Breast Cancer Res Treat (2011) 127:53–63
DOI 10.1007/s10549-010-0943-7

PRECLINICAL STUDY

DEAD box 1: a novel and independent prognostic marker for early recurrence in breast cancer

Devon R. Germain · Kathryn Graham · Darryl D. Glubrecht · Judith C. Hugh · John R. Mackey · Roseline Godbout

Received: 5 March 2010 / Accepted: 6 May 2010 / Published online: 25 May 2010

Abstract Breast cancer is a heterogeneous disease characterized by diverse molecular signatures and a variable response to therapy. Clinical management of breast cancer is guided by the expression of estrogen and progesterone receptors and HER2 amplification. New prognostic and predictive markers, as well as additional targets for therapy, are needed for more effective management of this disease. Gene expression microarrays were probed with RNAs from 176 primary breast cancer samples and tissue microarrays immunostained with anti-DDX1 antibody, an antibody to DEAD box protein DDX1, a putative RNA-RNA and RNA-DNA unwinding protein normally found in the nucleus. Half of the patient cohort had experienced early relapse despite standard adjuvant therapy, but were otherwise matched for estrogen receptor and HER2 status, stage and duration of follow-up. Here, we identify *DDX1* RNA overexpression as an independent prognostic marker for early recurrence in primary breast cancer, with a hazard ratio of 4.31 based on logrank analysis of Kaplan–Meier curves. Elevated levels of DDX1 protein in the cytoplasm also independently correlate with early recurrence with a hazard ratio of 1.90. In conclusion, our data indicate a strong and independent association between poor prognosis and deregulation of the DEAD box protein DDX1. We propose that elevated levels of *DDX1* RNA or the presence of DDX1 in the cytoplasm could serve as an effective prognostic biomarker for early recurrence in primary breast cancer.

Keywords Breast cancer · DEAD box 1 · Gene expression · Tissue microarray · Prognostic marker · Early recurrence

Introduction

Breast cancer is the second highest cause of cancer-related death in women, with approximately 1 million new cases diagnosed each year worldwide [1]. While there have been significant advances in the development of endocrine and chemotherapy-based therapies for the treatment of breast cancer, approximately 30% of women with early stage disease will eventually relapse [2], and those with distant metastases have less than a 3% chance of long-term survival [3–6]. The molecular pathways and events underlying recurrence in breast cancer are poorly understood. To compound this problem, breast cancer represents a conglomerate of many different clinical and pathological diseases characterized by different genetic alterations, growth properties, and responses to therapy.

A number of clinical and molecular parameters have traditionally been used to classify breast cancers including stage, grade (number of mitoses, nuclear architecture and tubule formation), estrogen receptor (ER) and progesterone receptor (PR) status, and HER2 (ERBB2) amplification. Recent molecular profiling based on hormone receptor status, HER2 amplification, and proliferation rates have resulted in the widely accepted classification of breast cancer into four major subtypes: luminal A (ER+ve,

An invited commentary to this article can be found at doi:10.1007/s10549-010-1105-7.

D. R. Germain · K. Graham · D. D. Glubrecht · J. R. Mackey · R. Godbout (✉)
Department of Oncology, School of Cancer/Engineering/Imaging Sciences, University of Alberta, Cross Cancer Institute, 11560 University Avenue, Edmonton, AB T6G 1Z2, Canada
e-mail: rgodbout@ualberta.ca

J. C. Hugh
Department of Lab Medicine and Pathology, University of Alberta, 11560 University Avenue, Edmonton, AB T6G 1Z2, Canada

PR+ve, low proliferation, and HER2−ve); luminal B (ER+ve, PR+ve, with either a higher proliferative index or HER2+ve); HER2-amplified (ER−ve, PR−ve and HER2+ve); and triple-negative or basal-like (ER−ve, PR−ve, and HER2−ve) [7–10]. Of these four subtypes, luminal A breast cancers have the best prognosis, with tumors responding well to adjuvant hormone therapy. In the absence of a specific target for therapy, triple-negative tumors have the worst prognosis [11]. While this molecular classification allows for more precise prognosis and treatment recommendations, there is still considerable response variation within each subtype [12].

Similarly, genome wide transcriptome analysis has defined multi-gene signatures reflecting breast cancer subtypes. Several multi-gene signatures with varying prognostic significance have been reported [13–15]. The 21-gene recurrence score assay and the 70-gene signature MammaPrint are currently being marketed as prognostic tools for breast cancer [16, 17]. Recent reports suggest that these multigene assays help identify which patients will benefit from chemotherapy [18]. In spite of these advances, it is clear that we need: (i) a better understanding of the events underlying early relapse in breast cancer, (ii) novel prognostic markers which can independently predict recurrence, and (iii) new approaches to the treatment of breast cancers with a poor prognosis.

DEAD box 1 (DDX1) is a member of the D(Asp)-E(Glu)-A(Ala)-D(Asp) box protein family of RNA unwinding proteins [19]. *DDX1* is amplified and over-expressed in a subset of retinoblastoma and neuroblastoma tumors [20–23] and has recently been reported to be involved in the development of testicular tumors [24]. *DDX1* is widely expressed in different cell types and tissues, albeit at different levels [20], and appears to be essential for embryonic development as knockout of the *DDX1* gene results in early embryonic lethality in both mice (our unpublished data) and *Drosophila melanogaster* [25]. A number of roles have been proposed for DDX1 including RNA processing, RNA transport from the nucleus to the cytoplasm and RNA clearance at sites of double-strand breaks [26–30]. Although DDX1 is predominantly a nuclear protein, it is also found in the cytoplasm of *DDX1*-amplified neuroblastomas and retinoblastoma cells [31].

Here, we examine DDX1 expression and subcellular location in gene expression microarrays and tissue microarrays designed to identify biomarkers associated with early recurrence in primary breast cancer. We show that over-expression of *DDX1* RNA (by as little as 40%) and elevated levels of DDX1 protein in the cytoplasm can both serve as prognostic markers of recurrence and death. Correlation of DDX1 with recurrence is independent of the commonly used breast cancer markers ER, PR, HER2 amplification, grade and stage, thus identifying DDX1 as a novel prognostic marker.

Materials and methods

Patient selection

Gene expression microarray analysis was performed on 176 primary, treatment- naive breast cancer samples and 10 normal breast tissue samples acquired from reduction mammoplasties through the Canadian Breast Cancer Foundation Tumor Bank. A flow chart depicting patient selection criteria is presented in Fig. 1. Patient information was collected under Research Ethics Board Protocol ETH-02-86-17. The tumor samples, collected at surgery, were frozen in liquid nitrogen within 20 min of devitalization. Evaluation of histology slides from tissue adjacent to the frozen samples indicated that at least 70% of the cells present were invasive tumor cells.

