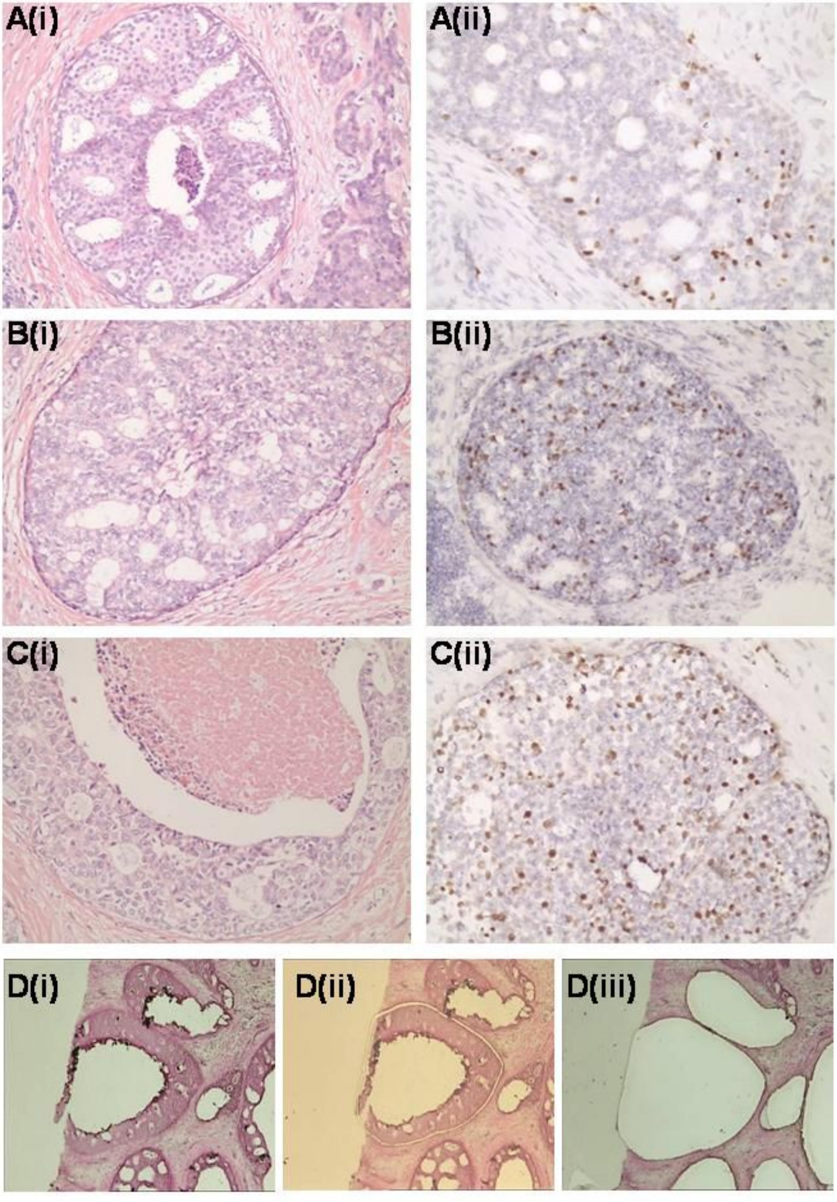
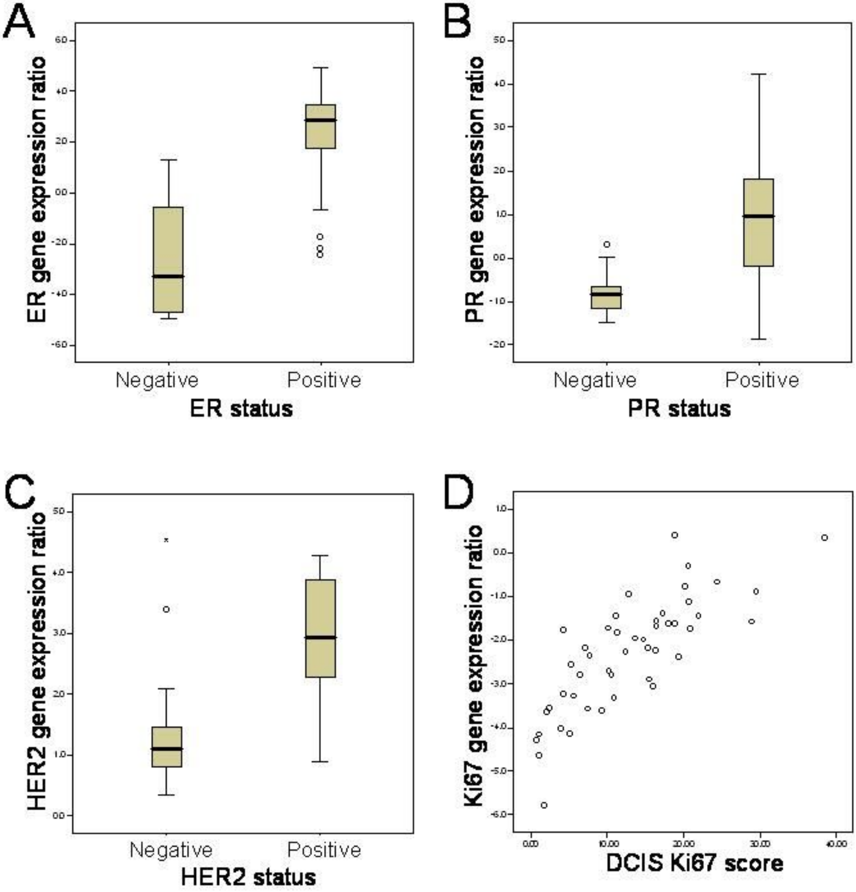
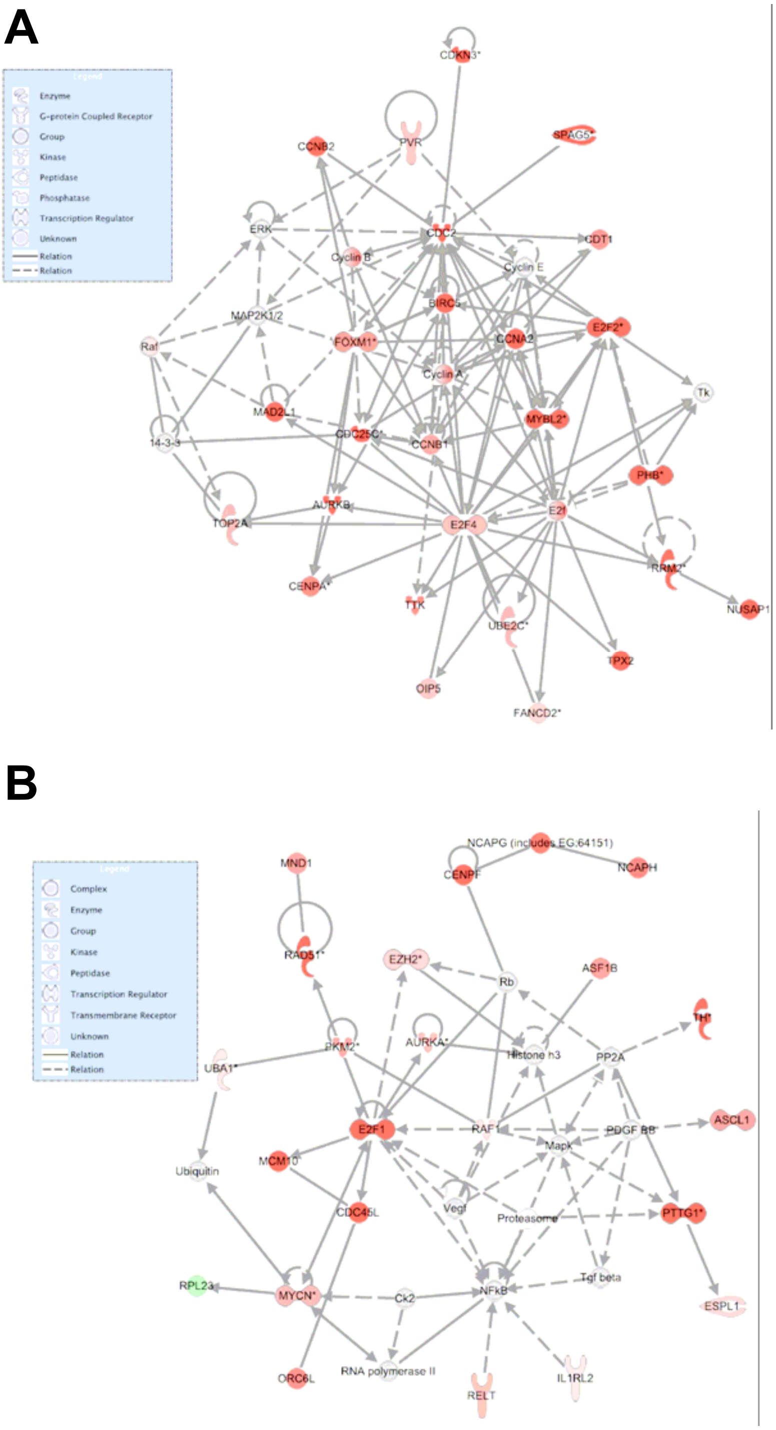


Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3

Supplementary Methods

Determination of tumor ER, PR, HER2 and p53 expression

Results from clinical assessment of tumor estrogen (ER) and progesterone receptor (PR) content were used. For 44 cases these were determined by enzyme immunoassay of tumor cytosol preparations according to a previously described method (1). For both ER and PR, measures of <10fmol/mg protein were regarded as negative. In the remaining 2 cases ER and PR had been determined by immunoperoxidase staining of frozen tissue or fine needle aspiration biopsy material.

