RESEARCH ARTICLE   **Open Access**

CrossMark

The long-term prognostic and predictive capacity of cyclin D1 gene amplification in 2305 breast tumours

Arian Lundberg[1], Linda S. Lindström[2], Jingmei Li[3,4], J. Chuck Harrell[5], Eva Darai-Ramqvist[6], Emmanouil G. Sifakis[1], Theodoros Foukakis[1], Charles M. Perou[7], Kamila Czene[3], Jonas Bergh[1,8] and Nicholas P. Tobin[1*] ![ORCID]

Abstract

Background: Use of cyclin D1 (*CCND1*) gene amplification as a breast cancer biomarker has been hampered by conflicting assessments of the relationship between cyclin D1 protein levels and patient survival. Here, we aimed to clarify its prognostic and treatment predictive potential through comprehensive long-term survival analyses.

Methods: *CCND1* amplification was assessed using SNP arrays from two cohorts of 1965 and 340 patients with matching gene expression array and clinical follow-up data of over 15 years. Kaplan-Meier and multivariable Cox regression analyses were used to determine survival differences between *CCND1* amplified vs. non-amplified tumours in clinically relevant patient sets, within PAM50 subtypes and within treatment-specific subgroups. Boxplots and differential gene expression analyses were performed to assess differences between amplified vs. non-amplified tumours within PAM50 subtypes.

Results: When combining both cohorts, worse survival was found for patients with *CCND1*-amplified tumours in luminal A (HR = 1.68; 95% CI, 1.15–2.46), luminal B (1.37; 1.01–1.86) and ER+/LN–/HER2– (1.66; 1.14–2.41) subgroups. In gene expression analysis, *CCND1*-amplified luminal A tumours showed increased proliferation ($P < 0.001$) and decreased progesterone ($P = 0.002$) levels along with a large overlap in differentially expressed genes when comparing luminal A and B-amplified vs. non-amplified tumours.

Conclusions: Our results indicate that *CCND1* amplification is associated with worse 15-year survival in ER+/LN–/HER2–, luminal A and luminal B patients. Moreover, luminal A *CCND1*-amplified tumours display gene expression changes consistent with a more aggressive phenotype. These novel findings highlight the potential of *CCND1* to identify patients that could benefit from long-term treatment strategies.

Keywords: CCND1 gene amplification, Gene expression, Breast cancer, PAM50, Luminal A, CDK4/6, BCSS

Background

Mammalian cyclin D1, first identified in 1991 [1, 2], mediates G1 to S-phase transition in the cell cycle along with its binding partners CDK4/6. Overexpression of its protein has been found in 50–70% of breast cancers [3–6] whilst amplification of its corresponding gene, *CCND1*, has been shown in approximately 9–30% of cases [7–11]. These figures indicate that processes other than gene amplification are also responsible for overexpression of the protein.

The *CCND1* gene maps to the 11q13 breast cancer risk locus and the majority of breast tumours bearing amplification of the gene are oestrogen receptor (ER) positive [12–15], are of luminal B subtype [12, 13], and overexpress cyclin D1 protein [9, 12, 14, 15]. Most notably, patients with ER-positive *CCND1*-amplified tumours also show reduced survival times [9, 10, 12, 15]. The use of *CCND1* amplification as a biomarker in a clinical setting has been hampered by conflicting assessments of the relationship between cyclin D1 protein levels and clinico-pathological parameters. To be explicit, overexpression of cyclin D1 protein has been linked to both better [9, 16–18] and worse prognosis [19, 20] in

* Correspondence: nick.tobin@ki.se
[1]Department of Oncology and Pathology, Karolinska Institutet and University Hospital, Stockholm, Sweden
Full list of author information is available at the end of the article

Lundberg *et al. Breast Cancer Research* (2019) 21:34

Page 2 of 10

breast cancer patients. These results are in clear contrast to the consistent prognostic message of tumour aggressively and reduced survival provided by *CCND1* amplification.

Among current research needs in breast cancer, biomarkers capable of helping to predict late recurrences are urgently needed [21]. In this regard, molecular biomarkers and in particular gene expression signatures may provide some utility [22–24]. Here, we aimed to determine if amplification of *CCND1* can function as a molecular biomarker for long-term breast cancer survival and more generally to comprehensively characterise its prognostic and treatment predictive capacity. In order to achieve this, we performed an integrative analysis combining gene amplification, gene expression and clinico-pathological data in two large cohorts of 1965 and 340 patients, respectively, with over 15 years of follow-up. We focus on clinically relevant patient subgroups including all, ER-positive/lymph node negative/human epidermal growth factor receptor 2 negative (ER +/LN−/HER2−), ER+/LN+/HER2−, HER2+ and triple negative breast cancers (TNBCs), as well as the PAM50 subtypes (luminal A, luminal B, HER2-enriched and basal-like) and treatment-specific subgroups (patients who received chemotherapy, endocrine therapy, both sequentially or untreated).

Materials and methods
Study population and specimens
Cohort 1 is comprised of tumours from the METABRIC study, and patient/tumour characteristics, treatments received and clinical endpoints have been previously described in detail [25, 26]. Briefly, this cohort consists of a total of 1992 primary breast cancers from patients in the UK and Canada with a median follow-up accounting for censoring of 10.2 years. Of the original 1992 patients, 1965 were included in our analysis and reasons for exclusion were duplicate samples ($n = 12$) or unclassified tumours ($n = 15$, ER−/PR+/HER2− tumours). METABRIC clinical and genomic data are publicly available from the EGA-archive (https://ega-archive.org) under study number EGAS00000000083.

Cohort 2 has also been previously extensively described [23]. Briefly, this cohort was derived from a nested case–control study and consists of 621 individuals diagnosed with primary breast tumours between January 1, 1997, and December 31, 2005. Of these, 340 were included in our analysis and reasons for exclusion were bilateral tumours ($n = 2$), unclassified tumours ($n = 14$, ER−/PR +/HER2− tumours), no matching SNP array ($n = 68$) and missing clinico-pathological data ($n = 197$). Median follow-up in this cohort is 14.4 years and is complete to January 10, 2015. Follow-up information was retrieved from the Stockholm–Gotland Breast Cancer Registry

using a national registration number unique to all Swedish citizens. The clinical endpoint for both studies was breast cancer-specific survival (BCSS) defined as patients who have not died from breast cancer in the study period from the date of surgery to end of follow-up. Exclusion criteria for both cohorts are shown in the CONSORT diagram in Fig. 1.

ER, PR and HER2
Cohort 1
Tumour ER and HER2 status was taken directly from the clinico-pathological data of the METABRIC study and is immunohistochemistry (IHC) based [25, 26]. PR status is based on a gene expression classification as described in the original publication [25].

Cohort 2
Primary breast tumour ER and PR status (assessed by IHC, > 10% cut-off for positivity) was collected from pathology reports. HER2 status was determined using chromogenic in situ hybridization (CISH) [27] on tissue microarrays (TMAs) and scored by a breast cancer pathologist. This study was performed and reported in accordance with the REporting recommendations for tumour MARKer prognostic studies (REMARK) guidelines [28], and biospecimen reporting for improved study quality (BRISQ) criteria for this cohort have been previously published [23].

