Check for updates

ORIGINAL ARTICLE

Copy Number Profiling of MammaPrint™ Genes Reveals Association with the Prognosis of Breast Cancer Patients

Areej Fatima*, Fomaz Tariq*, Muhammad Faraz Arshad Malik, Muhammad Qasim[1], Farhan Haq

Cancer Genetics and Epigenetics Lab, Department of Biosciences, COMSATS Institute of Information Technology, Islamabad, Pakistan; [1]Department of Pharmacology & Immune Network Pioneer Research Center, Ajou University School of Medicine, Suwon, Korea

Purpose: The MammaPrint™ gene signature, currently used in clinical practice, provides prognostic information regarding the recurrence and potential metastasis in breast cancer patients. However, the prognostic information of the 70 genes included can only be estimated at the RNA expression level. In this study, we investigated whether copy number information of MammaPrint™ genes at the DNA level can be used as a prognostic tool for breast cancer, as copy number variations (CNVs) are major contributors to cancer progression. **Methods:** We performed CNV profiling of MammaPrint™ genes in 59 breast cancer cell lines and 650 breast cancer patients, using publicly available data in The Cancer Genome Atlas (TCGA) database. Statistical analyses including Fisher exact test, chi-square test, and Kaplan-Meier survival analyses were performed. **Results:** All MammaPrint™ genes showed recurrent CNVs, particularly in TCGA cohort. CNVs of 32 and 36 genes showed significant associations with progesterone receptor and estrogen rector, respectively. No genes showed a significant association with human epidermal growth factor receptor 2 status and lymph node status. In addition, only six genes were associated with tumor stages. *RFC4*, *HRASLS*, *NMU*, *GPR126*, *SCUBE2*, *C20orf46*, and *EBF4* were associated with reduced survival and *RASSF7* and *ESM1* were associated with reduced disease-free survival. **Conclusion:** Based on these findings, a concordance of CNV-based genomic rearrangement with expression profiling of these genes and their putative roles in disease tumorigenesis was established. The results suggested that the CNV profiles of the MammaPrint™ genes can be used to predict the prognosis of breast cancer patients. In addition, this approach may lead to the development of new cancer biomarkers at the DNA level.

Key Words: *Biomarkers, Breast neoplasms, DNA copy number variations, Prognosis*

INTRODUCTION

Breast cancer is one of the most commonly diagnosed cancers in women worldwide and one of the leading causes of cancer-related death among women [1-4]. Because of its heterogeneity, breast cancer recurs in some patients but not in others [5]. Clinically, this heterogeneous disease is categorized based on three basic breast cancer biomarkers defined by immunohistochemistry of estrogen rector (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2

(HER2) [6,7].

A number of whole genome/exome and transcriptome profiling studies have identified different genes involved in breast cancer progression [8]. For instance, insulin growth factors have been implicated in breast tumorigenesis because of their ability to stimulate mitogenesis and promote differentiation [9]. Amplification and overexpression of *ERBB2* and *EGFR* genes were shown to be associated with increased tumor growth in breast cancer [10]. In addition, copy number variations (CNVs) in other genes, including *AKT2*, *MAP3K1*, *MAP3K13*, *NCOR1*, and *SMARCD1*, were also found to be associated with breast carcinogenesis [11,12].

Transcriptome profiling studies led to the identification of gene signatures for the early-stage diagnosis of breast cancer patients [13]. For instance, Oncotype DX® (24 genes), MammaPrint™ (70 genes), and Genomic Grade Index (97 genes) are first-generation gene signatures in breast cancer that predict the early recurrence of cancer (i.e., within 5 years) [14]. Additionally, gene signatures including Prosigna, EndoPredict®,

and Breast Cancer Index[SM] can predict the late recurrence of breast cancer [14]. However, no studies have identified gene signatures specifically using information at the copy number level. A previous study showed that CNVs play pivotal roles in the progression and prognosis of breast cancer patients [15].

Therefore, our main goal was to evaluate whether CNV-level information can be used to develop new breast cancer biomarkers. As a test study, we analyzed the MammaPrint™ gene signature. The well-known, 70-gene MammaPrint™ assay was developed as a prognostic tool for predicting the recurrence of breast cancer [16] and measures the mRNA expression of 70 genes to classify patients as having a lower risk or higher risk of metastasis. According to a previous study, the biological functions of all 70 MammaPrint™ genes are associated with the six hallmarks of cancer [1]. In this study, we evaluated CNV-level information of the MammaPrint™ genes in 59 breast cancer cell lines available from the Cancer Cell Line Encyclopedia database (CCLE) [17] and 650 breast cancer patients available in The Cancer Genome Atlas (TCGA) database [3]. We also analyzed the association of CNVs with different clinical features of breast cancer patients.

