



**HAL**  
open science

## **Glycodelin expression associates with differential tumour phenotype and outcome in sporadic and familial non-/ breast cancer patients**

Laura C. Hautala, Dario Greco, Riitta Koistinen, Tuomas Heikkinen, Päivi Heikkilä, Kristiina Aittomäki, Carl Blomqvist, Hannu Koistinen, Heli Nevanlinna

### ► To cite this version:

Laura C. Hautala, Dario Greco, Riitta Koistinen, Tuomas Heikkinen, Päivi Heikkilä, et al.. Glycodelin expression associates with differential tumour phenotype and outcome in sporadic and familial non-/ breast cancer patients. *Breast Cancer Research and Treatment*, 2010, 128 (1), pp.85-95. 10.1007/s10549-010-1065-y . hal-00615368

**HAL Id: hal-00615368**

**<https://hal.science/hal-00615368>**

Submitted on 19 Aug 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **Glycodelin expression associates with differential tumour phenotype and outcome in sporadic and familial non-*BRCA1/2* breast cancer patients**

Laura C Hautala<sup>1,5</sup>, Dario Greco<sup>1</sup>, Riitta Koistinen<sup>5</sup>, Tuomas Heikkinen<sup>1</sup>, Päivi Heikkilä<sup>2</sup>,  
Kristiina Aittomäki<sup>3</sup>, Carl Blomqvist<sup>4</sup>, Hannu Koistinen<sup>5</sup>, Heli Nevanlinna<sup>1</sup>

*Departments of<sup>1</sup>Obstetrics and Gynaecology, <sup>2</sup>Pathology, <sup>3</sup>Clinical Genetics and <sup>4</sup>Oncology, Helsinki University Central Hospital, Helsinki, Finland, <sup>5</sup>Department of Clinical Chemistry, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland*

*Corresponding author: Heli Nevanlinna, Helsinki University Central Hospital, Department of Obstetrics and Gynaecology, Biomedicum Helsinki, P.O. Box 700, FIN-00029 HUS, Finland; e-mail: [heli.nevanlinna@hus.fi](mailto:heli.nevanlinna@hus.fi); Tel. +358-9-4717-1750, Fax +358-9-47171-1751.*

## Abstract

**Introduction** Glycodelin (encoded by *PAEP* gene) is a secreted lipocalin protein mainly expressed in reproductive tissues, but also in several tumour types. In the breast, glycodelin is expressed both in normal epithelial and cancerous tissue. To investigate the association of glycodelin with clinicopathological features of breast cancer and outcome of patients we evaluated the protein expression of glycodelin in a large series of breast tumours.

**Methods** Immunohistochemical analysis of tissue microarrays was used to study glycodelin expression on 399 sporadic and 436 familial non-*BRCA1/2* tumours with strong family history. Gene expression analysis was used to define genes co-expressed with *PAEP* in sporadic and familial non-*BRCA1/2* breast tumours.

**Results** In the sporadic series, the glycodelin expression associated with low proliferation rate ( $P < 0.001$ ), with a tendency towards well-differentiated tumours (grades 1 and 2,  $P = 0.012$ ) and high cyclin D1 ( $P = 0.034$ ) expression. However, in familial non-*BRCA1/2* cases with strong family history glycodelin expression associated with a less favourable phenotype, i.e. positive lymph node status ( $P = 0.003$ ) and HER2-positive tumours ( $P = 0.009$ ). Moreover, the patients with glycodelin positive tumours had an increased risk for distant metastases ( $P = 0.001$ ) and in multivariate analysis glycodelin expression was an independent predictor of metastasis [hazard ratio (HR) = 2.22, 95% confidence interval (95% CI) = 1.22-4.03,  $P = 0.009$ ] in familial non-*BRCA1/2* breast cancer. Gene expression analysis further revealed different gene expression profiles correlating with the *PAEP* expression in the sporadic and familial non-*BRCA1/2* breast cancers.

**Conclusion** Our findings suggest differential progression pathways in the sporadic and familial non-*BRCA1/2* breast tumours expressing glycodelin.

## Introduction

Glycodelin (encoded by *PAEP* gene) is a secreted lipocalin protein mainly expressed in reproductive tissues [1]. It is also found in bone marrow and in many different glandular tissues of the human body. In the uterus, glycodelin is secreted in response to progesterone exposure from the endometrial glands during the secretory phase of the menstrual cycle [1]. In males, a differentially glycosylated form of glycodelin is expressed in the seminal vesicles [2]. Different glycoforms of glycodelin play important roles in regulation of sperm function [3]. Furthermore, glycodelin has immunosuppressive properties, which are thought to protect fetal semiallograft from maternal immune response [4, 5].

Several studies have shown that glycodelin is not only associated with differentiated epithelium, but it also induces cell differentiation in endometrial and breast cancer cells [5-11]. Glycodelin differentiates MCF-7 breast cancer cells towards normal, less malignant direction [6, 7]. When grown in Matrigel basement membrane matrix, the cells expressing glycodelin formed spheroids with acini, while the control cells formed more irregular structures. At the same time, glycodelin expression induced changes in the expression of genes, e.g., *CAVI* and *FGF2*, that are associated with normal mammary epithelium and inhibition of tumour growth [7]. In pre-clinical xenograft mouse model, the glycodelin-transfected MCF-7 cells formed significantly smaller tumours than the control cells. In addition, the glycodelin expressing tumours had lower proliferation rate (according to Ki-67 staining) than the control tumours [7]. Thus, in hormone-related cancers, glycodelin is considered to be a tumour growth suppressing factor.

