Table III. Continued.

Probe ID	Accession no.	Symbol	Gene name	Fold change (log)	P-value
A_32_P39944	AK095791	N/A		-3.82	1.26E-04
A_23_P217379	NM_033641	*COL4A6*	Collagen, type IV, α6	-3.8	1.26E-04
A_23_P407565	NM_001337	*CX3CR1*	Chemokine (C-X3-C motif) receptor 1	-3.76	1.26E-04
A_23_P373464	NM_002285	*AFF3*	AF4/FMR2 family, member 3	-3.75	1.26E-04
A_32_P183765	NM_005235	*ERBB4*	v-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian)	-3.75	1.26E-04
A_23_P145514	NM_014432	*IL20RA*	Interleukin 20 receptor, α	-3.75	1.26E-04
A_24_P870620	NM_002825	*PTN*	Pleiotrophin	-3.74	2.33E-04
A_32_P154361	N/A	N/A		-3.73	1.26E-04
A_24_P330633	NM_000353	*TAT*	Tyrosine aminotransferase	-3.72	1.26E-04
A_23_P360777	NM_013960	*NRG1*	Neuregulin 1	-3.72	1.26E-04
A_23_P253982	NM_002141	*HOXA4*	Homeobox A4	-3.69	1.26E-04
A_32_P114475	N/A	N/A		-3.68	1.26E-04
A_32_P221774	BX099483	N/A		-3.66	1.26E-04
A_23_P212608	NM_022131	*CLSTN2*	Calsyntenin 2	-3.66	2.33E-04
A_23_P254165	NM_021785	*RAI2*	Retinoic acid induced 2	-3.65	1.26E-04
A_24_P794447	NR_024430	*LOC399959*	Hypothetical LOC399959	-3.64	1.26E-04
A_23_P149517	NM_002644	*PIGR*	Polymeric immunoglobulin receptor	-3.64	1.26E-04
A_24_P904484	NR_024344	*LOC283174*	Hypothetical LOC283174	-3.62	1.26E-04
A_32_P194423	N/A	N/A		-3.62	1.26E-04
A_23_P371495	NM_175861	*TMTC1*	Transmembrane and tetratricopeptide repeat containing 1	-3.6	2.33E-04
A_23_P134162	NM_016356	*DCDC2*	Doublecortin domain containing 2	-3.58	1.26E-04
A_32_P232455	NM_178840	*C1orf64*	Chromosome 1 open reading frame 64	-3.58	1.26E-04
A_24_P318160	NM_014903	*NAV3*	Neuron navigator 3	-3.57	1.26E-04
A_23_P59388	NM_001723	*DST*	Dystonin	-3.56	1.26E-04
A_23_P399217	NM_153445	*OR5P3*	Olfactory receptor, family 5, subfamily P, member 3	-3.56	1.26E-04
A_23_P309739	NM_000125	*ESR1*	Estrogen receptor 1	-3.53	1.26E-04
A_24_P608007	AK022390	N/A		-3.53	1.26E-04
A_23_P501538	NM_153631	*HOXA3*	Homeobox A3	-3.52	1.26E-04
A_24_P602871	NM_001030060	*SAMD5*	Sterile α motif domain containing 5	-3.52	1.26E-04
A_23_P136433	N/A	N/A		-3.51	1.26E-04
A_23_P30294	NM_001801	*CDO1*	Cysteine dioxygenase, type I	-3.48	1.26E-04
A_23_P218928	NM_016613	*FAM198B*	Family with sequence similarity 198, member B	-3.47	1.26E-04
A_23_P154627	XM_002345419	*TSHZ2*	Teashirt zinc finger homeobox 2	-3.47	1.26E-04
A_23_P303833	NM_174934	*SCN4B*	Sodium channel, voltage-gated, type IV, β	-3.45	1.26E-04
A_24_P930088	XM_002342181	*LOC100286909*	Hypothetical protein LOC100286909	-3.45	1.26E-04
A_32_P81623	AA514833	N/A		-3.42	1.26E-04
A_24_P923028	BC020707	*TAT*	Tyrosine aminotransferase	-3.41	1.26E-04
A_23_P58869	NR_002932	*LOC442245*	Glutathione S-transferase mu 2 pseudogene	-3.4	1.26E-04
A_23_P2271	NM_198965	*PTHLH*	Parathyroid hormone-like hormone	-3.4	1.26E-04
A_32_P43664				-3.39	1.26E-04
A_32_P16007	NM_207355	*POTEB*	POTE ankyrin domain family, member B	-3.39	1.26E-04
A_23_P94840	NM_130897	*DYNLRB2*	Dynein, light chain, roadblock-type 2	-3.38	1.26E-04
A_24_P5153	NM_024817	*THSD4*	Thrombospondin, type I, domain containing 4	-3.38	1.26E-04

Table III. Continued.

Probe ID	Accession no.	Symbol	Gene name	Fold change (log)	P-value
A_32_P223675	N/A	N/A		-3.37	1.26E-04
A_24_P904845	AK095791	N/A		-3.37	1.26E-04
A_23_P403209	N/A	N/A		-3.36	1.26E-04
A_23_P215382	N/A	N/A		-3.35	3.41E-04
A_24_P209710	NM_004816	*FAM189A2*	Family with sequence similarity 189, member A2	-3.35	1.26E-04
A_23_P167168	NM_144646	*IGJ*	Immunoglobulin J polypeptide, linker protein for immunoglobulin α and mu polypeptides	-3.34	1.26E-04
A_24_P70183	NM_001040113	*MYH11*	Myosin, heavy chain 11, smooth muscle	-3.32	1.26E-04
A_23_P216361	NM_021110	*COL14A1*	collagen, type XIV, α1	-3.32	1.26E-04
A_23_P113351	NM_004684	*SPARCL1*	SPARC-like 1 (hevin)	-3.31	1.26E-04
A_32_P17145	N/A	N/A		-3.31	1.26E-04
A_23_P35414	NM_005398	*PPP1R3C*	Protein phosphatase 1, regulatory (inhibitor) subunit 3C	-3.29	1.26E-04
A_23_P31945	NM_033439	*IL33*	Interleukin 33	-3.27	1.26E-04
A_23_P204630	NM_021229	*NTN4*	Netrin 4	-3.26	1.26E-04
A_23_P501831	NM_032385	*C5orf4*	Chromosome 5 open reading frame 4	-3.26	1.26E-04
A_23_P200015	NM_174858	*AK5*	Adenylate kinase 5	-3.26	1.26E-04
A_24_P802145	NM_005544	*IRS1*	Insulin receptor substrate 1	-3.26	1.26E-04
A_24_P251969	NM_000800	*FGF1*	Fibroblast growth factor 1 (acidic)	-3.24	1.26E-04
A_32_P228618	NM_001003793	*RBMS3*	RNA binding motif, single stranded interacting protein	-3.23	1.26E-04
A_23_P125233	NM_001299	*CNN1*	Calponin 1, basic, smooth muscle	-3.22	2.33E-04
A_23_P500998	NM_152739	*HOXA9*	Homeobox A9	-3.19	2.33E-04
A_23_P83838	NM_004056	*CA8*	Carbonic anhydrase VIII	-3.19	1.26E-04
A_24_P911950	N/A	N/A		-3.17	1.26E-04
A_23_P159952	NM_018476	*BEX1*	Brain expressed, X-linked 1	-3.17	1.26E-04
A_23_P45185	NM_004469	*FIGF*	c-fos induced growth factor (vascular endothelial growth factor D)	-3.16	2.33E-04
A_23_P14083	NM_181847	*AMIGO2*	Adhesion molecule with Ig-like domain 2	-3.16	1.26E-04
A_24_P920366	N/A	N/A		-3.14	1.26E-04
A_24_P167668	NM_000428	*LTBP2*	Latent transforming growth factor β binding protein 2	-3.12	1.26E-04
A_32_P161033	BC043411	N/A		-3.11	1.26E-04
A_23_P348159	NM_020388	*DST*	Dystonin	-3.11	1.26E-04
A_32_P89415	N/A	N/A		-3.1	1.26E-04
A_23_P165778	NM_024101	*MLPH*	Melanophilin	-3.08	1.26E-04
A_32_P168701	N/A	N/A		-3.07	3.41E-04
A_32_P78491	NM_004956	*ETV1*	ets variant 1	-3.06	1.26E-04
A_24_P87036	NM_018043	*ANO1*	Anoctamin 1, calcium activated chloride channel	-3.06	1.26E-04
A_24_P912799	NM_003966	*SEMA5A*	Sema domain, seven thrombospondin repeats (type 1 and type 1-like), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 5A	-3.06	1.26E-04
A_23_P315364	NM_002089	*CXCL2*	Chemokine (C-X-C motif) ligand 2	-3.05	1.26E-04
A_24_P71341	NM_001461	*FMO5*	Flavin containing monooxygenase 5	-3.05	2.33E-04
A_32_P199796	NM_004023	*DMD*	Dystrophin	-3.05	2.33E-04
A_32_P179998	NM_033053	*DMRTC1*	DMRT-like family C1	-3.04	1.26E-04
A_32_P17984	N/A	N/A		-3.04	1.26E-04

Table III. Continued.

Probe ID	Accession no.	Symbol	Gene name	Fold change (log)	P-value
A_23_P138938	NM_000926	*PGR*	Progesterone receptor	-3.04	1.26E-04
A_23_P18559	NM_003866	*INPP4B*	Inositol polyphosphate-4-phosphatase, type II, 105 kDa	-3.03	1.26E-04
A_23_P124946	NM_153610	*CMYA5*	Cardiomyopathy associated 5	-3.03	1.26E-04
A_23_P212241	NM_006614	*CHL1*	Cell adhesion molecule with homology to L1CAM (close homolog of L1)	-3.03	1.26E-04
A_23_P156402	NM_003551	*NME5*	Non-metastatic cells 5, protein expressed in (nucleoside-diphosphate kinase)	-3.02	1.26E-04
A_23_P150053	NM_001613	*ACTA2*	Actin, α2, smooth muscle, aorta	-3.02	1.26E-04
A_32_P58912	N/A	N/A		-3.02	1.26E-04
A_32_P216841	NM_145263	*SPATA18*	Spermatogenesis associated 18 homolog (rat)	-3.01	2.33E-04
A_23_P257087	NM_002612	*PDK4*	Pyruvate dehydrogenase kinase, isozyme 4	-3.01	1.26E-04
A_23_P110686	NM_003714	*STC2*	Stanniocalcin 2	-3	1.26E-04
A_23_P369994	NM_004734	*DCLK1*	Doublecortin-like kinase 1	-2.99	2.33E-04
A_23_P422831	NM_004816	*FAM189A2*	Family with sequence similarity 189, member A2	-2.98	1.26E-04
A_24_P325992	NM_002310	*LIFR*	Leukemia inhibitory factor receptor α	-2.98	1.26E-04
A_23_P387000	NM_173683	*XKR6*	XK, Kell blood group complex subunit-related family, member 6	-2.98	3.41E-04
A_32_P83811	NM_001136570	*FAM47E*	Family with sequence similarity 47, member E	-2.98	1.26E-04
A_32_P44210	BX538299	N/A		-2.97	1.26E-04
A_24_P918317	NM_015881	*DKK3*	Dickkopf homolog 3 (*Xenopus laevis*)	-2.97	4.43E-04
A_23_P203957	NM_175861	*TMTC1*	Transmembrane and tetratricopeptide repeat containing 1	-2.96	3.41E-04
A_23_P30217	NM_052863	*SCGB3A1*	Secretoglobin, family 3A, member 1	-2.96	1.26E-04
A_23_P77066	NM_022807	*SNRPN*	Small nuclear ribonucleoprotein polypeptide N	-2.94	1.26E-04
A_32_P109242	AK055302	*CSRNP3*	Cysteine-serine-rich nuclear protein 3	-2.91	1.26E-04
A_24_P937265	N/A	N/A		-2.91	1.26E-04
A_32_P97968	N/A	N/A		-2.9	1.26E-04
A_32_P85684	AA069768	N/A		-2.89	1.26E-04
A_23_P385067	NM_053277	*CLIC6*	Chloride intracellular channel 6	-2.89	4.43E-04
A_23_P82868	NM_000930	*PLAT*	Plasminogen activator, tissue	-2.88	1.26E-04
A_32_P108396	N/A	N/A		-2.88	1.26E-04
A_23_P148345	NM_194463	*RNF128*	Ring finger protein 128	-2.87	1.26E-04
A_24_P314477	NM_178012	*TUBB2B*	Tubulin, β 2B	-2.87	1.26E-04
A_24_P895836	N/A	N/A		-2.87	1.26E-04
A_23_P171074	NM_004867	*ITM2A*	Integral membrane protein 2A	-2.85	1.26E-04
A_23_P9135	NM_033655	*CNTNAP3*	Contactin associated protein-like 3	-2.85	4.43E-04
A_23_P372234	NM_001218	*CA12*	Carbonic anhydrase XII	-2.83	1.26E-04
A_23_P393099	NM_003226	*TFF3*	Trefoil factor 3 (intestinal)	-2.82	2.33E-04
A_23_P113701	NM_002607	*PDGFA*	Platelet-derived growth factor α polypeptide	-2.82	1.26E-04
A_23_P10995	NM_014483	*RBMS3*	RNA binding motif, single stranded interacting protein	-2.82	1.26E-04
A_24_P269006	NM_001182	*ALDH7A1*	Aldehyde dehydrogenase 7 family, member A1	-2.81	1.26E-04
A_23_P415533	AK054879	N/A		-2.81	1.26E-04
A_23_P216225	NM_004430	*EGR3*	Early growth response 3	-2.8	1.26E-04
A_24_P101282	N/A	N/A		-2.8	1.26E-04
A_32_P72541	N/A	N/A		-2.8	2.33E-04
A_24_P299474	NM_001122679	*ODZ2*	odz, odd Oz/ten-m homolog 2 (*Drosophila*)	-2.8	1.26E-04

Table III. Continued.

