



HAL
open science

Extensive expression of craniofacial related homeobox genes in canine mammary sarcomas

Helena Wensman, Hanna Göransson, Karl-Johan Leuchowius, Sara Strömberg, Fredrik Pontén, Anders Isaksson, Gerard Roel Rutteman, Nils-Erik Heldin, Gunnar Pejler, Eva Hellmén

► **To cite this version:**

Helena Wensman, Hanna Göransson, Karl-Johan Leuchowius, Sara Strömberg, Fredrik Pontén, et al.. Extensive expression of craniofacial related homeobox genes in canine mammary sarcomas. Breast Cancer Research and Treatment, 2008, 118 (2), pp.333-343. 10.1007/s10549-008-0243-7 . hal-00535316

HAL Id: hal-00535316

<https://hal.science/hal-00535316>

Submitted on 11 Nov 2010

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

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

Extensive expression of craniofacial related homeobox genes in canine mammary sarcomas

Helena Wensman · Hanna Göransson · Karl-Johan Leuchowius · Sara Strömberg · Fredrik Pontén · Anders Isaksson · Gerard Roel Rutteman · Nils-Erik Heldin · Gunnar Pejler · Eva Hellmén

Received: 10 September 2008 / Accepted: 30 October 2008 / Published online: 2 December 2008
© Springer Science+Business Media, LLC. 2008

Abstract The global gene expression in three types of canine mammary tumors: carcinoma, fibrosarcoma and osteosarcoma were investigated by Affymetrix gene array technology. Unsupervised clustering analysis revealed a close clustering of the respective tumor types, with fibrosarcomas clustering close to the osteosarcomas and the carcinomas clustering closer to non-malignant mammary tissues (NMTs). A number of epithelial markers were expressed in both carcinomas and NMTs, whereas the sarcomas expressed genes related to mesenchymal differentiation. A comparison of the gene expression profile of the sarcomas versus carcinoma/NMTs revealed that the sarcomas, in particular the osteosarcomas, showed a striking upregulation of a panel of homeobox genes previously linked

to craniofacial bone formation. In line with this finding, osteosarcomas showed an upregulation of bone morphogenetic proteins (BMPs) and of genes associated with retinoic acid signaling. Increased homeobox gene expression in sarcomas was also confirmed at the protein level by immunohistochemical analysis of tumor tissue, and in an osteosarcoma cell line after stimulation by BMP-2. These findings suggest that the development of mammary sarcomas specifically involves triggering of a set of homeobox genes related to neural crest and craniofacial bone development.

Keywords Canine mammary tumors · Mammary sarcoma · Mammary osteosarcoma · Homeobox transcription factor · Gene expression profiling · Craniofacial

Electronic supplementary material The online version of this article (doi:10.1007/s10549-008-0243-7) contains supplementary material, which is available to authorized users.

H. Wensman (✉) · G. Pejler · E. Hellmén (✉)
Department of Anatomy, Physiology and Biochemistry,
Swedish University of Agricultural Sciences, P.O.Box 7011,
750 07 Uppsala, Sweden
e-mail: Helena.Wensman@afb.slu.se

E. Hellmén
e-mail: Eva.Hellmen@afb.slu.se

H. Göransson · A. Isaksson
Department of Medical Sciences, Uppsala University,
Uppsala, Sweden

K.-J. Leuchowius · S. Strömberg · F. Pontén · N.-E. Heldin
Department of Genetics and Pathology, Uppsala University,
Uppsala, Sweden

G. R. Rutteman
Department of Clinical Sciences of Companion Animals,
Faculty of Veterinary Medicine, Utrecht University, Utrecht,
The Netherlands

Abbreviations

NMT Non-malignant mammary tissues
BMP Bone morphogenetic protein

Introduction

Most research on human breast tumors has focused on carcinomas, since these are by far the most common type [1]. Breast sarcomas, and in particular osteosarcomas, are much more rare, which has hampered a profound characterization of these tumors in humans [2]. Notably, breast sarcomas can be very aggressive and it is therefore important to gain more knowledge concerning these forms of tumors [3, 4]. The biology of the carcinomas and sarcomas, including those of the breast, show important differences in morphology but also in metastatic behavior, with the vast majority of carcinomas metastasizing via the

lymphatic system and most sarcomas metastasizing directly via the blood vascular system [2, 5, 6].

In female dogs, spontaneous mammary tumors including sarcomas constitute the most common form of malignancy [7]. Overall, canine and human cancers have similar biological characteristics, such as the role of steroidal hormones in tumorigenesis, the correlation between grade of histological malignancy and rate of metastasis, and the organ-related pattern of metastasis [1, 8]. For mammary carcinomas the shift towards steroidal autonomy—based upon the lack of estrogen and progesterone receptors and absence of response to antihormonal treatment, appears to be more common in the dog than in the human [7], while a near complete lack of data exists on this issue in sarcomas. Mammary sarcomas, on the other hand, appear to be relatively more common in the dog as compared to the human. Thus, material for studying the disease is more readily available from canine than from human sources and investigations of canine sarcoma could provide useful comparative oncological insight that may be valuable for understanding the pathology also of human sarcomas. For example, dogs are relatively outbred based on single nucleotide polymorphism frequency and the background gene diversity between dog breeds is comparable with that in humans [9, 10]. Further, inter-species comparisons of a number of genes involved in tumor development (P53, MET, IGF1R, KIT and mTOR) have shown that the human and canine genes are more closely related to each other than the corresponding human and mouse genes [9, 11, 12]. Another major advantage of using the canine system as model is that the studied tumors occur spontaneously, in contrast to the experimentally induced tumors in various murine model systems.

Microarray technology has proven to be a powerful tool in classifying natural subtypes of breast cancer [13, 14]. Much effort has been put into finding a “minimal set of genes” to identify subclasses of tumors and there are examples of identified gene sets that have been successful in predicting outcome of disease [15–18]. So far, to our knowledge, there have been no mammary/breast sarcomas subjected to gene expression profiling, although soft tissue and osteosarcomas from other organ systems have been studied. Among the dysregulated genes in sarcomas are tyrosine kinases and transcription factors such as homeobox transcription factors normally involved in development [19, 20].

Homeobox transcription factors have well documented roles in embryo development, where they control embryo patterning and cell differentiation. A growing body of evidence indicates a connection between dysregulated homeobox genes with disease and congenital abnormalities [21]. Expression of several homeobox genes, for example *MSX-2*, has been identified during the development of the

mammary gland [22, 23] where they are candidates for the orchestration of the cell signaling that enables the gland to undergo repeated cycles of growth and functional differentiation [24]. Homeobox genes are also known to be upregulated in tumors, where they have been linked to differentiation and prognosis [24–27].

Here we used gene array technology to compare the global gene expression patterns in canine mammary simple carcinomas, fibrosarcomas and osteosarcomas. The canine genome has only recently been made available [10] and this is the first study where global gene expression is studied in primary canine mammary tumors. Our results reveal profound differences in the gene expression patterns of the various canine mammary tumors, and we provide evidence for a specific upregulation of homeobox genes in mammary sarcomas. Homeobox genes are currently receiving much attention for their potential role in metastasis [28–30]. Our results thus provide new insight into the expression of homeobox genes in mammary malignancies, and suggest that homeobox gene expression is a hallmark of mammary sarcoma/osteosarcoma.

