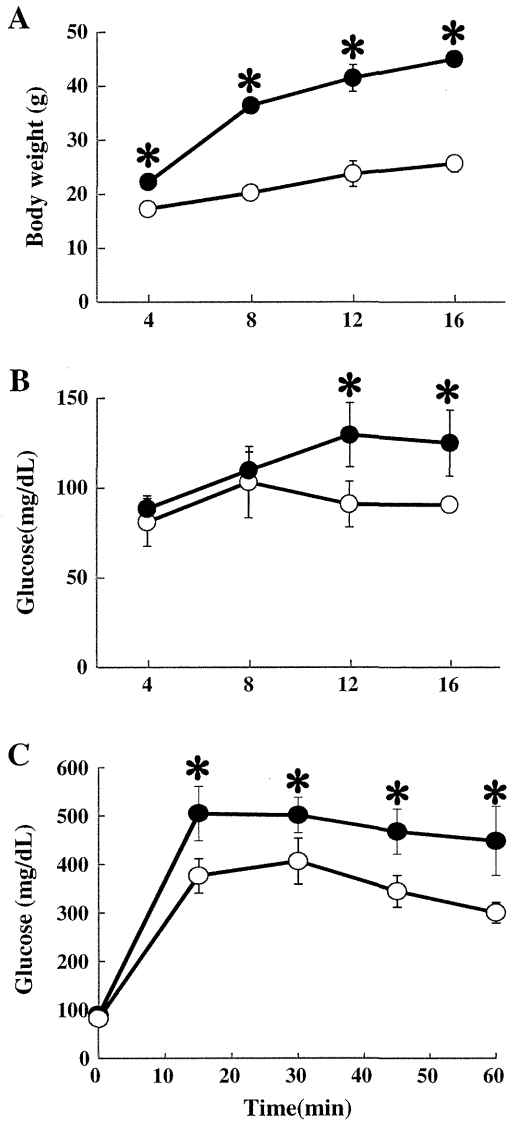

Fig. 1 – Characteristics of KK-A^y mice. Time course changes of body weight (A) and fasting blood glucose (B) in 4–16 weeks old KK-A^y (closed circles) and C57BL/6 mice (open circles). (C) Oral glucose tolerance test in 4-week-old KK-A^y and C57BL/6 mice. Plasma glucose levels at 15, 30, 45 and 60 min after oral glucose administration were measured. C57BL/6 mice are served as control. Each value represents mean ± SD (n = 4 for each group). *, $p < 0.05$ compared with the values of control.

compared to those of age-matched C57BL/6 mice (Fig. 1A) and the glucose concentrations were significantly elevated in 12- and 16-week-old KK-A^y mice compared with those of age-matched C57BL/6 mice (Fig. 1B). OGTT was analyzed at 4 weeks of age (Fig. 1C). The glucose concentrations during the course of glucose loading were significantly higher in 4-week-old KK-A^y mice than in C57BL/6 mice. This result indicates that glucose intolerance was initiated in 4-week-old KK-A^y mice although the fasting glucose concentrations did not show any differences between the 4-week-old KK-A^y mice and C57BL/6 mice. To obtain insights into the molecular differences from the early onset of T2DM, we investigated differential serum proteome analysis between 4-week-old male KK-A^y as T2DM and age-matched male C57BL/6 mice as wild-type mice. In the 2-D LC–MS/MS analysis, a total of 227 unique proteins were identified in the iTRAQ experiments at both the significance threshold $p < 0.05$ and the 5% false discovery rate threshold level (Supplementary Table 1). For the selection of differentially expressed proteins, we selected proteins using the following criteria: (1) the protein quantitation values between the KK-A^y and C57BL/6 groups were significantly different with a p value <0.05. (2) Proteins detected only in one group were also included as differentially expressed proteins. A total of 45 proteins were selected as differentially expressed proteins. Twenty five proteins among the 45 proteins were found to be up-regulated (Tables 1 and 2), and the other 20 proteins were down-regulated (Tables 3 and 4).

3.2. *Interaction networks and functional pathway analysis*

To gain insights into the biological changes in the KK-A^y versus C57BL/6 mice, the differentially expressed proteins were categorized according to the Gene Ontology (GO) classes "cellular component" and "molecular function". In the cellular component of GO analysis, differentially expressed proteins were located in the extracellular space (64%), membrane attack complex (10%), high-density lipoprotein particle (13%), and plasma lipoprotein particle (13%) (Fig. 2A). According to the molecular functions, most of the differentially expressed proteins were associated with peptidase inhibitor activity (50%), serine hydrolase activity (31%), and lipoprotein binding (19%) (Fig. 2B).

In the IPA analysis, 45 differentially expressed proteins were eligible for network analysis based on the IPA Knowledge Base criteria. IPA generated 3 interaction networks (Supplementary Table 2). Top-rated networks are related to: metabolic disease, endocrine system disorders, and energy production (Fig. 3), and lipid metabolism, molecular transport, and small molecule biochemistry (data not shown). Each network consists of 35 proteins, of which 13 were included in our list, and had a score of 23. Fig. 3 shows that 7 of 13 differentially expressed proteins are modulated directly by peroxisome proliferator activated receptor α, Ppara (p-value of overlap = 3.75×10^{-9}, Supplementary Table 3). Furthermore, 3 of 14 differentially expressed proteins are shown to be modulated directly by nuclear factor, erythroid derived 2, like 2, Nfe2l2 (p-value of overlap = 3.73×10^{-2}, Supplementary Table 3), which also interacted with Ppara.

3.3. *Validation of the selected proteins by MRM analysis*

In the iTRAQ experiments, we analyzed 3 biological replicates per group. This small number of samples could introduce many false positive findings. For the validation of the iTRAQ results, we carried out relative quantification of the identified proteins by MRM analysis using an independent sample set (the sera obtained from 4-week-old KK-A^y and C57BL/6 mice (n = 5/each group)) from that used in the iTRAQ experiments. In this work, proteotyptic peptides were chosen based on iTRAQ results and shotgun proteomic identification data using MRMPilot software. For MRM, 34 of

Table 1 – Increased serum proteins in KK-A^y mice compared to C57BL/6 mice.

Accession number	Entry name	Protein name	Relative protein ratio [a]		KK-A^y: C57BL/6	*p*-Value
			KK-A^y	C57BL/6		
Q61147	CERU_MOUSE	Ceruloplasmin	1.29 ± 0.07	1.05 ± 0.04	1.23	1.2E−02
P42703	LIFR_MOUSE	Leukemia inhibitory factor receptor	1.05 ± 0.03	0.83 ± 0.09	1.27	3.3E−02
P11589	MUP2_MOUSE	Major urinary protein 2	1.02 ± 0.09	0.8 ± 0.03	1.27	2.9E−02
Q06890	CLUS_MOUSE	Clusterin	1.19 ± 0.06	0.93 ± 0.06	1.28	1.4E−02
P08226	APOE_MOUSE	Apolipoprotein E	1.25 ± 0.16	0.9 ± 0.07	1.38	4.9E−02
Q8K182	CO8A_MOUSE	Complement component C8 alpha chain	1.18 ± 0.03	0.74 ± 0.08	1.59	1.7E−03
P19221	THRB_MOUSE	Prothrombin	1.36 ± 0.05	0.85 ± 0.11	1.60	3.8E−03
P06683	CO9_MOUSE	Complement component C9	1.38 ± 0.12	0.81 ± 0.09	1.69	5.5E−03
P20918	PLMN_MOUSE	Plasminogen	1.03 ± 0.05	0.61 ± 0.04	1.70	9.9E−04
P02762	MUP6_MOUSE	Major urinary protein 6	1.73 ± 0.24	0.74 ± 0.12	2.34	6.6E−03
P07759	SPA3K_MOUSE	Serine protease inhibitor A3K	1.53 ± 0.12	0.60 ± 0.05	2.56	6.0E−04
P09813	APOA2_MOUSE	Apolipoprotein A-II	1.83 ± 0.45	0.64 ± 0.17	2.85	2.6E−02
P04186	CFAB_MOUSE	Complement factor B	1.58 ± 0.09	0.52 ± 0.09	3.05	3.1E−04

[a] The relative protein ratios to an internal standard identified by iTRAQ are used and expressed as mean ± SD.

45 differentially expressed proteins were selected by the criteria described in the Materials and methods. Forty eight peptides representing 34 proteins were selected for MRM analysis. Finally, a total of 102 transitions were used for targeting 34 peptides of 34 proteins. The complete transition list is shown in Supplementary Table 4. For each of the 34 target proteins, three transitions were monitored for each peptide. Supplementary Fig. 2 displays a typical XIC (extracted ion chromatogram) overlay of the predetermined different transitions for the 34 proteins. The XIC peaks of the different transitions for each target protein were detected at the same retention time. Each resulting MRM peak was also examined by full scan MS/MS. The MS/MS spectrum for each MRM peak confirmed the sequence validation of the hypothesized peptide as shown in Supplementary Fig. 3. Individual serum samples were digested and analyzed in duplicate. Among them, 17 proteins showed significant differences between 4-week-old KK-A^y and C57BL/6 mice, which validated the iTRAQ results (Table 5). Retinol-binding protein 4 (RBP4) was found to be increased in 4-week-old KK-A^y mice compared with C57BL/6 mice in the MRM analysis although the expression of RBP4 was decreased in KK-A^y mice in the iTRAQ result. The inconsistency has arisen from the selection of the peptides for the quantitation in the iTRAQ experiments. Relative protein quantitation of RBP4 in C57BL/6 mice was calculated using 4 peptides in the iTRAQ experiments; however, only 2 of the 4 peptides were used for the quantitation in KK-A^y mice samples because other 2 peptides were miscleaved in KK-A^y mice samples. Taken together, we adopted the data obtained from the MRM analysis for the RBP4 protein quantitation. Furthermore, for further analyses of the candidate proteins, the levels of the 34 proteins in the sera of 12-week-old KK-A^y and C57BL/6 mice were also analyzed by MRM analysis because the glucose concentrations in KK-A^y mice were significantly elevated from the age of 12 weeks compared to the C57BL/6 group (Fig. 1B). Eight of the 34 proteins still showed significant differences between 12-week-old KK-A^y and C57BL/6 mice. Among the 8 proteins mentioned above, proteins with a significant increase included 7 proteins: apolipoprotein (apo) A-II, carboxypeptidase N catalytic chain, clusterin, inter-alpha-trypsin inhibitor heavy chain H3, RBP4, serine protease inhibitor A3K (SERPINA3K), and serum amyloid P-component; and another protein with a significant decrease was apo A–I (Table 5).

Table 2 – Identified and quantitated serum proteins only in KK-A^y mice.

Accession number	Entry name	Protein name	Relative protein ratio [a]
Q8BH35	CO8B_MOUSE	Complement component C8 beta chain	1.54 ± 0.10
Q8R121	ZPI_MOUSE	Protein Z-dependent protease inhibitor	1.21 ± 0.07
Q9ET66	PI16_MOUSE	Peptidase inhibitor 16	1.41 ± 0.15
Q9WVJ3	CBPQ_MOUSE	Carboxypeptidase Q	1.05 ± 0.04
Q9JJN5	CBPN_MOUSE	Carboxypeptidase N catalytic chain	1.01 ± 0.16
P31532	SAA4_MOUSE	Serum amyloid A-4 protein	1.43 ± 0.18
P00920	CAH2_MOUSE	Carbonic anhydrase 2	1.33 ± 0.15
P02535	K1C10_MOUSE	Keratin, type I cytoskeletal 10	1.72 ± 0.86
P12246	SAMP_MOUSE	Serum amyloid P-component	1.34 ± 0.20
P52480	KPYM_MOUSE	Pyruvate kinase isozymes M1/M2	1.14 ± 0.16
Q5FW60	MUP20_MOUSE	Major urinary protein 20	3.00 ± 0.75
Q61704	ITIH3_MOUSE	Inter-alpha-trypsin inhibitor heavy chain H3	1.14 ± 0.11

[a] The relative protein ratios to an internal standard identified by iTRAQ are used and expressed as mean ± SD.

3.4. *SERPINA3-induced changes in endothelial cell monolayer permeability*

Among the 8 differentially expressed proteins, SERPINA3K was chosen for further analyses. SERPINA3K was selected on the basis of the possible implications for diabetic retinopathy

Table 3 – Decreased serum proteins in KK-A^y mice compared to C57BL/6 mice.

