Award Number: W81XWH-11-1-0474

TITLE: **The Breast Cancer DNA Interactome**

PRINCIPAL INVESTIGATOR: Andrew R. Hoffman

CONTRACTING ORGANIZATION: Palo Alto Institute for Research and Education
Palo Alto, CA 94304

REPORT DATE: December 2014

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
December 2014	Final	30 Sep 2011 - 29 Sep 2014

4. TITLE AND SUBTITLE
The Breast Cancer DNA Interactome

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-11-1-0474

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Andrew R. Hoffman

Email: arhoffman@stanford.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
. Palo Alto Institute for Research and Education
Palo Alto, CA 94304

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Gene transcription may be regulated by remote enhancer regions through chromosome looping. The role that higher order chromatin organization plays in cancer progression and metastasis is not yet fully understood. Our major goal has been to characterize physical interactions among selected breast cancer gene loci in normal and malignant mammary cell lines. We used IGFBP3, a gene that has been involved in breast cancer pathogenesis, as bait in a 4C-seq experiment comparing normal breast cells (HMEC) and two different breast cancer cell lines (MCF7, an ER positive cell line and MDA-MB-231, a triple negative cell line). We found that in HMEC the genes BCAS1-4 (breast carcinoma amplified sequence 1-4) were found in the top 10 most significantly enriched regions for interactions with IGFBP3. We also found EGFR (epidermal growth factor receptor), a gene that has also been implicated in cancer, to interact significantly with IGFBP3 in all three samples. We also characterized IRAIN, a noncoding RNA important in long range interactions. These data suggest an important role for chromosomal interactions in the pathogenesis of breast cancer.

15. SUBJECT TERMS- nothing listed

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	39	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Introduction

Nuclear architecture is the new dimension of regulatory control, functioning in conjunction with genome organization and epigenetic marks. A full understanding of a cell's genetic repertoire cannot be discerned from linear sequence analysis alone. Instead, we must have a full understanding of the three dimensional nature of the human genome. Dynamic interactions occur among DNA elements, which can regulate gene expression over large genomic distances on a single chromosome, through DNA looping, or even between chromosomes. We propose that incorporating new knowledge regarding a breast cancer gene's spatial interactions (i.e., the nuclear neighborhood within which the genes reside) will yield novel and more accurate predictions of breast cancer susceptibility and suggest innovative therapeutic options.

Nuclear architecture is maintained through proteins and long noncoding RNAs that bind to DNA and stabilize loops and long range interactions.

Body

Task 1: Characterize physical interactions between selected breast cancer loci in normal and malignant mammary cell lines. (Months 1 - 36)

Insulin-like growth factor binding protein 3 (IGFBP3) has been implicated in breast cancer pathogenesis (1-5). IGFBP3 modulates cell growth and survival through binding to insulin-like growth factors I and II, and regulating their bioavailability (6). IGFBP3 has also been proposed to function independently of IGF and act as a growth modulator (7-9). While correlations between serum levels of IGFBP3 and breast cancer have yielded contradictory results (3-5, 10), increased levels of IGFBP3 in breast cancer tissue is correlated with a worse prognosis and poor clinical features (1,2).

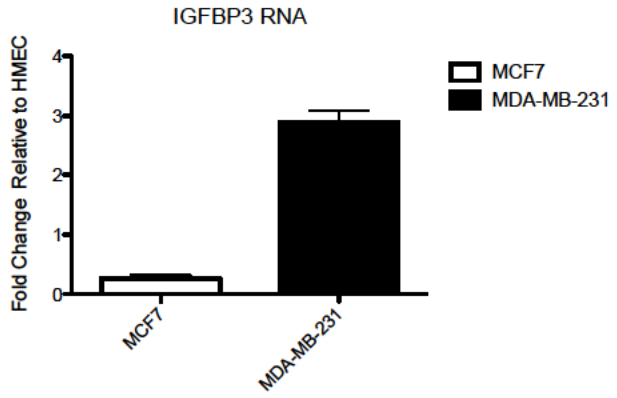
Dysregulation of IGFBP3 expression and hypermethylation of its promoter have been observed in many cancers (29). High levels of IGFBP3 expression was observed to increase survival of breast cancer cells exposed to environmental stress. We hypothesized that cancer-related changes in IGFBP3 regulation might coincide with altered spatial positioning and long-range DNA interactions contributing to breast cancer pathogenesis. We therefore used the IGFBP3 enhancer as bait in circular chromosome conformation capture with high throughput sequencing (4C-seq) in normal mammary epithelial cells (HMEC) and two breast cancer cell lines, MCF7 and MDA-MB-231, with opposite IGFBP3 expression profiles.

Expression of IGFBP3 is downregulated in MCF7, but upregulated in MDA-MB-231 relative to HMEC.

To better understand the role of IGFBP3 in breast cancer we analyzed its expression in primary breast cells, the estrogen receptor alpha (ERα) positive breast cancer cell line MCF7, and the triple-negative breast cancer cell line MDA-MB-231. IGFBP3 expression was increased nearly 3-fold in MDA-MB-231, and reduced 3.8-fold in MCF7, relative to HMEC (Figure 1A). To evaluate whether DNA methylation correlated with the changes in expression, we examined the methylation status of the IGFBP3 promoter by bisulfite pyrosequencing. The

IGFBP3 promoter was hypermethylated (91% CpG methylation) in MCF7 compared with 11% and 10% CpG methylation in HMEC and MDA-MB-231, respectively (figure 1B).

A



B

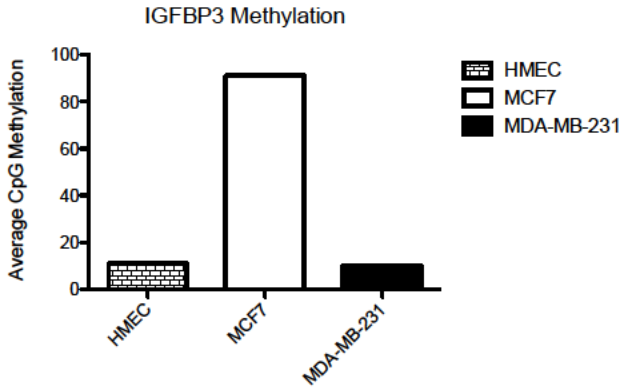


Figure 1. Expression and methylation status of IGFBP3
A) qRT-PCR: RNA levels of *IGFBP3* were measured in MCF-7, MDA-MB-231 and HMEC cells. Expression in cancer lines was plotted as fold change relative to HMEC. Data represent the SEM of three independent biological replicates. B) Percent methylation of CpG nucleotides in the *IGFBP3* promoter in HMEC, MCF-7 and MDA-MB-231. Bars represent the average percent methylation of 2-6 positions in the *IGFBP3* promoter.

EGFR interacts significantly with IGFBP3

To identify whether changes in IGFBP3 expression and methylation were accompanied by global alteration of its long-range chromatin interactions, we performed multiplex 4C-seq in HMEC, MCF7 and MDA-MB-231. We chose a region upstream of IGFBP3, classified as a strong enhancer in HMEC by chromatin profiling of several distinctive features including enrichment of the enhancer mark H3K4me1, as our bait . We obtained a combined total of approximately 12 million mapped reads for the three samples with the majority mapping in cis. The 4C-seq reads were binned into windows based on the number of mappable HindIII restriction sites ranging from 25 to 400. Regions with a FDR below 0.01 were considered significantly interacting. The significant long-range cis interactions for window size 100 in HMEC, MCF7 and MDA-MB-231 are diagrammed (Figure 2A). For every window size analyzed, MCF7 contained the largest number of significant interactions, followed by MDA-MB-231 and HMEC. Within a window size of 100, there were a total of 16 significant cis interactions in HMEC, 51 in MCF7 and 29 in MDA-MB-231. Of these interactions 8 were common to all samples.

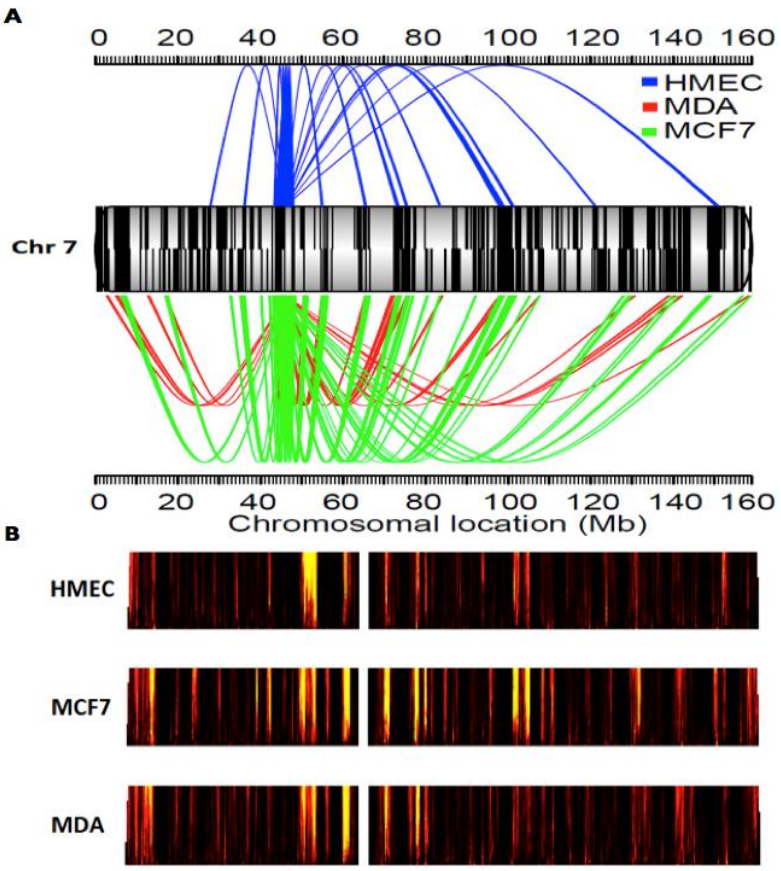


Figure 2.

Among the significant intrachromosomal interactions common to all samples, and across all window sizes, was an interaction with epidermal growth factor receptor (EGFR), another breast cancer related gene. EGFR is located approximately 9 Mb from IGFBP3 on chromosome 7. To examine this long-range interaction in more detail, we labeled gene pairs EGFR and IGFBP3 by 3D-FISH in HMEC and

breast cancer cell lines MCF7 and MDA-MB-231 (figure 3A). To quantitate differences in interaction frequencies at the cellular level, we measured the center-to-center distances between the closest pairs of labeled foci. In 88% of HMEC nuclei counted EGFR and IGFBP3 were within 1 micron (Figure 3B). This was reduced to 56% of MCF7 nuclei, and increased to 96% of MDA-MB-231 nuclei. To assess whether differences in spatial positioning were accompanied by changes in expression, we measured RNA levels of EGFR in HMEC, MCF7 and MDA-MB-231 by qRT-PCR (Figure 3C). Relative to HMEC, EGFR expression was unchanged in MDA-MB-231, yet was reduced 35-fold to nearly undetectable levels in MCF7 cells. In contrast to IGFBP3, the EGFR promoter had no change in CpG methylation (data not shown).

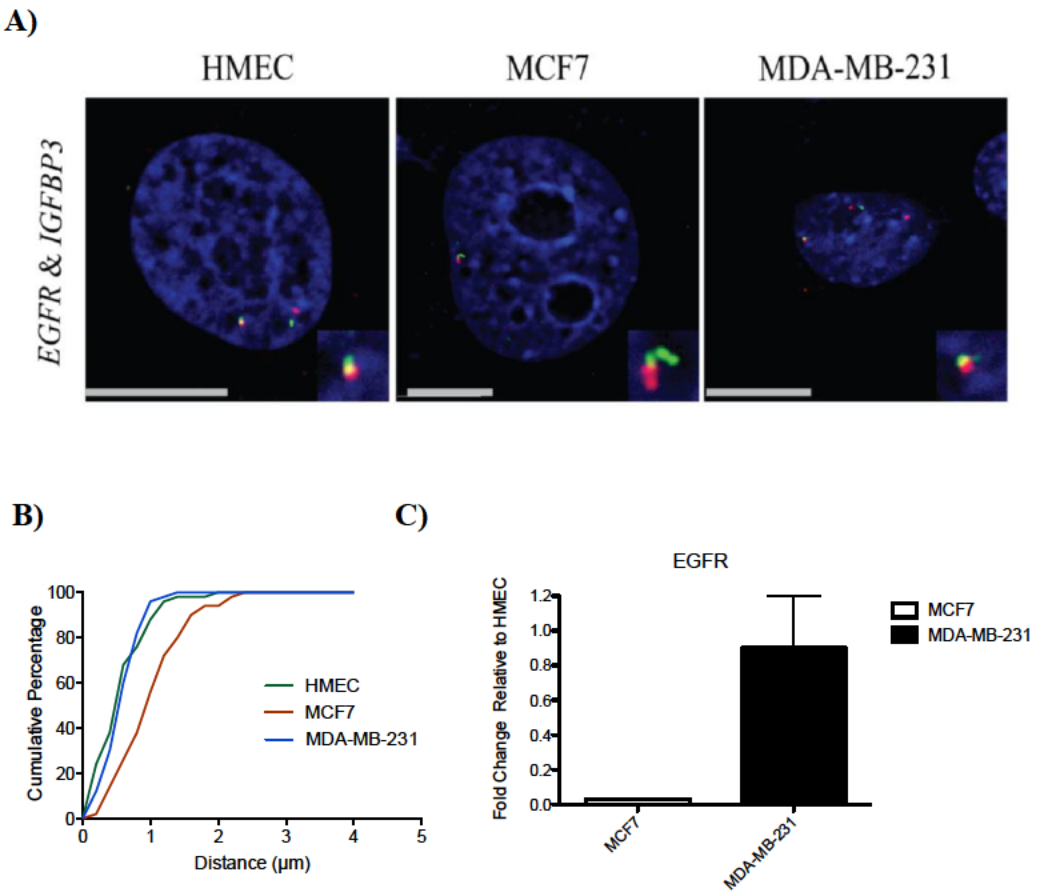


Figure 3.

We also discovered that recurrent breakpoints that map within HMEC 4C significant hits are also present within MCF-7 4C significant hits.

Some of the most significant 4C-seq interchromosomal interactions in HMEC included regions containing the genes *BCAS 1-4* located on chromosomes 1, 17 and 20. (10) All 4 of these genes were found among the 10 most significantly enriched regions in HMEC, and the region containing *BCAS1* and *ZNF217* was the overall top scoring window. These interactions were also enriched in MCF7, where they are frequently rearranged and amplified. We used 3D-FISH to investigate whether the *IGFBP3* interacting *BCAS* genes were also in close spatial proximity with one another prior to any oncogenic translocations (Figure 4). We performed dual and triple labeled 3D-FISH with probes for *IGFBP3*, *BCAS1*, *BCAS3* and *BCAS4* in primary HMEC cells (Figure 1A). Center-to-center distances were measured for the closest pairs of foci for each probe (Figure 1B). All probes targeting the *BCAS* genes were in close proximity, residing less than or equal to 1 micron to *IGFBP3* in at least 5% of nuclei. The *BCAS3-BCAS4* and *BCAS3-BCAS1* regions, which undergo translocations with one another in MCF7 were also within 1 micron in at least 4% of normal HMEC nuclei. These percentages are in line with reports of positive *trans* interacting loci identified using other molecular assays. This suggests spatial proximity of the *BCAS* genes in normal breast cells contributes to their frequent oncogenic translocations.

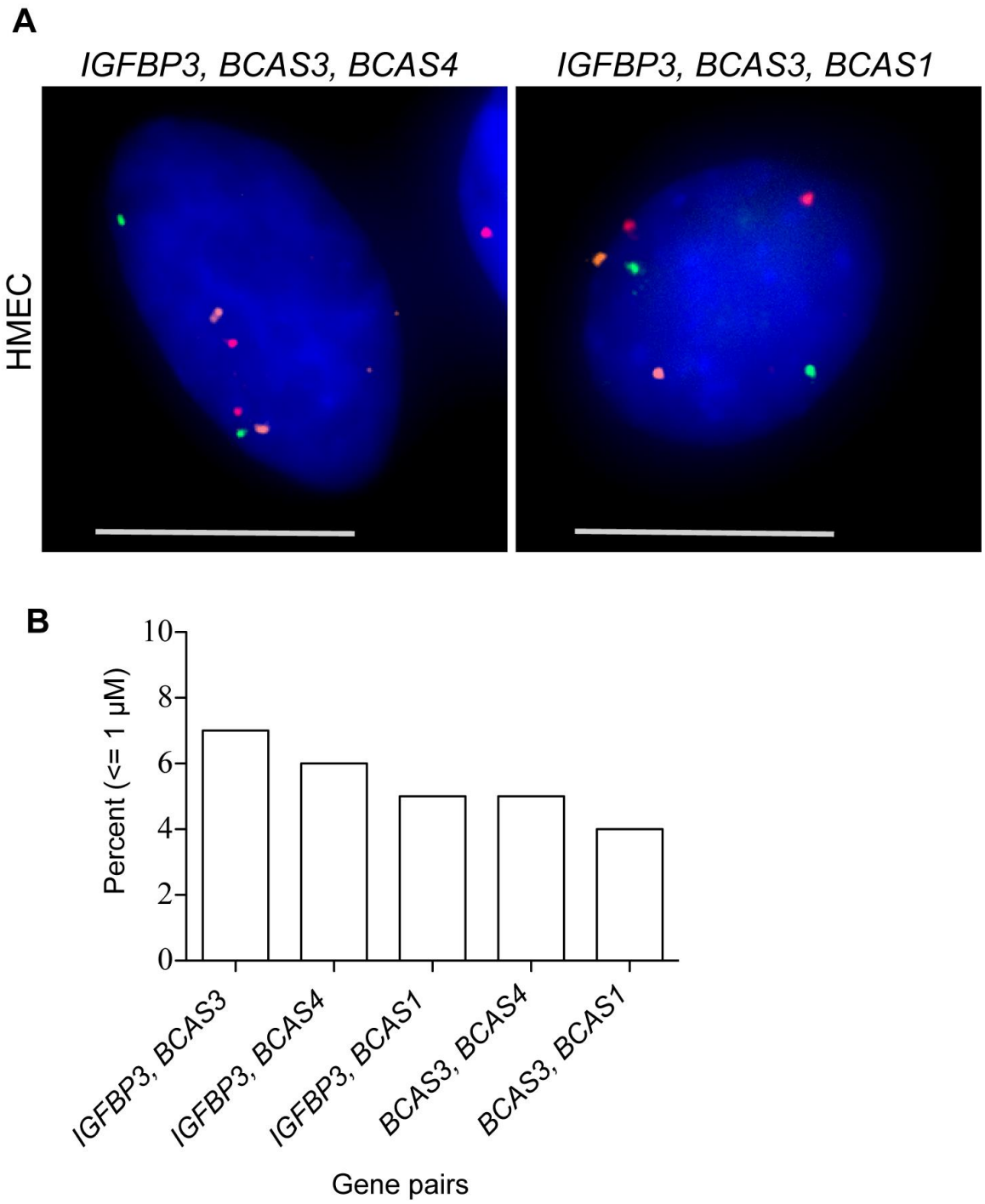


Figure 4. *IGFBP3* interacts with *BCAS* genes.
A, Representative triple labeled 3D-FISH, z-axis projection images of *IGFBP3, BCAS3, BCAS4* (left) and *IGFBP3, BCAS3, BCAS1* (right). Scale bar = 10 µm. B, Percentage of nuclei with the listed pair of gene loci within 1 micron of each other. Distances were measured between the closest two foci in each nucleus.

Our study demonstrates that long-range interactions of cancer-related loci, including *EGFR* and *IGFBP3*, are altered in breast cancer cells, and these alterations are frequently associated with epigenetic changes. Long-range interactions influence chromosomal translocations, and add an additional layer of

complexity to transcriptional and epigenetic regulation to coordinate gene expression. Therefore, a better understanding of aberrant chromatin interactions is needed to fully understand cancer pathology.

Task 2: Utilize a murine model of xenotropic tumor growth and metastasis to characterize the combinatorial contribution of multiple disease associated loci. (Months 12-24)

We were initially planning to study to use shRNA to knock out predicted genes in breast cancer cell lines and then use a mouse model to determine the effect of the knock out on tumor growth and metastasis. We were not able to successfully knock out the gene we were going to test (SAT1B) and because of limitations of time, we did not purse the murine model.

Therefore, we changed the scope of this task to explore a more fruitful avenue, and we elected to study another factor that stabilizes long-range chromatin interactions in breast cancer, a long noncoding RNA associated with the IGF1 receptor.

