Journal of Pathology
J Pathol 2004; **203**: 661–671
Published online 14 April 2004 in Wiley InterScience (www.interscience.wiley.com). **DOI:** 10.1002/path.1559

Original Paper

Expression of luminal and basal cytokeratins in human breast carcinoma

Dalia M Abd El-Rehim,[1] Sarah E Pinder,[1] Claire E Paish,[1] J Bell,[1] RW Blamey,[1] John FR Robertson,[2] Robert I Nicholson[3] and Ian O Ellis[1]*

[1]*Department of Histopathology, Breast Unit, Nottingham City Hospital NHS Trust and University of Nottingham, UK*
[2]*Department of Surgery, Breast Unit, Nottingham City Hospital NHS Trust and University of Nottingham, UK*
[3]*Tenovus Institute, Cardiff, UK*

**Correspondence to:*
Dr Ian O Ellis, Department of Histopathology, Nottingham City Hospital NHS Trust, Hucknall Road, Nottingham NG5 1PB, UK.
E-mail:
ian.ellis@nottingham.ac.uk

Received: 30 September 2003
Revised: 16 January 2004
Accepted: 26 January 2004

Abstract

We have examined basal and luminal cell cytokeratin expression in 1944 cases of invasive breast carcinoma, using tissue microarray (TMA) technology, to determine the frequency of expression of each cytokeratin subtype, their relationships and prognostic relevance, if any. Expression was determined by immunocytochemistry staining using antibodies to the luminal cytokeratins (CKs) 7/8, 18 and 19 and the basal markers CK 5/6 and CK 14. Additionally, assessment of α-smooth muscle actin (SMA) and oestrogen receptor status (ER) was performed. The vast majority of the cases showed positivity for CK 7/8, 18 and 19 indicating a differentiated glandular phenotype, a finding associated with good prognosis, ER positivity and older patient age. In contrast, basal marker expression was significantly related to poor prognosis, ER negativity and younger patient age. Multivariate analysis showed that CK 5/6 was an independent indicator for relapse free interval. We were able to subgroup the cases into four distinct phenotype categories (pure luminal, mixed luminal/basal, pure basal and null), which had significant differences in relation to the biological features and the clinical course of the disease. Tumours classified as expressing a basal phenotype (the combined luminal plus basal and the pure basal) were in a poor prognostic subgroup, typically ER negative in most cases. These findings provide further evidence that breast cancer has distinct differentiation subclasses that have both biological and clinical relevance.

Keywords: tissue microarray; immunohistochemistry; cellular phenotype; invasive breast cancer; cytokeratin; basal and luminal epithelium

Introduction

The cytoskeleton in all mammalian cells is composed of three types of filament: actin microfilaments (4–6 nm), intermediate filaments (10 nm) and the microtubules (25 nm) [1]. Expression of the intermediate filament proteins, particularly cytokeratins (CK), reflects the epithelial cell type, state of tissue growth and differentiation in addition to the functional status of the tissue [2]. In normal breast, both the luminal epithelial and the myoepithelial cells exhibit different and distinctive keratin phenotypes. CKs 7, 8, 18 and 19 are expressed in the luminal cells, while smooth muscle actin (SMA) and CKs 5, 14 and 17 are found in the myoepithelial/basal cells [1,3,4].

Within the normal mammary gland, a small number of cells that are positive for CK 5/6 but are negative for CK 8, 18, 19 and SMA have been identified. These cells are located in the luminal compartment and show the morphological features of stem cells that have the capacity to differentiate towards either the glandular (CK 8, 18 and 19 positive) or the basal (SMA positive) phenotype. Intermediate populations retaining CK 5/6 expression and showing positive reactivity with either CK 8, 18 and 19 or SMA have also been identified, indicating the onset of the luminal or the basal differentiation pathways respectively. Therefore, five distinct cell populations have been identified: committed stem cell (CK5+), glandular precursor cell (CK5+, CK8/18/19+), glandular end cell (CK8/18/19+), myoepithelial precursor cell (CK5+ SMA+) and myoepithelial end cell (SMA+) [5].

Recent studies using cDNA microarray technology applied to a relatively small series of breast cancers have identified four major groups of breast cancer. Two of them have gene expression patterns characteristic of the basal and the luminal phenotypes [6]. In a subsequent study, the same groups were distinguished and were found to have a statistically significant association with overall and relapse-free survival [7].

To investigate further the clinical significance of these findings, we have used TMA technology in a large retrospective immunohistochemical evaluation of 1944 cases of invasive breast cancer using

well-characterized, commercially available antibodies for both the luminal [CK 8, 18 and 19] and the basal [CK 5/6, CK 14 and SMA] phenotypes to determine the frequency of these subgroups of breast cancer and to assess their relationships with established patient and tumour variables, and with clinical outcome.

Materials and methods

Patients

A consecutive series of 1944 cases of primary operable invasive breast carcinoma from patients presenting between 1986 and 1998 and entered into the Nottingham Tenovus Primary Breast Carcinoma Series were used. This is a well-characterized series of primary breast cancer with a mean age of 53 (range 18–70 years) treated in a uniform way and has been used to study a wide range of potential prognostic factors and markers in breast cancer including histological grade [8], histological tumour type [9], vascular invasion [10], tumour size, lymph node stage and Nottingham Prognostic Index (NPI) [11–16]. Information on local, regional and distant metastases and survival is maintained on a prospective basis. Patients have been followed up at 3-month intervals initially, then 6-monthly, then annually. At the time of resection all tumours are managed in a standard fashion, incised and fixed immediately. The frequencies of histological types and the tumour grade distribution are presented in Tables 1 and 2. This research was approved by Nottingham Research Ethics Committee 2 under the title of 'Development of a molecular genetic classification of breast cancer'.

Table 1. Frequencies and percentages of histological tumour types

Tumour type	No.	%
Invasive NST	1094	56.5%
Tubular mixed	337	17.4%
Medullary	46	2.4%
Typical	5	
Atypical	41	
Lobular	220	11.4%
Classical	141	
Alveolar	2	
Solid	6	
Tubulo-lobular	6	
Mixed	65	
Tubular	79	4.1%
Mucinous	26	1.3%
Invasive cribriform	10	0.5%
Invasive papillary	7	0.4%
Mixed NST & lobular	65	3.4%
Mixed NST & special type	41	2.1%
Miscellaneous other types	14	0.7%
Adenoid cystic	5	
Metaplastic	3	
Spindle cell tumour	1	
Apocrine carcinoma	1	
NST with clear cell features	1	
NST with secretory features	1	
NST with spindle cell element	1	

NST, no special type.