Accession number	Entry name	Protein name	Relative protein ratio [a]		KK-A^y: C57BL/6	*p*-Value
			KK-A^y	C57BL/6		
Q00724	RET4_MOUSE	Retinol-binding protein 4	0.71 ± 0.06	1.40 ± 0.25	0.51	2.0E−02
P11859	ANGT_MOUSE	Angiotensinogen	0.73 ± 0.10	1.22 ± 0.19	0.60	3.1E−02
P23953	EST1C_MOUSE	Carboxylesterase 1C	0.62 ± 0.03	1.04 ± 0.06	0.60	9.6E−04
Q60994	ADIPO_MOUSE	Adiponectin	0.68 ± 0.08	1.09 ± 0.13	0.62	1.8E−02
P32261	ANT3_MOUSE	Antithrombin-III	0.75 ± 0.07	1.19 ± 0.05	0.63	1.5E−03
O08677	KNG1_MOUSE	Kininogen-1	0.61 ± 0.04	0.90 ± 0.10	0.67	1.7E−02
Q00623	APOA1_MOUSE	Apolipoprotein A-I	0.98 ± 0.01	1.42 ± 0.06	0.69	4.3E−04
P21614	VTDB_MOUSE	Vitamin D-binding protein	0.84 ± 0.03	1.18 ± 0.01	0.71	7.6E−05
Q8BND5	QSOX1_MOUSE	Sulfhydryl oxidase 1	0.80 ± 0.04	1.08 ± 0.04	0.74	2.9E−03
O89020	AFAM_MOUSE	Afamin	0.88 ± 0.05	1.18 ± 0.02	0.74	1.5E−03
Q61730	IL1AP_MOUSE	Interleukin-1 receptor accessory protein	0.63 ± 0.02	0.84 ± 0.06	0.74	9.8E−03
Q06770	CBG_MOUSE	Corticosteroid-binding globulin	1.38 ± 0.13	1.81 ± 0.12	0.76	2.7E−02
P26262	KLKB1_MOUSE	Plasma kallikrein	0.76 ± 0.07	0.98 ± 0.07	0.78	3.6E−02
Q91X72	HEMO_MOUSE	Hemopexin	0.79 ± 0.10	1.02 ± 0.06	0.78	4.9E−02
Q61129	CFAI_MOUSE	Complement factor I	0.82 ± 0.07	0.98 ± 0.04	0.84	4.4E−02
Q01339	APOH_MOUSE	Beta-2-glycoprotein 1	0.78 ± 0.04	0.91 ± 0.01	0.86	1.5E−02

[a] The relative protein ratios to an internal standard identified by iTRAQ are used and expressed as mean ± SD.

in rat models [19,20]. However, it remains to be elucidated whether human homologue, SERPINA3 is associated with the development of T2DM and/or diabetic retinopathy. Blood-retinal barrier breakdown contributes to macular edema, which occurs in over 25% of people with diabetes and correlated highly with visual impairment in people with diabetic retinopathy [21]. Therefore, we focused on the effects of SERPINA3 on endothelial cell monolayer permeability. To this end, we sought to investigate the impact of SERPINA3 on HRMVECs. Post-confluent serum-starved HRMVECs incubated with SERPINA3 exhibited a dose-dependent increase in endothelial permeability reflected as a decrease in transendothelial electrical resistance, whereas electrical impedance for the untreated HRMVECs was almost unchanged. Fig. 4A is a representative ECIS graph demonstrating the real-time loss of electrical resistance across the HRMVEC monolayer. Permeability increased steadily after SERPINA3 was added, and continued to decline up to 8 h after treatment. Fig. 4B–D is the typical real-time traces of the Rb, α, and Cm in the SERPINA3-treated HRMVECs. After HRMVECs were treated with SERPINA3 (time = 0 in Fig. 4B–D), the cell-to-cell resistances gradually decreased, and the cell-to-extracellular matrix interaction slightly decreased, whereas membrane capacitance was almost unchanged. These results indicate that SERPINA3 treatment mainly caused the decrease of HRMVEC cell-to-cell resistances. Consistent with the ECIS data, we observed that SERPINA3 treatment caused a massive disruption of HRMVEC monolayers (data not shown).

Table 4 – Identified and quantitated serum proteins only in C57BL/6 mice.

Accession number	Entry name	Protein name	Relative protein ratio [a]
P09581	CSF1R_MOUSE	Macrophage colony-stimulating factor 1 receptor	1.23 ± 0.07
Q62351	TFR1_MOUSE	Transferrin receptor protein 1	1.34 ± 0.36
Q8CIF4	BTD_MOUSE	Biotinidase	1.10 ± 0.13
P04939	MUP3_MOUSE	Major urinary protein 3	1.99 ± 0.49

[a] The relative protein ratios to an internal standard identified by iTRAQ are used and expressed as mean ± SD.

3.5. *SERPINA3-induced changes in endothelial cell proliferation and migration*

Proliferation and migratory capability of vascular endothelial cells have a pivotal role in the maintenance of microvascular integrity and angiogenesis. Proliferative capacity of HRMVECs treated with or without SERPINA3 was compared. We found that treatment with SERPINA3 barely affected the proliferative ability in HRMVECs (Fig. 5A).

To assess the effects of SERPINA3 on migratory capability of HRMVECs, an ECIS-based wound-healing assay was performed. We conducted a wounding experiment on single-point ECIS arrays with a single 250 μm-diameter working electrode from which the interrogating electrical current originates, then passes through the monolayer and finally reaches the surrounding counterelectrode. Again, monolayers were found to mature when measured electrically. Post-confluent serum-starved HRMVECs were treated with various doses of SERPINA3 for 6 h, and then monolayers were electrically removed from the central working electrode by driving 1400 μA at 60 kHz through the circuit for 20 s. When assessed 18 h after the SERPINA3 treatment, little differences were observed between the four conditions (Fig. 5B). After completion of the wound closure process, monolayer resistance was found to return to values similar to prewounding levels.

4. Discussion and conclusions

In this paper, we investigated the serum proteome of KK-A^y mice in a pre-diabetic state compared to wild C57BL/6 mice in an attempt to uncover early diagnostic markers of diabetes

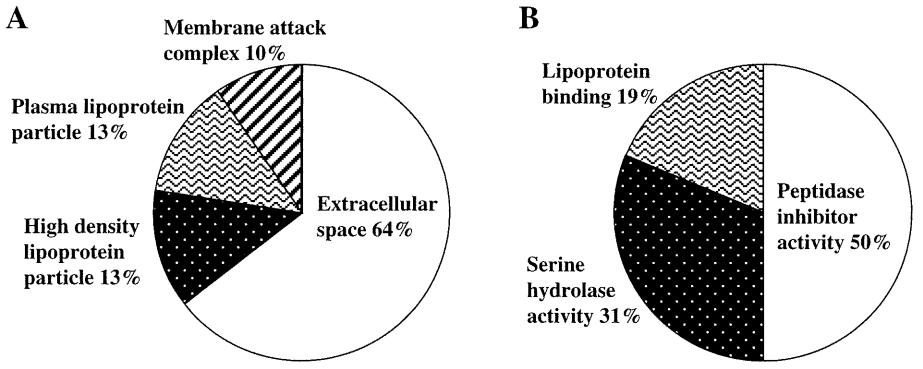

Fig. 2 – Gene ontology terms for cellular components (A), and molecular functions (B) of the differentially expressed proteins between the sera of KK-A^y and C57BL/6 mice.

that are maintained through a diabetic phenotype. We used iTRAQ-based two-dimensional LC–MS/MS serum profiling, and identified several differentially proteins at the pre-diabetic stage. The differential expression was confirmed by MRM analysis, which is fast gaining ground as a sensitive, specific, and cost-effective methodology for relative quantification of candidate proteins. Using these techniques, we have identified eight candidate proteins of interest including SERPINA3K, which may be important in the pathology of T2DM and/or diabetic retinopathy.

Several published studies have reported proteomic alterations in diabetes. During the preparation of this manuscript, Ressom et al. reported a similar approach for a large scale proteomic study aimed at identifying and validating changes

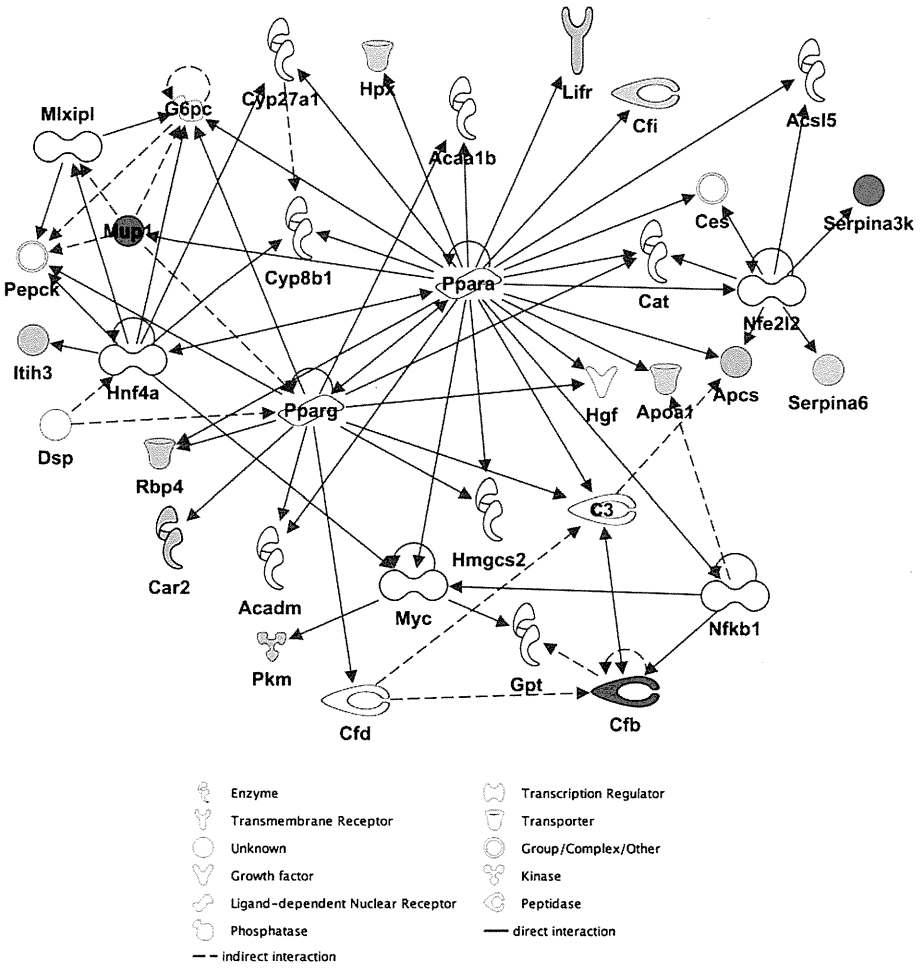

Fig. 3 – Interaction network analysis based on the IPA Knowledge Base criteria. The top-rated network is shown. This network is related to metabolic disease, endocrine system disorders, and energy production. Colored molecules: green; decreased in KK-A^y mice sera, and red; increased in KK-A^y mice sera.

Table 5 – Validation of differentially expressed proteins by MRM assay. Differentially expressed proteins in the iTRAQ experiments were analyzed by MRM assay. The values represent the normalized area under the most intense peak of the target peptides, and are represented as mean ± SD. Duplicate analyses were performed for each of the mouse serum samples.

Accession number	Protein name	Peptide sequence	4-week-old				12-week-old			
			C57BL/6 (B6)	KK-A^y (A^y)	A^y/B6	*p*-Value	C57BL/6 (B6)	KK-A^y (A^y)	A^y/B6	*p*-Value
Q60994	Adiponectin	AVLFTYDQYQEK	1.62 ± 0.71	1.19 ± 0.40	0.74	0.28	1.55 ± 0.69	0.37 ± 0.08	0.24	0.01*
O89020	Afamin	AAPITQYLK	26.36 ± 12.46	22.51 ± 7.73	0.85	0.57	8.41 ± 2.1	9.67 ± 3.20	1.15	0.48
P11859	Angiotensinogen	TLHDQLVLAAEK	2.30 ± 1.01	1.78 ± 0.16	0.78	0.29	0.64 ± 0.27	0.85 ± 0.23	1.32	0.24
P32261	Antithrombin-III	TEDGFSLK	1.38 ± 0.54	0.18 ± 0.25	0.13	2.E−03*	0.62 ± 0.32	0.66 ± 0.32	1.08	0.81
Q00623	Apolipoprotein A-I	TQVQSVIDK	315.84 ± 67.55	0.10 ± 0.11	3.E−04	6.E−06*	84.87 ± 56.38	7.17 ± 5.33	0.08	0.02*
P09813	Apolipoprotein A-II	THEQLTPLVR	60.59 ± 24.63	171.35 ± 63.93	2.83	0.01*	8.07 ± 4.36	60.09 ± 26.56	7.45	3.E−03*
P08226	Apolipoprotein E	LQAEIFQAR	28.41 ± 13.60	37.81 ± 14.93	1.33	0.33	3.63 ± 1.1	15.66 ± 4.37	4.32	3.E−04*
Q01339	Beta-2-glycoprotein 1	YTSFEYPK	36.76 ± 12.77	24.30 ± 18.78	0.66	0.25	17.04 ± 4.23	23.48 ± 4.21	1.38	0.04*
Q8CIF4	Biotinidase	GLSSGLVTAALYGR	0.24 ± 0.06	0.23 ± 0.09	0.96	0.85	0.11 ± 0.08	0.11 ± 0.03	0.97	0.92
P23953	Carboxylesterase 1C	EGASEEETNLSK	10.65 ± 6.01	6.62 ± 5.93	0.62	0.32	1.09 ± 0.59	3.90 ± 4.75	3.57	0.23
Q9JJN5	Carboxypeptidase N catalytic chain	AVIQWIR	0.79 ± 0.27	1.20 ± 0.26	1.52	0.04*	0.39 ± 0.28	0.74 ± 0.16	1.92	0.04*
Q61147	Ceruloplasmin	AGLQAFFQVR	5.46 ± 4.68	12.49 ± 2.27	2.29	0.02*	1.19 ± 1.1	2.23 ± 1.36	1.87	0.22
Q06890	Clusterin	ASGIIDTLFQDR	2.48 ± 0.64	4.28 ± 0.79	1.57	0.02*	1.26 ± 0.38	2.58 ± 0.12	2.24	4.E−05*
Q8K182	Complement component C8 alpha chain	TEGTTVDEVQK	0.09 ± 0.04	0.05 ± 0.05	0.61	0.25	0.02 ± 0.02	0.06 ± 0.06	2.37	0.27
Q8BH35	Complement component C8 beta chain	ALEEFQSEVSSC[CAM]R	0.21 ± 0.10	0.99 ± 0.64	4.78	0.03*	0.08 ± 0.02	0.42 ± 0.65	5.32	0.28
P06683	Complement component C9	AVEDYIDEFSTK	0.13 ± 0.05	1.43 ± 0.67	11.07	3.E−03*	0.13 ± 0.15	0.19 ± 0.07	1.42	0.48
Q06770	Corticosteroid-binding globulin	AGEQINNHVK	0.08 ± 0.04	0.07 ± 0.05	0.91	0.82	0.10 ± 0.07	0.08 ± 0.05	0.80	0.59
Q91X72	Hemopexin	LYVSSGR	124.18 ± 37.50	98.01 ± 27.86	0.79	0.25	109.93 ± 37.67	126.32 ± 33.71	1.15	0.49
Q61704	Inter-alpha-trypsin inhibitor heavy chain H3	SLPEGVVDGIEVYSTK	0.21 ± 0.18	5.25 ± 2.72	25.18	3.E−03*	0.10 ± 0.05	0.30 ± 0.11	3.11	0.01*
O08677	Kininogen-1	QFNPGVK	11.81 ± 4.26	8.73 ± 3.31	0.74	0.24	8.84 ± 4.79	8.30 ± 3.43	0.94	0.84
P42703	Leukemia inhibitory factor receptor	ITGLVGPR	3.08 ± 0.87	8.78 ± 1.83	2.85	2.E−04*	2.7 ± 1.38	1.61 ± 0.37	0.60	0.12
P09581	Macrophage colony-stimulating factor 1 receptor	ESTSTGIWLK	1.02 ± 0.36	0.55 ± 0.25	0.54	0.04*	0.47 ± 0.24	0.35 ± 0.10	0.73	0.31
P04939	Major urinary protein 3	ENIIDLTNVNR	2.65 ± 0.65	2.89 ± 0.63	1.09	0.58	2.31 ± 1.1	1.52 ± 0.25	1.22	0.55
Q9ET66	Peptidase inhibitor 16	AESPEAEAESPLSSEALVPVLPAQER	0.03 ± 0.01	0.05 ± 0.01	2.10	4.E−03*	0.02 ± 0.01	0.01 ± 3E−03	0.56	0.07
P26262	Plasma kallikrein	SADNLVSGFSLK	1.09 ± 0.85	0.77 ± 0.47	0.71	0.49	0.09 ± 0.04	0.15 ± 0.08	1.66	0.15
P20918	Plasminogen	TGIGNGYR	0.16 ± 0.07	0.58 ± 0.34	3.69	0.03*	0.11 ± 0.06	0.18 ± 0.08	1.60	0.17
Q8R121	Protein Z-dependent protease inhibitor	LILVDYVLFK	0.17 ± 0.07	0.39 ± 0.18	2.33	0.04*	0.57 ± 0.28	0.54 ± 0.47	0.95	0.91
P19221	Prothrombin	DNLSPPLGQC[CAM]LTER	1.62 ± 1.41	2.21 ± 1.62	1.36	0.56	0.25 ± 0.14	1.04 ± 0.38	4.19	2.E−03*
P52480	Pyruvate kinase isozymes M1/M2	LDIDSAPITAR	0.11 ± 0.06	0.44 ± 0.31	4.19	0.04*	0.07 ± 0.04	0.06 ± 0.03	0.85	0.59
Q00724	Retinol-binding protein 4	LQNLDGTC[CAM]ADSYSFVFSR	0.39 ± 0.17	0.79 ± 0.20	2.01	0.01*	0.05 ± 0.03	0.11 ± 0.03	2.26	0.01*
P07759	Serine protease inhibitor A3K	TLFPSQIEELNLPK	0.39 ± 0.18	68.06 ± 28.55	176.70	7.E−04*	1.23 ± 2.06	9.28 ± 2.16	7.54	3.E−04*
P12246	Serum amyloid P-component	SQSLFSYSVK	0.13 ± 0.14	1.52 ± 0.44	11.59	2.E−04*	2E−03 ± 9E−04	0.21 ± 0.17	85.66	0.04*
Q8BND5	Sulfhydryl oxidase 1	SYVQFFFGC[CAM]R	0.30 ± 0.20	0.22 ± 0.14	0.73	0.48	0.07 ± 0.02	0.09 ± 0.03	1.19	0.42
P21614	Vitamin D-binding protein	TQVPEVFLSK	101.77 ± 35.17	88.03 ± 19.21	0.87	0.47	54.49 ± 21.66	70.01 ± 17.57	1.28	0.25