Gene expression analysis

Total RNA was isolated from the frozen samples using Trizol and QIAGEN RNeasy columns. The RNA was quantified using a NanoDrop 1000 Spectrophotometer and its integrity evaluated using a Bioanalyzer 2100. RNA samples with RNA Integrity Numbers (RIN) greater than 7.0 were used.

The RNA was subjected to linear amplification and Cy3 labeling, then hybridized to Agilent Whole Human Genome Arrays using Agilent Technologies kits (One Color Low RNA Input Linear Amplification Kit Plus, One Color RNA Spike-In Kit and Gene Expression Hybridization Kit). The arrays were scanned using an Agilent Scanner. The data were extracted and quality-evaluated using Feature Extraction Software 9.5, and normalized and analyzed using GeneSpring GX 7.3 (Agilent Technologies).

Tissue microarray construction and immunohistochemical analysis

The TMA included three 0.6 mm cores from each of the samples and was constructed using a TMArrayer (Pathology Devices, Westminster, Maryland). TMAs were deparaffinized in xylene, re-hydrated and microwaved for 20 min in epitope retrieval buffer (10 mM citrate, 0.05% Tween-20; pH 6). TMAs were immunostained with rabbit anti-DDX1 antibody (1:2000) (batch 2910) [31] or mouse anti-Ki67 antibody (clone MIB-1; proliferation marker) (DakoCytomation, Carpinteria, California).

Scoring and quantification of immunohistochemical staining

Ki67 scoring to measure proliferative index was performed by a single pathologist (JH) blinded to outcomes, using the

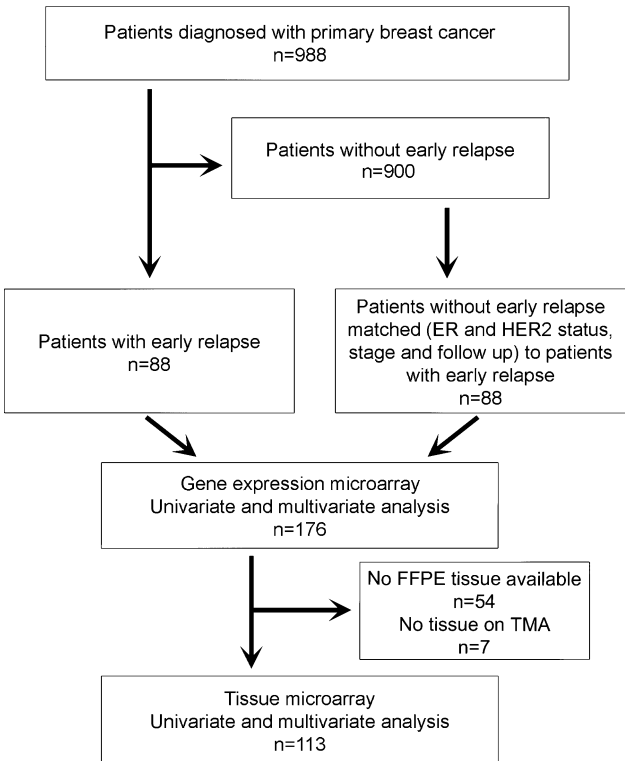


Fig. 1 Flow chart depicting patient selection for the study. Of the original population of 988 consented patients with treatment-naïve primary breast cancer, 88 had suffered an early relapse by September 30, 2009 when the data were locked. The women whose tumors were selected for this study received standardized guideline-based chemo- and hormonal therapies. These treatment guidelines recommend anthracycline chemotherapy for high risk node-negative disease, anthracycline with taxane chemotherapy for node-positive disease, hormonal therapy for all patients with ER+ve disease, and trastuzumab for those with HER2 positive tumors. Two groups of patients were selected for analysis, the first consisting of 88 patients who experienced an early relapse (less than 5 years after the initial treatment), and the second consisting of 88 patients who did not relapse. The two groups were matched for ER and HER2 status, stage and time of follow-up. The median duration of follow-up for surviving patients was 4.5 years. Abbreviation: FFPE, formalin fixed paraffin embedded

MIB1 antibody and dichotomized at 15% nuclear staining [10]. DDX1 protein was scored separately for nuclear and cytoplasmic levels. Each score was based on the average staining intensity throughout the tumor tissue on a scale of 0–3. Cytoplasmic staining was considered high if the score was 2–3 and low if the score was 0–1, while nuclear staining was considered high if the score was 3 and low if the score was 1 or 2 (no tumors were scored 0 as all tumors had some DDX1 in the nucleus). With few exceptions, staining intensity was consistent in all tumor cells throughout a single core. DDX1 staining was scored by DG with 63 random samples independently scored by JH. The agreement between the two sets had a Cohen's kappa value of 0.69 (substantial agreement) with complete agreement in 85% of cases for cytoplasmic intensity and a Cohen's kappa value of 0.57 (moderate agreement) with complete agreement in 78% of cases for nuclear intensity. Acquisition of images was performed using an Axioskop2 plus microscope with a 20× or 40× lens, a ZeissAxiocam and AxioVision software, version 4.7.1.0 (Carl Zeiss MicroImaging, Jena, Germany).

Statistical analysis

Statistical analyses were performed using MedCalc for Windows, version 11.1.0.0 (MedCalc Software, Mariakerke, Belgium). Rank correlation was performed to determine Spearman's Rho. Clinical/pathological variables as a function of DDX1 scores were assessed for both gene expression microarrays and TMAs using the Students's *t* test (continuous variables), Fisher's exact test (2 category variables), and chi-square test (3 or more category variables). Survival and recurrence-free survival were analyzed using the logrank test on Kaplan–Meier survival curves. Cox proportional-hazard regression was performed for univariate analysis using an enter model for survival and recurrence-free survival. Multivariate analysis was performed using a backward enter model with variable removal at $P > 0.10$ to test statistical significance and independence of factors shown to be significant by univariate analysis for survival or recurrence-free survival. This study complies with the Reporting Recommendation for Tumor Marker Prognostic Studies guidelines [32].

Results

Gene expression analysis of DDX1 in breast cancer

Gene expression microarrays were hybridized using RNAs isolated from samples obtained from 176 primary treatment-naive breast cancer patients (45 stage I, 117 stage IIA/IIB, and 14 stage IIIA/IIIB). Eighty-eight of the 176 patients experienced early relapse (recurrence within 5 years) and 57 patients had died when the study was locked (September 30 2009). Of the 176 tumors analysed, 31 were classified as luminal A, 45 as luminal B, 8 as HER2+ve and 56 as triple negative. Thirty-six samples were simply classified as luminal as their Ki67 status was not available (Table 1).

