TABLE I

FOXO1 mice are smaller in body weight and lean body mass

FOXO1 mice weighed less (body weight and lean body mass) than nontransgenic, age- and sex-matched controls, when measured at 5 months of age (line A1) and at 4 months of age (line A2). Fat content per body weight of control and FOXO1 mice did not differ significantly. Data on both male and female mice are shown. Food intake and blood analyses of these mice are also shown. Values represent means ± S.E.

Mice	Numbers	Sex	Age	Body weight	Lean body mass	Fat content	Food intake	Free fatty acid	Lactate	Glucose	Insulin
				g	*g*	%	*g/g/day*	*mEq/liter*	*mg/ml*	*mg/dl*	*pg/ml*
Control	4	Male	5 months	29.0 ± 1.0	24.1 ± 0.3	20.8 ± 1.6	0.18 ± 0.005	0.30 ± 0.025	53.0 ± 4.3	163 ± 2.9	1775 ± 700
A1	4	Male		24.5 ± 0.4[a]	20.3 ± 0.4[b]	20.8 ± 0.5	0.17 ± 0.004	0.34 ± 0.098	56.3 ± 8.3	173 ± 14	739 ± 139
Control	4	Female		21.6 ± 0.9	19.3 ± 0.9	12.9 ± 1.0	0.25 ± 0.017	0.39 ± 0.060	32.7 ± 3.1	158 ± 8.0	289 ± 14
A1	6	Female		18.4 ± 0.4[a]	16.4 ± 0.2[a]	15.8 ± 1.2	0.24 ± 0.017	0.38 ± 0.049	38.9 ± 2.5	163 ± 5.3	302 ± 5
Control	4	Male	4 months	24.3 ± 0.4	21.0 ± 0.4	15.2 ± 1.2	0.21 ± 0.011	0.40 ± 0.045	37.3 ± 3.7	160 ± 10	373 ± 19
A2	4	Male		19.4 ± 0.1[b]	17.4 ± 0.2[b]	15.0 ± 1.9	0.18 ± 0.017	0.33 ± 0.077	46.4 ± 6.6	184 ± 14	573 ± 109
Control	4	Female		19.9 ± 0.6	17.6 ± 0.7	12.8 ± 0.2	0.25 ± 0.027	0.45 ± 0.055	34.5 ± 1.9	144 ± 10	283 ± 10
A2	4	Female		17.3 ± 0.3[a]	15.1 ± 0.4[c]	13.3 ± 0.9	0.23 ± 0.041	0.58 ± 0.096	35.7 ± 5.9	143 ± 6.5	316 ± 10

[a] $p < 0.01$.
[b] $p < 0.001$.
[c] $p < 0.05$.

the presence of the FOXO1 protein in the skeletal muscle of FOXO1 mice (Fig. 1*D*). An ~2.2-fold (line A1) and 3-fold (line A2) increase in FOXO1 protein levels was observed. These increases were at the physiological level, since 24-h fasting has been shown to increase FOXO1 protein content by 2.5–3-fold (Ref. 53 and data not shown).

FOXO1 Mice Are Small—The apparent phenotype observed in FOXO1 mice was small stature and thinner legs than the control mice. Both male and female transgenic mice weighed about 10% less than the control mice at 5 weeks of age (not shown). We used DEXA to measure the lean body mass (body weight excluding fat weight) and the content of fat in the whole body of the A1 line (at 5 months of age) and the A2 line (at 4 months of age) in age- and sex-matched control mice (Table I). Both body weight and lean body mass were significantly lower in both male and female FOXO1 mice (both lines) than in control mice. However, the fat content per total body weight of both FOXO1 mouse lines was comparable with that of nontransgenic mice (Table I). Thus, the decrease in body weight of the FOXO1 mice is not caused by a decrease in body fat but by a decrease in lean body mass. Consistent with the data on decreased lean body mass, the skeletal muscles in FOXO1 mice were smaller in size and dry mass, as well as paler in color than those of control mice (Fig. 1*E*). Consumption of food per body weight was not significantly different between FOXO1 mice and control mice (Table I). Blood metabolite (free fatty acid, lactate, and glucose) and insulin levels did not differ significantly between FOXO1 mice and the controls (Table I).

Microarray Analysis—To obtain information on changes in gene expression in FOXO1 mice, we performed microarray analysis using RNA samples from skeletal muscle (quadriceps) of transgenic and control mice. Most interestingly, the largest category of genes with suppressed expression in the transgenic mice was those involved in cell structure. Namely, about half of the down-regulated genes were classified as cytoskeletal proteins (Table II). The FOXO1-induced genes were distributed throughout various categories (not shown, see Supplemental Material 1).

In the skeletal muscle of FOXO1 mice, there was a decrease in the expression levels of genes related to structural proteins of the type I fiber (slow twitch oxidative, red muscle), such as slow muscle isoforms of myosins (Table II, line numbers 1, 4, and 6), slow isoforms of troponins (Table II, line numbers 2, 5, and 7), α-tropomyosin slow type (Table II, line number 13), myoglobin (Table II, line number 12), and mtCK (Table II, line number 15), which are abundant in type I fibers (54). This is consistent with the observation that the skeletal muscles of FOXO1 mice are pale (Fig. 1*E*). In the microarray, the expression of mitochondrial oxidative metabolism genes, such as the electron transport system, did not differ between FOXO1 mice and controls (not shown). In large mammals such as humans, type I fibers are higher in mitochondrial content and more dependent on oxidative metabolism than type II fibers. In small mammals (*e.g.* mouse and rat), a large amount of mitochondria is seen in type II fibers as well as type I fibers (2). The large amount of mitochondria in both type I and type II fibers in mice would explain the unchanged gene expression of the mitochondrial electron transport system, although expression of type I fiber genes was markedly suppressed. In addition, the gene expression of type II fiber isoforms did not differ (not shown). Namely, expression of genes preferentially abundant in type I fibers appears to be suppressed in the skeletal muscle of FOXO1 mice.

Northern Blot Analysis of Representative Genes—We recognize the limitation of single microarray assays, as they can contain certain noise in the data. Thus, to verify the changes of gene expression found in the microarray analysis, we performed Northern blot analysis by using probes for several genes. In addition to representative genes in the list (Table II), we also analyzed several additionally selected genes of type I fiber or type II fiber markers or genes that may be involved in fiber differentiation. FOXO1 overexpression did not significantly affect mRNA levels of the other FOXO members, FOXO4 and FOXO3a (Fig. 2*A*). Consistent with the microarray data, a reduction in gene expression was confirmed for type I fiber proteins, such as troponin C (slow) (Table II, line number 2), MLC (slow) (Table II, line number 6), troponin T (slow) (Table II, line number 7), myoglobin (Table II, line number 12), and mtCK (Table II, line number 15) (Fig. 2*A*). On the other hand, expression levels of genes for components of the mitochondrial electron transport system, such as cytochrome *c* oxidase II and IV (COX II and IV), and the F_0,F_1-ATPase, were not markedly changed in the skeletal muscle of FOXO1 mice. Next, we examined type II fiber genes. The expression of genes for troponin I (fast), troponin T (fast), and MLC (fast) did not differ between FOXO1 mice and control mice. Thus, the results of the microarray analysis were confirmed by Northern blot analysis. In addition, given that Mef2, NFAT, CaMK, and PGC-1α have been implicated recently in regulating gene expression in type I fibers (14–18, 22), we also examined the level of their expression in skeletal muscle of control and FOXO1 mice. PGC-1α mRNA levels were slightly increased in the skeletal muscle of FOXO1 mice (line A2). Most interestingly, expression levels of Mef2c and CaMK were reduced in FOXO1 mice. FOXO1-mediated down-regulation of type I fiber genes may, in part, be regulated by Mef2c and CaMK.

Moreover, we examined the expression levels of genes whose expression levels are known to be changed during skeletal

TABLE II

Gene with decreased expression in the skeletal muscle of FOXO1 mice

The expression levels of 22 genes were significantly decreased in both the A1 and A2 lines of FOXO1 mice. The genes are listed in the order of greatest fold change in expression in skeletal muscle from line A1 mice relative to control mice. Fold change calculations were carried out as an indication of the relative change of each transcript represented on the probe array. The average difference value is a marker of abundance of each gene. Categories and subcategories are based on a previously established classification scheme (50) and literature review. Change (↓) indicates that expression is significantly decreased compared with control mice.

	GenBank™ accession no.	Gene description	Categories	Subcategories	Fold change (line A1)	Fold change (line A2)	Average difference (control)	Average difference (line A1)	Average difference (line A2)
1	AJ223362	Myosin, heavy polypeptide 7, cardiac muscle, β	Cell structure	Cytoskeletal	−57.6 ↓	−41.8 ↓	418	−97	10
2	M29793	Troponin C (cardiac/slow skeletal isoform)	Cell structure	Cytoskeletal	−30.4 ↓	−9.7 ↓	260	2	39
3	U88623	Aquaporin 4	Metabolism	Transport	−21 ↓	−10.2 ↓	259	10	21
4	X12972	Myosin alkali light chain (ventricular/slow muscle isoform)	Cell structure	Cytoskeletal	−6.8 ↓	−2.1 ↓	937	129	450
5	AJ242874	Troponin I, skeletal, slow 1	Cell structure	Cytoskeletal	−6.7 ↓	−3.2 ↓	313	48	109
6	M91602	Myosin light chain 2 (cardiac ventricle isoform)	Cell structure	Cytoskeletal	−5.8 ↓	−3 ↓	1038	165	316
7	AV213431	Troponin T1 (slow twitch isoform)	Cell structure	Cytoskeletal	−4.4 ↓	−3.1 ↓	772	174	246
8	M74570	Aldehyde dehydrogenase II	Metabolism	Sugar/glycolysis	−4 ↓	−2.8 ↓	567	140	209
9	U34277	Platelet-activating factor acetylhydrolase	Cell defense	Homeostasis	−3.1 ↓	−2 ↓	238	77	120
10	AI646638	Clone MGC:37615 IMAGE: 4989784, mRNA,	Not found in the list		−2.9 ↓	−2.2 ↓	150	51	68
11	D45203	Pentylenetetrazole-related mRNA PTZ-17	Not found in the list		−2.8 ↓	−3 ↓	1232	434	411
12	X04405	Myoglobin	Cell defense	Homeostasis	−2.8 ↓	−1.8 ↓	2484	812	1410
13	U04541	Tropomyosin 3 (slow twitch isoform)	Cell structure	Cytoskeletal	−2.7 ↓	−2.4 ↓	662	243	273
14	X92665	Ubiquitin-conjugating enzyme E2E1	Protein expression	Post-translational modification	−2.1 ↓	−1.7 ↓	298	187	175
15	AV250974	Creatine kinase, mitochondrial 2	Metabolism	Sugar/ glycolysis	−2 ↓	−1.8 ↓	671	300	339
16	X57349	Transferrin receptor	Cell defense	Homeostasis	−1.9 ↓	−2.8 ↓	276	121	82
17	L12447	Insulin-like growth factor-binding protein 5	Unclassified		−1.9 ↓	−1.9 ↓	2080	1111	1095
18	Z38015	Myotonin protein kinase	Cell signaling	Protein modification	−1.9 ↓	−1.8 ↓	684	361	374
19	AB010144	Mitsugumin29, a synaptophysin family	Cell structure	General	−1.8 ↓	−2.4 ↓	742	414	312
20	X63615	Calcium/calmodulin-dependent protein kinase II, β	Cell signaling	Protein modification	−1.8 ↓	−2.1 ↓	295	160	146
21	U00677	Syntrophin, acidic 1	Cell structure	Cytoskeletal	−1.7 ↓	−1.7 ↓	857	504	491
22	AF032099	Potassium voltage-gated channel	Cell signaling	Channel/transport	−1.5 ↓	−1.6 ↓	320	209	195

muscle atrophy such as caused by fasting, cachexia, and STZ-induced diabetes (55). Specifically, gene expression of atrogin 1/MAFbx, MuRF1 (both are ubiquitin ligases), and cathepsin L (a lysosomal protease) is up-regulated and IGFBP5 is down-regulated during skeletal muscle atrophy (55). In our Northern blot analysis, the level of atrogin 1 expression was increased in the A2 line of FOXO1 mice, which has less skeletal muscle, but not in the A1 line, which also has less skeletal muscle mass than nontransgenic controls. In both the A1 and A2 lines of FOXO1 mice, the expression of cathepsin L and IGFBP5 was increased and decreased, respectively. The MuRF1 mRNA level was not changed. Thus, atrophy-related gene expression changes including that of protein degradation likely occurred in the skeletal muscle of FOXO1 mice.

Western Blot Analysis of the Skeletal Muscle of FOXO1 Mice and PGC-1α Mice—We examined the expression of various gene products of FOXO1 mice at the protein level by Western blot analysis (Fig. 2*B*). Protein extracts from the skeletal muscle of FOXO1 mice (A1 and A2 lines) and wild-type control mice were used. For comparison, we analyzed protein extracts from the skeletal muscle of PGC-1α transgenic mice, which we previously analyzed (23). Protein levels of troponin I (slow) and myoglobin, which are rich in type I fibers, were increased in PGC-1α mice but decreased in FOXO1 mice (Fig. 2*B*). On the other hand, the protein level of troponin I (fast), which is rich in type II fibers, was decreased in PGC-1α mice but not in FOXO1 mice (Fig. 2*B*). Thus, Western blot analysis of the protein expression of genes for type I and type II fibers was consistent with the results of mRNA expression analysis.

