UNIVERSIDADE FEDERAL DO PARANÁ

KELIN GONÇALVES DE OLIVEIRA

ANÁLISE DE INTERATOMAS CENTRADOS NO RECEPTOR FGFR2 – POTENCIAIS ALVOS TERAPÊUTICOS NO CÂNCER DE MAMA.

CURITIBA

2018

KELIN GONÇALVES DE OLIVEIRA

ANÁLISE DE INTERATOMAS CENTRADOS NO RECEPTOR FGFR2 – POTENCIAIS ALVOS TERAPÊUTICOS NO CÂNCER DE MAMA.

Dissertação apresentada ao curso de Pós-Graduação em Bioinformática, Área de concentração em Bioinformática, Setor de Educação Profissional e Tecnológica da Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestre em Bioinformática.

Orientador: Mauro Antônio Alves Castro

CURITIBA

2018

Catalogação na Publicação
Sistema de Bibliotecas UFPR
Biblioteca de Educação Profissional e Tecnológica

Oliveira, Kelin Gonçalves de
O48 Análise de interatomas centrados no receptor FGFR2 – potenciais alvos terapêuticos no câncer de mama. - Curitiba, 2018.
72 f.; il.: tab.

Orientador: Mauro Antônio Alves Castro

Dissertação (Mestrado) - Universidade Federal do Paraná, Setor de Educação Profissional e Tecnológica, Curso de Pós-Graduação em Bioinformática.
Inclui Bibliografia.

1. Câncer – Aspectos Genéticos. 2. Mamas - Câncer. 3. Receptor 2 do Fator de Crescimento Fibroblástico (FGFR2). 4. PBX1. 5. EPHA2. 6.NFIB. 7. YBX1. 8. ESR1. 9. Bioinformática. I. Castro, Mauro Antônio Alves. II. Título. III. Universidade Federal do Paraná.

CDD 616.99449

Elaboração: Rafaela Paula Schmitz – CRB-9/1882.

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA

Pós-Graduação em Bioinformática WWW.BIOINFO.UFPR.BR
E-mail: bioinfo@ufpr.br Tel: 41 33614906

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em BIOINFORMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da dissertação de Mestrado de **KELIN GONÇALVES DE OLIVEIRA** intitulada: **"Análise de Interatomas Centrados no Receptor FGFR2 - Potenciais Alvos Terapêuticos no Câncer de Mama"**, após terem inquirido a aluna e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 20 de março de 2018.

Dr. Mauro Antonio Alves Castro
Presidente
Programa de Pós-graduação em Bioinformática – UFPR

Drª. Daniela Fiori Gradia
Avaliadora Externa
Departamento de Genética - UFPR

Dr. Dieval Guizelini
Avaliador Interno
Programa de Pós-graduação em Bioinformática - UFPR

Aos meus pais, Roberto e Rosânia, com todo o meu amor e gratidão.
De todo o amor do mundo, o de vocês é sempre o mais especial.

AGRADECIMENTOS

Ao meu orientador Prof. Dr. Mauro A. A. Castro, com muito carinho, por me ensinar tanto, pela paciência, pelo companheirismo e por ser um exemplo impecável de profissionalismo e humanidade a ser seguido.

À minha família por me dar todo o suporte necessário, todo o amor e por aguentar a distância intercalada com visitas apressadas. Amor e gratidão eternos é o que sinto por vocês todos os dias.

Ao coordenador do Programa de Pós Graduação em Bioinformática, Prof. Dr. Roberto Raittz, e aos professores por todo o trabalho de lapidação que fazem com cada aluno que ingressa nesta jornada; e à melhor secretária que existe, Suzana, por ser sempre tão amiga e solícita.

À todas as amizades incríveis, sejam as antigas que transcendem os anos ou aquelas que fiz nestes últimos dois: Jéssica, Franciele, Camila, Caroline, Josué, Clarice, Aryel, Carolina, Bruno, Alan, Mattheus, William e mais todos àqueles que fizeram deste período uma experiência tão marcante. Agradeço em especial àqueles que foram meus pilares nos momentos mais intensos: ao Vinícius por ter sido um verdadeiro tutor nos meus tempos de novata e pelas músicas ruins; à Sheyla pelas ótimas conversas que muito me guiaram e por todo carinho, ao Diônata pela amizade de ouro e por ser tão guerreiro (que toda luta se reverta numa fração da felicidade que você traz à turma) e ao Daniel por todo companheirismo e imensa ajuda com as aventuras futuras. Muito obrigada por aturarem a minha esquisitice diária. O aprendizado que tive com vocês é imensurável.

A natureza é sempre mais sutil, mais intrigante, mais elegante do que somos capazes de imaginar.

- Carl Sagan, O mundo assombrado pelos demônios

RESUMO

Em doenças heterogêneas, como o câncer de mama, muitos são os fatores que adicionam pontos de fragilidade que contribuem para o início e a progressão do tumor. Segundo estudos de associação de genomas (GWAS), uma parcela de tumores mamários apresenta polimorfismos de nucleotídeo único (SNPs) presentes no segundo íntron do gene do receptor 2 do fator de crescimento fibroblástico (FGFR2). Estes SNPs atuam como importantes fatores de risco para o câncer de mama e estão fortemente associados com a regulação do crescimento celular tumoral mediado por estrogênio via genes FOXA1, NFIB e YBX1. Muitos esforços são concentrados em compreender tal regulação, no entanto muitos transdutores ainda estão para serem elucidados. Neste âmbito, este trabalho visa esclarecer por meio de estudos *in silico* da biologia sistêmica humana inferida a partir de dados de transcriptoma, quais são outras importantes proteínas envolvidas na via associativa entre FGFR2 e o gene receptor de estrogênio (ESR1) que afetam a resposta terapêutica. Tal análise ambiciona encontrar mais pistas sobre o funcionamento de vias regulatórias específicas para tecido mamário tumoral, partindo de análises bioinformáticas na construção de redes biológicas e análises modulatórias, amparados por abordagens experimentais de superexpressão, *knockdown* e PDX. As análises permitiram selecionar transdutores da via de sinalização do FGFR2, tais como: EPHA2, PBX1, NFIB e YBX1. Estes genes se mostraram ser influentes nas respostas à diversos tratamentos anti tumorais, sendo capazes até de restaurar responsividade a tamoxifeno em certos perfis transcricionais.

Palavras-chave: Câncer de Mama, FGFR2, PBX1, EPHA2, NFIB, YBX1, ESR1

ABSTRACT

For heterogeneous diseases, such as breast cancer, many are the factors that add fragility points that contribute to the onset and progression of the tumour. According to genome wide association studies (GWAS), a set of mammary tumours present single nucleotide polymorphisms (SNPs) within the second intron of the fibroblast growth factor receptor 2 (FGFR2) gene. These SNPs act as important risk factors to breast cancer are strongly associated with growth regulation mediated by estrogen via FOXA1, NFIB, and YBX1 genes. Many efforts were put on trying to understand this regulation, but many transducers are yet to be elucidated. In this context, this current work aims to clarify, via *in silico* studies of human systemic biology inferred by transcriptomic data, which are other important proteins involved within that associative pathway between FGFR2 and the estrogen receptor gene (ESR1) capable of affecting therapeutic response. This analysis aims to find more clues on how tumorous breast-tissue specific regulatory pathways are given based on bioinformatical approaches to build biological networks and modulatory analysis, backed up by experimental data of overexpression, knockdown and PDX. The analysis allowed to select transducers from FGFR2 signaling pathway, such as: EPHA2, PBX1, NFIB and YBX1. These genes are shown to influence response to diverse anti-tumoural treatments, being capable to restore tamoxifen responsiveness for certain transcriptional profiles.

Keywords: Breast Cancer, FGFR2, PBX1, EPHA2, NFIB, YBX1, ESR1

LISTA DE FIGURAS

SUMÁRIO

APRESENTAÇÃO

Esta dissertação está estruturada em três partes. A Parte I apresenta uma introdução geral que dará embasamento teórico ao leitor quanto a anatomia da mama, a biologia do câncer de mama e o funcionamento, estrutura e importância dos receptores tirosina quinase dos fatores de crescimento fibroblásticos, que são contribuintes protagonistas na doença em questão. Esta parte também descreve os objetivos gerais e específicos deste estudo.

A parte II apresenta-se em dois capítulos que descrevem em detalhes os estudos feitos pela autora com o intuito de compreender pontos da complexa biologia tumoral, bem como apresenta o "loop"regulatório de estrogênio intermediado por FGFR2, e demais importantes transdutores envolvidos. Esta parte contém primeiramente o manuscrito principal da autora, a ser submetido, que sugere a inserção de novos genes à cascata regulatória descrita previamente. Posteriormente ao manuscrito, é apresentado um artigo já publicado em parceria com colaboradores do Instituto de Pesquisa em Câncer do Reino Unido, Cancer Research UK, o qual apresenta o envolvimento dos genes NFIB e YBX1 na regulação de estrogênio mediada pelo receptor em questão.

Por fim, a parte III apresenta o fechamento do trabalho, contendo discussão geral e conclusões com base nos resultados da parte II.

PARTE I

Introducção Geral

1 INTRODUÇÃO

1.1 CÂNCER DE MAMA

Derivados das glândulas sudoríparas, os arranjos primários da mama podem ser identificados a partir da sétima semana do desenvolvimento embrionário, na forma de um espessamento ectodérmico primitivo que, futuramente, proliferará cordões celulares na mesoderme para que se formem os canais lactíferos. Após desenvolvimento completo, a mama humana se consolida como um órgão glandular par, localizado anteriormente aos músculos peitoral e serrátil, lateral às extremidades do osso esterno. É composto internamente por tecidos glandulares, adiposos e fibrosos que se combinam para formar estruturas produtoras de leite bastante desenvolvidas em pessoas do sexo feminino e atrofiadas no sexo masculino (BERNARDES, 2010).

As estruturas internas são diversas e se dividem, de maneira simplificada, em lóbulos (unidade funcional produtora de leite), lobos (agrupamento de lóbulos), ductos (canais condutores do leite), linfonodos (tecido linfático responsável pela drenagem de líquidos excessivos), papila (protuberância resultante de junção dos ductos, antigo mamilo), ligamentos suspensores, além de redes vasculares e de inervação (JESINGER, 2014; DRAGHI et al., 2011; BERNARDES, 2010). Em uma mesma pessoa, o tamanho de cada mama e a distribuição estrutural podem variar consideravelmente e, a nível populacional, há grande diversificação destas características de acordo com o sexo, peso, etnia e outros fatores de influência (RAMSAY et al., 2005).

A mama feminina se destoa da masculina em muitos aspectos. Em relação à malefícios que a acometem, os seios são muito mais susceptíveis à formações malignas já que são estruturas de alta atividade biológica e influência hormonal. A presença de alterações nos níveis de esteróides sexuais, ainda que ocorram devido a eventos naturais como menopausa e gravidez, podem contribuir para maiores chances de desenvolvimento de uma das doenças de maior incidência na população feminina: o câncer de mama (CHEN, 2008).

O câncer é, de maneira geral, definido como um conjunto de células com acúmulo de

alterações genéticas e epigenéticas, capazes de escaparem do sistema endógeno de checagem e de proliferarem sem o controle sobre o seu ciclo celular (MARSICANO et al., 2014). Tal proliferação desregulada permite que uma massa amorfa se forme e cresça indevidamente no tecido de origem, atrapalhando as funções fisiológicas normais no local e podendo causar comprometimento metabólico, invalidez do órgão, início de tumores em outros tecidos e, em casos avançados, morte do paciente.

As células que constituem o tumor apresentam estratégias moleculares exclusivas à exemplo da resistência aos sinais apoptóticos e supressores de crescimento, que seriam naturalmente ativados no tecido saudável quando o organismo percebe que há anormalidades no microambiente celular. Adicionalmente, tais células também possuem potencial invasivo (metástase), sinalização proliferativa constitutiva e são capazes de ativarem processos extracelulares, como a angiogênese, para manter a nutrição e oxigenação, a fim de promover sua própria sobrevivência (HEJMADI, 2010; HANAHAN; WEINBERG, 2011). Quando a formação tumoral é iniciada em alguma estrutura mamária, dá-se origem ao câncer de mama.

Ambos os sexos podem desenvolver neoplasias no tecido mamário. No entanto, a proporção entre homens e mulheres afetados varia grandiosamente, sendo 1% da totalidade dos casos presentes em homens e o segundo mais ocorrente em mulheres, responsável por cerca de 59 mil novos casos e aproximadamente 14 mil mortes no Brasil no período de um ano (INCA, 2018). As estratégias para melhor lidar com esta endemia evolvem diagnóstico precoce e tratamentos eficientes e, em relação aos tratamentos adequados, a correta classificação do tipo de tumor de mama durante a triagem clínica é de suma importância por permitir adequar a terapia de acordo com a agressividade da patologia e oferecer melhor qualidade de vida à paciente (RAKHA et al., 2010; BØRRESEN-DALE et al., 2010).

Tumores de mama estão sujeitos a serem subtipados em diversos níveis, podendo ser de acordo com o tecido no qual se desenvolvem (tipo histopatológico) (WEIGELT et al., 2010), o grau de diferenciação e de atividade proliferativa (grau e estadiamento, respectivamente) (ELSTON; ELLIS, 1991) ou por avaliação do perfil molecular intrínseco. Nos anos 90, a classificação de maior peso era o subtipo histológico, dividindo os tumores de mama em (CRUK, 2018; NBCF, 2018):

- Carcinoma
  - Ductal (in situ ou invasivo)
  - Lobular (in situ ou invasivo)
- Tumor Inflamatório

- Doença de Paget
- Tumor Filóide
- Angiosarcoma

Os subtipos emergentes mais comuns segundo o tipo tecidual são os carcinomas ductais (tanto *in situ*, quanto invasivo) e carcinomas lobulares invasivos (ACS, 2018). Alguns tipos tumorais permanecem sub-representados nos estudos gerados, por serem de difícil coleta e de menor ocorrência. Por outro lado, há grupos de grande representação, como o carcinoma ductal sem tipo especial (IDC-NST), que compreende tumores de mama não classificados com maior especificidade, não se encaixando nas características das categorias existentes de tipagem histológica (WEIGELT et al., 2010). Apesar destas divisões serem feitas com base nos tipos específicos de células afetadas pelo câncer, em algumas vezes, um único tumor pode se apresentar como uma combinação de diferentes características clínicas. Portanto, a fim de complementar a caracterização anatomo-clínica e especificar melhor o tumor do paciente para melhor encaminhamento terapêutico, outros fatores como grau e estadiamento tumorais são considerados.

O sistema de Gradeamento de Nottingham (NGS) e o estadiamento tumoral são as categorizações adicionais à classificação histopatológica recomendados por grandes organizações de saúde, pois melhoram significantemente o direcionamento terapêutico, sendo combinados entre si para formar o Índice Prognóstico de Nottingham (NPI). O NPI atribui pesos para o grau de diferenciação celular, presença e nível de invasão aos linfonodos adjacentes e tamanho da massa (RAKHA et al., 2010). Segundo esta metodologia de agrupamento, o grau de diferenciação das células tumorais varia entre os pesos 1 a 3, considerando células bem ou pouco diferenciadas, respectivamente. Tumores grau um possuem melhor prognóstico, pois as células são semelhantes às do tecido saudável, ao passo que, tumores compostos por células menos diferenciadas tendem à um pior prognóstico por indicarem elevada taxa de divisão celular, tendo peso dois para indiferenciação média ou três caso seja severa.

Nos anos 2000, a revolução da subtipagem por meio da classificação à partir da expressão gênica foi iniciada, quando Perou e colaboradores (PEROU et al., 2000) geraram uma assinatura inicial de 496 genes considerados intrínsecos à tumores de mama avançados, que apresentavam pouca variância entre as amostras pertencentes ao grupo, mas grande variância em relação à outros tipos de tumores. A assinatura se mostrou capaz de separar os subtipos com base nos receptores hormonais e estratificar estes em mais outros 4 subtipos: Luminal A, Luminal B, HER2-Enriched e Basal-Like (PEROU; BORRESEN-DALE, 2011). A quantidade de genes intrínsecos foi reduzida e a categorização aperfeiçoada pelo método de Predição de

Amostra Única (SSP) desenvolvida por Hu e colaboradores (HU et al., 2006) e aplicada por Parker (PARKER et al., 2009), resultando no PAM50: 50 genes que classificam os 4 subtipos tumorais e um adicional, denominado normal-like, que compartilha características semelhantes ao tecido mamário saudável.

Atualmente, a combinação desta identificação de padrões moleculares intrínsecos juntamente com as demais metodologias citadas anteriormente é a melhor abordagem disponível para identificar o tipo do tumor, prever o desenvolvimento deste e estabelecer o tratamento mais adequado, por permitir detectar características gênicas que funcionam de marcadores dos processos biológicos da patologia, além de combiná-las com características teciduais (MATSUMOTO et al., 2016). Os marcadores moleculares primeiramente analisados para a categorização de um câncer de mama são: o Receptor de Estrogênio (ER) e o Receptor de Progesterona (PR). Estes genes podem ser utilizados, por exemplo, para indicar se o tumor é de alta ou baixa classe, ainda que tais receptores não apresentem especificidade ao tecido mamário (também podem ser detectados em carcinomas de endométrio e útero, por exemplo) (GOWN et al., 2016). Cânceres de baixa classe apresentam, em maioria, um número menor de alterações genéticas e possuem ambos ou um dos receptores na superfície de suas células, a medida que os de alta classe são negativos para ER e PR e apresentam uma quantidade maior de mutações nos oncogenes (GIANCOTTI, 2006). Apesar de pouco específicos, o receptor de estrogênio é expresso em 75% dos casos de tumores de mama invasivos, enquanto que 55% dos casos apresentam o receptor de progesterona. A presença de um ou ambos receptores indica maior probabilidade de sucesso para aqueles pacientes que optarem pelo uso de terapias que alterem o padrão homonal, já que estes tumores dependem do hormônio como ligante para progredirem (GIANCOTTI, 2006).

Os métodos de classificações de doenças heterogêneas, como o câncer, estarão sempre em processo de melhoria, uma vez que novos conhecimentos gerados a respeito dos fatores influenciadores adicionam esclarecimento quanto à biologia da enfermidade, permitindo a inclusão de novos marcadores de probabilidade de desenvolvimento patológico, progressão e predição clínica. Neste âmbito, diversos genes já foram inferidos como fatores de risco para o desenvolvimento de tumores mamários. Genes como o de Susceptibilidade ao Câncer de Mama 1 e 2 (BRCA1 e BRCA2) bem como o gene supressor tumoral p53 são biomarcadores clássicos na predição de risco. Em alguns pacientes, uma mutação não incapacitante de BRCA1 pode levar a superexpressão de p53 (PENG et al., 2016), enquanto que em certos pacientes diagnosticados, a mutação em BRCA1 causa uma perda completa da expressão, podendo significar uma predisposição maior a metástase para os linfonodos (KIM et al., 2016). Um prognóstico ruim também é identificado pela presença do antígeno marcador Ki-67, uma

proteína nuclear com função de promover a proliferação celular. Pacientes Ki-67 positivos possuem uma expectativa de vida menor, bem como menores chances de cura, caso venham a desenvolver um tumor (ABUBAKAR et al., 2016).

Tumores mamários positivos para o receptor de estrogênio (ER+) têm sido frequentemente relacionados com mutações no locus do receptor 2 do fator de crescimento fibroblástico, sendo descrito em diversos estudos como mais um fator de risco importante a ser considerado no desenvolvimento desta doença. Apesar da função específica da proteína ser bastante estudada, a ação global de interação e mecanismos no qual este receptor participa para conseguir conferir risco tumoral ainda está para ser melhor esclarecido, assim como diversos outros que permanecem incógnitas ao entendimento humano.

1.2 RECEPTORES DE FATOR DE CRESCIMENTO FIBROBLÁSTICO

Os fatores de crescimento fibroblástico (FGF) compreendem uma família de 18 polipeptídeos diferentes (FGF 1 ao 10 e FGF 16 ao 23) agrupados em 6 famílias, que se ligam aos seus respectivos Receptores de Fator de Crescimento Fibroblástico (FGFR), cada um com especificidade e afinidade diferentes, estimulando vias proliferativas, de sobrevivência e homeostásicas (ORNITZ; ITOH, 2015). Os receptores FGFRs formam uma família de 4 proteínas tirosina-quinases (FGFR1 a FGFR4) de aproximadamente 800 aminoácidos, caracterizadas pela presença de dois a três domínios extracelulares de loops semelhantes à imunoglobulinas (Ig-like) denominadas I ou D1, II ou D2 e III ou D3; uma caixa acídica rica em serina e localizada entre os dois primeiros loops; uma região transmembrana e dois domínios citoplasmáticos de atividade fosforilativa (LU et al., 2003) (**Figura 1 e Figura 2**).

O receptor FGFR2 é também conhecido como Bek, Cek ou Kgfr e é codificado pelo gene de mesmo nome, localizado no locus 10q26. Este apresenta tamanho aproximado de 92118Da e, assim como os demais membros da família, é indispensável a célula por regular processos celulares envolvidos no controle proliferativo, diferenciação, morte, migração celular e angiogênese (TURNER; GROSE, 2010; CUI et al., 2016). Mutações no FGFR2 podem alterar o padrão de sinalização molecular e gerar uma variedade de anomalias sindrômicas, como a Síndrome de Apert, Crouzon, Saethe-Chotzen, Pfeiffer, Jackson-Weiss e LADD.

Para executar a ativação da sinalização celular, é necessário que o receptor esteja na sua forma dimerizada (**Figura 2b**), o que pode ocorrer tanto na presença quanto na ausência do ligante, fator que afetará a intensidade fosforilativa (SARABIPOUR; HRISTOVA, 2016). No caso do ligante estar presente, o FGF interage com moduladores proteoglicanos

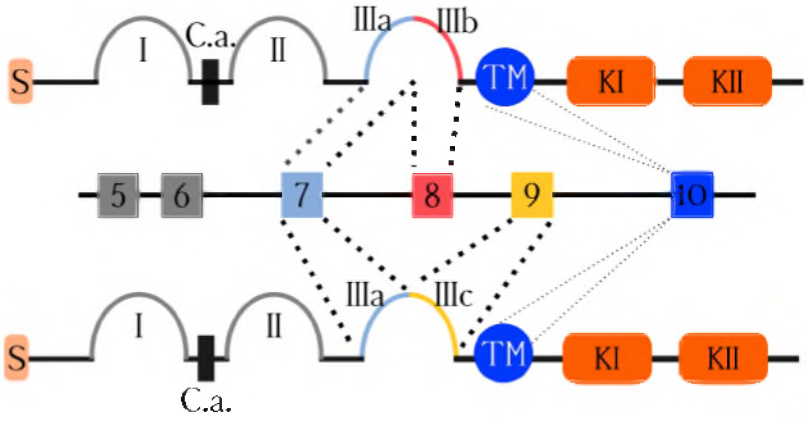


Figura 1: Características estruturais dos FGFRs. S-Peptídeo sinal. Éxons 5 e 6 são responsáveis pelo loop I e II, que são intercalados pela estrutura da caixa acídica (C.a.). De acordo com o splicing feito entre os exons 7, 8 e 9, a estrutura do loop D3 será definida. Éxon 10 codifica região transmembrana (TM) que separará a região extracelular dos domínios kinases (KI e KII) intracelulares. A autora, baseado no esquema apresentado em (LIN; WANG, 2010)

de sulfato-heparina a fim de ativar a porção extracelular do FGFR, dimerizando-o, gerando mudanças conformacionais no receptor e induzindo cascatas sinalizadoras por intermédio de transfosforilação de resíduos de tirosina presente em moléculas como DAG, IP3, PI3K, RAS e outras, responsáveis pela iniciação de migração celular, invasão e diferenciação (YEUNG et al., 2008; KATOH, 2016; ZHANG et al., 2017). O receptor também atua em proteínas como Fosfolipase C Gama 1 , FRS2 e P21-proteína quinase ativada, além de exercer auto-fosforilação para ativar-se (LU et al., 2003; ESWARAKUMAR et al., 2005; LUO et al., 2009).

Os receptores de FGF se apresentam em diversas isoformas consequentes de recombinações do mRNA endógeno, gerados através de splicing alternativo (TIONG et al., 2013). A recombinação no loop D3, um fator essencial para determinar a especificidade do ligante ao FGFR, pode gerar uma de duas formas, determinando qual substrato irá se ligar de acordo com o tipo celular no qual está presente e, consequentemente, definindo qual via este irá ativar. Para o receptor FGFR2 (**Figura 1**), que será o foco neste trabalho, a sua forma 2b possui os exons 7 e 8 como formadores do loop D3 (IIIa + IIIb) e apresenta especificidade pelos substratos FGF 1, 3, 7, 10 e 22 em células epiteliais, além de alta afinidade pelo fator de crescimento de keratinócitos (KGF); já a forma FGFR2c, que apresenta os exons 7 e 9 como estrutura D3 (IIIa + IIIc), possui afinidade por FGF 1, 2, 4, 6, 9, 17, 18 e 19 como ligantes, majoritariamente em células mesenquimais. Já a recombinação da região carboxi-terminal de FGFR2 resulta em formas completas ou encurtadas da proteína (TANNHEIMER et al., 2000).

Quatro cascatas de segundos mensageiros são majoritariamente ativadas tanto pelo FGFR2 quanto pelos demais receptores da família: RAS-MAPK, fosfolipaseC-gama (PLC),

PI3K-AKT e STATS, sendo a primeira a mais comumente estimulada e a qual regula a ativação de fatores de transcrição ETS. A cascata AKT, por sua vez, atua inibitoriamente no fator de transcrição FOXO1 (um efetor pró-apoptótico) e no gene TSC2, ambos a fim de promover proliferação celular. Já a ativação da via PLC causa aumento nos níveis de cálcio intracelular e gera ativação da proteína quinase C, enquanto que a via STAT atua como contrabalanço das vias proliferativas ativadas por FGFRs, uma vez que é cascata supressora de proliferação (ORNITZ; ITOH, 2015).