Tumor HER2 and p53 expression were determined by immunohistochemical staining of 4μm formalin-fixed paraffin embedded tumor sections that had been stored at 4°C. For HER2, mouse monoclonal anti-human cerbB2 clone CB11 was used at a concentration of 1:40 (Novocastra, Newcastle on Tyne UK), and for p53 mouse monoclonal anti-human p53 BP53.12 at 1:100 (Zymed, CA, USA). For each stain, antigen retrieval was performed by autoclaving at 121°C 15 psi in 0.01M sodium citrate pH 6. Following treatment with H_2O_2 (3% v/v) to quench endogenous peroxidase and blocking with normal goat serum (50% v/v in phosphate buffered saline), sections were incubated for one hour in a humidification chamber with the primary antibody in 0.5% triton-X 100 in phosphate buffered saline. Sections were subsequently incubated with biotinylated goat-anti-mouse immunoglobulins (Dako, Glostrup Denmark) and avidin-biotin complex (Zymed, CA, USA). Positive results were revealed by treatment with diaminobenzadine (1 mg/ml (w/v), 0.02% (v/v) H_2O_2). A positive control section was included in each staining run. For HER2 staining, a negative control section with the primary antibody omitted was included for each case and minor non-specific staining on these sections was taken into account in interpretation of test sections.

HER2 expression was evaluated according to the Dako *HerceptTest* scoring protocol (accessed at www.dakousa.com) and +++ staining was designated positive. For p53 the proportion of tumor cells positive was estimated and the pattern of staining was noted (extensive or scattered). Cases were designated p53 positive if >50% of cell nuclei were positively stained in an extensive pattern.

RNA and DNA extraction and amplification

RNA was extracted from microdissected tissue using the *Absolutely RNA Microprep Kit* (Stratagene, CA, USA) and quantitated using the *RiboGreen RNA Quantitation Reagent* (Molecular Probes, OR, USA). RNA quality was assessed by running RNA isolated from each tissue sample on the *Agilent 2100 Bioanalyzer* (Agilent Technologies, Victoria, Australia). This showed a degree of RNA degradation in all samples but 18S and 28S ribosomal bands were clearly distinguished in 39/46 (84.8%) cases. Depending on tissue sample size, 100ng, 50ng or ≤20ng of RNA was subject to two rounds of amplification using the *RiboAmp RNA Amplification Kit* (Arcturus, CA, USA). Two-round amplified Stratagene Universal Human Reference RNA (Stratagene, CA, USA) was the reference in gene expression microarray experiments.

DNA was extracted from laser microdissected material using the *QIAamp DNA Micro Kit* (Qiagen Inc., CA, USA) and quantitated using the *PicoGreen dsDNA Quantitation Reagent* (Molecular Probes) both according to the manufacturers instructions. Approximately 10ng of DNA was amplified using the *GenomiPhi DNA Amplification Kit* (Amersham Biosciences, NJ, USA) according to the manufacturer’s instructions. Amplified Promega Normal Male DNA (Promega, WI, USA) was used as the reference for array CGH.

RNA labeling

Due to the antisense orientation of the amplified RNA samples, a labeling protocol based on a method published by Schlingemann *et al* (2) was used to generate fluorescent labeled antisense cDNA for hybridization to sense-oriented oligonucleotide microarrays.

One microgram of amplified RNA was incubated for 5 minutes at room temperature in the presence of 0.5µg of pdN6 (Amersham Biosciences) in a total volume of 5.5µL then chilled on ice. Reverse transcription mixture was added, yielding final concentrations of 1X First Strand Buffer (Invitrogen, CA, USA), 10mM DTT (Invitrogen), 500µM each of dATP, dCTP, dGTP, dTTP (Amersham Biosciences), 2U/µL RNasin ribonuclease inhibitor (Promega) and 10U/µL Superscript II reverse transcriptase (Invitrogen) in a total volume of 10µL. The following temperature profile was employed for reverse transcription: 37^{o}C 20 minutes, 42^{o}C 20 minutes, 50^{o}C 10 minutes, 55^{o}C 10 minutes, 65^{o}C 15 minutes; reactions were then held at 37^{o}C. RNase H (Promega) digestion (1U per reaction) was carried out for 30 minutes at 37^{o}C, followed by 2 minutes at 95^{o}C to degrade enzymes. Samples were chilled to 4^{o}C. cDNA labeling by Klenow fragment and amino-allyl incorporation was performed using the Bioprime Kit with a modified protocol. The 10µL cDNA sample (unpurified) was mixed with 90uL of Klenow mixture to yield a reaction mixture that contained 1X random primer solution (Invitrogen), 650µM each of dATP, dCTP, dGTP, 400µM dTTP (Amersham Biosciences), 260µM amino-allyl dUTP (Sigma, MO, USA) and 1.0U/µL Klenow fragment (Invitrogen). DNA polymerization was carried out at 37^{0}C for 16 hours. cDNA was purified using the QIAquick PCR purification kit (Qiagen Inc., CA, USA). Samples were eluted in 30µL DEPC-water (warmed to 50^{0}C) and dried in a speed vac on medium heat. Samples were re-suspended in DEPC-H$_2$O and coupled for 1 hour in the dark to the appropriate NHS ester Cy dye (Cy3 or Cy5, Amersham Biosciences) that had been prepared in DMSO and bicarbonate buffer (pH 8.5). Uncoupled dye was removed by cleaning the samples using the QIAquick PCR purification kit. Cleaned probes were eluted in 100µL elution buffer (Qiagen). Individual labeling reactions (Cy3 and Cy5) were analyzed using spectrophotometry by measuring absorbance at 260, 550 (Cy3) and 650nM (Cy5). An arbitrary wavelength (750nM) was also measured to normalize the baselines between samples. The labeled Cy3 and Cy5 probes pairs were combined in an equimolar ratio and concentrated in a speedvac to 18µL. Purified, labeled cDNA was mixed with 1X hybridization buffer (50% formamide, 10XSSC, 0.4% SDS) and 50µg Cot1 DNA and 100µg yeast tRNA were added to block repetitive sequence elements. Probes were denatured at 95^0C for 5 minutes and pre-blocked for 60 minutes at 42^0C.

DNA labeling

Three micrograms of amplified test and reference DNA were directly labeled with either Cy3-dUTP or Cy5-dUTP. Bioprime random primer mix (Invitrogen) was added (1X final concentration) to 3µg of diluted DNA; the samples were vortexed, denatured at 98^oC for 5 minutes and cooled on ice. Labeling master mix was added containing 120µM each of dATP, dCTP, dGTP, 60µM dTTP (Amersham Biosciences), 2.5nmol Cy3-dUTP or Cy5-dUTP (Amersham Biosciences), 0.8U/µL Klenow Fragment (Invitrogen). Samples were vortexed and incubated in a thermocycler at 37^0C for 2 hours. DNA was purified with the QIAquick PCR purification kit and eluted in 30µL Buffer EB (warmed to 50^0C) (Qiagen). The labeled Cy3 and Cy5 probes pairs were combined and concentrated in a speedvac to 18µL. A hybridization mix (warmed to 42^0C) containing 50µg Cot1 DNA, 100µg yeast tRNA, 50% formamide, 5X SSC and 0.2% SDS was added to each sample. The probe was denatured at 95^oC for 5 minutes and pre-blocked by incubation at 42^oC for 60 minutes in a water bath.