Relative *DDX1* RNA levels in the 176 tumors ranged from 0.497 to 3.437. In comparison, relative *DDX1* RNA levels in 10 normal breast tissue samples ranged from 0.804 to 1.094. ROC curve analysis in relation to recurrence defined a relative RNA level of more than 1.365 as the most appropriate cut-off point (sensitivity 23%, selectivity 91%). Of 176 patients, 28 (16%) had relative *DDX1* RNA levels of >1.365, with the remaining 148 (84%)

Table 1 Clinicopathologic features of the patients included in the gene expression microarray analysis

		Univariate analysis			
		Recurrence		Death	
		HR	*P*	HR	*P*
Number of patients	176				
Age at diagnosis					
Median	52 years				
Range	26–89 years				
Recurrence					
Events	88 (50%)				
Average time to	818 days				
Death					
Events	57 (32%)				
Average time to	1,056 days				
ER status					
Positive	112 (64%)	0.61	**0.03**	0.35	**<0.001**
Negative	64 (36%)				
PR status					
Positive	94 (53%)	0.55	**<0.01**	0.43	**<0.01**
Negative	82 (47%)				
HER2 status					
Amplified	30 (17%)	1.02	0.95	0.98	0.96
Non-amplified	146 (83%)				
Grade					
3	120 (68%)	1.68	**0.04**	1.84	**0.05**
1–2	56 (32%)				
Stage					
I	45 (26%)	1.17	0.46	1.38	0.21
IIA/IIB	117 (66%)				
IIIA/IIIB	14 (8%)				
Subtype					
Luminal A	31 (18%)				
Luminal B	45 (26%)				
Luminal undefined	36 (20%)				
HER2 amplified	8 (5%)				
Triple negative	56 (32%)				
Gene expression microarray score for *DDX1*					
Range	0.497–3.437	2.74	**<0.0001**	2.09	**0.02**
≤1.365	148 (84%)				
>1.365	28 (16%)				

P values correspond to univariate Cox regression analysis. Percentages may not equal 100% due to rounding. *HR* hazard ratio

P values in bold indicate statistical significance ($P < 0.05$)

having relative RNA levels of ≤1.365. Univariate Cox regression analysis showed a significant correlation to both recurrence and death for negative ER status, negative PR status, high grade (defined as grade 3) and elevated *DDX1* RNA levels. There was no correlation to death or recurrence for HER2 status or stage (Table 1).

DDX1 RNA levels correlate with death, recurrence, ER negative status, PR negative status, and high grade

Fisher's exact tests, chi square tests or Student's *t* tests were performed to determine if elevated *DDX1* RNA levels correlated with known prognostic indicators and clinical outcomes. High *DDX1* RNA levels were found to correlate with recurrence, death, negative ER status, negative PR status, and high grade (Fig. 2a, b). There was no correlation between relative *DDX1* RNA levels of >1.365 and HER2 amplification, breast cancer family history, menopause status, stage, and tumor size (Fig. 2b).

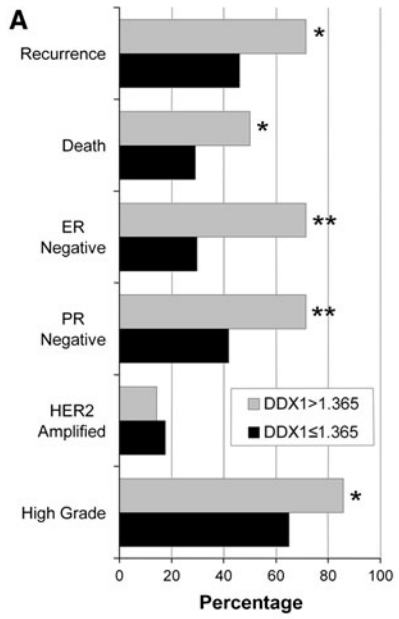
Kaplan–Meier survival curve analysis of high *DDX1* RNA levels showed a higher risk of recurrence with a hazard ratio of 4.31 (95% CI 2.22–9.19, $P < 0.0001$) (Fig. 2c) and a higher risk of death with a hazard ratio of 2.58 (95% CI 1.22–5.61, $P = 0.014$) (Fig. 2d). Similar data were obtained upon analysis of recurrence within systemic therapy subgroups (+ or − adjuvant chemotherapy; + or − adjuvant hormone therapy). High *DDX1* RNA levels were significantly associated with recurrence in patients who received chemotherapy, with a hazard ratio of 8.45 (95% CI 3.38–21.05, $P < 0.0001$) and in patients who received hormone therapy, with a hazard ratio of 14.68 (95% CI 3.49–61.65, $P = 0.0002$). Although not significant, there was also a trend towards increased recurrence in patients who did not receive chemotherapy and in patients who did not receive hormone therapy.

DDX1 protein subcellular localization in breast cancer tissue

A TMA was generated using breast cancer tissue samples from 120 (of the original 176) patients, of which seven were discarded because of insufficient tissue left on the TMA. The TMA also included cores from six normal breast tissue. TMAs were immunostained with anti-DDX1 antibody. Sixty-two of the 113 patients represented in the TMAs had recurred at the time of analysis, and 32 patients had died (Table 2).

As different levels of DDX1 protein were observed in the cytoplasm and nucleus, cytoplasmic and nuclear DDX1 protein were individually scored (Figs. 3b–g). Nuclear staining was scored on a relative scale of 1–3, while cytoplasmic staining was scored on a relative scale of 0–3. There was a non-random distribution of nuclear to cytoplasmic staining intensity ($P < 0.001$), with an inverse relationship between cytoplasmic and nuclear levels (rho = −0.28, $P = 0.0027$) (Fig. 3a). Thirty-seven of 113 (33%) tumors had a cytoplasmic score of 2 or 3. In contrast, all six normal breast tissues had cytoplasmic scores of 0 or 1 and nuclear scores of 3 (Fig. 3g).