Table 2. Frequencies and percentages of tumour grades

Grade	No.	%
1	367	18.9%
2	647	33.4%
3	925	47.7%

Construction of the tissue microarray blocks

Breast cancer tissue microarrays were prepared as described previously [17–19]. Each case was sampled twice from the centre and the periphery of the tumour to form an array of 100 cases per block.

Immunohistochemistry

Immunohistochemical staining was performed using the strepavidin–biotin complex method. The sections (except for SMA) were microwaved in citrate buffer, pH 6 for a total 20 min. Luminal differentiation was demonstrated using monoclonal antibodies (MAbs) specific for CK 7/8 (clone CAM 5.2, Becton Dickinson, diluted 1 : 2), CK 18 (clone DC10, Dako, diluted 1 : 50) and CK 19 (clone Bck 108, Dako, diluted 1 : 100), while myoepithelial differentiation was assessed with MAbs against CK 5/6 (clone D5/16134, Boehringer Biochemica, diluted 1 : 100), CK 14 (clone LL002, Novocastra, diluted 1 : 100) and SMA (clone 1A4, Dako, diluted 1 : 20 then 1 : 100 to use).

Controls

Adult breast tissue parenchyma was used as positive control and entrapped normal breast tissues within the cores used as an internal positive control. Negative controls were obtained by omitting the primary antibodies.

Immunohistochemistry scoring

The modified Histo-score (H-score) [20] was used as it includes a semi-quantitative assessment of both the intensity of staining and the percentage of positive cells. The range of possible scores is thus 0–300. H-score and similar semi-quantitative scoring systems have been successfully used for TMA evaluation [21–23]. By using such a score, we were able to explore rationalization of our cases into biologically relevant groups depending on different levels of expression, which could not be obtained on using simpler scoring methods (eg positive vs negative).

Two cores were evaluated from each tumour. Each core was scored individually, then the mean of the two readings was calculated. If one core was uninformative (either lost or contained no tumour tissue), the overall score applied was that of the remaining core. Previous studies have validated the use of one core to study the expression of tumour markers that have a heterogeneous distribution (18,19). One observer scored all cases, which were rechecked randomly by the same

Table 3. Association of luminal marker expression with clinical and pathological parameters

	CK7/8					CK18					CK 19				
Feature	**−**	**W+**	**M+**	**S+**	***p*-value**	**−**	**W+**	**M+**	**S+**	***p*-value**	**−**	**W+**	**M+**	**S+**	***p*-value**
Grade															
1	1	6	44	280		4	23	78	198		10	19	121	181	
2	3	10	126	472		15	66	170	298		21	71	256	262	
3	26	158	262	447	<0.001	172	174	188	303	<0.001	100	215	363	214	<0.001
Total	30	174	432	1199		191	263	436	799		131	305	740	657	
LN															
1	21	122	258	763		126	170	261	505		86	201	452	423	
2	9	26	129	359		43	76	129	242		32	73	229	189	
3	1	26	44	75	<0.001	22	17	44	52	0.048	14	31	57	44	0.120
Total	31	174	431	1197		191	263	434	799		132	305	738	656	
Size															
≤1.5 cm	4	45	138	467		44	82	146	316		32	100	265	255	
>1.5 cm	27	129	294	733	<0.001	147	181	290	484	<0.001	100	205	475	403	0.010
Total	31	174	432	1200		191	263	436	800		132	305	740	658	
NPI															
Good	3	11	102	490		12	58	147	335		21	60	225	299	
Moderate	18	121	246	557		137	158	212	352		77	187	406	270	
Poor	9	42	83	148	<0.001	42	47	75	110	<0.001	33	58	107	85	<0.001
Total	30	174	431	1195		191	263	434	797		131	305	738	654	
ER															
Negative	29	141	178	200		151	129	102	120		86	153	205	102	
Positive	1	27	228	974	<0.001	36	125	326	670	<0.001	42	134	509	545	<0.001
Total	30	168	406	1174		187	254	428	790		128	287	714	647	
VI															
No	24	120	289	833		135	184	279	555		94	211	478	479	
Yes	7	53	138	356	0.625	55	78	150	239		38	93	250	175	0.023
Total	31	173	427	1189		190	262	429	494	0.264	132	304	728	654	
LR															
No	28	155	392	1111		171	240	406	734		120	278	678	608	
Yes	2	15	35	79	0.598	18	17	29	54	0.577	9	22	53	46	0.998
Total	30	170	427	1190		189	257	435	788		129	300	731	654	
RR															
No	27	154	392	1129		168	242	411	748		122	277	686	614	
yes	3	16	35	61	0.032	21	15	24	40	0.017	7	23	45	40	0.761
Total	30	170	427	1190		189	257	435	788		129	300	731	654	
DM															
No	26	138	360	1081		157	227	385	711		110	254	651	587	
Yes	4	32	67	107	<0.001	32	30	49	76	0.043	19	46	80	65	0.063
Total	30	170	427	1188		189	257	434	787		129	300	731	652	
Death															
No	25	139	364	1106		155	227	394	730		110	258	660	603	
Yes	5	31	63	84	<0.001	34	30	41	58	<0.001	19	42	71	51	0.007
Total	30	170	427	1190		189	257	435	788		129	300	731	654	
Age															
≤35	2	16	22	30		21	14	11	22		7	19	34	10	
36–45	6	46	82	189		50	52	78	122		23	65	141	99	
46–55	11	59	129	390		64	76	142	260		52	86	228	222	
>55	12	53	200	593	<0.001	56	121	206	398	<0.001	50	135	338	329	<0.001
Total	31	174	433	1202		191	263	437	802		132	305	741	660	

LN, lymph node status (1 = negative, 2 = positive < 4, 3 = positive ≥ 4); VI, vascular invasion; LR, local recurrence; RR, regional recurrence; DM, distant metastases; positive expression (W, weak; M, moderate; S, strong).

investigator after a period of time. A good correlation was found between the two estimations.

The cut-off points for expression were assigned by using a frequency distribution histogram of all H-scores for each antibody; the cut-off was selected corresponding to the troughs between different apparent populations. Positivity was defined as detection of any invasive malignant cells positive for both CK 5/6 and CK 14 and a cut-off point of 10 was used for SMA staining. For CK 7/8, the cut-offs were 50, 120, 240. For both CK 18 and CK 19, cut-offs at H scores of 20, 100 and 200 were utilized. For ER immunoreactivity, the cut-off was 20 to divide cases into negative and positive groups. For the luminal markers CK 7, 8, 18 and 19, cut-offs were selected to identify negative, weak, moderate and strong positive expression.

