Figure 1

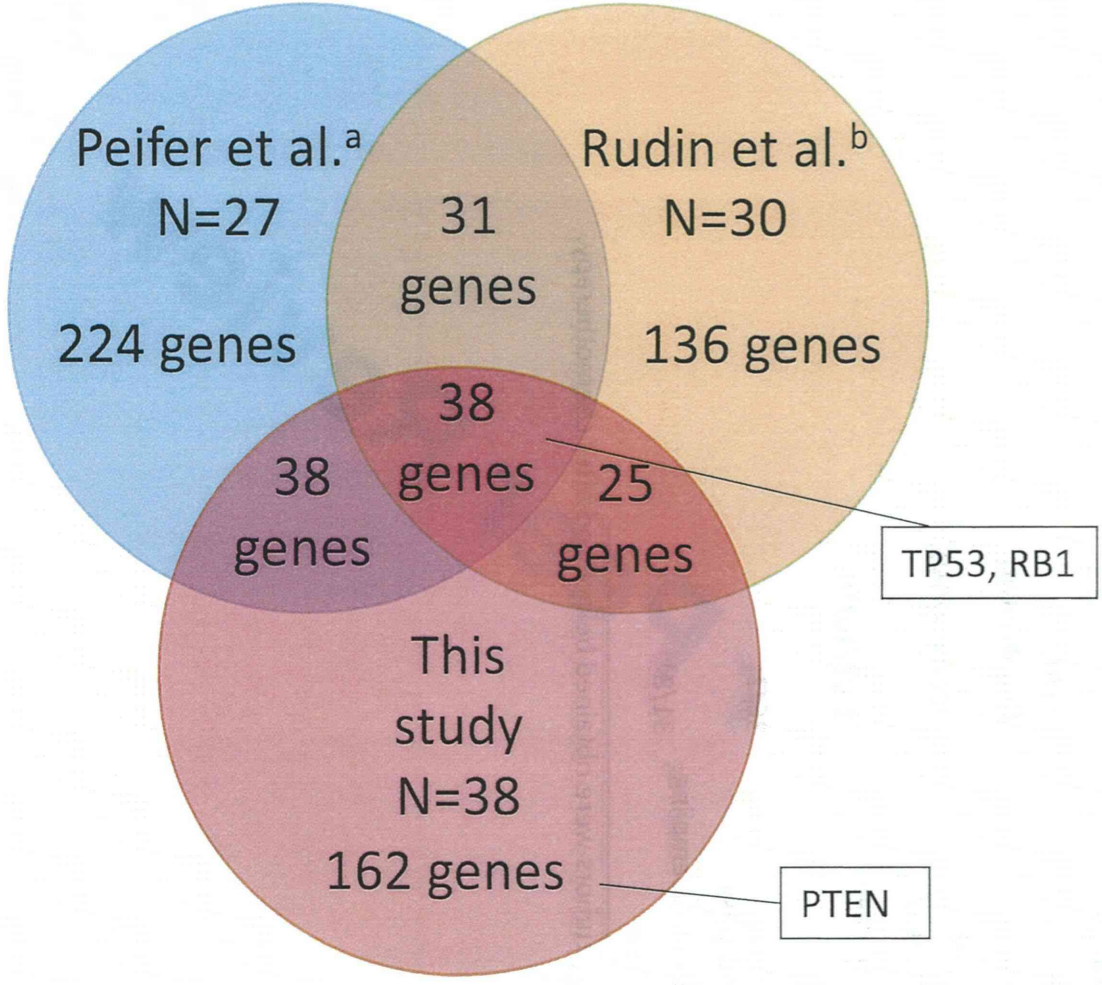

Peifer et al.[a]
N=27
224 genes
31
genes
Rudin et al.[b]
N=30
136 genes
38
genes
38
genes
25
genes
TP53, RB1
This
study
N=38
162 genes
PTEN

Figure 2

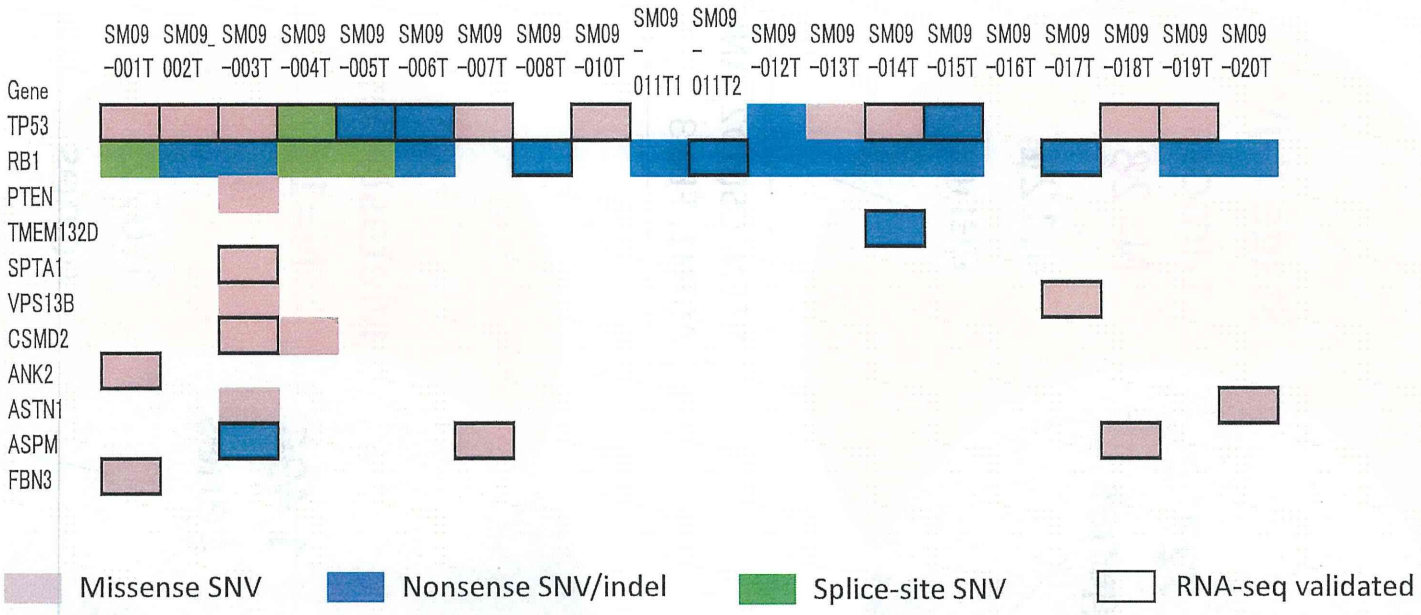
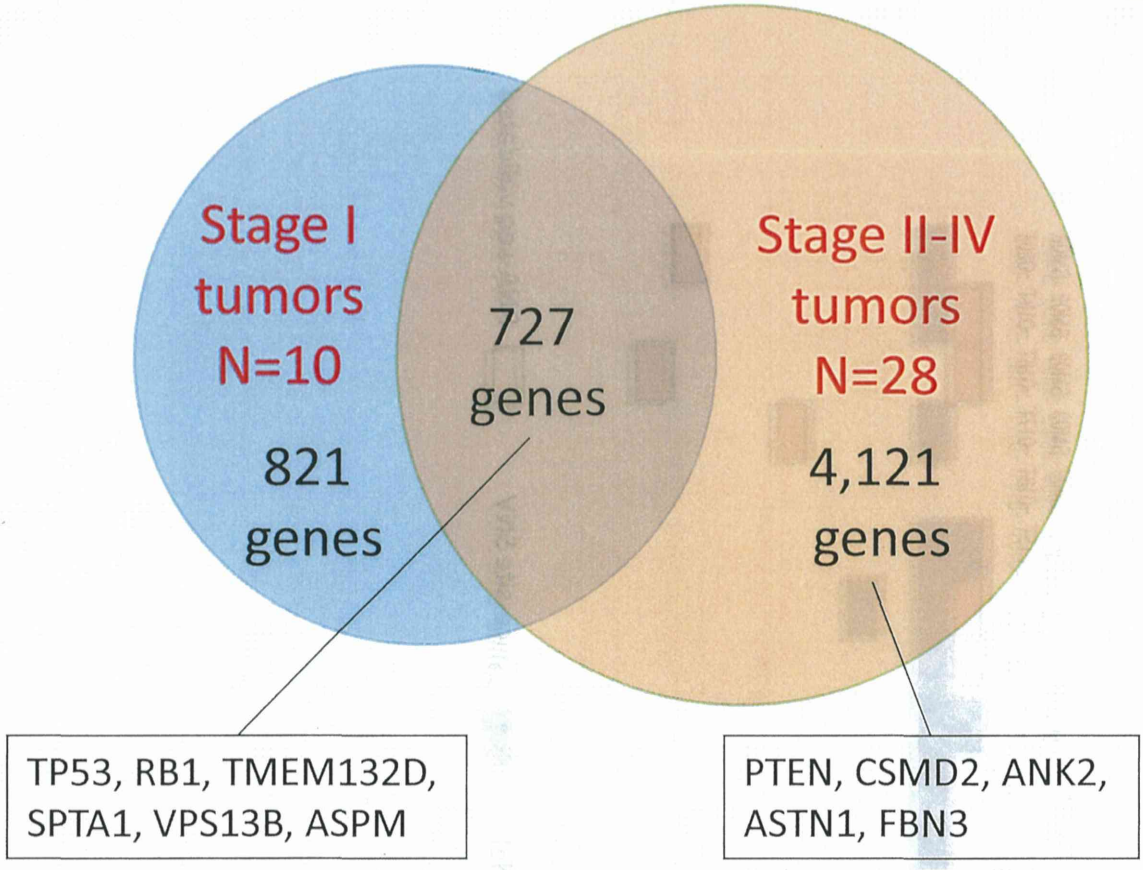


