosaccharides in beluga (*Delphinapterus leucas*) and Minke whale (*Balaenoptera acutorostrata*) milk, *Comp. Biochem. Physiol. B* 132 (2002) 611–624.
- [22] J. Amano, M. Messer, K. Kobata, Structures of the oligosaccharides isolated from milk of the platypus, *Glycoconj. J.* 2 (1985) 121–135.

- [23] B. Finke, B. Stahl, A. Pfenninger, M. Karas, H. Daniel, G. Sawatzki, Analysis of high-molecular-weight oligosaccharides from human milk by liquid chromatography and MALDI-MS, *Anal. Chem.* 71 (1999) 3755–3762.
- [24] A. Pfenninger, M. Karas, B. Finke, B. Stahl, G. Sawatzki, Mass spectrometric investigations of human milk oligosaccharides, *Adv. Exp. Med. Biol.* 501 (2001) 279–284.
- [25] V.N. Reinhold, B.B. Reinhold, C.E. Costello, Carbohydrate molecular weight profiling, sequence, linkage, and branching data: ES-MS and CID, *Anal. Chem.* 67 (1995) 1772–1784.
- [26] W. Morelle, J.C. Michalski, Glycomics and mass spectrometry, *Curr. Pharm. Des.* 11 (2005) 2615–2645.
- [27] W. Chai, V. Piskarev, A.M. Lawson, Negative-ion electrospray mass spectrometry of neutral underivatized oligosaccharides, *Anal. Chem.* 73 (2001) 651–657.
- [28] W. Chai, V. Piskarev, A.M. Lawson, Branching pattern and sequence analysis of underivatized oligosaccharides by combined MS/MS of singly and doubly charged molecular ions in negative-ion electrospray mass spectrometry, *J. Am. Soc. Mass Spectrom.* 13 (2002) 670–679.
- [29] J.E. Hodge, B.T. Hofreiter, Determination of reducing sugars and carbohydrates, *Methods Carbohydr. Chem.* 1 (1962) 308–394.
- [30] K.R. Anumula, S.T. Dhume, High resolution and high sensitivity methods for oligosaccharide mapping and characterization by normal phase high performance liquid chromatography following derivatization with highly fluorescent anthranilic acid, *Glycobiology* 8 (1998) 685–694.
- [31] M. Nakano, K. Kakehi, M.H. Tsai, Y.C. Lee, Detailed structural features of glycan chains derived from α 1-acid glycoproteins of several different animals: the presence of hypersialylated, O-acetylated sialic acids but not disialyl residues, *Glycobiology* 14 (2004) 431–441.
- [32] R. Naka, S. Kamoda, A. Ishizuka, M. Kinoshita, K. Kakehi, Analysis of total N-glycans in cell membrane fractions of cancer cells using a combination of serotonin affinity chromatography and normal phase chromatography, *J. Proteome Res.* 5 (2006) 88–97.
- [33] T. Saito, T. Itoh, S. Adachi, Presence of two neutral disaccharides containing N-acetylhexosamine in bovine colostrum as free forms, *Biochim. Biophys. Acta* 801 (1984) 147–150.
- [34] T. Saito, T. Itoh, S. Adachi, Chemical structure of three neutral trisaccharides isolated in free form from bovine colostrum, *Carbohydr. Res.* 165 (1987) 43–51.
- [35] T. Urashima, Y. Kusaka, T. Nakamura, T. Saito, N. Maeda, M. Messer, Chemical characterization of milk oligosaccharides of the brown bear, *Ursus arctos yesoensis*, *Biochim. Biophys. Acta* 1334 (1997) 247–255.
- [36] K. Brew, T.C. Vanaman, R.L. Hill, The role of α -lactalbumin and the A protein in lactose synthetase: a unique mechanism for the control of a biological reaction, *Proc. Natl. Acad. Sci. USA* 59 (1968) 491–497.
- [37] B. Rajput, N.L. Shaper, J.H. Shaper, Transcriptional regulation of murine β 1,4-galactosyltransferase in somatic cells: analysis of a gene that serves both a housekeeping and a mammary gland-specific function, *J. Biol. Chem.* 271 (1996) 5131–5142.
- [38] D.T. Davis, C. Holt, W.W. Christie, *Biochemistry of Lactation*, Elsevier, Amsterdam, 1983.
- [39] P.K. Gopal, H.S. Gill, Oligosaccharides and glycoconjugates in bovine milk and colostrums, *Br. J. Nutr.* 84 (Suppl. 1) (2000) S69–S74.
- [40] S. Haeuw-Fievre, J.M. Wieruszkeski, Y. Plancke, J.C. Michalski, J. Montreuil, G. Strecker, Primary structure of human milk octa-, dodeca-, and tridecasaccharides determined by a combination of ^1H -NMR spectroscopy and fast-atom-bombardment mass spectrometry: Evidence for a new core structure, the *para*-lacto-N-octaose, *Eur. J. Biochem.* 215 (1993) 361–371.
- [41] H. Kawashima, K. Yamamoto, T. Osawa, T. Irimura, Purification and characterization of UDP-GlcNAc:Gal β 1–4Glc(NAc) β 1,3-N-acetylglucosaminyltransferase (poly-N-acetylglucosamine extension enzyme) from calf serum, *J. Biol. Chem.* 268 (1993) 27118–27126.
- [42] F. Piller, J.P. Cartron, UDP-GlcNAc:Gal β 1–4Glc(NAc) β 1–3-N-acetylglucosaminyltransferase: identification and characterization in human serum, *J. Biol. Chem.* 258 (1983) 12293–12299.
- [43] D.H. van den Eijnden, A.H. Koenderman, W.E. Schiphorst, Biosynthesis of blood group i-active polyglucosaminoglycans: partial purification and properties of an UDP-GlcNAc:N-acetylglucosaminide β 1–3-N-acetylglucosaminyltransferase from Novikoff tumor cell ascites fluid, *J. Biol. Chem.* 263 (1988) 12461–12471.
- [44] M. Ujita, A.K. Misra, J. McAuliffe, O. Hindsgaul, M. Fukuda, Poly-N-acetylglucosamine extension in N-glycans and core 2- and core 4-branched O-glycans is differentially controlled by i-extension enzyme and different members of the β 1,4-galactosyltransferase gene family, *J. Biol. Chem.* 275 (2000) 15868–15875.
- [45] G.Y. Chen, N. Kurosawa, T. Muramatsu, A novel variant form of murine β 1,6-N-acetylglucosaminyltransferase forming branches in poly-N-acetylglucosamines, *Glycobiology* 10 (2000) 1001–1011.
- [46] N. Inaba, T. Hiruma, A. Togayachi, H. Iwasaki, X.H. Wang, Y. Furukawa, R. Sumi, T. Kudo, K. Fujimura, T. Iwai, M. Gotoh, M. Nakamura, H. Narimatsu, A novel i-branching β 1,6-N-acetylglucosaminyltransferase involved in human blood group I antigen expression, *Blood* 101 (2003) 2870–2876.
- [47] A.D. Magnet, M. Fukuda, Expression of the large I antigen forming β 1,6-N-acetylglucosaminyltransferase in various tissues of adult mice, *Glycobiology* 7 (1997) 285–295.
- [48] B. Domon, C.E. Costello, Structure elucidation of glycosphingolipids and gangliosides using high-performance tandem mass spectrometry, *Biochemistry* 27 (1988) 1534–1543.

