Received: 6 August 2020 | Revised: 24 January 2021 | Accepted: 29 January 2021

DOI: 10.1002/ijc.33512

TUMOR MARKERS AND SIGNATURES

IJC International Journal of Cancer

A novel signature predicts recurrence risk and therapeutic response in breast cancer patients

Quynh Hoa Tran[1] | Van Thai Than[2,3] | Phuc Loi Luu[4] | Declan Clarke[5,6] | Hanh Ngoc Lam[7] | Thanh-Giang Tan Nguyen[8] | Dinh-Truong Nguyen[8] | Phan Q. Duy[9] | Dung Phung[10] | Minh Nam Nguyen[11]

[1]Department of Biotechnology, Ho Chi Minh City University of Food Industry, Ho Chi Minh City, Vietnam

[2]Faculty of Biotechnology, Chemistry and Environmental Engineering, PHENIKAA University, Hanoi, Vietnam

[3]PHENIKAA Research and Technology Institute (PRATI), A&A Green Phoenix Group JSC, Hanoi, Vietnam

[4]Epigenetics Research Laboratory, Genomics and Epigenetics Division, Garvan Institute of Medical Research, Darlinghurst, New South Wales, Australia

[5]Program in Computational Biology and Bioinformatics, Yale University, New Haven, Connecticut

[6]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut

[7]Department of Microbiology and Environmental Toxicology, University of California, Santa Cruz, California

[8]School of Biotechnology, Tan Tao University, Tan Duc E.City, Vietnam

[9]Medical Scientist Training Program, Yale University School of Medicine, New Haven, Connecticut

[10]School of Medicine, Griffith University, Southport, Queensland, Australia

[11]School of Medicine, Vietnam National University Ho Chi Minh City, Ho Chi Minh City, Vietnam

Correspondence
Minh Nam Nguyen, School of Medicine, Vietnam National University Ho Chi Minh City, Linh Trung Ward, Thu Duc District, Hochiminh City, Vietnam.
Email: nmnam@medvnu.edu.vn; minhnam1984@gmail.com

Funding information
Ho Chi Minh City University of Food Industry, HCM City, Vietnam, Grant/Award Number: 145/HĐ-DCT; NIH Medical Scientist Training Program Training Grant, Grant/Award Number: T32GM136651

Abstract

Acetylserotonin O-methyltransferase (ASMT) is a key enzyme in the synthesis of melatonin. Although melatonin has been shown to exhibit anticancer activity and prevents endocrine resistance in breast cancer, the role of *ASMT* in breast cancer progression remains unclear. In this retrospective study, we analyzed gene expression profiles in 27 data sets on 7244 patients from 11 countries. We found that *ASMT* expression was significantly reduced in breast cancer tumors relative to healthy tissue. Among breast cancer patients, those with higher levels of *ASMT* expression had better relapse-free survival outcomes and longer metastasis-free survival times. Following treatment with tamoxifen, patients with greater *ASMT* expression experienced longer periods before relapse or distance recurrence. Motivated by these results, we devised an ASMT gene signature that can correctly identify low-risk cases with a sensitivity and specificity of 0.997 and 0.916, respectively. This signature was robustly validated using 23 independent breast cancer mRNA array data sets from different platforms (consisting of 5800 patients) and an RNAseq data set from TCGA (comprising 1096 patients). Intriguingly, patients who are classified as high-risk by the signature benefit from adjuvant chemotherapy, and those with grade II tumors who are classified as low-risk exhibit improved overall survival and distance relapse-free

Abbreviations: AR, androgen receptor; ASMT, acetylserotonin O-methyltransferase; CCP, compound covariate predictor; DFS, disease-free survival; DMFS, distance metastasis-free survival; ER, estrogen receptor; HER, human epidermal growth factor receptor 2; MFS, metastasis-free survival; OS, overall survival; PR, progesterone receptor; RFS, relapse-free survival; RNA, ribonucleic acid; TCGA, array express, and the cancer genome atlas.

Quynh Hoa Tran and Van Thai Than contributed equally to this study.

outcomes following endocrine therapy. Together, our findings more clearly elucidate the roles of *ASMT*, provide strategies for improving the efficacy of tamoxifen treatment and help to identify those patients who may maximally benefit from adjuvant or endocrine therapies.

KEYWORDS

ASMT, breast cancer, endocrine therapy, gene signature, recurrence risk prediction

1 | INTRODUCTION

Breast cancer constitutes the most common form of cancer in women and is the second-leading cause of cancer death worldwide.[1,2] About 75% of breast cancer patients are estrogen receptor (ER)-positive, and the ER is a target for endocrine therapy such as tamoxifen administration.[3-5] Tamoxifen is a leading medication that is used for treating patients or preventing recurrence of ER-positive breast cancer. It substantially reduces recurrence rates (by ~39%) and mortality rates (by ~33%) in ER-positive breast cancer patients.[4,6-8] However, among patients who initially respond to tamoxifen, 21% to 30% relapse within 4 to 14 years, even after 5 years of continuous administration.[4,6,7] Furthermore, 30% to 50% of ER-positive patients exhibit immediate resistance, whereas most patients develop resistance after initially responding to the drug.[5] Therefore, clinical resistance to tamoxifen presents a major challenge to successfully treating breast cancer.

Primarily synthesized in the pineal gland,[5] melatonin sensitizes the response of breast cancer cells to tamoxifen in vitro and in vivo. Melatonin exerts cytostatic and cytotoxic apoptotic effects in breast cancer by engaging the MT1 and MT2 receptors.[5] Pretreatment with melatonin improves tamoxifen efficacy by more than 100-fold in MCF-7 breast cancer cells.[9] It has been shown to sensitize breast tumors to tamoxifen and tumor regression in a xenograft model.[10] Supplementing tamoxifen with melatonin slows the progression of metastatic breast cancer[11] and increases response and survival rates in ER-negative metastatic breast cancer patients.[12]

Acetylserotonin O-methyltransferase (ASMT) is a pivotal enzyme in melatonin synthesis. It has previously been combined with *CYP1B1* as a two-gene prognostic index of human glioma.[13] However, the biological roles of *ASMT* in cancer remain unknown. Furthermore, its prognostic potential in breast cancer and its clinical utility in enhancing endocrine therapy are not understood.