Materials and methods

Tumor and control specimens

Primary canine mammary tumors, osteosarcomas ($n = 5$), fibrosarcomas ($n = 6$), simple carcinomas ($n = 7$) and normal mammary tissue controls ($n = 4$) were obtained from patients subjected to surgery, snap frozen in liquid nitrogen and stored in -80°C . See Table 1 for further information and classification of the tumors.

RNA preparation

RNA was prepared from tumor specimens by pulverizing 250 mg of the tumor under liquid nitrogen in a mortar followed by addition of 6 ml Trizol (Invitrogen). The RNA quality was evaluated with the Agilent 2100 Bioanalyzer system. RNA was purified on Qiagen mini columns and RNA concentration measured with a NanoDrop ND-1000 spectrophotometer.

Microarray expression analysis

From each sample of total RNA, 2 μg were used to prepare biotinylated fragmented cDNA according to the GeneChip[®] Expression Analysis Technical Manual (Rev. 5, Affymetrix Inc., Santa Clara, CA). Affymetrix GeneChip[®] expression arrays (Canine Genome 2.0 Array) were hybridized for 16 h in a 45°C incubator, rotated at 60 rpm. The arrays were then washed and stained using the Fluidics

Table 1 Data of the dogs and tumors included in the study

Tumor	Dog	Pathology	Follow-up	Metastases
6B	Mixed breed, 13 years	Osteosarcoma	Lung metastases	Yes
143B	Airedale terrier, 11 years	Osteosarcoma	Metastases in mediastinal lnn, lungs, myocardium and kidneys	Yes
144A	Giant Schnauzer, 11 years	Osteosarcoma	Metastases in sternal lnn, lungs, pericardium, myocardium, pleura, spleen and kidneys	Yes
353H1A	Papillion, 15 years	Combined osteosarcoma	Metastases in lungs and kidneys	Yes
P136	Poodle, 12 years	Combined osteosarcoma	Lung metastases	Yes
83B	Great Dane, 10 years	Fibrosarcoma	Autopsied 6 months post-surgery	No
105A	Mixed breed, 8 years	Fibrosarcoma	Euthanised due to local recurrence	ND
184C	Boxer, 10 years	Fibrosarcoma	Lung metastases	Yes
195B	Long-haired Dachshund, 12 years	Fibrosarcoma	Euthanised due to local recurrence	ND
P138	Mechelen Belgian shepherd (or malinois), 9 years	Fibrosarcoma	Euthanised due to local recurrence	No
P222	Toy Dachshund, 12 years	Fibrosarcoma	Metastases in lymph vessels, no metastases at follow up 3 month after surgery	ND
203B	Shih Tzu 12 years	Simple carcinoma, lipid-rich	Lung metastases (X-ray)	Yes
346A	Dachshund, 13 years	Simple carcinoma, solid	Euthanised due to local recurrence	ND
P132	Greyhound, 13 years	Simple carcinoma, tubulopapillary	Metastases in regional lnn and lungs	Yes
P152	Dalmatian, 6 years	Simple carcinoma, anaplastic	Metastases in lymph vessels and lungs	Yes
P187	Dobermann Pinscher, 12 years	Simple carcinoma, anaplastic and tubulopapillary	Metastases in lymph vessels euthanised due to local recurrence	Yes
P208	Mixed-breed, 9 years	Simple carcinoma, tubulopapillary, squamous cell metaplasia, lymphocytes	Metastases in lymph vessels and lungs	Yes
P225	Engl Cocker Spaniel, 9 years	Simple carcinoma, tubulopapillary	No metastases at 6 weeks after surgery, then lost to follow up	ND

ND not determined

Station 450 and finally scanned using the GeneChip® Scanner 3000 7G.

Microarray data analysis

Analysis of the gene expression data was carried out in the statistical computing language R (<http://www.r-project.org>) using packages available from the Bioconductor project (www.bioconductor.org). The raw data were normalized using the robust multi-array average (RMA) [31] background-adjusted, normalized and log-transformed summarized values. In order to search for the differentially expressed genes, an empirical Bayes moderated *t*-test was applied [32]. To address the problem with multiple testing, the *P* values were adjusted according to Benjamini and Hochberg [33]. The gene clusterings were performed with the program Genesis, version 1.7.1 [34].

Immunohistochemistry

Tissues were fixed in 4% buffered formalin and sections were baked in 37°C, overnight followed by 1 h in 60°C. Heat induced antigen retrieval in decloaking chamber

(Biocare Medical) and PT Module Buffer 1 (100× citrate buffer pH 6.0) was performed (Labvision). Slides were washed in TBS with 0.3% Tween-20. Polyclonal affinity purified antibodies against MSX2 (HPA005652), SIP1/ZEB2 (HPA003456) and SatB2 (HPA001042) generated within the Swedish Human protein atlas project (HPA) www.proteinatlas.org were used [35]. The antibodies were generated using an antibody-based proteomics approach [36]. For detection, the ABC-Elite system (Vector laboratories, CA, USA) and DAB were used. For every immunohistochemical analysis, a negative control was used in addition to canine mammary gland and human tissue arrays as positive controls for the antibodies.

BMP cell stimulation

The mammary osteosarcoma cell line 353H1A clone 1 was stimulated with BMP-2. Cells were incubated overnight in RPMI 1640 with 0.5% fetal bovine serum. Cells were stimulated with 250 ng/ml BMP-2 (Peprotech, London, UK) in PBS/1.3% bovine serum albumin (BSA) for 1 h (37°C). Negative controls were incubated with PBS/1.3% BSA. The cells were trypsinized and cytopsin slides were prepared.

Immunofluorescence with stimulated cells

Mammary osteosarcoma cells were fixed with 4% buffered formalin and antigen retrieval performed. The cells were blocked with 1% horse serum/1% donkey serum in TBS with 0.05% Tween-20 and then incubated with the polyclonal primary antibodies diluted in the blocking solution (P-Smad 1/5 [37]), SIP1/ZEB2 (HPA003456) and Goosecoid (Santa Cruz). The primary antibodies were detected by a Cy3-labeled donkey anti-rabbit antibody (Jackson ImmunoResearch). The slides were mounted with Vectashield (Vector laboratories) and imaged with an epifluorescence microscope (Axioplan II, Zeiss).

Results and discussion

This is the first study that addresses the gene expression profiles of various canine mammary tumor types. We used the Affymetrix Canine Genome 2.0 expression array to compare the global gene expression patterns of canine mammary simple carcinomas, osteosarcomas and fibrosarcomas (see Table 1 for classification of the used tumors). As a control, non-malignant mammary tissue (NMT) was used. An unsupervised clustering of the gene expression data showed that the different tissues clustered well together in the four separate groups (Fig. 1). As expected, the osteosarcoma and fibrosarcoma groups showed an adjacent clustering. The carcinomas and NMTs clustered closely, in line with predominant epithelial gene expression of both tissues. Since the canine genome has not yet been fully annotated with gene ontology terms, a detailed gene ontology analysis of the data is not possible at this stage.