Glycodelin has been found to be expressed in several tumour types, but to date its clinical significance is unknown. In ovarian cancer, glycodelin staining has been found to be associated with favourable prognosis [12]. In breast, glycodelin is expressed both in normal epithelial and cancerous tissue, where it is localized in the cytoplasm [13]. In mucinous breast cancer, an additional diffuse staining of the surrounding mucin may also exist. Previously, the expression of glycodelin has been reported to be associated with non-metastatic breast tumours in a series of 38 tumours as well as oestrogen and progesterone receptor positive tumours in 121 invasive breast cancers [14, 15]. However, a more extensive study is needed to elucidate the role of glycodelin expression in breast cancer. Furthermore, glycodelin expression has not yet been investigated in familial breast cancer. In this study, we evaluated the expression of glycodelin in extensive series of sporadic and familial breast cancers in order to investigate the association of glycodelin with clinicopathological features of the tumours and the outcome in breast cancer, also with respect of the familial background of the patients. We further used gene expression analysis of breast tumours to investigate correlation of gene expression profiles with the *PAEP* expression in sporadic and familial breast tumours.

## **Materials and methods**

### *Patients*

A series of 884 unselected breast cancer patients was recruited at the Department of Oncology, Helsinki University Central Hospital, during the years 1997-1998 and 2000 [16, 17]. An additional familial breast cancer patient series ( $N = 546$ ) was recruited at the Departments of Oncology and Clinical Genetics [18]. For the tissue microarrays, altogether 1356 invasive breast cancer tumours were available. Of these tumours, 423 were derived from patients who did not have a familial background of breast cancer (i.e. sporadic), 818 were

from patients with family history and 56 from *BRCA1* and 59 from *BRCA2* mutation positive families. Of the cases with family history, 481 tumours were from patients with strong family history of breast cancer (at least three first or second degree relatives with breast or ovarian cancer, including the proband) and 337 with two affected first degree relatives (including the proband). *BRCA1* and *BRCA2* mutations had been screened from the familial patient series as previously described [19-21].

The clinicopathological data of the patients, including tumour diameter, nodal status, distant metastasis and oestrogen receptor (ER) and progesterone receptor (PR) status, were collected from the pathology reports. In addition, all the tumours were re-evaluated for histological diagnosis and graded by a pathologist (P.H.). The grading was performed using the Scarff-Bloom-Richardson method, modified by Elston and Ellis [22]. The tissue microarray sections were immunohistochemically stained for p53 expression [23] and considered as positive when 20% or more of the cancer cells were positively stained. HER2 protein expression was analyzed by immunohistochemical staining (samples with <10% of the cancer cells stained were considered as negative and >90% as positive) and gene amplification by using chromogenic in situ hybridization (6 or more copies in a single nucleus was considered as positive result) [24]. The HER2 gene amplification was considered as the primary result and when not available, the result from the immunohistochemical staining was used. The expression of Ki-67 was evaluated using Ki-67 antibody and the tumour was considered as strong positive if 30% or more of the cancer cells were stained, moderate if 20-29% stained, weak positive if 5-19% stained and negative if less than 5% of the cancer cells were stained [25]. Cyclin D1 and cyclin E expressions were detected by immunohistochemical staining and evaluated as described by Aaltonen et al. [24]. The tumours expressing cyclins above mean

expression of all tumours (9.1% for cyclin D1 and 6.8% for cyclin E) were considered as high-expressing and those below mean as low-expressing tumours.

The tumours were further divided into subgroups by hormone receptor and HER2 status with 1) positive ER or PR status and no HER2 over-expression, 2) positive ER or PR status and HER2 overexpression, 3) HER2-positive, ER- and PR-negative tumours and 4) triple-negative tumours (for ER, PR and HER2 status) as previously described [26, 27]. The patient and tumour characteristics are shown in Supplementary Table 1.

A total of 1356 invasive tumours were scored for glycodeclin staining. The median age of the patients at diagnosis was 54 years (range 22-96 years). For survival analysis, the 10-year breast cancer-specific survival was defined as the time from the date of diagnosis to the date of death due to breast cancer and for the 5-year metastasis-free survival as the time from the date of diagnosis to the date of detection of distant metastasis. Information on death due to breast cancer was obtained from the Finnish Cancer Registry and on distant metastasis from the patient records. The median follow-up time was 109 months (range 1-540 months). Of all the 1098 patients in the survival analysis, 179 died from breast cancer and 193 relapsed with distant metastases during the follow-up time of 10 or 5 years, respectively.

This research was approved by the Ethics Committee of the Helsinki University Central Hospital and the Ministry of Social Affairs and Health in Finland. The study was performed according to reporting recommendations for tumour marker prognostic study (REMARK) [28].

### *Immunohistochemical analysis of glycodelin on tissue microarrays*

Tissue microarrays were constructed as previously described [23, 29]. Briefly, haematoxylin and eosin stained sections of the paraffin-embedded tumours were screened and the morphologically most representative sites of the tumours were selected by a pathologist (P.H.). Four tissue cores (0.6 mm in diameter) from these regions of each donor block were transferred into the recipient block. Five  $\mu\text{m}$  thick sections were cut from the tissue microarray blocks and placed onto glass slides. For immunohistochemical staining, the tissue microarray sections were deparaffinised. To enhance the antigen retrieval the sections were incubated in citrate buffer (pH 6.0) in a microwave oven for 10 minutes. Rabbit (#1) anti-glycodelin IgG [12] was used as the primary antibody (3.0  $\mu\text{g}/\text{ml}$ ) and the IgG fraction from proimmune serum of the same rabbit (3.0  $\mu\text{g}/\text{ml}$ ) was used as a negative control. Biotinylated swine anti-rabbit immunoglobulin E0353 (DAKOCytomation, Glostrup, Denmark) was used as the secondary antibody. Inactivation of the endogenous peroxidase was carried out using 2%  $\text{H}_2\text{O}_2$  in methanol. ABC complex (Vectastain<sup>®</sup> ABC kit, Vectorlaboratories Inc., Burlingame, CA) detection system and AEC+ high sensitivity chromogen (DAKO, Carpinteria, CA) were used for antigen detection. Counterstaining was performed with Mayer's haematoxylin (Reagen, Kuopio, Finland). Glycodelin staining was scored either negative or positive by two observers (L.H. and R.K.).