Histological Analysis of Skeletal Muscle of FOXO1 Mice—We examined the relationship between the change in type I fiber gene expression and actual muscle fiber morphology in the skeletal muscle (soleus) of transgenic mice using light microscopy and histochemical procedures (A1 line, 4 months after birth; A2 line, 3 months after birth). Distinction between type I and type II fibers can be made by myosin ATPase staining at different pH values. Specifically, at pH 10.5, type II fibers are well stained but not type I fibers, and at pH 4.3, type I fibers are well stained but not type II fibers (2). ATPase staining revealed that skeletal muscle cells (both type I and type II fibers) in the FOXO1 mice are smaller than those of the control mice (average cross-sectional area of muscle fibers; A1 line, 11.5 ± 0.8 μm^2 in FOXO1 mice and 20.0 ± 2.7 μm^2 in control mice; A2 line, 9.8 ± 0.5 μm^2 in FOXO1 mice and 14.1 ± 1.9 μm^2 in control mice) and had fewer type I fibers than those in the control mice (average; A1 line, 28.6 ± 1.3% in FOXO1 mice and

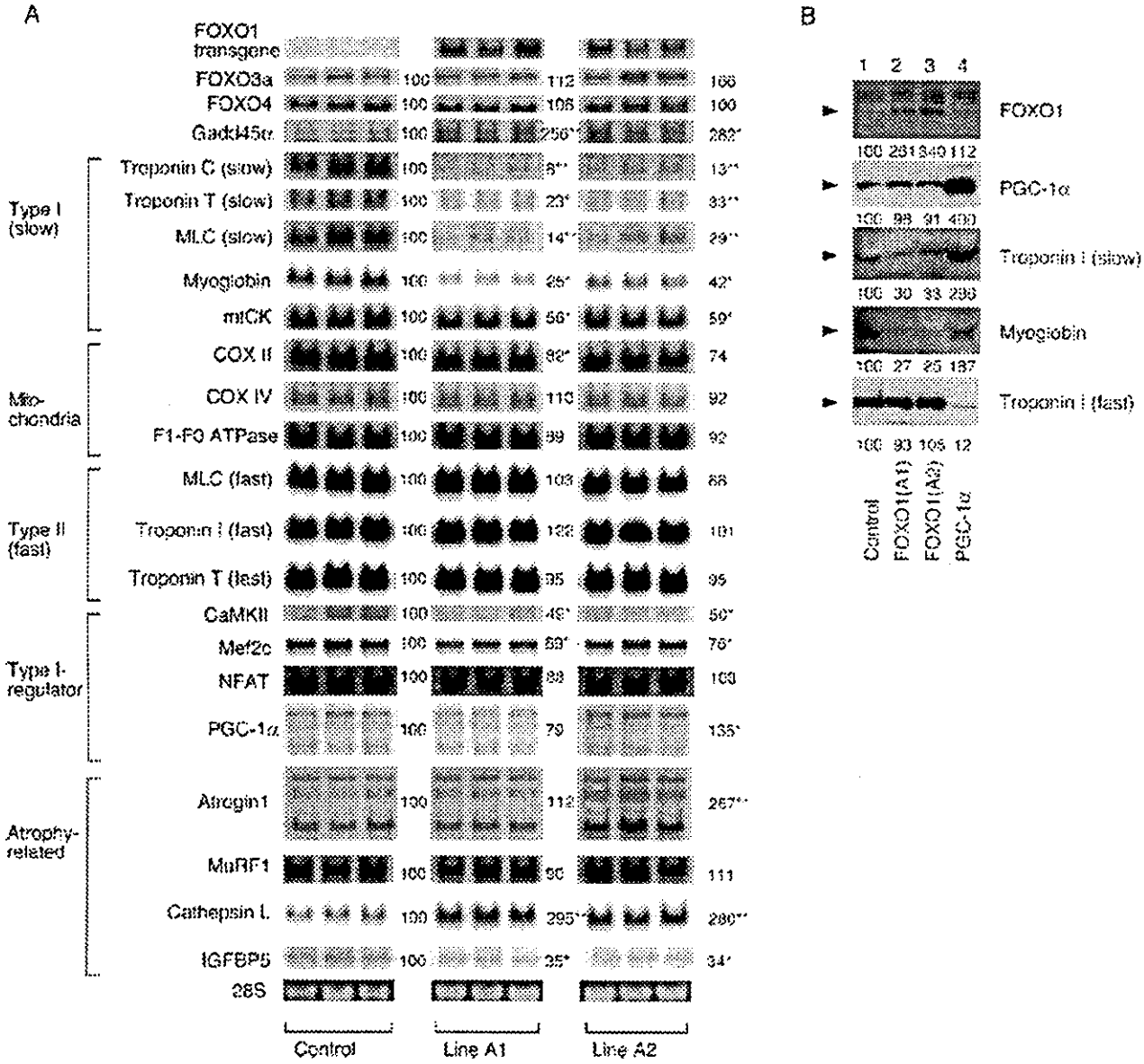


FIG. 2. **Gene product levels in the skeletal muscle of FOXO1 mice.** *A*, Northern blot analysis was performed on total RNA (20 μg per lane) isolated from skeletal muscle (quadriceps) of FOXO1 mice (*line A1* and *line A2*) and nontransgenic control mice. The same RNA sample sets were blotted onto multiple membranes and hybridized with the indicated probes. The names of genes examined are on the *left* of the autoradiograms, and average densitometric ratios (the control was set as *100*) are on the *right* (*, $p < 0.05$; **, $p < 0.01$). Equal sample loading was confirmed by ethidium bromide staining of 28 S ribosomal RNA. Each lane represents a sample from an individual mouse. *B*, Western blot analysis was performed on protein extracts from the skeletal muscle of *FOXO1* mice (*A1* and *A2* lines), PGC-1α mice, and control mice. Antibodies against FOXO1, PGC-1α, troponin I (slow), myoglobin, and troponin I (fast) were used. A typical autoradiogram, representative of three independent experiments with similar results, is shown. *Numbers* below the panels are values of the densitometric ratios (the signal of the control for each sample was set as *100*). Corresponding bands are indicated by *arrowheads*. The approximate estimated molecular sizes are as follows: FOXO1, 70 kDa; PGC-1α, 90 kDa; troponin I (slow), 30 kDa; myoglobin, 30 kDa; and troponin (fast), 40 kDa.

37.8 ± 2.2% in control; A2 line, 20.2 ± 2.3% in FOXO1 mice and 40.4 ± 2.0% in control) (Fig. 3*A*). Immunohistochemistry with antibodies to myoglobin (present at high concentrations in type I fibers) confirmed the reduction in the number of type I fibers in the skeletal muscle of FOXO1 mice (not shown). Skeletal muscle samples from FOXO1 mice had no structural abnormalities such as mitochondrial abnormalities, glycogen accumulation, vacuolar formation, and muscle fiber degeneration (not shown).

Running Wheel Activity of FOXO1 Mice—The mass and fiber composition of skeletal muscle are important for physical exercise. Type I fibers are more resistant to fatigue than type II fibers (2). As the FOXO1 mice had decreased total skeletal muscle mass and fewer type I fibers, they may have a low capacity for endurance, such as that needed in a marathon. We then compared the running wheel activity (spontaneous locomotive activity) in FOXO1 mice and control mice. Mice were transferred to cages with a running wheel and monitored daily for the number of wheel revolutions made for 6 days. Both lines of FOXO1 mice showed significantly fewer wheel revolutions (Fig. 3*B*). The decrease in running wheel activity suggested that FOXO1 mice were less able to sustain continuous muscle

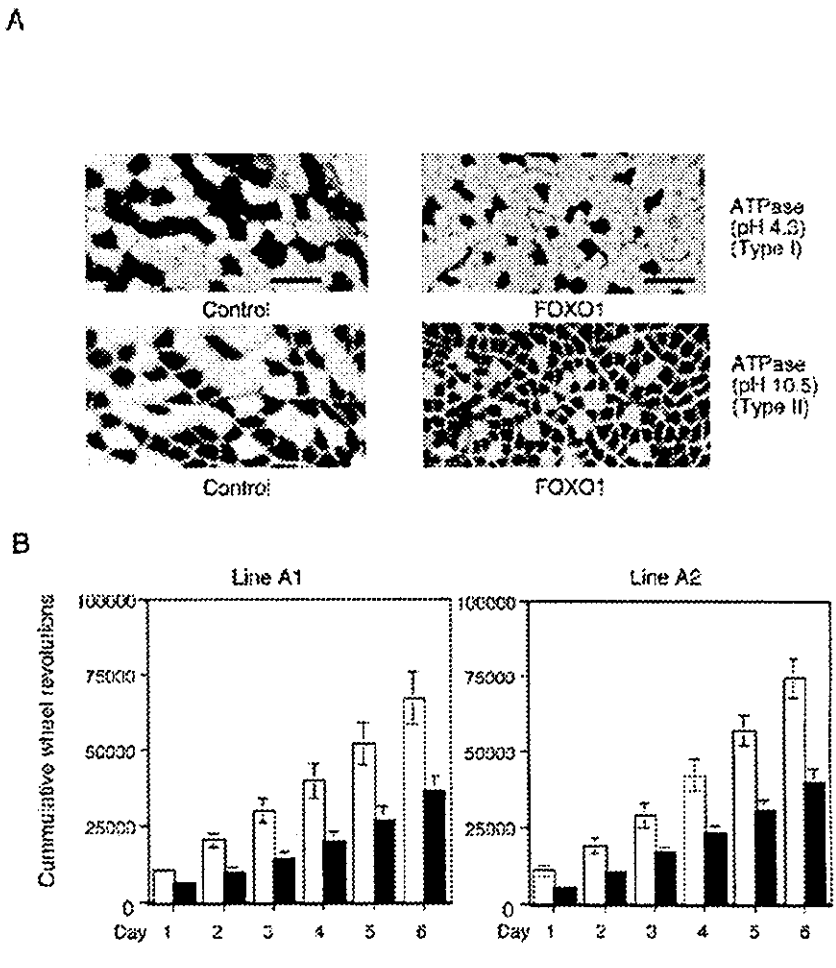


FIG. 3. *A*, histological analysis of skeletal muscle. Light microscopy of ATPase (pH 4.3 for type I fibers and pH 10.5 for type II fibers)-stained transverse sections of skeletal muscle (soleus) specimens from FOXO1 mice (line A2) and control littermates at 3 months of age. *Bars*, 50 μm. Skeletal muscle fibers of *FOXO1* mice were thinner and contained fewer type I fibers than that of control mice. *B*, running wheel activity of FOXO1 mice. Mice were housed individually in cages equipped with a running wheel (20 cm in diameter). The number of revolutions made was recorded daily for 6 days, and the cumulative values are shown. *Open column*, control; *closed column*, FOXO1 mice. Running wheel activity was significantly ($p < 0.05$) reduced in FOXO1 mice (*line A1*, *left*; *line A2*, *right*) compared with control mice. Mice used were females at 10 weeks (*line A1*) and 9 weeks (*line A2*) of age. Numbers of animals used are as follows: line A1, control, $n = 6$; FOXO1 mice, $n = 5$; line A2, control, $n = 4$; FOXO1 mice, $n = 3$. Because male mice responded similarly, only the data from female mice are shown. *C* and *D*, oral glucose tolerance tests (*C*) and insulin tolerance tests (*D*) on FOXO1 mice. For the oral glucose tolerance test, mice were fasted overnight and given D-glucose (1 mg/g body weight) orally by a stomach tube. Blood glucose levels were determined at the times indicated. For the insulin tolerance test, mice were allowed free access to food and then given 0.75 milliunits of human insulin/g of body weight. Blood glucose levels were measured at the indicated time points. Mice used were males at 10 weeks (line A1) and 9 weeks (line A2) of age. The numbers of animals used were: *line A1*, control, $n = 6$; FOXO1 mice, $n = 5$; *line A2*, control, $n = 5$; FOXO1 mice, $n = 4$.

contractions than control mice, which is consistent with the reduction in the mass of skeletal muscle and the number of type I fibers.

Oral Glucose Tolerance Test and Insulin Tolerance Test on FOXO1 Mice—Skeletal muscle is important for glucose metabolism. To examine whether the decreased skeletal muscle mass of FOXO1 mice is affecting their systemic glucose homeostasis, we examined oral glucose tolerance and insulin tolerance in FOXO1 mice. Glucose tolerance was impaired in both lines of FOXO1 mice, namely peak blood glucose values in FOXO1 mice were elevated significantly above those of the control mice (Fig. 3*C*). The insulin tolerance test clearly demonstrated that the glucose-lowering effects of insulin were impaired in both the A1 and A2 lines of FOXO1 mice, compared with those in age- and sex-matched control mice (Fig. 3*D*). FOXO1 mice showed a low capacity for glucose metabolism and decreased insulin sensitivity. Adipose tissue, another organ playing a role in glucose metabolism, appears not to be involved in this impaired glycemic control because 1) body fat did not differ between FOXO1 mice and control mice (Table I), and 2) gene expression of glucose transporter 4, which is a rate-limiting molecule of insulin-dependent glucose intake (56), was not decreased in adipose tissue of FOXO1 mice (see Supplemental Material 2). FOXO1 mice may therefore represent a certain type of diabetic state in humans.

Change in Endogenous FOXO1 Expression by Physical Inactivity—We performed Northern blot analysis with RNA from the skeletal muscle of mice maintained under a long period of physical inactivity. The right hindlimbs of wild-type mice were immobilized in plaster casts, and the left hindlimbs were left freely moving for the control sample. After 3 weeks in the plaster casts, skeletal muscle (gastrocnemius) weight of the right hindlimbs was significantly decreased compared with that in the controls (average, 88 ± 12 mg for immobilized and 149 ± 6 mg for freely moving controls, $n = 3$, $p < 0.05$). As shown in Fig. 4, the gene expression of troponin C (slow), myoglobin, and mtCK but not MLC (fast) and troponin T (fast) was markedly decreased in the plaster-casted muscle. At the same time, endogenous FOXO1 mRNA was increased in the immobilized muscle (Fig. 4). Furthermore, Gadd45α was increased in the same sample. In addition, cathepsin L, but not atrogin 1 and MuRF1, were increased (Fig. 4). Thus, mRNAs of endogenous FOXO1, Gadd45α, and cathepsin L were increased; skeletal muscle mass was decreased, and the expression of type I fiber genes but not type II fiber genes were decreased. The gene expression changes observed in the plaster-casted skeletal muscle were similar to the changes observed in the FOXO1 mice (Fig. 2*A*). These results further support the involvement of FOXO1 in the negative regulation of skeletal muscle mass and the expression of type I fiber genes.