In order to analyze differences in the gene expression patterns of the different types of tumors, we performed a supervised analysis with the two sarcoma groups, osteosarcomas and fibrosarcomas versus the simple carcinomas (suppl. Fig. 1). Overall, genes related to extracellular matrix and mesenchymal genes dominated in the sarcomas whereas characteristic epithelial genes dominated in the

carcinomas. For example, several genes involved in cell–cell adhesion and genes related with epithelial differentiation were expressed at higher levels in the carcinomas than in sarcomas, e.g., various cytokeratins, E-cadherins and occludin (suppl. Fig. 1, suppl. Table 1). The carcinomas and the NMT showed similar expression patterns, e.g., expressing several genes associated with epithelial morphology and cell–cell adhesion (suppl. Fig. 1, suppl. Table 1).

In Table 2, the 21 genes that showed the highest upregulation in sarcomas versus carcinomas are displayed. These include a number of connective tissue-related genes, e.g., procollagen type V, matrix metalloprotease 16, collagen-binding protein 2 and procollagen C-endopeptidase enhancer. Further, an upregulation of leprecan, a proteoglycan found in basement membranes, was seen and there was also an upregulation of sarcoglycan, the latter being a transmembrane glycoprotein found predominantly in muscle tissue. We also note that Elastin microfibril interfacier, a gene associated with focal adhesion, was highly expressed and that genes characteristic for mesenchymal morphology were expressed: FK506 binding protein precursor, Microtubule-associated protein 1A and Discoidin domain receptor tyrosine kinase 2, the latter being implicated in chondrocyte proliferation. Further, sarcomas expressed high levels of cysteine dioxygenase, an enzyme involved in Cys metabolism and that has been linked to neurological disorders. Interestingly, pleiotrophin, an angiogenic factor that has been implicated in the angiogenic switch in different cancer models in vivo [38], was significantly higher expressed in the sarcomas than in carcinomas. Sarcomas also expressed high levels of reticulocalbin 3, an ER resident Ca^{2+} -binding protein, as well as olfactomedin-like protein 3, glutathione S-transferase theta 2, ubiquitin carboxyterminal hydrolase L1, serine (or cysteine) proteinase inhibitor clade H member 1 and aldehyde dehydrogenase.

A number of transcription factors were highly expressed in the sarcomas: early B-cell factor (EBF1), transcription factor COE1, zinc finger CCHC domain containing 12 and two homeobox transcription factors: Paired mesoderm homeobox 1 isoform (PMX-1b) and SIP1/ZEB2. Moreover, when comparing the expression profile of osteosarcomas versus carcinomas, additional homeobox factors were found among the genes showing the highest degree of upregulation, e.g., homeobox protein DLX-5, homeobox A4, paired mesoderm homeobox protein 2 and Short stature homeobox 2. Taken together, the individual tumor types thus express high levels of a number of genes characteristic of the respective tissue, indicating a good correlation between gene expression profile and histological classification.

Considering the potential involvement of homeobox genes in differentiation and in malignancy, we decided to

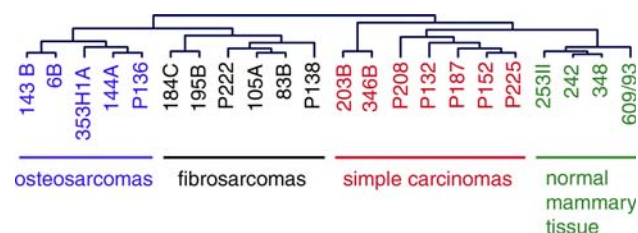


Fig. 1 Unsupervised clustering of gene expression data from the Affymetrix Canine Genome 2.0 Array, all tumors and normal mammary tissue

Table 2 Genes with increased expression in sarcomas compared to carcinomas

Gene	Increased expression in sarcomas		
	Fold	Intensity	
		Carcinomas	Sarcomas
Pleiotrophin (PTN)	43.5	27	1,175
Cysteine dioxygenase, type I	17.0	72	1,224
Matrix metalloproteinase 16 (MMP16)	9.8	63	620
Olfactomedin-like 3	9.8	124	1,219
Elastin microfibril interfacier	7.5	57	425
Procollagen, type 5, alpha 1	7.3	66	485
Paired mesoderm homeobox 1 isoform (PMX-1b)	7.3	96	702
Delta-sarcoglycan isoform	7.1	27	191
Glutathione S-transferase theta 2B	6.4	69	442
Aldehyde dehydrogenase 1 family, member L2	6.3	33	207
Procollagen C-endopeptidase enhancer	5.9	217	1,284
FKBP10 FK506 binding protein 10, 65 kDa	5.8	80	467
Early B-cell factor (EBF1)	5.8	166	965
Reticulocalbin 3, EF-hand calcium binding domain	5.8	100	582
Leucine proline-enriched proteoglycan (leprecan) 1	5.3	159	839
Ubiquitin carboxyl-terminal esterase L1	5.1	92	469
Discoidin domain receptor tyrosine kinase 2 (DDR2)	4.9	86	422
Serine (or cysteine) proteinase inhibitor, clade H, member 1 precursor	4.9	197	959
Zinc finger, CCHC domain containing 12	4.7	73	340
Microtubule-associated protein 1A	4.3	78	338
Zinc finger E-box binding homeobox 2 (ZEB2) (SIP1)	4.1	160	648

Genes with a difference of greater than fourfold in expression between the groups and with adjusted *P* value <0.0005 are shown

further investigate the homeobox gene expression in the various mammary tumors. Another reason for addressing the homeobox gene expression is that they can act as master genes, i.e., regulating a number of additional genes [28]. A supervised clustering of this group of genes indeed revealed a striking upregulation of a large number of homeobox genes in the sarcomas, especially in the osteosarcomas, when compared to the expression levels in the carcinomas (Fig. 2). For example, SIP1/ZEB2, paired related homeobox 1 and paired like transcription factor 2 isoform C were upregulated in both osteo and fibrosarcomas. Further, MSX-2, special AT-rich sequence-binding protein 2 (SatB2), Goosecoid, DLX-5 and DLX-6, all belonging to the homeobox gene family, were markedly upregulated in osteosarcomas as opposed to NMT/carcinomas. Interestingly, two of the simple carcinomas, 203B and P208, clustered together with the sarcomas in terms of homeobox gene expression (Fig. 2). For example, 203B showed expression of the homeobox genes DLX-5, DLX-6 and hox A5, whereas P208 expressed Paired related homeobox 1 and MSX-2. These findings indicate a distinct phenotype of 203B and P208 in comparison with the remaining carcinomas. This is also supported by lower expression of the luminal cytokeratins 7, 8 and 18 in the P208 tumor as compared with the other carcinomas (suppl. Fig. 1). Together with expression of cytokeratin 5, a basal

cell marker, these results indicate that the P208 tumor may constitute a basal cell carcinoma and this notion is also supported by morphological criteria (see Fig. 3i). The 203B carcinoma differed from the other carcinomas in being lipid-rich (Table 1). Possibly, the different homeobox gene expression profile of this carcinoma might be related to this feature.

