Identification and validation of DOCK4 as a potential biomarker for risk of bone metastasis development in patients with early breast cancer.

<table>
<thead>
<tr>
<th>Journal:</th>
<th>The Journal of Pathology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>18-325.R1</td>
</tr>
<tr>
<td>Wiley - Manuscript type:</td>
<td>Original Research Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>n/a</td>
</tr>
</tbody>
</table>
| Complete List of Authors: | Westbrook, Jules; University of Sheffield, Academic Unit of Clinical Oncology, Department of Oncology and Metabolism; University of Leeds, Clinical and Biomedical Proteomics Group, Leeds Institute of Cancer and Pathology  
Wood, Steven; University of Sheffield, Academic Unit of Clinical Oncology, Oncology and Metabolism; University of Leeds, Clinical and Biomedical Proteomics Group, Leeds Institute of Cancer and Pathology  
Cairns, David; University of Leeds, Clinical and Biomedical Proteomics Group, Leeds Institute of Cancer and Pathology  
McMahon, Kathryn; University of Leeds, Leeds Institute of Cancer and Pathology  
Gahlaut, Renu; University of Leeds, Leeds Institute of Cancer and Pathology  
Thygesen, Helene; University of Leeds, Leeds Institute of Cancer and Pathology  
Shires, Mike; University of Leeds, Leeds Institute of Cancer and Pathology  
Roberts, Stephanie; University of Leeds, Leeds Institute of Cancer and Pathology  
Marshall, Helen; University of Leeds, Clinical Trials Research Unit, Leeds Institute of Clinical Trials Research  
Oliva, Maria; University of Sheffield, Academic Unit of Clinical Oncology, Department of Oncology and Metabolism  
Dunning, Mark; University of Sheffield, Neurosciences  
Hanby, Andrew; University of Leeds, Leeds Institute of Cancer and Pathology (LICAP)  
Selby, Peter; University of Leeds, Clinical and Biomedical Proteomics Group, Leeds Institute of Cancer and Pathology  
Speirs, Valerie; University of Leeds, Leeds Institute of Cancer and Pathology  
Mavria, Georgia; University of Leeds, Leeds Institute of Cancer and Pathology  
Coleman, Robert; University of Sheffield, Academic Unit of Clinical Oncology, Department of Oncology and Metabolism  
Brown, Janet; University of Sheffield Department of Oncology and Metabolism, Oncology and Metabolism |
<table>
<thead>
<tr>
<th>Tissue:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathology:</td>
</tr>
<tr>
<td>Technique:</td>
</tr>
</tbody>
</table>
Identification and validation of DOCK4 as a potential biomarker for risk of bone metastasis development in patients with early breast cancer

Running Title: DOCK4 in breast cancer bone metastasis

Jules A. Westbrook¹,², Steven L. Wood¹,², David A. Cairns²,⁴, Kathryn McMahon³, Renu Gahlaut¹, Helene Thygesen³, Mike Shires³, Stephanie Roberts², Helen Marshall⁴, Maria R. Oliva¹, Mark J. Dunning⁴, Andrew M. Hanby³, Peter J. Selby²,³, Valerie Speirs³,⁶, Georgia Mavria³, Robert E. Coleman¹, Janet E. Brown¹,².

¹Academic Unit of Clinical Oncology, Department of Oncology and Metabolism, University of Sheffield, UK; ²Clinical and Biomedical Proteomics Group, Leeds Institute of Cancer and Pathology, University of Leeds, UK; ³Leeds Institute of Cancer and Pathology, University of Leeds, UK; ⁴Clinical Trials Research Unit, Leeds Institute of Clinical Trials Research, University of Leeds, UK; ⁵Institute for Translational Neuroscience, Department of Neuroscience, University of Sheffield, UK; ⁶Institute of Medical Sciences, University of Aberdeen, UK.

Conflicts: Prof Janet Brown received fees from Novartis and Amgen for Advisory Boards and Speakers Bureaux; Prof Rob Coleman received consultancy fees from Amgen, Astellas and Boehringer Ingelheim, and lecture fees from Amgen. All other authors declared no conflicts.
Corresponding Author: Janet E Brown (j.e.brown@sheffield.ac.uk), Academic Unit of Clinical Oncology, Department of Oncology and Metabolism, University of Sheffield, Weston Park Hospital, Whitham Road, Sheffield S10 2SJ, UK.
ABSTRACT

Skeletal metastasis occurs in around 75% of advanced breast cancers, with the disease incurable once cancer cells disseminate to bone, but there remains an unmet need for biomarkers to identify patients at high risk of bone recurrence. This study aimed to identify such a biomarker and to assess its utility in predicting response to adjuvant zoledronic acid.

We used quantitative proteomics (SILAC-MS), to compare protein expression in a bone-homing variant (BM1) of the human breast cancer cell line MDA-MB-231 with parental non-bone-homing cells to identify novel biomarkers for risk of subsequent bone metastasis in early breast cancer. SILAC-MS showed that Dedicator of cytokinesis protein4 (DOCK4) was upregulated in bone-homing BM1 cells, confirmed by Western blotting. BM1 cells also had enhanced invasive ability compared with parental cells which could be reduced by DOCK4-shRNA.

In a training Tissue Microarray (TMA) comprising 345 patients with early breast cancer, immunohistochemistry followed by Cox regression revealed that high DOCK4 expression correlated with histological grade (p=0.004) but not oestrogen receptor status (p=0.19) or lymph node involvement (p=0.15). A clinical validation TMA used tissue samples and the clinical database from the large AZURE adjuvant study (n=689). Adjusted Cox regression analyses showed that high DOCK4 expression in the control arm (no zoledronic acid) was significantly prognostic for first recurrence in bone (HR 2.13, 95%CI 1.06-4.30, p=0.034). No corresponding association was found in patients who received zoledronic acid (HR 0.812, 95%CI 0.176-3.76, p=0.790), suggesting that treatment with zoledronic acid may counteract the higher risk for bone relapse from high DOCK4-expressing tumours.
High DOCK4 expression was not associated with metastasis to non-skeletal sites when these were assessed collectively. In conclusion, high DOCK4 in early breast cancer is significantly associated with aggressive disease and with future bone metastasis and is a potentially useful biomarker for subsequent bone metastasis risk.

Keywords: DOCK4; Bone metastasis; breast cancer; biomarker; proteomics
INTRODUCTION

Despite substantial progress in early detection and treatment, breast cancer still accounts for 15% of female cancer-related deaths, with skeletal metastasis occurring in approximately 75% of patients with advanced disease [1]. Relapse in bone typically occurs years after apparently successful treatment of early breast cancer and a period of tumour dormancy. Bone-targeted agents such as bisphosphonates [2] and denosumab [3] are widely used to treat the skeletal complications of established bone metastases, but have also recently been the focus of several large adjuvant studies in early breast cancer to assess their potential to reduce the frequency of relapse in bone and subsequent breast cancer mortality. The phase III AZURE trial (BIG01/04-ISRCTN79831382) recruited 3360 patients with stage II/III breast cancer randomised (1:1) to 5 years of standard adjuvant therapy alone (control) or standard therapy with zoledronic acid (zoledronate[4]. Although there was no significant difference in invasive disease-free survival in the overall population, zoledronate improved disease outcomes for women who were >5 years postmenopausal at diagnosis and a meta-analysis of 26 randomised trials (N=18,766), demonstrated that bone recurrences (HR=0.72; 95%CI 0.60, 0.86, 2p=0.0002) and breast cancer deaths (HR 0.82; 95%CI 0.73, 0.93, 2p=0.002) were reduced by adjuvant bisphosphonates in post-menopausal women [5]. Breast cancer practice has changed as a result of these studies, but they also highlight the unmet need for biomarkers to identify patients with early breast cancer who are most at risk of developing bone recurrence, thus permitting tailoring of treatment to patients most likely to benefit and sparing patients who would not benefit, from potential complications [6]. Tissue samples from AZURE trial patients and the associated clinical database, prospectively designed to support biomarker
studies, offer a major opportunity for clinical validation of novel potential biomarkers for bone metastasis and other disease-related outcomes.

Proteomic studies are yielding key information about breast cancer metastasis to bone [7-9] and, in a recent proteomics-based study, validated in 571 patients, we showed that the proteins CAPG and GIPC1 had both prognostic and predictive potential as biomarkers of bone metastasis [10]. In the current study, we hypothesized that proteins, identified by proteomics and up-regulated in breast cancer cells which have a propensity to home to bone, would be potential biomarkers for metastasis and could play key mechanistic roles in the process of metastatic dissemination to bone.

MATERIALS AND METHODS

Proteomic studies (Further details to those below are given in the text and in Figure S1 of Supporting Information)

Cell culture and SILAC (Stable Isotope Labelling by Amino acids in Cell culture)

The human breast cancer cell line MDA-MB-231 (“PCC”, parental control cells, obtained originally from ATCC) and a bone-homing variant (“BM1”, bone metastatic cells) [11] (the latter supplied by Prof. Joan Massague, Sloan-Kettering Institute, New York), were used in a ‘classical’ SILAC experiment [12,13]. PCC and BM1 cells were cultured in ‘heavy’ SILAC medium and ‘light’ medium for ~10 doublings to ensure >95% isotope label incorporation (Heavy media), and all cells were used within 10 doublings. ‘Heavy’ SILAC media consisted of DMEM containing L-arginine ($^{13}$C$_6$, $^{15}$N$_4$) and L-lysine ($^{13}$C$_6$, $^{15}$N$_2$) (R10K8, DMEM-15, Dundee Cell Products, Dundee, UK) supplemented with dialysed bovine serum (D-FBS100,
Dundee Cell Products, Dundee, UK). ‘Light’ media was DMEM but without the heavy isotopes (R0K0).