IGF1R is one of the most abundantly phosphorylated receptor tyrosine kinases in tumors (11). The insulin-like growth factor system, including the type I IGF receptor *IGF1R* and the mitogenic ligands IGF-I and IGF-II, is frequently dysregulated in breast cancer and is known to contribute to disease progression and metastasis (12). IGF-I and IGF-II promote cell growth and survival via the IGF1R receptor-mediated signal transduction through intracellular tyrosine kinase linked to the phosphatidyl-inositol-3 kinase (PI3K)-Akt-mammalian target of rapamycin (mTOR) pathway. Overexpression of *IGF1R* activates the PI3-K and MAPK signal cascades, resulting in cell proliferation and resistance to chemotherapeutic agents, radiation, and targeted therapies using Tamoxifen and Herceptin (13). Therapeutic agents targeting *IGF1R* are currently in clinical development(12, 14), including those that inhibit the IGF1R tyrosine kinase using monoclonal antibodies and small molecules. However, the clinical development of various IGF1R inhibitors has been put on hold due to lack of sufficient clinical efficacy. Thus, the regulation of this pathway needs to be further defined to aid in the development of next generation regimens.

Currently, the molecular mechanisms underlying the dysregulation of the *IGF1R* pathway in tumors remain unknown. Using a recently-developed R3C (**R**NA-guided **C**hromatin **C**onformation **C**apture) technique (15), we recently identified a novel non-coding RNA (lncRNA) *IRAIN* within the *IGF1R* locus (16). *IRAIN* is transcribed from an intragenic promoter located in the first intron of *IGF1R*. *IRAIN* lncRNA is transcribed in an antisense orientation compared with the IGF1R gene, and it is expressed exclusively from the paternal allele, with the maternal allele being silenced. Interestingly, this lncRNA interacts with chromatin DNA and is involved in the formation of an intrachromosomal enhancer/promoter loop. In addition, *IRAIN* was downregulated in leukemia cell lines and in leukocytes from patients with high-risk AML. These data suggested that *IRAIN*

might play a role in the dysregulation of the IGF pathway in hematopoietic malignancies.

However, the function of this noncoding RNA in other malignancies remains to be explored. The *IGF1R* pathway is frequently dysregulated in breast cancer. It is unclear if *IRAIN* lncRNA is aberrantly imprinted in breast cancer patients. In these experiments, we characterize the allelic expression of *IRAIN* lncRNA in a cohort of breast cancer samples.

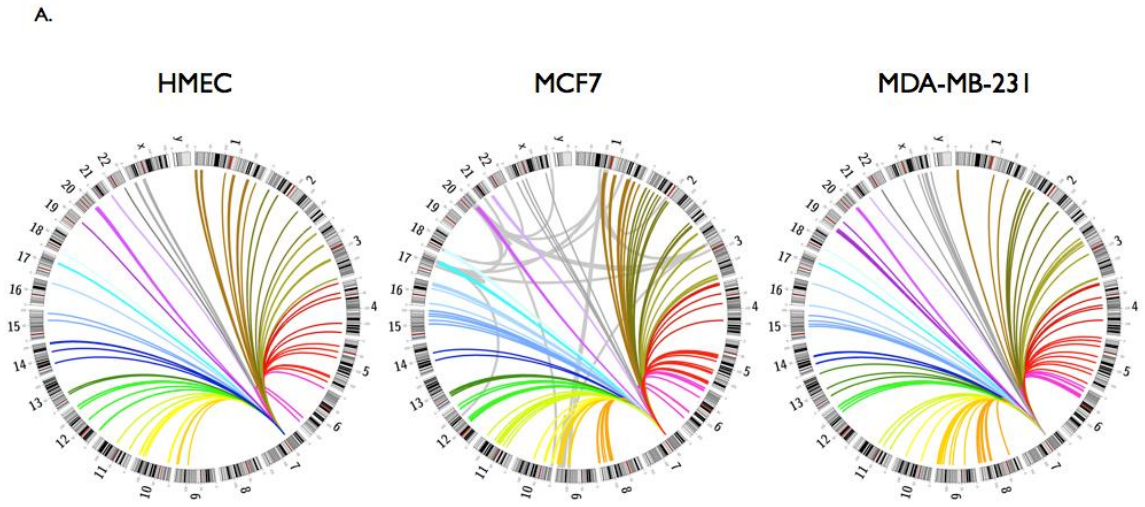
In breast cancer tissues, we found that IRAIN lncRNA was transcribed from an intronic promoter in an antisense direction as compared to the IGF1R coding mRNA. Unlike the IGF1R coding RNA, this noncoding RNA was imprinted, with monoallelic expression from the paternal allele. In breast cancer tissues that were informative for SNP rs8034564, there was an imbalanced expression of the two parental alleles, where the “G” genotype was favorably imprinted over the “A” genotype. In breast cancer patients, IRAIN was aberrantly imprinted in both tumors and peripheral blood leukocytes, exhibiting a pattern of allele-switch: the allele expressed in normal tissues was inactivated and the normally imprinted allele was expressed. Epigenetic analysis revealed that there was extensive DNA demethylation of CpG islands in the gene promoter. These data identify IRAIN lncRNA as a novel imprinted gene that is aberrantly regulated in breast cancer.

Task 3: Identify additional genomic sites that interact with our selected 3DAS loci. (Months 12-24)

Recurrent breakpoints that map within HMEC 4C significant hits are also present within MCF-7 4C significant hits

We constructed a circus plot to highlight the significant interchromosomal interactions involving the IGFBP3 enhancer in HMEC, MCF7 and MDA-MB-231 that fell within a window size of 200 (Figure 5A). There were a total of 87 significant interactions in HMEC, 194 in MCF7 and 115 in MDA-MB-231. Of these interactions only 11 were common to all samples (figure 5B). Because a large proportion of the significant 4C windows fell within chromosomes prone to rearrangements, fusions and amplifications, we compared the locations of a list of 157 breakpoints mapped in MCF7 cells to our significant interchromosomal 4C windows. The breakpoints could be categorized as 2 distinct types. The first category contained the majority of breakpoints, which were dispersed throughout the genome in regions of low copy repeats. The second category included MCF7 breakpoints falling within four highly amplified regions located on chromosomes 1, 3, 17 and 20. We found that breakpoints falling within our 4C windows were almost exclusively in the latter category. We considered a subset of 74 breakpoints, described as interchromosomal rearrangements , and determined how many of these fell within significant 4C windows in MCF-7. As a comparison we also mapped these breakpoints to our significant 4C windows in HMEC. A total of 29 breakpoint ends mapped within significant windows in HMEC, as compared to 61 in the MCF-7 line. Interestingly, all but 1 of the breakpoints within HMEC 4C windows was also present within MCF-7 4C hits. Also, when we

compared the number of breakpoints of which both ends of the breakpoint mapped to a 4C hit, the percentage was nearly twice as many in the breast cancer cell line MCF-7 as in HMEC.



B)

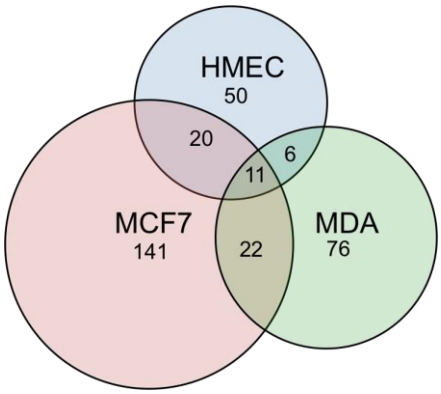


Figure 5

KEY RESEARCH ACCOMLISHMENTS

- Development of 4C-seq assays for breast cancer cells
- Demonstration that breast cancer cells differ from normal cells and from each other in their “interactome”
- Discovery of an imprinted long noncoding RNA (*IRAIN*) that is crucial in forming a loop between the IGF1-receptor promoter and enhancer and which is dysregulated in breast cancer.

REPORTABLE OUTCOMES

- Manuscripts:

Zeitz MJ, Ay F, Heidmann JD, Lerner PL, Noble WS, Steelman BN and Hoffman AR. Genomic interaction profiles in breast cancer reveal altered chromatin architecture. ***PLoS One*** Sep 3;8(9): e73974. doi: 10.1371/journal.pone.0073974. 2013

Kang L, Sun J, Wen X, Cui J, Wang G, Hoffman AR, Hu JF and Li W. Aberrant allele-switch imprinting of a novel *IGF1R* intragenic antisense non-coding RNA in breast cancers. ***Eur J Cancer*** 51: 260-270, 2015.

- Licenses: none
- Degrees obtained: n/a
- Development of cell lines, tissue or serum repositories: none
- Informatics: new sets of data regarding interchromosomal interactions
- Funding applied for based on this award: none
- Employment or research opportunities: none

CONCLUSIONS

Physical contact is a prerequisite for chromosomal translocations. Both cytogenetic and molecular evidence suggests spatial proximity influences recurrent chromosomal translocations. Long noncoding RNAs help stabilize long-range intra- and inter chromosomal interactions. We have described a novel long noncoding RNA derived from the IGF-1 receptor locus that stabilizes a loop between the IGF1R promoter and its enhancer, which is dysregulated in breast cancer.

Our data also demonstrate that there are numerous breast cancer genes present within significantly interacting regions in normal breast cells. These data suggest the possibility that certain loci in the genome form “hubs” of preferentially interacting loci. These hubs may have a functional purpose, such as being co-transcribed in “transcription factories.” It is likely that these interacting genes regulate each other’s’ transcription and that changes in long range interactions in cancer may lead to detrimental changes in gene expression. Breakpoint analysis suggests that when an interacting region undergoes a translocation an additional interaction detectable by 4C is gained. Overall, our data from multiple lines of evidence suggest an important role for long-range chromosomal interactions in

the pathogenesis of breast cancer, and it is possible that new gene targets for diagnosis or therapeutics may become evident from the study of interactome informatics.

REFERENCES

1. Rocha RL, Hilsenbeck SG, Jackson JG, VanDenBerg CL, Weng C, Lee AV, et al. Insulin-like growth factor binding protein-3 and insulin receptor substrate-1 in breast cancer: correlation with clinical parameters and disease-free survival. Clin Cancer Res. 1997;3:103-9.
2. Yu H, Levesque MA, Khosravi MJ, Papanastasiou-Diamandi A, Clark GM, Diamandis EP. Insulin-like growth factor-binding protein-3 and breast cancer survival. Int J Cancer. 1998;79:624-8.
3. Sugumar A, Liu YC, Xia Q, Koh YS, Matsuo K. Insulin-like growth factor (IGF)-I and IGF-binding protein 3 and the risk of premenopausal breast cancer: a meta-analysis of literature. Int J Cancer. 2004;111:293-7.
4. Baglietto L, English DR, Hopper JL, Morris HA, Tilley WD, Giles GG. Circulating insulin-like growth factor-I and binding protein-3 and the risk of breast cancer. Cancer Epidemiol Biomarkers Prev. 2007;16:763-8.
5. Key TJ, Appleby PN, Reeves GK, Roddam AW. Insulin-like growth factor 1 (IGF1), IGF binding protein 3 (IGFBP3), and breast cancer risk: pooled individual data analysis of 17 prospective studies. Lancet Oncol. 2010;11:530-42.
6. Firth SM, Baxter RC. Cellular actions of the insulin-like growth factor binding proteins. Endocr Rev. 2002;23:824-54.
7. Oh Y, Muller HL, Pham H, Rosenfeld RG. Demonstration of receptors for insulin-like growth factor binding protein-3 on Hs578T human breast cancer cells. J Biol Chem. 1993;268:26045-8.
8. Rajah R, Valentinis B, Cohen P. Insulin-like growth factor (IGF)-binding protein-3 induces apoptosis and mediates the effects of transforming growth factor-beta1 on programmed cell death through a p53- and IGF-independent mechanism. J Biol Chem. 1997;272:12181-8.
9. Mohan S, Baylink DJ. IGF-binding proteins are multifunctional and act via IGF-dependent and -independent mechanisms. J Endocrinol. 2002;175:19-31.
10. Tanner MM, Tirkkonen M, Kallioniemi A, Holli K, Collins C, et al. Amplification of chromosomal region 20q13 in invasive breast cancer: prognostic implications. Clinical cancer research : an official journal of the American Association for Cancer Research 1995; 1: 1455–1461.
11. Ge NL, Rudikoff S. Insulin-like growth factor I is a dual effector of multiple myeloma cell growth. Blood 2000;96(8):2856-2861.
12. Seccareccia E, Brodt P. The role of the insulin-like growth factor-I receptor in malignancy: an update. Growth Horm IGF Res 2012;22(6):193-199.
13. Chapuis N, Tamburini J, Cornillet-Lefebvre P, Gillot L, Bardet V, Willems L, et al. Autocrine IGF-1/IGF-1R signaling is responsible for constitutive PI3K/Akt activation in acute myeloid leukemia: therapeutic value of neutralizing anti-IGF-1R antibody. Haematologica 2010;95(3):415-423.

14. Karamouzis MV, Papavassiliou AG. Targeting insulin-like growth factor in breast cancer therapeutics. Crit Rev Oncol Hematol 2012;84(1):8-17.
15. Zhang H, Zeitz MJ, Wang H, Niu B, Ge S, Li W, et al. Long noncoding RNA-mediated intrachromosomal interactions promote imprinting at the Kcnq1 locus. J Cell Biol 2014;204(1):61-75.
16. Sun J, Li W, Sun Y, Yu D, Wen X, Wang H, et al. A novel antisense long noncoding RNA within the IGF1R gene locus is imprinted in hematopoietic malignancies. Nucleic Acids Res 2014;42(15):9588-9601.

APPENDICES: manuscripts

OPEN ACCESS Freely available online

PLOS ONE

Genomic Interaction Profiles in Breast Cancer Reveal Altered Chromatin Architecture

Michael J. Zeitz[1]*, Ferhat Ay[2], Julia D. Heidmann[1], Paula L. Lerner[1], William S. Noble[2,3☯], Brandon N. Steelman[1☯], Andrew R. Hoffman[1☯]

1 Department of Medicine, Veterans Affairs Palo Alto Health Care System, Stanford University Medical School, Palo Alto, California, United States of America, **2** Department of Genome Sciences, University of Washington, Seattle, Washington, United States of America, **3** Department of Computer Science and Engineering, University of Washington, Seattle, Washington, United States of America

Abstract

Gene transcription can be regulated by remote enhancer regions through chromosome looping either in *cis* or in *trans*. Cancer cells are characterized by wholesale changes in long-range gene interactions, but the role that these long-range interactions play in cancer progression and metastasis is not well understood. In this study, we used *IGFBP3*, a gene involved in breast cancer pathogenesis, as bait in a 4C-seq experiment comparing normal breast cells (HMEC) with two breast cancer cell lines (MCF7, an ER positive cell line, and MDA-MB-231, a triple negative cell line). The *IGFBP3* long-range interaction profile was substantially altered in breast cancer. Many interactions seen in normal breast cells are lost and novel interactions appear in cancer lines. We found that in HMEC, the breast carcinoma amplified sequence gene family (*BCAS*) 1–4 were among the top 10 most significantly enriched regions of interaction with *IGFBP3*. 3D-FISH analysis indicated that the translocation-prone *BCAS* genes, which are located on chromosomes 1, 17, and 20, are in close physical proximity with *IGFBP3* and each other in normal breast cells. We also found that epidermal growth factor receptor (*EGFR*), a gene implicated in tumorigenesis, interacts significantly with *IGFBP3* and that this interaction may play a role in their regulation. Breakpoint analysis suggests that when an *IGFBP3* interacting region undergoes a translocation an additional interaction detectable by 4C is gained. Overall, our data from multiple lines of evidence suggest an important role for long-range chromosomal interactions in the pathogenesis of cancer.

Citation: Zeitz MJ, Ay F, Heidmann JD, Lerner PL, Noble WS, et al. (2013) Genomic Interaction Profiles in Breast Cancer Reveal Altered Chromatin Architecture. PLoS ONE 8(9): e73974. doi:10.1371/journal.pone.0073974

Editor: Sumitra Deb, Virginia Commonwealth University, United States of America

Received April 2, 2013; **Accepted** July 24, 2013; **Published** September 3, 2013

Funding: This work was supported by a DOD Breast Cancer Research Program Idea Award (BC102122) and the Research Service of the Department of Veterans Affairs. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E mail: mjzeitz@stanford.edu

☯ These authors contributed equally to this work.

Introduction

It is now widely recognized that the spatial organization of the genome and not only its linear sequence is essential for normal genome function [1]. Recent breakthroughs combining high throughput DNA sequencing and molecular assays have revolutionized our understanding of chromatin organization [2,3,4]. Three dimensional chromatin structure is important in the regulation of transcription [5], and in the control of epigenetic states (including the regulation of imprinted genes) by means of chromosome looping between distant regulatory regions on the same or on different chromosomes [6,7]. Dynamic, long range interactions have been observed to regulate gene expression, contribute to the developmental processes of T cell differentiation and X inactivation, and may play a role in tumorigenesis [7,8,9,10,11]. The interchromosomal interaction between the *Ifng* promoter on chromosome 10 and the T_H2 cytokine gene locus on chromosome 11 in naive T cells maintains both loci in a configuration poised for rapid transcription and is thought to facilitate the developmental choice between T_H1 or T_H2 cells [8]. Transient homologous pairing of X inactivation centers early in development is crucial for correct X chromosome dosage compensation in mammalian females [9,10]. We have shown that *Igf2* on chromosome 7 interacts with the *Wsb1/Nf1* locus on chromosome 11, and disruption of this interaction results in decreased expression of *Wsb1* and *Nf1* [7]. We also observed a substantial alteration in chromatin structure within human cancers that have lost *IGF2* imprinting, resulting in a striking loss of long range interactions across the *IGF2/H19* locus [11]. These studies indicate that a better understanding of intricate 3D chromatin organization is crucial to understanding human diseases, particularly cancer, in which genomic instability and dysregulation are widespread.

Breast cancer is a complex disease that involves alterations in both genetic and epigenetic factors [12,13,14]. While numerous genetic mutations, translocations and aberrant DNA methylation have been reported in breast cancer, the role of long range interactions during cancer progression remains elusive. Recent evidence suggests that genome organization is altered early in breast tumorigenesis [15]. Cancer related genes were observed to change their radial positions in a cell culture model of early breast tumor development [15]. Changes in radial position of cancer related genes were also observed in breast tumor tissue samples, and were not caused by genomic instability [16].

Insulin like growth factor binding protein 3 (*IGFBP3)* has been implicated in breast cancer pathogenesis [17,18,19,20,21]. IGFBP3 modulates cell growth and survival by binding to insulin like growth factors I and II, and regulating their bioavailability [22]. IGFBP3 has also been proposed to function independently of IGF I or IGF II and act as a growth modulator [23,24,25]. While correlations between serum levels of IGFBP3 and breast cancer have yielded contradictory results [19,20,21,26], increased levels of IGFBP3 in breast cancer tissue is correlated with a worse prognosis and poor clinical features [17,18].

Dysregulation of *IGFBP3* expression and hypermethylation of its promoter have been observed in many cancers [27]. Increased *IGFBP3* expression has been shown to enhance survival of breast cancer cells exposed to environmental stress [28]. Alternatively, a mouse model of prostate cancer crossed with a knockout of *Igfbp3* displayed significant increase in metastasis in double mutant animals. *In vitro* assays of prostate cell lines derived from these mouse lines also indicated a more aggressive cancer phenotype in IGFBP3 deficient cells [29]. We sought to explore global differences of *IGFBP3* long range interaction profiles between normal breast cells and breast cancer cell lines. We hypothesized that cancer related changes in *IGFBP3* regulation and epigenetic modification might coincide with altered spatial positioning and long range DNA interactions contributing to breast cancer pathogenesis. We therefore used the *IGFBP3* enhancer as bait in circular chromosome conformation capture with high throughput sequencing (4C seq) in normal human mammary epithelial cells (HMEC) and two breast cancer cell lines, MCF7 and MDA MB 231. MCF7 and MDA MB 231 represent distinct breast cancer subtypes. MCF7 is a human breast adenocarcinoma cell line positive for estrogen receptor alpha, and MDA MB 231 is a human breast carcinoma cell line negative for estrogen and progesterone receptors as well as *HER2*. The *IGFBP3* promoter displays hypermethylation, and there is reduced *IGFBP3* expression in MCF7, while in MDA MB 231, the promoter is relatively hypomethylated, and *IGFBP3* is over expressed compared to HMEC.

In this study, we examined *IGFBP3* long range interactions and show that the three dimensional structure of the genome changes dramatically in breast cancer. Our data suggest a possible role for long range chromatin interactions in the pathogenesis of breast cancer as well as in the formation of translocations often seen in malignant cells.