Univariate analysis of the 113 tumor samples showed a significant correlation with both recurrence and death



B

DDX1 RNA	>1.365	≤1.365	p
Total	28	148	
Recurrence	20	68	**p=0.011**
Death	14	43	**p=0.028**
Negative ER status	20	44	**p<0.001**
Negative PR status	20	62	**p<0.01**
High grade	24	96	**p=0.024**
HER2 amplified	4	26	p=0.45
Breast cancer family history	11	66	p=0.38
Pre-menopausal	11	49	
Peri-menopausal	3	12	p=0.98
Post-menopausal	14	87	
Stage I	8	37	
Stage IIA/B	17	100	p=0.74
Stage IIIA/B	3	11	
Average tumor size	2.78 cm	2.62 cm	p=0.62

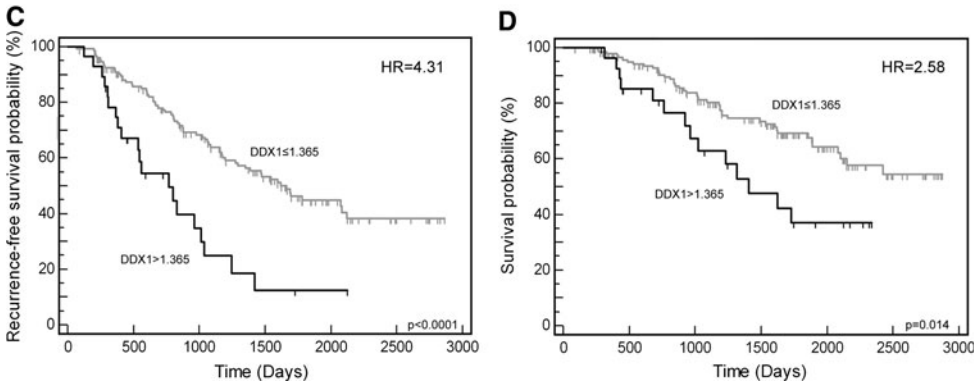


Fig. 2 **a**, **b** Occurrence of clinical/pathological features in patients with relative *DDX1* RNA levels of >1.365 ($n = 28$) compared to patients with relative *DDX1* RNA levels of ≤1.365 ($n = 148$). High *DDX1* RNA levels were found to correlate with recurrence [71% of cases with elevated *DDX1* RNA (>1.365) showed recurrence compared to 46% of cases with low *DDX1* RNA (≤1.365)], death (50% compared to 29%), negative ER status (71% compared to 29%), negative PR status (71% compared to 42%), and high grade (86% compared to 65%). There was no correlation between relative *DDX1* RNA levels of >1.365 and HER2 amplification (14% compared to 18%), breast cancer family history (39% compared to 45%), menopause status (39% pre-, 50% post-, and 11% peri-menopausal compared to 33% pre-, 59% post-, and 8% peri-menopausal), stage (29% stage I, 61% stage IIA/B and 11% stage IIIA/B compared to 25% stage I, 68% stage IIA/B, and 7% stage IIIA/B), and tumor size (average size of 2.8 cm compared to 2.6 cm). * Indicates $P < 0.05$, ** indicates $P < 0.01$. *P* values in bold indicate statistical significance ($P < 0.05$). **c**, **d** Kaplan–Meier survival curves of patients with relative *DDX1* RNA levels >1.365 ($n = 28$) compared to patients with relative *DDX1* RNA levels ≤1.365 ($n = 148$). **c** Recurrence-free survival. **d** Survival. *HR* hazard ratio

for: (i) negative PR status and (ii) elevated levels of DDX1 in the cytoplasm. A significant correlation was also observed between grade and recurrence, but not grade and death. A negative ER status was correlated with death only, whereas decreased levels of DDX1 protein in the nucleus was correlated with recurrence only (Table 2).

Subcellular DDX1 protein localization correlates with death, recurrence, ER negative status, and PR negative status

Statistical analysis was performed as described for *DDX1* RNA to determine if elevated levels of DDX1 protein in the cytoplasm (defined by a score of 2 or 3) compared to low levels in the cytoplasm (defined by a score of 0 or 1) correlated with known prognostic indicators and clinical outcomes. Elevated levels of DDX1 protein in the cytoplasm were found to correlate with recurrence, death, negative ER status, negative PR status, and high grade (Figs. 4a, b). There was no correlation between elevated levels of cytoplasmic DDX1 protein and HER2 amplification, breast cancer family history, menopause status, stage, and tumor size (Fig. 4b). Kaplan–Meier survival curve analysis of cytoplasmic DDX1 localization showed a higher risk of recurrence with a hazard ratio of 1.90 (95% CI 1.09–3.34, $P = 0.0237$) (Fig. 4c) and death with a

Table 2 Clinicopathologic features of the patients included in the TMA analysis

		Univariate analysis			
		Recurrence		Death	
		HR	*P*	HR	*P*
Number of patients	113				
Age at diagnosis					
Median	51 years				
Range	26–89 years				
Recurrence					
Events	62 (55%)				
Average time to	755 days				
Death					
Events	32 (29%)				
Average time to	910 days				
ER status					
Positive	70 (62%)	0.65	0.10	0.35	**<0.01**
Negative	43 (38%)				
PR status					
Positive	58 (51%)	0.53	**0.01**	0.42	**0.02**
Negative	55 (49%)				
HER2 status					
Amplified	21 (19%)	0.90	0.77	1.00	0.99
Non-amplified	92 (81%)				
Grade					
3	78 (69%)	1.88	**0.03**	2.36	0.06
1–2	35 (31%)				
Stage					
I	29 (26%)	0.96	0.84	1.41	0.25
IIA/IIB	74 (65%)				
IIIA/IIIB	10 (9%)				
Subtype					
Luminal A	34 (30%)				
Luminal B	36 (32%)				
HER2 amplified	6 (5%)				
Triple negative	37 (33%)				
Cytoplasmic DDX1 intensity					
0	28 (25%)				
1	48 (42%)				
2	28 (25%)	0–1 vs. 2–3		0–1 vs. 2–3	
3	9 (8%)	1.79	**0.03**	2.51	**<0.01**
Nuclear DDX1 intensity					
1	22 (19%)				
2	31 (27%)	1 vs. 2 vs. 3		1 vs. 2 vs. 3	
3	60 (53%)	0.72	**0.04**	0.75	0.18

P values correspond to univariate Cox regression analysis. Percentages may not equal 100% due to rounding. *HR* hazard ratio

P values in bold indicate statistical significance ($P < 0.05$)

hazard ratio of 2.79 (95% CI 1.32–5.89, $P = 0.0073$) (Fig. 4d).

There was also correlation between low levels of nuclear DDX1 protein (score of 1 or 2 in the nucleus) and known prognostic markers (negative ER status, negative PR status and high grade) but not clinical outcome based on the number of events (data not shown). However, we did observe a non-significant trend between low levels of DDX1 protein in the nucleus and recurrence, but not death, based on logrank analysis of Kaplan–Meier curves (HR = 1.65, 95% CI 0.99–2.76, $P = 0.055$) and univariate Cox regression analysis (HR = 1.62, 95% CI 0.99–2.68, $P = 0.058$).

DDX1 RNA levels and protein localization predict recurrence independently of common markers

Multivariate Cox regression analysis of factors shown to be significant in our univariate analysis (*DDX1* RNA levels, DDX1 localization, ER status, PR status, and grade) was performed to determine if either *DDX1* RNA levels or protein localization, or both parameters, were independently predictive of death and recurrence (Table 3).