**LC-MS/MS**

Equal amounts of protein (40 µg) from light and heavy labelled samples were combined, reduced, alkylated, and separated on a 1-D SDS-PAGE gel. LC-MS-MS was performed by Dundee Cell Products, Dundee, Scotland, UK. Ten slices of the gel-resolved proteins were cut and proteins were digested to peptides using trypsin. Tryptic peptides were separated using a nanoflow LC-System coupled to an LTQ-Orbitrap mass spectrometer (ThermoFisher Scientific).

**Quantitation and bioinformatics analysis**

Quantitation was performed using the software Max Quant (http://www.maxquant.org/downloads.htm), with peptide ratios calculated for each arginine- and/or lysine-containing peptide as the peak area of labelled arginine/lysine divided by the peak area of non-labelled arginine/lysine for each single-scan mass spectrum. Peptide ratios for all arginine- and lysine-containing peptides sequenced for each protein were averaged. Data output from Max Quant were analysed further using Excel and R (v. 3.2, [http://www.r-project.org/](http://www.r-project.org/)) to select differentially expressed proteins.

**Western blotting**

The differential expression of DOCK4, and confirmation of DOCK4 knockdown, was assessed in cell lysates using Western blotting and an infra-red (IR) immunodetection system (LI-COR Biosciences), as well as by Enhanced ChemiLuminescence (ECL). Primary antibodies used were: DOCK4 (Abcam ab56743, mouse monoclonal, 0.1µg/mL); beta-tubulin (Abcam ab6046, rabbit polyclonal, 1/5000
dilution). Secondary antibodies were: #925-68070 IRDye 680RD goat anti-mouse IgG (H+L), LI-COR Biosciences, 1/5000 dilution, for detection of DOCK4 in the 700nm channel (red); #925-32211 IRDye 800CW goat anti-rabbit IgG (H+L), LI-COR Biosciences, 1/5000 dilution, for detection of beta-tubulin in the 800nm channel (green). For ECL secondary antibodies were Goat-anti-mouse-HRP (Abcam, ab 6789) and goat anti-rabbit-HRP (Abcam, Ab6721, both 1:2,500). Normalised densitometric data from six replicate runs of the fluorescent-antibody immunoprobed samples, (and three blots from the ECL-visualized samples), were tested for significance using Student’s t-test.

**Generation of cell lines with stable DOCK4 knockdown and 3D invasion assay**

DOCK4 expression was knocked down in the BM1 and PCC cell lines by lentiviral delivery of a validated shRNA targeting DOCK4 (shR), with the use of an empty vector (cv) as control [14]. Untransduced, wild-type (wt) cells were also studied as control. Inclusion of a GFP marker protein within the lentiviral vectors enabled FACS separation of vector-bearing clones and subsequent culture of pure cell populations. Infected cells were selected by FACS sorting 48 hrs after lentiviral infection. Confirmation of knockdown of DOCK4 protein expression was carried out by Western blotting as described above. Quantification of the blots was performed by densitometric scanning.

To study the effects of reduced expression of DOCK4 protein in PCC and BM1, an invasion assay was carried out using the IncuCyte® platform (Essen BioScience) following the manufacturer's instructions for a ‘scratch-wound’ procedure. Briefly, triplicate wells of an ImageLock 96-well plate (Essen BioScience 4379) were overlaid with thin Matrigel onto which wt, cv, and shR cells for PCC and BM1 were
seeded (50k cells/well and 100k cells/well, respectively), and a homogenous scratch was created using the Incucyte® WoundMaker tool when the cell monolayer was confluent. Cell invasion was assessed by measuring the closure of the scratch introduced on the confluent cell monolayer. Experiments were performed in triplicate using independent cultures of cells. Data were analysed using IncuCyte® software and Excel.

**Patients, samples and immunohistochemistry:**

All analyses on patient samples were performed with Ethics approval (Leeds training set, 06/Q1206/180; AZURE validation set 55/03/182).

Initial studies of DOCK4 expression were performed in a training TMA comprising 345 specimens with available tumour type, grade, ER, lymph node (LN) and overall survival (OS) data, from breast tumours diagnosed at the Leeds Teaching Hospitals NHS Trust (1987 – 2005). Samples were stained with a rabbit polyclonal DOCK4 antibody (1:100, Bethyl Laboratories Inc;A302-263A) and corroborated in a smaller cohort with a mouse monoclonal antibody (1:50, Abcam;Ab56743).

The main analyses for correlations with risk of bone metastasis were performed on TMAs constructed from primary tumours from a sub-set of patients within the overall AZURE trial (n = 689). Due to the relatively high prevalence of bone metastatic outcomes and the long follow-up (median 84 months [interquartile range 66-93]), these TMAs provide an excellent resource for validation of protein biomarkers emerging from our proteomics studies. Protein expression was assessed using the Bethyl DOCK4 antibody (A302-263A). DOCK4 specificity for this antibody was confirmed by immunohistochemistry using FFPE cell pellets from BM1 cells with and without DOCK4 knockdown (Figure S2, Supporting Information). Full gel
images for the antibodies using ECL as the visualization method with use of positive and negative control cell-lines are included (Figure S3, Supporting Information).

Digital images were created using a digital scanner (Aperio Scan Scope XT). Cytoplasmic staining assessment in the invasive margins of primary breast tumours was carried out independently by two trained operators, blinded to outcome data, under the supervision of an experienced breast histopathologist (AMH) who also adjudicated discrepant scores and the level of agreement of the two scores was measured using Cohen’s kappa coefficient. All scores were based upon intensity of staining and not the number of positive cells. Staining intensity was ranked based on a three tier ordinal categorical system used to rank the tumours based on intensity of cytoplasmic staining [15,16] where 1 = weak staining; 2 = moderate, easily perceived staining; 3 = strong/intense staining, i.e. the scoring was based on staining intensity only.

Statistical analyses

All immunohistochemical analyses followed REMARK guidelines [17]. Statistical analyses evaluated the associations between protein expression and relevant clinical and pathological variables (e.g. ER/PR/HER2 status) using Fisher’s Exact test (categorical variables) and the Kruskal-Wallis test (continuous variables), before assessing prognostic and predictive associations with time-to-event data (time to first distant recurrence, time to first skeletal recurrence, time to first non-skeletal recurrence) using Cox proportional hazards regression, the Kaplan-Meier estimate of the survival function and the log-rank test. Time to first distant recurrence was defined as the time from the date of randomization to the date of the distant
recurrence. In analyses, other types of events were censored, e.g. if a local recurrence occurred prior to any distant recurrence, the patient would be censored at the date of the local recurrence. Time to first skeletal recurrence and first non-skeletal recurrence were defined similarly. Time to first skeletal recurrence irrespective of all other previous recurrences was also investigated. Time to event analysis was first performed within treatment arms to identify prognostic associations with the biomarkers. The predictive heterogeneity of effect between treatment arms for time to distant events was assessed in multivariable analysis by including an interaction term in the Cox proportional hazard regressions for treatment arm and biomarker (while adjusting for systemic therapy plan, ER status and lymph node involvement). All significance tests were two-sided and were designated significant at the 5% level.

RESULTS

Proteomic identification of proteins specifically associated with breast cancer bone metastasis

Proteins identified in the ‘forward’ and ‘reciprocal’ labelling experiments (Figure S1) were aligned and the most robust, highest quality, data were extracted for analysis using the following stringent selection criteria: protein identified in both datasets; ≥2 so-called razor+unique peptides assigned to identifications (clarifies assignment of protein identification); equal numbers of peptides and razor+unique peptides assigned to identifications. This resulted in 2006 proteins taken forward for analysis out of a total of 2999 identified in the complete data set (Full proteomic data are available on the publicly accessible database ORDA (https://orda.shef.ac.uk/)).
The aligned dataset was filtered further by using a 1.75-fold cut-off to distinguish change (up or down) in protein expression, resulting in 48 proteins up-regulated in BM1 relative to PCC cells (Table S1, Supporting Information). These were prioritized for further study using literature evidence of relevance to bone metastasis, the magnitude of the differential expression fold change and evidence (where available) of non-association with lung metastases, based on correlation with our other proteomic datasets that included a MDA-MB-231 variant which specifically homes to lung (10, 11). This allowed us to identify proteins likely to be involved in breast cancer metastasis to bone. Consequently, four proteins had potential for further consideration: Dedicator of cytokinesis protein 4 (DOCK4, fold change 2.7), SerpinB2 (fold change 15.6), cell-division cycle protein 20 homolog (CDC20, fold change 3.7), and pericentrin (fold change 2.5).

While all four have published evidence linking them to breast cancer we focused on DOCK4 for further investigation and clinical validation based on its published role in cell migration including breast cancer cell migration, metastasis [18-21] and tumour angiogenesis (14), processes known to be integral to dissemination of tumour cells and development of bone metastases. Moreover, DOCK4 functions as a guanine nucleotide exchange factor for the GTPase Rac1, a key regulator of motility [14,19,20] and localises at actin-rich protrusions in migrating breast cancer cells [19], whilst SNPs within the promoter region of DOCK4 have been detected in breast cancer [22].

**Confirmation of DOCK4 upregulation in BM1 cells**
Analysis of DOCK4 expression levels by Western blotting in the BM1 and PCC cell-lines showed a 2-fold increase in DOCK4 expression within the bone homing BM1 cell-line compared to parental PCC (Figure 1A,C). Increased expression of DOCK4 within BM1 cells was also confirmed using ECL-based visualization.