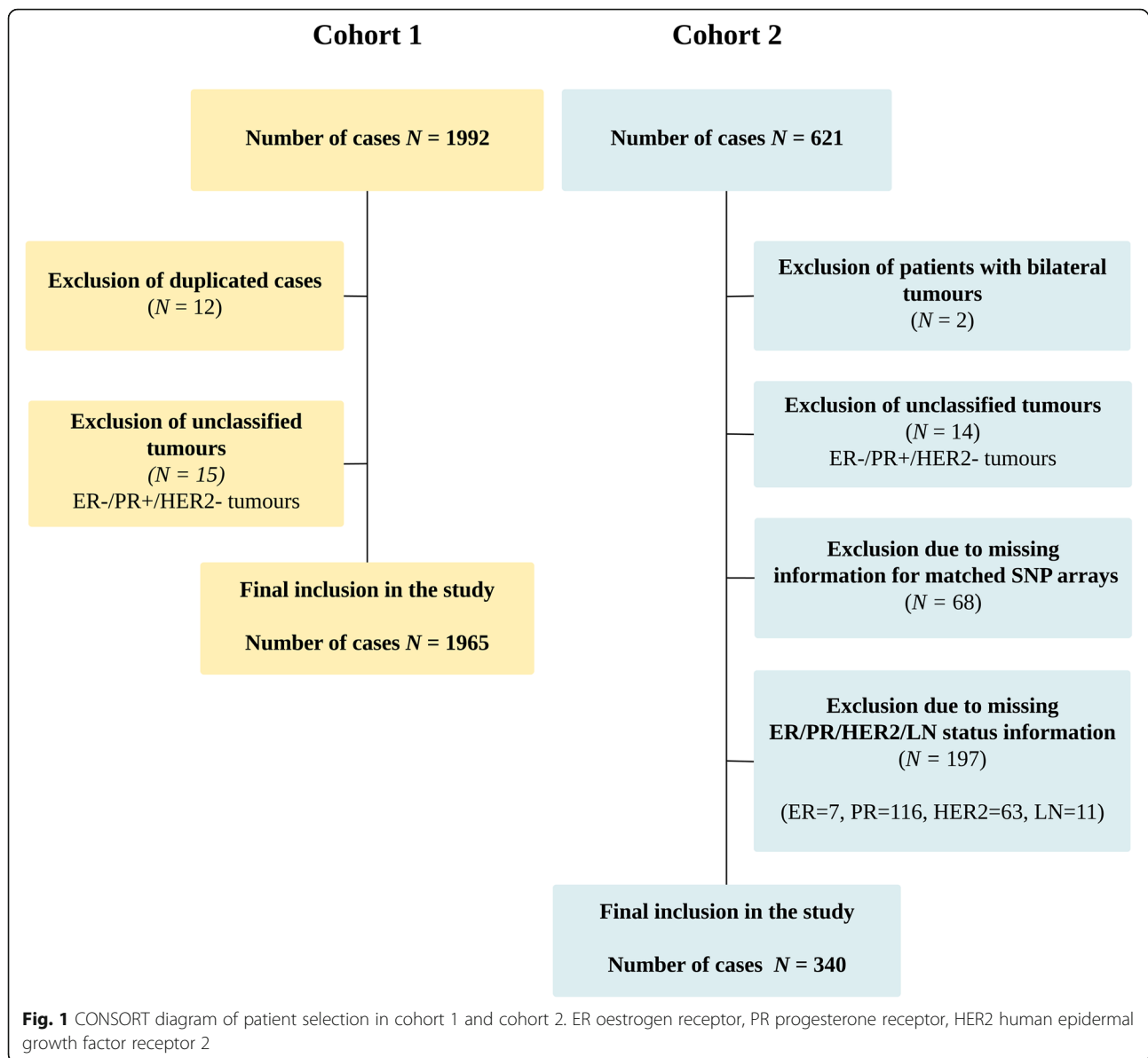### Genomic profiling
Cohort 1
Genomic profiling of METABRIC tumours was performed using whole genome gene expression (Illumina HT-12-v3 platform) and SNP arrays (Affymetrix SNP 6.0 platform) as detailed in the original publication [25].

Cohort 2

RNA gene expression profiling The Qiagen RNeasy Mini Kit (Qiagen, Germany) was used for RNA extraction from frozen tumours, and samples were subsequently profiled for whole genome gene expression through hybridization to the HRSTA-2.0 custom human Affymetrix array GPL10379; see further details here [23].

DNA genotyping Extraction was carried out using the QIAamp DNA mini kit (Qiagen, Germany) on frozen tumours, and genotyping was performed using the Human1M-Duo BeadChip (Illumina, CA, USA).

Both gene expression and genotyping studies in cohort 2 were approved by the ethics committee at Karolinska Institutet (Stockholm, Sweden).

Cohort 1   **Cohort 2**

Number of cases *N* = 1992   Number of cases *N* = 621

Exclusion of duplicated cases
(*N* = 12)

                                      Exclusion of patients with bilateral
                                      tumours
                                      (*N* = 2)

Exclusion of unclassified
tumours
(*N* = 15)
ER-/PR+/HER2- tumours

                                      Exclusion of unclassified tumours
                                      (*N* = 14)
                                      ER-/PR+/HER2- tumours

Final inclusion in the study

Number of cases *N* = 1965

                                      Exclusion due to missing
                                      information for matched SNP arrays
                                      (*N* = 68)

                                      Exclusion due to missing
                                      ER/PR/HER2/LN status information
                                      (*N* = 197)

                                      (ER=7, PR=116, HER2=63, LN=11)

                                      Final inclusion in the study

                                      Number of cases *N* = 340

Fig. 1 CONSORT diagram of patient selection in cohort 1 and cohort 2. ER oestrogen receptor, PR progesterone receptor, HER2 human epidermal growth factor receptor 2

CCND1 gene copy number analysis
Cohort 1
Normalised copy number variation (CNV) data and segmentation files (derived using circular binary segmentation—CBS) for METABRIC were downloaded from the EGA-archive.

Cohort 2
CNV data was generated from SNP array files using the CNVpartition (version 3.2.0) plugin from within GenomeStudio software (version 2011.1, Illumina, CA, USA). Similar to cohort 1, segmentation was performed using CBS as part of the R *DNAcopy* package [29]. Copy number alteration (CNA) calls for both cohorts were then derived by importing CNV, segmentation (derived from CBS) and platform-specific marker files to the Genomic Identification of Significant Targets in Cancer 2.0 (GISTIC) module [30]. GISTIC CNA amplification/deletion thresholds were set to ± 0.3, and regions with a *q* value of < 0.25 were considered significant. Of note, both of these cut-off values are in line with those used by The Cancer Genome Atlas for breast tumours [31]. *CCND1* amplification/deletion calls for each tumour in both cohorts were taken from the gene-level output files of the GISTIC algorithm and matched to clinico-pathological data.

Genomic classifiers and differential gene expression
Cohort 1
PAM50 molecular subtype calls for METABRIC were taken from the original publication. Normalised

Lundberg *et al. Breast Cancer Research*   (2019) 21:34

Page 4 of 10

microarray gene expression data for the METABRIC cohort was downloaded from the EGA-archive and subsequently used for differential gene expression (DGE) analysis. DGE was assessed between *CCND1*-amplified and non-amplified luminal A and luminal B tumours using the R package *Limma* [32], and only genes with an adjusted *P* value < 0.05 were considered significant.