We investigated the significance of *ASMT* expression in human breast cancer. We found that low *ASMT* expression is characteristic of breast tumors, and elevated levels of *ASMT* expression improve relapse-free survival (RFS) outcomes and metastasis-free survival (MFS) times. Relative to other patients, those with high *ASMT* expression exhibited fewer relapses or longer distance recurrence outcomes following tamoxifen treatment. Furthermore, we devised a novel gene signature that can robustly predict recurrence risk. Patients predicted to be high-risk benefitted significantly from adjuvant chemotherapy. Following endocrine therapy, low-risk patients showed better overall survival (OS) and distance relapse-free outcomes.

What's new?

Melatonin is associated with delayed breast cancer progression and improved survival in tamoxifen-treated, estrogen receptor (ER)-negative breast cancer patients. Whether acetylserotonin O-methyltransferase (ASMT), which has a key role in melatonin synthesis, influences these effects of melatonin on breast cancer remains unclear. Here, ASMT expression was found to be significantly reduced in human breast cancer tumors. Moreover, patients with relatively high ASMT expression experienced improved relapse-free survival outcomes and increased metastasis-free survival times, particularly following tamoxifen therapy. These findings suggest that in certain subsets of ER-negative patients, knowledge of ASMT expression could be used to better tailor treatment strategies.

2 | MATERIALS AND METHODS

2.1 | Patients and gene expression profiles

Clinical and gene expression data were collected from gene expression omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/), TCGA (http://www.cbioportal.org/index.do) and EMBL-EBI (https://www.ebi.ac.uk/arrayexpress/) under the following criteria: (a) sample sizes must be greater than 100, and (b) patients for whom data on disease progression are available were included (such as relapse time, survival status, etc). Raw data were normalized as described previously.[14] Twenty-seven data sets were used with a total of 7328 samples (Table S1). Six hundred and thirty-five samples were excluded, either because they were not tumor samples or because follow-up information was unavailable.

2.2 | The ASMT-associated gene signature

Data sets for building the *ASMT* gene signature were selected using the following criteria: (a) only patients with the status of RFS as

patient endpoints were included, and (b) samples from these patients had to be assayed using Affymetrix HG-U133A chips. Three data sets (GSE1456, GSE2034, and GSE7390) with a total of 643 patients were used as a discovery set. Broadly, the gene signature was constructed in five steps. First, patients in the discovery data set were separated into two groups (high- and low-expressing *ASMT* groups), based on median *ASMT* expression. Second, a two-sample *t*-test identified genes that exhibit significant differential expression between the two groups ($P < .001$). Third, a univariate Cox proportional hazard regression ($P < .001$) was performed to identify the RFS-correlated genes. Fourth, common genes among different platforms were identified. Last, gene expression levels in the common gene set and RFS were used as features to build a survival risk classifier using principal components analysis.[15] This classifier uses the principal component from the discovery data set to produce a prognostic index for each patient. The index is computed as $\sum_i w_i x_i$, where w_i and x_i designate the weight and $\log_2$-transformed gene expression for the *i*th gene, respectively. The robustness of the classifier was evaluated using leave-one-out cross-validation. A patient was predicted to be high (low)-risk if their prognostic index was greater than (less than or equal to) the median prognostic index of 0.782.

2.3 | Validation of the prognostic signature

Validation was performed on 23 independent mRNA array data sets (4795 patients) and one RNA-seq data set from TCGA (1096 patients). The median was subtracted from each gene individually. The compound covariate predictor (CCP) was used as a class prediction algorithm to further refine this model and to substratify outcome predictions.[14,16-18]

Kaplan-Meier survival analyses were performed after samples were classified into two risk groups, and log-rank tests were used to evaluate risk. Uni- and multi-variate Cox proportional hazard regression analyses were conducted to evaluate independent prognostic factors. The gene signature, tumor grade, age and molecular marker status were used as covariates.

2.4 | Statistical methods of microarray data

We used the R language environment for all statistical analyses. A survival curve was analyzed and drawn using the Kaplan-Meier method, and *P* values were evaluated using log-rank tests. The Cox proportional hazard regression model was used to perform univariate and multivariate tests. Cluster analyses were conducted with Cluster and Tree View.[19] *P* values less than .05 were considered to be statistically