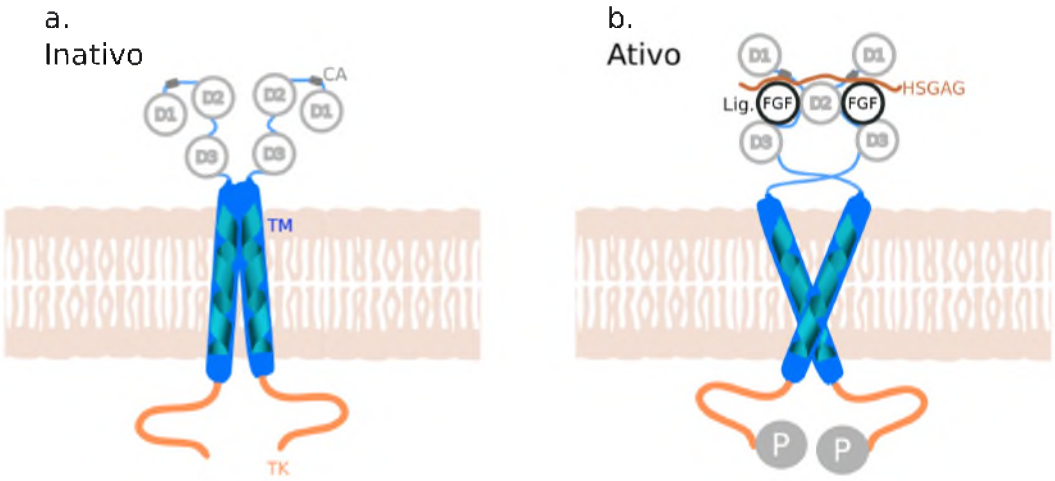


Figura 2: Esquema de mudança conformacional de FGFRs mediante ligação de FGF. D1-Domínio IgG-like I; D2-Domínio IgG-like II; D3-Domínio IgG-like III; CA-Caixa acídica; TM-Região transmembrana; TK-Domínios tirosina quinase; HSGAG-Glicosaminoglicano sulfato heparina; Lig-Ligante FGF; P-Proteína Fosforilada. A autora, baseado no esquema apresentado em (BOCHAROV et al., 2013)

O FGFR2 já foi relatado por inúmeros estudos como amplificado em diversos tipos de tumores (próstata, gástrico e mamário), além de se apresentar translocado em colangiosarcomas (ORNITZ; ITOH, 2015). A relação íntima entre alterações neste receptor e maior risco de desenvolvimento tumoral foi descrito em estudos de associação de genomas (GWAS), os quais apresentam diversos variantes intrônicos no gene do FGFR2 associados ao aumento no risco de desenvolvimento de câncer de mama. Para parte destes polimorfismos, o risco é conferido para a doença do tipo ER+. Apesar de estudos funcionais anteriores terem relacionados que tais variantes afetam os níveis de expressão do gene, pouco conseguiu ser concluído em relação a estas afirmações quando estudos de correlação genotípica foram realizados, uma vez que a expressão não apareceu alterada. Pesquisas posteriores puderam, então, correlacionar os variantes com vias alteradas de sinalização mediada por FGFR2 resultantes em atividade anormal da rede transcricional associada à regulação de estrogênio (FLETCHER et al., 2013).

A partir da criação de redes regulatórias transcricionais incrementadas com dados

de GWAS, foi possível observar que não só o FGFR2, mas também cerca de 36 fatores de transcrição se apresentavam enriquecidos com variantes de susceptibilidade e tinham suas vias também clusterizadas em regulação de crescimento celular mediada pelo receptor de estrogênio (CASTRO et al., 2015). A presença de tantas vias culminando para um mesmo mecanismo regulatório, o papel conhecido do receptor 2 do fator de crescimento fibroblástico e a relação deste com o câncer de mama fazem do FGFR2 e das vias de estrogênio pontos importantes a serem estudados. Estes fatores sugerem o estabelecimento de caminhos conectados para a regulação do crescimento tumoral e podem apresentar alvos interessantes para melhorar tanto subtipagem quanto tratamento se pensarmos que, mediante o perfil de expressão de um paciente, uma suposta regulação negativa do receptor de estrogênio que confere um tumor mais agressivo possa ser restabelecido por alteração de transdutores específicos, permitindo alteração do quadro tumoral.

1.3 JUSTIFICATIVA

O locus do FGFR2 foi diversas vezes relacionado, através de estudos experimentais e de associação de genomas completos (GWAS), como um importante fator de risco para o desenvolvimento do câncer de mama (CAMPBEL et al., 2016). Alguns pesquisadores atribuem este risco a um silenciamento genético, que resulta em ativação de vias de crescimento celular dependentes de estrogênio e, portanto, este locus confere o risco de desenvolvimento de tumores de mama ER+ (ZHU et al., 2010). Outras pesquisas, por outro lado, propõem que há, na verdade, amplificação do receptor 2 do fator de crescimento fibroblástico, o que aumenta a atividade de cascatas proliferativas e, por isso, uma menor expressão deste culminaria em menores números de células propagadoras de tumor (ORNITZ; ITOH, 2015; KIM et al., 2013). Por outro lado, foi observado que os níveis de expressão do FGFR2 podem não se apresentar alterados e, portanto, o efeito seria por meio de alterações dos processos regulatórios normais como afinidade e interações proteicas.

Em função destas controvérsias, é razoável assumir que os mecanismos de risco e propagação de sinais tumorais dependentes de FGFR2 ainda não estão devidamente esclarecidos. Sendo assim, estudos exploratórios de biologia de sistemas são extremamente necessários para auxiliar no entendimento dos processos integrativos pelos quais este receptor contribui para a doença, sendo abordagens poderosas para auxiliar desde o conhecimento básico ao entendimento clínico. Um melhor entendimento de como as cascatas tumorais se dão pode auxiliar na tomada de decisões terapêuticas que tornem o tratamento mais efetivo.

1.4 OBJETIVO GERAL

Este trabalho tem por objetivo geral realizar buscas por potenciais alvos terapêuticos dentro da nuvem de interação do receptor 2 do fator de crescimento fibroblástico, bem como visa auxiliar no entendimento da contribuição deste elemento na biologia do câncer de mama.

1.5 OBJETIVOS ESPECÍFICOS

- Construir rede de interação proteína-proteína referente ao FGFR2 em tecidos saudáveis
- Avaliar quais proteínas podem ter papel modulatório de fatores de transcrição relacionados previamente com risco ao câncer de mama
- Integrar estudos experimentais de perturbação ao receptor a fim de analisar quais fatores de interação são consistentes e, portanto, possíveis candidatos à participantes da cascata tumoral.
- Avaliar resposta terapêutica em coorte de PDX estratificada por nível de expressão dos genes candidatos.
- Avaliar se genes candidatos estão presentes na via regulatória de estrogênio
- Propor candidatos a alvos terapêuticos

PARTE II

Análise do Interatoma de FGFR2

2 MANUSCRITO

O atual capítulo apresenta o manuscrito que descreve o estudo exploratório entitulado "EPHA2 and PBX1 mediate FGFR2 signalling associated with breast cancer risk regulators and modulate tamoxifen responsiveness"a ser submetido. Este manuscrito será submetido para publicação no primeiro semestre de 2018 e constitui o principal estudo desta dissertação, descrevendo os genes PBX1 e EPHA2 como transdutores da via de sinalização do FGFR2.

Bioinformatic Article
doi.XXXXXX/bioinf/xxxxxx
Advance Access Publication Date: day Month Year
Applications Note

Original Research

EPHA2 and PBX1 mediate FGFR2 signalling associated with breast cancer risk regulators and modulate tamoxifen responsiveness.

Kelin G. de Oliveira [1,*], Sheyla Trefflich [2], Clarice S. Groeneveld [1] and Mauro A. A. Castro [1,*]

[1]Bioinformatics and Systems Biology Group, Federal University of Paraná, Curitiba, 12345678, Brazil and
[2]Bioinformatics, Federal University of Minas Gerais, Belo Horizonte, 12345678, Brazil.

*To whom correspondence should be addressed.

Associate Editor: XXXXXXX

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Abstract

Motivation: Sporadic breast cancer has a polygenic profile, in which many genes act together with small contributions to give rise to the disease phenotype. Several genome-wide association studies have identified over 150 loci associated with breast cancer risk, being eight of those found within the fibroblast growth factor receptor 2 (FGFR2) locus. The presence of the rs2981578 SNP inside FGFR2 gene is believed to increase significantly the risk for estrogen receptor-positive (ER+) breast cancer. It has been recently demonstrated that FGFR2 has an important role on estrogen regulation. However, the extent of the impact of the receptor's activity in tumourigenic pathways and how transducers lead to such ER regulation are still poorly known. We therefore performed an exploratory systems biology approach to scrutinize the interaction between FGFR2 and other genes that might have influence onto the breast cancer risk regulators, aiming to find new core players in the FGFR2 signalling cascade. This analysis makes use of FGFR2-centered approach to explore candidate proteins that help layer understanding about this disease and searches for possible therapeutic targets.

Results: Integrative analysis of experimental data showed that many of the FGFR2-interacting proteins within breast tissue act as modulators of transcriptional regulators previously described as associated with increased breast cancer risk. From the 700+ FGFR2-interacting proteins we have retrieved from several knowledge databases, 434 are mapped to breast tissue, and 52 showed consistent response to FGFR2-perturbation experiments using a panel of breast cancer cell lines stimulated by different FGFR2 ligands. Also, all these 52 FGFR2-mediator candidates presented modulatory activity on transcription factors associated with breast cancer risk, of which 8 were seen to respond promptly to siRNA against FGFR2 mRNA, strongly indicating that they might have high involvement with FGFR2-dependent core regulatory pathways. Patient derived xenografts data analysis showed that PBX1 and EPHA2 genes impact the effectiveness of FGFR2 inhibitory drugs for cancer treatment and might be tagged as possible therapeutic targets. It was seen that both genes are involved in estrogen regulation mediated by FGFR2 activity.

Contact: kelin.g.oliveira@gmail.com

Supplementary information: Supplementary data are available at *XXXX* online.

1 Introduction

Breast cancer (BC) is a heterogeneous disease of high mortality and with multiple triggers leading to the same outcome (World Health Organisation (WHO),

2018). The incidence of approximately 59 new cases per 100.000 women categorizes breast cancer as a major public health problem (INCA, 2018). Among the several triggers that act in the development of this condition, some factors play an active role in the establishment or progression of the breast tumour, such as alcohol consumption, hormonal dysregulation or hormonal therapies, exposure to radiation, and other causes of endocrine, genetic and environmental origins (Nickels *et al.*, 2013).

Genetic history might carry risk variables that partially explain breast cancer progression by intensifying genomic instabilities, deepening the deregulation of several pathways and impacting treatment type and intensity. The presence of mutations within key genes for the development and maintenance of the breast tissue, such as BRCA1, BRCA2 and HER2, increases risk significantly, but they are believed to account only for up to seven percent of total breast cancer incidence (Campbel *et al.*, 2016). The risk of breast cancer development is estimated to be doubled for groups of people with family predisposition and to increase even more for patients who have more than two first-degree (mother or sister) women with the pathology (Brewer *et al.*, 2017).

The complexity of breast cancer follows a polygenic profile, in which genetic variations linked to elevated risk are more likely to be spread across the genome, and combinations of common polymorphisms of small effects contribute with the trait (Boyle *et al.*, 2017). Among several molecular factors described in the literature as strong contributors to tumour susceptibility so far, the Fibroblast Growth Factor Receptor 2 gene (FGFR-2), located in chromosome 10, has been reported altered in different types of diseases by several studies, including those of breast cancer. In general, high FGFR2 expression has been linked to poorer prognosis and lower survival rates for breast cancer patients (Sun *et al.*, 2012), however, the type of impairment to the receptor and the developmental stage within the tissue will define its contribution to the phenotype (Campbel *et al.*, 2016).

The unaltered form of the Fibroblast Growth Factor Receptor 2 actively participates in cellular cycle modulation (e.g. cellular division, growth, differentiation and apoptosis) and is a pro-proliferative gene, important for bone and mammary development, as well as tumourigenesis (Zhang *et al.*, 2017). Its encoded protein is a transmembrane tyrosine kinase receptor highly sensitive to extracellular signalling that triggers response via activation of multiple pro-proliferative pathways such as RAS-MAPK, PLC-Y and JAK-STATS (Etokebe *et al.*, 2009). In its active state, the phosphorylated tyrosine residues of the FGF receptor can function as docking regions for adapter proteins and enhance signalling through cascades (Turner and Grose, 2010).

In breast cancer, leading variations inside intron 2 of the FGFR2 gene were identified by multiple genome-wide association studies (GWAS) (Easton *et al.*, 2007; Hunter *et al.*, 2007) and these are believed to contribute with up to 16% of breast cancer incidence, being a significant player in susceptibility (Campbel *et al.*, 2016). The variants are eight single nucleotide polymorphisms (rs35054928, rs2981578, rs2912778, rs2912781, rs35393331, rsl0736303, rs7895676, rs33971856) close to promoter region, which suggests effect through impaired binding of interacting regulatory elements (Zhang *et al.*, 2017). Some of these risk SNPs inside the FGFR2 gene, such as the rs2981578, according to Robbez-Masson *et al.* (2013) and Cui *et al.* (2016), are strongly associated with increased risk for ER+ breast tumour development. However, for the same variant, Zhou *et al.* (2012) reported a decrease in BC risk, which shows a divergence around SNP effect (Zhang *et al.*, 2017).

When assessing shared pathways associated with risk, in order to clarify how these multiple variants combine, Castro *et al.* (2015) encountered a total of 36 overlapping transcription factors, along with their putative target genes - referred to as regulon - enriched with breast cancer risk loci. Most regulons clustered within two opposing groups, indicating that they share regulatory mechanisms. For one of these clusters, the participating TFs are important for FGFR2 and estrogen signalling and their respective induced targets are highly expressed in estrogen-receptor (ER) positive tumour subtypes (such as luminal A and B subtypes, for example). The stronger members of this group comprises YPEL3, BRD8, AFF3, RARA, SNAPC2, ZNF587, AR, ARNT2, ZNF552, MYB, GATA3, ESR1, FOXA1, SPDEF, XBP1 and MZF1. The other cluster of regulons holds regulators such as TBX19, NFIB, TRIM29, SOX10, CEBPB, CBFB and YBX1 that have their overexpressed positive targets associated with more basal-like, therapeutically challenging ER negative breast tumours (Castro *et al.*, 2015).

Even though it is known that FGFR2 is a key factor for estrogen regulation in BC, the full cloud of interactions, the deep effects of FGFR2 cascade transducers involved in tumour progression, estrogen regulation and drug response are yet not completely clarified.

Here we have applied a systems biology approach, backed up by analysis of multiple experimental data, to explore the functional consequences of FGFR2 activity, aiming to identify FGFR2-mediator candidates associated with breast cancer risk. The combination of this method plus regulatory approaches and patient datasets shed light into the breast cancer circuitry and pointed to possible new therapeutic targets.

2 Materials and methods

2.1 Data mining of FGFR2 PPI Network

Protein-protein interaction (PPI) information between FGFR2 and other proteins was retrieved from public databases such as APID, BioGrid, DIP, STRING, HPRD, IntAct, Mint, I2D, Spectra, GIANT, Mentha, HumanBase and IID (list of all retrieved sources is available at Table 1 and Supplementary Material). Two types of data were retrieved from these sources: information of FGFR2 interacting proteins for whole human organism (from experimental and literature-based curation) and experimentally validated interactions specific for the mammary tissue. The list of unique names was standardized to HGNC (HUGO Gene Nomenclature Committee) approved symbols, which were updated according with HGNC's Multi-symbol checker tool (available at: `www.genenames.org/cgi-bin/symbol_checker`). Results from the data mining approach are organized in an occurrence binary matrix available at Supplementary Online Material. Visual representation of the two data-mined (whole-body and mammary tissue) FGFR2 interactomes was built with RedeR R package, available in Bioconductor Repository (`http://bioconductor.org/packages/release/bioc/html/RedeR.html`)(Castro *et al.*, 2012).

2.2 Conditional analysis on FGFR2 downstream regulators

FGFR2 interacting proteins retrieved for breast tissue were set as the input list of potential candidate modulators within RTN's R package conditional analysis (tni.conditional function) (package available at: `http://bioconductor.org/packages/release/bioc/html/RTN.html`). This conditional modulation method, based on Mindy algorithm (Wang *et al.*, 2009), identifies genes able to change a TF's activity by looking at a large expression set and detecting changes in mutual information pattern between the transcription factor and its targets when conditioned to expression values of a candidate modulator gene. The method ranks the matrix according to expression of a candidate modulator gene (M), subsets it to contain samples with either high or low expression of M and then re-sorts samples based on TF expression (Campbell *et al.*, 2016). After multiple hypothesis testing corrections, the algorithm decides whether the input gene is capable or not of modulating the regulon, either by stimulating TF's activity or by suppressing it. Conditional

step was performed for the preprocessed transcriptional network built by Fletcher (Fletcher *et al.*, 2013; Curtis *et al.*, 2012) with 434 candidate modulators and 23 transcription factors associated with breast cancer risk (Castro *et al.*, 2015). P-value cutoff was set to 0.01 and p-value adjustment method used was bonferroni.

2.3 FGFR2 perturbation experiments

The list of differentially expressed genes for different constructions of FGFR2 cascade activation in MCF-7 breast cancer cell line previously described by Fletcher *et al.* (2013) was surveyed in order to retrieve modulation candidate genes that are responsive to FGFR2 signalling perturbation. The three different model systems - simulation of endogenous receptor activation (Exp1), -pathway activation by synthetic ligand molecule (Exp2) and overexpression of FGFR2 (Exp3) - were built to contain specific FGFR-regulated genes derived from the contrast between estradiol only control treatment (E2) versus response to E2 supplementation followed by receptor stimulation (via FGF10 for Exp1, AP20187 for Exp2 and tetracycline for Exp3) for times 6hrs, 12hrs or 24hrs. Additionally, Campbel *et al.* (2016) datasets on FGFR2 super-expression, designed to demonstrated how signalling counteracts estrogen activation were also examined. For this methodology, five estrogen receptor positive (ER+) breast cancer cell lines were used (MCF7, ZR751, BT474, T47D and SUM52PE) and each received a 6 hours treatment of 1 nM estradiol plus 100 ng/ml FGF10. The differential analysis was performed setting contrast between the treatment and a control group that received only 1nM E2. The list of modulators differentially expressed for Fletcher's and Campbell's approach were selected to proceed to knockdown analysis. The knockdown approach described by (Campbel *et al.*, 2016) comprised transfection of MCF-7 cells with FGFR2 directed siRNA, following stimulation. Response was assessed by absence of supposed FGFR2-related genes when compared to scrambled siRNA control group.

2.4 Breast tumour PDX response to FGFR2 inhibitors

A set of Novartis' patient derived xenografts (Gao *et al.*, 2015) results was explored in order to evaluate whether genes that respond to knockdown perturbation had perceptible impact on breast cancer treatment with 4 different drugs: FGFR2 inhibitors BGJ398 and LLM871, Tamoxifen (an anti-estrogen drug) and the chemotherapeutic Paclitaxel. The RNA Seq data was stratified according to FGFR2 gene expression (low and not-low), followed by later sub stratification of expression values for the gene of interest. Treatments were compared according to tumour volume (mm3) over time. Repeated Measures One Way Anova (RM ANOVA) was applied as statistical method followed by Bonferroni's correction.

3 Results and Discussion

3.1 Approach

To find important mediators of FGFR2 signalling within breast cancer, this research followed the work-flow represented in **Fig 1**. Data-mining (Fig 1a) was used to retrieve predicted and validated interactors of FGFR2 gene product from the thirteen databases listed in **Table 1**. For the entire healthy human organism, the final list of FGFR2-interacting proteins comprised 756 proteins, of which 434 were mapped to breast tissue by up to four of these databases, thus constituting a tissue-specific network.

Gene names of breast-specific mediators were set as input for Mindy algorithm (Fig 1b) to identify which of those were possible modulators for the activity of transcription factors highly associated with breast cancer risk. In this step, 422 out of 434 genes were seen to modulate at least one of the 23 tested TFs.

The search on key modulators within FGFR2 cascade was followed by analysis of multiple data on differentially expressed genes for different perturbations to the receptor (Fig 1c), which revealed that 52 genes from previous step responded to FGFR2 activation. Moreover, only eight of those responded to receptor knockdown. Lastly, the patient derived xenografts dataset (Fig 1d) allowed to identify two genes capable of assisting therapeutic response to tamoxifen according to their expression profile.

3.2 Data-mining of FGFR2 PPI network reflects FGFR2 broad impact

Normalized outcome of the different databases allowed to build protein-protein interaction networks. Firstly, an overall human-FGFR2 network was built for the 756 proteins that were seen to establish interaction with the receptor along all human tissues (Fig. 2a) according to experimentally validated and literature-based informations. Secondly, an experimentally validated healthy tissue-specific network of was built for the 434 interactors of the fibroblast growth factor receptor 2 retrieved when considering normal breast-tissue only (Fig. 2b). The most reported interactions are shown as highlighted and named elements of each PPI network in figure 2.

3.3 Conditional analysis reveals the receptor's downstream regulators

Modulation methodology was an interesting method to help infer key players of FGFR2-mediated response inside breast tumour biology. The technique interrogated each of the 434 FGFR2-interacting proteins inside a gene expression profile to identify significant candidate modulators for the activity of at least one of the chosen TFs (Wang *et al.*, 2009). For this, proteins within healthy breast tissue interactome were assessed for 23 Transcription Factors (TFs) previously described as usually enriched with BC risk SNPs (Castro *et al.*, 2015). In the outcome, partially presented at **Supplementary Table 1** and fully available at Online Supplementary Material, it was possible to observe that most genes present inside the PPI were capable of modulating at least one TF.

3.4 Some modulatory genes respond to experimental perturbation onto FGFR2

Due to the strong relation between FGFR2 mediators and TFs associated with ER response, collaboration datasets provided by Fletcher *et al.* (2013)

Table 1. Details on consulted databases. Pred - Predicted Interactions. Val - Experimentally Validated Interactions. PPI-size - Number of interactors retrieved for FGFR2 for all human tissues

Name	Version	Evidence	PPIsize
StringDB	10.5	Pred/Val	272
IntAct	4.2.10	Val	28
BioGrid	3.4	Val	86
APID	0	Val	87
DIP	0	Val	10
HPRD	9	Val	21
MINT	0	Val	15
I2D	2.9	Pred/Val	58
Spectra	0	Val	41
GIANT	0	Val	94
Mentha	2.5	Val	45
HumanBase	0	Pred/Val	95
IID	2017-04	Pred/Val	385

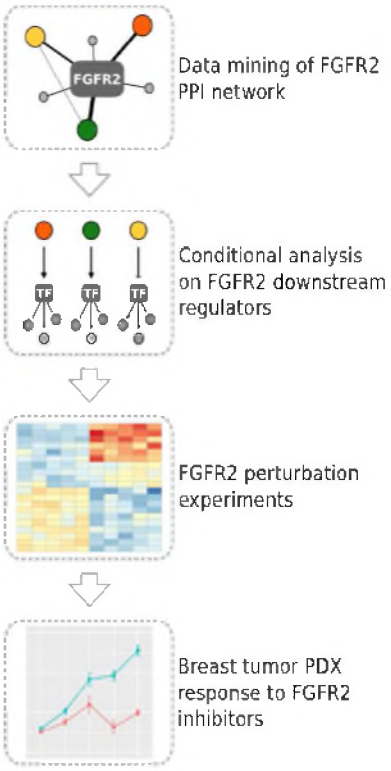


Fig. 1: **Schematic diagram of applied methodology**. **Step 1:** Thirteen databases were consulted and a breast tissue specific protein-protein network centered on FGFR2 was built after preprocessing the data-mined information. **Step 2:** Universe of interactions was set as input for modulatory analysis via Mindy algorithm implemented inside the RTN R package for 23 transcription factors. **Step 3:** Modulators had importance assessed by analysis of experimental data. Only differentially expressed genes that were responsive to FGFR2 knockdown experiment were kept for further scrutiny. **Step 4:** The remaining corroborated genes were checked concerning their behavior within PDX datasets and their capability of impacting drug response.

and Campbel *et al.* (2016) of FGFR2 perturbation in ER+ cellular lines were included to filter the list of modulators to reach those of true validated effect. The two sets of experimental data include endogenous simulation and overexpression of FGFR2 pathways stimulated by estradiol supplementation in MCF7 breast cancer cell line (Fig. 3A), as well as overexpression construction for 5 different ER+ cell lines (Fig. 3B) along with MCF-7 FGFR2 knockdown results (Fig. 3C). Experiments with MCF-7 express the most reliable results for understanding FGFR2-mediated ER response once this cell line was designed have optimal signalling of estrogen-dependent oncogenic pathways (Holliday and Speirs, 2011). Also, MCF-7 cell line have good FGFR2 expression levels, similar to ZR751 and T47D, not as high as SUM52PE nor as low as BT474 (Campbel *et al.*, 2016). Fifty two modulators genes were shown as differentially expressed for both over-expression experiments (Fig. 3A and 3B). It can be observed that level of expression is consistent for most genes throughout tested cell lines: from the range of gene names going

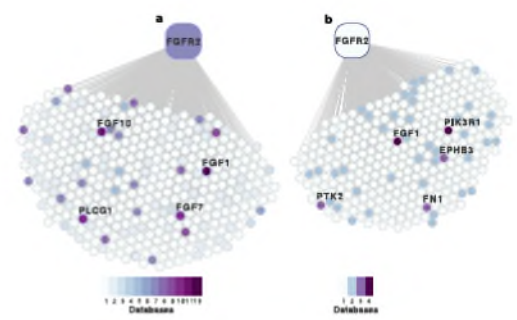


Fig. 2: **Data-mined Protein Protein Interaction Networks: a)** Proteins that interact with FGFR2 within a healthy human body. **b)** Interactions that occur inside healthy breast tissue. Keys for color range represents the amount of times a given protein within the total universe of 13 databases.

from FOXN3 up to PTPN6, are those genes usually suppressed for such experiments. From ERRFI1 to PHLDA1, are the genes presented mostly as super-expressed for most constructions, indicating the robustness of the methods. A deeper gene selection was performed to observe which ones respond to FGFR2 depletion by siRNA knockdown (Fig. 3C). After depletion, from the 52 differentially expressed genes, 8 could be detected within control (scrambled siRNA) construction with consistent expression levels when compared to previous experiments, and lost their signal in anti-FGFR2 siRNA group, indicating that depletion of the receptor directly impacts them. They are PBX1, ERRFI1, SHB, EPHA2, TNIK, S100A14, TGFBR2 and MAP3K5.

The eight lasting genes, are those that showed robust relation with the FGF kinase receptor: they are inside the initial PPI network, are strong candidate modulators of breast cancer risk transcription factors and respond to all fibroblast growth factor receptor perturbation experiments. Thus, they are believed to participate as transducers for this cascade and, more importantly, we suspect that some might play role in tumour progression via estrogen regulation.

3.5 EPHA2 and PBX1 are the most responsive genes for PDX analysis of drug response

To assess how the remaining genes act inside a more realistic environment of breast cancer, Patient Derived Xenograft (PDXs) data was searched in order to evaluate which genes have impact over different BC treatments. The PDX approach uses immunodeficient mice implanted with cells from a given primary tumour to mimic a closer-to-reality environment. It is a promising strategy once the animal model keeps most of the major genetic variations coming from patients, enabling to perform multiple pre-clinical drug screenings to evaluate treatment response and cancer behavior with high accuracy, offering results that can be transposed to back to patient by guiding decision-making (Gao *et al.*, 2015).