Cohort 2
Microarray data was pre-processed and normalised using the R *aroma.affymetrix* package [33]; see further details here [23, 34]. Data has been deposited at the NCBI Gene Expression Omnibus under the accession number GSE48091. PAM50 was applied as described in the original publication [35].

Statistical analysis
All statistical analyses were performed using R statistical software version 3.4.3 [36]. To assess differences between clinico-pathological variables and *CCND1* amplified/non-amplified tumours statistical tests were chosen based on the variable class being compared: nominal versus nominal—χ^2; ordinal versus nominal—Mann–Whitney. Student's *t* test was used to test for differences in mean gene expression between *CCND1*-amplified vs. non-amplified tumours within PAM50 subtypes. Similarly, ANOVA with post hoc Tukey was used for determining differences in mean gene expression across PAM50 subtypes. All tests were two-sided, and a *P* value of < 0.05 was considered statistically significant. Kaplan–Meier analysis was performed for *CCND1*-amplified and non-amplified tumours with 15-year breast cancer-specific survival as the clinical endpoint. Similarly, Cox multivariable proportional hazard analyses were used to determine survival differences between *CCND1*-amplified and non-amplified tumours with the latter as reference group and the same clinical endpoint. Every multivariable analysis was adjusted for tumour size, tumour grade, nodal status, endocrine treatment, PAM50 subtype and patient cohort.

Results
Clinico-pathological characteristics of *CCND1*-amplified tumours
In keeping with our aim to comprehensively determine the long-term prognostic and treatment predictive capacity of *CCND1* amplification, we analysed two cohorts of 1965 and 340 breast cancer patients respectively, with matching gene expression, SNP arrays and long-term (15 years) clinical follow-up. A CONSORT diagram of exclusion criteria for both cohorts is shown in Fig. 1, and clinico-pathological characteristics for both cohort split by *CCND1* amplification status are shown in Table 1.

Twenty-two percent (426/1965, 22%) and 35% (119/340, 35%) of tumours in cohorts 1 and 2, respectively, were found to harbour *CCND1* amplifications, in line with previously published figures [7–11] (Table 1). Of note, as cohort 2 is enriched for patients with aggressive metastatic tumours, a higher percentage of *CCND1* amplifications was anticipated. In general, *CCND1*-amplified tumours were more likely to be ER-positive and of luminal subtype relative to non-amplified tumours (Table 1). It is, however, worth noting that amplified tumours were also present in ER-negative and non-luminal breast cancer subtypes.

CCND1 amplification predicts poor long-term survival in ER+ breast cancer patient subgroups (cohort 1)
Next, in cohort 1, we examined the relationship between *CCND1* amplification and long-term BCSS in clinically relevant patient subgroups defined by IHC/nodal status (all, ER+/LN–/HER2, ER+/LN+/HER2–, HER2+ or TNBCs), PAM50 gene expression subtype (luminal A/B, HER2-enriched or basal-like) or treatment received (endocrine therapy, chemotherapy, both sequentially or untreated). In Kaplan–Meier analysis *CCND1*-amplifed patients were found to have a worse 15-year BCSS in the IHC/nodal subgroups ER+/LN–/HER2– and ER+/LN +/HER2– (Additional file 1: Figure S1B and C, *P* < 0.001 and *P* = 0.016, respectively). Similarly, *CCND1* amplification was associated with poorer survival in luminal A—with a notable deviation between survival curves after 5 years, endocrine-treated and untreated breast cancer patients (Additional file 2: Figure S2A, E, and H, *P* = 0.019, 0.007 and 0.014, respectively). In multivariable Cox regression analysis, this statistical significance remained for ER+/LN–/HER2– (HR = 1.72, 95% CI, 1.14–2.59, Fig. 2a) patients only; however, trends were observed for luminal A (HR = 1.55, 95% CI, 0.99–2.45) and untreated subgroups (HR = 1.52, 95% CI, 0.95–2.44).

Similar results in a second smaller cohort (cohort 2) and when combining both cohorts (cohorts 1 + 2)
Similar trends were found for *CCND1*-amplified tumours in Kaplan–Meier (Additional file 3: Figure S3 and Additional file 4: Figure S4) and multivariable (Fig. 2b) analyses of a second independent cohort of 340 patients; however, smaller patient numbers resulted in wider confidence intervals and reduced statistical power. In multivariable analysis, patients with ER+/LN+/HER2– *CCND1*-amplified tumours were found to have worse 15-year BCSS (HR = 2.07, 95% CI, 1.12–3.83), and comparable trends were noted in luminal A (HR = 1.97, 95% CI, 0.95–4.08) and luminal B (HR = 2.01, 95% CI, 0.94–4.30) subgroups. *CCND1*-amplified ER+/LN –/HER2– patients did not show poorer survival in this cohort (HR = 1.07, 95% CI, 0.38–3.01) (Fig. 2b);

Table 1 Clinico-pathological characteristics of patients in cohort 1 and 2 split by *CCND1* amplification status

Variable	Cohort 1 (*n* = 1965)		*P*	Cohort 2 (*n* = 340)		*P*
	Non-Amp	Amp		Non-Amp	Amp	
	n (%)	*n* (%)		*n* (%)	*n* (%)	
	1539 (78)	426 (22)		221 (65)	119 (35)	
ER						
Positive	1129 (73)	377 (88)	< 0.001	158 (71)	93 (78)	0.182
Negative	410 (27)	49 (12)		63 (29)	26 (22)	
PR						
Positive	794 (52)	231 (54)	0.364	140 (63)	71 (60)	0.504
Negative	745 (48)	195 (46)		81 (37)	48 (40)	
HER2						
Positive	190 (12)	57 (13)	0.626	40 (18)	30 (25)	0.122
Negative	1349 (88)	369 (87)		181 (82)	89 (75)	
Elston–Ellis grade						
I	148 (10)	22 (5)	0.200*	22 (10)	9 (8)	0.400*
II	625 (43)	142 (34)		102 (47)	47 (40)	
III	690 (47)	252 (61)		93 (43)	62 (52)	
Missing cases = 86				Missing cases = 5		
LN status						
Positive	748 (49)	188 (44)	0.114	116 (52)	81 (68)	0.005
Negative	791 (51)	238 (56)		105 (48)	38 (32)	
Tumour size						
< 20 mm	497 (33)	120 (28)	0.108	101 (47)	51 (43)	0.492
≥ 20 mm	1026 (67)	303 (72)		115 (53)	68 (57)	
Missing cases = 19				Missing cases = 5		
Age						
≤ 45	208 (14)	41 (10)	0.006	11 (12)	7 (10)	0.825
45–55	312 (20)	69 (16)		78 (83)	53 (82)	
≥ 55	1019 (66)	316 (74)		5 (5)	5 (8)	
				Missing cases = 181		
IHC subgroups						
ER+/LN−/HER2−	581 (38)	190 (45)	< 0.001	64 (29)	26 (23)	0.142
ER+/LN+/HER2−	482 (31)	145 (34)		75 (34)	47 (39)	
HER2+	190 (12)	57 (13)		40 (18)	30 (25)	
TN (ER−/PR−/HER2−)	286 (19)	34 (8)		42 (19)	16 (13)	
PAM50						
Luminal A	602 (39)	116 (27)	< 0.001	77 (35)	38 (32)	0.051
Luminal B	291 (19)	197 (46)		36 (16)	32 (27)	
HER2-enriched	184 (12)	54 (14)		30 (14)	20 (17)	
Basal-like	289 (19)	31 (7)		55 (25)	24 (20)	
Normal-like	168 (11)	27 (6)		23 (10)	5 (4)	
Missing cases = 6						