**DOCK4 is upregulated during breast cancer cell-motility**

There was significant knockdown of DOCK4 protein expression in the BM cell type following lentiviral delivery of DOCK4 shRNA (shR) compared with the empty vector (cv) control (Figure 1B,D). In the wound healing assay, BM1 cells (control vector) had enhanced invasive ability compared with the PCC cells, with 13% and 39% wound closure at 6 and 12h respectively, compared to 9 and 20% for PCC at 6 and 12hrs (Figure 1E,F). Invasion through Matrigel was reduced in cells with DOCK4 knockdown (shR) compared to control, with significant differences in wound closure observed for both cell types at 6 hours and 12 hours, and with a greater effect seen at 6 hrs.

**Association between DOCK4 expression and tumour grade in local breast tumour array**

DOCK4 expression was initially assessed in a local breast tumour array of 345 unselected breast tumours (88% ductal, 9% lobular, 3% other) with patient data available on tumour grade (18% Grade 1, 44% Grade 2, 38% Grade 3) oestrogen receptor (ER), and axillary lymph node (LN) involvement. Examples of typical staining patterns are shown in Figure 2. Analysis revealed a significant association between DOCK4 expression and histological tumour type (p = 0.002) and tumour grade (p=0.004) with 86.4% Grade 3 and 77.3% grade 2 ductal carcinomas expressing

http://mc.manuscriptcentral.com/jpath
For Peer Review

moderate/high DOCK4 (as opposed to 62.5% grade 1 carcinomas), but no association with ER status (positive versus negative, p=0.185) or LN status (involved versus not involved, p=0.15). These data suggested that DOCK4 expression was associated with tumour aggressiveness and further supported our selection of DOCK4 for clinical validation in the AZURE patient cohort.

**Clinical validation of DOCK4 expression in breast cancer patients in the AZURE study.**

**Patients.**

Independent scoring of stained TMAs by two trained operators (JW and SR) under the supervision of AMH yielded a Cohen’s kappa coefficient value of 0.86, signifying excellent agreement. Possible associations between clinical outcomes and immunohistochemistry scores for DOCK4 were tested for 689 patients in the AZURE study (330 control arm, 359 zoledronate arm). Demographic data for these patients are similar to those of the whole AZURE population (Table 1). DOCK4 had no significant associations with age, lymph node involvement, ER status, tumour grade (though this approached significance, p = 0.062), menopausal status, systemic therapy, chemotherapy and statin use. HER2 status was not a mandated assessment but was available for 307 participants. DOCK4 was significantly associated with HER2 status with a smaller proportion of patients with low DOCK4 being HER2 positive (P<0.001). However, adjustment for HER2 status had no impact on subsequent statistical analyses in either control or zoledronate arms.

**AZURE patients: Association of DOCK4 expression with distant event recurrence.**
**Control arm.** Initial analyses considered possible associations between DOCK4 expression (scored as 1, 2 or 3) and disease-free survival (DFS) for distant recurrence components of DFS. Whilst we found no statistically significant association between DOCK4 and first event in non-skeletal sites, taken as a group (p = 0.08, Figure 3a) or in any distant site (p = 0.475, Figure 3d) higher DOCK4 was significantly associated with increased risk of developing first event as skeletal recurrence whether only in bone (p = 0.043, Figure 3c) or in bone and other distant sites concurrently (p = 0.033, Figure 3b).

Subsequently, dichotomised TMA expression scores were used where high DOCK4 expression (score of 3) was compared with low DOCK4 expression (score of 1 or 2). Using these categories, Kaplan-Meier estimates of the survival function for time to distant recurrence confirmed that high DOCK4 is significantly prognostic for first distant recurrence involving bone only (HR 2.1; 95% CI 1.09, 4.15, p = 0.024, Figure 4a). For first distant event involving both skeletal and other site(s) concurrently, although the same trend was observed, the association did not reach significance (HR 1.6; 95%CI 0.88, 3.05, p = 0.113, Figure 4c).

Where the first distant event was in non-skeletal sites (taken as a group), the corresponding dichotomised analyses (Figure 4e), suggested a non-significant reduced risk of non-skeletal events with high DOCK4 (HR = 0.2; 95%CI 0.03, 1.51, p = 0.12). In the dichotomised analyses, we found no association between high DOCK4 and first skeletal event, whether or not other distant events had occurred first, (HR = 1.3; 95%CI 0.73, 2.33, p = 0.373), suggesting that the inclusion of metastasis in other
sites, before eventual spread to bone, counters the significant association seen with the bone only analyses.

**Zoledronate arm.**

In the zoledronate arm, dichotomised analyses revealed no association between DOCK4 and bone only distant events (HR 0.6; 95%CI 0.14, 2.62, p = 0.488, Figure 4b), ie the increased risk of bone-only first event posed by high DOCK4 in the control arm was abolished by zoledronate. This suggests that DOCK4 may act as a predictive biomarker for prevention of bone metastases by zoledronate. This was confirmed by examining Kaplan-Meier plots for DOCK4 high and DOCK4 low comparing control and zoledronate patients for the various events (Supporting Information, Figure S5). Figure S5(b) clearly shows the reduction in bone metastasis-only risk as a first event (HR 0.1; 95% CI 0.03, 0.5, p=0.003) in patients with high DOCK4 treated with zoledronate.

For the first event involving bone and other sites concurrently, there was no significant association seen in the zoledronate group (Figure 4d).

**Cox proportional hazards regressions adjusted for systemic therapy plan, ER status, HER2 status and lymph node involvement**

**Control arm.** Adjusted analyses (Table 2) confirmed that high DOCK4 expression was significantly prognostic for skeletal only events (HR 2.13, 95% CI 1.06, 4.30, p = 0.034). Similar analyses for first recurrence events involving the skeleton and other site(s) concurrently showed no significant association (HR 1.63, 95%CI [0.86, 3.12], p = 0.137; Table 2). No significant association was found in these adjusted analyses.
for non-skeletal distant events when assessed collectively. Analysis of relationships to other individual metastatic sites was not possible with this data-set.

Zoledronate arm. The significant prognostic effect of DOCK4 for skeletal only metastases seen in the control arm patients was not observed in the zoledronate arm (HR 0.812, 95%CI [0.18, 3.76], p = 0.790) (Table 2), suggesting again that the increased risk for skeletal metastasis in patients with high DOCK4 levels at baseline, may be counteracted by treatment with zoledronate. This was tested formally by including an interaction term in the Cox proportional hazards regressions for treatment arm and DOCK4 level for time to first event, skeletal only. These analyses suggested a predictive effect for treatment with zoledronate (HR 0.12 95%CI 0.03, 0.56; likelihood ratio test p=0.063, [hazard ratio <1 indicates improvement with zoledronate]), though this did not reach significance.

**Overall survival and menopausal status**

As shown in Figure 4 (i) and (j) and Supplementary Figure S5 (i) and (j), although there appeared to be a trend towards high DOCK4 producing worse outcome, DOCK4, expression level did not impact significantly on overall survival in either the control or the zoledronate arms. When post-menopausal and pre-menopausal patients were analysed separately, although similar associations with bone metastasis to the full group were observed, there was a loss of statistical significance due to the smaller numbers involved.

**DISCUSSION**
Breast cancer bone metastasis causes significant morbidity and biomarkers which can predict the development of bone metastases are badly needed. In this translational study, SILAC-MS-based comparison of bone homing and non-homing cell variants and functional in vitro work, coupled with rigorous clinical validation enabled us to identify DOCK4 as a potential biomarker for this purpose.

Mechanistically, there have been few studies identifying differentially expressed proteins within breast cancer bone metastasis; as such the current study provides a novel perspective on homing of breast cancer to bone sites of distant relapse. SILAC is an established proteomics approach, which gives quantitative information on the relative expression of proteins when comparing two or more cell types. This approach was successful in terms of the overall numbers of proteins identified with high confidence in the total data set (n = 2999).

Our discovery science used MDA-MB-231 and its bone homing variant BM1. The MDA-MB-231 cell line is regarded as ‘gold standard’ for this type of research as it was derived from a pleural effusion of a breast cancer patient with widespread metastasis, many years after resection of the primary tumour [23]. Furthermore, the validity of this approach has already been proven in our previous work in which the importance of the proteins CAPG and GIPC1 as biomarkers was discovered and subsequently clinically validated (10). Here, we focused on DOCK4, a key guanine nucleotide exchange factor (GEF) regulating the activation of the small GTPase Rac1 (19,20,24,25). DOCK4-mediated activation of Rac1 has been demonstrated to promote actin reorganization and the formation of lamellipodia at the leading edge of
breast cancer cells [19] as well as the formation of lateral filopodia, and blood vessel lumen morphogenesis within tumour angiogenesis [14].

DOCK4 was expressed in parental MDA-MB-231 cells (PCC), but was more abundant in the bone-homing variant (BM1). DOCK4 knockdown inhibited the migration of both PCC and BM cell-lines with inhibition being greater at 6 hours than 12 hours post assay-initiation, suggesting that DOCK4-mediated cell invasion may be important in the earlier stages of breast cancer cell migration.

However, enhanced cell-migration and invasiveness resulting from high DOCK4 expression is only one aspect of the bone-homing cells which makes them bone homing. In particular, c-MAF targets and other proteins elevated in the bone homing cells might drive the bone-homing phenotype as well as DOCK4. Our observation that, in patients, high DOCK4 is specifically associated with first distant metastasis in bone, may be linked to these factors, including those in the bone micro-environment. Notably in this regard, the specific association of high DOCK4 with bone metastasis at any time, is lost once metastasis has occurred elsewhere. This presumably indicates the substantially altered metastatic environment influencing bone metastasis once non-bone metastases have occurred. Association with bone as first metastatic site is lost in zoledronate-treated patients suggesting that zoledronate treatment reduces the risk of bone metastasis to a level similar to that in non-high DOCK4 patients.

