J Breast Cancer. 2019 Dec;22(4):548-561
https://doi.org/10.4048/jbc.2019.22.e55
pISSN 1738-6756·eISSN 2092-9900

Journal of Breast Cancer JBC

Original Article

Check for updates

Prognostic Role of *KRAS* mRNA Expression in Breast Cancer

Ki-Tae Hwang [1], **Byoung Hyuck Kim** [2], **Sohee Oh** [3], **So Yeon Park** [4], **Jiwoong Jung** [5], **Jongjin Kim** [1], **In Sil Choi** [6], **Sook Young Jeon** [7], **Woo-Young Kim** [8]

[1]Department of Surgery, Seoul Metropolitan Government-Seoul National University Boramae Medical Center, Seoul, Korea
[2]Department of Radiation Oncology, Seoul Metropolitan Government-Seoul National University Boramae Medical Center, Seoul, Korea
[3]Department of Biostatistics, Seoul Metropolitan Government-Seoul National University Boramae Medical Center, Seoul, Korea
[4]Department of Pathology, Seoul National University College of Medicine, Seoul, Korea
[5]Department of Surgery, Seoul Medical Center, Seoul, Korea
[6]Department of Internal Medicine, Seoul Metropolitan Government-Seoul National University Boramae Medical Center, Seoul, Korea
[7]Department of Surgery, Graduate School, Kyung Hee University, Seoul, Korea
[8]College of Pharmacy, Sookmyung Women's University, Seoul, Korea

OPEN ACCESS

Received: Aug 15, 2019
Accepted: Nov 15, 2019

Correspondence to
Ki-Tae Hwang
Department of Surgery, Seoul Metropolitan Government-Seoul National University Boramae Medical Center, 20 Boramae-ro 5-gil, Dongjak-gu, Seoul 07061, Korea.
E-mail: kiterius@snu.ac.kr

ORCID iDs
Ki-Tae Hwang
https://orcid.org/0000-0001-6597-3119
Byoung Hyuck Kim
https://orcid.org/0000-0002-6156-0744
Sohee Oh
https://orcid.org/0000-0002-3010-448X
So Yeon Park
https://orcid.org/0000-0002-0299-7268
Jiwoong Jung
https://orcid.org/0000-0003-4146-6475
Jongjin Kim
https://orcid.org/0000-0001-5234-7856
In Sil Choi
https://orcid.org/0000-0002-8494-584X
Sook Young Jeon
https://orcid.org/0000-0003-2888-1623
Woo-Young Kim
https://orcid.org/0000-0002-2895-4995

ABSTRACT

Purpose: We investigated the prognostic role of *KRAS* mRNA expression in breast cancer using The Cancer Genome Atlas (TCGA) and Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) databases.
Methods: Clinical and biological data of 1,093 breast cancers from TCGA database and 1,904 breast cancers from METABRIC database were analyzed. Overall survival (OS) and breast cancer-specific survival (BCSS) were determined.
Results: The group with high *KRAS* mRNA expression showed worse survival than the group with low *KRAS* mRNA expression regarding both OS ($p = 0.012$ in TCGA, $p < 0.001$ in METABRIC) and BCSS ($p = 0.001$ in METABRIC). According to multivariate analysis, the level of *KRAS* mRNA expression was an independent prognostic factor in both TCGA (hazard ratio [HR], 1.570; 95% confidence interval [CI], 1.026–2.403; $p = 0.038$) and METABRIC (HR, 1.254; 95% CI, 1.087–1.446; $p = 0.002$) databases. The prognostic impact of mRNA expression was effective only for luminal A subtype ($p < 0.001$ in METABRIC). Positive correlation was observed between mRNA expression and copy number alteration (CNA) ($r = 0.577$, $p < 0.001$ in TCGA; $\rho = 0.343$, $p < 0.001$ in METABRIC). Methylation showed negative correlations with both mRNA expression and CNA ($r = -0.272$, $p < 0.001$ in TCGA). The expression of mRNA had little association with the mutation status in breast cancers, having a mutation frequency of approximately 0.6%.
Conclusion: *KRAS* mRNA expression was significantly associated with breast cancer prognosis. It was found to be an independent prognostic factor for breast cancer. Prognostic role of *KRAS* mRNA expression was effective only in luminal A subtype. Further studies are needed to validate the prognostic role of *KRAS* mRNA expression in breast cancer, thus paving a way for clinical application of *KRAS* in practice.

Keywords: Breast neoplasms; Prognosis; ras Proteins; RNA

Funding

This work was supported by grant number 03-2016-0400 (2016-1203) from Seoul National University Hospital Research Fund and grant number NRF-2017R1D1A1B03029392 from National Research Foundation.

Conflict of Interest

The authors declare that they have no competing interests.

Author Contributions

Conceptualization: Hwang KT, Choi IS; Data curation: Hwang KT, Kim J, Jeon SY; Formal analysis: Hwang KT, Kim BH, Oh S, Park SY, Kim J; Funding acquisition: Hwang KT; Investigation: Hwang KT, Kim BH, Oh S, Jung J; Methodology: Hwang KT, Kim BH, Oh S, Park SY, Jung J, Choi IS, Jeon SY, Kim WY; Project administration: Hwang KT; Supervision: Hwang KT, Park SY, Choi IS, Kim WY; Validation: Hwang KT, Jung J; Visualization: Hwang KT; Writing - original draft: Hwang KT; Writing - review & editing: Hwang KT, Kim BH, Oh S, Park SY, Jung J, Kim J, Choi IS, Jeon SY, Kim WY.

INTRODUCTION

In humans, *KRAS* proto-oncogene is located at 12p12.1 and consists of 6 exons. It encodes guanosine triphosphatase (GTPase) KRAS isoform, which is a 21-kDa protein with 188 or 189 amino acids. KRAS protein is a member of the RAS superfamily of small GTPases [1]. A single amino acid substitution in KRAS is responsible for an activating mutation. It is known to be associated with human malignancies such as pancreas ductal carcinoma [2], colorectal carcinoma [3], lung adenocarcinoma [4], and various other human malignancies [5,6].

Little is known about the role of *KRAS* in breast cancer. Some of the previous studies have reported the potential role of *KRAS* mutation in endocrine resistance of luminal breast cancer [7,8]. Genome-wide screening results showing an association between *KRAS* status and breast cancer have been reported [9-11]. Other studies have shown the association of *KRAS* 3′ untranslated region (UTR) variants with increased risk of breast cancer [12-14]. The prognostic role of *KRAS* in breast cancer and the role of *KRAS* mRNA expression status *per se* in human malignancy remain unclear. Although many previous studies have determined the impact of the activating mutation of *KRAS* on human malignancies, little is known about the significance of other molecular statuses such as copy number alteration (CNA), methylation, mRNA expression, and protein expression. Therefore, the following 2 well-organized public databases were utilized for this study. The Cancer Genome Atlas (TCGA) Research Network has profiled a large number of human tumors including breast cancer to discover molecular aberrations at the DNA, RNA, protein, and epigenetic levels [15]. Regarding breast cancer, TCGA Research Network has previously reported the comprehensive molecular portraits of invasive ductal and invasive lobular carcinomas [9,10]. The Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) has released genomic and transcriptomic architecture of 2,000 breast cancers [16,17]. METABRIC database is believed to provide landscapes for understanding how somatic CNA affects gene expression and reveal novel subgroups that could be the target for future investigation in breast cancer.