Table 1 Clinico-pathological characteristics of patients in cohort 1 and 2 split by *CCND1* amplification status *(Continued)*

Variable	Cohort 1 (*n* = 1965)		*P*	Cohort 2 (*n* = 340)		*P*
	Non-Amp	Amp		Non-Amp	Amp	
	n (%)	*n* (%)		*n* (%)	*n* (%)	
	1539 (78)	426 (22)		221 (65)	119 (35)	
Treatments						
Endocrine therapy	773 (50)	249 (58)	< 0.001	59 (27)	33 (27)	0.950
Chemotherapy	196 (13)	25 (6)		60 (27)	30 (25)	
Both	148 (10)	41 (10)		99 (45)	55 (47)	
None	422 (27)	111 (26)		3 (1)	1 (1)	

Correlations were calculated using X^2 test unless otherwise specified
ER oestrogen receptor alpha, *PR* progesterone receptor, *HER2* human epidermal growth factor 2 receptor, *TN* triple negative (ER−/PR−/HER2−), *LN* lymph node status, *Amp/non-Amp CCND1* amplified/non-amplified, *Both* patients sequentially received chemotherapy and endocrine therapy
* = Wilcoxon/Mann–Whitney

however, the size of this subgroup is considerably less (*n* = 90) than cohort 1 (*n* = 771).

Combining both cohorts to increase patient numbers demonstrated poorer BCSS for *CCND1*-amplified ER +/LN−/HER2− (HR = 1.66, 95% CI, 1.14–2.41, Fig. 2c), luminal A (HR = 1.68, 95% CI, 1.15–2.46) and luminal B (HR = 1.37, 95% CI, 1.01–1.86) subgroups, along with trends for ER+/LN+/HER2− (HR = 1.32, 95% CI, 0.98–1.78) and endocrine-treated patients (HR = 1.29, 95% CI, 0.98–1.71). These results highlight the ability of *CCND1* amplification status to select a group of patients with poor 15-year breast cancer-specific survival. Of note, results were more ambiguous for systemically treated *CCND1*-amplified patients (see group "Both (Chemo + Endo)") as no difference in survival was noted in cohort 1 (HR = 1.04, 95% CI, 0.52–2.10, Fig. 2a) and significantly worse survival was found in cohort 2 and combined cohorts (HR = 2.24, 95% CI, 1.28–3.91, Fig. 2b and HR = 1.67, 95% CI, 1.11–2.53, Fig. 2c, respectively).

CCND1-amplified luminal A tumours display gene expression changes consistent with more aggressive tumours
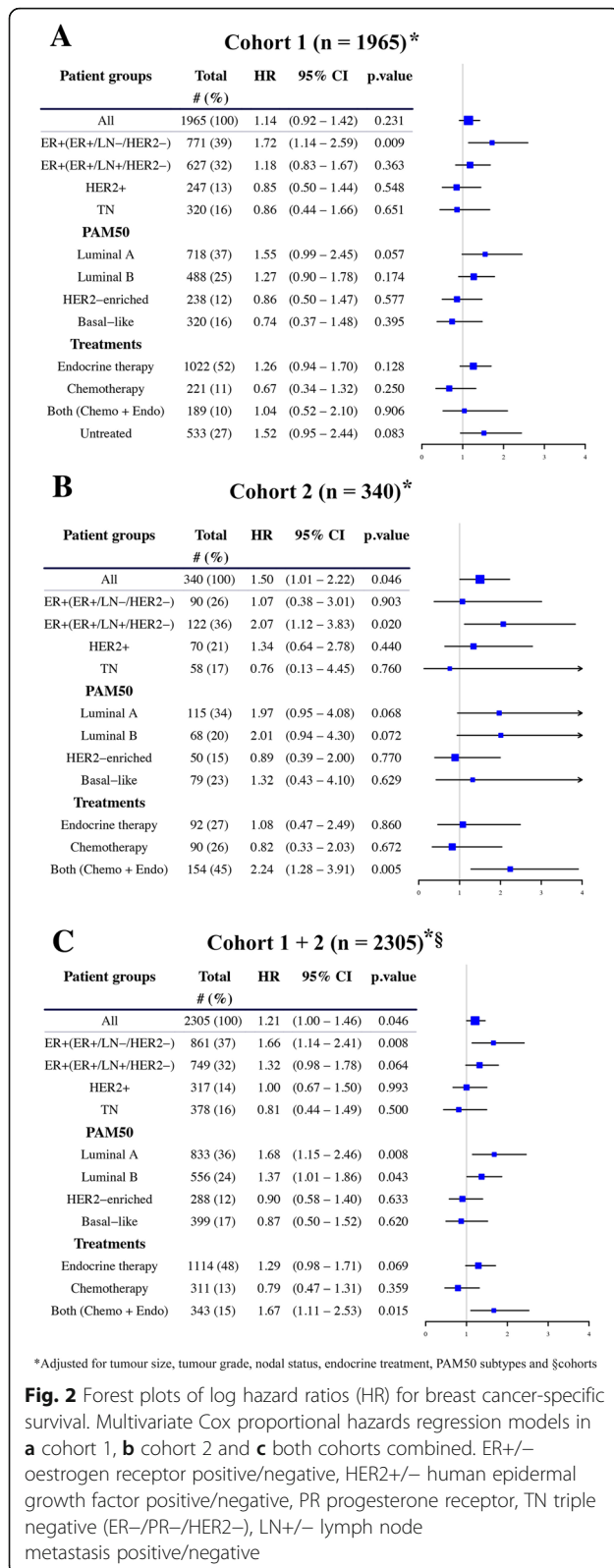
In order to understand why *CCND1* amplification confers a worse survival in luminal A tumours, we first examined the expression of genes related to the cell cycle and cell proliferation across all tumours of cohort 1 within the context of the PAM50 subtypes. *CCND1* gene expression was highest in luminal A/B and lowest in basal-like tumours, while the opposite was true for *CDK4/6* and the proliferation marker gene *MKI67* (Additional file 5: Figure S5). Similarly to *CCND1*, expression of the cell cycle-related genes *RB1* (retinoblastoma 1) and *AR* (androgen receptor) were also higher in luminal tumours relative to basal-like (Additional file 5: Figure S5, see "AR" and "RB1", *P* < 0.001 for luminal A vs. basal-like comparisons). Taken together, these findings suggest that the high levels of proliferation seen in basal-like tumours are unlikely driven by traditional *CCND1/RB1* signalling. As expected, expression of the

oestrogen (*ESR1*)/progesterone (*PGR*) and HER2 (*ERBB2*) genes were highest in the luminal and HER2-enriched subtypes respectively, and as shown by others [37], expression of *PGR* was lower in luminal B tumours relative to luminal A (Additional file 5: Figure S5, "PGR", *P* < 0.001). Examining the same genes in *CCND1*-amplified vs. non-amplified luminal A tumours showed increased *MKI67* and decreased *PGR* gene expression in amplified tumours (Fig. 3, see "MKI67" and "PGR", *P* < 0.001 and *P* = 0.002 respectively), consistent with a more aggressive phenotype. This may partially explain why luminal A *CCND1*-amplified tumours demonstrate poorer survival relative to their non-amplified counterparts. Interestingly, *CDK4* gene expression was higher in luminal B and basal-like *CCND1*-amplified tumours only (Fig. 3, "CDK4" *P* = 0.004 and *P* = 0.029 respectively) whilst *CDK6* expression was lower in luminal A, luminal B and HER2-enriched *CCND1*-amplified tumours (Fig. 3, "CDK6" *P* < 0.001 for all comparisons). No differences were found between *CCND1*-amplified vs. non-amplified tumours in any subtype for the *AR* or *RB1* genes (Additional file 6: Figure S6).