Results

Expression of IGFBP3 is Downregulated in MCF7, but Upregulated in MDA-MB-231 Relative to HMEC

To better understand the role of *IGFBP3* in breast cancer, we analyzed its expression in primary breast cells, the estrogen receptor alpha (ERα) positive breast cancer cell line MCF7, and the triple negative breast cancer cell line MDA MB 231. *IGFBP3* expression was increased nearly 3 fold in MDA MB 231, and reduced 3.8 fold in MCF7, relative to HMEC (Figure 1A). To evaluate whether DNA methylation correlated with the changes in expression, we examined the methylation status of the *IGFBP3* promoter by bisulfite pyrosequencing. The *IGFBP3* promoter was hypermethylated (91% CpG methylation) in MCF7 compared with 11% and 10% CpG methylation in HMEC and MDA MB 231, respectively (Figure 1B).

EGFR Interacts Significantly with IGFBP3

To identify whether changes in *IGFBP3* expression and methylation were accompanied by global alteration of its long range chromatin interactions, we performed multiplex 4C seq in HMEC, MCF7 and MDA MB 231. We chose as our bait a region upstream of *IGFBP3* classified as a strong enhancer in HMEC by chromatin profiling of several distinctive features, including enrichment of the enhancer marks H3K4me1 and H3K4me2 and the active regulatory H3K9ac and H3K27ac marks (Figure S1) [30]. We obtained a combined total of approximately 12 million mapped reads from the three cell lines with the majority mapping in *cis* (Table S1). The 4C seq reads were binned into windows based on the number of mappable HindIII restriction sites, ranging from 25 to 400. Regions with a false discovery rate (FDR) below 0.01 (see Methods) were considered to be significantly interacting. The significant long range *cis* interactions for window size 100 in HMEC, MCF7 and MDA MB 231 are diagrammed in Figure 2A. For every window size analyzed, MCF7 contained the largest number of significant long range intrachromosomal interactions, followed by MDA MB 231 and HMEC. Using a window size of 100, there were a total of 16 significant *cis* long range interactions in HMEC, 51 in MCF7 and 29 in MDA MB 231. Of these interactions, 8 were common to all 3 cell lines, indicating a 50% conservation of all high confidence long range interactions from HMEC (Figure 2B). Numerous novel long range interactions were observed in each cancer cell line, and some long range interactions found in normal cells were lost in each cancer cell line.

Among the significant intrachromosomal interactions common to all samples, and across all window sizes, was an interaction with epidermal growth factor receptor (*EGFR*), another breast cancer related gene. *EGFR* is located approximately 9 Mb from *IGFBP3* on chromosome 7. To examine this long range interaction in more detail, we labeled gene pairs *EGFR* and *IGFBP3* by 3D FISH in HMEC and breast cancer cell lines MCF7 and MDA MB 231 (Figure 3A). To quantitate differences in interaction frequencies at the cellular level, we measured the center to center distances between the closest pairs of labeled foci. In 88% of HMEC nuclei counted, *EGFR* and *IGFBP3* were within 1 micron of each other, indicating frequent interactions (Figure 3B). This interaction frequency was only 56% in MCF7 nuclei, but was 96% in MDA MB 231 nuclei. To assess whether differences in spatial positioning were accompanied by changes in expression, we measured RNA levels of *EGFR* in HMEC, MCF7 and MDA MB 231 by qRT PCR (Figure 3C). Relative to HMEC, *EGFR* expression was unchanged in MDA MB 231, yet it was reduced 35 fold to nearly undetectable levels in MCF7 cells. In contrast to *IGFBP3*, the expression change in *EGFR* was not accompanied by a change in CpG methylation in the *EGFR* promoter among the three cell lines (data not shown). This suggests the difference in *EGFR* expression could be driven in part by chromatin architecture rather than methylation. In MCF7, the reduction in long range interaction frequency with *EGFR* provides the opportunity for *IGFBP3* to form additional contacts. This may partially explain the gain of 35 unique intrachromosomal interactions in MCF7 cells compared to HMEC.

Interchromosomal Rearrangements Involving IGFBP3 Interacting Regions Facilitate an Increase in Long-Range Interactions in MCF7

We constructed circos plots to highlight the significant interchromosomal interactions involving the *IGFBP3* enhancer in HMEC, MCF7 and MDA MB 231 that fell within a window size of 200 (Figure 4A, Figure S2). There were a total of 87 significant interactions in HMEC, 194 in MCF7 and 115 in MDA MB 231. Of these interactions only 11 were common to all samples (Figure 4B, Table 1). Because a large proportion of the significant

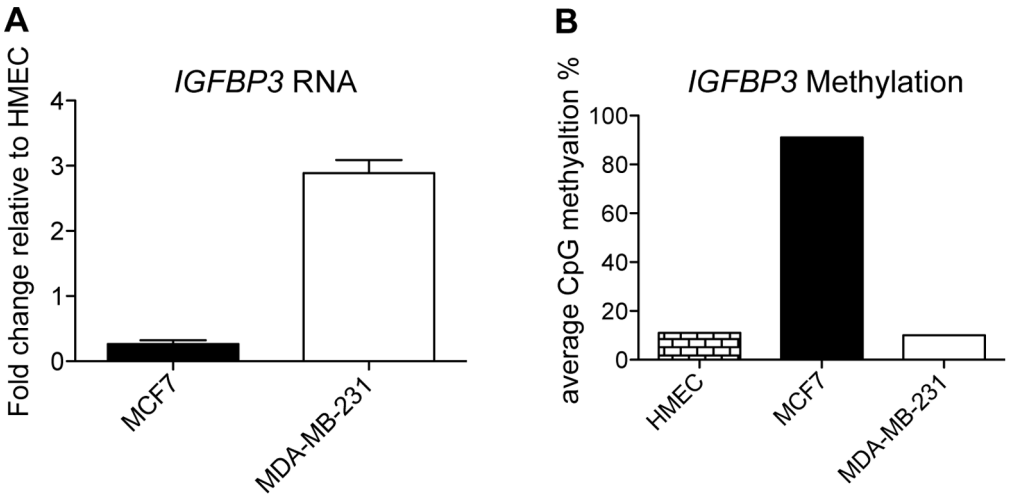


Figure 1. Expression and methylation status of *IGFBP3*. A, qRT PCR: RNA levels of *IGFBP3* were measured in MCF7, MDA MB 231 and HMEC cells. Expression in cancer lines was plotted as fold change relative to HMEC. Data represent the SEM of three independent biological replicates. B, Percent methylation of CpG nucleotides in the *IGFBP3* promoter in HMEC, MCF7 and MDA MB 231. Bars represent the average percent methylation of 4 positions in the *IGFBP3* promoter.
doi:10.1371/journal.pone.0073974.g001

4C windows fell within chromosome regions prone to rearrange ments, fusions and amplifications, we compared the locations of 157 breakpoints mapped in MCF7 cells [31] to the list of regions that participated in significant interchromosomal interactions. We have limited our analysis to the relationship between interactions in normal HMEC and known breakpoints in MCF7 since it was the only cell line with comprehensive breakpoint data available. This allows for the correlation of interactions pre and post breakage. The MCF7 breakpoints could be categorized as 2 distinct types. The first category contains the majority of breakpoints, which are dispersed throughout the genome in regions of low copy repeats. The second category includes MCF7 breakpoints falling within four highly amplified regions located on chromosomes 1, 3, 17 and 20. We found that breakpoint regions that also participated in interchromosomal interactions were almost exclusively in the latter category. We then considered a subset of 74 MCF7 breakpoints, described as interchromosomal rearrangements, and determined how many were associated with long range chromatin interactions in HMEC and MCF7 cell lines (Table 2). A total of 29 breakpoint ends mapped within significant windows in HMEC, as compared to 61 in the MCF7 line. All but one of the breakpoints within HMEC 4C windows was also present within MCF7 4C windows. Importantly, when we compared the number of breakpoints for which both ends of the breakpoint mapped to a 4C hit, the percentage was nearly twice as many in the breast cancer cell line MCF7 as in HMEC. This suggests that when an *IGFBP3* interacting region undergoes a translocation involving a different chromosome, the *IGFBP3* interaction is not lost, but instead the translocation brings into proximity an additional interaction detectable by 4C.

Breast Carcinoma Amplified Sequence (BCAS1-4) Genes Interact Significantly with IGFBP3 and Each Other in Normal Breast Cells

Some of the most significant 4C seq interchromosomal inter actions in HMEC included regions containing the genes *BCAS 1 4* located on chromosomes 1, 17 and 20. All 4 of these genes were found among the 10 most significantly enriched regions in HMEC, and the region containing *BCAS1* and *ZNF217* was the overall top scoring window. These interactions were also enriched in MCF7, where they are frequently rearranged and amplified (Table 3). We used 3D FISH to investigate whether the *IGFBP3* interacting *BCAS* genes were also in close spatial proximity with one another prior to any oncogenic translocations (Figure 5). We performed dual and triple labeled 3D FISH with probes for *IGFBP3*, *BCAS1*, *BCAS3* and *BCAS4* in primary HMEC cells (Figure 5A). Center to center distances were measured for the closest pairs of foci for each probe (Figure 5B). All probes targeting the *BCAS* genes were in close proximity, residing less than or equal to 1 micron to *IGFBP3* in at least 5% of nuclei. The *BCAS3 BCAS4* and *BCAS3 BCAS1* regions, which undergo translocations with one another in MCF7 [31], were also within 1 micron in at least 4% of normal HMEC nuclei. These percentages are in line with reports of positive *trans* interacting loci identified using other molecular assays [32,33]. This suggests spatial proximity of the *BCAS* genes in normal breast cells contributes to their frequent oncogenic translocations.

Methylated Promoters in Breast Cancer Disproportionally Fall within 4C Windows

Using genome wide CpG methylation data from Sproul *et al.* [34], we analyzed the distribution of methylated promoters in our 4C data sets. CpG sites with a value equal or greater than 0.8 were considered methylated. Consistent with an increase in global CpG methylation in breast cancer, the total number of methylated sites was greater in MCF7 (3847 sites) and MDA MB 231 (3282 sites), compared with HMEC (374 sites). There is a significant increase in the proportion of methylated promoters that participated in long range interactions with *IGFBP3* in both breast cancer cell lines relative to HMEC. This increase was more pronounced in MCF7 cells where *IGFBP3* itself is hypermethylated (Table S2). After correcting for the total number of methylated sites, there was a 3.77 fold (Fisher's exact test, one sided p value 4.742×10^{-9}) and 2.85 fold (Fisher's exact test, one sided p value 1.122×10^{-5}) increase in methylated promoters located within our 4C windows in MCF7 and MDA MB 231, respectively.

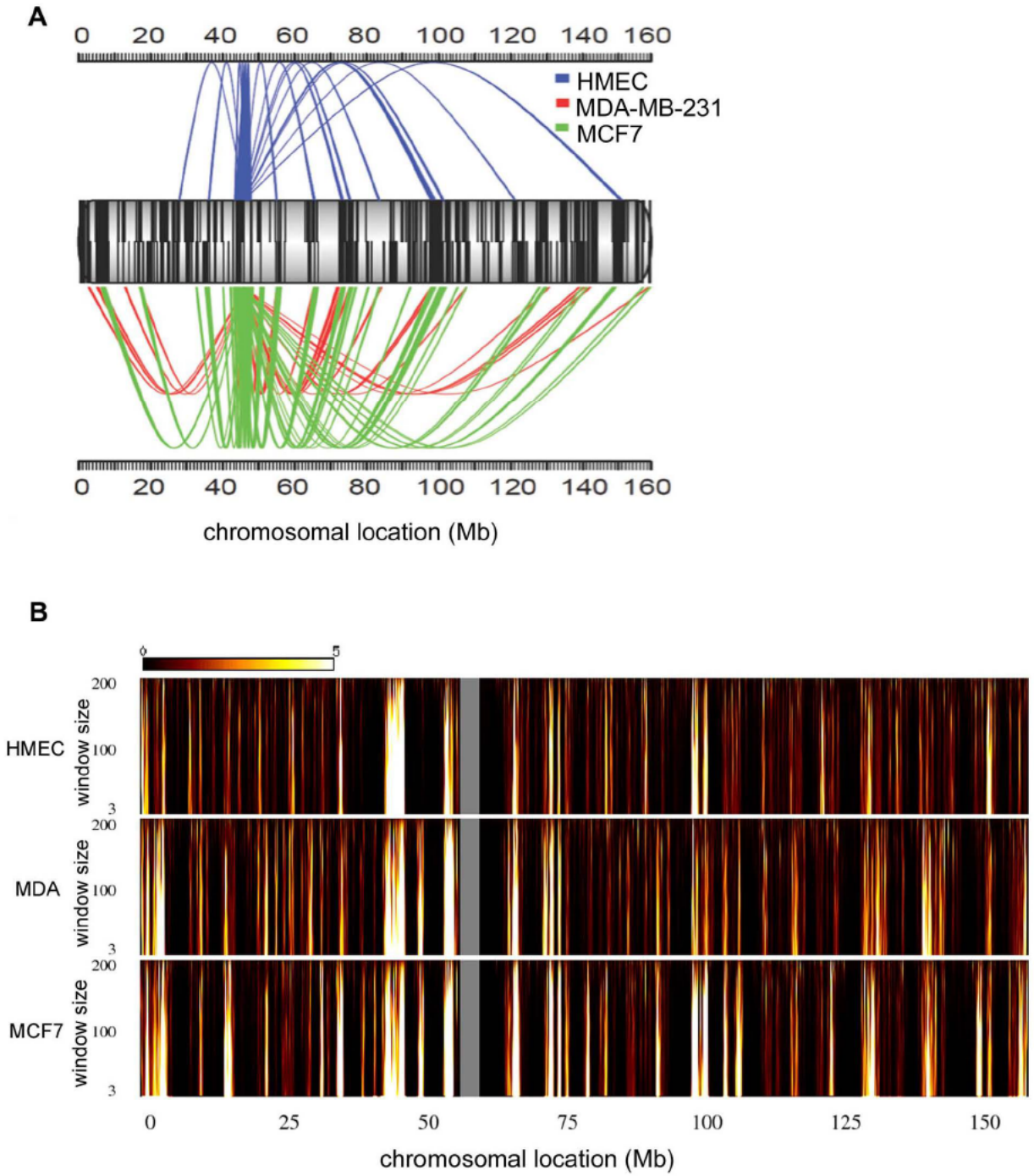


Figure 2. Intrachromosomal interaction profile of *IGFBP3*. A, Spider plot showing the significant long range interactions of the *IGFBP3* enhancer across chromosome 7 for a window size of 100 consecutive restriction fragments in HMEC (blue), MDA MB 231 (red), and MCF7 (green). Mb position is plotted. Tick marks on chromosome 7 represent gene locations with positive strand genes on top and negative strand genes on bottom. B, Domainograms illustrating the significance of intrachromosomal interactions for window sizes ranging from 3 to 200 consecutive fragments for each cell line. The color represents −log(p value) of the calculated significance score ranging from black (not significant) to white (most significant). The gray region corresponds to the centromere of chromosome 7, which lacks HindIII cut sites.
doi:10.1371/journal.pone.0073974.g002

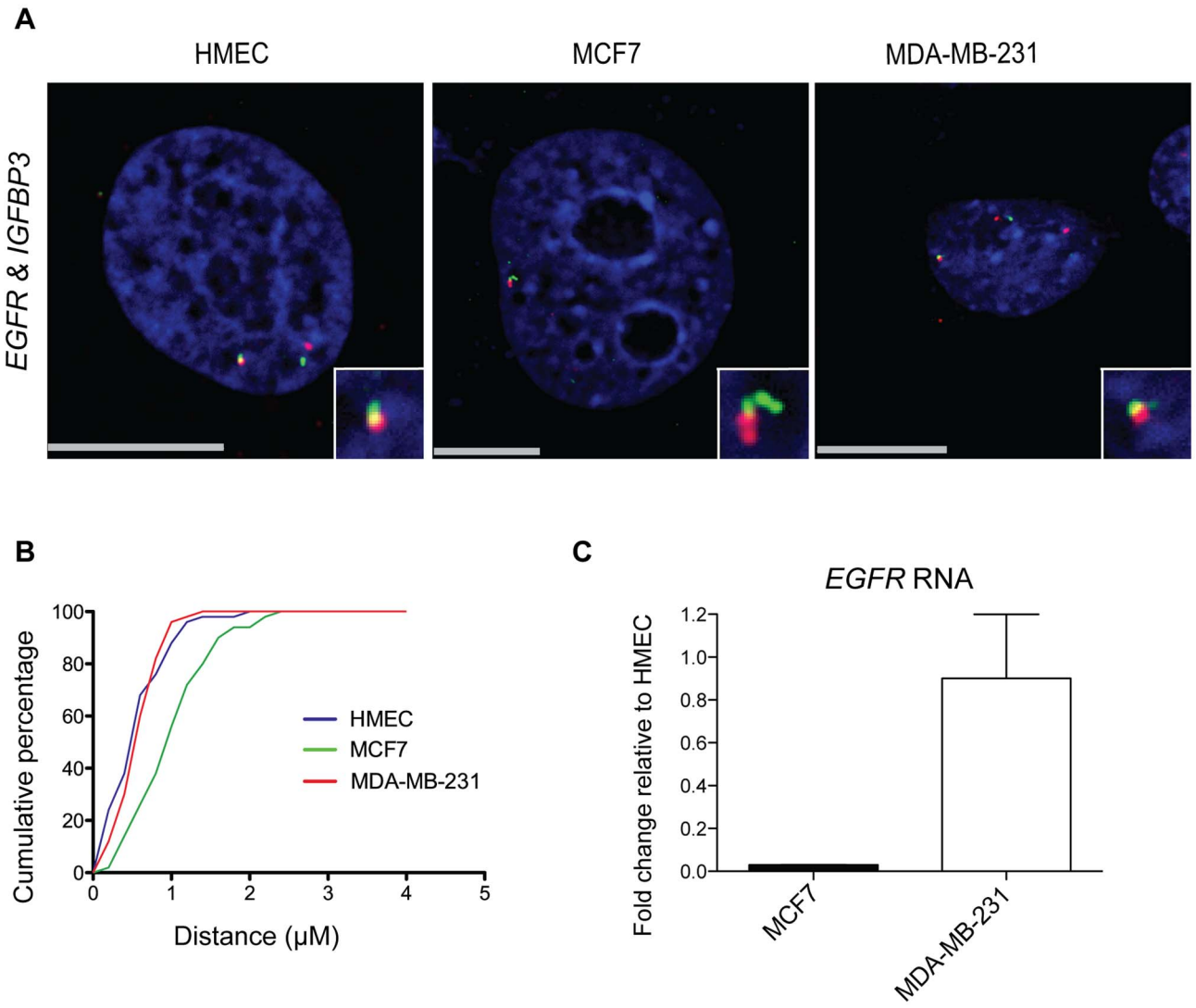


Figure 3. Interaction frequency of *IGFBP3* with the breast cancer related gene *EGFR* by 3D FISH. A, 3D FISH labeling of breast cancer related loci in HMEC, MCF7, MDA MB 231. BAC probe combinations: *IGFBP3* (green) and *EGFR* (red) n = 50, DAPI DNA stain (blue), boxes in lower right corner contain a magnified view of each interaction. Scale bar = 10 μm. B, Cumulative percentage of distances between *IGFBP3* and *EGFR* loci. Distances were measured between the closest two foci in each nucleus. C, qRT PCR: RNA levels of *EGFR* measured in MCF7, MDA MB 231 and HMEC cells. Expression in cancer lines plotted as fold change relative to HMEC. Data represent the SEM of three independent biological replicates.
doi:10.1371/journal.pone.0073974.g003

Discussion

Chromatin structure plays a key role in establishing and maintaining tissue specific gene expression profiles throughout development. Epigenetic modification of chromatin can influence DNA packaging and accessibility to *trans* acting regulatory factors. Active regulatory regions are maintained in open chromatin, characterized by nucleosome depletion and DNase I hypersensitivity [35]. A vast number of transcription factor binding sites are situated far from any transcription start site, and interactions occurring among distant regulatory elements can regulate gene expression [30]. Long range interactions between active regulatory elements may therefore provide a means to fine tune gene activity.

The importance of long range interactions may be especially relevant in cancer where genomic instability and extensive epigenetic modification of chromatin is common. Rickman *et al.*, for example, found that overexpression of an oncogenic transcription factor in normal cells leads to large scale changes in chromatin organization [36]. We have seen that there is a dramatic change in long range interactions in cancer cells compared with cells derived from normal tissues. We have previously shown that loss of IGF2 imprinting in cancer is accompanied by loss of normal long range intrachromosomal interactions involving the IGF2/H19 locus [11]. In this study we have expanded our view of long range interactions in cancer by exploring the genome wide interaction profile of *IGFBP3*.