Next, we used immunohistochemistry to confirm the upregulated homeobox expression at the protein level, by analyzing three different homeobox transcription factors: SIP1/ZEB2, MSX-2 and SatB2. Overall, the immunohistochemical analysis correlated well with the gene expression data, showing a higher percentage of homeobox-positive cells in the sarcomas, in particular the osteosarcomas, than in carcinomas (Fig. 3, Table 3). Further, the homeobox transcription factor staining intensity was generally higher in the osteosarcomas than in the carcinomas. The fibrosarcomas showed strong staining for SIP1/ZEB2, in agreement with the high mRNA levels in the fibrosarcomas (Fig. 2) and, moreover, the weak staining for SatB2 in the fibrosarcomas is in line with the low SatB2 mRNA levels. SatB2 staining was predominantly seen in undifferentiated cells in the osteosarcomas, whereas the more differentiated cells showed lower expression (Fig. 3c). This is well in line with the reported expression of SatB2 in undifferentiated cells from calvarial sutures

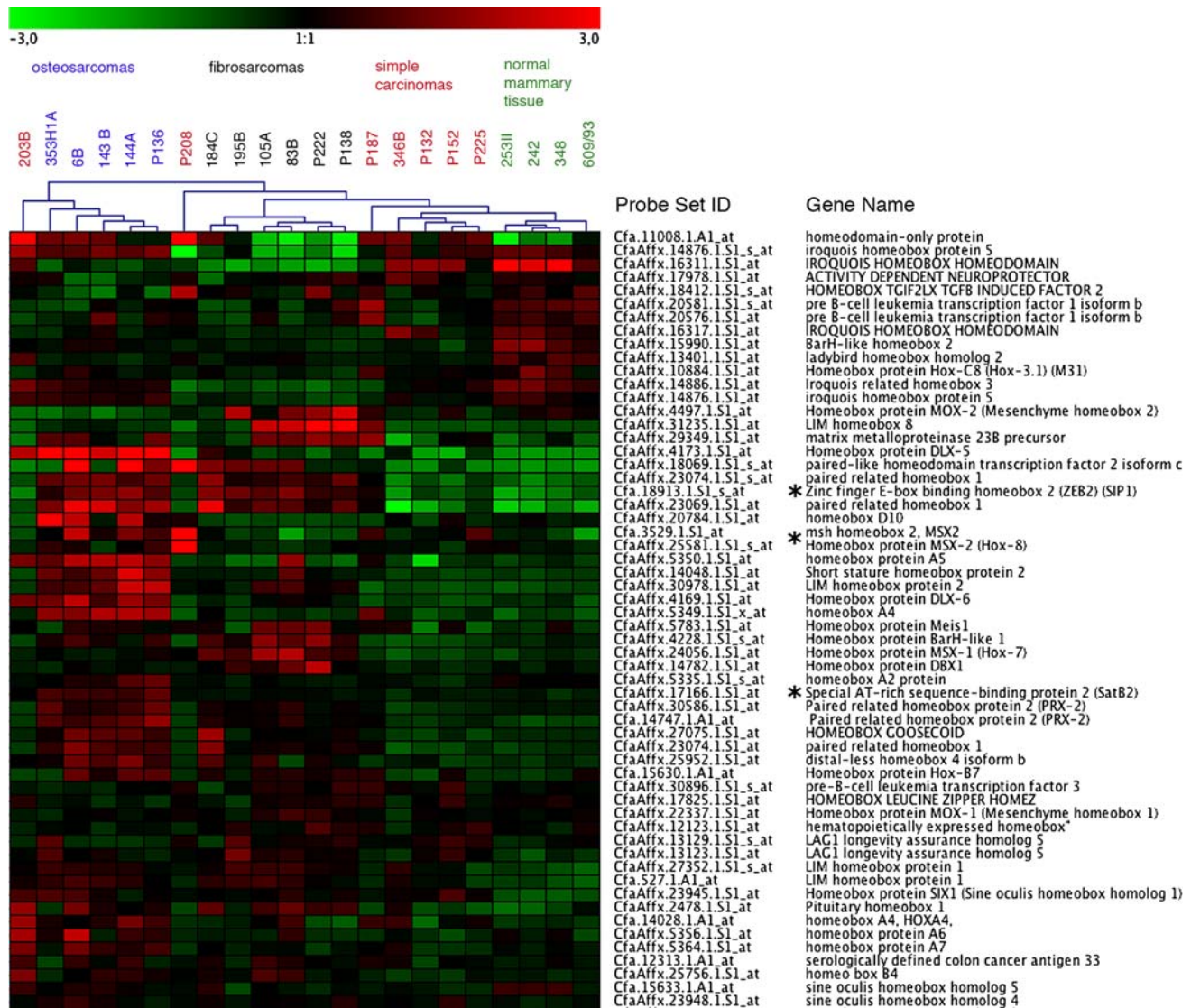


Fig. 2 Clustering of homeobox genes, in total 206 genes. Genes with visible change of expression are shown. Genes used for verification with immunohistochemistry (see Fig. 3 and Table 3) are marked with asterisks

[39]. Only two out of six studied fibrosarcomas were positive for MSX-2 (Table 3), in good agreement with the relatively low levels of MSX-2 mRNA in fibrosarcomas (Fig. 2). Although the majority of the carcinomas were negative or only weakly positive for MSX-2 protein, one of the carcinomas (P208) was positive (Fig. 3i), with the latter being well in agreement with the mRNA expression profile (Fig. 2). It should also be emphasized that the carcinomas, despite low SIP1/ZEB2 mRNA expression (see Fig. 2), showed positive staining for SIP1/ZEB2, although the number of positive cells and the intensity of staining was much lower than in the osteosarcomas (Fig. 3, Table 3). Taken together, the immunohistochemistry data thus support an upregulation of homeobox expression in sarcomas, particularly in osteosarcomas.

Retinoic acid is well known to regulate homeobox genes in development [40, 41]. In addition, bone morphogenetic proteins (BMPs) are known to upregulate homeobox gene expression and are also implicated in other aspects of homeobox transcription factor function [22, 42–44]. Hence, we can hypothesize that upregulated homeobox gene expression in the sarcomas may be accompanied by expression of genes related to BMP and retinoic acid signaling. Next, we investigated this hypothesis. Indeed, expression of several BMPs, BMP-1, -2, -4, -5 and -6, was evident in the osteosarcomas. For example, BMP-1 was expressed in the majority of the osteo and fibrosarcomas and expression of BMP-2 and BMP-5 were seen in several of the fibrosarcomas as well as in the carcinomas P208 and P225 (Fig. 4a). Further, BMPs were among the genes that

Fig. 3 Expression of homeobox proteins verified by immunohistochemistry. Osteosarcoma 143B (a), 144A (b) and 6B (c). Fibrosarcoma 184C (d), 184C (e) and 83B (f). Simple carcinoma P132 (g), 346B (h), P208 (i) and P132 (j). Antibodies: SIP1/ZEB2 (a, d, g), MSX-2 (b, e, h, i), SatB2 (c, f, j)