This work has shown that DOCK4 has a similar prognostic and predictive profile to CAPG and GIPC1 for the prognosis of skeletal-only relapse within control arm patients.
DOCK4 was also similar to the previously discovered markers in not being predictive of non-skeletal recurrence events within control arm patients. Treatment with zoledronate abolished the association of high levels of all three of these proteins in development of skeletal-only metastasis.

DOCK4 expression is induced by the cytokine transforming growth factor-β (TGFβ) acting via the Smad pathway, and this is a key step in TGFβ’s pro-metastatic effect [21]. GIPC1 is also a key scaffolding protein which functions to transmit signals from the TGFβ-receptor and its co-receptor endoglin to downstream Smad-phosphorylation [26]. TGFβ is a regulator of numerous steps within metastasis including intravasation, extravasation and cancer-cell survival at distant organ sites [27]. The transcription factor c-MAF has recently been identified as a key regulator of breast cancer bone metastasis [28,29]. c-MAF expression is induced by TGFβ and a recent patent application showed that DOCK4-expression correlates with MAF-expression within primary tumours [30]. Examination of our quantitative proteomic data set identified 36 proteins also present in the c-MAF gene set (a panel of 109 genes in total). Within the 36 proteins quantified by proteomics, 15 proteins displayed the same change (increase or decrease of expression within bone-homing cells compared to parental cells) as the relevant gene transcripts in response to c-MAF expression. DOCK4 may therefore be a component of a protein panel which responds to elevated c-MAF expression within bone homing breast cancer cells.

As well as being a prognostic biomarker, our data also provide evidence that DOCK4 is a potential predictive biomarker in terms of the treatment effect of zoledronate for bone as first metastatic site, since the addition of zoledronate appears to reduce the
risk of patients with high DOCK4 levels to that of patients with lower DOCK4 levels. There was also a substantial HR in favour of a treatment effect when an interaction term was included in the Cox proportional hazards regression, though this fell short of statistical significance and additional testing in a further patient cohort is needed before DOCK4 can be confirmed as a predictive biomarker.

Our current studies are supported further through analysis of DOCK4 gene expression within a publicly available database where high DOCK4 transcript levels predicted distant bone metastatic spread of breast cancer (Figure S4, Supporting Information).

Interestingly, DOCK4 was not prognostic for DFS in terms of metastasis at non-skeletal sites, taken as a group (N=28). Indeed, data for the control arm in Figure 4 and Table 2 (and in Figure S5) suggest that high DOCK4 may be associated with a reduction in the occurrence of non-skeletal metastasis, an effect which is not seen in the zoledronate arm. Whilst zoledronate is advantageous in terms of preventing skeletal metastasis in DOCK4 high patients, it may also remove a potentially advantageous effect for non-skeletal metastases. Whilst our data do not point to any particular mechanistic explanation for this effect, it is possible that zoledronate may have effects other than on the skeleton and, in this respect, we note the negative impact of zoledronate on overall survival in pre-menopausal women recently reported, where in non-postmenopausal patients with MAF-positive tumours, zoledronate was associated with worse invasive-disease-free survival and overall survival [28]. Clearly, these factors need to be borne in mind in the consideration of DOCK4 as a predictive biomarker for zoledronate response in bone metastasis.
prevention. DOCK4 may play a functional role in the process of breast cancer metastasis to the bone, as it was also previously shown to play a role in the process of extravasation of lung adenocarcinoma cells to the liver [21]. However, DOCK4 may not be necessary for metastasis to all secondary sites, so that other metastatic sites such as the lung, may not require elevated DOCK4 expression for metastasis. This may be related to the role of DOCK4 for cancer cell extravasation and vascular differences at different metastatic sites, or differential requirement for metastatic growth in different organ microenvironments. Because our analyses included non-skeletal sites as a single group, they do not exclude the possibility of DOCK4 association with a less common site for metastatic spread than the skeleton.

There are some limitations to the current study. Although the number of patients available for analysis from the AZURE study considerably exceeds that required for statistical powering, there is currently no equivalent independent sample set available for further validation. Also, breast cancer metastasis to bone involves numerous autocrine and paracrine signalling events [31], and these and the directionality and recruitment of key signalling pathways that migration involves, is clearly not replicated within the invasion assay used in the current study. More complex tools to assess this are not currently available.

Further studies of the mechanistic role of DOCK4 in breast cancer bone metastasis and implications for pharmacological inhibition is justified by our work. In addition, the potential for the DOCK4 mediated signalling pathway to function as a target for pharmacological inhibition within breast cancer bone metastasis is an exciting possibility. This study suggests that DOCK4 may have clinical utility as a potential
prognostic biomarker for assessment of risk of cancer progression to the skeleton, possibly in combination with other biomarkers, since the specificity and sensitivity of a biomarker test can be improved by the incorporation of multiple markers into a diagnostic/prognostic panel of markers [27]. DOCK4 may therefore form part of a future panel-based test (possibly including CAPG, GIPC1 and the MAF gene [28]) used for informing patient treatment options. Future validation studies of DOCK4 as a biomarker of bone metastasis in other datasets with denosumab treatment or other use of bisphosphonates would be useful, if samples from such datasets become available in the future.

Acknowledgements:

We thank the patients in the AZURE study. We thank Cancer Research UK for the award of a Clinician Scientist Fellowship (to JEB, C18605/A 10048) and for support to Sheffield ECMC, Yorkshire Cancer Research for a project award (S315 to JEB, REC, VS), Breast Cancer Now for pilot and project grant awards (2012MaySP047 and 2016MayPR74 to GM and VS) and the UK Medical Research Council for a Career Development Fellowship [G0802416] to DAC. We also thank Adam Davison and Liz Straszynsky for assistance with cell sorting, and Paul Morris for assistance with immunohistochemistry.

Contributions:

JW, SW, KM, RG, HT, MO, MD, SR, AH, PS, VS GM, RC and JB conceived and/or carried out the experiments. JW, DC, SW, HM, AH, MD, PS, VS, GM, RC and JB analysed the data, all authors were involved in writing the manuscript and had final approval of the submitted version.

Supporting Information
A file has been uploaded for on-line access only containing Supporting Information. This file contains further details on the methods used in this study, and further relevant data in table and figure format. This table (Table S1) and the figures (Figures S1-S5) have been cited in the main article.
REFERENCES


Legends to Figures

**Figure 1.** Schematic showing the key steps in the SILAC proteomic approach used.

MDA-MB-231 (MDA231) and bone-homed variant (BM) cell lines were incubated with media containing ‘heavy’ and ‘light’ isotopically labelled amino acids arginine (R) and lysine (K). The heavy medium contained $^{13}\text{C}_6,^{15}\text{N}_4$-L-Arginine and $^{13}\text{C}_6,^{15}\text{N}_2$-L-Lysine, i.e. stable isotopes of $^{12}\text{C}_6,^{14}\text{N}_4$-L-Arginine and $^{12}\text{C}_6,^{14}\text{N}_2$-L-Lysine, respectively. Two experiments were performed involving ‘forward’ and ‘reverse’ labelling (reciprocal labelling) in which the cell lines were incubated in both media types and then combined in heavy:light pairs in a 1:1 ratio (based on extracted protein assay) prior to separation by molecular weight using 1-DE. The entire lane of gel-separated proteins was cut into equal slices ($n=10$). The proteins in each gel slice were reduced to peptides by enzymatic digestion (using trypsin) and were further separated using high-performance liquid chromatography (HPLC) and analysed using high-resolution mass spectrometry (MS).

**Figure 1:** A and C. Higher DOCK4 expression in BM cell type; B and D

Knock-down of DOCK4 expression by shRNA. DOCK4 was confirmed to be of higher expression in BM1 cell type by Western blotting (2.0-fold induction, $p = 0.027$), and 81% and 88% knock-down of DOCK4 protein expression was achieved in PCC and BM1 cells, respectively. Representative whole-gel lane images are shown.

wt = wild type (no vector involvement); cv = control vector; sh = shRNA.

E and F: Invasion assay. The ability of PCC and BM1 cells to move through Matrigel™ matrix was assessed using a scratch-wound assay. Significant differences
between the control vector and DOCK4 knock-down cells were seen at the 6 and 12hr
time points.

**Figure 2.** Examples of staining profiles for DOCK 4 in patients from the
AZURE study.

Composite of differential protein expression intensities for DOCK4 protein
expression as revealed by IHC and visualised at magnification x 20. The scoring is
based on the intensity of staining in the cytoplasmic compartment in the tumour cells
only. Scale bar = 200µm.

**Figure 3.** Association of DOCK4 with DFS events. Kaplan-Meier estimates of
the relationship between expression of DOCK4 and (a) non-skeletal DFS events (b)
skeletal DFS events (where other distant events may have been recorded at the same
time); (c) solely skeletal DFS events (where no other distant event was recorded at
the same time) and (d) any DFS events (where first event was recorded at any distant
site) in patients in the control arm of the AZURE trial (n = 434). P-value is from the
logrank test for testing equality of survival functions.