First, we carried out multivariate analysis on the four variables found to be significantly associated with survival or recurrence in the 176-patient gene expression microarray study: relative *DDX1* RNA levels, ER status, PR status, and grade. We used the backward stepwise method to remove variables at each step based on a 0.1 level of significance. Only one variable was retained when survival was modeled: ER status (HR = 0.35 95% CI 0.21–0.60, $P = 0.0001$). Upon modelling recurrence-free survival in the same cohort, all four factors were retained, *DDX1* RNA levels (HR = 2.61 95% CI 1.50–4.54, $P = 0.0007$), ER status (HR = 1.92 95% CI 0.89–4.12, $P = 0.10$), PR status (HR = 0.43 95% CI 0.23–0.83, $P = 0.01$), and grade (HR = 1.67 95% CI 0.96–2.90, $P = 0.07$). Three of the four variables retained in our model for recurrence had hazard ratios similar to those calculated using univariate analysis (*DDX1* RNA level, PR status, and grade) suggesting that they have independent prognostic value. The hazard ratio for ER status is significantly different in the univariate analysis (0.61) compared to the multivariate analysis (1.92), suggesting that ER status does not confer an independent prognostic value. This is expected as our patient cohort (relapsed vs. non-relapsed) was controlled for ER status.

Second, we carried out multivariate analysis with the four variables found to be significantly associated with survival or recurrence in the 113-patient TMA study: cytoplasmic DDX1 protein, ER status, PR status, and grade. Upon modelling survival, only cytoplasmic DDX1

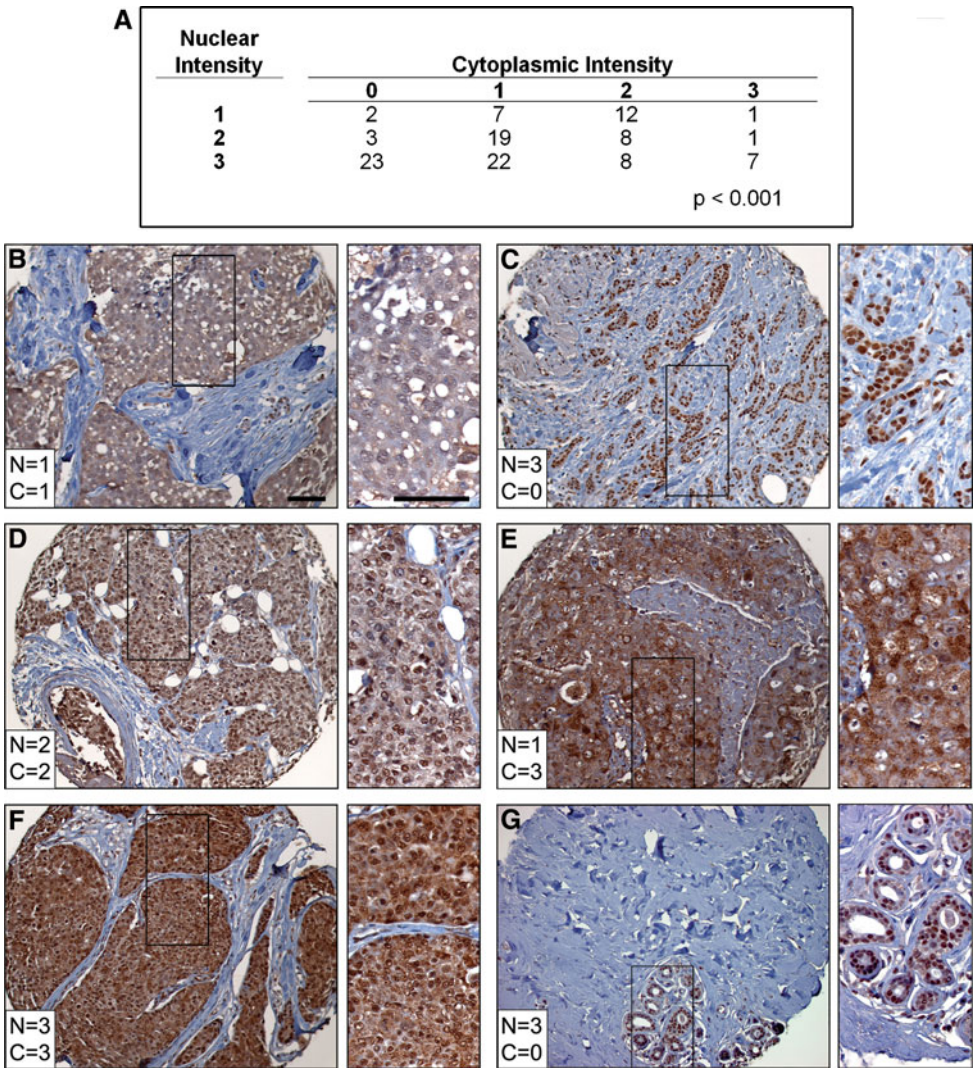


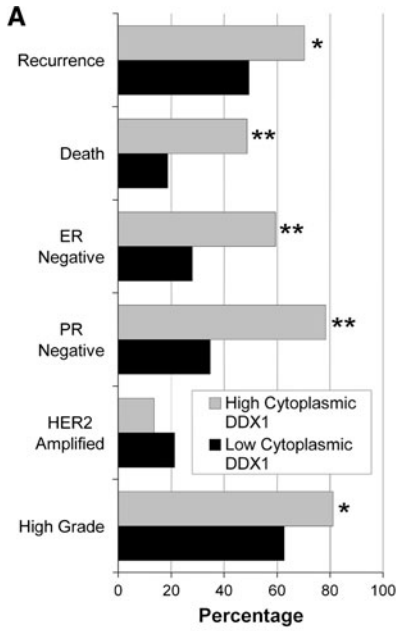
Fig. 3 DDX1 protein subcellular location in breast cancer TMAs. **a** The intensity of cytoplasmic DDX1 protein is plotted against the intensity of nuclear DDX1 protein. The distribution of DDX1 nuclear and cytoplasmic intensities is non-random with a p value of <0.001 based on chi square distribution analysis. **b** Breast cancer tissue from patient MT861 had a nuclear DDX1 score of 1 ($N = 1$) and a cytoplasmic DDX1 score of 1 ($C = 1$). The boxed area is magnified in the right panel. **c–f** The nuclear and cytoplasmic DDX1 scores for patient GT178 (**c**), patient MT340 (**d**), patient GT226 (**e**), and patient MT604 (**f**) are indicated at the bottom left of each figure. **g** Normal breast tissue had a nuclear DDX1 score of 3 and a cytoplasmic DDX1 score of 0. Scale bars = 60 μm

protein (HR = 1.97 95% CI 0.96–4.06 $P = 0.067$) and ER status (HR = 0.43 95% CI 0.21–0.89, $P = 0.024$) were retained. The model for recurrence-free survival retained only cytoplasmic DDX1 protein (HR = 1.73 95% CI 1.04–2.88, $P = 0.036$) and grade (HR = 1.82 95% CI 1.02–3.27, $P = 0.045$). Cytoplasmic DDX1 protein and grade were retained with hazard ratios similar to those generated by univariate analysis suggesting that they both provide independent prognostic significance.