Development of Defective and Persistent Sendai Virus Vector A UNIQUE GENE DELIVERY/EXPRESSION SYSTEM IDEAL FOR CELL REPROGRAMMING^{*,§}

Received for publication, September 9, 2010, and in revised form, December 5, 2010. Published, JBC Papers in Press, December 7, 2010, DOI 10.1074/jbc.M110.183780

Ken Nishimura,^{a,b} Masayuki Sano,^a Manami Ohtaka,^{a,c} Birei Furuta,^d Yoko Umemura,^a Yoshiro Nakajima,^a Yuzuru Ikehara,^e Toshihiro Kobayashi,^{f,g} Hiroaki Segawa,^{h,1} Satoko Takayasu,^{a,c} Hideyuki Sato,^{f,g} Kaori Motomura,^{a,c} Eriko Uchida,ⁱ Toshie Kanayasu-Toyoda,^d Makoto Asashima,^{a,j} Hiromitsu Nakauchi,^{f,g} Teruhide Yamaguchi,^d and Mahito Nakanishi^{a,2}

From the ^aResearch Center for Stem Cell Engineering, ^bResearch Center for Medical Glycoscience, and ^hOrgan Development Research Laboratory, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Central 4, Tsukuba, Ibaraki 305-8562, ^cPRESTO, Japan Science and Technology Agency, 4-1-8 Hon-cho, Kawaguchi, Saitama 332-0012, ^eJapan Biological Informatics Consortium, TIME24 Building, 2-4-32 Aomi, Koto-ku, Tokyo 135-8073, ^dDivision of Biological Chemistry and Biologicals and ⁱDivision of Cellular and Gene Therapy Products, National Institute of Health Sciences, 1-18-1 Kami-Yoga, Setagaya-ku, Tokyo 158-8501, ^fDivision of Stem Cell Therapy, Center for Stem Cell Biology and Regenerative Medicine, the Institute of Medical Science, the University of Tokyo, and the ^gNakauchi Stem Cell and Organ Regeneration Project, ERATO, Japan Science and Technology Agency, 4-6-1 Shirokane-dai, Minato-ku, Tokyo 108-8639, and the ^jDepartment of Life Sciences (Biology), Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan

The ectopic expression of transcription factors can reprogram differentiated tissue cells into induced pluripotent stem cells. However, this is a slow and inefficient process, depending on the simultaneous delivery of multiple genes encoding essential reprogramming factors and on their sustained expression in target cells. Moreover, once cell reprogramming is accomplished, these exogenous reprogramming factors should be replaced with their endogenous counterparts for establishing autoregulated pluripotency. Complete and designed removal of the exogenous genes from the reprogrammed cells would be an ideal option for satisfying this latter requisite as well as for minimizing the risk of malignant cell transformation. However, no single gene delivery/expression system has ever been equipped with these contradictory characteristics. Here we report the development of a novel replication-defective and persistent Sendai virus (SeVdp) vector based on a non-cytopathic variant virus, which fulfills all of these requirements for cell reprogramming. The SeVdp vector could accommodate up to four exogenous genes, deliver them efficiently into various mammalian cells (including primary tissue cells and human hematopoietic stem cells) and express them stably in the cytoplasm at a prefixed balance. Furthermore, interfering with viral transcription/replication using siRNA could erase the genomic RNA of SeVdp vector from the target cells quickly and thoroughly. A SeVdp vector installed with *Oct4/Sox2/Klf4/c-Myc* could reprogram mouse primary fibroblasts quite efficiently; ~1% of the cells were reprogrammed to Nanog-positive induced pluripotent stem cells without chromosomal gene integration. Thus, this SeVdp vector has poten-

tial as a tool for advanced cell reprogramming and for stem cell research.

The generation of induced pluripotent stem (iPS)³ cells by reprogramming tissue cells with defined factors opened the door for realizing the medical application of patient-derived engineered stem cells (1). iPS cells were established originally by the ectopic expression of multiple transcription factors (e.g. Oct3/4, Sox2, Klf4, and c-Myc) using a retroviral vector (1). Since then, researchers have established iPS cells by several different approaches (and by their combination), including gene transfer, protein transduction, and treatment with chemical compounds (2). However, because of superior reproducibility and efficacy, ectopic expression of reprogramming factors by gene transfer is still the primary method of choice.

Various lines of evidence indicate that efficient cell reprogramming requires the sustained and simultaneous expression of several (usually 4) exogenous factors for at least 10–20 days (3). On the other hand, after reprogramming has been completed, these exogenous factors should be replaced promptly with their endogenous counterparts if the cells are to acquire autoregulated pluripotency (3). For this reason, retroviral and lentiviral vectors have been used preferentially; chromosomal insertion of the vector genome allows for stable gene expression, whereas epigenetic modification of the viral promoter shuts off the vector-mediated gene expression after cell reprogramming has been accomplished. Nevertheless, cell reprogramming with these insertional vectors has a crucial disadvantage in that silencing and reactivation of the inte-

* This work was supported in part by a grant from the New Energy and Industrial Technology Development Organization of Japan, by the Program for Promotion of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation, and by JST PRESTO Program.

§ The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Tables S1 and S2 and Figs. S1–S4.

¹ Present address: Fuji-Gotemba Research Laboratories, Chugai Pharmaceutical Co. Ltd., 1-135, Komakado, Gotemba, Shizuoka, 412-8513, Japan.

² To whom correspondence should be addressed. Tel.: 81-29-861-3040; Fax: 81-29-861-2798; E-mail: mahito-nakanishi@aist.go.jp.

³ The abbreviations used are: iPS, induced pluripotent stem; ES, embryonic stem; SeV, Sendai virus; SeVdp, defective and persistent Sendai virus; Bs, blasticidin S; Zeo, zeocin; Zeo, phleomycin-binding protein gene; Bsr, blasticidin S deaminase gene; Hyg, hygromycin B phosphotransferase gene; EGFP, enhanced green fluorescent protein; KO, Kusabira Orange; KR, Keima Red; MEF, mouse embryonic fibroblast; NP, nucleocapsid protein; CFU, colony-forming unit.

grated reprogramming genes is often unmanageable, which might affect the differentiation potency of iPS cells and the safety of the iPS-derived cells. Thus, investigators have focused on generating iPS cells carrying no exogenous genetic materials either by repetitive transient gene expression (4, 5), by passive elimination of stable episomal DNA (6), or by recombinase-mediated excision of integrated genes from the chromosome (Refs. 7 and 8; for review, see Ref. 9). However, all of these approaches are not only inefficient but also laborious in practice, and development of a simpler gene delivery/expression system suitable for cell reprogramming is needed.

Sendai virus (SeV) is a nonsegmented negative-strand RNA virus belonging to the *Paramyxoviridae* (10). As SeV can infect various animal cells with an exceptionally broad host range and is not pathogenic to humans, various applications have been explored for SeV as a recombinant viral vector capable of transient but strong gene expression (11). We have demonstrated the potential of SeV as a tool for stable gene expression through an analysis of the Cl.151 strain (12). This unique variant was originally isolated as a mutant capable of persistent infection at a nonpermissive temperature (38 °C) (13). We cloned the entire genome of SeV Cl.151 and determined that more than two genetic elements were responsible independently for the establishment of stable persistent infections (12). We also demonstrated that SeV Cl.151 installed with a single exogenous gene could express it stably without chromosomal insertion (12). As this characteristic is advantageous for cell reprogramming, we planned to optimize this gene delivery/expression system through a more extensive analysis of SeV-mediated stable gene expression.

Here we describe the replication-defective and persistent Sendai virus (SeVdp) vector, a novel gene transfer/expression system based on SeV Cl.151, with the following characteristics, 1) efficient, harmless, and simultaneous delivery of up to four exogenous genes installed on a single vector, 2) stable and reproducible expression of installed genes at a pre-fixed balance without chromosomal integration, and 3) quick and complete erasure of the vector genome by interfering with viral RNA-dependent RNA polymerase using siRNA. We also demonstrated that an SeVdp vector installed with *Oct4/Sox2/Klf4/c-Myc* could reprogram mouse primary fibroblasts efficiently. These characteristics should make SeVdp a universal tool for stem cell research, especially for advanced cell reprogramming.