METHODS

Study design

The overall study design is presented in Supplementary Figure 1 (available online). Complete genomic and functional information was available for 62 genes of the 70 MammaPrint™ (Agilent, Santa Clara, USA) genes.

For data analysis of 59 breast cancer cell lines, we downloaded logR values generated by the CCLE database (https://portals.broadinstitute.org/ccle). Because the study was CNV-based, we selected stringent criteria to minimize the false-positive rate (i.e., logR values ≥ 1 were considered as amplifications, while logR values ≤ -0.7 were considered as deletions). Additionally, CNV values of breast cancer data for 773 patients was downloaded from TCGA database [3]. The platform used in the cell line and TCGA studies was the Affymetrix SNP 6.0 array (Santa Clara, USA). Clinical information of TCGA dataset was available for only 650 patients, and thus we excluded an additional 123 tumor samples. Clinical information included age, ethnicity, ER status, HER2 status, PR status, TNM staging, recurrence, and overall survival (Table 1).

Statistical analysis

All statistical tests were performed using SPSS version 20.0 (IBM Corp., Armonk, USA), including Fisher exact test, chi-square test, and Kaplan-Meier overall survival and disease-free survival analyses. p-values were two-sided and $p < 0.05$
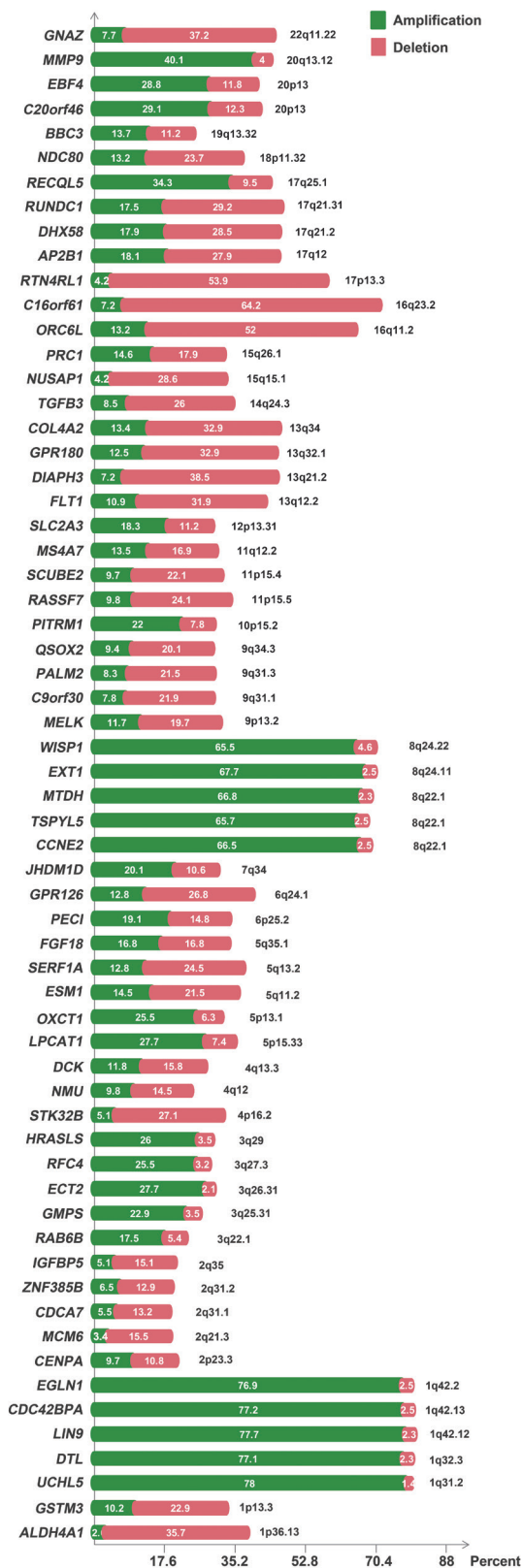

Figure 1. Copy number profiling of MammaPrint™ genes in 650 breast cancer patients.