Finally, we further characterised luminal A *CCND1*-amplified tumours by determining if any shared genes exist between amplified tumours in luminal A vs. luminal B molecular subgroups. Differential gene expression analyses were performed between *CCND1*-amplified vs. non-amplified tumours in luminal A and B tumours separately, and the top most changed genes are shown in Additional file 7: Table S1. Remarkably, the top three differentially expressed genes are identical in both luminal A and B subtypes and 12 of the top 20 genes also overlap and are consistent with an 11q amplification event (Additional file 7: Table S1). These results highlight the similarities between luminal A and B *CCND1*-amplified tumours.

Discussion

In this study, we integrated DNA copy number data from SNP assays, RNA expression data from whole genome

A   **Cohort 1 (n = 1965)***

Patient groups	Total # (%)	HR	95% CI	p.value
All	1965 (100)	1.14	(0.92 – 1.42)	0.231
ER+(ER+/LN–/HER2–)	771 (39)	1.72	(1.14 – 2.59)	0.009
ER+(ER+/LN+/HER2–)	627 (32)	1.18	(0.83 – 1.67)	0.363
HER2+	247 (13)	0.85	(0.50 – 1.44)	0.548
TN	320 (16)	0.86	(0.44 – 1.66)	0.651
PAM50				
Luminal A	718 (37)	1.55	(0.99 – 2.45)	0.057
Luminal B	488 (25)	1.27	(0.90 – 1.78)	0.174
HER2–enriched	238 (12)	0.86	(0.50 – 1.47)	0.577
Basal–like	320 (16)	0.74	(0.37 – 1.48)	0.395
Treatments				
Endocrine therapy	1022 (52)	1.26	(0.94 – 1.70)	0.128
Chemotherapy	221 (11)	0.67	(0.34 – 1.32)	0.250
Both (Chemo + Endo)	189 (10)	1.04	(0.52 – 2.10)	0.906
Untreated	533 (27)	1.52	(0.95 – 2.44)	0.083

B   **Cohort 2 (n = 340)***

Patient groups	Total # (%)	HR	95% CI	p.value
All	340 (100)	1.50	(1.01 – 2.22)	0.046
ER+(ER+/LN–/HER2–)	90 (26)	1.07	(0.38 – 3.01)	0.903
ER+(ER+/LN+/HER2–)	122 (36)	2.07	(1.12 – 3.83)	0.020
HER2+	70 (21)	1.34	(0.64 – 2.78)	0.440
TN	58 (17)	0.76	(0.13 – 4.45)	0.760
PAM50				
Luminal A	115 (34)	1.97	(0.95 – 4.08)	0.068
Luminal B	68 (20)	2.01	(0.94 – 4.30)	0.072
HER2–enriched	50 (15)	0.89	(0.39 – 2.00)	0.770
Basal–like	79 (23)	1.32	(0.43 – 4.10)	0.629
Treatments				
Endocrine therapy	92 (27)	1.08	(0.47 – 2.49)	0.860
Chemotherapy	90 (26)	0.82	(0.33 – 2.03)	0.672
Both (Chemo + Endo)	154 (45)	2.24	(1.28 – 3.91)	0.005

C   **Cohort 1 + 2 (n = 2305)*§**

Patient groups	Total # (%)	HR	95% CI	p.value
All	2305 (100)	1.21	(1.00 – 1.46)	0.046
ER+(ER+/LN–/HER2–)	861 (37)	1.66	(1.14 – 2.41)	0.008
ER+(ER+/LN+/HER2–)	749 (32)	1.32	(0.98 – 1.78)	0.064
HER2+	317 (14)	1.00	(0.67 – 1.50)	0.993
TN	378 (16)	0.81	(0.44 – 1.49)	0.500
PAM50				
Luminal A	833 (36)	1.68	(1.15 – 2.46)	0.008
Luminal B	556 (24)	1.37	(1.01 – 1.86)	0.043
HER2–enriched	288 (12)	0.90	(0.58 – 1.40)	0.633
Basal–like	399 (17)	0.87	(0.50 – 1.52)	0.620
Treatments				
Endocrine therapy	1114 (48)	1.29	(0.98 – 1.71)	0.069
Chemotherapy	311 (13)	0.79	(0.47 – 1.31)	0.359
Both (Chemo + Endo)	343 (15)	1.67	(1.11 – 2.53)	0.015

*Adjusted for tumour size, tumour grade, nodal status, endocrine treatment, PAM50 subtypes and §cohorts