The BRCA PDX samples, reorganized by expression levels of these genes, when tested for 4 treatments (BGJ891, LLM871, Tamoxifen and Paclitaxel), showed that LLM871 seemed to be as efficient as Paclitaxel for these samples (Fig. 4 and Supplementary Figures 2-9), indicating that FGFR2/4 inhibition might have similar therapeutic effects when compared to microtubule inhibition chemotherapy, with possibly less side effects,

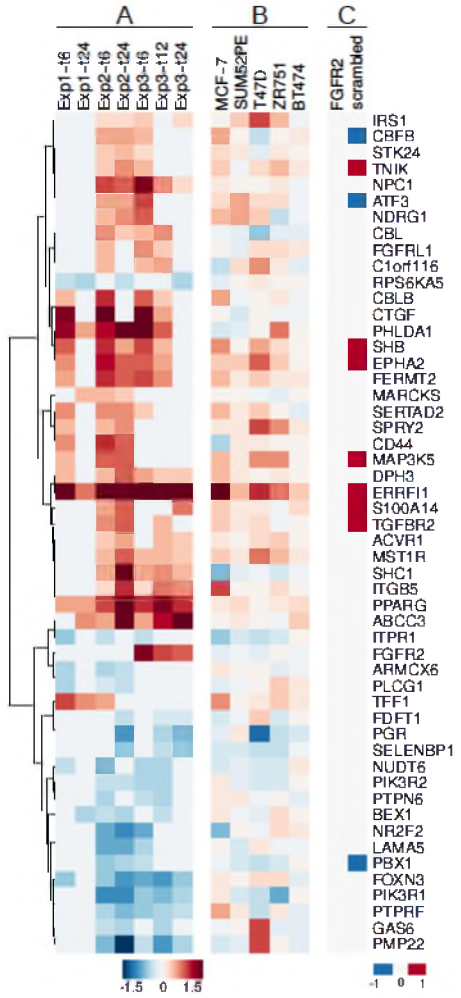


Fig. 3: **Experimental Data Analysis.** **(A)** Fletcher's dataset of MCF-7 cell lines tested for 3 constructions (see Methods) and **(B)** Campbell's data on 5 breast cancer ER+ cell lines, both induced with estradiol, had their differential expression profile assessed to retrieve the experimental condition of genes that yielded in conditional analysis. Only modulatory-FGFR2-interacting genes that had significant alterations on expression levels in, at least three experimental times in (A), were kept. Differential profile of these 52 genes were retrieved from Campbell's investigation in order to evaluate how they show for other cell lines. Even though a few differences can be spotted, Exp1-t6 construction in (A) is equivalent to MCF-7 construction in (B). Direct FGFR2 transducers were picked through Campbell's FGFR knockdown (C) analysis in MCF-7. Genes that lost signal of differential expression after siRNA FGFR2 knockdown (C - FGFR2 column) compared to scrambled, and were consistent with previous MCF-7 expression, were considered good possible targets for further analysis.

once it has a less broad effect, but further investigation is needed to evaluate whether this pattern is significant for other breast tumours.

After cohort stratification and assessment of drug responses, we also found that the PBX homeobox 1 (PBX1) and the EPH Receptor A2 (EPHA2) genes respond promptly to FGFR2 perturbation and seems to impact breast cancer treatment in xenografts data. The PBX1 gene, is a pioneer factor fit to start cellular-fate changes. It encodes a nuclear TALE (Three Aminoacid Loop Extension) protein known to participate in hematopoiesis (Ficara *et al.*, 2009), as well as multiple organogenesis processes (Schnabel *et al.*, 2001, 2003) by acting as cofactor for TFs, with the purpose of increasing chromatin interaction with silenced genome areas and enhance transcriptional activation (Grebbin and Schulte, 2017). It was previously shown that this gene has the higher expression levels in primary breast tumours than other TALE proteins (Crijns *et al.*, 2007) and mediates estrogen regulation (Magnani *et al.*, 2011). Indeed, besides LLM781, high expression levels of PBX1 were seen to significantly ($p<0.005$) increase treatment response of tamoxifen (Fig. 4a), an ER-blocker drug, when compared to control untreated group. Low levels of this gene resulted in tamoxifen-treated samples having response curves equivalent to untreated group, with no statistical difference between them.

Modulation analysis (Fig. 4c) suggests that PBX1, when not suppressed by FGFR2, represses both cluster 1 and cluster 2 TFs, being a negative modulator of, among others, ESR1 and NFIB genes, going accordingly with the concept that it has participation in estrogen regulation. The hypothesis is that PBX1 participates within - or in a similar way of - a FGFR2-driven regulatory loop, as was recently reported by Campbell *et al.* (2018). In this loop, FGFR2 plays within two cascades that lead to inhibition of ESR1 transcription, thus promoting an aggressive basal-like ER- breast cancer phenotype. One pathway uses YBX1 DNA-binding protein to directly contact and repress ESR1, whereas the second pathway has FGFR2 to activate the nuclear factor NFIB, also leading to repressed estrogen regulon intermediated by FOXA1 signalling.

Even though high expression of PBX1 might lead to a better drug response, the relation between PBX1, ESR1 and patient's outcome is not yet understood. Poorer patient outcome was found associated with PBX1-dependent pathways in breast cancer by Magnani *et al.* (2011), who also stated that the homeobox protein guides ERa to specific binding sites and is able to control over 70% of estrogen response. This regulation is consistent with the findings that, for the MCF-7 cell lines used in this study, the activation of the fibroblast growth factor receptor 2 was seen to lower PBX1 levels leading to unresponsiveness of PDX samples to tamoxifen treatment, indicating that transcription of PBX1 might act to increase ER and/or estrogen-sensitive proliferative pathways. Nevertheless, our study also showed that the homeobox gene tends to be a negative modulator of ESR1 regulon, making unclear by which means PBX1 overexpression can lead to increased ESR1-mediated activity while negatively modulating its putative targets. One point to consider is whether the PBX1 locus in this PDX cohort corresponds to its ancestor form or if the locus presents any of the eleven intronic SNPs strongly related with early breast cancer onset as reported by Rafiq *et al.* (2013), such as the rs1387389 variant, causing regulation to be changed. Furthermore, Wang *et al.* (2017) associated the homeobox over expression with significant up-regulation of lipid metabolism (LiMe) genes both for benign breast tissue as well as for Oncomine ER+ breast cancer set, seeming that increased lipid metabolism would assist a better prognosis, contrary to the idea that lipid rafts increase tumour cell proliferation and leads to worst outcomes (Beloribi-Djefaflia *et al.*, 2016).

The ephrin type-A receptor 2 (EPHA2) gene, the second most responsive gene within PDX dataset (Fig. 4b), was found to be induced by FGFR2 expression and modulates positively six transcription factors from risk-related clusters (Fig. 4d), mostly counteracting PBX1 modulatory action, once its mode of action overlaps with 5 PBX1-modulated TFs.

Previous researches with EPHA2 strongly indicated this genes plays role in tumoural stem cell maintenance for glioblastomas and human lung cancer (Binda *et al.*, 2012; Song *et al.*, 2014) and presents a more direct

relation with patient outcome, once its high expression levels in breast cancer tissue are predicted to increase metastatic capability, poor patient prognosis and reduced survival rates (Blan *et al.*, 2017; Edwards *et al.*, 2017).

Here, we showed that, in concordance with other authors, high expression of EPHA2 leads to reduced dependence on estrogen for tumorous growth, resulting in less sensitivity to tamoxifen treatment (Gökmen-Polar *et al.*, 2011; Tandon *et al.*, 2011). Low gene expression indicated that, besides restored tamoxifen response, response to FGFR2 inhibition via LLM781 drug gets enhanced (compared with untreated group), but treated group responds significantly to LLM781 even when EPHA2 levels are increased (Figure 4b). Tandon *et al.* (2011) and Song *et al.* (2017) also showed, respectively, that lower EPHA2 activity can lead to better response with Her2 antibody therapy (trastuzumab) and reduces proliferation in triple-negative breast cancers. To the best of our knowledge, there are no other studies so far stating a loop regulation between FGFR2 and EPHA2, which seem to be an interesting dual of therapeutic targets.

4 conclusion

Throughout the studies, it was observed that increase in FGFR2-related risk for breast cancer development results from a coordinated repercussion of combined susceptibility factors that build a risk haplotype within the second intron of the gene (Robbez-Masson *et al.*, 2013). Consecutive Genome Wide Association Sudies (GWAS) perfomed for different populations (Raskin *et al.*, 2008; Shan *et al.*, 2012; Barnholtz-Sloan *et al.*, 2011; Han *et al.*, 2011) displayed strong correlation with mammary tumours and FGFR2 risk alleles. The consequences of such alterations is a deep impaired cross-talk between cascades and inter connectors that are not yet fully understood. Taking this interconnectivity into consideration, our work was able to show that PPI-approach backed up by experimental data was capable to help inference of important novel transducers of FGFR2-mediated estrogen regulation loop in tumour biology, be them players inside the previously reported regulation or within a still unknown alternative pathway. The PBX1 and EPHA2 transducers were seen to respond directly to changes in FGFR2 levels within breast cancer cell lines experiments with estradiol supplementation and FGFR2 knockdown. These genes are already known to be highly involved on the onset and progression of breast tumour, but experimental validation is needed to pin-point their respective timing, interactions, mode of action and actual contributions to survival and progression within the cancerous pathways via fibroblast growth factor receptor 2.

The Eph type-A Receptor is an interesting target to be inhibited and used as a co-treatment with either tamoxifen or trastuzumab, to restore estrogen sensitivity within tamoxifen unresponsive BCs due to impaired ESR1 activity or to further improve the quality for triple negative breast cancer therapy. Such co-treatments can be an interesting way to address major problems that patients go through such as treatment specificity, efficiency and recurrence, once this approach might offer ways to reactivate pathways deregulated by the cancer mechanisms and to diminish the establishment of tumour stem cells, by targeting these genes that have high expression specially in breast tissue tumours.

PBX1 gene is thought to be part of an even more complex cascade of estrogen and FGFR2 regulation, once its high expression seems to also restore tamoxifen response through an indirect pathway even though the gene acts as a negative modulator of the drug's direct target, the ERS1. More study is necessary to know whether a higher expression of the homeobox gene leads to a better overall survival, response and prognosis.

We suggest that further experiments should take place now to establish the steps of FGFR2-mediated response up to ERS1 regulation. Two-hybrid screening or FRET experiments could be useful approaches to analyze direct protein-protein interactions between FGFR2, EPHA2 and PBX1 and implementation of cell line and PDX studies could be performed in order to evaluate the effects of concomitant inhibition of EPHA2 and FGFR2 over tumorous cells and clinical outcomes.

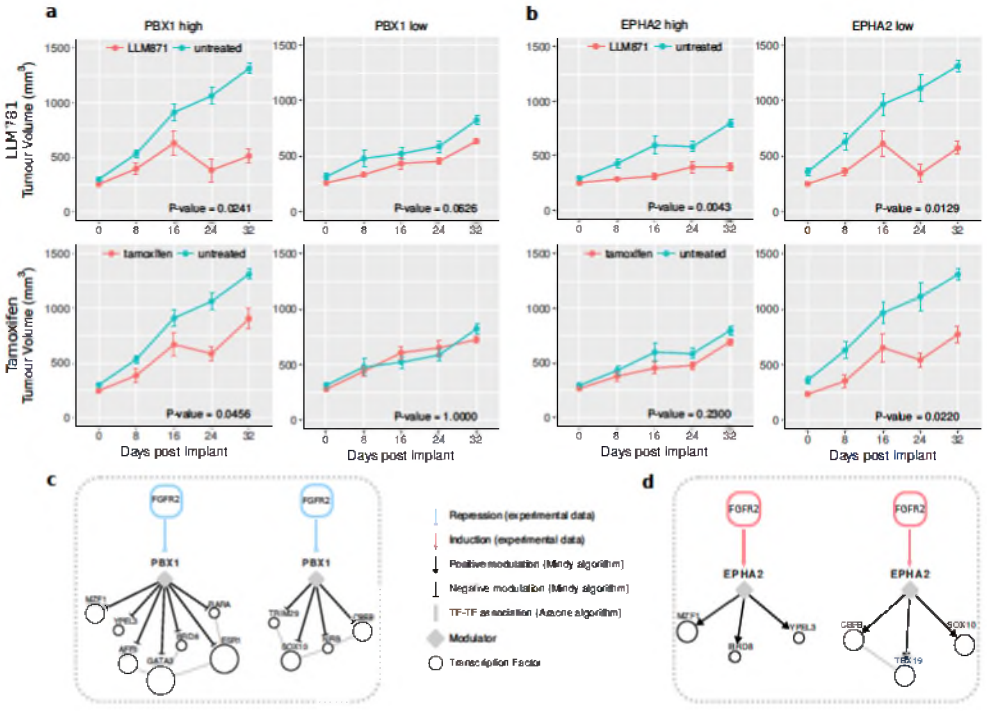


Fig. 4: **Patient Derived Xenograft Analysis**. Breast cancer PDX cohort stratified according to **a)** PBX1 or **b)** EPHA2 expression levels (High and Low). For each level, treatment response was plotted for 2 different anti-cancer drugs: the FGFR2/4 inhibitor LLM781 and the ER-binding Tamoxifen. PDXs plots correlates the tumour volume (Y-axis) along time after primary tumour implant into imunosuppressed mice (X-axis) to compare the control untreated group (blue line) with treated group (red line). Statistical test applied to generate pvalues was Repeated Measures One Way ANOVA, adjusted with Bonferroni corrections. **c)** Modulation analysis with Mindy algorithm for PBX1 and **d)** for EPHA2. Blue container indicates that FGFR2 has a suppressive regulation onto PBX1 transcription (diamond shaped) inferred by differential expression and red container indicates induction of EPHA2 gene expression. RTN's conditional algorithm allowed to construct the representation of modulation for risk transcription factors (round forms) present in cluster 1 (left plot of the box) and cluster 2 (right plot). Arrows connecting modulators to TFs indicate modulatory mode of action and gray lines interconnecting TFs represents the existence of mutual information between them.

5 Reproductibility

Interactive plots of protein-protein interaction networks and modulatory analysis presented in this article can be reproduced with FGFR2interactome package, a data package built in R language provided at Github repository: `https://github.com/kelgoncalves/FGFR2interactome`. Full tutorial on how to install and access the R data package can be found inside readme file.

Acknowledgements

We kindly thank the Federal University of Paraná (UFPR), the Technological and Professional Education Sector (SEPT) facilities and the collaborators that provided experimental results for this research.

Funding

This work was supported by the Coordination for the Improvement of Higher Education Personnel (CAPES) and the National Council for Scientific and Technological Development (CNPq).

References

Barnholtz-Sloan, J. S. *et al.* (2011). Replication of GWAS "hits" by race for breast and prostate cancers in European Americans and African Americans. *Frontiers in Genetics*, **2**(JULY).

Beloribi-Djefaflia, S. *et al.* (2016). Lipid metabolic reprogramming in cancer cells. *Oncogenesis*, **5**(1), e189.

Bian, S. *et al.* (2017). EphA2 , EphrinA1 expression in breast cancer and its relationship with clinical pathological study . **28**(13), 6022–6027.

Binda, E. *et al.* (2012). The EphA2 Receptor Drives Self-Renewal and Tumorigenicity in Stem-like Tumor-Propagating Cells from Human Glioblastomas. *Cancer Cell*, **22**(6), 765–780.

Boyle, E. A. *et al.* (2017). Leading Edge Perspective An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell*, **169**(7), 1177–1186.

Brewer, H. R. *et al.* (2017). Family history and risk of breast cancer: an analysis accounting for family structure. *Breast Cancer Research and Treatment*, **165**(1), 193–200.

Campbel, T. M. *et al.* (2016). FGFR2 risk SNPs confer breast cancer risk by augmenting oestrogen responsiveness. *Carcinogenesis*, **37**(8), 741–750.

Campbell, T. M. *et al.* (2016). Identification of post-transcriptional modulators of breast cancer transcription factor activity using MINDy. *PLoS ONE*, **11**(12).

Campbell, T. M. *et al.* (2018). ER alpha binding by transcription factors NFIB and YBX1 enables FGFR2 signalling to modulate estrogen responsiveness in breast cancer. *Cancer Research*, **78**(2), canres.1153.2017.

Castro, M. A. A. *et al.* (2012). RedeR: R/Bioconductor package for representing modular structures, nested networks and multiple levels of hierarchical associations. *Genome Biology*, **13**(4).

Castro, M. A. A. *et al.* (2015). Regulators of genetic risk of breast cancer identified by integrative network analysis. *Nature Genetics*, **48**(November), 12–21.

Crijns, A. P. *et al.* (2007). MEIS and PBX homeobox proteins in ovarian cancer. *European Journal of Cancer*, **43**(17), 2495–2505.

Cui, F. *et al.* (2016). Variants of FGFR2 and their associations with breast cancer risk: a HUGE systematic review and meta-analysis. *Breast Cancer Research and Treatment*, **155**(2), 313–335.

Curtis, C. *et al.* (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*, **486**, 346–352.

Easton, D. F. *et al.* (2007). Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature*, **447**(7148), 1087–1093.

Edwards, D. N. *et al.* (2017). The receptor tyrosine kinase EphA2 promotes glutamine metabolism in tumors by activating the transcriptional coactivators YAP and TAZ. *Science Signaling*, **10**(508).

Etokebe, G. E. *et al.* (2009). Single-Nucleotide Polymorphisms in Genes Encoding Toll-Like Receptor -2, -3, -4, and -9 in CaseâĹ"Control Study with Breast Cancer. *Genetic Testing and Molecular Biomarkers*, **13**(6), 729–734.

Ficara, F. *et al.* (2009). By Maintaining Their Quiescence. *Cell*, **2**(5), 484–496.

Fletcher, M. N. *et al.* (2013). Master regulators of FGFR2 signalling and breast cancer risk. *Nature Communications*, **4**.

Gao, H. *et al.* (2015). High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response. *Nature Medicine*, **21**(11), 1318–1325.

Gökmen-Polar, Y. *et al.* (2011). Dual targeting of EphA2 and ER restores Tamoxifen sensitivity in ER/EphA2-positive breast cancer. *Breast Cancer Research and Treatment*, **127**(2), 375–384.

Grebbin, B. M. and Schulte, D. (2017). PBX1 as Pioneer Factor: A Case Still Open. *Frontiers in Cell and Developmental Biology*, **5**(February), 1–6.

Han, W. *et al.* (2011). Common genetic variants associated with breast cancer in Korean women and differential susceptibility according to intrinsic subtype. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, **20**, 793–798.

Holliday, D. L. and Speirs, V. (2011). Choosing the right cell line for breast cancer research. *Breast cancer research : BCR*, **13**, 215.

Hunter, D. J. *et al.* (2007). A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nature Genetics*, **39**(7), 870–874.

INCA (2018). Instituto Nacional de Câncer José Alencar Gomes da Silva.

Magnani, L. *et al.* (2011). PBX1 genomic pioneer function drives ERα signaling underlying progression in breast cancer. *PLoS Genetics*, **7**(11).

Nickels, S. *et al.* (2013). Evidence of Gene-Environment Interactions between Common Breast Cancer Susceptibility Loci and Established Environmental Risk Factors. *PLoS Genetics*, **9**(3).

Rafiq, S. *et al.* (2013). Identification of inherited genetic variations influencing prognosis in early-onset breast cancer. *Cancer Research*, **73**(6), 1883–1891.

Raskin, L. *et al.* (2008). FGFR2 is a breast cancer susceptibility gene in Jewish and Arab Israeli populations. *Cancer Epidemiology Biomarkers and Prevention*, **17**(5), 1060–1065.

Robbez-Masson, L. J. *et al.* (2013). Functional analysis of a breast cancer-associated FGFR2 single nucleotide polymorphism using zinc finger mediated genome editing. *PLoS ONE*, **8**(11).

Schnabel, C. A. *et al.* (2001). Expression of Pbx1 during mammalian organogenesis. *Mech Dev*, **100**, 131–135.

Schnabel, C. A. *et al.* (2003). Pbx1 is Essential for Adrenal Development and Urogenital Differentiation. *Genesis*, **37**(3), 123–130.

Shan, J. *et al.* (2012). Genome-Wide Association Studies (GWAS) breast cancer susceptibility loci in Arabs: Susceptibility and prognostic implications in Tunisians. *Breast Cancer Research and Treatment*, **135**(3), 715–724.

Song, W. *et al.* (2014). JNK Signaling Mediates EPHA2-Dependent Tumor Cell Proliferation, Motility, and Cancer Stem Cell-like Properties in Non-Small Cell Lung Cancer. *Cancer Research*, **74**(9), 2444–2454.

Song, W. *et al.* (2017). Targeting EphA2 impairs cell cycle progression and growth of basal-like/triple-negative breast cancers. *Oncogene*, **36**(40), 5620–5630.

Sun, S. *et al.* (2012). Increased expression of fibroblastic growth factor receptor 2 is correlated with poor prognosis in patients with breast cancer. *Journal of Surgical Oncology*, **105**(8), 773–779.

Tandon, M. *et al.* (2011). *Emerging strategies for EphA2 receptor targeting for cancer therapeutics*, volume 15.

Turner, N. and Grose, R. (2010). Fibroblast growth factor signalling: From development to cancer.

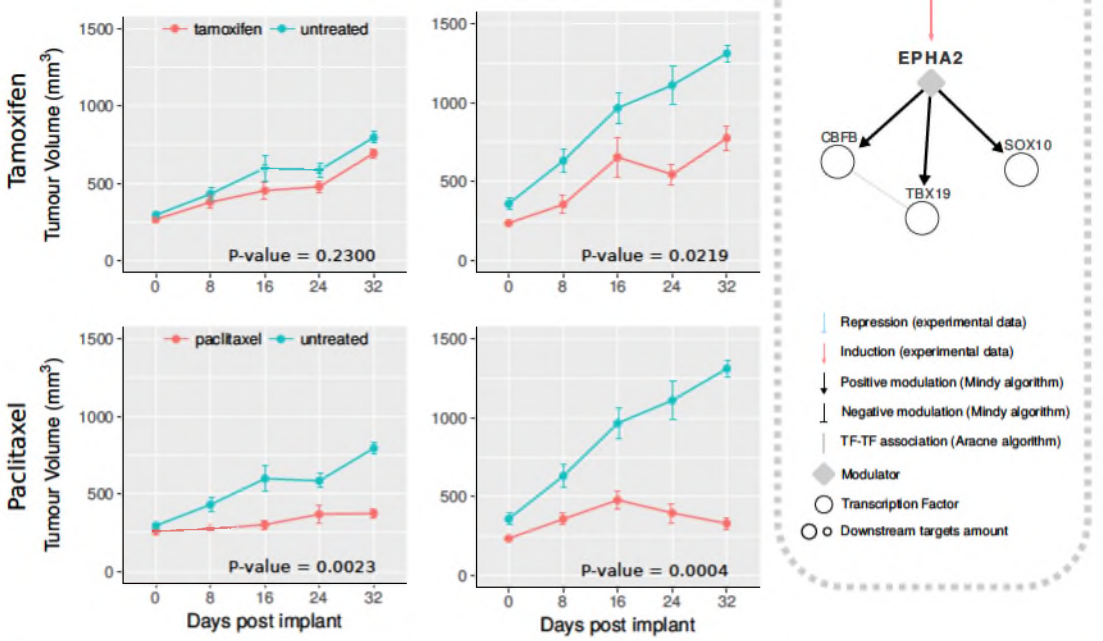
Wang, J. *et al.* (2017). Overexpression of lipid metabolism genes and PBX1 in the contralateral breasts of women with estrogen receptor-negative breast cancer. *International Journal of Cancer*, **140**(11), 2484–2497.

Wang, K. *et al.* (2009). Genome-wide identification of post-translational modulators of transcription factor activity in human B cells. *Nature Biotechnology*, **27**(9), 829–837.

World Health Organisation (WHO) (2018). Breast cancer: prevention and control.

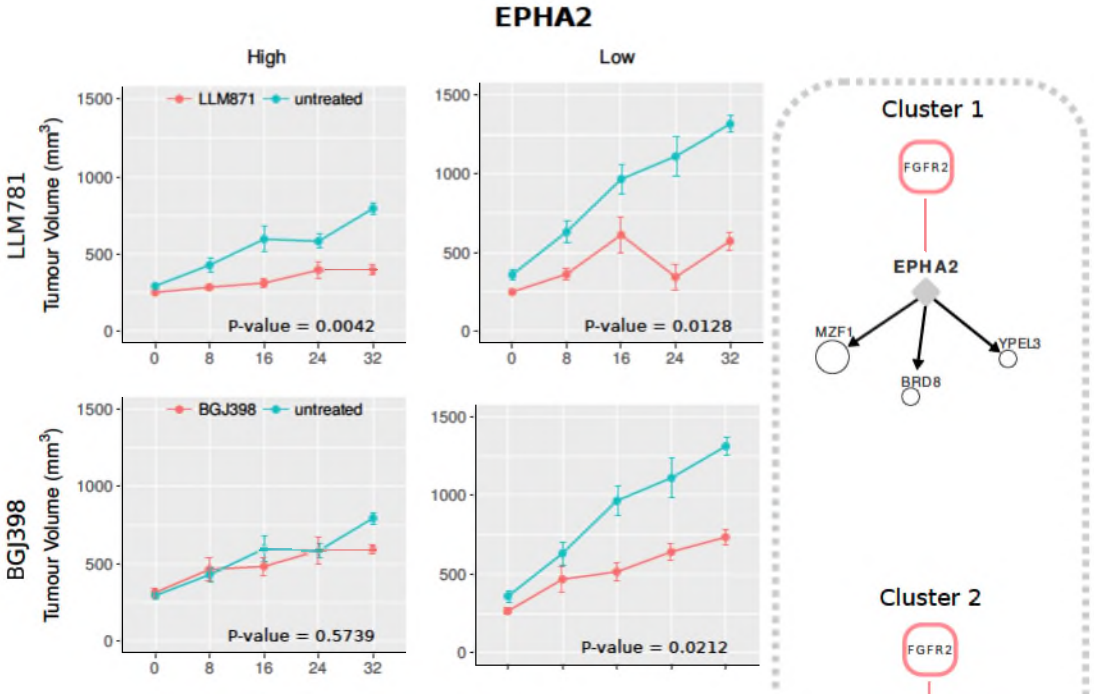
Zhang, Y. *et al.* (2017). Association between FGFR2 (rs2981582, rs2420946 and rs2981578) polymorphism and breast cancer susceptibility: A meta-analysis. *Oncotarget*, **8**(2).