IGFBP3 plays a major role in IGF signaling through binding the majority of circulating IGF I and IGF II, and it may also function independently in a growth stimulating or inhibitory fashion depending on the system studied. We observed that *IGFBP3* interacts with epidermal growth factor receptor (*EGFR*) in all 3 cell lines. EGFR is a receptor tyrosine kinase whose dysregulation can promote tumorigenesis, and nuclear EGFR has been shown to function as a transcription factor to activate genes required for cell

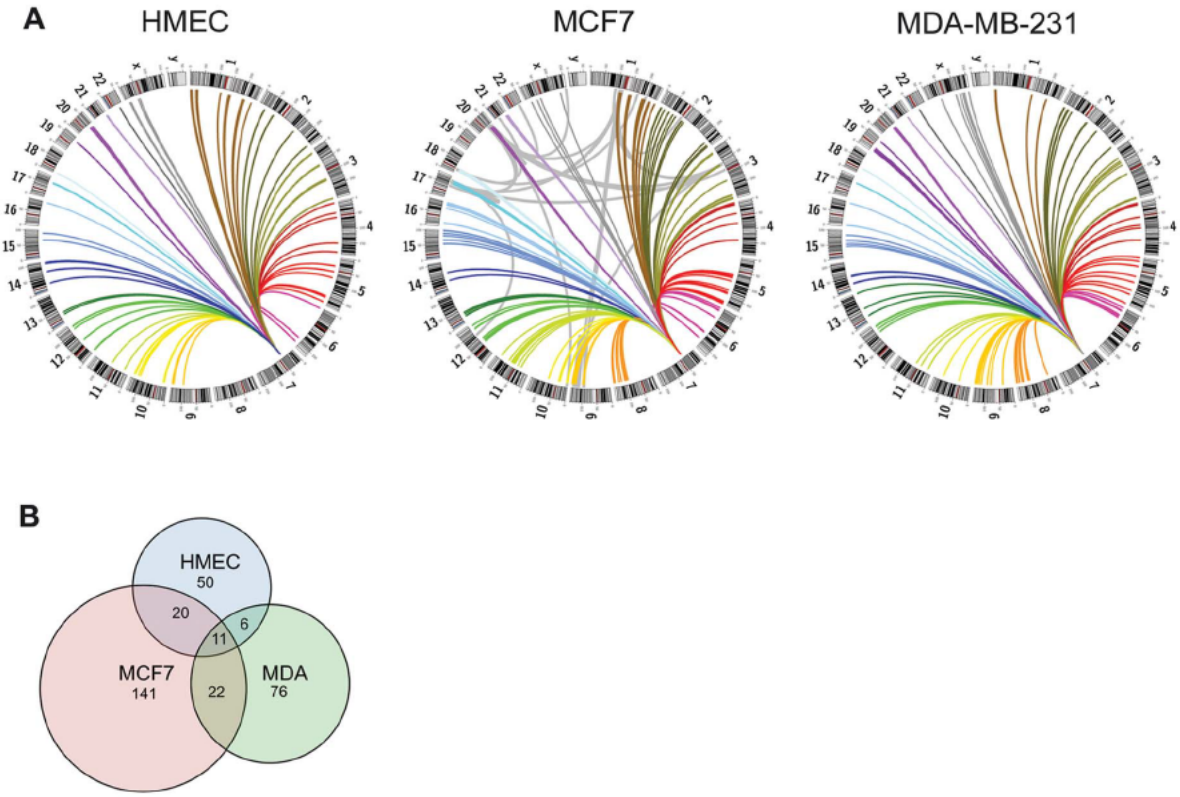


Figure 4. Interchromosomal interaction profile of *IGFBP3*. A, Circos plots showing the distribution of significant interchromosomal interactions involving *IGFBP3* in HMEC, MCF7 and MDA MB 231. Grey lines in MCF7 plot represent interchromosomal translocations, adapted from Hampton *et al.* [31], falling within windows of significant 4C interactions. B, Venn diagram showing the number of unique and overlapping significant interchromosomal interactions for a window size of 200 consecutive restriction fragments.
doi:10.1371/journal.pone.0073974.g004

Table 1. Common *trans* interactions among all samples.

window	genes
chr20:51905176 52752692	*BCAS1, ZNF217,TSHZ2, SUMO1P1, MIR4756*
chr17:58766326 59643391	*BCAS3, TBX2, C17orf82, TBX4*
chr20:48615985 49541607	*BCAS4, LINC00651, UBE2V1, TMEM189,CEBPB, LOC284751, PTPN1, MIR645, FAM65C, PARD6B, ADNP*
chr20:46815739 47725424	*LINC00494, PREX1, ARFGEF2, CSE1L*
chr1:144919188 145816358	*PDE4DIP, SEC22B, NOTCH2NL, NBPF10, HFE2, TXNIP, POLR3GL, ANKRD34A, LIX1L, RBM8A, GNRHR2, PEX11B, ITGA10, ANKRD35, PIAS3, NUDT17, POLR3C, RNF115, CD160, PDZK1, GPR89A*
chr20:45119530 45995741	*ZNF334, OCSTAMP, SLC13A3, TP53RK, SLC2A10, EYA2, MIR3616, ZMYND8, LOC100131496*
chr3:196975652 197787067	*DLG1, MIR4797, DLG1 AS1, BDH1, LOC220729, KIAA0226, MIR922, FYTTD1, LRCH3, IQCG, RPL35A, LMLN, ANKRD18DP*
chr1:200591661 201448561	*DDX59, CAMSAP2, GPR25, C1orf106, KIF21B, CACNA1S, ASCL5, TMEM9, IGFN1, PKP1, TNNT2, LAD1, TNNI1, PHLDA3*
chr2:24898227 25798560	*NCOA1, PTRHD1, CENPO, ADCY3, DNAJC27, DNAJC27 AS1, EFR3B, POMC, DNMT3A, MIR1301, DTNB*
chr4:1134384 2497968	*SPON2, LOC100130872, CTBP1, CTBP1 AS1, MAEA, UVSSA, CRIPAK, FAM53A, SLBP, TMEM129, TACC3, FGFR3, LETM1, WHSC1, SCARNA22, WHSC2, MIR943, C4orf48, NAT8L, POLN, HAUS3, MXD4, MIR4800, ZFYVE28, LOC402160, RNF4*
chr9:132166602 133421900	*LOC100506190, C9orf50, NTMT1, ASB6, PRRX2, PTGES, TOR1B, TOR1A, C9orf78, USP20, FNBP1, GPR107, NCS1, ASS1*

doi:10.1371/journal.pone.0073974.t001

Table 2. Distribution of MCF7 translocation breakpoints.

	HMEC	MCF7
4C windows containing at least one breakpoint end	11.5%	13.4%
Total number of breakpoint ends mapping to 4C windows	29	61
Number of breakpoint ends common to HMEC and MCF7	28	28
Breakpoints with both ends in 4C windows	34.5%	68.9%

doi:10.1371/journal.pone.0073974.t002

Table 3. *BCAS* gene loci are located in significantly interacting 4C windows.

Cell Line	*BCAS1* chr20	*BCAS2* chr1	*BCAS3* chr17	*BCAS4* chr20
HMEC	1	10	3	5
MCF7	4	1	7	14
MDA MB 231	1	NA	8	5

Numbers represent rank by p value with 1 being the most significant interaction.

doi:10.1371/journal.pone.0073974.t003

proliferation [37,38]. Recently, the cancer genome atlas network identified four major subtypes of breast cancer based on extensive genomic analyses. They found high level EGFR and phosphor ylated EGFR to be associated with a subset of breast cancers with HER2 enrichment, suggesting possible targets for combined therapy [39]. Crosstalk exists between insulin like growth factor 1 receptor (IGF1R), and other signaling receptors including EGFR. Inhibiting either IGF1R or EGFR results in activation of the reciprocal receptor, suggesting that combined inhibition of both

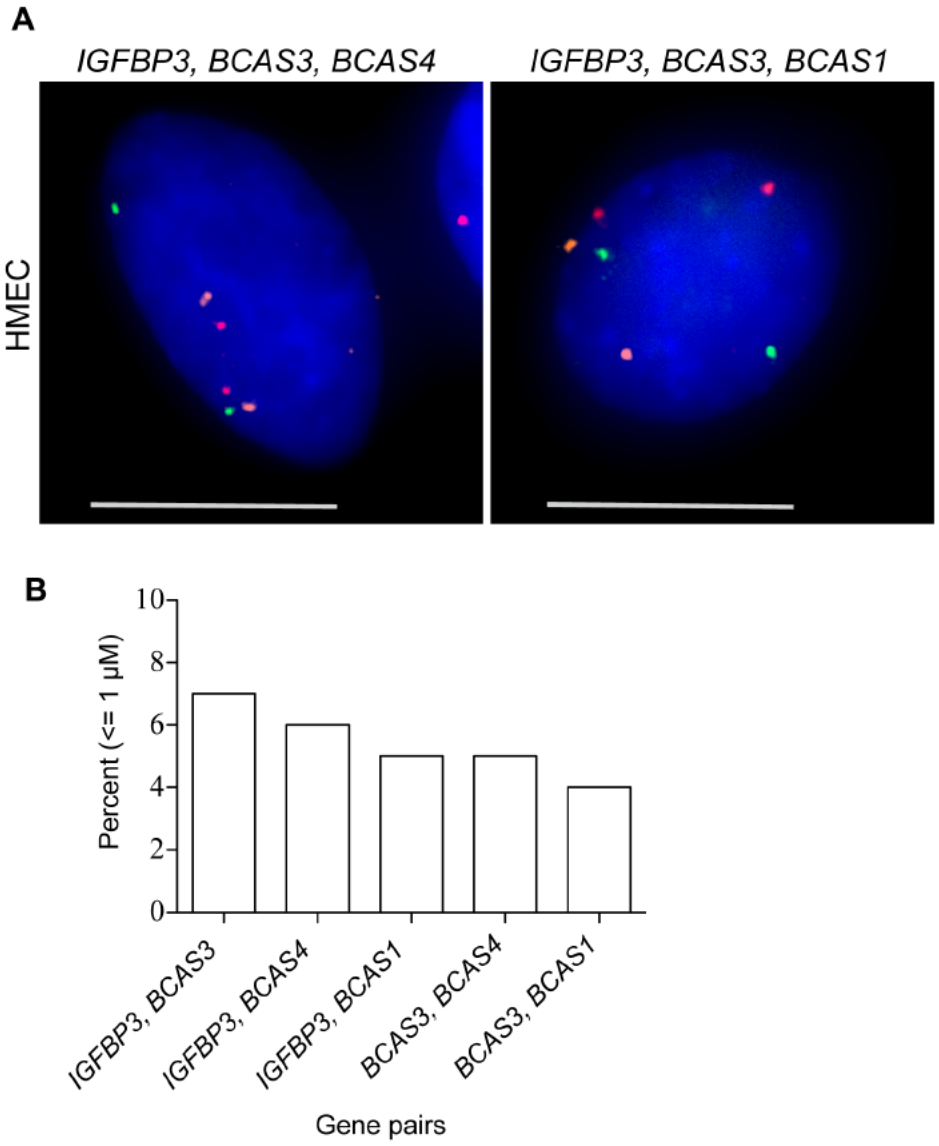


Figure 5. ***IGFBP3*** **interacts with** ***BCAS*** **genes.** A, Representative triple labeled 3D FISH, z axis projection images of *IGFBP3, BCAS3, BCAS4* (left) and *IGFBP3, BCAS3, BCAS1* (right). Scale bar = 10 µm. B, Percentage of nuclei with the listed pair of gene loci within 1 micron of each other. Distances were measured between the closest two foci in each nucleus.

doi:10.1371/journal.pone.0073974.g005

pathways may yield enhanced tumor therapy [40]. There is also interplay between IGFBP3 and EGFR in cancer cells. The initial reaction of ER positive T47D breast cancer cells to IGFBP3 is inhibitory, yet prolonged expression of *IGFBP3* cDNA stimulates growth. Chronic exposure of cells to IGFBP3 over many passages *in vitro* also led to an increase in EGFR protein levels, and enhanced the response to EGF as demonstrated by an increase in both phosphorylated EGFR and DNA synthesis. Furthermore, xenograft tumors in mice that expressed *IGFBP3* showed enhanced growth and increased levels of EGFR [41]. Conversely, overexpression of *EGFR* in primary keratinocytes resulted in 4.4 fold induction of *IGFBP3* [42]. Our 4C seq and 3D FISH data indicate *IGFBP3* and *EGFR*, separated by 9 Mb, are often in close spatial proximity (Figure 3). Spatial proximity of loci residing on the same chromosome is influenced to some extent by their linear separation in base pairs; it is therefore difficult to make comparisons between studies of loci with differing amounts of linear separation. Nonetheless, the number of nuclei scored by 3D FISH containing *IGFBP3* and *EGFR* in close proximity can be considered high in our cell lines, especially HMEC and MDA MB 231, where nearly all cells had at least one allele demonstrating proximity within 1 micron. Importantly, this high interaction frequency was not due solely to linear distance between the genes, as a large number of interactions occurred monoallelically. This can be observed in HMEC and MDA MB 231 3D FISH images in Figure 3A.

Mounting evidence suggests eukaryotic transcription occurs in localized factories [43,44]. Transcription factories may exist to provide coordinated expression of coregulated genes. By uniting distant regions of DNA they may also serve as sites to share specific or limiting regulatory factors, and may be required for high levels of transcription. We observed that in cell lines with increased *IGFBP3* mRNA there is also an increase in the interaction frequency of *IGFBP3* with *EGFR*. The relationship between interaction frequency and expression is nonlinear, and we expect other factors are modulating expression such as the observed hypermethylation of the *IGFBP3* promoter. Additional factors may include crosstalk between IGFBP3 and EGFR signaling pathways and tumor heterogeneity. Our data suggest that *IGFBP3* and *EGFR* may share a common transcriptional hub or factory, and disruption of these interactions could play a role in tumor progression. Reduction of the *IGFBP3 EGFR* interaction may not only affect these genes, but could result in new long range interactions.

Cytogenetic and molecular evidence suggests spatial proximity influences recurrent chromosomal translocations [45,46,47,48]. In response to genotoxic stress, oncogenic translocations could potentially form when DNA breaks occur within an interacting "hub". This was demonstrated in prostate cancer cells where irradiation led to translocations among genes with hormone induced proximity [49,50].

From our 4C data, we found that the breast carcinoma amplified sequence family of genes (*BCAS1*, *BCAS2*, *BCAS3* and *BCAS4*) interacts with *IGFBP3*. *BCAS1* has been found amplified in primary breast tumors [51] and associated with a poor prognosis [52]. *BCAS2* can function as a transcriptional coactivator of estrogen receptor [53] as well as a negative regulator of *P53* [54]. *BCAS3* is overexpressed and associated with impaired response to tamoxifen in ER positive premenopausal breast cancers [55]. Fine mapping of breakpoints in MCF7 revealed *BCAS3* to be located in a rearrangement hotspot, where 7 breakpoints were observed within *BCAS3* and 19 in the surrounding region of the gene [31]. One of the translocation partners of *BCAS3* is *BCAS4*, and fusion transcripts have been detected in MCF7 and HCT116 colon cancer cells [31,56]. Additionally, *BCAS4* was found overexpressed in nine out of 13 different breast cancer cell lines [57]. The *BCAS* genes are frequently amplified and some have been found to translocate with each other in breast tumors, such as *BCAS4 BCAS3* and *BCAS1 BCAS3*. Interestingly, using 3D FISH in normal breast cells, we found *BCAS4 BCAS3* and *BCAS1 BCAS3* to interact with one another as well as *IGFBP3*, supporting the role of spatial proximity in oncogenic translocations. All pairwise interactions, defined as being equal to or less than 1 micron, occurred in 4% or greater of HMEC nuclei. This is similar to the association levels measured for loci participating in interchromosomal interactions identified using the tethered chromosome conformation capture assay [33]. It is also similar to colocalization levels of genes that occupy specialized transcription factories in mouse erythroid nuclei [32]. We chose to verify 4C interactions with 3D FISH, as the interacting regions can be large, consisting of windows of 100 or 200 restriction sites. 3C would provide better resolution, but doing so on such large regions would be quite challenging considering the number of primers that would be needed since detecting an interaction between two specific elements alone with 3C is not technically sound.

Although all interactions were present within the population of cells, there was not a simultaneous association of all three loci. This suggests the long range interactions of the *BCAS* genes with *IGFBP3* and with one another are dynamic in nature, and illustrates the heterogeneity of chromatin architecture within a cell population. Chromatin displays rapid constrained motion over distances of $\sim$ 1 micron and longer directional movement of chromatin domains has been associated with gene expression [58]. We note that 3D FISH experiments were performed in cycling cells. Since this data is limited to interphase cells we don't expect it to have a major effect on our results. As the field progresses we will likely see 4D studies incorporating cell cycle stages; there have already been correlations drawn between Hi C data and replication timing [59].

It remains to be seen what role *trans* acting factors play in mediating these long range interactions. In the case of prostate cancer, the androgen receptor was shown to rapidly induce long range interactions both in *cis* and in *trans* following ligand binding [49,50]. Estrogen was also shown to induce rapid interchromosomal interactions among estrogen receptor α (ERα) regulated genes [5,60]. In addition to nuclear receptor mediated long range interactions, increased expression of the architectural protein SATB1, which participates in chromatin loop formation, alters the expression of over 1000 genes and is associated with aggressive breast cancer [61]. Whatever the mechanism governing long range interactions, it is likely to involve a combination of chromatin remodeling complexes and possibly nuclear motor proteins. Along these lines, chromatin interacting with *IGFBP3* in the breast cancer cell lines was significantly enriched for methylated promoters relative to HMEC, with MCF7 showing the greatest fold increase. The *IGFBP3* promoter is hypermethylated in MCF7, and this may indicate a preference for chromatin domains with similar modifications to associate.

It is notable that a large proportion of the MCF7 translocation breakpoints fall within 4C windows. To rule out artifacts due to an interaction with a breakpoint near our bait, we checked for breakpoints proximal to *IGFBP3*, but found none within ± 5 Mb. It is important to note that MCF7 breakpoints mapped in HMEC reflect areas of potential translocations. It is interesting that all HMEC 4C windows containing translocation breakpoints were also present in MCF7, where breakage had occurred. There was also an increase in breakpoints with both ends mapping to 4C windows in MCF7 as compared to HMEC. In these instances, the

IGFBP3 long range interactions present in normal breast cells were maintained in the tumor cells, and additional interactions with the reciprocal breakpoints were formed due to rearrangements in the cancer cell line. This indicates a preference for gaining *trans* interactions even after large scale genomic aberrations occur.

Our study demonstrates that long range interactions of cancer related loci, including *EGFR* and *IGFBP3*, are altered in breast cancer cells, and these alterations are frequently associated with epigenetic changes. Long range interactions influence chromosomal translocations, and add an additional layer of complexity to transcriptional and epigenetic regulation to coordinate gene expression. Therefore, a better understanding of aberrant chromatin interactions is needed to fully understand cancer pathology.

Methods

Cell Culture

Primary human mammary epithelial cells, HMEC (Life Technologies, Grand Island, NY) were cultured in HuMEC Ready Medium (Gibco, Grand Island, NY) with 1% penicillin streptomycin (Gibco). Human breast cancer cell lines MCF7 and MDA MB 231 (ATCC, Manassas, VA) were grown in Dulbecco's Modified Eagle Medium (DMEM) with high glucose, sodium pyruvate, GlutaMAX media supplemented with 10% fetal bovine serum, 1% penicillin streptomycin (Gibco) at 37°C in 5% CO_2.

Circular Chromosome Conformation Capture (4C) Sequencing Assay

4C was performed as in Gheldof *et al.* with minor modifications [62]. HMEC, MCF7 and MDA MB 231 cells (2×10^7) were fixed in 2% formaldehyde in fresh medium for 10 min at room temperature, followed by quenching with 0.125 M glycine. Fixed cells were scraped from culture plates, spun, (750×g for 10 min), and the frozen pellets were stored at −80°C until lysis. Cells were resuspended in ice cold lysis buffer (0.2% IGEPAL CA 630, 10 mM NaCl, 10 mM Tris HCl) with SigmaFast complete protease inhibitor tablet (Sigma Aldrich, St. Louis, MO) and lysed for 30 min on ice. After recovery of nuclei by centrifugation (2000×g for 5 minutes), nuclei were washed twice in cold 1.2× NEB buffer 2 and resuspended in the same buffer. Nuclei were incubated in the presence of 0.3% SDS for 1 h at 37°C with shaking at 950 rpm, followed by the addition of Triton X 100 to 1.8% for 1 h at 37°C with shaking at 950 rpm. Nuclei were digested with 1500 U of HindIII (New England Biolabs Ipswich, MA) overnight at 37°C with shaking at 950 rpm. 200 µl of digested nuclei were removed for assessing digestion efficiency by qPCR. The restriction enzyme was inactivated by the addition of 1.6% SDS and was incubated at 65°C for 20 min. The digested nuclei were diluted in 7 ml of 1.1× T4 DNA ligase buffer in the presence of 1% Triton X 100 and incubated for 1 h at 37°C. Ligation was performed by adding 800 U of T4 DNA Ligase (2,000,000 U/ml; New England Biolabs) to the diluted mixture of digested nuclei and incubating in a 16°C H_2O bath for 4 hours followed by a 30 min incubation at room temperature. To reverse cross links, proteinase K was added to a final concentration of 100 µg/ml and incubated overnight at 65°C. Samples were incubated with 0.5 µg/ml of RNase A at 37°C for 1 h and purified by phenol chloroform extraction followed by ethanol precipitation. DNA concentration was measured using a Qubit® 2.0 Fluorometer (Life Technologies).