**Figure 4.** Univariate associations of distant recurrence outcomes with biomarker
expression in control and zoledronate arms. (Estimates are from Cox proportional
hazards regressions) Kaplan-Meier estimates of the survival function for time to
distant recurrence (DR) and overall survival for control and zoledronate arms.
Numbers 1 to 3 refer to the DOCK4 staining intensity scores. These were
dichotomised, ie DOCK4 low (1 and 2); DOCK4 high (3). Comparisons shown to be
significant are also significant in analyses adjusting for the effect of systemic therapy.
plan, ER status and lymph node involvement. (a) and (b): Skeletal only; (c) and (d): skeletal and other; (e) and (f): Non-skeletal; (g) and (h): First skeletal irrespective of whether other distant events have occurred previously (ie bone metastasis-free survival). (i) and (j): Overall Survival (OS). P-values refer to the logrank test. For definitions of non-skeletal, skeletal and other and skeletal only see legend to Table 1.
Table 1: Characteristics of the patients whose tissue was assessed on TMAs in this study (as at baseline on the AZURE study) and first disease-free survival (DFS) events. Non-skeletal (first distant recurrence event does not include any skeletal component); Skeletal and other (first distant recurrence event reported includes both skeletal and other sites of metastasis, as well as skeletal only); Skeletal only (first distant recurrence event only skeletal - this group is a subset of those classified as skeletal and other).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>DOCK4 dataset</th>
<th>Full AZURE trial population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>zoleodrantone</td>
<td>Control</td>
</tr>
<tr>
<td></td>
<td>(n=359)</td>
<td>(n=330)</td>
</tr>
<tr>
<td>Age (years) Median (range)</td>
<td>50 (26, 75)</td>
<td>51 (32, 79)</td>
</tr>
<tr>
<td>Axillary lymph nodes - no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>4 (1.2)</td>
<td>4 (1.2)</td>
</tr>
<tr>
<td>1-3</td>
<td>223 (67.6)</td>
<td>223 (67.6)</td>
</tr>
<tr>
<td>4</td>
<td>103 (31.2)</td>
<td>103 (31.2)</td>
</tr>
<tr>
<td>Tumour stage - no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>115 (32.0)</td>
<td>116 (35.2)</td>
</tr>
<tr>
<td>T2</td>
<td>196 (54.6)</td>
<td>163 (49.4)</td>
</tr>
<tr>
<td>T3</td>
<td>37 (10.3)</td>
<td>43 (13.3)</td>
</tr>
<tr>
<td>T4</td>
<td>11 (3.1)</td>
<td>8 (2.4)</td>
</tr>
<tr>
<td>Histological grade - no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>26 (7.2)</td>
<td>23 (7.0)</td>
</tr>
<tr>
<td>2</td>
<td>144 (40.1)</td>
<td>130 (39.4)</td>
</tr>
<tr>
<td>3</td>
<td>187 (52.1)</td>
<td>174 (52.7)</td>
</tr>
<tr>
<td>Not specified</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>2 (0.6)</td>
<td></td>
</tr>
<tr>
<td>ER status - no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER positive</td>
<td>255 (77.3)</td>
<td>255 (77.3)</td>
</tr>
<tr>
<td>ER negative</td>
<td>80 (22.3)</td>
<td>72 (21.8)</td>
</tr>
<tr>
<td>ER unknown</td>
<td>1 (0.3)</td>
<td>3 (0.9)</td>
</tr>
<tr>
<td>PR status - no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PR positive</td>
<td>126 (35.1)</td>
<td>102 (30.9)</td>
</tr>
<tr>
<td>PR negative</td>
<td>63 (17.5)</td>
<td>73 (22.1)</td>
</tr>
<tr>
<td>PR unknown</td>
<td>169 (47.1)</td>
<td>153 (46.4)</td>
</tr>
<tr>
<td>Missing</td>
<td>1 (0.3)</td>
<td>2 (0.6)</td>
</tr>
<tr>
<td>HER2 status - no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HER2 positive</td>
<td>38 (10.6)</td>
<td>48 (12.7)</td>
</tr>
<tr>
<td>HER2 negative</td>
<td>106 (29.5)</td>
<td>84 (26.0)</td>
</tr>
<tr>
<td>HER2 unknown/not measured</td>
<td>209 (58.3)</td>
<td>197 (61.2)</td>
</tr>
<tr>
<td>Missing</td>
<td>6 (1.7)</td>
<td>1 (0.3)</td>
</tr>
<tr>
<td>Menopausal status - no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-menopausal</td>
<td>167 (46.5)</td>
<td>151 (45.8)</td>
</tr>
<tr>
<td>≤ 5 years since menopause</td>
<td>53 (14.8)</td>
<td>54 (16.4)</td>
</tr>
<tr>
<td>&gt; 5 years since menopause</td>
<td>112 (31.2)</td>
<td>97 (29.4)</td>
</tr>
<tr>
<td>Menopausal status unknown</td>
<td>27 (7.5)</td>
<td>28 (8.5)</td>
</tr>
<tr>
<td>Planned systemic therapy - no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endocrine therapy alone</td>
<td>24 (6.7)</td>
<td>17 (5.2)</td>
</tr>
<tr>
<td>Chemotherapy alone</td>
<td>79 (22.0)</td>
<td>72 (21.8)</td>
</tr>
<tr>
<td>Endocrine therapy plus chemotherapy</td>
<td>256 (71.3)</td>
<td>241 (73.0)</td>
</tr>
<tr>
<td>Type of chemotherapy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anthracyclins - no. (%)</td>
<td>328 (91.4)</td>
<td>307 (93.0)</td>
</tr>
<tr>
<td>Taxanes - no. (%)</td>
<td>48 (13.4)</td>
<td>41 (12.4)</td>
</tr>
<tr>
<td>Timing of chemotherapy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neo-adjuvant</td>
<td>14 (3.3)</td>
<td>11 (3.3)</td>
</tr>
<tr>
<td>Post-operative</td>
<td>345 (96.7)</td>
<td>319 (96.7)</td>
</tr>
<tr>
<td>Statin use - no. (%)</td>
<td>19 (5.3)</td>
<td>15 (4.5)</td>
</tr>
<tr>
<td>Type of first disease-free survival event - no. (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loco-regional recurrence</td>
<td>26 (7.2)</td>
<td>17 (5.2)</td>
</tr>
<tr>
<td>Distant recurrence</td>
<td>68 (18.9)</td>
<td>74 (22.4)</td>
</tr>
<tr>
<td>Distant and loco-regional recurrence</td>
<td>5 (1.4)</td>
<td>3 (0.9)</td>
</tr>
<tr>
<td>Death without prior recurrence</td>
<td>11 (3.1)</td>
<td>13 (3.9)</td>
</tr>
<tr>
<td>First distant recurrence is nonskeletal - no. (%)</td>
<td>46 (12.8)</td>
<td>28 (8.5)</td>
</tr>
<tr>
<td>-------------------------------------------------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>First distant recurrence includes skeletal and other - no. (%)</td>
<td>27 (7.5)</td>
<td>49 (14.8)</td>
</tr>
<tr>
<td>First distant recurrence is skeletal only - no. (%)</td>
<td>16 (4.5)</td>
<td>38 (11.5)</td>
</tr>
</tbody>
</table>
Table 2 Cox regression analysis for associations between protein IHC and distant recurrence events by AZURE trial arm.

Reference category in each multivariable model is DOCK4 low (1 or 2) Comparisons shown to be significant are also significant in analyses adjusting for the effect of systemic therapy plan, ER status, HER2 status and lymph node involvement. Time to first bone metastasis is defined as time to the first skeletal distant recurrence, irrespective of whether other distant recurrence has occurred earlier (ie bone metastasis-free survival). For definitions of non-skeletal, skeletal plus other and skeletal only, see legend to Table 1. NB: n = number of events and N = number at risk.

<table>
<thead>
<tr>
<th></th>
<th>Standard Treatment</th>
<th></th>
<th>Standard Treatment + zoledronic Acid</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (at risk)</td>
<td>n (events)</td>
<td>HR (95% CI)</td>
<td>p-value</td>
</tr>
<tr>
<td>Nonskeletal distant recurrence</td>
<td>Unadj</td>
<td>330</td>
<td>28</td>
<td>0.154 [0.021,1.131]</td>
</tr>
<tr>
<td></td>
<td>Adj</td>
<td>329</td>
<td>28</td>
<td>0.201 [0.027,1.506]</td>
</tr>
<tr>
<td>Distant recurrence including skeletal</td>
<td>Unadj</td>
<td>330</td>
<td>49</td>
<td>1.642 [0.883,3.052]</td>
</tr>
<tr>
<td></td>
<td>Adj</td>
<td>329</td>
<td>49</td>
<td>1.634 [0.855,3.121]</td>
</tr>
<tr>
<td>Skeletal only distant recurrence</td>
<td>Unadj</td>
<td>330</td>
<td>38</td>
<td>2.121 [1.085,4.148]</td>
</tr>
<tr>
<td></td>
<td>Adj</td>
<td>329</td>
<td>38</td>
<td>2.133 [1.058,4.304]</td>
</tr>
<tr>
<td>Bone metastasis at any time</td>
<td>Unadj</td>
<td>330</td>
<td>62</td>
<td>1.302 [0.728,2.328]</td>
</tr>
<tr>
<td></td>
<td>Adj</td>
<td>329</td>
<td>62</td>
<td>1.344 [0.734,2.46]</td>
</tr>
</tbody>
</table>
Supporting Information.