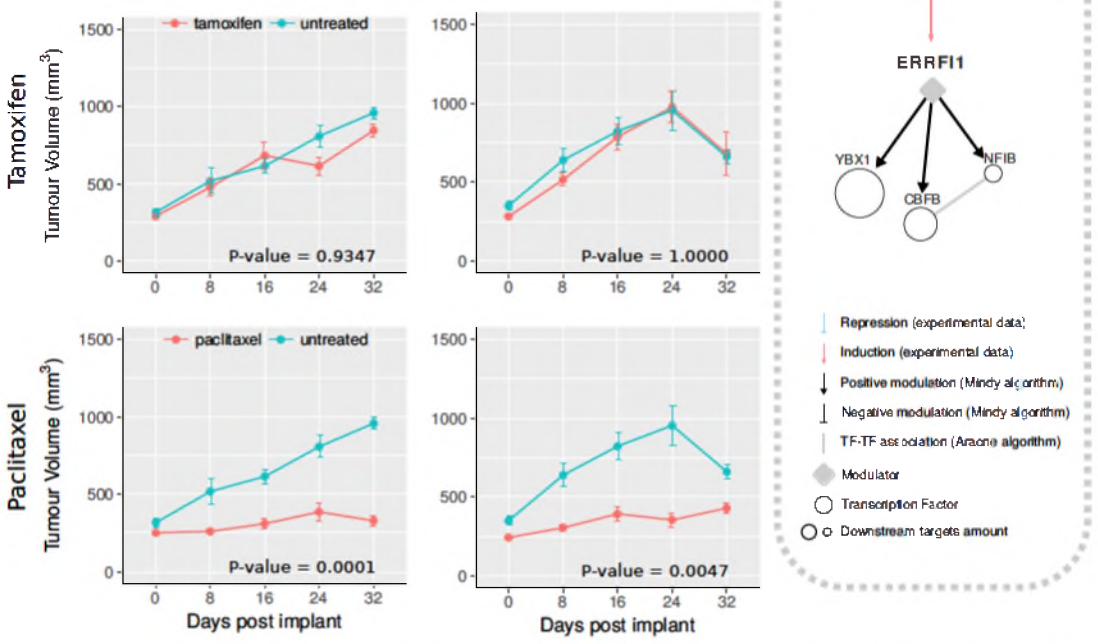
Zhou, L. *et al.* (2012). Three novel functional polymorphisms in the promoter of FGFR2 gene and breast cancer risk: A HuGE review and meta-analysis. *Breast Cancer Research and Treatment*, **136**(3), 885–897.



EPHA2 Pannel. Breast cancer PDX samples of 4 different treatments were stratified according to EPHA2 gene expression levels. Dotted box shows that FGFR2 activity induces EPHA2 expression, which positively modulates cluster 1 and cluster 2 regulons.

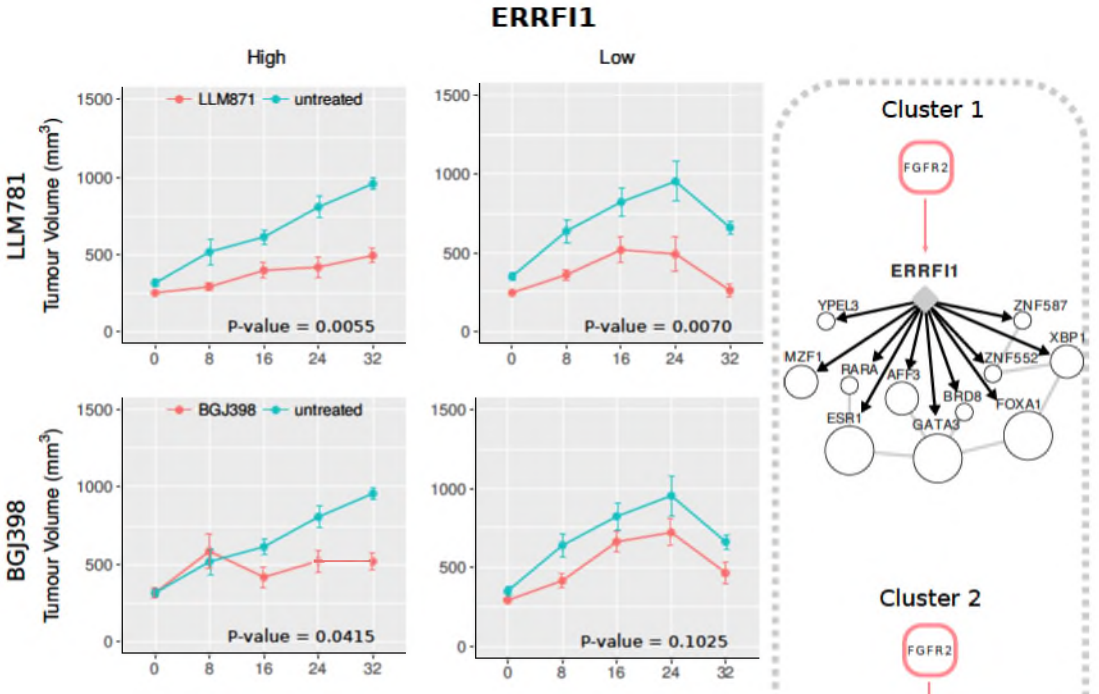
Supplementary Figure 1

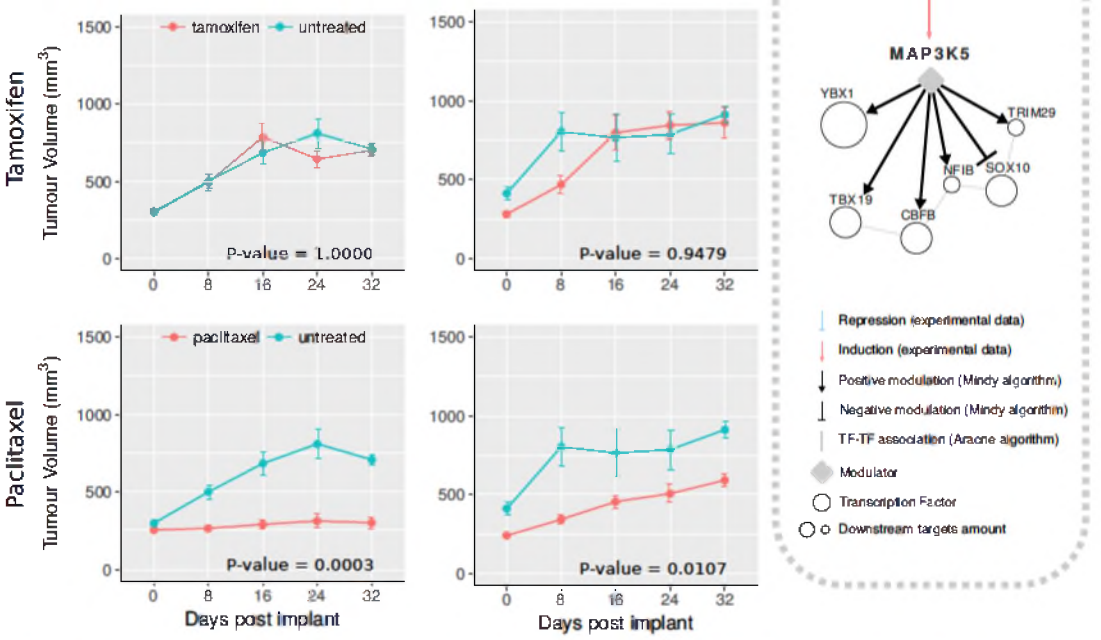




ERRFI1 Pannel. Breast cancer PDX samples of 4 different treatments were stratified according to ERRFI1 gene expression levels. Dotted box shows that FGFR2 activity induces ERRFI1 expression, which positively modulates cluster 1 and cluster 2 regulons.

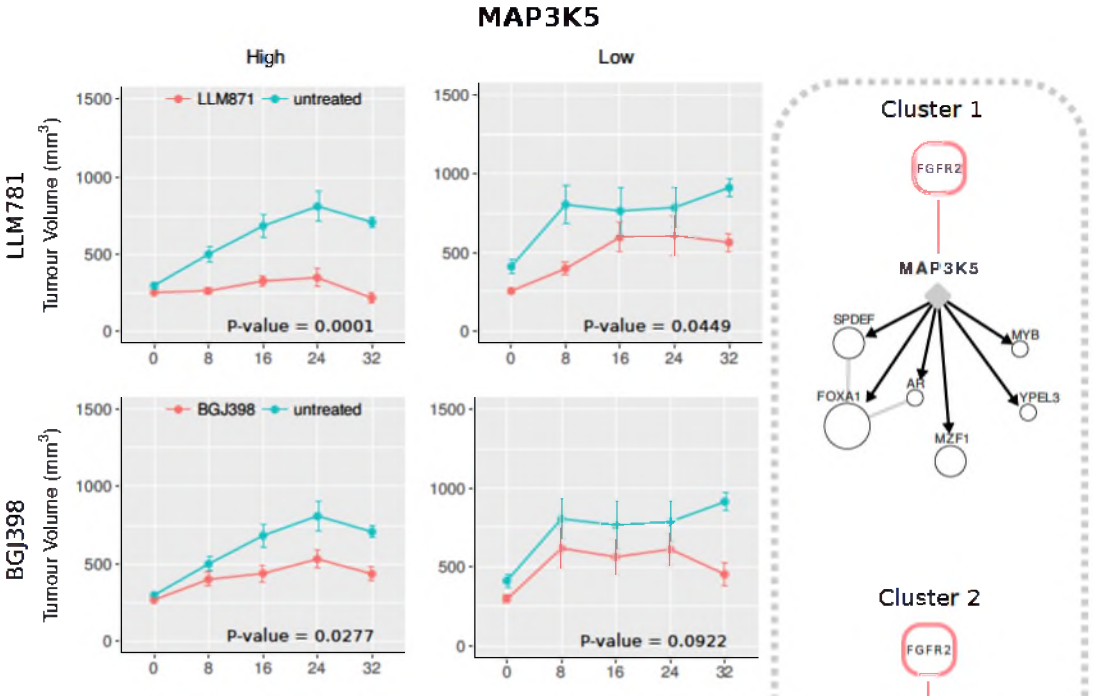
Supplementary Figure 2

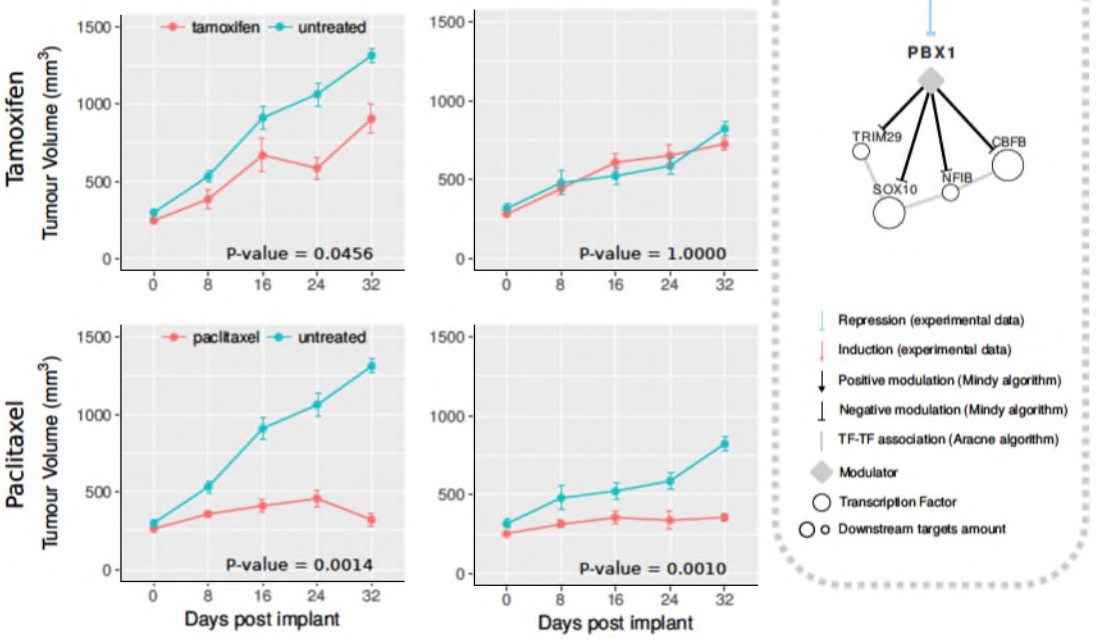




MAP3K5 Pannel. Breast cancer PDX samples of 4 different treatments were stratified according to MAP3K5 gene expression levels. Dotted box shows that FGFR2 activity induces MAP3K5 expression, which positively modulates cluster 1 and has mixed activity for cluster 2 regulons.

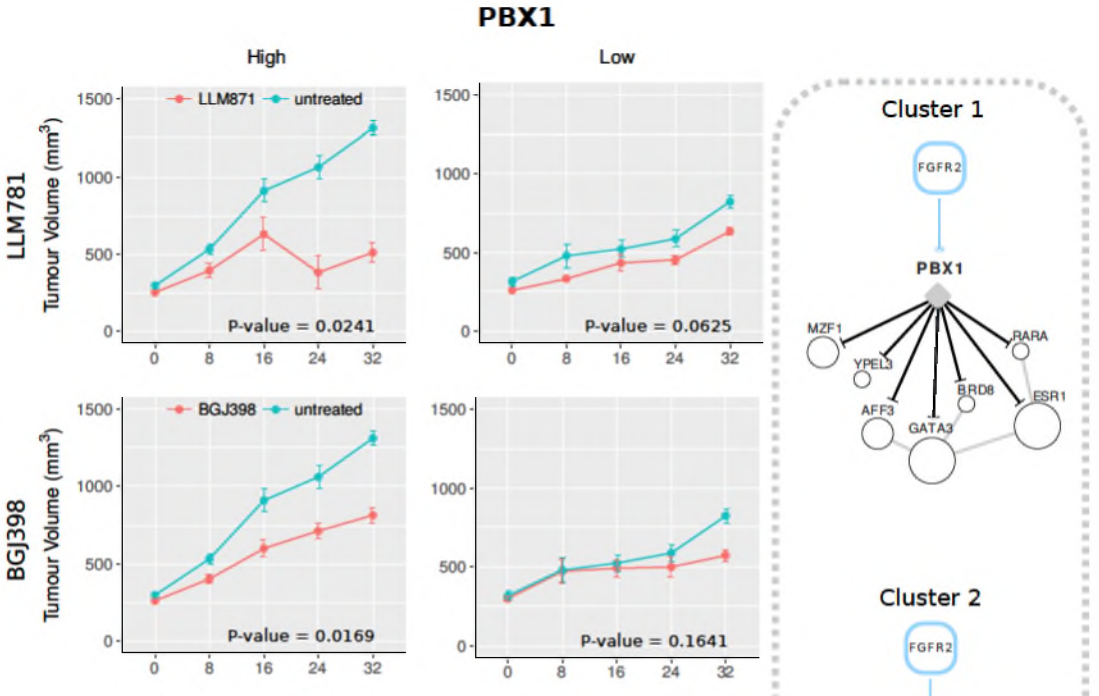
Supplementary Figure 3

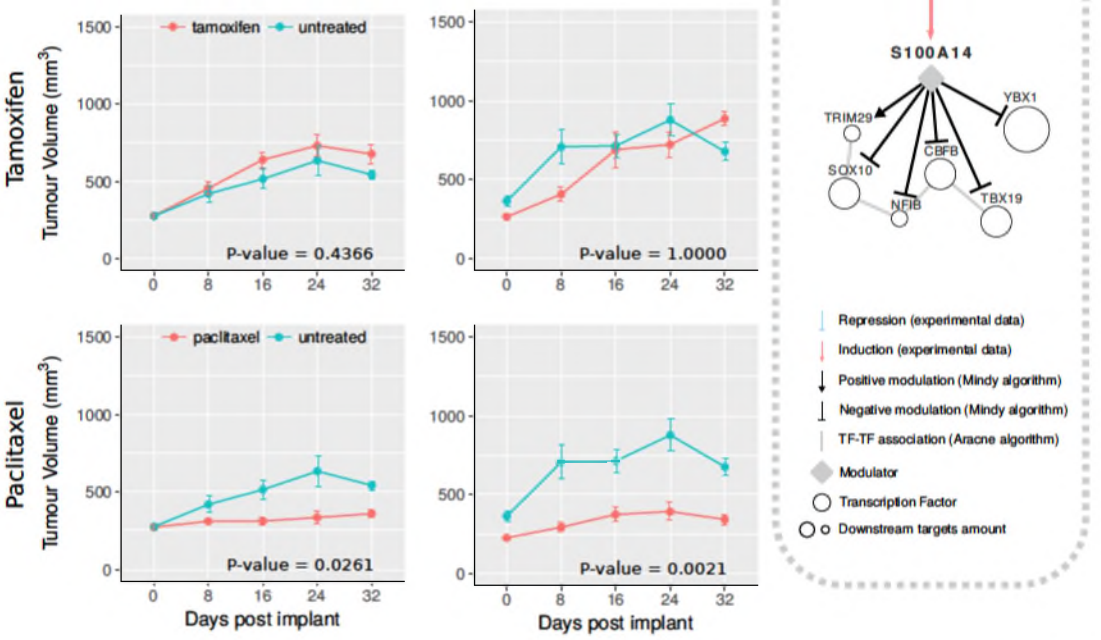




PBX1 Pannel. Breast cancer PDX samples of 4 different treatments were stratified according to PBX1 gene expression levels. Dotted box shows that FGFR2 activity represses PBX1 expression, which negatively modulates cluster 1 and cluster 2 regulons.

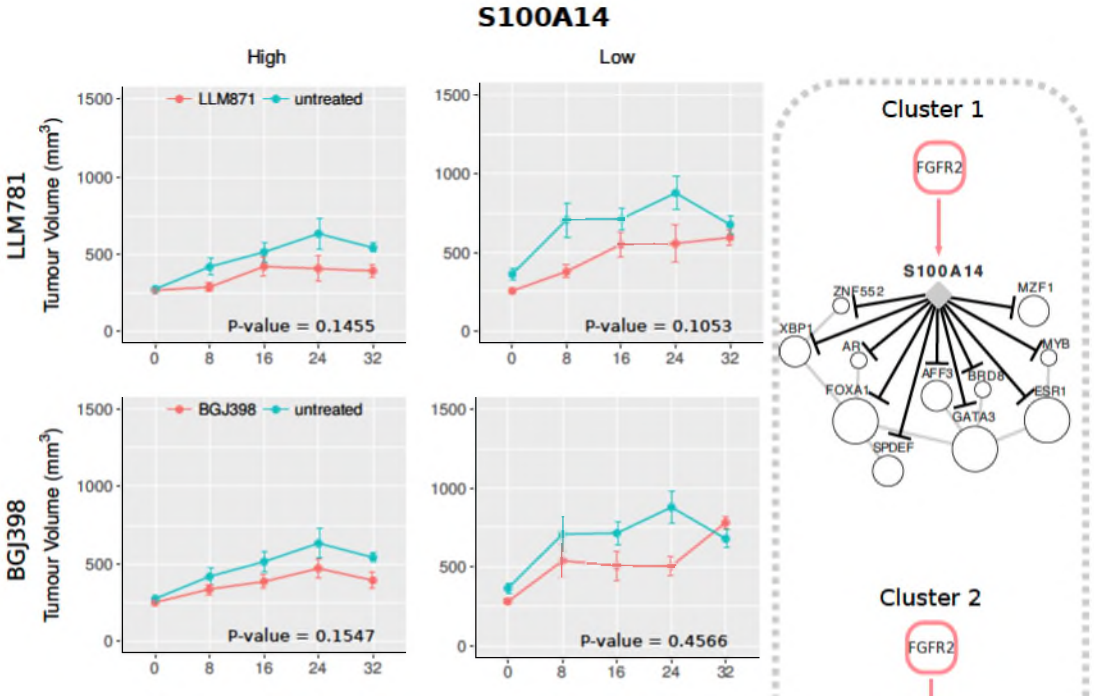
Supplementary Figure 4

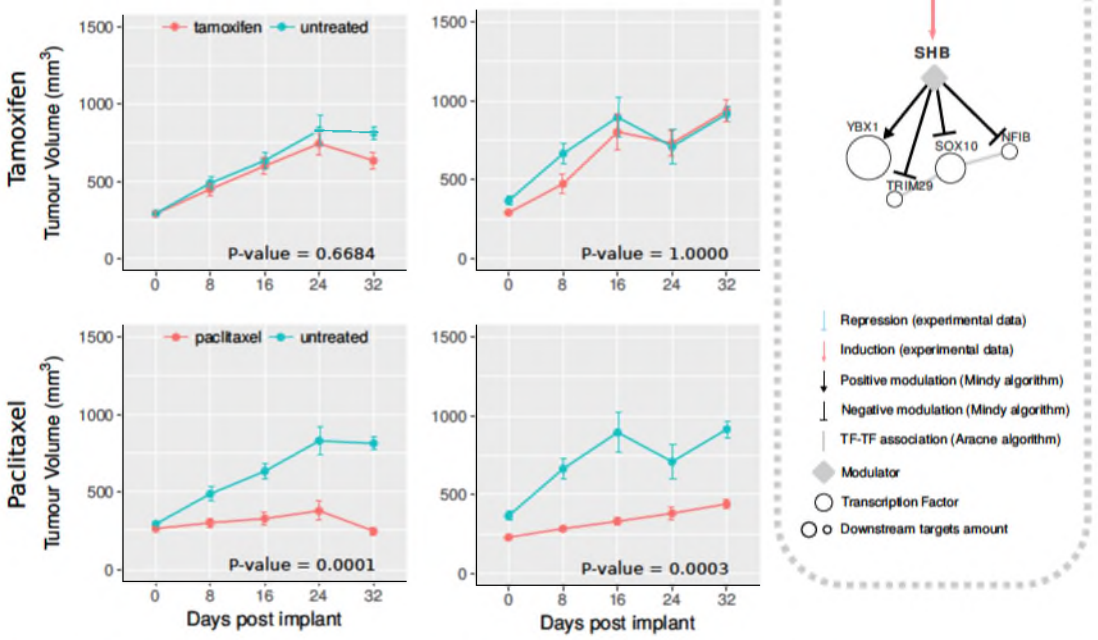




S100A14 Pannel. Breast cancer PDX samples of 4 different treatments were stratified according to S100A14 gene expression levels. Dotted box shows that FGFR2 activity induces S100A14 expression, which negatively modulates cluster 1 and has mixed activity for cluster 2 regulons.

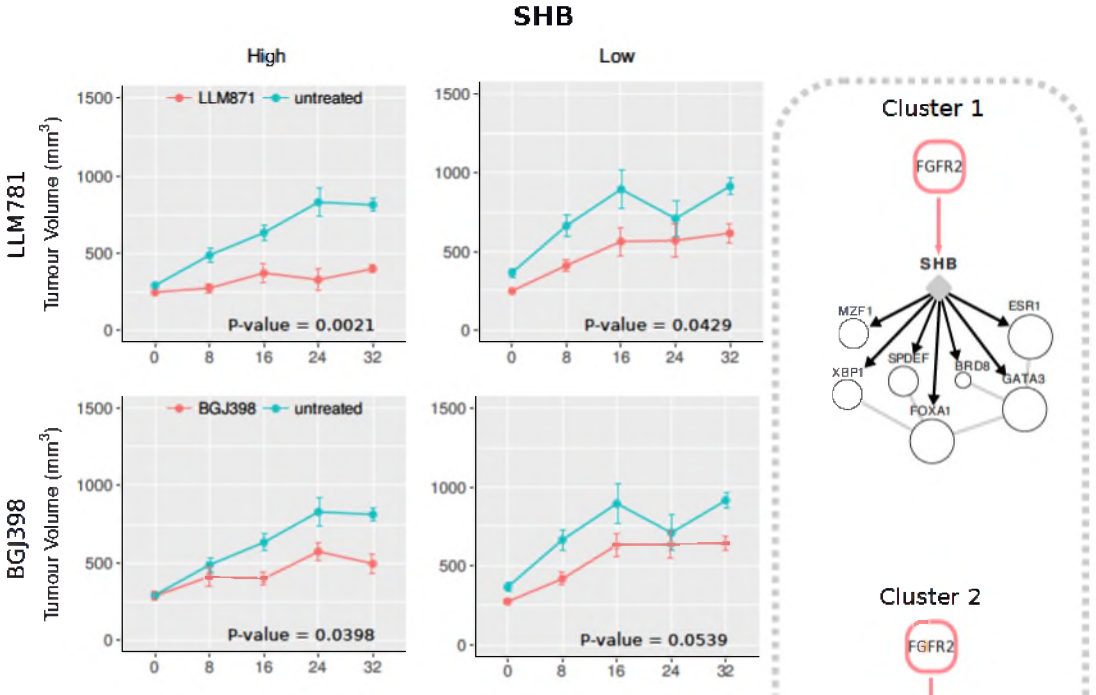
Supplementary Figure 5

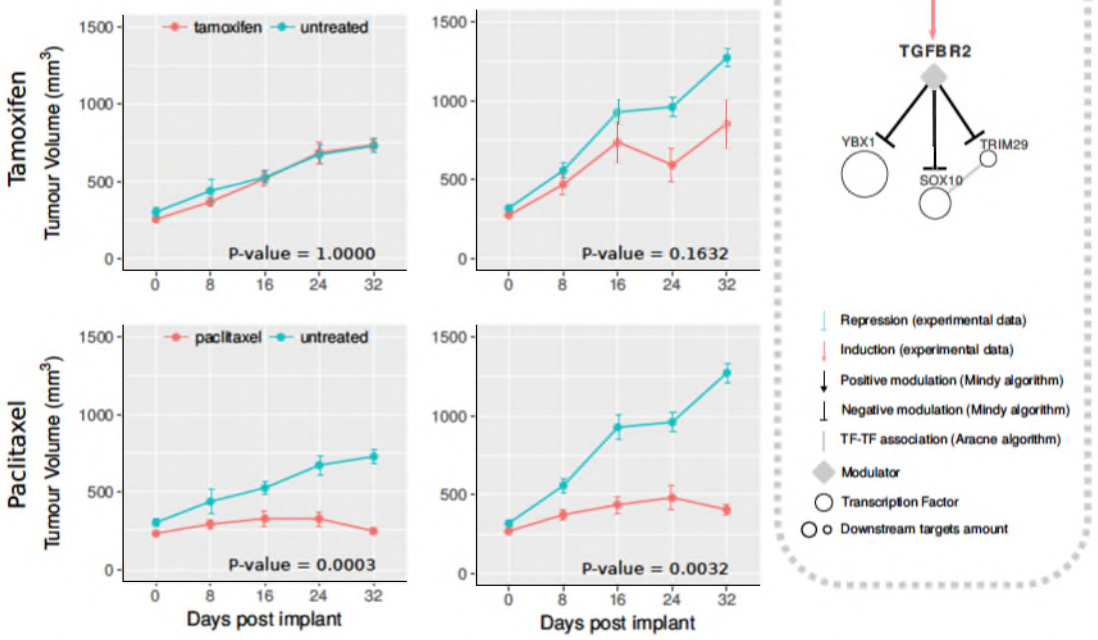




SHB Pannel. Breast cancer PDX samples of 4 different treatments were stratified according to SHB gene expression levels. Dotted box shows that FGFR2 activity induces SHB expression, which positively modulates cluster 1 and has mixed activity for cluster 2 regulons.