Figure 3

A



B

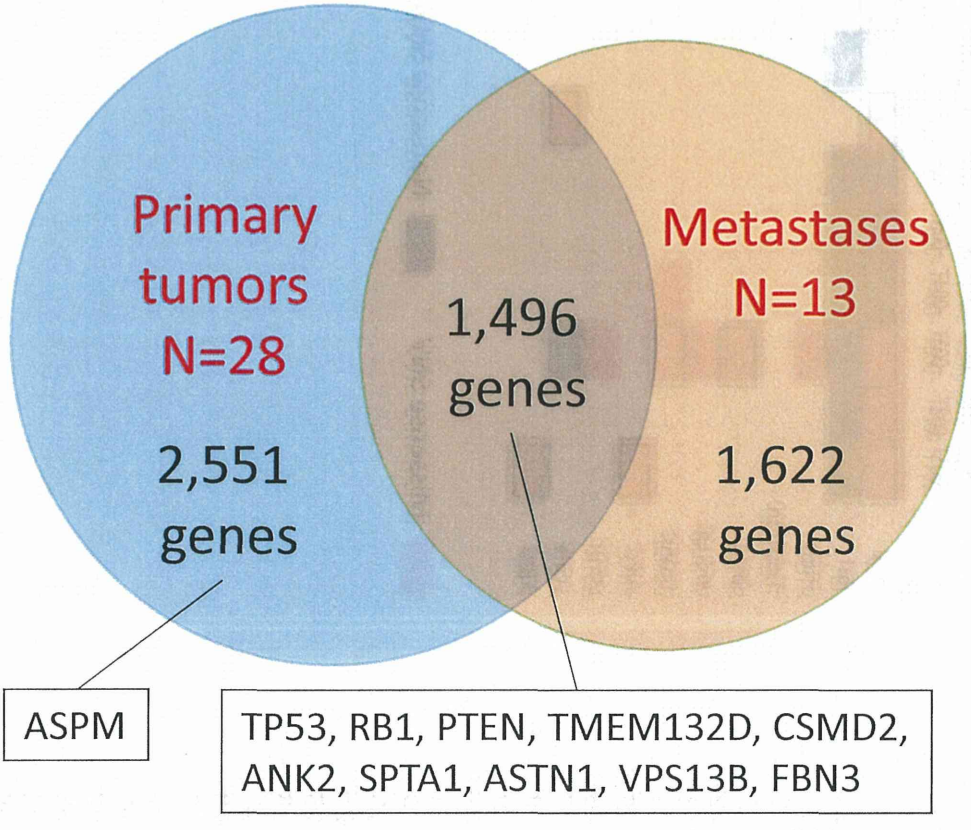
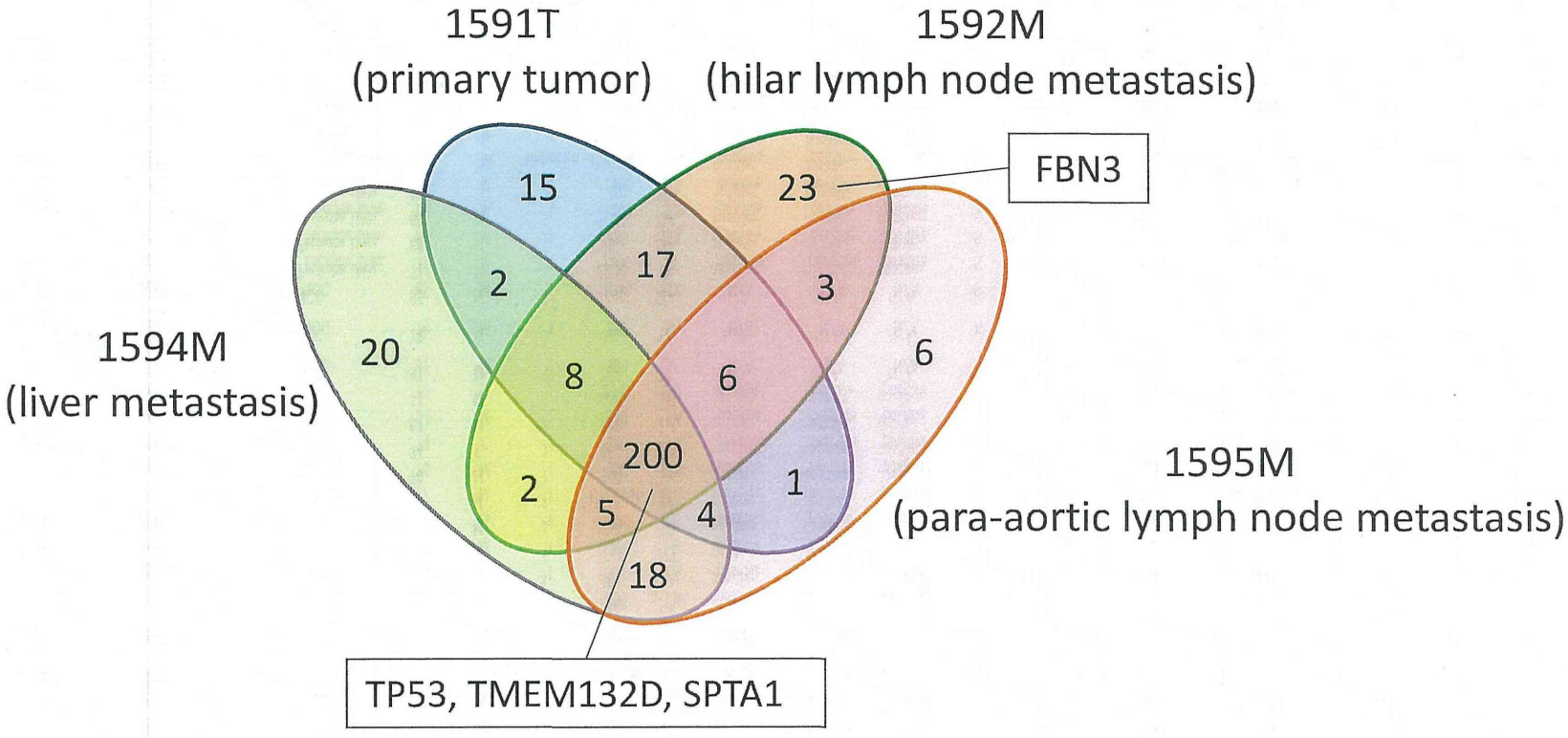


Figure 4



Supplementary Table S1. Sample information

	Sample ID	Origin of Tumor	Source	Age at Diagnosis	Gender Male = M Female = F	Smoking Never = 0 Ever = 1	Brinkman Index	Stage	Chromogranin A	Synaptophysin	NCAM	Obtained at Surgery = S Autopsy = A	Treatment Before Sampling	RNA-Seq	Chemotherapy regimen	No. of courses	Response for chemotherapy	OS (month)	Outcome Dead = 1
1	SM09-001T	Lung (P)	NCC	75	F	Unknown	Unknown	IIA	positive	positive	positive	S	No	Yes	Non	–	–	20.0	1
2	SM09-002T	Lung (P)	NCC	89	M	1	1020	IA	negative	negative	negative	S	No	Yes	Non	–	–	26.6	1
3	SM09-003T	Lung (P)	NCC	79	F	1	680	IIIB	positive	positive	positive	S	No	Yes	CBDCA+VP16	3	–	64.5	0
4	SM09-004T	Lung (P)	NCC	70	F	1	1000	IIIB	positive	positive	positive	S	No	Yes	Non	–	–	19.7	1
5	SM09-005T	Lung (P)	NCC	59	M	1	660	IA	positive	positive	positive	S	No	Yes	Non	–	–	29.6	1
6	SM09-006T	Lung (P)	NCC	75	M	1	800	IIB	negative	negative	positive	S	No	Yes	Non	–	–	1.1	0
7	SM09-007T	Lung (P)	NCC	65	M	1	1350	IA	positive	positive	positive	S	No	Yes	Non	–	–	1.5	1
8	SM09-008T	Lung (P)	NCC	64	M	1	1000	IIA	positive	positive	positive	S	No	Yes	CDDP+CPT-11	4	–	9.1	1
9	SM09-010T	LN (M)	NCC	67	F	Unknown	Unknown	IIA	positive	positive	positive	S	No	Yes	CDDP+CPT-11	4	–	67.0	0
10	SM09-011T1/011T2	Lung (P)/Rib (M)	NCC	66	M	1	900	IIIA	positive	positive	positive	S	No	Yes	CDDP+CPT-11	4	–	19.7	1
11	SM09-012T	Lung (P)	NCC	63	M	1	400	IA	positive	positive	positive	S	No	Yes	Non	–	–	97.7	0
12	SM09-013T	Lung (P)	NCC	76	M	1	301	IA	positive	positive	positive	S	No	Yes	Non	–	–	26.2	0
13	SM09-014T	Lung (P)	NCC	57	M	1	1560	IB	positive	positive	positive	S	No	Yes	Non	–	–	92.0	0
14	SM09-015T	Lung (P)	NCC	67	M	1	1290	IIB	positive	positive	positive	S	No	Yes	Non	–	–	9.7	0
15	SM09-016T	Lung (P)	NCC	47	M	1	Unknown	IIIA	positive	positive	positive	S	No	Yes	CDDP+CPT-11	4	–	47.8	1
16	SM09-017T	Lung (P)	NCC	70	M	1	800	IIIA	positive	positive	positive	S	No	Yes	Non	–	–	78.8	1
17	SM09-018T	Lung (P)	NCC	74	F	1	980	IA	negative	negative	positive	S	No	Yes	CDDP+CPT-11	4	–	67.7	0
18	SM09-019T	Lung (P)	NCC	67	M	1	940	IA	positive	negative	positive	S	No	Yes	CDDP+CPT-11	4	–	54.8	0
19	SM09-020T	Lung (P)	NCC	72	M	1	2040	IIB	positive	positive	positive	S	No	Yes	CBDCA+VP16	4	–	22.9	0
20	1581M/1582M	Liver (M)/LN (M)	NCC	61	M	1	800	IV	N/A	N/A	N/A	A	CT + RT	No	CDDP+VP16	4	PR	9.0	1
21	1591T/1592M/1594M/1595M	Lung (P)/Hilar LN (M) /Liver (M)/Para-aortic LN (M)	NCC	78	M	1	360	IV	N/A	N/A	N/A	A	CT	No	CBDCA	3	PR	10.8	1
22	1601T/1602M	Lung (P)/LN (M)	NCC	65	M	1	1840	IIIA	N/A	N/A	N/A	S	No	No	Non	–	–	1.1	1
23	THB-Lu-1T	Lung (P)	Tsukuba Univ.	71	M	1	2000	IA	negative	negative	positive	S	No	No	Non	–	–	28.4	1
24	THB-Lu-2T	Lung (P)	Tsukuba Univ.	68	M	1	960	IIIA	positive	positive	positive	S	No	No	CBDCA + VP16	4	–	28.4	1
25	THB-Lu-3T	Lung (P)	Tsukuba Univ.	68	M	1	1200	IIIA	positive	positive	positive	S	No	No	Non	–	–	5.1	1
26	THB-Lu-4T	Lung (P)	Tsukuba Univ.	63	M	1	1380	IIIA	positive	positive	positive	S	No	No	CDDP+VP16	4	–	25.4	1
27	THB-Lu-5T	Lung (P)	Tsukuba Univ.	66	M	Unknown	Unknown	IIIA	positive	positive	positive	S	CT	No	CBDCA+PEM	2	Unknown	30.4	0
28	SBM-T04	Brain (M)	Saitama Univ.	73	M	1	1060	IV	negative	positive	N/A	S	No	No	Non	–	–	4.6	1
29	SBM-T08	Brain (M)	Saitama Univ.	58	M	1	1000	IV	negative	positive	N/A	S	No	No	Non	–	–	0.6	1
30	SBM-T17	Brain (M)	Saitama Univ.	66	M	1	1380	IV	N/A	negative	N/A	S	No	No	Non	–	–	22.3	0
31	SBM-T37	Brain (M)	Saitama Univ.	66	F	1	520	IV	positive	positive	N/A	S	No	No	Non	–	–	24.1	1
32	SBM-T40	Brain (M)	Saitama Univ.	76	F	Unknown	Unknown	IV	positive	positive	N/A	S	No	No	Non	–	–	0.5	0
33	M153T	Brain (M)	NCC	72	F	Unknown	Unknown	IV	N/A	N/A	N/A	S	CT + S +RT	No	CDDP+CPT-11/AMR	1/4	CR	31.9	1
34	M189T	Brain (M)	NCC	56	F	1	Unknown	IV	positive	positive	positive	S	CT + RT	No	CDDP+CPT-11/AMR/PEI	3/1/5	CR/PD/PD	29.6	1
35	M288T	Brain (M)	NCC	58	M	1	1170	IV	positive	positive	positive	S	RT + S + CT	No	CDDP+CPT-11	4	Unknown	22.7	1
36	11169T	Lung (P)	US	76	M	1	1220	IB	N/A	N/A	N/A	S	No	No	Non	–	–	101.7	1
37	12878T	Lung (P)	US	70	M	1	1160	IIB	N/A	positive	N/A	S	No	No	Non	–	–	11.3	1
38	19100T	Lung (P)	US	58	F	1	1140	IIA	N/A	N/A	N/A	S	CT	No	Unknown	Unknown	Unknown	14.5	1