DISCUSSION

To gain insight into the role of FOXO1 in skeletal muscle *in vivo*, we established transgenic mice overexpressing human FOXO1. The *FOXO1* transgene was predominantly expressed in the skeletal muscle, and the increase in FOXO1 protein expression was within physiological levels. Most interestingly, the skeletal muscle of FOXO1 mice weighed less and was paler

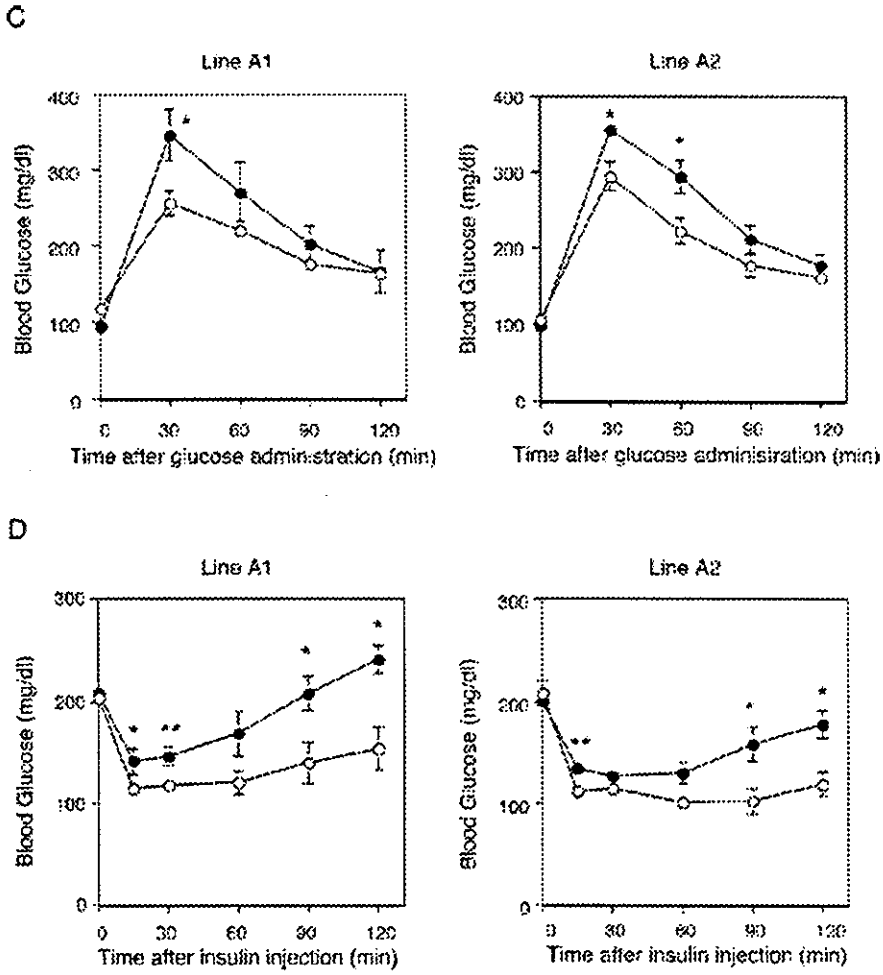


FIG. 3—*continued*

in color. The results of gene expression analyses showed that type I (red muscle) fiber-related gene expression was decreased in the skeletal muscle of FOXO1 mice. In addition, histological examinations showed that the skeletal muscle of FOXO1 mice had fewer type I fibers and smaller type I and type II fibers. Consistently, under long time physical inactivity by immobilizing skeletal muscle in plaster casts, an increased expression of endogenous FOXO1 mRNA and a markedly decreased expression of genes related to type I fibers were observed. These results suggest that FOXO1 is a negative regulator of skeletal muscle mass and expression of type I fiber-related genes. Moreover, FOXO1 mice showed poor glycemic control and low capacity for physical exercise, which involves the skeletal muscles, especially type I fibers. These phenotypes are consistent with the decreased mass of skeletal muscle including type I fibers in FOXO1 mice.

How does FOXO1 affect the skeletal muscle, including the reduction of mass of both type I and type II fibers and the suppressed expression of type I fiber genes? In the following, we discuss the possibility of involvement of FOXO1 in 1) growth, 2) protein degradation, and 3) differentiation of skeletal muscle.

1) FOXO1 may suppress increase of skeletal muscle mass. A genetic study of *C. elegans* showed that DAF16, the worm counterpart of FOXO, functions as a suppressor of insulin receptor-like signaling (44). Thus, the FOXO family might act negatively in mammals as a downstream player in insulin or IGF signaling. IGF-1 stimulates the proliferation of skeletal muscle satellite cells (57). Mature skeletal muscle fibers are not able to proliferate. Skeletal muscle satellite cells, mononuclear cells located between the basement membrane and the plasma membrane of myofibers in mature cells, are important in postnatal skeletal muscle hypertrophy because of their ability to add new myonuclei into growing myofibers. Machida *et al.* (58) showed that FOXO1 inhibited IGF-1-mediated skeletal muscle cell proliferation. In primary skeletal muscle satellite cells, FOXO1 activates the promoter of p27 Kip1, an inhibitor of the cell cycle at the G_1 stage, which leads to inhibition of cell proliferation, and addition of IGF-1 reverses the FOXO1-mediated activation of the p27 Kip1 promoter (58). Unexpectedly, p27 Kip1 mRNA expression was unchanged in the skeletal muscle of FOXO1 mice compared with that of controls (not shown). As the ratio of satellite cells is very small in total skeletal muscle, the increased expression of p27 Kip1 in satellite cells may not have been detected in our assay. On the other hand, we showed enhanced expression of Gadd45α, an inhibitor of the cell cycle at the G_2 stage (51, 52), in the skeletal muscle of FOXO1 mice (Figs. 1 and 2*A*). As a 0.7-kb stretch of the rat skeletal muscle α-actin promoter is active in skeletal muscle satellite cells (59), the *FOXO1* transgene, driven by a 2-kb stretch of the human skeletal muscle α-actin promoter (45), is likely to be expressed in the skeletal muscle satellite cells of the FOXO1 mice. Thus, the increased amount of Gadd45α and possibly p27 Kip1 in the skeletal muscle satellite cells of FOXO1 mice may have suppressed the proliferation of satellite cells and caused a decrease in skeletal muscle mass (size).

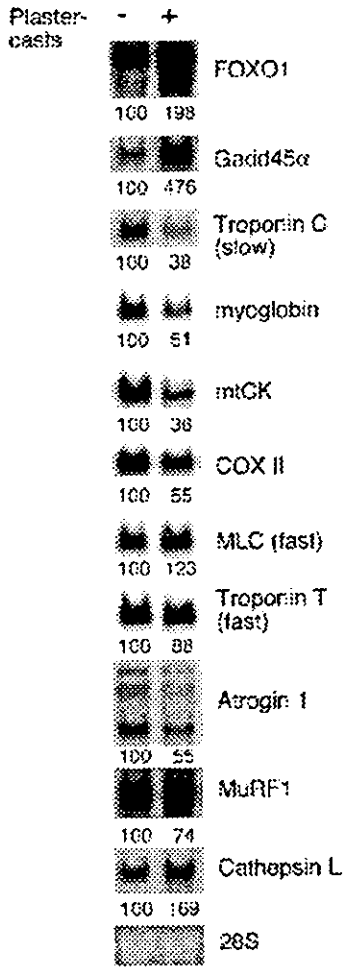


FIG. 4. **Gene expression in skeletal muscle immobilized in plaster casts.** The right hindlimbs of mice at 9 weeks of age were immobilized in plaster casts, and left hindlimbs of the mice were kept free for the control sample. After 3 weeks of immobilization in plaster casts, Northern analysis was performed on total RNA (20 μg per lane) isolated from the skeletal muscle (gastrocnemius) of right hindlimbs and left hindlimbs. *Plus* and *minus* denote with or without immobilization, respectively. The names of the genes examined are on the *right* of the autoradiograms. A typical autoradiogram, representative of three independent mice with similar results, is shown. The densitometric ratio is shown *below* the autoradiograms (the control was set as *100*).

2) FOXO1 may increase the degradation rate of skeletal muscle proteins. Gene expression of atrogin 1, MuRF1 (both are ubiquitin ligases), and cathepsin L (a lysosomal protease) is up-regulated and IGFBP5 is down-regulated during skeletal muscle atrophy caused by fasting, cachexia, STZ-induced diabetes, and other diseases (55). After we submitted our manuscript, a member of the FOXO family, FOXO3a, was reported to activate the gene expression of atrogin 1, and addition of IGF-1 was found to reverse the FOXO3a-mediated activation of the atrogin 1 promoter (60). Overexpression of an active form of FOXO3a reduces the size of skeletal muscle fibers, both *in vivo* and *in vitro* (60). In addition, another group reported that overexpression of an active form of FOXO1 in C2C12 muscle cells did not change the base-line expression of atrogin 1 and MuRF1, but the active form of FOXO1 suppresses IGF-1-mediated repression of atrogin 1 and MuRF1 expression induced by glucocorticoids (61). This suggests that FOXO1 expression is not sufficient for inducing atrophy-related genes, but FOXO1 is negatively involved in IGF-1-mediated suppression of atrophy of skeletal muscle. In our Northern blot analysis, the level of atrogin 1 was increased in the A2 line but not in the A1 line of FOXO1 mice, although both had less skeletal muscle mass than the nontransgenic controls. In both the A1 and A2 lines of FOXO1 mice, the expression of cathepsin L and IGFBP5 was increased and decreased, respectively. MuRF1 mRNA levels were not altered in both lines. Thus, atrophy-related protein degradation probably occurs in the skeletal muscle of FOXO1 mice and could explain, in part, the decrease in skeletal muscle mass of the FOXO1 mice. However, the increase in atrogin 1 is unlikely to be enough to cause the decrease in skeletal muscle mass of FOXO1 mice, because the expression level did not change in the A1 line of FOXO1 mice. This is consistent with the description by Sandri *et al.* (60) that overexpression of atrogin 1 alone does not cause myotube or muscle atrophy. On the other hand, IGFBP5 is reported to modulate the activity of IGF-1 (62), and hence decreased expression of IGFBP5 may contribute to the decrease in skeletal muscle mass by affecting IGF-1 action. *FOXO1* transgene expression was observed in both type I fiber-rich soleus and type II fiber-rich EDL. Thus, changes in the expression of atrophy-related genes may be an alternative molecular explanation for the decreased skeletal muscle mass, including the size of both type I and type II fibers of FOXO1 mice.

3) Does FOXO1 inhibit the differentiation of type I fibers? The *FOXO1* transgene is expressed in muscles rich in both type I and type II fibers. How does it cause the selective reduction of gene expression in type I fibers but not in type II fibers? It is possible that FOXO1 suppresses the function of a factor(s) that is preferentially expressed in type I fibers and therefore activates gene expression only in type I fibers. One candidate for such a factor is PGC-1α, which is known to be preferentially expressed in type I fibers and enhances type I fiber gene expression (22). As the FOXO1 protein can interact with the PGC-1α protein (43), FOXO1 may affect certain functions of PGC-1α. FOXO1 may inhibit PGC-1α function via its binding to PGC-1α. FOXO1 itself is a transcription factor. In addition, several reports (27–30) have shown that FOXO1 acts as a corepressor of nuclear receptors, whereas PGC-1α can activate many nuclear receptors (21, 63). Although to our knowledge nuclear receptors have not been shown to be involved in type I fiber-specific gene expression, a certain nuclear receptor(s) and transcription factor(s), which can interact with both FOXO1 and PGC-1α, may be involved in a process positively and negatively regulated by PGC-1α and FOXO1, respectively. Further studies are required to examine this possibility. Besides, although PGC-1α stimulates the differentiation of type I fibers, in FOXO1 mice, gene expression was reduced in type I fibers but was not affected in type II fibers. Thus, fiber differentiation (switching) from type I to type II is not likely to occur in FOXO1 mice, and FOXO1 appears not to be involved in fiber differentiation.

Calcineurin (14, 17) and CaMK (15), downstream molecules of calcium signaling (13), the transcription factors Mef2c (14–16, 18) and NFAT (14, 15, 17), as well as the nuclear receptor coactivator PGC-1α (22) are known to promote type I fiber differentiation and type I fiber gene expression. In skeletal muscle of FOXO1 mice, mRNA levels of Mef2c and CaMK are significantly decreased (Fig. 2*A*). FOXO1 may reduce gene expression in type I fiber by suppressing gene expression of Mef2c and CaMK.

FOXO1 mice showed a clear phenotype related to the function of skeletal muscle. Specifically, spontaneous locomotor activity was lower in FOXO1 mice than in control mice (Fig. 3*B*). In addition, FOXO1 mice had impaired oral glucose tolerance and impaired insulin-mediated glucose-lowering effects (Fig. 3, *C* and *D*). Elderly humans have been reported to show

a progressive loss of muscle fibers associated with diabetes, obesity, and decreased physical activity (sarcopenia). Overexpression of IGF-1 in skeletal muscle prevents the age-related decline in muscle mass (11, 57). As described above, the reduced skeletal muscle mass in FOXO1 mice may be caused by the suppression of IGF signaling during skeletal muscle formation, and FOXO1 may therefore be involved in age-related sarcopenia in humans. FOXO1 mice may be valuable as a model for human diseases related to loss of muscle fibers. Further analysis of the molecular mechanisms of FOXO1 action in skeletal muscle is important from a clinical as well as a sports science perspective.

Acknowledgments—We thank H. Meguro for technical assistance, Dr. S. Machida (University of Missouri, Columbia) for valuable comments, and Dr. H. A. Popiel for proofreading.