Probe ID	Accession no.	Symbol	Gene name	Fold change (log)	P-value
A_23_P416395	NM_003714	*STC2*	Stanniocalcin 2	-2.8	1.26E-04
A_23_P40415	NM_007038	*ADAMTS5*	ADAM metallopeptidase with thrombospondin type 1 motif, 5	-2.8	1.26E-04
A_32_P3545	XM_002345868	*LOC100131504*	Hypothetical LOC100131504	-2.79	4.43E-04
A_23_P106405	NM_002487	*NDN*	Necdin homolog (mouse)	-2.79	1.26E-04
A_23_P405129	NM_000428	*LTBP2*	Latent transforming growth factor β binding protein 2	-2.79	1.26E-04
A_24_P237804	NM_174981	*POTED*	POTE ankyrin domain family, member D	-2.78	1.26E-04
A_23_P89780	NM_198129	*LAMA3*	Laminin, α3	-2.78	1.26E-04
A_23_P213415	NM_003966	*SEMA5A*	Sema domain, seven thrombospondin repeats (type 1 and type 1-like), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 5A	-2.77	3.41E-04
A_24_P397386	NM_002310	*LIFR*	Leukemia inhibitory factor receptor α	-2.77	1.26E-04
A_23_P73297	NM_004742	*MAGI1*	Membrane associated guanylate kinase, WW and PDZ domain containing 1	-2.77	1.26E-04
A_23_P165783	NM_024101	*MLPH*	Melanophilin	-2.76	1.26E-04
A_23_P212061	NM_007289	*MME*	Membrane metallo-endopeptidase	-2.76	1.26E-04
A_23_P75056	NM_001002295	*GATA3*	GATA binding protein 3	-2.76	1.26E-04
A_24_P748377	CR749529			-2.75	2.33E-04
A_24_P810476		*NTRK3*	Neurotrophic tyrosine kinase, receptor, type 3	-2.74	3.41E-04
A_32_P60606	AL713753	*DKFZp667F0711*	Hypothetical protein DKFZp667F0711	-2.74	1.26E-04
A_32_P200697	NM_181709	*FAM101A*	Family with sequence similarity 101, member A	-2.73	4.43E-04
A_24_P84220	NR_027995	*LOC284232*	Ankyrin repeat domain 20 family, member A2 pseudogene	-2.73	1.26E-04
A_23_P157914	NM_153267	*MAMDC2*	MAM domain containing 2	-2.71	1.26E-04
A_24_P393596	N/A	N/A		-2.71	1.26E-04
A_32_P25419	N/A	N/A		-2.7	1.26E-04
A_24_P169873	N/A	N/A		-2.7	1.26E-04
A_24_P358534	N/A	N/A		-2.69	3.41E-04
A_32_P34750	AV702101	N/A		-2.69	1.26E-04
A_32_P9941	NM_007191	*WIF1*	WNT inhibitory factor 1	-2.68	2.33E-04
A_23_P335143	U81001	*SNRPN*	Small nuclear ribonucleoprotein polypeptide N	-2.67	1.26E-04
A_23_P56855	NM_001137671	*POTEC*	POTE ankyrin domain family, member C	-2.67	1.26E-04
A_32_P59837	AK091914	N/A		-2.65	1.26E-04
A_24_P737553	AK023774	N/A		-2.65	2.33E-04
A_23_P204286	NM_000900	*MGP*	Matrix Gla protein	-2.65	1.26E-04
A_24_P725895	BE218249	N/A		-2.63	1.26E-04
A_32_P4337	N/A	N/A		-2.63	1.26E-04
A_23_P154400	NM_001042467	*MLPH*	Melanophilin	-2.62	1.26E-04
A_23_P29800	NM_005602	*CLDN11*	Claudin 11	-2.61	1.26E-04
A_23_P156025	NM_033267	*IRX2*	Iroquois homeobox 2	-2.61	1.26E-04
A_32_P193091	N/A	N/A		-2.61	1.26E-04
A_23_P83857	NM_000240	*MAOA*	Monoamine oxidase A	-2.6	1.26E-04
A_32_P355396	NM_014844	*TECPR2*	Tectonin β-propeller repeat containing 2	-2.6	1.26E-04
A_32_P214565	BU928689	N/A		-2.6	1.26E-04

Table III. Continued.

Probe ID	Accession no.	Symbol	Gene name	Fold change (log)	P-value
A_24_P468950	AK021439	N/A		-2.6	1.26E-04
A_24_P683583	N/A	N/A		-2.6	1.26E-04
A_23_P203558	NM_000518	*HBB*	Hemoglobin, β	-2.6	2.33E-04
A_32_P140153	N/A	N/A		-2.6	1.26E-04
A_32_P124461	AK129743	N/A		-2.59	1.26E-04
A_23_P136026	AK128476	N/A		-2.59	1.26E-04
A_23_P28295	NM_004525	*LRP2*	Low density lipoprotein-related protein 2	-2.59	4.43E-04
A_24_P586712	NM_198485	*TPRG1*	Tumor protein p63 regulated 1	-2.58	1.26E-04
A_23_P139500	NM_030762	*BHLHE41*	Basic helix-loop-helix family, member e41	-2.58	1.26E-04
A_23_P121480	NM_001004196	*CD200*	CD200 molecule	-2.58	1.26E-04
A_23_P32577	NM_080759	*DACH1*	Dachshund homolog 1 (*Drosophila*)	-2.58	1.26E-04
A_23_P315815	NM_004495	*NRG1*	Neuregulin 1	-2.58	1.26E-04
A_23_P93772	NM_019102	*HOXA5*	Homeobox A5	-2.58	1.26E-04
A_32_P150748	CR749529	N/A		-2.58	1.26E-04
A_32_P204959	N/A	N/A		-2.58	1.26E-04
A_23_P363149	N/A	N/A		-2.57	4.43E-04
A_23_P41487	NM_015130	*TBC1D9*	TBC1 domain family, member 9 (with GRAM domain)	-2.57	1.26E-04
A_23_P257296	NM_003226	*TFF3*	Trefoil factor 3 (intestinal)	-2.56	3.41E-04
A_23_P250735	NM_175709	*CBX7*	Chromobox homolog 7	-2.56	1.26E-04
A_24_P189516	NM_001609	*ACADSB*	acyl-coenzyme A dehydrogenase, short/ branched chain	-2.56	1.26E-04
A_23_P253012	NM_017577	GRAMD1C	GRAM domain containing 1C	-2.56	1.26E-04
A_24_P179244	XM_001723863	*LOC100128979*	Hypothetical protein LOC100128979	-2.55	1.26E-04
A_32_P117846	N/A	N/A		-2.55	1.26E-04
A_32_P42224	BX097190	N/A		-2.55	2.33E-04
A_24_P119665	NM_001128933	*SYNPO2*	Synaptopodin 2	-2.54	1.26E-04
A_32_P105825	NM_001584	*MPPED2*	Metallophosphoesterase domain containing 2	-2.54	3.41E-04
A_24_P225679	NM_005544	*IRS1*	Insulin receptor substrate 1	-2.54	1.26E-04
A_32_P226907	N/A	N/A		-2.54	1.26E-04
A_23_P356581	NM_022370	*ROBO3*	Roundabout, axon guidance receptor, homolog 3 (*Drosophila*)	-2.53	1.26E-04
A_32_P221096	AW015426	N/A		-2.53	1.26E-04
A_23_P106016	NM_002742	*PRKD1*	Protein kinase D1	-2.52	1.26E-04
A_32_P210193	N/A	N/A		-2.52	1.26E-04
A_32_P38436	N/A	N/A		-2.52	1.26E-04
A_24_P512775	N/A	N/A		-2.52	1.26E-04
A_23_P151529	NR_023938	*C14orf132*	Chromosome 14 open reading frame 132	-2.52	1.26E-04
A_32_P235568	AK125221	N/A		-2.52	1.26E-04
A_23_P71270	NM_001185	*AZGP1*	α-2-glycoprotein 1, zinc-binding	-2.52	4.43E-04
A_24_P650425	N/A	N/A		-2.51	1.26E-04
A_23_P71328	NM_030583	*MATN2*	Matrilin 2	-2.51	2.33E-04
A_24_P153803	NM_020663	*RHOJ*	ras homolog gene family, member J	-2.51	1.26E-04
A_24_P912730	N/A	N/A		-2.51	1.26E-04
A_24_P347624	NM_022804	*SNURF*	SNRPN upstream reading frame	-2.5	1.26E-04
A_32_P52785	NM_015345	*DAAM2*	Dishevelled associated activator of morphogenesis 2	-2.5	3.41E-04
A_23_P61042	N/A	N/A		-2.5	1.26E-04

Table III. Continued.

Probe ID	Accession no.	Symbol	Gene name	Fold change (log)	P-value
A_23_P67661	NM_001864	*COX7A1*	Cytochrome c oxidase subunit VIIa polypeptide 1 (muscle)	-2.49	1.26E-04
A_23_P213486	N/A	*PARP8*	Poly(ADP-ribose) polymerase family, member 8	-2.49	1.26E-04
A_23_P18713	NM_004827	*ABCG2*	ATP-binding cassette, sub-family G (WHITE), member 2	-2.48	4.43E-04
A_23_P76658	NM_052818	*N4BP2L1*	NEDD4 binding protein 2-like 1	-2.48	1.26E-04
A_23_P96590	NM_014710	*GPRASP1*	G protein-coupled receptor associated sorting protein 1	-2.48	1.26E-04
A_24_P460763	AK022443	N/A		-2.48	1.26E-04
A_23_P85672	NM_006610	*MASP2*	Mannan-binding lectin serine peptidase 2	-2.48	1.26E-04
A_24_P416489	N/A	N/A		-2.47	1.26E-04
A_24_P321525	NM_032918	*RERG*	RAS-like, estrogen-regulated, growth inhibitor	-2.47	1.26E-04
A_24_P256526	BC005914	*SP2*	Sp2 transcription factor	-2.47	1.26E-04
A_24_P261417	NM_015881	*DKK3*	Dickkopf homolog 3 (*Xenopus laevis*)	-2.47	1.26E-04
A_23_P98369	NM_000829	*GRIA4*	Glutamate receptor, ionotrophic, AMPA 4	-2.47	1.26E-04
A_23_P6818	NM_020163	*SEMA3G*	Sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3G	-2.46	3.41E-04
A_32_P100379	N/A	N/A		-2.46	1.26E-04
A_23_P30163	NR_026804	*FLJ13197*	Hypothetical FLJ13197	-2.46	1.26E-04
A_24_P206328	NM_005020	*PDE1C*	Phosphodiesterase 1C, calmodulin-dependent 70 kDa	-2.46	1.26E-04
A_24_P93948	AB210045	N/A		-2.46	1.26E-04
A_32_P52414	N/A	N/A		-2.45	1.26E-04
A_23_P123228	NM_000111	*SLC26A3*	Solute carrier family 26, member 3	-2.45	1.26E-04
A_24_P666553	N/A	N/A		-2.45	1.26E-04
A_24_P916816	N/A	N/A		-2.44	1.26E-04
A_23_P134734	NM_017786	*GOLSYN*	Golgi-localized protein	-2.44	1.26E-04
A_24_P296772	NM_033256	*PPP1R14A*	Protein phosphatase 1, regulatory (inhibitor) subunit 14A	-2.43	1.26E-04
A_24_P267523	NM_144613	*COX6B2*	Cytochrome c oxidase subunit VIb polypeptide 2 (testis)	-2.43	1.26E-04
A_23_P133517	NM_002310	*LIFR*	Leukemia inhibitory factor receptor α	-2.43	1.26E-04
A_24_P787680	N/A	N/A		-2.43	1.26E-04
A_32_P52829	N/A	N/A		-2.43	3.41E-04
A_23_P162047	NM_015881	*DKK3*	Dickkopf homolog 3 (*Xenopus laevis*)	-2.43	1.26E-04
A_32_P185140	BX648171	*TPM1*	Tropomyosin 1 (α)	-2.43	1.26E-04
A_24_P319892	NM_198274	*SMYD1*	SET and MYND domain containing 1	-2.43	1.26E-04
A_24_P226322	NM_031469	*SH3BGRL2*	SH3 domain binding glutamic acid-rich protein like 2	-2.42	1.26E-04
A_23_P86012	NM_001017402	*LAMB3*	Laminin, β3	-2.42	1.26E-04
A_23_P62255	NM_005314	*GRPR*	Gastrin-releasing peptide receptor	-2.41	1.26E-04
A_24_P141520	N/A	N/A		-2.41	2.33E-04
A_23_P114883	NM_002023	*FMOD*	Fibromodulin	-2.41	1.26E-04
A_23_P300033	NM_006206	*PDGFRA*	Platelet-derived growth factor receptor, α polypeptide	-2.41	2.33E-04
A_24_P108311	NM_015277	*NEDD4L*	Neural precursor cell expressed, developmentally downregulated 4-like	-2.41	1.26E-04

Table III. Continued.