The objective of this study was to investigate the prognostic role of *KRAS* mRNA expression in breast cancer using TCGA and METABRIC databases. Molecular regulation of *KRAS* was also investigated, including mRNA expression, methylation, CNA, and mutation.

METHODS

Data acquisition

TCGA and METABRIC databases were acquired from the following websites: TCGA (https://cancergenome.nih.gov/), Synapse (https://www.synapse.org/), and cBioPortal (http://www.cbioportal.org/). Currently, cBioPortal contains datasets of 233 cancer genomics studies including 14 breast cancer studies. TCGA dataset on breast invasive carcinoma (TCGA, Provisional) and METABRIC dataset on breast cancer (METABRIC, Nature 2012) were utilized for this study (access date: July 11, 2018). The Institutional Review Board of Seoul Metropolitan Government-Seoul National University Boramae Medical Center approved this study (approval number: 17-2018-23). The informed consent of this study was waived.

Clinicopathologic parameters

Patient age was defined as the age at the time of initial diagnosis of primary breast cancer. TNM categories and anatomic stage groups were described according to the breast cancer

staging system of the American Joint Committee on Cancer. The status of estrogen receptor or progesterone receptor was described based on the result of immunohistochemical test for each receptor. Hormone receptor (HRc) status was defined according to the statuses of estrogen and progesterone receptors. Human epidermal growth factor receptor 2 (HER2) status was defined according to the results of immunohistochemical test and *in situ* hybridization assay. Regarding TCGA dataset, the breast cancer subtypes were classified according to the statuses of HRc and HER2. Regarding METABRIC dataset, the breast cancer subtypes were described according to the PAM50 classification.

Biological parameters

RNA sequencing (RNA seq) data were obtained from two different datasets: RNA seq median value by Illumina RNA seq version 2 RSEM and RNA seq z-score. RNA microarray (RNA mic) data were provided as median values of raw data by Agilent microarray analysis. Methylation data were provided as beta values of raw data. CNA linear (CNA_lin) data were provided as relative linear values for each gene by Affymetrix SNP 6. CNA non-linear (CNA non) data were provided as non-linear values by Genomic Identification of Significant Targets in Cancer (GISTIC) 2.0 with values of −2, −1, 0, 1, and 2, which represented homozygous deletion, hemizygous deletion, neutral or no change, gain, and high-level amplification, respectively. Mutation data were generated by whole exome sequencing and provided in a mutation annotation format. Biological data types and sample numbers utilized in this study are described in **Supplementary Table 1**. Biological parameters were classified into low or high based on the mean value of each parameter. RNA seq z-scores were classified into down-regulation, normal-regulation, and up-regulation with cut-off values of −2 and +2, respectively.

Statistical analyses

Two-sample *t*-test was used to determine the difference in expression levels of biological parameters while Pearson's χ^2 test was used to determine the difference in clinicopathologic characteristics between groups. Pearson correlation coefficient (r) was used to evaluate the bivariate correlation between biological parameters and continuous variables. Spearman correlation coefficient (ρ) was used for biological parameters with ordinal variables. Survival analyses were carried out regarding overall survival (OS) and breast cancer-specific survival (BCSS). Kaplan-Meier estimation was used to analyze survival rates while log-rank test was used to determine the significance of differences between 2 survival curves. Cox-proportional hazards model was used for univariable and multivariable analyses. Hazard ratio (HR) was calculated using 95% confidence interval (CI). All statistical analyses were conducted using IBM SPSS Statistics version 20.0 (IBM Inc., Armonk, USA) and R software version 3.3.2 (R Foundation for Statistical Computing, Vienna, Austria). Statistical significance was considered when p-value was less than 0.05.

RESULTS

Clinicopathologic characteristics of study subjects according to expression level of *KRAS* mRNA

Total number of study subjects was 1,093 from the TCGA dataset and 1,904 from the METABRIC dataset. The mean follow-up period regarding OS was 40.9 ± 39.2 months (range, 0–283 months) from the TCGA dataset and 125.0 ± 76.3 months (range, 0–355 months) from the METABRIC dataset. The mean follow-up period regarding BCSS was 98.2 ± 60.4 months (range, 0–307 months) from the METABRIC dataset. Baseline clinicopathologic

https://doi.org/10.4048/jbc.2019.22.e55

Table 1. Clinicopathologic characteristics according to expression level of *KRAS* mRNA in TCGA and METABRIC databases

Group	TCGA (RNA seq, median value)					METABRIC (RNA mic, median value)				
	Subgroup	Low	High	Sum	p^*	Subgroup	Low	High	Sum	p^*
All		657 (60.1)	436 (39.9)	1,093 (100.0)			1,026 (53.9)	878 (46.1)	1,904 (100.0)	
Age (yr)	≤ 50	198 (30.1)	131 (30.0)	329 (30.1)	0.717	≤ 50	236 (23.0)	199 (22.7)	435 (22.8)	0.861
	> 50	458 (69.7)	305 (70.0)	763 (69.8)		> 50	790 (77.0)	679 (77.3)	1,469 (77.2)	
	Unknown	1 (0.2)	0 (0.0)	1 (0.1)		Unknown	0 (0.0)	0 (0.0)	0 (0.0)	
T category	1	172 (26.2)	107 (24.5)	279 (25.5)	0.198	≤ 2 cm	459 (44.7)	362 (41.2)	821 (43.1)	0.172
	2	379 (57.7)	254 (58.3)	633 (57.9)		> 2 cm	553 (53.9)	508 (57.9)	1,061 (55.7)	
	3	87 (13.2)	51 (11.7)	138 (12.6)		Unknown	14 (1.4)	8 (0.9)	22 (1.2)	
	4	17 (2.6)	23 (5.3)	40 (3.7)						
	Unknown	2 (0.3)	1 (0.2)	3 (0.3)						
N category	0	317 (48.2)	198 (45.4)	515 (47.1)	0.791	0	537 (52.3)	446 (50.8)	983 (51.6)	0.714
	1	215 (32.7)	146 (33.5)	361 (33.0)		1	321 (31.3)	282 (32.1)	603 (31.7)	
	2	67 (10.2)	53 (12.2)	120 (11.0)		2	113 (11.0)	107 (12.2)	220 (11.6)	
	3	45 (6.8)	32 (7.3)	77 (7.0)		3	46 (4.5)	39 (4.4)	85 (4.5)	
	Unknown	13 (2.0)	7 (1.6)	20 (1.8)		Unknown	9 (0.9)	4 (0.5)	13 (0.7)	
M category	0	529 (80.5)	380 (87.2)	909 (83.2)	0.005	0	764 (74.5)	630 (71.8)	1,394 (73.2)	0.086
	1	12 (1.8)	10 (2.3)	22 (2.0)		1	2 (0.2)	7 (0.8)	9 (0.5)	
	Unknown	116 (17.7)	46 (10.6)	162 (14.8)		Unknown	260 (25.3)	241 (27.4)	501 (26.3)	
Anatomic stage group	1	113 (17.2)	68 (15.6)	181 (16.6)	0.569	1	264 (25.7)	211 (24.0)	475 (24.9)	0.203
	2	380 (57.8)	242 (55.5)	622 (56.9)		2	440 (42.9)	360 (41.0)	800 (42.0)	
	3	144 (21.9)	107 (24.5)	251 (23.0)		3	58 (5.7)	57 (6.5)	115 (6.0)	
	4	12 (1.8)	10 (2.3)	22 (2.0)		4	2 (0.2)	7 (0.8)	9 (0.5)	
	Unknown	8 (1.2)	9 (2.1)	17 (1.6)		Unknown	262 (25.5)	243 (27.7)	505 (26.5)	
Estrogen receptor	Negative	136 (20.7)	101 (23.2)	237 (21.7)	0.371	Negative	205 (20.0)	240 (27.3)	445 (23.4)	< 0.001
	Positive	494 (75.2)	312 (71.6)	806 (73.7)		Positive	821 (80.0)	638 (72.7)	1,459 (76.6)	
	Unknown	27 (4.1)	23 (5.3)	50 (4.6)		Unknown	0 (0.0)	0 (0.0)	0 (0.0)	
Progesterone receptor	Negative	207 (31.5)	135 (31.0)	342 (31.3)	0.289	Negative	452 (44.1)	443 (50.5)	895 (47.0)	0.005
	Positive	422 (64.2)	276 (63.3)	698 (63.9)		Positive	574 (55.9)	435 (49.5)	1,009 (53.0)	
	Unknown	28 (4.3)	25 (5.7)	53 (4.8)		Unknown	0 (0.0)	0 (0.0)	0 (0.0)	
HER2	Negative	465 (70.8)	311 (71.3)	776 (71.0)	0.289	Negative	926 (90.3)	742 (84.5)	1,668 (87.6)	< 0.001
	Positive	98 (14.9)	75 (17.2)	173 (15.8)		Positive	100 (9.7)	136 (15.5)	236 (12.4)	
	Unknown	94 (14.3)	50 (11.5)	144 (13.2)		Unknown	0 (0.0)	0 (0.0)	0 (0.0)	
Subtype	HRc+HER2−	375 (57.1)	237 (54.4)	612 (56.0)	0.387	Luminal A	397 (38.7)	282 (32.1)	679 (35.7)	< 0.001
	HRc+HER2+	77 (11.7)	59 (13.5)	136 (12.4)		Luminal B	204 (19.9)	257 (29.3)	461 (24.2)	
	HRc−HER2+	21 (3.2)	15 (3.4)	36 (3.3)		HER2	106 (10.3)	114 (13.0)	220 (11.6)	
	HRc−HER2−	89 (13.5)	73 (16.7)	162 (14.8)		Basal	93 (9.1)	106 (12.1)	199 (10.5)	
	Unknown	95 (14.5)	52 (11.9)	147 (13.4)		Others and unknown	226 (22.0)	119 (13.6)	345 (18.1)	