P = primary tumor
M = metastatic tumor
LN = lymph node

NCC = National Cancer Center
US = United States

CT = Chemotherapy
RT = Radiotherapy

CBDCA = Carboplatin
VP16 = Etoposide
CDDP = Cisplatin
PEM = Pemetrexed
AMR = Amrubicin

Supplementary Table S3. List of genes frequently mutated in small cell lung cancer

		Total	N=95	This study	N=38	Peifer et al. (ref. 6)	N=27	Rudin et al. (ref. 7)	N=30
No	Gene	#of mutated case	Freq (%)	#of mutated case	Freq (%)	#of mutated case	Freq (%)	#of mutated case	Freq (%)
1	TP53	79	83.2	30	78.9	24	88.9	25	83.3
2	RB1	63	66.3	28	73.7	18	66.7	17	56.7
3	PTEN	9	9.5	5	12.5	2	7.4	2	6.7
4	LRP1B	41	43.2	14	36.8	13	48.1	14	46.7
5	TTN	35	36.8	18	47.4	10	37.0	7	23.3
6	CSMD3	35	36.8	12	31.6	10	37.0	13	43.3
7	USH2A	34	35.8	15	39.5	12	44.4	7	23.3
8	MUC17	26	27.4	10	26.3	8	29.6	8	26.7
9	SYNE1	24	25.3	10	26.3	6	22.2	8	26.7
10	CSMD2	20	21.1	9	23.7	7	25.9	4	13.3
11	COL11A1	20	21.1	5	13.2	8	29.6	7	23.3
12	FLG	19	20.0	7	18.4	7	25.9	5	16.7
13	PKHD1	19	20.0	6	15.8	9	33.3	4	13.3
14	SPHKAP	19	20.0	7	18.4	8	29.6	4	13.3
15	SI	18	18.9	6	15.8	7	25.9	5	16.7
16	GRIN2A	17	17.9	7	18.4	5	18.5	5	16.7
17	TNN	16	16.8	6	15.8	7	25.9	3	10.0
18	FAM135B	16	16.8	7	18.4	5	18.5	4	13.3
19	XIRP2	16	16.8	9	23.7	4	14.8	3	10.0
20	NAV3	15	15.8	6	15.8	4	14.8	5	16.7
21	COL12A1	15	15.8	5	13.2	5	18.5	5	16.7
22	COL22A1	15	15.8	4	10.5	5	18.5	6	20.0
23	DNAH9	15	15.8	4	10.5	6	22.2	5	16.7
24	PLXNA4	15	15.8	9	23.7	3	11.1	3	10.0
25	CDH10	15	15.8	6	15.8	6	22.2	3	10.0
26	ANK2	14	14.7	5	13.2	5	18.5	4	13.3
27	TMEM132D	14	14.7	4	10.5	7	25.9	3	10.0
28	CDH18	14	14.7	7	18.4	4	14.8	3	10.0
29	DNAH5	14	14.7	6	15.8	3	11.1	5	16.7
30	SPTA1	14	14.7	7	18.4	4	14.8	3	10.0
31	HCN1	13	13.7	4	10.5	5	18.5	4	13.3
32	LAMA2	13	13.7	4	10.5	5	18.5	4	13.3
33	LRRC7	12	12.6	4	10.5	4	14.8	4	13.3
34	GPR112	11	11.6	5	13.2	3	11.1	3	10.0
35	ASTN1	11	11.6	4	10.5	3	11.1	4	13.3
36	ASPM	11	11.6	4	10.5	4	14.8	3	10.0
37	VPS13B	11	11.6	4	10.5	4	14.8	3	10.0
38	FBN3	10	10.5	4	10.5	3	11.1	3	10.0
39	MYH1	10	10.5	4	10.5	3	11.1	3	10.0

Supplementary Table S4. Types of TP53, RB1 and PTEN gene mutations detected in 38 SCLC cases

No	Sample name	TP53	RB1	PTEN
1	SM09-001T	NM_000546:exon5:c.G473T:p.R158L	NM_000321:exon16:c.1422-2A>T	—
2	SM09-002T	NM_000546:exon5:c.A536G:p.H179R	NM_000321:exon20:c.2033_2034insT:p.H678fs	—
3	SM09-003T	NM_000546:exon8:c.G814A:p.V272M	NM_000321:exon17:c.1621delA:p.N541fs	NM_000314:exon5:c.T463C:p.Y155H
4	SM09-004T	NM_000546:exon6:c.376-1G>T	NM_000321:exon24:c.2520+1G>T	—
5	SM09-005T	NM_000546:exon6:c.G610T:p.E204X	NM_000321:exon16:c.1422-2A>G	—
6	SM09-006T	NM_000546:exon4:c.262_263insA:p.A88fs	NM_000321:exon20:c.G2043A:p.W681X	—
7	SM09-007T	NM_000546:exon7:c.G746T:p.R249M	—	—
8	SM09-008T	—	NM_000321:exon13:c.G1267T:p.G423X	—
9	SM09-010T	NM_000546:exon5:c.A536T:p.H179L, exon6:c.559+1G>C	—	—
10	SM09-011T1/ SM09-011T2	—	NM_000321:exon24:c.C2501G:p.S834X	—
11	SM09-012T	NM_000546:exon6:c.652delG:p.V218fs	NM_000321:exon11:c.C1072T:p.R358X	—
12	SM09-013T	NM_000546:exon7:c.G733C:p.G245R	NM_000321:exon23:c.2436_2439del:p.812_813del	—
13	SM09-014T	NM_000546:exon6:c.A583T:p.I195F	NM_000321:exon22:c.2218_2219del:p.740_740del	—
14	SM09-015T	NM_000546:exon5:c.C548G:p.S183X	NM_000321:exon9:c.G937T:p.E313X	—
15	SM09-016T	—	—	—
16	SM09-017T	—	NM_000321:exon17:c.1580_1581insT:p.D527fs	—
17	SM09-018T	NM_000546:exon8:c.G818C:p.R273P	—	—
18	SM09-019T	NM_000546:exon8:c.G818T:p.R273L	NM_000321:exon3:c.370_371del:p.124_124del	—
19	SM09-020T	—	NM_000321:exon19:c.1868_1883del:p.623_628del	—
20	1581M/1582M	—	—	—
21	1591T/1592M/ 1594M/1595M	NM_000546:exon6:c.G592T:p.E198X	—	—
22	1601T/1602M	NM_000546:exon4:c.96+1G>T	—	NM_000314:exon9:c.C1126T:p.H376Y, exon7:c.711delG:p.K237fs
23	THB_Lu_1T	NM_000546:exon6:c.T660G:p.Y220X	NM_000321:exon17:c.1695+1G>T	—
24	THB_Lu_2T	NM_000546:exon4:c.C112T:p.Q38X	NM_000321:exon19:c.1917_1918insA:p.Q639fs	—
25	THB_Lu_3T	NM_000546:exon4:c.162_163insC:p.T55fs	NM_000321:exon24:c.2490-1G>A	—
26	THB_Lu_4T	NM_000546:exon6:c.G661T:p.E221X	—	—
27	THB_Lu_5T	NM_000546:exon7:c.G743T:p.R248L	NM_000321:exon13:c.G1318T:p.E440X	NM_000314:exon5:c.G448T:p.E150X
28	SBM_T04	NM_000546:exon7:c.A707G:p.Y236C	NM_000321:exon12:c.1128-2A>T	—
29	SBM_T08	—	NM_000321:exon13:c.A1234T:p.K412X	NM_000314:exon6:c.G518T:p.R173L
30	SBM_T17	NM_000546:exon6:c.G638T:p.R213L	NM_000321:exon10:c.1011delG:p.L337fs, exon10:c.G1011T:p.L337F	—
31	SBM_T37	NM_000546:exon8:c.782+1G>C	NM_000321:exon2:c.G157T:p.E53X	—
32	SBM_T40	NM_000546:exon5:c.529_546del:p.177_182del	NM_000321:exon17:c.G1675T:p.E559X	—
33	M153T	NM_000546:exon6:c.G592T:p.E198X	—	—
34	M189T	NM_000546:exon9:c.C991T:p.Q331X	NM_000321:exon23:c.2330_2331insT:p.P777fs	—
35	M288T	NM_000546:exon7:c.G733T:p.G245C	NM_000321:exon22:c.C2273A:p.S758X	NM_000314:exon5:c.G316T:p.E106X
36	11169T	NM_000546:exon8:c.G839T:p.R280I	NM_000321:exon14:c.1333-1G>C	—
37	12878T	NM_000546:exon5:c.G469T:p.V157F	—	—
38	19100T	NM_000546:exon5:c.G469T:p.V157F	NM_000321:exon21:c.2211+1G>A	—