### *Statistical analyses*

The data was analyzed using SPSS for Windows v15.0 (SPSS Inc., Chicago, IL). The correlation of the glycodelin expression with different variables was done using Pearson's chi-square ( $\chi^2$ ) test. When one or more groups had less than five cases, the Fisher's exact test was used. Kaplan-Mayer survival analysis and log-rank test were used to estimate the effect of glycodelin expression on 10-year breast cancer-specific and 5-year metastasis-free survival



rates of different patient series. Cox's proportional hazards regression model (backward stepwise, 95% CI) was used to estimate the association of glycodeilin expression with survival and also to multivariate comparison of hazard ratios after adjusting for other prognostic factors (tumour diameter, nodal status, distant metastases, grade, and expression of oestrogen and progesterone receptors, Ki-67, p53, HER2). Glycodeilin was considered as a categorical variable. To account for the multiple clinicopathological variables tested,  $P$ -values  $< 0.01$  were considered significant. All  $P$ -values are two-sided.

#### *Microarray data analysis*

Total RNA was extracted from 64 familial non-*BRCA1/2* and 106 sporadic primary breast cancer tumours (D. Greco et al, in preparation). The samples were labelled and hybridized to HumanHT-12 v3 Expression BeadChips (Illumina, San Diego, CA), containing 24660 Entrez Gene entities, according to the instructions provided by the manufacturer (<http://www.illumina.com>). After hybridization, the arrays were washed and scanned as recommended by the manufacturer.

Microarray raw data were imported into R (<http://cran.r-project.org>) and processed by the methods included in the BioConductor facilities [30]. Briefly, after quality control [31], the data was normalized using the quantile method [32] and the gene expression matrix was obtained by averaging the probes mapping to the same Entrez Gene IDs [33]. The Pearson's correlation of each gene expression profile with the expression of glycodeilin (*PAEP*) was calculated and the nominal  $P$ -values were estimated for the familial and the sporadic breast tumours separately. The genes with correlation  $P$ -value  $< 0.01$  were further analyzed. Functional annotation was performed on the *PAEP*-correlated gene lists using the DAVID annotation tools [34]. The categories with Fisher's exact test  $P$ -value  $< 0.05$  were considered

to be significantly enriched.

## Results

Altogether, 1222 samples were scored for glycodeilin staining. The remaining 10% ( $N = 134$ ) of the samples were not analyzed due to either unrepresentative or missing tissue samples. Of the analyzed samples, 399 were from the sporadic series, 746 samples were from familial non-*BRCA1/2* patients, of which 436 had strong family history and 310 had two affected first degree relatives, and 38 were *BRCA1* and 39 *BRCA2* mutation positive. Of all analyzed samples, 26% were positive for glycodeilin expression. In sporadic and familial non-*BRCA1/2* series with either two affected first degree relatives or strong family history the glycodeilin staining was positive in 22%, 34% and 25% of the cases, respectively. The numbers of cases included in different analyses are shown in Supplementary Table 2.

In the sporadic series, the glycodeilin expression was significantly associated with low proliferation rate ( $P < 0.001$ ), with also a tendency towards low grade (grades 1 and 2) tumours ( $P = 0.012$ ; Table 1) and high cyclin D1 ( $P = 0.034$ ) expression. The expression of glycodeilin was not significantly associated either with 10-year breast cancer-specific [cumulative survival (CS) = 80.6% vs. 80.6% for cases with glycodeilin-positive vs. -negative tumours,  $P = 0.990$ ] or 5-year metastasis-free survival of the patients (CS = 77.1% vs. 81.1% for cases with glycodeilin-positive vs. -negative tumours,  $P = 0.478$ ; Fig. 1a and b). Similarly, glycodeilin expression had no significant effect on survival when the analysis was stratified by different systemic adjuvant therapy (Supplementary Table 3).

In the familial non-*BRCA1/2* patient series with strong family history, the expression of glycodeclin was associated with positive nodal status ( $P = 0.003$ ) and positive HER2 status ( $P = 0.009$ ; Table 1). Moreover, the expression of glycodeclin was associated with the subgroup of HER2-positive, ER- and PR-negative tumours ( $P = 0.009$ ). The expression of glycodeclin was also associated with poor 10-year breast cancer-specific and 5-year metastasis-free survival rates of the patients (CS = 76.8% vs. 86.1% for cases with glycodeclin-positive vs. -negative tumours,  $P = 0.030$  and CS = 73.0% vs. 86.3%,  $P = 0.001$ , respectively; Fig 1c and d). The survival effect by glycodeclin expression was not influenced by different systemic adjuvant therapy (Supplementary Table 3).

In univariate analysis performed with Cox's regression model, glycodeclin expression was associated with increased risk for death due to breast cancer (HR = 1.78, 95% CI = 1.05-3.02,  $P = 0.032$ ) and distant metastasis (HR = 2.24, 95% CI = 1.37-3.66,  $P = 0.001$ ) (Table 2a). Furthermore, the multivariate analysis on 5-year metastasis-free survival showed that glycodeclin was an independent predictor of distant metastasis (HR = 2.22, 95% CI = 1.22-4.03,  $P = 0.009$ ). In addition to glycodeclin, large tumour diameter, positive nodal status, negative ER status and high Ki-67 expression were independent prognostic factors in the analysis in the familial non-*BRCA1/2* breast cancer patient series (Table 2b).

In the tumours from familial non-*BRCA1/2* patients having only two affected first degree relatives, glycodeclin expression was marginally associated with positive nodal status ( $P = 0.018$ ) as was seen also in the familial patient series with strong family history. Glycodeclin expression did not significantly associate with either the 10-year breast cancer-specific or the 5-year metastasis-free survival rates of the patients (CS = 83.7%, 95% CI = 0.74-0.86 vs. CS = 80.3%, 95% CI = 0.74-0.93 for cases with glycodeclin-positive vs. -negative tumours,  $P =$

0.454 and CS = 85.3%, 95% CI = 0.75-0.87 vs. CS = 80.9%, 95% CI = 0.78-0.92,  $P = 0.378$ , respectively). No significant association of glycodelin expression with any clinicopathological features in patient series with *BRCA1* or *BRCA2* mutation ( $N = 38$  and  $N = 39$ , respectively) was detected (data not shown).