* $p < 0.05$ to age-matched C57BL/6 mice.

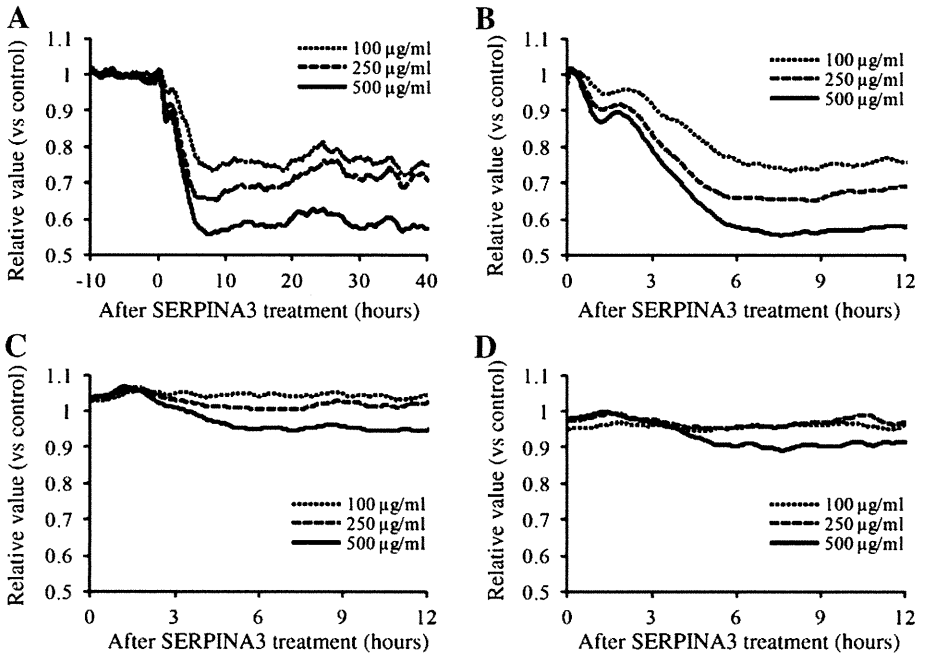

Fig. 4 – Monolayer barrier function of human retinal microvascular endothelial cells (HRMVECs) was notably decreased by SERPINA3 treatment. Post-confluent serum-starved HRMVECs were treated with SERPINA3 (100, 250, and 500 μg/ml) for 40 h. HRMVECs treated without SERPINA3 were used as control. (A) Impedance measurements at 16 kHz were obtained from HRMVECs treated with SERPINA3. (B–D) ECIS Zθ measurements were modeled to obtain the Rb (modeled barrier function, B), α (modeled cell-to-extracellular matrix interactions, C), and Cm (modeled membrane capacitance, D). Each trace of the impedance, Rb, α, and Cm is shown as an average of three replicate wells and representative of three independent experiments. Each value was normalized to the control.

in protein abundance among human serum obtained from a cohort of nondiabetic control subjects and patients diagnosed with T2DM [22]. The study reported here uses a similar approach nevertheless we analyzed the serum proteome analysis using T2DM mice model. However, it is novel in terms of quantitative profiling and comprehensive validation of the target proteins in individual samples. A limitation of this study is the use of T2DM mouse with obesity. We cannot delineate the differential markers of obesity and T2DM. Consequently, further validation analyses are required using human healthy and T2DM serum samples, which are now under investigation.

Among the eight serum proteins validated by MRM analysis, several published studies have also reported the differential expression of apo-AI, apo-AII, carboxypeptidase N catalytic chain, clusterin, inter-alpha-trypsin inhibitor heavy chain H3, SERPINA3K, and serum amyloid P-component in patients with diabetes and/or microvasculopathy [22–24]. Identification of RBP4 has not been reported in iTRAQ-based serum proteomic analysis, however, elevated levels of RBP4 have been reported in patients with obesity and metabolic syndrome [25–27]. The levels of RBP4 seem to be regulated in visceral adipose tissue. In the current study, SERPINA3K was chosen for further analyses because it remains to be elucidated whether human homologue, SERPINA3 is associated with the development of T2DM and/or diabetic retinopathy.

SERPINA3K was originally identified as a SERPIN family member with specific inhibitory effect on tissue kallikrein [28]. It is expressed in the liver, kidney, pancreas, and retina [29,30]. The previous studies have shown decreased retinal levels of SERPINA3K in a diabetic retinopathy rat model [19]. SERPINA3K attenuates inflammation in the retina with retinopathy [31] and has been found to inhibit ischemia-induced retinal neovascularization [20]. Recently, Zhang et al. have reported that rat SERPINA3K inhibited angiogenesis through Wnt/β-catenin signaling pathway as LRP6 antagonist [32]. In the present study, we found that SERPINA3K was consistently induced in KK-A^y mice sera compared with C57BL/6 mice. We also investigated the potential association between human SERPINA3 levels and T2DM. The serum SERPINA3 levels were approximately 180 μg/mL in healthy subjects, and the serum SERPINA3 levels were higher in T2DM and diabetic retinopathy patients compared to those in healthy control subjects (data not shown). In the IPA network analysis, SERPINA3 is shown to be modulated directly by Nfe2l2 (Fig. 2) [33]. Nfe2l2, also known as nuclear factor E2-related factor 2 (Nrf2) is a transcription factor that functions as a master regulator of the cellular adaptive response to oxidative stress [34]. Hyperglycemia is the metabolic hallmark of diabetes and leads to widespread cellular damage through oxidative stress. Recently, Slitt et al. reported that enhanced Nrf2 activation inhibited lipid accumulation in white adipose tissue, suppressed adipogenesis, induced insulin resistance and glucose intolerance, and increased hepatic steatosis in diabetic $Lep^{ob/ob}$ mice [35]. We analyzed the tissue distribution of SERPINA3 in KK-A^y mice and found that SERPINA3 is highly expressed in liver and also detected in pancreas, white adipose tissue, skeletal muscle, kidney, and heart (data not shown). We

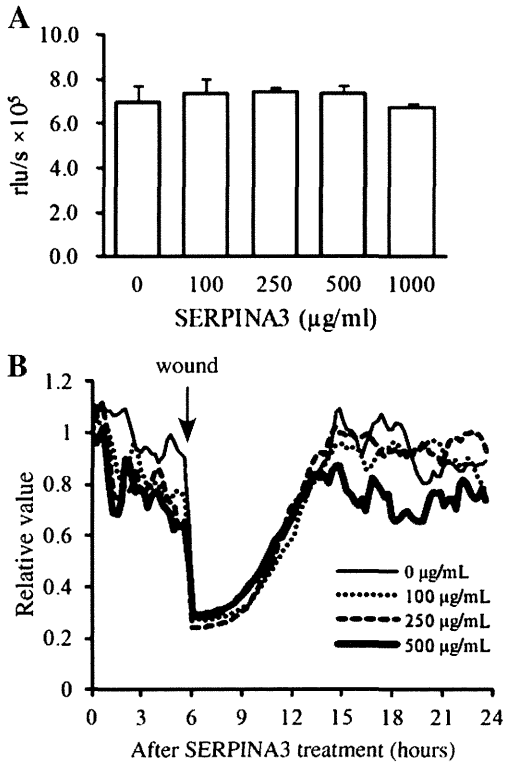

Fig. 5 – SERPINA3 did not significantly affect the proliferative ability or migratory capability of HRMVECs. (A) BrdU incorporation in HRMVECs. The cells were treated with or without SERPINA3 (100, 250, 500, 1000 μg/mL) and labeled with BrdU from 18 to 24 h after treatment. Data are expressed as mean ± SD values of triplicate samples. (B) Electrical wound-healing assay by ECIS Zθ. Post-confluent serum-starved HRMVECs were treated with SERPINA3 for 6 h, after which an elevated voltage pulse was applied to the electrode (marked with an arrow). Impedance values were normalized by dividing the values by those measured just prior to the addition of SERPINA3 (at time = 0).

speculate that enhanced activation of Nrf2 in liver may cause the increase in serum SERPINA3 from the early onset of T2DM although further studies are required to identify the precise mechanisms.

To gain insights into the function of SERPINA3, we assessed the effects of SERPINA3 on endothelial cell monolayer permeability. An ECIS assay was carried out by varying SERPINA3 concentrations ranging from 100 to 500 μg/mL, which cover the physiological concentration range in healthy subjects and T2DM patients. An *in vitro* assay revealed that SERPINA3 increased transendothelial permeability of HRMVECs (Fig. 5). Diabetic maculopathy, a leading cause of vision loss in patients with T2DM is characterized by hyperpermeability of retinal blood vessels with subsequent formation of macular edema and hard exudates. Epidemiological studies have suggested that glycemic control plays a major role in the development of vascular complications of diabetes [36]. SERPINA3 may become a downstream glycemic target that contributes to increased retinal vascular permeability that could be targeted therapeutically, and a blood biomarker of diabetic retinopathy although further investigations are required. In addition, our present study showed that SERPINA3 showed little proliferative ability and little effect of migratory capability of HRMVECs, which is inconsistent with the previous reports [20,32]. The discrepancy may have risen from interspecies differences. Rat SERPINA3K shares high degree of sequence identity with mouse SERPINA3K (70–71%) and human SERPINA3 (50–55%). However, mouse SERPINA3K and human SERPINA3 are known to be induced by acute inflammation in contrast to rat SERPINA3K which is a negative acute-phase protein [37]. Further studies are necessary to clarify the precise mechanism of action of SERPINA3 for cell proliferation.

In the current study, 11 of 45 differentially expressed proteins could not be validated by MRM analysis using an independent sample set. In MRM, there are several possible reasons why not all targeted proteins were detected. They include 1) the absence of appropriate peptide regions of the targeted proteins which meet the criteria described in the Materials and methods, and 2) the MS peak of the target protein was hard to be detected due to high background noise. For the validation of the 11 candidate proteins, we need to analyze the serum levels by other methods such as ELISA or Western blotting. We now investigate the potential association between the levels of the 11 candidate proteins and T2DM using human healthy and T2DM serum samples. We would examine the functional properties of the validated proteins including the 7 MRM-validated proteins in the future studies.