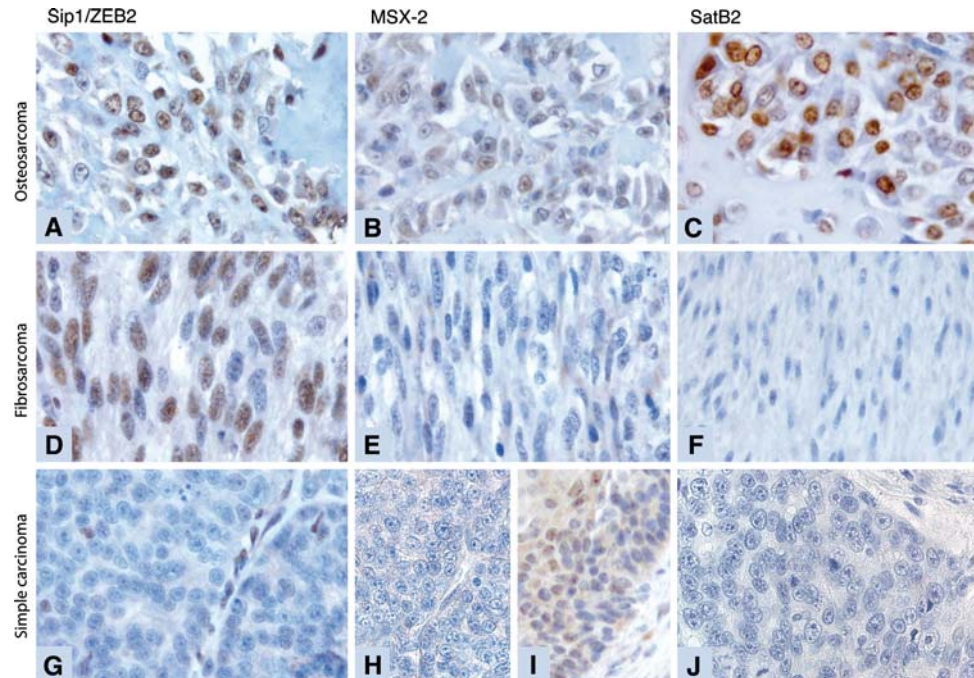


Table 3 Nuclear immunolabeling pattern of the tumor cells in the studied mammary tumors

Antibody	SIP1/ZEB2	MSX-2	SatB2
Osteosarcomas			
6B	4, +	1, +	3, ++
P136	4, ++	3, ++	2, ++
143B	4, ++	4, +	4, ++
144A	4, ++	4, ++	4, ++
353H1A	4, ++	0	3, ++
Fibrosarcomas			
83B	3, ++	0	0
105A	4, ++	0	ND
P138	3, ++	0	0
184C	3, ++	3, +	0
195B	3, +++	2, ++	0
P222	3, ++	0	0
Carcinomas			
P132	1, ++	0	0
P152	2, +	0	0
P187	Anaplastic: 2, ++; Ductal: 3, +	Anaplastic: 1, ++; Ductal: 2, +	0; 0
203B	1, ++	0	0
P208	3, ++	3, +	0
P225	2, ++	1, +	0
346B	1, +	0	2, +

Number of labeled cells estimated in the intervals: 1 = 1–25%; 2 = 26–50%; 3 = 51–75%; 4 = 76–100%. Intensity of immunolabeling: 0 = none; + = weak; ++ = moderate; +++ = strong

ND not determined

showed the highest extent of upregulation when comparing the global gene expression pattern of osteosarcomas with that of both carcinomas and NMT (data not shown). It was also evident that a number of genes related to retinoic acid signaling, e.g., retinoic acid receptor beta (RAR-beta) and Retinoic acid receptor responder 1 isoform 1, were upregulated in osteosarcomas (Fig. 4b). Noteworthy, the latter genes were also highly expressed in two of the carcinomas, 346B and P208. Moreover, retinoic acid binding protein 1 was highly expressed in the fibrosarcomas (Fig. 4b).

The presented results implicate increased homeobox transcription factor expression in mammary osteosarcomas. Next, we studied the homeobox protein expression and subcellular localization in response to growth factor stimulation. An osteosarcoma cell line (derived from the primary osteosarcoma 353H1A included in this study) was stimulated by BMP-2, a growth factor fundamental for development and bone formation [45]. BMPs are known to activate the Smad pathway, leading to Smad phosphorylation [46], and previous studies have shown that Smad proteins can interact with the homeobox transcription factor SIP1/ZEB2 [47]. We therefore studied Smad and homeobox protein expression and cellular localization in response to BMP stimulation. As shown in Fig. 5, non-stimulated cells were only weakly positive for phosphorylated Smad 1/5, (Fig. 5a). However, upon stimulation with BMP-2, a markedly increased staining for phosphorylated Smad 1/5 was seen and, interestingly, the staining was predominantly localized to the nucleus (Fig 5b).