EXPERIMENTAL PROCEDURES

Reconstitution of SeVdp Vector by Reverse Genetics—All recombinant DNA experiments were performed according to our institutional guidelines and under the permission of the institutional recombinant DNA experiment committee of the National Institute of Advanced Industrial Science and Technology and of the National Institutes of Health Sciences. Replication-competent SeV was reconstituted as described (12). Full-length SeVdp vector genomic cDNA for SeV (Cl.151 strain and Nagoya strain) and for installed genes was constructed on the lambda Dash II vector as described in supplemental Fig. S1. In brief, the *M*, *F*, and *HN* genes were replaced with exogenous genes cloned between *KasI* and *MluI* restric-

tion sites (for *M*), between *BglII* sites (for *F*), and between *NheI* and *SphI* sites (for *HN*). Additional extra genes were inserted into an *NheI/NotI* site created between the *P/C/V* and *M* genes. cDNAs encoding blasticidin S deaminase (*Bsr*), phleomycin-binding protein (*Zeo*), enhanced green fluorescent protein (*EGFP*), *Cypridina noctiluca* luciferase (*CLuc*), humanized Kusabira Orange (*KO*), and human gp91phox (*CYBB*) were amplified by polymerase chain reaction using pCX4-*bsr* (14), pUT58 (15), pEGFP-1 (Takara Bio, Otsu, Japan), pCLm (ATTO, Tokyo, Japan), pHKO1-MN1 (Medical & Biological Laboratories, Nagoya, Japan), and gp91phox-pCI-neo as templates, respectively. cDNA encoding humanized Keima Red (*KR*) was synthesized by GenScript (Piscataway, NJ), according to a published peptide sequence (16).

The reconstructed cDNA plasmids (2 µg) and the expression vector plasmids for SeV nucleocapsid protein (NP), P/C, and L proteins (1 µg each) and pSRD-HN-Fmut (17) (2 µg) were transfected into BHK/T7/151M(SE) cells using Lipofectamine LTX Plus reagent (Invitrogen). BHK/T7/151M(SE) cells were established by expressing humanized T7 RNA polymerase and the M protein of the SeV Cl.151 strain stably in BHK-21 cells and cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (FCS). The SeVdp vector was reconstituted in cells from positive-strand antigenome RNA transcribed from this cDNA and the SeV NP, P/C, L, Fmut, HN, and M (Cl.151) proteins. The vector-packaging cells harboring the SeVdp vector were established by selecting with antibiotics (blasticidin S at 10 µg/ml, zeocin at 500 µg/ml, or hygromycin B at 200 µg/ml) except for SeVdp(*c-Myc/Klf4/Oct4/Sox2*). The SeVdp vector was rescued by transient expression of the SeV *Fmut*, *HN*, and *M* (Cl.151) genes (driven by the SRα promoter derived from pcDL-SRα) (18) in the packaging cells as described above and recovered into the culture supernatant after incubation at 32 °C for 4 days. *Fmut*, a modified F gene for expressing the protease-susceptible SeV F protein, was generated as described (17). The supernatant was filtered through 0.45-µm cellulose acetate filters and stored in small aliquots at -80 °C. Titers of SeVdp vectors were determined by examining LLCMK₂ cells infected with a diluted SeVdp vector suspension using indirect immunofluorescence microscopy with an anti-NP rabbit polyclonal antibody.

Cell Culture, Fluorescence Microscopy, and Flow Cytometry—Long term stability of gene expression mediated by the SeVdp vectors was examined in LLCMK₂ cells and in human primary fibroblasts (TIG3) cultured in Eagle's minimum essential medium supplemented with 10% FCS. For examining stability under antibiotic selection, the cells were incubated in the presence of blasticidin S (Bs) (5 µg/ml), hygromycin B (200 µg/ml), or a mixture of blasticidin S (5 µg/ml) and zeocin (*Zeo*) (100 µg/ml). For examining stability without selection, the cells were preselected with antibiotics and then cultured without selection. The presence of SeVdp was determined by detecting SeV NP antigen with indirect immunofluorescence microscopy, counterstained with DAPI. EGFP, KO, and KR were detected by fluorescence microscopy (Zeiss, Oberkochen, Germany) using specific filters customized for these proteins. Flow cytometry was performed using a

Novel Sendai Virus Vector Ideal for Cell Reprogramming

FACSCalibur (BD Biosciences; see Figs. 2B and 3, E and F) and with FISHMAN R (On-chip Biotechnologies, Tokyo, Japan; Fig. 3, G and H) according to the standard procedures provided by the manufacturers.

Gene Delivery to Human Hematopoietic Stem Cells—All experiments using human resources were performed according to National Institute of Advanced Industrial Science and Technology and National Institutes of Health Sciences guidelines. OP9 cells (19) (provided by the RIKEN BioResource Center through the National Bio-Resource Project of the Ministry of Education, Culture, Sports, Science, and Technology of Japan) were cultured with minimal essential medium- α containing 20% FCS and 4 mM L-glutamine. Human umbilical cord blood was collected after a normal pregnancy and delivery after obtaining informed consent from the mothers. Human mononuclear cells were isolated from the cord blood using Lymphoprep (Axis-Shield, Oslo, Norway) according to the protocol provided by the manufacturer. CD133/1(+) cells were prepared from mononuclear cells using CD133 microbead kits and an AutoMACS system (Miltenyi Biotec, Bergisch Gladbach, Germany) with a purity greater than 90%, confirmed by flow cytometry using a FACSCalibur.

The purified CD133/1(+) cells were infected with SeVdp(Bsr/ Δ F/KO) at a multiplicity of infection of 4 at 37 °C for 2 h. For examining the efficiency of gene delivery, the infected cells were cultured for 10 days in Iscove's modified Dulbecco's medium supplemented with 20% FCS, 8 mM L-glutamine, 50 μ M 2-mercaptoethanol, 50 ng/ml human stem cell factor, 5 ng/ml interleukin-6 (IL-6), 5 ng/ml IL-3, 25 ng/ml flt-3 ligand (all from PeproTech, Rocky Hill, NJ), and 50 ng/ml human thrombopoietin (Kirin, Tokyo, Japan), and the fraction of KO-positive cells was determined by flow cytometry (Fig. 2B). For long term culture-initiating cell assays (20), 20–200 SeVdp-infected cells were seeded on 1.25×10^4 γ -ray-irradiated OP9 cells in 96-well plates in Iscove's modified Dulbecco's medium with 12.5% FCS, 12.5% horse serum, 8 mM L-glutamine, 50 μ M 2-mercaptoethanol, and 1 μ M hydrocortisone for 5 weeks; half of the medium was exchanged every week. Then whole cells in a well were recovered and cultured in 300 μ l of semisolid colony-forming cell assay medium (MethoCult GF+ H4433, StemCell Technologies, Vancouver, Canada) in 48-well plates. After culturing for 2 weeks, the numbers of KO-positive colonies were determined.

RNA Interference Analysis—The sequences of siRNAs against SeV NP, P, and L mRNAs (designed and synthesized by iGENE Therapeutics, Tokyo, Japan) used in this study are listed in supplemental Table S1. For examining the effect of siRNAs on the infection of replication-competent SeV vectors (supplemental Fig. S4A), 2×10^4 HeLa cells were seeded in 48-well plates with DMEM containing 10% FCS on day 0. On day 1, the cells were treated with siRNAs (100 nM) mixed with Lipofectamine 2000 (Invitrogen) for 6 h, then infected with SeV Cl.151(EGFP) at a multiplicity of infection of 100. The medium was replaced on day 2, and the cells were examined using fluorescence microscopy on day 4. For examining the effect of siRNAs on the removal of SeVdp vector from BHK-21 cells expressing the T7 RNA polymerase and F pro-

tein of the SeV Nagoya strain constitutively (BHK/T7/NaF cells) (supplemental Fig. S4B), 1×10^4 cells harboring SeVdp(M/EGFP/Bsr) were seeded in 48-well plates on day 0. On day 1 the cells were treated with siRNA (100 nM) mixed with Lipofectamine 2000 as described above. On day 5, the cells were examined using fluorescence microscopy. In both of these experiments, siRNA against firefly luciferase (21) was used as a control.