Table 1. Clinicopathological characteristics of 650 breast cancer patients

Characteristic	No. (%)
Diagnosis age (yr)	
<45	100 (15.3)
≥45	550 (84.7)
Ethnic group	
Not Hispanic or Latino	500 (76.9)
Hispanic or Latino	23 (3.5)
NA	127 (19.5)
Primary tumor site	
Left	345 (53.1)
Right	305 (46.9)
T-stage	
T1	177 (27.3)
T2	371 (57.1)
T3	73 (11.2)
T4	28 (4.3)
TX	1 (0.2)
Lymph node stage	
N0	301 (46.31)
N1	222 (34.2)
N2	75 (11.5)
N3	39 (6.0)
NX	13 (2.0)
M-stage	
M0	583 (89.7)
M1	13 (2.0)
MX	54 (8.3)
Neoplasm histologic type	
Infiltrating ductal	526 (81.0)
Infiltrating lobular	64 (9.9)
Mucinous carcinoma	6 (1.0)
Medullary carcinoma	4 (0.7)
Other	49 (7.5)
NA	1 (0.0)
ER	
Positive	476 (73.2)
Negative	139 (21.4)
Indeterminate	2 (0.3)
NA	33 (5.1)
PR status	
Positive	416 (64.0)
Negative	196 (30.2)
Indeterminate	4 (0.6)
NA	34 (5.2)
HER2	
Positive	100 (15.4)
Negative	333 (51.2)
Equivocal	105 (16.5)
Indeterminate	7 (1.1)
NA	105 (16.2)
Overall survival status	
Living	555 (85.4)
Deceased	95 (14.6)

NA=not applicable; TX=tumor unknown; NX=node unknown; MX=metastasis unknown; ER=estrogen receptor; PR=progesterone receptor; HER2=human epidermal growth factor receptor 2.

was considered statistically significant.

RESULTS

Copy number variations of MammaPrint™ genes in 59 breast cancer cell lines

By cell line data analysis, numerous recurrent amplifications and deletions were observed (Supplementary Figure 2, available online). Genes present in the 1q31.2-1q42.2, 3q26.31-3q29, 5p13.1-5p15.33, 8q22.1-8q24.22, 17q25.1, and 20q13.12 chromosomal regions showed amplifications in >5% of breast cancer cell lines. In addition, genes present in the 1p13.3-1p36.13, 4p16.2, 6p25.2, 6q24.1, 11p15.5-11p15.4, 11q12.2, 12p13.31, 13q12.2-13q34, 16q11.2-16q23.2, 17p13.3, 17q12-17q21.31, 18p11.32, and 20p13 chromosomal regions showed deletions in >5% of breast cancer cell lines (Supplementary Figure 2, available online). The most recurrently amplified gene was *EXT1*, which was present in the 8q24 region, with amplification in 23% of cell lines. The most commonly deleted gene was *DIAPH3*, which was present in the 13q21 region, with deletions in 20% of cell lines (Supplementary Figure 2, available online).

Copy number variations of MammaPrint™ genes in 650 breast cancer patients

Interestingly, compared to cell line data, tumor data showed significant differences in the frequencies of amplifications and deletions (Figure 1). According to the results of TCGA data, genes present in the 1q31.2-1q42.2, 3q22.1-3q29, 5p13.1-5p15.33, 6p25.2, 7q34, 8q22.1-8q24.22, 10p15.2, 12p13.31, 17q25.1, 19q13.32, 20p13, and 20q13.12 chromosomal regions showed recurrent amplifications (Figure 1). In addition, genes present in the 1p13.3-1p36.13, 2p23.3, 2q21.3-2q35, 4p16.2, 4q12-4q13.3, 5q11.2-5q13.2, 6q24.1, 9p13.2, 9q31.1-9q34.3, 11p15.5-11p15.4, 11q12.2, 13q12.2-13q34, 14q24.3, 15q15.1-15q26.1, 16q11.2-16q23.2, 17p13.3, 17q12-17q21.31, 18p11.32, and 22q11.22 chromosomal regions showed recurrent deletions. The 8q region containing *WISP1*, *EXT1*, *MTDH*, *TSPYL5*, and *CCNE2* was the most recurrently amplified region, while the 16q region containing *C16orf61* and *ORC6L* was the most recurrently deleted region. *DIAPH3* showed the highest number of deletions in the cell line data and a significant number of deletions (38.5%) in TCGA data.