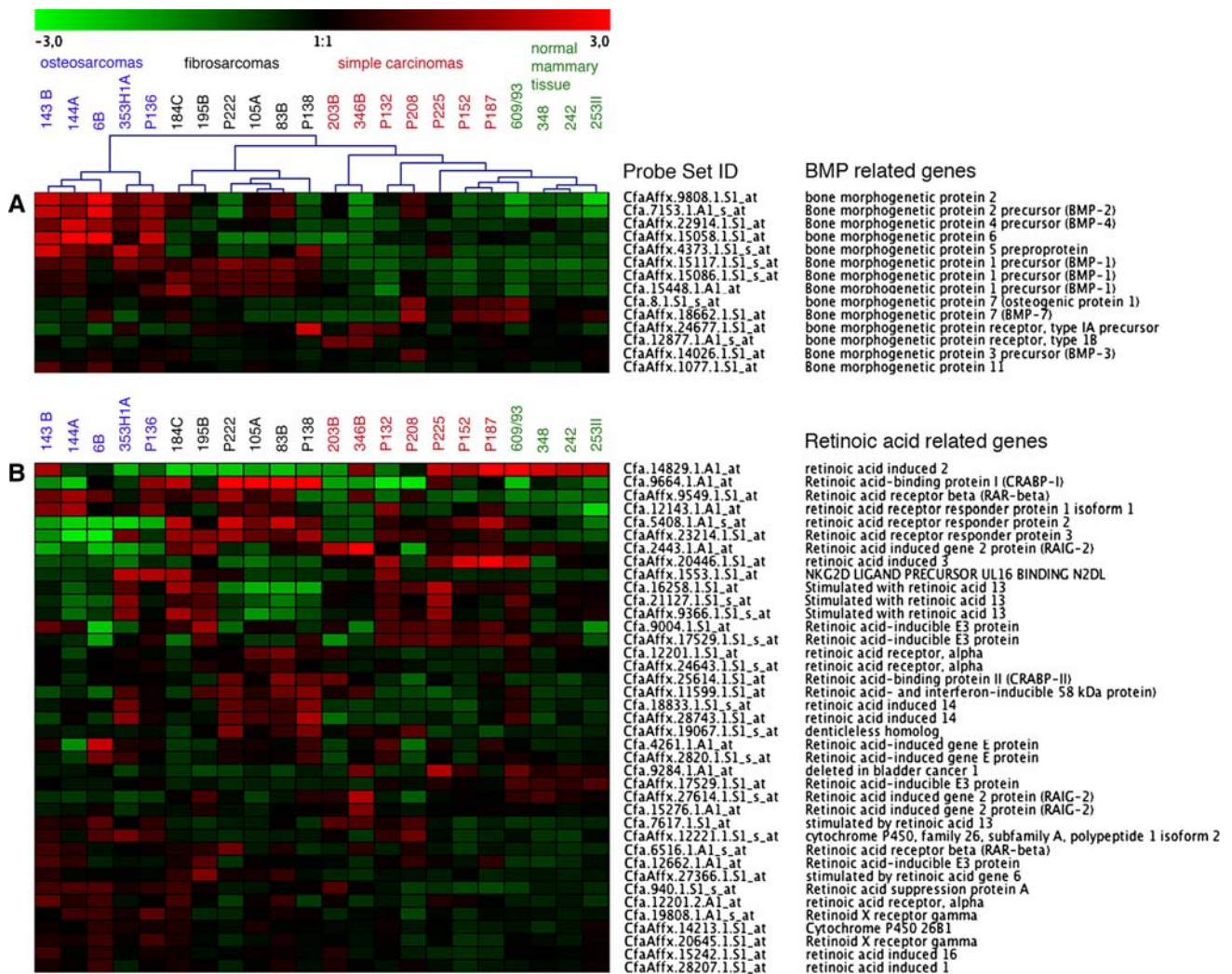


Fig. 4 Supervised clustering of genes related to BMP (a) and retinoic acid signaling (b). Genes with visible change of expression are shown

Non-stimulated cells were also positive for SIP1/ZEB2, being located in the nucleus (Fig. 5c). However, upon BMP-2 stimulation, an increased nuclear staining for SIP1/ZEB2 was seen, together with the appearance of SIP1/ZEB2 staining in the cytoplasm and the plasma membrane (Fig. 5d). Thus, BMP stimulation leads to increased SIP1/ZEB2 protein expression accompanied by an altered cellular localization, implicating SIP1/ZEB2 in the BMP signal transduction pathway in this osteosarcoma cell line. The osteosarcoma cell line was also positive for Goosecoid, another homeobox transcription factor, in agreement with the high levels of Goosecoid mRNA seen in the various osteosarcomas (Fig. 2). Goosecoid staining was predominantly seen in the nucleus (Fig. 5e, f) and was not increased upon BMP-2 stimulation, indicating constitutive expression.

It is striking to note that the majority of the homeobox genes that were upregulated in the osteosarcomas are known to be involved in craniofacial bone development

(suppl. Table 2), e.g., Goosecoid [48], SatB2 [21], MSX1 and MSX2 [49, 50], DLX5 and DLX6 [51], paired related homeobox 1 [52] and SIP1/ZEB2 [53]. The cells forming the craniofacial bone are neural crest-derived [44, 54] and it is worth noting that neural crest genes have been found to be upregulated in synovial sarcomas [55, 56]. Interestingly, the same homeobox and BMP genes (MSX2, Iroquois homeobox protein 5, BMP-2 and -7) together with retinoic acid signaling genes identified in synovial sarcomas [55], were also expressed in the sarcomas used in this study. Induction of neural crest genes typically involves signaling through the retinoic acid pathway and the latter is crucial for migrating neural cells in development. The osteosarcomas were found to express genes involved in retinoic acid signaling (Fig 4b) as well as an upregulation of growth factors of the BMP family (Fig. 4a), well in line with the known involvement of BMPs in bone formation as well as with the connection between BMPs and retinoic acid signaling seen in neural crest development. In fact, the

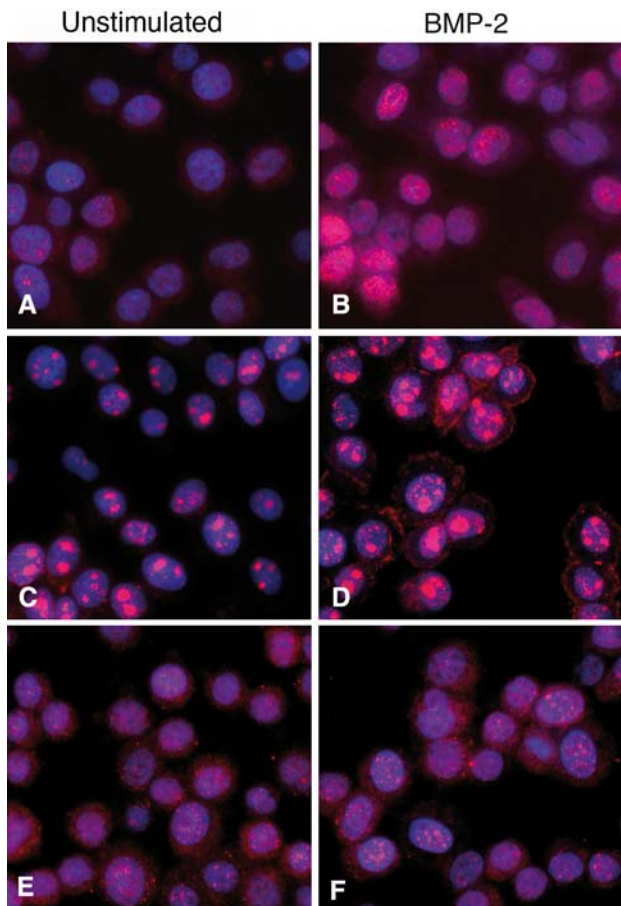


Fig. 5 Immunofluorescence staining of a mammary osteosarcoma cell clone. Unstimulated cells were stained for phosphorylated Smad 1/5 (a), SIP1/ZEB2 (c) and Goosecoid (e). BMP-2 stimulated cells were stained for phosphorylated Smad 1/5 (b), SIP1/ZEB2 (d) and Goosecoid (f)

significant expression of genes involved in craniofacial bone and neural crest development can explain why we see a number of neural genes expressed, e.g., Microtubule-associated protein 1A and transcription factor COE1, which are involved in dendritic branching and development of medium spiny neurons, respectively (see suppl. Fig 1 and Table 2). It is also noteworthy that the sarcomas expressed tooth markers (dentin sialophosphoprotein, amelogenin), and glial (meteorin) genes (data not shown). Taken together, we may thus suggest that the development of sarcoma specifically involves triggering of a set of homeobox genes related to neural crest and craniofacial bone development.

Acknowledgments We would like to thank Maria Rydåker for technical assistance. This work was financed by a Strategic funding from the Swedish University of Agricultural Sciences (SLU), Uppsala, Sweden to E. Hellmén.

Author contribution H. Wensman, E. Hellmén, H. Göransson, A. Isaksson, G.R. Rutteman and N.E. Heldin designed the study.

E. Hellmén did the pathological evaluation of the tumors. H. Wensman prepared the RNA and RNA quality controls. H. Göransson performed the array data normalization, background adjustments and statistical tests. H. Wensman set up the main direction for the manuscript and array data analysis and G. Pejler, H. Göransson, N.E. Heldin, E. Hellmén and A. Isaksson contributed to the data analysis. S. Strömberg and F. Pontén contributed with novel, validated antibodies. H. Wensman performed the IH verification and H. Wensman, E. Hellmén, S. Strömberg and F. Pontén evaluated the IH results and controls. K. Leuchowius performed the immunofluorescence and the evaluation of those results. H. Wensman and G. Pejler wrote the manuscript. All Authors approved the final version of the manuscript.