For examining the effect of siRNA on the removal of SeVdp vector using luciferase activity as a quantitative index (Fig. 4), 1.5×10^5 (Fig. 4B) or 3×10^4 (Fig. 4C) HeLa cells harboring SeVdp(KO/Hyg/EGFP/Luc2CP) were seeded with siRNAs (40 nM) mixed with Lipofectamine RNAiMAX (Invitrogen) as described above on day 0 in a 6-well plate (Fig. 4B) or in a 24-well plate (Fig. 4C), respectively. siRNA against *Renilla reniformis* luciferase (22) was used as a control. The cells were passaged with fresh siRNA on days 3 and 7, and the culture medium was replaced on the next day at each point. Firefly luciferase activity in the cell extract was determined on the indicated day using a luciferase assay system (Promega, Madison, WI). Specific luciferase activity was determined by normalizing against the amount of protein, determined using a Bradford protein assay kit (Bio-Rad). The cell lysates prepared on days 3, 7, and 12 were also analyzed by Western blotting using affinity-purified anti-SeV L protein rabbit antibody (2 μ g/ml). For certifying complete removal of the SeVdp genome, the cells harboring SeVdp(KO/Hyg/EGFP/Luc2CP) and treated with siRNA as described above were cultured in the absence of siRNA for 4 weeks. Then, 1×10^4 of the cells were seeded in a 6-well plate and cultured in the presence of hygromycin B (100 μ g/ml) for 10 days. The surviving cells were fixed, then stained with 0.01% crystal violet.

Biochemical Assays—SDS-PAGE and protein blotting were performed as described (23) using SuperSignal West Dura Extended Duration substrate (Thermo Fisher Scientific, Waltham, MA). For filter trap assays to detect the NP antigen (Table 1), culture supernatants of the cells harboring SeVdp vectors were passed through 0.45- μ m cellulose acetate filters, trapped onto supported nitrocellulose membranes (0.2 μ m, Bio-Rad) by vacuum filtration, and probed with an anti-SeV NP monoclonal mouse antibody.

Cell Reprogramming—Isolation and culture of mouse embryonic fibroblasts (MEFs) from a Nanog/GFP knock-in mouse (provided by the Riken BioResource Center) (MEF/Nanog-GFP), reprogramming with retroviral vectors and culture of mouse iPS cells were performed as described previously (24). Retroviral vectors installed separately with *Oct3/4*, *Sox2*, *Klf4*, and *c-Myc* (RvMX4) were prepared as described (1) using template DNA obtained from Addgene (Cambridge, MA). For reprogramming with SeVdp vectors installed with *Oct3/4*, *Sox2*, *Klf4*, and *c-Myc*, 1.25×10^5 of MEF/Nanog-GFP-expressing cells were infected with SeVdp vectors at 32 °C for 14 h. Then 1.0×10^3 of infected cells were seeded onto the feeder cells in 6-well plates and cultured as indicated in the legend to Fig. 5. The numbers of iPS colonies expressing GFP were determined using fluorescent microscopy. At 10 days after SeVdp infection, GFP-positive clones were isolated and treated with siRNA L527 as described above.

Characterization of iPS Cells—Semi-quantitative RT-PCR Assays were performed using GoTaq qPCR Master Mix (Promega) and the primer sets listed in supplemental Table S2. Template cDNA was synthesized with random primers using SuperScript III reverse transcriptase (Invitrogen) from 2 μ g of total cellular RNA isolated using ISOGEN (Nippon Gene, Tokyo, Japan). Bisulfite sequencing analysis was performed using the EpiTect Bisulfite kit (Qiagen, Hilden, Germany) with primer sets listed in supplemental Table S2. The PCR DNA fragments were cloned into pCR2.1 vector (Invitrogen), and sequenced by TAKARA BIO INC. (Shiga, Japan). Telomerase activity was determined using the Quantitative Telomerase Detection kit (Allied Biotech, Vallejo, CA). Teratoma formation was performed by subcutaneous injection of 1×10^6 SeVdp-iPS cells (clone #13) into SCID mice. Tumors recovered at necropsy after 6 weeks were processed for fixing and paraffin wax embedding, sectioned (4 μ m), and stained with hematoxylin and eosin. Histological findings were evaluated using a DM3000 microscope (Leica, Wetzlar, Germany). Chimera animals were generated by microinjection of iPS cells into eight-cell or morula stage embryos. The embryos were collected in Medium 2 (Millipore, Billerica, MA) from oviduct and uterus of ICR female mice 2.5 days post-coitum. These embryos were transferred into potassium simplex optimized medium with amino acids (KSOM-AA, Millipore) and cultured for 1–2 h. iPS cells were trypsinized and suspended in iPS cell culture medium. A piezo-driven micro-manipulator (Prime Tech, Tokyo, Japan) was used to drill zona pellucida under the microscope, and 10–15 iPS cells were introduced into the subzonal space of individual 8-cell or morula-stage embryos. After injection, embryos underwent follow-up culture in KSOM-AA for 24 h (until blastomere stage) and then were transferred into the uteri of pseudopregnant recipient ICR female mice.

RESULTS

Basic Design of the SeVdp Vector—The SeV genome consists of six independent cistrons (NP, P/C/V, M, F, HN, and L), encoding eight proteins (10). Each cistron is preceded by a gene-start signal (3'-UCCCNNUUUC) and is followed by a gene-end signal (3'-AUUCUUUUU), which are the only essential *cis*-elements for transcription (25). This simple structure of each cistron makes it easier to design a defective viral vector by gene replacement. The NP, P, and L genes of SeV encode a major NP and two subunits of RNA-dependent RNA polymerase (P and L), respectively. All of these are indispensable for viral transcription and replication (10). We revealed previously that the L gene of the SeV Cl.151 strain with four missense mutations (V981I, S1088A, C1207S, and V1618L) contributes to long term persistence by providing the mechanism to escape from interferon β (IFN β) induction (12). Among these mutations, V1618L is most critical; SeV with a mutant L protein (V1618L) is defective in IFN β induction as with the Cl.151 strain.⁴ We also found that uncapped read-through transcripts synthesized in an early stage of infection with wild-type SeV were only barely detectable in cells in-

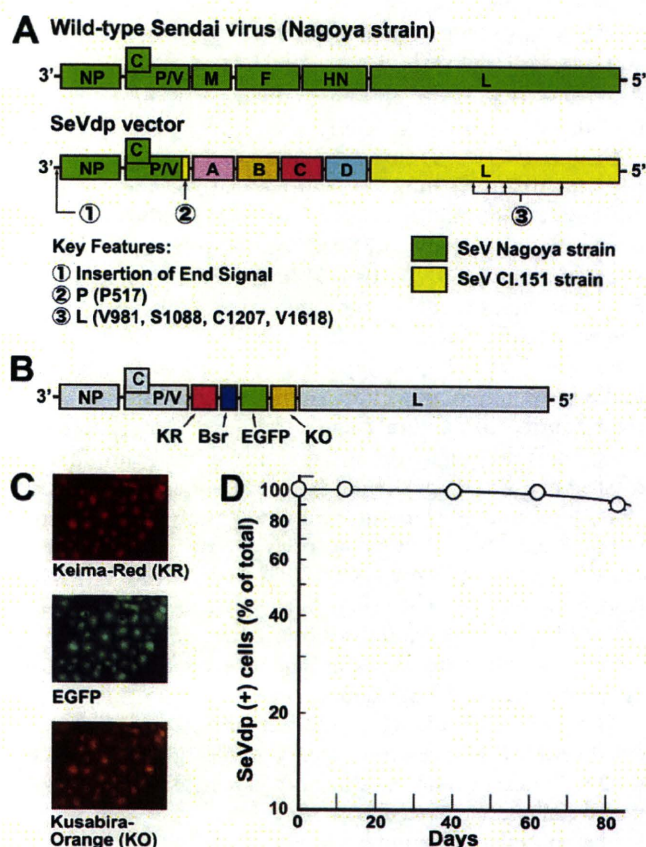


FIGURE 1. Design and characteristics of the SeVdp vectors. A, genome structure of the Sendai virus and an SeVdp vector is shown. Exogenous genes installed on the SeVdp vector are indicated as A–D. B, genome structure of SeVdp(KR/Bsr/EGFP/KO) is shown. C, expression of the fluorescent marker genes installed on the SeVdp vector is shown. LLCMK₂ cells were infected with SeVdp(KR/Bsr/EGFP/KO) at a multiplicity of infection of 0.1, selected with blasticidin S (5 μ g/ml), and examined by fluorescence microscopy with dye-specific filters. D, stability of gene expression induced by the SeVdp vector is shown. LLCMK₂ cells were infected with SeVdp(KR/Bsr/EGFP/KO) and selected with blasticidin S as described in C. The cells were then cultured for the indicated period in the absence of blasticidin S. The ratio of SeV NP antigen-positive cells in the total cells was determined by fluorescence microscopy, as described under “Experimental Procedures.”

fectured with SeV/L (V1618L).⁴ We hypothesized that the defect in IFN β induction might be correlated with the altered transcription of uncapped read-through RNA by the mutant SeV RNA polymerase.