Supplementary Table S5. Significantly mutated genes in three studies

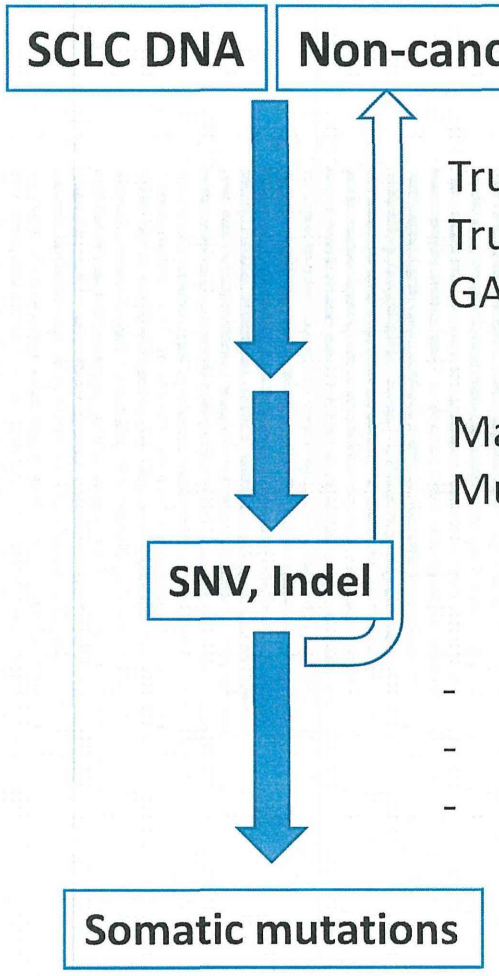
No	Gene	This study	Peifer et al. (ref. 6)	Rudin et al. (ref. 7)
1	TP53	v	v	v
2	RB1	v	v	v
3	PTEN	v		
4	PRDM9		v	
5	TFAP2D		v	
6	TKTL2		v	
7	OR10G7		v	
8	OR11L1		v	
9	ZNF536		v	
10	TMEM132D		v	v
11	ZNF521		v	
12	TUBA3C		v	
13	HCN1		v	
14	SLURP1		v	
15	TNN		v	
16	FAM133A		v	
17	CDH10		v	
18	SPHKAP		v	
19	COL11A1		v	
20	NCAM2		v	
21	XPNPEP2		v	
22	OR5B17		v	
23	FOXE1		v	
24	COL22A1			v
25	ELAVL2			v
26	RASSF8			v
27	CNTNAP2			v
28	BCLAF1			v
29	GRM8			v
30	KIF21A			v
31	GRIK3			v
32	C17orf108			v
33	RUNX1T1			v
34	PLSCR4			v
35	CDYL			v
36	RIMS2			v
37	ZDBF2			v
38	KHSRP			v
39	SATB2			v
40	COL4A2			v
41	DIP2C			v
42	ADCY1			v

Supplementary Table S6. Mutation spectra of significantly mutated genes

Gene	Name	Location	Alterations	SM09-001T	SM09-002T	SM09-003T	SM09-004T	SM09-005T	SM09-006T	SM09-007T	SM09-008T	SM09-010T	SM09-011T1/011T2	SM09-012T	SM09-013T	SM09-014T	SM09-015T	SM09-016T	SM09-017T	SM09-018T	SM09-019T	SM09-020T	THB_Lu_1T	THB_Lu_2T	THB_Lu_3T	THB_Lu_4T	THB_Lu_5T	11169T	12878T	19100T	M153T	M189T	M288T	SBM_T04	SBM_T08	SBM_T17	SBM_T37	SBM_T40-	1581M/1582M	1591T/1592M/1594M/1595M	1601T/1602M
TP53	tumor protein p53	17p13.1	Mutation	p.R158L	p.H179R	p.V272M	c.376-1G	p.E204X	p.A88fs	p.R249M		c.559+1G		p.Y218fs	p.G245R	p.I195F	p.S183X			p.R273P	p.R273L		p.Y220X	p.Q38X	p.T55fs	p.E221X	p.R248L	p.R280I	p.V157F	p.V157F	p.E198X	p.Q331X	p.G245C	p.Y238C		p.R213L	c.782+1G	p.177_182del		p.E198X	c.96+1G
			Variant frequency	0.91	0.89	0.88	0.89	0.38	0.87	1.00		0.45, 0.56		0.82	0.87	1.00	1.00			0.84	0.67		0.39	0.58	0.85	0.13	0.95	0.98	0.91	0.91	0.13	0.75	0.97	0.49		0.83	0.73	0.57		0.89	0.85
			Copy number	LOH	LOH	LOH	LOH	LOH	LOH	LOH	LOH	Gain	LOH	LOH	LOH	LOH	LOH		LOH	LOH	LOH	LOH	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
RB1	retinoblastoma 1	13q14.2	Mutation	c.1422-2_	p.H678fs	p.N541fs	c.2520+1	c.1422-2_	p.W681X		p.Q423X		p.S834X	p.R358X	p.812_813	p.740_740	p.E313X		p.D527fs		p.124_124	p.823_823	c.1695+1	p.Q639fs	c.2490-1		p.E440X	c.1333-1		c.2211+1		p.P777fs	p.S758X	c.1128-2	p.X412X	p.L337fs	p.E53X	p.E559X			
			Variant frequency	0.87	0.85	0.63	0.79	0.41	0.89		0.92		0.88	0.89	0.56	0.45	0.98		0.58		0.57	0.86	0.41	0.34	0.85		0.91	0.98		0.76		0.57	1.00	0.46	0.83	0.79, 0.26	0.66	0.63			
			Copy number	LOH	LOH	LOH	LOH		LOH	LOH	LOH	LOH	LOH	LOH	LOH	LOH	Gain	LOH	LOH		LOH	LOH	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PTEN	phosphatase and tensin homolog	10q23.3	Mutation			p.Y155H																					p.E150X					p.E106X		p.R173L						p.H376Y,
			Variant frequency			0.62																					0.92					0.98		0.70						0.80, 0.83
			Copy number	LOH		LOH	LOH		LOH		Loss		T2:LOH	LOH			LOH					Gain	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
TMEM132D	transmembrane protein 132D	12q24.33	Mutation													p.Q763X							p.A694T														p.M614I		p.P1038L	
			Variant frequency													0.37							0.35														0.39		1.00	
			Copy number						Gain				LOH			Gain	LOH					Gain	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SPTA1	spectrin, alpha, erythrocytic 1	1q21	Mutation			p.J2231M																	p.A2025E		p.J1326T		p.D2084E					p.V1685G				c.1678-1			p.P1040L	
			Variant frequency			0.37																	0.34		0.56		0.45					0.39				0.56			0.43	
			Copy number		Gain	Gain	Gain		Gain			Gain	T2:Gain	Gain	Gain	Gain		Gain	Gain			Gain	Gain	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
VPS13B	vacuolar protein sorting 13 homolog B	8q22.2	Mutation			p.A2041S													p.Y324C				p.L631V											p.G1637C							
			Variant frequency			0.40													0.44				0.31											0.57, 0.21	0.24						
			Copy number				Gain	Gain	LOH	Gain			T1:Gain	Gain		Gain	Gain		Gain		Gain		NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

inactivating mutation (nonsense mutations, frameshift mutations and mutations at exon-intron boundary)
variant frequency > 50%
LOH detected by SNP array analysis
Gain detected by SNP array analysis