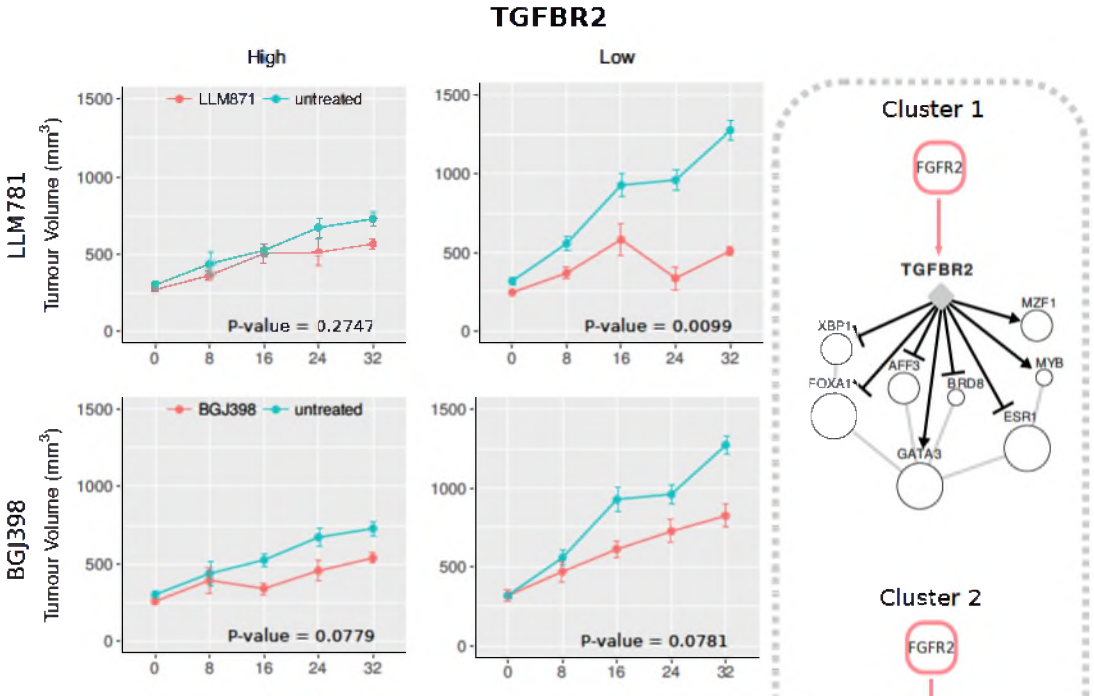
Supplementary Figure 6

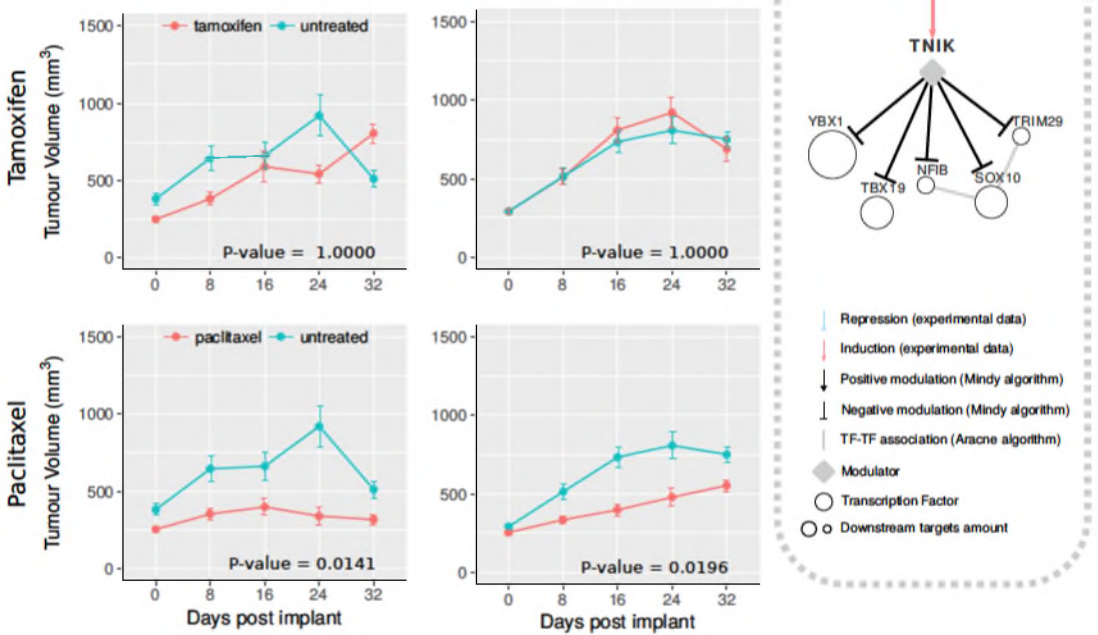




TGFBR2 Pannel. Breast cancer PDX samples of 4 different treatments were stratified according to TGFBR2 gene expression levels. Dotted box shows that FGFR2 activity induces TGFBR2 expression, which has mixed activity for cluster 1 and negatively modulates cluster 2 regulons.

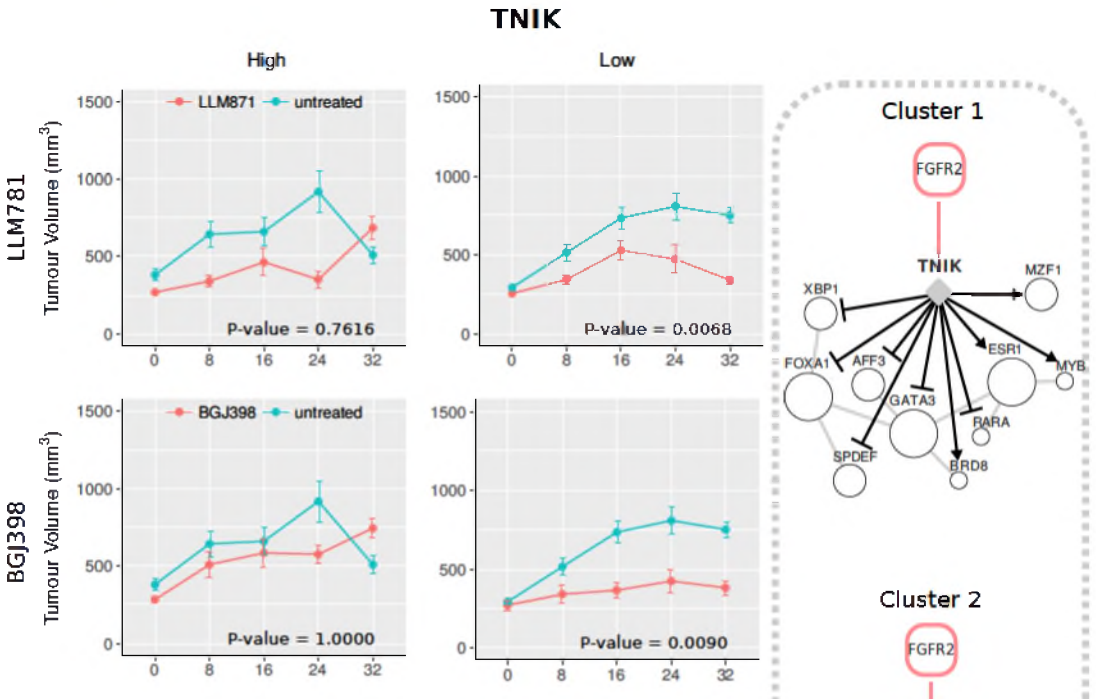
Supplementary Figure 7





TNIK Pannel. Breast cancer PDX samples of 4 different treatments were stratified according to TNIK gene expression levels. Dotted box shows that FGFR2 activity induces TNIK expression, which has mixed activity for cluster 1 and negatively modulates cluster 2 regulons.

Supplementary Figure 8



Supplementary Online Material

Tables

The constructed interactome tables "BodyFGFR2interactome.csv" and "BreastFGFR2interactome.csv" are available at Online Methods section.

Author's disclosure

All datasets were obtained from the databases mentioned at the section below. The group generated only the "Total" column in each interactome table, indicating the total amount of times a given interaction was found among the databases. Binary values inside these tables represent either presence (1) or absence (0) of a given interaction within the given database.

HUGO Gene Nomenclature Committee Multi Symbol Checker Tool

HGNC: `http://www.genenames.org/cgi-bin/symbol_checker`

Databases addresses

APID: `http://cicblade.dep.usal.es:8080/APID/init.action`
BioGrid: `https://thebiogrid.org/`
DIP: `http://dip.doe-mbi.ucla.edu/dip/Main.cgi`
STRING: `http://string-db.org/`
HPRD: `http://www.hprd.org/`
IntAct: `http://www.ebi.ac.uk/intact/`
Mint: `http://mint.bio.uniroma2.it/`
OPHID I2D: `http://ophid.utoronto.ca/ophidv2.204/`
Spectra: `https://alpha.dmi.unict.it/spectra/`
GIANT: `http://giant.princeton.edu/`
Mentha: `http://mentha.uniroma2.it/`
HumanBase: `http://hb.flatironinstitute.org/`
IID: `http://iid.ophid.utoronto.ca/SearchPPIs/protein/`

3 ARTIGO PUBLICADO NO PERIÓDICO CIENTÍFICO CANCER RESEARCH

Este capítulo mostra um estudo em colaboração com o Cancer Research UK, já publicado no jornal científico Cancer Research, distribuído pela Associação Americana de Pesquisa em Câncer. Neste estudo, abordagens experimentais foram combinadas com predições computacionais e demonstram a existência de um loop regulatório envolvendo NFIB, YBX1, ESR1, FOXA1 e FGFR2.

Tumor and Stem Cell Biology

Cancer Research

ER Alpha Binding by Transcription Factors NFIB and YBX1 Enables FGFR2 Signaling to Modulate Estrogen Responsiveness in Breast Cancer

Thomas M. Campbell[1], Mauro A. A. Castro[2], Kelin Gonçalves de Oliveira[2], Bruce A. J. Ponder[1], and Kerstin B. Meyer[1]

Abstract

Two opposing clusters of transcription factors (TF) have been associated with the differential risks of estrogen receptor positive or negative breast cancers, but the mechanisms underlying the opposing functions of the two clusters are undefined. In this study, we identified NFIB and YBX1 as novel interactors of the estrogen receptor (ESR1). NFIB and YBX1 are both risk TF associated with progression of ESR1-negative disease. Notably, they both interacted with the ESR1-FOXA1 complex and inhibited the transactivational potential of ESR1. Moreover, signaling through FGFR2, a known risk factor in breast cancer development, augmented these interactions and further repressed ESR1 target gene expression. We therefore show that members of two opposing clusters of risk associated with ESR1 positive and negative breast cancer can physically interact. We postulate that this interaction forms a toggle between two developmental pathways affected by FGFR2 signaling, possibly offering a junction to exploit therapeutically. *Cancer Res; 1–14.* ©2017 AACR.

Introduction

The estrogen receptor (ESR1) is the key driver and therapeutic target of breast cancer (1) and plays a critical role in determining the risk of developing this disease (2–4). Using a systems biology approach, we have examined transcriptional networks in breast cancer affecting ESR1 activity and have identified two distinct and opposing clusters of transcription factors (TFs) associated with enhanced breast cancer risk (5). The "cluster 1" risk TFs are associated with estrogen receptor-positive (ER^+) breast cancer risk and comprise TFs such as ESR1, FOXA1, and GATA3 whereas the "cluster 2" risk TFs appear to be associated with estrogen receptor-negative (ER^-), basal-like breast cancer (BLBC). Two of the TFs located in the cluster associated with ER^- disease are NFIB and YBX1. Here, we examine the molecular mechanisms underlying the opposing functions of the two groups of TFs by studying protein–protein interactions between TFs and their functional consequences. We also examine the effect of cell signaling, in particular by fibroblast growth factor receptor 2 (FGFR2), on the relative activity of the two groups of TFs.

The nuclear factor I (NFI) family of TFs consists of four members, NFIA, NFIB, NFIC, and NFIX, which can all bind DNA as homo- or heterodimers (6). They are particularly important during developmental stages (7, 8), and NFIB is crucial for normal lung and brain development (9). NFIB commonly has an increased copy number in small cell lung cancer, indicating a role as an oncogene (10). In BLBC, both copy number and expression levels of NFIB are also increased (11, 12). In addition, NFIB is important in the regulation of expression of mammary gland-specific genes, specifically those associated with lactation such as Whey acidic protein and α-lactalbumin (13). NFIB has been shown to modulate androgen receptor target genes in prostate cancer cells via an interaction with FOXA1 (14, 15). An investigation into whether similar modulation of estrogen receptor (ER) occurs in the breast has yet to be carried out.

Y-box binding protein 1 (YBX1) is a member of a family of DNA- and RNA-binding proteins with an evolutionarily ancient and conserved cold shock domain. It is a multifunctional protein that certainly does not follow the classical "one protein-one function" rule, but rather has disordered structure suggesting many different functions (16). It has been extensively studied in cancer, and its overexpression is associated with many hallmarks of the disease. It is expressed in many breast cancer cell lines regardless of subtype. However, there are higher levels of phosphorylated YBX1 in BLBC cell lines (17, 18). YBX1 expression in inversely correlated with ER, PR, and HER2 expressions and is positively correlated with the MAPK signaling cascade, a pathway important in BLBC (19, 20). YBX1 is highly expressed in 70% of BLBC cases and many of its target genes are associated with a basal-like signature (18, 20). Higher expression of YBX1 correlates with poor survival, drug resistance, and a high rate of relapse in all subtypes (18, 19, 21–23). Suppression of YBX1 reduces 2D cell growth and growth in mammospheres (18, 20). There is also evidence to suggest that YBX1 binds ESR1 in ER^+ breast cancer cell nuclei (24, 25).

[1]Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK. [2]Bioinformatics and Systems Biology Lab, Federal University of Paraná (UFPR), Polytechnic Center, Curitiba, Brazil. Current address for K.B. Meyer: Wellcome Trust Sanger Institute, Wellcome Genome Campus, Cambridge, UK.

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Corresponding Author: Kerstin B Meyer, University of Cambridge, Robinson Way, Cambridge CB2 0RE, UK. Phone: 44-1223-769651; Fax: 44-1223-769880; E-mail: kerstin.meyer@cruk.cam.ac.uk

doi: 10.1158/0008-5472.CAN-17-1153

A locus within the second intron of the FGFR2 gene is consistently identified as the genetic locus most strongly associated with ER^+ breast cancer risk by independent genome-wide association studies (GWAS; ref. 26). We have shown previously that the top three risk single nucleotide polymorphisms (SNP; refs. 27, 28) act to reduce FGFR2 gene expression and enhance the estrogen response (29). Increased FGFR2 stimulation repressed estrogen signaling in ER^+ breast cancer cell lines. However, the underlying molecular mechanism remains undear.

Here, we demonstrate that two members of the cluster 2 TFs, NFIB, and YBX1, both physically interact with ESR1, repress its activity, and drive breast cancer cells toward a less estrogen-dependent cancer phenotype. FGFR2 signaling augments this interaction and subsequent repression of ESR1 target gene expression. Our evidence suggests that FGFR2 has wide-ranging effects on driving breast cancer cells toward a more basal-like phenotype and that inhibiting FGFR2 signaling in ER^+ breast cancer sensitizes cells to antiestrogen therapies.

Materials and Methods

Cell culture

MCF-7 human breast cancer cells and HeLa cells were cultured in DMEM (Invitrogen) supplemented with 10% FBS and antibiotics. ZR751 human breast cancer cells were cultured in RPMI (Invitrogen) supplemented with 10% FBS and antibiotics. SUM52PE human breast cancer cells were cultured in Ham/F-12 (Invitrogen) supplemented with 10% FBS, 5 μg/mL insulin, 1 μg/mL hydrocortisone and antibiotics. All cells were maintained at 37°C, 5% CO_2, obtained from the CRUK Cambridge Institute collection and authenticated by STR genotyping.

Quantitative RT-PCR

1 μg of total RNA was reverse transcribed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) and qRT-PCR performed using cDNA obtained from 10 ng of total RNA. qRT-PCR was performed using a QuantStudio6 system (Life Technologies). Amplification and detection were carried out in 384-well Optical Reaction Plates (Applied Biosystems) with Power SYBR Green Fast 2× qRT-PCR Mastermix (Applied Biosystems). All expression data were normalized to DGUOK expression. The specificity of primers (Supplementary Table 1) was confirmed through generation of single peaks in a melt-curve analysis. Data analysis was performed using the $2^{-\Delta\Delta CT}$ method (30).

Western immunoblotting

Cells were grown in 10 cm Petri dishes, washed in PBS, and lysed on ice in RIPA buffer with complete Mini EDTA-free protease inhibitor cocktail (Roche). Resulting cell lysates were passed through a fine-gauge syringe needle several times, centrifuged at 10,000 g for 1 minute and left at −80°C at least overnight. Protein samples were separated by SDS-PAGE using 4% to 12% Bis-Tris gels (Novex) for 2.5 hours (30 minutes at 60 V, 120 minutes at 120 V) and transferred by electrophoresis using an iBlot (Novex) for 7 minutes onto a nitrocellulose membrane (iBlot Gel Transfer Stacks; Novex). Successful transfer of protein was confirmed using Ponceau S Solution (Sigma). Membranes were "blocked" at room temperature for 1 hour with 5% dried milk in Tris-buffered saline (TBS) with 0.1% Tween-20 (TTBS), washed 3× with TTBS and probed with the relevant primary antibody (Supplementary Table S2) in blocking solution at 4°C overnight. Membranes were then rewashed with TTBS 3× and incubated with appropriate HRP-conjugated secondary antibody (Supplementary Table S2) in blocking solution at room temperature for 90 minutes. Following further washing with TTBS, blots were treated with SuperSignal West Chemiluminescent Substrate (Thermo Scientific) and immunoreactive proteins detected by exposure to film (FUJIFILM).

Rapid Immunoprecipitation Mass Spectrometry of Endogenous Proteins (RIME)

Rapid Immunoprecipitation Mass Spectrometry of Endogenous Proteins was performed on the ESR1 protein (ERα) in MCF-7 and ZR751 ER^+ breast cancer cells, as described previously (31–33). Briefly, cells were crosslinked for 8 minutes at room temperature in media containing 1% formaldehyde. Crosslinking was quenched by adding glycine to a final concentration of 0.2 M. Cells were washed with ice-cold PBS, harvested in PBS, and the resulting cell pellet was washed in PBS. The nuclear fraction was extracted from the samples by first suspending the pellet in 10 mL LB1 buffer (50 mmol/L HEPES-KOH pH7.5, 140 mmol/L NaCl, 1 mmol/L EDTA, 10% glycerol, 0.5% NP-40, and 0.25% Triton X-100) for 10 minutes at 4°C. Cells were then pelleted, resuspended in 10 mL LB2 buffer (10 mmol/L Tris-HCl pH8.0, 200 mmol/L NaCl, 1 mmol/L EDTA, and 0.5 mmol/L EGTA) and mixed at 4°C for 5 minutes. Cells were then pelleted and resuspended in 300 μL of LB3 buffer (10 mmol/L Tris-HCl pH8.0, 100 mmol/L NaCl, 1 mmol/L EDTA, 0.5 mmol/L EGTA, 0.1% Na-deoxycholate, and 0.5% N-lauroylsarcosine) and sonicated in a water bath sonicator (Diagenode). The resulting supernatant was incubated with protein A Dynabeads (Invitrogen) prebound with ESR1 antibody (Santa Cruz sc-543 X), and immunoprecipitation (IP)was performed at 4°C overnight. The beads were washed 10× in RIPA buffer and twice in 100 mmol/L AMBIC solution. Tryptic digestion of bead-bound protein and mass spectrometry was performed by the Proteomics Core Facility at The CRUK Cambridge Institute using an LTQ Velos-Orbitrap MS (Thermo Scientific) coupled to an Ultimate RSLCnano-LC system (Dionex). Full RIME data are given in Supplementary Table S3.

Table 1. RIME analysis shows ESR1 binds to NFIB and YBX1

	E2		E2 + FGF10		IgG	
	ZR751	MCF-7	ZR751	MCF-7	ZR751	MCF-7
ESR1	8	8	7	7	0	0
GATA3	3	3	2	3	0	0
FOXA1	1	1	1	0	0	0
NFIB	1	0	1	2	0	0
YBX1	2	4	4	7	0	1

NOTE: Values indicate the number of unique peptides identified by MS for the TFs listed in the left column, in ZR751 and MCF-7 ER^+ breast cancer cells, following nuclear immunoprecipitation with an ESR1 antibody after treatment with 1 nmol/L E2 or 1 nmol/L E2 plus 100 ng/mL FGF10 (E2 + FGF10) for 90 minutes, or with an IgG control antibody after E2 treatment.

Coimmunoprecipitation

Cells from five 15 cm Petri dishes were harvested after washing with PBS. The cellular nuclear fraction was then obtained using a nuclear extraction kit (Affymetrix), according to manufacturer's protocol. The resulting nuclear fraction was precleared for 60 minutes with protein A Dynabeads (Invitrogen). IP was then performed with 5 μg of antibody prebound to protein A Dynabeads. Each IP was coupled with a corresponding IgG^- control of the same species. IP was performed overnight and the beads were washed with wash buffer (50 mmol/L Tris pH7.4, 140 nmol/L NaCl, 2 mmol/L EGTA, and 0.1% Tween-20). Beads were then boiled at 95°C for 15 minutes in LDS loading buffer and Western immunoblot analysis performed.

Molecular Cloning

The plasmid constructs used for the fluorescence resonance energy transfer (FRET) were developed as follows from mCerulean3 (mCer3)-C1 and mVenus-C1 vectors kindly donated by Magdalena Grabowska (14). The ESR1-Cerulean construct was created by amplifying the gene encoding ESR1 (from RC213277; OriGene) and performing sequential digestion/ligation of the product and mCer3-C1 vector using NheI and AgeI restriction enzymes. The NFIB/YBX1-Cerulean and NFIB/YBX1-Venus constructs were created similarly (from plasmids RC231275 (NFIB) and RC209835 (YBX1); OriGene). The FOXA1-Venus construct was kindly donated by Magdalena Grabowska (14). All primer sequences are given in Supplementary Table S4. The orientation and sequence of all plasmids were confirmed by DNA sequencing (GATC Biotech).

FRET HeLa cells were transiently transfected with plasmid DNA encoding the tagged TFs described above. 15,000 cells were seeded into each well of a μ-Slide 8 Well-chambered coverslip (ibidi) and cultured for 24 to 48 hours. Samples were then fixed with 4% paraformaldehyde for 20 minutes at room temperature, washed in PBS, and stored in PBS. FRET imaging was performed using a Leica TCS SP5 confocal microscope (Leica Microsystems). Data were analyzed by FRET Acceptor Photobleaching (34) using the Leica LAS imaging software (Leica Microsystems). A total of 20 to 30 cells/well were quantified for FRET efficiency, and the experiments were repeated in at least three cellular preparations. FRET efficiency was calculated as follows:

$$Efficiency = \frac{Donor_{post-bleach} - Donor_{pre-bleach}}{Donor_{post-bleach}}$$

Luciferase reporter assay

MCF-7 cells stably expressing a luciferase reporter gene under the transcriptional control of an upstream ESR1 and FOXA1 binding site, cloned from the human RARα gene (kindly donated by the lab of Jason Carroll), were plated at 50,000 cells/well in 24-well dishes and left in complete medium until 50% to 70% confluent. Cells were then transfected with the relevant siRNA/expression plasmids and a β-galactosidase construct using FuGENE HD Transfection Reagent (Promega), according to manufacturer's protocol (DNA:FuGENE ratio = 1 μg:4 μL). After 24 hours at 37°C, 5% CO_2, cells were lysed with Reporter Lysis Buffer (Promega) and luciferase, and β-galactosidase assays were performed on a PHERAstar FS Microplate Reader (BMG LABTECH) using the appropriate assay kits (Promega), according to manufacturer's protocol. Each assay was performed in triplicate and a total of three assays were performed on three separate days.

Transient transfection of siRNA

MCF-7 cells were transfected with ON-TARGETplus SMARTpool siRNA (Dharmacon) directed against ESR1 (L-003401-00), FOXA1 (L-010319-00), NFIB (L-008456-00), YBX1 (L-010213-00), FGFR2 (L-003132-00), and a control nontargeting pool (D-001810-10) using Lipofectamine RNAiMax Reagent (Invitrogen), according to manufacturer's protocol. Following addition of the transfection complexes, cells were incubated at 37°C, 5% CO_2 for at least 24 hours before experiments were performed.

Transient transfection of plasmid DNA

Cells were plated at 50,000 cells/well in 24-well dishes and grown in complete medium until 50% to 70% confluent, transiently transfected with plasmid using FuGENE HD Transfection Reagent (Promega), according to manufacturer's protocol (DNA: FuGENE ratio = 1 μg:4 μL), and maintained for 24 to 48 hours at 37°C, 5% CO_2 in complete medium prior to conducting experiments.

Generation of stable cell lines

MCF-7 cells stably expressing FLAG-tagged NFIB (RC231275; OriGene) and YBX1 (RC209835; OriGene) were generated via transfection of the NFIB and YBX1 constructs, as described above. The day following cell transfection, cell culture medium was changed to fresh medium containing 1.5 mg/mL geneticin (G418; Invitrogen). Cells were grown and passaged, with media changed every other day until mass cell death was observed. Clonal populations of cells were selected by transferring well-isolated single dumps of cells into a 24-well plate. Cells were expanded under antibiotic selection.

Proliferation assay

Cells were plated at 4,000 cells per well into 96-well plates and cell numbers monitored in real time by *in vitro* microimaging using an IncuCyte incubator (Essen BioScience), allowing for monitoring of cell proliferation by observing cell confluence. Images were taken every 3 hours and data consisted of an average of four separate images taken for each well. Assays were performed in eight separate wells on three separate occasions.

RNA collection and RNA sequencing

Total RNA was extracted from cells using the miRNeasy Mini Kit (QIAGEN) and quality checked using an RNA 6000 Nano Chip on a 2100 Bioanalyzer (Agilent). mRNA-seq libraries were prepared from three biological replicates of each stable overexpression system using the TruSeq Stranded mRNA Library Prep Kit (Illumina), according to manufacturer's protocol. Single-end 50 bp reads generated on the Illumina HiSeq 4000 were aligned to the human genome version GRCh37.75. Read counts were then obtained using Subread v1.5.1 (35), normalized and tested for differential gene expression using the Bioconductor package DESeq2 (36, 37). Multiple testing correction was applied using the Benjamini–Hochberg method. The full mRNA-seq data set has been deposited in GEO under accession GSE95299.

Two-tailed gene set enrichment analysis (GSEA)

Two-tailed gene set enrichment analysis (GSEA; ref. 38) was performed as described previously (29). *P* values derived from DESeq analyses of the mRNA-seq data were $-\log_{10}$ transformed and then signed according to whether genes were up- or down-regulated compared with control samples. These values were then

used for ranking and weighting of genes in subsequent GSEA analyses.

Survival analysis

Analysis of breast cancer patient survival stratified by YBX1 expression was carried out using the KM plotter (39).

Patient-derived xenograft (PDX) analysis

A subset of breast cancer samples from Novartis' PDX dataset (40) was stratified according to YBX1, NFIB, and FGFR2 expression levels. Clinical tamoxifen response was assessed by comparison of tumor volume between treated versus untreated groups. *P* values were generated with Repeated Measures one-way ANOVA (RM-ANOVA) statistical test. ESR1 gene expression levels were also compared between groups using Kruskal–Wallis.