Probe ID	Accession no.	Symbol	Gene name	Fold change (log)	P-value
A_23_P345746	NM_199261	*TPTE*	Transmembrane phosphatase with tensin homology	-2.41	1.26E-04
A_23_P418083	NM_181714	*LCA5*	Leber congenital amaurosis 5	-2.41	1.26E-04
A_32_P208341	N/A	N/A		-2.41	1.26E-04
A_24_P930337	N/A	N/A		-2.41	1.26E-04
A_24_P915095	NM_017577	*GRAMD1C*	GRAM domain containing 1C	-2.4	1.26E-04
A_32_P4792	AK057820	N/A		-2.4	1.26E-04
A_24_P82032	NM_020663	*RHOJ*	ras homolog gene family, member J	-2.39	2.33E-04
A_23_P204296	NM_032918	*RERG*	RAS-like, estrogen-regulated, growth inhibitor	-2.38	1.26E-04
A_24_P920712	N/A	N/A		-2.38	2.33E-04
A_24_P401185	NM_001042784	*CCDC158*	Coiled-coil domain containing 158	-2.38	1.26E-04
A_32_P109604	XM_001715342	*LOC100132733*	Similar to FLJ00310 protein	-2.37	1.26E-04
A_24_P131173	NM_024709	*C1orf115*	Chromosome 1 open reading frame 115	-2.37	2.33E-04
A_24_P64241	NM_001012421	*ANKRD20A2*	Ankyrin repeat domain 20 family, member A2	-2.37	1.26E-04
A_32_P58437	N/A	N/A		-2.37	1.26E-04
A_24_P602348	N/A	N/A		-2.37	1.26E-04
A_24_P135856	NM_016124	*RHD*	Rh blood group, D antigen	-2.37	1.26E-04
A_23_P333038	NM_025145	*C10orf79*	Chromosome 10 open reading frame 79	-2.37	2.33E-04
A_23_P352266	NM_000633	*BCL2*	B-cell CLL/lymphoma 2	-2.36	1.26E-04
A_23_P207699	NM_016835	*MAPT*	Microtubule-associated protein tau	-2.36	1.26E-04
A_23_P392529	NR_027270	*C21orf81*	Ankyrin repeat domain 20 family, member A3 pseudogene	-2.36	1.26E-04
A_23_P904	NM_024603	*BEND5*	BEN domain containing 5	-2.36	1.26E-04
A_23_P115785	NM_145235	*FANK1*	Fibronectin type III and ankyrin repeat domains 1	-2.35	1.26E-04
A_32_P146844	N/A	N/A		-2.35	1.26E-04
A_23_P26865	NM_002470	*MYH3*	Myosin, heavy chain 3, skeletal muscle, embryonic	-2.35	1.26E-04
A_32_P100641	XM_001714998	*LOC100128139*	Hypothetical LOC100128139	-2.35	2.33E-04
A_24_P930727	AK091677	N/A		-2.35	1.26E-04
A_23_P406341	NM_001001936	*AFAP1L2*	Actin filament associated protein 1-like 2	-2.35	1.26E-04
A_24_P54863	NM_152400	*C4orf32*	Chromosome 4 open reading frame 32	-2.34	1.26E-04
A_23_P133120	NM_018342	*TMEM144*	Transmembrane protein 144	-2.34	1.26E-04
A_32_P86705	BC040577	N/A		-2.34	1.26E-04
A_24_P833256	N/A	N/A		-2.33	1.26E-04
A_23_P401106	NM_002599	*PDE2A*	Phosphodiesterase 2A, cGMP-stimulated	-2.33	1.26E-04
A_24_P102119	AF264623	N/A		-2.33	1.26E-04
A_23_P358714	NM_020775	*KIAA1324*	KIAA1324	-2.32	1.26E-04
A_32_P162494	N/A	N/A		-2.32	3.41E-04
A_23_P326931	NM_145170	*TTC18*	Tetratricopeptide repeat domain 18	-2.32	1.26E-04

N/A, not annotated; P-value, Benjamini-Hochberg false discovery rate of random permutation test; log fold change, between groups. Gene symbol, accession number and gene name were exported from GeneSpring (from the NCBI databases).

showing significant knockdown effects. FACS analysis revealed that depleting *ASPM* caused a cell cycle arrest at the G2/M phase in HCC1937 cells (si*EGFP*:si*ASPM*, 24.4:34.0%) at 2 days after transfection, and a subsequent increase in the

Table IV. Genes specifically expressed in TNBC, but not expressed in normal human vital organs.

Probe ID	Accession no.	Symbol	Gene name	Fold change (log)	P-value
A_23_P118834	NM_001067	*TOP2A*	Topoisomerase (DNA) IIα 170 kDa	4.76	1.26E-04
A_32_P119154	BE138567	N/A		4.75	1.26E-04
A_23_P35219	NM_002497	*NEK2*	NIMA (never in mitosis gene a)-related kinase 2	4.67	1.26E-04
A_23_P166360	NM_206956	*PRAME*	Preferentially expressed antigen in melanoma	4.64	1.26E-04
A_24_P332314	NM_198947	*FAM111B*	Family with sequence similarity 111, member B	4.63	1.26E-04
A_24_P413884	NM_001809	*CENPA*	Centromere protein A	4.59	1.26E-04
A_23_P68610	NM_012112	*TPX2*	TPX2, microtubule-associated, homolog (Xenopus laevis)	4.58	1.26E-04
A_23_P401	NM_016343	*CENPF*	Centromere protein F, 350/400 ka (mitosin)	4.44	1.26E-04
A_23_P57379	NM_003504	*CDC45L*	CDC45 cell division cycle 45-like (*S. cerevisiae*)	4.44	1.26E-04
A_23_P356684	NM_018685	*ANLN*	Anillin, actin binding protein	4.29	1.26E-04
A_23_P52017	NM_018136	*ASPM*	asp (abnormal spindle) homolog, microcephaly associated (*Drosophila*)	4.17	1.26E-04
A_32_P199884	NM_032132	*HORMAD1*	HORMA domain containing 1	4.13	2.33E-04
A_23_P259586	NM_003318	*TTK*	TTK protein kinase	4.09	1.26E-04
A_23_P200310	NM_017779	*DEPDC1*	DEP domain containing 1	4.08	1.26E-04
A_23_P115872	NM_018131	*CEP55*	Centrosomal protein 55 kDa	4.03	1.26E-04
A_24_P911179	NM_018136	*ASPM*	asp (abnormal spindle) homolog, microcephaly associated (*Drosophila*)	4.02	1.26E-04
A_24_P96780	NM_016343	*CENPF*	Centromere protein F, 350/400 ka (mitosin)	3.92	1.26E-04
A_24_P14156	NM_006101	*NDC80*	NDC80 homolog, kinetochore complex component (*S. cerevisiae*)	3.86	1.26E-04
A_23_P254733	NM_024629	*MLF1IP*	MLF1 interacting protein	3.85	1.26E-04
A_23_P74115	NM_003579	*RAD54L*	RAD54-like (*S. cerevisiae*)	3.84	1.26E-04
A_23_P50108	NM_006101	*NDC80*	NDC80 homolog, kinetochore complex component (*S. cerevisiae*)	3.84	1.26E-04
A_23_P155815	NM_022346	*NCAPG*	Non-SMC condensin I complex, subunit G	3.82	1.26E-04
A_23_P51085	NM_020675	*SPC25*	SPC25, NDC80 kinetochore complex component, homolog (*S. cerevisiae*)	3.81	1.26E-04
A_32_P62997	NM_018492	*PBK*	PDZ binding kinase	3.8	1.26E-04
A_23_P256956	NM_005733	*KIF20A*	Kinesin family member 20A	3.79	1.26E-04
A_23_P212844	NM_006342	*TACC3*	Transforming, acidic coiled-coil containing protein 3	3.78	1.26E-04
A_24_P254705	NM_020394	*ZNF695*	Zinc finger protein 695	3.76	1.26E-04
A_23_P432352	NM_001017978	*CXorf61*	Chromosome X open reading frame 61	3.73	1.26E-04
A_23_P48669	NM_005192	*CDKN3*	Cyclin-dependent kinase inhibitor 3	3.71	1.26E-04
A_23_P94571	NM_004432	*ELAVL2*	ELAV (embryonic lethal, abnormal vision, *Drosophila*)-like 2 (Hu antigen B)	3.67	1.26E-04
A_23_P150667	NM_031217	*KIF18A*	Kinesin family member 18A	3.64	1.26E-04
A_32_P68525	BC035392	N/A		3.58	1.26E-04
A_24_P319613	NM_002497	*NEK2*	NIMA (never in mitosis gene a)-related kinase 2	3.53	1.26E-04
A_23_P10385	NM_016448	*DTL*	Denticleless homolog (*Drosophila*)	3.53	1.26E-04
A_23_P94422	NM_014791	*MELK*	Maternal embryonic leucine zipper kinase	3.5	1.26E-04
A_23_P340909	BC013418	*SKA3*	Spindle and kinetochore associated complex subunit 3	3.48	1.26E-04
A_23_P124417	NM_004336	*BUB1*	Budding uninhibited by benzimidazoles 1 homolog (yeast)	3.47	1.26E-04
A_24_P257099	NM_018410	*HJURP*	Holliday junction recognition protein	3.43	1.26E-04

Table IV. Continued.

Probe ID	Accession no.	Symbol	Gene name	Fold change (log)	P-value
A_23_P74349	NM_145697	*NUF2*	NUF2, NDC80 kinetochore complex component, homolog (*S. cerevisiae*)	3.36	1.26E-04
A_24_P302584	NM_003108	*SOX11*	SRY (sex determining region Y)-box 11	3.36	4.43E-04
A_24_P68088	NR_002947	*TCAM1*	Testicular cell adhesion molecule 1 homolog (mouse)	3.35	2.33E-04
A_24_P366033	NM_018098	*ECT2*	Epithelial cell transforming sequence 2 oncogene	3.34	1.26E-04
A_23_P93258	NM_003537	*HIST1H3B*	Histone cluster 1, H3b	3.33	1.26E-04
A_23_P149668	NM_014875	*KIF14*	Kinesin family member 14	3.29	1.26E-04
A_23_P34325	NM_033300	*LRP8*	Low density lipoprotein receptor-related protein 8, apolipoprotein E receptor	3.28	1.26E-04
A_32_P56154	N/A	N/A		3.28	1.26E-04
A_23_P138507	NM_001786	*CDC2*	Cell division cycle 2, G1→S and G2→M	3.24	1.26E-04
A_23_P49972	NM_001254	*CDC6*	Cell division cycle 6 homolog (*S. cerevisiae*)	3.22	1.26E-04
A_24_P306896	XR_040656	*LOC283711*	Hypothetical protein LOC283711	3.22	1.26E-04
A_23_P44684	NM_018098	*ECT2*	Epithelial cell transforming sequence 2 oncogene	3.21	1.26E-04
A_23_P100344	NM_014321	*ORC6L*	Origin recognition complex, subunit 6 like (yeast)	3.2	1.26E-04
A_23_P163481	NM_001211	*BUB1B*	Budding uninhibited by benzimidazoles 1 homolog β (yeast)	3.17	1.26E-04
A_32_P87849	N/A	N/A		3.16	1.26E-04
A_24_P397107	NM_001789	*CDC25A*	Cell division cycle 25 homolog A (S. pombe)	3.15	1.26E-04
A_23_P209200	NM_001238	*CCNE1*	Cyclin E1	3.15	1.26E-04
A_32_P16625	N/A	N/A		3.15	1.26E-04
A_24_P37903	N/A	*LOX*	Lysyl oxidase	3.12	1.26E-04
A_24_P313504	NM_005030	*PLK1*	Polo-like kinase 1 (*Drosophila*)	3.07	1.26E-04
A_23_P252292	NM_006733	*CENPI*	Centromere protein I	3.04	1.26E-04
A_23_P161474	NM_182751	*MCM10*	Minichromosome maintenance complex component 10	2.99	1.26E-04
A_23_P253762	N/A	N/A		2.94	1.26E-04
A_24_P225534	NM_017821	*RHBDL2*	Rhomboid, veinlet-like 2 (*Drosophila*)	2.94	1.26E-04
A_24_P412088	NM_182751	*MCM10*	Minichromosome maintenance complex component 10	2.94	1.26E-04
A_32_P109296	NM_152259	*C15orf42*	Chromosome 15 open reading frame 42	2.91	1.26E-04
A_24_P76521	AK056691	*GSG2*	Germ cell associated 2 (haspin)	2.83	1.26E-04
A_23_P126212	NM_022111	*CLSPN*	Claspin homolog (*Xenopus laevis*)	2.83	1.26E-04
A_23_P60120	NM_031415	*GSDMC*	Gasdermin C	2.81	2.33E-04
A_24_P902509	NM_018193	*FANCI*	Fanconi anemia, complementation group I	2.8	1.26E-04
A_23_P155969	NM_014264	*PLK4*	Polo-like kinase 4 (*Drosophila*)	2.79	1.26E-04
A_32_P183218	NM_153695	*ZNF367*	Zinc finger protein 367	2.77	1.26E-04
A_23_P46118	NM_001821	*CHML*	Choroideremia-like (Rab escort protein 2)	2.76	2.33E-04
A_23_P327643	N/A	N/A		2.75	1.26E-04
A_23_P215976	NM_057749	*CCNE2*	Cyclin E2	2.72	2.33E-04
A_32_P151800	NM_207418	*FAM72D*	Family with sequence similarity 72, member D	2.7	1.26E-04
A_23_P34788	NM_006845	*KIF2C*	Kinesin family member 2C	2.7	1.26E-04
A_23_P133956	NM_002263	*KIFC1*	Kinesin family member C1	2.69	1.26E-04
A_23_P88630	NM_000057	*BLM*	Bloom syndrome, RecQ helicase-like	2.68	1.26E-04
A_24_P276102	NM_183404	*RBL1*	Retinoblastoma-like 1 (p107)	2.68	1.26E-04
A_23_P23303	NM_003686	*EXO1*	Exonuclease 1	2.67	1.26E-04
A_23_P88691	NM_000745	*CHRNA5*	Cholinergic receptor, nicotinic, α5	2.67	1.26E-04
A_32_P72341	NM_173084	*TRIM59*	Tripartite motif-containing 59	2.62	1.26E-04

Table IV. Continued.