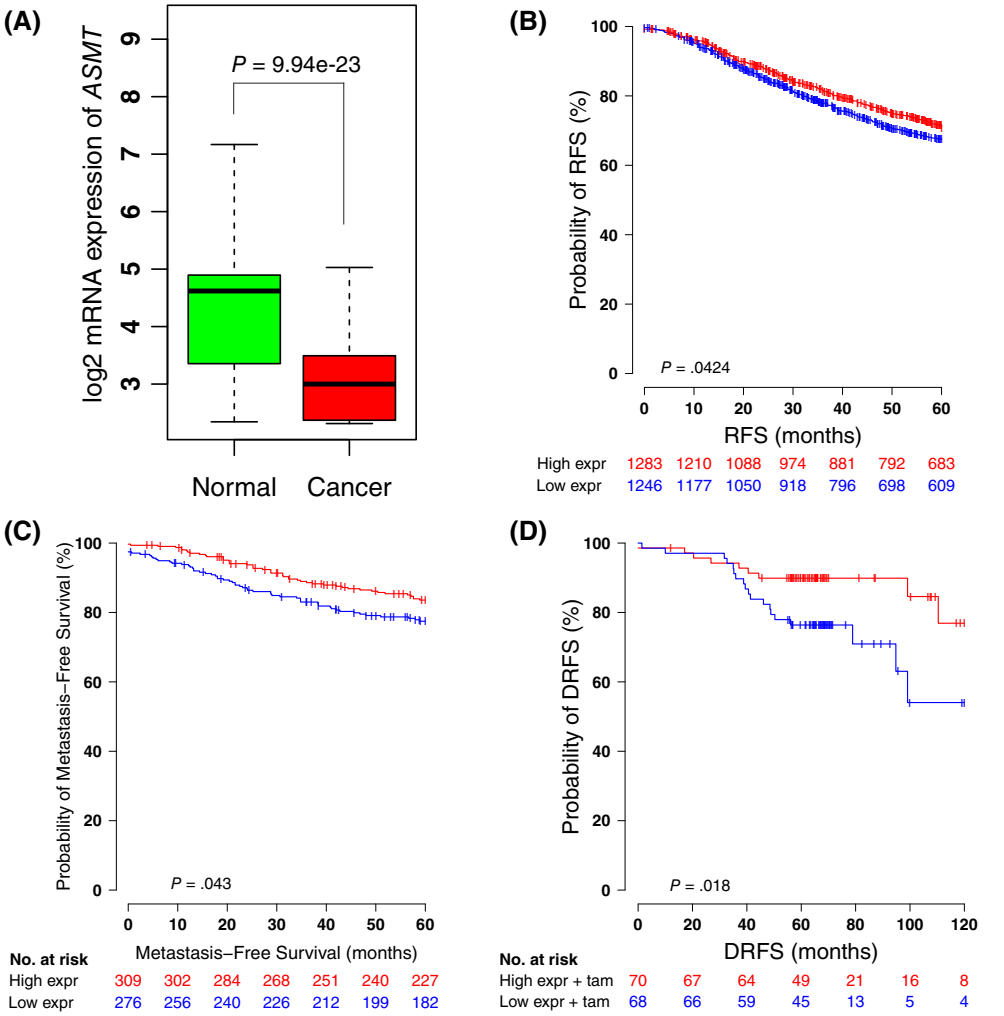

FIGURE 1 Upregulation of *ASMT* increases survival rates and tamoxifen sensitivity in breast cancer patients. A, Log_2-transformed *ASMT* mRNA expression in normal and breast cancer tissues. Breast cancer patients were segregated into two groups based on to their *ASMT* expression, with median expression being used as the threshold to demarcate groups. Kaplan-Meier survival curves for that two groups were compared using RFS (B) and metastasis-free survival (C). D, A Kaplan-Meier plot (using DRFS) of the two groups of patients who received tamoxifen treatment. DRFS, distant recurrence-free survival; RFS, recurrence-free survival [Color figure can be viewed at wileyonlinelibrary.com]

significant. We note that not all survival curves start at 100%. In such cases, this is because a number of patients had already developed metastasis by the time of diagnosis (ie, by the starting point on such survival curves). Thus, in these cases, not all patients were recorded as being metastasis-free during the start time.

3 | RESULTS

3.1 | Upregulation of ASMT is associated with improved clinical outcomes

To test whether *ASMT* expression is different between cancerous and normal breast tissue, we compared mRNA in 270 normal and 395 breast tumor samples (Table S2). On average, *ASMT* mRNA levels were significantly lower in tumor tissues (Figure 1A). To evaluate the relationship between *ASMT* expression and clinical outcomes, we classified 2529 patients into low and high expression groups (relative to the median *ASMT* expression). High expression of *ASMT* was positively correlated with better RFS outcomes (Figure 1B). Notably, a Kaplan-Meier plot showed that breast cancer patients with high *ASMT* expression had longer MFS times (Figure 1C). These results indicate that high expression of *ASMT* is not only positively related to reduced breast cancer risk, but is also associated with greater RFS and MFS times.

3.2 | Elevated expression of ASMT increases tamoxifen sensitivity

To investigate whether *ASMT* expression is related to tamoxifen efficacy, we analyzed the GSE2990, GSE6532 and GSE9893 data sets in which information on tamoxifen treatment is available. Patients with higher *ASMT* expression exhibited improved responses to tamoxifen treatment and longer survival times and distance recurrence after tamoxifen treatment (Figures 1D and S1, respectively). *ASMT* expression is thus associated with improved tamoxifen efficacy.

3.3 | Deletion of ASMT downregulates protein expression of ER

To better elucidate how ASMT expression may promote tamoxifen efficacy, we identified proteins for which expression was altered by ASMT. Specifically, we found that ER, HER3, androgen receptor