In summary, iTRAQ and MRM based discovery-through-verification strategy led to the identification of several differentially protein including SERPINA3K in T2DM. With the identified expressed proteins, our proteomics study could provide valuable clues to better understand the underlying mechanisms associated with T2DM.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jprot.2013.03.014.

Conflict of interest

The authors report no conflict of interest.

Acknowledgments

We thank W. Iwata, T. Okamura, M. Goto and the technical staff of the Division of Animal Models at the National Center for Global Health and Medicine for their technical assistances. This work was supported by a Grant-in-Aid (Research on Biological Markers for New Drug Development, H20-009) from the Ministry of Health, Labor, and Welfare of Japan, a Grant-in-Aid for Scientific Research (KAKENHI, 23500870) from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and a grant from the National Center for Global Health and Medicine (23S104).

REFERENCES

[1] Korc M. Diabetes mellitus in the era of proteomics. Mol Cell Proteomics 2003;2:399–404.

[2] Gerich JE. The genetic basis of type 2 diabetes mellitus: impaired insulin secretion versus impaired insulin sensitivity. Endocr Rev 1998;19:491–503.
[3] Klein R, Klein BEK, Moss SE, Cruickshanks KJ. The Wisconsin epidemiologic study of diabetic retinopathy: XVII. The 14-year incidence and progression of diabetic retinopathy and associated risk factors in type 1 diabetes. Ophthalmology 1998;105:1801–15.
[4] Horikawa Y, Yamasaki T, Nakajima H, Shingu R, Yoshiuchi I, Miyagawa J, et al. Identification of a novel variant in the phosphoenolpyruvate carboxykinase gene promoter in Japanese patients with type 2 diabetes. Horm Metab Res 2003;35:308–12.
[5] Vendrell J, Fernandez-Real JM, Gutierrez C, Zamora A, Simon I, Bardaji A, et al. A polymorphism in the promoter of the tumor necrosis factor-α gene (-308) is associated with coronary heart disease in type 2 diabetic patients. Atherosclerosis 2003;167:257–64.
[6] Lindgren CM, Widen E, Tuomi T, Li H, Almgren P, Kanninen T, et al. Contribution of known and unknown susceptibility genes to early-onset diabetes in Scandinavia: evidence for heterogeneity. Diabetes 2002;51:1609–17.
[7] Rao AA, Sridhar GR, Das UN. Elevated butyrylcholinesterase and acetylcholinesterase may predict the development of type 2 diabetes mellitus and Alzheimer's disease. Med Hypotheses 2007;69:1272–6.
[8] Rao AA, Sridhar GR, Srinivas B, Das UN. Bioinformatics analysis of functional protein sequences reveals a role for brain-derived neurotrophic factor in obesity and type 2 diabetes mellitus. Med Hypotheses 2008;70:424–9.
[9] On YK, Park HK, Hyon MS, Jeon ES. Serum resistin as a biological marker for coronary artery disease and restenosis in type 2 diabetic patients. Circ J 2007;71:868–73.
[10] Zhang R, Barker L, Pinchev D, Marshall J, Rasamoelisolo M, Smith C, et al. Mining biomarkers in human sera using proteomic tools. Proteomics 2004;4:244–56.
[11] Dayal B, Ertel NH. ProteinChip technology: a new and facile method for the identification and measurement of high-density lipoproteins apoA-I and apoA-II and their glycosylated products in patients with diabetes and cardiovascular disease. J Proteome Res 2002;1:375–80.
[12] Sundsten T, Eberhardson M, Goransson M, Bergsten P. The use of proteomics in identifying differentially expressed serum proteins in humans with type 2 diabetes. Proteome Sci 2006;4:22.
[13] Suto J, Matsuura S, Imamura K, Yamanaka H, Sekikawa K. Genetics of obesity in KK mouse and effects of A^y allele on quantitative regulation. Mamm Genome 1998;9:506–10.
[14] Unoki-Kubota H, Yamagishi S, Takeuchi M, Bujo H, Saito Y. Pyridoxamine, an inhibitor of advanced glycation end product (AGE) formation ameliorates insulin resistance in obese, type 2 diabetic mice. Protein Pept Lett 2010;17:1177–81.
[15] Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4:44–57.
[16] Okumura A, Ohta H, Inoue Y, Enami I. Identification of functional domains of the extrinsic 12 kDa protein in red algal PSII by limited proteolysis and directed mutagenesis. Plant Cell Physiol 2001;42:1331–7.
[17] Giaever I, Keese CR. Micromotion of mammalian cells measured electrically. Proc Natl Acad Sci U S A 1991;88:7896–900.
[18] Nishimura M. Breeding of mice strains for diabetes mellitus. Exp Anim 1969;18:147–57.
[19] Hatcher HC, Ma JX, Chao J, Chao L, Ottlecz A. Kallikrein-binding protein levels are reduced in the retinas of streptozotocin-induced diabetic rats. Invest Ophthalmol Vis Sci 1997;38:658–64.
[20] Gao G, Shao C, Zhang SX, Dudley A, Fant J, Ma J-X. Kallikrein-binding protein inhibits retinal neovascularization and decreases vascular leakage. Diabetologia 2003;46:689–98.
[21] Klein R, Klein BEK, Moss SE, Cruickshanks KJ. The Wisconsin epidemiologic study of diabetic retinopathy. Opthalmology 1995;102:7–16.
[22] Kaur P, Rizk NM, Ibrahim S, Younes N, Uppal A, Dennis K, et al. iTRAQ-based quantitative protein expression profiling and MRM verification of markers in type 2 diabetes. J Proteome Res 2012;11:5527–39.
[23] Overgaard AJ, Thingholm TE, Larsen MR, Tarnow L, Rossing P, McGuire JN, et al. Quantitative iTRAQ-based proteomic identification of candidate biomarkers for diabetic nephropathy in plasma of type 1 diabetic patients. Clin Proteomics 2010;6:105–14.
[24] García-Ramírez M, Canals F, Hernández C, Colomé N, Ferrer C, Carrasco E, et al. Proteomic analysis of human vitreous fluid by fluorescence based difference gel electrophoresis (DIGE): a new strategy for identifying potential candidates in the pathogenesis of proliferative diabetic retinopathy. Diabetologia 2007;50:1294–303.
[25] Klöting N, Graham TE, Berndt J, Kralisch S, Kovacs P, Wason CJ, et al. Serum retinol-binding protein is more highly expressed in visceral than in subcutaneous adipose tissue and is a marker of intra-abdominal fat mass. Cell Metab 2007;6:79–87.
[26] Fernández-Real JM, Moreno JM, Ricart W. Circulating retinol-binding protein-4 concentration might reflect insulin resistance-associated iron overload. Diabetes 2008;57:1918–25.
[27] Raila J, Henze A, Spranger J, Möhlig M, Pfeiffer AF, Schweigert FJ. Microalbuminuria is a major determinant of elevated plasma retinol-binding protein 4 in type 2 diabetic patients. Kidney Int 2007;72:505–11.
[28] Chao J, Chai KX, Chen LM, Xiong W, Chao S, Woodley-Miller C, et al. Tissue kallikrein-binding protein is a SERPIN. I. Purification, characterization, and distribution in normotensive and spontaneously hypertensive rats. J Biol Chem 1990;265:16394–401.
[29] Chao J, Tillman DM, Wang MY, Margolius HS, Chao L. Identification of a new tissue-kallikrein-binding protein. Biochem J 1986;239:325–31.
[30] Gettins PG. SERPIN structure, mechanism, and function. Chem Rev 2002;102:4751–803.
[31] Zhang B, Hu Y, Ma JX. Anti-inflammatory and antioxidant effects of SERPINA3K in the retina. Invest Ophthalmol Vis Sci 2009;50:3943–52.
[32] Zhang B, Abreu JG, Zhou K, Chen Y, Hu Y, Zhou T, et al. Blocking the Wnt pathway, a unifying mechanism for an angiogenic inhibitor in the serine proteinase inhibitor family. Proc Natl Acad Sci U S A 2010;107:6900–5.
[33] Kwak MK, Wakabayashi N, Itoh K, Motohashi H, Yamamoto M, Kensler TW. Modulation of gene expression by cancer chemopreventive dithiolethiones through the Keap1-Nrf2 pathway: identification of novel gene clusters for cell survival. J Biol Chem 2003;278:8135–45.
[34] Ishii T, Itoh K, Ruiz E, Leake DS, Unoki H, Yamamoto M, et al. Role of Nrf2 in the regulation of CD36 and stress protein expression in murine macrophages: activation by oxidatively modified LDL and 4-hydroxynonenal. Circ Res 2004;94:609–16.
[35] Xu J, Kulkarni SR, Donepudi AC, More VR, Slitt AL. Enhanced Nrf2 activity worsens insulin resistance, impairs lipid accumulation in adipose tissue, and increases hepatic steatosis in leptin-deficient mice. Diabetes 2012;61:3208–18.
[36] American Diabetes Association. Implications of the diabetes control and complications trial. American Diabetes Association. Diabetes 1993;42:1555–8.
[37] Ohkubo K, Ogata S, Misumi Y, Takami N, Ikehara Y. Molecular cloning and characterization of rat contrapsin-like protease inhibitor and related proteins. J Biochem 1991;109:243–50.

JJCO Japanese Journal of Clinical Oncology

Jpn J Clin Oncol 2013;43(2)103–109
doi:10.1093/jjco/hys200
Advance Access Publication 16 December 2012

Review Articles

Proteomic Approaches to the Discovery of Cancer Biomarkers for Early Detection and Personalized Medicine

Kazufumi Honda*, Masaya Ono, Miki Shitashige, Mari Masuda, Masahiro Kamita, Nami Miura and Tesshi Yamada

Department of Chemotherapy and Clinical Research, National Cancer Center Research Institute, Tokyo, Japan

*For reprints and all correspondence: Kazufumi Honda, Department of Chemotherapy and Clinical Research, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan.
E-mail: khonda@ncc.go.jp

Received July 29, 2012; accepted October 31, 2012

Cancer biomarkers for the early detection of malignancies and selection of therapeutic strategies have been requested in the clinical field. Accurate and informative cancer biomarkers hold significant promise for improvements in the early detection of disease and in the selection of the most effective therapeutic strategies. Recently, significant progress in the comprehensive analysis of the human genome, epigenome, transcriptome, proteome and metabolome has led to revolutionary changes in the discovery of cancer biomarkers. The Human Proteome Organization has launched a global Human Proteome Project to map the entire human protein set. The Human Proteome Project research group has focused on three working proteomic pillars—mass spectrometry-based, antibody-based and knowledge-based proteomics—and each of these technologies is advancing rapidly. In this review, we introduce the proteomic platforms that are currently being used for cancer biomarker discovery, and describe examples of novel cancer biomarkers that were identified with each proteomic technology.

Key words: cancer biomarker – personalized medicine – early detection – proteomics

INTRODUCTION

Despite progress in four important areas (imaging technology, surgical management, therapeutic modalities and molecular-targeted therapies), cancer has remained a leading cause of mortality in Japan. Biomarkers, when used for the early detection of cancer and selection of therapeutic strategy, are powerful tools that can improve the outcomes of cancer treatments and reduce cancer-related mortalities. In the past decade, significant progress in the comprehensive analyses of the human genome, epigenome, transcriptome, proteome and metabolome has led to the discovery of a variety of biomarkers that can be used to detect early-stage cancers and predict tumor progression, drug response, clinical outcome or some combination of these three trajectories (1). Moreover, bioinformatics technology based on computational science has greatly facilitated the discovery of biomarkers that are strongly associated with the pathophysiology of particular cancers. For example, multiple dynamic alterations—including protein phosphorylation, protein trafficking and localization, and protein–protein interactions—that have secondary effects are frequently observed in cancer cells. Specific post-translation modifications (e.g. phosphorylation) and/or the status (e.g. nuclear localization) of particular proteins in cancer cells may be meaningful in clinical situations as potential cancer biomarkers. Recently, proteomic analyses have greatly facilitated the comprehensive cataloging of protein expression profiles in not only cell lines but also clinical samples, including serum/plasma, urine, spinal fluid, synovial fluid and tissues. In this review, we focus on the proteomic approaches used in the discovery of

cancer biomarkers, and describe the current state of these proteomic studies.

PLATFORM FOR PROTEOMIC ANALYSIS

Recently, and following the successful completion of the Human Genome Project, the Human Proteome Organization (HUPO) (http://www.hupo.org/) officially launched a global Human Proteome Project (HPP) (http://www.hupo.org/research/hpp/), which aims to map the entire human protein set (2). A systematic global effort will be necessary to achieve this goal with respect to protein abundance, distribution and subcellular localization, as well as with respect to protein interactions with other biomolecules and protein functions at specific time points. As a general experimental strategy, the HPP research group has focused on three working pillars—mass spectrometry (MS)-based, antibody-based and knowledge-based proteomic strategies. In this section, we describe the main proteomic platforms used for cancer biomarker discovery (2).

MS-BASED PROTEOMICS

Recently, MS has evolved to a level such that it can be used to assess the complexity of the human proteome (3). MS has played a critical role in protein biomarker discovery in research on cancer and other diseases.

Initially, proteomic studies were based on two-dimensional gel electrophoresis and subsequent MS. This sequential approach greatly facilitated the identification of peptide sequences in proteins that were present in differential abundance on gels (4–6). Subsequently, proteomic pattern analysis was developed as one approach to biomarker discovery. Mass spectra and pattern recognition have enabled researchers and clinicians to use bioinformatics to distinguish cancer patients from healthy subjects (7). One such MS method is surface-enhanced laser-desorption ionization time-of-flight MS (SELDI-TOF-MS). This technique can be used to analyze the masses of protein directly without digestion using the enzyme (8,9).