Fig. 2 Forest plots of log hazard ratios (HR) for breast cancer-specific survival. Multivariate Cox proportional hazards regression models in **a** cohort 1, **b** cohort 2 and **c** both cohorts combined. ER+/– oestrogen receptor positive/negative, HER2+/– human epidermal growth factor positive/negative, PR progesterone receptor, TN triple negative (ER–/PR–/HER2–), LN+/– lymph node metastasis positive/negative
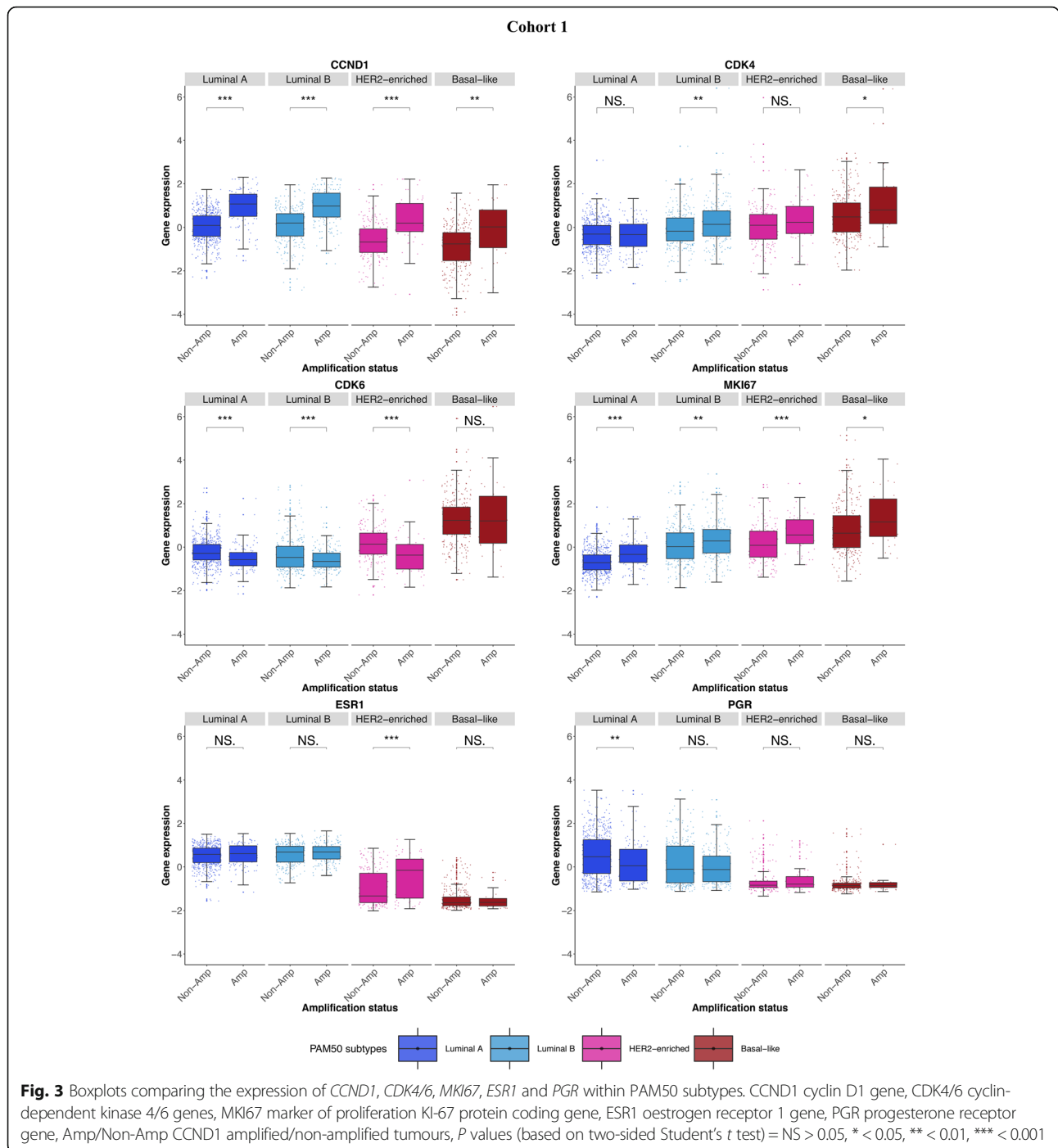
transcriptome arrays and long-term survival data from over 2305 breast cancer patients, with the central goal of determining the prognostic and treatment predictive capacity of cyclin D1 gene amplification. In what is, to our knowledge, the largest and most comprehensive analysis of CCND1 amplification to date, two main novel findings were observed. First, ER+/LN–/HER2–, luminal A and B breast cancer patients with a CCND1-amplified tumour show worse 15-year BCSS relative to non-amplified patients. Similar statistical trends were observed for ER+/LN+/HER2– and endocrine-treated patients. Second, luminal A CCND1-amplified tumours display gene expression changes consistent with more aggressive tumours, specifically increased MKI67 and decreased PGR gene expression in addition to an overlap in genes differentially expressed in CCND1-amplified luminal B tumours.

These findings are supported by the work of others who have also shown poor survival for patients with ER-positive CCND1-amplified tumours [8–10, 12, 15], albeit with shorter clinical follow-up relative to our cohorts. In particular, results from another large study— TransATAC ($n = 1155$)—are in line with our conclusions and show an increased risk of recurrence at 10 years in ER-positive endocrine-treated patients with CCND1-amplified tumours [9]. Focusing on luminal A tumours, Holm et al. also noted worse 10-year overall survival for CCND1-amplified luminal A tumours in univariate analysis [13] ($n = 12$ CCND1-amplified and 78 non-amplified) as did Chin et al. when examining the amplification status of the 11q13 CCND1 amplicon with a median overall survival follow-up of 6.6 years [38]. Patient numbers in both of these studies were however too low to perform multivariable analyses.

It has been previously been suggested that CCND1 amplification could serve as a biomarker for the prediction of response to cyclin-dependent kinase 4/6 inhibitors [39]. When comparing CCND1-amplified vs. non-amplified tumours, we found that CCND1-amplified tumours show higher expression of the proliferation gene MKI67 across all PAM50 subtypes, higher CDK4 expression in luminal B and basal-like subtypes only and lower CDK6 expression in luminal A, luminal B and HER2-enriched subtypes. These findings imply that the increase in proliferation found in CCND1-amplified tumours is unlikely to be dependent on the upregulation of CDK4/6. As such, our data indicate that CCND1 amplification may perform poorly as a predictive biomarker in this setting. This hypothesis is supported by patient data from the PAMOLA-1 study showing no benefit of palbociclib in patients whose tumours were CCND1-amplified [40].

There are a number of strengths to our analyses; first, as this is the largest study of its kind, the number of patients in clinically relevant and PAM50 subtypes has allowed us to comprehensively characterise the

Fig. 3 Boxplots comparing the expression of *CCND1*, *CDK4/6*, *MKI67*, *ESR1* and *PGR* within PAM50 subtypes. CCND1 cyclin D1 gene, CDK4/6 cyclin-dependent kinase 4/6 genes, MKI67 marker of proliferation KI-67 protein coding gene, ESR1 oestrogen receptor 1 gene, PGR progesterone receptor gene, Amp/Non-Amp CCND1 amplified/non-amplified tumours, P values (based on two-sided Student's t test) = NS > 0.05, * < 0.05, ** < 0.01, *** < 0.001

prognostic and predictive potential of *CCND1* amplification using multivariable adjusted statistics; second, multivariable results are generally analogous across two independent breast cancer cohorts with a 15-year BCSS endpoint; third, matching gene expression and SNP array data has meant that we have been able to provide biological insight as to why *CCND1*-amplified tumours may confer worse survival; and fourth, we have kept methods for amplification calls (using CBS, GISTIC) as

similar as possible between both cohorts in order to ensure the reproducibility and consistency of our findings. The limitations are as follows: our analyses are retrospective in nature, we have not performed adjustment for multiple testing and our second cohort is substantially smaller than our first and enriched for more aggressive tumours, ultimately resulting in wider confidence intervals and statistical trends rather than formal significance for some subgroups.

Conclusions

In summary, we show that assessment of *CCND1* amplification status can provide long-term independent prognostic information in patients with ER+/LN−/HER2− tumours, and novelly, within luminal A and luminal B tumours. These findings highlight the potential of *CCND1* to pinpoint patients with poor long-term survival that could benefit from more aggressive clinical treatment strategies.