Supplementary Figure S1. Exome Sequencing and Data Processing



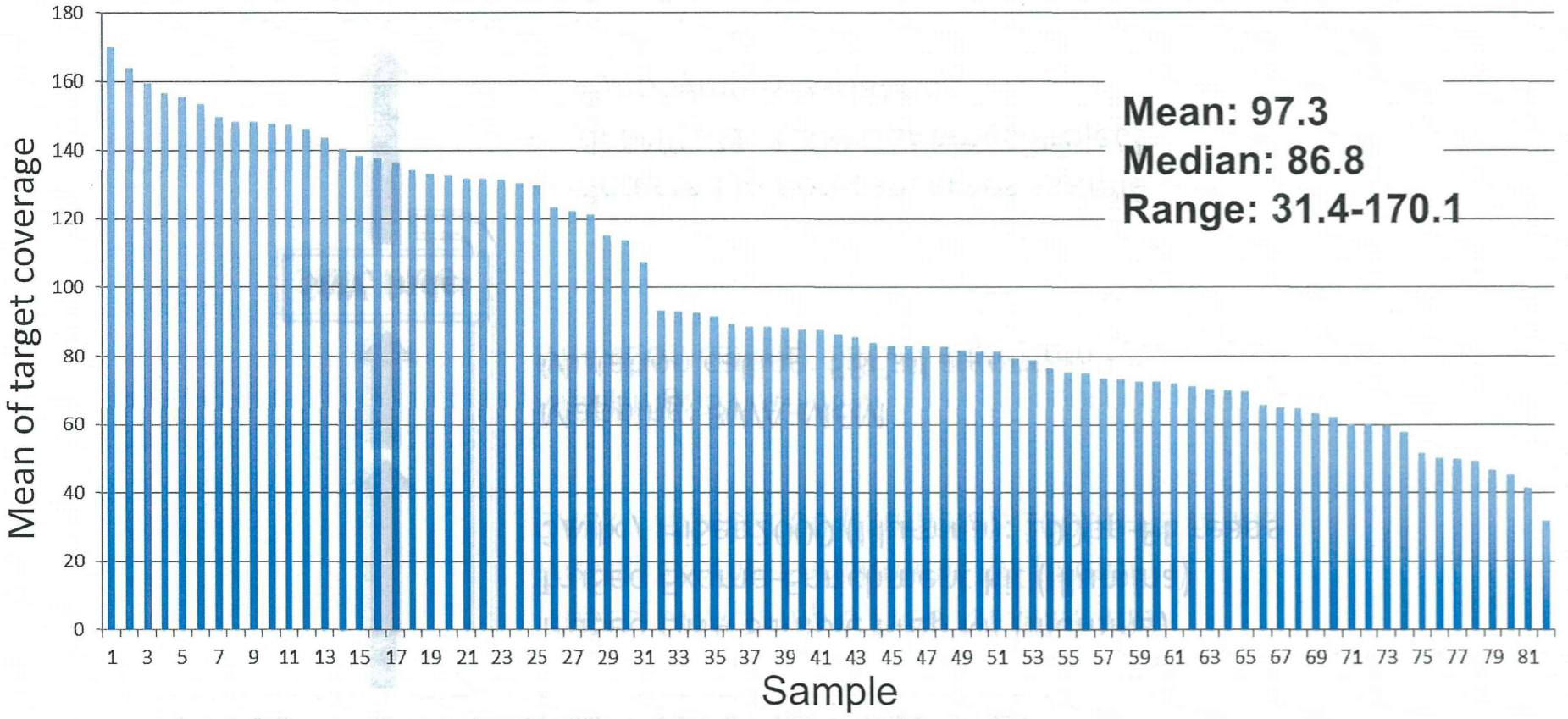
TruSeq DNA Sample Prep Kit (Illumina)
TruSeq Exome-Enrichment Kit (Illumina)
GAIIx / HiSeq2000 (Illumina): 100bp-PE reads

Mapping: BWA-MEM
Mutation calling: EBCall algorithm[1]

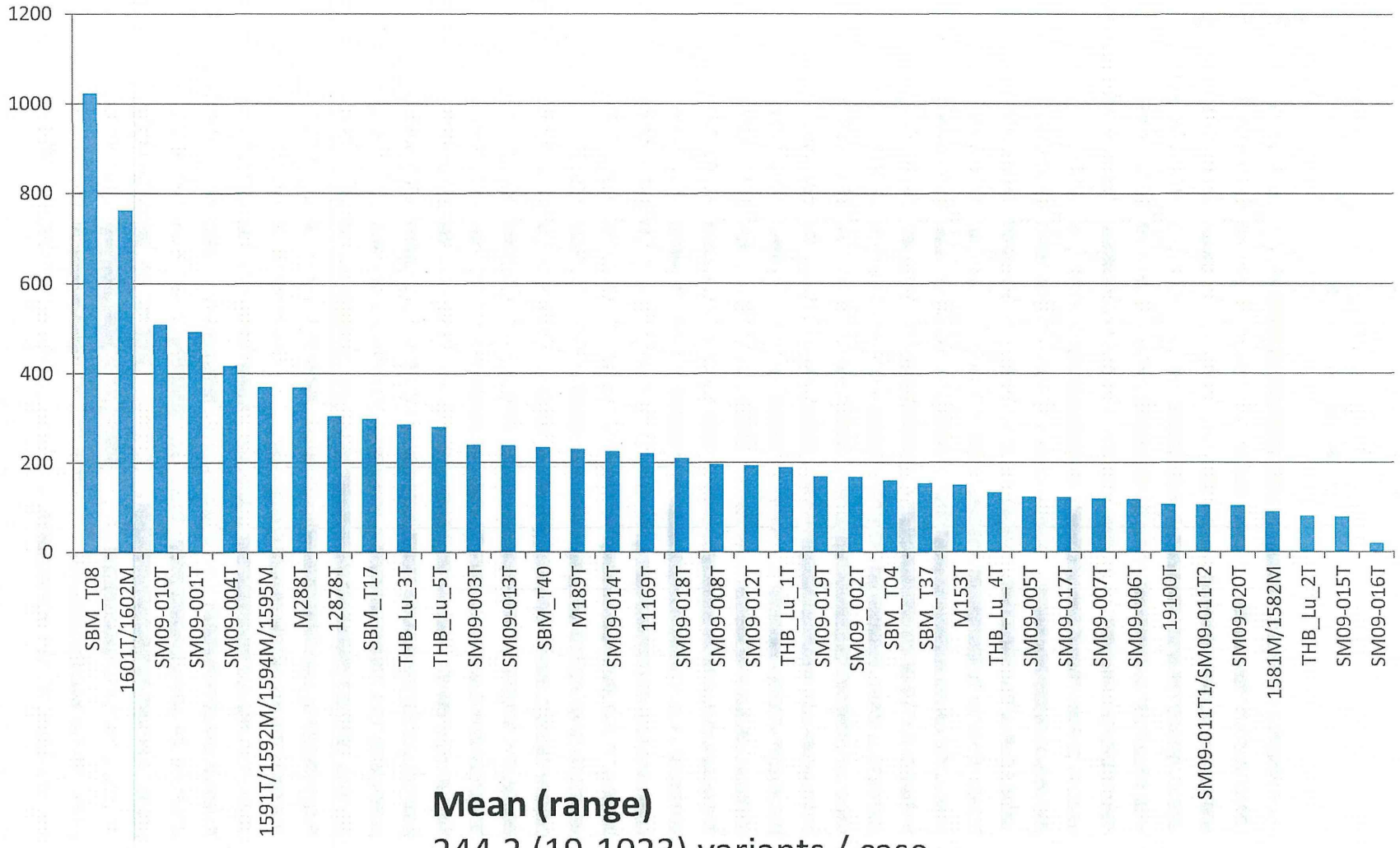
- detected in non-cancerous sample
- dbSNP131, 1000 Genomes Project
- Synonymous variations

(1), Shiraishi Y, et al., An empirical Bayesian framework for somatic mutation detection from cancer genome sequencing data. Nucleic Acids Res. 2013;41:e89.

Supplementary Figure S2. Mean of Target Coverage in 38 SCLC Cases



Supplementary Figure S3. Numbers of Somatic Mutations Detected in 38 SCLC Cases

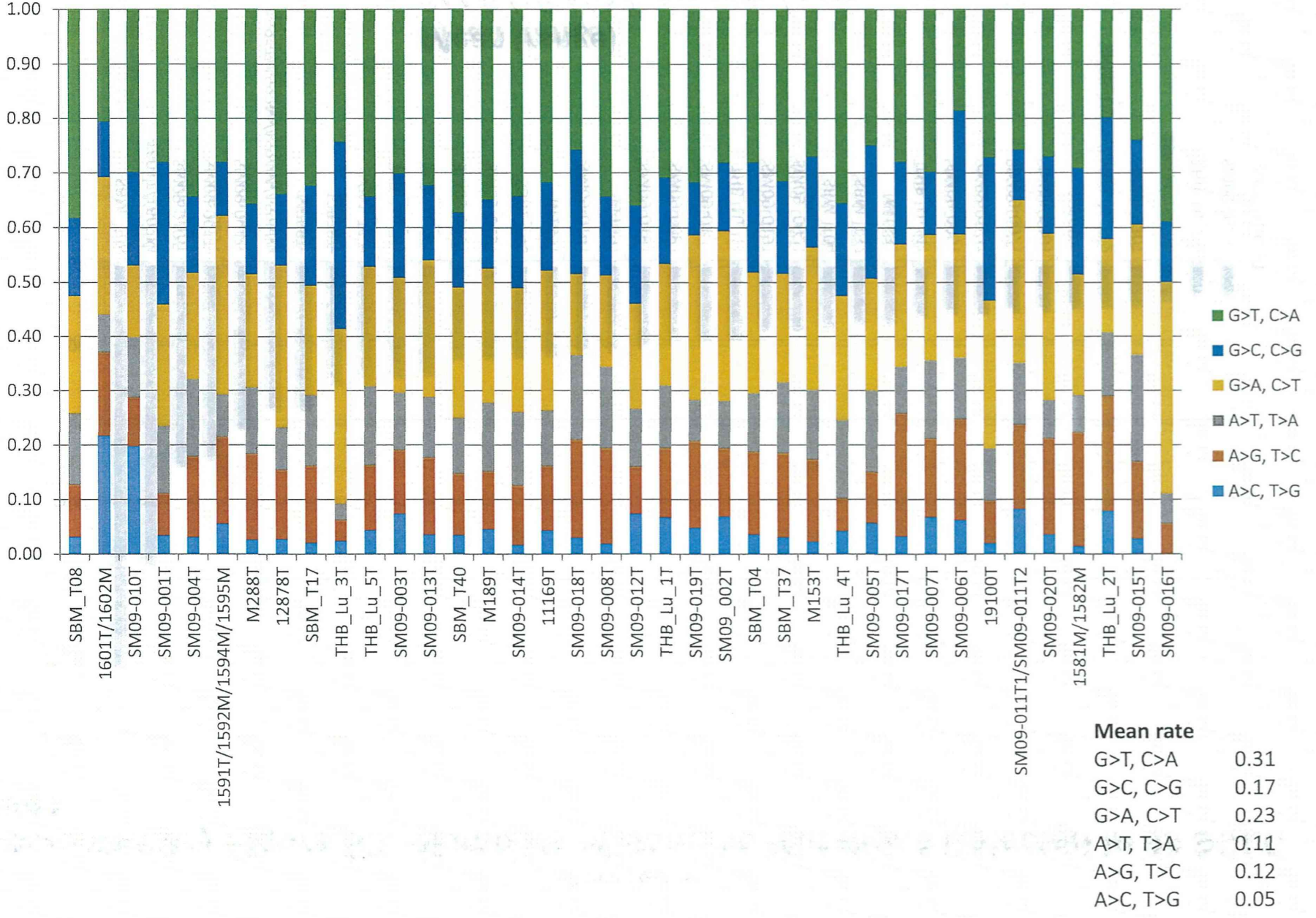


Mean (range)