3C templates were digested with 200 U MspI (New England Biolabs) overnight at 37°C with shaking at 500 rpm, followed by heat inactivation at 65°C for 20 min. Digestion products were purified by phenol chloroform extraction and ethanol precipitation. Ligations were performed in 14 ml of 1× T4 DNA ligase buffer with 2000 U of T4 DNA ligase. Circular ligation products were purified by phenol chloroform extraction and ethanol precipitation followed by clean up with Ampure beads (Beckman Coulter, Brea, CA). A total of 16 inverse PCR reactions with 200 ng input per 4C template were performed for each library with primers that included Illumina adapter sequences and custom barcodes. All PCR reactions were performed with Expand Long Template PCR system (Roche, Indianapolis, IN). Excess primers were removed by gel extraction. HMEC, MCF7 and MDA MB 231 4C libraries were analyzed on a MultiNA microchip electrophoresis system (Shimadzu Columbia, MD) and mixed in equimolar amounts. Multiplex sequencing was performed on an Illumina genome analyzer IIx (Illumina, San Diego, CA). Illumina sequencing data have been submitted to the GEO database accession number: GSE49521.

Mapping and Filtering of 4C Reads

We first de multiplexed the 76 bp single end reads using barcodes for each cell line. We only retained the reads that contained one of the valid barcodes followed by the primer sequence and a HindIII cleavage site and truncated them to obtain the prey sequence. We mapped the truncated reads to the human genome (UCSC hg19) using the short read alignment mode of BWA (v0.5.9) with default parameter settings. We post processed the alignment results to extract the reads that satisfied the following three criteria: (i) mapped uniquely to one location in the reference genome, (ii) mapped with an alignment quality score of at least 30 (which corresponds to 1 in 1000 chance that mapping is incorrect), (iii) mapped with an edit distance of at most 3. We assigned the qualified reads to the nearest HindIII cleavage site using their mapping coordinates. We then identified the restriction fragments interacting (those flanking the cleavage sites with a read count of at least one) with the bait region. We discarded ±50 kb region around the bait from further analysis.

Statistical Analysis of 4C Data

We first identified all the HindIII sites in the genome (~840 k) and eliminated the ones with no MspI site within 2 kb downstream of the HindIII site, resulting in ~470 k restriction fragments for downstream analysis. In order to avoid PCR artifacts, we binarized the interactions counts as was done previously in other 4C analysis pipelines [63]. This processing resulted in 23,559, 19,876 and 16,387 restriction fragments that interact with the bait region for HMEC, MCF7 and MDA MB 231 cell lines, respectively. In order to account for the difference in the number of interacting fragments between cell types and the effect of genomic distance on the intrachromosomal interaction probability, we applied a statistical significance assignment procedure similar to the one described in Splinter *et al* [63]. We first separated interactions into four groups depending on the linear distance of interacting loci to the bait.

1. Bait region interactions: Intrachromosomal interactions below 50 kb distance to the bait and are excluded from our analysis.
2. Proximal intrachromosomal interactions: Intrachromosomal interactions between 50 kb to 2 Mb distance from the bait.
3. Long range intrachromosomal interactions: Intrachromosomal interactions above 2 Mb distance from the bait.
4. Interchromosomal interactions: Interactions that are on chromosomes other than the bait chromosome (chr 7).

We then combined multiple consecutive restriction fragments with window sizes that are appropriate for each of the groups

above. This step is necessary due to limited resolution of current 4C methods and enables us to assign statistical confidences for interactions at varying resolutions. We used window sizes of 10, 20 and 40 for group 2; 50, 100 and 200 for group 3; 100, 200 and 400 for group 4 interactions. For each group of interactions, we counted the number of interacting fragments within a window for each window size. We then generated a background distribution by randomly shuffling the interacting and non interacting fragments for each group and repeating this randomization 100 times. For intrachromosomal interactions, we take into account the linear distance of each region to the bait when generating the background. For interchromosomal interactions, we generated the background by aggregating all chromosomes (unlike Splinter et al [63] who generate one background per each chromosome) to preserve the information from possible chromosome territory associations that include chromosome 7. Similar to Splinter *et al*, [63] we calculated the z value threshold at which the false discovery rate (FDR) is 0.01 to determine the windows that significantly interact with the 4C bait (4C enriched windows/regions). To determine cell line specific 4C enriched regions, at a given window size, we simply take the list of regions that are deemed interacting at FDR 0.01 in one cell line and not in the other.

3D-fluorescence *in situ* Hybridization

Cells grown on 12 mm coverslips were fixed in 4% paraformaldehyde (PFA) for 10 min, made permeable with 0.5% Triton X 100 for 5 min, incubated in 20% glycerol/1× PBS for at least 40 min, freeze thawed in liquid nitrogen four times, and treated with 0.1 N HCl for 5 min. Cells were then treated with RNase A for 45 min at 37°C. Coverslips were then stored in 50% formamide/2× SSC at 4°C until denaturation at 75°C for 7 min in 70% formamide/2× SSC followed by immersion in ice cold 50% formamide/2× SSC.

BAC probes: RP11 89E8, RP11 1083I7, RP11 55E1, RP11 1115J10, RP11 705A3, RP11 805G4, RP11 185P21, RP11 1058F18, RP11 937E18, RP11 5P14 (Roswell Park Cancer Institute, Buffalo, NY) were labeled with dinitrophenol 11 dUTP (PerkinElmer, Waltham, MA), Alexa488 dUTP or Alexa594 dUTP (Life Technologies) by nick translation (Roche). Probes in 50% formamide/2× SSC/10% dextran sulfate were denatured for 8 10 min at 75°C. Probes were cooled on ice and hybridized for 36 48 h at 37°C, followed by three post hybridization washes with 50% formamide/2× SSC/0.05% Tween 20, 2× SSC/0.05% Tween 20, and 1× SSC for 30 min each at 37°C. Detection of BAC probes was performed by reaction with rabbit anti DNP (Life Technologies) diluted (1:1000) and secondary goat anti rabbit (1:200) conjugated to Alexa594 or Alexa647 (Life Technologies). Following labeling, indirect immunofluorescence was detected with Chroma filter sets using an Olympus BX41 upright microscope (100× UPLSAPO, oil, 1.4 NA) equipped with motorized *z* axis controller (Prior Scientific, Rockland, MA) and Slidebook 5.0 software (Intelligent Imaging Innovations, Denver, CO). Optical sections of 0.5 μm were collected, deconvolved using a NoNeighbor algorithm operating within Slidebook 5.0, and 3D distances were measured from the center of each FISH focus.

CpG Methylation by Bisulfite Pyrosequencing

Genomic DNA from HMEC, MCF7 and MDA MB 231 were treated with bisulfite using the EZ DNA Methylation kit (ZYMO Research, Irvine, CA). The locus of interest was amplified using a combination of forward and biotinylated reverse primers (see Table S3 for primer sequences). 40 ng bisulfite treated DNA was used for each 25 μl PCR reaction with 2G Robust polymerase (KAPA Biosystems, Woburn, MA) following KAPA's recommended cycling conditions. Pyrosequencing of the resulting amplicons was performed at the PAN facility, Stanford University using a Qiagen Pyromark instrument. Assays were designed using Pyromark Assay Design software (Qiagen, Valencia, CA). The methylation indices were calculated as the average percent methylation of successive CpG dinucleotides between the primers.

RNA Extraction and Quantitative RT-PCR

RNA was extracted from HMEC, MCF7 and MDA MB 231 cells using the RNeasy Mini Kit and QIAshredder mini column (Qiagen) according to the manufacturer's instructions. DNA was digested on a column using RNase free DNase set (Qiagen). 1 μg of RNA was reverse transcribed with Superscript III first strand synthesis supermix for qRT PCR (Life Technologies). qRT PCR was performed using KAPA SYBR Fast ABI PRISM qPCR mix (KAPA) on an ABI 7900HT Real Time PCR System (Applied Biosystems). Primers were purchased from RealTimePrimers.com. The most stable reference genes (*ACTB* and *GAPD*) were selected from a set of 10 using geNorm software [64]. Reaction efficiency for each primer set was calculated using Real time PCR Miner [65] and fold change of target genes relative to HMEC was calculated using the Pfaffl method [66].

Supporting Information

Figure S1 ***IGFBP3* 4C-Seq Bait.** The bait sequence, top (red bar) flanks a HindIII site upstream of *IGFBP3* in a region classified as a strong enhancer (orange bar). Image generated with UCSC genome browser, hg19.
(JPG)

Figure S2 Distribution of the significant 200 restriction site interchromosomal windows for HMEC, MCF7 and MDA-MB-231. Percent of total interactions per cell line are plotted for each chromosome.
(JPG)

Table S1 Sequence read distribution (not corrected for local interactions).
(JPG)

Table S2 Distribution of methylated promoter CpG nucleotides relative to HMEC.
(DOCX)

Table S3 Methylation assay primer sequences.
(DOCX)

Author Contributions

Conceived and designed the experiments: MJZ JDH PLL BNS. Performed the experiments: MJZ JDH PLL. Analyzed the data: MJZ JDH PLL FA WSN ARH. Wrote the paper: MJZ JDH PLL FA. Designed and conducted bioinformatics and statistical analysis: FA WSN.

References

1. Misteli T (2007) Beyond the sequence: cellular organization of genome function. Cell 128: 787 800.

2. Duan Z, Andronescu M, Schutz K, McIlwain S, Kim YJ, et al. (2010) A three-dimensional model of the yeast genome. Nature 465: 363 367.

3. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, et al. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326: 289 293.
4. Noordermeer D, Leleu M, Splinter E, Rougemont J, De Laat W, et al. (2011) The dynamic architecture of Hox gene clusters. Science 334: 222 225.
5. Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, et al. (2009) An oestrogen-receptor-alpha-bound human chromatin interactome. Nature 462: 58 64.
6. Cope NF, Fraser P, Eskiw CH (2010) The yin and yang of chromatin spatial organization. Genome biology 11: 204.
7. Ling JQ, Li T, Hu JF, Vu TH, Chen HL, et al. (2006) CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. Science 312: 269 272.
8. Spilianakis CG, Lalioti MD, Town T, Lee GR, Flavell RA (2005) Interchromosomal associations between alternatively expressed loci. Nature 435: 637 645.
9. Xu N, Tsai CL, Lee JT (2006) Transient homologous chromosome pairing marks the onset of X inactivation. Science 311: 1149 1152.
10. Bacher CP, Guggiari M, Brors B, Augui S, Clerc P, et al. (2006) Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. Nature cell biology 8: 293 299.
11. Vu TH, Nguyen AH, Hoffman AR (2010) Loss of IGF2 imprinting is associated with abrogation of long-range intrachromosomal interactions in human cancer cells. Human molecular genetics 19: 901 919.
12. Shah SP, Roth A, Goya R, Oloumi A, Ha G, et al. (2012) The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature 486: 395 399.
13. Esteller M (2007) Cancer epigenomics: DNA methylomes and histone-modification maps. Nature reviews Genetics 8: 286 298.
14. Stephens PJ, McBride DJ, Lin ML, Varela I, Pleasance ED, et al. (2009) Complex landscapes of somatic rearrangement in human breast cancer genomes. Nature 462: 1005 1010.
15. Meaburn KJ, Misteli T (2008) Locus-specific and activity-independent gene repositioning during early tumorigenesis. The Journal of cell biology 180: 39 50.
16. Meaburn KJ, Gudla PR, Khan S, Lockett SJ, Misteli T (2009) Disease-specific gene repositioning in breast cancer. The Journal of cell biology 187: 801 812.
17. Rocha RL, Hilsenbeck SG, Jackson JG, VanDenBerg CL, Weng C, et al. (1997) Insulin-like growth factor binding protein-3 and insulin receptor substrate-1 in breast cancer: correlation with clinical parameters and disease-free survival. Clinical cancer research : an official journal of the American Association for Cancer Research 3: 103 109.
18. Yu H, Levesque MA, Khosravi MJ, Papanastasiou-Diamandi A, Clark GM, et al. (1998) Insulin-like growth factor-binding protein-3 and breast cancer survival. International journal of cancer Journal international du cancer 79: 624 628.
19. Sugumar A, Liu YC, Xia Q, Koh YS, Matsuo K (2004) Insulin-like growth factor (IGF)-I and IGF-binding protein 3 and the risk of premenopausal breast cancer: a meta-analysis of literature. International journal of cancer Journal international du cancer 111: 293 297.
20. Baglietto L, English DR, Hopper JL, Morris HA, Tilley WD, et al. (2007) Circulating insulin-like growth factor-I and binding protein-3 and the risk of breast cancer. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 16: 763 768.
21. Key TJ, Appleby PN, Reeves GK, Roddam AW (2010) Insulin-like growth factor 1 (IGF1), IGF binding protein 3 (IGFBP3), and breast cancer risk: pooled individual data analysis of 17 prospective studies. The lancet oncology 11: 530 542.
22. Firth SM, Baxter RC (2002) Cellular actions of the insulin-like growth factor binding proteins. Endocrine Reviews 23: 824 854.
23. Oh Y, Muller HL, Pham H, Rosenfeld RG (1993) Demonstration of receptors for insulin-like growth factor binding protein-3 on Hs578T human breast cancer cells. The Journal of biological chemistry 268: 26045 26048.
24. Rajah R, Valentinis B, Cohen P (1997) Insulin-like growth factor (IGF)-binding protein-3 induces apoptosis and mediates the effects of transforming growth factor-beta1 on programmed cell death through a p53- and IGF-independent mechanism. The Journal of biological chemistry 272: 12181 12188.
25. Mohan S, Baylink DJ (2002) IGF-binding proteins are multifunctional and act via IGF-dependent and -independent mechanisms. The Journal of endocrinology 175: 19 31.
26. Rinaldi S, Peeters PH, Berrino F, Dossus L, Biessy C, et al. (2006) IGF-I, IGFBP-3 and breast cancer risk in women: The European Prospective Investigation into Cancer and Nutrition (EPIC). Endocrine-related cancer 13: 593 605.
27. Tomii K, Tsukuda K, Toyooka S, Dote H, Hanafusa T, et al. (2007) Aberrant promoter methylation of insulin-like growth factor binding protein-3 gene in human cancers. International journal of cancer Journal international du cancer 120: 566 573.
28. Grkovic S, O'Reilly VC, Han S, Hong M, Baxter RC, et al. (2012) IGFBP-3 binds GRP78, stimulates autophagy and promotes the survival of breast cancer cells exposed to adverse microenvironments. Oncogene 32(19): 2412 2420.
29. Mehta HH, Gao Q, Galet C, Paharkova V, Wan J, et al. (2011) IGFBP-3 is a metastasis suppression gene in prostate cancer. Cancer research 71: 5154 5163.
30. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, et al. (2011) Mapping and analysis of chromatin state dynamics in nine human cell types. Nature 473: 43 49.
31. Hampton OA, Den Hollander P, Miller CA, Delgado DA, Li J, et al. (2009) A sequence-level map of chromosomal breakpoints in the MCF-7 breast cancer cell line yields insights into the evolution of a cancer genome. Genome research 19: 167 177.
32. Schoenfelder S, Sexton T, Chakalova L, Cope NF, Horton A, et al. (2010) Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. Nature genetics 42: 53 61.
33. Kalhor R, Tjong H, Jayathilaka N, Alber F, Chen L (2012) Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. Nature biotechnology 30: 90 98.
34. Sproul D, Nestor C, Culley J, Dickson JH, Dixon JM, et al. (2011) Transcriptionally repressed genes become aberrantly methylated and distinguish tumors of different lineages in breast cancer. Proceedings of the National Academy of Sciences of the United States of America 108: 4364 4369.
35. Giresi PG, Kim J, McDaniell RM, Iyer VR, Lieb JD (2007) FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. Genome research 17: 877 885.
36. Rickman DS, Soong TD, Moss B, Mosquera JM, Dlabal J, et al. (2012) Oncogene-mediated alterations in chromatin conformation. Proceedings of the National Academy of Sciences of the United States of America.
37. Lin SY, Makino K, Xia W, Matin A, Wen Y, et al. (2001) Nuclear localization of EGF receptor and its potential new role as a transcription factor. Nature cell biology 3: 802 808.
38. Salomon DS, Brandt R, Ciardiello F, Normanno N (1995) Epidermal growth factor-related peptides and their receptors in human malignancies. Critical reviews in oncology/hematology 19: 183 232.
39. Cancer Genome Atlas Network (2012) Comprehensive molecular portraits of human breast tumours. Nature 490: 61 70.
40. Buck E, Eyzaguirre A, Rosenfeld-Franklin M, Thomson S, Mulvihill M, et al. (2008) Feedback mechanisms promote cooperativity for small molecule inhibitors of epidermal and insulin-like growth factor receptors. Cancer research 68: 8322 8332.
41. Butt AJ, Martin JL, Dickson KA, McDougall F, Firth SM, et al. (2004) Insulin-like growth factor binding protein-3 expression is associated with growth stimulation of T47D human breast cancer cells: the role of altered epidermal growth factor signaling. The Journal of clinical endocrinology and metabolism 89: 1950 1956.
42. Takaoka M, Harada H, Andl CD, Oyama K, Naomoto Y, et al. (2004) Epidermal growth factor receptor regulates aberrant expression of insulin-like growth factor-binding protein 3. Cancer research 64: 7711 7723.
43. Jackson DA, Hassan AB, Errington RJ, Cook PR (1993) Visualization of focal sites of transcription within human nuclei. The EMBO journal 12: 1059 1065.
44. Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, et al. (2012) Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell 148: 84 98.
45. Zhang Y, McCord RP, Ho YJ, Lajoie BR, Hildebrand DG, et al. (2012) Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. Cell 148: 908 921.
46. Neves H, Ramos C, da Silva MG, Parreira A, Parreira L (1999) The nuclear topography of ABL, BCR, PML, and RARalpha genes: evidence for gene proximity in specific phases of the cell cycle and stages of hematopoietic differentiation. Blood 93: 1197 1207.
47. Rocha PP, Micsinai M, Kim JR, Hewitt SL, Souza PP, et al. (2012) Close Proximity to Igh Is a Contributing Factor to AID-Mediated Translocations. Molecular cell 47(6): 873 885.
48. Roix JJ, McQueen PG, Munson PJ, Parada LA, Misteli T (2003) Spatial proximity of translocation-prone gene loci in human lymphomas. Nature genetics 34: 287 291.
49. Lin C, Yang L, Tanasa B, Hutt K, Ju BG, et al. (2009) Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. Cell 139: 1069 1083.
50. Mani RS, Tomlins SA, Callahan K, Ghosh A, Nyati MK, et al. (2009) Induced chromosomal proximity and gene fusions in prostate cancer. Science 326: 1230.
51. Kallioniemi A, Kallioniemi OP, Piper J, Tanner M, Stokke T, et al. (1994) Detection and mapping of amplified DNA sequences in breast cancer by comparative genomic hybridization. Proceedings of the National Academy of Sciences of the United States of America 91: 2156 2160.
52. Tanner MM, Tirkkonen M, Kallioniemi A, Holli K, Collins C, et al. (1995) Amplification of chromosomal region 20q13 in invasive breast cancer: prognostic implications. Clinical cancer research : an official journal of the American Association for Cancer Research 1: 1455 1461.
53. Qi C, Zhu YT, Chang J, Yeldandi AV, Rao MS, et al. (2005) Potentiation of estrogen receptor transcriptional activity by breast cancer amplified sequence 2. Biochemical and biophysical research communications 328: 393 398.
54. Kuo PC, Tsao YP, Chang HW, Chen PH, Huang CW, et al. (2009) Breast cancer amplified sequence 2, a novel negative regulator of the p53 tumor suppressor. Cancer research 69: 8877 8885.
55. Gururaj AE, Holm C, Landberg G, Kumar R (2006) Breast cancer-amplified sequence 3, a target of metastasis-associated protein 1, contributes to tamoxifen resistance in premenopausal patients with breast cancer. Cell cycle 5: 1407 1410.
56. Ruan Y, Ooi HS, Choo SW, Chiu KP, Zhao XD, et al. (2007) Fusion transcripts and transcribed retrotransposed loci discovered through comprehensive

transcriptome analysis using Paired-End diTags (PETs). Genome research 17: 828 838.
57. Barlund M, Monni O, Weaver JD, Kauraniemi P, Sauter G, et al. (2002) Cloning of BCAS3 (17q23) and BCAS4 (20q13) genes that undergo amplification, overexpression, and fusion in breast cancer. Genes, chromosomes & cancer 35: 311 317.
58. Soutoglou E, Misteli T (2007) Mobility and immobility of chromatin in transcription and genome stability. Current Opinion in Genetics & Development 17: 435 442.
59. Ryba T, Hiratani I, Lu J, Itoh M, Kulik M, et al. (2010) Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. Genome research 20: 761 770.
60. Hu Q, Kwon YS, Nunez E, Cardamone MD, Hutt KR, et al. (2008) Enhancing nuclear receptor-induced transcription requires nuclear motor and LSD1-dependent gene networking in interchromatin granules. Proceedings of the National Academy of Sciences of the United States of America 105: 19199 19204.
61. Han HJ, Russo J, Kohwi Y, Kohwi-Shigematsu T (2008) SATB1 reprogrammes gene expression to promote breast tumour growth and metastasis. Nature 452: 187 193.
62. Gheldof N, Leleu M, Noordermeer D, Rougemont J, Reymond A (2012) Detecting long-range chromatin interactions using the chromosome conformation capture sequencing (4C-seq) method. Methods in molecular biology 786: 211 225.
63. Splinter E, de Wit E, van de Werken HJ, Klous P, de Laat W (2012) Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: From fixation to computation. Methods 58(3): 221 230.
64. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, et al. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome biology 3: RESEARCH0034.
65. Zhao S, Fernald RD (2005) Comprehensive algorithm for quantitative real-time polymerase chain reaction. Journal of computational biology : a journal of computational molecular cell biology 12: 1047 1064.
66. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic acids research 29: e45.