Discussion

Breast cancer is increasingly managed on the basis of molecular classification. There is widespread consensus that ER+/PR+/low proliferation tumors are associated with a good outcome while HER2-positive and triple-negative tumors are associated with a poor outcome. Adjuvant hormonal therapy in ER+ve breast cancers, which constitute ~70% of breast cancers in developed countries, reduces the relative risk of death by approximately 22% and the risk of recurrence by 42% [33]. Nonetheless, a significant number of ER+ve tumors will relapse. At issue are the diverse nature of breast cancer and the complexity and multitude of events leading to tumor formation and progression.

Here, we use gene expression and immunohistochemical analysis to investigate DDX1 expression in 176 primary breast cancers, half of which were selected for early recurrence. We demonstrate a highly significant correlation between recurrence and increases in *DDX1* RNA levels, with a hazard ratio of 4.31. We also observe a significant correlation between recurrence and elevated DDX1 protein in the cytoplasm, and a non-significant trend between recurrence and low levels of DDX1 in the nucleus. Furthermore, analysis of systemic therapy subgroups suggests that elevated levels of *DDX1* RNA is a prognostic factor for all treatment subpopulations. Assessment of DDX1's



B

Cytoplasmic DDX1	2 - 3	0 - 1	p
Total	37	76	
Recurrence	25	37	**p=0.045**
Death	17	15	**p<0.01**
Negative ER status	22	21	**p<0.01**
Negative PR status	29	26	**p<0.0001**
High grade	30	48	**p=0.041**
HER2 amplified	5	16	p=0.24
Breast cancer family history	18	29	p=0.21
Pre-menopausal	12	27	
Peri-menopausal	22	42	p=0.91
Post-menopausal	3	7	
Stage I	8	21	
Stage IIA/B	23	51	p=0.15
Stage IIIA/B	6	4	
Average tumor size	3.31 cm	2.71 cm	p=0.17

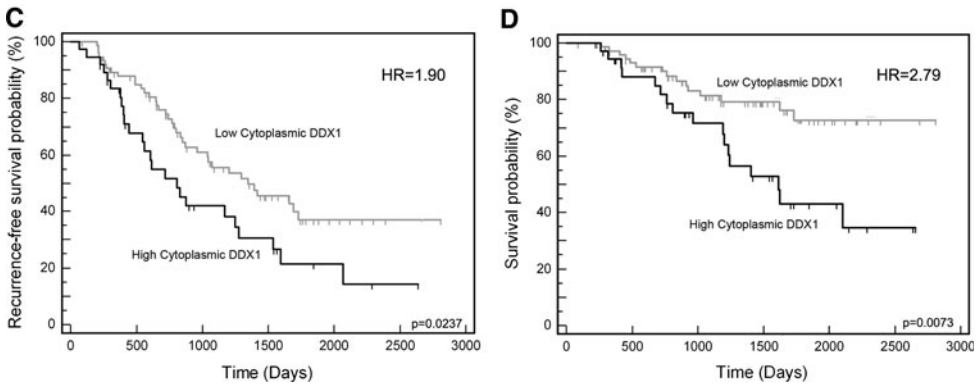


Fig. 4 **a**, **b** Occurrence of clinical/pathological features in patients with high levels of DDX1 protein in the cytoplasm (DDX1 scores of 2 or 3) ($n = 37$) compared to patients with low levels of DDX1 in the cytoplasm (DDX1 scores of 0 or 1) ($n = 76$). Elevated levels of DDX1 protein in the cytoplasm were found to correlate with death (46% of cases with elevated cytoplasmic DDX1 protein levels compared to 20% of cases with low cytoplasmic DDX1 protein levels,), recurrence (68% compared to 49%), negative ER status (59% compared 28%), negative PR status (78% compared to 34%, $P < 0.0001$) and high grade (81% compared to 63%). There was no correlation between elevated levels of cytoplasmic DDX1 protein and HER2 amplification (14% compared to 21%), breast cancer family history (49% compared to 38%), menopause status (32% pre-, 59% post-, and 8% peri-menopausal compared to 36% pre-, 55% post-, and 9% peri-menopausal), stage (22% stage I, 62% stage IIA/B and 16% stage IIIA/B compared to 28% stage I, 67% stage IIA/B, and 5% stage IIIA/B), and tumor size (average size of 3.3 compaired to 2.7 cm). * Indicates $P < 0.05$, ** indicates $P < 0.01$. *P* values in bold indicate statistical significance ($P < 0.05$). **c**, **d** Kaplan–Meier survival curves of patients with cytoplasmic DDX1 scores of 2 and 3 ($n = 37$) compared to patients with cytoplasmic DDX1 scores of 0 and 1 ($n = 76$). **c** Recurrence-free survival. **d** Survival. *HR* hazard ratio

predictive value for treatment outcome in primary breast cancer will require evaluation of DDX1 in the context of a randomized clinical trial.

Cox multivariate analysis of high *DDX1* RNA levels and DDX1 cytoplasmic localization indicate that both *DDX1* RNA levels and cytoplasmic localization are independent markers of recurrence. In both cases, the hazard ratio remained relatively unchanged between univariate and multivariate analysis, demonstrating that additional factors (ERα, grade) did not significantly modulate the effect of *DDX1* RNA or protein localization. Cytoplasmic DDX1 localization, but not *DDX1* RNA levels, was also independently correlated with death. These results suggest that DDX1 analysis refines prognostic assessments using standard clinicopathologic parameters (stage, grade, hormone receptor and HER2 status) in a population receiving guideline-based standardized adjuvant therapy.