In addition, we identified a missense mutation of the P gene (P517H) that is also essential for establishing long term persistency (Fig. 1A, supplemental Fig. S2A). The P protein makes a complex with the NP and L proteins, and the C terminus of P protein (amino acids, 479–568) has been assigned as a binding region for the NP (26). However, the precise role of this alteration in long term persistency remains to be determined. On the other hand, the 3'-distal region (nucleotides 1–2870) of the genome of SeV Cl.151 consists of the whole NP gene and part of the P/C/V genes but does not contribute to viral persistency (12) (Fig. 1A, supplemental Fig. S2A). Rather, we found that replacement of this 3'-distal region with that of the wild-type Nagoya strain significantly improved the recovery of the recombinant SeV from full-length genomic cDNA (supplemental Fig. S2B). We also inserted a

⁴ K. Nishimura and M. Nakanishi, unpublished information.

Novel Sendai Virus Vector Ideal for Cell Reprogramming

gene-end signal just upstream of the gene-start signal of the NP gene (Fig. 1A, supplemental Fig. S2C). This modification further stabilized SeV-mediated gene expression through more stringent control of IFN β induction by forced termination of uncapped read-through transcripts.⁴

We then planned to expand the capacity of the SeV Cl.151-based vector by deleting all the viral genes dispensable for stable gene expression. We have revealed previously that mutations within the central region (nucleotides 2871–9594) of the SeV Cl.151 genome contributed to viral persistence independently from the altered *L* gene (12). The region consists of *M*, *F*, and *HN* genes, which encode a matrix protein underlining the viral envelope (*M*) and envelope glycoproteins (*F* and *HN*), respectively. These structural genes are essential for production of infectious virions but are dispensable for viral transcription/replication, as SeV vectors with all of the *M*, *F*, and *HN* genes deleted could be generated successfully (27). However, as these defective vectors could not support stable gene expression, the role of these structural genes in long term persistency remains obscure.

We found previously that cells infected with SeV Cl.151 expressed large quantities of *F* and *HN* proteins on their surface (28). This observation suggested that the accumulation of the structural gene products (proteins and/or mRNAs) might interfere with the lytic infection cycle by a negative feedback mechanism, as proposed previously for the *M* protein (29). To examine this premise directly, we characterized recombinant SeVs with a dysfunction in each of the *M*, *F*, or *HN* genes either by deletion or by nonsense mutation. A selective marker gene (*Bsr*) conferring resistance to blasticidin S was used for estimating persistency rapidly (supplemental Fig. S3A). We found that all of these single-gene defective viruses established stable Bs-resistant colonies (supplemental Fig. S3A). Moreover, any two of these structural genes could be replaced with exogenous genes without affecting the persistent phenotype (supplemental Fig. S3A). To examine the role(s) of the *M*, *F*, and *HN* proteins further, we coexpressed these proteins directly from the cloned cDNAs and found that those derived from wild-type SeV strains induced much stronger cytopathic effects than did those derived from the SeV Cl.151 strain (supplemental Fig. S3B). From these results, we conclude that all of the *M*, *F*, and *HN* genes of SeV Cl.151 are dispensable and can be replaced with exogenous genes without disturbing viral persistency. We also succeeded in expanding the capacity of the vector by inserting an extra gene cassette between the *P/C/V* genes and the *M* gene without affecting viral persistency (Fig. 1A).

Dysfunction of structural genes is also essential for preventing self-replication of the vector. As the vectors used for cell reprogramming are installed with tumorigenic genes, such as *c-Myc* and *LIN28*, avoiding the production of secondary infectious particles from the gene-transferred cells is important not only to observe the regulation of recombinant DNA experiments but also to guarantee the safety of the vector in any therapeutic application. In the case of SeV, cultured cells infected with SeV variants defective in single structural genes produced significant amounts of virus-like particles (30, 31), suggesting that a dysfunction in single genes is insufficient for

TABLE 1
Determination of infectious virions and the NP protein in the culture supernatant of the cells harboring SeV Cl.151-based vectors
All the vectors were installed with the *Bsr* gene encoding blasticidine S deaminase. Aliquots of 10⁶ of LLCMK₂ cells harboring each SeVdp vector were seeded in 90-mm wells with 8 ml of medium. After culturing for 3 days, culture supernatant was recovered and filtered through 0.45- μ m cellulose acetate membranes. NP protein was determined by blotting 0.04–20 μ l of the supernatant on nitrocellulose membranes as described under “Experimental Procedures.” The supernatant was also incubated with 10⁶ uninfected LLCMK₂ cells for 14 h and then cultured in the presence of Bs (10 μ g/ml) for 7 days. The numbers of cell colonies resistant to Bs were determined by staining with crystal violet.

Structural genes			NP protein	Number of Bs ^r colonies
			ng/day/10 ⁵ cells	
M ^a	F ^a	HN ^a	8.75	> 10 ⁶
– ^b	F	HN	5.36	366
M	– ^b	HN	7.62	2
M	F	– ^b	72.05	10
M	– ^b	– ^b	2.51 ^c	0
– ^b	F	– ^b	2.76 ^c	0
– ^b	– ^b	HN	2.61 ^c	0
– ^b	– ^b	– ^b	2.56 ^c	0

^a Replication-competent vector.
^b Corresponding genes were deleted or replaced with exogenous genes.
^c Background caused by spontaneous cell lysis.

the complete blockage of self-replication. To reexamine this phenomenon, we determined the numbers of infectious particles and the amount of NP antigen in the culture supernatant of cells infected with various SeV Cl.151-derived defective viruses carrying the *Bsr* gene (Table 1). We found that cells infected with SeVs bearing a defect in one of the *F*, *HN*, or *M* genes produced significant amounts of NP antigen as well as infectious particles capable of transmitting Bs resistance to naïve cells (Table 1). This phenomenon was not observed when the viruses carried defects in at least two of the structural genes (Table 1). Therefore, we conclude that all three structural genes should be eliminated for maximizing the safety of the SeVdp vector through abolishing self-replication and for maximizing the vector capacity for installing exogenous genes.

In summary, we have designed the basic genome structure of the SeVdp vector. This consists of three separate genetic elements (Fig. 1A) as follows. 1) The 3'-terminal structure comprises the *NP* and *P/C/V* genes derived from the Nagoya strain with an alteration for supporting stable gene expression. 2) Internal gene cassettes capable of installing up to four exogenous genes, created by deletion/replacement/insertion of *M*, *F*, and *HN* genes. 3) The *L* gene and the 5'-terminal structure derived from the Cl.151 strain with four missense mutations necessary for stable gene expression and for escaping from IFN β induction.