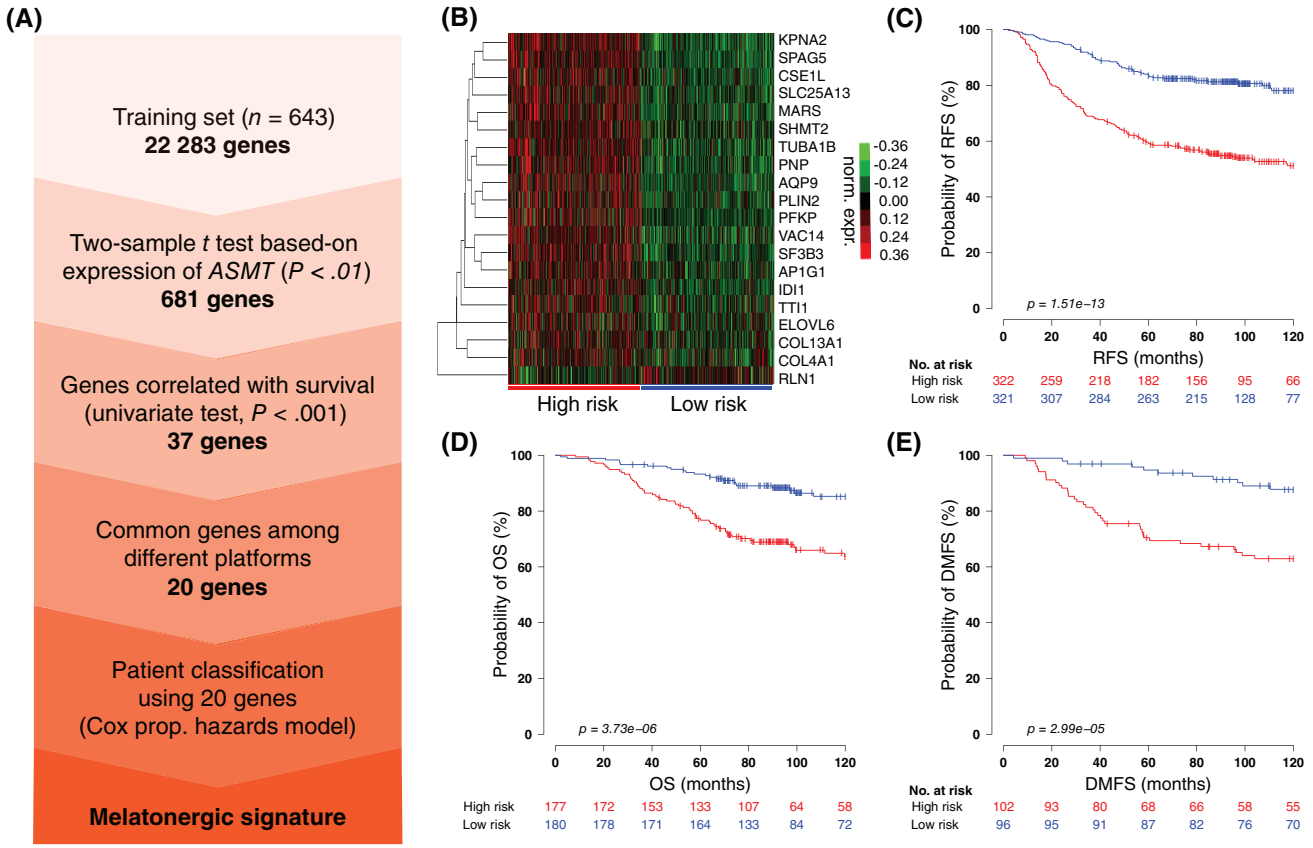

FIGURE 2 Discovering the genes associated with the melatonergic signature. A, Schematic overview of the procedure used to construct the melatonergic signature based on gene expression profiles. B, Hierarchical cluster analysis of the melatonergic signature in two risk groups of the discovery data set. Kaplan-Meier plots for RFS (C), OS (D) and DMFS (E) of the two risk groups in the discovery data set. *P* values were computed using a log-rank test. DMFS, distance metastasis-free survival; OS, overall survival; RFS, recurrence-free survival [Color figure can be viewed at wileyonlinelibrary.com]

(AR) and progesterone receptor (PR) were among the most down-regulated (Figure S2). ER and ER-pS118 expression was significantly reduced in patients with deep deleted *ASMT* (Figure S3).

3.4 | The ASMT signature for accurate risk prediction

Expression profiles were analyzed by relating *ASMT* expression to empirical survival outcomes (Figure 2A). We identified 681 probes that were differentially expressed between the high- and low-expression groups (Figure S4). Thirty-seven genes were then identified using a survival univariate test ($P < .001$). Among these, 20 genes (which were shared between the different platforms) were selected as a gene signature and incorporated into a classifier based on a Cox proportional hazards model (Table S3). This 20-member set constitutes our melatonergic signature. We used this signature to classify patients in the training set as being high- (n = 322) and low-risk (n = 321). Note that the 322/321 split refers to low and high expression (relative to the median expression of the entire 20-gene signature), and so it does not refer to the initial partitioning by median ASMT expression. There were significant differences between predicted groups in RFS (Figure 2C), OS (Figure 2D) and distance metastasis-free survival (DMFS; Figure 2E).

3.5 | The melatonergic signature is associated with clinical features

We investigated whether the melatonergic signature is correlated with clinicopathological characteristics (such as age at diagnosis, tumor grade, histology and follow-up time; Table S4). Tumor grade, ER status, histology and follow-up times were significantly associated with melatonergic signature classification, whereas age showed no such association. Univariate and multivariate Cox regression analyses were performed on the discovery data set to compare the prognostic value of the melatonergic signature with other prognostic covariates (Table 1). Interestingly, the melatonergic signature showed stronger prognostic predictive ability than these other clinical variables (Table 1). For both the univariate and multivariate analyses, the melatonergic signature was significant in OS and RFS. These data suggest that the melatonergic signature is a significant predictor of disease-free survival (DFS), OS and DMFS, and it is generally independent of age, grade and ER status.

3.6 | Validation of the melatonergic signature using independent data sets

To evaluate the robustness of the melatonergic signature, we performed validation using 23 independent breast cancer mRNA array data sets and one RNA-seq data set. The procedure used to validate the external data sets is shown in Figure 3A. Using the CCP classifier, the sensitivity and specificity for correctly predicting low risk were 0.997 and 0.916, respectively. The melatonergic signature successfully classified patients in different platforms into the two-risk groups (Table S5). Kaplan Meier analysis

TABLE 1 Univariate and multivariate Cox proportional hazard regression analyses of clinical variables in the training data set