Further experimental details of proteomics methods

**Cell culture and SILAC (Stable Isotope Labelling by Amino acids in Cell culture)**

The human breast cancer cell line MDA-MB-231 (“PCC”, parental control cells, obtained originally from ATCC) and a bone-homing variant (“BM1”, bone metastatic cells) (1) (the latter supplied by Prof. Joan Massague, Sloan-Kettering Institute, New York), were used in a ‘classical’ SILAC experiment’ (2, 3). Quarterly Mycoplasma checks and annual STR profiling confirmed cell provenance. PCC and BM1 cells were cultured in ‘heavy’ SILAC medium and ‘light’ medium for ~10 doublings to ensure >95% isotope label incorporation (Heavy media), and all cells were used within 10 doublings. ‘Heavy’ SILAC media consisted of DMEM containing L-arginine ($^{13}$C$_6$, $^{15}$N$_4$) and L-lysine ($^{13}$C$_6$, $^{15}$N$_2$) (R10K8, DMEM-15, Dundee Cell Products, Dundee, UK) supplemented with dialysed bovine serum (D-FBS100, Dundee Cell Products, Dundee, UK). ‘Light’ media was DMEM but without the heavy isotopes (R0K0). Essentially, two experiments were performed involving ‘forward’ and ‘reverse’ labelling (i.e. reciprocal labelling) in which the cell lines were incubated in both media types and then combined in heavy:light pairs in a 1:1 ratio (based on extracted protein amount) prior to separation by molecular weight using 1-DE. For proteomic analysis, cells were extracted for total protein content using Laemmli buffer (S-3401 Sigma-Aldrich).

**LC-MS/MS**

Equal amounts of protein (40 µg) from light and heavy labelled samples were combined, reduced, alkylated, and separated on a 1-D gel. LC-MS-MS was performed by Dundee Cell products, Dundee, Scotland, UK. Ten slices of the gel-resolved proteins were cut and proteins were digested to peptides using trypsin. Tryptic peptides were separated using a nanoflow LC-System coupled to an LTQ-Orbitrap mass spectrometer (ThermoFisher
Scientific). Survey full scan MS spectra (m/z 300-1700) were acquired in the Orbitrap with a resolution of 60,000 at m/z 400 after accumulation of 1,000,000 ions. The five most intense ions from the preview survey scan delivered by the Orbitrap were sequenced by collision induced dissociation (normalized collision energy = 40%) after accumulation of 5,000 ions concurrently to full scan acquisition in the Orbitrap. Data were acquired using the Xcalibur software.

**Quantitation and bioinformatics analysis**

Quantitation was performed using the software Max Quant (http://www.maxquant.org/downloads.htm), with peptide ratios calculated for each arginine- and/or lysine-containing peptide as the peak area of labelled arginine/lysine divided by the peak area of non-labelled arginine/lysine for each single-scan mass spectrum. Peptide ratios for all arginine- and lysine-containing peptides sequenced for each protein were averaged. Data output from Max Quant were analysed further using Excel and R (v. 3.2, http://www.r-project.org/) to select differentially expressed proteins for consideration as candidates for further verification and validation.

**Western blotting**

The differential expression of DOCK4, and confirmation of DOCK4 knockdown, was assessed in cell lysates using Western blotting and an infra-red (IR) immunodetection system (LI-COR Biosciences) as well as by Enhanced Chemiluminescence (ECL). The primary antibodies used were: DOCK4 (Abcam ab56743, mouse monoclonal, 0.1µg/mL; beta-tubulin loading control (Abcam ab6046, rabbit polyclonal, 1/5000 dilution. Secondary antibodies were: #925-68070 IRDye 680RD goat anti-mouse IgG (H+L), LI-COR Biosciences, 1/5000 dilution, for detection of DOCK4 in the 700nm channel (red); #925-32211 IRDye 800CW goat anti-rabbit IgG (H+L), LI-COR Biosciences, 1/5000 dilution, for detection of beta-
tubulin in the 800nm channel (green). Normalised densitometric data from six replicate runs of the immunoprobed samples were tested for significance using Student’s t-test. For ECL the secondary antibodies used were: goat-anti-mouse HRP (Abcam, Ab6789) and goat anti-rabbit-HRP (Abcam, Ab 6721) both used a 1 in 2500 dilution followed by visualization with ECL-Reagent (Promega).

**Further experimental details of immunohistochemistry**

Bone homing MDA-MB-231 cells (BM1) were subjected to lentiviral transfection of either control plasmid (control vector) or plasmid expressing anti-DOCK4 miRNA (DOCK4 miRNA) as described in the main manuscript, Materials and Methods. Cells were cultured in DMEM + 10% (v/v) FBS until confluent and then harvested by trypsinization, cells pelleted at 1000 x g and the cell pellet washed twice with PBS. Cell pellets were then resuspended in 10% formalin in Neutral Buffered Saline and fixed for 48 hours at 4°C. In preparation for sectioning, 100µl of 10% (w/v) molten agarose was allowed to set in eppendorf tubes forming an agarose plug. Following fixation of the cells in 10% (v/v) formalin in neutral buffered saline they were pelleted at 1000 x g and then resuspended in 300µl (10% w/v) molten agarose. Cells in molten agarose were introduced into the pre-prepared Eppendorf tubes with 100µl agarose plugs and allowed to set overnight at 4°C. The next day cell-pellets were progressively dehydrated in graded alcohol followed by xylene and then finally molten wax treatment. Cell-pellets were embedded in wax blocks and sectioned at 5µm onto Superfrost Slides (Manufacturer: Menzel-Gläser), then dried for approximately 48 hours at 37°C. Sections were dewaxed in xylene, and rehydrated in graded alcohol and then endogenous peroxidase-blocked with 10% (v/v) hydrogen peroxide. Antigen retrieval was done using citrate buffer (pH 6.0) using a microwave for 20 minutes. Sections were blocked in 10% (v/v) goat serum and then primary antibody (Bethyl anti-DOCK4 antibody, Cat No: A302-263A) added overnight at a dilution of 1 in 100. Sections were developed using
biotinylated secondary antibody (Vector labs) at a dilution of 1 in 200 for 1 hour. Sections were rinsed and developed using the Vector Labs ABC kit (Cat No: PK-6100, standard) for 30 minutes. DAB was applied for 1 minute and then sections rinsed and counter stained with Gills Haematoxylin. Sections were then rehydrated with graded alcohols and mounted using DPX. Slides were scanned using a Pannoramic 250 slide scanner (3DHistech) and images analysed for DAB staining intensity within the QuPath software program.

Testing antibody specificity: Western blotting:

Total cell lysates from bone-homing MDA-MB-231 cells (BM1) and parental controls (PCC) were prepared in 1x Laemmlie Sample Buffer. In addition cell lysates were also prepared from bone homing (BM1) cells following lentiviral transfection of either control vector or DOCK4 miRNA). All cell-lysates were assayed for protein content using the Bio-Rad RC-DC protein assay kit. 50µg of total cell protein was loaded onto each lane of a 10-lane Bio-Rad Mini-Protean 4-20% TGX SDS-PAGE gel and blotted onto nitrocellulose membrane. Following blocking of the membranes overnight at 4°C with 5% (w/v) milk powder in phosphate-buffered saline (PBS) they were probed with either: Abcam anti-DOCK4 (ab56743) at a dilution of 1 in 3400 or the Bethyl anti-DOCK4 antibody (A302-263A) at a 1 in 2000 dilution for 1 hour. Blots were washed and secondary antibody applied (for Abcam ab56743, mouse secondary was applied and for Bethyl anti-DOCK4 a rabbit secondary, both secondaries were used at a 1 in 2500 dilution) for one hour. Blots were then washed with PBS-tween and developed with ECL-reagent (Promega, Cat. No: W1001) prior to exposure to x-ray film.

Gene-expression analysis for DOCK4:

Gene expression data for Wang et al (4) were obtained from the breastCancerVDX (5) Bioconductor package, and sample metadata from the GEO submission (GSE2034) was used
to determine which samples had metastasis to the bone. A recursive partitioning (RP) analysis 
(6) was performed to determine if the samples can be split into two groups based on the 
expression data of DOCK4. Difference in the time to metastasis in the two groups was then 
assessed using the “survival” R package (7).

References for Supporting Information

multigenic program mediating breast cancer metastasis to bone. Cancer Cell. 2003;3(6):537-
49.

isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to 

3. Ong SE, Mann M. A practical recipe for stable isotope labeling by amino acids in cell 

expression profiles to predict distant metastasis of lymph-node-negative primary breast 

5. Schroeder M H-K-B. breastCancerVDX: Gene expression datasets published by 


7. Therneau TMaG, P.M. Modeling Survival Data: Extending the Cox Model. Springer., 
Supplementary Table S1: Proteomic Results for up-regulated proteins within the SILAC comparison of parental (PCC) and bone-homing (BM1) MDA-MB-231 cells:

<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Gene</th>
<th>UniProt</th>
<th>BM1 / PCC Fold change</th>
<th>Number of Peptides</th>
<th>Mol. Weight [kDa]</th>
<th>Sequence Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interstitial collagenase</td>
<td>MMP1</td>
<td>P03956</td>
<td>22.56</td>
<td>12</td>
<td>54.006</td>
<td>469</td>
</tr>
<tr>
<td>Sperm protein associated with the nucleus on the X chromosome B1</td>
<td>SPANXB1</td>
<td>Q9N525</td>
<td>15.58</td>
<td>6</td>
<td>11.826</td>
<td>103</td>
</tr>
<tr>
<td>Plasminogen activator inhibitor 2</td>
<td>SERPINB2</td>
<td>P05120</td>
<td>15.56</td>
<td>12</td>
<td>46.596</td>
<td>415</td>
</tr>
<tr>
<td>Protein EVI2B</td>
<td>EVI2B</td>
<td>P34910</td>
<td>7.43</td>
<td>4</td>
<td>50.583</td>
<td>463</td>
</tr>
<tr>
<td>Creatine kinase B-type</td>
<td>CKB</td>
<td>P12277</td>
<td>6.92</td>
<td>6</td>
<td>42.644</td>
<td>381</td>
</tr>
<tr>
<td>Regulator of G-protein signaling 10</td>
<td>RG510</td>
<td>O43665</td>
<td>4.70</td>
<td>7</td>
<td>21.21</td>
<td>181</td>
</tr>
<tr>
<td>Neutral cholesterol ester hydrolase 1</td>
<td>NCEH1</td>
<td>Q6PIU2</td>
<td>4.69</td>
<td>9</td>
<td>49.881</td>
<td>448</td>
</tr>
<tr>
<td>Protein-glutamine gamma-glutamyltransferase 2</td>
<td>TGM2</td>
<td>P21980</td>
<td>4.40</td>
<td>24</td>
<td>77.328</td>
<td>687</td>
</tr>
<tr>
<td>Annexin A3</td>
<td>ANXA3</td>
<td>P12429</td>
<td>4.08</td>
<td>14</td>
<td>36.417</td>
<td>323</td>
</tr>
<tr>
<td>Cell division cycle protein 20 homolog</td>
<td>CDC20</td>
<td>Q12834</td>
<td>3.67</td>
<td>4</td>
<td>54.722</td>
<td>499</td>
</tr>
<tr>
<td>CD44 antigen</td>
<td>CD44</td>
<td>P16070</td>
<td>3.22</td>
<td>10</td>
<td>81.553</td>
<td>742</td>
</tr>
<tr>
<td>Formin-like protein 1</td>
<td>FMNL1</td>
<td>Q95466</td>
<td>3.19</td>
<td>14</td>
<td>122.41</td>
<td>1104</td>
</tr>
<tr>
<td>Alpha-taxillin</td>
<td>TXLNA</td>
<td>P40222</td>
<td>2.76</td>
<td>6</td>
<td>61.89</td>
<td>546</td>
</tr>
<tr>
<td>Dedicator of cytokinesis protein 4</td>
<td>DOCK4</td>
<td>Q8N110</td>
<td>2.69</td>
<td>6</td>
<td>225.15</td>
<td>1966</td>
</tr>
<tr>
<td>Niemann-Pick C1 protein</td>
<td>NPC1</td>
<td>O15118</td>
<td>2.68</td>
<td>9</td>
<td>142.17</td>
<td>1278</td>
</tr>
<tr>
<td>Membrane-associated progesterone receptor component 2</td>
<td>PGRMC2</td>
<td>O15173</td>
<td>2.53</td>
<td>7</td>
<td>26.17</td>
<td>247</td>
</tr>
<tr>
<td>Pericentrin</td>
<td>PCNT</td>
<td>Q95613</td>
<td>2.52</td>
<td>6</td>
<td>378.08</td>
<td>3336</td>
</tr>
<tr>
<td>Four and a half LIM domains protein 2</td>
<td>FHL2</td>
<td>Q14192</td>
<td>2.42</td>
<td>3</td>
<td>44.205</td>
<td>389</td>
</tr>
<tr>
<td>Ferritin light chain</td>
<td>FTL</td>
<td>P02792</td>
<td>2.38</td>
<td>6</td>
<td>20.019</td>
<td>175</td>
</tr>
<tr>
<td>Vimentin</td>
<td>VIM</td>
<td>P08670</td>
<td>2.37</td>
<td>46</td>
<td>53.651</td>
<td>466</td>
</tr>
<tr>
<td>Serine/threonine-protein phosphatase CPPED1</td>
<td>CPPED1</td>
<td>Q9BRF8</td>
<td>2.36</td>
<td>4</td>
<td>35.576</td>
<td>314</td>
</tr>
<tr>
<td>CUB domain-containing protein 1</td>
<td>CDCP1</td>
<td>Q9HSV8</td>
<td>2.33</td>
<td>4</td>
<td>92.901</td>
<td>836</td>
</tr>
<tr>
<td>Tyrosine-protein phosphatase non-receptor type 1</td>
<td>PTPN1</td>
<td>P18031</td>
<td>2.31</td>
<td>12</td>
<td>49.966</td>
<td>435</td>
</tr>
<tr>
<td>Nicotinamide phosphoribosyltransferase</td>
<td>NAMPT</td>
<td>P43490</td>
<td>2.29</td>
<td>26</td>
<td>55.52</td>
<td>491</td>
</tr>
<tr>
<td>Xaa-Pro aminopeptidase 1</td>
<td>XPNPE1</td>
<td>Q9NQW7</td>
<td>2.27</td>
<td>8</td>
<td>74.797</td>
<td>666</td>
</tr>
<tr>
<td>Integrin alpha-6</td>
<td>ITGA6</td>
<td>P23229</td>
<td>2.25</td>
<td>22</td>
<td>126.63</td>
<td>1130</td>
</tr>
<tr>
<td>Cytosolic non-specific dipeptidase</td>
<td>CNDP2</td>
<td>Q96KP4</td>
<td>2.21</td>
<td>20</td>
<td>52.878</td>
<td>475</td>
</tr>
<tr>
<td>Fascin</td>
<td>FSCN1</td>
<td>Q16658</td>
<td>2.21</td>
<td>21</td>
<td>55.135</td>
<td>500</td>
</tr>
<tr>
<td>Enoyl-CoA delta isomerase 2, mitochondrial</td>
<td>EC12</td>
<td>Q75521</td>
<td>2.16</td>
<td>2</td>
<td>43.585</td>
<td>394</td>
</tr>
<tr>
<td>Coiled-coil domain-containing protein 124</td>
<td>CCDC124</td>
<td>Q96CTT</td>
<td>2.11</td>
<td>6</td>
<td>25.835</td>
<td>223</td>
</tr>
<tr>
<td>Glutathione S-transferase Mu 1</td>
<td>GSTM1</td>
<td>P09488</td>
<td>2.04</td>
<td>3</td>
<td>25.712</td>
<td>218</td>
</tr>
<tr>
<td>Cap-specific mRNA</td>
<td>CMTR1</td>
<td>Q8N1G2</td>
<td>2.01</td>
<td>5</td>
<td>95.32</td>
<td>835</td>
</tr>
<tr>
<td>Nucleobindin-1</td>
<td>NUCB1</td>
<td>Q02188</td>
<td>2.00</td>
<td>5</td>
<td>53.879</td>
<td>461</td>
</tr>
<tr>
<td>Adenosine kinase</td>
<td>ADK</td>
<td>P5S263</td>
<td>1.97</td>
<td>13</td>
<td>40.545</td>
<td>362</td>
</tr>
<tr>
<td>Unconventional myosin-Ixβ</td>
<td>MYO9B</td>
<td>Q13459</td>
<td>1.96</td>
<td>13</td>
<td>243.56</td>
<td>2158</td>
</tr>
<tr>
<td>Gene Name</td>
<td>Accession</td>
<td>Score</td>
<td>Rank</td>
<td>Log2 Fold Change</td>
<td>Score</td>
<td></td>
</tr>
<tr>
<td>--------------------------------------------------------------------------</td>
<td>-----------</td>
<td>-------</td>
<td>------</td>
<td>-----------------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>Stathmin</td>
<td>STMN1</td>
<td>P16949</td>
<td>1.96</td>
<td>7</td>
<td>17.302</td>
<td>149</td>
</tr>
<tr>
<td>Non-syndromic hearing impairment protein 5</td>
<td>DFNAS</td>
<td>O60443</td>
<td>1.96</td>
<td>6</td>
<td>54.554</td>
<td>496</td>
</tr>
<tr>
<td>C-1-tetrahydrofolate synthase, cytoplasmic</td>
<td>MTHFD1</td>
<td>P11586</td>
<td>1.92</td>
<td>49</td>
<td>101.56</td>
<td>935</td>
</tr>
<tr>
<td>Tryptophan--tRNA ligase, cytoplasmic</td>
<td>WARS</td>
<td>P23381</td>
<td>1.92</td>
<td>15</td>
<td>53.165</td>
<td>471</td>
</tr>
<tr>
<td>Protein CutA</td>
<td>CUTA</td>
<td>O60888</td>
<td>1.90</td>
<td>2</td>
<td>20.925</td>
<td>198</td>
</tr>
<tr>
<td>Epidermal growth factor receptor substrate 15-like 1</td>
<td>EPS15L1</td>
<td>Q9UBC2</td>
<td>1.89</td>
<td>9</td>
<td>99.606</td>
<td>910</td>
</tr>
<tr>
<td>Cdc42-interacting protein 4</td>
<td>TRIP10</td>
<td>Q15642</td>
<td>1.88</td>
<td>8</td>
<td>68.351</td>
<td>601</td>
</tr>
<tr>
<td>Mitochondrial import inner membrane translocase subunit Tim9</td>
<td>TIMM9</td>
<td>Q9Y5J7</td>
<td>1.88</td>
<td>3</td>
<td>10.378</td>
<td>89</td>
</tr>
<tr>
<td>Neutral amino acid transporter B</td>
<td>SLC1A5</td>
<td>Q15758</td>
<td>1.81</td>
<td>12</td>
<td>56.598</td>
<td>541</td>
</tr>
<tr>
<td>Chloride intracellular channel protein 4</td>
<td>CLC4</td>
<td>Q9Y696</td>
<td>1.80</td>
<td>17</td>
<td>28.772</td>
<td>253</td>
</tr>
<tr>
<td>Copine-1</td>
<td>CPNE1</td>
<td>BOQZ18</td>
<td>1.79</td>
<td>12</td>
<td>59.717</td>
<td>542</td>
</tr>
<tr>
<td>E3 ubiquitin-protein ligase Itchy homolog</td>
<td>ITCH</td>
<td>Q96102</td>
<td>1.76</td>
<td>4</td>
<td>102.8</td>
<td>903</td>
</tr>
<tr>
<td>Nucleosome assembly protein 1-like 1</td>
<td>NAP1L1</td>
<td>P55209</td>
<td>1.75</td>
<td>14</td>
<td>45.374</td>
<td>391</td>
</tr>
</tbody>
</table>
Legends to Figures in Supporting Information

Figure S1. Schematic showing the key steps in the SILAC proteomic approach used.