Genetic alterations of MammaPrint™ genes and clinicopathological features

Next, the association of CNVs of MammaPrint™ genes was evaluated against multiple clinical and pathological factors, including age, hormone receptors, tumor stages, recurrence,
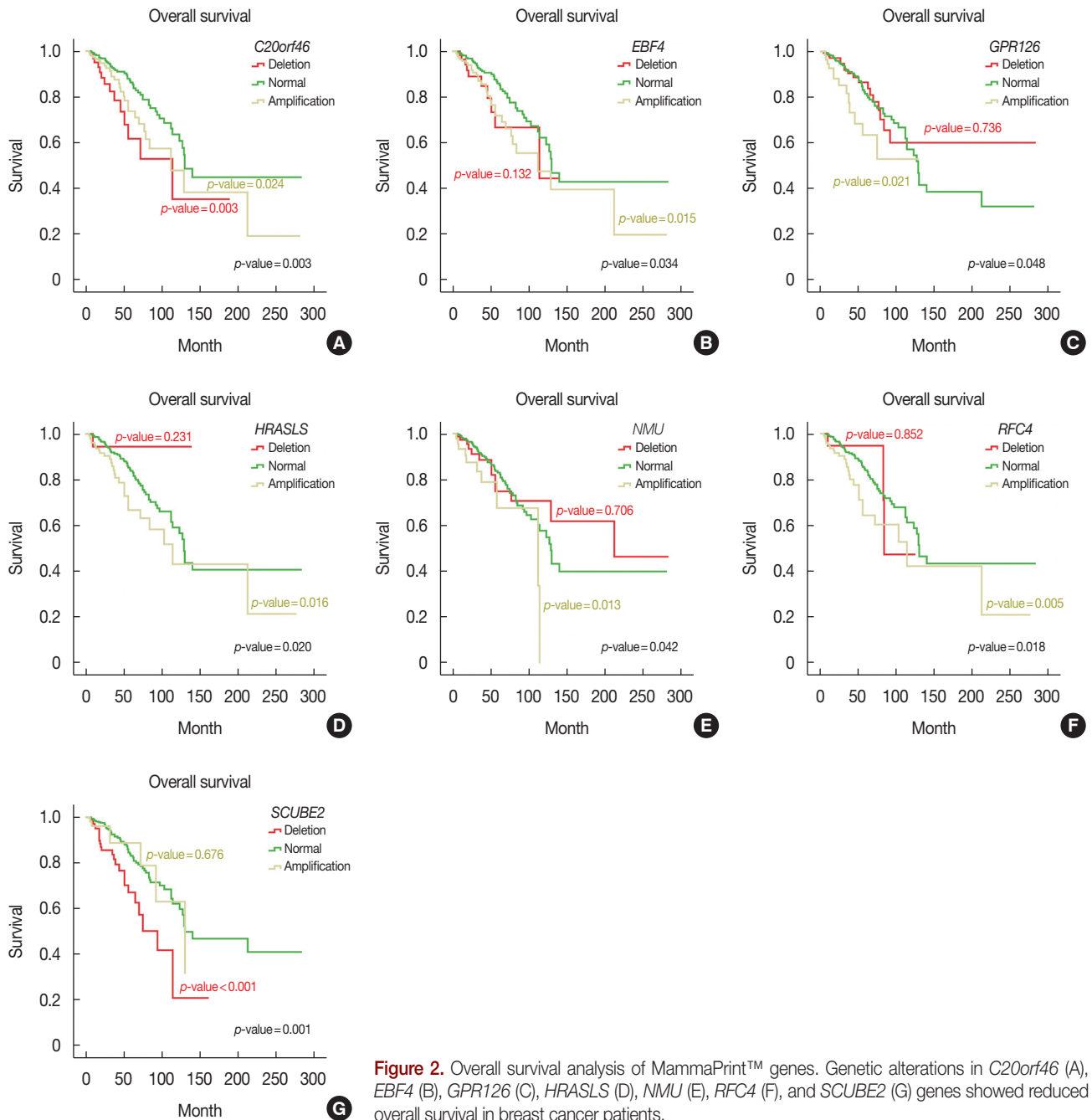
Figure 2. Overall survival analysis of MammaPrint™ genes. Genetic alterations in *C20orf46* (A), *EBF4* (B), *GPR126* (C), *HRASLS* (D), *NMU* (E), *RFC4* (F), and *SCUBE2* (G) genes showed reduced overall survival in breast cancer patients.

and survival. The median age of the 650 breast cancer patients was 59 years (range, 26–90 years) and all patients were female. Of the 650 patients, 476 (73.2%) were ER positive, 139 (21.4%) were ER negative, 100 (15.4%) were HER2 positive, 333 (51.2%) were HER2 negative, 416 (64.0%) were PR positive, and 196 were PR negative (30.2%). A total of 583 (89.7%) had M0 stage breast cancer and 13 (2.0%) had M1 stage breast cancer. In addition, 177 patients (27.3%) had T1 stage, 371 (57.1%) had T2 stage, 73 (11.2%) had T3 stage, and 28 (4.3%) had T4

stage. Five hundred patients (76.9%) were not of Hispanic or Latino origin, while 23 (3.5%) were of Hispanic or Latino origin. According to survival data, 555 patients (85.4%) were alive and 95 (14.6%) were dead at the time of the last follow-up (Table 1).