In the gene expression analysis, a total of 267 and 109 gene expression profiles were found to be correlated with the *PAEP* expression with  $P$ -value  $< 0.01$  in sporadic and familial non-*BRCA1/2* tumours, respectively. Of these, 3 genes were found in both groups (*PAEP*, *OR4F15* and *ATP6V1B1*; Supplementary Table 4 and Supplementary Table 5). The genes whose expression correlated with *PAEP* expression were enriched in 13 and 49 biological themes (i.e. terms, Table 3) in familial non-*BRCA1/2* and sporadic tumours, respectively. None of these themes were present in both sporadic and familial non-*BRCA1/2* groups.

## Discussion

To date, this is the largest study evaluating the association of glycodelin expression with tumour phenotype and outcome in breast cancer patients, as well as the first to specifically investigate the glycodelin expression on familial non-*BRCA1/2* breast tumours. We found that, in sporadic breast cancer patients, glycodelin expression was associated with a non-aggressive tumour phenotype with low proliferation rate and a tendency towards well-differentiated tumours. This is in keeping with previous studies showing that glycodelin is associated with well-differentiated tumours having negative lymph node status [14]. However, in the breast cancers of patients with strong family history glycodelin expression was associated with a less favourable phenotype of positive lymph node and HER2 status. Furthermore, glycodelin expression was also an independent prognostic factor with over

twofold increased risk of distant metastasis for breast cancer patients in *BRCA1/2*-mutation negative breast cancer families.

In the sporadic series, the association of glycodeclin with low proliferation rate and tendency towards well-differentiated tumours is in keeping with our previous pre-clinical mouse model study showing that the glycodeclin-transfected MCF-7 human breast carcinoma cells form significantly smaller tumours and express less Ki-67 than the control tumours and that glycodeclin induces cancer cell differentiation towards non-malignant direction [7]. Furthermore, the expression of glycodeclin in ER-positive MCF-7 breast cancer xenograft tumours decreased cyclin D1 levels [7]. In human breast cancer, we have previously shown that high cyclin D1 expression associates with high proliferation rate in ER-positive, but with low proliferation rate in ER- negative breast cancer [35]. In this study, positive glycodeclin expression associated marginally with high cyclin D1 expression and the association was seen only in ER-positive tumours ( $P = 0.048$ ; data not shown). The number of ER-negative tumours was relatively small, which might make it more difficult to reach statistical significance.

However, in familial non-*BRCA1/2* tumours of the patients with strong family history, glycodeclin expression was associated with a less favourable phenotype, i.e. lymph node positive tumours, and ER- or PR-positive and HER2-positive tumours which phenotype has been shown to associate with worst long term survival in breast cancer [27]. Furthermore, patients with glycodeclin-positive tumours had reduced 5-year metastasis-free survival also when adjusted for other conventional prognostic factors. Interestingly, contrary to breast cancer cells in vitro, in melanoma cells glycodeclin transfection has been found to cause more aggressive phenotype [36]. The association of glycodeclin expression with the less favourable

tumour phenotype and poor survival in familial non-*BRCA1/2* cases with a strong family history is intriguing and suggests that in breast cancer families the tumour phenotype and survival may be affected by genetic factors underlying the genetic predisposition to breast cancer. Such effects may lead to different, still unknown, molecular mechanisms, which regulate glycodefin expression or alter its effects and may lead to differential progression pathway(s) with effect also on survival. This hypothesis is supported by the gene expression analysis which revealed different gene expression profiles correlating with the *PAEP* expression in the sporadic and familial non-*BRCA1/2* breast cancers. Furthermore, these genes were enriched in different biological themes based on their functions. We have previously found also differential expression of cyclin E and D1 in sporadic and familial breast tumours and shown that cyclin E and D1 expression distinguishes familial non-*BRCA1/2* tumours both from sporadic, and *BRCA1*- and *BRCA2*-associated tumours [24]. Tumours from especially *BRCA1* but to some extent also *BRCA2* mutation carriers display distinct tumour characteristics and gene expression or genomic copy number patterns and associate predominantly with the basal or luminal A or B subtypes, respectively. Some studies have also shown phenotypic and genomic characteristics for non-*BRCA1/2* breast tumours although this group of tumours is heterogeneous at large and fall in similar intrinsic subtypes as also the sporadic breast tumours [37-39]. This is consistent with the heterogeneous genetic background in the non-*BRCA1/2* familial breast cancer which may result from joint effects of multiple low penetrance alleles or a large number of rare mutations with moderate or even high penetrance [40, 41]. However, the results derived from defined breast cancer families, like the different gene expression profiles correlating with the *PAEP* expression in the sporadic and familial non-*BRCA1/2* breast cancers in the present study may reflect the different predisposition and pathogenesis in the familial breast tumours, suggesting also

distinct genetic background and progression pathways in familial, non-*BRCA1/2* breast cancer.

In conclusion, we have shown that glycodeclin expression is associated with a less aggressive tumour phenotype in sporadic breast cancer patients, while in familial non-*BRCA1/2* breast cancer patients glycodeclin associated with a less favourable tumour phenotype and was an independent predictor of distant metastasis for breast cancer patients in *BRCA1/2*-mutation negative breast cancer families. Gene expression analysis further revealed different gene expression profiles correlating with the *PAEP* expression in the sporadic and familial non-*BRCA1/2* breast cancers. These findings suggest differential progression pathways in the sporadic and familial non-*BRCA1/2* breast tumours expressing glycodeclin and provide insight into the genes and biological themes associated with *PAEP* expression in the sporadic and familial breast cancer.