Table 4. Association of basal marker expression with clinical and pathological parameters

	CK5/6			CK14			SMA		
Feature	**−**	**+**	***p*-value**	**−**	**+**	***p*-value**	**−**	**+**	***p*-value**
Grade									
1	289	42		287	38		296	30	
2	560	51		556	45		550	45	
3	663	230	<0.001	717	170	<0.001	724	160	<0.001
Total	1512	323		1560	253		1570	235	
LN									
1	950	216		959	186		993	158	
2	445	77		474	48		448	62	
3	117	29	0.127	125	19	0.001	128	15	0.438
Total	1512	322		1558	253		1569	235	
Size									
≤1.5 cm	571	79		575	70		559	78	
>1.5 cm	942	245	<0.001	986	184	0.004	1013	157	0.478
Total	1513	324		1561	254		1572	235	
NPI									
Good	545	64		533	60				
Moderate	743	197		776	159				
Poor	222	60	<0.001	247	33	<0.001			
Total	1510	321		1556	252				
ER									
Negative	327	217		396	144		417	121	
Positive	1132	92	<0.001	1110	99	<0.001	1100	112	<0.001
Total	1459	309		1506	243		1517	233	
VI									
No.	1031	230		1047	196		1081	161	
Yes	468	91	0.311	500	54	0.001	474	74	0.755
Total	1499	321		1547	250		1555	235	
LR									
No	1393	289		1438	226		1439	217	
Yes	102	32	0.050	104	27	0.026	113	17	1.000
Total	1495	321		1542	253		1552	234	
RR									
No	1417	284		1455	228		1451	221	
Yes	78	37	<0.001	87	25	0.010	101	13	0.579
Total	1495	321		1542	253		1552	234	
DM									
No	1343	262		1375	211		1381	197	
Yes	150	59	<0.001	165	42	0.007	170	37	0.031
Total	1493	321		1540	253		1551	234	
Death									
No	1387	266		1402	213		1405	201	
Yes	128	55	<0.001	140	40	0.001	147	33	0.028
Total	1495	321		1542	253		1552	234	
Age									
≤35	53	18		60	10		52	17	
36–45	249	77		274	49		269	51	
46–55	482	102		487	93		502	68	
>55	732	127	0.001	743	102	0.173	751	100	0.006
Total	1516	324		1564	254		1574	236	

LN, lymph node status (1 = negative, 2 = positive < 4, 3 = positive ≥ 4); VI, vascular invasion; LR, local recurrence; RR, regional recurrence; DM, distant metastases.

Statistical analysis

Association between the immunoreactivity with the various antibodies used and different clinicopathological parameters was evaluated by chi-squared test. A p-value of <0.05 was considered to reflect a significant correlation. Survival curves were calculated by the Kaplan–Meier method. The differences between survivals were estimated using the log rank test. Multivariate Cox regression analysis was used to evaluate any independent prognostic effect of the variables on relapse-free interval (RFI) and the overall survival (OS).

Results

Clinical outcomes

Complete clinical follow-up information was available for 1917 patients. The median follow-up period was

58 months (range 1–192 months). During this period, a total of 188 (9.8%) patients died from breast cancer. At the time of the primary diagnosis, 1231 (63.6%) of the patients had lymph node negative disease and 705 (36.4%) had positive lymph nodes (549 cases with one to three positive nodes, 156 cases with four or more positive). ER status was estimated immunohistochemically in 1805 of the tumours; 553 (30.6%) were negative for ER expression, while 1252 (69.4%) carcinomas were ER positive.

Luminal and basal marker expression and distribution

After excluding the uninformative cores from the study, 1841, 1693, 1838, 1840, 1818 and 1810 tumours were available for CK 7/8, 18 and 19, CK5/6, 14 and SMA analyses respectively.

A very high proportion of cases demonstrated positivity for the luminal cytokeratins 7/8, 18 and 19 (98.3%, 88.7% and 92.8% respectively). In contrast, a lower proportion showed expression of the basal markers (17.6% for CK5/6, 14% for CK 14 and 13% for SMA). There was a highly significant inverse correlation between the luminal (CK 7/8, CK18 and CK19) and the basal (CK 5/6, CK14 and SMA) immunoprofiles ($p < 0.001$).

Marker expression in relation to prognostic data

Tables 3 and 4 show the associations between luminal and basal markers and key prognostic and outcome variables. In essence, an inverse correlation was found between all luminal phenotype markers (assessed on a four-point scale as negative, weak, moderate and strong expression) and histological grade, tumour size, NPI, regional recurrence (CK 19 excluded), distant metastasis (CK 19 excluded) and death from breast cancer. Conversely, positive correlations with ER expression and patient age were noted.

In contrast, all the basal markers showed significant positive correlations with histological grade, tumour size (except SMA), NPI, local and regional recurrence (except SMA), distant metastasis and patient death. Significant inverse correlations were identified with ER status and patient age (except for CK 14).

Kaplan–Meier survival analyses demonstrated that the absence of a detectable basal phenotype, as identified by CK 5/6, CK 14 or SMA immunoreactivity, was significantly associated with a more favourable overall survival compared with the presence of a basal phenotype (log rank p-value $= 0.0001$ for CK 5/6 (Figure 3B), $p = 0.0220$ for CK 14 and $p = 0.0198$ for SMA). In addition, survival analyses showed that tumours with moderate or high levels of luminal marker expression had a significantly longer overall survival compared with those showing low or no expression (log rank p-value < 0.0001, $p = 0.0002$ and $p = 0.0030$ for CK 7/8 (Figure 3A), CK 18 and CK 19 respectively).

Relapse-free interval was significantly longer in cases that showed moderate or high expression of CK 7/8 (log rank $p < 0.0001$) (Figure 3C) or CK 18 (log rank $p = 0.0009$). Tumours devoid of expression of the basal markers CK 5/6 and CK 14 had a significantly longer relapse-free interval compared with their positive counterparts (log rank $p < 0.0001$ for CK 5/6 (Figure 3D) and $p = 0.0009$ for CK14).

Mutivariate Cox regression analyses estimated that the prognostic effect of CK5/6 in relation to DFS was independent of grade, lymph node stage and tumour size ($p = 0.003$) (Table 5).

Table 5. Cox multivariate analyses for CK 5/6 and other variables related to disease free survival

Prognostic variable	Hazard ratio	*p*-value
Grade		<0.001
2 versus 1	1.126 (0.773–1.640)	0.535
3 versus 1	1.850 (1.305–2.623)	0.001
LN stage		<0.001
2 versus 1	1.047 (0.812–1.350)	0.772
3 versus 1	3.159 (2.357–4.235)	<0.001
Size	1.604 (1.228–2.095)	0.001
CK 5/6	1.447 (1.130–1.851)	0.003

Marker expression in relation to histological tumour type

The expression of the luminal and basal markers in relation to histological tumour type is summarized in Tables 6 and 7. Interestingly, 3/5 and 20/40 cases of typical and atypical medullary carcinoma, all cases of adenoid cystic carcinoma, 8/8 cases of grade 3 carcinoma NST with extensive central necrosis and 5/5 cases of grade 3 carcinoma NST with squamous differentiation were positive for at least one of the basal markers. In contrast, expression of the luminal markers was predominantly observed in special type cancers: lobular, tubular mixed, tubular and invasive cribriform.