According to the analysis, PR positive status and ER positive status were associated with 32 and 36 MammaPrint™ genes, respectively (Table 2). Deletions in 20 genes were associated with PR positive and amplification of 12 genes were associated with PR positive status. Similarly, deletion of 25 genes showed

Table 2. Association of copy number variations with biomarkers

Serial No.	Gene	ER	PR	HER2
1	LPCAT1	NA	Positive-amplification	NA
2	OXCT1	Positive-amplification	Positive-amplification	NA
3	DTL	Positive -amplification	Positive-amplification	NA
4	MMP9	Positive-amplification	Positive-amplification	NA
5	COL4A2	Positive-deletion	Positive-deletion	NA
6	DHX58	Positive-deletion	Positive-deletion	NA
7	NDC80	Positive-deletion	NA	NA
8	ZNF385B	Positive-deletion	Positive-deletion	NA
9	GSTM3	Positive-deletion	Positive-deletion	Negative-deletion
10	ECT2	Positive-amplification	Positive-amplification	NA
11	RFC4	Positive-amplification	Positive-amplification	NA
12	JHDM1D	Positive-amplification	NA	Negative-amplification
13	AP2B1	Positive-deletion	Deletion-negative	NA
14	SLC2A3	NA	Negative-amplification	NA
15	EBF4	Positive-amplification	Positive-amplification	NA
16	GMPS	Positive-amplification	NA	NA
17	MELK	Positive-deletion	NA	NA
18	CENPA	Positive-deletion	Positive-deletion	NA
19	SERF1A	Negative-deletion	Positive-amplification	NA
20	DIAPH3	NA	Positive-deletion	NA
21	RUNDC1	Positive-deletion	Negative-deletion	NA
22	RASSF7	Positive-deletion	Positive-deletion	NA
23	GPR180	Positive-deletion	Positive-deletion	NA
24	ORC6L	Positive-deletion	NA	NA
25	GPR126	Positive-deletion	Positive-deletion	NA
26	ALDH4A1	NA	Positive-deletion	NA
27	SCUBE2	Positive-deletion	Positive-deletion	NA
28	ESM1	Positive-amplification	Positive-amplification	NA
29	BBC3	NA	Positive-amplification	NA
30	GNAZ	Positive-deletion	Positive-deletion	NA
31	IGFBP5	Positive-deletion	NA	NA
32	PITRM1	Negative-amplification	Negative-amplification	NA
33	QSOX2	Positive-deletion	Positive-deletion	NA
34	MCM6	Positive-deletion	Positive-deletion	NA
35	NUSAP1	Positive-deletion	Positive-deletion	NA
36	TGFB3	Positive-deletion	Positive-deletion	NA
37	PRC1	Positive-deletion	Positive-deletion	NA
38	NMU	Positive-deletion	NA	NA
39	CDCA7	Positive-deletion	Positive-deletion	NA
40	FGF18	Positive-amplification	Positive-amplification	NA
41	PECI	Positive-deletion	Positive-deletion	Negative-deletion
42	C16orf61	Positive-deletion	Positive-deletion	NA
43	C20orf46	Positive-amplification	Positive-amplification	NA

Genes with no association include *EXT1, TSPYL5, CCNE2, MTDH, WISP1, CSC42BPA, LIN9, EGLN1, RECQL5, HRASLS, UCHL5, RTN4RL1, DCK, STK32B, C9orf30, FLT1, MS4A7, RAB6B, PALM2*.
ER = estrogen receptor; PR = progesterone receptor; HER2 = human epidermal growth factor receptor 2; NA = not applicable.

an association with ER positive and amplification of 11 genes showed an association with ER positive status. No gene was associated with HER2 positive status. In addition, PR negative, ER negative, and HER2 negative statuses were associated with only four, two, and three genes, respectively (Table 2).