Probe ID	Accession no.	Symbol	Gene name	Fold change (log)	P-value
A_24_P227091	NM_004523	*KIF11*	Kinesin family member 11	2.61	1.26E-04
A_23_P136805	NM_014783	*ARHGAP11A*	Rho GTPase activating protein 11A	2.6	1.26E-04
A_23_P63402	NM_013296	*GPSM2*	G-protein signaling modulator 2 (AGS3-like, *C. elegans*)	2.6	1.26E-04
A_23_P35871	NM_024680	*E2F8*	E2F transcription factor 8	2.58	1.26E-04
A_23_P207307	N/A	N/A		2.58	1.26E-04
A_24_P399888	NM_001002876	*CENPM*	Centromere protein M	2.58	1.26E-04
A_23_P155989	NM_022145	*CENPK*	Centromere protein K	2.57	1.26E-04
A_23_P411335	NM_152524	*SGOL2*	Shugoshin-like 2 (S. pombe)	2.54	1.26E-04
A_23_P43484	NM_058197	*CDKN2A*	Cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)	2.52	1.26E-04
A_32_P28704	N/A	N/A		2.52	1.26E-04
A_24_P351466	NM_020890	*KIAA1524*	KIAA1524	2.5	1.26E-04
A_24_P334248	NM_014996	*PLCH1*	Phospholipase C, eta 1	2.48	1.26E-04
A_23_P88331	NM_014750	*DLGAP5*	Discs, large (*Drosophila*) homolog-associated protein 5	2.47	1.26E-04
A_32_P31021	N/A	N/A		2.46	1.26E-04
A_23_P361419	NM_018369	*DEPDC1B*	DEP domain containing 1B	2.45	1.26E-04
A_23_P397341	NM_152341	*PAQR4*	Progestin and adipoQ receptor family member IV	2.42	1.26E-04
A_23_P140316	NM_001099652	*GPR137C*	G protein-coupled receptor 137C	2.42	1.26E-04
A_23_P217049	NM_014286	*FREQ*	Frequenin homolog (*Drosophila*)	2.41	2.33E-04
A_32_P35839	N/A	N/A		2.4	1.26E-04
A_24_P857404	NM_001093725	*MEX3A*	mex-3 homolog A (*C. elegans*)	2.4	1.26E-04
A_24_P323598	NM_001017420	*ESCO2*	Establishment of cohesion 1 homolog 2 (*S. cerevisiae*)	2.36	1.26E-04
A_23_P112673	NM_017975	*ZWILCH*	Zwilch, kinetochore associated, homolog (*Drosophila*)	2.33	1.26E-04
A_24_P296254	NM_014783	*ARHGAP11A*	Rho GTPase activating protein 11A	2.32	1.26E-04

N/A, not annotated; P-value, Benjamini-Hochberg false discovery rate of random permutation test; log fold change, between groups. Gene symbol, accession number and gene name were exported from GeneSpring (from the NCBI databases).

sub-G1 population (si*EGFP*:si*ASPM*, 9.86:43.68%) at 6 days (Fig. 5A). On the other hand, reduced *CENPK* expression resulted in an increase in the proportion of G0/G1 phase cells (si*EGFP*:si*CENPK*, 56.49:72.2%) in MDA-MB-231 after 2 days of transfection, and a subsequent increase in the sub-G1 population (si*EGFP*:si*CENPK*, 12.73:30.96%) at 6 days (Fig. 5B). Interestingly, we observed an enlarged size of HCC1937 cells, which was likely due to abnormal tubulin formation due to decreased *ASPM* expression (Fig. 5C, arrowheads). In addition, we observed a disruption in the structural integrity of tubulin in *CENPK*-depleted MDA-MB-231 cells (Fig. 5D, arrowheads), compared with those in si*EGFP*-transfected cells.

These results suggest that the absence of *ASPM* and *CENPK* caused an arrest in the G2/M and G0/G1 phases, respectively, and then induced cell death. Taken together, these findings strongly suggest that *ASPM* and *CENPK* have indispensable roles in cell proliferation and mitosis, especially in the G2/M and G0/G1 phases, in TNBC cells.

Discussion

TNBC patients do not benefit from endocrine therapy and trastuzumab. Conventional chemotherapy is currently the mainstay of systemic medical treatment, although TNBC patients have a worse outcome after chemotherapy than patients with other breast subtypes. In particular, because cytotoxic drugs often cause severe adverse effects, it is obvious that thoughtful selection of novel target molecules based on the detailed molecular mechanisms of TNBC carcinogenesis

Table V. Genes listed in cluster 1 and cluster 2.

No. of genes	Genes
Cluster 1 (enrichment score, 29.90)	
87	*BLM, CKS1B, CKS2, CHEK1, E2F1, E2F2, E2F8, FANCA, FANCI, H2AFX, HORMAD1, HJURP, MAD2L1, NDC80, NEK2, NUF2, OIP5, PBK, RAD51, RAD54L, SPC25, TPX2, TTK, ZWINT ZWILCH, ANLN, ASPM, AURKA, BIRC5, BUB1, BUB1B, CASC5, CDC25A, CDC6, CDCA2, CDCA5, CDCA8, CENPA, CENPF, CEP55, CHAF1B, SKA3, C13orf34, CIT, CLSPN, CCNA2, CCNB1, CCNE1, CCNE2, CDKN2A, CDKN2C, CDKN3, DSCC1, DLGAP5, ESCO2, EXO1, FAM83D, GSG2, INHBA, KIF11, KIF14, KIF18A, KIF18B, KIF20A, KIF23, KIF2C, KIFC1, LMNB1, MND1, NCAPG, NUSAP1, PTTG1, PLK1, PLK4, PKMYT1, PRC1, RBL1, SGOL2 SPAG5, STMN1, SMC4, TMSB15A, TOP2A, TACC3, TUBB3, UBE2C, UHRF1*
Cluster 2 (enrichment score, 6.43)	
45	*ADAMTS5, MAMDC2, SPARCL1, WIF1, AZGP1, APOD, FIGF, CHL1, CCL28, CXCL2, COL4A6, COL14A1, COL17A1, CNTNAP3, DKK3, DST, FGF1, FMOD, HS3ST4, IGJ, IL33, LAMA3, LAMAB, LTBP2, LIFR, LRP2, MASP2, MATN2, MGP, NTN4, NRG1, PTHLH, PI15, PLAT, PDGFA, PTN, PIGR, PIP, SCGB1D1, SCGB1D2, SCGB3A1, SEMA3G, STC2, THSD4, TFF3*

Genes enriched in cluster 1 and cluster 2 according to DAVID.

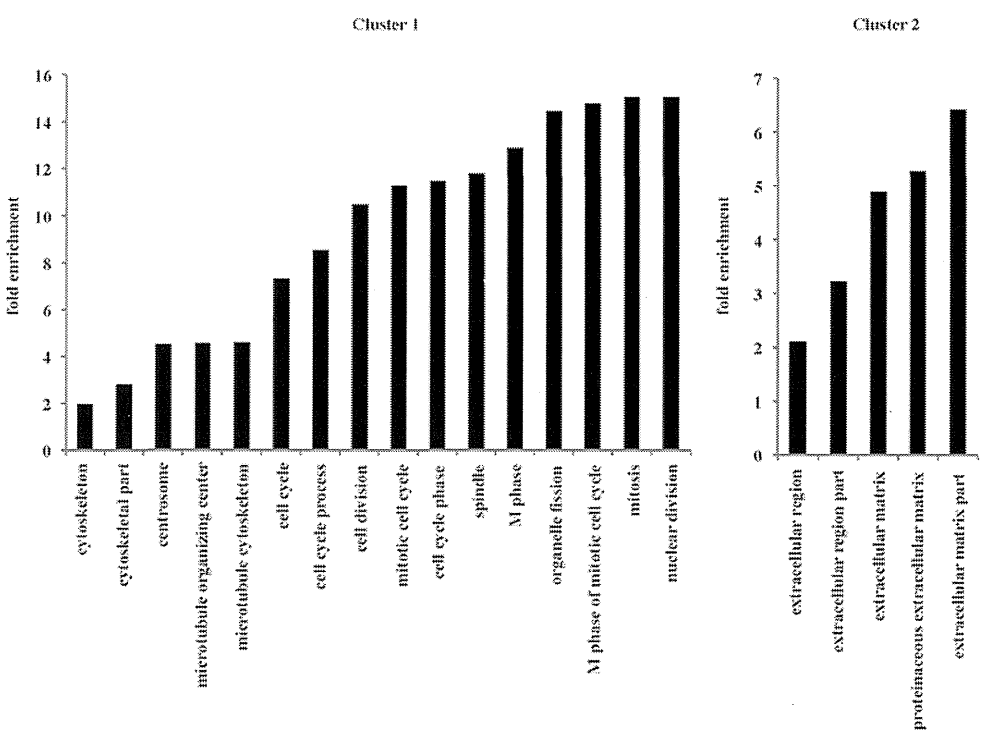

Figure 2. Gene annotation enrichment analysis based on DAVID was performed to elucidate the biological processes and pathways characterized in TNBC. Functional annotation terms are shown in bar plots; the value of the vertical axis represents the fold enrichment score of each term.

should be very helpful to develop effective anticancer drugs with a minimum risk of side effects. To this end, we performed DNA microarray using the microdissected TNBC and normal ductal cells, and normal human vital organs including the

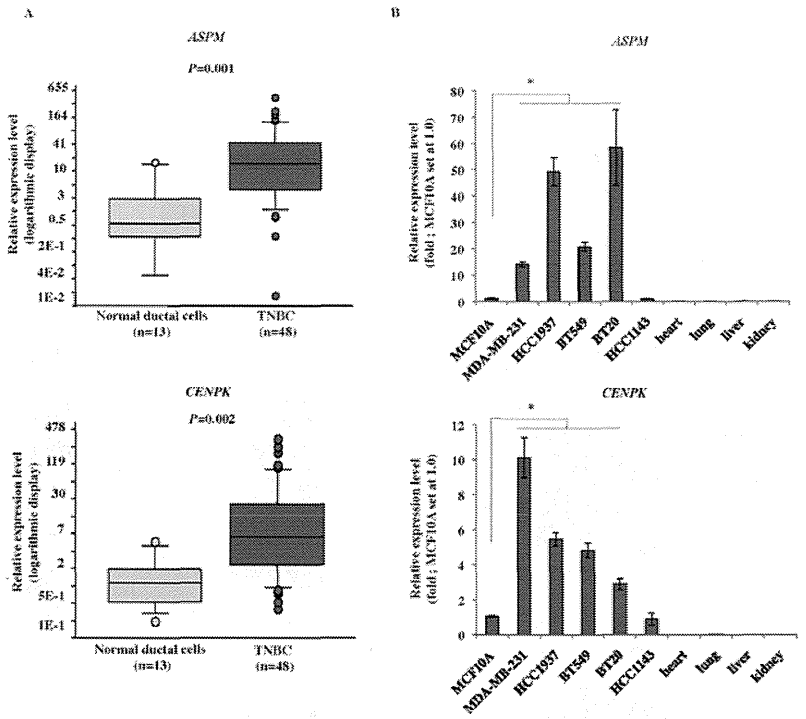

Figure 3. *ASPM* and *CENPK* expression profiles. (A) qRT-PCR results of *ASPM* and *CENPK* in microdissected tumor cells from 48 TNBC tissues and 13 normal ductal cells (Mann-Whitney t-test). (B) qRT-PCR results of *ASPM* and *CENPK* in five TNBC cell lines, MCF10A cells (human normal mammary epithelial cell line) and various normal organs (Student's two-sided t-test: $^{*}P<0.05$).

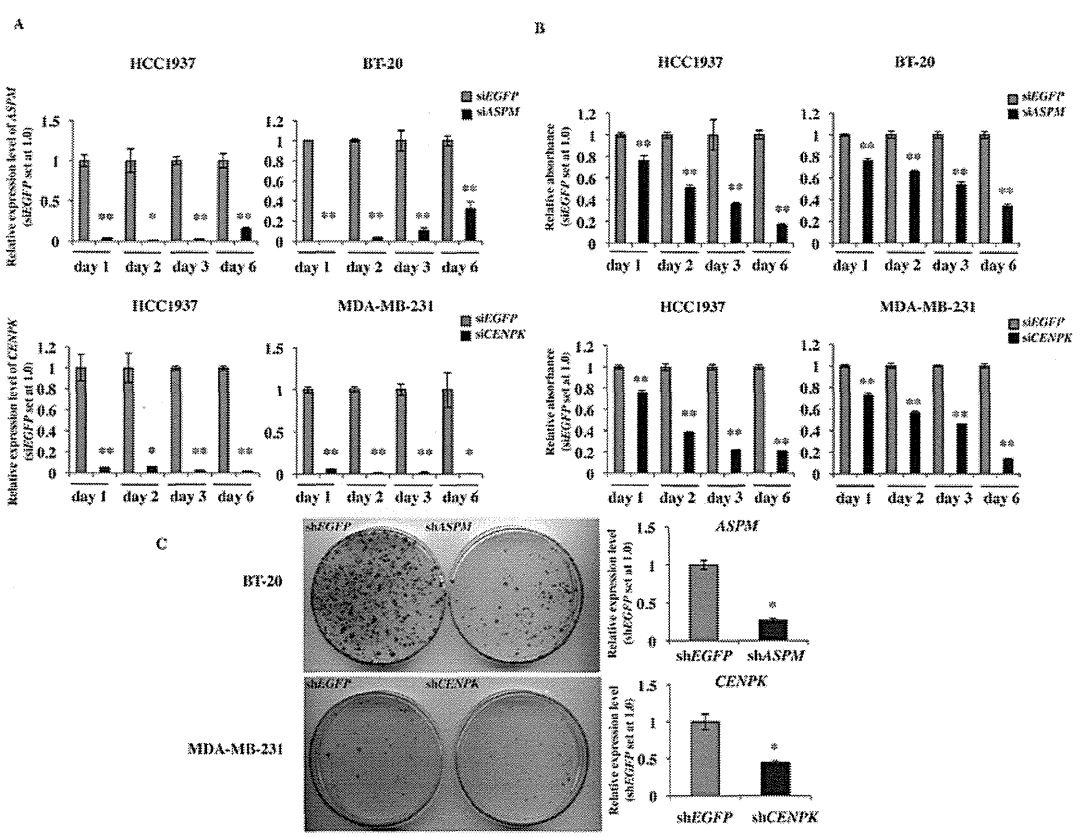

Figure 4. siRNA-mediated growth inhibitory effects in TNBC cells. (A) siRNA-mediated knockdown of *ASPM* in HCC1937 and BT-20 cells, and *CENPK* in HCC1937 and MDA-MB-231 cells was validated by qRT-PCR analysis (Student's two-sided t-test: $^{*}P<0.05$, $^{**}P<0.01$). (B) The MTT assay showing a decrease in the number of cells upon *ASPM* knockdown in HCC1937 and BT-20 cells and *CENPK* knockdown in HCC1937 and MDA-MB-231 cells (Student's two-sided t-test: $^{*}P<0.05$, $^{**}P<0.01$). (C) Colony formation assay (left) demonstrating a decrease in the number of colonies upon *ASPM* and *CENPK* knockdown (right) (Student's two-sided t-test: $^{*}P<0.05$).