MS can be coupled with liquid chromatography separation (LC); this combination is called LC/MS. Typically, in an LC/MS analysis, whole proteins that are present in complex biological samples are digested to peptide fragments by enzyme, and LC/MS is then used to identify the thousands of proteins in the biological samples; these samples can be tissues, serum, plasma, urine or protein elution of cell lines. The LC/MS-based methods, which can be used to comprehensively analyze digested peptides, are collectively called shotgun proteomics (10). There are two major approaches to shotgun proteomics, and each method is designed for accurate quantification of the peptide amount. One method is a labeling method that uses non-radio isotope labels and the other is a non-labeling method. We developed a third method that uses non-labeling shotgun proteomics, named two-dimensional image-converted analysis of LC and MS (LC/MS 2DICAL) (11,12).

ANTIBODY-BASED PROTEOMICS

The human proteome consists of ~20 500 non-redundant proteins that are defined as representative isoforms from every protein-coding gene locus. There are ongoing efforts in the field of proteomics to generate specific antibodies that recognize each component of the human proteome. Antibody-based proteomics plays a pivotal role in the cancer biomarker discovery and validation pipeline; specifically, it facilitates high-throughput evaluation of cancer biomarkers and provides a logical strategy for the systematic generation and use of specific antibodies to explore the proteome. The Human Protein Atlas project has been set up to systematically generate specific antibodies on a global scale and to utilize these antibodies for studies of the corresponding proteins and protein isoforms.

The use of antibodies for protein profiling on a global scale is an intuitive approach that should facilitate the system examination of the cancer proteome; antibody-based approaches can be used in conjunction with a wide range of high-throughput assays such as immunohistochemistry (IHC) on tissue microarrays (TMAs) and protein microarrays (13–15).

TMA is a method of assembling multiple tissue samples from an individual paraffin block to simultaneously evaluate multiple biomarkers using IHC; TMA can potentially become an accelerated molecular method for using a large-scale library of antibodies to examine the association between molecular biomarkers and clinical outcomes (16).

Protein microarray formats can be divided into two major classes: forward-phase protein arrays (FPPAs) and reverse-phase protein arrays (RPPAs) (17). In the FPPA formats, the analytes are captured from the solution phase by a capture molecule, usually an antibody that is immobilized on a substratum and acts as a bait molecule. Each spot on an array contains one type of immobilized antibody or bait protein. Each array is incubated with one test sample (e.g. a cellular lysate or serum sample) that represents a specific treatment condition, and multiple analytes from that sample are measured simultaneously. In contrast, RPPA immobilizes an aliquot of an individual complex test sample (e.g. serum, plasma, protein elution from a tissue sample or a cell lysate) in each array spot. In RPPA formats, each array is incubated with one detection protein, which is usually an antibody, and a single analyte is measured and directly compared across multiple samples (18).

RPPAs are a direct descendant of the miniaturized immunoassay, and the RRPA technology platform is designed for quantitative, multiplexed analysis of specific protein modifications (phosphorylation, glycosylation, processing or some combination thereof), or total forms of cellular proteins from a limited amount of samples. In addition, RPPAs

can be used to interrogate samples of tissues, cells, plasma, serum or body fluids (19,20). Moreover, we have developed a quantitative assay that is based on RPPAs and fluorescent immunoassays (21,22).

Knowledge-based Proteomics

Molecular-level information has been accumulated on a large scale to understand the molecular biology of human cells and the causes of human diseases. The initial goals of these large-scale efforts focused on sequencing the entire human genome and mapping human transcriptomes. Recently, information about human proteomes has attracted increasing attention. The molecular and functional complexity of human proteomes poses challenges to researchers, and this complexity requires bioinformatic resources specifically aimed at capturing, integrating and maintaining up-to-date available knowledge. The HPP will deliver a comprehensive atlas of human proteins in their biological context. It will generate publicly accessible data and informational resources that will, in turn, support further exploration of the human proteome by basic and clinical scientists. The HPP will be built on a knowledge-based database to integrate information derived from the major technological pillars, which include shotgun and targeted MS and polyclonal and monoclonal antibodies (2). For knowledge-based proteomic approaches, the HPP working group has decided to draw upon the UniProtKB/Swiss-Prot (23), PRIDE (24), PeptodeAtras (25), GPMDB (26) and Human Protein Atlas (14) databases.

IDENTIFICATION OF CANCER BIOMARKERS VIA PROTEOMIC ANALYSIS

Plasma and Serum Biomarkers for Cancer

Human plasma proteome analyses have the potential to significantly improve disease diagnosis and therapeutic monitoring. However, analysis of plasma and serum proteomes is a significant obstacle, as these samples contain extremely complex human-derived proteomes and other tissue proteomes as subsets. At the high abundance end of the protein-content spectrum, the concentration of serum albumin is 35–50 mg/ml; in contrast at the low abundance end, the normal range of interleukin 6 in plasma is 0–5 pg/ml. These two clinically useful proteins differ in plasma abundance by a factor of 10^{10}. Given the low abundance in serum and plasma of known cancer biomarkers, proteomic technologies that provide sufficient depth of analysis for biomarker discovery are now urgently needed (27). Moreover, post-translational modifications of proteins are important in many biological processes, and post-translational changes have relevance to disease and cancer. We recently identified a novel and specific post-translational modification of alpha-fibrinogen in plasma samples from pancreatic cancers using 2DICAL; specifically, we found that a proline on alpha-fibrinogen was hydroxylated to form 4-hydroxlated alpha-fibrinogen. We generated a specific antibody that recognizes 4-hydroxlated alpha-fibrinogen, and validated the accuracy of assays based on this antibody for the diagnosis of pancreatic cancer. The levels of serum 4-hydroxlated alpha-fibrinogen were significantly higher in Stage-I patients with pancreatic cancer than in control subjects (11). We are currently developing a sandwich enzyme-linked immunosorbent assay (ELISA) system for prospective validation of 4-hydroxlated alpha-fibrinogen-positive cancer samples.

In addition, we used SELDI-TOF-MS to identify another plasma biomarker for the early detection of pancreatic cancer (28); the C-termini of the apolipoprotein AII homodimer is specifically and identifiably modified in patients with pancreatic cancer and pancreatitis (29,30).

To validate the diagnostic accuracy of biomarker candidates, it is important that the verification studies are conducted with a very large number of samples. As the tool for the rapid validation of plasma and serum biomarkers, we developed an RPPA on which 380 plasma samples were individually printed. Those plasma samples included samples from healthy controls and patients with pancreatic cancer, benign pancreatic cysts and tumors, chronic pancreatitis, esophageal cancer, hepatocellular carcinoma, gastric cancer, colorectal cancer or cholangiocarcinoma (21). This tool can rapidly validate the usefulness of individual antibodies that recognize respective plasma components to decide whether an individual antibody (biomarker candidate) should be subjected to further development for use in clinical situations. We used a combination of 2DICAL and plasma-RPPA to successfully identify plasma and serum biomarkers that can be used for the early detection of colorectal cancers (Fig. 1). Plasma and serum protein profiling is also useful for identifying predictive biomarkers; predictive biomarkers are used to assess the efficacy and side effects of a particular therapeutic modality. Gemcitabine monotherapy is currently the standard of care for patients with advanced pancreatic cancer, but the occurrence of severe neutropenia and thrombocytopenia can sometimes be life-threatening. In order to predict the hematologic toxicities of gemcitabine, we used serum plasma profiling to assess the state of predictive biomarkers. We found that low levels of haptoglobin, a predictive biomarker, are associated with the gemcitabine-induced hematologic toxicities (31).

High-dose interleukin-2 (IL-2) induces durable therapeutic responses in a small subset of patients with metastatic melanoma or with renal cell carcinoma, but simple pretreatment predictors of response have not been identified. To identify predictive biomarkers of clinical response to IL-2 treatments, sera that were obtained from patients who were treated with high-dose IL-2 were collected and analyzed using a customized antibody array. Elevated levels of vascular endothelial growth factor (VEGF) or fibronectin in serum were identified as independent predictors of positive responses for high-dose IL-2 therapy (32).

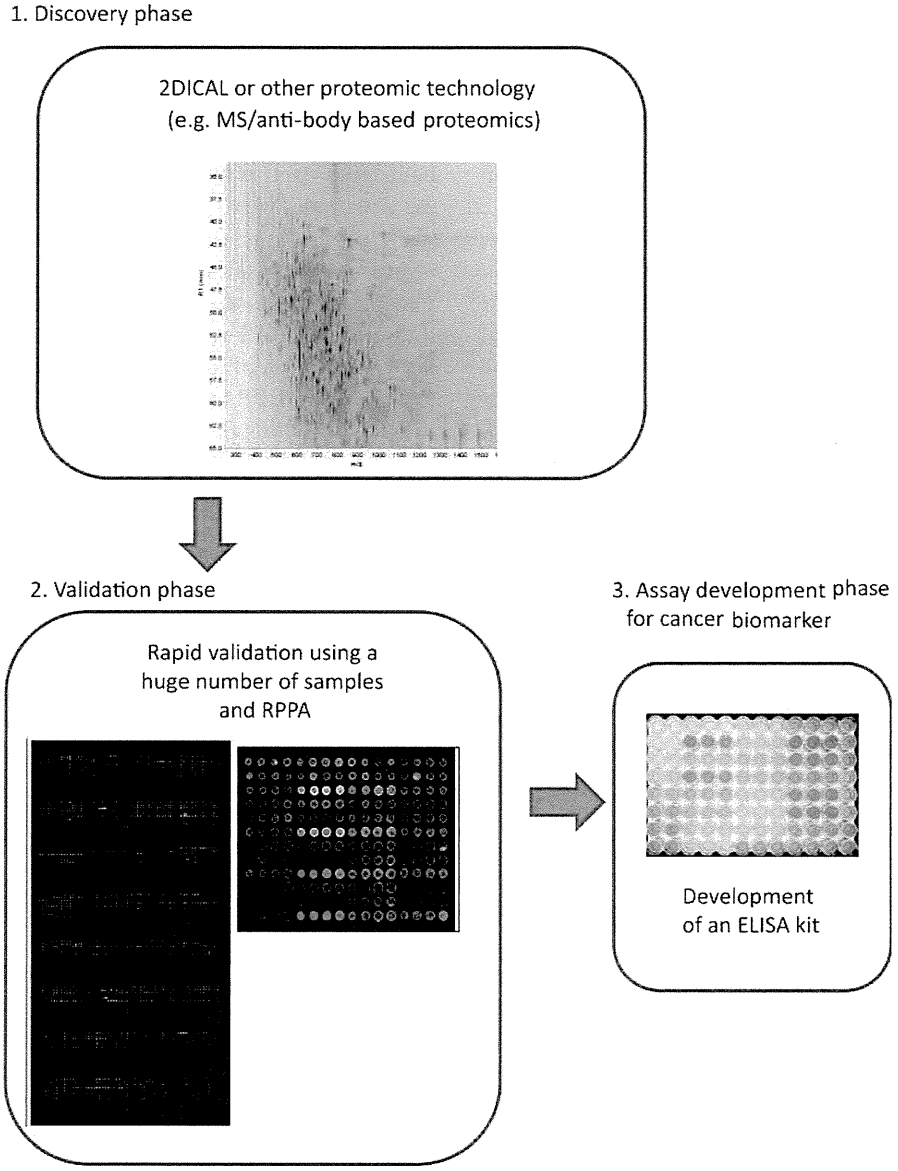

Figure 1. The pipeline for the discovery of a plasma biomarker. There are three phases in the discovery of a plasma biomarker: discovery, validation and assay development. In the discovery phase, MS/antibody-based proteomic approaches are used to screen candidates for plasma biomarkers of cancer. During the validation phase, reverse phase plasma arrays (RPPAs) containing 380 tightly printed plasma samples are used for rapid validation studies. Finally, during the assay development phase, biomarker based-assays are developed in concert with clinical trials for those biomarkers that were validated in the validation phase (assay development phase).

Identification of Therapeutic Targets and Companion Biomarkers

Results from genomic and proteomic studies are eagerly awaited for selecting patients, avoiding the use in non-targeted situation, and reducing the cost of development of the drug. One of the major contributions that proteomics has made to medical communities is the identification of potential targets for effective cancer treatments. Many cancers are characterized by particular alterations in certain signaling pathways, and the identification of a particular alteration in a particular patient could facilitate the selection of an effective therapy targeted at that specific pathway. For example, pathogenesis of non-small cell lung cancer often depends on the activation of the epithelial growth factor receptor (EGFR) signaling pathway, and gefitinib and erlotinib are used for patients who have clinical indicators of activated EGFR signaling (33,34). Profiles of protein expression and post-translational modifications are powerful tools for identifying specific therapeutic targets in cancers. Proteomic technology greatly facilitates the comprehensive analysis of protein expression and post-translational modifications.

One of the most important issues in oncogenesis with regard to post-translational modifications of proteins is aberrant phosphorylation; consequently, inhibition of aberrant

protein phosphorylation has been a focus of molecular-targeted therapies. It is important that the profile of protein phosphorylation within a tumor sample is determined in detail when selecting an appropriate therapeutic strategy. In addition, new molecular targets may be identified during analyses of the phosphorylation protein profiles of cancer cells.

Recently, proteomic analyses have given rise to new technologies for the analysis of phosphoprotein expression profiles. Knowledge of the players in signal transduction mechanisms has accumulated mainly through the study of individual molecules in specific pathways. The emergence of high-throughput transcriptome technologies has resulted in a more detailed and objective view of the downstream transcriptional changes that follow respective stimuli. However, many critical events involved in cellular response are mediated by changes in post-translational protein modification rather than changes in the pattern of transcription.