References

1. Young JL Jr, Ward KC, Wingo PA et al (2004) The incidence of malignant non-carcinomas of the female breast. *Cancer Causes Control* 15:313–319. doi:10.1023/B:CACO.0000024224.70386.d1
2. Silver SA, Tavassoli FA (1998) Primary osteogenic sarcoma of the breast: a clinicopathologic analysis of 50 cases. *Am J Surg Pathol* 22:925–933. doi:10.1097/00000478-199808000-00002
3. Adem C, Reynolds C, Ingle JN et al (2004) Primary breast sarcoma: clinicopathologic series from the Mayo Clinic and review of the literature. *Br J Cancer* 91:237–241
4. Blanchard DK, Reynolds CA, Grant CS et al (2003) Primary nonphylloides breast sarcomas. *Am J Surg* 186:359–361. doi:10.1016/S0002-9610(03)00269-1
5. Gullett NP, Delman K, Folpe AL et al (2007) National surgical patterns of care: regional lymphadenectomy of breast sarcomas. *Am J Clin Oncol* 30:461–465. doi:10.1097/COC.0b013e31804b40f4
6. Fong Y, Coit DG, Woodruff JM et al (1993) Lymph node metastasis from soft tissue sarcoma in adults. Analysis of data from a prospective database of 1772 sarcoma patients. *Ann Surg* 217:72–77. doi:10.1097/0000658-199301000-00012
7. Lana SE, Rutteman GR, Withrow SJ (2007) Tumors of the mammary gland. In: Withrow SJ, Vail DM (eds) *Withrow & MacEwen's small animal clinical oncology*, 4th edn. Saunders, Missouri, pp 619–636
8. Bostock DE (1986) Canine and feline mammary neoplasms. *Br Vet J* 142:506–515
9. Paoloni M, Khanna C (2008) Translation of new cancer treatments from pet dogs to humans. *Nat Rev Cancer* 8:147–156. doi:10.1038/nrc2273
10. Lindblad-Toh K, Wade CM, Mikkelsen TS et al (2005) Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* 438:803–819. doi:10.1038/nature04338
11. Hoffman MM, Birney E (2007) Estimating the neutral rate of nucleotide substitution using introns. *Mol Biol Evol* 24:522–531. doi:10.1093/molbev/msl179
12. Chu LL, Rutteman GR, Kong JM et al (1998) Genomic organization of the canine p53 gene and its mutational status in canine mammary neoplasia. *Breast Cancer Res Treat* 50:11–25. doi:10.1023/A:1006010526813
13. Sorlie T, Perou CM, Tibshirani R et al (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 98:10869–10874. doi:10.1073/pnas.191367098
14. Sorlie T, Tibshirani R, Parker J et al (2003) Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA* 100:8418–8423. doi:10.1073/pnas.0932692100