Characterization of SeVdp Vector-mediated Gene Expression—We then prepared the SeVdp vectors installed with four exogenous genes and characterized vector-mediated gene expression. We first constructed SeVdp(KR/*Bsr*/EGFP/*KO*) installed with *Bsr* and three marker genes encoding KR, EGFP, and *KO* (Fig. 1B). All the cells infected with this vector expressed the three marker genes stably after selection with Bs (Fig. 1C). Furthermore, even in the absence of selection, 98.2% of cells retained the vectors for 62 days (Fig. 1D). Stability of gene expression induced by SeVdp vectors was solely dependent on the vector backbone described above and was not affected either by the installed genes or by the characteris-

TABLE 2

Stability of gene expression induced by SeVdp vectors

Cells harboring the SeVdp vectors were cultured in the presence of Bs (10 μ g/ml) to certify that 100% of the cells were SeVdp (+). On day 0 the cells were set up in the medium either with Bs (Bs(+)) or without Bs (Bs(-)). Expression of SeV NP antigen was recorded periodically, and the day on which 100% ($T_{100\%}$), 95% ($T_{95\%}$), or 80% ($T_{80\%}$) of the cells expressed NP is indicated. The gene cassette no. corresponds to those shown in Fig. 1A. *Bsr*, blasticidin S deaminase; *KO*, Kusabira Orange; *EGFP*, enhanced green fluorescent protein; *CLuc*, *Cypridina noctiluca* luciferase; *CYBB*, gp91 phox; *aGal*, human α -galactosidase; *KR*, Keima Red; ND, not determined because of the limited lifespan of the cells.

Gene cassette no.				Bs (-)		Bs (+)
A	B	C	D	$T_{95\%}$	$T_{80\%}$	$T_{100\%}$
				Days		
- ^a	<i>Bsr</i>	- ^a	<i>KO</i>	80 ^b	205 ^b	>180 ^b
- ^a	<i>Bsr</i>	- ^a	<i>KO</i>	>80 ^c	ND ^c	ND ^c
- ^a	<i>Bsr</i>	<i>EGFP</i>	<i>CLuc</i>	65 ^b	105 ^b	>180 ^b
- ^a	<i>Bsr</i>	<i>EGFP</i>	<i>CYBB</i>	65 ^b	170 ^b	>180 ^b
- ^a	<i>Bsr</i>	<i>EGFP</i>	α - <i>Gal</i>	60 ^b	112 ^b	>180 ^b
- ^a	<i>Bsr</i>	<i>EGFP</i>	<i>KO</i>	70 ^b	195 ^b	>180 ^b

^a No exogenous gene was installed.

^b Determined in LLCMK₂ cells.

^c Determined in normal human fibroblasts.

tics of host cells (Table 2). Under selection with antibiotics, nearly 100% of cells could retain the expression of all the marker genes for at least 6 months (Table 2). Reflecting the characteristics of its parental virus, the SeVdp vector could deliver and express the installed genes stably in various host cells, including cell lines derived from the mouse (NIH3T3), hamster (CHO, BHK-21), monkey (LLCMK₂, CV-1, COS-7), and human (HeLa, U937) as well as human and mouse primary fibroblasts (12). Thus, we proved that the SeVdp vectors had preserved the same characteristics of the parental SeV Cl.151 to establish stable persistent infection after it had been modified with four exogenous genes.

We then examined the feasibility of using the SeVdp vectors in stem cell research, focusing on their biological inertness. Most gene delivery/expression systems using either recombinant viruses or physical/chemical means often trigger cellular defense systems against pathogenic microbes (32). The stimulated cells secrete various cytokines, which affect the proliferation, differentiation, and survival of stem cells. Wild-type SeV and conventional SeV vectors based on the wild-type Z strain powerfully induced the production of IFN β , IFN γ , TNF α , IL-1 β , IL-6, IL-8, and many other cytokines (33–35), resulting in the apoptotic death of target cells. On the other hand, we revealed previously that SeV Cl.151 has a defect in inducing these inflammatory cytokines (12), suggesting strongly that the SeV Cl.151-based vector is biologically inert. To verify this under more stringent experimental conditions, we examined the effect of SeVdp-mediated gene transfer/expression on human hematopoietic stem cells (HSCs) by long term culture-initiating cell assays (20).

We isolated a CD133 (+) HSC-enriched fraction from human cord blood, infected it with the SeVdp vector bearing the *KO* gene (SeVdp(*Bsr*/Δ*F*/*KO*)) on day 0, and cultured it further in standard conditions. More than 90% of the cells in the HSC-enriched fraction were susceptible to the SeVdp vector under this infection protocol and sustained strong *KO* expression on days 3 (Fig. 2A) and 10 (Fig. 2B). Seven weeks after culturing on OP9 stromal cells, all kinds of myeloid lineage colonies derived from human HSCs, including colony-form-

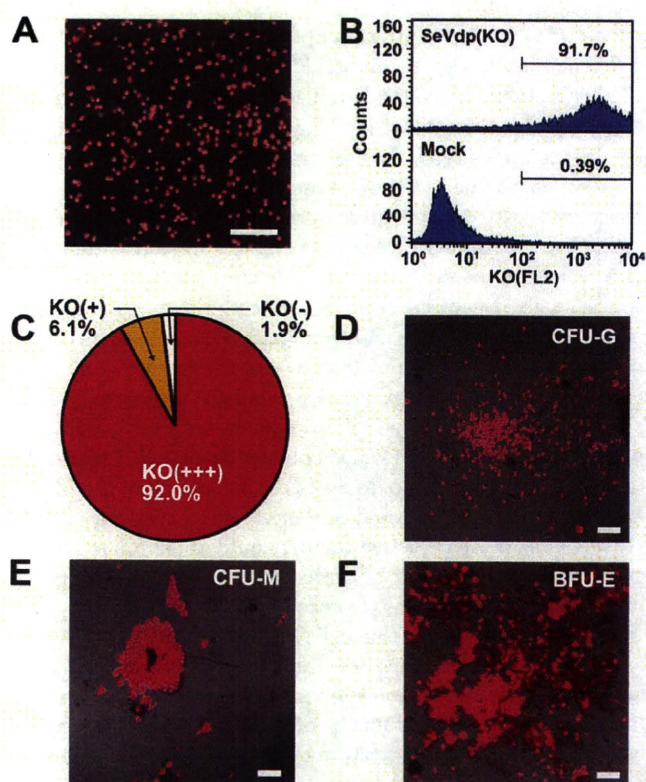


FIGURE 2. Expression of *KO* in human hematopoietic stem cells and in their descendant cells. A and B, expression of *KO* in CD133 (+) cord blood cells is shown. CD133 (+) cells were purified with magnetic beads conjugated with anti-CD133 antibody (Miltenyi Biotech). The cells were infected with SeVdp(*Bsr*/Δ*F*/*KO*) at a multiplicity of infection of 4 at 37 °C for 2 h. The cells were then cultured for 3 days (A) and 10 days (B) and examined using fluorescence and phase-contrast microscopy (A) and with flow cytometry using a FACSCalibur (BD Biosciences) (B), respectively. C–F, expression of *KO* in descendant colonies differentiated *in vitro* is shown. Cells infected with SeVdp(*Bsr*/Δ*F*/*KO*) as described above were cultured on OP9 cells in a 96-well plate for 5 weeks for lineage commitment. The cells in each well were then harvested, cultured in semisolid medium for 2 weeks, and examined for the expression of *KO* using fluorescence microscopy. C, the ratio of *KO*-positive colonies; 2931 differentiated colonies were examined. *KO* (+++), colonies expressing *KO* strongly; *KO* (+), colonies expressing *KO* weakly or heterogeneously; *KO* (–), colonies with no detectable *KO* expression. D–F, fluorescence and phase-contrast micrographs of typical colonies representative of each lineage. D, *CFU-G*, CFU-granulocytes. E, *CFU-M*, CFU-macrophages. F, *BFU-E*, burst-forming unit-erythroid cells. Scale bar, 100 μ m.

ing unit (CFU)-granulocytes (*CFU-G*), CFU-macrophages (*CFU-M*), CFU-granulocyte-macrophage, and burst-forming unit-erythroid (*BFU-E*) cells were readily detectable (Fig. 2D–F). Most importantly, 92% of these colonies still expressed *KO* very strongly on the seventh week (Fig. 2C), indicating that the SeVdp vector can deliver the gene quite efficiently into HSCs and is inert enough to sustain gene expression without affecting the differentiation of multipotent HSCs.