244.2 (19-1023) variants / case

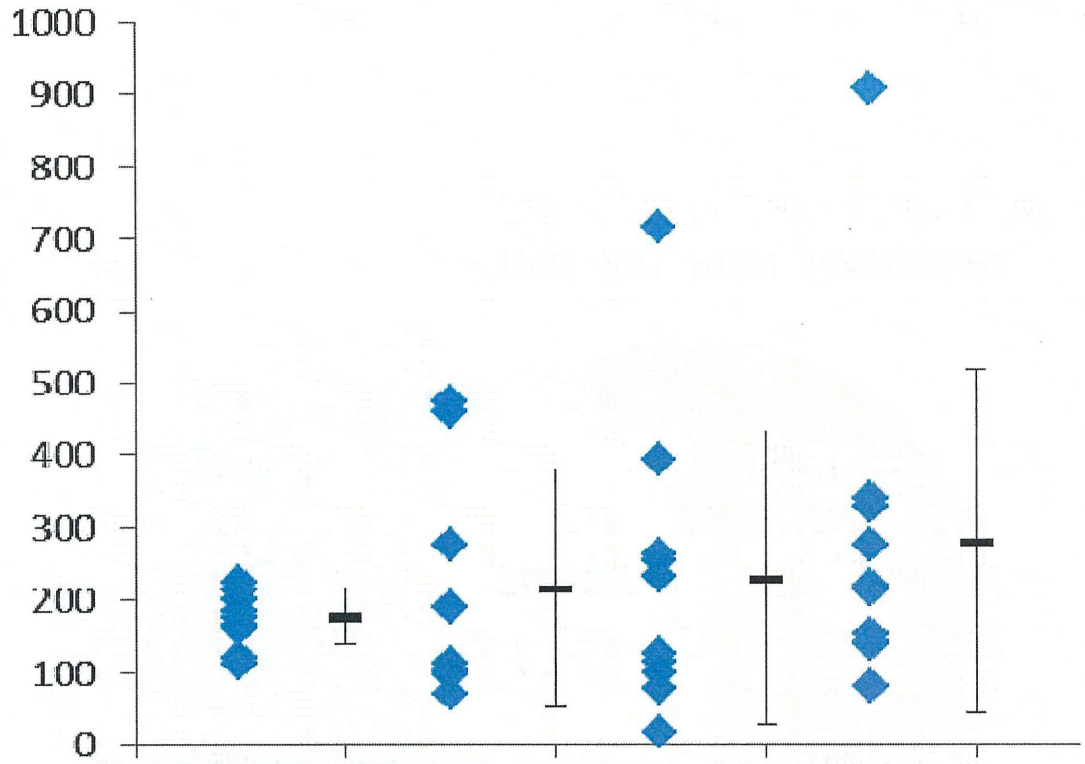
7.4 (0.6-30.8) variants / Mb

Supplementary Figure S4. Base-level Transitions and Transversions in 38 SCLC Cases



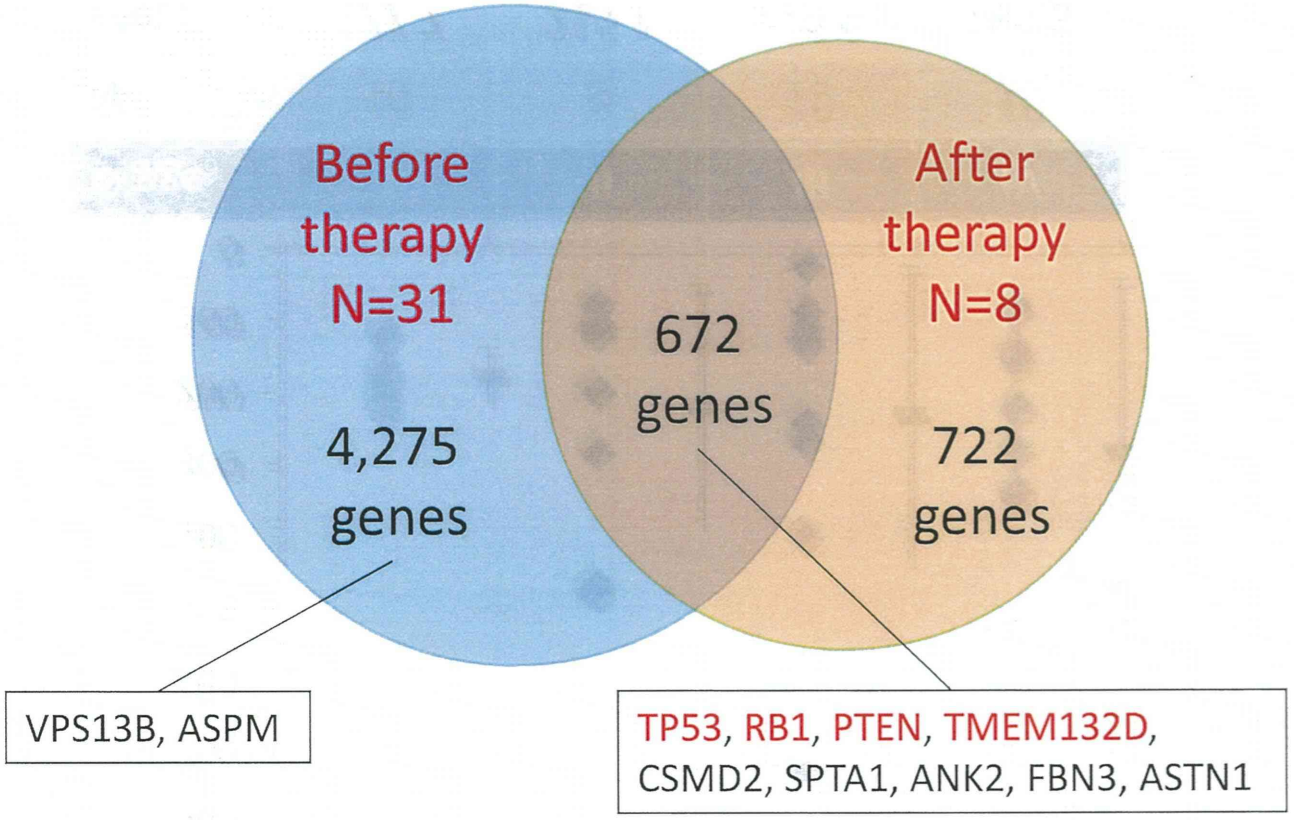
Mean rate	
G>T, C>A	0.31
G>C, C>G	0.17
G>A, C>T	0.23
A>T, T>A	0.11
A>G, T>C	0.12
A>C, T>G	0.05

Supplementary Figure S5. Numbers of Somatically Mutated Genes in SCLC According to the TNM Staging



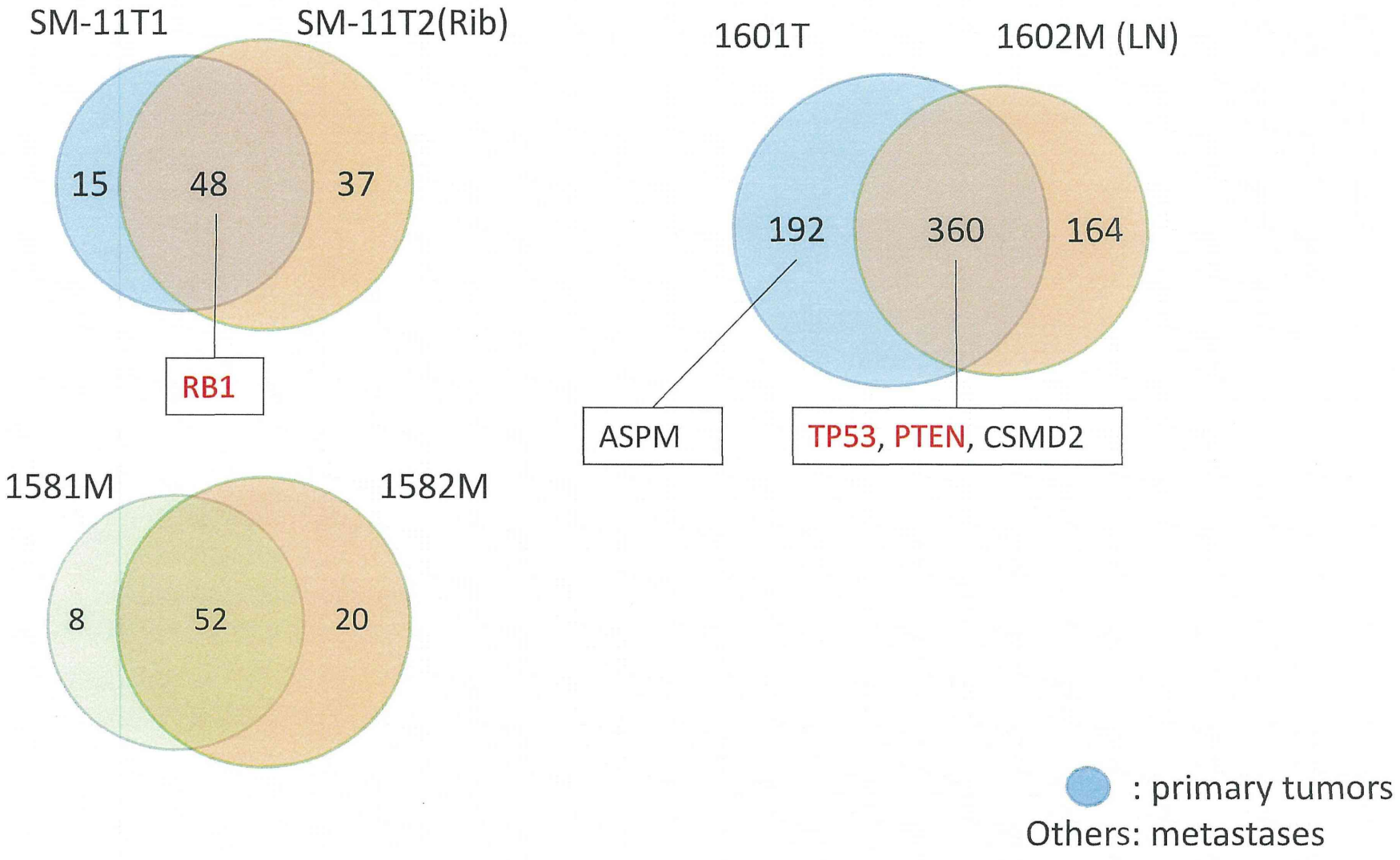
Stage	I	II	III	IV
N	10	8	10	10
Mean	177.7	215.1	229.9	280.6