Additional files

Additional file 1: Figure S1. Survival analysis (Kaplan–Meier estimates) with breast cancer-specific survival (BCSS) as a clinical end point for all patients of cohort 1 split by IHC subtypes. (A) All patients, (B) ER+/LN−/HER2−, (C) ER+/LN+/HER2−, (D) HER2+, and (E) TN. ER oestrogen receptor alpha, PR progestrone receptor, HER2 human epidermal growth factor receptor 2, LN lymph node metastasis, TN triple negative (ER−/PR−/HER2−). *P* values refer to log-rank tests. (TIF 3090 kb)

Additional file 2: Figure S2. Survival analysis (Kaplan–Meier estimates) with breast cancer-specific survival (BCSS) as a clinical end point for all patients of cohort 1 split by PAM50 subtypes and treatments. (A) Luminal A, (B) luminal B, (C) HER2-enriched, (D) basal-like, (E) endocrine therapy, (F) chemotherapy, (G) both chemotherapy and endocrine combined, (H) untreated patients. *P* values refer to log-rank tests. (TIF 3311 kb)

Additional file 3: Figure S3. Survival analysis (Kaplan–Meier estimates) with breast cancer-specific survival (BCSS) as a clinical end point for all patients of cohort 2 split by IHC subtypes. (A) All patients, (B) ER+/LN−/HER2−, (C) ER+/LN+/HER2−, (D) HER2+, (E) TN. ER oestrogen receptor alpha, PR progestrone receptor, HER2 human epidermal growth factor receptor 2, LN lymph node metastasis, TN triple negative (ER−/PR−/HER2−). *P* values refer to log-rank tests. (TIF 2991 kb)

Additional file 4: Figure S4. Survival analysis (Kaplan–Meier estimates) with breast cancer-specific survival (BCSS) as a clinical end point for all patients of cohort 2 split by PAM50 subtypes and treatments. (A) Luminal A, (B) luminal B, (C) HER2-enriched, (D) basal-like, (E) endocrine therapy, (F) chemotherapy, (G) both chemotherapy and endocrine combined, (H) untreated patients. *P* values refer to log-rank tests. (TIF 3194 kb)

Additional file 5: Figure S5. Box plots comparing the expression of *CCND1*, *CDK4/6*, *MKI67*, *RB1*, *AR*, *ESR1*, *PGR* and *ERBB2* genes split by PAM50 subtypes in cohort 1. CCND1 cyclin D1 gene, CDK4/6 cyclin-dependent kinase 4/6 genes, MKI67 marker of proliferation KI-67 protein-coding gene, RB1 retinoblastoma 1 gene, AR androgen receptor gene, ESR1 oestrogen receptor 1 gene, PGR progestrone receptor gene, ERBB2 human epidermal growth factor receptor 2 gene, *P* values (based on ANOVA with post hoc Tukey HSD test) = NS > 0.05, * < 0.05, ** < 0.01, *** < 0.001 (JPG 4941 kb)

Additional file 6: Figure S6. Box plots comparing the expression of *ERBB2*, *AR* and *RB1* genes split by PAM50 subtypes in CCND1-amplified/non-amplified tumours in cohort 1. CCND1 cyclin D1 gene, ERBB2 human epidermal growth factor receptor 2 gene, AR androgen receptor gene, RB1 retinoblastoma 1 gene, *P* values (based on two-sided Student's *t* test) = NS > 0.05, * < 0.05, ** < 0.01, *** < 0.001 (JPG 2924 kb)

Additional file 7: Table S1. Common differentially expressed genes among the top 20 hits of cohort 1 (luminal A, luminal B). Common genes are shown in bold; Chr.loc, chromosomal location; FDR Adj *P*, the false discovery rate adjusted *P* values derived from Benjamini–Hochberg correction; logFC, logarithmic fold change; *Complementary information derived from The HUGO Gene Nomenclature Committee (HGNC) and Gene-Card databases. (PDF 66 kb)

Abbreviations
AR: Androgen receptor; BCSS: Breast cancer-specific survival; BRISQ: Biospecimen reporting for improved study quality; CBS: Circular binary segmentation; CDK4/6: Cyclin-dependent kinases 4/6; CISH: Chromogenic in situ hybridization; CNA: Copy number alteration; CNV: Copy number

variation; DGE: Differential gene expression analysis; ER: Oestrogen receptor alpha; ESR1: Oestrogen receptor 1; GISTIC: Genomic Identification of Significant Targets in Cancer; HER2: Human epidermal growth factor 2; HR: Hazard ratio; IHC: Immunohistochemistry; LN: Lymph node; MKI67: Marker Of proliferation Ki-67 ; PGR: Progesterone; RB1: Retinoblastoma 1; REMARK: REporting recommendations for tumour MARKer prognostic studies; TMA: Tissue microarray; TNBC: Triple-negative breast cancer

Availability of data and materials
Cohort 1 (METABRIC) clinical and genomic data are publicly available from the EGA-archive (https://ega-archive.org) under study number EGAS00000000083. Cohort 2 microarray data has been deposited at the NCBI Gene Expression Omnibus under the accession number GSE48091.

Authors' contributions
AL and NPT contributed the study concept and design. AL contributed to the acquisition and analysis of data. All authors interpreted the data and did the manuscript drafting and critical revision. NPT did the study supervision. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Cohort 1 (METABRIC), publicly available data.
Cohort 2, both gene expression and genotyping studies were approved by the ethics committee at Karolinska Institutet (Stockholm, Sweden).

Consent for publication
Not applicable

Competing interests
JB receives research funding from Merck, paid to Karolinska Institutet and from Amgen, Roche, Sanofi-Aventis and Bayer and paid to Karolinska University Hospital. CMP is an equity stock holder, and a board of director member, of BioClassifier LLC and University Genomics. CMP is also listed as an inventor on a patent application on the PAM50 molecular assay. The other authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Department of Oncology and Pathology, Karolinska Institutet and University Hospital, Stockholm, Sweden. [2]Department of Biosciences and Nutrition, Karolinska Institutet and University Hospital, Stockholm, Sweden. [3]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. [4]Human Genetics, Genome Institute of Singapore, Singapore, Singapore. [5]Department of Pathology, Virginia Commonwealth University, Richmond, VA, USA. [6]Department of Pathology and Cytology, Karolinska Institutet and University Hospital, Stockholm, Sweden. [7]Department of Genetics, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [8]Department of Public Health, Oxford University, Oxford, UK.