### **Conflict of interest**

The author declare no conflict of interest.

### **Acknowledgements**

We thank Drs. Kirsimari Aaltonen and Karl von Smitten and RN Hanna Jäntti for their help with the patient data and specimens. The Finnish Cancer Registry is gratefully acknowledged for the cancer data. This work was supported by Helsinki University Central Hospital Research Fund, Academy of Finland (132473, 135937), the Finnish Cancer Society, the Sigrid Juselius Foundation and Magnus Ehrnrooth Foundation.

## References

1. Seppälä M, Taylor RN, Koistinen H et al (2002) Glycodelin: a major lipocalin protein of the reproductive axis with diverse actions in cell recognition and differentiation. *Endocr Rev* 23:401-430
2. Morris HR, Dell A, Easton RL et al (1996) Gender-specific glycosylation of human glycodelin affects its contraceptive activity. *J Biol Chem* 271:32159-32167
3. Yeung WS, Lee KF, Koistinen R et al (2009) Effects of glycodelins on functional competence of spermatozoa. *J Reprod Immunol* 83:26-30
4. Alok A, Karande AA (2009) The role of glycodelin as an immune-modulating agent at the feto-maternal interface. *J Reprod Immunol* 83:124-127
5. Seppälä M, Koistinen H, Koistinen R et al (2009) Glycodelin in reproductive endocrinology and hormone-related cancer. *Eur J Endocrinol* 160:121-133
6. Kamäräinen M, Seppälä M, Virtanen I et al (1997) Expression of glycodelin in MCF-7 breast cancer cells induces differentiation into organized acinar epithelium. *Lab Invest* 77:565-573
7. Hautala LC, Koistinen R, Seppälä M et al (2008) Glycodelin reduces breast cancer xenograft growth in vivo. *Int J Cancer* 123:2279-2284
8. Koistinen H, Seppälä M, Nagy B et al (2005) Glycodelin reduces carcinoma-associated gene expression in endometrial adenocarcinoma cells. *Am J Obstet Gynecol* 193:1955-1960
9. Uchida H, Maruyama T, Nagashima T et al (2005) Histone deacetylase inhibitors induce differentiation of human endometrial adenocarcinoma cells through up-regulation of glycodelin. *Endocrinology* 146:5365-5373



10. Ohta K, Maruyama T, Uchida H et al (2008) Glycodelin blocks progression to S phase and inhibits cell growth: a possible progesterone-induced regulator for endometrial epithelial cell growth Glycodelin inhibits G1/S progression. *Mol Hum Reprod* 14:17-22
11. Koistinen H, Hautala LC, Seppälä M et al (2009) The role of glycodelin in cell differentiation and tumor growth. *Scand J Clin Lab Invest* 69:452-459
12. Mandelin E, Lassus H, Seppälä M et al (2003) Glycodelin in ovarian serous carcinoma: association with differentiation and survival. *Cancer Res* 63:6258-6264
13. Kämäräinen M, Halttunen M, Koistinen R et al (1999) Expression of glycodelin in human breast and breast cancer. *Int J Cancer* 83:738-742.
14. Jeschke U, Mylonas I, Kunert-Keil C et al (2005) Expression of glycodelin protein and mRNA in human ductal breast cancer carcinoma in situ, invasive ductal carcinomas, their lymph node and distant metastases, and ductal carcinomas with recurrence. *Oncol Rep* 13:413-419
15. Scholz C, Toth B, Barthell E et al (2009) Immunohistochemical expression of glycodelin in breast cancer correlates with estrogen-receptor alpha and progesterone-receptor A positivity. *Histol Histopathol* 24:467-471
16. Syrjäkoski K, Vahteristo P, Eerola H et al (2000) Population-based study of BRCA1 and BRCA2 mutations in 1035 unselected Finnish breast cancer patients. *J Natl Cancer Inst* 92:1529-1531
17. Kilpivaara O, Bartkova J, Eerola H et al (2005) Correlation of CHEK2 protein expression and c.1100delC mutation status with tumor characteristics among unselected breast cancer patients. *Int J Cancer* 113:575-580
18. Eerola H, Blomqvist C, Pukkala E et al (2000) Familial breast cancer in southern Finland: how prevalent are breast cancer families and can we trust the family history reported by patients? *Eur J Cancer* 36:1143-1148

19. Vehmanen P, Friedman LS, Eerola H et al (1997) Low proportion of BRCA1 and BRCA2 mutations in Finnish breast cancer families: evidence for additional susceptibility genes. *Hum Mol Genet* 6:2309-2315
20. Vahteristo P, Eerola H, Tamminen A et al (2001) A probability model for predicting BRCA1 and BRCA2 mutations in breast and breast-ovarian cancer families. *Br J Cancer* 84:704-708
21. Vahteristo P, Bartkova J, Eerola H et al (2002) A CHEK2 genetic variant contributing to a substantial fraction of familial breast cancer. *Am J Hum Genet* 71:432-438
22. Elston CW, Ellis IO (1991) Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 19:403-410
23. Tommiska J, Eerola H, Heinonen M et al (2005) Breast cancer patients with p53 Pro72 homozygous genotype have a poorer survival. *Clin Cancer Res* 11:5098-5103
24. Aaltonen K, Blomqvist C, Amini RM et al (2008) Familial breast cancers without mutations in BRCA1 or BRCA2 have low cyclin E and high cyclin D1 in contrast to cancers in BRCA mutation carriers. *Clin Cancer Res* 14:1976-1983
25. Ahlin C, Aaltonen K, Amini RM et al (2007) Ki67 and cyclin A as prognostic factors in early breast cancer. What are the optimal cut-off values? *Histopathology* 51:491-498
26. Heikkinen T, Kärkkäinen H, Aaltonen K et al (2009) The breast cancer susceptibility mutation PALB2 1592delT is associated with an aggressive tumor phenotype. *Clin Cancer Res* 15:3214-3222
27. Blows FM, Driver KE, Schmidt MK et al (2010) Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. *PLoS Med* (in press)