European Journal of Cancer (2015) 51, 260 270

Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Aberrant allele-switch imprinting of a novel *IGF1R* intragenic antisense non-coding RNA in breast cancers

Lihua Kang [a], Jingnan Sun [a,b], Xue Wen [a], Jiuwei Cui [a], Guanjun Wang [a], Andrew R. Hoffman [b], Ji-Fan Hu [a,b,*,1], Wei Li [a,*,1]

[a] *Stem Cell and Cancer Center, First Affiliated Hospital, Jilin University, Changchun, Jilin 130061, PR China*
[b] *Stanford University Medical School, Stanford, CA 94305, USA*

Received 17 July 2014; received in revised form 14 October 2014; accepted 27 October 2014
Available online 20 November 2014

KEYWORDS
Genomic imprinting
IRAIN
Non coding RNA
Breast cancer
IGF1R
Allelic expression
Gene regulation
DNA methylation

Abstract The insulin like growth factor type I receptor (*IGF1R*) is frequently dysregulated in breast cancers, yet the molecular mechanisms are unknown. A novel intragenic long non coding RNA (lncRNA) *IRAIN* within the *IGF1R* locus has been recently identified in haematopoietic malignancies using RNA guided chromatin conformation capture (R3C). In breast cancer tissues, we found that *IRAIN* lncRNA was transcribed from an intronic promoter in an antisense direction as compared to the *IGF1R* coding mRNA. Unlike the *IGF1R* coding RNA, this non coding RNA was imprinted, with monoallelic expression from the paternal allele. In breast cancer tissues that were informative for single nucleotide polymorphism (SNP) rs8034564, there was an imbalanced expression of the two parental alleles, where the 'G' genotype was favorably imprinted over the 'A' genotype. In breast cancer patients, *IRAIN* was aberrantly imprinted in both tumours and peripheral blood leucocytes, exhibiting a pattern of allele switch: the allele expressed in normal tissues was inactivated and the normally imprinted allele was expressed. Epigenetic analysis revealed that there was extensive DNA demethylation of CpG islands in the gene promoter. These data identify *IRAIN* lncRNA as a novel imprinted gene that is aberrantly regulated in breast cancer.

1. Introduction

Despite recent advances in molecular therapeutics, breast cancer remains a highly lethal malignancy worldwide [1]. Anti-human epidermal growth factor receptor 2 (HER2) antibody therapy using Herceptin has been successful in the treatment of HER2-positive early stage

* *Corresponding authors at:* Stem Cell and Cancer Center, First Affiliated Hospital, Jilin University, Changchun, Jilin 130061, PR China. Tel: +86 431 8878 2047; fax: +86 431 8878 6102.
E mail addresses: Jifan@stanford.edu (J. F. Hu), drweili@yahoo.com (W. Li).
[1] These authors are senior authors of this report.

http://dx.doi.org/10.1016/j.ejca.2014.10.031
0959 8049/

and metastatic breast cancers [2,3]. However, resistance to Herceptin therapy has become an obstacle for treatment of HER2-positive breast cancer patients [4,5]. The activation of alternative growth factor pathways, particularly via the insulin-like growth factor 1 receptor (*IGF1R*), represents a common feature of Herceptin-refractory cells [6].

IGF1R is one of the most abundantly phosphorylated receptor tyrosine kinases in tumours [7 9]. The insulin-like growth factor system, including the type I IGF receptor *IGF1R* and the mitogenic ligands IGF-I and IGF-II, is frequently dysregulated in breast cancer and is known to contribute to disease progression and metastasis [10 14]. IGF-I and IGF-II promote cell growth and survival via the IGF1R receptor-mediated signal transduction through intracellular tyrosine kinase linked to the phosphatidyl-inositol-3 kinase (PI3K)-Akt-mammalian target of rapamycin (mTOR) pathway. Overexpression of *IGF1R* activates the PI3-K and mitogen-activated protein kinases (MAPK) signal cascades, resulting in cell proliferation and resistance to chemotherapeutic agents, radiation, and targeted therapies using Tamoxifen and Herceptin [15 17]. Therapeutic agents targeting *IGF1R* are currently in clinical development [10 14,18 23], including those that inhibit the IGF1R tyrosine kinase using monoclonal antibodies and small molecules [24]. However, the clinical development of various IGF1R inhibitors has been put on hold due to lack of sufficient clinical efficacy. Thus, the regulation of this pathway needs to be further defined to aid in the development of next generation regimens.

Currently, the molecular mechanisms underlying the dysregulation of the *IGF1R* pathway in tumours remain unknown. Using a recently-developed R3C (**R**NA-guided **C**hromatin **C**onformation **C**apture) technique [25], we recently identified a novel long non-coding RNA (lncRNA) *IRAIN* within the *IGF1R* locus [26]. *IRAIN* is transcribed from an intragenic promoter located in the first intron of *IGF1R*. *IRAIN* lncRNA is transcribed in an antisense orientation compared with the IGF1R gene, and it is expressed exclusively from the paternal allele, with the maternal allele being silenced. Interestingly, this lncRNA interacts with chromatin DNA and is involved in the formation of an intrachromosomal enhancer/promoter loop. In addition, *IRAIN* was downregulated in leukaemia cell lines and in leucocytes from patients with high-risk AML [26]. These data suggested that *IRAIN* might play a role in the dysregulation of the IGF pathway in haematopoietic malignancies.

However, the function of this non-coding RNA in other malignancies remains to be explored. The *IGF1R* pathway is frequently dysregulated in breast cancer. It is unclear if *IRAIN* lncRNA is aberrantly imprinted in breast cancer patients. In this communication, we characterise the allelic expression of *IRAIN* lncRNA in a cohort of breast cancer samples.

2. Materials and methods

2.1. Breast cancer cell lines and tissues

Breast cancer cell lines (MCF7 and MDA-MB-231) used in this study were purchased from ATCC. Cells were grown in RP1640 Media, supplemented with 10% foetal bovine serum (FBS), 100 U/ml penicillin and 100 μg/ml streptomycin.

Breast tumour specimens were collected from 74 female patients with invasive breast cancer who were treated at The First Hospital of Jilin University between 2007 and 2010. All tumour samples were obtained from patients with invasive ductal carcinomas (IDC) (n 74) (Table S1). Normal breast tissues (n 9) were collected as control samples from the patients who either underwent prophylactic mastectomy (n 3) or in whom normal breast tissue was removed at a site distant from the primary tumour (n 6). The protocol was approved by the Human Medical Ethical Review Committee from Jilin University First Hospital and informed consent was obtained from each breast cancer patient and normal subject.

Samples were snap frozen in liquid nitrogen at the time of the pre-therapeutic biopsy or surgical treatment and were stored at 80 °C for total RNA and genomic DNA extraction. The pathological diagnosis was made in accordance with the histological classification of tumours developed by the World Health Organisation. Tumour stage was defined according to American Joint Committee on Cancer/International Union Against Cancer tumour, node, metastasis (TNM) classification system. Tumours were histologically graded according to the Elston and Ellis method. Molecular markers, including the oestrogen receptor (ER), the progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2) and the mitotic index Ki67, were examined by using immunohistochemical (IHC) methods. Patients with Her2/neu2^{+} were tested for gene amplification using fluorescence in situ hybridisation (FISH) for validation. Breast cancer molecular subtype was defined by IHC receptor status of breast cancer according to St Gallen International Breast Cancer Conference (2013). We divided the patients into three groups: (1) Triple negative breast cancer (TNB, ER^{-}, PR^{-}, HER2^{-}), (2) HER2^{+} (ER^{-}, PR^{-}, and HER2^{+}), (3) luminal ['Luminal A-like': ER^{+} and PR^{+} ($\geqslant$20%), HER2 , Ki67 < 14%; 'Luminal B-like (HER2 negative)': ER^{+}, HER2 , and at least one of: Ki-67 $\geqslant$14%, PR^{-}, PR^{+} (<20%); 'Luminal B-like (HER2 positive)': ER^{+}, HER2^{+}, Any Ki-67, Any PR)]. After pathologic diagnosis, patients were treated according to standard clinical protocols. Clinical data such as date of birth, sex, date of surgery were extracted from the computerised clinical database.

Peripheral blood leucocytes (PBL) collected from breast cancer patients were isolated by Ficoll-Hypaque

(Sigma, MO) centrifugation and then cryopreserved for DNA and RNA analyses.

2.2. Reverse transcription-polymerase chain reaction (PCR) analysis

As previously described [25,27], total RNA was extracted from tissues by TRI-REAGENT (Sigma, CA) and cDNA was synthesised with RNA reverse transcriptase. Briefly, 1 μg of total RNA was used, and PCR was carried out under liquid wax in a 6 μl reaction mixture containing 2 μl of 3 × Klen-T*aq* I Mix, 2 μl cDNA and 1 μl of each 2.5 μM primer. After incubation at 95 °C for 2 min, *IRAIN* cDNA was amplified by 32 cycles of 95 °C for 30 s, 65 °C for 30 s of annealing and 72 °C for 35 s of extension, and finally with extension at 72 °C for 5 min.

For real-time quantitative polymerase chain reaction (qPCR), cDNA samples were amplified using CFX96™ real-time system (BIO-RAD) by SYBR PrimeScript™ RT-PCR Kit (Takara, Japan). The mRNA expression level of *IRAIN* and *IGF1R* was quantitated by normalising with β-actin (housekeeping gene) as previously described [27,28]. PCR primers used for qPCR and RT-CPR are listed in Table S2.

2.3. Gene strand-specific RT-PCR

The orientation of *IRAIN* was mapped with a strand-specific PCR (SSRT) assay [26,27]. Total RNA was extracted from tissues by TRI-REAGENT (Sigma, MA). Total RNA (400 ng) was reverse transcribed with the *IRAIN* 5′- or 3′-primers using Maxima Reverse Transcriptase (Thermo Fisher Scientific, CA) at 60 °C for 50 min, followed by 85 °C for 5 min to inactivate the transcriptase. After 10-fold dilution, PCR was carried out under liquid wax in a 6 μl reaction containing 2 μl of 3 × Klen-T*aq* I Mix, 2 μl cDNA and 1 μl of each 2.5 μM downstream PCR primer set. After initial denaturing at 95 °C for 2 min, *IRAIN* cDNA was amplified by 32 cycles at 95 °C for 30 s, 65 °C for 30 s of annealing and 72 °C for 35s of extension, followed by incubation at 72 °C for 5 min. PCR primers used for strand-specific PCR are listed in Table S2.

2.4. Allelic expression

Quantitation of allelic expression requires the presence of heterozygous single nucleotide polymorphisms (SNPs) to distinguish the two parental alleles. We extracted genomic DNAs from breast cancer samples and screened for heterozygosity of SNPs in *IRAIN* lncRNA [26]. Only those SNP-informative breast cancer samples were used for *IRAIN* imprinting analysis. However, for imprinting assessment, the data of allelic distribution of the parents were needed to track which parental allele was expressed. In these studies, only those cases with parental information available were included. To examine differential allelic expression of *IRAIN* lncRNA between tumours and peripheral blood leucocytes (PBL), only those informative cohort cases with available blood samples were included in the study.

Total RNA extraction and cDNA synthesis were performed as previously described [29,30]. Allelic expression of *IRAIN* was examined using the same program as in the RT-PCR but using primers specific for polymorphic restriction enzymes. Allelic expression of *IRAIN* was assessed by polymorphic restriction enzymes Nde1. In some cases, DNA sequencing of genomic DNA and cDNA PCR products was used to determine allelic expression of *IRAIN*. PCR primers used to assess allelic expression covering the Nde1 polymorphic site were JH891 and JH892, and primers for allelic sequencing using SNP rs8034564 were JH248 and JH781 (listed in Table S2).

2.5. DNA methylation analysis

Genomic DNA was extracted from cells and tumours with Perfect gDNA kit (Eppendorf, NY). Genomic DNA (1 μg) was used for bisulphite conversion with EZ DNA Methylation-Gold™ Kit (ZYMO Research, CA) according to the manufacturer's instructions. PCR reactions were performed using Kantaq1 DNA polymerase (Ab Peptides, MO). Bisulphite-sequencing PCR (BSP) was used to analyse DNA methylation status. PCR conditions were 97 °C for 10 min followed by 35 cycles of 96 °C for 20 s, 64 °C for 30 s of annealing, 72 °C for 30 s of extension, and completing the reaction at 72 °C for 10 min. The primers used for assessing DNA methylation were JH852 and JH855 (Table S2) [26]. PCR products were separated by gel electrophoresis and purified with Axygen DNA Gel Extraction kit (Axygen, CA). The PCR products were cloned into CloneJET vector using PCR Cloning Kit (Thermo Scientific # K1231, MA), and sequenced for analysis of CpG methylation.

2.6. Statistical analysis

All experiments were performed in triplicate, and the data are expressed as mean ± SD. The data were analysed by one-way analysis of variance, and results were considered statistically significant at $P \leqslant 0.05$.

3. Results

3.1. Ubiquitous expression of IGF1R intragenic non-coding RNA in breast cancers

In studying the tumour-specific dysregulation of *IGF1R*, we recently identified a novel 5366 bp *IGF1R*-intragenic long non-coding RNA (*IRAIN*) that is associated with haematopoietic malignancies [26]. To

determine if this non-coding RNA is aberrantly regulated in breast cancer, we first used RT-PCR to determine the presence of this non-coding RNA in breast cancer tissues. We found that *IRAIN* lncRNA was ubiquitously expressed at various levels in the breast cancer samples (Fig. 1A).

We then grouped the breast cancer patients into three groups: (1) Triple-negative breast cancer, (2) $HER2^{+}$ and (3) Luminal (luminal A-like, luminal B1-like, luminal B2-like) as described in Section 2. As shown in Fig. 1B, *IRAIN* lncRNA was downregulated in both TNB and $HER2^{+}$ groups ($P < 0.05$).

3.2. IRAIN is transcribed antisense to IGF1R in breast cancers

We then used a strand-specific RT-PCR (SSRT) method to map the orientation of gene transcription. SSRT cDNA was synthesised by Thermo-stable reverse transcriptase utilising a 5′-specific oligonucleotide or a 3′-specific oligonucleotide, respectively. After SSRT, a pair of downstream PCR primers was used to amplify the strand-specific cDNA (Fig. 2A).

As seen in Fig. 2B, *IRAIN* RNA was detected only when cDNA was synthesised using 5′-oligonucleotides (#513, #400) (lanes 1, 4, 7, 10). No PCR products were amplified when the 3′ oligonucleotides were used (#514, #401, lanes 2, 5, 8, 11) or in the RT-minus controls (lanes 3, 6, 9, 12), indicating that *IRAIN* was transcribed in the antisense direction as compared with the *IGF1R* coding RNA.

3.3. Monoallelic expression of IRAIN in breast cancer tissues

In the mouse, the gene transcribing the Type 2 IGF receptor (*Igf2r*) is associated with an lncRNA *Airn* that is transcribed antisense to *Igf2r*. These transcripts are reciprocally imprinted, with *Airn* transcribed from the paternal allele only. The transcription of the antisense lncRNA *Airn* regulates in *cis* the allelic expression of the *Igf2r* coding RNA [31 34]. In leukaemia cells, we showed that *IRAIN* was expressed solely from the paternal allele [26]. To learn if *IRAIN* uses a similar epigenetic mechanism to regulate genes locally in breast cancers, we examined if *IRAIN* lncRNA was monoallelically expressed in the MCF7 breast cancer cell line, which is heterozygous for the polymorphic Nde1 restriction site. Two alleles, termed 'A' and 'G', were detected in genomic DNA (Fig. 3A, lanes 2 3). In cDNA samples, however, only the 'A' allele was detected (lanes 5, 6). The other parental allele (G), in contrast, was totally suppressed. These data indicate that *IRAIN* lncRNA is monoallelically transcribed in the MCF7 breast cancer cell line.

We then examined the allelic expression of *IRAIN* lncRNA in breast cancer tissue samples using SNP rs8034564 to distinguish the two parental alleles. As this SNP does not contain a restriction enzyme site, PCR sequencing was used to determine the allelic expression of *IRAIN*. In three breast cancer tissues that were heterogeneous for this SNP, both the 'A' and 'G' alleles were observed in genomic DNAs (gDNA). However, in all cDNA samples tested, only a single parental allele (A)

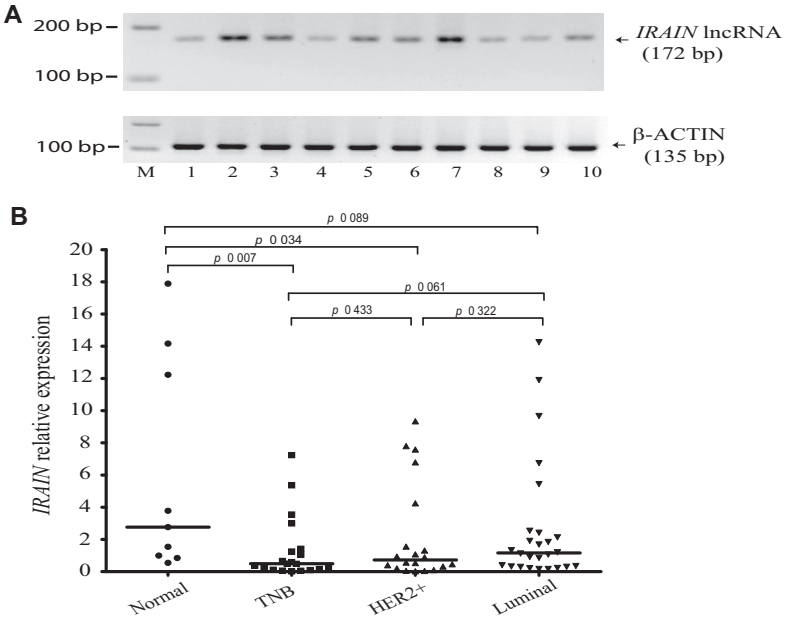


Fig. 1. Downregulation of *IRAIN* long non coding RNA (lncRNA) in breast cancer. (A) Ubiquitous expression of *IRAIN* in breast cancer tissues. Expression of *IRAIN* lncRNA was analysed by reverse transcription polymerase chain reaction (RT PCR). β Actin was used as the internal control. Lanes 1 4: HER2+; lanes 5 7: Luminal; lanes 8 10: TNB. (B) Dysregulation of *IRAIN* lncRNA in breast cancer subtypes. TNB, Triple negative breast cancer (TNB, ER , PR , HER2); $HER2^{+}$: (ER , PR , and $HER2^{+}$); Luminal: ['Luminal A like': ER^{+} and PR^{+} ($\geqslant$20%), HER2 , Ki67 <14%; 'Luminal B like (HER2 negative)': ER^{+}, HER2 , and at least one of: Ki 67 $\geqslant$14%, PR , PR^{+} (<20%); 'Luminal B like (HER2 positive)': ER^{+}, $HER2^{+}$, Any Ki 67, Any PR)]. $^{*}p < 0.05$ between the groups.