REFERENCES

1. Zurlo, F., Larson, K., Bogardus, C., and Ravussin, E. (1990) *J. Clin. Investig.* **86,** 1423–1427
2. Berchtold, M. W., Brinkmeier, H., and Muntener, M. (2000) *Physiol. Rev.* **80,** 1215–1265
3. Proctor, D., Balagopal, P., and Nair, K. (1998) *J. Nutr.* **128,** S351–S355
4. Fitts, R., Riley, D., and Widrick, J. (2001) *J. Exp. Biol.* **204,** 3201–3208
5. Hickey, M. S., Carey, J. O., Azevedo, J. L., Houmard, J. A., Pories, W. J., Israel, R. G., and Dohm, G. L. (1995) *Am. J. Physiol.* **268,** E453–E457
6. Gaster, M., Staehr, P., Beck-Nielsen, H., Schroder, H. D., and Handberg, A. (2001) *Diabetes* **50,** 1324–1329
7. Tanner, C. J., Barakat, H. A., Dohm, G. L., Pories, W. J., MacDonald, K. G., Cunningham, P. R., Swanson, M. S., and Houmard, J. A. (2002) *Am. J. Physiol.* **282,** E1191–E1196
8. Frost, R. A., and Lang, C. H. (2003) *Minerva Endocrinol.* **28,** 53–73
9. DeVol, D., Rotwein, P., Sadow, J., Novakofski, J., and Bechtel, P. (1990) *Am. J. Physiol.* **259,** E89–E95
10. Zdanowicz, M., Moyse, J., Wingertzahn, M., O'Connor, M., Teichberg, S., and Slonim, A. (1995) *Endocrinology* **136,** 4880–4886
11. Barton-Davis, E. R., Shoturma, D. I., Musaro, A., Rosenthal, N., and Sweeney, H. L. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 15603–15607
12. Florini, J., Ewton, D., and Coolican, S. (1996) *Endocr. Rev.* **17,** 481–517
13. Stull, J. T. (2001) *J. Biol. Chem.* **276,** 2311–2312
14. Chin, E. R., Olson, E. N., Richardson, J. A., Yang, Q., Humphries, C., Shelton, J. M., Wu, H., Zhu, W., Bassel-Duby, R., and Williams, R. S. (1998) *Genes Dev.* **12,** 2499–2509
15. Wu, H., Naya, F. J., McKinsey, T. A., Mercer, B., Shelton, J. M., Chin, E. R., Simard, A. R., Michel, R. N., Bassel-Duby, R., Olson, E. N., and Williams, R. S. (2000) *EMBO J.* **19,** 1963–1973
16. Yan, Z., Serrano, A. L., Schiaffino, S., Bassel-Duby, R., and Williams, R. S. (2001) *J. Biol. Chem.* **276,** 17361–17366
17. Chakkalakal, J. V., Stocksley, M. A., Harrison, M. A., Angus, L. M., Deschenes-Furry, J., St-Pierre, S., Megeney, L. A., Chin, E. R., Michel, R. N., and Jasmin, B. J. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 7791–7796
18. Karasseva, N., Tsika, G., Ji, J., Zhang, A., Mao, X., and Tsika, R. (2003) *Mol. Cell. Biol.* **23,** 5143–5164
19. Kamei, Y., Xu, L., Heinzel, T., Torchia, J., Kurokawa, R., Gloss, B., Lin, S. C., Heyman, R. A., Rose, D. W., Glass, C. K., and Rosenfeld, M. G. (1996) *Cell* **85,** 403–414
20. Glass, C. K., Rose, D. W., and Rosenfeld, M. G. (1997) *Curr. Opin. Cell Biol.* **9,** 222–232
21. Puigserver, P., Wu, Z., Park, C. W., Graves, R., Wright, M., and Spiegelman, B. M. (1998) *Cell* **92,** 829–839
22. Lin, J., Wu, H., Tarr, P. T., Zhang, C. Y., Wu, Z., Boss, O., Michael, L. F., Puigserver, P., Isotani, E., Olson, E. N., *et al.* (2002) *Nature* **418,** 797–801
23. Miura, S., Kai, Y., Ono, M., and Ezaki, O. (2003) *J. Biol. Chem.* **278,** 31385–31390
24. Anderson, M. J., Viars, C. S., Czekay, S., Cavenee, W. K., and Arden, K. C. (1998) *Genomics* **47,** 187–199
25. Kaestner, K. H., Knochel, W., and Martinez, D. E. (2000) *Genes Dev.* **14,** 142–146
26. Galili, N., Davis, R. J., Fredericks, W. J., Mukhopadhyay, S., Rauscher, F. J., 3rd, Emanuel, B. S., Rovera, G., and Barr, F. G. (1993) *Nat. Genet.* **5,** 230–235
27. Schuur, E. R., Loktev, A. V., Sharma, M., Sun, Z., Roth, R. A., and Weigel, R. J. (2001) *J. Biol. Chem.* **276,** 33554–33560
28. Zhao, H. H., Herrera, R. E., Coronado-Heinsohn, E., Yang, M. C., Ludes-Meyers, J. H., Seybold-Tilson, K. J., Nawaz, Z., Yee, D., Barr, F. G., Diab, S. G., Brown, P. H., Fuqua, S. A. W., and Osborne, C. K. (2001) *J. Biol. Chem.* **276,** 27907–27912
29. Dowell, P., Otto, T. C., Adi, S., and Lane, M. D. (2003) *J. Biol. Chem.* **278,** 45485–45491
30. Hirota, K., Daitoku, H., Matsuzaki, H., Araya, N., Yamagata, K., Asada, S., Sugaya, T., and Fukamizu, A. (2003) *J. Biol. Chem.* **278,** 13056–13060
31. Ayala, J. E., Streeper, R. S., Desgrosellier, J. S., Durham, S. K., Suwanichkul, A., Svitek, C. A., Goldman, J. K., Barr, F. G., Powell, D. R., and O'Brien, R. M. (1999) *Diabetes* **48,** 1885–1889
32. Barthel, A., Schmoll, D., Kruger, K. D., Bahrenberg, G., Walther, R., Roth, R. A., and Joost, H. G. (2001) *Biochem. Biophys. Res. Commun.* **285,** 897–902
33. Nakae, J., Kitamura, T., Silver, D. L., and Accili, D. (2001) *J. Clin. Investig.* **108,** 1359–1367
34. Nadal, A., Marrero, P. F., and Haro, D. (2002) *Biochem. J.* **366,** 289–297
35. Dijkers, P. F., Medema, R. H., Pals, C., Banerji, L., Thomas, N. S. B., Lam, E. W. F., Burgering, B. M. T., Raaijmakers, J. A. M., Lammers, J. W. J., Koenderman, L., and Coffer, P. J. (2000) *Mol. Cell. Biol.* **20,** 9138–9148
36. Medema, R. H., Kops, G. J. P. L., Bos, J. L., and Burgering, B. M. T. (2000) *Nature* **404,** 782–787
37. Brunet, A., Bonni, A., Zigmond, M. J., Lin, M. Z., Juo, P., Hu, L. S., Anderson, M. J., Arden, K. C., Blenis, J., and Greenberg, M. E. (1999) *Cell* **96,** 857–868
38. Kops, G. J. P. L., Dansen, T. B., Polderman, P. E., Saarloos, I., Wirtz, K. W. A., Coffer, P. J., Huang, T. T., Bos, J. L., Medema, R. H., and Burgering, B. M. T. (2002) *Nature* **419,** 316–321
39. Nakae, J., Kitamura, T., Kitamura, Y., Biggs, W. H., III, Arden, K. C., and Accili, D. (2003) *Dev. Cell* **4,** 119–129
40. Bois, P. R. J., and Grosveld, G. C. (2003) *EMBO J.* **22,** 1147–1157
41. Nakae, J., Biggs, W. H., III, Kitamura, T., Cavenee, W. K., Wright, C. V., Arden, K. C., and Accili, D. (2002) *Nat. Genet.* **32,** 245–253
42. Kamei, Y., Mizukami, J., Miura, S., Suzuki, M., Takahashi, N., Kawada, T., Taniguchi, T., and Ezaki, O. (2003) *FEBS Lett.* **536,** 232–236
43. Puigserver, P., Rhee, J., Donovan, J., Walkey, C. J., Yoon, J. C., Oriente, F., Kitamura, Y., Altomonte, J., Dong, H., Accili, D., and Spiegelman, B. M. (2003) *Nature* **423,** 550–555
44. Ogg, S., Paradis, S., Gottlieb, S., Patterson, G. I., Lee, L., Tissenbaum, H. A., and Ruvkun, G. (1997) *Nature* **389,** 994–999
45. Brennan, K. J., and Hardeman, E. C. (1993) *J. Biol. Chem.* **268,** 719–725
46. Nagy, T. R., and Clair, A. L. (2000) *Obes. Res.* **8,** 392–398
47. Hahn, C. G., and Covault, J. (1990) *Anal. Biochem.* **190,** 193–197
48. Ogilvie, R. W., and Feeback, D. L. (1990) *Stain Technol.* **65,** 231–241
49. Takahashi, M., Tsuboyama-Kasaoka, N., Nakatani, T., Ishii, M., Tsutsumi, S., Aburatani, H., and Ezaki, O. (2002) *Am. J. Physiol.* **282,** G338–G348
50. Adams, M. D., Kerlavage, A. R., Fleischmann, R. D., Fuldner, R. A., Bult, C. J., Lee, N. H., Kirkness, E. F., Weinstock, K. G., Gocayne, J. D., White, O., *et al.* (1995) *Nature* **377,** 3–174
51. Furukawa-Hibi, Y., Yoshida-Araki, K., Ohta, T., Ikeda, K., and Motoyama, N. (2002) *J. Biol. Chem.* **277,** 26729–26732
52. Tran, H., Brunet, A., Grenier, J. M., Datta, S. R., Fornace, A. J., Jr., DiStefano, P. S., Chiang, L. W., and Greenberg, M. E. (2002) *Science* **296,** 530–534
53. Furuyama, T., Kitayama, K., Yamashita, H., and Mori, N. (2003) *Biochem. J.* **375,** 365–371
54. Garry, D. J., Ordway, G. A., Lorenz, J. N., Radford, N. B., Chin, E. R., Grange, R. W., Bassel-Duby, R., and Williams, R. S. (1998) *Nature* **395,** 905–908
55. Lecker, S., Jagoe, R., Gilbert, A., Gomes, M., Baracos, V., Bailey, J., Price, S., Mitch, W., and Goldberg, A. (2004) *FASEB J.* **18,** 39–51
56. Ezaki, O. (1997) *Biochem. Biophys. Res. Commun.* **241,** 1–6
57. Chakravarthy, M., Davis, B., and Booth, F. (2000) *J. Appl. Physiol.* **89,** 1365–1379
58. Machida, S., Spangenburg, E., and Booth, F. (2003) *J. Cell. Physiol.* **196,** 523–531
59. Yamanouchi, K., Soeta, C., Suzuki, S., Hasegawa, T., Naito, K., and Tojo, H. (2000) *J. Vet. Med. Sci.* **62,** 1213–1216
60. Sandri, M., Sandri, C., Gilbert, A., Skurk, C., Calabria, E., Picard, A., Walsh, K., Schiaffino, S., Lecker, S., and Goldberg, A. (2004) *Cell* **117,** 399–412
61. Stitt, T., Drujan, D., Clarke, B., Panaro, F., Timofeyva, Y., Kline, W., Gonzalez, M., Yancopoulos, G., and Glass, D. (2004) *Mol. Cell* **14,** 395–403
62. Schneider, M., Wolf, E., Hoeflich, A., and Lahm, H. (2002) *J. Endocrinol.* **172,** 423–440
63. Kamei, Y., Ohizumi, H., Fujitani, Y., Nemoto, T., Tanaka, T., Takahashi, N., Kawada, T., Miyoshi, M., Ezaki, O., and Kakizuka, A. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 12378–12383

Exp. Anim. 53(2), 103–111, 2004

Chromosomal Mapping and Zygosity Check of Transgenes Based on Flanking Genome Sequences Determined by Genomic Walking

Akira NOGUCHI, Naho TAKEKAWA, Thorbjorg EINARSDOTTIR, Minako KOURA, Yoko NOGUCHI, Kaoru TAKANO, Yoshie YAMAMOTO, Junichiro MATSUDA, and Osamu SUZUKI

Department of Veterinary Science, National Institute of Infectious Diseases, 1–23–1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan

Abstract: *Transgenes can affect transgenic mice via transgene expression or via the so-called positional effect. DNA sequences can be localized in chromosomes using recently established mouse genomic databases. In this study, we describe a chromosomal mapping method that uses the genomic walking technique to analyze genomic sequences that flank transgenes, in combination with mouse genome database searches. Genomic DNA was collected from two transgenic mouse lines harboring pCAGGS-based transgenes, and adaptor-ligated, enzyme restricted genomic libraries for each mouse line were constructed. Flanking sequences were determined by sequencing amplicons obtained by PCR amplification of genomic libraries with transgene-specific and adaptor primers. The insertion positions of the transgenes were located by BLAST searches of the Ensembl genome database using the flanking sequences of the transgenes, and the transgenes of the two transgenic mouse lines were mapped onto chromosomes 11 and 3. In addition, flanking sequence information was used to construct flanking primers for a zygosity check. The zygosity (homozygous transgenic, hemizygous transgenic and non-transgenic) of animals could be identified by differential band formation in PCR analyses with the flanking primers. These methods should prove useful for genetic quality control of transgenic animals, even though the mode of transgene integration and the specificity of flanking sequences needs to be taken into account.*

Key words: *chromosomal mapping, flanking primers, genomic walking, zygosity check*

Introduction

Transgenes can exert effects in transgenic animals by transgene expression and by the so-called positional effect. Transgene expression often varies among multiple lines that have been derived from different founders. This variation may result from the position of the transgene integration site as well as the copy

(Received 22 August 2003 / Accepted 4 November 2003)

Address corresponding: O. Suzuki, Department of Veterinary Science, National Institute of Infectious Diseases, 1–23–1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan

number of the transgene. A transgene integrated into the X chromosome would exhibit sex-linked inheritance. Side effects, such as unexpected gene knockouts, may be caused by the disruption of an endogenous gene due to transgene insertion [e.g., 11, 14]. Thus, information on transgene position is important for analyses of transgenic animals.

Genes are typically localized by *in situ* hybridization methods, such as FISH [e.g., 3, 6]. In recent years, mouse genome databases have expanded rapidly and offer highly integrated records of genomic sequences, chromosome maps, and related information. These databases can be used to map transgenes on mouse chromosomes with great accuracy if the flanking sequences of the transgene can be determined.

Here, we describe a rapid, simple method using genomic walking [12] to determine genomic sequences that flank pCAGGS-based transgenes (Fig. 1A) [8]. The strategy of the genomic walking used in this study is outlined in Fig. 1B. Genomic sequences flanking the transgenes were determined by two consecutive PCR amplifications with multiple enzyme-digested, adaptor-ligated genomic libraries of transgenic animals (Fig. 1B). Determination of the nucleotide sequences flanking the transgene enables the transgene to be mapped in chromosomes by searching mouse genome databases. In addition, the sequence information can be used to construct locus-specific zygosity check systems with flanking primer methods (Fig. 1C). However, our experience suggests that modes of transgene integration should be considered as well as the flanking genomic sequences.