Development of global and quantitative methods for elucidating dynamic phosphorylation events is therefore essential for a systematic understanding of cancer biology. MS has matured to become a powerful technology for phosphoproteomic analysis. Olsen et al. (35) identified the site-specific phosphorylation dynamics in signal networks under EGF signaling globally. They stimulated HeLa cells with EGF, and then quantified comprehensively the phosphorylation status of phosphopeptides that had been eluted from the cells. These assays involved a combination of the techniques including stable isotope labeling using amino acids in cell culture (SILAC), chromatography of strong-cation exchange (SCX) and titanium dioxide (TiO_2) for phosphopeptide enrichment, and high-accuracy multistage MS. This technique allowed this group to detect 6600 phosphorylation sites on 2244 proteins and to determine the temporal dynamics of the peptide phosphorylation that followed stimulation with EGF (36). More recently, this group applied high-resolution MS-based proteomics to investigate the proteome and phosphoproteome of human cells over the course of the cell cycle, and quantified 6027 proteins, identified 20 443 unique phosphorylation sites and monitored dynamics of the phosphoproteins (35). Individual antibodies that recognize specific phosphopeptides have been generated by many individual scientists and companies to identify the phosphorylation status of particular proteins. RPPAs that employ site-specific antibodies for protein phosphorylations can directly measure the level of phosphorylated protein isoforms, and screens that combine RPPA and antibodies have been used to generate extensive phosphoproteomic profiles. As an alternative to conventional, target-oriented drug discovery, Sevecka et al. (37) reported that RPPA could be used to identify compounds on the basis of the state that they induce in a signaling network. Using a high-throughput strategy, this group screened 84 kinase and phosphatase inhibitors for the ability to induce different states under EGFR signaling.

The assay developed by Sevecka et al. (19,38) allowed these researchers not only to identify interesting compounds that were overlooked in target-oriented screens, but also to uncover functional connections between proteins in signaling networks. RPPAs were originally developed to quantitatively measure numerous proteins extracted from a small number of cells obtained from tissue microdissection. RPPAs have been also used with the samples that were resected from patients with ovary (39), lung (40), prostate (41) or head and neck carcinomas (42). VanMeter et al. used RPPA and microdissected samples from lung cancers to quantify six phosphorylation sites on EGFR to evaluate whether EGFR mutation status *in vivo* was associated with the coordinated phosphorylation of multiple specific phosphorylation sites on EGFR. This RPPA-based analysis revealed simultaneous increases in the phosphorylation of two EGFR residues (^{1148}Tyr and ^{1068}Tyr) and the decreases in phosphorylation of EGFR ^{1045}Tyr, of human epidermal growth factor receptor type 2 ^{1248}Tyr and of insulin receptor substrate (IRS-1) ^{612}Ser. These researchers concluded that a higher proportion of the EGFR mutant carcinoma cells exhibited activation of the phosphatidylinositol 3-kinase/protein kinase B/mammalian target of rapamycin pathway through ^{1148}Tyr and ^{1068}Tyr of EGFR.

In theory, protein-specific affinity reagents that specifically recognize each and every human protein could potentially be generated based on genome sequence information, and such validated affinity reagents could be used to explore protein expression patterns in various normal and cancer tissues. The Human Protein Atlas (HPA) project has started to generate the protein-specific affinity reagents toward all human proteins (15). A major effort in the HPA program was to use affinity-purified mono-specific antibodies for immunohistochemical analysis of a standardized set of TMAs that comprises 48 normal tissues and 216 different cancer specimens, which represent 20 different cancer types (43). The results are being made available as high-resolution images in an Internet-based Human Protein Atlas (http://www.proteinatlas.org/). With the release of the Human Protein Atlas version 7.0, an important objective was reached with the inclusion of 10 118 protein-coding genes corresponding to over 50% of the 19 627 human entries as defined by UniProt. With this high-throughput protein profiling effort, previously unrecognized associations between novel molecules and certain tumor types can be identified and analyzed. In order to search for proteins expressed in human melanocytes and melanomas, Stromberg et al. (44) used an antibody-based proteomics strategy to screen for proteins expressed in a TMA that contained normal tissue and cancer tissues. They reported that the expression level of the syntaxin-7 protein was inversely correlated with clinical stage of malignant melanoma.

CONCLUDING REMARKS

Over the past two decades, concerted scientific efforts have aimed at mapping the human genome. Recently, proteomic

technologies—including MS-based, antibody-based and knowledge-based proteomics—have matured. These proteomic approaches have the potential to greatly facilitate the discovery of new cancer biomarkers that can be applied to clinically applicable assays. It is hoped that this branch of proteomics will accelerate the development of biomarkers that can be used for the early detection of cancer and personalized therapeutic strategies in clinical situations.

Funding

This work was supported in part by the National Cancer Center Research and Development Fund (23-A-38 and 23-A-11).

Conflict of interest statement

None declared.

References

1. Ludwig JA, Weinstein JN. Biomarkers in cancer staging, prognosis and treatment selection. *Nat Rev Cancer* 2005;5:845–56.
2. Legrain P, Aebersold R, Archakov A, et al. The human proteome project: current state and future direction. *Mol Cell Proteomics* 2011;10:M111.009993.
3. Lamond AI, Uhlen M, Horning S, et al. Advancing cell biology through proteomics in space and time (PROSPECTS). *Mol Cell Proteomics* 2012;11:O112 017731.
4. Govorun VM, Archakov AI. Proteomic technologies in modern biomedical science. *Biochemistry (Mosc)* 2002;67:1109–23.
5. Zhou G, Li H, DeCamp D, et al. 2D differential in-gel electrophoresis for the identification of esophageal scans cell cancer-specific protein markers. *Mol Cell Proteomics* 2002;1:117–24.
6. Kondo T, Hirohashi S. Application of highly sensitive fluorescent dyes (CyDye DIGE Fluor saturation dyes) to laser microdissection and two-dimensional difference gel electrophoresis (2D-DIGE) for cancer proteomics. *Nat Protoc* 2006;1:2940–56.
7. Petricoin EF, Zoon KC, Kohn EC, et al. Clinical proteomics: translating benchside promise into bedside reality. *Nat Rev Drug Discov* 2002;1:683–95.
8. Wulfkuhle JD, Liotta LA, Petricoin EF. Proteomic applications for the early detection of cancer. *Nat Rev Cancer* 2003;3:267–75.
9. Issaq HJ, Veenstra TD, Conrads TP, et al. The SELDI-TOF MS approach to proteomics: protein profiling and biomarker identification. *Biochem Biophys Res Commun* 2002;292:587–92.
10. Chen EI, Yates JR, III. Cancer proteomics by quantitative shotgun proteomics. *Mol Oncol* 2007;1:144–59.
11. Ono M, Shitashige M, Honda K, et al. Label-free quantitative proteomics using large peptide data sets generated by nanoflow liquid chromatography and mass spectrometry. *Mol Cell Proteomics* 2006;5:1338–47.
12. Negishi A, Ono M, Handa Y, et al. Large-scale quantitative clinical proteomics by label-free liquid chromatography and mass spectrometry. *Cancer Sci* 2009;100:514–9.
13. Uhlen M, Ponten F. Antibody-based proteomics for human tissue profiling. *Mol Cell Proteomics* 2005;4:384–93.
14. Uhlen M, Bjorling E, Agaton C, et al. A Human Protein Atlas for normal and cancer tissues based on antibody proteomics. *Mol Cell Proteomics* 2005;4:1920–32.
15. Uhlen M, Oksvold P, Fagerberg L, et al. Towards a knowledge-based human protein atlas. *Nat Biotechnol* 2010;28:1248–50.
16. Brennan DJ, O'Connor DP, Rexhepaj E, et al. Antibody-based proteomics: fast-tracking molecular diagnostics in oncology. *Nat Rev Cancer* 2010;10:605–17.
17. Espina V, Mehta AI, Winters ME, et al. Protein microarrays: molecular profiling technologies for clinical specimens. *Proteomics* 2003;3:2091–100.
18. Chandra H, Reddy PJ, Srivastava S. Protein microarrays and novel detection platforms. *Expert Rev Proteomics* 2011;8:61–79.
19. Petricoin EF, III, Bichsel VE, Calvert VS, et al. Mapping molecular networks using proteomics: a vision for patient-tailored combination therapy. *J Clin Oncol* 2005;23:3614–21.
20. Grubb RL, Calvert VS, Wulkuhle JD, et al. Signal pathway profiling of prostate cancer using reverse phase protein arrays. *Proteomics* 2003;3:2142–6.
21. Murakoshi Y, Honda K, Sasazuki S, et al. Plasma biomarker discovery and validation for colorectal cancer by quantitative shotgun mass spectrometry and protein microarray. *Cancer Sci* 2011;102:630–8.
22. Matsubara J, Ono M, Honda K, et al. Survival prediction for pancreatic cancer patients receiving gemcitabine treatment. *Mol Cell Proteomics* 2010;9:695–704.
23. Consortium TU. Ongoing and future developments at the universal protein resource. *Nucleic Acids Res* 2011;39:D214–9.
24. Martens L, Hermjakob H, Jones P, et al. PRIDE: the proteomics identifications database. *Proteomics* 2005;5:3537–45.
25. Fenyo D, Eriksson J, Beavis R. Mass spectrometric protein identification using the global proteome machine. *Methods Mol Biol* 2010;673:189–202.
26. Craig R, Cortens JP, Beavis RC. Open source system for analyzing, validating, and storing protein identification data. *J Proteome Res* 2004;3:1234–42.
27. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* 2002;1:845–67.
28. Honda K, Hayashida Y, Umaki T, et al. Possible detection of pancreatic cancer by plasma protein profiling. *Cancer Res* 2005;65:10613–22.
29. Ehmann M, Felix K, Hartmann D, et al. Identification of potential markers for the detection of pancreatic cancer through comparative serum protein expression profiling. *Pancreas* 2007;34:205–14.
30. Honda K, Okusaka T, Felix K, et al. Altered plasma apolipoprotein modifications in patients with pancreatic cancer: protein characterization and multi-institutional validation. *PLoS One* 2012;7:e46908.
31. Matsubara J, Ono M, Negishi A, et al. Identification of a predictive biomarker for hematologic toxicities of gemcitabine. *J Clin Oncol* 2009;27:2261–8.
32. Sabatino M, Kim-Schulze S, Panelli MC, et al. Serum vascular endothelial growth factor and fibronectin predict clinical response to high-dose interleukin-2 therapy. *J Clin Oncol* 2009;27:2645–52.
33. Pao W, Miller VA. Epidermal growth factor receptor mutations, small-molecule kinase inhibitors, and non-small-cell lung cancer: current knowledge and future directions. *J Clin Oncol* 2005;23:2556–68.
34. McDermott U, Settleman J. Personalized cancer therapy with selective kinase inhibitors: an emerging paradigm in medical oncology. *J Clin Oncol* 2009;27:5650–9.
35. Olsen JV, Vermeulen M, Santamaria A, et al. Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Sci Signal* 2010;3:ra3.
36. Olsen JV, Blagoev B, Gnad F, et al. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 2006;127:635–48.
37. Sevecka M, MacBeath G. State-based discovery: a multidimensional screen for small-molecule modulators of EGF signaling. *Nat Methods* 2006;3:825–31.
38. Pierobon M, Belluco C, Liotta LA, et al. Reverse phase protein microarrays for clinical applications. *Methods Mol Biol* 2011;785:3–12.
39. Annunziata CM, Walker AJ, Minasian L, et al. Vandetanib, designed to inhibit VEGFR2 and EGFR signaling, had no clinical activity as monotherapy for recurrent ovarian cancer and no detectable modulation of VEGFR2. *Clin Cancer Res* 2010;16:664–72.
40. VanMeter AJ, Rodriguez AS, Bowman ED, et al. Laser capture microdissection and protein microarray analysis of human non-small cell lung cancer: differential epidermal growth factor receptor (EGPR) phosphorylation events associated with mutated EGFR compared with wild type. *Mol Cell Proteomics* 2008;7:1902–24.

41. Rapkiewicz A, Espina V, Zujewski JA, et al. The needle in the haystack: application of breast fine-needle aspirate samples to quantitative protein microarray technology. *Cancer* 2007;111:173–84.
42. Frederick MJ, VanMeter AJ, Gadhikar MA, et al. Phosphoproteomic analysis of signaling pathways in head and neck squamous cell carcinoma patient samples. *Am J Pathol* 2011;178:548–71.
43. Ponten F, Schwenk JM, Asplund A, et al. The Human Protein Atlas as a proteomic resource for biomarker discovery. *J Intern Med* 2011;270:428–46.
44. Stromberg S, Agnarsdottir M, Magnusson K, et al. Selective expression of Syntaxin-7 protein in benign melanocytes and malignant melanoma. *J Proteome Res* 2009;8:1639–46.

Cancer Cell
Article

Ablation of Fbxw7 Eliminates Leukemia-Initiating Cells by Preventing Quiescence

Shoichiro Takeishi,[1,2] Akinobu Matsumoto,[1,2] Ichiro Onoyama,[1,2] Kazuhito Naka,[3] Atsushi Hirao,[2,3] and Keiichi I. Nakayama[1,2,*]

[1]Department of Molecular and Cellular Biology, Medical Institute of Bioregulation, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka, Fukuoka 812-8582, Japan

[2]CREST (Core Research for Evolutional Science and Technology), Japan Science and Technology Agency, Kawaguchi, Saitama 332-0012, Japan

[3]Division of Molecular Genetics, Center for Cancer and Stem Cell Research, Cancer Research Institute, Kanazawa University, Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan

*Correspondence: nakayak1@bioreg.kyushu-u.ac.jp

http://dx.doi.org/10.1016/j.ccr.2013.01.026

SUMMARY

Imatinib eradicates dividing progenitor cells of chronic myeloid leukemia (CML) but does not effectively target nondividing leukemia-initiating cells (LICs); thus, the disease often relapse after its discontinuation. We now show that Fbxw7 plays a pivotal role in maintenance of quiescence in LICs of CML by reducing the level of c-Myc. Abrogation of quiescence in LICs by Fbxw7 ablation increased their sensitivity to imatinib, and the combination of Fbxw7 ablation with imatinib treatment resulted in a greater depletion of LICs than of normal hematopoietic stem cells in mice. Purging of LICs by targeting Fbxw7 to interrupt their quiescence and subsequent treatment with imatinib may thus provide the basis for a promising therapeutic approach to CML.