15. Van't Veer LJ, Dai H, Van De Vijver MJ et al (2002) Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415:530–536. doi:[10.1038/415530a](https://doi.org/10.1038/415530a)
16. Buyse M, Loi S, Van't Veer L et al (2006) Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J Natl Cancer Inst* 98:1183–1192
17. Paik S, Shak S, Tang G et al (2004) A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 351:2817–2826. doi:[10.1056/NEJMoa041588](https://doi.org/10.1056/NEJMoa041588)
18. Fan C, Oh DS, Wessels L et al (2006) Concordance among gene-expression-based predictors for breast cancer. *N Engl J Med* 355:560–569. doi:[10.1056/NEJMoa052933](https://doi.org/10.1056/NEJMoa052933)
19. Baird K, Davis S, Antonescu CR et al (2005) Gene expression profiling of human sarcomas: insights into sarcoma biology. *Cancer Res* 65:9226–9235. doi:[10.1158/0008-5472.CAN-05-1699](https://doi.org/10.1158/0008-5472.CAN-05-1699)
20. Francis P, Namlos HM, Muller C et al (2007) Diagnostic and prognostic gene expression signatures in 177 soft tissue sarcomas: hypoxia-induced transcription profile signifies metastatic potential. *BMC Genomics* 8:73. doi:[10.1186/1471-2164-8-73](https://doi.org/10.1186/1471-2164-8-73)
21. Dobreva G, Chahrour M, Dautzenberg M et al (2006) SATB2 is a multifunctional determinant of craniofacial patterning and osteoblast differentiation. *Cell* 125:971–986. doi:[10.1016/j.cell.2006.05.012](https://doi.org/10.1016/j.cell.2006.05.012)
22. Phippard DJ, Weber-Hall SJ, Sharpe PT et al (1996) Regulation of *Msx-1*, *Msx-2*, *Bmp-2* and *Bmp-4* during foetal and postnatal mammary gland development. *Development* 122:2729–2737
23. Satoh K, Hovey RC, Malewski T et al (2007) Progesterone enhances branching morphogenesis in the mouse mammary gland by increased expression of *Msx2*. *Oncogene* 26:7526–7534. doi:[10.1038/sj.onc.1210555](https://doi.org/10.1038/sj.onc.1210555)
24. Chen H, Sukumar S (2003) Role of homeobox genes in normal mammary gland development and breast tumorigenesis. *J Mammary Gland Biol Neoplasia* 8:159–175. doi:[10.1023/A:1025996707117](https://doi.org/10.1023/A:1025996707117)
25. Elloul S, Elstrand MB, Nesland JM et al (2005) Snail, Slug, and Smad-interacting protein 1 as novel parameters of disease aggressiveness in metastatic ovarian and breast carcinoma. *Cancer* 103:1631–1643. doi:[10.1002/cncr.20946](https://doi.org/10.1002/cncr.20946)
26. Zhai Y, Kuick R, Nan B et al (2007) Gene expression analysis of preinvasive and invasive cervical squamous cell carcinomas identifies *HOXC10* as a key mediator of invasion. *Cancer Res* 67:10163–10172. doi:[10.1158/0008-5472.CAN-07-2056](https://doi.org/10.1158/0008-5472.CAN-07-2056)
27. Barnes GL, Javed A, Waller SM et al (2003) Osteoblast-related transcription factors *Runx2* (*Cbfa1/AML3*) and *MSX2* mediate the expression of bone sialoprotein in human metastatic breast cancer cells. *Cancer Res* 63:2631–2637
28. Han HJ, Russo J, Kohwi Y et al (2008) *SATB1* reprogrammes gene expression to promote breast tumour growth and metastasis. *Nature* 452:187–193. doi:[10.1038/nature06781](https://doi.org/10.1038/nature06781)
29. Hartwell KA, Muir B, Reinhardt F et al (2006) The *Spemann* organizer gene, *Gooseoid*, promotes tumor metastasis. *Proc Natl Acad Sci USA* 103:18969–18974. doi:[10.1073/pnas.0608636103](https://doi.org/10.1073/pnas.0608636103)
30. Comijn J, Berx G, Vermassen P et al (2001) The two-handed E box binding zinc finger protein *SIP1* downregulates *E-cadherin* and induces invasion. *Mol Cell* 7:1267–1278. doi:[10.1016/S1097-2765\(01\)00260-X](https://doi.org/10.1016/S1097-2765(01)00260-X)
31. Irizarry RA, Hobbs B, Collin F et al (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4:249–264. doi:[10.1093/biostatistics/4.2.249](https://doi.org/10.1093/biostatistics/4.2.249)
32. Smyth Gk (2004) Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3:Article3
33. Hochberg Y, Benjamini Y (1990) More powerful procedures for multiple significance testing. *Stat Med* 9:811–818. doi:[10.1002/sim.4780090710](https://doi.org/10.1002/sim.4780090710)
34. Sturn A, Quackenbush J, Trajanoski Z (2002) Genesis: cluster analysis of microarray data. *Bioinformatics* 18:207–208. doi:[10.1093/bioinformatics/18.1.207](https://doi.org/10.1093/bioinformatics/18.1.207)
35. Uhlen M, Bjorling E, Agaton C et al (2005) A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol Cell Proteomics* 4:1920–1932. doi:[10.1074/mcp.M500279-MCP200](https://doi.org/10.1074/mcp.M500279-MCP200)
36. Uhlen M, Ponten F (2005) Antibody-based proteomics for human tissue profiling. *Mol Cell Proteomics* 4:384–393. doi:[10.1074/mcp.R500009-MCP200](https://doi.org/10.1074/mcp.R500009-MCP200)
37. Persson U, Izumi H, Souchelnytskyi S et al (1998) The *L45* loop in type I receptors for TGF-beta family members is a critical determinant in specifying Smad isoform activation. *FEBS Lett* 434:83–87. doi:[10.1016/S0014-5793\(98\)00954-5](https://doi.org/10.1016/S0014-5793(98)00954-5)
38. Perez-Pinera P, Berenson JR, Deuel TF (2008) Pleiotrophin, a multifunctional angiogenic factor: mechanisms and pathways in normal and pathological angiogenesis. *Curr Opin Hematol* 15:210–214. doi:[10.1097/MOH.0b013e3282fd6c9e](https://doi.org/10.1097/MOH.0b013e3282fd6c9e)
39. Coussens Ak, Hughes Ip, Wilkinson Cr et al (2008) Identification of genes differentially expressed by prematurely fused human sutures using a novel in vivo-in vitro approach. *Differentiation* 76(5):531–545
40. Daftary GS, Taylor HS (2006) Endocrine regulation of *HOX* genes. *Endocr Rev* 27:331–355. doi:[10.1210/er.2005-0018](https://doi.org/10.1210/er.2005-0018)
41. Williams SS, Mear JP, Liang HC et al (2004) Large-scale reprogramming of cranial neural crest gene expression by retinoic acid exposure. *Physiol Genomics* 19:184–197. doi:[10.1152/physiolgenomics.00136.2004](https://doi.org/10.1152/physiolgenomics.00136.2004)
42. Liem KF Jr, Tremml G, Roelink H et al (1995) Dorsal differentiation of neural plate cells induced by BMP-mediated signals from epidermal ectoderm. *Cell* 82:969–979. doi:[10.1016/0092-8674\(95\)90276-7](https://doi.org/10.1016/0092-8674(95)90276-7)
43. Tribulo C, Aybar MJ, Nguyen VH et al (2003) Regulation of *Msx* genes by a *Bmp* gradient is essential for neural crest specification. *Development* 130:6441–6452. doi:[10.1242/dev.00878](https://doi.org/10.1242/dev.00878)
44. Trainor PA, Melton KR, Manzanares M (2003) Origins and plasticity of neural crest cells and their roles in jaw and craniofacial evolution. *Int J Dev Biol* 47:541–553
45. Bandyopadhyay A, Tsuji K, Cox K et al (2006) Genetic analysis of the roles of *BMP2*, *BMP4*, and *BMP7* in limb patterning and skeletogenesis. *PLoS Genet* 2:e216. doi:[10.1371/journal.pgen.0020216](https://doi.org/10.1371/journal.pgen.0020216)
46. Von Bubnoff A, Cho KW (2001) Intracellular BMP signaling regulation in vertebrates: pathway or network? *Dev Biol* 239:1–14. doi:[10.1006/dbio.2001.0388](https://doi.org/10.1006/dbio.2001.0388)
47. Verschuere K, Remacle JE, Collart C et al (1999) *SIP1*, a novel zinc finger/homeodomain repressor, interacts with Smad proteins and binds to 5'-CACCT sequences in candidate target genes. *J Biol Chem* 274:20489–20498. doi:[10.1074/jbc.274.29.20489](https://doi.org/10.1074/jbc.274.29.20489)
48. Yamada G, Mansouri A, Torres M et al (1995) Targeted mutation of the murine *gooseoid* gene results in craniofacial defects and neonatal death. *Development* 121:2917–2922
49. Ishii M, Han J, Yen HY et al (2005) Combined deficiencies of *Msx1* and *Msx2* cause impaired patterning and survival of the cranial neural crest. *Development* 132:4937–4950. doi:[10.1242/dev.02072](https://doi.org/10.1242/dev.02072)
50. Hosokawa R, Urata M, Han J et al (2007) TGF-beta mediated *Msx2* expression controls occipital somites-derived caudal region of skull development. *Dev Biol* 310:140–153. doi:[10.1016/j.ydbio.2007.07.038](https://doi.org/10.1016/j.ydbio.2007.07.038)
51. Robledo RF, Rajan L, Li X et al (2002) The *Dlx5* and *Dlx6* homeobox genes are essential for craniofacial, axial, and

- appendicular skeletal development. *Genes Dev* 16:1089–1101. doi:[10.1101/gad.988402](https://doi.org/10.1101/gad.988402)
52. Lu MF, Cheng HT, Kern MJ et al (1999) prx-1 functions cooperatively with another paired-related homeobox gene, prx-2, to maintain cell fates within the craniofacial mesenchyme. *Development* 126:495–504
53. Van De Putte T, Maruhashi M, Francis A et al (2003) Mice lacking ZFHX1B, the gene that codes for Smad-interacting protein-1, reveal a role for multiple neural crest cell defects in the etiology of Hirschsprung disease-mental retardation syndrome. *Am J Hum Genet* 72:465–470. doi:[10.1086/346092](https://doi.org/10.1086/346092)
54. Golding JP, Trainor P, Krumlauf R et al (2000) Defects in pathfinding by cranial neural crest cells in mice lacking the neuregulin receptor ErbB4. *Nat Cell Biol* 2:103–109. doi:[10.1038/35000058](https://doi.org/10.1038/35000058)
55. Nielsen TO, West RB, Linn SC et al (2002) Molecular characterisation of soft tissue tumours: a gene expression study. *Lancet* 359:1301–1307. doi:[10.1016/S0140-6736\(02\)08270-3](https://doi.org/10.1016/S0140-6736(02)08270-3)
56. Tschöep K, Kohlmann A, Schlemmer M et al (2007) Gene expression profiling in sarcomas. *Crit Rev Oncol Hematol* 63:111–124. doi:[10.1016/j.critrevonc.2007.04.001](https://doi.org/10.1016/j.critrevonc.2007.04.001)