Values are presented as number of patients (%).
TCGA = The Cancer Genome Atlas; METABRIC = Molecular Taxonomy of Breast Cancer International Consortium; HER2, human epidermal growth factor receptor 2; HRc, hormone receptor, RNA mic, RNA microarray; RNA seq, RNA sequencing.
*χ^2 test was used.

characteristics of study subjects are summarized in **Table 1**. In TCGA dataset, there was no significant difference in clinicopathologic factors except M category. In METABRIC dataset, proportions of subjects with positive estrogen receptor or positive progesterone receptor were higher in the group with low *KRAS* expression compared to those in the group with high *KRAS* expression while the group with high *KRAS* expression group showed higher proportion of subjects with positive HER2. Luminal A subtype was more prevalent in the group with low *KRAS* expression.

Survival analyses according to expression level of *KRAS* mRNA

In TCGA dataset, the group with high *KRAS* mRNA expression showed worse OS compared to that shown by the group with low *KRAS* mRNA expression ($p = 0.012$, **Figure 1A**, cut-off value of RNA seq median value = 1,356.9). The group with up-regulated expression of *KRAS* mRNA also showed lower OS rate compared to that shown by the group with normal-

https://doi.org/10.4048/jbc.2019.22.e55

regulated *KRAS* expression ($p < 0.001$, **Figure 1B**). In METABRIC dataset, the group with high *KRAS* mRNA expression showed worse survival rate than the group with low *KRAS* mRNA expression regarding OS ($p < 0.001$, **Figure 1C**) and BCSS ($p = 0.001$, **Figure 1D**). Similarly, the third and fourth quartiles (Q3 and Q4), according to the expression level of *KRAS* mRNA using RNA mic data, showed worse survival rates than the first or second quartiles (Q1 or Q2) regarding both OS and BCSS (**Supplementary Figure 1**). In METABRIC dataset, the group with high *KRAS* mRNA expression showed lower OS rate for luminal A subtype ($p < 0.001$, **Figure 2A**), but not for other subtypes (**Figures 2B-D**). Detailed survival rates are shown in **Supplementary Tables 2** and **3**. In TCGA dataset, the expression of *KRAS* mRNA was lower in luminal A subtype compared to that in any other subtypes (all $p < 0.001$, **Supplementary Figure 2**). There were no significant differences in clinicopathologic factors according to mRNA expression level for luminal A subtype (**Supplementary Table 4**).