Lundberg *et al. Breast Cancer Research*   (2019) 21:34

Page 10 of 10

References

1. Matsushime H, Roussel MF, Sherr CJ. Novel mammalian cyclins (CYL genes) expressed during G1. Cold Spring Harb Symp Quant Biol. 1991;56:69–74.
2. Matsushime H, Roussel MF, Ashmun RA, Sherr CJ. Colony-stimulating factor 1 regulates novel cyclins during the G1 phase of the cell cycle. Cell. 1991;65:701–13.
3. Santarius T, Shipley J, Brewer D, Stratton MR, Cooper CS. A census of amplified and overexpressed human cancer genes. Nat Rev Cancer. 2010;10:59–64.
4. Arnold A, Papanikolaou A. Cyclin D1 in breast cancer pathogenesis. J Clin Oncol. 2005;23:4215–24.
5. Gillett C, Fantl V, Smith R, Fisher C, Bartek J, Dickson C, et al. Amplification and overexpression of cyclin D1 in breast cancer detected by immunohistochemical staining. Cancer Res. 1994;54:1812–7.
6. Bartkova J, Lukas J, Müller H, Lützhøft D, Strauss M, Bartek J. Cyclin D1 protein expression and function in human breast cancer. Int J Cancer. 1994;57:353–61.
7. Keilty D, Buchanan M, Ntapolias K, Aleynikova O, Tu D, Li X, et al. RSF1 and not cyclin D1 gene amplification may predict lack of benefit from adjuvant tamoxifen in high-risk pre-menopausal women in the MA.12 randomized clinical trial. PLoS ONE. 2013;8:e81740 Burns JS, editor.
8. Quintayo MA, Munro AF, Thomas J, Kunkler IH, Jack W, Kerr GR, et al. GSK3β and cyclin D1 expression predicts outcome in early breast cancer patients. Breast Cancer Res Treat. 2012;136:161–8.
9. Lundgren K, Brown M, Pineda S, Cuzick J, Salter J, Zabaglo L, et al. Effects of cyclin D1 gene amplification and protein expression on time to recurrence in postmenopausal breast cancer patients treated with anastrozole or tamoxifen: a TransATAC study. Breast Cancer Res. 2012;14:R57.
10. Bostner J, Ahnström Waltersson M, Fornander T, Skoog L, Nordenskjöld B, Stål O. Amplification of CCND1 and PAK1 as predictors of recurrence and tamoxifen resistance in postmenopausal breast cancer. Oncogene. 2007;26:6997–7005.
11. Choschzick M, Heilenkötter U, Lebeau A, Jaenicke F, Terracciano L, Bokemeyer C, et al. MDM2 amplification is an independent prognostic feature of node-negative, estrogen receptor-positive early-stage breast cancer. Cancer Biomarkers. 2011;8:53–60.
12. Ortiz AB, Garcia D, Vicente Y, Palka M, Bellas C, Martin P. Prognostic significance of cyclin D1 protein expression and gene amplification in invasive breast carcinoma. PLoS One. 2017;12:e0188068.
13. Holm K, Staaf J, Jönsson G, Vallon-Christersson J, Gunnarsson H, Arason A, et al. Characterisation of amplification patterns and target genes at chromosome 11q13 in CCND1-amplified sporadic and familial breast tumours. Breast Cancer Res Treat. 2011;133:583–94.
14. Roy PG, Pratt N, Purdie CA, Baker L, Ashfield A, Quinlan P, et al. High CCND1 amplification identifies a group of poor prognosis women with estrogen receptor positive breast cancer. Int J Cancer. 2010;127:355–60.
15. Elsheikh S, Green AR, Aleskandarany MA, Grainge M, Paish CE, Lambros MBK, et al. CCND1 amplification and cyclin D1 expression in breast cancer and their relation with proteomic subgroups and patient outcome. Breast Cancer Res Treat. 2008;109:325–35.
16. Gillett C, Smith P, Gregory W, Richards M, Millis R, Peters G, et al. Cyclin D1 and prognosis in human breast cancer. Int J Cancer. 1996;69:92–9.
17. Hwang TS, Han HS, Hong YC, Lee HJ, Paik N-S. Prognostic value of combined analysis of cyclin D1 and estrogen receptor status in breast cancer patients. Pathol Int. 2003;53:74–80.
18. Pelosio P, Barbareschi M, Bonoldi E, Marchetti A, Verderio P. Caffo O, et al. Clinical significance of cyclin D l expression in patients with node-positive breast carcinoma treated with adjuvant therapy. 2018;9.
19. Guo L-L, Gao P, Wu Y-G, Jian W-C, Hao C-Y, Li H, et al. Alteration of cyclin D1 in Chinese patients with breast carcinoma and its correlation with Ki-67, pRb, and p53. Arch Med Res. 2007;38:846–52.
20. Rudas M, Lehnert M, Huynh A, Jakesz R, Singer C, Lax S, et al. Cyclin D1 expression in breast cancer patients receiving adjuvant tamoxifen-based therapy. Clin Cancer Res. 2008;14:1767–74.
21. Cardoso F, Harbeck N, Barrios CH, Bergh J, Cortés J, El Saghir N, et al. Research needs in breast cancer. Ann Oncol. 2017;28:208–17 Available from: http://academic.oup.com/annonc/article/28/2/208/2676891.
22. Buus R, Sestak I, Kronenwett R, Denkert C, Dubsky P, Krappmann K, et al. Comparison of EndoPredict and EPclin with oncotype DX recurrence score

23. for prediction of risk of distant recurrence after endocrine therapy. J Natl Cancer Inst. 2016;108.
23. Lundberg A, Lindström LS, Harrell JC, Falato C, Carlson JW, Wright PK, et al. Gene expression signatures and immunohistochemical subtypes add prognostic value to each other in breast cancer cohorts. Clin Cancer Res. 2017;23:7512–20.
24. Tobin NP, Lundberg A, Lindström LS, Harrell JC, Foukakis T, Carlsson L, et al. PAM50 provides prognostic information when applied to the lymph node metastases of advanced breast cancer patients. Clin Cancer Res. 2017;23:7225–31.
25. Curtis C, Shah SP, Chin S-F, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012;486:346–52.
26. Pereira B, Chin S-F, Rueda OM, Vollan H-KM, Provenzano E, Bardwell HA, et al. The somatic mutation profiles of 2,433 breast cancers refine their genomic and transcriptomic landscapes. Nat Commun. 2016. 7:11479 Available from: http://www.nature.com/articles/ncomms11479. [cited 2018 Jun 13].
27. Hanna WM, Kwok K. Chromogenic *in-situ* hybridization: a viable alternative to fluorescence *in-situ* hybridization in the HER2 testing algorithm. Mod Pathol. 2006;19:481–7 Available from: https://www.nature.com/articles/3800555. [cited 2018 Mar 14].
28. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM, et al. Reporting recommendations for tumor marker prognostic studies (REMARK). J Natl Cancer Inst. 2005;97:1180–4.
29. Venkatraman E, Olhsen S, Olshen A. DNAcopy: DNA copy number data analysis. R package. (version 1.54.0).
30. Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, et al. The landscape of somatic copy-number alteration across human cancers. Nature. 2010;463:899–905.
31. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490:61–70.
32. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43:e47 Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4402510/.
33. Bengtsson H, Simpson K, Bullard J, Hansen K. aroma.affymetrix: a generic framework in R for analyzing small to very large Affymetrix data sets in bounded memory. Berkeley: Department of Statistics, University of California; 2008.
34. Cunha SI, Bocci M, Lövrot J, Eleftheriou N, Roswall P, Cordero E, et al. Endothelial ALK1 is a therapeutic target to block metastatic dissemination of breast cancer. Cancer Res. 2015;75:2445–56.
35. Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. JCO. 2009;27:1160–7 Available from: http://jco.ascopubs.org/content/27/8/1160.
36. R Development Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2008. ISBN 3–900051–07-0, URL http://www.R-project.org
37. Prat A, Cheang MCU, Martín M, Parker JS, Carrasco E, Caballero R, et al. Prognostic significance of progesterone receptor–positive tumor cells within immunohistochemically defined luminal A breast cancer. J Clin Oncol. 2013;31:203–9 Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3532392/. [cited 2018 Jun 19].
38. Chin K, DeVries S, Fridlyand J, Spellman PT, Roydasgupta R, Kuo W-L, et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. Cancer Cell. 2006;10:529–41.
39. Musgrove EA, Caldon CE, Barraclough J, Stone A, Sutherland RL. Cyclin D as a therapeutic target in cancer. Nat Rev Cancer. 2011;11:558–72 Available from: http://www.nature.com/nrc/journal/v11/n8/full/nrc3090.html. [cited 2017 Jan 20].
40. Finn RS, Crown JP, Lang I, Boer K, Bondarenko IM, Kulyk SO, et al. The cyclin-dependent kinase 4/6 inhibitor palbociclib in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/TRIO-18): a randomised phase 2 study. Lancet Oncol. 2015;16:25–35.