28. McShane LM, Altman DG, Sauerbrei W et al (2005) Reporting recommendations for tumor marker prognostic studies. *J Clin Oncol* 23:9067-9072
29. Eerola H, Heikkilä P, Tamminen A et al (2005) Relationship of patients' age to histopathological features of breast tumours in BRCA1 and BRCA2 and mutation-negative breast cancer families. *Breast Cancer Res* 7:R465-469
30. Gentleman RC, Carey VJ, Bates DM et al (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5:R80
31. Du P, Kibbe WA, Lin SM (2008) lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 24:1547-1548
32. Bolstad BM, Irizarry RA, Astrand M et al (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19:185-193
33. Tatusova T (2010) Genomic databases and resources at the national center for biotechnology information. *Methods Mol Biol* 609:17-44
34. Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4:44-57
35. Aaltonen K, Amini RM, Landberg G et al (2009) Cyclin D1 expression is associated with poor prognostic features in estrogen receptor positive breast cancer. *Breast Cancer Res Treat* 113:75-82
36. Ren S, Liu S, Howell PM, Jr et al (2009) Functional characterization of the progesterone-associated endometrial protein gene in human melanoma. *J Cell Mol Med*. doi: 10.1111/j.1582-4934.2009.00922.x
37. Palacios J, Robles-Frias MJ, Castilla MA et al (2008) The molecular pathology of hereditary breast cancer. *Pathobiology* 75:85-94

38. Melchor L, Benitez J (2008) An integrative hypothesis about the origin and development of sporadic and familial breast cancer subtypes. *Carcinogenesis* 29:1475-1482
39. Waddell N, Arnold J, Cocciardi S et al (2009) Subtypes of familial breast tumours revealed by expression and copy number profiling. *Breast Cancer Res Treat.* doi: 10.1007/s10549-009-0653-1
40. Fletcher O, Houlston RS (2010) Architecture of inherited susceptibility to common cancer. *Nat Rev Cancer* 10:353-361
41. Meindl A, Hellebrand H, Wiek C et al (2010) Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. *Nat Genet* 42:410-414

**Table 1** Association of glycodeclin expression with the clinicopathological features in the tumours of sporadic and familial non-*BRCA1/2* patient series

Category	Sporadic series					Familial non- <i>BRCA1/2</i> series				
	Gd neg <i>N</i> (%)	Gd pos <i>N</i> (%)	<i>P</i>	OR	(95% CI)	Gd neg <i>N</i> (%)	Gd pos <i>N</i> (%)	<i>P</i>	OR	(95% CI)
<b>T</b>										
1	167 (54.6)	56 (62.9)	0.162 1 vs. 2, 3, and 4	0.71	(0.44-1.15)	209 (65.3)	64 (59.3)	0.258 1 vs. 2, 3, and 4	1.29	(0.83-2.03)
2	121 (39.5)	27 (30.3)				93 (29.1)	37 (34.3)			
3	6 (2.0)	3 (3.4)				8 (2.5)	5 (4.6)			
4	12 (3.9)	3 (3.4)				10 (3.1)	2 (1.9)			
<b>N</b>										
neg	160 (52.3)	45 (50.6)	0.774	1.07	(0.67-1.72)	203 (63.4)	50 (47.2)	0.003	1.94	(1.25-3.03)
pos	146 (47.7)	44 (49.4)				117 (36.6)	56 (52.8)			
<b>M</b>										
neg	299 (96.5)	85 (95.5)	0.752	1.28	(0.40-4.12)	315 (98.4)	106 (98.1)	1.000	1.19	(0.23-6.22)
pos	11 (3.5)	4 (4.5)				5 (1.6)	2 (1.9)			
<b>ER</b>										
neg	59 (19.4)	12 (13.6)	0.216	0.66	(0.33-1.28)	71 (23.7)	18 (17.6)	0.201	0.69	(0.39-1.22)
pos	245 (80.6)	76 (86.4)				228 (76.3)	84 (82.4)			
<b>PR</b>										
neg	99 (32.5)	22 (25.0)	0.182	0.69	(0.40-1.19)	98 (32.9)	37 (37.0)	0.452	1.20	(0.75-1.92)
pos	206 (67.5)	66 (75.0)				200 (67.1)	63 (63.0)			
<b>Grade</b>										
1	70 (22.8)	28 (31.8)	0.012 1 and 2 vs. 3	0.48	(0.27-0.86)	75 (23.7)	27 (25.2)	0.689 1 and 2 vs. 3	0.90	(0.55-1.49)
2	135 (44.0)	43 (48.9)				155 (49.1)	53 (49.5)			
3	102 (33.2)	17 (19.3)				86 (27.2)	27 (25.2)			
<b>Histology</b>										
Ductal	216 (69.7)	67 (75.3)	0.305 Ductal vs. non-ductal 0.314 Lobular vs. non-lobular	1.33	(0.77-2.27)	230 (70.3)	79 (72.5)	0.670 Ductal vs. non-ductal 0.677 Lobular vs. non-lobular	1.11	(0.69-1.80)
Lobular	52 (16.8)	11 (12.4)				66 (20.2)	20 (18.3)			
Medullar	2 (0.6)	0 (0.0)				7 (2.1)	0 (0.0)			
Other	40 (12.9)	11 (12.4)				24 (7.3)	10 (9.2)			
<b>p53</b>										
neg	231 (80.2)	71 (86.6)	0.188	0.63	(0.31-1.26)	252 (80.0)	84 (77.8)	0.622	1.14	(0.67-1.94)
pos	57 (19.8)	11 (13.4)				63 (20.0)	24 (22.2)			
<b>HER2</b>										
neg	261 (87.6)	74 (85.1)	0.537	1.24	(0.63-2.45)	284 (91.0)	84 (81.6)	0.009	2.29	(1.22-4.31)
pos	37 (12.4)	13 (14.9)				28 (9.0)	19 (18.4)			
<b>Ki-67</b>										
0 (<5%)	57 (18.9)	23 (25.8)	<0.001 0 and 1 vs. 2 and 3	0.37	(0.21-0.64)	96 (29.8)	33 (30.6)	0.102 0 and 1 vs. 2 and 3	1.48	(0.92-2.36)
1 (5-19%)	123 (40.3)	48 (53.9)				142 (44.1)	38 (35.2)			
2 (20-29%)	64 (21.0)	11 (12.4)				41 (12.7)	22 (20.4)			
3 (≥30%)	61 (20.0)	7 (7.9)				43 (13.4)	15 (13.9)			
<b>Type</b>										
ER/PR pos, HER2 neg	218 (74.7)	68 (79.1)	0.946 ER/PR pos, HER2 pos vs. other	0.97	(0.38-2.48)	210 (73.9)	70 (72.2)	0.009 ER/PR pos, HER2 pos vs. other	2.65	(1.25-5.60)
ER/PR pos, HER2 pos	21 (7.2)	6 (7.0)				17 (6.0)	14 (14.4)			
ER/PR neg, HER2 pos	16 (5.5)	7 (8.1)	0.075 Triple-neg vs. other	0.43	(0.16-1.12)	10 (3.5)	5 (5.2)	0.044 Triple-neg vs. other	0.45	(0.21-1.00)
Triple-neg	37 (12.7)	5 (5.8)				47 (16.5)	8 (8.2)			
<b>Cyclin E</b>										
high	63 (21.5)	14 (16.9)	0.356 0.74	(0.39-1.40)	34 (12.1)	10 (11.8)	0.934 0.97	(0.46-2.05)		
low	230 (78.5)	69 (83.1)			247 (87.9)	75 (88.2)				
<b>Cyclin D1</b>										
high	141 (48.0)	53 (60.9)	0.034 0.59	(0.36-0.96)	153 (54.3)	50 (60.2)	0.335 0.78	(0.48-1.29)		
low	153 (52.0)	34 (39.1)			129 (45.7)	33 (39.8)				