Oligonucleotide microarray hybridization and analysis

Microarrays were pre-blocked in 5X SSC, 1% BSA and 0.2% SDS at 42^oC for 60 minutes. Slides were washed in 2 changes of dH_2O and 100% isopropanol for 1 minute each prior to drying by centrifugation. Probes were applied to the oligonucleotide microarrays mounted in a Bio-Micro Maui Hybstation and hybridized for 20 hours at 45^oC. Following hybridization slides were washed in 0.5XSSC, 0.05% SDS; 2 changes of 0.5XSSC and 0.1XSSC for 5 minutes each and then dried by centrifugation.

Microarray slides were scanned on an Agilent DNA Microarray Scanner. Image analyses were using DeArray software (Scanalytics, VA, USA) (3). The open source statistical environment R (www.r-project.org) was used for statistical analysis. Gene expression microarrays were loess normalized using the limma package (4). Supervised analysis was performed using the limma package from Bioconductor. Except where otherwise noted, gene lists were defined using a false discovery rate (FDR) of 0.05.

Microarray based CGH analysis

Scanning and image analysis of CGH microarrays were as for the gene expression arrays. No normalization was performed.

Segmentation provides a means of “smoothing” the data while maintaining its underlying genomic structure. As part of the segmentation process the “center” of the data, representing the modal copy number, was determined by performing multiple segmentation runs using different offsets for the raw data; the offset that produced the smallest number of probes counted as gained or lost was chosen as the “center” of the data. Doing so corresponds to the hypothesis that the most parsimonious “center” of the data is that which produces the fewest number of aberrant probes. Note that we cannot determine the absolute copy number easily from the CGH data themselves and then aberrations are then measured relative to the centre of the data as defined above.

References

1. Graham JD, Yeates C, Balleine RL, Harvey SS, Milliken JS, Bilous AM, et al. Characterization of progesterone receptor A and B expression in human cancer. Cancer Res 1995;55:5063-8.

2. Schlingemann J, Thuerigen O, Ittrich C, Toedt G, Kramer H, Hahn M, et al. Effective transcriptome amplification for expression profiling on sense-oriented oligonucleotide microarrays. Nucleic Acids Res 2005;33:e29.

3. Chen Y, Dougherty ER, Bittner ML. Ratio-based decisions and the quantitative analysis of cDNA microarray images. J Biomed Opt 1997;2:364-374.

4. Smyth GK, Michaud J, Scott HS. Use of within-array replicate spots for assessing differential expression in microarray experiments. Bioinformatics 2005;21:2067-75.

Supplementary Table 1: Detailed cohort characteristics for 46 cases of invasive breast cancer. Concomitant DCIS was present in 45 cases and one case was associated with LCIS only. In six cases with apparently distinct foci of invasive cancer,

Case #	Sample Type	Patient Age	Tumor Size	Histological Type	Grade	Tubules	Nuclear Pleomorphism	Mitotic Score	Lymphovascular Invasion	Lymphocytic Infiltrate	LN Status	LN Pos/Neg	Mixed DCIS Morphologies	DCIS Nuclear Grade	DCIS Necrosis	DCIS Calcification	DCIS Cell Polarization
1	HG DCIS	4	4	1	3	3	3	3	2	2	2	1	2	3	2	2	1
1.1	IG DCIS													2	1	1	2
2	DCIS	4	4	1	1	2	1	1	2	2	2	1	1	1	1	2	2
2.1	LCIS																
2.2	ADH																
3	DCIS	4	1	1					1	2	0	0	2	2	2	2	2
3.1	Benign epithelium																
4	DCIS	3	1	1	1	1	2	1	1	1	1	1	2	2	2	2	2
5	DCIS	2	4	1	2	3	2	1	2	1	3	1	1	2	2	2	2
5.1	Benign epithelium																
6	DCIS	5	3	4	1	2	2	1	2	1	1	1	1	1	1	2	2
7	DCIS	3	3	1	3	3	3	3	2	2	1	1	2	3	2	2	2
8	DCIS	2	3	1	2	2	3	1	2	3	3	1	1	3	2	2	2
8.1	Benign epithelium																
9	DCIS	3	3	1	3	3	3	2	2	2	1	1	2	3	2	2	1
9.1	Benign epithelium																
9.2	PDWA																
10	DCIS	6	4	4	1	2	2	1	1	1	1	1	2	1	2	2	2
10.1	ADH																
11	DCIS	3	1	1					2	2	1	1	1	3	2	2	2
12	DCIS	3	3	1	1	2	2	1	1	2			1	1	1	1	2
13	DCIS	3	5	1	2	2	2	2	2	2	3	1	2	2	2	2	2
14	DCIS	6	2	1	2	3	2	1	1	1	0	0	2	2	2	2	2
15	DCIS	1	4	1	2	3	3	1	2	1	2	1	1	3	2	2	2
16	DCIS	4	3	1	1	2	1	1	2	2	1	1	2	1	2	1	2
17	DCIS	4	4	1	2	2	3	2	2	2	0	0	1	3	2	1	1
18	DCIS	3	4	1	2	3	2	2	2	1	0	0	1	1	2	2	2
18.1	ADH																
19	DCIS	3	3	1	2	2	2	2	2	3	0	0	1	3	2	1	2
20	DCIS	3	3	1	2	2	3	2	2	2	0	0	1	3	2	1	1
20.1	Benign epithelium																
21	DCIS	6	2	1	3	3	3	3	2	3	1	1	1	3	2	2	2
22	DCIS	4	4	1	1	2	2	1	2	2	1	1	1	2	2	1	2
23	DCIS	3	5	1	2	3	2	1	2	1	1	1	1	2	1	2	1
24	DCIS	4	4	1	1	2	2	1	2	2	1	1	2	3	2	2	2
25	DCIS	3	1	1	1	2	1	1	2	1	0	0	2	2	1	1	2
26	DCIS	5	4	1	2	2	3	2	2	3	1	1	1	3	2	1	2
27	DCIS	4	3	1	1	1	2	2	1	1	0	0	1	2	2	2	2
28	DCIS	3	4	1	2	3	2	1	2	1	1	1	1	2	2	2	1
28.1	ADH																
29	DCIS	4	2	1	3	3	3	2	2	2	0	0	1	3	2	2	2
29.1	Benign epithelium																
30	DCIS	4	2	1	3	3	3	3	2	2	1	1	2	3	2	2	2
31	DCIS	2	4	1	3	3	3	3	2	2	1	1	1	3	2	2	2
32	DCIS	4	4	1	1	1	2	1	2	2	0	0	1	2	2	1	2
34	DCIS	4	2	3	1	1	1	1	1	1	0	0	1	1	1	2	2
35	DCIS	5	3	1	2	3	2	2	2	2	1	1	1	2	1	1	2
37	IG DCIS	5	4	1	1	2	2	1	2	2	2	1	2	2	1	2	2
37.1	LG DCIS													1	2	1	2
38	DCIS	3	1	5					2	2	0	0	1	2	2	2	2
39	DCIS	3	4	1	3	3	3	3	2	2	0	0	1	3	2	2	1
40	DCIS	2	3	1	2	2	2	2	1	2	1	1	1	2	2	1	2
41	DCIS	3	3	1	2	3	2	1	2	2	1	1	2	2	2	2	2
42	DCIS	4	5	1	3	3	3	3	2	3	2	1	1	3	2	2	1
43	DCIS	5	3	1	1	1	1	1	1	1	1	1	1	1	1	2	2
44	DCIS	1	3	1	2	3	2	1	1	2	2	1	1	1	2	2	2
45	DCIS	3	4	1	2	2	3	1	2	3	1	1	1	3	1	1	2
45.1	Benign epithelium																
46	LCIS	2	4	2	3	3	3	3	2	2	1	1	1	2	2	2	1
48	DCIS	2	3	1	2	3	2	1	2	2	1	1	1	2	2	2	2
50	DCIS	4	3	1	2	2	3	2	2	1	1	1	1	3	2	2	2

Abbreviations: DCIS, ductal carcinoma in situ; HG, high nuclear grade; IG, intermediate nuclear grade; LG, low nuclear grade; LCIS, lobular carcinoma in situ; ADH, atypical ductal hyperplasia; PDWA, proliferative disease without atypia

Key:

Patient age (years)- 1, 20-29; 2, 30-39; 3, 40-49; 4, 50-59; 5, 60-69; 6, 70-79
Tumor size (mm)- 1,≤5; 2, >5-≤10; 3, >10-≤20; 4, >20-≤50; 5, >50 (represents the size of the invasive lesion only)
Histological type- 1, ductal NOS; 2, lobular pleomorphic; 3, cribriform; 4, mixed histological types; 5, too small to type
Invasive Cancer Grade- 1, grade 1; 2, grade 2; 3, grade 3
Lymphovascular Invasion- 1, Absent; 2, Present
Lymphocytic Infiltrate- 1, Scant/Absent; 2, Moderate; 3, Prominent
Lymph Node (LN) Status- 0, no involved nodes; 1, 1-3 positive nodes; 2, 4-9 positive nodes; 3≥10 positive nodes
Lymph node (LN) Pos/Neg- 0, Negative; 1, Positive

Mixed DCIS Morphologies- 1, Uniform; 2, Mixed
DCIS Nuclear Grade- 1, low nuclear grade; 2, intermediate nuclear grade; 3, high nuclear grade
DCIS Necrosis- 1, Absent or Punctate necrosis; 2, Comedo necrosis
DCIS Calcification- 1, Absent; 2, Present (coarse or fine/punctate calcification)
DCIS Cell Polarization- 1, Absent; 2, Present
DCIS Architectural Pattern- 1, polarized architecture only (1 or more of cribriform, micropapillary, papillary); 2, mixed polarized and non-polarized architecture; 3, non-polarized architecture only (1 or more of solid, cribriform)

ER (determined by enzyme immunoassay)- 0, negative; 1, positive
ER level (fmol/mg protein)- 0,≤10; 1, 11-50; 2, 51-150; 3, >150
PR (determined by enzyme immunoassay)- 0, negative; 1, positive
PR level (fmol/mg protein)- 0,≤10; 1, 11-50; 2, 51-150; 3, >150
(Note: For 44 cases ER & PR was determined by enzyme immunoassay of tumor cytosol preparations according to a previously described method [1]. For both ER and PR, measures of <10fmol/mg protein were regarded as negative. In the remaining 2 cases ER and PR had been determined by immunoperoxidase staining of frozen tissue or fine needle aspiration biopsy material.)

1. Graham, J.D., et al., *Characterization of progesterone receptor A and B expression in human cancer.* Cancer Research, 1995. **55**(21): p. 5063-8.

HER2 status- 0, negative; 3, positive
p53 status- 0, negative; 1, positive

DCIS Ki67 score (determined by immunohistochemistry)
(Note: Values in red represent Ki67 scores derived from paraffin block sections as DCIS had cut out of the frozen tissue block)
DCIS Ki67 pos cell count- number of positive cells counted
DCIS Ki67 total cell count- total number of cell counted for Ki67 score (including positive and negative cells)
(Note: For each score a minimum of 1500 cells were counted if possible or else all available cells were counted with a minimum of 500 cells required).

Gene Expression Group (clustering) - two cluster groups defined by hierarchical clustering of the top 100 grade associated oligonucleotide probes
GGI- '*gene expression grade index'* calculated from expression of 173 grade associated oligonucleotideprobes (see Methods and Supplementary Methods for details of GGI calculation)
Molecular Grade (MG) Group- 0, low molecular grade; 1, high molecular grade (Molecular grade groups were defined using a GGI cut-off of 0)
High level DNA amplification- 'high level amplification' was defined as a discrete amplification exceeding a threshold corresponding to the 95% of the distribution of segment means (for all tumors, all segments) from the segmentation algorithm.

invasive cancer features were derived from the lesion associated with *in situ* carcinoma sampled for molecular analysis.

DCIS Architectural Pattern	ER	ER level	PR	PR level	HER2 status	p53	DCIS Ki67	DCIS Ki67 pos cell count	DCIS Ki67 total cell count	Gene Expression Group (clustering)	GGI	MG Group	High Level Amplifications	ng of RNA amplified with RiboAmp	NIH Oligo Array Expt #	NIH Oligo Array Label	ng of DNA amplified with GenomiPhi	aCGH Expt #	aCGH Array Label
3	0	0	0	0	3	0	38.54	822	2133	2	3.29	1	10	100	19	HV3_35_20	10.0	6	HV3_27_54
1							19.35	264	1364	2	0.73	1	5	100	19	HV3_35_22	10.0	6	HV3_27_55
1	1	1	1	2	0	0	0.99	15	1510	1	-1.37	0		100	19	HV3_35_23			
										2	0.29	1		100	19	HV3_35_21			
										1	-1.51	0		20	25	HV3_35_58			
1	0	0	0	0	0	1	10.26	156	1521	2	0.22	1	0	100	20	HV3_35_27	10.0	3	HV3_27_46
										1	-1.85	0		20	25	HV3_35_59			
1	1	3	1	3	0	0	10.91	189	1732	1	-0.86	0	2	100	20	HV3_35_28	10.0	6	HV3_27_56
2	1	1	1	2	0		20.68	375	1813	2	1.04	1	2	100	20	HV3_35_29	10.0	7	HV3_27_57
										1	-1.36	0		<20	26	HV3_35_70			
1	1	3	1	1	0	0	9.27	166	1791	1	-0.62	0	5	100	22	HV3_35_33	10.0	7	HV3_27_58
2	1	1	0	0	3	0	11.18	170	1521	2	1.57	1	6	100	23	HV3_35_38	10.0	7	HV3_27_59
2	1	1	1	1	3	1	18.79	296	1575	2	1.28	1	8	100	26	HV3_35_60	10.0	7	HV3_27_60
										1	-1.53	0		20	26	HV3_35_71			
3	1	1	1	2	0	1	18.85	284	1507	2	3.50	1	10	100	22	HV3_35_34	10.0	8	HV3_27_61
										1	-1.91	0		20	26	HV3_35_63			
										2	-0.11	0		20	27	HV3_35_79			
2	1	2	1	2	0	0	2.11	36	1703	1	-1.34	0	1	100	22	HV3_35_35	10.0	8	HV3_27_62
										1	-1.98	0	0	20	26	HV3_35_64	10.0	17	HV3_27_97
2	0	0	0	0	3	0	29.52	498	1687	2	1.57	1	11	100	26	HV3_35_61	10.0	5	HV3_27_51
1	1	1	1	3	0	0	5.39	89	1651	1	-0.71	0	1	20	19	HV3_35_24	6.0	8	HV3_27_63
1	1	2	1	3	0	0	21.94	375	1709	2	0.96	1	3	100	26	HV3_35_62	10.0	8	HV3_27_64
1	1	3	1	3	0	0	12.45	222	1783	1	-0.67	0	0	100	26	HV3_35_86	10.0	9	HV3_27_65
2	1	2	1	2	0	0	13.74	235	1710	2	1.07	1	7	100	22	HV3_35_36	10.0	9	HV3_27_66
1	1	3	1	1	0	0	4.23	64	1512	1	-1.28	0	1	100	26	HV3_35_87	10.0	9	HV3_27_67
3	0	0	0	0	3	0	16.47	273	1658	2	0.49	1	20	100	26	HV3_35_68	10.0	9	HV3_27_68
2	1	2	1	3	0	0	2.42	39	1611	1	-0.81	0	0	100	25	HV3_35_52	10.0	10	HV3_27_69
										1	-0.69	0	1	<20	27	HV3_35_80	0.6	16	HV3_27_93
2	1	1	1	3	3	0	11.23	189	1683	2	0.44	1	2	100	23	HV3_35_40	10.0	17	HV3_27_98
3	0	0	1	1	0	1	15.26	98	642	2	1.37	1	0	100	22	HV3_35_37	10.0	10	HV3_27_70
										1	-0.41	0		20	26	HV3_35_65			
2	1	1	0	0	0	0	20.62	386	1872	2	2.15	1	5	100	23	HV3_35_39	10.0	10	HV3_27_71
2	1	3	1	2	0	0	4.04	72	1781	1	-1.64	0	1	100	25	HV3_35_53	10.0	10	HV3_27_72
3	1	2	1	2	0	0	12.81	217	1694	2	0.25	1	5	100	23	HV3_35_41	10.0	11	HV3_27_73
2	1	1	1	1	3	1	14.72	236	1603	2	1.12	1	10	100	23	HV3_35_42	10.0	11	HV3_27_74
2	1	2	1	3	0	0	7.43	135	1818	1	-1.03	0	1	100	24	HV3_35_44	10.0	6	HV3_27_53
2	0	0	0	0	0	0	16.47	272	1651	2	0.94	1	6	100	24	HV3_35_45	10.0	19	HV3_27_42
1	1	1	1	3	0	0	7.14	101	1414	1	-0.84	0	2	50	24	HV3_35_46	10.0	11	HV3_27_75
3	1	1	1	3	0	0	17.98	280	1557	2	1.26	1	1	50	23	HV3_35_43	10.0	17	HV3_27_99
										1	-0.62	0	5	<20	27	HV3_35_81	1.4	16	HV3_27_94
2	1	3	0	0	0	0	6.5	129	1984	2	1.35	1	3	100	24	HV3_35_47	10.0	11	HV3_27_76
										1	-1.75	0		<20	27	HV3_35_82			
2	1	2	1	1	0	0	20.22	333	1647	2	1.38	1	6	100	24	HV3_35_49	10.0	12	HV3_27_77
2	1	1	1	1	0	0	10.23	163	1594	2	1.92	1	2	100	24	HV3_35_50	10.0	12	HV3_27_78
2	1	2	1	3	0	0	5.64	97	1720	1	-1.36	0	1	20	25	HV3_35_54	10.0	12	HV3_27_79
1	1	2	1	2			1.74	23	1324	1	-2.15	0	1	50	25	HV3_35_55	10.0	12	HV3_27_80
2	1	3	1	3	0	0	7.7	117	1519	1	-0.52	0	2	<20	25	HV3_35_56	10.0	13	HV3_27_81
2	1	3	1	2	0	0	4.36	59	1352	1	-0.27	0	0	<20	26	HV3_35_69	1.7	13	HV3_27_82
1							1.02	16	1562	1	-1.43	0	0	<20	25	HV3_35_57	10.0	13	HV3_27_83
2	1	2	1	3	0	0	10.59	160	1511	1	-0.32	0	2	100	19	HV3_35_25	10.0	5	HV3_27_52
3	1	3	1	3	0	1	28.94	257	888	2	1.22	1	11	20	20	HV3_35_30	10.0	13	HV3_27_84
1	1	3	1	3	0	0				1	-0.19	0	7	<20	27	HV3_35_72	10.0	16	HV3_27_95
1	1	1	0	0	0	0	5.07	79	1559	1	-0.38	0	8	20	20	HV3_35_31	10.0	14	HV3_27_85
3	1	3	1	1	0	1	17.29	153	885	2	2.37	1	30	20	27	HV3_35_73	5.9	16	HV3_27_96
2	1	1	1	2	0	0	0.78	13	1670	1	-1.32	0	1	<20	20	HV3_35_32	3.7	14	HV3_27_86
1	1	2	1	3	0	0	16	296	1850	1	-0.46	0	5	50	27	HV3_35_74	10.0	14	HV3_27_87
2	0	0	1	2	3	1	16.41	147	896	2	0.32	1	1	<20	27	HV3_35_75	1.9	14	HV3_27_88
										1	-1.87	0		20	27	HV3_35_83			
3	1	2	1	2	0	0	15.56	239	1536	1	-0.41	0	10	100	27	HV3_35_76	10.0	15	HV3_27_89
2	1	1	1	1	0	0	24.47	391	1598	2	1.57	1	7	50	27	HV3_35_77	10.0	15	HV3_27_91
2	1	1	0	0	3	0	20.84	373	1790	2	0.38	1	9	100	27	HV3_35_78	10.0	15	HV3_27_92