For efficient and reproducible cell reprogramming, it is also important to express the various reprogramming genes at a fixed balance in each target cell (36–39). The SeVdp vector installed with four exogenous genes could deliver these genes simultaneously, so is theoretically superior to those systems delivering the genes separately. To clarify this issue further, we prepared an SeVdp vector installed with *KO* and *EGFP* together (SeVdp(*KO*/*Hyg*/*EGFP*/*Luc2CP*)) (Fig. 3A) and two

Novel Sendai Virus Vector Ideal for Cell Reprogramming

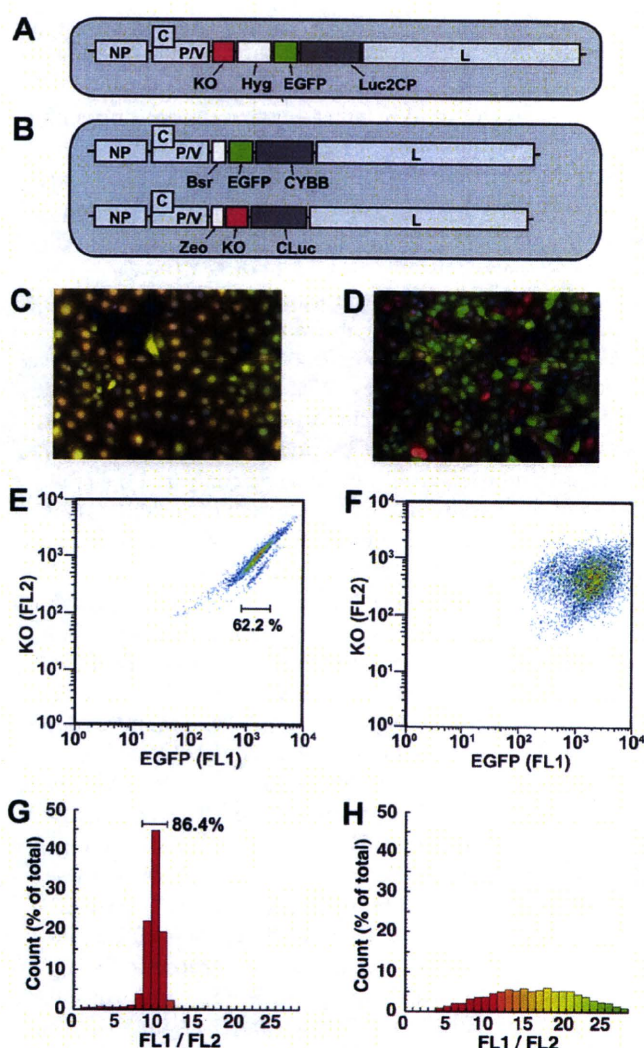


FIGURE 3. Compatibility of two independent SeVdp vectors in a single cell. *A*, genome structure of SeVdp(KO/Hyg/EGFP/Luc2CP) is shown. *B*, genome structure of SeVdp(Bsr/EGFP/CYBB) and SeVdp(Zeo/KO/CLuc) coexisting in a single cell is shown. *C* and *D*, fluorescence and phase-contrast micrographs of the cells described in schema *C* and in schema *B* (*D*) are shown. LLCMK₂ cells were infected with the SeVdp vectors as described in Fig. 1C and selected with hygromycin B (200 μ g/ml) (*C*) or with zeocin (100 μ g/ml) and blasticidin S (5 μ g/ml) (*D*). Fluorescence images of KO and of EGFP were obtained separately with specific filter sets, converted to artificial color (green for EGFP and red for KO), and merged using iVision software (BioVision Technologies, Exton, PA). *E–H*, quantitative analysis of EGFP and KO expression by flow cytometry is shown. The cells shown in *C* (*E* and *G*) or in *D* (*F* and *H*) were harvested as single-cell suspensions with trypsin, and the fluorescent signals were analyzed using a FACSCalibur (BD Biosciences) for quantifying the signals of EGFP (FL1, 515–545 nm) and KO (FL2, 564–606 nm) after compensation (*E* and *F*) and analyzed with FISHMAN R (On-tip Biotechnologies) for determining the ratio of the signals of EGFP and KO in each cell as a histogram (*G* and *H*).

others installed with KO and EGFP separately on different SeVdp vectors (SeVdp(Bsr/EGFP/CYBB) and SeVdp(Zeo/KO/CLuc)) (Fig. 3*B*). We then characterized the expression levels of KO and EGFP induced by a single infection with SeVdp(KO/Hyg/EGFP/Luc2CP) and by coinfection with SeVdp(Bsr/EGFP/CYBB) and SeVdp(Zeo/KO/CLuc) (Fig. 3). When the cells were infected solely with SeVdp(KO/Hyg/EGFP/Luc2CP) (Fig. 3*A*), they expressed both KO (red) and EGFP (green) at a constant balance, shown by a uniform yellow

low color in merged microscopy images (Fig. 3*C*). In contrast, the cells coinfecting with SeVdp(Bsr/EGFP/CYBB) and SeVdp(Zeo/KO/CLuc) (Fig. 3*B*) expressed KO and EGFP at a significantly different balance even after selection under Zeo plus Bs conditions (Fig. 3*D*). We then examined these cells quantitatively by flow cytometry (Fig. 3, *E–H*). When coinfecting with SeVdp(Bsr/EGFP/CYBB) and SeVdp(Zeo/KO/CLuc), nearly 100% of the infected cells expressed both KO and EGFP after antibiotic selection (Fig. 3*F*), but the balance of expression varied widely (Fig. 3, *F* and *H*). In contrast, 86.4% of the cells infected with SeVdp(KO/Hyg/EGFP/Luc2CP) expressed KO and EGFP at a fixed balance (Fig. 3, *E* and *G*) and at a constant level (62.2% of the cells expressed EGFP and KO within a 3-fold range) (Fig. 3*E*). From these results, we conclude that only the SeVdp vector installed with all the genes required to be expressed from a single genome can express these genes reproducibly at a fixed balance, thus providing a significant advantage for cell reprogramming.

Elimination of SeVdp Vector with siRNA—The last hurdle for efficient cell reprogramming is to establish a method for eliminating the vector genome from those cells harboring it stably. Although the viral family *Paramyxoviridae* includes major human pathogens (e.g. measles virus and respiratory syncytial virus), there is no specific small-molecule antiviral drug available. Instead, siRNAs against viral genes have been investigated with the aim of interfering with viral replication (40). However, the effects of siRNAs on stable persistent infections such as the SeVdp system have not been established. Therefore, we examined the effect of knocking down the viral replication machinery on the stability of the SeVdp genome using specific siRNAs. We designed siRNAs against each of the NP, P, and L genes (Fig. 4*A*) and examined their effects on the infection of a replication-competent SeV Cl.151 installed with the EGFP gene (SeV Cl.151(EGFP)). When the cells were treated with these siRNAs just before infection, the replication of SeV Cl.151(EGFP) was blocked almost completely (supplemental Fig. S4*A*). However, the effects of these siRNAs on the cells already harboring an SeVdp vector stably were quite different; siRNA against the L gene was most effective, and that against the NP gene showed almost no effect (supplemental Fig. S4*B*). This phenomenon might simply reflect the relative abundance of the target gene products; NP mRNA is about 34 times more abundant than L mRNA (41). Otherwise, suppression of a catalytic subunit of RNA polymerase (L protein) might interfere with the replication of the SeVdp genome more profoundly.