According to TNM staging results, deletions in *ZNF385B*

($p = 0.011$), *CDCA7* ($p = 0.002$), *CENPA* ($p = 0.035$), *COL4A2* ($p = 0.006$), *GPR180* ($p = 0.013$), and *DIAPH3* ($p = 0.041$) were associated with T2 stage, and deletions in *CDCA7* ($p = 0.009$) and *COL4A2* ($p = 0.043$) were associated with N1 stage. In addition, a deletion in *NDC80* was associated with the M0 stage of breast cancer ($p = 0.027$) (Supplementary Figure 3, available
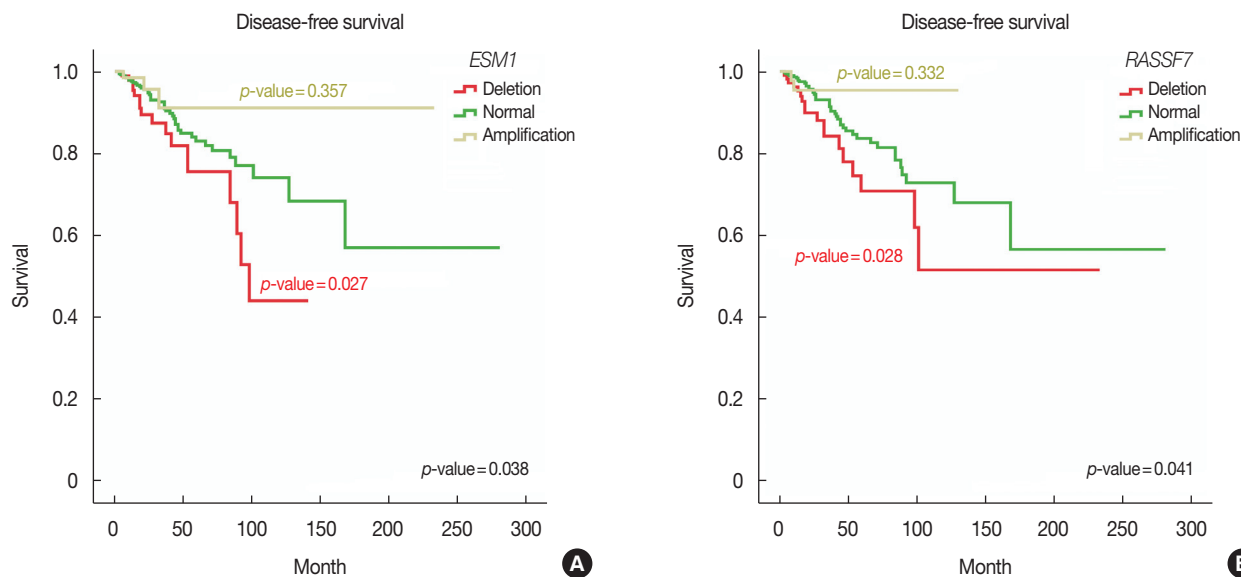
Figure 3. Disease-free survival analysis of MammaPrint™ genes. Genetic alterations in *ESM1* (A) and *RASSF7* (B) genes showed poor disease-free survival in breast cancer patients.

online). According to the tumor primary site, deletions in *FLT1* ($p = 0.019$) and amplifications in *LPCAT1* ($p = 0.034$), *PITRM1* ($p = 0.023$), *C20orf46* ($p = 0.042$), and *MMP9* ($p = 0.003$) were associated with the left tumor site of breast cancer (Supplementary Figure 3, available online).

Overall survival and disease-free survival

Furthermore, we performed Kaplan-Meier survival analysis of all MammaPrint™ genes to calculate overall and disease-free survival. Interestingly, amplifications in *RFC4*, *HRASLS*, *NMU*, *EBF4*, and *GPR126* genes indicated poor overall survival, with Mantel-Cox *p*-values of 0.018, 0.020, 0.042, and 0.048, respectively. Deletion of *SCUBE2*, *C20orf46*, and *EBF4* was associated with poor survival, with Mantel-Cox *p*-values of 0.001, 0.003, and 0.034, respectively (Figure 2). In addition, deletions in *RASSF7* and *ESM1* were associated with poor disease-free survival of breast cancer patients, with Mantel-Cox *p*-values of 0.041 and 0.038, respectively (Figure 3).

These analyses revealed that the CNV profiles of all MammaPrint™ genes contained recurrent amplifications and deletions. In addition, MammaPrint™ genes were associated with hormone receptor status, recurrence, and survival of breast cancer patients.

DISCUSSION

To understand breast cancer tumor heterogeneity, routinely used clinical and pathological markers are gradually being replaced with genetic assays. Based on these genomic tests, con-

sensus regarding the classification of breast cancer subtypes, prognosis, and therapeutic strategies were also rationalized at the 14th St Gallen International Breast Cancer Conference [18, 19]. The Breast Cancer Index[SM], MammaPrint™, and Oncotype DX® kits were developed to predict breast cancer patient prognosis. The MammaPrint™ is the only Food and Drug Administration approved kit used to assess chances of disease recurrence [1]. Based on differential expressional profiling of 70 genes, prospective estimation of relapse, chemotherapy response, and optional treatment plans may also be devised. A probable concordance of genomic rearrangements (amplification/deletion) for these 70 genes with their clinicopathological features was assessed. CNVs have emerged as major contributors to cancer progression [19].