Supplementary Table 2: List of oligonucleotide probes significantly differentially expressed between DCIS lesions associated with grade 1 and grade 3 invasive cancer

Rank	oligo	symbol_from_entrez_gene	description_from_entrez_gene	chrom	unigene	M	A	t	P.Value	adj.P.Val	B
1	H300003937	EME1	essential meiotic endonuclease 1 homolog 1 (S. pombe)	17	Hs.514330	3.206	11.028	7.306	0.000	0.0002	10.107
2	H200010726	E2F1	E2F transcription factor 1	20	Hs.558118||	2.321	13.643	6.738	0.000	0.0005	8.447
3	H300004854	NA	NA	15	NA	2.113	10.082	6.684	0.000	0.0005	8.289
4	H300020138	FLJ10719	hypothetical protein FLJ10719	15	Hs.513126	1.451	11.746	6.276	0.000	0.0014	7.080
5	H200012382	GTSE1	G-2 and S-phase expressed 1	22	Hs.386189	2.272	9.529	6.144	0.000	0.0014	6.688
6	H300001120	C8orf38	chromosome 8 open reading frame 38	8	NA	1.874	11.119	6.113	0.000	0.0014	6.597
7	H300006537	NA	NA	15	Hs.112160	2.289	8.769	6.051	0.000	0.0014	6.413
8	H200011914	CDC45L	CDC45 cell division cycle 45-like (S. cerevisiae)	22	NA	1.828	5.676	6.166	0.000	0.0014	6.328
9	H300003796	BIRC5	baculoviral IAP repeat-containing 5 (survivin)	17	Hs.514527	2.198	13.796	6.002	0.000	0.0015	6.266
10	H200015433	E2F2	E2F transcription factor 2	1	Hs.194333	2.236	11.184	5.970	0.000	0.0015	6.170
11	H200019974	RAD51	RAD51 homolog (RecA homolog, E. coli) (S. cerevisiae)	15	NA	1.707	12.069	5.921	0.000	0.0016	6.026
12	H300022790	ADC	arginine decarboxylase	1	Hs.101807	2.400	6.926	5.892	0.000	0.0016	5.929
13	H200012628	DTL	denticleless homolog (Drosophila)	1	Hs.126774	1.422	10.725	5.874	0.000	0.0016	5.886
14	H200006808	MKI67	antigen identified by monoclonal antibody Ki-67	10	Hs.80976	2.392	11.832	5.850	0.000	0.0016	5.815
15	H200005673	KIF2C	kinesin family member 2C	1	Hs.69360	2.142	12.496	5.745	0.000	0.0021	5.502
16	H200014659	CCNB2	cyclin B2	15	NA	2.229	12.525	5.672	0.000	0.0024	5.285
17	H300002397	PTTG1	pituitary tumor-transforming 1	8	Hs.350966||	1.660	12.403	5.667	0.000	0.0024	5.270
18	H200008012	TROAP	trophinin associated protein (tastin)	12	Hs.524399	1.287	7.461	5.651	0.000	0.0024	5.221
19	H200010755	TM4SF19	transmembrane 4 L six family member 19	3	Hs.135997	1.371	9.383	5.591	0.000	0.0026	5.045
20	H300010830	ISG20L1	interferon stimulated exonuclease gene 20kDa-like 1	15	Hs.436102||	1.544	10.715	5.581	0.000	0.0026	5.014
21	H200015825	C9orf100	chromosome 9 open reading frame 100	9	Hs.277026	1.920	8.700	5.571	0.000	0.0026	4.984

22	H200000307	BRRN1	barren homolog (Drosophila)	2	Hs.308045	1.466	8.934	5.563	0.000	0.0026	4.960
23	H200004844	EXO1	exonuclease 1	1	Hs.498248	2.453	11.054	5.553	0.000	0.0026	4.932
24	H200000064	FOXM1	forkhead box M1	12	Hs.239||Hs.	1.314	11.322	5.491	0.000	0.0029	4.747
25	H200005853	KIF20A	kinesin family member 20A	5	Hs.73625	2.139	13.401	5.490	0.000	0.0029	4.745
26	H200008454	AURKB	aurora kinase B	17	NA	1.822	12.400	5.486	0.000	0.0029	4.732
27	H200013545	MGC2408	hypothetical protein MGC2408	3	NA	1.624	10.382	5.467	0.000	0.0029	4.675
28	H200004950	ORC6L	origin recognition complex, subunit 6 homolog-like (yeast)	16	Hs.49760	1.424	8.539	5.463	0.000	0.0029	4.662
29	H300007595	PTTG1	pituitary tumor-transforming 1	4	Hs.350966	1.673	12.615	5.433	0.000	0.0031	4.574
30	H300013861	NA	NA	12	Hs.37044	1.703	9.302	5.349	0.000	0.0038	4.328
31	H300013227	MYBL2	v-myb myeloblastosis viral oncogene homolog (avian)-like 2	20	Hs.179718	2.836	12.649	5.343	0.000	0.0038	4.307
32	H300020211	KIF2C	kinesin family member 2C	1	Hs.69360	1.824	11.182	5.331	0.000	0.0038	4.274
33	H200017540	NUSAP1	nucleolar and spindle associated protein 1	15	Hs.511093	2.165	12.775	5.322	0.000	0.0038	4.247
34	H200006072	RRM2	ribonucleotide reductase M2 polypeptide	2	Hs.226390	2.148	13.435	5.311	0.000	0.0038	4.213
35	H200008371	MYBL2	v-myb myeloblastosis viral oncogene homolog (avian)-like 2	20	Hs.179718	2.898	11.737	5.309	0.000	0.0038	4.207
36	H300012511	NA	NA	12	NA	0.839	9.340	5.290	0.000	0.0040	4.152
37	H300002127	C8orf16	chromosome 8 open reading frame 16	8	Hs.59115	-1.984	7.950	-5.244	0.000	0.0045	4.017
38	H300008715	NA	NA	11	Hs.375110	-0.632	6.980	-5.227	0.000	0.0046	3.966
39	H300001422	CDT1	DNA replication factor	16	Hs.122908	1.341	9.928	5.219	0.000	0.0046	3.944
40	H200007139	CCNA2	cyclin A2	4	Hs.85137	2.004	12.061	5.173	0.000	0.0051	3.807
41	H200010558	UBE2C	ubiquitin-conjugating enzyme E2C	20	Hs.93002	1.076	12.574	5.171	0.000	0.0051	3.800
42	H200003981	CDCA3	cell division cycle associated 3	7	NA	1.295	7.799	5.133	0.000	0.0057	3.689
43	H300006651	DEFB127	defensin, beta 127	20	Hs.99362	1.435	8.040	5.114	0.000	0.0059	3.632
44	H200003232	CDCA5	cell division cycle associated 5	11	Hs.434886	1.257	12.841	5.132	0.000	0.0061	3.565
45	H300001057	NA	NA	20	NA	1.691	7.516	5.042	0.000	0.0071	3.421
46	H300005482	NUDT1	nudix (nucleoside diphosphate linked moiety X)-type motif 1	7	NA	1.202	12.823	5.035	0.000	0.0071	3.401