	OS				DMFS				RFS			
	Univariate		Multivariate		Univariate		Multivariate		Univariate		Multivariate	
Variable	HR (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Age > 45 vs ≤45	0.97 (0.53-1.78)	.93	1.06 (0.57-1.97)	.86	1.06 (0.60-1.9)	.84	1.22 (0.68-2.20)	.51	0.90 (0.57_1.41)	.64	0.98 (0.61-1.57)	.94
ER 1 vs 0	0.33 (0.18-0.59)	2.98×10^{-4}	0.49 (0.24-1.00)	.05	0.40 (0.22-0.71)	.002	0.70 (0.36-1.38)	.31	0.73 (0.53-1.00)	.05	0.89 (0.51-1.55)	.68
Grade 2 vs 1	4.37 (1.03-18.60)	<.05	2.88 (0.66-12.58)	.16	1.8 (1.2-2.8)	.01	3.34 (0.78-14.41)	.11	2.52 (1.06-5.95)	.04	2.52 (1.06-6.00)	.04
ASMT expression	1.1 (0.7-1.7)	.73	1.01 (0.54-1.87)	.98	1.3 (0.76-2.3)	.32	1.05 (0.59-1.88)	.86	1.1 (0.87-1.5)	.36	0.88 (0.56-1.39)	.58
Risk low vs high	0.25 (0.12-0.53)	2.42×10^{-4}	0.27 (0.12-0.64)	.003	0.26 (0.13-0.52)	1.04×10^{-4}	0.28 (0.13-0.63)	.002	0.35 (0.26-0.46)	1.7×10^{-12}	0.27 (0.15-0.49)	1.1×10^{-5}

Note: *P* values were obtained using Cox proportional hazard regression analysis.

Abbreviations: ASMT, acetylserotonin O-methyltransferase; CI, confidence interval; DMFS, distance metastasis-free survival; ER, estrogen receptor; HR, hazard ratio; OS, overall survival; RFS, recurrence-free survival.

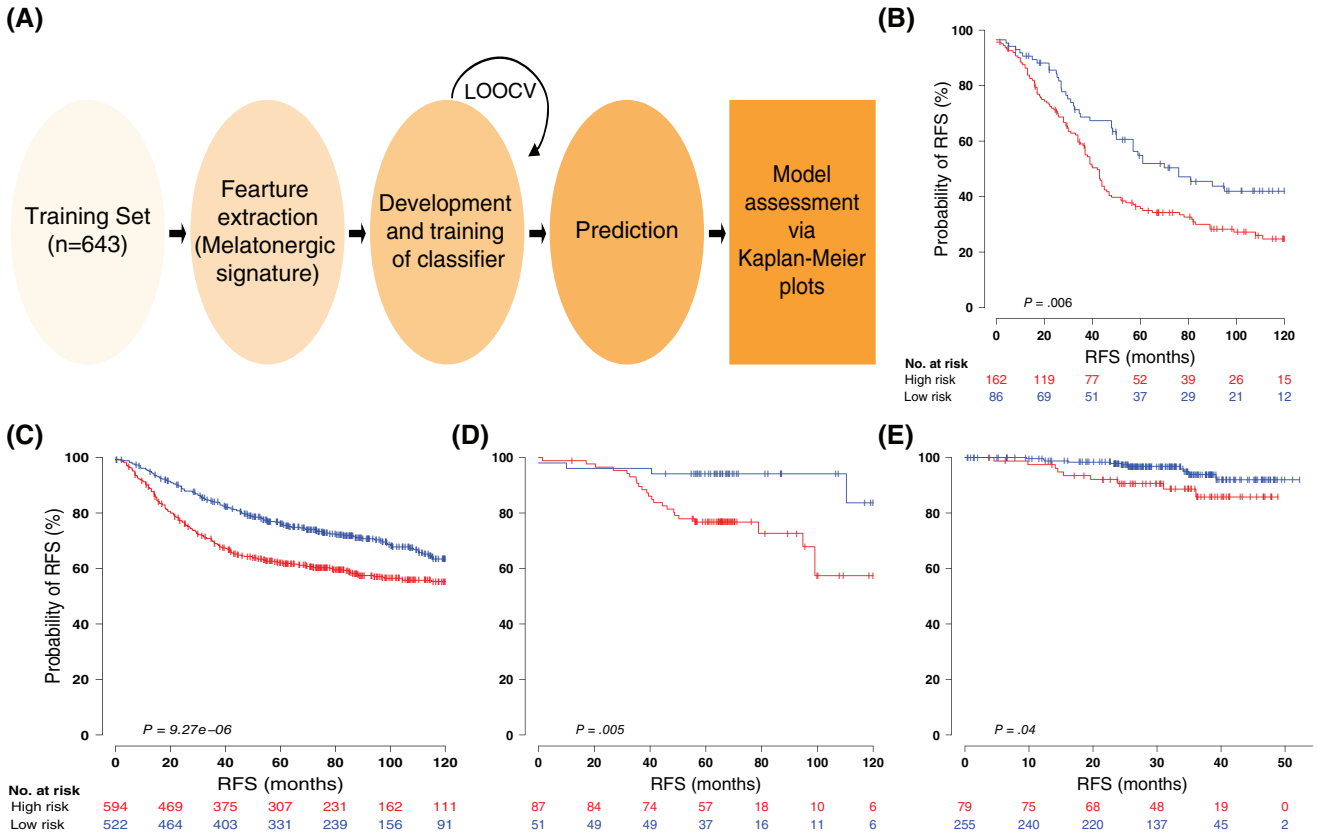

FIGURE 3 Prognostic significance of the melatonergic signature in independent validation cohorts. A, Generating the risk prediction model and evaluating risk outcomes. Kaplan-Meier survival analyses for RFS of the validation datasets with the Human Genome U133A (B), Human Genome U133 Plus 2.0 (C), MLRG Human 21 K V12.0 (D) and Illumina HumanRef-8 v3.0 expression beadchip (E) platforms. *P* values were computed using a log-rank test. RFS, recurrence-free survival [Color figure can be viewed at wileyonlinelibrary.com]

showed a significant difference between the two risk groups in platform HG-U136A, HG-U133_Plus_2, MLRG Human 21K, Illumina, Agilent, HG_U95Av2 and TCGA (Figures 3B-E and S5-S7). The melatonergic signature correctly classified a significant number of patients across independent platforms.

3.7 | The melatonergic signature correctly classifies patients with different tumor grades

A low-vs-high risk classification scheme can also be applied to patients with different tumor grades. The melatonergic signature accurately classified patients across all grades into low- and high-risk groups (Figure S8 and Table S6).