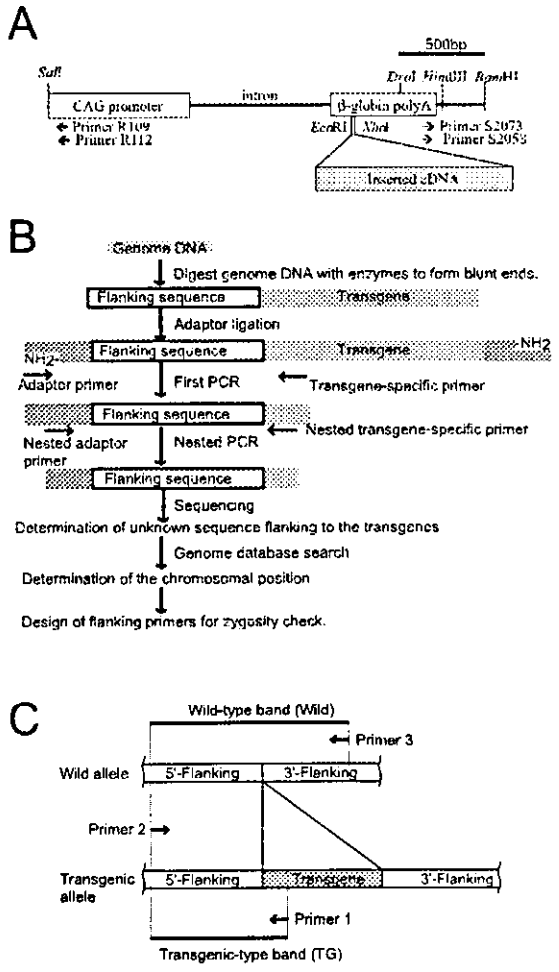


Fig. 1. Transgene construct (A) and strategies for chromosome mapping (B) and zygosity check (C) of the transgene. Transgenes were excised from pCAGGS-plasmids that contained mouse cDNA. Two sets of primers, R109 and R112, and S2058 and S2072, were used as transgene-specific primers for genomic walking to 5'-flanking and 3'-flanking regions, respectively. Figure 1B summarizes the procedure of the genomic walking in this study (modified from a chart in a manual of the Universal GenomeWalker kit). In brief, six enzyme-digested, adaptor-ligated genomic libraries were constructed from transgenic mice. Fragments flanking the transgenes were amplified by two consecutive PCR amplifications with adaptor and transgene-specific primers. Genomic sequences flanking the transgenes were determined by sequencing the fragments. Integration sites of the transgenes were determined by genome database searches. Genomic sequences flanking the other end of the transgene were determined by genomic walking from the 5'-flanking region to the 3'-flanking region with genomic libraries from non-transgenic mice. Flanking primers for a zygosity check were designed on the basis of the genomic sequences at both ends of the integration sites (See text for details). Figure 1C summarizes the strategy for zygosity check. Zygosity was judged by the differential bands produced with flanking primers, which were designed to distinguish between wild-type (Wild) and transgenic (TG) alleles by amplicon length.

Materials and Methods

Transgenic animals

This study used two lines of transgenic mice, 4c30 and CK35, at the eighth and sixth generations, respectively, and their non-transgenic parent strain, C57BL/6CrSlc (B6). The transgenic lines were produced in

our laboratory by zygote microinjection of transgene constructs based on a pCAGGS plasmid (Fig. 1A) [8]. Line 4c30 received a plasmid that contained mouse α2,3-sialyltransferase type II [5] cDNA, whereas line CK35 received a plasmid that contained human acid β-galactosidase [9] cDNA. Constructs containing a cDNA insert at the multi-cloning site (MCS) were excised from *Sal*I to *Bam*HI (approximately 3.5 kb long) and from *Sal*I to *Hin*dIII (approximately 4.5 kb long) from the plasmids and used for production of 4c30 and CK35, respectively. Southern blot analyses confirmed integration of the transgenes in the two lines. Animal experiments were performed according to the Guides for Animal Experiments Performed at NIID.

Determination of flanking sequences using genomic walking for chromosomal mapping of the transgene

The genomic sequences that flanked the transgenes were determined by genomic walking using the Universal GenomeWalker™ kit (BD Bioscience Clontech, Palo Alto, CA) with a slight modification. Adaptor-ligated genomic DNA libraries of the transgenic lines were constructed with tail DNA digested with six restriction enzymes: *Dra*I, *Sca*I, *Pvu*II, *Eco*RV, *Ssp*I, and *Stu*I (Takara Bio Inc., Tokyo, Japan). The genomic walk consisted of two PCR amplifications. The primary PCR amplification was performed with an outer transgene-specific primer (R112, 5'-CCA GGC GGG CCA TTT ACC GTA AGT TAT-3') and the outer adaptor primer provided in the kit. The primary PCR mixture was diluted and used as the template for a nested PCR amplification with a nested transgene-specific primer (R109, 5'-GGC GGG CCA TTT ACC GTA AGT TAT GT-3') and the nested adaptor primer provided in the kit. All PCR amplifications were performed with a hot-start DNA polymerase (HotStarTaq; Qiagen K.K., Tokyo, Japan) in a Hybaid PCR Express Thermal Cycler (Thermo Hybaid, Ashford, Middlesex, UK). The hot-start DNA polymerase requires a 15-min incubation at 95°C for activation as the first step of PCR amplification. The primary PCR amplification was performed at 95°C for 15 min, followed by 40 cycles at 94°C for 2 s and 68°C for 5 min. The nested PCR amplification was performed at 95°C for 15 min, followed by 30 cycles at 94°C for 2 s and 68°C for 5 min. The nested PCR products were separated by electrophoresis on a 2% agarose gel in TAE buffer. The nested PCR bands were visualized with ethidium bromide staining and extracted from the gels using a MinElute Gel Extraction kit (QIAGEN). The nucleotide sequences of the PCR bands were determined by direct sequencing using a dye-primer method (Thermo Sequenase Primer Cycle Sequencing kit with 7-deaza GTP; Amersham Biosciences, Piscataway, NJ) with a DSQ-2000L DNA Sequencer (Shimadzu Corp., Kyoto, Japan). Among the nested PCR products, the DNA fragment containing the 5'-flanking sequence of the transgene was determined by identifying a nucleotide sequence that was not found in the transgene and that continued to the 5'-transgene sequence. Other than the adaptor primers provided in the kit, all primers were designed using the Primer3 program [10]. The transgene insertion sites in chromosomes were determined using a BLAST search of the flanking sequences in the Ensembl genome database [2] via the Internet (http://www.ensembl.org).

Genomic walking analyses that proceeded backwards from the tails of the transgene were performed with two primers S2058 (5'-GTA TAT GAA ACA GCC CCC TGC TGT CCA-3') and S2072 (5'-CCC CTG CTG TCC ATT CCT TAT TCC ATA G-3') in combination with adapter primers provided in the kit.

Zygosity check of the transgene by a flanking primer method

For locus-specific zygosity checks, we performed PCR analyses using flanking primers with sequences based on the flanking sequences of the transgenes (Fig. 1C). The flanking sequence around a transgene was determined by performing two consecutive genomic walking experiments. The first walk from the transgene to the 5'-flanking region was performed with genomic libraries of the transgenic line. A second walk from the 5'- to 3'-flanking region was performed with non-transgenic genomic libraries to determine the 3'-flanking sequence of the transgene. Two flanking primers, designated as Primer 2 and Primer 3 in Fig. 1C, were designed with the Primer3 program [10] using 5'- and 3'-flanking sequences at the transgene insertion site of transgenic lines 4c30 and CK35. Primer R109 was used as Primer 1, as shown in Fig. 1C. The zygosity of these transgenic animals was determined using PCR analyses with their tail DNA and the flanking primers under the following thermal conditions: 95°C for 15 min, followed by 30 cycles of 94°C for 2 s, 55°C for 15 s, and 72°C for 1

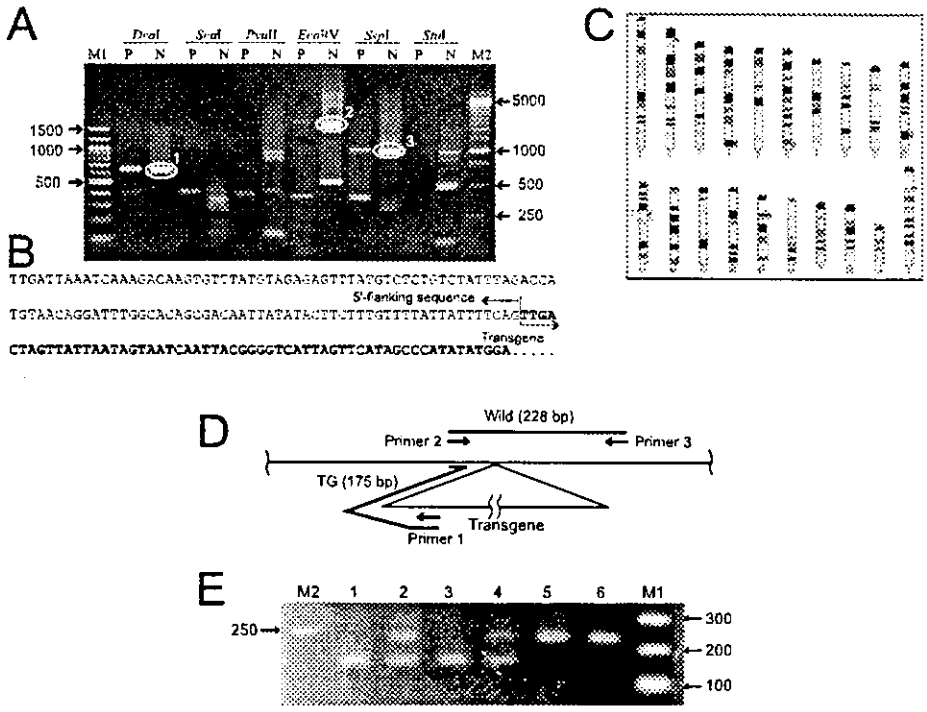


Fig. 2. Chromosomal mapping and zygosity check in transgenic mouse line 4c30. The products of primary (P) and nested (N) PCR amplification with 4c30 genomic libraries were separated by agarose gel electrophoresis (A). Sequence information revealed that bands #1 and #2 in Fig. 2A were derived from the tandem repeat of the transgene. The sequence (B) flanking the transgene was determined from the ~1,000-bp *Stu*I band (band #3 in Fig. 2A). A search of the Ensembl genome database with the flanking sequence suggested that the transgene is located on chromosome 11 (boxed in panel C). Primers were designed for the zygosity check of 4c30 (D). Zygosity of six mice was determined by the differential bands on a gel (E) . Lanes 1 and 3: homozygous transgenic; Lanes 2 and 4: hemizygous; Lanes 5 and 6: non-transgenic; M1: 100-bp DNA ladder; M2: 250-bp DNA ladder.

min. The PCR products were separated by electrophoresis on a 2% agarose gel in TAE buffer, and bands were detected by ethidium bromide staining with ultraviolet illumination. The zygosities of all the animals used for the zygosity check in this study were confirmed in advance using Southern blot analysis, breeding tests, and zygosity-specific symptoms (homozygous-specific cardiac hypertrophy) in 4c30, and the enzyme activity of the transgene products in CK35 tissues.

Results

Mapping and genotyping of mouse line 4c30

Figure 2A shows the electrophoretic patterns of primary and nested PCR amplification products from the first genomic walking experiment. The gel contained three major bands consistent with both primary (P) and nested (N) PCR amplifications: ~650 bp of *Dra*I library, ~1,500 bp of *Eco*RV, and ~1,000 bp of *Ssp*I library. Sequencing analysis revealed that the first two bands were derived from the tandem repeat of the transgene, and that only the last band contained the 5'-flanking sequence (Fig. 2B). A BLAST search with the 5'-flanking sequence revealed that the transgene was integrated into Chromosome 11 (Fig. 2C).

After the second genomic walking experiment, two primers, labeled Primer 2 and Primer 3 in Fig. 1C, were designed so that the primers in combination with Primer R109 (labeled Primer 1 in Fig. 1A) would produce wild-type and transgenic alleles as 228-bp (Wild) and 175-bp (TG) bands, respectively (Fig. 2D). As shown in Fig. 2E, analysis with these primers revealed

a clear genotyping of six line 4c30 mice as homozygous (lanes 1 and 3) with a Wild band only, wild-type (lanes 5 and 6) with a TG band only, and hemizygous (lanes 2 and 4) with both bands. These genotypings matched those determined using the Southern blot analysis perfectly.

Mapping and genotyping of mouse line CK35

More bands were detected in the electrophoretic patterns of nested PCR products from mouse line CK35, as compared with the patterns observed for line 4c30 (Fig. 3A). Of these CK35 bands, only one band of ~700 bp in the *Ssp*I library contained the 5'-flanking genomic sequence (Fig. 3B). A BLAST search of this sequence in the Ensembl genome databases indicated that the transgene was mapped to chromosome 3 (Fig. 3C).

Flanking primers 2 and 3 for the mouse line CK35 were designed from two consecutive genomic walking experiments, as described for mouse line 4c30 (Fig. 3D). The expected Wild (329-bp) and transgene (213-bp) bands are shown as solid lines in Fig. 3D. A zygosity check was performed using PCR amplification with these primers and R109 as Primer 1. The results revealed unexpected bands of about the same length as the anticipated Wild band in homozygous mice (lanes 1 and 2 in Fig. 3E), in which zygosity had been confirmed by Southern blot analysis, yet only single bands were produced when primer pairs for wild-type (Primer 2 and Primer 3) and transgenic (Primer 1 and Primer 2) alleles were used separately (Fig. 3F). Sequencing of the unexpected PCR products revealed that the nucleotide sequence of the transgene at the 3'-flanking region was identical to the nucleotide sequence at the 5'-flanking region. Since Primer 1 could bind to two regions, one additional band, shown as a dotted line in Fig. 3D, was produced when all three primers were simultaneously used for PCR amplification. Based on these facts, new flanking primers were designed so that two TG bands derived from 5'- and 3'-junctions would move to about the same position on a gel and yet be clearly distinct from a Wild band (Fig. 3G). With these new primers, clear genotyping of line CK35 was achieved as shown in Fig. 3H. As expected, two TG bands formed one bold band on a gel (lanes 1 and 2 in Fig. 3H).

Backward genomic walking in mouse line 4c30

Figure 4 shows electrophoretic patterns of nested PCR products obtained by backward genomic walking with B6 and 4c30 strains. Note that a considerable number of amplicons were found even in genomic libraries of non-transgenic mice as well as in 4c30 libraries.

Discussion

Our results with two examples show that the chromosomal mapping and design of flanking primers for a zygosity check can be achieved by the genomic walking technique in combination with a genome database search. While this method is quite powerful, the mode of transgene integration and specificity of flanking sequences need to be taken into account.

Figures 2 and 3 summarize two examples of successful chromosomal mapping using genomic walking. The locations of transgene insertions were determined by our method on chromosomes 11 and 3 in 4c30 mice and CK35 mice, respectively (Figs. 2C and 3C). Transgenes are often integrated into a host genome in tandem. In 4c30, the transgenes contained a *Dra*I site in the pCAGGS backbone (Fig. 1A) and an *Eco*RV site in the inserted cDNA sequence. If head-to-tail tandem repeats of the transgene exist in a host genome, then amplicons are produced in *Dra*I- and *Eco*RV-digested genome libraries. As expected, approximately 650-bp *Dra*I (band #1) and 1,700-bp *Eco*RV (band #2) bands were seen in the example (Fig. 2A). Tandem repeats were also found in CK35 mice (band #1 in *Dra*I library). Thus, genomic walking can be used for detection of tandem repeats, although this should be confirmed by sequencing.