47	H200001767	TPX2	TPX2, microtubule-associated, homolog (Xenopus laevis)	20	Hs.244580	1.882	12.817	5.024	0.000	0.0071	3.368
48	H300008454	PTTG1	pituitary tumor-transforming 1	5	Hs.511755	1.468	9.866	5.021	0.000	0.0071	3.361
49	H200013932	SLC35B1	solute carrier family 35, member B1	17	NA	1.708	13.987	4.986	0.000	0.0078	3.257
50	H300009764	FLJ34790	hypothetical protein FLJ34790	17	Hs.121692	-0.936	10.695	-4.981	0.000	0.0078	3.244
51	H200007186	FAM64A	family with sequence similarity 64, member A	17	Hs.404323	1.765	8.583	4.957	0.000	0.0082	3.174
52	H300012966	SLC35B4	solute carrier family 35, member B4	7	Hs.490181	1.150	11.709	4.955	0.000	0.0082	3.167
53	H300000019	ASCL1	achaete-scute complex-like 1 (Drosophila)	12	Hs.524672	1.271	8.308	4.943	0.000	0.0083	3.131
54	H300006630	GPR35	G protein-coupled receptor 35	2	Hs.239891	1.074	7.547	4.926	0.000	0.0086	3.084
55	H300015294	NA	NA	12	NA	1.001	11.421	4.817	0.000	0.0121	2.765
56	H200003654	ASF1B	ASF1 anti-silencing function 1 homolog B (S. cerevisiae)	19	Hs.558081	1.299	9.405	4.804	0.000	0.0124	2.727
57	H200002977	SPBC25	spindle pole body component 25 homolog (S. cerevisiae)	2	Hs.421956	1.074	10.331	4.790	0.000	0.0127	2.686
58	H300021897	MGC3130	hypothetical protein MGC3130	17	Hs.437059	1.301	8.053	4.774	0.000	0.0131	2.640
59	H300004086	LOC440406||U	similar to Ubiquitin-conjugating enzyme E2S (Ubiquitin-conjugating enzyme E2-24 kDa) (Ubiquitin-protein ligase) (Ubiquitin carrier protein) (E2-EPF5)||ubiquitin-conjugating enzyme E2S	17	Hs.462306	1.634	8.559	4.761	0.000	0.0133	2.601
60	H200000012	CEACAM8	carcinoembryonic antigen-related cell adhesion molecule 8	19	Hs.41	0.763	9.121	4.760	0.000	0.0133	2.601
61	H300000739	MGC40170	hypothetical LOC253128	12	Hs.408739	0.859	7.236	4.741	0.000	0.0139	2.544
62	H300013069	ANLN	anillin, actin binding protein (scraps homolog, Drosophila)	7	Hs.62180	1.396	7.198	4.685	0.000	0.0155	2.383
63	H300000453	Pfs2	DNA replication complex GINS protein PSF2	16	NA	1.647	13.382	4.680	0.000	0.0155	2.369

64	H200007088	CDKN3	cyclin-dependent kinase inhibitor 3 (CDK2-associated dual specificity phosphatase)	14	Hs.84113	2.018	12.229	4.679	0.000	0.0155	2.365
65	H200014800	PKM2	pyruvate kinase, muscle	15	Hs.198281||	1.229	11.090	4.673	0.000	0.0155	2.348
66	H300013182	SPAG5	sperm associated antigen 5	17	Hs.514033	1.799	12.280	4.670	0.000	0.0155	2.338
67	H300018527	MCM10	MCM10 minichromosome maintenance deficient 10 (S. cerevisiae)	10	Hs.198363	2.124	9.354	4.669	0.000	0.0155	2.335
68	H300022393	CDC25C	cell division cycle 25C	5	Hs.558020	2.039	11.793	4.668	0.000	0.0155	2.334
69	H300022231	MYCN	v-myc myelocytomatosis viral related oncogene, neuroblastoma derived (avian)	2	Hs.25960	1.045	7.529	4.670	0.000	0.0155	2.333
70	H300022044	ASPM	asp (abnormal spindle)-like, microcephaly associated (Drosophila)	1	Hs.121028	2.634	10.095	4.669	0.000	0.0159	2.301
71	H300013149	CDKN3	cyclin-dependent kinase inhibitor 3 (CDK2-associated dual specificity phosphatase)	14	Hs.84113	1.646	8.237	4.623	0.000	0.0174	2.204
72	H300001740	ZWINT	ZW10 interactor	10	Hs.42650	1.860	9.197	4.610	0.000	0.0175	2.167
73	H200013875	ESPL1	extra spindle poles like 1 (S. cerevisiae)	12	Hs.153479	0.700	12.199	4.604	0.000	0.0175	2.151
74	H300008186	SNAI1	snail homolog 1 (Drosophila)	20	Hs.48029	1.197	10.426	4.603	0.000	0.0175	2.146
75	H200002555	SPAG5	sperm associated antigen 5	17	Hs.514033	2.034	11.071	4.601	0.000	0.0175	2.140
76	H200011605	ZNF536	zinc finger protein 536	19	Hs.378901	0.667	6.555	4.600	0.000	0.0175	2.137
77	H200004431	MLF1IP	MLF1 interacting protein	4	Hs.481307	1.828	11.960	4.596	0.000	0.0175	2.128
78	H200001011	UBE2T	ubiquitin-conjugating enzyme E2T (putative)	1	NA	1.587	13.188	4.568	0.000	0.0187	2.045
79	H300021654	TACC3	transforming, acidic coiled-coil containing protein 3	4	Hs.104019	1.311	11.308	4.567	0.000	0.0187	2.044
80	H300019472	DCC1	defective in sister chromatid cohesion homolog 1 (S. cerevisiae)	8	Hs.315167	1.603	9.457	4.533	0.000	0.0206	1.946
81	H200015853	PCGF5	polycomb group ring finger 5	10	NA	1.087	6.849	4.520	0.000	0.0212	1.909
82	H200008314	TH	tyrosine hydroxylase	11	Hs.435609	1.890	5.732	4.516	0.000	0.0212	1.830
83	H300016689	UBE2C	ubiquitin-conjugating enzyme E2C	20	Hs.93002	1.447	7.273	4.500	0.000	0.0218	1.852
84	H300003832	FLJ38723	hypothetical protein FLJ38723	15	Hs.370479	2.392	7.193	4.496	0.000	0.0218	1.842
85	H200002312	STARD13	START domain containing 13	13	Hs.507704	-1.756	10.138	-4.506	0.000	0.0218	1.837
86	H300019042	ZWINT	ZW10 interactor	10	Hs.42650	1.347	11.403	4.493	0.000	0.0218	1.831