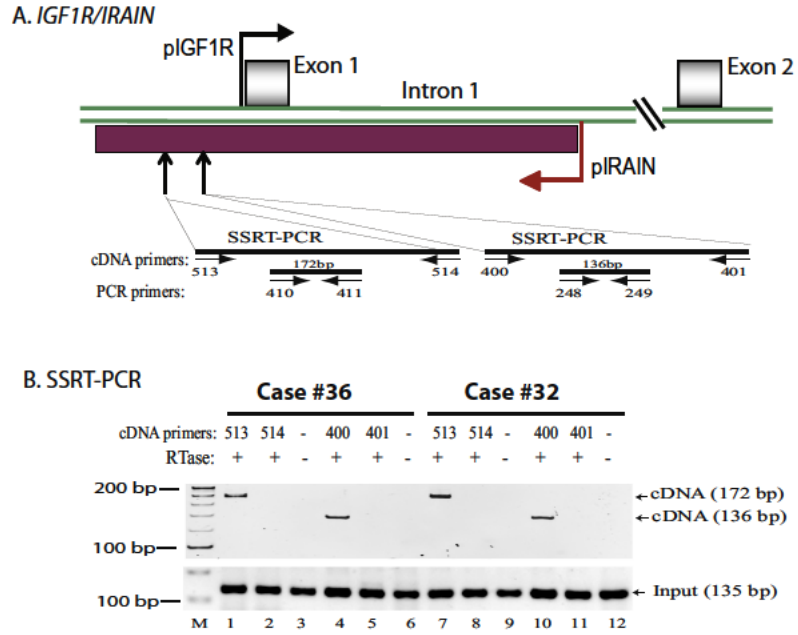


Fig. 2. *IRAIN* is an antisense long non coding RNA (lncRNA). (A) Diagram of the *IRAIN*/*IGF1R* locus. pIRAIN: *IRAIN* lncRNA promoter that is transcribed in antisense; pIGF1R: *IGF1R* coding RNA promoter that is transcribed in sense. Horizontal arrows: SSRT PCR primers used to map the orientation of *IRAIN* lncRNA. (B) *IRAIN* lncRNA is an antisense lncRNA. The strand specific cDNAs were synthesised using either the 5′ or the 3′ oligonucleotides. A pair of polymerase chain reaction (PCR) primers located between two cDNA oligonucleotides was then used to determine the transcription orientation of the *IRAIN* lncRNA. M: 100 bp marker; input: total RNA collected before SSRT PCR.

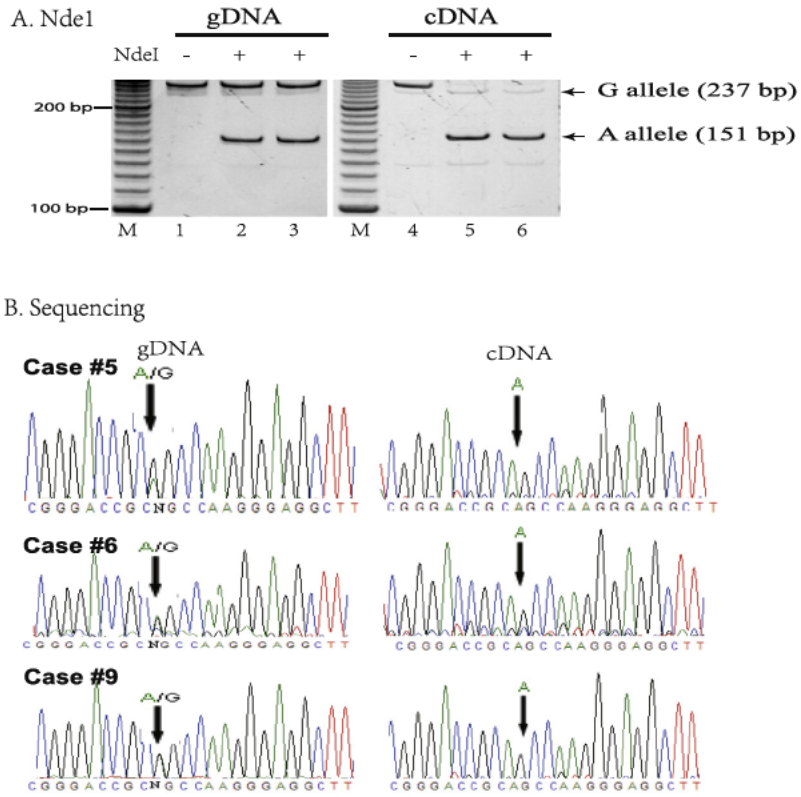


Fig. 3. Monoallelic expression of *IRAIN* long non coding RNA (lncRNA). (A) Monoallelic expression of *IRAIN* lncRNA in MCF7 breast cancer cell line. gDNA: heterozygous genomic DNA. Note only the single 'A' allele of *IRAIN* lncRNA was detected in cDNA samples. (B) Monoallelic expression of *IRAIN* lncRNA in breast cancer tissues. In breast cancers that are heterogeneous for the *IRAIN* polymorphic site in genomic DNA, only the 'A' allele was expressed. (C) Parental imprinting of *IRAIN* lncRNA. Genomic DNA and cDNA from peripheral blood leucocytes (PBL) were amplified and PCR products were sequenced for the A/G alleles. Note the *IRAIN* lncRNA was expressed from the paternal allele.

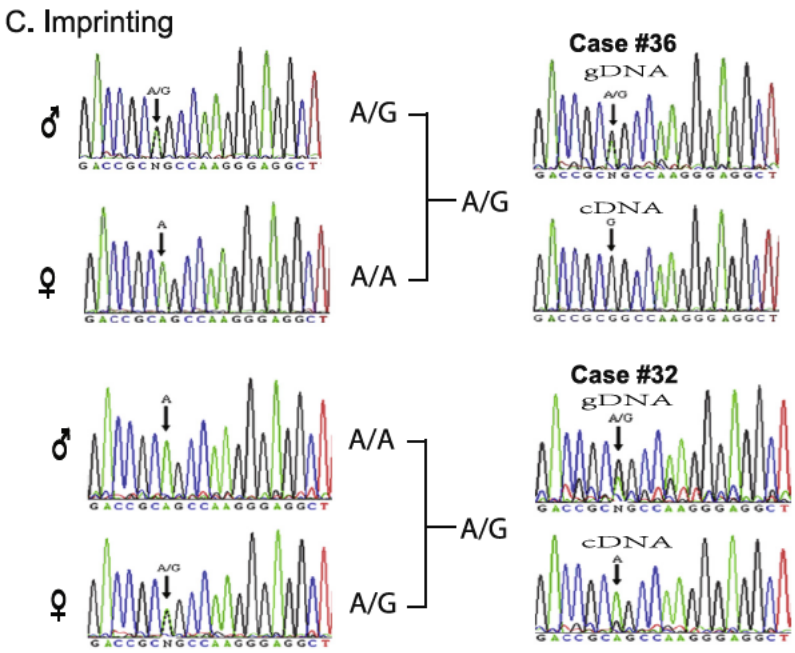


Fig 3. *(continued)*

was detected (Fig. 3B), suggesting that *IRAIN* is monoallelically expressed in breast cancer.

In examining allelic expression, it was interesting to note the expression of the 'A' allele was favored over the 'G' allele. In 18 breast cancers that were heterozygous for the polymorphic SNP, 16 tumour cDNAs expressed the 'A' allele alone (Table 1).

3.4. *IRAIN* is imprinted in breast cancer tissues

To determine if *IRAIN* is imprinted, we tracked the expression from the paternal or maternal allele using peripheral blood leucocytes from two patients whose parents had also donated blood samples. In Case #36, the father was heterozygous for the A and G alleles while the mother was homozygous for the A allele. The patient was informative at the polymorphic site, carrying both the A and G alleles in the genomic DNA. In the cDNA sample, however, we detected the expression of *IRAIN* lncRNA only from the G allele that was inherited from the father (Fig. 3C, left panel), demonstrating that *IRAIN* is paternally expressed and maternally suppressed. We also confirmed the paternal expression in Case #32. The patient was heterozygous for the A and G alleles. In the cDNA sample, only the paternal A allele was expressed (right panel). Thus, *IRAIN* is maternally imprinted, in agreement with our previous finding in leukaemia samples [26].

Table 1
The favored 'A' monoallelic expression of *IRAIN* in breast cancer tissues.

Case	DNA genotype	cDNA allelic expression	
		G	A
5	AG		+
6	AG		+
9	AG		+
11	AG		+
12	AG		+
14	AG		+
16	AG		+
18	AG		+
19	AG		+
23	AG		+
28	AG		+
29	AG		+
32	AG		+
36	AG	+	
37	AG		+
39	AG		+
42	AG	+	
43	AG		+
Total		2 (11%)	16 (89%)

3.5. *DNA methylation in the IRAIN promoter*

The *IRAIN* promoter is very rich in CpG dinucleotides. In peripheral blood leucocytes, the promoter CpG islands were semi-methylated [26]. We analysed the status of DNA methylation in the *IRAIN* promoter in our breast cancer specimens and cell lines (Fig. 4A). Using bisulphite sequencing, we found that the *IRAIN* promoter is totally unmethylated in two breast cancer cell lines (MCF-7, MDA-MB-231) (Fig. 4B).

We also examined DNA methylation in three breast cancer tissues that show monoallelic expression. In the breast cancer specimens, we observed a hemi-methylation pattern in the *IRAIN* promoter in two breast cancer samples (Cases #5, #9) (Fig. 4C). However, in Case #6, the *IRAIN* promoter was almost totally unmethylated, as was seen in the two breast cancer cell lines. Thus,

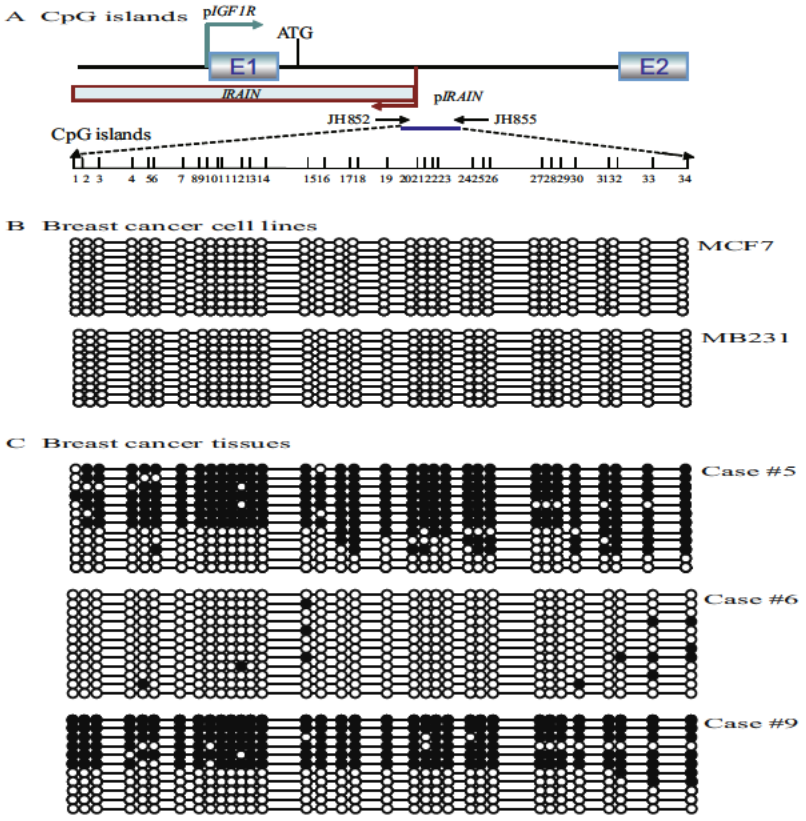


Fig. 4. DNA methylation in the regulation of the *IRF1R/IRAIN* locus. (A) Schematic diagram of CpG islands in the *IRAIN* promoter. Vertical lines: location of CpG islands. (B) DNA methylation of the *IRAIN* promoter in breast cancer cell lines (MCF7 and MDA MB 231). Genomic DNAs were extracted from breast cancer cells. After treated with sodium sulphite, the *IRAIN* promoter DNA was amplified and sequenced. Open circles: unmethylated CpGs; solid circles: methylated CpGs. (C) CpG methylation of the *IRAIN* promoter in three breast cancer patients.

compared to peripheral blood leucocytes, breast cancer specimens had aberrant DNA methylation in the *IRAIN* promoter.

3.6. Aberrant imprinting of IRAIN lncRNA in breast cancers

Loss of *IGF2* imprinting is a very common epigenetic abnormality in cancer and may even be a prognostic biomarker [35]. In order to compare the status of *IRAIN* imprinting in peripheral blood leucocytes and breast cancer tissues, we studied five patients for whom both blood and surgery samples were available to track allelic expression.

As seen in Fig. 5A, all peripheral blood leucocyte cDNAs showed monoallelic expression of *IRAIN*. Four cases (#32, #36, #42, #37) expressed *IRAIN* lncRNA monoallelically from the 'A' allele, while case #39 expressed the 'G' allele. Surprisingly, in contrast to peripheral blood leucocytes, we found that in two cases (#36, #42), *IRAIN* expression switched to the 'G' allele in breast cancer specimens. In case #37, the breast cancer expressed the 'A' allele, while metastatic tumours switched to 'G' allele expression. These data suggest that in breast cancer tissues, *IRAIN* can undergo allelic switch, expressing the opposite allele as compared with that in circulating cells.

The *IRAIN* promoter was aberrantly unmethylated in breast cancer samples (Fig. 5B). Intriguingly, this aberrant demethylation pattern was also observed in peripheral blood leucocytes. Thus, it seems that breast cancer patients may undergo extensive alterations in promoter epigenotype. In the human *IGF2* gene, DNA demethylation is also a common epigenetic mutation observed in many human tumours [36 40].

4. Discussion

As the *IGF1R* signalling pathway is often aberrantly activated in tumours, including breast cancers, treatments using small molecule inhibitors and antibodies to block the tyrosine kinase activity have been advanced in preclinical and clinical testing [24]. In this communication, we have characterised *IRAIN*, a novel 5.4 kb intragenic non-coding RNA within the *IGF1R* locus in clinical samples collected from breast cancer patients. In cancer tissues, *IRAIN* is expressed in an antisense orientation within the *IGF1R* locus. A unique characteristic of this non-coding RNA is its monoallelic expression in breast cancers (Fig. 3A and B). By tracking the allelic expression in patient families, we demonstrate that *IRAIN* is transcribed from the paternal allele, while the copy from the maternal allele is silenced or imprinted (Fig. 3C), in agreement with the data in haematopoietic

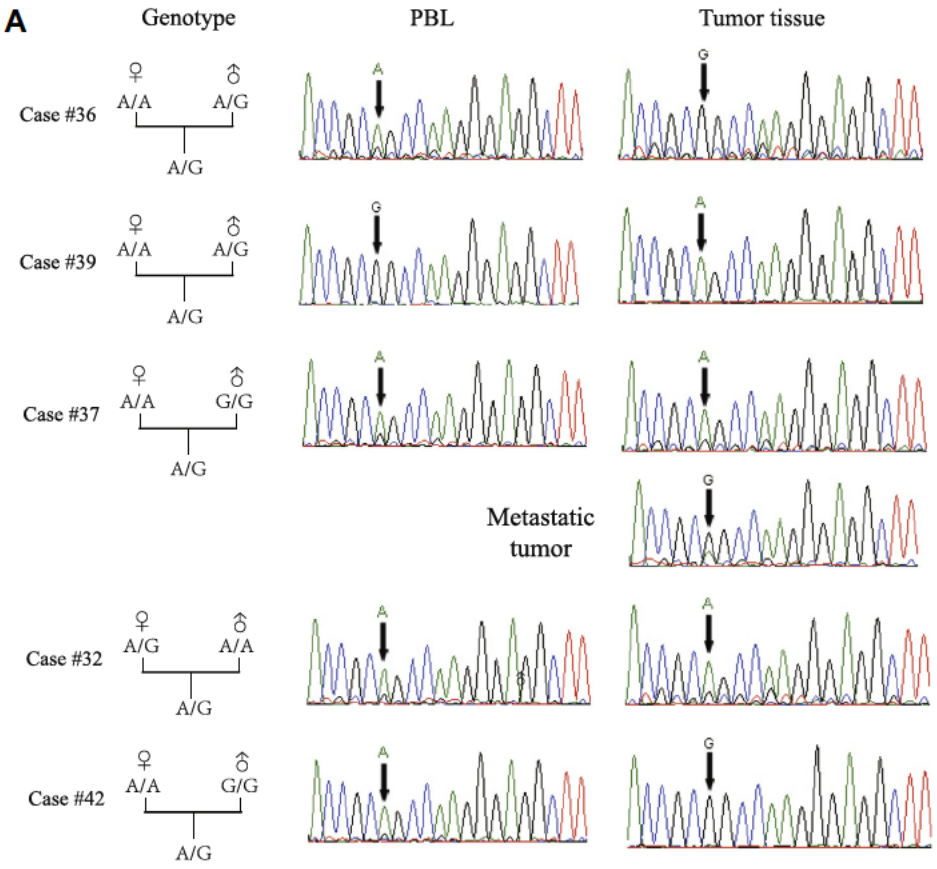


Fig. 5. Aberrant imprinting of IRAIN in breast cancer patients. (A) Allelic switch of *IRAIN* long non coding RNA (lncRNA) imprinting. Allelic expression of *IRAIN* was examined by DNA sequencing of the single nucleotide polymorphism (SNP) rs8034564. Note the allele switch of *IRAIN* lncRNA between the breast cancer tissues and peripheral blood leucocytes. PBL: peripheral blood leucocytes. (B) Aberrant DNA demethylation in the *IRAIN* promoter in allelic switching tumours. DNA methylation in CpG islands in the *IRAIN* promoter was quantitated by bisulphite sequencing. Open circles: unmethylated CpGs; solid circles: methylated CpGs. Note that in allelic switching tumours, there is extensive DNA demethylation of the *IRAIN* promoter in both breast cancer specimens and peripheral blood leucocytes.

malignancies [26]. Together, our data validate *IRAIN* as a new member of the family of imprinted genes [41].

By comparing the genotypes in informative tissues, it is interesting to note that *IRAIN* lncRNA seems to prefer the 'A' genotype expression (Table 1). In 18 heterozygous breast tissues examined, 89% expressed the 'A' allele. The 'G' allele, however, is rarely used by the host machinery for transcription. Similar cases have been reported in the *TMPRSS2* gene in prostate cancer stem cells [42]. In addition, stochastic monoallelic expression is also widespread in mammalian genomes, including olfactory receptor, *V1rb2* receptor, T-cell receptor and immunoglobulin genes, pheromone receptors, p120 catenin, odorant receptors, and protocadherins [43 47]. Allele-biased expression has also been observed in a number of putative schizophrenia (SZ) and autism spectrum disorder (ASD) SZ and ASD candidate genes, including *A2BP1* (*RBFOX1*), *ERBB4*, *NLGN4X*, *NRG1*, *NRG3*, *NRXN1*, and *NLGN1* [48]. Random monoallelic expression in the brain is related to epidemiological features of neuropsychiatric disorders [49,50]. In this study, however, it is still unclear if the preferential 'A' allele expression is associated with the function of this non-coding RNA.