Successful zygosity checks were also accomplished in both examples by PCR amplification with the flanking primers and tail DNA from transgenic and non-transgenic mice. The genotypes determined using the flanking primers matched those determined in advance using the Southern blot analysis. As shown in Fig. 2D, three primers for the 4c30 line were designed so that amplification would produce wild-type and transgenic alleles as 228 bp (Wild) and 175 bp (TG) bands, respectively (Fig. 2D). Analyses with these primers clearly provided the genotypes of the six mice (Fig. 2E). As shown in Fig. 3, flanking primer design for the CK35 line was also achieved, although certain elaborations were needed, because the mode of transgene integration in CK35 was more complicated than in 4c30. This topic will be discussed in detail

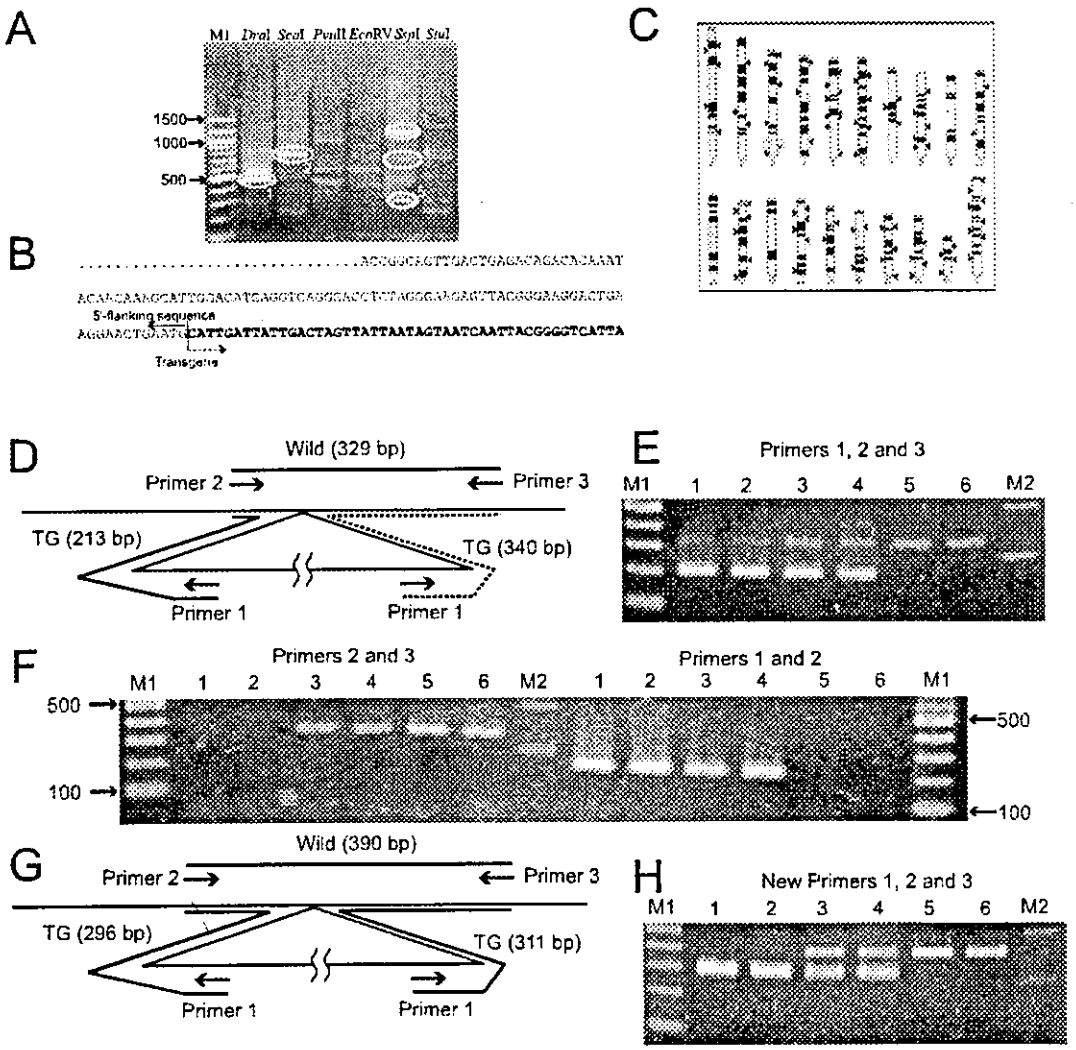


Fig. 3. Chromosomal mapping and zygosity check in transgenic mouse line CK35. The products of the nested PCR amplification of genomic libraries from CK35 genomic DNA were separated by agarose gel electrophoresis (A). Band #1 was derived from the tandem repeat of the transgene. Bands #2, #3 and #5 in Fig. 3A were not related to the transgene. The sequence flanking the transgene (B) was determined from the ~700-bp *SspI* band (band #4 in Fig. 3A). A search of the Ensembl genome database with the flanking sequence suggested that the transgene is located on chromosome 3 (boxed in panel C). Flanking primers for zygosity check of CK35 were designed as follows (panels D-H). DNA samples for Lanes 1 and 2 had been pre-confirmed as homozygous transgenic, lanes 3 and 4 as hemizygous transgenic, and lanes 5 and 6 as non-transgenic by Southern blot analysis. D) Flanking primer design for the zygosity check in line CK35 at the first attempt. Three primers were designed so that PCR amplification produced results in 329-bp bands (Wild) for wild-type alleles and 213-bp bands (TG) for transgenic allele, as indicated by solid lines. The dotted line shows unexpected TG bands produced by Primer 1 and Primer 3. E) The zygosity check of transgenic line CK35 with primers shown in panel D. Band patterns indicated that zygosities seen in lanes 1 and 2 were not relevant to our Southern blot analysis, while the patterns in the other lanes were the same. Sequence analysis of bands at a position similar to Wild bands in lanes 1 and 2 revealed that the bands were produced from the 3'-junction of the transgene by primer pair 1 and 3 as shown in a dotted line in panel D. F) The zygosity check of transgenic line CK35 with two sets of primer pairs in panel D. TG and wild-type alleles were clearly observed using primer pairs separately. G) The principle of the zygosity check by PCR with re-designed flanking primers for line CK35. The new primers were designed so that two TG bands, 296 bp and 311 bp long, derived from the 5'- and 3'-junctions of the transgene, respectively, would move to about the same position in the gel and yet be clearly distinguishable from a Wild band (390 bp). H) The zygosity check of transgenic line CK35 with new primers shown in panel G. Zygosity was clearly determined by band patterns, as already confirmed by Southern blot analysis. In panels E, F and H, M1 and M2 indicate 100-bp and 250-bp DNA ladders, respectively.

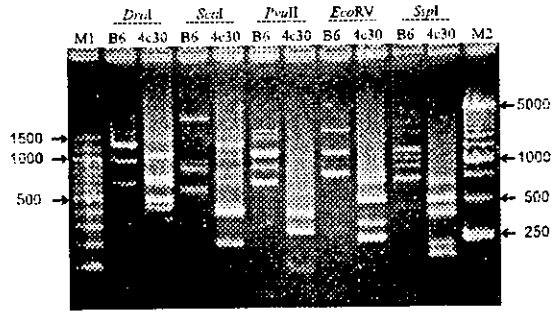


Fig. 4. Electrophoretic patterns of nested PCR products from backward genomic walking with B6 and 4c30 strains. Note that a considerable number of amplicons were found even in non-transgenic mouse genomic libraries as well as in 4c30 libraries. M1: 100-bp DNA ladder; M2: 250-bp DNA ladder.

below. PCR amplification with flanking primers has the potential to identify the transgene copy number from PCR product length using Primer 2 and Primer 3 shown in Fig. 1C. Nevertheless, in some cases this might be impossible if only a few repeats of the transgene exceed the length of DNA that routine PCR can amplify.

The mode of transgene integration can complicate flanking sequence determination. As shown in Fig. 4, our experiments revealed that the 5'-flanking region of pCAGGS-based transgenes was easier to determine than the 3'-flanking region, because primers with greater specificity could be designed for the 5'-flanking sequence. Primers for the CMV enhancer region are highly specific for detecting transgenes in the mouse genome, because this region is derived from a non-mammalian genome. Conversely, primers for the β-globin polyadenylation signal sequence residing downstream from the MCS of a pCAGGS-based transgene produced many non-specific bands (Fig. 4), probably because many highly homologous regions exist in the mouse genome. Efficient 3'-end genomic walking may require specific tag sequences at the 3'-end of the transgene. In addition, the mode of tandem repeat of the transgene should be considered when determining flanking sequences. Figure 5 illustrates the three patterns of transgene integration into the host genome. Using 5'-end genomic walking, one and two

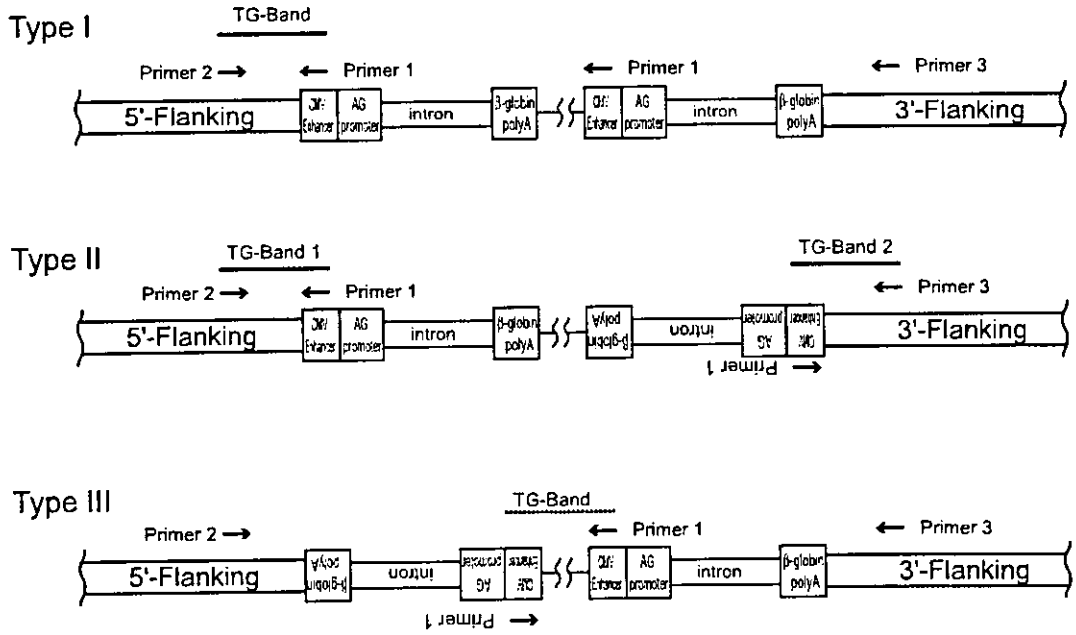


Fig. 5. Integration patterns of transgenes into a host genome. The three possible types of transgene integration at 5'- and 3'-junctions are: Type I, head and tail; Type II, head and head; Type III, tail and tail. In type I, only one expected band (TG-band) is produced by PCR amplification, while two bands (TG-bands 1 and 2) are amplified in type II. No band or an unexpected band (TG-band shown by a dotted line) may be produced by PCR in type III. See discussion for details.

PCR amplicons are obtained in types I and II, respectively. In type III, a total of three amplicons (two TG and one Wild band) should be considered for a zygosity check. In our study, the CK35 line turned out to be an example of the type II integration of the transgene (Fig. 3G). In type III (i.e., tail and tail orientation), a genome-specific sequence cannot be determined by our method, which captures the 5'-flanking sequence of pCAGGS-based transgenes. PCR amplification with Primer 1 might produce a band seen as a "TG-band" in type III, if a head-to-head concatenation of transgenes has occurred, but the band would not be informative with respect to the flanking genomic sequence. In addition, unknown sequences, such as fragmented copies of the transgene, might also be incorporated in the insertion site. Therefore, careful judgment should be exercised when determining which PCR bands contain sequences flanking the transgene.

The redundancy of genomic sequences also complicates localization of the insertion position of the transgene. Higher animals like mice have large genomes with complex nucleotide sequence conformations, such as gene duplications, pseudogenes, interspersed-type repeats such as SINE and LINE, etc. Consequently, similar nucleotide sequences are disseminated throughout the mouse genome. If the transgenes are integrated in or near such redundant sequences, it will be quite difficult to locate the transgene at a single location based only on the flanking sequence information. Our preliminary experiments showed that a sequence more than 200 bases long was insufficient to determine a single site in the mouse genome in certain cases, because many identical sequences were found in multiple chromosomes in the genome database (data not shown). Thus, sequence information is a powerful tool, but not always sufficient for practical genotyping.

Our mapping method has both advantages and disadvantages in comparison with the widely used cytogenetic mapping using the FISH technique. Mapping with FISH is more versatile, since it can be applied to any transgenes and does not depend on the flanking sequence information or the mode of integration of the transgenes. The resolution of mapping with FISH is lower than the resolution of our method, which can locate transgenes at the sequence level, whereas FISH does so at the chromosome band level. Although our method involves many steps, it can be carried out in three days using commercially available kits, which, in our opinion, is comparable to genomic mapping using FISH in terms of ease of performance.

Transgenes are not always integrated at a single site [6]. Our method can be applied to mice harboring transgenes in multiple sites, since multiple PCR products derived from multiple integration sites can be separated in a gel. Another possible strategy is to map transgenes in descendants in later generations, in which transgenes integrated at multiple sites are segregated. No triple-loci integration was observed in EGFP transgenic mice [6], suggesting that only two integration sites need be considered using descendants, even when the founder mice have been judged to have multiple transgene integrations using methods such as Southern blot analysis. The original composition of the transgene integration sites in founder mice could be re-constituted afterwards using a hereditary analysis of multiple sublines. Chromosome translocation was also observed in transgenic mice [6]. If chromosomes are rearranged at transgene integrations, flanking sequence information would be inappropriate for locating the position of the site, depending on the scale of the rearrangement. A zygosity check system could still be designed from flanking sequences determined using our method, but our method would not be applicable for precise mapping of the transgene.