We then examined the time course with which an SeVdp vector would be eliminated by siRNA against the L gene (Fig. 4). To monitor elimination quantitatively, we used a cell line harboring the SeVdp vector installed with a destabilized firefly luciferase gene (SeVdp(KO/Hyg/EGFP/Luc2CP)) and determined luciferase activity as a faithful marker of gene expression from the SeVdp vector. As shown in Fig. 4*B*, the siRNA blocked expression of the L protein quite efficiently after day 3. In parallel with this suppression, the SeVdp was eliminated at a half-life of 17.5 h after a short time lag: the luciferase activity fell below the detection limit after day 8 (Fig. 4*C*). This elimination was irreversible; when the cells

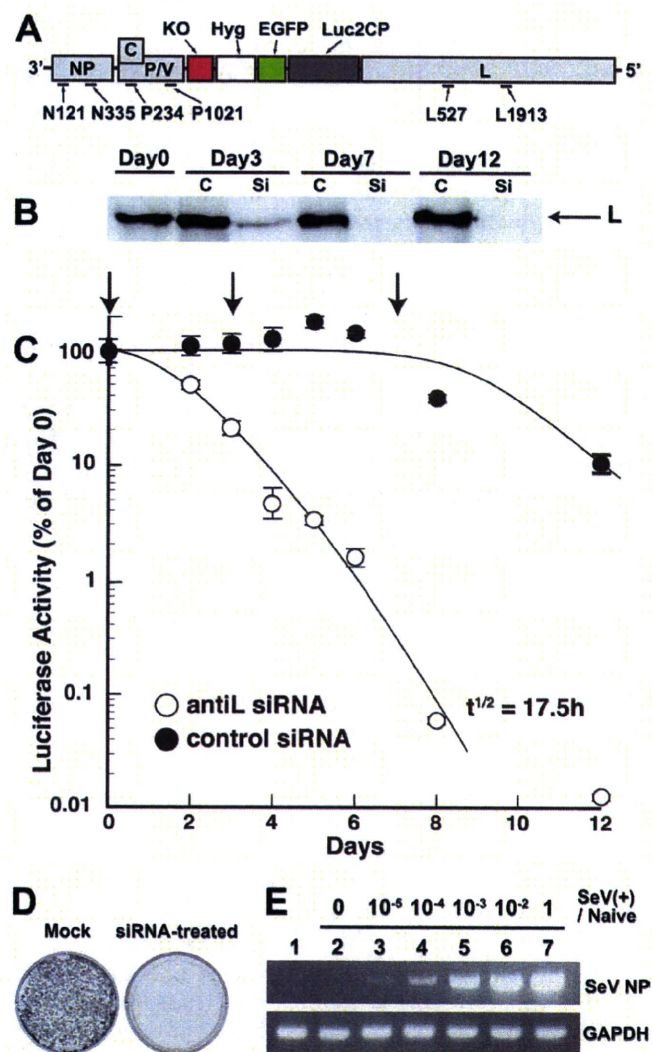


FIGURE 4. Elimination of SeVdp vectors from the cells with specific siRNAs. **A**, the genome structure of SeVdp(KO/Hyg/EGFP/Luc2CP) and target sites of siRNAs is shown. **B**, quantitative analysis of L protein by Western blotting is shown. HeLa cells carrying SeVdp(KO/Hyg/EGFP/Luc2CP) were cultured in the absence of hygromycin B and treated with siL527 (Si) or with control siRNA (C) complexed with Lipofectamine RNAiMAX (Invitrogen) on days 0, 3, and 7. The cells were harvested periodically as indicated, and 50- μ g aliquots of cell extracts were separated on SDS-PAGE. The amount of L protein was determined by Western blotting probed with an anti-SeV L protein rabbit polyclonal antibody. **C**, quantitative analysis of SeVdp-mediated gene expression is shown. The HeLa cells carrying SeVdp(KO/Hyg/EGFP/Luc2CP) were treated with siRNA as described in **B**. The cells were harvested periodically as indicated, and the specific firefly luciferase activity was determined as described under "Experimental Procedures." Open circles, treated with siL527; closed circles, treated with control siRNA; vertical arrows, the day of siRNA treatment. **D**, detection of the cells carrying the SeVdp vector after treatment with siL527 is shown. HeLa cells carrying SeVdp(KO/Hyg/EGFP/Luc2CP) and treated with siL527 for 8 days as described in **C** were further cultured for 4 weeks in the absence of the siRNA and of hygromycin B. Aliquots of 1×10^4 cells were then seeded into 6-well plates with medium containing hygromycin B (100 μ g/ml) and cultured for 10 days. The cells were fixed then stained with 0.01% crystal violet. **E**, detection of SeVdp by semiquantitative RT-PCR is shown. cDNAs were prepared by using 2- μ g aliquots of total cellular RNAs as indicated, and the cDNA corresponding to 10^4 cells was analyzed by RT-PCR to determine SeV NP mRNA as described under "Experimental Procedures." Lane 1, HeLa cells carrying SeVdp(KO/Hyg/EGFP/Luc2CP) and treated with siL527 as described in **D**; lanes 2–7, naïve HeLa cells containing SeVdp(+) cells at the ratio indicated.

harboring SeVdp(KO/Hyg/EGFP/Luc2CP) had been treated with the siRNA for 8 days and were further cultured for 4 weeks in the absence of the siRNA, all the cells became susceptible to hygromycin B, indicating that the SeVdp genome was no longer present in the cells (Fig. 4D). We further confirmed the complete erasure of the SeVdp genome by sensitive RT-PCR analysis; we could not detect any SeV genome under conditions capable of detecting a single SeV(+) cell among 10^5 naïve cells (Fig. 4E). From these data, we conclude that siRNA against the L gene is an effective tool for erasing the SeVdp genome thoroughly from the cells.

Generation of Mouse iPS Cells with SeVdp Vectors Installed with Reprogramming Genes—We then constructed an SeVdp vector installed with four reprogramming genes (SeVdp(*c-Myc/Klf4/Oct4/Sox2*)) (Fig. 5A) and examined its potential to reprogram mouse fibroblasts. First, we compared the efficiency of reprogramming by an infection of SeVdp(*c-Myc/Klf4/Oct4/Sox2*) with that produced by the coinfection of ecotropic retrovirus vectors installed with the same genes separately (RvMX4), which is a current standard approach for iPS generation (24) (Fig. 5C). For assessing the expression of Nanog, a well known marker of fully reprogrammed iPS cells, we used embryonic fibroblasts derived from a Nanog-GFP knock-in mouse (MEF/Nanog-GFP) (24) and monitored the expression of GFP. We found that SeVdp(*c-Myc/Klf4/Oct4/Sox2*) and pMX4 reprogrammed 0.83 and 0.01% of MEF cells to Nanog(+) iPS cell-like colonies, respectively, on day 14 after vector infection (Fig. 5C). As the efficiency of emergence of Nanog-GFP(+) colonies from MEF with retrovirus vectors was consistent with that reported previously (24), we concluded that the SeVdp vector could reprogram MEF about 100 times more efficiently than standard procedures using retrovirus vectors. We also prepared SeVdp vectors installed with *Klf4/Oct4/Sox2* genes (SeVdp(*Klf4/Oct4/Sox2*)) and with a *c-Myc* gene (SeVdp(*Zeo/hKO/c-Myc*)) (Fig. 5B) separately. We found that the coinfection with these two SeVdp vectors reprogrammed MEF/Nanog-EGFP much less efficiently than did the single infection with SeVdp(*c-Myc/Klf4/Oct4/Sox2*) (Fig. 5D). These results clearly proved our assumption that installing all the necessary genes on a single vector was critical for maximizing the potential of the vector to reprogram cells.

After treating these Nanog(+) cells with siRNA L527, we obtained SeV antigen-free iPS cells with typical characteristics (Fig. 5). First, they expressed ES/iPS cell markers detectable by fluorescence microscopy (*Nanog*, *Sox2*, and *SSEA-1*) (Fig. 5E) as well as by RT-PCR (*Nanog*, *Sox2*, *Oct4*, *c-Myc*, *ECAT1*, and *FGF4*) (Fig. 5F). The primer sets used in this RT-PCR assay detected the expression of mouse genes but not that of human genes installed on the SeVdp vector. Some of the endogenous iPS marker genes (*Nanog*, *Oct4*, and *Sox2*) were detectable as early as on day 5 after gene transfer (Fig. 5F, lane 3), suggesting rapid cell reprogramming by the SeVdp vector. Second, the promoter regions of the *Oct4* and *Nanog* genes were epigenetically remodeled similar to ES cells (Fig. 5G). Third, telomerase activity was increased by 50–200-fold to the same level as in ES cells (Fig. 5H). Fourth, they differentiated into derivatives of all three germ layers in teratomas (Fig.

Novel Sendai Virus Vector Ideal for Cell Reprogramming