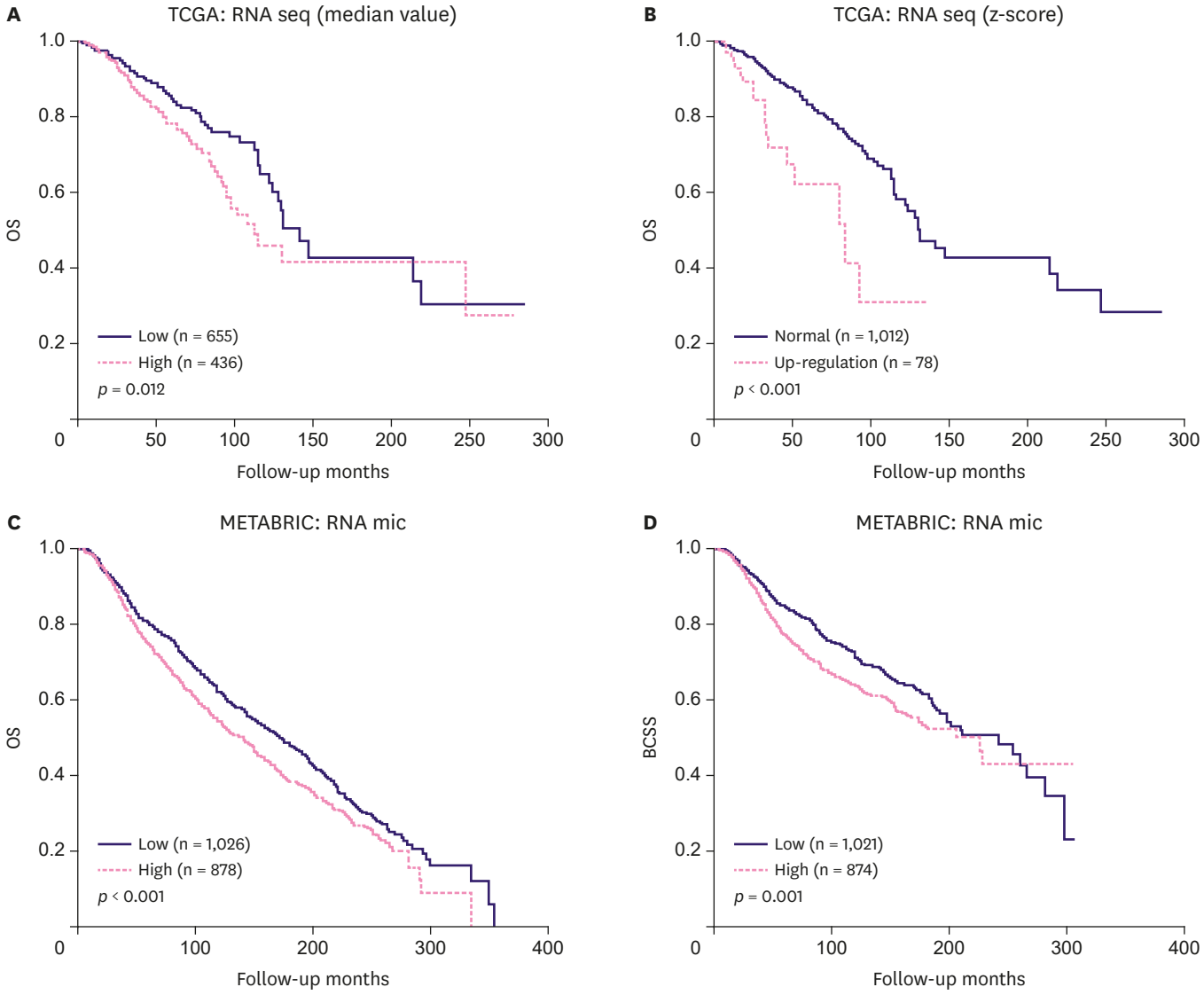

Figure 1. Survival curves according to the expression level of *KRAS* mRNA using TCGA and METABRIC databases. OS curves according to RNA seq median value (A) and RNA seq z-score (B) in TCGA database. OS curves (C) and BCSS curves (D) according to RNA mic in METABRIC database.
OS = overall survival; TCGA = The Cancer Genome Atlas; METABRIC = Molecular Taxonomy of Breast Cancer International Consortium; RNA seq = RNA sequencing; BCSS = breast cancer-specific survival; RNA mic = RNA microarray.

https://doi.org/10.4048/jbc.2019.22.e55

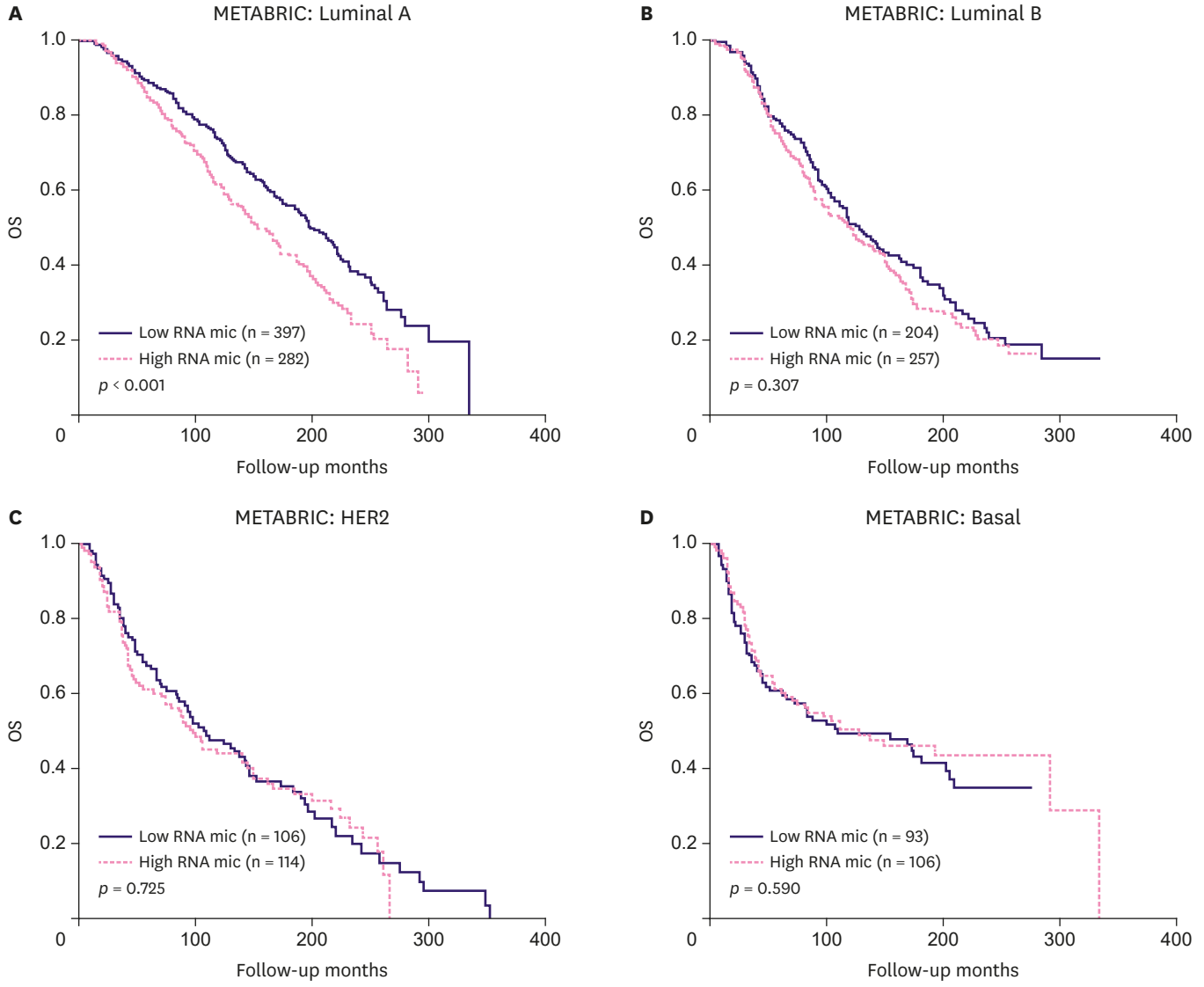

Figure 2. OS curves according to the expression level of *KRAS* mRNA in each subtype of breast cancer using METABRIC database. OS curves in luminal A (A), luminal B (B), HER2 (C), and basal (D) subtypes.
OS = overall survival; METABRIC, Molecular Taxonomy of Breast Cancer International Consortium; HER2 = human epidermal growth factor receptor 2; RNA mic, RNA microarray.

Univariable and multivariable analyses

Expression level of *KRAS* mRNA was a significant factor in univariable analysis of both TCGA (HR, 1.501; 95% CI, 1.091–2.067; $p = 0.013$) and METABRIC (HR, 1.238; 95% CI, 1.100–1.394; $p < 0.001$) datasets. It was also an independent significant prognostic factor for breast cancer according to multivariable analysis of both TCGA (HR, 1.570; 95% CI, 1.026–2.403; $p = 0.038$) and METABRIC (HR, 1.254; 95% CI, 1.087–1.446; $p = 0.002$) datasets after adjusting for age, T category, N category, M category, estrogen receptor, progesterone receptor, and HER2 (**Table 2**).