Supplementary Figure S6. Comparison of Mutated Genes Between Tumors Obtained Before and After Chemotherapy/Radiotherapy



Mutated genes / sample (average)	230.7	195.0	(*P*=0.50 by Student's *t*-test)

Supplementary Figure S7. Heterogeneity in the Accumulated Mutations among Multiple Tumors Obtained from Same Patients



Comprehensive Exploration of Novel Chimeric Transcripts in Clear Cell Renal Cell Carcinomas Using Whole Transcriptome Analysis

Masahiro Gotoh,[1†] Hitoshi Ichikawa,[2†] Eri Arai,[1] Suenori Chiku,[3] Hiromi Sakamoto,[2] Hiroyuki Fujimoto,[4] Masaki Hiramoto,[5] Takao Nammo,[5] Kazuki Yasuda,[5] Teruhiko Yoshida,[2] and Yae Kanai[1]*

[1]Division of Molecular Pathology, National Cancer Center Research Institute, Tokyo, Japan
[2]Division of Genetics, National Cancer Center Research Institute, Tokyo, Japan
[3]Science Solutions Division, Mizuho Information and Research Institute, Inc., Tokyo, Japan
[4]Department of Urology, National Cancer Center Hospital, Tokyo, Japan
[5]Department of Metabolic Disorder, Diabetes Research Center, National Center for Global Health and Medicine, Tokyo, Japan

The aim of this study was to clarify the participation of expression of chimeric transcripts in renal carcinogenesis. Whole transcriptome analysis (RNA sequencing) and exploration of candidate chimeric transcripts using the deFuse program were performed on 68 specimens of cancerous tissue (T) and 11 specimens of non-cancerous renal cortex tissue (N) obtained from 68 patients with clear cell renal cell carcinomas (RCCs) in an initial cohort. As positive controls, two RCCs associated with Xp11.2 translocation were analyzed. After verification by reverse transcription (RT)-PCR and Sanger sequencing, 26 novel chimeric transcripts were identified in 17 (25%) of the 68 clear cell RCCs. Genomic breakpoints were determined in five of the chimeric transcripts. Quantitative RT-PCR analysis revealed that the mRNA expression levels for the *MMACHC, PTER, EPC2, ATXN7, FHIT, KIFAP3, CPEB1, MINPP1, TEX264, FAM107A, UPF3A, CDC16, MCCC1, CPSF3,* and *ASAP2* genes, being partner genes involved in the chimeric transcripts in the initial cohort, were significantly reduced in 26 T samples relative to the corresponding 26 N samples in the second cohort. Moreover, the mRNA expression levels for the above partner genes in T samples were significantly correlated with tumor aggressiveness and poorer patient outcome, indicating that reduced expression of these genes may participate in malignant progression of RCCs. As is the case when their levels of expression are reduced, these partner genes also may not fully function when involved in chimeric transcripts. These data suggest that generation of chimeric transcripts may participate in renal carcinogenesis by inducing dysfunction of tumor-related genes.

INTRODUCTION

Clear cell renal cell carcinoma (RCC) is the most common histological subtype of adult kidney cancer (Ljungberg et al., 2011). In general, RCCs at an early stage are curable by nephrectomy. However, some RCCs relapse and metastasize to distant organs. Even though molecular targeting agents have been developed for treatment of RCCs, their effectiveness for relapsed or metastasized RCCs after nephrectomy is very limited. To improve prognostication and the effectiveness of targeting therapy in patients with RCCs, the molecular background of renal carcinogenesis should be further elucidated.

We and other groups have revealed both genetic and epigenetic events during renal carcinogenesis (Arai and Kanai, 2010). Especially, recent developments in high-throughput sequence capture methods and next-generation sequencing technologies have made exome sequencing technically feasible. Such whole exome analyses have revealed that renal carcinogenesis involves inactivation of histone-modifying genes such as *SETD2* (Dalgliesh et al., 2010), *KDM5C* (Dalgliesh et al., 2010), *UTX* (van Haaften et al., 2009), and *PBRM1*

Additional Supporting Information may be found in the online version of this article.

Supported by: the Program for Promotion of Fundamental Studies in Health Sciences (10-42 and 10-43) of the National Institute of Biomedical Innovation (NiBio), Japan and the National Cancer Center Research and Development Fund (23-A-1), Japan.

†Masahiro Gotoh and Hitoshi Ichikawa contributed equally to this work.

*Correspondence to: Yae Kanai; Division of Molecular Pathology, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan. E-mail: ykanai@ncc.go.jp

Received 17 April 2014; Accepted 7 August 2014

DOI 10.1002/gcc.22211

Published online 18 September 2014 in Wiley Online Library (wileyonlinelibrary.com).

TABLE 1. The Clinicopathological Parameters of Clear Cell Renal Carcinomas Belonging to the Initial and Second Cohorts

Clinicopathological parameters	Initial cohort ($n = 68$)	Second cohort ($n = 26$)	P
Age (mean ± SD)	62.25 ± 11.00	57.12 ± 10.80	0.078[a]
Sex			
Male	49	17	0.616[b]
Female	19	9	
Tumor diameter (cm, mean ± SD)	5.55 ± 3.21	5.86 ± 2.84	0.407[a]
Macroscopic configuration[c]			
Type 1	25	13	0.562[b]
Type 2	17	5	
Type 3	26	8	
Predominant histological grades[d,e]			
G1	36	12	0.696[b]
G2	21	8	
G3	9	4	
G4	2	2	
Highest histological grades[f]			
G1	5	1	0.105[b]
G2	32	6	
G3	16	11	
G4	15	8	
Vascular involvement[g]			
Negative	38	13	0.649[b]
Positive	30	13	
Predominant growth pattern[e]			
Expansive	61	24	1.000[b]
Infiltrative	7	2	
Most aggressive growth pattern[f]			
Expansive	43	21	0.139[b]
Infiltrative	25	5	
Tumor necrosis			
Negative	51	16	0.212[b]
Positive	17	10	
Renal pelvic invasion			
Negative	61	23	1.000[b]
Positive	7	3	
Distant metastasis			
Negative	58	24	0.500[b]
Positive	10	2	
Pathological TNM stage[h]			
Stage I	33	13	0.531[b]
Stage II	4	3	
Stage III	19	8	
Stage IV	12	2	

[a]Mann–Whitney *U* test.
[b]Fisher's exact test. No significant differences of clinicopathologcial parameters were observed between two cohorts.
[c]Macroscopic configuration was evaluated on the basis of previously described criteria (Arai et al., 2006).
[d]All the tumors were graded on the basis of previously described criteria (Fuhrman et al., 1982).
[e]If the tumor showed heterogeneity, findings in the predominant area were described.
[f]If the tumor showed heterogeneity, the most aggressive features of the tumor were described.
[g]The presence or absence of vascular involvement was examined microscopically on slides stained with hematoxylin-eosin and elastica van Gieson.
[h]All the tumors were classified according to the pathological Tumor-Node-Metastasis classification (Sobin et al., 2009).

(Varela et al., 2011). Moreover, it is well known that clear cell RCCs are characterized by inactivation of the *VHL* tumor suppressor gene encoding a component of the protein complex that possesses ubiquitin ligase E3 activity (Baldewijns et al., 2010). Frequent mutation of a further component of the ubiquitin-mediated proteolysis pathway gene, *BAP1* (Guo et al., 2012), and *VHL*-associated transcription elongation factor, *TCEB1* (Sato et al., 2013), has also been demonstrated on the basis of exome analyses. However, only a limited number of reports have described next-generation sequencing-based whole transcriptome analysis (RNA sequencing) of RCCs, and the molecular background of renal carcinogenesis has not been fully elucidated.