Different cellular immunoprofiles and their relation to clinicopathological parameters

By combining the results of the luminal markers together with those of basal marker expression, we subdivided the cases into four different cellular phenotypes as follows:

1. Luminal phenotype: 1323 (71.4% of cases) (tumours which expressed one or more of the luminal markers only).
2. Combined luminal and basal phenotype: 508 (27.4% of cases) (tumours which were positive for one or more of the luminal markers together with one or more of the basal markers).
3. Basal phenotype: 15 (0.8% of cases) (tumours which only expressed one or more of the basal markers).
4. Null phenotype: 6 (0.4% of cases) (tumours that were negative for both luminal and basal markers).

Table 6. Association between luminal marker expression and histological tumour type

	CK7/8					CK18					CK19				
Tumour type	**–**	**W+**	**M+**	**S+**	**Total**	**–**	**W+**	**M+**	**S+**	**Total**	**–**	**W+**	**M+**	**S+**	**Total**
Invasive NST	22	143	265	615	1045	155	181	224	401	961	94	226	413	310	1042
Tubular mixed	0	0	48	270	318	2	21	72	197	292	6	15	134	164	319
Medullary															
Typical	0	2	3	0	5	1	4	0	0	5	2	1	2	0	5
Atypical	4	17	14	5	40	16	11	4	6	37	10	17	8	5	40
Lobular															
Classical	0	2	33	94	129	0	18	49	56	123	3	8	64	53	128
Alveolar	0	0	2	0	2	0	1	1	0	2	0	0	2	0	2
Solid	0	1	2	3	6	1	1	2	1	5	2	1	2	1	6
Tubulolobular	0	1	2	3	6	1	1	1	3	6	1	0	3	2	6
Mixed	0	0	23	39	62	2	5	25	26	58	0	13	26	22	61
Tubular	0	0	6	62	68	0	2	14	42	58	0	4	20	45	69
Mucinous	0	1	6	15	22	0	3	8	9	20	0	5	10	8	23
Cribriform	0	0	0	10	10	1	0	5	4	10	0	0	4	6	10
Papillary	0	0	2	4	6	1	1	1	3	6	0	0	3	3	6
NST + lobular	0	2	16	47	65	2	7	21	32	62	2	8	35	19	64
NST + special type	1	0	8	30	39	2	4	6	20	32	2	6	12	19	39
Miscellaneous															
Adenoid cystic	1	3	0	1	5	3	2	0	0	5	4	1	0	0	5
Metaplastic	2	0	0	1	3	2	0	1	0	3	2	0	0	1	3
Spindle cell	1	0	0	0	1	0	0	0	0	0	1	0	0	0	1
Apocrine car	0	0	1	0	1	0	0	1	0	1	0	0	0	1	1
NST + clear cell	0	0	1	0	1	0	1	0	0	1	1	0	0	0	1
NST + secretory	0	1	0	0	1	1	0	0	0	1	0	0	0	0	0
NST + spindle cell elements	0	1	0	0	1	1	0	0	0	1	1	0	0	0	1

Table 7. Association between the basal markers expression and histological tumour types

	CK5/6			CK14			SMA		
Tumour type	**–**	**+**	**Total**	**–**	**+**	**Total**	**–**	**+**	**Total**
Invasive NST	829	218	1047	872	169	1041	866	166	1032
Tubular mixed	284	34	318	282	28	310	281	31	312
Medullary									
Typical	3	2	5	5	0	5	3	2	5
Atypical	20	20	40	28	12	40	31	8	39
Lobular									
Classical	123	7	130	113	9	122	121	2	123
Alveolar	2	0	2	2	0	2	2	0	2
Solid	5	1	6	5	1	6	6	0	6
Tubulolobular	4	1	5	6	0	6	6	0	6
Mixed	57	3	60	59	2	61	65	3	68
Tubular	65	4	69	59	5	64	18	5	23
Mucinous	16	6	22	19	6	25	18	5	23
Cribriform	10	0	10	10	0	10	10	0	10
Papillary	4	2	6	4	2	6	4	2	6
NST + lobular	54	10	64	57	8	65	59	4	63
NST + special type	31	8	39	33	5	38	31	6	37
Miscellaneous									
Adenoid cystic	0	5	5	0	5	5	0	5	5
Metaplastic	2	1	3	3	0	3	3	0	3
Spindle cell tumour*	0	1	1	0	1	1	0	1	1
Apocrine carcinoma	1	0	1	1	0	1	1	0	1
NST + clear cell features	1	0	1	1	0	1	1	0	1
NST + secretory features	1	0	1	1	0	1	1	0	1
NST + spindle cell elements	0	1	1	0	1	1	0	1	1

* Spindle cell tumour associated with metaplastic features.