Correlation analyses among mRNA expression, CNA, and methylation of *KRAS*

In TCGA dataset, the mRNA expression was higher when the value of CNA non increased (**Figure 3A**). The expression of mRNA was positively correlated with the value of CNA_lin ($r =$

https://doi.org/10.4048/jbc.2019.22.e55

Table 2. Univariable and multivariable analyses regarding OS using TCGA and METABRIC databases

Characteristics	TCGA				Characteristics	METABRIC			
	Univariable analysis		Multivariable analysis*			Univariable analysis		Multivariable analysis*	
	HR (95% CI)	*p*	HR (95% CI)	*p*		HR (95% CI)	*p*	HR (95% CI)	*p*
KRAS (RNA seq, median)					KRAS (RNA mic, median)				
Low	Reference		Reference		Low	Reference		Reference	
High	1.501 (1.091–2.067)	0.013	1.570 (1.026–2.403)	0.038	High	1.238 (1.100–1.394)	< 0.001	1.254 (1.087–1.446)	0.002
Age (yr)					Age (yr)				
≤ 50	Reference		Reference		≤ 50	Reference		Reference	
> 50	1.540 (1.080–2.197)	0.017	1.861 (1.146–3.022)	0.012	> 50	1.762 (1.501–2.069)	< 0.001	1.750 (1.445–2.119)	< 0.001
T category		< 0.001		0.007	Tumor size				
T1	Reference		Reference		≤ 2 cm	Reference		Reference	
T2	1.292 (0.859–1.943)	0.219	0.918 (0.533–1.580)	0.757	> 2 cm	1.756 (1.552–1.987)	< 0.001	1.515 (1.302–1.763)	< 0.001
T3	1.579 (0.938–2.661)	0.086	1.391 (0.656–2.950)	0.390					
T4	3.979 (2.145–7.383)	< 0.001	3.906 (1.535–9.940)	0.004					
N category		< 0.001		0.003	N category		< 0.001		< 0.001
N0	Reference		Reference		N0	Reference		Reference	
N1	1.852 (1.253–2.737)	0.002	1.505 (0.879–2.577)	0.136	N1	1.363 (1.190–1.561)	< 0.001	1.298 (1.101–1.530)	0.002
N2	2.741 (1.640–4.583)	< 0.001	3.093 (1.624–5.894)	0.001	N2	2.226 (1.862–2.663)	< 0.001	2.072 (1.669–2.572)	< 0.001
N3	4.103 (2.268–7.423)	< 0.001	3.189 (1.393–7.301)	0.006	N3	4.380 (3.397–5.649)	< 0.001	3.554 (2.552–4.950)	< 0.001
M category					M category				
M0	Reference		Reference		M0	Reference		Reference	
M1	4.889 (2.918–8.191)	< 0.001	2.829 (1.245–6.426)	0.013	M1	3.753 (1.867–7.545)	< 0.001	1.456 (0.702–3.023)	0.313
Anatomic stage group		< 0.001			Anatomic stage group		< 0.001		
Stage I	Reference				Stage I	Reference			
Stage II	1.600 (0.926–2.765)	0.092			Stage II	1.791 (1.524–2.105)	< 0.001		
Stage III	3.004 (1.693–5.329)	< 0.001			Stage III	3.323 (2.584–4.274)	< 0.001		
Stage IV	8.622 (4.3251–7.186)	< 0.001			Stage IV	5.914 (2.9131–2.009)	< 0.001		
Estrogen receptor					Estrogen receptor				
Negative	Reference		Reference		Negative	Reference		Reference	
Positive	0.718 (0.498–1.035)	0.076	0.410 (0.202–0.832)	0.014	Positive	0.849 (0.737–0.978)	0.023	0.899 (0.733–1.102)	0.306
Progesterone receptor					Progesterone receptor				
Negative	Reference		Reference		Negative	Reference		Reference	
Positive	0.737 (0.526–1.032)	0.076	0.935 (0.493–1.774)	0.837	Positive	0.788 (0.700–0.887)	< 0.001	0.944 (0.799–1.115)	0.497
HER2					HER2				
Negative	Reference		Reference		Negative	Reference		Reference	
Positive	1.122 (0.681–1.850)	0.651	1.063 (0.595–1.898)	0.836	Positive	1.450 (1.218–1.725)	< 0.001	1.287 (1.033–1.604)	0.025
Subtype		0.042			Subtype		< 0.001		
HR+HER2–	Reference				Luminal A	Reference			
HR+HER2+	1.244 (0.691–2.239)	0.468			Luminal B	1.508 (1.294–1.757)	< 0.001		
HR–HER2+	1.488 (0.538–4.117)	0.444			HER2	1.706 (1.412–2.062)	< 0.001		
HR–HER2–	1.961 (1.234–3.118)	0.004			Basal	1.349 (1.090–1.668)	< 0.001		

OS = overall survival; TCGA = The Cancer Genome Atlas; METABRIC = Molecular Taxonomy of Breast Cancer International Consortium; HR = hazard ratio; CI = confidence interval; HER2, human epidermal growth factor receptor 2; HRc, hormone receptor; RNA mic, RNA microarray; RNA seq, RNA sequencing.
*KRAS expression was adjusted for 7 factors including age, T category, N category, M category, estrogen receptor, progesterone receptor, and HER2.

0.577, $p < 0.001$, **Figure 3B**). In METABRIC dataset also, the mRNA expression was higher when the value of CNA non increased (**Figure 3C**). The expression of mRNA was positively correlated with CNA non ($\rho = 0.343$, $p < 0.001$, **Figure 3D**). In TCGA dataset, the mRNA expression was negatively correlated with methylation ($r = 0.272$, $p < 0.001$, **Figure 4**). Positive correlation was observed between mRNA expression and CNA_lin. Methylation showed negative correlation with both mRNA expression and CNA_lin (**Supplementary Table 5**).

Analyses of *KRAS* mutation

In TCGA dataset, 6 (0.61%) out of 977 samples showed *KRAS* mutations. All the 6 were missense mutations. In METABRIC dataset, 12 (0.64%) out of 1,871 samples showed *KRAS* mutations, all of which were missense mutations. Details are described in **Supplementary Table 6**.