MDA-MB-231 (MDA231) and bone-homed variant (BM) cell lines were incubated with media containing ‘heavy’ and ‘light’ isotopically labelled amino acids arginine (R) and lysine (K). The heavy medium contained $[^{13}C_6, ^{15}N_4]$-L-Arginine and $[^{13}C_6, ^{15}N_2]$-L-Lysine, i.e. stable isotopes of $[^{12}C_6, ^{14}N_4]$-L-Arginine and $[^{12}C_6, ^{14}N_2]$-L-Lysine, respectively. Two experiments were performed involving ‘forward’ and ‘reverse’ labelling (reciprocal labelling) in which the cell lines were incubated in both media types and then combined in heavy:light pairs in a 1:1 ratio (based on extracted protein assay) prior to separation by molecular weight using 1-DE. The entire lane of gel-separated proteins was cut into equal slices (n=10). The proteins in each gel slice were reduced to peptides by enzymatic digestion (using trypsin) and were further separated using high-performance liquid chromatography (HPLC) and analysed using high-resolution mass spectrometry (MS).

Figure S2: Testing antibody specificity by immunohistochemistry within FFPE-cells:

(A) Western blotting confirmation of DOCK4 knockdown in DOCK4 miRNA transfected MDA-MB-231 cells,

(B) Representative microscopy fields from FFPE-embedded bone-homing MDA-MB-231 cells (BM1) stably transfected with either control lentivirus (Control vector) or DOCK4-miRNA-expressing lentivirus (DOCK4-miRNA) and subjected to immunohistochemical staining using the Bethyl anti-DOCK4-antibody with Haematoxylin counterstaining. Two representative fields are shown for each cell-type.
(C) Quantification of mean cellular DAB staining level across five replicates of each cell-type (mean ± SEM, n = 5).

Figure S3: Testing antibody specificity by western blotting of cell-lysates:

(A) Full length ECL images of Western blots of 50µg total cell lysate from PCC-cells (lane 1) and BM1 cells (lane 2). Western blotting was performed using the Abcam anti-DOCK4 antibody (ab56743) as described. Tubulin loading control is also shown. Quantification of the normalised DOCK4 band intensity is depicted in the attached histogram (n = 3 replicate gels, mean band intensity ± SEM).

(B) Full length ECL images of Western blots of 50µg total cell lysate from PCC (lane 1) and BM1-cells (lane 2) probed with the Bethyl anti-DOCK4 antibody. Tubulin loading control included. Histogram depicts the quantification of the normalised DOCK4 band intensity (mean ± SEM, n = 3, replicate gels).

Figure S4: Gene expression analysis of DOCK4 expression and time to bone metastasis:

Time to bone metastasis analysis of DOCK4 expression level within breast cancer patients from Wang et al. with high (> 8.53) and low (< 8.53) levels of DOCK4-gene expression.

Figure S5 Univariate associations of distant recurrence outcomes with biomarker expression for DOCK4 low and DOCK4 high. (Estimates are from Cox proportional hazards regressions).

Kaplan-Meier estimates of the survival function for time to distant recurrence (DR) and overall survival for control and zoledronate arms for dichotomised DOCK4 low (1 and 2) and
high (3). Numbers 1 to 3 refer to the DOCK4 staining intensity scores. Comparisons shown
to be significant are also significant in analyses adjusting for the effect of systemic therapy
plan, ER status and lymph node involvement. (a) and (b): Skeletal only; (c) and (d): Skeletal
and other; (e) and (f): Non-skeletal; (g) and (h): First skeletal irrespective of whether other
distant events have occurred previously (ie bone metastasis-free survival). (i) and (j): Overall
Survival (OS). P-values refer to the logrank test. For definitions of non-skeletal, skeletal and
other and skeletal only see legend to Table 1.
Figure S1.
Figure S2:

(A) DOCK4 WB and Tubulin WB

(B) BM1 Cells (Control Vector)

(C) BM1 Cells – DOCK4 miRNA

p-value = 0.013

Mean Cell DAB O.D.
Figure S3:

(A)

Abcam Anti-DOCK4 (ab56743)

MW (kDa) 1 2 1 2 1 2

DOCK4

Tubulin

Normarlised DOCK4 Band Intensity

PCC BM1

p = 0.004

(B)

Bethyl Anti-DOCK4 (A302-263A)

MW (kDa) 1 2 1 2 1 2

DOCK4

Tubulin

Normarlised DOCK4 Band Intensity

PCC BM1

p = 0.004
Figure S4

DOCK4, p = 0.001

Probability of Freedom from Metastasis

Time to Metastasis

- DOCK4 > 8.53 n = 215
- DOCK4 ≤ 8.53 n = 129
Figure S5

**DOCK4 = 1 and 2**

- (a) %DR (skeletal only)
  - HR=0.5 95%CI (0.26, 0.97)
  - adj. HR=0.5 95%CI (0.28, 1.1)
  - P=0.036

- (c) %DR (skeletal and other)
  - HR=0.6 95%CI (0.33, 0.96)
  - adj. HR=0.6 95%CI (0.34, 1)
  - P=0.033

- (e) %DR (nonskeletal)
  - HR=1.2 95%CI (0.73, 1.97)
  - adj. HR=1.2 95%CI (0.62, 1.7)
  - P=0.479

- (g) %first skeletal
  - HR=0.8 95%CI (0.51, 1.19)
  - adj. HR=0.8 95%CI (0.51, 1.2)
  - P=0.25

- (i) %OS
  - HR=0.9 95%CI (0.64, 1.2)
  - adj. HR=0.9 95%CI (0.59, 1.2)
  - P=0.537

**DOCK4 = 3**

- (b) %DR (skeletal only)
  - HR=0.1 95%CI (0.03, 0.62)
  - adj. HR=0.1 95%CI (0.03, 0.5)
  - P=0.003

- (d) %DR (skeletal and other)
  - HR=0.3 95%CI (0.12, 0.91)
  - adj. HR=0.2 95%CI (0.08, 0.7)
  - P=0.025

- (f) %DR (nonskeletal)
  - HR=0.9 95%CI (1.17, 71.3)
  - adj. HR=0.9 95%CI (1.34, 96.8)
  - P=0.01

- (h) %first skeletal
  - HR=0.5 95%CI (0.2, 1.13)
  - adj. HR=0.4 95%CI (0.16, 1)
  - P=0.087

- (j) %OS
  - HR=1.3 95%CI (0.66, 2.44)
  - adj. HR=1.1 95%CI (0.56, 2.2)
  - P=0.468
Figure 1: A and C. Higher DOCK4 expression in BM cell type; B and D Knock-down of DOCK4 expression by shRNA. DOCK4 was confirmed to be of higher expression in BM1 cell type by Western blotting (2.0-fold induction, p = 0.027), and 81% and 88% knock-down of DOCK4 protein expression was achieved in PCC and BM1 cells, respectively. Representative whole-gel lane images are shown. wt = wild type (no vector involvement); cv = control vector; sh = shRNA. E and F: Invasion assay. The ability of PCC and BM1 cells to move through Matrigel™ matrix was assessed using a scratch-wound assay. Significant differences between the control vector and DOCK4 knock-down cells were seen at the 6 and 12hr time points. 

209x259mm (300 x 300 DPI)
Figure 2. Examples of staining profiles for DOCK 4 in patients from the AZURE study. Composite of differential protein expression intensities for DOCK4 protein expression as revealed by IHC and visualised at magnification x 20. The scoring is based on the intensity of staining in the cytoplasmic compartment in the tumour cells only. Scale bar = 200µm.
Figure 3. Association of DOCK4 with DFS events. Kaplan-Meier estimates of the relationship between expression of DOCK4 and (a) non-skeletal DFS events (b) skeletal DFS events (where other distant events may have been recorded at the same time); (c) solely skeletal DFS events (where no other distant event was recorded at the same time) and (d) any DFS events (where first event was recorded at any distant site) in patients in the control arm of the AZURE trial (n = 434). P-value is from the logrank test for testing equality of survival functions.

190x107mm (300 x 300 DPI)
Figure 4. Univariate associations of distant recurrence outcomes with biomarker expression in control and zoledronate arms. (Estimates are from Cox proportional hazards regressions) Kaplan-Meier estimates of the survival function for time to distant recurrence (DR) and overall survival for control and zoledronate arms. Numbers 1 to 3 refer to the DOCK4 staining intensity scores. These were dichotomised, ie DOCK4 low (1 and 2); DOCK4 high (3). Comparisons shown to be significant are also significant in analyses adjusting for the effect of systemic therapy plan, ER status and lymph node involvement. (a) and (b): Skeletal only; (c) and (d): skeletal and other; (e) and (f): Non-skeletal; (g) and (h): First skeletal irrespective of whether other distant events have occurred previously (ie bone metastasis-free survival). (i) and (j): Overall Survival (OS). P-values refer to the logrank test. For definitions of non-skeletal, skeletal and other and skeletal only see legend to Table 1.