3.8 | Leveraging the melatonergic signature to stratify patients into different subtypes

A number of features associated with a particular cancer are often used to stratify breast cancer tumors into specific subtypes. However, considerable heterogeneity exists even within cancers that share similar features. We thus investigated whether the melatonergic signature can further stratify breast cancer patients. We incorporated the melatonergic signature with ER, PR, HER2, lymph node, p53 status and canonical intrinsic subtypes (luminal A,B, HER2, Basal, Normal-like). The melatonergic signature was found to be associated with ER+, HER−, PR+, PR−, Node+, Node−, wild-type p53 (Figure S9), as well as the luminal A subtype (Figure S10). However, we did not find significant differences between the two risk groups in ER−, HER+, p53-mutated (Figure S9), and luminal B, HER2, Basal and Normal-like subtypes (Figure S10). The melatonergic signature thus captures heterogeneity in patients of the same subtype, suggesting that it may help to overcome the limitations of classification schemes that use standard molecular markers.

3.9 | The melatonergic signature can be used to better select patients for adjuvant chemotherapy and endocrine therapy

We evaluated whether the signature may be used to identify patients who benefit from adjuvant chemotherapy and endocrine therapy. We selected patients who did not receive any treatment and those

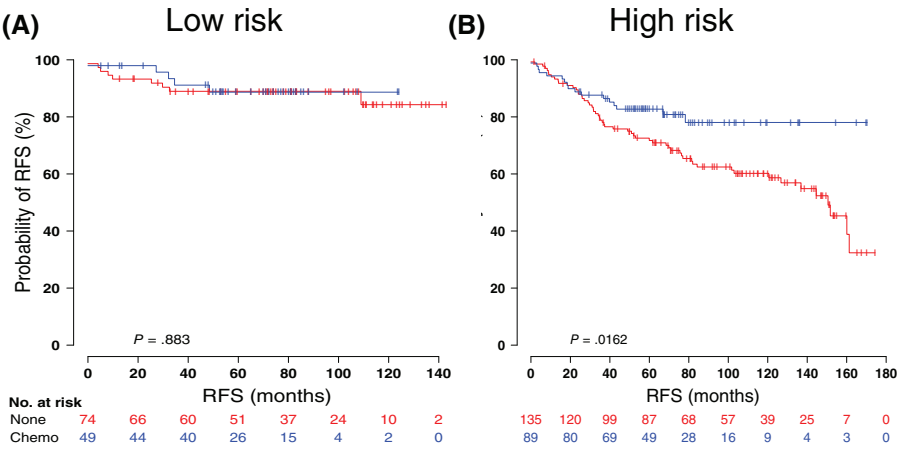

FIGURE 4 Kaplan-Meier survival analysis of the melatonergic signature and adjuvant chemotherapy. Patients were classified into different risk groups. Chemotherapy along with the conferred treatment in low-risk (A) and high-risk (B) patient groups evaluated by Kaplan-Meier analyses. *P* values were computed using a log-rank test [Color figure can be viewed at wileyonlinelibrary.com]

treated with only adjuvant chemotherapy or endocrine therapy. To eliminate confounding factors of potential synergistic effects, patients treated with a combination of both treatments (in addition to those treated with radiotherapy) were excluded. High-risk patients who received adjuvant chemotherapy exhibited more pronounced improvements relative to those who did not receive this treatment, though there was no significant difference among low-risk patients (Figure 4A,B). A plot of the associated results (with a very small number of low-risk patients) appears in Figure S11. Low-risk patients (in grade II) had better OS and distance RFS outcomes after endocrine therapy (Figure S11A,C).

4 | DISCUSSION

The role of ASMT in cancer has long been unclear. Our work reveals that *ASMT* expression is lower in breast cancer than in healthy tissue. Patients with higher *ASMT* expression exhibited improved RFS outcomes, longer MFS times and better responses to tamoxifen treatment.

Supplementing tamoxifen with melatonin has been found to prolong the premetastatic phase in breast cancer patients.[11] Consistent with this, we found that patients with high *ASMT* expression levels tend to be more sensitive to tamoxifen treatment and experience fewer relapses or longer phases before disease recurrence. Among patients with reduced *ASMT* expression, we found that ER, PR, HER and AR expression levels were reduced. This may help to explain why tamoxifen exerts less of an effect in patients with reduced *ASMT* expression.

Our findings may pave the way for novel therapeutic options for ER-negative patients. Patients with high AMST expression may represent a putative target cohort for antihormonal treatment (Tamoxifen), though we note that this concept would need to be tested and investigated with mechanistic data and appropriately designed clinical trials. Though the results presented in our study suggest that *ASMT* likely plays a role in improved responses to endocrine therapy in the treatment of ER-negative breast cancer patients, the precise biological roles by which ASMT may confer improved responses have not been explored in detail here. However, we tentatively conjecture a role whereby elevated levels of melatonin (which result from enhanced expression of ASMT) help to promote improved responses to endocrine therapy in an ER-independent fashion. However, support for this conjecture about any specific means by which such ER-independent effects actually materialize in tumor cells is outside the scope of this work, and would thus need to be investigated in future experimental studies.

We devised a signature that was validated in 5891 patients. This signature indicated that high-risk patients may benefit from adjuvant chemotherapy, and low-risk patients (in grade II) have better OS and distance RFS outcomes following endocrine therapy.

This signature consists of many genes that play important roles in migration/invasion, survival and proliferation of breast cancer cells (including *PFKP*[20,21] and *COL4A1*). Several others (such as *AP1G1* and *SF3B3*[22,23]) have been suggested as novel targets for overcoming drug resistance. *SHMT2*, *COL4A1*, *SF3B3*, *PLIN2*, *Elov16* and *SPAG5* serve as independent prognostic factors in breast cancer patients, and high expression levels for these genes were found to be significantly correlated with poor survival outcomes.[23-29] Many novel genes (such as *CSE1L*, *MARS*, *TTI1*, *VAC14* and *TUBA1B*) were included in the signature, suggesting that it contains new candidates which may play critical roles in breast cancer. These may serve as novel diagnostic biomarkers and therapeutic targets.