According to high-throughput sequencing data of recent cancer genomes, CNVs are associated with many clinicopathological factors [20]. However, no CNV-related gene signatures have been identified in cancer. The major reason for this is that copy number segments are independent of gene length (ranging from a few kilobases to several megabases). Therefore, it is very challenging to identify driver genes among the many genes in a CNV segment [21]. Thus, we selected a set of genes shown to be associated with the prognosis of breast cancer at the RNA level and evaluated their association between their CNV information and different clinicopathological factors.

CNV analysis of these MammaPrint™ genes suggested that copy number information can be used to develop new cancer-related gene signatures. Interestingly, in TCGA dataset, all

genes showed recurrent genetic alterations (at least 10% amplifications/deletions). The most recurrently amplified genes were present in the 8q and 1q regions, while the most recurrently deleted genes were present in the 16q and 17p regions (Figure 1). Notably, more than 50% of genes were found to be associated with ER positive and PR positive status, but not a single gene was associated with HER2 positive status.

Kaplan-Meier analysis revealed numerous genes associated with reduced survival and recurrence of breast cancer. For instance, *EBF4* and *HRASLS* were associated with reduced overall survival. Previously, involvement of *EBF4* and *HRASLS* in oncogenic transformation and self-sufficiency for primary tumor proliferation was established [22,23]. Data related to molecular cross-talk involving *EBF4* remains limited. However, based on sequence homology (90%–95%) with the EBF family, an influence on cell differentiation during embryogenesis was reported [24]. *RFC4* overexpression with advanced tumor stage and grade and poor prognosis was also established in colorectal cancer patients. As a multimeric protein, *RFC4* is responsible for loading the clamp during DNA replication and cell proliferation. Increased CNV values of *RCF4* indicate its dysregulated expression in breast cancer patients [25]. Similarly, in human dermal microvascular endothelial cells, conditional *GPR126* expression mediates angiogenesis via *VEGF2* activation [26]. Interestingly, *GPR126* upregulation was also reported in acute myeloid leukemia-affected patients suffering from variable mixed-lineage leukemia translocation, indicating its subtyping-specific tumorigenic role [27]. Furthermore, these results are consistent with our finding that inactivation of *SCUBE2* is associated with poor prognosis of breast and colorectal carcinoma [28].

Similarly, deletion of *ESM1* with early recurrence has also been observed in prostate cancer [29]. *RASSFF* contains the conserved motif *RalGDS/AF6* Ras association domain and is responsible for the suppression of guanosine-5´-triphosphate-Ras-induced cell proliferation [30]. Thus, the deletion observed in the given cohort supports the notion of disease recurrence.

Based on these findings, a concordance in CNV based genomic rearrangement with expression profiling of these genes and their putative roles in disease tumorigenesis was established. Therefore, the CNV profiles of RNA expression gene signatures can reveal cancer driver chromosomal regions at the DNA level. This approach will ultimately lead to the development of different therapeutic strategies against these cancer driver regions.

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

REFERENCES

1. Tian S, Roepman P, Van't Veer LJ, Bernards R, de Snoo F, Glas AM. Biological functions of the genes in the mammaprint breast cancer profile reflect the hallmarks of cancer. Biomark Insights 2010;5:129-38.

2. Ellsworth RE, Decewicz DJ, Shriver CD, Ellsworth DL. Breast cancer in the personal genomics era. Curr Genomics 2010;11:146-61.

3. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature 2012;490:61-70.

4. Sharma GN, Dave R, Sanadya J, Sharma P, Sharma KK. Various types and management of breast cancer: an overview. J Adv Pharm Technol Res 2010;1:109-26.

5. Hwang KT, Han W, Cho J, Lee JW, Ko E, Kim EK, et al. Genomic copy number alterations as predictive markers of systemic recurrence in breast cancer. Int J Cancer 2008;123:1807-15.

6. Dai X, Xiang L, Li T, Bai Z. Cancer hallmarks, biomarkers and breast cancer molecular subtypes. J Cancer 2016;7:1281-94.

7. Onitilo AA, Engel JM, Greenlee RT, Mukesh BN. Breast cancer subtypes based on ER/PR and Her2 expression: comparison of clinicopathologic features and survival. Clin Med Res 2009;7:4-13.

8. Verigos J, Magklara A. Revealing the complexity of breast cancer by next generation sequencing. Cancers (Basel) 2015;7:2183-200.