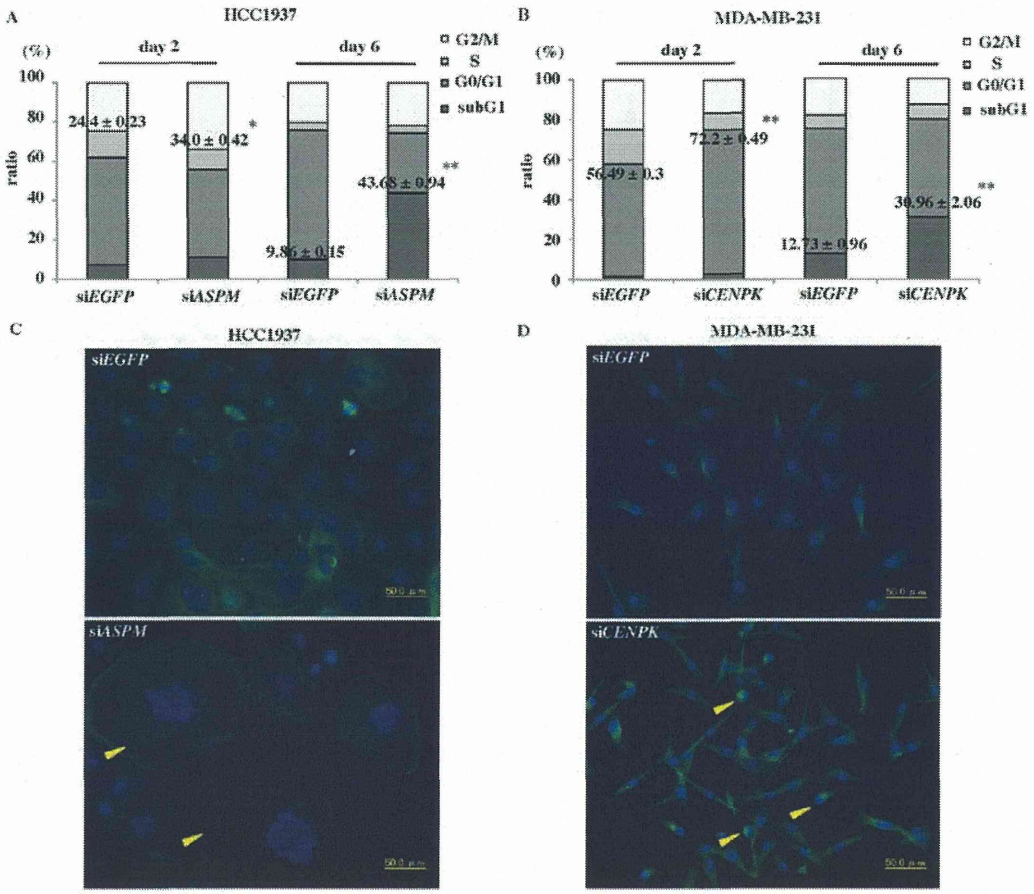

Figure 5. Alteration of the cell cycle and changes in cancer cell morphology upon *ASPM* and *CENPK* knockdown in TNBC cells. (A) FACS analysis at each time-point. The proportion of cells at the G2/M phase was elevated 2 days after si*ASPM* transfection followed by sub-G1 induction at 6 days in HCC1937 cells. (B) Upon *CENPK* knockdown, the proportion of cells at the G0/G1 phase was elevated in MDA-MB-231 cells at 2 days after si*CENPK* transfection, followed by sub-G1 induction at 6 days after transfection. A total of 10,000 cells were counted (Student's two-sided t-test: $^{*}P<0.05$, $^{**}P<0.01$). (C) Immunocytochemical staining analysis of α/β-tubulin at 48 h after siRNA transfection. Enlarged si*ASPM*-treated HCC1937 cells (arrowhead). Control cells that entered metaphase are indicated by the arrow. (D) Disruption of the structural integrity of tubulin in si*CENPK*-treated MDA-MB-231 cells (arrowhead). α/β-tubulin and nuclei staining are shown as green and blue, respectively. Scale bars, 50 μm.

heart, lung, liver and kidney and identified 104 genes that were significantly upregulated in TNBC compared to normal duct cells, but not expressed in normal human vital organs. They included cancer specific kinases, such as *NEK2*, *PBK*, and *MELK*, which might serve as druggable targets for new therapeutic agents against TNBC.

NEK2, a member of the NIMA-related serine/threonine kinase family, is involved in cell division and the mitotic regulation by centrosome splitting, and is upregulated in a wide variety of human cancers including breast cancer (40). siRNA-mediated depletion of *NEK2* expression results in growth suppression of breast and colorectal cancers (29,30). *PBK*, a mitotic serine/threonine kinase, is significantly upregulated in the majority of breast cancers. siRNA-mediated knockdown of PBK expression also results in significant suppression of cell growth due to cytokinetic failure (31). *MELK*, a member of the snf1/AMPK serine-threonine kinase family, is involved in mammalian embryonic development and is also frequently upregulated in breast cancers and brain tumors (33,41). Suppression of *MELK* expression by siRNA significantly inhibits the growth of human breast cancer cells (33). These findings strongly suggest that these cancer-specific kinases, *NEK2, PBK* and *MELK*, are promising therapeutic targets for TNBC.

Furthermore, we performed a gene-annotation enrichment analysis using DAVID based on gene expression profiling to elucidate the biological processes and pathways associated with each gene cluster. We found that the vast majority of genes upregulated in TNBC are functionally responsible for cell cycle progression involved in nuclear division, microtubule organization, kinetochore, and chromosome segregation, and that most inactivated functions closely related to TNBC progression are involved in cell-cell or cell-matrix interactions, which is consistent with epithelial mesenchymal transition (EMT) features as a phenotype of TNBC (42).

To further the development of novel anticancer drugs with minimum adverse effects, we focused on the cancer-specific cell-cycle associated genes *ASPM* and *CEPNK* as novel molecular targets for TNBC therapy. *ASPM* has been reported to play an essential role in nucleating microtubules at centrosomes, to localize to the spindle poles during mitosis (39) and to contribute to glioblastoma cell growth (43), but has not been associated with breast carcinogenesis, especially

TNBC. Here, we confirmed that *ASPM* is upregulated in clinical samples and TNBC cell lines (Fig. 3) and that siRNA-mediated knockdown of endogenous *ASPM* results in the loss of nucleating microtubules through mitosis by impeding centrosome function, resulting in G2/M cell cycle arrest and subsequent apoptosis. These results suggest that aberrant *ASPM* expression might be involved in the carcinogenesis of TNBC and that *ASPM* targeting might be an attractive therapeutic option with less adverse effects. *CENPK* is known to be a subunit of the *CENPH-I* complex, and essential for proper kinetochore assembly (39), but little is known about the roles of *CENPK* in human cancer growth, progression, and carcinogenesis. We also confirmed that *CENPK* is upregulated in clinical samples and TNBC cell lines, and that siRNA-mediated knockdown also causes cell growth inhibition through G0/G1 cell cycle arrest due to a loss of correct tubulin structures (Figs. 3-5). Interestingly, we determined that other centromere or kinetochore-associated proteins, *CENPA, CENPF, CENPI, CENPM, NDC80* and *HJURP*, were also significantly overexpressed in TNBC cases, but not expressed in normal vital organs (Fig. 1C and Table IV). Human *CENPA* was first identified based on autoantibodies found in patients suffering from scleroderma (44) and is over-expressed in colorectal cancers (45). *CENPF* is also reportedly upregulated in head and neck squamous cell carcinomas and pancreatic ductal carcinomas (46,47). *NDC80* and *HJURP* are reportedly overexpressed in breast cancers and associated with tumor grade and poor prognosis (48,49). These findings suggest that aberrant regulation of kinetochore assembly and centromere function through mitosis might contribute to the carcinogenesis of TNBC and that destroying one component of the kinetochore, such as targeting *CENPK*, might be a novel molecular target for TNBC treatment.

TNBC is a heterogeneous subgroup of breast cancers; therefore oncologists, pathologists, and geneticists had tried to clarify TNBC by means of gene expression profiling and immunohistochemical analyses. We also applied unsupervised 2-dimensional hierarchical clustering analysis to groups of genes based on similarities in the expression pattern, but there is no clustering for TNBC based on gene expression patterns, probably due to the small sample size (data not shown). However, the information provided in this study will facilitate the development of novel and attractive molecular drug targets without adverse events.

Acknowledgements

We thank Dr Tomoya Fukawa and Dr Le Tan Dat for helpful and constructive discussions and Ms. Hitomi Kawakami for technical assistance in microdissection. This work was supported in part by a grant from Health Labour Research Grant 'Third Term Comprehensive Control Research for Cancer (H24-3rd-Gan-Ippan-006), and Kobayashi Foundation for cancer Research (2009) (TK).

References

1. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T and Thun MJ: Cancer Statistics, 2008. CA Cancer J Clin 58: 71-96, 2008.
2. Di Cosimo S and Baselga J: Management breast cancer with targeted agents: importance of heterogeneity. Nat Rev Clin Oncol 7: 139-147, 2010.
3. Rahman M, Pumphrey JG and Lipkowitz S: The TRAIL to targeted therapy of breast cancer. Adv Cancer Res 103: 43-73, 2009.
4. Smith I, Procter M, Gelber RD, Guillaume S, Feyereislova A, Dowsett M, Goldhirsch A, Untch M, Mariani G, Baselga J, Kaufmann M, Cameron D, Bell R, Bergh J, Coleman R, Wardley A, Harbeck N, Lopez RI, Mallmann P, Gelmon K, Wilcken N, Wist E, Sánchez Rovira P and Piccart-Gebhart MJ: HERA study team: 2-year follow-up of trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer: a randomised controlled trial. Lancet 369: 29-36, 2007.
5. Romond EH, Perez EA, Bryant J, Suman VJ, Geyer CE Jr, Davidson NE, Tan-Chiu E, Martino S, Paik S, Kaufman PA, Swain SM, Pisansky TM, Fehrenbacher L, Kutteh LA, Vogel VG, Visscher DW, Yothers G, Jenkins RB, Brown AM, Dakhil SR, Mamounas EP, Lingle WL, Klein PM, Ingle JN and Wolmark N: Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. N Engl J Med 353: 1673-1684, 2005.
6. Joensuu H, Kellokumpu-Lehtinen PL, Bono P, Alanko T, Kataja V, Asola R, Utriainen T, Kokko R, Hemminki A, Tarkkanen M, Turpeenniemi-Hujanen T, Jyrkkiö S, Flander M, Helle L, Ingalsuo S, Johansson K, Jääskeläinen AS, Pajunen M, Rauhala M, Kaleva-Kerola J, Salminen T, Leinonen M, Elomaa I and Isola J, for the FinHer Study Investigators: Adjuvant docetaxel or vinorelbine with or without trastuzumab for breast cancer. N Engl J Med 354: 809-820, 2006.
7. Foulkes WD, Smith IE and Reis-Filho JS: Triple-negative breast cancer. N Engl J Med 363: 1938-1948, 2010.
8. Liedtke C, Mazouni C, Hess KR, André F, Tordai A, Mejia JA, Symmans WF, Gonzalez-Angulo AM, Hennessy B, Green M, Cristofanilli M, Hortobagyi GN and Pusztai L: Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. J Clin Oncol 26: 1275-1281, 2008.
9. Petricoin EF III, Hackett JL, Lesko LJ, Puri RK, Gutman SI, Chumakov K, Woodcock J, Feigal DW Jr, Zoon KC and Sistare FD: Medical applications of microarray technologies: a regulatory science perspective. Nat Genet 32: 474-479, 2002.
10. Huang DW, Sherman BT and Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nat Protoc 4: 44-57, 2009.
11. Huang DW, Sherman BT and Lempicki RA: Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37: 1-13, 2009.
12. Hao JM, Chen JZ, Sui HM, Si-Ma XQ, Li GQ, Liu C, Li JL, Ding YQ and Li JM: A five-gene signature as a potential predictor of metastasis and survival in colorectal cancer. J Pathol 220: 475-489, 2010.
13. Ueki T, Park JH, Nishidate T, Kijima K, Hirata K, Nakamura Y and Katagiri T: Ubiquitination and downregulation of BRCA1 by ubiquitin-conjugating enzyme E2T overexpression in human breast cancer cells. Cancer Res 69: 8752-8760, 2009.
14. Loussouarn D, Campion L, Leclair F, Campone M, Charbonnel C, Ricolleau G, Gouraud W, Bataille R and Jézéquel P: Validation of UBE2C protein as a prognostic marker in node-positive breast cancer. Br J Cancer 101: 166-173, 2009.
15. Arumugam T and Logsdon CD: S100P: a novel therapeutic target for cancer. Amino Acids 41: 893-899, 2011.
16. Xiang T, Li L, Yin X, Yuan C, Tan C, Su X, Xiong L, Putti TC, Oberst M, Kelly K, Ren G and Tao Q: The ubiquitin peptidase UCHL1 induces G0/G1 cell cycle arrest and apoptosis through stabilizing p53 and is frequently silenced in breast cancer. PLoS One 7: e29783, 2012.
17. Yoon CH, Kim MJ, Lee H, Kim RK, Lim EJ, Yoo KC, Lee GH, Cui YH, Oh YS, Gye MC, Lee YY, Park IC, An S, Hwang SG, Park MJ, Suh Y and Lee SJ: PTTG1 oncogene promotes tumor malignancy via epithelial to mesenchymal transition and expansion of cancer stem cell population. J Biol Chem 287: 19516-19527, 2012.
18. Jin W, Liu Y, Xu SG, Yin WJ, Li JJ, Yang JM and Shao ZM: UHRF1 inhibits MDR1 gene transcription and sensitizes breast cancer cells to anticancer drugs. Breast Cancer Res Treat 124: 39-48, 2010.
19. Ford HL, Landesman-Bollag E, Dacwag CS, Stukenberg PT, Pardee AB and Seldin DC: Cell cycle-regulated phosphorylation of the human SIX1 homeodomain protein. J Biol Chem 275: 22245-22254, 2000.