A zygosity check by PCR amplification with flanking primers is powerful, because it is site-specific. Zygosity checking in transgenic animals is a critical step for analysis of transgene effects. All transgenic mice should be genotyped before experiments are performed. For production and maintenance of transgenic mouse lines, essentially all animals used as breeding pairs should be genotyped as well. Thus, large-scale genotyping should be routinely employed, and this requires a simple and rapid genotyping method. Several techniques are used for this purpose [reviewed in, e.g., 1]. Although Southern and dot blot analyses are commonly used, these methods are technically demanding and often provide ambiguous results. Moreover, the use of radioisotopes may cause problems in some laboratories. Although genotyping by FISH [4, 7] and real-time quantitative PCR amplification [13] have been reported, these methods do not offer the simplicity required for practical genotyping on a large scale. PCR amplification with flanking primers would be a rapid,

simple, and position-specific method for zygosity checking of the transgene. As shown in Fig. 1C, PCR amplification of a mouse genome with three primers enables zygosity to be judged by differential band formation. A drawback of the method is the need for nucleotide sequences that flank the transgene, which typically has an unpredictable insertion site. As described in this paper, the problem can be solved by genomic walking in most cases and, therefore, the flanking primer method can be practically applied. In addition, site-specific genotyping with flanking primers enables one to check the zygosity of each transgenic allele even when animals inherit transgenes from different transgenic lines with the same transgenes in different chromosomal positions (i.e., alleles). Since this checking system distinguishes between multiple transgenic lines containing the same transgene(s) inserted at different loci, multi-cross hybrids between lines can be produced with confirmed zygosity at individual sites. This results in the production of mice harboring more copies of the transgene than the original lines, a strategy that is useful if the gene dosage is a major concern.

Chromosome position and the zygosity of the transgene are important factors in the analysis of transgene expression in transgenic animals. In particular, the presence and chromosomal position of the transgene should be monitored for quality control of transgenic animals as genetically certified transgenic lines. As described above, determination of the flanking sequence of the transgene by genomic walking is useful, not only for chromosome mapping, but also for constructing zygosity check systems. Genomic walking is a powerful tool for genomic studies, especially with rapidly expanding genome databases like Ensembl [2]. For example, the second genomic walk for a zygosity check may be replaced by a genome database search. Thus, the rapid, simple method described here should facilitate the study of transgene effects.

References

1. Hogan, B., Beddington, R., Constantini, F., and Lacy, E. 1994. Identifying homozygous transgenic mice or embryos, pp. 305–308. *In*: Manipulating the mouse embryo: a laboratory manual. Cold Spring Harbor Laboratory Press, Woodbury.
2. Hubbard, T., Barker, D., Birney, E., Cameron, G., Chen, Y., Clark, L., Cox, T., Cuff, J., Curwen, V., Down, T., Durbin, R., Eyras, E., Gilbert, J., Hammond, M., Huminiecki, L., Kasprzyk, A., Lehvaslaiho, H., Lijnzaad, P., Melsopp, C., Mongin, E., Pettett, R., Pocock, M., Potter, S., Rust, A., Schmidt, E., Searle, S., Slater, G., Smith, J., Spooner, W., Stabenau, A., Stalker, J., Stupka, E., Ureta-Vidal, A., Vastrik, I., and Clamp, M. 2002. The Ensembl genome database project. *Nucleic Acids Res.* 30: 38–41.
3. Ichikawa, S., Ozawa, K., and Hirabayashi, Y. 1998. Assignment1 of a UDP-glucose:ceramide glucosyltransferase gene (*Ugcg*) to mouse chromosome band 4B3 by in situ hybridization. *Cytogenet. Cell Genet.* 83: 14–15.
4. Kulnane, L.S., Lehman, E.J., Hock, B.J., Tsuchiya, K.D., and Lamb, B.T. 2002. Rapid and efficient detection of transgene homozygosity by FISH of mouse fibroblasts. *Mamm. Genome* 13: 223–226.
5. Lee, Y.C., Kurosawa, N., Hamamoto, T., Nakaoka, T., and Tsuji, S. 1993. Molecular cloning and expression of Galβ1,3GalNAc β2,3-sialyltransferase from mouse brain. *Eur. J. Biochem.* 216: 377–385.
6. Nakanishi, T., Kuroiwa, A., Yamada, S., Isotani, A., Yamashita, A., Tairaka, A., Hayashi, T., Takagi, T., Ikawa, M., Matsuda, Y., and Okabe, M. 2002. FISH analysis of 142 EGFP transgene integration sites into the mouse genome. *Genomics* 80: 564–574.
7. Nishino, H., Herath, J.F., Jenkins, R.B., and Sommer, S.S. 1995. Fluorescence in situ hybridization for rapid differentiation of zygosity in transgenic mice. *Biotechniques* 19: 587–590, 592.
8. Niwa, H., Yamamura, K., and Miyazaki, J. 1991. Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* 108: 193–199.
9. Oshima, A., Tsuji, A., Nagao, Y., Sakuraba, H., and Suzuki, Y. 1988. Cloning, sequencing, and expression of cDNA for human β-galactosidase. *Biochem. Biophys. Res. Commun.* 157: 238–244.
10. Rozen, S. and Skaletsky, H.J. Primer3. Code available at http://www-genome.wi.mit.edu/genome_software/other/primer3.html. 1998.
11. Schnieke, A., Harbers, K., and Jaenisch, R. 1983. Embryonic lethal mutation in mice induced by retrovirus insertion into the alpha 1(I) collagen gene. *Nature* 304: 315–320.
12. Siebert, P.D., Chenchik, A., Kellogg, D.E., Lukyanov, K.A., and Lukyanov, S.A. 1995. An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res.* 23: 1087–1088.
13. Tesson, L., Heslan, J.M., Menoret, S., and Anegon, I. 2002. Rapid and accurate determination of zygosity in transgenic animals by real-time quantitative PCR. *Transgenic Res.* 11: 43–48.
14. Xiang, X., Benson, K.F., and Chada, K. 1990. Mini-mouse: disruption of the pygmy locus in a transgenic insertional mutant. *Science* 247: 967–969.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry
ELSEVIER
Bioorganic & Medicinal Chemistry 12 (2004) 995–1002

Convenient synthesis and evaluation of glycosidase inhibitory activity of α- and β-galactose-type valienamines, and some *N*-alkyl derivatives

Seiichiro Ogawa,[a,*] Yuko Sakata,[a] Naoyuki Ito,[a] Maiko Watanabe,[a] Kazuya Kabayama,[a] Masayoshi Itoh[b] and Takashi Korenaga[b]

[a]*Department of Biosciences and Informatics, Faculty of Science and Technology, Keio University, Hiyoshi, Kohoku-ku, Yokohama, 223-8522 Japan*
[b]*Department of Chemistry, Tokyo Metropolitan University, Minami-Ohsawa, Hachioji, Tokyo, 192-0397 Japan*

Received 17 November 2003; accepted 11 December 2003

Abstract—Valienamine analogues having α- and β-galactose-type structures were synthesized by racemic modification from (1*SR*,2*RS*,3*SR*)-6-methylenecyclohex-4-ene-1,2,3-triol. Four *N*-alkyl derivatives of the β-anomer were readily prepared selectively by treatment of key intermediate 2,6-di-*O*-acetyl-3,4-*O*-isopropylidene-5a-carba-α- and β-L-*arabino*-hex-5(5a)-enopyranosyl bromides with alkyl amines. All compounds were assayed for inhibitory activity against six glycosidases, and the *N*-dodecyl derivative was shown to be a very strong inhibitor of β-galactosidase (IC_{50} 0.01 μM, bovine liver).

1. Introduction

Mutant forms of enzyme proteins have been shown to be labile and rapidly degraded in somatic cells from patients with lysosomal storage diseases.[1] However, they can be stabilized and transported to the lysosomes by competitive inhibitors of low molecular weight (chemical chaperones) for therapeutic purposes. This phenomenon was confirmed for the mutant enzyme causing Fabry disease (α-galactosidase deficiency),[2] and very recently this strategy has been extended to two other diseases involving β-galactosidase deficiency in the central nervous system: GM1-gangliosidosis and Morquio B disease. Some unsaturated carbaglycosylamine derivatives have recently been found with remarkable effects: galactose-type[3] [*N*-octyl-β-D-5aCGal(5,5a)enamine,† GalX **1**] and glucose-type *N*-octyl-β-valienamine[4,5] [*N*-octyl-β-D-5aCGlc(5,5a)enamine, GlcX **2**] for β-galactosidase and β-glucosidase, respectively (Fig. 1). Compound **1** was actually demonstrated[3] to be a very potent inhibitor of human β-galactosidase (IC_{50} = 0.3 μM), and has been extensively studied as a candidate novel therapeutic agent for treatment of several human genetic diseases. Thus, such unsaturated 5a-carba-sugars are now regarded as important lead compounds.

2. Results and discussion

We report here a sequence worked out by modification of the route[6] for 5a-carba-α-fucopyranosylamines[7] with nucleophilic substitution of the primary bromo group of 2,3,4-tri-*O*-acetyl-6-bromo-6-deoxy-5a-carba-α- and β-DL-*arabino*-hex-5(5a)-enopyranosyl bromides‡ (**5**α,β), derived from the alkadiene[8] **4** (Scheme 1). Selective preparation of **4** was also here achieved in a 72% yield

* Corresponding author. Tel.: +81-045-566-1559; fax: +81-045-566-1551; e-mail: ogawa@bio.keio.ac.jp

† For convenience, we herewith propose abbreviations for naming the carba sugar and unsaturated carba sugars as follows: 5a-Carba-α-D-glucopyranose: α-D-5aCGlc; 2-Acetamido-2-deoxy-5a-carba-α-D-glucopyranose: α-D-5aCGlcNAc; 5a-Carba-α-D-glucopyranosylamine (validamine): α-D-5aCGlcamine; 5a-Carba-α-D-*xylo*-hex-(5,5a)-enopyranosylamine (valienamine): α-D-5aCGlc(5,5a)enamine not β-L-5aCIdo(5,5a)enamine. As exemplified above, the (5,5a)-unsaturated 5a-carba-sugar is a named derivative of the parent D-hexopyranose.

‡ In the text use of carba-sugar nomenclature following the IUPAC-IUBMB Nomenclature of Carbohydrates (Recommendation 1996: *Carbohydr. Res.*, **1997**, *297*, 1-92) is discussed. However, in the experimental section, IUPAC nomenclature for bi- and tri-cyclic compounds was used throughout for the sake of general understanding.

0968-0896/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2003.12.016

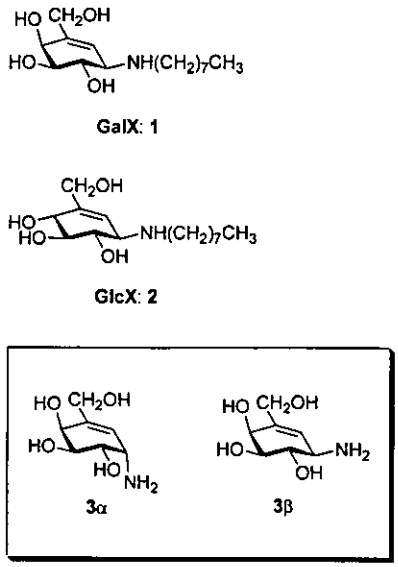
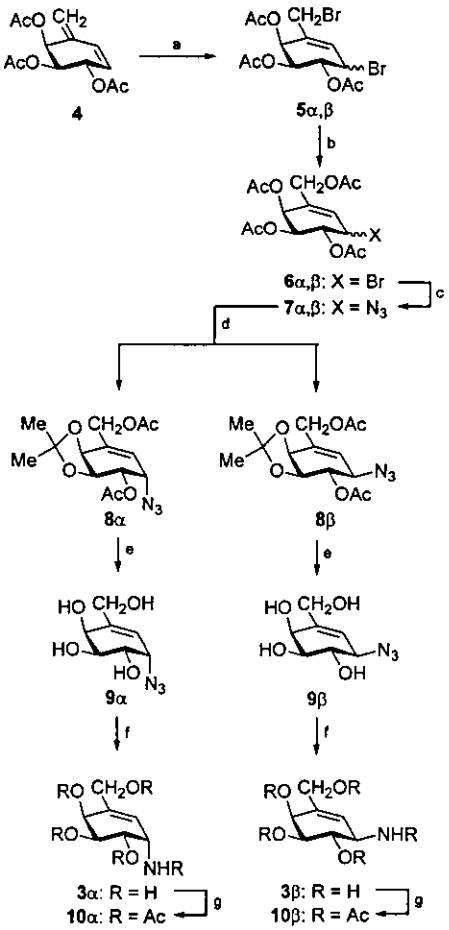


Figure 1.



Scheme 1. Reagents and conditions: (a) Br_2 (molar equiv), CCl_4, 1 h, rt; (b) NaOAc (molar equiv), DMF, 20 h, rt; (c) NaN_3, DMF, r.t.; (d) 1M NaOMe-MeOH, 1 h, rt; $(MeO)_2CMe_2$;, DMF, TsOH hydrate, 1 day, rt; Ac_2O, pyridine; (e) 4 M HCl:THF (1:1), 1 h, reflux; f) Ph_3P (3 molar equiv), 50% aq THF, 1 day, rt; Dowex 50 W×2 (H^+) resin, 1% aq NH_3; g) Ac_2O, pyridine.

by treatment of (1*SR*,2*SR*,3*RS*,4*SR*,6*RS*)-1,2,3-triacetoxy-4-bromo-6-(bromomethyl)cyclohexane[8] with sodium acetate in HMPA at 120 °C.