INTRODUCTION

Cancer-initiating cells (CICs) are thought to constitute a minor subpopulation of cancer cells that is required for the initiation and maintenance of cancer (Clevers, 2011; Huntly and Gilliland, 2005). This notion is based largely on the characterization of leukemia-initiating cells (LICs), a rare subpopulation of cells that propagates leukemia (Lapidot et al., 1994). LICs were recently shown to share many properties, including self-renewal, pluripotency, and quiescence, with normal hematopoietic stem cells (HSCs) (Clevers, 2011; Huntly and Gilliland, 2005). A fundamental problem in treating leukemia is that the quiescent LIC subpopulation is particularly resistant to conventional chemotherapy and radiation, both of which target cells undergoing DNA replication and are therefore not effective against quiescent (noncycling) cells (Clevers, 2011; Huntly and Gilliland, 2005). Failure to eradicate quiescent LICs may result in reinitiation of malignancy after a period of latency. The development of therapeutic approaches that target quiescent CICs might therefore be expected to have a profound impact on cancer eradication.

Chronic myeloid leukemia (CML) in humans is characterized by the presence of the Philadelphia chromosome, which is generated by a chromosomal translocation that joins the *BCR* gene on chromosome 22 to the *ABL* gene on chromosome 9 (de Klein et al., 1982; Rowley, 1973). CML is a biphasic myeloproliferative disorder, which initially assumes a chronic phase before progressing to an accelerated phase and finally to blast crisis. Given that individuals with CML in blast crisis have a poor prognosis associated with a short survival time, it is critical to treat CML patients during the chronic phase. Several lines

Significance

Most cancer-initiating cells (CICs) are quiescent and therefore resistant to anticancer drugs that preferentially target dividing cells. CICs that survive therapy are a potential cause of relapse. Elucidation of the mechanism by which CICs maintain quiescence is thus critical for the elimination of cancer. Here, we show that Fbxw7 plays a pivotal role in maintenance of quiescence in leukemia-initiating cells (LICs) of chronic myeloid leukemia. Our findings reveal that ablation of Fbxw7 in LICs results in deregulated activation of c-Myc and impaired maintenance of quiescence followed by p53-dependent apoptosis and consequent cell exhaustion. Moreover, they provide a rationale for Fbxw7-targeted therapy to sensitize LICs to currently available drugs by interrupting their quiescence, potentially resulting in a substantial survival benefit.

of evidence indicate that LICs of CML emerge as a result of expression of BCR-ABL in normal HSCs (Pear et al., 1998), supporting the notion that CML is a "stem cell disease."

The BCR-ABL fusion protein possesses constitutive tyrosine kinase activity and triggers molecular events that result in the expansion of malignant hematopoiesis (Deininger et al., 2000). The recent development of the tyrosine kinase inhibitor (TKI) imatinib represented a breakthrough in treatment of the chronic phase of CML, resulting in a marked improvement in the prognosis of CML patients (Druker et al., 2001; Kantarjian et al., 2002). The French CML Intergroup Stop Imatinib study recently found that ~40% of CML patients in complete molecular remission for >2 years while on treatment with imatinib did not relapse within 12 months after discontinuation of imatinib treatment (Mahon et al., 2010). However, this observation suggests that LICs of CML persist in more than half of patients treated with imatinib alone, resulting in relapse after discontinuation of imatinib treatment. Several mechanisms of resistance of CML LICs to imatinib therapy have been suggested, including the maintenance of quiescence (Holtz et al., 2007) and the lack of addiction to BCR-ABL in these cells (Corbin et al., 2011). Although more potent TKIs such as nilotinib and dasatinib have been developed, these drugs also do not target quiescent LICs of CML (Copland et al., 2006; Jørgensen et al., 2007). Therapy with these TKIs thus serves to suppress, not to eliminate, the disease. Moreover, quiescence in CML LICs is thought not only to contribute to TKI resistance but also to be essential for their long-term maintenance. Elucidation of the molecular mechanism by which LICs maintain quiescence is therefore expected to provide a basis for the development of approaches to sensitize CML LICs to TKI therapy, thereby allowing efficient eradication of leukemia cells, prevention of relapse, and increased patient survival. Although several key molecules and signaling pathways have been implicated in LIC maintenance (Ito et al., 2008; Naka et al., 2010; Zhao et al., 2007), the mechanism by which LICs maintain quiescence has been poorly understood.

c-Myc is one of the best characterized proteins found to determine the state of cell proliferation or quiescence (Laurenti et al., 2009). Regulation of the abundance of c-Myc is achieved at several levels, one of which is control of protein stability mediated by posttranslational modification. We and others have shown that the F-box protein Fbxw7 (also known as Fbw7, Sel-10, hCdc4, or hAgo), the substrate-recognition subunit of an SCF-type ubiquitin ligase complex, interacts with and mediates the ubiquitylation of c-Myc (Nakayama and Nakayama, 2006). The ubiquitin-dependent degradation of c-Myc mediated by Fbxw7 has been found to be essential for maintenance of the quiescence and reconstitution capacity of normal HSCs (Matsuoka et al., 2008; Reavie et al., 2010; Thompson et al., 2008). Given that LICs share many properties with normal HSCs, we hypothesized that Fbxw7 might also be required for the maintenance of LICs and that the Fbxw7–c-Myc axis might be a promising target for leukemia therapy.

RESULTS

Fbxw7 Is Required for Maintenance of Quiescence in LICs

To examine whether Fbxw7 expression is modulated during leukemogenesis, we first measured the amount of Fbxw7 mRNA at various stages of the differentiation of leukemic cells in a mouse model of CML caused by the human BCR-ABL fusion protein (Pear et al., 1998). For generation of the model, a c-Kit$^+$Sca-1$^+$Lin$^-$ (KSL) fraction of bone marrow cells, which represents immature hematopoietic cells, was infected with a retrovirus encoding both p210$^{BCR-ABL}$ and green fluorescent protein (GFP) and was subsequently transplanted into syngeneic recipients (Figure 1A). Recipient mice developed signs of CML, including decreased activity, weight loss, an increased number of myeloid cells in peripheral blood, and splenomegaly, and all mice died within 4 weeks after transplantation. Bone marrow cells were collected from the recipient mice after they began to show such signs of CML. Reverse transcription (RT) and real-time PCR analysis revealed that Fbxw7 mRNA was highly abundant in the LIC compartment (GFP$^+$KSL population), whereas it was present in much smaller amounts in the leukemic progenitor compartment (GFP$^+$c-Kit$^+$Sca-1$^-$Lin$^-$ population) and its abundance decreased further as cell differentiation progressed (Figure 1B). Similar results were obtained by quantification of the copy number of Fbxw7 mRNA per cell in these various compartments (Figure S1A available online). These data thus suggested that Fbxw7 expression during leukemogenesis is regulated at least in part at the transcriptional level.

We next examined the role of Fbxw7 in the maintenance of LIC quiescence by conditional disruption of the Fbxw7 gene in this CML mouse model. KSL cells from Mx1-Cre;*Fbxw7*$^{+/+}$ (control) and Mx1-Cre;*Fbxw7*$^{F/F}$ mice were infected with the retrovirus encoding BCR-ABL and GFP and were then transplanted into syngeneic wild-type mice. The donor mice harbored wild-type (+) or floxed (F) alleles of *Fbxw7*, as well as a transgene for Cre recombinase under the control of the Mx1 gene promoter. The number of white blood cells in peripheral blood of the recipient mice was determined every 5 days; when it had increased to >20,000/μl, GFP$^+$KSL cells (2×10^4) were collected from the recipients and transplanted into additional recipient mice (first bone marrow transplantation [BMT]). These recipients were then injected with polyinosinic:polycytidylic acid (pIpC) beginning the day after the first BMT to activate the Mx1-Cre transgene and thereby to delete the floxed *Fbxw7* allele in Mx1-Cre;*Fbxw7*$^{F/F}$ leukemia cells (to yield the *Fbxw7*$^{\Delta/\Delta}$ genotype) (Figure 1C). One week after the final injection of pIpC, we confirmed that almost all floxed alleles of *Fbxw7* were inactivated in each fraction of the targeted leukemia cells (Figure S1B). Analysis of the cell cycle status of leukemic cells from the recipients of the first BMT by flow cytometry revealed that the frequency of Hoechst^{low}pyronin Y^{low} cells, which represent cells in G_0 phase (quiescence), was significantly smaller in the GFP$^+$KSL compartment of *Fbxw7*$^{\Delta/\Delta}$ leukemia cells than in the corresponding compartment of pIpC-treated Mx1-Cre;*Fbxw7*$^{+/+}$ (control) leukemia cells (Figure 1D). Given that gene ablation induced by pIpC is mediated by interferon, which has been shown to act directly on HSCs to induce cell cycle progression (Essers et al., 2009), recipients of control bone marrow cells as well as those of Mx1-Cre;*Fbxw7*$^{F/F}$ cells were injected with pIpC to eliminate any bias attributable to interferon action. In contrast to LICs, most cells in the GFP$^+$c-Kit$^+$Sca-1$^-$Lin$^-$ compartment were actively cycling, and the proportion of quiescent cells in this compartment did not differ between the two genotypes (Figure 1D), consistent with our observation that

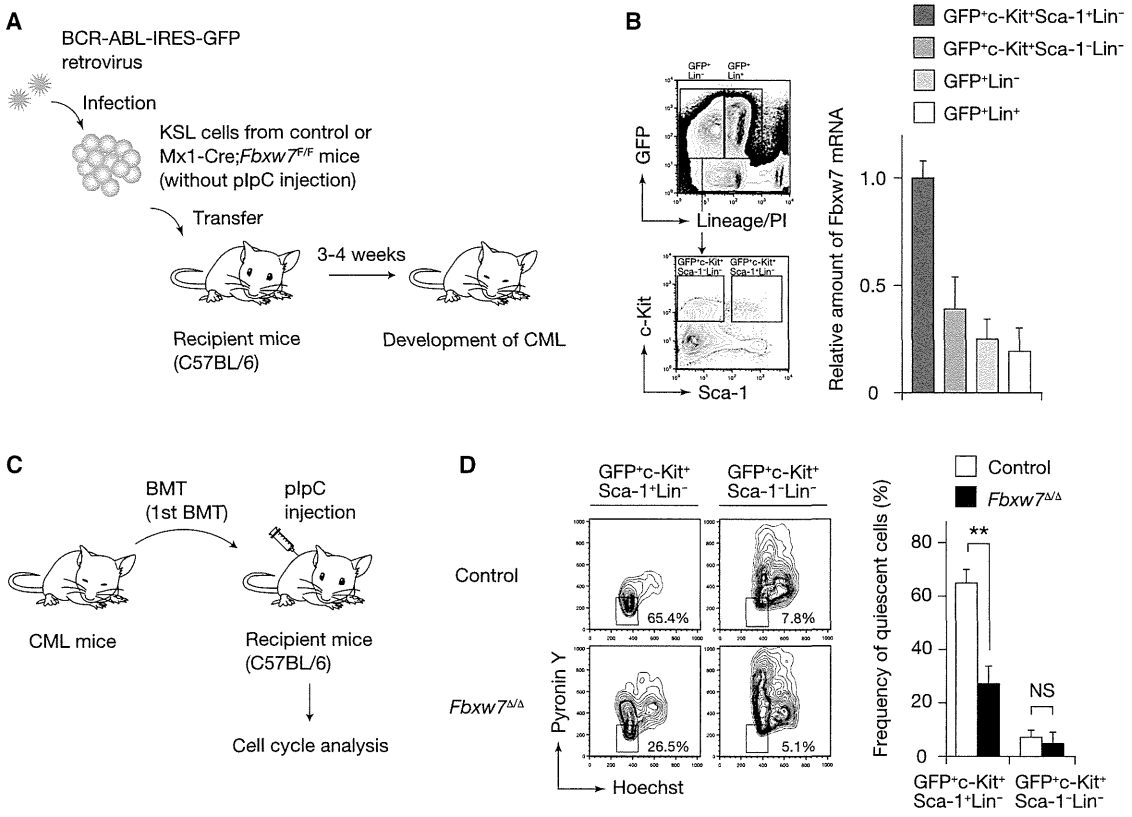

Figure 1. Loss of Fbxw7 in LICs Promotes Cell Cycle Progression

(A) Experimental strategy for generation of a mouse model of CML.

(B) GFP$^+$ cells of control CML bone marrow were fractionated by FACS as indicated (left panels; PI, propidium iodide) and then assayed for Fbxw7 mRNA by RT and real-time PCR analysis (right panel; n = 3).

(C) Experimental strategy for deletion of the floxed *Fbxw7* allele in leukemic cells.

(D) CML bone marrow cells from recipients of the first BMT were subjected to flow cytometry (left panels) for determination of the percentage of quiescent cells in the indicated fractions (right panel; n = 3).

Data are means ± SD. **p < 0.01; NS, not significant. See also Figure S1.

these cells express *Fbxw7* at a low level (Figure 1B). Together, these results suggested that Fbxw7 is expressed predominantly in the LIC fraction of leukemia cells and is required for maintenance of quiescence in LICs.