T, tumour size; N, nodal status; M, primary metastasis; Gd, glycodeclin; OR, odds ratio.

**Table 2a** Univariate analysis (Cox's proportional hazards model) of prognostic factors and glycodeclin expression in the tumours of familial non-*BRCA1/2* patient series

Variables <sup>a</sup>	10-year breast cancer-specific survival			Variables	5-year metastasis-free survival		
	<i>P</i>	HR	(95% CI)		<i>P</i>	HR	(95% CI)
T				T			
1 reference				1 reference			
2	< 0.001	2.93	(1.74-4.94)	2	< 0.001	2.78	(1.68-4.63)
3	< 0.001	6.02	(2.47-14.65)	3	< 0.001	4.91	(2.03-11.87)
4	< 0.001	16.26	(7.60-34.76)	4	< 0.001	11.00	(5.01-24.16)
N				N			
positive	< 0.001	5.16	(3.02-8.82)	positive	< 0.001	3.63	(2.22-5.94)
M				M			
positive	< 0.001	17.10	(7.78-37.59)	positive	NA	NA	NA
ER				ER			
negative	< 0.001	2.56	(1.58-4.15)	negative	< 0.001	2.49	(1.54-4.01)
PR				PR			
negative	0.010	1.86	(1.16-2.97)	negative	0.011	1.83	(1.15-2.91)
Grade				Grade			
1 reference				1 reference			
2	0.277	1.52	(0.71-3.25)	2	0.005	4.36	(1.55-12.31)
3	< 0.001	4.54	(2.19-9.42)	3	< 0.001	9.32	(3.30-26.37)
p53				p53			
positive	0.059	1.65	(0.98-2.77)	positive	0.175	1.45	(0.85-2.46)
HER2				HER2			
positive	< 0.001	3.02	(1.72-5.30)	positive	0.003	2.40	(1.36-4.25)
Ki-67				Ki-67			
0 reference				0 reference			
1	0.357	1.34	(0.72-2.51)	1	0.119	1.68	(0.88-3.24)
2	0.083	1.96	(0.92-4.18)	2	0.020	2.46	(1.16-5.23)
3	0.002	3.03	(1.51-6.07)	3	< 0.001	3.82	(1.89-7.74)
Gd				Gd			
positive	0.032	1.78	(1.05-3.02)	positive	0.001	2.24	(1.37-3.66)

<sup>a</sup> The number of tumours included in each category is shown in Supplementary Table 1

**b** Multivariate analysis (Cox's proportional hazards model, final step) of prognostic factors in the tumors of familial non-*BRCA1/2* patient series

Variables	10-year breast cancer-specific survival (N = 354)			Variables	5-year metastasis-free survival (N = 351)		
	<i>P</i>	HR	(95% CI)		<i>P</i>	HR	(95% CI)
T				T			
1 reference				1 reference			
2	0.074	1.83	(0.94-3.53)	2	0.068	1.88	(0.95-3.71)
3	0.053	3.57	(0.98-12.99)	3	0.063	2.97	(0.94-9.38)
4	< 0.001	18.27	(6.77-49.31)	4	< 0.001	24.02	(8.69-66.39)
N				N			
positive	0.001	3.45	(1.69-7.02)	positive	< 0.001	3.55	(1.77-7.12)
M				ER			
positive	< 0.001	9.20	(3.28-25.84)	negative	0.006	2.45	(1.29-4.68)
ER				Ki-67			
negative	0.003	2.54	(1.37-4.70)	0 reference			
HER2				1	0.098	2.32	(0.86-6.28)
positive	0.044	1.99	(1.02-3.88)	2	0.006	4.37	(1.51-12.66)
				3	0.024	3.39	(1.18-9.75)
				Gd			
				positive	0.009	2.22	(1.22-4.03)

T, tumour size; N, nodal status; M, primary metastasis; Gd, glycodeclin.