Stimulation of FGFR2 signaling

Cells in which FGFR2 signaling was stimulated were first left in complete medium overnight. Cell synchronization via estrogen-starvation was then carried out for 3 days in estrogen-free media (phenol red-free media supplemented with 5% charcoal dextran-treated FBS and 2 mmol/L L-glutamine), with media changes every 24 hours. Estrogen-deprived cells were stimulated with 1 nmol/L β-estradiol (E2; Sigma) or 100 ng/mL FGF10 (Invitrogen) in combination with 1 nmol/L E2, for 6 hours.

Results

ESR1 interacts with NFIB and YBX1

Previously, we have shown that FGFR2 signaling reduces estrogen responsiveness in breast cancer cells (29) but has little effect on ESR1 expression levels. We therefore tested whether FGFR2 signaling affects the interaction of ESR1 with its protein-binding partners. To this end, we performed a RIME analysis on the ESR1 protein (Table 1). Unique peptides for ESR1, as well as its known binding partners, FOXA1 and GATA3, were detected. The only other cluster 1 or 2 TFs for which unique peptides were detected in the RIME analysis were NFIB and YBX1. YBX1 has previously been reported to interact with ESR1 (24, 25), whereas NFIB appears to be a novel interacting partner. RIME cannot be considered a truly quantitative technique. Nevertheless, the number of unique peptides for NFIB and YBX1 detected by mass spectrometry increases when both MCF-7 and ZR751 are stimulated with FGF10, the most potent agonist of the FGFR2 receptor (41, 42). This suggests that FGFR2 signaling in ER^+ breast cancer cell lines might augment the interaction of ESR1 with the two cluster 2 risk TFs.

To confirm the exploratory RIME experiments, coimmunoprecipitation experiments were performed in order to test if NFIB and YBX1 could be confirmed as ESR1 binding partners by Western immunoblotting. Following IP of the nuclear fraction of both MCF-7 and ZR751 cells with an ESR1 antibody (Fig. 1A), ESR1, FOXA1, and GATA3 protein bands could be resolved by Western immunoblotting, as expected. Moreover, NFIB and YBX1 were also present in the ESR1 immunoprecipitates, while being absent in the IgG control pull downs, suggesting that both NFIB and YBX1 physically interact with the ESR1 protein in the nucleus of these ER^+ breast cancer cells. As control experiments, blots were also performed for TFs that are not expected to bind to ESR1 (E2F2, SP1, and YY1), and no protein bands were detected. The inverse pull-down experiments were also performed, in which the nuclear fractions of MCF-7 and ZR751 cells were immunoprecipitated with an NFIB (Fig. 1B) or YBX1 (Fig. 1C) antibody. In both cases, the ESR1 protein was detected in the immunoprecipitate.

RIME data suggested that FGFR2 signaling in MCF-7 and ZR751 cells might increase the association of ESR1 with both NFIB and YBX1. Therefore, coimmunoprecipitation experiments were also carried out in MCF-7 cells that had been stimulated with estrogen alone or with a combination of estrogen and FGF10 (Fig. 1D and E). Densitometry analysis of the Western immunoblots against NFIB and YBX1 following pull down of ESR1 shows that stimulation of MCF-7 cells with FGF10 appears to augment the interaction of the two ER^- risk TFs with ESR1, without affecting protein levels (Fig. 1F and G). Moreover, FGFR2 signaling in ER^+ breast cancer cells increases the level of phosphorylated YBX1 (demonstrated in MCF-7 cells), while FGFR2 inhibition reduces it (demonstrated in SUM52PE cells, which carry an FGFR2 gene amplification; Fig. 1H and I). Our finding that YBX1 can bind to ESR1 is consistent with recent reports of an interaction between these two proteins (24, 25).

FRET, which is facilitated by tagging proteins of interest with fluorescent proteins as reporters, is an imaging technique useful for studying protein– interactions (43). FRET only occurs when the fluorescent proteins are within very close proximity of each other (<10 nm), thereby allowing for the measurement of the proximity of proteins of interest (Fig. 2A). Here, we tagged FOXA1, NFIB, YBX1, and ESR1 with either a donor (mCerulean3) or acceptor (mVenus) fluorescent protein and performed FRET in HeLa cells expressing the constructs (Fig. 2B; Supplementary Fig. S1). Consistent with previous reports of ESR1 and FOXA1 interactions (44), cotransfected ESR1-Cer and FOXA1-Venus emitted a strong FRET signal (Fig. 2B) with an efficiency of 0.139 ± 0.011 (Supplementary Fig. S2). To determine whether NFIB and FOXA1 are also able to interact directly, cells were cotransfected with NFIB-Cer donor and FOXA1-Venus acceptor constructs. The pairing resulted in a positive FRET signal with a FRET efficiency of 0.055 ± 0.007. On the other hand, the ESR1-Cer and NFIB-Venus pairing did not result in FRET (efficiency of 0), suggesting that these proteins do not interact directly. To test the hypothesis that FOXA1 can bridge the interaction between ESR1 and NFIB, we cotransfected cells with ESR1-Cer, NFIB-Venus, and untagged FOXA1. The FRET efficiency of ESR1-Cer and NFIB-Venus was increased to 0.018 ± 0.003. This result suggests that FOXA1 serves as an intermediary "bridge" to bring ESR1 and NFIB together. The same experiments were carried out with YBX1 FRET constructs, demonstrating that YBX1 is able to bind to ESR1 directly, without requiring FOXA1 (Fig. 2B; Supplementary Fig. S2).

NFIB and YBX1 suppress ESR1 activity

Having established that both NFIB and YBX1 interact with the ESR1/FOXA1 TF complex, we asked whether NFIB and YBX1 are able to influence the transcriptional activity of ESR1. When NFIB or YBX1 were transiently overexpressed in MCF-7 cells, expression of the ESR1-target gene, pS2, was significantly reduced compared with the control cells (Fig. 3A). Conversely, reduction of NFIB or YBX1 levels via siRNA transfection resulted in increased pS2 expression. The same results were also obtained for other ESR1-target genes (Supplementary Fig. S3). Similarly, when NFIB or YBX1 were transiently overexpressed in MCF-7 cells stably expressing a luciferase reporter gene under the transcriptional control of an upstream ESR1/FOXA1 binding site, luciferase

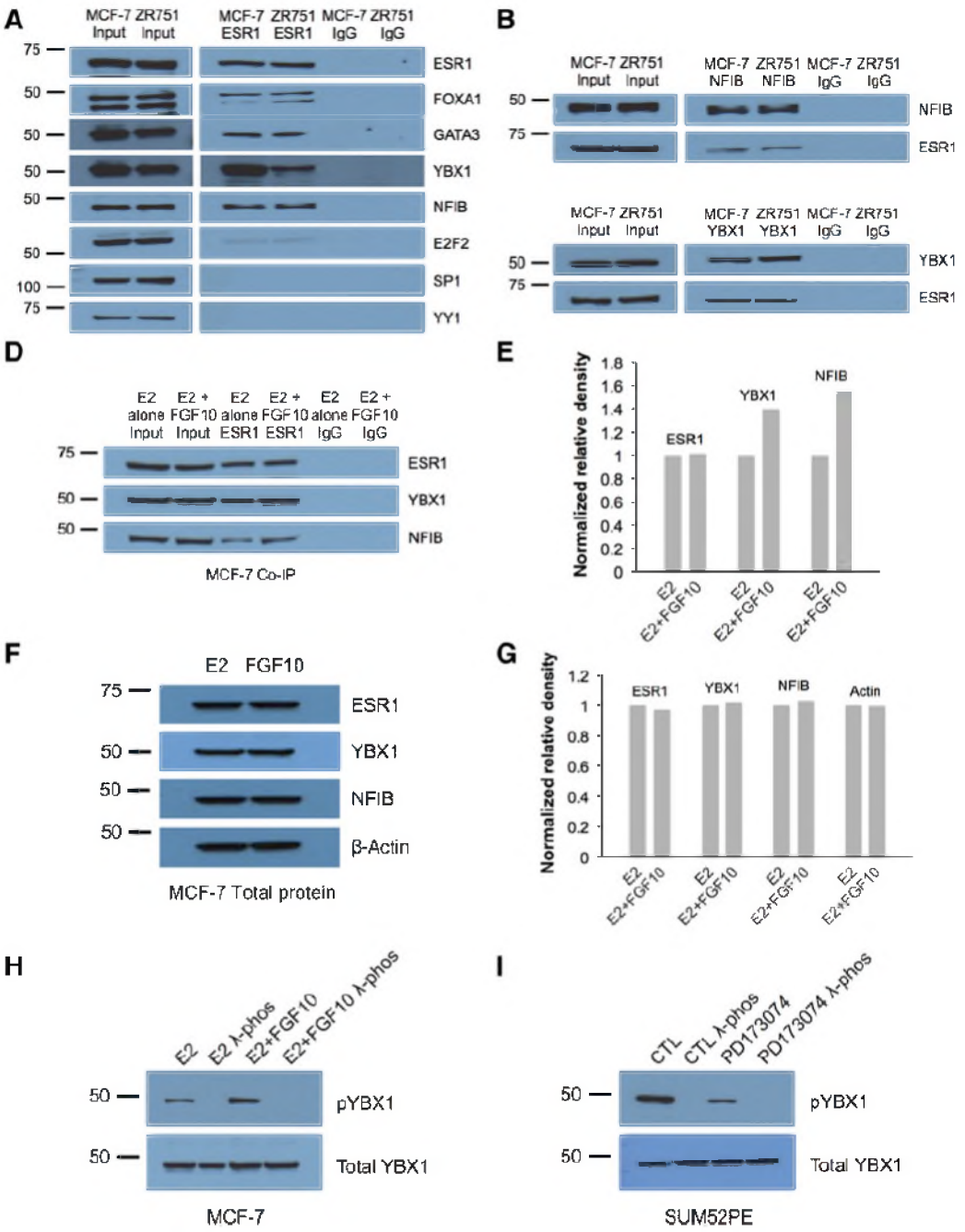


Figure 1.

ESR1 protein binds to the TFs NFIB and YBX1 in ER$^+$ breast cancer cells. **A–C,** Coimmunoprecipitation assays were performed in MCF-7 and ZR751 cells, as indicated. Antibodies used in each immunoprecipitation are shown above the panels, antibodies used to develop Western immunoblots to the right of each panel. **D,** Coimmunoprecipitation assays carried out in MCF-7 cells following treatment of the cells with 1 nmol/L E2 or 1 nmol/L E2 plus 100 ng/mL FGF10 for 90 minutes. **E,** Densitometry analysis of the Western immunoblots displayed in **D**. **F,** Representative Western immunoblots showing expression of ESR1, NFIB, YBX1, and β-actin proteins in MCF-7 cells following treatment of the cells with E2 or E2 plus FGF10 (as above). **G,** Densitometry analysis of the Western immunoblots displayed in **F**. **H,** Representative Western immunoblots showing expression of phosphorylated YBX1 (pYBX1) and total YBX1 in MCF-7 cells following treatment of the cells with E2 or E2 plus FGF10 (as above). **I,** Representative Western immunoblots showing expression of pYBX1 and total YBX1 in SUM52PE cells (carrying an FGFR2 gene amplification) following treatment of the cells with 100 ng/mL PD173074 (FGFR inhibitor) for 90 minutes. λ-phosphatase treatment of cell lysates was performed to demonstrate antibody specificity for phosphorylated (Ser102) YBX1. $n = 3$ for all blots.

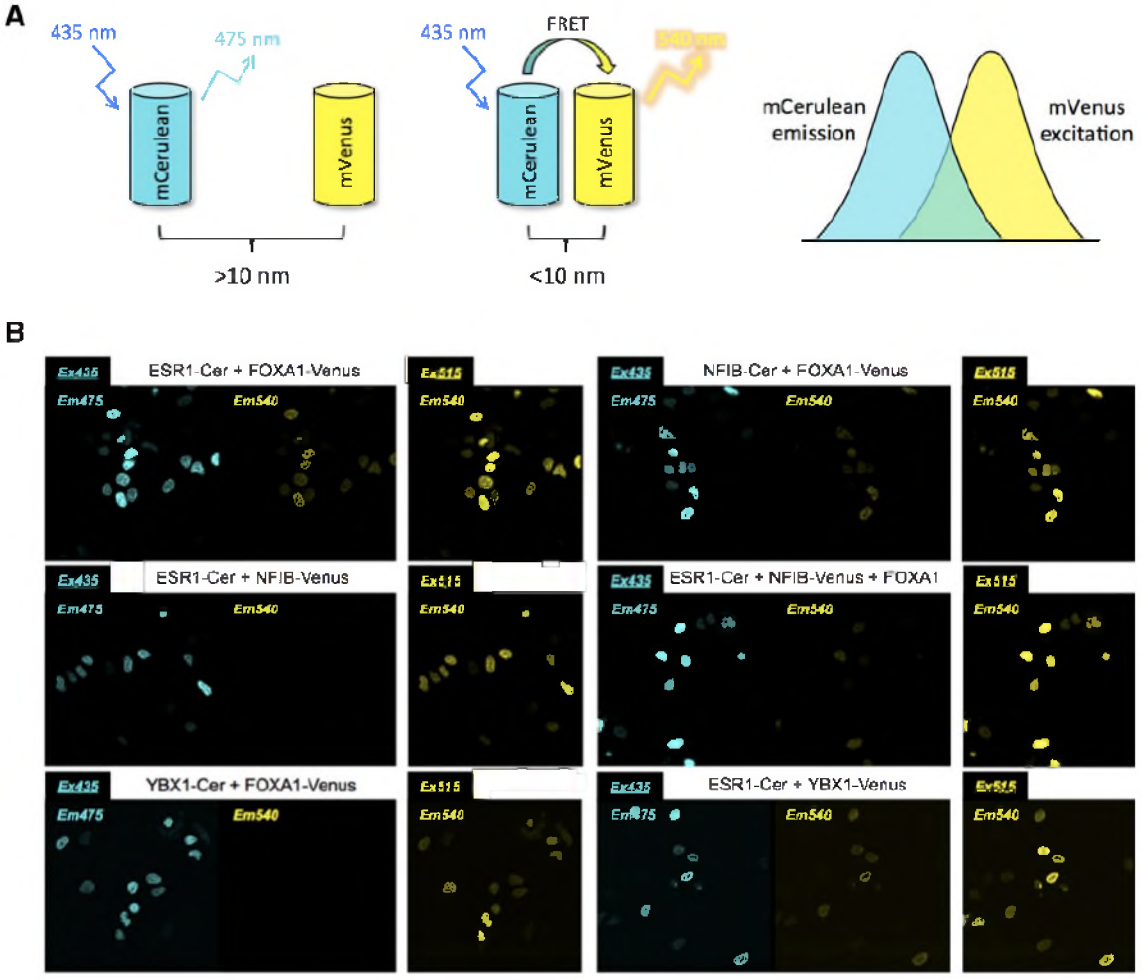


Figure 2.

FRET demonstrates protein–protein interactions between ESR1 and NFIB/YBX1. **A,** Schematic showing how FRET microscopy works. If the FRET donor and acceptor fluorophores are >10 nm apart no FRET occurs and donor fluorescence is observed. If the FRET donor and acceptor fluorophores are within ~10 nm of one another, then energy transfer can occur from the donor to the acceptor. After excitation at 435 nm (Cerulean excitation), fluorescence at 540 nm (Venus emission) is only observed if the two FRET fluorophores are in very close proximity to one another (<10 nm), owing to the spectral overlap of the two fluorophores. **B,** Representative images of FRET in HeLa cells transfected with FRET constructs, as listed above the panels. As expected, ESR1 and FOXA1 interact. FOXA1 also interacts with NFIB and facilitates the association of NFIB with ESR1. YBX1 interacts directly with ESR1 without interacting with FOXA1.

expression was significantly reduced compared with control cells (Fig. 3B). These data suggest that both NFIB and YBX1 are able to inhibit ESR1-mediated transcriptional activity.

To investigate further the possible role of NFIB and YBX1 on ESR1 activity, MCF-7 cell lines stably overexpressing FLAG-tagged NFIB or YBX1 were generated (Supplementary Fig. S4). For each TF, three independent clones were expanded, mRNA-seq data generated, and the regulatory network examined. We previously defined regulons (set of target genes) for all TFs by measuring the similarities in gene expression patterns of the TF of interest and all possible target genes in gene expression data from breast tumor samples (5). Here, we carried out a two-tailed GSEA (5, 29) to assay the activity of the ESR1 regulon in the stably transfected cells. As a control, we show the behavior of the ESR1 regulon in response to estrogen stimulation. As expected, positive targets of ESR1 are induced and negative targets of ESR1 are repressed in the parental MCF-7 cells (Fig. 4A). Overexpression of both NFIB and YBX1 leads to a relative repression of the ESR1 regulon (Fig. 4B and C), with negative ESR1 targets being upregulated and positive targets showing lower expression. These experiments confirm that both NFIB and YBX1 are able to inhibit ESR1 function.

When the MCF-7 cells stably overexpressing NFIB or YBX1 were estrogen starved, they were able to proliferate faster than estrogen-starved parental MCF-7 cells (Fig. 4D and E). A study by Shibata and colleagues reported that YBX1 is able to reduce the stability of ESR1 protein (25). However, Western immunoblots of cell extracts demonstrate that full-length ESR1 protein levels are not altered by either NFIB or YBX1 overexpression in our system

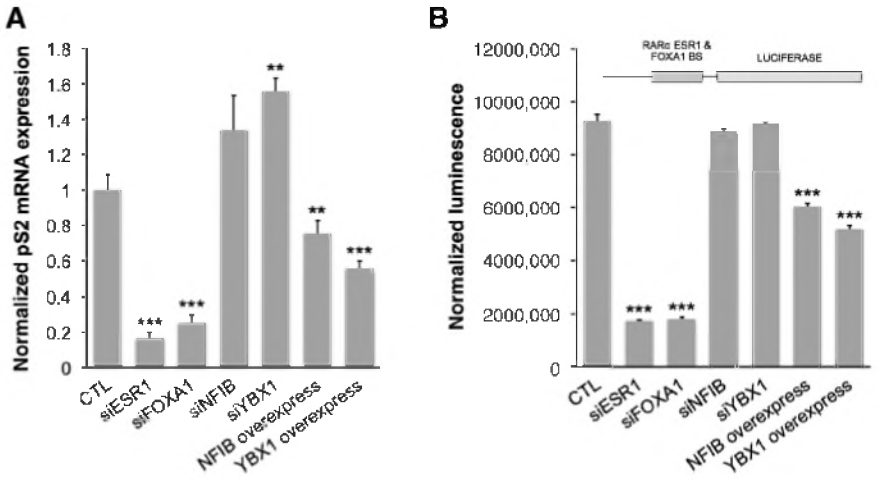


Figure 3.

NFIB and YBX1 repress the transcriptional activity of ESR1. **A,** Relative mRNA expression of the ESR1-target gene, pS2, in MCF-7 breast cancer cells following transfection with siRNA directed against ESR1, FOXA1, NFIB, and YBX1, and with plasmids overexpressing NFIB and YBX1, compared with a scrambled control siRNA transfection. All data were normalized to DGUOK expression ($n = 10$, two separate experiments, $P < 0.01$ (**); $P < 0.001$ (***), one-way ANOVA and SNK correction; error bars, SEM). **B,** Luciferase luminescence in MCF-7 cells stably expressing a luciferase reporter gene under the transcriptional control of an upstream ESR1/FOXA1 binding site, cloned from the human RARα gene, 24 hours posttransfection with siRNA directed against ESR1, FOXA1, NFIB, and YBX1, and with plasmids overexpressing NFIB and YBX1, compared with a scrambled control siRNA transfection, normalized to β-galactosidase expression ($n = 9$, three separate experiments, $P < 0.001$ (***), one-way ANOVA and SNK correction; error bars, SEM). Inset, schematic depiction of the stably expressed reporter construct used in the luciferase reporter assays.

(Supplementary Fig. S4). Our results suggest that overexpression of these cluster 2 risk TFs is able to drive ER$^+$ breast cancer cells toward a more ER$^-$, basal-like cancer phenotype in which estrogen dependency is reduced.

FGFR2 signaling and breast cancer regulon activity

To further assess the shift from luminal to a more basal-like phenotype, we extended our two-tailed GSEA to all regulons and visualized the results in a tree and leaf diagram, where regulons are represented as leaves, and the branching between them is a measure of their relatedness (5). Using this approach, a gene signature derived from ER$^+$ versus ER$^-$ tumors showed a positive enrichment in the regulons of cluster 1 risk TFs and a negative enrichment of cluster 2 risk TFs (Fig. 5A). A basal gene signature showed the inverse (Fig. 5B). Interestingly, we found that a FGFR2 signaling gene signature was able to activate the NFIB and YBX1 regulons, as well as almost all TF regulons that are associated with ER$^-$ disease (Fig. 5C), mimicking very closely the results obtained with the basal gene signature. A reduction of FGFR2 gene expression via siRNA transfection has the opposite effect, increasing the activity of ESR1 and other cluster 1 TFs (Fig. 5D), supporting and extending our earlier findings that FGFR2 signaling opposes estrogen signaling.

The fact that FGFR2 signaling inhibits estrogen signaling in ER$^+$ breast cancer cells, possibly via an increased association of ESR1 with the ER$^-$ risk TFs, NFIB, and YBX1, led us to test the hypothesis that the inhibition of FGFR2 signaling in ER$^+$ breast cancer cells sensitizes cells to antiestrogen therapies. When three different ER$^+$ breast cancer cell lines (MCF-7, ZR751, and T47D), which all express NFIB and YBX1 (Fig. 6A), are treated with the FGFR2 inhibitors, AZD4547 and PD173074, their growth, as measured in an IncuCyte incubator, is more sensitive to the antiestrogen tamoxifen (Fig. 6B–D; Supplementary Fig. S5). This suggests that anti-FGFR2 treatments make breast cancer cells more reliant on estrogen signaling for growth and could therefore be used in combination with antiestrogen therapies to treat breast cancer. When MCF-7 cells stably overexpressing either NFIB or YBX1 are treated with siRNA against NFIB/YBX1, they become significantly less sensitive to the combined drug treatment when compared with nontransfected control cells (Fig. 6E and F; Supplementary Fig. S5), suggesting that NFIB and YBX1 do indeed play an important role in the FGFR2-driven estrogen activity/sensitivity of breast cancer cells. Much more work is needed to determine if the effect of FGFR2 signaling on a breast cancer cell's reliance on estrogen signaling is primarily mediated by NFIB and YBX1. However, it is interesting to note that overexpression of YBX1 in breast cancer is associated with poorer survival, even when tested just in ER$^+$ breast cancer (Supplementary Fig. S6). Furthermore, in PDX models of breast tumors (40), we find that tamoxifen treatment is not very effective in PDXs with high YBX1 expression (Supplementary Fig. S7), although this group is likely to contain ER$^-$ tumors. In contrast, tamoxifen efficacy is greater in YBX1-low PDXs and even greater in PDXs that express both low FGFR2 and YBX1, further supporting the notion that inhibition of FGFR2 may increase a tumor's response to tamoxifen.

Discussion

In this study, we demonstrate that in ER$^+$ breast cancer the TFs NFIB and YBX1 interact with ESR1, the key driver of luminal breast cancer. We examine the functional consequences of this interaction and find that NFIB and YBX1 are each able to repress transcriptional activation by ESR1. This can be observed in

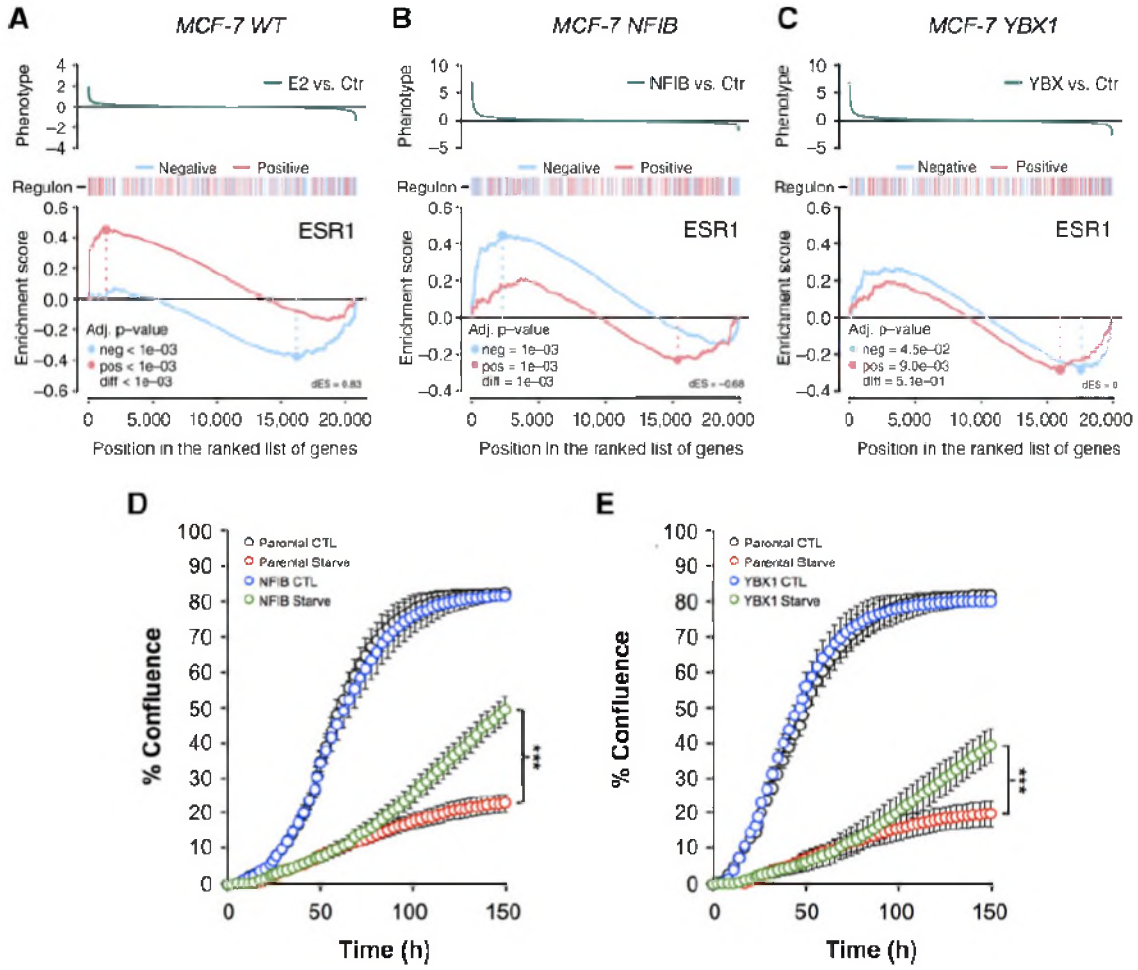


Figure 4.