20. Shimo A, Nishidate T, Ohta T, Fukuda M, Nakamura Y and Katagiri T: Elevated expression of protein regulator of cytokinesis 1, involved in the growth of breast cancer cells. Cancer Sci 98: 174-181, 2007.
21. Di Leo A and Isola J: Topoisomerase II alpha as a marker predicting the efficacy of anthracyclines in breast cancer: are we at the end of the beginning? Clin Breast Cancer 4: 179-186, 2003.
22. Nakagawa M, Bando Y, Nagao T, Morimoto M, Takai C, Ohnishi T, Honda J, Moriya T, Izumi K, Takahashi M, Sasa M and Tangoku A: Expression of p53, Ki-67, E-cadherin, N-cadherin and TOP2A in triple-negative breast cancer. Anticancer Res 31: 2389-2393, 2011.
23. Adélaïde J, Finetti P, Bekhouche I, Repellini L, Geneix J, Sircoulomb F, Charafe-Jauffret E, Cervera N, Desplans J, Parzy D, Schoenmakers E, Viens P, Jacquemier J, Birnbaum D, Bertucci F and Chaffanet M: Integrated profiling of basal and luminal breast cancers. Cancer Res 67: 11565-11575, 2007.
24. Liu RZ, Graham K, Glubrecht DD, Germain DR, Mackey JR and Godbout R: Association of FABP5 expression with poor survival in triple-negative breast cancer: implication for retinoic acid therapy. Am J Pathol 178: 997-1008, 2011.
25. Kalashnikova EV, Revenko AS, Gemo AT, Andrews NP, Tepper CG, Zou JX, Cardiff RD, Borowsky AD and Chen HW: ANCCA/ATAD2 overexpression identifies breast cancer patients with poor prognosis, acting to drive proliferation and survival of triple-negative cells through control of B-Myb and EZH2. Cancer Res 70: 9402-9412, 2010.
26. Parris TZ, Danielsson A, Nemes S, Kovács A, Delle U, Fallenius G, Möllerström E, Karlsson P and Helou K: Clinical implications of gene dosage and gene expression patterns in diploid breast carcinoma. Clin Cancer Res 16: 3860-3874, 2010.
27. Ai L, Tao Q, Zhong S, Fields CR, Kim WJ, Lee MW, Cui Y, Brown KD and Robertson KD: Inactivation of Wnt inhibitory factor-1 (WIF1) expression by epigenetic silencing is a common event in breast cancer. Carcinogenesis 27: 1341-1348, 2006.
28. Cheng CJ, Lin YC, Tsai MT, Chen CS, Hsieh MC, Chen CL and Yang RB: SCUBE2 suppresses breast tumor cell proliferation and confers a favorable prognosis in invasive breast cancer. Cancer Res 69: 3634-3641, 2009.
29. Tsunoda N, Kokuryo T, Oda K, Senga T, Yokoyama Y, Nagino M, Nimura Y and Hamaguchi M: Nek2 as a novel molecular target for the treatment of breast carcinoma. Cancer Sci 100: 111-116, 2009.
30. Suzuki K, Kokuryo T, Senga T, Yokoyama Y, Nagino M and Hamaguchi M: Novel combination treatment for colorectal cancer using Nek2 siRNA and cisplatin. Cancer Sci 101: 1163-1169, 2010.
31. Park JH, Lin ML, Nishidate T, Nakamura Y and Katagiri T: PDZ-binding kinase/T-LAK cell-originated protein kinase, a putative cancer/testis antigen with an oncogenic activity in breast cancer. Cancer Res 66: 9186-9195, 2006.
32. Ueki T, Nishidate T, Park JH, Lin ML, Shimo A, Hirata K, Nakamura Y and Katagiri T: Involvement of elevated expression of multiple cell-cycle regulator, DTL/RAMP (denticleless/RA-regulated nuclear matrix associated protein), in the growth of breast cancer cells. Oncogene 27: 5672-5683, 2008.
33. Lin ML, Park JH, Nishidate T, Nakamura Y and Katagiri T: Involvement of maternal embryonic leucine zipper kinase (MELK) in mammary carcinogenesis through interaction with Bcl-G, a pro-apoptotic member of the Bcl-2 family. Breast Cancer Res 9: R17, 2007.
34. Shimo A, Tanikawa C, Nishidate T, Lin ML, Matsuda K, Park JH, Ueki T, Ohta T, Hirata K, Fukuda M, Nakamura Y and Katagiri T: Involvement of kinesin family member 2C/mitotic centromere-associated kinesin overexpression in mammary carcinogenesis. Cancer Sci 99: 62-70, 2008.
35. Chan SH, Yee Ko JM, Chan KW, Chan YP, Tao Q, Hyytiainen M, Keski-Oja J, Law S, Srivastava G, Tang J, Tsao SW, Chen H, Stanbridge EJ and Lung ML: The ECM protein LTBP-2 is a suppressor of esophageal squamous cell carcinoma tumor formation but higher tumor expression associates with poor patient outcome. Int J Cancer 129: 565-573, 2011.
36. Sathyanarayana UG, Maruyama R, Padar A, Suzuki M, Bondaruk J, Sagalowsky A, Minna JD, Frenkel EP, Grossman HB, Czerniak B and Gazdar AF: Molecular detection of noninvasive and invasive bladder tumor tissues and exfoliated cells by aberrant promoter methylation of laminin-5 encoding genes. Cancer Res 64: 1425-1430, 2004.
37. Senchenko VN, Krasnov GS, Dmitriev AA, Kudryavtseva AV, Anedchenko EA, Braga EA, Pronina IV, Kondratieva TT, Ivanov SV, Zabarovsky ER and Lerman MI: Differential expression of CHL1 gene during development of major human cancers. PLoS One 6: e15612, 2011.
38. do Carmo Avides M and Glover DM: Abnormal spindle protein, Asp, and the integrity of mitotic centrosomal microtubule organizing centers. Science 283: 1733-1735, 1999.
39. Cheeseman IM, Hori T, Fukagawa T and Desai A: KNL1 and the CENP-H/I/K complex coordinately direct kinetochore assembly in vertebrates. Mol Biol Cell 19: 587-594, 2008.
40. Hayward DG and Fry AM: Nek2 kinase in chromosome instability and cancer. Cancer Lett 237: 155-166, 2006.
41. Nakano I, Masterman-Smith M, Saigusa K, Paucar AA, Horvath S, Shoemaker L, Watanabe M, Negro A, Bajpai R, Howes A, Lelievre V, Waschek JA, Lazareff JA, Freije WA, Liau LM, Gilbertson RJ, Cloughesy TF, Geschwind DH, Nelson SF, Mischel PS, Terskikh AV and Kornblum HI: Maternal embryonic leucine zipper kinase is a key regulator of the proliferation of malignant brain tumors, including brain tumor stem cells. J Neurosci Res 86: 48-60, 2008.
42. Jeong H, Ryu YJ, An J, Lee Y and Kim A: Epithelial-mesenchymal transition in breast cancer correlates with high histological grade and triple-negative phenotype. Histopathology 60: E87-E95, 2012.
43. Horvath S, Zhang B, Carlson M, Lu KV, Zhu S, Felciano RM, Laurance MF, Zhao W, Qi S, Chen Z, Lee Y, Scheck AC, Liau LM, Wu H, Geschwind DH, Febbo PG, Kornblum HI, Cloughesy TF, Nelson SF and Mischel PS: Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target. Proc Natl Acad Sci USA 103: 17402-17407, 2006.
44. Moroi Y, Peebles C, Fritzler MJ, Steigerwald J and Tan EM: Autoantibody to centromere (kinetochore) in scleroderma sera. Proc Natl Acad Sci USA 77: 1627-1631, 1980.
45. Tomonaga T, Matsushita K, Yamaguchi S, Oohashi T, Shimada H, Ochiai T, Yoda K and Nomura F: Overexpression and mistargeting of centromere protein-A in human primary colorectal cancer. Cancer Res 63: 3511-3516, 2003.
46. de la Guardia C, Casiano CA, Trinidad-Pinedo J and Báez A: CENP-F gene amplification and overexpression in head and neck squamous cell carcinomas. Head Neck 23: 104-112, 2001.
47. Grützmann R, Pilarsky C, Ammerpohl O, Lüttges J, Böhme A, Sipos B, Foerder M, Alldinger I, Jahnke B, Schackert HK, Kalthoff H, Kremer B, Klöppel G and Saeger HD: Gene expression profiling of microdissected pancreatic ductal carcinomas using high-density DNA microarrays. Neoplasia 6: 611-622, 2004.
48. Bièche I, Vacher S, Lallemand F, Tozlu-Kara S, Bennani H, Beuzelin M, Driouch K, Rouleau E, Lerebours F, Ripoche H, Cizeron-Clairac G, Spyratos F and Lidereau R: Expression analysis of mitotic spindle checkpoint genes in breast carcinoma: role of NDC80/HEC1 in early breast tumorigenicity, and a two-gene signature for aneuploidy. Mol Cancer 10: 23, 2011.
49. Hu Z, Huang G, Sadanandam A, Gu S, Lenburg ME, Pai M, Bayani N, Blakely EA, Gray JW and Mao JH: The expression level of HJURP has an independent prognostic impact and predicts the sensitivity to radiotherapy in breast cancer. Breast Cancer Res 12: R18, 2010.

Journal of Human Genetics (2012), 1–6

www.nature.com/jhg

ORIGINAL ARTICLE

A genome-wide association study identifies a genetic variant in the *SIAH2* locus associated with hormonal receptor-positive breast cancer in Japanese

Seham Elgazzar[1], Hitoshi Zembutsu[1], Atsushi Takahashi[2], Michiaki Kubo[3], Fuminori Aki[4], Koichi Hirata[5], Yuichi Takatsuka[6], Minoru Okazaki[7], Shozo Ohsumi[8], Takashi Yamakawa[9], Mitsunori Sasa[10], Toyomasa Katagiri[11], Yoshio Miki[12] and Yusuke Nakamura[1]

In Japan, breast cancer is the most common cancer among women and the second leading cause of cancer death among women worldwide. To identify genetic variants associated with the disease susceptibility, we performed a genome-wide association study (GWAS) using a total of 1086 Japanese female patients with hormonal receptor-positive (HRP) breast cancer and 1816 female controls. We selected 33 single-nucleotide polymorphisms (SNPs) with suggestive associations in GWAS (*P*-value of $<1\times10^{-4}$) as well as 4 SNPs that were previously implicated their association with breast cancer for further replication by an independent set of 1653 cases and 2797 controls. We identified significant association of the disease with a SNP rs6788895 ($P_{combined}$ of 9.43×10^{-8} with odds ratio (OR) of 1.22) in the *SIAH2* (*intron of seven in absentia homolog 2*) gene on chromosome 3q25.1 where the involvement in estrogen-dependent diseases was suggested. In addition, rs3750817 in intron 2 of the *fibroblast growth factor receptor 2* gene, which was reported to be associated with breast cancer susceptibility, was significantly replicated with $P_{combined}$ of 8.47×10^{-8} with OR = 1.22. Our results suggest a novel susceptibility locus on chromosome 3q25.1 for a HRP breast cancer.
Journal of Human Genetics advance online publication, 6 September 2012; doi:10.1038/jhg.2012.108

Keywords: breast cancer in the Japanese population; *FGFR2* gene; GWAS; hormonal receptor-positive breast cancer; *SIAH2* gene; 3q25.1 locus; 10q26 locus

INTRODUCTION

Nearly 70% of breast cancer is known to be hormone dependent, as estrogen and progesterone have key roles both in the development and progression of the disease.[1,2] The exposures to higher level and/or for longer period of estrogen such as early menarche, late menopause, late age at first pregnancy, nulliparity, postmenopausal obesity and high serum estrogen level in postmenopausal women is considered to be risk factors for breast cancer.[3–5] Furthermore, progestin, synthetic progesterone, was shown to markedly increase the risk of breast cancer in postmenopausal women when this hormonal therapy was provided for >10 years.[6] In Japan, breast cancer is the most common cancer among women and its incidence has been doubled in both pre- and postmenopausal women in the last 20 years, mainly as an estrogen receptor-positive subgroup.[7] Although hormone therapy and radiotherapy are effective, cancer cells often become resistant to these treatments; nearly half of estrogen receptor-positive breast cancer patients at an advanced stage suffer from recurrence[8–10] and only one-third of hormonal receptor-positive (HRP) patients with metastatic disease respond to radiotherapy.[11] Therefore, new therapeutic options for the disease are eagerly awaited.

The aim of this study is to identify the genetic factors susceptible to HRP breast cancer in the Japanese population and should facilitate the development of novel approaches to prevent and/or treat breast cancer.