**Table 3** Functional terms over-represented among the *PAEP*-correlated genesFamilial non-*BRCA1/2* tumours

Category	Term	Count <sup>a</sup>	P
GOTERM_BP_ALL	GO:0050877~neurological system process	11	0.009
INTERPRO	IPR002290:Serine/threonine protein kinase	5	0.012
INTERPRO	IPR000961:Protein kinase, C-terminal	3	0.017
SP_PIR_KEYWORDS	sensory transduction	7	0.017
SMART	SM00133:S_TK_X	3	0.019
UP_SEQ_FEATURE	domain:AGC-kinase C-terminal	3	0.020
PIR_SUPERFAMILY	PIRSF000613:RAC serine/threonine-protein kinase	2	0.021
GOTERM_BP_ALL	GO:0007606~sensory perception of chemical stimulus	6	0.032
GOTERM_BP_ALL	GO:0007600~sensory perception	8	0.032
GOTERM_BP_ALL	GO:0003008~system process	11	0.033
INTERPRO	IPR008271:Serine/threonine protein kinase, active site	5	0.034
SP_PIR_KEYWORDS	receptor	12	0.038
SP_PIR_KEYWORDS	Serine/threonine-protein kinase	5	0.046

## Sporadic tumours

Category	Term	Count <sup>a</sup>	P
SP_PIR_KEYWORDS	ubl conjugation	12	0.002
KEGG_PATHWAY	hsa04080:Neuroactive ligand-receptor interaction	8	0.004
GOTERM_CC_ALL	GO:0000785~chromatin	8	0.005
GOTERM_CC_ALL	GO:0005624~membrane fraction	17	0.006
INTERPRO	IPR007125:Histone core	4	0.007
GOTERM_CC_ALL	GO:0000786~nucleosome	5	0.007
GOTERM_BP_ALL	GO:0015672~monovalent inorganic cation transport	10	0.008
SP_PIR_KEYWORDS	Chromosomal protein	6	0.008
SP_PIR_KEYWORDS	Nucleosome core	4	0.010
GOTERM_BP_ALL	GO:0006334~nucleosome assembly	5	0.012
SP_PIR_KEYWORDS	membrane	69	0.014
GOTERM_MF_ALL	GO:0031420~alkali metal ion binding	7	0.015
SP_PIR_KEYWORDS	ion transport	12	0.016
GOTERM_MF_ALL	GO:0004725~protein tyrosine phosphatase activity	5	0.018
SP_PIR_KEYWORDS	DNA binding	9	0.018
GOTERM_BP_ALL	GO:0031497~chromatin assembly	5	0.019
GOTERM_CC_ALL	GO:0044427~chromosomal part	9	0.022
INTERPRO	IPR013099:Ion transport 2	3	0.026
SP_PIR_KEYWORDS	potassium transport	5	0.026
GOTERM_MF_ALL	GO:0004553~hydrolase activity, hydrolyzing O-glycosyl compounds	5	0.027
GOTERM_CC_ALL	GO:0000267~cell fraction	18	0.028
GOTERM_BP_ALL	GO:0007187~G-protein signaling, coupled to cyclic nucleotide second messenger	5	0.030
GOTERM_BP_ALL	GO:0045817~positive regul. of global transcription from RNA polymerase II promoter	2	0.030
SP_PIR_KEYWORDS	acetyly sine	2	0.030
GOTERM_MF_ALL	GO:0030955~potassium ion binding	5	0.031
GOTERM_MF_ALL	GO:0043167~ion binding	54	0.031
SP_PIR_KEYWORDS	transmembrane	58	0.032
SP_PIR_KEYWORDS	potassium	5	0.035
GOTERM_MF_ALL	GO:0046873~metal ion transmembrane transporter activity	8	0.035
SP_PIR_KEYWORDS	transport	23	0.036
GOTERM_BP_ALL	GO:0019935~cyclic-nucleotide-mediated signaling	5	0.036
GOTERM_BP_ALL	GO:0006325~establishment and/or maintenance of chromatin architecture	8	0.038
GOTERM_BP_ALL	GO:0006358~regulation of global transcription from RNA polymerase II promoter	2	0.039
GOTERM_CC_ALL	GO:0005789~endoplasmic reticulum membrane	10	0.041
UP_SEQ_FEATURE	topological domain:Cytoplasmic	38	0.041
GOTERM_MF_ALL	GO:0004871~signal transducer activity	34	0.041
GOTERM_MF_ALL	GO:0060089~molecular transducer activity	34	0.041

GOTERM_BP_ALL	GO:0006323~DNA packaging	8	0.042
GOTERM_MF_ALL	GO:0005261~cation channel activity	7	0.043
SP_PIR_KEYWORDS	Ionic channel	8	0.044
GOTERM_MF_ALL	GO:0043169~cation binding	49	0.045
GOTERM_CC_ALL	GO:0044425~membrane part	76	0.045
GOTERM_CC_ALL	GO:0042175~nuclear envelope-endoplasmic reticulum network	10	0.046
GOTERM_BP_ALL	GO:0019953~sexual reproduction	8	0.046
GOTERM_MF_ALL	GO:0019901~protein kinase binding	4	0.048
SP_PIR_KEYWORDS	calcium transport	4	0.049
GOTERM_MF_ALL	GO:0019842~vitamin binding	5	0.049
GOTERM_MF_ALL	GO:0016798~hydrolase activity, acting on glycosyl bonds	5	0.049
GOTERM_BP_ALL	GO:0006812~cation transport	11	0.050

---

Category, annotation database; Term, functional term; Count, number of *PAEP*-correlated genes annotated in the term.

<sup>a</sup>The same genes may be present in several categories.



## Legend to figure

**Fig. 1** Kaplan-Meier cumulative survival curves of **(a, b)** sporadic and **(c, d)** familial non-*BRCA1/2* breast cancer patients series for **(a, c)** 10-year breast cancer-specific and **(b, d)** for 5-year metastasis-free survival. Gd, glycodeclin