Effect of overexpressing NFIB and YBX1 on the estrogen response. **A–C,** GSEA of the ESR1 regulon using gene signatures derived from starved and estrogen-stimulated parental MCF-7 cells **(A)** and MCF-7 cells stably overexpressing NFIB **(B)** or YBX1 **(C)**. **D** and **E,** Effect of NFIB **(D)** and YBX1 **(E)** overexpression on MCF-7 cell proliferation in the absence of estrogen ($n = 16$, two separate experiments, $P < 0.001$ (***), one-way ANOVA and SNK correction; error bars, SEM).

reporter assays, at the level of endogenous estrogen-regulated genes such as pS2, and also in the reduction of the overall activity of the ESR1 regulon. The interaction between YBX1 and ESR1 is direct, while NFIB requires FOXA1 as a bridging protein that allows the interaction. The complex formation we observe between NFIB or YBX1 and ESR1 may explain the opposing action that NFIB/YBX1 and ESR1 have on shared target genes (5). In addition to repressing ESR1, NFIB and YBX1 are also able to drive proliferation: while proliferation of parental MCF-7 cells is strictly dependent on the presence of estrogen and hence nuclear ESR1, MCF-7 cells overexpressing either NFIB or YBX1 are able to grow in estrogen-depleted medium.

To date, NFIB and YBX1 have primarily been associated with ER$^-$ breast cancer, where both factors contribute to increased aggressiveness and metastatic potential (12, 45). We now report that these two TFs repress ESR1 activity, suggesting that they may play a similar role in ER$^+$ breast cancer. Although ER$^+$ breast cancer has better patient outcomes, in large part driven by the effectiveness of hormone deprivation therapy, relapse and resistance to therapy are relatively common and can occur many years after the primary tumor was diagnosed and treated (46). Our previous work suggests that patient outcomes are strongly affected by the relative activity of TFs driving ER$^+$ (cluster 1, e.g., ESR1, GATA3, and FOXA1) versus ER$^-$ disease (cluster 2, e.g., YBX1 and NFIB). We found that, in an ER$^+$ patient cohort, patients with a repressed ESR1 regulon have worse prognosis (5). We now show that NFIB and YBX1 can both function to repress the activity of the ESR1 regulon. In line with this observation, we found that in clinical samples from patients with ER$^+$ disease, higher YBX1 expression is associated with reduced survival. As a corollary, interventions that increase the activity of the ESR1 regulon may improve patient outcomes, since the tumor is likely to have increased sensitivity to estrogen deprivation therapy.

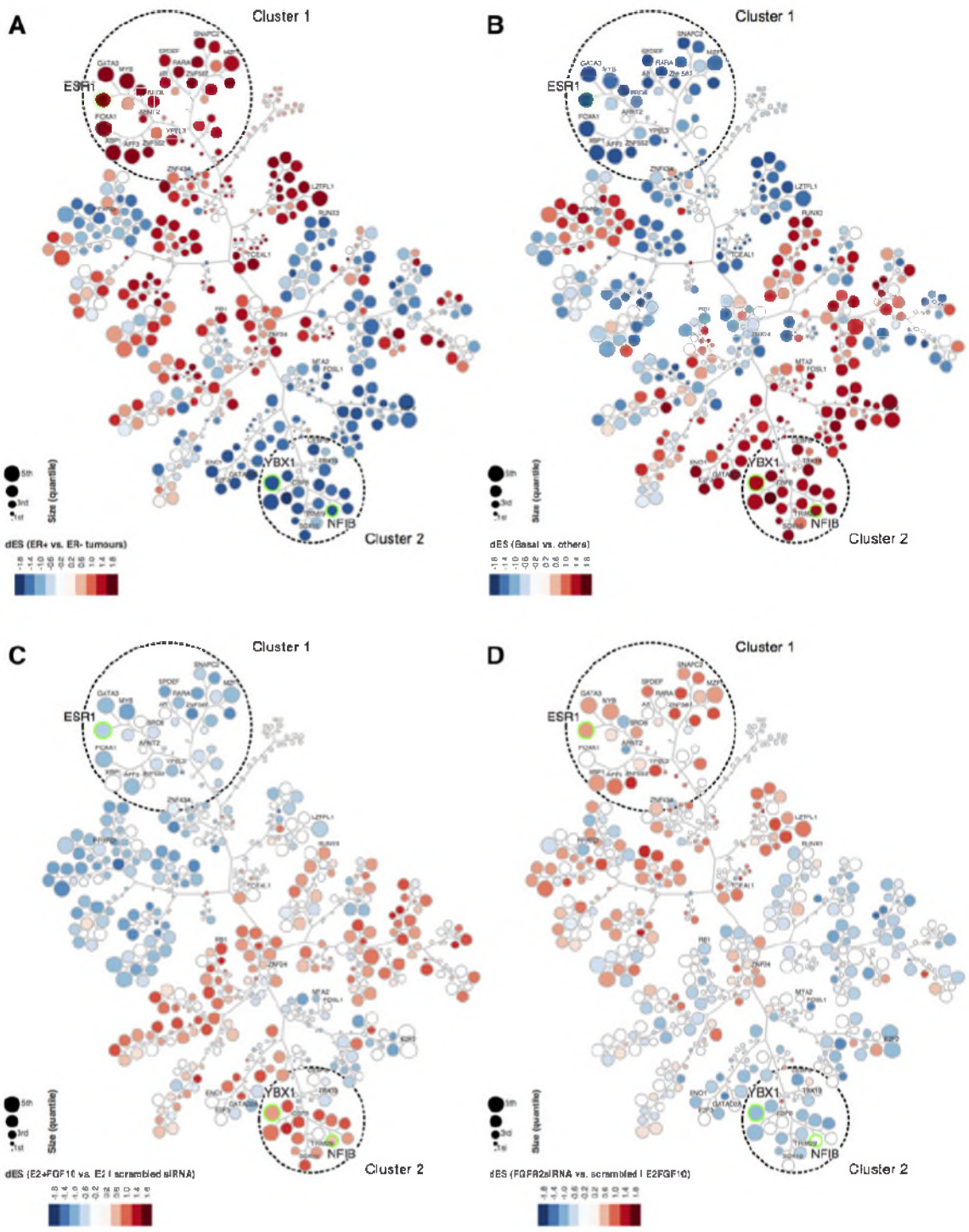


Figure 5.

FGFR2 signaling and breast cancer regulon activity. Tree-and-leaf representations of breast cancer regulon activity in ER^+ versus ER^- tumors **(A)**, in basal versus nonbasal tumors **(B)**, in MCF-7 cells stimulated with FGF10 versus nontreated cells **(C)**, and in MCF-7 cells transfected with siRNA directed against FGFR2 versus MCF-7 cells transfected with a nontargeting control siRNA **(D)**. Generation of the tree-and-leaf diagrams, representing the breast cancer risk TF network, has been described previously (5). The size of the regulons is represented by circle size, and differential enrichment score (dES), as determined by GSEA, is represented by color. Data for A and B were from the METABRIC data set (57). Data for **C** and **D** were from microarray analysis (deposited in GEO under the SuperSeries number GSE74663).

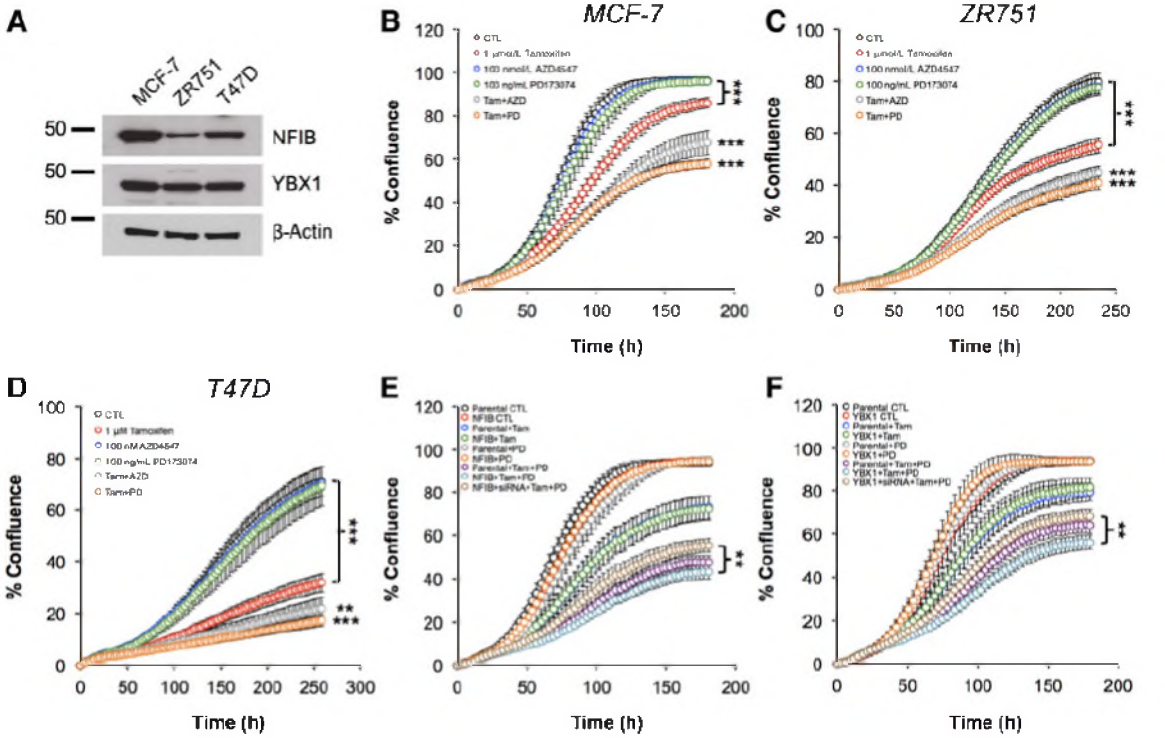


Figure 6.

FGFR2 inhibition sensitizes ER$^+$ breast cancer cells to antiestrogen therapies. **A,** Representative Western immunoblots showing expression of NFIB, YBX1, and β-actin proteins in MCF-7, ZR751, and T47D cells ($n = 3$ for all blots). **B–D,** Growth curves for MCF-7 **(B)**, ZR751 **(C)** and T47D **(D)** cells following treatment with 1 μmol/L tamoxifen, 100 nmol/L AZD4547 (FGFR inhibitor), 100 ng/mL PD173074 (FGFR inhibitor), 1 μmol/L tamoxifen plus 100 nmol/L AZD4547 (Tam + AZD), or 1 μmol/L tamoxifen plus 100 ng/mL PD173074 (Tam + PD). **E** and **F**, Growth curves for parental MCF-7 cells versus MCF-7 cells stably overexpressing NFIB **(E)** or YBX1 **(F)** following treatment with 1 μmol/L tamoxifen (Tam), 100 ng/mL PD173074 (PD), 1 μmol/L tamoxifen plus 100 ng/mL PD173074 (Tam + PD), or siRNA directed against NFIB **(E)**/YBX1 **(F)** plus 1 μmol/L tamoxifen plus 100 ng/mL PD173074 (siRNA + Tam + PD). $n = 16$ for all growth curves, two separate experiments, $P < 0.01$ (**); $P < 0.001$ (***), one-way ANOVA and SNK correction; error bars, SEM. Statistical comparison for the Tam + AZD/PD treatments in **B–D** is against the tamoxifen alone treatment shown in red. To avoid overlapping curves in **E** and **F**, data from parental and stably expressing cell lines are presented separately in Supplementary Fig. S5.

We have previously demonstrated that the risk gene FGFR2 can influence the way in which a cell responds to estrogen, with FGFR2 signaling leading to reduced activity of the ESR1 regulon (29). We have now extended our analysis and found that FGFR2 signaling not only affects the ESR1 regulon, but alters the activity of many TFs: the activity of TFs highly expressed in luminal A or B tumors is decreased, while the activity of TFs highly expressed in BLBC, such as NFIB and YBX1, is increased. A link between FGFR2 signaling and the activity of specific TFs has previously been reported. For example, in MCF-7 cells it causes degradation of the progesterone receptor, leading to increased proliferation and cell migration (47). FGFR2 mediated activation of TFs associated with ER$^-$ disease has not been studied directly, but indirect evidence exists. Signaling through FGFR2 leads to phosphorylation of RSK2, a mediator of anchorage independent growth and motility (48), which in turn activates YBX1 by phosphorylation (49). Our data here indicates that FGFR2 signaling also increases the affinity of YBX1 for ESR1. Taken together these observations suggest that FGFR2 signaling increases the ability of YBX1 to activate target genes associated with BLBC, while at the same time increasing its ability to repress ESR1 target genes.

A role for FGFR2 in promoting a basal-like phenotype is consistent with previous findings. Functional studies of FGFR2 risk variants have demonstrated that a decrease in FGFR2 expression is associated with an increased risk in ER$^+$, but not ER$^-$ breast cancer (28). Conversely, FGFR2 amplifications, although infrequent (4%; ref. 50), occur primarily in ER$^-$ breast cancer. ER$^-$ breast cancer cell lines tend to express higher levels of FGFR2 than ER$^+$ breast cancer cell lines (51) and are more sensitive to FGFR2 inhibitors such as PD173074. In clinical samples, FGFR2 expression was higher in ER$^-$ tumors and associated with poor patient outcome (51). However, inhibition of FGFR2 signaling may also be effective in ER$^+$ tumors. We hypothesized that inhibition of FGFR2 signaling would make cells more dependent on estrogen (through upregulation of the ESR1 regulon) and therefore more sensitive to estrogen deprivation therapy. We tested this in cell lines and found that MCF-7, ZR751, and T47D cells treated with the FGFR2 inhibitors PD173074 or AZD4547 became more sensitive to treatment with tamoxifen.

FGFR inhibitors have been used effectively in the treatment of a variety of cancers, particularly those carrying FGFR amplifications (52, 53). In breast cancer, the FGFR1 gene is amplified in about

13% of all breast cancer cases, while other FGFR genes are only rarely amplified (FGFR2, 1.5%; FGFR3, 0.5%; FGFR4, 1.5%) and are not frequently mutated. In line with our findings for FGFR2, activation of both FGFR1 (by amplification) and FGFR3 (*in vitro*) is associated with a reduced response to endocrine therapy (54, 55). This suggested clinical trials of FGFR inhibitors in combination with estrogen deprivation therapy. Not surprisingly, such trials have focused on patients with amplifications in the FGFR pathway and gave encouraging results, but were ultimately inconclusive due to the small number of patients carrying the relevant genomic alteration (56). Our work here suggests that rather than just focusing on FGFR amplification, alternative biomarkers such as the presence of activated YBX1 could be used to select patients that may benefit from FGFR2 inhibition. Consistent with this suggestion, we find that high expression of YBX1 in ER^+ disease is associated with worse outcome. In the future, this link needs to be further explored and activated YBX1 protein measured in ER^+ tumor samples. Alternatively, treatment could be focused on downstream events, preventing the interaction of YBX1 or NFIB with ESR1. If this interaction is dependent on posttranslational modifications, the inhibition of the relevant enzymes may be effective. As a first step toward moving our findings to the clinic, we envisage the use of PDX models of breast cancer to confirm synergy between FGFR2 inhibition and estrogen deprivation treatment in preventing tumor growth.

In conclusion, we demonstrate that signaling by FGFR2 pushes cells toward a more basal phenotype, which is at least in part mediated by facilitating the interaction between NFIB and YBX1, and ESR1. The regulatory loop between NFIB/YBX1 and ESR1 may be a promising target for developing new therapeutic strategies.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: T.M. Campbell, M.A.A. Castro, B.A.J. Ponder, K.B. Meyer

Development of methodology: T.M. Campbell, M.A.A. Castro, K.B. Meyer

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): T.M. Campbell

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): T.M. Campbell, M.A.A. Castro, K.G.d. Oliveira, B.A.J. Ponder, K.B. Meyer

Writing, review, and/or revision of the manuscript: T.M. Campbell, M.A.A. Castro, K.G.d. Oliveira, B.A.J. Ponder, K.B. Meyer

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): M.A.A. Castro, K.B. Meyer

Study supervision: M.A.A. Castro, B.A.J. Ponder, K.B. Meyer

Acknowledgments

This work was funded by The Breast Cancer Research Foundation (BCRF) and by Cancer Research UK (CRUK). We thank the Genomics, Proteomics, Bioinformatics, Light Microscopy and Research Instrumentation Core Facilities at The CRUK Cambridge Institute for their help and expertise. We are grateful to Magdalena Grabowska for the gift of plasmids for the FRET experiments, and Kelly Holmes for the gift of the luciferase reporter cell line. B.A.J. Ponder is a Gibb Fellow of CRUK.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received April 19, 2017; revised September 22, 2017; accepted November 6, 2017; published OnlineFirst xx xx, xxxx.

References

1. Jia M, Dahlman-Wright K, Gustafsson JA. Estrogen receptor alpha and beta in health and disease. Best Pract Res Clin Endocrinol Metab 2015;29:557–68.
2. Clemons M, Goss P. Estrogen and the risk of breast cancer. N Eng J Med 2001;344:276–85.
3. Advani P, Moreno-aspitia a: current strategies for the prevention of breast cancer. Breast Cancer (Dove Medical Press) 2014;6:59–71.
4. Ban KA, Godellas CV. Epidemiology of breast cancer. Surg Oncol Clin North America 2014;23:409–22.
5. Castro MA, de Santiago I, Campbell TM, Vaughn C, Hickey TE, Ross E, et al. Regulators of genetic risk of breast cancer identified by integrative network analysis. Nat Genet 2016;48:12–21.
6. Kruse U, Sippel AE. Transcription factor nuclear factor I proteins form stable homo- and heterodimers. FEBS Lett 1994;348:46–50.
7. Campbell CE, Piper M, Plachez C, Yeh YT, Baizer JS, Osinski JM, et al. The transcription factor Nfix is essential for normal brain development. BMC Dev Biol 2008;8:52.
8. Qian F, Kruse U, Lichter P, Sippel AE. Chromosomal localization of the four genes (NFIA, B, C, and X) for the human transcription factor nuclear factor I by FISH. Genomics 1995;28:66–73.
9. Steele-Perkins G, Plachez C, Butz KG, Yang G, Bachurski CJ, Kinsman SL, et al. The transcription factor gene Nfib is essential for both lung maturation and brain development. Mol Cell Biol 2005;25:685–98.
10. Dooley AL, Winslow MM, Chiang DY, Banerji S, Stransky N, Dayton TL, et al. Nuclear factor I/B is an oncogene in small cell lung cancer. Gen Develop 2011;25:1470–5.
11. Han W, Jung EM, Cho J, Lee JW, Hwang KT, Yang SJ, et al. DNA copy number alterations and expression of relevant genes in triple-negative breast cancer. Genes Chromosomes Cancer 2008;47:490–99.
12. Moon HG, Hwang KT, Kim JA, Kim HS, Lee MJ, Jung EM, et al. NFIB is a potential target for estrogen receptor-negative breast cancers. Mol Oncol 2011;5:538–44.
13. Murtagh J, Martin F, Gronostajski RM. The Nuclear Factor I (NFI) gene family in mammary gland development and function. J Mammary Gland Biol Neoplasia 2003;8:241–54.
14. Grabowska MM, Elliott AD, DeGraff DJ, Anderson PD, Anumanthan G, Yamashita H, et al. NFI transcription factors interact with FOXA1 to regulate prostate-specific gene expression. Mol Endocrinol (Baltimore, Md) 2014;28:949–64.
15. Grabowska MM, Kelly SM, Reese AL, Cates JM, Case TC, Zhang J, et al. Nfib regulates transcriptional networks that control the development of prostatic hyperplasia. Endocrinology 2015:en20151312.
16. Lyabin DN, Eliseeva IA, Ovchinnikov LP. YB-1 protein: functions and regulation. Wiley Interdiscip Rev RNA 2014;5:95–110.
17. Bergmann S, Royer-Pokora B, Fietze E, Jurchott K, Hildebrandt B, Trost D, et al. YB-1 provokes breast cancer through the induction of chromosomal instability that emerges from mitotic failure and centrosome amplification. Cancer Res 2005;65:4078–87.
18. Finkbeiner MR, Astanehe A, To K, Fotovati A, Davies AH, Zhao Y, et al. Profiling YB-1 target genes uncovers a new mechanism for MET receptor regulation in normal and malignant human mammary cells. Oncogene 2009;28:1421–31.
19. Dahl E, En-Nia A, Wiesmann F, Krings R, Djudjaj S, Breuer E, et al. Nuclear detection of Y-box protein-1 (YB-1) closely associates with progesterone receptor negativity and is a strong adverse survival factor in human breast cancer. BMC Cancer 2009;9:410.
20. Davies AH, Reipas KM, Pambid MR, Berns R, Stratford AL, Fotovati A, et al. YB-1 transforms human mammary epithelial cells through chromatin remodeling leading to the development of basal-like breast cancer. Stem Cells 2014;32:1437–50.
21. Habibi G, Leung S, Law JH, Gelmon K, Masoudi H, Turbin D, et al. Redefining prognostic factors for breast cancer: YB-1 is a stronger predictor of relapse and disease-specific survival than estrogen receptor or HER-2 across all tumor subtypes. Breast Cancer Res 2008;10:R86.

22. Huang J, Tan PH, Li KB, Matsumoto K, Tsujimoto M, Bay BH. Y-box binding protein, YB-1, as a marker of tumor aggressiveness and response to adjuvant chemotherapy in breast cancer. Int J Oncol 2005;26:607–13.
23. Maciejczyk A, Szelachowska J, Ekiert M, Matkowski R, Halon A, Lage H, et al. Elevated nuclear YB1 expression is associated with poor survival of patients with early breast cancer. Anticancer Res 2012;32:3177–84.
24. Tarallo R, Bamundo A, Nassa G, Nola E, Paris O, Ambrosino C, et al. Identification of proteins associated with ligand-activated estrogen receptor alpha in human breast cancer cell nuclei by tandem affinity purification and nano LC-MS/MS. Proteomics 2011;11:172–9.
25. Shibata T, Watari K, Izumi H, Kawahara A, Hattori S, Fukumitsu C, et al. Breast cancer resistance to antiestrogens is enhanced by increased ER degradation and ERBB2 expression. Cancer Res 2017;77:545–56.
26. Fachal L, Dunning AM. From candidate gene studies to GWAS and post-GWAS analyses in breast cancer. Curr Opinion Genet Develop 2015; 30:32–41.
27. Edwards SL, Beesley J, French JD, Dunning AM. Beyond GWASs: illuminating the dark road from association to function. Am J Hum Genet 2013;93:779–97.
28. Meyer KB, O'Reilly M, Michailidou K, Carlebur S, Edwards SL, French JD, et al. Fine-scale mapping of the FGFR2 breast cancer risk locus: putative functional variants differentially bind FOXA1 and E2F1. Am J Hum Genet 2013;93:1046–60.
29. Campbell TM, Castro MA, de Santiago I, Fletcher MN, Halim S, Prathalingam R, et al. FGFR2 risk SNPs confer breast cancer risk by augmenting oestrogen responsiveness. Carcinogenesis 2016;37:741–50.
30. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001;25:402–8.
31. Mohammed H, Carroll JS. Approaches for assessing and discovering protein interactions in cancer. Mol Cancer Res 2013;11:1295–302.
32. Mohammed H, D'Santos C, Serandour AA, Ali HR, Brown GD, Atkins A, et al. Endogenous purification reveals GREB1 as a key estrogen receptor regulatory factor. Nat Protoc 2013;3:342–9.
33. Mohammed H, Taylor C, Brown GD, Papachristou EK, Carroll JS, D'Santos CS. Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. Cell Rep 2016;11:316–26.
34. Gastard M.FRET acceptor photobleaching (AB) in LASAF. Available from: http://bticornelledu/wp-content/uploads/2014/04/AppNote_FRET_ABpdf2006.
35. Liao Y, Smyth GK, Shi W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res 2013;41:e108.
36. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol 2010;11:R106.
37. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.
38. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Nat Acad Sci U S A 2005;102:15545–50.
39. Gyorffy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q, et al. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. Breast Cancer Res Treat 2010;123:725–31.
40. Gao H, Korn JM, Ferretti S, Monahan JE, Wang Y, Singh M, et al. High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response. Nat Med 2015;21:1318–25.
41. Itoh N, Ohta H. Fgf10: a paracrine-signaling molecule in development, disease, and regenerative medicine. Curr Mol Med 2014;14:504–9.
42. Zhang X, Martinez D, Koledova Z, Qiao G, Streuli CH, Lu P. FGF ligands of the postnatal mammary stroma regulate distinct aspects of epithelial morphogenesis. Development (Cambridge, England) 2014;141:3352–62.
43. Piston DW, Kremers GJ. Fluorescent protein FRET: the good, the bad and the ugly. Trends Biochem Sci 2007;32:407–14.
44. Jozwik KM, Carroll JS. Pioneer factors in hormone-dependent cancers. Nat Rev Cancer 2012;12:381–5.
45. Lasham A, Print CG, Woolley AG, Dunn SE, Braithwaite AW. YB-1: oncoprotein, prognostic marker and therapeutic target? Biochem J 2013; 449:11–23.
46. EBCTCG. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet (London, England) 2005;365:1687–717.
47. Piasecka D, Kitowska K, Czaplinska D, Mieczkowski K, Mieszkowska M, Turczyk L, et al. Fibroblast growth factor signalling induces loss of progesterone receptor in breast cancer cells. Oncotarget 2016;7:86011–25.
48. Czaplinska D, Turczyk L, Grudowska A, Mieszkowska M, Lipinska AD, Skladanowski AC, et al. Phosphorylation of RSK2 at Tyr529 by FGFR2-p38 enhances human mammary epithelial cells migration. Biochimica Et Biophysica Acta 2014;1843:2461–70.
49. Stratford AL, Reipas K, Hu K, Fotovati A, Brough R, Frankum J, et al. Targeting p90 ribosomal S6 kinase eliminates tumor-initiating cells by inactivating Y-box binding protein-1 in triple-negative breast cancers. Stem Cells (Dayton, Ohio) 2012;30:1338–48.
50. Turner N, Lambros MB, Horlings HM, Pearson A, Sharpe R, Natrajan R, et al. Integrative molecular profiling of triple negative breast cancers identifies amplicon drivers and potential therapeutic targets. Oncogene 2010;29:2013–23.
51. Sharpe R, Pearson A, Herrera-Abreu MT, Johnson D, Mackay A, Welti JC, et al. FGFR signaling promotes the growth of triple-negative and basal-like breast cancer cell lines both in vitro and in vivo. Clin Cancer Res 011; 17:5275–86.
52. Chae YK, Ranganath K, Hammerman PS, Vaklavas C, Mohindra N, Kalyan A, et al. Inhibition of the fibroblast growth factor receptor (FGFR) pathway: the current landscape and barriers to clinical application. Oncotarget 2016.
53. De Luca A, Frezzetti D, Gallo M, Normanno N. FGFR-targeted therapeutics for the treatment of breast cancer. Expert Opinion Invest Drugs 2017; 26:303–11.
54. Turner N, Pearson A, Sharpe R, Lambros M, Geyer F, Lopez-Garcia MA, et al. FGFR1 amplification drives endocrine therapy resistance and is a therapeutic target in breast cancer. Cancer Res 2010;70:2085–94.
55. Tomlinson DC, Knowles MA, Speirs V. Mechanisms of FGFR3 actions in endocrine resistant breast cancer. Int J Cancer 2012;130:2857–66.
56. Musolino A, Campone M, Neven P, Denduluri N, Barrios CH, Cortes J, et al. Phase II, randomized, placebo-controlled study of dovitinib in combination with fulvestrant in postmenopausal patients with HR+, HER2- breast cancer that had progressed during or after prior endocrine therapy. Breast Cancer Res 2017;19:18.
57. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012;486:346–52.