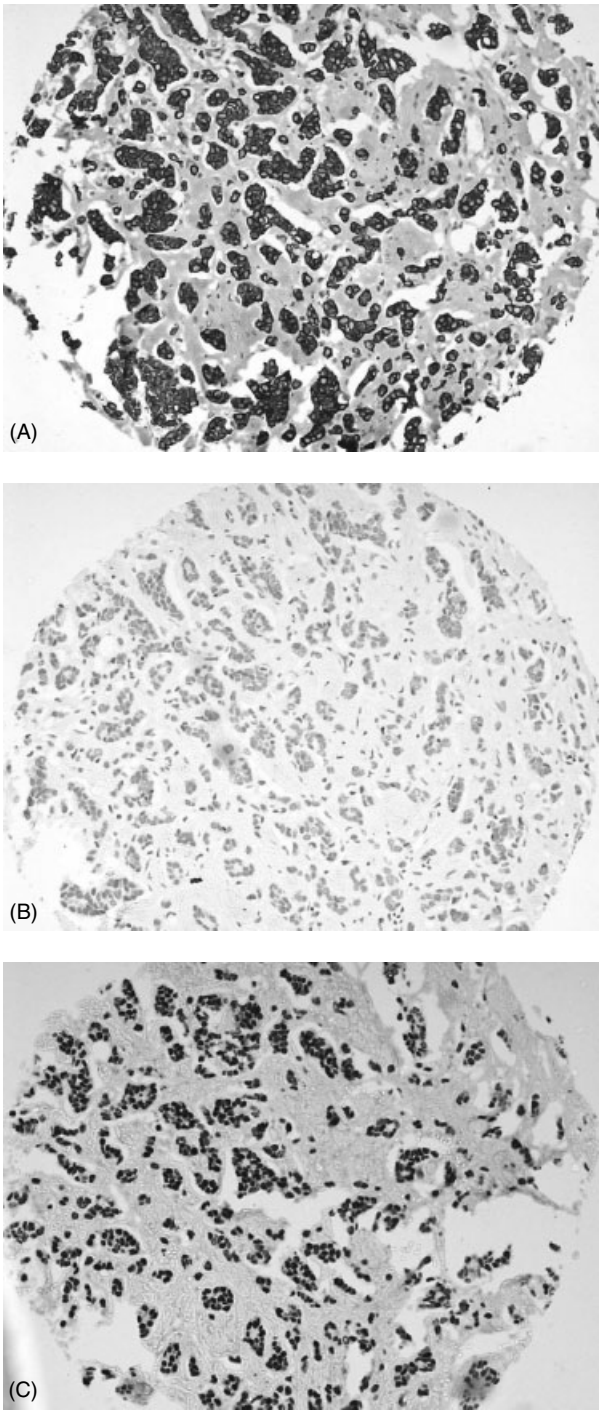

Figure 1. A histological grade 2 invasive carcinoma of ductal/NST showing expression of the luminal marker CK 18 (A), and ER (C), but no reactivity for the basal marker CK 5/6 (B)

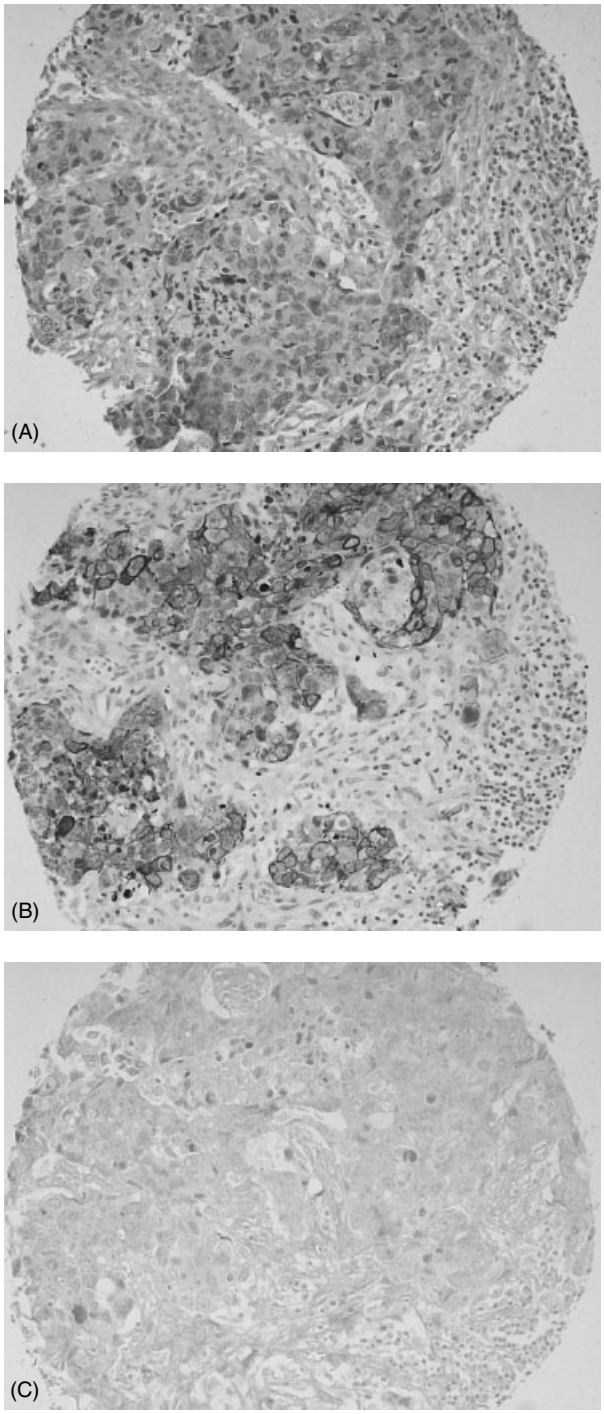

Figure 2. A histological grade 3 invasive carcinoma of ductal/NST showing combined expression of CK18 (A) and CK 5/6 (B) but no expression of ER (C)

When these immunoprofiles were compared in relation to different clinicopathological variables, significant associations were identified between the different phenotype groups and histological grade, tumour size, ER status (Figures 1 and 2), NPI, distant metastases and death from breast cancer (Table 8).

Kaplan–Meier analyses of these four groups revealed significant differences regarding the OS and the RFI (log rank $p = 0.0023$ and $p = 0.0003$ respectively). Highly significant differences were noted between the luminal group and the combined luminal and basal group for overall survival and relapse-free survival (log rank $p = 0.0002$ and $p < 0.0001$ respectively) (Figure 3E and F). No significant differences were identified amongst other groups.

Discussion

These results demonstrate that a high proportion of invasive breast carcinomas express only the luminal epithelial cell cytokeratins and have a pure luminal

Table 8. Cellular profiles in relation to pathological and clinical variables

Feature	Luminal	Luminal and basal	Basal	Null	*p*-value
Grade					
1	265 (20.1%)	72 (14.2%)	1 (7.1%)	0	
2	522 (39.6%)	94 (18.5%)	1 (7.1%)	0	
3	532 (40.3%)	342 (67.3%)	12 (85.7%)	6 (100%)	<0.001
Total	1319	508	14	6	
Size					
≤1.5 cm	515 (39%)	144 (28.4%)	2 (13.3%)	0	
>1.5 cm	806 (61%)	363 (71.6%)	13 (86.7%)	6 (100%)	<0.001
Total	1321	507	15	6	
NPI					
Good	503 (38.2%)	111 (22%)	2 (14.3%)	0	
Moderate	620 (47.1%)	311 (61.6%)	8 (57.1%)	5 (83.3%)	
Poor	194 (14.7%)	83 (16.4%)	4 (28.6%)	1 (16.7%)	<0.001
Total	1317	505	14	6	
ER					
Negative	253 (19.9%)	274 (55.5%)	15 (100%)	5 (100%)	
positive	1021 (80.1%)	220 (44.5%)	0	0	<0.001
Total	1274	494	15	5	
Distant metastases					
No	1175 (90.3%)	422 (83.6%)	13 (92.9%)	5 (83.3%)	
Yes	126 (9.7%)	83 (16.4%)	1 (7.1%)	1 (16.7%)	0.001
Total	1301	505	14	6	
Death					
No	1197 (91.9%)	430 (85.1%)	12 (85.7%)	5 (83.3%)	
Yes	106 (8.1%)	75 (14.9%)	2 (14.3%)	1 (16.7%)	<0.001
Total	1303	505	14	6	

cell phenotype, with combination of luminal and basal markers expression (ie a mixed luminal and basal phenotype) forming the second main group. Exclusive expression of the basal epithelial markers (ie a pure basal phenotype) was restricted to a very small subset, and a very rare group of cases showing no expression of either luminal or basal markers (a null phenotype) was also identified. Previous studies have also recorded that the luminal CKs 7, 8, 18 and 19 are predominantly expressed in breast cancer [24,25]. Expression of the basal markers CK 5/6, CK14 and SMA has been reported in invasive breast carcinoma, ranging from 4% to 16% of cases [4,26–32]. Our study is the largest series to date, and is composed of a consecutive series from a single centre with long-term follow-up and thus provides robust data on the proportions of these phenotypes of breast cancer in an unselected population.