MATERIALS AND METHODS

Samples

Characteristics of study subjects are shown in Table 1. Most of the breast cancer cases and all the controls in this study were registered in the BioBank Japan, which begun in 2003 with the goal of collecting DNA and serum

[1]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan; [2]Laboratory for Statistical Analysis, RIKEN Center for Genomic Medicine, Yokohama, Japan; [3]Laboratory for Genotyping Development, RIKEN Center for Genomic Medicine, Yokohama, Japan; [4]Department of Surgery, Itoh Surgery and Breast Clinic, Kochi, Japan; [5]First Department of Surgery, Sapporo Medical University, Sapporo, Japan; [6]Department of Breast Surgery, Kansai Rosai Hospital, Hyogo, Japan; [7]Department of Surgery, Sapporo Breast Surgical Clinic, Sapporo, Japan; [8]Department of Breast Oncology, Shikoku Cancer Center, Ehime, Japan; [9]Department of Surgery, Yamakawa Breast Clinic, Kochi, Japan; [10]Department of Surgery, Tokushima Breast Care Clinic, Tokushima, Japan; [11]Division of Genome Medicine, Institute for Genome Research, The University of Tokushima, Tokushima, Japan and [12]Genome Center, The Cancer Institute, Japanese Foundation for Cancer Research, Tokyo, Japan
Correspondence: Dr Y Nakamura, Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan.
E-mail: yusuke@ims.u-tokyo.ac.jp
Received 3 April 2012; revised 5 August 2012; accepted 8 August 2012

Table 1 Characterization of samples used in hormonal receptor-positive breast cancer

	GWAS	*Replication*
Case		
Number of subjects	1086	1653
Mean age at interview (±s.d.)	66.7 (18.5)	60.7 (9.3)
Mean age of menarche	12.4	12.2
Mean age of menopause	48.3	47.9
Cases with DCIS	52	207
Cases with invasion	1034	1446
Body mass index prime	1.08	1.03
Platform	Illumina HumanHap 610K	Invader assay
Source	BioBank Japan Collaborative hospitals[a]	BioBank Japan Collaborative hospitals[a]
Control		
Number of subjects	1816	2797
Mean age at interview (±s.d.)	61.3 (12.6)	65.9 (13.2)
Body mass index prime	1.06	1.02
Platform	Illumina HumanHap 610K	Illumina HumanHap 610K
Source	BioBank Japan	BioBank Japan
Diseases in control[b]	MRC healthy volunteer Hepatitis B Keloid Drug eruption Pulmonary tuberculosis Peripheral artery disease Arrhythmias Stroke Myocardial infarction	Rheumatoid arthritis Amyotrophic lateral sclerosis Liver cirrhosis

Abbreviations: DCIS, ductal carcinoma in situ; GWAS, genome-wide association study.
[a]Tokushima Breast Care Clinic, Yamakawa Breast Clinic, Shikoku Cancer Center, and Itoh Surgery and Breast Clinic, Kansai Rosai Hospital, Sapporo Breast Surgical Clinic and Sapporo Medical University Hospital.
[b]The control groups from BioBank Japan consisted of female individuals without cancer also without any disease related to breast cancer.

samples, along with clinical information from 300 000 individuals who were diagnosed to have any of 47 different diseases from a collaborative network of 66 hospitals in Japan. All cases were diagnosed to have a HRP breast cancer by the following examinations: examination of breast tissue (biopsy or cytology), estrogen receptor and progesterone receptor positivities were evaluated by immunohistochemistry. For the genome-wide association study (GWAS) study, 1086 subjects with HRP breast cancer had been selected as cases (Table 1); 846 samples were collected from the BioBank Japan and the remaining 240 samples were collected from collaborative hospitals. Controls for the GWAS consisted of 1816 females including 231 healthy volunteers from the Midosuji Rotary Club, Osaka, Japan. In addition, we also used genome-wide screening data of 1585 female samples for 8 diseases registered in the BioBank Japan (Table 1). In the replication stage, 1547 cases were obtained from BioBank Japan and 105 cases from the collaborative hospitals. In all, 2797 female controls were registered in BioBank Japan and were genotyped in GWAS for other diseases (Table 1).

For re-sequencing analysis, we selected 2266 cases with HRP breast cancer from the BioBank Japan. We used 497 female controls with 4 diseases (hepatitis B, keloid, drug eruption and pulmonary tuberculosis) from the BioBank Japan as well as 231 healthy volunteers from the Midosuji Rotary Club, Osaka, Japan. All participating subjects provided written informed consent to participate in the study in accordance with the process approved by Ethical Committee at each of the Institute of Medical Science of the University of Tokyo and the Center for Genomic Medicine of RIKEN.

SNP genotyping

For the first stage, we genotyped 1086 female individuals with HRP breast cancer and 1816 female controls using the Illumina HumanHap 610 Genotyping BeadChip (Illumina, San Diego, CA, USA). We applied our single-nucleotide polymorphism (SNP) quality control standard (call rate of $\geqslant 0.99$ in both cases and controls, and Hardy–Weinberg equilibrium test of $P < 1.0 \times 10^{-6}$ in controls). A total of 453 627 SNPs on autosomal chromosomes and 10 525 SNPs on X chromosome passed the quality control filters and were further analyzed. All control samples for the replication stage were genotyped using the Illumina HumanHap 610 BeadChip (female samples of three diseases as controls). All cluster plots were checked by visual inspection by trained personnel, and SNPs with ambiguous calls were excluded. For cases in the replication study, we used the multiplex PCR-based Invader assay (Third Wave Technologies).[12] In addition, 22 variations resulted from re-sequencing analysis were selected and genotyped in 2266 cases and 728 female controls also using the multiplex PCR-based Invader assay (Third Wave Technologies, Madison, WI, USA).

Statistical analysis

Associations of SNPs were tested by employing the Cochran–Armitage trend test in both the GWA and replication stages. For the combined study, the simple combined method was applied. In the replication analyses, significance level was applied to be *P*-value of $< 1.35 \times 10^{-3}$ (calculated as 0.05/37) by Bonferroni correction. Odds ratios (ORs) and confidence intervals were calculated using the non-susceptible allele as a reference. Heterogeneity between the GWAS and replication sets was examined using the Breslow–Day test. The genomic inflation factor (λGC) was calculated from the median of the Cochran–Armitage trend test statistics. The quantile–quantile plot of the logarithms of the genome-wide *P*-values was generated by the 'snpMatrix' package in R program v2.10.0 (see URLs), and the Manhattan plot was generated using Haploview v4.1 (see URLs). Haplotype analysis was performed by the use of Haploview v4.1 by considering genotyped SNPs located within 500 kb upstream or downstream of the marker SNP. *In silico* prediction of functional consequences of SNP was done by the use of the SNP info web server (see URLs). (Haploview software was used to analyze linkage disequilibrium (LD) values, visualize haplotype.)

Imputation

Imputation was performed by referring to the genotype data of Japanese (JPT) individuals as deposited in the Phase II HapMap database using MACH v1.0 (see URLs). Genotypes of SNPs that are located in the genomic region within 500 kb upstream or downstream of the marker SNP (the SNP that showed the strongest association with HRP breast cancer) were imputed. In the process of imputation, 50 Markov chain iterations were implemented. Imputed SNPs with an imputation quality score of $r^2 < 0.3$ were excluded from the subsequent analysis.

Re-sequencing analysis

Initially, we carried out SNP discovery by using DNA samples of 96 cases with HRP breast cancer. We designed 98 sets of primers (Supplementary Table 1) using the genomic sequence information from UCSC Genome Bioinformatics data base (NM_005067) to amplify the 22 353 bps (two exons, one intron, 5′-UTR and 3′-UTR) of the genomic region corresponding to the *SIAH2* (*intron of seven in absentia homolog 2*) gene. For each of the 96 DNA samples, PCRs were performed by using GeneAmp PCR system 9700 (Applied Biosystems, Foster City, CA, USA). We performed direct sequencing of the PCR products with the 96-capillary 3730 × l DNA Analyzer (Applied Biosystems) with Big Dye Terminators (Applied Biosystems) according to standard protocols. All amplified fragments were sequenced by two pairs of sequencing primers. Then SNPs were detected by Sequecher software v4.8 (Gene Codes, Ann Arbor, MI, USA).

RESULTS

To identify genetic variants susceptible to HRP breast cancer in the Japanese population, we performed a GWAS using 1086 female patients and 1816 female controls with Illumina HumanHap 610k BeadChip (Table 1). After the quality check of SNP genotyping data, a total of 453 627 SNPs were selected for further analysis. Principal component analysis revealed that all the subjects participating in this study were clustered in the Hapmap Asian population (Supplementary Figure 1S). A quantile–quantile plot for this GWAS is shown in Supplementary Figure 2S. The genomic inflation factor (λGC) of the test statistic in this study was 1.053 indicating a very low possibility of false-positive associations resulted from the population stratification. Although no SNP achieved genome-wide significance level, 46 SNPs in various chromosomes showed suggestive association (P-values $< 1 \times 10^{-4}$) as illustrated in Figure 1.

Among these 46 SNPs, we excluded SNPs possessing strong LD ($r^2 > 0.8$) and selected 33 SNPs for replication analysis as well as 4 additional SNPs that were previously reported their association with breast cancer and showed P-value of $< 1.0 \times 10^{-2}$ in GWAS analysis, using an independent set of 1653 female patients and 2797 female controls. Among 37 SNPs analyzed in the replication study, an SNP rs6788895 was successfully replicated with the P-value of $< 1.35 \times 10^{-3}$ even after the Bonferroni correction (0.05/37) as shown in Table 2 and Supplementary Table 2S. Combined analysis of the results of the GWAS and the replication study suggested strong association of the locus of the *SIAH2* gene on chromosome 3q25.1 (rs6788895, P_{combined} of 9.43×10^{-8} with OR of 1.22, 95% confidence interval 1.13–1.31) without any significant heterogeneity between the two studies ($P_{\text{heterogeneity}} = 2.33 \times 10^{-01}$).

The SNP rs6788895 was further examined its association with the subgroups of breast cancer, an invasive papilloductal breast cancer group and a HER2-negative breast cancer group, and found significant associations with them ($P_{\text{combined}} = 3.61 \times 10^{-07}$, 6.78×10^{-06}, OR = 1.23, 1.21, respectively) although they did not reach to the genome-wide significant level (Supplementary Table 3S). Imputation analysis of this locus identified nine additional SNPs in strong LD (r^2 of > 0.8) that showed similar levels of association with rs6788895 (Figure 2a). The subsequent logistic regression analysis revealed no significant association of these nine SNPs when we accounted the effect of SNP rs6788895. The haplotype analysis found no haplotype revealing stronger association than the single SNP (Supplementary Table 4S). Although *in silico* prediction of the functional effect of rs6788895 identified no possible biological effect, one SNP rs2018246 showing strong LD with rs6788895 ($r^2 = 0.94$), which was located about 0.7 kb upstream from the transcription initiation site of *SIAH2*, was indicated to be present within the binding site of multiple transcription factors such as STAT1, LEF1, PAX2, which were reported to have some implication to breast cancer.[13–16] The re-sequencing of 22 353 bps corresponding to the *SIAH2* gene identified 10 novel genetic variations in addition to 37 genetic variations reported previously. We further genotyped 22 of the 47 variations after the exclusion of SNPs showing strong LD with the marker SNP (r^2 of > 0.8). As a result, we identified no genetic variant showing significant association in HRP breast cancer (Supplementary Table 5S and Supplementary Table 6S)

Furthermore, we examined the association of 37 previously reported SNPs with the HRP breast cancer[17–26] using our sample sets (Supplementary Table 7S) and found very moderate association of four genetic variants, rs1292011, rs3803662, rs2981579 and rs3750817, with HRP breast cancer in the GWAS phase ($P_{GWAS} = 5.89 \times 10^{-02}$, 6.95×10^{-03}, 8.68×10^{-04} and 5.03×10^{-04}, respectively). Further analysis of these four SNPs identified significant

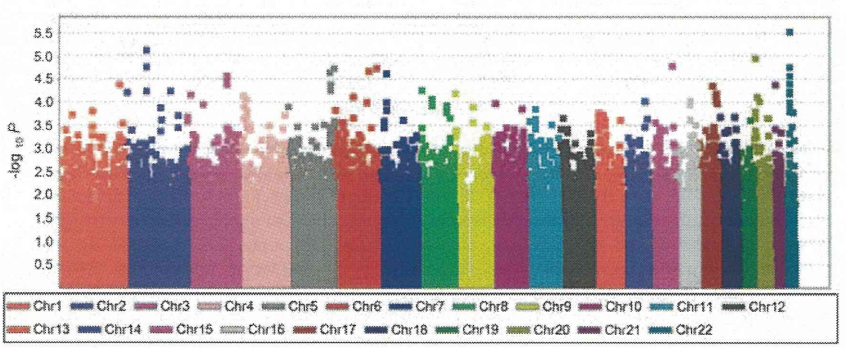

Figure 1 Manhattan plot for the genome-wide association study (GWAS) of hormonal receptor-positive breast cancer indicating −log10*P* of the Cochran–Armitage trend test for 453 627 single-nucleotide polymorphisms (SNPs) plotted against their respective positions on each chromosome.

Table 2 Association of SNP rs6788895 on chromosome 3q25.1 with hormonal receptor-positive breast cancer

					Case				*Control*							
Chr.	*Chrloc.*	*SNP*	*RA*	*Stage*	*11*	*12*	*22*	*RAF*	*11*	*12*	*22*	*RAF*	P_{assoc}[a]	*OR*	*(95% CI)*	*Phet*[b]
3	151950498	rs6788895	G	GWAS	106	456	524	0.69	242	832	742	0.64	2.34E–05	1.28	(1.14–1.43)	
				Rep	164	694	786	0.69	337	1265	1195	0.65	5.77E–04	1.18	(1.07–1.29)	2.33E–01
				Combined	270	1150	1310	0.69	579	2097	1937	0.65	9.43E–08	1.22	(1.13–1.31)	

Abbreviations: Chr., chromosome; chrloc., chromosomal location (bp); CI, confidence interval; GWAS, genome-wide association study; OR, odds ratio (calculated based on the risk allele); RA, risk allele; RAF, risk allele frequency; Rep, replication; SNP, single-nucleotide polymorphism; 11, homozygous non-risk genotype; 12, heterozygous genotype; 22, homozygous risk genotype.
[a]Passoc, *P*-value for the GWAS and replication study obtained from the Cochran–Armitage trend test and *P*-value for the combined study obtained from the simple combined test.
[b]Phet, *P*-value for heterogeneity test obtained from the Breslow–Day test.