It is not clear to what extent increased *DDX1* RNA levels correlates with increased DDX1 protein levels in the breast cancer tissues analysed. Although there was considerable overlap between those recurrences characterized by elevated *DDX1* RNA levels and those characterized by

Table 3 Cox multivariate analysis of survival and recurrence-free survival based on the gene expression microarray data and TMA data

Factor	Survival			Recurrence-free survival		
	HR	95% CI	*P*	HR	95% CI	*P*
Microarray $n = 176$						
DDX1	N/S	N/S	N/S	2.61	1.50–4.54	<0.001
ER	0.35	0.21–0.60	0.0001	1.92	0.89–4.12	0.10
PR	N/S	N/S	N/S	0.43	0.23–0.83	0.01
Grade	N/S	N/S	N/S	1.67	0.96–2.90	0.07
TMA $n = 113$						
DDX1	1.97	0.96–4.06	0.07	1.73	1.04–2.88	0.04
ER	0.43	0.21–0.89	0.02	N/S	N/S	N/S
PR	N/S	N/S	N/S	N/S	N/S	N/S
Grade	N/S	N/S	N/S	1.82	1.02–3.27	0.04

HR hazard ratio, *CI* confidence interval

elevated levels of DDX1 protein in the cytoplasm, it seems unlikely that the relatively small increases in *DDX1* RNA levels (>40%) detected by gene expression microarray analysis could account for the considerable increases in cytoplasmic DDX1 protein levels observed by TMA analysis We postulate that deregulation of DDX1, be it at the expression or subcellular distribution level, is at the heart of its association with recurrence. In support of this idea, analyses of a wide variety of tissues and cell lines demonstrate that DDX1 is primarily a nuclear protein [26, 31], with the exception of *MYCN/DDX1*-amplified retinoblastoma and neuroblastoma tumor cells which show equal distribution of DDX1 protein in the nucleus and cytoplasm [31]. Furthermore, we have not been able to stably alter DDX1 protein levels in either cell lines or transgenic mice (our unpublished data), and knock-out of *DDX1* in fruit flies produces an embryonic lethal phenotype [25]. Combined, these data suggest that levels of DDX1 and its subcellular distribution are tightly controlled and that it is only when cells become tumorigenic that this regulation is relaxed.

DDX1 is a DEAD box protein that can bind and unwind DNA/RNA and RNA/RNA duplexes in vitro [29]. Roles proposed for DDX1 include RNA processing [26, 34], transcription regulation [24], DNA double-strand break repair [29], and RNA transport [27, 28]. While the first three roles are strictly dependent on the presence of DDX1 protein in the nucleus, RNA transport involves shuttling of molecules between the nucleus and cytoplasm, and to specific regions of the cytoplasm. Deregulation of DDX1 could result in altered subcellular localization of RNAs, which in turn could affect the availability of specific RNAs for translation. Thus, breast cancer cells with elevated levels of cytoplasmic DDX1 protein may exhibit alterations in their complement of translated proteins. As increases in cytoplasmic DDX1 protein are associated with a worse prognosis, reduction of its extraneous cytoplasmic activity represents an attractive therapeutic option. One possibility is to target DDX1 with small molecule inhibitors, as reported for two other DEAD box proteins [35, 36].

The mechanisms of action of several members of the DEAD box protein family implicated in cancer have been investigated. For example, p68 (DDX5) and p72 (DDX17) have been shown to interact with ERα and to alter ERα transcription activity in breast cancer cells [37, 38]. Expression of p72 in ERα-positive breast cancers is associated with longer recurrence-free survival and overall survival, and is inversely correlated with HER2 expression. DDX6 (RCK/p54), is over-expressed in colorectal cancer, and may be deregulating proliferation by activating the Wnt pathway [39]. Finally, DDX53 (CAGE), normally specific to the testis, is expressed in a variety of cancers, including lung, cervical, and colon [40]. The wide spectrum of associations between DEAD box proteins and cancer define this family of proteins as an attractive target for future therapies.

In summary, we show that increased *DDX1* RNA levels and cytoplasmic localization of DDX1 protein both correlate with increased risk of recurrence in breast cancer, independently of commonly used markers such as ERα and grade. Future study will involve determining whether DDX1 can serve as a prognostic marker for all subtypes of breast cancer, and to assess DDX1's potential as a predictive biomarker and breast cancer therapeutic target.

Acknowledgements We are grateful to Lillian Cook and Cheryl Santos for their excellent technical assistance, Adrian Driga for his help with database analysis and Sunita Ghosh for her assistance with statistical analyses. This study was supported by grants from the Alberta Cancer Board—Research Initiative Program (R.G.), Alberta Cancer Research Institute (J.R.M.) and an Alberta Cancer Foundation studentship (D.R.G.).

References

1. Porter P (2008) "Westernizing" women's risks? Breast cancer in lower-income countries. N Engl J Med 358:213–216
2. Jones SE (2008) Metastatic breast cancer: the treatment challenge. Clin Breast Cancer 8:224–233
3. Bergh J, Jonsson PE, Glimelius B, Nygren P (2001) A systematic overview of chemotherapy effects in breast cancer. Acta Oncol 40:253–281
4. Lopez-Tarruella S, Martin M (2009) Recent advances in systemic therapy: advances in adjuvant systemic chemotherapy of early breast cancer. Breast Cancer Res 11:204
5. Barnadas A, Gil M, Sanchez-Rovira P, Llombart A, Adrover E, Estevez LG, De la Haba J, Calvo L (2008) Neoadjuvant endocrine therapy for breast cancer: past, present and future. Anticancer Drugs 19:339–347
6. Fossati R, Confalonieri C, Torri V, Ghislandi E, Penna A, Pistotti V, Tinazzi A, Liberati A (1998) Cytotoxic and hormonal treatment for metastatic breast cancer: a systematic review of published randomized trials involving 31,510 women. J Clin Oncol 16:3439–3460