https://doi.org/10.4048/jbc.2019.22.e55

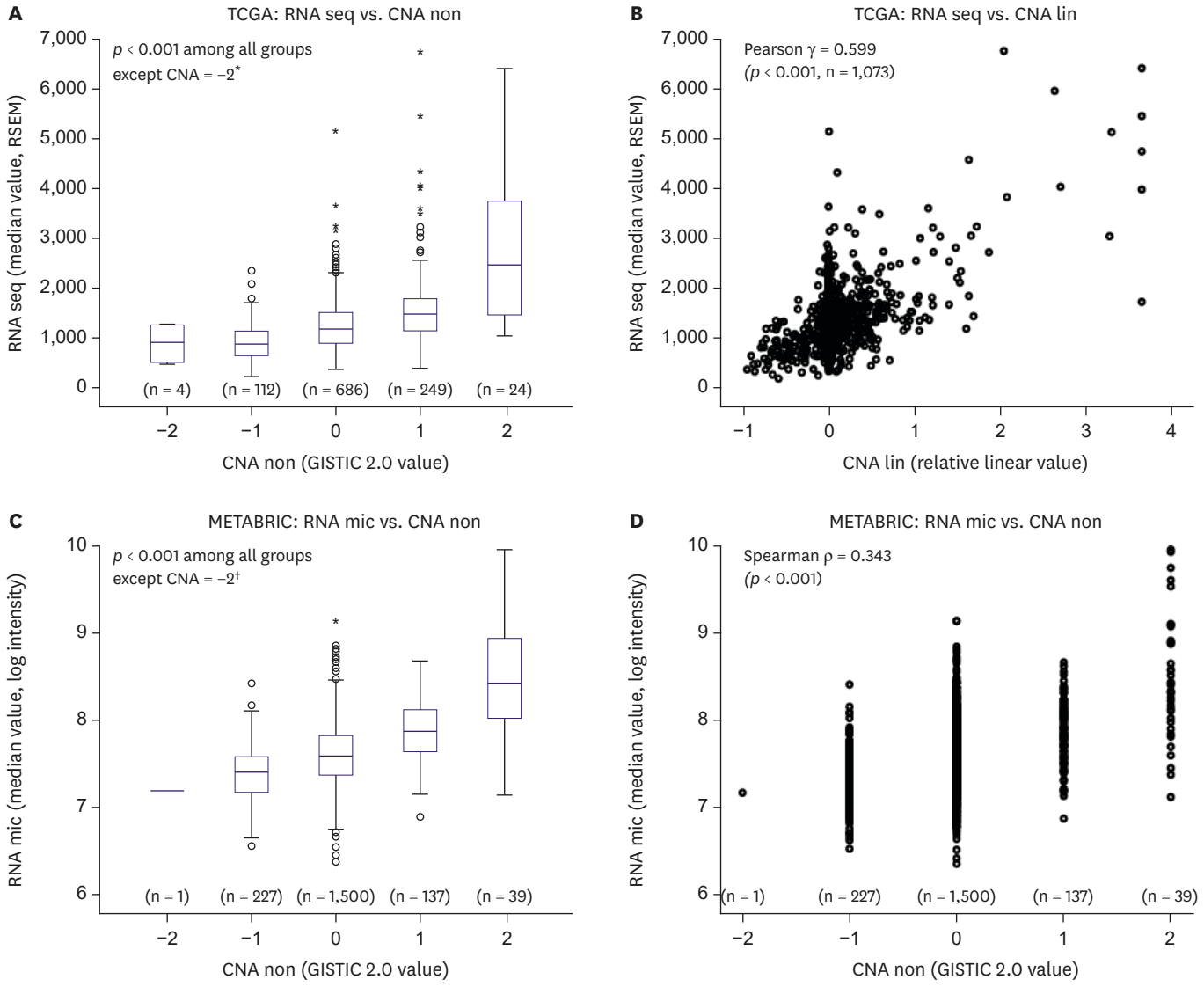

Figure 3. Boxplots and scatter plots for correlation between mRNA expression and CNA of *KRAS*. Boxplots depicting the correlation between RNA seq and nonlinear data of CNA (A), and scatter plot showing the correlation between RNA seq and linear data of CNA (B) using TCGA database. Boxplots depicting the correlation between RNA mic and nonlinear data of CNA (C), and scatter plot showing the correlation between RNA mic and nonlinear data of CNA (D) using METABRIC database.

CNA = copy number alteration; RNA seq = RNA sequencing; TCGA = The Cancer Genome Atlas; RNA mic = RNA microarray; METABRIC = Molecular Taxonomy of Breast Cancer International Consortium; CNA_lin = copy number alteration linear; CNA non = copy number alteration nonlinear; GISTIC = Genomic Identification of Significant Targets in Cancer; RSEM = RNA sequencing by expectation maximization.

*p values by t-test between CNA = −2 and the other groups were 0.890 (CNA = −1), 0.133 (CNA = 0), 0.039 (CNA = 1), and < 0.001 (CNA = 2); †p values by t-test between CNA = −2 and the other groups were 0.513 (CNA = −1), 0.218 (CNA = 0), 0.044 (CNA = 1), and 0.069 (CNA = 2).

DISCUSSION

In 1964, Jennifer Harvey reported that a preparation of murine leukemia virus, later unveiled to carry *Hras* oncogene, caused sarcoma in newborn mice. It was the first report in the history of RAS research [18,19]. In 1970, Werner Kirsten reported the existence of Kirsten murine sarcoma virus, later shown to carry *Kras* oncogene, by serial passage of murine leukemia viruses in Wister-Furth rats [20]. In 1982, human nucleotide sequences of *HRAS* and *KRAS* oncogenes from T24 and EJ bladder carcinoma cell lines were finally reported [21,22].

https://doi.org/10.4048/jbc.2019.22.e55

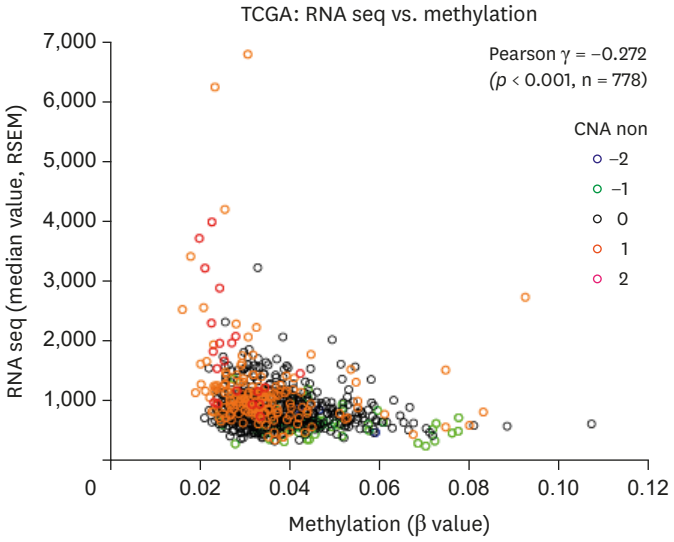

Figure 4. Scatter plot showing correlation between mRNA expression and methylation of *KRAS* using TCGA database.
TCGA = The Cancer Genome Atlas; RNA seq = RNA sequencing; RSEM = RNA sequencing by expectation maximization; CNA non = copy number alteration nonlinear.

Currently, three RAS members (*HRAS*, *KRAS*, and *NRAS*) are clinically the most notable founding members of the RAS subfamily. The RAS superfamily comprises over 150 members [23]. The most characterized RAS signaling pathway is the RAS-RAF-MEK-ERK cascade, which is also known as MAPK/ERK pathway. RAS can also activate several other effector pathways, including PI3K, RalGEF-Ral, and PLCε pathways [1]. Recent studies have reported that there are almost 140 genes whose intragenic mutations can contribute to human cancers [24]. More than 500 human cancer genes have been identified [6]. RAS genes are still the most frequently mutated oncogenes in human cancers followed by *TP53*, *BRAF*, and *PIK3CA*. *KRAS* mutation is dominant among all RAS subfamily members [6].

The results of this study revealed that breast cancer patients with high *KRAS* mRNA expression were significantly associated with worse survival regarding both OS and BCSS. The expression level of *KRAS* mRNA was a significant prognostic factor in breast cancer regarding OS according to both univariable and multivariable analyses. These findings were cross-validated by analyzing two different and independent datasets from TCGA and METABRIC databases. In METABRIC dataset, the group with high *KRAS* mRNA expression showed higher proportions of negative estrogen receptor, negative progesterone receptor, and positive HER2. As a result, the group with high *KRAS* mRNA expression showed lower proportion of luminal A subtype but higher proportions of other subtypes. The expression level of *KRAS* mRNA was a significant prognostic factor by univariable analysis. It remained an independent significant factor after adjusting for main clinicopathologic factors including estrogen receptor, progesterone receptor, HER2, and subtype. This is the first study to reveal the association between mRNA expression of *KRAS* and breast cancer prognosis. Few studies have reported the prognostic role of mRNA or protein expression of *KRAS* in human malignancies. A previous study has investigated the prognostic role of *KRAS* expression using RNA seq data of 1,017 patients with non-small cell lung cancer enrolled in TCGA program [25]. It reported that *KRAS* mRNA expression *per se* was not significantly correlated with OS. However, mRNA expression of surrogate signature genes for *KRAS* deletion was significantly associated with OS (high vs. low expression: HR, 2.3; 95% CI, 1.8–2.9) [25].