On studying the association between the expression of these markers with different clinical and pathological parameters, we found that expression of luminal markers was associated with good prognostic tumour characteristics and outcome, in contrast to the expression of basal markers, which was associated with poor prognostic features and behaviour. These findings are in accordance with previous studies where an inverse association between CK 8, 18 expression and tumour grade, recurrence rate and ER negative status has been reported [33–35]. They also support previous research which has shown that highly metastatic cell lines are associated with loss of CK 18 expression [34].

Expression of the luminal markers was significantly related to overall survival in this series; in particular, patients with high or moderate expression showed better overall survival compared with those with low or no expression of these luminal markers. The converse was observed in tumours that labelled with the basal markers, where positive cases were associated with poor outcome particularly with CK 5/6 expression, which proved to be an independent prognostic predictor for RFI. Previous studies have reported that CK 8 is associated with better overall survival (although this did not reach statistical significance) and is an independent prognostic indicator of relapse-free survival [33]; and that CK 18 expression is an independent prognostic factor in predicting overall survival [34]. In contrast, a significant association has been reported between poor overall survival and expression of the basal markers, CK 5/6 and 17, and these markers also had an independent prognostic impact in patients without nodal metastasis [32]. In this study we also observed an association between expression of the luminal markers and patient age. Positivity for both CK 5/6 and SMA was reciprocally related to age, supporting the finding that breast cancer in younger women is more aggressive, with lower hormone receptor levels, higher proliferation and a worse prognosis compared with those in women of older age [36–38].

On studying the expression of basal markers in different morphological tumour types, we found that tumours of special type were either absolutely negative or expressed these markers in a small proportion of

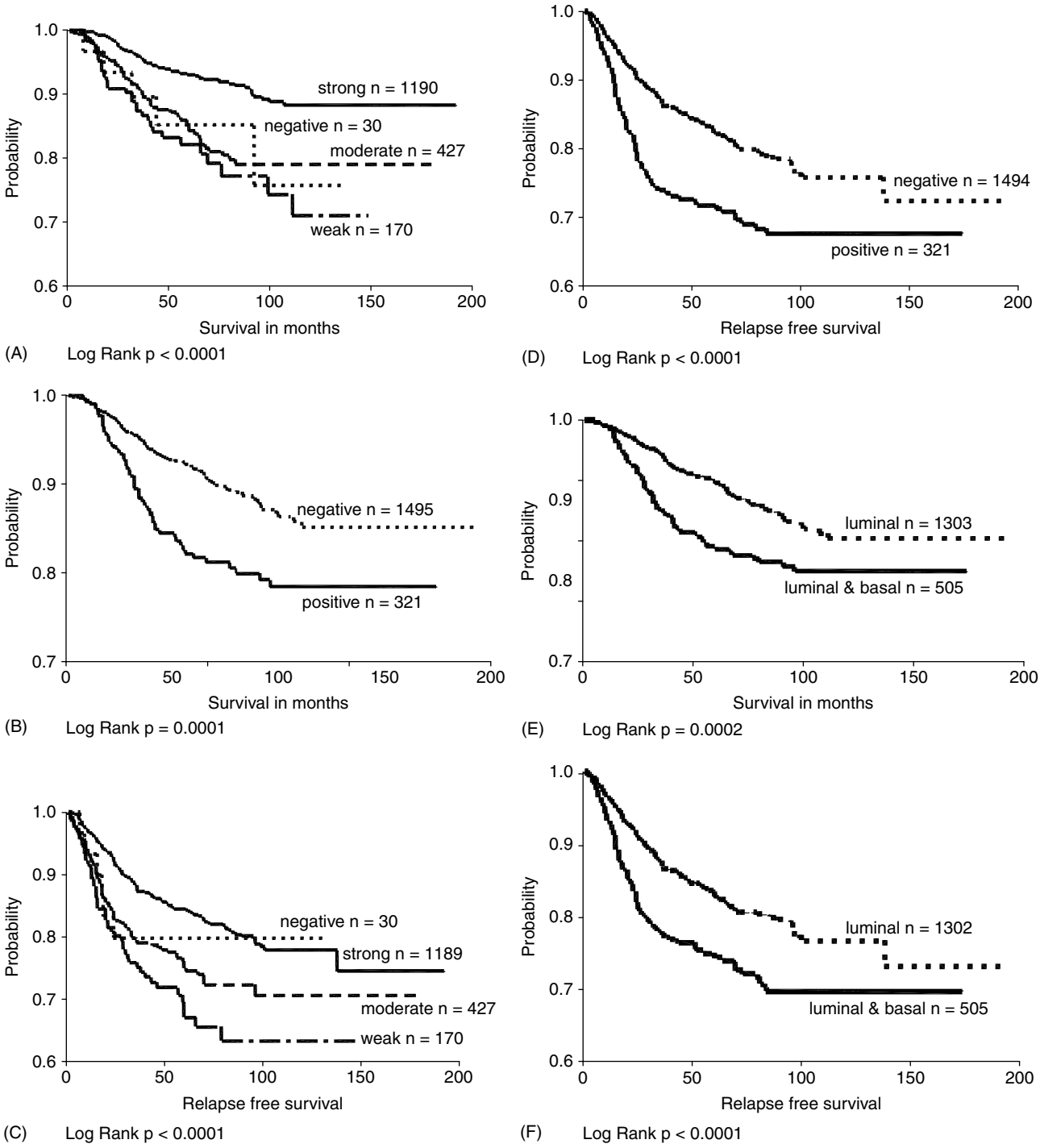

Figure 3. (A) CK 7/8 expression in relation to overall survival. (B) CK 5/6 expression in relation to overall survival. (C) CK 7/8 expression in relation to relapse-free survival. (D) CK 5/6 expression in relation to relapse free survival. (E) Luminal/combined luminal and basal phenotypes in relation to overall survival. (F) Luminal/combined luminal and basal phenotypes in relation to relapse free survival

cases. The exception was medullary [typical and atypical] carcinomas, a great proportion of which were positive for one or more of the basal markers, despite their good prognosis. The same finding has been reported in a previous study [39]. We also noticed that all adenoid cystic carcinomas were positive for at least one of the basal markers, confirming their differentiation towards the myoepithelial pathway [40]. We found that invasive grade 3 carcinomas of no special type associated with extensive central necrosis had a basal phenotype and were ER negative. These findings are in keeping with two previous reports that reported a high prevalence of the basal phenotype in ductal carcinoma with extensive necrosis and their aggressive behaviour [41,42]. We also noted the expression of the basal phenotype in one case of metaplastic carcinoma and one case of spindle cell tumour, as previously reported [43,44]. Overall, tumours expressing the basal phenotype (either combined luminal and basal, or basal markers only) were more often grade 3 tumours and