9. Deming SL, Ren Z, Cai Q, Shu XO, Wen W, Long JR, et al. IGF-I and IGF-II genetic variation and breast cancer risk in Chinese women: results from the Shanghai Breast Cancer Study. Cancer Epidemiol Biomarkers Prev 2008;17:255-7.

10. Hwangbo W, Lee JH, Ahn S, Kim S, Park KH, Kim CH, et al. EGFR gene amplification and protein expression in invasive ductal carcinoma of the breast. Korean J Pathol 2013;47:107-15.

11. Chen Y, Hao J, Jiang W, He T, Zhang X, Jiang T, et al. Identifying potential cancer driver genes by genomic data integration. Sci Rep 2013; 3:3538.

12. Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, et al. The landscape of cancer genes and mutational processes in breast cancer. Nature 2012;486:400-4.

13. Urquidi V, Goodison S. Genomic signatures of breast cancer metastasis. Cytogenet Genome Res 2007;118:116-29.

14. Győrffy B, Hatzis C, Sanft T, Hofstatter E, Aktas B, Pusztai L. Multigene prognostic tests in breast cancer: past, present, future. Breast Cancer Res 2015;17:11.

15. Ping Z, Siegal GP, Almeida JS, Schnitt SJ, Shen D. Mining genome sequencing data to identify the genomic features linked to breast cancer histopathology. J Pathol Inform 2014;5:3.

16. Wittner BS, Sgroi DC, Ryan PD, Bruinsma TJ, Glas AM, Male A, et al. Analysis of the MammaPrint breast cancer assay in a predominantly postmenopausal cohort. Clin Cancer Res 2008;14:2988-93.

17. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 2012;483:603-7.

18. Esposito A, Criscitiello C, Curigliano G. Highlights from the 14(th) St Gallen International Breast Cancer Conference 2015 in Vienna: dealing with classification, prognostication, and prediction refinement to personalize the treatment of patients with early breast cancer. Ecancermedicalscience 2015;9:518.

19. Shlien A, Malkin D. Copy number variations and cancer. Genome Med

2009;1:62.

20. Ziogas D, Roukos DH. Genetics and personal genomics for personalized breast cancer surgery: progress and challenges in research and clinical practice. Ann Surg Oncol 2009;16:1771-82.

21. Jang H, Hur Y, Lee H. Identification of cancer-driver genes in focal genomic alterations from whole genome sequencing data. Sci Rep 2016;6: 25582.

22. Eirew P, Steif A, Khattra J, Ha G, Yap D, Farahani H, et al. Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. Nature 2015;518:422-6.

23. Li J, Wang J, Chen Y, Yang L, Chen S. A prognostic 4-gene expression signature for squamous cell lung carcinoma. J Cell Physiol. Epub 2017 Feb 4. https://doi.org/10.1002/jcp.25846.

24. Liberg D, Sigvardsson M, Akerblad P. The EBF/Olf/Collier family of transcription factors: regulators of differentiation in cells originating from all three embryonal germ layers. Mol Cell Biol 2002;22:8389-97.

25. Xiang J, Fang L, Luo Y, Yang Z, Liao Y, Cui J, et al. Levels of human replication factor C4, a clamp loader, correlate with tumor progression and predict the prognosis for colorectal cancer. J Transl Med 2014;12:320.

26. Cui H, Wang Y, Huang H, Yu W, Bai M, Zhang L, et al. GPR126 protein regulates developmental and pathological angiogenesis through modulation of VEGFR2 receptor signaling. J Biol Chem 2014;289:34871-85.

27. Maiga A, Lemieux S, Pabst C, Lavallée VP, Bouvier M, Sauvageau G, et al. Transcriptome analysis of G protein-coupled receptors in distinct genetic subgroups of acute myeloid leukemia: identification of potential disease-specific targets. Blood Cancer J 2016;6:e431.

28. Song Q, Li C, Feng X, Yu A, Tang H, Peng Z, et al. Decreased expression of SCUBE2 is associated with progression and prognosis in colorectal cancer. Oncol Rep 2015;33:1956-64.

29. Chen CM, Lin CL, Chiou HL, Hsieh SC, Lin CL, Cheng CW, et al. Loss of endothelial cell-specific molecule 1 promotes the tumorigenicity and metastasis of prostate cancer cells through regulation of the TIMP-1/MMP-9 expression. Oncotarget 2017;8:13886-97.

30. van der Weyden L, Adams DJ. The Ras-association domain family (RASSF) members and their role in human tumourigenesis. Biochim Biophys Acta 2007;1776:58-85.