https://doi.org/10.4048/jbc.2019.22.e55

In this study, the prognostic impact of *KRAS* mRNA expression was effective only for luminal A subtype. Since there were no significant differences in clinicopathologic features between the two subgroups of luminal A according to mRNA expression level, the difference in *KRAS* mRNA expression level could be a crucial factor to differentiate the two. *KRAS* mRNA expression was significantly lower in luminal A compared to that in other subtypes. This could partly explain the favorable prognosis of luminal A subtype among all breast cancer subtypes. Although little is known about the prognostic role of *KRAS* mRNA expression in subtypes of breast cancer, some studies have reported the association of *KRAS* with endocrine resistance in luminal breast cancer subtype [9]. One study has shown that endocrine-resistant advanced breast tumors harbor oncogenic mutations in the MAPK signaling pathway including *KRAS*. These mutations are typically not detected in pre-treatment of primary tumors [8]. A genome-wide functional screening study has reported a gene set including *KRAS* whose silencing can cause sensitivity to endocrine therapy in breast cancer [11]. Another study has analyzed three independent gene expression datasets from GEO database and reported that RAS pathway activation is strongly associated with poor survival of patients having luminal breast cancers [26]. Further studies are needed to validate the prognostic role of *KRAS* in each breast cancer subtype. Some studies have reported the association between *KRAS* and breast cancer subtypes. One study has demonstrated that *KRAS* has a crucial role in the maintenance of mesenchymal phenotypes and metastatic ability of basal-type breast cancer by molecular experiments using breast cancer cell lines [27]. TCGA Research Network has reported that many components of the PI3K and RAS-RAF-MEK pathways are amplified, but not typically mutated, in basal-like breast cancer subtypes, including *PIK3CA* (49%), *KRAS* (32%), *BRAF* (30%), and *EGFR* (23%) [9].

In the present study, mRNA expression was positively correlated with CNA status but negatively correlated with methylation status. The positive correlation between mRNA expression and CNA was cross-validated using TCGA and METABRIC datasets while the negative correlation between mRNA expression and methylation was only proved in the TCGA dataset due to the lack of methylation data in METABRIC dataset. Little is known about the molecular regulation of *KRAS* at DNA, RNA, and protein levels, and others in breast cancer. TCGA Research Network has reported the landscape of genetic alterations in 10 canonical oncogenic signaling pathways including the RTK-RAS pathway using mutations, CNA, mRNA expression, gene fusions, and DNA methylation from 9,125 human malignancies profiled by TCGA. *KRAS* was the most frequently altered gene (9%), followed by *BRAF* (7%) and *EGFR* (4%) across all human cancers. *KRAS* alterations were the most commonly observed alterations in pancreatic carcinoma (72%), genomically stable colorectal cancer (69%), and lung adenocarcinoma (33%) [28].

Breast invasive carcinoma has been reported to have a rare frequency of *KRAS* mutation (0.7%), although it is the most common one in pancreatic ductal adenocarcinoma (97.7%) followed by colorectal adenocarcinoma (44.7%) and lung adenocarcinoma (30.9%) [6]. This study showed similar *KRAS* mutation frequencies (0.61% and 0.64% in TCGA and METABRIC datasets, respectively). Other studies have also reported rare frequencies of *KRAS* mutation in invasive breast cancers [5,29]. A previous study has shown that frequencies of *KRAS* mutations as: G12 (83%), G13 (14%), Q61 (2%), and others (1%) [6,30]. In the present study, G12 mutation accounted for 83.3% (5 out of 6 in a TCGA dataset and 10 out of 12 in a METABRIC dataset), similar to the previous report. Some previous studies have shown that rs61764370, an inherited variant residing in a *KRAS* 3' UTR microRNA binding site, is associated with increased risk of ovarian cancer and breast cancer and increased

https://doi.org/10.4048/jbc.2019.22.e55

risk of reduced survival [12,13]. Ovarian Cancer Association Consortium, Breast Cancer Association Consortium, and Consortium of Modifiers of BRCA1 and BRCA2 have revealed that rs61764370 is not associated with risk of ovarian or breast cancer or clinical outcome for patients with these cancers by analyzing data of 140,012 women enrolled in these consortia [14]. In the present study, *KRAS* mutation seemed to have little association with mRNA expression having a mutation frequency of approximately 0.6%.

This study has some limitations. First, we could not analyze the role of KRAS protein expression because of insufficient information in TCGA and METABRIC databases. The TCGA dataset has information on protein expression for only 74 subjects (6.8%) while the METABRIC dataset has no information regarding protein expression. Second, although this study revealed the prognostic role of the level of *KRAS* mRNA expression in breast cancer, the roles of other related genes were not analyzed. Third, we could not analyze the recurrence or metastasis pattern according to the level of *KRAS* mRNA expression due to lack of information. Further studies are needed to elucidate associations between *KRAS* mRNA expression and breast cancer prognosis.

In conclusion, high level of *KRAS* mRNA expression was associated with worse prognosis compared to low level of *KRAS* mRNA expression in breast cancer regarding both OS and BCSS. The expression level of *KRAS* mRNA was an independent significant prognostic factor in breast cancer regarding OS. These findings were cross-validated by analyzing two different and independent datasets from TCGA and METABRIC databases. The prognostic role of *KRAS* mRNA expression was effective only for luminal A subtype, which showed significantly lower level of mRNA expression compared to that in other subtypes. The expression level of mRNA was positively correlated with CNA status but negatively correlated with methylation status. The expression of mRNA had little association with mutation status in breast cancer having a mutation frequency of approximately 0.6%. Further studies are needed to validate the prognostic role of *KRAS* mRNA expression in breast cancer, thus paving a way for clinical application of *KRAS* in practice.

ACKNOWLEDGMENTS

We appreciate valuable discussion from the members of the Boramae hospital Breast cancer Study group (BBS). Participant researcher is follow as: Ki-Tae Hwang, MD, PhD[1]; Bo Kyung Koo, MD, PhD[2]; Byoung Hyuck Kim[3], MD, Young A Kim, MD, PhD[4]; Jongjin Kim, MD[1]; Eun Youn Roh, MD, PhD[5]; Sejung Maeng, PhD[1]; Sung Bae Park, MD[6]; Jin Hyun Park, MD, MS[2]; Cho Won Park, RN[1]; Bumjo Oh, MD[7]; So Won Oh, MD, PhD[8]; Sohee Oh, PhD[9]; Jong Yoon Lee, MD, MS[10]; Ji Hyun Chang, MD, PhD[3]; Se Hee Jung, MD, PhD[11]; Young Jun Chai, MD, MS[1]; In Sil Choi, MD, PhD[2]; A Jung Chu, MD[10]; Kyu Ri Hwang, MD, PhD[12]

[1]Department of Surgery; [2]Department of Internal Medicine; [3]Department of Radiation Oncology; [4]Department of Pathology; [5]Department of Laboratory Medicine; [6]Department of Neurosurgery; [7]Department of Family Medicine; [8]Department of Nuclear Medicine; [9]Department of Biostatistics; [10]Department of Radiology; [11]Department of Rehabilitation Medicine; [12]Department of Obstetrics and Gynecology.

All BBS members are from Seoul Metropolitan Government-Seoul National University Boramae Medical Center (20 Boramae-ro 5-gil, Dongjak-gu, Seoul 07061, Korea).

The results of this study are in whole or part based upon data generated by the TCGA Research Network (http://cancergenome.nih.gov/). This study makes use of data generated by the Molecular Taxonomy of Breast Cancer International Consortium. Funding for this project was provided by Cancer Research UK and the British Columbia Cancer Agency Branch. All data used for this study from TCGA and METABRIC datasets are publically available without restrictions.