87	H300013620	FLJ20171	hypothetical protein FLJ20171	8	Hs.487471	1.270	9.602	4.477	0.000	0.0225	1.785
88	H200005718	C20orf129	chromosome 20 open reading frame 129	20	NA	2.277	11.702	4.475	0.000	0.0225	1.780
89	H300003454	LOC138412||S	similar to ADP,ATP carrier protein, liver isoform T2 (ADP/ATP translocase 3) (Adenine nucleotide translocator 3) (ANT 3)||solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 6	9	Hs.534548	0.933	9.613	4.468	0.000	0.0227	1.761
90	H300000073	C18orf56	chromosome 18 open reading frame 56	18	NA	1.247	9.546	4.452	0.000	0.0237	1.715
91	H300002712	LOC283299	hypothetical protein LOC283299	11	Hs.501825	-1.669	6.941	-4.446	0.000	0.0238	1.698
92	H200001248	GPR56	G protein-coupled receptor 56	16	Hs.513633	1.066	11.795	4.434	0.000	0.0238	1.664
93	H300002144	APOBEC3B	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3B	22	Hs.226307	1.249	11.695	4.433	0.000	0.0238	1.661
94	H200016159	NA	NA	6	Hs.555894	1.176	7.482	4.432	0.000	0.0238	1.658
95	H200006228	PGD	phosphogluconate dehydrogenase	1	Hs.464071	1.328	10.860	4.427	0.000	0.0238	1.644
96	H200011999	OIP5	Opa interacting protein 5	15	NA	0.896	8.720	4.426	0.000	0.0238	1.642
97	H300002100	DERL1	Der1-like domain family, member 1	8	Hs.241576	1.249	8.553	4.422	0.000	0.0238	1.631
98	H300015925	DNAH8	dynein, axonemal, heavy polypeptide 8	6	Hs.520106	0.889	13.014	4.421	0.000	0.0238	1.626
99	H200004071	MKL1	megakaryoblastic leukemia (translocation) 1	22	Hs.474905	0.799	9.985	4.420	0.000	0.0238	1.623
100	H300016661	WDR68	WD repeat domain 68	17	Hs.410596	0.926	8.820	4.405	0.000	0.0247	1.581
101	H200013113	NEIL3	nei endonuclease VIII-like 3 (E. coli)	4	Hs.405467	1.127	9.840	4.397	0.000	0.0251	1.559
102	H300004556	NRM	nurim (nuclear envelope membrane protein)	6	Hs.519993	1.537	10.053	4.390	0.000	0.0253	1.539
103	H200009956	KIF18A	kinesin family member 18A	11	Hs.301052	1.248	9.717	4.386	0.000	0.0253	1.527
104	H300015612	NA	NA	3	NA	0.902	11.367	4.385	0.000	0.0253	1.525
105	H300018824	FLJ23554	hypothetical protein FLJ23554	11	Hs.164705	-1.371	7.579	-4.382	0.000	0.0253	1.517

106	H300017497	PGD	phosphogluconate dehydrogenase	1	Hs.464071	0.897	14.737	4.374	0.000	0.0257	1.493
107	H200017889	SPON2	spondin 2, extracellular matrix protein	4	NA	-0.415	7.061	-4.359	0.000	0.0266	1.452
108	H300017624	NA	NA	20	NA	1.914	4.614	4.914	0.000	0.0266	0.019
109	H300004245	C6orf173	chromosome 6 open reading frame 173	6	NA	1.375	11.217	4.348	0.000	0.0271	1.421
110	H200000213	SOX2	SRY (sex determining region Y)-box 2	3	Hs.518438	1.632	9.614	4.341	0.000	0.0273	1.401
111	H200010931	ATP6V0A4	ATPase, H+ transporting, lysosomal V0 subunit a isoform 4	7	Hs.98967	2.305	9.694	4.340	0.000	0.0273	1.399
112	H300018716	FLJ20171	hypothetical protein FLJ20171	8	Hs.487471	1.330	10.654	4.326	0.000	0.0283	1.358
113	H200006364	CENPF	centromere protein F, 350/400ka (mitosin)	1	Hs.497741	1.635	13.014	4.318	0.000	0.0283	1.337
114	H300021140	RECQL4	RecQ protein-like 4	8	Hs.31442	1.584	13.026	4.318	0.000	0.0283	1.335
115	H200007867	TTK	TTK protein kinase	6	Hs.169840	1.438	10.324	4.317	0.000	0.0283	1.333
116	H200019588	CDC2	cell division cycle 2, G1 to S and G2 to M	10	Hs.334562	1.667	8.993	4.310	0.000	0.0285	1.313
117	H200017465	PIR	pirin (iron-binding nuclear protein)	23	NA	1.858	11.959	4.309	0.000	0.0285	1.310
118	H300005079	NA	NA	8	Hs.291342	-1.827	6.064	-4.307	0.000	0.0285	1.287
119	H300003913	TOMM34	translocase of outer mitochondrial membrane 34	20	Hs.517066	1.020	8.340	4.300	0.000	0.0289	1.285
120	H200008233	NA	NA	1	NA	1.471	9.581	4.292	0.000	0.0293	1.264
121	H200017742	PRO1855	hypothetical protein PRO1855	17	Hs.370927	1.425	14.644	4.289	0.000	0.0294	1.254
122	H300006435	NRM	nurim (nuclear envelope membrane protein)	6	Hs.519993	1.445	9.642	4.282	0.000	0.0298	1.234
123	H300017569	FANCD2	Fanconi anemia, complementation group D2	3	Hs.208388	0.712	9.893	4.264	0.000	0.0312	1.185
124	H200003347	CCNB1	cyclin B1	5	Hs.23960	1.224	10.427	4.253	0.000	0.0319	1.152
125	H300017969	Kua	ubiquitin-conjugating enzyme variant Kua	20	NA	1.059	13.593	4.253	0.000	0.0319	1.152
126	H200006074	PHB	prohibitin	17	NA	1.661	9.488	4.246	0.000	0.0323	1.133
127	H200009953	CABLES2	Cdk5 and Abl enzyme substrate 2	20	Hs.301040	1.389	12.897	4.233	0.000	0.0333	1.098
128	H200004242	ISYNA1	myo-inositol 1-phosphate synthase A1	19	Hs.405873||	1.316	11.177	4.225	0.000	0.0337	1.075
129	H300003438	HSHIN1	HIV-1 induced protein HIN-1	12	Hs.270851	0.626	6.956	4.225	0.000	0.0337	1.074

130	H200006619	MAD2L1	MAD2 mitotic arrest deficient-like 1 (yeast)	4	NA	1.714	10.335	4.215	0.000	0.0343	1.048
131	H300005780	E2F4	E2F transcription factor 4, p107/p130-binding	16	Hs.108371	1.020	7.489	4.215	0.000	0.0343	1.046
132	H200014130	CLUAP1	clusterin associated protein 1	16	Hs.155995	-1.328	12.210	-4.203	0.000	0.0351	1.013
133	H300009402	VPS13C	vacuolar protein sorting 13C (yeast)	15	Hs.511668||	0.603	6.940	4.202	0.000	0.0351	1.011
134	H200006714	TNFRSF19L	tumor necrosis factor receptor superfamily, member 19-like	11	Hs.533720||	0.998	7.280	4.198	0.000	0.0353	0.999
135	H300021601	FANCB	Fanconi anemia, complementation group B	23	Hs.445921||	0.961	9.292	4.194	0.000	0.0354	0.989
136	H200006122	NP	nucleoside phosphorylase	14	Hs.75514	0.929	9.102	4.187	0.000	0.0359	0.970
137	H300020810	TPCN2	two pore segment channel 2	11	Hs.387396	1.448	9.654	4.183	0.000	0.0361	0.959
138	H200014576	HCAP-G	chromosome condensation protein G	4	Hs.479270	1.549	10.415	4.181	0.000	0.0361	0.951
139	H300020785	RAD52B	RAD52 homolog B (S. cerevisiae)	17	Hs.194411	2.297	10.666	4.183	0.000	0.0367	0.934
140	H200009627	GAJ	GAJ protein	4	Hs.294088	1.240	8.772	4.166	0.000	0.0372	0.911
141	H200006371	EZH2	enhancer of zeste homolog 2 (Drosophila)	7	NA	0.803	9.451	4.155	0.000	0.0378	0.881
142	H200020461	RAF1	v-raf-1 murine leukemia viral oncogene homolog 1	3	Hs.159130	0.593	10.981	4.153	0.000	0.0378	0.875
143	H300004838	ZFP276	zinc finger protein 276 homolog (mouse)	16	Hs.290154	1.713	11.426	4.152	0.000	0.0378	0.871
144	H300000820	FLJ23657	hypothetical protein FLJ23657	4	Hs.24510	2.287	4.879	4.694	0.000	0.0378	-0.038
145	H300021996	DEPDC1	DEP domain containing 1	1	Hs.445098	1.131	10.100	4.134	0.000	0.0397	0.821
146	H300005578	RPL23	ribosomal protein L23	17	Hs.512542	-0.694	15.173	-4.128	0.000	0.0402	0.804
147	H300001915	ADCY3||MGC1	adenylate cyclase 3||hypothetical protein MGC11266	2	Hs.546428	1.100	12.013	4.125	0.000	0.0403	0.796
148	H200014873	C14orf105	chromosome 14 open reading frame 105	14	Hs.146040	0.749	8.362	4.122	0.000	0.0403	0.789
149	H200005265	DOC2A	double C2-like domains, alpha	16	Hs.355281	-2.418	11.314	-4.120	0.000	0.0403	0.783
150	H200008215	NA	NA	3	Hs.506115	-1.415	6.895	-4.106	0.000	0.0419	0.743
151	H300018882	DAP	death-associated protein	5	Hs.75189	0.773	9.406	4.099	0.000	0.0425	0.725
152	H300015276	BM039	uncharacterized bone marrow protein BM039	16	NA	1.052	10.374	4.092	0.000	0.0430	0.706