Several genes undergo aberrant imprinting in cancers [51 53]. The most extensively studied example is the paternally-expressed *IGF2* gene. In many tumours, both parental copies of the *IGF2* gene may become fully expressed [54 56]. Reactivation of the normally-suppressed *IGF2* (imprinted) maternal allele, known as loss of imprinting (LOI), is a hallmark of many human tumours, especially childhood tumours [54 61] and cancer stem cells [62]. In this study, we did not observe biallelic expression of *IRAIN* in tumours. However, we did show an *IRAIN* epigenetic abnormality in breast tumour specimens. In normal tissues, *IRAIN* lncRNA is expressed from the maternal allele. However, in breast cancer tissues, the expression of *IRAIN* lncRNA switches to the alternative parental allele (Fig. 5A). The mechanisms underlying this allele-switch in breast cancer are not known. It is also unclear if this aberrant allelic switch will affect the activity of the *IGF1R* signal pathway in breast cancer. CRISPR Cas9 RNA genome editing has been recently used to study gene function [63,64], and it

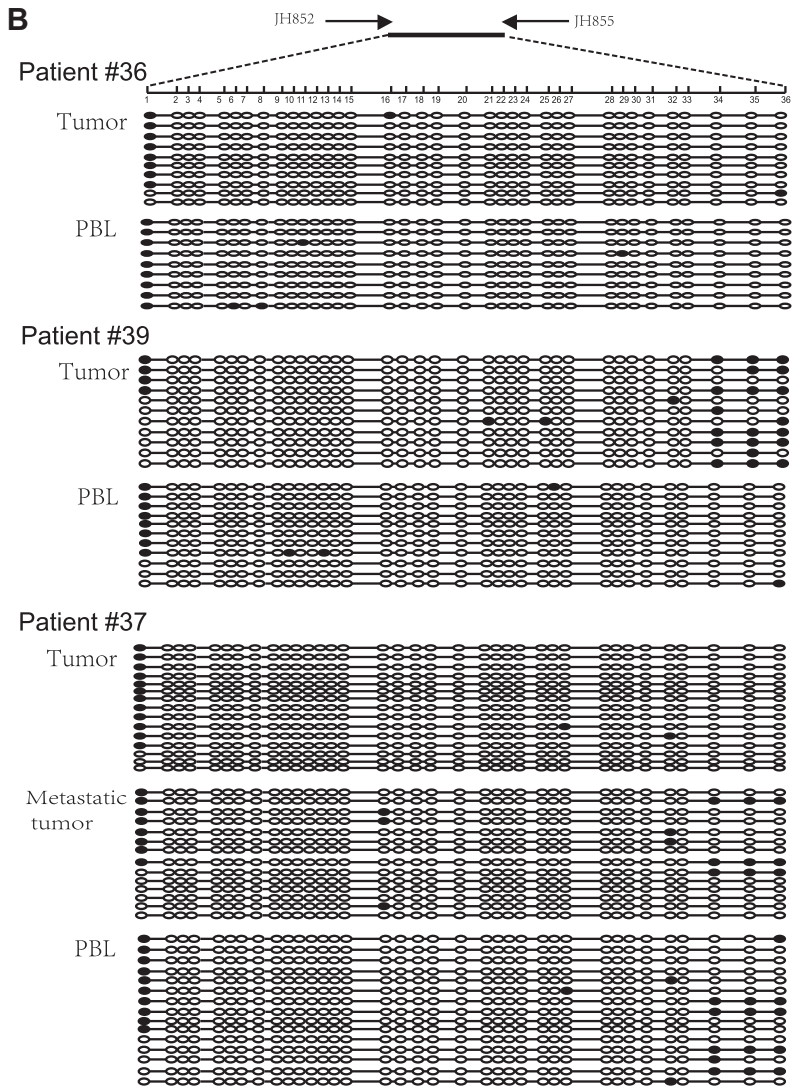


Fig 5. *(continued)*

would be interesting to learn if knockdown of *IRAIN* using this approach would affect *IGF1R* expression and thus the activity of the IGF signal pathway in tumours.

Allelic expression of sense and antisense RNAs is usually coupled via a *cis* transcription competition mechanism. A typical example is the mouse sense *Igf2r* coding RNA and the *Airn* antisense non-coding RNA, which are reciprocally imprinted [65 67] and tightly coordinated by DNA methylation in the *Airn* promoter [68]. The maternal *Airn* promoter is silenced by CpG island hypermethylation. Lack of the *Airn* lncRNA *cis*-competition leads to the expression of *Igf2r* from the maternal allele. In contrast, the unmethylated paternal *Airn* promoter leads to lncRNA expression, silencing the *Igf2r* promoter using a *cis* regulation mechanism [68,69]. In this study, however, we found that allelic expression between the *IRAIN* lncRNA and the *IGF1R* coding RNA is totally uncoupled. While the *IRAIN* lncRNA is monoallelically expressed (Fig. 3), the *IGF1R* coding mRNA is known to be biallelically expressed [26,70,71]. However, the fact that both *IRAIN* antisense lncRNA and *IGF1R* sense RNA are transcribed from the paternal chromosome without transcription competition or inhibition may provide a unique model to study imprinting mechanisms [66,67].

In summary, we have identified *IRAIN* as a novel maternally imprinted lncRNA located within the human *IGF1R* locus. In breast cancers, *IRAIN* undergoes aberrant allelic switching. However, many questions remain to be explored regarding this aberrant imprinting. For example, what is the impact of *IRAIN* expression/imprinting in the development of breast cancers? Could the aberrant allele-switch of *IRAIN* lncRNA be a prognostic biomarker? Is *IRAIN* lncRNA a predictive marker for *IGF1R* targeted therapies? Does the down-regulation of the *IRAIN* lncRNA in TNB and HER2+ samples correlate with clinic outcomes? Future studies are needed to address these questions. Detection of aberrant IGF2 imprinting in circulating leucocytes represents a valuable biomolecular marker for predicting individuals with high risk for colorectal cancer [38]. It would be interesting to learn whether the aberrant

allele-switch of *IRAIN* lncRNA can be utilised as a prognostic biomarker to assess breast cancer risk.

Conflict of interest statement

None declared.

Acknowledgements

This work was supported by California Institute of Regenerative Medicine (CIRM) grant (RT2-01942), Jilin International Collaboration Grant (#20120720), the National Natural Science Foundation of China grant (#81272294, #31430021) to J.F.H; the National Natural Science Foundation of China grants (#81071920, #81372835) to W.L.; the grant of Key Project of Chinese Ministry of Education (#311015) to C.J.; and grant BC102122 from the Breast Cancer Research Program Idea Award of the Department of Defense to A.R.H.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ejca.2014.10.031.

References

[1] Blackman DJ, Masi CM. Racial and ethnic disparities in breast cancer mortality: are we doing enough to address the root causes? J Clin Oncol 2006;24(14):2170 8.
[2] Harbeck N, Beckmann MW, Rody A, Schneeweiss A, Muller V, Fehm T, et al. HER2 dimerization inhibitor pertuzumab mode of action and clinical data in breast cancer. Breast Care (Basel) 2013;8(1):49 55.
[3] Slamon DJ, Leyland Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overex presses HER2. N Engl J Med 2001;344(11):783 92.
[4] Barok M, Joensuu H, Isola J. Trastuzumab emtansine: mechanisms of action and drug resistance. Breast Cancer Res 2014;13(2):R46.
[5] Nahta R, Yu D, Hung MC, Hortobagyi GN, Esteva FJ. Mechanisms of disease: understanding resistance to HER2 targeted therapy in human breast cancer. Nat Clin Pract Oncol 2006;3(5):269 80.
[6] Sharma SV, Lee DY, Li B, Quinlan MP, Takahashi F, Maheswaran S, et al. A chromatin mediated reversible drug tolerant state in cancer cell subpopulations. Cell 2010;141(1):69 80.
[7] Ge NL, Rudikoff S. Insulin like growth factor I is a dual effector of multiple myeloma cell growth. Blood 2000;96(8):2856 61.
[8] Sprynski AC, Hose D, Caillot L, Reme T, Shaughnessy Jr JD, Barlogie B, et al. The role of IGF 1 as a major growth factor for myeloma cell lines and the prognostic relevance of the expression of its receptor. Blood 2009;113(19):4614 26.
[9] Sprynski AC, Hose D, Kassambara A, Vincent L, Jourdan M, Rossi JF, et al. Insulin is a potent myeloma cell growth factor through insulin/IGF 1 hybrid receptor activation. Leukemia 2010;24(11):1940 50.
[10] Seccareccia E, Brodt P. The role of the insulin like growth factor I receptor in malignancy: an update. Growth Horm IGF Res 2012;22(6):193 9.
[11] Pollak M. The insulin and insulin like growth factor receptor family in neoplasia: an update. Nat Rev Cancer 2012;12(3):159 69.
[12] Ozkan EE. Plasma and tissue insulin like growth factor I receptor (IGF IR) as a prognostic marker for prostate cancer and anti IGF IR agents as novel therapeutic strategy for refractory cases: a review. Mol Cell Endocrinol 2011;344(1 2):1 24.
[13] Capoluongo E. Insulin like growth factor system and sporadic malignant melanoma. Am J Pathol 2011;178(1):26 31.
[14] Gallagher EJ, LeRoith D. The proliferating role of insulin and insulin like growth factors in cancer. Trends Endocrinol Metab 2010;21(10):610 8.
[15] Chapuis N, Tamburini J, Cornillet Lefebvre P, Gillot L, Bardet V, Willems L, et al. Autocrine IGF 1/IGF 1R signaling is responsible for constitutive PI3K/Akt activation in acute myeloid leukemia: therapeutic value of neutralizing anti IGF 1R anti body. Haematologica 2010;95(3):415 23.
[16] Grandage VL, Gale RE, Linch DC, Khwaja A. PI3 kinase/Akt is constitutively active in primary acute myeloid leukaemia cells and regulates survival and chemoresistance via NF kappaB, Map kinase and p53 pathways. Leukemia 2005;19(4):586 94.
[17] Xu Q, Simpson SE, Scialla TJ, Bagg A, Carroll M. Survival of acute myeloid leukemia cells requires PI3 kinase activation. Blood 2003;102(3):972 80.
[18] Karamouzis MV, Papavassiliou AG. Targeting insulin like growth factor in breast cancer therapeutics. Crit Rev Oncol Hematol 2012;84(1):8 17.
[19] Tognon CE, Sorensen PH. Targeting the insulin like growth factor 1 receptor (IGF1R) signaling pathway for cancer therapy. Expert Opin Ther Targets 2012;16(1):33 48.
[20] Rowinsky EK, Schwartz JD, Zojwalla N, Youssoufian H, Fox F, Pultar P, et al. Blockade of insulin like growth factor type 1 receptor with cixutumumab (IMC A12): a novel approach to treatment for multiple cancers. Curr Drug Targets 2011;12(14):2016 33.
[21] Buck E, Mulvihill M. Small molecule inhibitors of the IGF 1R/IR axis for the treatment of cancer. Expert Opin Invest Drugs 2011;20(5):605 21.
[22] Heidegger I, Pircher A, Klocker H, Massoner P. Targeting the insulin like growth factor network in cancer therapy. Cancer Biol Ther 2011;11(8):701 7.
[23] Jones RA, Campbell CI, Wood GA, Petrik JJ, Moorehead RA. Reversibility and recurrence of IGF IR induced mammary tumors. Oncogene 2009;28(21):2152 62.
[24] Moreau P, Cavallo F, Leleu X, Hulin C, Amiot M, Descamps G, et al. Phase I study of the anti insulin like growth factor 1 receptor (IGF 1R) monoclonal antibody, AVE1642, as single agent and in combination with bortezomib in patients with relapsed multiple myeloma. Leukemia 2011;25(5):872 4.
[25] Zhang H, Zeitz MJ, Wang H, Niu B, Ge S, Li W, et al. Long noncoding RNA mediated intrachromosomal interactions pro mote imprinting at the Kcnq1 locus. J Cell Biol 2014;204(1):61 75.
[26] Sun J, Li W, Sun Y, Yu D, Wen X, Wang H, et al. A novel antisense long noncoding RNA within the IGF1R gene locus is imprinted in hematopoietic malignancies. Nucleic Acids Res 2014;42(15):9588 601.
[27] Wang H, Li W, Guo R, Sun J, Cui J, Wang G, et al. An intragenic long noncoding RNA interacts epigenetically with the RUNX1 promoter and enhancer chromatin DNA in hematopoietic malig nancies. Int J Cancer 2014;135(12):2783 94.
[28] Li T, Chen H, Li W, Cui J, Wang G, Hu X, et al. Promoter histone H3K27 methylation in the control of IGF2 imprinting in human tumor cell lines. Hum Mol Genet 2014;23:117 28.
[29] Hu JF, Vu TH, Hoffman AR. Promoter specific modulation of insulin like growth factor II genomic imprinting by inhibitors of DNA methylation. J Biol Chem 1996;271(30):18253 62.
[30] Zhang S, Zhong B, Chen M, Yang G, Li Y, Wang H, et al. Epigenetic reprogramming reverses the malignant epigenotype of the MMP/TIMP axis genes in tumor cells. Int J Cancer 2014;134(7):1583 94.
[31] Stoger R, Kubicka P, Liu CG, Kafri T, Razin A, Cedar H, et al. Maternal specific methylation of the imprinted mouse Igf2r locus

identifies the expressed locus as carrying the imprinting signal. Cell 1993;73(1):61 71.
[32] Wutz A, Smrzka OW, Schweifer N, Schellander K, Wagner EF, Barlow DP. Imprinted expression of the Igf2r gene depends on an intronic CpG island. Nature 1997;389(6652):745 9.
[33] Hu JF, Balaguru KA, Ivaturi RD, Oruganti H, Li T, Nguyen BT, et al. Lack of reciprocal genomic imprinting of sense and antisense RNA of mouse insulin like growth factor II receptor in the central nervous system. Biochem Biophys Res Commun 1999;257(2):604 8.
[34] Hu JF, Pham J, Dey I, Li T, Vu TH, Hoffman AR. Allele specific histone acetylation accompanies genomic imprinting of the insulin like growth factor II receptor gene. Endocrinology 2000;141(12):4428 35.
[35] Cui H, Cruz Correa M, Giardiello FM, Hutcheon DF, Kafonek DR, Brandenburg S, et al. Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. Science 2003;299(5613):1753 5.
[36] Moore T, Constancia M, Zubair M, Bailleul B, Feil R, Sasaki H, et al. Multiple imprinted sense and antisense transcripts, differ ential methylation and tandem repeats in a putative imprinting control region upstream of mouse Igf2. Proc Natl Acad Sci USA 1997;94(23):12509 14.
[37] Sullivan MJ, Taniguchi T, Jhee A, Kerr N, Reeve AE. Relaxation of IGF2 imprinting in Wilms tumours associated with specific changes in IGF2 methylation. Oncogene 1999;18(52):7527 34.
[38] Cui H, Onyango P, Brandenburg S, Wu Y, Hsieh CL, Feinberg AP. Loss of imprinting in colorectal cancer linked to hypome thylation of H19 and IGF2. Cancer Res 2002;62(22):6442 6.
[39] Chen HL, Li T, Qiu XW, Wu J, Ling JQ, Sun ZH, et al. Correction of aberrant imprinting of IGF2 in human tumors by nuclear transfer induced epigenetic reprogramming. EMBO J 2006;25(22):5329 38.
[40] Li T, Hu JF, Qiu X, Ling J, Chen H, Wang S, et al. CTCF regulates allelic expression of Igf2 by orchestrating a promoter polycomb repressive complex 2 intrachromosomal loop. Mol Cell Biol 2008;28(20):6473 82.
[41] Skaar DA, Li Y, Bernal AJ, Hoyo C, Murphy SK, Jirtle RL. The human imprintome: regulatory mechanisms, methods of ascer tainment, and roles in disease susceptibility. ILAR J 2012;53(3 4):341 58.
[42] Polson ES, Lewis JL, Celik H, Mann VM, Stower MJ, Simms MS, et al. Monoallelic expression of TMPRSS2/ERG in prostate cancer stem cells. Nat Commun 2013;4:1623.
[43] Gimelbrant AA, Ensminger AW, Qi P, Zucker J, Chess A. Monoallelic expression and asynchronous replication of p120 catenin in mouse and human cells. J Biol Chem 2005;280(2):1354 9.
[44] Esumi S, Kakazu N, Taguchi Y, Hirayama T, Sasaki A, Hirabayashi T, et al. Monoallelic yet combinatorial expression of variable exons of the protocadherin alpha gene cluster in single neurons. Nat Genet 2005;37(2):171 6.
[45] Magklara A, Yen A, Colquitt BM, Clowney EJ, Allen W, Markenscoff Papadimitriou E, et al. An epigenetic signature for monoallelic olfactory receptor expression. Cell 2011;145(4):555 70.
[46] Roppolo D, Vollery S, Kan CD, Luscher C, Broillet MC, Rodriguez I. Gene cluster lock after pheromone receptor gene choice. EMBO J 2007;26(14):3423 30.
[47] Rhoades KL, Singh N, Simon I, Glidden B, Cedar H, Chess A. Allele specific expression patterns of interleukin 2 and Pax 5 revealed by a sensitive single cell RT PCR analysis. Curr Biol 2000;10(13):789 92.
[48] Lin M, Hrabovsky A, Pedrosa E, Wang T, Zheng D, Lachman HM. Allele biased expression in differentiating human neurons: implications for neuropsychiatric disorders. PLoS One 2012;7(8):e44017.
[49] Bonora E, Graziano C, Minopoli F, Bacchelli E, Magini P, Diquigiovanni C, et al. Maternally inherited genetic variants of CADPS2 are present in autism spectrum disorders and intellectual disability patients. EMBO Mol Med 2014;6 [Epub ahead of print].
[50] Kopsida E, Mikaelsson MA, Davies W. The role of imprinted genes in mediating susceptibility to neuropsychiatric disorders. Horm Behav 2011;59(3):375 82.
[51] Feinberg AP, Tycko B. The history of cancer epigenetics. Nat Rev Cancer 2004;4(2):143 53.
[52] Holm TM, Jackson Grusby L, Brambrink T, Yamada Y, Rideout 3rd WM, Jaenisch R. Global loss of imprinting leads to widespread tumorigenesis in adult mice. Cancer Cell 2005;8(4):275 85.
[53] Soejima H, Higashimoto K. Epigenetic and genetic alterations of the imprinting disorder Beckwith Wiedemann syndrome and related disorders. J Hum Genet 2013;58(7):402 9.
[54] Ogawa O, Eccles MR, Szeto J, McNoe LA, Yun K, Maw MA, et al. Relaxation of insulin like growth factor II gene imprinting implicated in Wilms' tumour. Nature 1993;362(6422):749 51.
[55] Rainier S, Johnson LA, Dobry CJ, Ping AJ, Grundy PE, Feinberg AP. Relaxation of imprinted genes in human cancer. Nature 1993;362(6422):747 9.
[56] Feinberg AP. Genomic imprinting and gene activation in cancer. Nat Genet 1993;4(2):110 3.
[57] Randhawa GS, Cui H, Barletta JA, Strichman Almashanu LZ, Talpaz M, Kantarjian H, et al. Loss of imprinting in disease progression in chronic myelogenous leukemia. Blood 1998;91(9):3144 7.
[58] Zhang L, Zhou W, Velculescu VE, Kern SE, Hruban RH, Hamilton SR, et al. Gene expression profiles in normal and cancer cells. Science 1997;276(5316):1268 72.
[59] Sohda T, Iwata K, Soejima H, Kamimura S, Shijo H, Yun K. In situ detection of insulin like growth factor II (IGF2) and H19 gene expression in hepatocellular carcinoma. J Hum Genet 1998;43(1):49 53.
[60] Takeda S, Kondo M, Kumada T, Koshikawa T, Ueda R, Nishio M, et al. Allelic expression imbalance of the insulin like growth factor 2 gene in hepatocellular carcinoma and underlying disease. Oncogene 1996;12(7):1589 92.
[61] Ulaner GA, Vu TH, Li T, Hu JF, Yao XM, Yang Y, et al. Loss of imprinting of *Igf2* and *H19* in osteosarcoma is accompanied by reciprocal methylation changes of a CTCF binding site. Hum Mol Genet 2003;12(5):535 49.
[62] Hofmann WK, Takeuchi S, Frantzen MA, Hoelzer D, Koeffler HP. Loss of genomic imprinting of insulin like growth factor 2 is strongly associated with cellular proliferation in normal hemato poietic cells. Exp Hematol 2002;30(4):318 23.
[63] Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. Sci ence 2013;339(6121):819 23.
[64] Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA guided human genome engineering via Cas9. Science 2013;339(6121):823 6.
[65] Barlow DP, Stoger R, Hermann BG, Saito K, Schweifer N. The mouse insulin like growth factor type 2 receptor is imprinted and closely linked to the *Tme* locus. Nature 1991;349:84 7.
[66] Barlow DP. Gametic imprinting in mammals. Science 1995;270(5242):1610 3.
[67] Barlow DP. Competition a common motif for the imprinting mechanism? EMBO J 1997;16(23):6899 905.
[68] Hu JF, Oruganti H, Vu TH, Hoffman AR. Tissue specific imprinting of the mouse insulin like growth factor II receptor gene correlates with differential allele specific DNA methylation. Mol Endocrinol 1998;12(2):220 32.
[69] Wutz A, Barlow DP. Imprinting of the mouse Igf2r gene depends on an intronic CpG island. Mol Cell Endocrinol 1998;140(1 2):9 14.
[70] Howard TK, Algar EM, Glatz JA, Reeve AE, Smith PJ. The insulin like growth factor 1 receptor gene is normally biallelically expressed in human juvenile tissue and tumours. Hum Mol Genet 1993;2(12):2089 92.
[71] Ogawa O, McNoe LA, Eccles MR, Morison IM, Reeve AE. Human insulin like growth factor type I and type II receptors are not imprinted. Hum Mol Genet 1993;2(12):2163 5.