Fbxw7 Is Required for Leukemic Stemness

The critical role of Fbxw7 in the maintenance of quiescence in LICs and the fact that quiescence is thought to be essential for maintenance of normal HSCs prompted us to examine whether Fbxw7 is essential for LIC maintenance. To test this possibility, we first performed long-term culture-initiating cell assays. Control and *Fbxw7*$^{\Delta/\Delta}$ GFP$^+$KSL cells were isolated from recipients of the first BMT, cultured on OP-9 stromal cells for 2 or 6 weeks, and then subjected to colony-formation assays (Figure 2A). In these assays, the number of colony-forming cells arising after short-term culture (2 weeks) on OP-9 cells mainly reflects progenitor function, whereas that arising after long-term culture (6 weeks) reflects stem cell function (Matsumoto et al., 2011b). We confirmed that the floxed alleles of *Fbxw7* were indeed inactivated efficiently in the *Fbxw7*$^{\Delta/\Delta}$ leukemic cells cultured on OP-9 cells (Figure S2A). The number of cells derived from *Fbxw7*$^{\Delta/\Delta}$ GFP$^+$KSL cells during culture on OP-9 cells was smaller than that derived from the corresponding control cells (Figure S2B). Furthermore, whereas the number of colonies derived from the cells cultured for the short term (2 weeks) did not differ between the two genotypes (Figure 2B), the number of colonies formed by *Fbxw7*$^{\Delta/\Delta}$ GFP$^+$KSL cells was significantly smaller than that formed by the control cells after long-term culture (6 weeks), suggesting that increased cycling of *Fbxw7*$^{\Delta/\Delta}$ LICs eventually results in their exhaustion. We also performed serial replating assays and confirmed that almost all floxed alleles of *Fbxw7* were inactivated in the targeted leukemic cells after serial replating (Figure S2C). Whereas the number of colonies formed did not differ between control and Fbxw7-deficient LICs in the first plating, the number of colonies derived from Fbxw7-deficient LICs was significantly smaller than that derived from control LICs in the second plating (Figure S2D).

To assess the repopulating ability of *Fbxw7*$^{\Delta/\Delta}$ LICs in vivo, we performed serial BMT experiments in which we collected control and *Fbxw7*$^{\Delta/\Delta}$ GFP$^+$KSL cells (2×10^4) from the recipients of the first BMT and transplanted these cells into new recipients (second BMT) (Figure 2C). Almost all floxed alleles of *Fbxw7* were inactivated in the targeted leukemic cells isolated from the recipients of the second BMT (Figure S2E). Whereas the

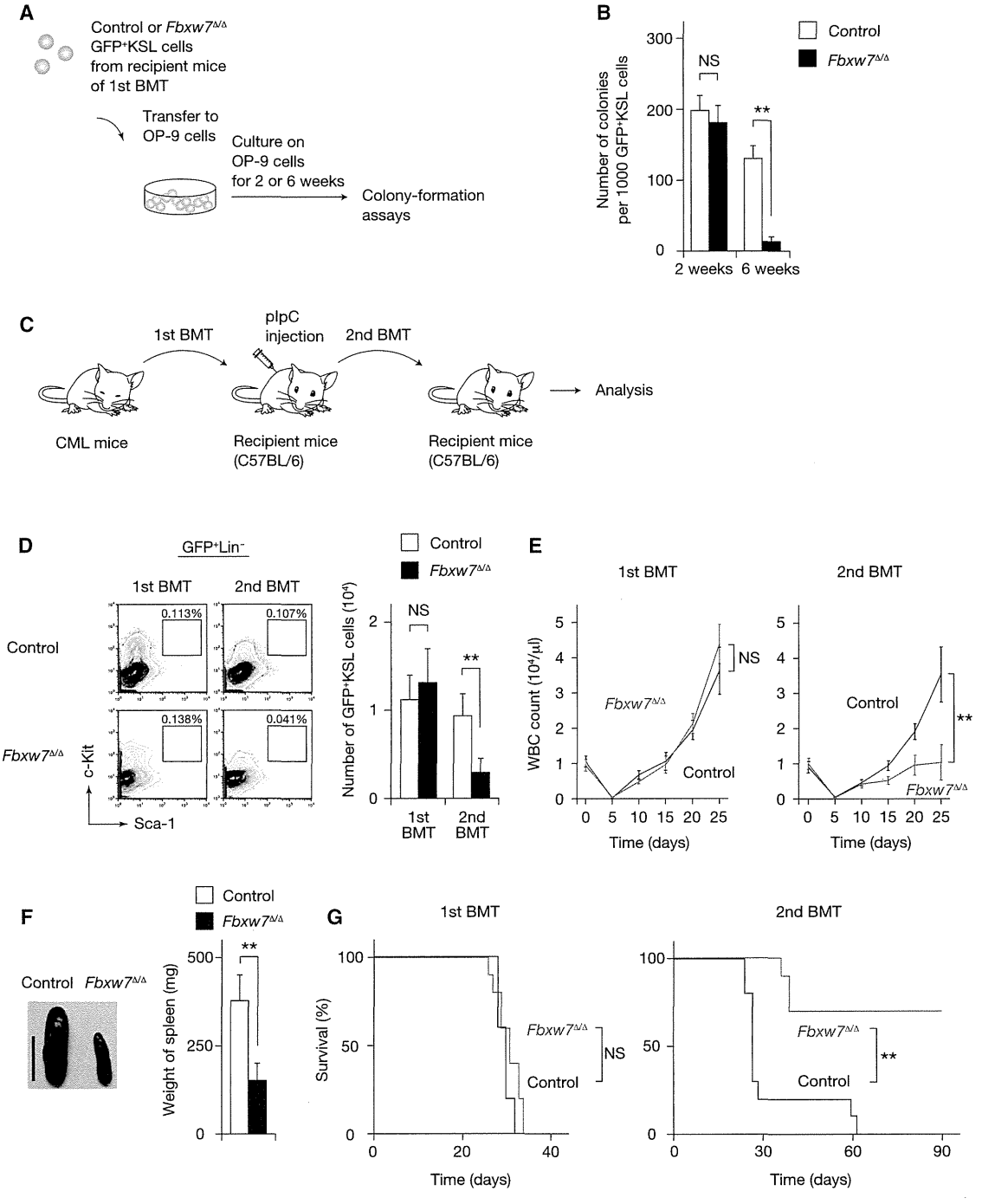

Figure 2. Loss of Fbxw7 in LICs Results in Cell Exhaustion

(A) Experimental strategy for colony formation assays.

(B) Colony formation by BCR-ABL-transduced KSL cells of the indicated genotypes after 2 or 6 weeks of culture on OP-9 cells (n = 3).

(C) Experimental strategy for serial BMT experiments.

(D) Flow-cytometric determination of the absolute number of GFP$^+$KSL cells in bone marrow of recipients (n = 5) after the first and second BMTs.

(E) Numbers of white blood cells (WBC) in peripheral blood of recipient mice (n = 10) after the first and second BMTs.

(F) Representative appearance (scale bar, 10 mm) and weight (n = 5) of the spleen in recipients of the second BMT.

(G) Survival of recipient mice (n = 10) after the first and second BMTs.

Data are means ± SD. **$p < 0.01$; NS, not significant. See also Figure S2.

number of GFP$^+$KSL cells isolated from recipients of the first BMT did not differ between the two genotypes, that of GFP$^+$KSL cells isolated from recipients of the second BMT was greatly reduced in the case of *Fbxw7*$^{\Delta/\Delta}$ donor cells compared with that for control donor cells (Figure 2D). These results thus suggested that the loss of Fbxw7 results in disruption of quiescence, followed by eventual exhaustion of LICs. We also observed that both the proportion and absolute number of Fbxw7-deficient leukemic progenitors were significantly smaller than those of control leukemic progenitors in recipients of the second BMT (Figures S2F and S2G). Given that Fbxw7 deficiency did not affect the cell cycle status of the progenitors (Figure 1D), these latter results were likely attributable to a functional defect in Fbxw7-deficient LICs. Consistent with these observations, the timing of CML development did not differ between the two types of recipients of the first BMT (Figure 2E). In contrast, after the second BMT, Fbxw7 deficiency prevented the propagation of leukemic cells in peripheral blood (Figure 2E) and the spleen (Figure 2F). Furthermore, whereas most recipients of control LICs died of CML at ~30 days after the second BMT, ~70% of mice receiving *Fbxw7*$^{\Delta/\Delta}$ LICs survived for >120 days (Figure 2G; data not shown). In a similar mouse model of CML, LICs were also shown to give rise to acute lymphocytic leukemia (ALL) with a long latency (Pear et al., 1998). Notably, we did not observe development of ALL or CML in recipients of *Fbxw7*$^{\Delta/\Delta}$ LICs later than 40 days after the second BMT. Flow cytometric analysis and histological examination revealed the almost complete absence of leukemic cells in peripheral blood and no infiltration of leukemic cells in the spleen, liver, or lungs of the recipients of Fbxw7-deficient LICs that survived for >90 days after the second BMT (Figure S2H), suggesting that Fbxw7-deficient LICs lose their potential to generate malignancies. Fbxw7 thus appears to be essential for the long-term maintenance of leukemia-initiating potential.

Accumulation of c-Myc Is Responsible for Loss of Leukemic Stemness

We next investigated the mechanism underlying LIC exhaustion associated with Fbxw7 deficiency. Fbxw7 targets many proteins related to HSC maintenance, including c-Myc, Notch1 intracellular domain (NICD1), and mammalian target of rapamycin (mTOR) (Nakayama and Nakayama, 2006). To determine whether these substrates accumulate in Fbxw7-deficient LICs, we examined their abundance in LICs isolated from recipients of the first BMT and found that the abundance of c-Myc was increased in *Fbxw7*$^{\Delta/\Delta}$ GFP$^+$KSL cells compared with control GFP$^+$KSL cells (Figure 3A). In contrast, the expression levels of NICD1 and mTOR did not differ between control and Fbxw7-deficient GFP$^+$KSL cells (Figures S3A and S3B).

To determine whether c-Myc accumulation is responsible for the phenotype of *Fbxw7*$^{\Delta/\Delta}$ LICs, we first cultured control and *Fbxw7*$^{\Delta/\Delta}$ GFP$^+$KSL cells isolated from recipients of the first BMT with the c-Myc inhibitor 10058-F4 (Follis et al., 2009). A colony-formation assay performed after culture of GFP$^+$KSL cells for 2 or 6 weeks in the presence of 10058-F4 revealed that the number of colonies did not differ between the two genotypes (Figure 3B). To confirm that 10058-F4 indeed inhibited c-Myc function in these cells, we measured the abundance of mRNAs for cyclin D2 and ornithine decarboxylase 1 (ODC1), the genes for which are direct targets of c-Myc. The abundance of these mRNAs was increased in Fbxw7-deficient LICs compared with that in control cells, and each increase was attenuated by 10058-F4, suggesting that 10058-F4 indeed inhibits c-Myc activity in these cells (Figure 3C). To demonstrate further that c-Myc accumulation contributes to the phenotype of Fbxw7-deficient LICs, we next generated Mx1-Cre;*Fbxw7*$^{F/F}$;*c-Myc*$^{+/F}$ mice in order to analyze the cell cycle status and colony-forming ability of *Fbxw7*$^{\Delta/\Delta}$;*c-Myc*$^{+/\Delta}$ LICs. The loss of quiescence and impaired colony-forming ability apparent for *Fbxw7*-deficient LICs were normalized by the additional deletion of one allele of the c-Myc gene (Figures 3D and 3E). In contrast, neither a γ-secretase inhibitor, *N*-[*N*-(3,5-difluorophenacetyl)-L-alanyl]-*S*-phenylglycine *t*-butyl ester (DAPT), which antagonizes Notch signaling, nor rapamycin, which antagonizes mTOR, mimicked the effects of 10058-F4 or c-Myc depletion (Figures S3C and S3D), suggesting that neither Notch nor mTOR contributes to the phenotype of Fbxw7-deficient LICs. Collectively, these results thus indeed suggested that the phenotype of *Fbxw7*$^{\Delta/\Delta}$ LICs is attributable to increased activity of c-Myc.

We next compared the amount of BCR-ABL between control and Fbxw7-deficient LICs. Immunoblot analysis revealed no substantial difference in the level of BCR-ABL between these cells (Figure S3E). We further examined whether Fbxw7 deficiency affects signaling downstream of BCR-ABL in LICs by analyzing the phosphorylation of Stat5 (on Tyr^{694}), Crkl (on Tyr^{207}), and Akt (on Ser^{473}). Intracellular flow cytometric analysis revealed that the proportions of cells expressing the phosphorylated forms of Stat5 or Crkl were similar for control and Fbxw7-deficient LICs (Figures S3F and S3G). In contrast, the frequency of cells positive for phosphorylated Akt was greater for Fbxw7-deficient LICs than for control LICs (Figure S3H). We previously showed that Akt phosphorylation is inhibited by transforming growth factor–β (TGF-β) signaling in CML LICs (Naka et al., 2010), and a recent study indicated that TGF-β is activated by a niche for HSCs (Yamazaki et al., 2011). These observations suggest that Fbxw7-deficient LICs might enter the cell cycle and cease to interact with a LIC niche, resulting in a decrease in TGF-β signaling and an increase in Akt phosphorylation. Consistent with this notion, we found that the frequency of cells positive for phosphorylated Smad2/3 was smaller for *Fbxw7*$^{\Delta/\Delta}$ GFP$^+$KSL cells than for control cells (Figure S3I). Moreover, TGF-β1 treatment reversed the increase in the proportion of Fbxw7-deficient LICs positive for phosphorylated Akt (Figure S3J), suggesting that downregulation of TGF-β signaling indeed contributes to this increase. Collectively, our results exclude the possibility that the phenotype of Fbxw7-deficient LICs is attributable to downregulation of BCR-ABL itself or of signaling downstream of BCR-ABL.

To further exclude the possibility that Fbxw7 deficiency in LICs impairs their homing ability, we collected the same number of control and *Fbxw7*$^{\Delta/\Delta}$ GFP$^+$KSL cells from recipients of the first BMT, transplanted the cells into new recipient mice, and determined the proportion of GFP$^+$ cells among bone marrow cells by flow cytometry at 12 hr after transplantation. We found that the frequency of GFP$^+$ cells did not differ significantly between the two genotypes (Figure S3K), suggesting that the phenotype of mice receiving Fbxw7-deficient LICs was not likely a consequence of impaired homing.