TABLE 2. The 33 Chimeric Transcripts[a] from the 61 Genes Verified by Reverse Transcription-PCR and Sanger Sequencing

		Exon boundaries[b]								
		5′-partner gene			3′-partner gene					
Subtype	Sample	Symbol	RNA accession no.	Exon (chromosome, position)	Symbol	RNA accession no	Exon (chromosome, position)	Distance (bp)	Flanking sequences[c]	In-frame transcripts
(A) Initial cohort										
Clear	K1	*ANTXR1*	NM_032208	17 (ch 2, 69420547)	*GKN1*	NM_019617	2 (ch 2, 69204627)	215920	TCTGTGATGCGTCCACAGCCAGGAGACACG|ATCGTCTTTGCTGGACTTCTTGGAGTCTTT	+
		B4GALT2	NM_001005417	3 (ch 1, 44450537)	*ZSWIM5*	NR_024270	1 (ch 1, 45769599)	1319062	CTTGGCCCGGTTGAAGGTGTCCTCACCATG|GGTTTCGGAGTAAGGGGAGCGGCCCGGCGA	−
		C9orf3	NM_001193329	12 (ch 9, 97767898)	*PRUNE2*	NM_015225	19 (ch 9, 79229516)	18538382	CGCAGGGAGAAGGAAGAGGTGTTTGAAAAG|TATTGACTTGAAGCTGAAAGAAAAGCCTTA	−
		CTNNB1	NM_001904	1 (ch 3, 41241161)	*PLAG1*	NM_002655	3 (ch 8, 57083748)		CCCGCGGCGGGAGGAGCCTGTTCCCCTGAG|ATTGGCCAAAATGGGAAGGATTGGATTCCA	−
		ERBB2	NM_001005862	15 (ch 17, 37872192)	*LTBP4*	NM_001042545	12 (ch 19, 41116438)		ACACTGCCAACCGGCCAGAGGACGAGTGTG|ATGTGGACGAGTGCCACCGCGTGCCGCCGC	+
		MMACHC	NM_015506	1 (ch 1, 45966085)	*BX004987.7*		(ch 1, 143529235)	97563150	CCTTTTGGCTTCGAGGTTTACCCCTTCCAG|GTTATATGCAGTACTGAAGAGCAACTTCAG	−
		MMACHC	NM_015506	2 (ch 1, 45973222)	*BX004987.7*		(ch 1, 143529235)	97556013	GTGGCCTACCATCTGGGCCGTGTTAGAGAG|GTTATATGCAGTACTGAAGAGCAACTTCAG	−
		PHF3	NM_015153	1 (ch 6, 64356700)	*PTP4A1*	NM_003463	1 (ch 6, 64237724)	118976	ATCCCAATTTCCAGATGCCTTGTTCAACAG|AGACTTGCTTCAACACCAGGAAGAACATGA	−
		PTER	NM_001001484	5 (ch 10, 16547159)	*MSRB2*	NM_012228	4 (ch 10, 23399171)	6852012	CATGCCTGATGATAACAAAAGAATTAGAAG|CCTTTCAGTGGGATCTACCTGAATAACAAG	−
Clear	K2	*DNAH11*	NM_003777	79 (ch 7, 21934617)	*RAPGEF5*	NM_012294	7 (ch 7, 22306631)	372014	TTATCCTTCTACTTATGGCCTAGCCCAGTG|GTGCAGAGAGAGCTAGCAGCTGTTATTGCT	−
Clear	K3	*PLOD2*	NM_000935	1 (ch 3, 145878668)	*CCNYL1*	NM_001142300	4 (ch 2, 208602135)		CTCGGAGAAGCCCTCGAGCATCCCCACAG|TGTGACCTTAGCAATATATTACCACATAAAG	−
Clear	K4	*EPC2*	NM_015630	2 (ch 2, 149447942)	*RP11–715D1.1*		(ch 3, 108828611)		AACAGCCAAAACAGTTCATTCATATTCAGC|GAGAGAGCTGCCTGCAGAGAGCGTGAGTCC	−
Clear	K5	*ATXN7*	NM_000333	5 (ch 3, 63938159)	*FHIT*	NM_001166243	2 (ch 3, 61186339)	2751820	TCAAACCGCAGGCATTTCAATCACATTATG|CTTTGAAGCTCAGGAAAGAAGAGAAATCCA	−
		RP1–45C12.1		(ch 1, 171137243)	*KIFAP3*	NM_001204514	21 (ch 1, 169890922)	1246321	AGATGACAATTCATGTTGTTTTCAGGAAAG|CCTTGGAATGGATGGCTTTGGCCAACCAGT	−
Clear	K6	*POLR2G*	NM_002696	2 (ch 11, 62529376)	*CYP1A2*	NM_000761	7 (ch 15, 75047132)		CACCGAGGTGGAGGGGACCTGCACAGGGAA|AGAGCTGTGGGAGGACCCCTCTGAGTTCCG	+
Clear	K7	*AC010724.1*		(ch 15, 83207631)	*CPEB1*	NM_001079533	8 (ch 15, 83218408)	10777	AGAGAAAGGAGAGACAATTATGTTCCTGAG|GTGCAGGTGATCCCCTGGGTATTAGCCGAC	+
Clear	K8	*RIC8A*	NM_021932	1 (ch 11, 211349)	*RP11–34P13.7*		(ch 1, 92240)		CTCATCTTCCTAGAGAAGCGTTTGCACAAG|ACCAAACCAATGCAGACCAAACCAATGCAG	−
Clear	K9	*SEMA6A*	NM_020796	18 (ch 5, 115803279)	*CAMK4*	NM_001744	11 (ch 5, 110825288)	4977991	TGTCTTCCCATAATCACCAAGACAAGAAGG|AGGTCCACCAGGGCAAAATAGAAGCCTTGC	−
Clear	K10	*RP11–322M19.1*		(ch 10, 89005872)	*MINPP1*	NM_001178118	2 (ch 10, 89268093)	262221	CCTCAGCCTCCAAAGTAGCTGAGACTACAG|ATATGGAGTTTGGACCTCCAACAGTTAATG	−
Clear	K11	*TEX264*	NM_001129884	3 (ch 3, 51708578)	*FAM107A*	NM_001076778	1 (ch 3, 58594984)	6886406	ATCGCTGTCTACTATGACAACCCCCACATG|AAAAAGAAATGAAGGCCCAGACACGTTACG	−
Clear	K12	*UPF3A*	NM_023011	5 (ch 13, 115052104)	*CDC16*	NM_001078645	15 (ch 13, 115027362)	24742	AGATGGAGGCGAAGACAAGAGAGCTCATTG|GTAACAGTTGACAAATGGGAACCTTTGTTG	−
Clear	K13	*ASAP1*	NM_001247996	29 (ch 8, 131072825)	*ADCY8*	NM_001115	10 (ch 8, 131862049)	789224	CCCGTACCACTGCCCAGAAAAATCAATACG|AGTGATGCCAATGACCATCCAGTTCTCCAT	−

TABLE 2. (Continued)

		Exon boundaries[b]								
		5′-partner gene			3′-partner gene					
Subtype	Sample	Symbol	RNA accession no.	Exon (chromosome, position)	Symbol	RNA accession no	Exon (chromosome, position)	Distance (bp)	Flanking sequences[c]	In-frame transcripts
Clear	K14	*DCUN1D1*	NM_020640	(ch 3, 182679014)	*MCCC1*	NM_020166	18 (ch 3, 182735125)	56111	TTGCAAAGAATCCAGGACAAAAAGGATTAG| GTGTTTGTCAAAGCTGGAGACAAAGTGAAA	−
Clear	K15	*CPSF3*	NM_016207	7 (ch 2, 9576490)	*ASAP2*	NM_001135191	23 (ch 2, 9533611)	42879	GAAGGGCTCAGGAGCTGCTCTTGATTCTAG| ATCCCCTGACCCCCACGCCGCCCCCACCCG	+
Clear	K16	*ADAMTS2*	NM_014244	2 (ch 5, 178770768)	*RP11–798K23.4*		(ch 5, 178930741)	159973	TCTGTGGCGCTCAGCAACTGCGATGGGCTG| ACCTCTGAATAAGTCGTGGGAGCCCTCGGG	−
Clear	K17	*TPPP*	NM_007030	2 (ch 5, 677865)	*TERT*	NM_198253	3 (ch 5, 1282739)	604874	TGACGTGGACATCGTCTTCAGCAAGATCAA| GGGTTGGCTGTGTTCCGGCCGCAGAGCACC	−
(B) Positive controls										
Xp11.2	K69	*EEF2*	NM_001961	1 (ch 19, 3985376)	*ENHO*	NM_198573	2 (ch 9, 34521854)		TGGGAGAATCCACCGCCATCCGCCACCATG| GCTCAGGACTGCAGGTAGACATCTCCACTG	−
		NONO	NM_001145408	10 (ch X, 70517788)	*TFE3*	NM_006521	6 (ch X, 48891766)	21626022	GAAGGATTCAAGGGAACCTTCCCTGATGCG| CTGCCTGTGTCAGGGAATCTGCTTGATGTG	+
		PARG	NM_003631	8 (ch 10, 51093249)	*BMS1*	NM_014753	6 (ch 10, 43287075)	7806174	GCACTCTGTCTGCCAAATATTTGCACCCAG| GGTGTCAAGCTGTTCTACCTTTCTGGAATG	+
		RAGE	NM_014226	1 (ch 14, 102771299)	*EML1*	NM_001008707	2 (ch 14, 100317190)	2454109	TGTCACCGGCTTCCGCATCCAAGATGAAGA| ATGACAGCGCTTCTGCTGCAAGTAGCATGG	+
Xp11.2	K96	*DPP6*	NM_001039350	1 (ch 7, 153584819)	*ACTR3B*	NM_020445	4 (ch 7, 152498706)	1086113	AAGACCGCTAAGATGCAGGGGAACGTGATG| GTATACCAAGCTTGGCTACGCAGGCAACAC	+
		TFE3	NM_006521	4 (ch X, 48895722)	*RBM10*	NM_001204468	18 (ch X, 47044454)	1851268	CTCACCATCGGGTCCAGCTCAGAGAAGGAG| ATTGCCAAGGACATGGAACGCTGGGCCCG	+
		RBM10	NM_001204468	17 (ch X, 47041725)	*TFE3*	NM_006521	5 (ch X, 48895639)	1853914	GAGAAGCACAAGACCAAGACAGCTCAACAG| ATTGATGATGTCATTGATGAGATCATCAGC	+

[a]Chimeric transcripts (33) included the transcripts (*MMACHC-BX004987.7* and *TFE3-RBM10*) consisting of the same partner gene sets with a different exon boundary or different transcriptional direction, and the transcripts sharing a partner gene, *TFE3*.
[b]National Center for Biotechnology Information Database (Genome Build 37).
[c]The exon boundaries are indicated by a vertical bar. Clear, clear cell renal cell carcinoma; Xp11.2, renal cell carcinoma associated with Xp11.2 translocation; ch, chromosome; +, positive; −, negative.