ER negative with poor outcome. These findings have been reported previously where the bimodal phenotype (combined basal and luminal) was identified in 62% of poorly differentiated breast cancers. These cases were more often steroid receptor negative, tended to have more frequent metastases, shorter RFI and poorer OS [25]. A previous genetic study on grade 3 carcinoma of no special type with basal phenotype identified specific genetic alterations and the aggressive nature of that subset [45].

Recent cDNA gene expression analysis and TMA immunohistochemical studies have proposed two distinguishable groups with luminal and basal phenotypes that have different cytogenetic alterations and protein expression patterns [6,46]. Despite the complexity of expression of the markers used in the present study, we were able to identify four profiles: luminal, combined luminal and basal, basal and null/no expression. This supports the finding of studies that have, similarly, reported cases of breast cancer with a pure luminal phenotype, a basal phenotype or a combined luminal and basal phenotype [4,25,30]. Two previous studies have reported cases of breast cancer which were negative for both the luminal and basal markers [27,30]. Both studies used frozen section material in which immunoreactivity was optimally preserved. One could argue that cases having no demonstrable phenotype may be a consequence of loss of reaction due to differences in tissue handling. However, all cases in our study were handled in a similar way and optimally fixed in formalin. We identified six cases that were devoid of expression of both the luminal and basal markers. All were grade 3 and were ER negative. It is possible that this phenotype could reflect non-epithelial derivation or de-differentiation of an epithelial-derived tumour to a more primitive subclass.

Our findings indicate that there are several cellular profiles in breast cancer; each one may reflect alternative pathways of epithelial differentiation during carcinogenesis. These data provide supportive evidence for subgrouping breast cancer into different phenotypes: a stem cell phenotype (CK 5/6+), an intermediate glandular phenotype (CK 5/6+, CK 8/18+) and a differentiated glandular phenotype (CK 8/18+) [46]. Although we identified broadly similar patterns, we observed more complex phenotypes in some cases owing to different combinations of the luminal and the basal markers expressed.

In summary, we have identified, in this large series of invasive breast cancers, distinct subclasses based on luminal and basal epithelial marker expression. These subclasses have significant differences in tumour characteristics and in clinical outcome. Our findings provide further evidence that breast cancer has distinct differentiation subclasses which have both biological and clinical relevance and which may reflect different mechanisms of histogenesis or development along different lineage pathways.

Acknowledgements

We thank Mr Peter and Mr Jone Roman for technical assistance in tissue array construction and immunohistochemical staining.

References

1. Taylor-Papdimitriou J, Lane EB. Keratin expression in the mammary gland. In *The Mammary Gland: Development Regulation and Function*, Daniel CW (ed). Plenum Press: New York, 1987; 181–215.
2. Cooper D, Schermer A, Sun TT. Classification of human epithelia and their neoplasms using monoclonal antibodies to keratins: strategies, applications, and limitations. *Lab Invest* 1985; **52**: 243–256.
3. Debus E, Weber K, Osborn M. Monoclonal cytokeratin antibodies that distinguish simple from stratified squamous epithelia: characterization on human tissues. *Embo J* 1982; **1**: 1641–1647.
4. Nagle RB, Bocker W, Davis JR, *et al*. Characterization of breast carcinomas by two monoclonal antibodies distinguishing myoepithelial from luminal epithelial cells. *J Histochem Cytochem* 1986; **34**: 869–881.
5. Bocker W, Moll R, Poremba C, *et al*. Common adult stem cells in the human breast give rise to glandular and myoepithelial cell lineages: a new cell biological concept. *Lab Invest* 2002; **82**: 737–746.
6. Perou CM, Sorlie T, Eisen MB, *et al*. Molecular portraits of human breast tumours. *Nature* 2000; **406**: 747–752.
7. Sorlie T, Perou CM, Tibshirani R, *et al*. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 2001; **98**: 10869–10874.
8. Elston CW, Ellis IO. 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* 1991; **19**: 403–410.
9. Ellis IO, Galea M, Broughton N, Locker A, Blamey RW, Elston CW. Pathological prognostic factors in breast cancer. II. Histological type: relationship with survival in a large study with long-term follow-up. *Histopathology* 1992; **20**: 479–489.
10. Pinder SE, Ellis IO, Galea M, O'Rouke S, Blamey RW, Elston CW. Pathological prognostic factors in breast cancer. III. Vascular invasion: relationship with recurrence and survival in a large study with long-term follow-up. *Histopathology* 1994; **24**: 41–47.
11. Galea MH, Blamey RW, Elston CE, Ellis IO. The Nottingham Prognostic Index in primary breast cancer. *Breast Cancer Res Treat* 1992; **22**: 207–219.
12. Blamey RW, Davies CJ, Elston CW, Johnson J, Haybittle JL, Maynard PV. Prognostic factors in breast cancer: the formation of a prognostic index. *Clin Oncol* 1979; **5**: 227–236.
13. Elston CW, Gresham GA, Rao GS, *et al*. The cancer research campaign (King's/Cambridge trial for early breast cancer: clinicopathological aspects. *Br J Cancer* 1982; **45**: 655–669.
14. Haybittle JL, Blamey RW, Elston CW, *et al*. A prognostic index in primary breast cancer. *Br J Cancer* 1982; **45**: 361–366.
15. Du Toit RS, Locker AP, Ellis IO, Elston CW, Blamey RW. Evaluation of the prognostic value of triple node biopsy in early breast cancer. *Br J Surg* 1990; **77**: 163–167.
16. Kollias J, Murphy CA, Elston CW, Ellis IO, Robertson JF, Blamey RW. The prognosis of small primary breast cancers. *Eur J Cancer* 1999; **35**: 908–912.
17. Kononen J, Bubendorf L, Kallioniemi A, *et al*. Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med* 1998; **4**: 844–847.
18. Camp RL, Charette LA, Rimm DL. Validation of tissue microarray technology in breast carcinoma. *Lab Invest* 2000; **80**: 1943–1949.
19. Torhorst J, Bucher C, Kononen J, *et al*. Tissue microarrays for rapid linking of molecular changes to clinical endpoints. *Am J Pathol* 2001; **159**: 2249–2256.