Breast cancer is a heterogeneous disease. Though individuals with the same breast cancer subtype generally receive similar therapies,[30] this heterogeneity results in only a subset of patients who substantively benefit from treatment.[4-8] Most (60%-75%) cannot clearly be classified into a particular subtype on the basis of canonical features.[31] Integrating the melatonergic signature with molecular subtypes may allow for more precise subclassification, leading to more personalized therapies.

Histological grade and lymph node status are well-established prognostic factors.[31-33] However, grade II invasive ductal carcinoma (without a clearly designated subtype) presents challenges in clinical decision making.[34] Although chemotherapy has been suggested to be

advantageous for grade III patients, there is no significant benefit (in terms of DFS or OS) with a single cycle of systemic treatment.[33] Furthermore, standard chemotherapy offers the same proportional reductions in recurrence and mortality among younger node-positive or node-negative patients.[7] Thus, histological grade and lymph node status inadequately reflect heterogeneity. The melatonergic signature more finely stratifies patients in each histological grade and lymph node status group into prognostic subgroups.

Similar therapeutic approaches are generally applied to most patients,[30] yet only a fraction of them benefit.[7,35] Limited understanding of disease heterogeneity limits the ability to identify patients who respond to therapy or exhibit recurrence. We demonstrate that patients identified as high-risk recur earlier than those identified as low-risk. Moreover, patients in the high-risk group may benefit from adjuvant therapy, whereas those in the low-risk group may be spared unnecessary treatment. Patients with grade II and low-risk have better OS and distance RFS outcomes after endocrine therapy. An estimated ~30% of all breast cancer patients are considered to be overdiagnosed and overtreated.[36] Although we currently have a very small sample size of low-risk patients (Figure S11), our preliminary results suggest that it would be very interesting to further investigate the possibility or survival disparities once a larger number of low-risk patients becomes available in the future. Thus, having a greater sample size of patients for a such follow-up studies would enable us to more reliably make these types of comparisons, thereby pointing out future avenues of investigation. Moreover, our findings suggest that the melatonergic signature is at least equivalent in its predictive power to that of other predictors (Figure S12). The melatonergic signature is thus a powerful tool that may help to prevent unnecessary treatment.

Although our multivariate regression analyses included multiple features of interest (Table 1), it would also be valuable to investigate additional clinical variables as part of this type of analysis (such features may include histology and lymph node status, for instance). As data sets that include such variables on growing numbers of patients become available and provide sufficient data on necessary patient sample sizes, follow-up studies may provide an even fuller picture of the disease characteristics that are most strongly correlated with the melatonergic signature.

We elucidate the role of *ASMT* in breast cancer progression, highlighting the importance of *ASMT* as a prognostic marker and target for personalized therapy. We demonstrate that the melatonergic signature is a powerful predictor for understanding the heterogeneous landscape of breast cancer, and that it can identify patients who optimally benefit from adjuvant and endocrine therapy. This signature has also identified a number of novel breast cancer-associated genes. In addition to these promising new avenues for future study, this signature provides an accurate and easy-to-implement tool with prognostic and clinical utility in the age of precision medicine. Thus, given the prognostic predictive strength demonstrated by this 20-member melatonergic gene signature, it is our hope that it may find direct applications in clinical practice, and that it will serve as a valuable complement to the repertoire of existing therapies and approaches for treating breast cancer.

ACKNOWLEDGMENTS

We thank Emily Xiao (Yale University) for comments that greatly improved the manuscript. We also thank Prof. Lajos Pusztai (Yale Cancer Center) as well as anonymous reviewers for valuable feedback. Q. H. T. was supported by Ho Chi Minh City University of Food Industry, HCM City, Vietnam (145/HĐ-DCT). P. Q. D. was supported by NIH Medical Scientist Training Program Training Grant (T32GM136651). The funders of this study had no role in study design, data collection, data analysis, data interpretation or writing of the report. The corresponding author had full access to the data in the study and had final responsibility for the decision to submit for publication.

CONFLICT OF INTEREST

All authors have read the journal's policy on disclosure of potential conflicts of interest and declare no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

DATA AVAILABILITY STATEMENT

Only publicly available data were used in this study, and data sources and handling of these data are described in the Section 2 and in Tables S1 and S2. Further information is available from the corresponding author upon request.

ORCID

Minh Nam Nguyen https://orcid.org/0000-0001-5596-6222

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin*. 2018;68:7-30.
2. Stewart B, Wild C. *World Cancer Report 2014*. Lyon, France: International Agency for Research on Cancer; 2014:16-53. https://publications.iarc.fr/Non-Series-Publications/World-Cancer-Reports/World-Cancer-Report-2014.
3. Turner NC, Ro J, Andre F, et al. Palbociclib in hormone-receptor-positive advanced breast cancer. *N Engl J Med*. 2015;373:209-219.
4. Early Breast Cancer Trialists' Collaborative Group, Davies C, Godwin J, et al. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *Lancet*. 2011;378:771-784.
5. Hill SM, Belancio VP, Dauchy RT, et al. Melatonin: an inhibitor of breast cancer. *Endocr Relat Cancer*. 2015;22:R183-R204.
6. Davies C, Pan H, Godwin J, et al. Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. *Lancet*. 2013;381:805-816.
7. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet*. 2005;365:1687-1717.
8. Paridaens RJ, Dirix LY, Beex LV, et al. Phase III study comparing exemestane with tamoxifen as first-line hormonal treatment of metastatic breast cancer in postmenopausal women: the European