7. Cheang MC, Chia SK, Voduc D, Gao D, Leung S, Snider J, Watson M, Davies S, Bernard PS, Parker JS, Perou CM, Ellis MJ, Nielsen TO (2009) Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. J Natl Cancer Inst 101:736–750
8. Fan C, Oh DS, Wessels L, Weigelt B, Nuyten DS, Nobel AB, van't Veer LJ, Perou CM (2006) Concordance among gene-expression-based predictors for breast cancer. N Engl J Med 355:560–569
9. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO, Botstein D (2000) Molecular portraits of human breast tumours. Nature 406:747–752
10. Hugh J, Hanson J, Cheang MC, Nielsen TO, Perou CM, Dumontet C, Reed J, Krajewska M, Treilleux I, Rupin M, Magherini E, Mackey J, Martin M, Vogel C (2009) Breast cancer subtypes and response to docetaxel in node-positive breast cancer: use of an immunohistochemical definition in the BCIRG 001 trial. J Clin Oncol 27:1168–1176
11. Kreike B, van Kouwenhove M, Horlings H, Weigelt B, Peterse H, Bartelink H, van de Vijver MJ (2007) Gene expression profiling and histopathological characterization of triple-negative/basal-like breast carcinomas. Breast Cancer Res 9:R65
12. Brenton JD, Carey LA, Ahmed AA, Caldas C (2005) Molecular classification and molecular forecasting of breast cancer: ready for clinical application? J Clin Oncol 23:7350–7360
13. Chiuri VE, Leo G, Lorusso V (2007) Clinical and therapeutic perspectives of gene expression profiling for breast cancer. Ann Oncol 18(Suppl 6):Vi58–Vi62
14. Oakman C, Bessi S, Zafarana E, Galardi F, Biganzoli L, Di Leo A (2009) Recent advances in systemic therapy: new diagnostics and biological predictors of outcome in early breast cancer. Breast Cancer Res 11:205
15. Stadler ZK, Come SE (2009) Review of gene-expression profiling and its clinical use in breast cancer. Crit Rev Oncol Hematol 69:1–11
16. Slodkowska EA, Ross JS (2009) MammaPrint 70-gene signature: another milestone in personalized medical care for breast cancer patients. Expert Rev Mol Diagn 9:417–422
17. Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, Ravdin P, Bugarini R, Baehner FL, Davidson NE, Sledge GW, Winer EP, Hudis C, Ingle JN, Perez EA, Pritchard KI, Shepherd L, Gralow JR, Yoshizawa C, Allred DC, Osborne CK, Hayes DF (2010) Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. Lancet Oncol 11:55–65
18. Mook S, Schmidt MK, Weigelt B, Kreike B, Eekhout I, van de Vijver MJ, Glas AM, Floore A, Rutgers EJ, van 't Veer LJ (2009) The 70-gene prognosis signature predicts early metastasis in breast cancer patients between 55 and 70 years of age. Ann Oncol. doi:10.1093/annonc/mdp388
19. Linder P (2006) Dead-box proteins: a family affair—active and passive players in RNP-remodeling. Nucleic Acids Res 34:4168–4180
20. Godbout R, Squire J (1993) Amplification of a DEAD box protein gene in retinoblastoma cell lines. Proc Natl Acad Sci USA 90:7578–7582
21. Squire JA, Thorner PS, Weitzman S, Maggi JD, Dirks P, Doyle J, Hale M, Godbout R (1995) Co-amplification of MYCN and a DEAD box gene (DDX1) in primary neuroblastoma. Oncogene 10:1417–1422
22. Manohar CF, Salwen HR, Brodeur GM, Cohn SL (1995) Co-amplification and concomitant high levels of expression of a DEAD box gene with MYCN in human neuroblastoma. Genes Chromosomes Cancer 14:196–203
23. George RE, Kenyon RM, McGuckin AG, Malcolm AJ, Pearson AD, Lunec J (1996) Investigation of co-amplification of the candidate genes ornithine decarboxylase, ribonucleotide reductase, syndecan-1 and a DEAD box gene, DDX1, with N-myc in neuroblastoma. United Kingdom Children's Cancer Study Group. Oncogene 12:1583–1587
24. Tanaka K, Okamoto S, Ishikawa Y, Tamura H, Hara T (2009) DDX1 is required for testicular tumorigenesis, partially through the transcriptional activation of 12p stem cell genes. Oncogene 28:2142–2151
25. Zinsmaier KE, Eberle KK, Buchner E, Walter N, Benzer S (1994) Paralysis and early death in cysteine string protein mutants of Drosophila. Science 263:977–980
26. Bleoo S, Sun X, Hendzel MJ, Rowe JM, Packer M, Godbout R (2001) Association of human DEAD box protein DDX1 with a cleavage stimulation factor involved in 3′-end processing of pre-MRNA. Mol Biol Cell 12:3046–3059
27. Kanai Y, Dohmae N, Hirokawa N (2004) Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. Neuron 43:513–525
28. Fang J, Kubota S, Yang B, Zhou N, Zhang H, Godbout R, Pomerantz RJ (2004) A DEAD box protein facilitates HIV-1 replication as a cellular co-factor of Rev. Virology 330:471–480
29. Li L, Monckton EA, Godbout R (2008) A role for DEAD box 1 at DNA double-strand breaks. Mol Cell Biol 28:6413–6425
30. Li L, Roy K, Katyal S, Sun X, Bleoo S, Godbout R (2006) Dynamic nature of cleavage bodies and their spatial relationship to DDX1 bodies, Cajal bodies, and gems. Mol Biol Cell 17:1126–1140
31. Godbout R, Packer M, Bie W (1998) Overexpression of a DEAD box protein (DDX1) in neuroblastoma and retinoblastoma cell lines. J Biol Chem 273:21161–21168
32. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM (2006) REporting recommendations for tumor MARKer prognostic studies (REMARK). Breast Cancer Res Treat 100:229–235
33. Early Breast Cancer Trialists' Collaborative Group (EBCTCG) (2005) Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 365:1687–1717
34. Sunden Y, Semba S, Suzuki T, Okada Y, Orba Y, Nagashima K, Umemura T, Sawa H (2007) Identification of DDX1 as a JC virus transcriptional control region-binding protein. Microbiol Immunol 51:327–337
35. Bordeleau ME, Matthews J, Wojnar JM, Lindqvist L, Novac O, Jankowsky E, Sonenberg N, Northcote P, Teesdale-Spittle P, Pelletier J (2005) Stimulation of mammalian translation initiation factor eIF4A activity by a small molecule inhibitor of eukaryotic translation. Proc Natl Acad Sci USA 102:10460–10465
36. Erkizan HV, Kong Y, Merchant M, Schlottmann S, Barber-Rotenberg JS, Yuan L, Abaan OD, Chou TH, Dakshanamurthy S, Brown ML, Uren A, Toretsky JA (2009) A small molecule blocking oncogenic protein EWS-FLI1 interaction with RNA helicase A inhibits growth of Ewing's sarcoma. Nat Med 15:750–756
37. Fuller-Pace FV, Ali S (2008) The DEAD box RNA helicases p68 (Ddx5) and p72 (Ddx17): novel transcriptional co-regulators. Biochem Soc Trans 36:609–612
38. Wortham NC, Ahamed E, Nicol SM, Thomas RS, Periyasamy M, Jiang J, Ochocka AM, Shousha S, Huson L, Bray SE, Coombes RC, Ali S, Fuller-Pace FV (2009) The DEAD-box protein p72 regulates ERalpha-/oestrogen-dependent transcription and cell growth, and is associated with improved survival in ERalpha-positive breast cancer. Oncogene. doi: 10.1038/onc.2009.261

39. Lin F, Wang R, Shen JJ, Wang X, Gao P, Dong K, Zhang HZ (2008) Knockdown of RCK/p54 expression by RNAi inhibits proliferation of human colorectal cancer cells in vitro and in vivo. Cancer Biol Ther 7:1669–1676

40. Cho B, Lim Y, Lee DY, Park SY, Lee H, Kim WH, Yang H, Bang YJ, Jeoung DI (2002) Identification and characterization of a novel cancer/testis antigen gene CAGE. Biochem Biophys Res Commun 292:715–726