20. McCarty KS Jr, Miller LS, Cox EB, Konrath J, McCarty KS Sr. Estrogen receptor analyses: correlation of biochemical and immunohistochemical methods using monoclonal antireceptor antibodies. *Arch Pathol Lab Med* 1985; **109**: 716–721.
21. Rimm DL, Camp RL, Charette LA, Olsen DA, Provost E. Amplification of tissue by construction of tissue microarrays. *Exp Mol Pathol* 2001; **70**: 255–264.
22. Callagy G, Cattaneo E, Daigo Y, *et al*. Molecular classification of breast carcinomas using tissue microarrays. *Diagn Mol Pathol* 2003; **12**: 27–34.
23. Ginestier C, Charafe-Jauffret E, Bertucci F, *et al*. Distinct and complementary information provided by use of tissue and DNA microarrays in the study of breast tumor markers. *Am J Pathol* 2002; **161**: 1223–1233.
24. Rejthar A, Nenutil R. The intermediate filaments and prognostically oriented morphological classification in ductal breast carcinoma. *Neoplasma* 1997; **44**: 370–373.
25. Malzahn K, Mitze M, Thoenes M, Moll R. Biological and prognostic significance of stratified epithelial cytokeratins in infiltrating ductal breast carcinomas. *Virchows Arch* 1998; **433**: 119–129.
26. Altmannsberger M, Dirk T, Droese M, Weber K, Osborn M. Keratin polypeptide distribution in benign and malignant breast tumors: subdivision of ductal carcinomas using monoclonal antibodies. *Virchows Arch B Cell Pathol Incl Mol Pathol* 1986; **51**: 265–275.
27. Dairkee SH, Puett L, Hackett AJ. Expression of basal and luminal epithelium-specific keratins in normal, benign, and malignant breast tissue. *J Natl Cancer Inst* 1988; **80**: 691–695.
28. Jarasch ED, Nagle RB, Kaufmann M, Maurer C, Bocker WJ. Differential diagnosis of benign epithelial proliferations and carcinomas of the breast using antibodies to cytokeratins. *Hum Pathol* 1988; **19**: 276–289.
29. Wetzels RH, Holland R, van Haelst UJ, Lane EB, Leigh IM, Ramaekers FC. Detection of basement membrane components and basal cell keratin 14 in noninvasive and invasive carcinomas of the breast. *Am J Pathol* 1989; **134**: 571–579.
30. Wetzels RH, Kuijpers HJ, Lane EB, *et al*. Basal cell-specific and hyperproliferation-related keratins in human breast cancer. *Am J Pathol* 1991; **138**: 751–763.
31. Santini D, Ceccarelli C, Taffurelli M, Pileri S, Marrano D. Differentiation pathways in primary invasive breast carcinoma as suggested by intermediate filament and biopathological marker expression. *J Pathol* 1996; **179**: 386–391.
32. van de Rijn M, Perou CM, Tibshirani R, *et al*. Expression of cytokeratins 17 and 5 identifies a group of breast carcinomas with poor clinical outcome. *Am J Pathol* 2002; **161**: 1991–1996.
33. Takei H, Iino Y, Horiguchi J, *et al*. Immunohistochemical analysis of cytokeratin #8 as a prognostic factor in invasive breast carcinoma. *Anticancer Res* 1995; **15**: 1101–1105.
34. Schaller G, Fuchs I, Pritze W, *et al*. Elevated keratin 18 protein expression indicates a favorable prognosis in patients with breast cancer. *Clin Cancer Res* 1996; **2**: 1879–1885.
35. Brotherick I, Robson CN, Browell DA, *et al*. Cytokeratin expression in breast cancer: phenotypic changes associated with disease progression. *Cytometry* 1998; **32**: 301–308.
36. Adami HO, Malker B, Holmberg L, Persson I, Stone B. The relation between survival and age at diagnosis in breast cancer. *N Engl J Med* 1986; **315**: 559–563.
37. Host H, Lund E. Age as a prognostic factor in breast cancer. *Cancer* 1986; **57**: 2217–2221.
38. Ferno M, Borg A, Johansson U, *et al*. Estrogen and progesterone receptor analyses in more than 4000 human breast cancer samples: a study with special reference to age at diagnosis and stability of analyses. Southern Swedish Breast Cancer Study Group. *Acta Oncol* 1990; **29**: 129–135.
39. Tot T. The cytokeratin profile of medullary carcinoma of the breast. *Histopathology* 2000; **37**: 175–181.
40. Lamovec J, Us-Krasovec M, Zidar A, Kljun A. Adenoid cystic carcinoma of the breast: a histologic, cytologic, and immunohistochemical study. *Semin Diagn Pathol* 1989; **6**: 153–164.
41. Tsuda H, Takarabe T, Hasegawa T, Murata T, Hirohashi S. Myoepithelial differentiation in high-grade invasive ductal carcinomas with large central acellular zones. *Hum Pathol* 1999; **30**: 1134–1139.
42. Tsuda H, Takarabe T, Hasegawa F, Fukutomi T, Hirohashi S. Large, central acellular zones indicating myoepithelial tumor differentiation in high-grade invasive ductal carcinomas as markers of predisposition to lung and brain metastases. *Am J Surg Pathol* 2000; **24**: 197–202.
43. Reis-Filho JS, Milanezi F, Paredes J, *et al*. Novel and classic myoepithelial/stem cell markers in metaplastic carcinomas of the breast. *Appl Immunohistochem Mol Morphol* 2003; **11**: 1–8.
44. Popnikolov NK, Ayala AG, Graves K, Gatalica Z. Benign myoepithelial tumors of the breast have immunophenotypic characteristics similar to metaplastic matrix-producing and spindle cell carcinomas. *Am J Clin Pathol* 2003; **120**: 161–167.
45. Jones C, Nonni AV, Fulford L, *et al*. CGH analysis of ductal carcinoma of the breast with basaloid/myoepithelial cell differentiation. *Br J Cancer* 2001; **85**: 422–427.
46. Korsching E, Packeisen J, Agelopoulos K, *et al*. Cytogenetic alterations and cytokeratin expression patterns in breast cancer: integrating a new model of breast differentiation into cytogenetic pathways of breast carcinogenesis. *Lab Invest* 2002; **82**: 1525–1533.