Organisation for Research and Treatment of Cancer Breast Cancer Cooperative Group. *J Clin Oncol*. 2008;26:4883-4890.
9. Wilson ST, Blask DE, Lemus-Wilson AM. Melatonin augments the sensitivity of MCF-7 human breast cancer cells to tamoxifen in vitro. *J Clin Endocrinol Metab*. 1992;75:669-670.
10. Dauchy RT, Xiang S, Mao L, et al. Circadian and melatonin disruption by exposure to light at night drives intrinsic resistance to tamoxifen therapy in breast cancer. *Cancer Res*. 2014;74:4099-4110.
11. Lissoni P, Barni S, Meregalli S, et al. Modulation of cancer endocrine therapy by melatonin: a phase II study of tamoxifen plus melatonin in metastatic breast cancer patients progressing under tamoxifen alone. *Br J Cancer*. 1995;71:854-856.
12. Lissoni P, Ardizzoia A, Barni S, et al. A randomized study of tamoxifen alone versus tamoxifen plus melatonin in estrogen receptor-negative heavily pretreated metastatic breast-cancer patients. *Oncol Rep*. 1995;2:871-873.
13. Kinker GS, Oba-Shinjo SM, Carvalho-Sousa CE, et al. Melatonergic system-based two-gene index is prognostic in human gliomas. *J Pineal Res*. 2016;60:84-94.
14. Nguyen MN, Choi TG, Nguyen DT, et al. CRC-113 gene expression signature for predicting prognosis in patients with colorectal cancer. *Oncotarget*. 2015;6:31674-31692.
15. Bair E, Tibshirani R. Semi-supervised methods to predict patient survival from gene expression data. *PLoS Biol*. 2004;2:E108.
16. Matondo A, Jo YH, Shahid M, et al. The prognostic 97 chemoresponse gene signature in ovarian cancer. *Sci Rep*. 2017;7:9689.
17. Shahid M, Choi TG, Nguyen MN, et al. An 8-gene signature for prediction of prognosis and chemoresponse in non-small cell lung cancer. *Oncotarget*. 2016;7:86561-86572.
18. Shahid M, Cho KM, Nguyen MN, et al. Prognostic value and their clinical implication of 89-gene signature in glioma. *Oncotarget*. 2016;7:51237-51250.
19. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A*. 1998;95:14863-14868.
20. Prasad CP, Sodergren K, Andersson T. Reduced production and uptake of lactate are essential for the ability of WNT5A signaling to inhibit breast cancer cell migration and invasion. *Oncotarget*. 2017;8:71471-71488.
21. Kim NH, Cha YH, Lee J, et al. Snail reprograms glucose metabolism by repressing phosphofructokinase PFKP allowing cancer cell survival under metabolic stress. *Nat Commun*. 2017;8:14374.
22. Tao X, Lu Y, Qiu S, Wang Y, Qin J, Fan Z. AP1G1 is involved in cetuximab-mediated downregulation of ASCT2-EGFR complex and sensitization of human head and neck squamous cell carcinoma cells to ROS-induced apoptosis. *Cancer Lett*. 2017;408:33-42.
23. Gokmen-Polar Y, Neelamraju Y, Goswami CP, et al. Expression levels of SF3B3 correlate with prognosis and endocrine resistance in estrogen receptor-positive breast cancer. *Mod Pathol*. 2015;28:677-685.
24. Feng YH, Chen WY, Kuo YH, et al. Elovl6 is a poor prognostic predictor in breast cancer. *Oncol Lett*. 2016;12:207-212.
25. Bernhardt S, Bayerlova M, Vetter M, et al. Proteomic profiling of breast cancer metabolism identifies SHMT2 and ASCT2 as prognostic factors. *Breast Cancer Res*. 2017;19:112.
26. Thangavelu PU, Krenacs T, Dray E, Duijf PH. In epithelial cancers, aberrant COL17A1 promoter methylation predicts its misexpression and increased invasion. *Clin Epigenet*. 2016;8:120.
27. Abdel-Fatah TMA, Agarwal D, Liu DX, et al. SPAG5 as a prognostic biomarker and chemotherapy sensitivity predictor in breast cancer: a retrospective, integrated genomic, transcriptomic, and protein analysis. *Lancet Oncol*. 2016;17:1004-1018.
28. Lucenay KS, Doostan I, Karakas C, et al. Associates with the lipogenic enzyme ATP-citrate lyase to enable malignant growth of breast cancer cells. *Cancer Res*. 2016;76:2406-2418.
29. Yamashita Y, Nishiumi S, Kono S, Takao S, Azuma T, Yoshida M. Differences in elongation of very long chain fatty acids and fatty acid metabolism between triple-negative and hormone receptor-positive breast cancer. *BMC Cancer*. 2017;17:589.
30. Gradishar WJ, Anderson BO, Balassanian R, et al. NCCN guidelines insights: breast cancer, version 1.2017. *J Natl Compr Canc Netw*. 2017;15:433-451.
31. Rakha EA, Reis-Filho JS, Baehner F, et al. Breast cancer prognostic classification in the molecular era: the role of histological grade. *Breast Cancer Res*. 2010;12:207.
32. Rakha EA, El-Sayed ME, Lee AH, et al. Prognostic significance of Nottingham histologic grade in invasive breast carcinoma. *J Clin Oncol*. 2008;26:3153-3158.
33. Pinder SE, Murray S, Ellis IO, et al. The importance of the histologic grade of invasive breast carcinoma and response to chemotherapy. *Cancer*. 1998;83:1529-1539.
34. Geyer FC, Marchio C, Reis-Filho JS. The role of molecular analysis in breast cancer. *Pathology*. 2009;41:77-88.
35. Anderson WF, Chatterjee N, Ershler WB, Brawley OW. Estrogen receptor breast cancer phenotypes in the surveillance, epidemiology, and end results database. *Breast Cancer Res Treat*. 2002;76:27-36.
36. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. *N Engl J Med*. 2012;367:1998-2005.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Tran QH, Than VT, Luu PL, et al. A novel signature predicts recurrence risk and therapeutic response in breast cancer patients. *Int. J. Cancer*. 2021;148:2848–2856. https://doi.org/10.1002/ijc.33512