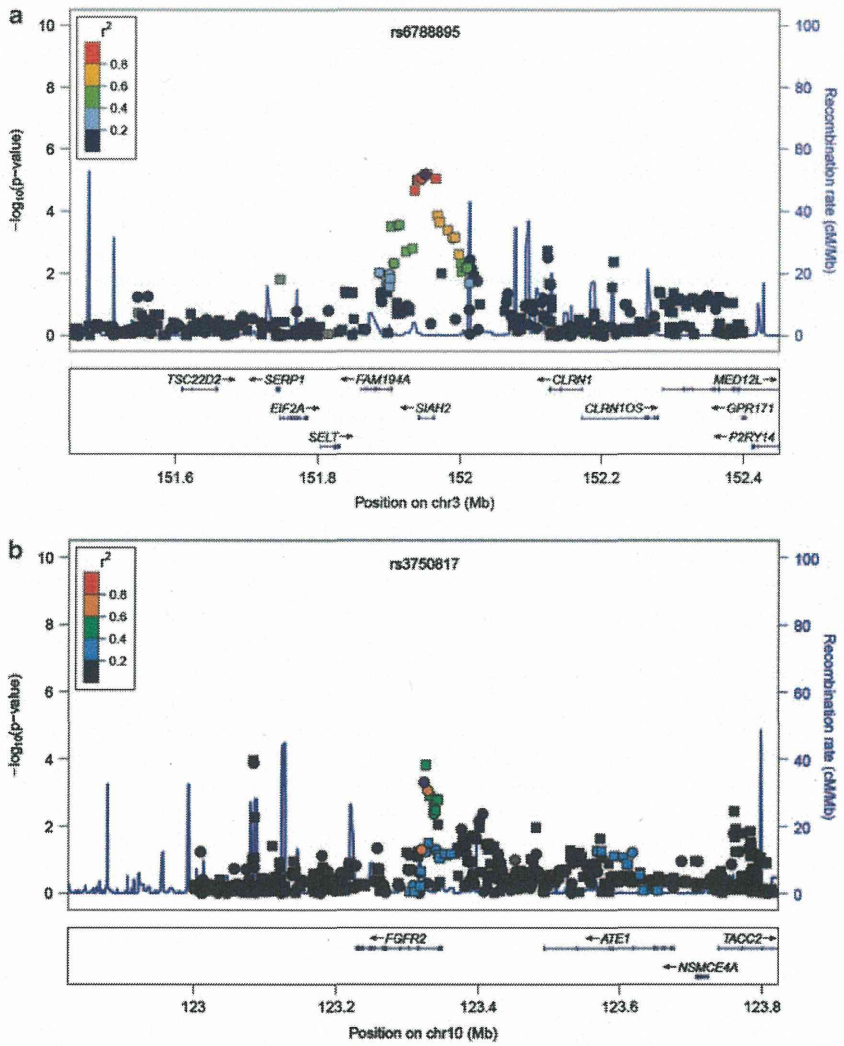

Figure 2 (**a**) Regional association plots of the locus associated with hormonal receptor-positive breast cancer on chromosomes 3q25.1 (*intron of seven in absentia homolog 2* (*SIAH2*)). (**b**) Regional association plots of the locus associated with hormonal receptor-positive breast cancer on chromosomes 10q26 (*fibroblast growth factor receptor 2* (*FGFR2*)). For each plot, $-\log_{10}P$ of the Cochran–Armitage trend test of single-nucleotide polymorphisms (SNPs) in the genome-wide association study (GWAS) was plotted against relative chromosomal locations. The square and rounded signs represent imputed and genotyped SNPs, respectively. All SNPs are color coded as red ($r^2 = 0.8$–1.0), orange ($r^2 = 0.6$–0.8), green ($r^2 = 0.4$–0.6), light blue ($r^2 = 0.4$–0.6), and dark blue ($r^2 < 0.2$) according to their pair wise r^2 to the marker SNP. The marker SNP is represented in purple color. SNP positions followed NCBI build 36 coordinates. Estimated recombination rates (cM/Mb) are plotted as a blue line.

replication of two SNPs, rs3750817 ($P_{\text{replication}} = 5.39 \times 10^{-5}$, OR $= 1.22$) and rs2981579 ($P_{\text{replication}} = 1.21 \times 10^{-3}$, OR $= 1.20$). Both SNPs are located within intron 2 of the fibroblast growth factor receptor 2 *(FGFR2)* genes. The combined analysis of the GWAS and replication phases of rs3750817 revealed strong association with $P_{\text{combined}} = 8.47 \times 10^{-08}$ (OR $= 1.22$) and that of rs2981579 was 1.77×10^{-06} (OR $= 1.20$) (Table 3). Imputation analysis of this locus identified three additional SNPs, rs9420318, rs11199914 and rs10736303 that showed similar levels of association with rs3750817 (Figure 2b).

DISCUSSION

We reported here GWA and replication studies using a total of 2730 female breast cancer cases and 4613 female controls in the Japanese population to identify common genetic variants susceptible to the HRP breast cancer. The SNP rs6788895 located in the intronic region of the *SIAH2* gene on chromosome 3q25.1 revealed a significant association with the HRP breast cancer (P_{combined} of 9.43×10^{-08} with OR of 1.22, 95% confidence interval of 1.13–1.31). We further examined the association of rs6788895 with the subgroups of breast cancer. The analysis of two histological subgroups, an invasive papilloductal breast cancer group and a HER2-negative breast cancer group, indicated suggestive associations with P_{combined} of 3.61×10^{-07} (OR $= 1.24$) and with P_{combined} of 6.78×10^{-06} (OR $= 1.21$), respectively (Supplementary Table 3S). However, rs6788895 showed no association in the GWAS with the hormonal receptor-negative group (P_{trend} of 1.03×10^{-01}) or with the HER2-positive breast cancer group (P_{trend} of 1.15×10^{-01}).

For further characterization of the chromosome 3q25.1 locus, we imputed genotypes of SNPs that were not genotyped in the GWAS

Table 3 rs2981579 and rs3750817 in different population

SNPs	Minor/major allele	MAF	OR	P-trend	Population
rs2981579 (FGFR2)	A/G	0.42	1.43	3.60×10^{-31}	UK[20]
rs2981579	A/G	0.44	1.31	2.60×10^{-09}	American[25]
rs2981579	A/G	0.47	1.20	1.77×0^{-06}	Japanese
rs3750817 (FGFR2)	T/C	0.49	1.22	8.47×10^{-08}	Japanese
rs3750817	T/C	0.37	0.78	8.20×10^{-08}	American[25]

Abbreviations: FGFR2, fibroblast growth factor receptor 2; MAF, minor allele frequency; OR, odds ratio (calculated based on the non susceptible allele) except rs3750817 in American population OR, calculated based on the susceptible allele); SNP, single-nucleotide polymorphism.

and then examined their associations with HRP breast cancer, but found no SNP showing stronger association than the marker SNP rs6788895 although several SNPs having strong LD with rs6788895 ($r^2 > 0.8$) showed similar levels of associations (Figure 2a). Previous reports implicated possible roles of *SIAH2* in breast carcinogenesis and described that SIAH2 expression was highly associated with estrogen receptor levels.[9,27–29] In addition, SIAH2 protein was indicated to have an essential role in the hypoxic response by regulating the hypoxia-inducible factor-α.[30]

Moreover, *SIAH2* was known to induce ubiquitin-mediated degradation of many substrates, including proteins involved in transcriptional regulation (POU2AF1, PML and NCOR1), a cell surface receptor (DCC) and an anti-apoptotic protein (BAG1). These proteins were reported to have some relations to breast cancer by different mechanisms.[31–35] Recent genetic studies showed that the chromosome 3q25.1 region might have a critical role in some estrogen-dependent diseases such as development of peritoneal leioyomatosis.[36,37]

We also examined the association of previously reported loci with the breast cancer[17–26] using our sample sets and found very moderate association of four genetic variants in our GWAS. Further analysis of these four SNPs identified significant replication of two SNPs, rs3750817 and rs2981579 ($P_{combined} = 8.47 \times 10^{-8}$ and 1.77×10^{-06} with OR = 1.22 and OR = 1.20, respectively). A T allele for rs3750817 is a protective allele for both Japanese and American populations with comparable ORs (Table 3).

For characterization of the chromosome 10q26 locus, we imputed genotypes of SNPs that were not genotyped in the GWAS, and examined the associations of these SNPs with HRP breast cancer. As a result, three additional SNPs, rs9420318, rs11199914 and rs10736303 were found to have similar levels of association with rs3750817 (Figure 2b). The most strongly associated SNPs are located in intron 2 of the *FGFR* gene. The intron 2 region contains a highly conserved region and possess the transcription factor binding sites possibly related to the estrogen receptor signaling pathway.[38] FGFR2 encodes a receptor tyrosine kinase and has an important role in human mammary epithelial-cell transformation,[39,40] suggesting that FGFR2 is a good candidate for breast cancer susceptibility. Subsequent functional analyses are thus essential to pinpoint the causal variants and genes associated with HRP breast cancer. In addition, because breast cancer is multi factorial disease, we could not exclude the possibility that some subjects with undiagnosed early stage of cancers or undiagnosed hormonal-dependent diseases or subject have diseases related to breast cancer might have been included as controls. Hence, this study might not have sufficient power to detect SNPs having very modest effects on susceptibility to HRP breast cancer. In conclusion, our findings, the verification of the association of the FGFR2 to the risk of breast cancer in the Japanese population and the novel identification of significant association of genetic variations in the *SIAH2* gene, should contribute to the better understanding of the susceptibility to HRP breast cancer.

URLS

The Leading Project for Personalized Medicine, http://biobankjp.org/;
EIGENSTRATsoftwarev2.0, http://genepath.med.harvard.edu/~reich/Software.htm;
R project v2.10.0, http://www.r-project.org/;
Haploview v4.1, http://www.broadinstitute.org/haploview/haploview;
MACH v1.0, http://www.sph.umich.edu/csg/yli/mach/index.html;
PLINK statistical software v1.06, http://pngu.mgh.harvard.edu/~purcell/plink/;
SNP info web server, http://manticore.niehs.nih.gov/index.html.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We extend our heartfelt gratitude to all individuals who kindly participated in the current study. We also thank the BioBank Japan for providing DNA samples and clinical information and the members of the Rotary Club of Osaka-Midosuji District 2660 Rotary International in Japan for supporting this study. We extend our appreciation to Dr Siew-Kee Low for her useful discussion in the statistical analysis and Dr Osman Wael Mohammed for his insightful advices. Finally, we would like to express our gratefulness to technical staff of the Laboratory for Genotyping Development in RIKEN, Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, the University of Tokyo and technical staff of collaborative hospitals for their assistances.

1 Massarweh, S. & Schiff, R. Resistance to endocrine therapy in breast cancer: exploiting estrogen receptor/growth factor signaling crosstalk. *Endocr. Relat. Cancer* **13,** S15–S24 (2006).
2 Xu, W. -H., Shu, X. O., Long, J., Lu, W., Cai, Q., Zheng, Y. *et al.* Relation of FGFR2 genetic polymorphisms to the association between oral contraceptive use and the risk of breast cancer in Chinese women. *Am. J. Epidemiol.* **173,** 923–931 (2011).
3 Hu, Y.-H., Nagata, C., Shimizu, H., Kaneda, N. & Kashiki, Y. Association of body mass index, physical activity, and reproductive histories with breast cancer: a case-control study in Gifu, Japan. *Breast Cancer Res. Treat.* **43,** 65–72 (1997).
4 Tung, H. T., Tsukuma, H., Tanaka, H., Kinoshita, N., Koyama, Y., Ajiki, W. *et al.* Risk factors for breast cancer in Japan, with special attention to anthropometric measurements and reproductive history. *Jpn. J. Clin. Oncol.* **29,** 137–146 (1999).
5 Miyoshi, Y., Tanji, Y., Taguchi, T., Tamaki, Y. & Noguchi, S. Association of serum estrone levels with estrogen receptor-positive breast cancer risk in postmenopausal Japanese women. *Clin. Cancer Res.* **9,** 2229–2233 (2003).
6 Colditz, G. A. & Rosner, B. Cumulative risk of breast cancer to age 70 years according to risk factor status: data from the Nurses' Health Study. *Am. J. Epidemiol.* **152,** 950–964 (2000).
7 Yamashita, H., Iwase, H., Toyama, T., Takahashi, S., Sugiura, H., Yoshimoto, N. *et al.* Estrogen receptor-positive breast cancer in Japanese women: trends in incidence, characteristics, and prognosis. *Ann. Oncol.* **22,** 1318–1325 (2011).
8 Ao, A., Morrison, B. J., Wang, H., López, J. A., Reynolds, B. A. & Lu, J. Response of estrogen receptor-positive breast cancer tumorspheres to antiestrogen treatments. *PLoS ONE* **6,** e18810 (2011).
9 Jansen, M., Ruigrok-Ritstier, K., Dorssers, L. C., van Staveren, I. L., Look, M. P., Meijer-van Gelder, M. E. *et al.* Downregulation of SIAH2, an ubiquitin E3 ligase, is associated with resistance to endocrine therapy in breast cancer. *Breast Cancer Res. Treat.* **116,** 263–271 (2009).
10 Kiyotani, K., Mushiroda, T., Tsunoda, T., Morizono, T., Hosono, N., Kubo, M. *et al.* A genome-wide association study identifies locus at 10q22 associated with clinical outcomes of adjuvant tamoxifen therapy for breast cancer patients in Japanese. *Hum. Mol. Genet.* **21,** 1665–1672 (2012).
11 Prat, A. & Baselga, J. The role of hormonal therapy in the management of hormonal-receptor-positive breast cancer with co-expression of HER2. *Nat. Clin. Prac. Oncol.* **5,** 531–542 (2008).
12 Ohnishi, Y., Tanaka, T., Ozaki, K., Yamada, R., Suzuki, H. & Nakamura, Y. A high-throughput SNP typing system for genome-wide association studies. *J. Hum. Genet.* **46,** 471–477 (2001).