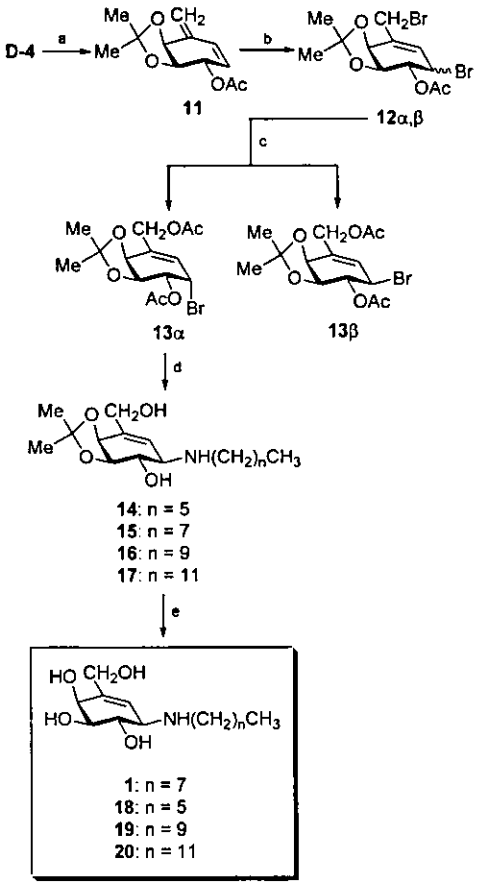
Treatment of **4** with a slight excess of bromine in carbon tetrachloride gave a 90% yield of approximately 1:1 mixture of the 1,4-addition products. The mixture was separable by silica gel chromatography to afford the dibromides[8] **5α** and **5β**. Compound **5α** was treated with sodium acetate to give a mixture of 2,3,4,6-tetra-*O*-acetyl-5a-carba-DL-*arabino*-hex-5(5a)-enopyranosyl bromides (**6α,β**). These results suggest that, although the allylic secondary bromo group remains unchanged, epimerization at C-1 occurs through nucleophilic attack by bromide ions generated in situ. Therefore, the mixture of **5α,β** was directly converted into the bromides **6α,β**, which was without purification treated with sodium azide in DMF at room temperature to give an inseparable 1:1 mixture of the azides[8] **7α,β** in 85% yield. In this case, neighboring group participation by the 2-acetoxyl group at C-1 was first anticipated to give rise to the β-azide **7β** selectively through formation of an intermediate 1,2-acetoxonium ion. However, the allylic carbon atom seems sufficiently active to suffer rear-side attack by an azide anion. Therefore, the mixture of **6α,β** was treated with sodium azide for conversion into a mixture of **7α,β** (85%), which was subsequently subjected to reduction with triphenylphosphine to generate free bases. The free bases were then converted into the *N*-acetyl and penta-*N*,*O*-acetyl derivatives in the usual manner. However, none of the above anomeric pairs could be separated by conventional silica gel chromatography.

O-Deacetylation of **7α,β** under Zemplén conditions gave the tetrols **9α,β**, which were treated with 2,2-dimethoxypropane-TsOH in DMF to afford, after acetylation, a mixture of the 2,3-*O*-isopropylidene derivatives **8α,β** selectively. These compounds were found to be separable on a silica gel column with 1:6 EtOAc/hexane as an eluent, giving **8α** (40%) and **8β** (42%), the ^1H NMR spectra of which showed a doublet of doublets (δ 4.25, $J = 3.8$ and 4.1 Hz) and a broad doublet (δ 3.97, $J = 8.9$ Hz) due to the pseudo-equatorial and axial protons on carbon atoms attached to the azido functions, respectively, supporting the proposed structures. Removal of the acetyl and isopropylidene groups of **8α** and **8β** was effected by treatment with 4 M hydrochloric acid at reflux temperature to give the respective azides **9α** (80%) and **9β** (74%). Reduction of the azido group of **9α** and **9β** with triphenylphosphine in 70% aqueous THF afforded, after purification over a column of Dowex 50 W×2 (H^+) resin with 5% aqueous NH_3 as eluent, the free amines[9] **3α** and **3β** in 86 and 86% yields, respectively. Their structures were further verified with ^{13}C and ^1H NMR spectra of the respective penta-*N*,*O*-acetyl derivatives[10] **10α** and **10β** obtained by conventional acetylation.

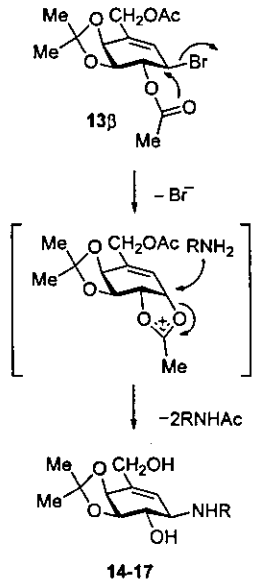
Incorporation of an alkylamino function at C-1 of the allyl bromides **6α,β** was attempted by treatment with an excess of *n*-octylamine. However, a complex mixture of products was formed. Then, the two *cis*-hydroxyl

groups of 6α,β were protected, with the aim of restricting conformational flexibility and avoiding a possible participation of the 4-hydroxyl (Scheme 2). Thus, the optically pure diene[8,11] **D-4** was first converted into the 3,4-*O*-isopropylidene derivative **11**, which was subsequently treated with bromine to afford the 1,4-addition products, ca. 1.7:1 mixture (76%) of **12α** and **12β**. Treatment of **12α,β** with sodium acetate gave selectively a mixture of the bromides **13α,β**, which was separable by silica gel chromatography, affording **13α** (48%) and **13β** (23%). The ^{1}H NMR signals due to C-1 appeared as a doublet of doublets (δ 4.83, J = 3.9 and 7.7 Hz) and a broad doublet (δ 4.50, J 8.3 Hz), respectively. Reaction of the α-bromide **13α** with 4 molar equiv of octylamine proceeded smoothly in an S_N2 fashion to give the protected *N*-octyl derivative as a single β-amine **15** (68%), which was then treated with aqueous acetic acid and subsequently purified on a column of Dowex 50W×2 (H^+) resin with aqueous ammonia, affording **1**[3] (56%). Similar treatment with hexyl-, decyl- and dodecyl-amines produced the corresponding *N*-alkyl derivatives[12] **14**, **16**, and **17**, which were deprotected to provide the β-amines **18**, **19**, and **20** in 30–65% total yields.

It is interesting to note that, on treatment with alkyl amines, the β-bromide **13β** also gave the β-amines as sole products, conceivably through neighboring group participation with the 2-acetoxyl or hydroxyl group (Scheme 3). Therefore, synthesis of **1** and its analogues would be much improved by use of an intact mixture of **13α,β** as the starting material.



Scheme 2. Reagents and conditions: (a) 1 M NaOMe-MeOH, 1 h, rt; $(MeO)_2Me_2$, DMF, TsOH hydrate, 1 day, rt; Ac_2O, pyridine; (b) Br_2 (molar equiv), CCl_4; rt; (c) NaOAc (1.4 molar equiv), 2 days, rt; (d) for example; decylamine (6 molar equiv), 2-propanol, 3 days, rt; 80% aq AcOH, 8 h, 80 °C, 8 h; Dowex 50 W×2 (H^+) resin, 1% NH_3-MeOH.



Scheme 3.

3. Biological assay

Results of biological assays[13] for inhibitory activity toward several glycohydrolases are listed in Table 1. None of the compounds showed any inhibitory activity against α-fucosidase (bovine kidney), α-glucosidase (Baker's yeast), or α-mannosidase (Jack beans). Although, for synthetic reasons,[3] only the *N*-octyl derivative **1** of **3β** has so far received attention as a β-galactosidase inhibitor, the present work provides the first description of inhibitory activity against glycohydrolases obtained by a series of N-substituted derivatives of **3β**. As expected,[4,6] *N*-alkylation dramatically improved the inhibitory activity against α- and β-galactosidases. It is worthy of note that the β-galactose-type valienamines **1** and **18**–**20** have both been shown to be very strong inhibitors of β-galactosidase and β-glucosidase, with no specificity regarding the 4-epimeric structures of the substrates. This characteristic is in good accordance with the cases of isofagomine[14] and calystegins.[15] Very recently, compound **1** has extensively been studied[16] as an important candidate for generation of novel therapeutic agents for treatment of GM_1-gangliosidosis and β-galactosidosis. Development of such enzyme-inhibitors might advantageously be accelerated by provision of various N-substituted derivatives, including **19** and **20**, readily prepared by use of the versatile precursors **13α,β**.

The present work describes a convenient synthetic route for the β-galactose-type valienamine **3β** and some *N*-alkyl derivatives thereof, demonstrating that enzyme-inhibitory activity can be significantly increased by a suitable N-substitution.

Table 1. Inhibitory activity (IC_{50}, μM) of compounds **1**, **3α,β**, and **18–20** against four glycosidases[a]

Compd	IC_{50} (μM)			
	α-Galactosidase (Green coffee beans)	β-Galactosidase (Bovine liver)	β-Glucosidase (Almonds)	α-Mannosidase (Jack beans)
1	3.1	0.87	3.1	NI
3α	56	NI	NI	370
3β	12	NI	NI	190
18	2.7	2.3	1.2	NI
19	1.9	0.13	2.5	NI
20	4.4	0.01	0.87	NI
DMJ	NT	NT	NT	150

[a] Compounds **3α** and **3β** are racemic; DMJ: deoxymannonojirimycin; NI: $IC_{50} > 0.1$ mg/mL; NT: Not tested.

4. Experimental

4.1. General methods

Optical rotations were measured with a JASCO DIP-370 polarimeter, and $[\alpha]_D$ values are given in 10^{-1} deg cm^2 g^{-1}. ^1H NMR spectra were recorded for solutions in deuteriochloroform and deuteriomethanol with internal tetramethylsilane (TMS) as a reference with a JEOL JNM LAMDA-300 (300 MHz) instrument. ^{13}C NMR spectra were recorded with the same instrument (75 MHz). IR spectra were recorded with a JASCO IR-810 or HITACHI Bio-Rad Digital Lab FTS-65 spectrometer. Mass spectra were determined with HITACHI M-8000 ion trap mass spectrometer using electrospray ionization (ESI). TLC was performed on silica gel 60 F-254 (E. Merck, Darmstadt). The silica gel used for a column chromatography was Wakogel C-300 (Wako Junyaku Kogyo Co., Osaka, 200–300 mesh) or silica gel 60 KO (Katayama Kagaku Kogyo Co., Osaka, 70–230 mesh). Organic solutions were dried over anhydrous Na_2SO_4 and concentrated at >45 °C under diminished pressure.

4.1.1. (1*SR*,2*RS*,3*SR*)-1,2,3-Triacetoxy-4-methylenecyclohex-5-ene (4). A mixture of (1*RS*,2*RS*,3*RS*,4*RS*,6*RS*)-1,2,3-triacetoxy-4-bromo-6-(bromomethyl)cyclohexane[8] (1.64 g, 3.81 mmol) and anhydrous sodium acetate (1.25 g, 15.2 mmol) in HMPA (65 mL) was stirred for 2 h at 120 °C. After cooling, the mixture was diluted with ethyl acetate (210 mL), and the solution was washed thoroughly with water, dried, and evaporated. The residue was chromatographed on a silica gel column (90 g, 1:8 acetone/hexane) to give the conjugated diene **4** (0.73 g, 72%) as a syrup, TLC: R_f 0.45 (1:3 acetone/hexane); ^1H NMR (300 MHz, $CDCl_3$) δ 6.25 (br d, 1H, J = 10.0 Hz, H-6), 5.79 (d, 1H, J = 2.8 Hz, H-3), 5.70 (br d, 1H, J = 10.0 Hz, H-5), 5.64 (br d, 1H, J = 7.6 Hz, H-1), 5.33 and 5.28 (2 s, each 1H, CH_2), 5.17 (dd, 1H, J = 2.8 and 7.6 Hz, H-2), 2.10 and 2.10 (3 s, each 3H, 3 Ac). This compound was identified with an authentic sample[8] on comparison with spectral data.

4.1.2. (1*RS*,2*RS*,3*SR*,6*SR*)- and (1*RS*,2*RS*,3*SR*,6*RS*)-1,2,3-Triacetoxy-6-bromo-4-(bromomethyl)cyclohex-4-ene [2,3,4-tri-*O*-acetyl-6-bromo-6-deoxy-5a-carba-α and β-DL-*arabino*-hex-5(5a)-enopyranosyl bromide] (5α and 5β). To a solution of the diene **4** (1.96 g, 7.34 mmol) in carbon tetrachloride (40 mL) was added dropwise bromine (1.2 g, 7.6 mmol) for 1 h at room temperature. The mixture was then diluted with chloroform (400 mL) and the solution was washed thoroughly with saturated aqueous sodium thiosulfate, aqueous sodium hydrogen carbonate, and water, dried, and evaporated. The residue was chromatographed on a silica gel column (280 g, 6:1 ethyl acetate/hexane) to give about 1:1 mixture of the dibromides **5α** and **5β** (2.82 g, 90%) as colorless crystals.

The mixture of the products obtained from **4** (320 mg) was carefully fractionated by chromatography on silica gel (10:1 EtOAc/hexane) to give pure **5α** (120 mg) and **5β** (43 mg), together with **5α,β** (220 mg): ^1H NMR (300 MHz, $CDCl_3$) **5α**: δ 6.22 (d, 1H, J = 4.6 Hz, H-5), 5.94 (d, 1H, J = 4.1 Hz, H-3), 5.48 (dd, 1H, J = 4.1 and 10.3 Hz, H-2), 5.10 (dd, 1H, J = 3.7 and 10.3 Hz, H-1), 5.08 (dd, 1H, J = 3.7 and 4.6 Hz, H-6), 3.93 (s, 2H, CH_2Br); **5β**: δ 6.18 (d, 1H, J = 2.7 Hz, H-5), 5.88 (d, 1H, J = 3.6 Hz, H-3), 5.69 (dd, 1H, J = 7.6 and 10.5 Hz, H-1), 5.10 (dd, 1H, J = 3.6 and 10.5 Hz, H-2), 4.61 (dd, 1H, J = 2.7 and 7.6 Hz, H-6), 3.91 (s, 2H, CH_2Br).

These compounds were identified with authentic samples[8] on comparison with spectral data.

4.1.3. (1*RS*,2*RS*,3*SR*,6*SR*)- and (1*RS*,2*RS*,3*SR*,6*RS*)-1,2,3-Triacetoxy-4-(acetoxymethyl)-6-bromocyclohex-4-ene [2,3,4,6-tetra-*O*-acetyl-5a-carba-α and β-DL-*arabino*-hex-5(5a)-enopyranosyl bromide] (6α and 6β). A ca. 1:1 mixture (912 mg, 2.13 mmol) of the dibromides **5α** and **5β**, anhydrous sodium acetate (175 mg, 2.13 mmol) in DMF (14 mL) was stirred for 20 h at room temperature. The mixture was then diluted with ethyl acetate (180 mL), and the solution was washed with saline (3×60 mL), dried, and evaporated. The residue was chromatographed on a silica gel column (80 g, 1:3 ethyl acetate/hexane) to give a 1:1 inseparable mixture (743 mg, 86%) of the bromides **6α** and **6β** as colorless crystals, TLC: R_f 0.35 (1:2 EtOAc/hexane). This mixture of the compounds was identified with an authentic sample[8] on comparison with spectral data.

Compound **5α** (45 mg) was similarly treated with sodium acetate (8.6 mg) in DMF (1 mL) to give products, which were shown to be a mixture of **6α,β** by ^1H NMR spectrum.