SUPPLEMENTARY MATERIALS

Supplementary Table 1
Biological data types and sample numbers utilized from TCGA and METABRIC databases

Click here to view

Supplementary Table 2
Detailed survival rates according to mRNA expression level of *KRAS* using TCGA and METABRIC databases

Click here to view

Supplementary Table 3
Detailed survival rates according to mRNA expression level of *KRAS* in each subtype of breast cancer using METABRIC database

Click here to view

Supplementary Table 4
Clinicopathologic characteristics according to mRNA expression level of *KRAS* in luminal A subtype from METABRIC databases

Click here to view

Supplementary Table 5
Correlation analysis among mRNA expression, methylation and CNA of *KRAS* using TCGA database

Click here to view

Supplementary Table 6
List of tumor samples with *KRAS* mutation in TCGA and METABRIC databases

Click here to view

Supplementary Figure 1
Survival curves according to the quartiles of mRNA expression level of *KRAS* in METABRIC databases. OS curves (A) and BCSS curves (B). Q1 is the lowest mRNA expression group and Q4 is the highest mRNA expression group.

Click here to view

Supplementary Figure 2

Box plots showing mRNA expression of *KRAS* in each subtype of breast cancer using METABRIC database.

Click here to view

REFERENCES

1. Goitre L, Trapani E, Trabalzini L, Retta SF. Chapter 1. The Ras superfamily of small GTPases: the unlocked secrets. In: Trabalzini L, Retta SF, editors. Ras Signaling: Methods and Protocols. Methods in Molecular Biology. Vol 1120. Totowa (NJ): Humana Press; 2014, p.1-18.
CROSSREF
2. Biankin AV, Waddell N, Kassahn KS, Gingras MC, Muthuswamy LB, Johns AL, et al. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. Nature 2012;491:399-405.
PUBMED | CROSSREF
3. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. Nature 2012;487:330-7.
PUBMED | CROSSREF
4. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. Nature 2014;511:543-50.
PUBMED | CROSSREF
5. Prior IA, Lewis PD, Mattos C. A comprehensive survey of Ras mutations in cancer. Cancer Res 2012;72:2457-67.
PUBMED | CROSSREF
6. Cox AD, Fesik SW, Kimmelman AC, Luo J, Der CJ. Drugging the undruggable RAS: mission possible? Nat Rev Drug Discov 2014;13:828-51.
PUBMED | CROSSREF
7. McGlynn LM, Kirkegaard T, Edwards J, Tovey S, Cameron D, Twelves C, et al. Ras/Raf-1/MAPK pathway mediates response to tamoxifen but not chemotherapy in breast cancer patients. Clin Cancer Res 2009;15:1487-95.
PUBMED | CROSSREF
8. Razavi P, Chang MT, Xu G, Bandlamudi C, Ross DS, Vasan N, et al. The Genomic landscape of endocrine-resistant advanced breast cancers. Cancer Cell 2018;34:427-438.e6.
PUBMED | CROSSREF
9. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature 2012;490:61-70.
PUBMED | CROSSREF
10. Ciriello G, Gatza ML, Beck AH, Wilkerson MD, Rhie SK, Pastore A, et al. Comprehensive molecular portraits of invasive lobular breast cancer. Cell 2015;163:506-19.
PUBMED | CROSSREF
11. Mendes-Pereira AM, Sims D, Dexter T, Fenwick K, Assiotis I, Kozarewa I, et al. Genome-wide functional screen identifies a compendium of genes affecting sensitivity to tamoxifen. Proc Natl Acad Sci U S A 2012;109:2730-5.
PUBMED | CROSSREF
12. Ratner E, Lu L, Boeke M, Barnett R, Nallur S, Chin LJ, et al. A KRAS-variant in ovarian cancer acts as a genetic marker of cancer risk. Cancer Res 2010;70:6509-15.
PUBMED | CROSSREF
13. Paranjape T, Heneghan H, Lindner R, Keane FK, Hoffman A, Hollestelle A, et al. A 3′-untranslated region KRAS variant and triple-negative breast cancer: a case-control and genetic analysis. Lancet Oncol 2011;12:377-86.
PUBMED | CROSSREF
14. Ovarian Cancer Association Consortium, Breast Cancer Association Consortium, and Consortium of Modifiers of BRCA1 and BRCA2, Hollestelle A, van der Baan FH, Berchuck A, Johnatty SE, Aben KK, et al. No clinical utility of KRAS variant rs61764370 for ovarian or breast cancer. Gynecol Oncol 2016;141:386-401.
PUBMED | CROSSREF

15. Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet 2013;45:1113-20.
PUBMED | CROSSREF
16. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012;486:346-52.
PUBMED | CROSSREF
17. Pereira B, Chin SF, Rueda OM, Vollan HK, Provenzano E, Bardwell HA, et al. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. Nat Commun 2016;7:11479.
PUBMED | CROSSREF
18. Harvey JJ. An unidentified virus which causes the rapid production of tumors in mice. Nature 1964;204:1104-5.
PUBMED | CROSSREF
19. Malumbres M, Barbacid M. RAS oncogenes: the first 30 years. Nat Rev Cancer 2003;3:459-65.
PUBMED | CROSSREF
20. Kirsten WH, Mayer LA. Morphologic responses to a murine erythroblastosis virus. J Natl Cancer Inst 1967;39:311-35.
PUBMED | CROSSREF
21. Shih C, Weinberg RA. Isolation of a transforming sequence from a human bladder carcinoma cell line. Cell 1982;29:161-9.
PUBMED | CROSSREF
22. Goldfarb M, Shimizu K, Perucho M, Wigler M. Isolation and preliminary characterization of a human transforming gene from T24 bladder carcinoma cells. Nature 1982;296:404-9.
PUBMED | CROSSREF
23. Wennerberg K, Rossman KL, Der CJ. The Ras superfamily at a glance. J Cell Sci 2005;118:843-6.
PUBMED | CROSSREF
24. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. Science 2013;339:1546-58.
PUBMED | CROSSREF
25. Nagy Á, Pongor LS, Szabó A, Santarpia M, Győrffy B. KRAS driven expression signature has prognostic power superior to mutation status in non-small cell lung cancer. Int J Cancer 2017;140:930-7.
PUBMED | CROSSREF
26. Wright KL, Adams JR, Liu JC, Loch AJ, Wong RG, Jo CE, et al. Ras signaling is a key determinant for metastatic dissemination and poor survival of luminal breast cancer patients. Cancer Res 2015;75:4960-72.
PUBMED | CROSSREF
27. Kim RK, Suh Y, Yoo KC, Cui YH, Kim H, Kim MJ, et al. Activation of KRAS promotes the mesenchymal features of basal-type breast cancer. Exp Mol Med 2015;47:e137.
PUBMED | CROSSREF
28. Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, et al. Oncogenic signaling pathways in The Cancer Genome Atlas. Cell 2018;173:321-337.e10.
PUBMED | CROSSREF
29. Sánchez-Muñoz A, Gallego E, de Luque V, Pérez-Rivas LG, Vicioso L, Ribelles N, et al. Lack of evidence for KRAS oncogenic mutations in triple-negative breast cancer. BMC Cancer 2010;10:136.
PUBMED | CROSSREF
30. Hobbs GA, Der CJ, Rossman KL. RAS isoforms and mutations in cancer at a glance. J Cell Sci 2016;129:1287-92.
PUBMED | CROSSREF