153	H200014581	GSG2	germ cell associated 2 (haspin)	17	NA	1.475	8.746	4.090	0.000	0.0430	0.700
154	H300012266	MGC3130	hypothetical protein MGC3130	17	Hs.437059	0.963	9.955	4.098	0.000	0.0430	0.700
155	H300016532	UBE1	ubiquitin-activating enzyme E1 (A1S9T and BN75 temperature sensitivity complementing)	23	NA	0.507	10.255	4.086	0.000	0.0431	0.688
156	H200015577	HOXD9	homeo box D9	2	Hs.236646	0.765	7.665	4.081	0.000	0.0435	0.675
157	H200016261	STK6	serine/threonine kinase 6	20	Hs.250822	1.013	13.012	4.065	0.000	0.0453	0.632
158	H300021781	CENPA	centromere protein A, 17kDa	2	NA	1.436	10.905	4.058	0.000	0.0461	0.611
159	H300022861	ISGF3G	interferon-stimulated transcription factor 3, gamma 48kDa	14	Hs.1706	0.965	10.506	4.054	0.000	0.0464	0.599
160	H200012451	SHCBP1	SHC SH2-domain binding protein 1	16	Hs.123253	1.360	10.649	4.052	0.000	0.0464	0.595
161	H300018657	C14orf93	chromosome 14 open reading frame 93	14	Hs.255874	0.704	9.762	4.048	0.000	0.0467	0.583
162	H300002155	K-ALPHA-1||TU	tubulin alpha 6||tubulin, alpha 3||tubulin, alpha, ubiquitous	11	Hs.524390	1.070	13.576	4.046	0.000	0.0467	0.578
163	H300006913	UBE2S	ubiquitin-conjugating enzyme E2S	17	Hs.462306	1.681	11.387	4.040	0.000	0.0470	0.562
164	H200014857	EPN3	epsin 3	17	Hs.165904	1.299	7.471	4.040	0.000	0.0470	0.561
165	H200007997	PVR	poliovirus receptor	19	Hs.171844	0.988	8.420	4.036	0.000	0.0472	0.550
166	H300005525	RRM2	ribonucleotide reductase M2 polypeptide	23	Hs.226390	0.719	8.757	4.032	0.000	0.0475	0.542
167	H200000527	UBE1	ubiquitin-activating enzyme E1 (A1S9T and BN75 temperature sensitivity complementing)	23	Hs.533273	0.901	12.591	4.015	0.000	0.0496	0.492
168	H300020183	TOP2A	topoisomerase (DNA) II alpha 170kDa	17	Hs.156346	1.050	7.226	4.013	0.000	0.0496	0.487
169	H300005394	FTL	ferritin, light polypeptide	20	Hs.495896	0.852	14.247	4.019	0.000	0.0496	0.482
170	H300004609	FLJ35801	hypothetical protein FLJ35801	22	Hs.120446	1.102	8.531	4.009	0.000	0.0496	0.478
171	H200015043	C22orf18	chromosome 22 open reading frame 18	22	Hs.208912	0.891	9.600	4.006	0.000	0.0496	0.469

172	H300021877	DDX11||DDX1	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 11 (CHL1-like helicase homolog, S. cerevisiae)||DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 12 (CHL1-like helicase homolog, S. cerevisiae)	12	Hs.504688	1.035	9.192	4.005	0.000	0.0496	0.465
173	H200011199	IL1RL2	interleukin 1 receptor-like 2	2	Hs.469520	0.458	8.593	4.005	0.000	0.0496	0.465

Column Name	Description
oligo	Oligo identifier from manufacturer
symbol_from_entrez_gene	The gene symbol extracted from NCBI (should be the HUGO gene name in most cases)
unigene	Unigene identifier
M	Estimated log2(fold change) for ER+ versus ER- (for example)
A	log2(average intensity) for the probe (reported but not used)
t	The t-statistic (moderated) (reported but not used)
P.Value	The raw p-value (reported but not used)
adj.P.Val	The false-discovery-rate (FDR), so corrected for multiple testing
B	The log(odds) of differential expression (reported, but not used)

Supplementary Table 3: Gene ontology enrichment analysis of differentially expressed genes

System	Category	EASE Score (P-value)[a]	List Hits (LH)[b]	List Total (LT)[c]	Population Hits (PH)[d]	Population Total (PT)[e]
Biological Process						
	mitotic cell cycle	1.36E-18	25	77	266	8700
	cell cycle	3.71E-17	31	77	550	8700
	mitosis	4.91E-15	16	77	101	8700
	M phase of mitotic cell cycle	6.65E-15	16	77	103	8700
	nuclear division	9.11E-15	17	77	128	8700
	M phase	1.79E-14	17	77	133	8700
	cell proliferation	3.47E-13	32	77	826	8700
	regulation of cell cycle	7.17E-09	17	77	310	8700
	cell growth and/or maintenance	1.19E-08	48	77	2639	8700
	cellular physiological process	0.00000019	50	77	3070	8700
	DNA metabolism	0.00000123	16	77	395	8700
	regulation of mitosis	0.00000166	6	77	24	8700
	DNA replication and chromosome cycle	0.0000038	10	77	143	8700
	cytokinesis	0.00000608	8	77	81	8700
	cellular process	0.0000797	59	77	4767	8700
	G2/M transition of mitotic cell cycle	0.000468	5	77	42	8700
	cell cycle checkpoint	0.0022	4	77	30	8700
	response to DNA damage stimulus	0.00235	7	77	156	8700
	response to endogenous stimulus	0.00275	7	77	161	8700
	chromosome condensation	0.0047	3	77	12	8700
	DNA repair	0.00546	6	77	130	8700
	G1 phase of mitotic cell cycle	0.0064	3	77	14	8700
	DNA replication	0.0155	5	77	110	8700
	S phase of mitotic cell cycle	0.0159	5	77	111	8700
	regulation of CDK activity	0.0392	3	77	36	8700
	transcription from Pol II promoter	0.0464	8	77	378	8700
Cellular Component						
	microtubule cytoskeleton	0.00000289	10	67	144	7813
	nucleus	0.0000161	37	67	2298	7813
	spindle	0.000249	5	67	37	7813
	intracellular	0.000273	53	67	4519	7813
	chromosome, pericentric region	0.000592	4	67	20	7813
	chromosome	0.00829	6	67	149	7813
	cytoskeleton	0.0179	10	67	476	7813
	microtubule associated complex	0.0279	4	67	78	7813

Molecular Function						
	ATP binding	0.0264	13	68	900	9264
	DNA binding	0.0275	18	68	1457	9264
	adenyl nucleotide binding	0.0287	13	68	911	9264
	motor activity	0.0482	4	68	113	9264

[a] EASE Score (P-value) represents the level of confidence that this term is over-represented in the DCIS discriminative gene list

[b] LH - number of genes with this term in the DCIS discriminative gene list

[c] LT - number of genes in the DCIS discriminative gene list mapped to any term in this ontology ("system")

[d] PH - number of genes with this gene ontology term on the background gene list (ie. the entire oligonucleotide microarray)

[e] PT - number of genes on the entire oligonucleotide microarray mapped to any term in this ontology ('system")