Supplementary figure legends

Supplementary Figure 1. Nuclear localisation and fluorescent properties of the individual FRET constructs, as listed above the panels, transfected into HeLa cells.

Supplementary Figure 2. Quantification of FRET efficiency by FRET Acceptor Photobleaching (n=20-30 cells per transfection, and the experiments were repeated in at least three cellular preparations, P<0.001 (***), Student's t-test, error bars = SEM).

Supplementary Figure 3. Relative mRNA expression of the ESR1-target genes, GREB1 (A) and MYC (B), in MCF-7 breast cancer cells following transfection with siRNA directed against ESR1, FOXA1, NFIB and YBX1, and with plasmids overexpressing NFIB and YBX1, compared with a scrambled control siRNA transfection (CTL). All data were normalised to DGUOK expression (n=10, two separate experiments, P<0.05 (*), P<0.001 (***), one-way ANOVA and SNK correction, error bars = SEM).

Supplementary Figure 4. Characterisation of MCF-7 clones stably overexpressing NFIB and YBX1. Relative mRNA expression of FLAG-tagged NFIB (A) and YBX1 (B) in three separate stable clones of MCF-7 breast cancer cells. All data were normalised to DGUOK expression (n=10, two separate experiments, error bars = SEM). Insets: representative Western immunoblots showing expression of FLAG-tagged NFIB and YBX1, respectively, in the three stable MCF-7 clones (n=3 for all blots). (C) Representative Western immunoblots showing expression of ESR1 and β-actin protein in the NFIB- and YBX1-overexpressing clones (n=3 for all blots).

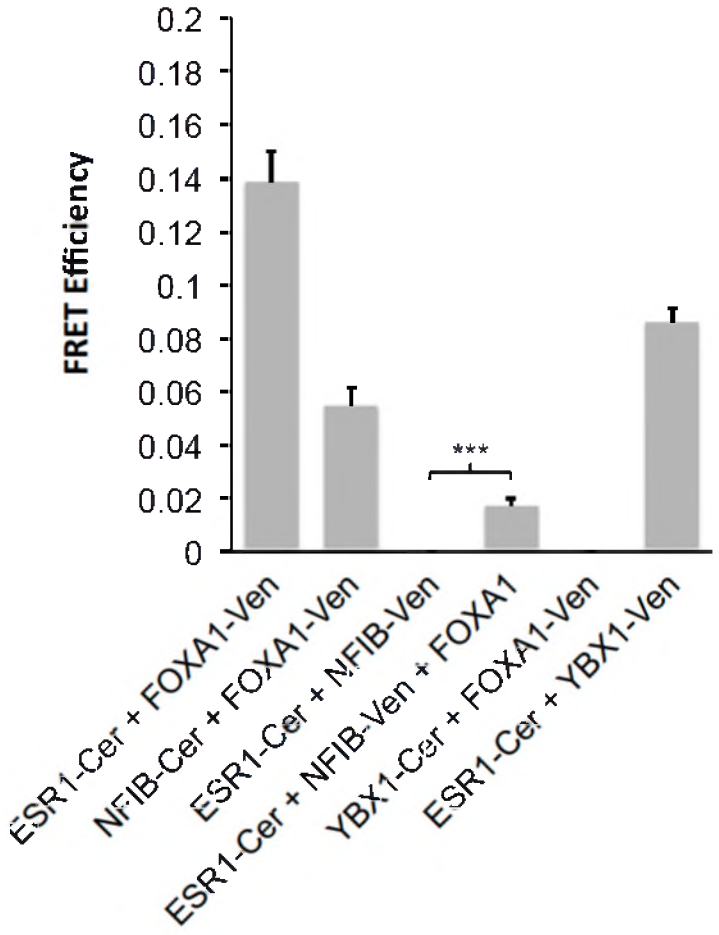
Supplementary Figure 5. Separated growth curves for parental MCF-7 cells (A,C) versus MCF-7 cells stably overexpressing NFIB (B) or YBX1 (D) following treatment with 1 μM tamoxifen (Tam), 100 ng/mL PD173074 (PD), 1 μM tamoxifen plus 100 ng/mL PD173074 (Tam+PD) or siRNA directed against NFIB (B)/YBX1 (D) plus 1 μM tamoxifen plus 100 ng/mL PD173074 (siRNA+Tam+PD). *n*=16 for all growth curves, two separate experiments, *P*<0.01 (**), *P*<0.001 (***), one-way ANOVA and SNK correction, error bars = SEM.

Supplementary Figure 6. Kaplan-Meier survival curves stratified for median expression of YBX1, as assayed by probe 208628_s_at, in all breast cancer cases (A), ER^+ breast cancer cases (B), and luminal A breast cancer cases (C). Kaplan-Meier survival analysis was carried out using the KM plotter website (http://kmplot.com/analysis/). Except for the sample selection shown in A-C, all parameters were left on default settings.

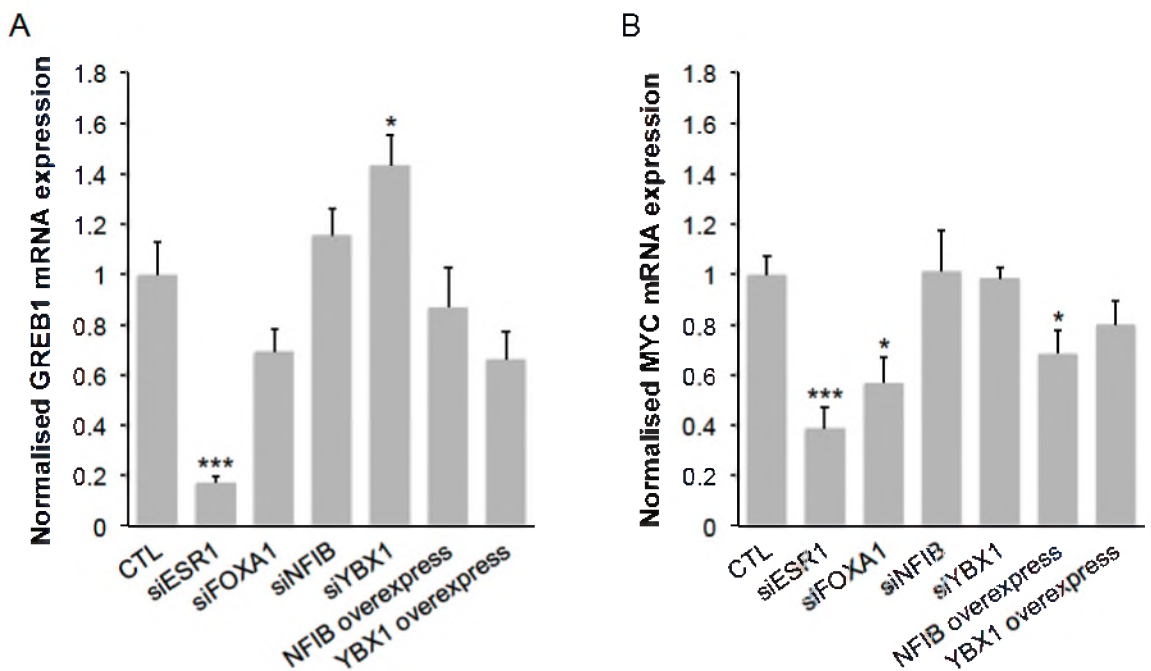
Supplementary Figure 7. Response of breast cancer PDX models to tamoxifen treatment. A total of 42 PDX models were included in the analysis. Gene expression profiles of PDXs were used to split the cohort into (A) upper YBX1 quartile containing 257 samples, (B) lower YBX1 quartile containing 175 samples and (C) lower YBX1 and lower FGFR2 quartiles containing 40 samples. Difference between treatment and control groups was assessed by RM-ANOVA test (for further details on the PDX models, please refer to Gao *et al.*, 2015). ESR1 expression levels were also compared between the three groups in the analysis and we found no statistical difference (*P* value = 0.176).

Supplementary Figure 1

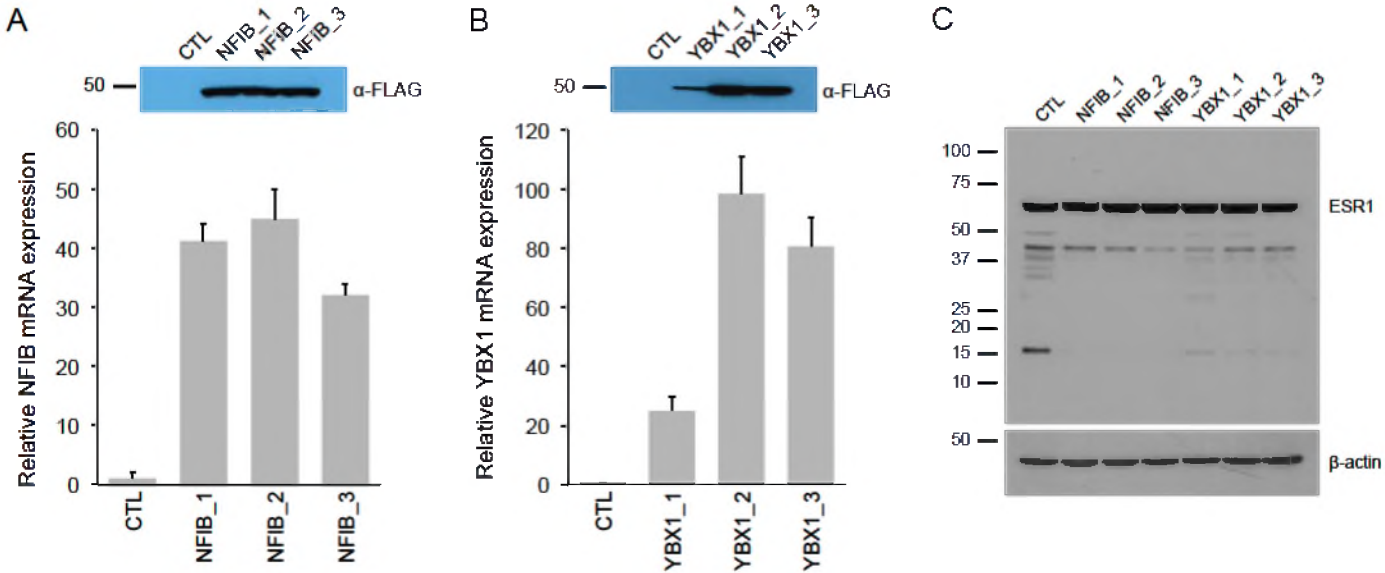
Supplementary Figure 2



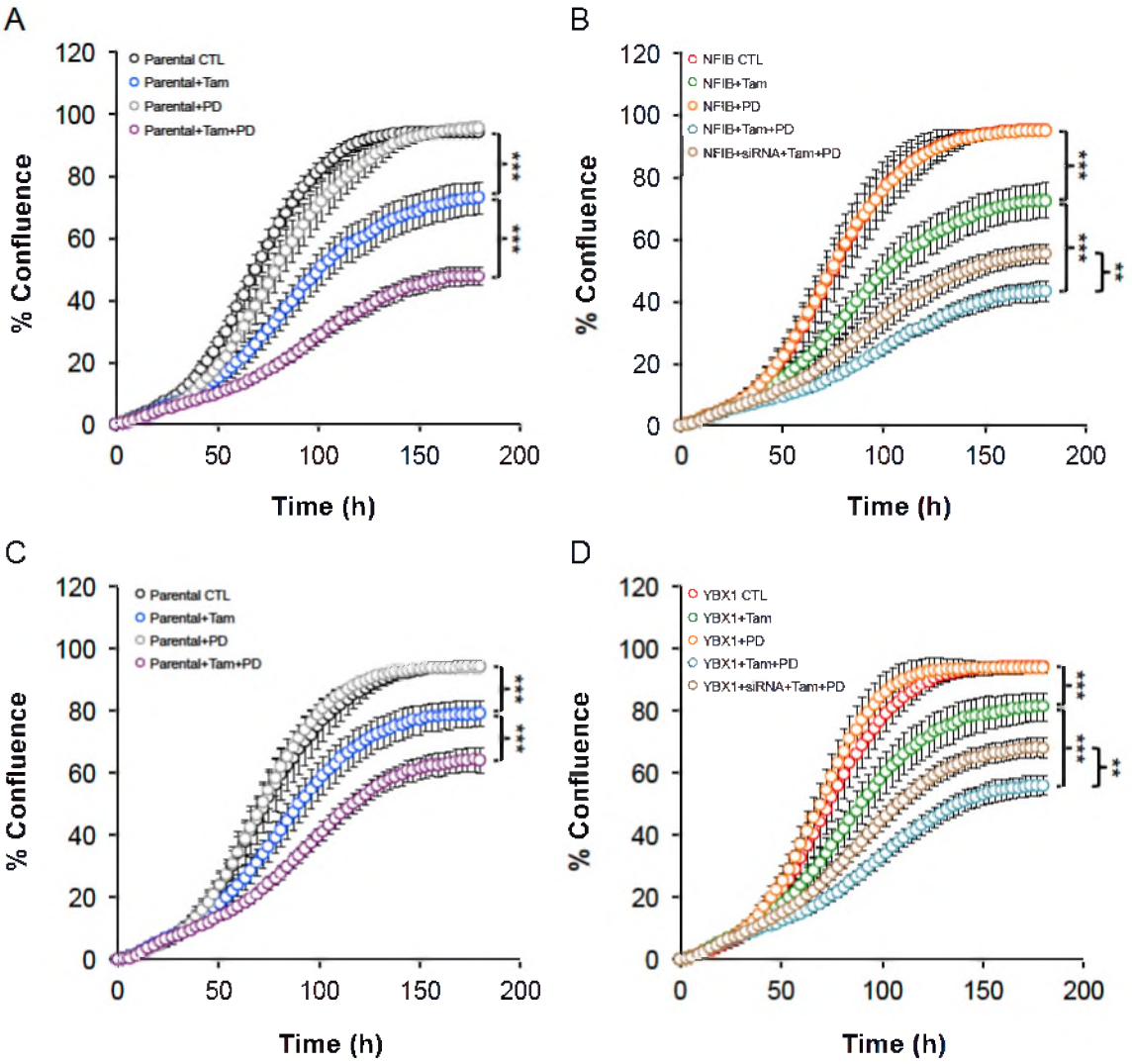
Supplementary Figure 3



Supplementary Figure 4



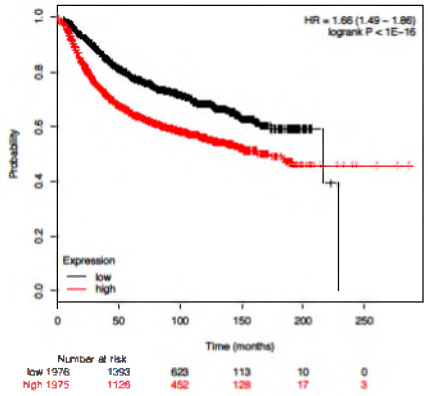
Supplementary Figure 5



Supplementary Figure 6

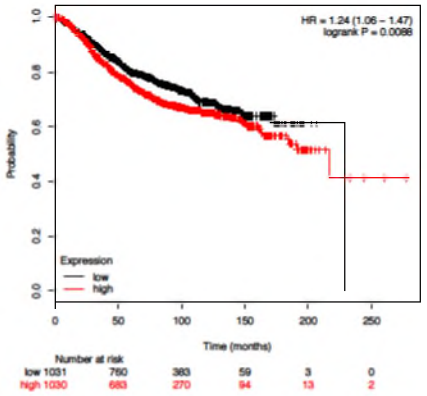
A

All BC



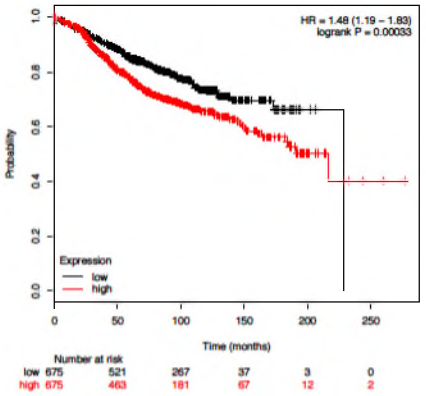
B

ER+

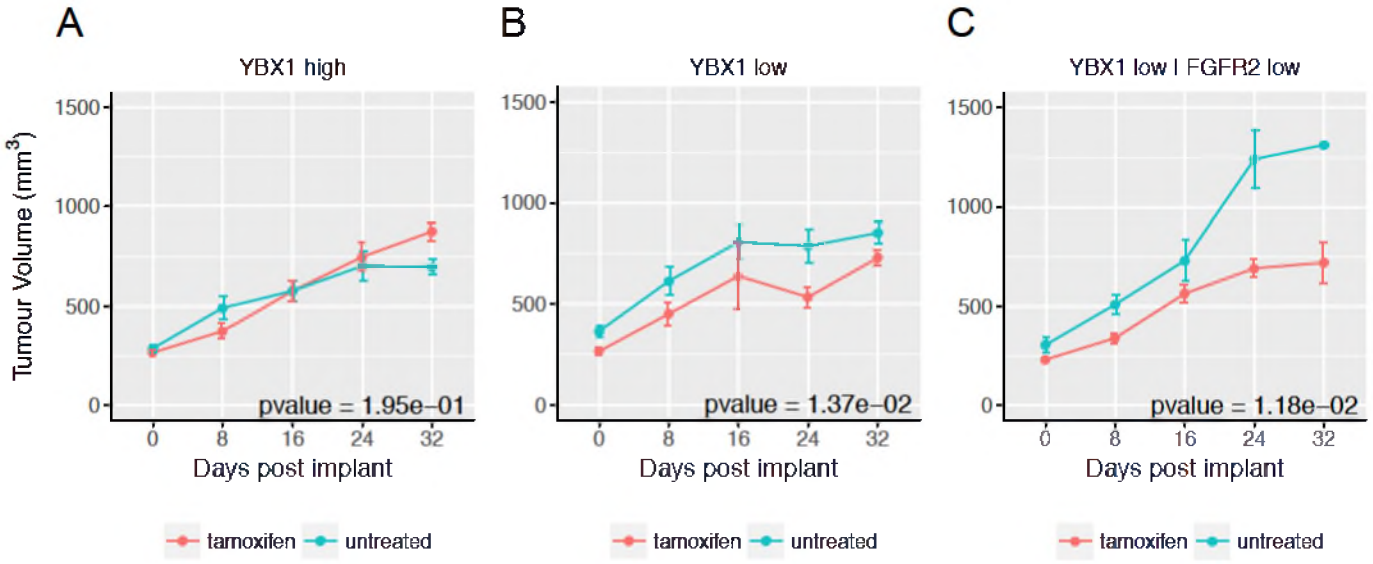


C

Luminal A



Supplementary Figure 7



Supplementary tables

Supplementary Table 1. Primers used in qRT-PCR to determine mRNA expression.

Gene	*Forward primer (5'-3')*	*Reverse primer (5'-3')*
DGUOK	GCTGGTGTTGGATGTCAATG	GCCTGAACTTCATGGTATTGG
NFIB	AAACCCAGCACTTTGTGTCC	TCTTGGGGAAGAATCCTGTG
YBX1	GGAGTTTGATGTTGTTGAAGGA	AACTGGAACACCACCAGGAC
pS2	GTGTCACGCCCTCCCAGT	GGACCCCACGAACGGTG
GREB1	CAAAGAATAACCTGTTGGCCC	GACATGCCTGCGCTCTCATACTTA
MYC	CCTCCACTCGGAAGGACTATC	TGTTCGCCTCTTGACATTCTC

Supplementary Table 2. Antibodies used in Western immunoblot experiments.

Protein of interest	*Antibody*
Primary antibodies:	
β-actin	Cell signaling 4970S (1:5000)
E2F2	Santa Cruz sc-632 (1:1000)
ESR1	Santa Cruz sc-543 X (1:5000)
FLAG	OriGene TA50011 (1:2000)
FOXA1	Abcam ab5089 (1:1000)
GATA3	Santa Cruz sc-268 X (1:5000)
NFIB	Active Motif 39091 (1:500)
pYBX1 (Ser102)	Cell Signaling 2900S (1:1000)
SP1	Santa Cruz sc-14027 X (1:1000)
YBX1	Abcam ab12148 (1:500)
YY1	Santa Cruz sc-7341 (1:1000)
Secondary antibodies:	
Anti-rabbit-HRP	Amersham NA934V (1:10,000)
Anti-mouse-HRP	Dako P0447 (1:2000)

Supplementary Table 3. MCF-7 and ZR751 cell line RIME data.

(ATTATCHED FILE: RIME DATA.xls)

Supplementary Table 4. Primers used to generate FRET constructs.

Gene	*Forward primer (5'-3')*	*Reverse primer (5'-3')*
ESR1	AAAGCTAGCATGACCATGACCCTCCACACC	AAAACCGGTGACCGTGGCAGGGAAACCCT
NFIB	AAAGCTAGCATGATGTATTCTCCCATCTGT	AAAACCGGTGCCCAGGTACCAGGACTGTTG
YBX1	AAAGCTAGCATGAGCAGCGAGGCCGAGACC	AAAACCGGTCTCAGCCCCGCCCTGCTCAG

PARTE III

Discussão Geral e Conclusões

4 DISCUSSÃO GERAL E CONCLUSÕES

A cada ano, cerca de 59 mil brasileiras são diagnosticadas com câncer de mama, uma doença heterogênea com múltiplos fatores de risco (INCA, 2018). Em diversos estudos, SNPs presentes no gene do receptor 2 do fator de crescimento fibloblástico foram reportados como intimamente relacionados com maior risco à doença (FLETCHER et al., 2013). O gene FGFR2 codifica receptores tirosina-quinase atuantes em processos pró- e anti- proliferativos mediante propagação de sinais via fatores de transcrição, que modificam eventos nucleares e remodelam a expressão gênica (SARABIPOUR; HRISTOVA, 2016). Dentre os fatores de transcrição relacionados com câncer de mama, 36 destes foram recentemente identificados como enriquecidos com variantes de risco e, assim como o FGFR2, participam da regulação de cascatas proliferativas mediadas por estrogênio (CASTRO et al., 2015), um hormônio protagonista na biologia deste tipo de tumor, que permite inferir subtipagem e direcionamento terapêutico (FLETCHER et al., 2013).

O presente trabalho buscou encontrar elementos de transdução entre o FGFR2 e os reguladores mestre do risco do câncer de mama através de duas etapas: predição computacional e corroboração experimental.

No primeiro artigo apresentado, a predição computacional permitiu a construção de redes PPI contendo possíveis alvos do receptor em tecido mamário saudável, os quais foram analizados quanto a capacidade de modulação de fatores de transcrição relacionados com risco ao cancer de mama. Consecutivamente, este conjunto de genes foi filtrado através da investigação de dados de perturbação ao receptor de membrana. Como resultado, oito genes candidatos se apresentaram diferencialmente expressos nos dados de perturbação experimental por superexpressão e knockdown do gene FGFR2 e foram analisados quanto a capacidade de direcionarem a resposta terapêutica em coortes de PDX. Desta maneira, identificamos dois genes que estabelecem interação com o receptor FGFR2, cujos níveis de expressão são capazes de alterar a resposta à medicamentos anti-tumorais na coorte, sendo estes PBX1 e EPHA2.

O gene PBX1 foi identificado como um modulador inibitório de diversos fatores

de transcrição relacionados com resposta a estrogênio. Altos níveis de expressão de PBX1 mostraram um bom perfil de resposta aos medicamentos inibidores de FGFRs, bem como para tamoxifeno, o que indica que a atividade de PBX1 pode culminar na ativação de ESR1, ainda que PBX1 tenha sido apontado como um modulador inibitório da via de estrogênio. Ao contrário de PBX1, a resposta terapêutica foi restaurada quando a coorte de PDX apresentava baixos níveis de expressão de EPHA2, indicando que a expressão deste gene possa levar a insensibilidade a estrogênio.

No segundo artigo apresentado neste estudo, dois outros transdutores da cascata de sinalização existente entre FGFR2 e ESR1 foram identificados: YBX1 e NFIB. A partir de interação física com FGFR2, ambos atuam para reprimir a atividade transcricional de ESR1, seja através da ação direta de repressão mediada por YBX1 ou indireta, via NFIB e FOXA1. Análises de PDX também indicaram que a indução de YBX1 através de FGFR2, resulta em irresponsividade ao tratamento com tamoxifeno, sendo que tal resposta pôde ser restaurada em amostras que possuiam ambos os genes suprimidos.

Em conclusão, a expressão de FGFR2 tem efeito indutório para EPHA2, NFIB e YBX1, enquanto que é inibitório para PBX1, demonstrando que as células tumorais podem se utilizar da expressão de FGFR2 a fim de ativar múltiplas vias de inibição de ESR1, fazendo com que células inicialmente ER+ sejam alteradas para proliferarem de maneira mais parecida as células basal-like, ou seja, insensíveis a presença de estrogênio no meio. Este perfil onde a atividade do receptor de estrogênio não tem participação confere um pior prognóstico ao paciente devido ao fenótipo tumoral mais agressivo e piores respostas aos tratamentos. No entanto, terapias conjuntas com inibidores de FGFR2 e tamoxifeno podem ser estratégias aplicáveis para aqueles que se encaixam nesta categoria de falsos ER-. Alternativamente a isso, novas estratégias terapêuticas poderiam trabalhar na combinação entre o aumento dos níveis de PBX1 ou na redução dos níveis de EPHA2, NFIB ou YBX1 combinados com tamoxifeno para aumentar a efetividade terapêutica.

5 PERSPECTIVAS

Futuramente, corroborações experimentais poderão adicionar pontos de esclarecimento entre a relação entre FGFR2-PBX1-ERS1 e FGFR2-EPHA2-ERS1, bem como estas vias se correlacionam com as anteriormente propostas contendo NFIB e YBX1.

Além disso, a visão panorâmica sobre os conjuntos de variantes atuantes como fatores de risco pode ser melhor explorada por meio de buscas de polimorfismos que estejam relacionados aos genes candidatos e, talvez, adicionem novos pontos a serem considerados anteriormente a estratificação tumoral e a escolha do tratamento mais adequado para pacientes.

REFERÊNCIAS

ABUBAKAR, M. et al. Prognostic value of automated KI67 scoring in breast cancer: A centralised evaluation of 8088 patients from 10 study groups. **Breast Cancer Research**, Breast Cancer Research, v. 18, n. 1, p. 1–13, 2016. ISSN 1465542X. Disponível em: <http://dx.doi.org/10.1186/s13058-016-0765-6>.

ACS. **American Cancer Society**. 2018. Disponível em: <https://www.cancer.org/cancer/breast-cancer.html>.

BERNARDES, A. Anatomia da mama feminina. In: **Manual de Ginecologia - Volume II**. [S.l.: s.n.], 2010. p. 167–174.

BOCHAROV, E. V. et al. Structure of FGFR3 transmembrane domain dimer: Implications for signaling and human pathologies. **Structure**, Elsevier Ltd, v. 21, n. 11, p. 2087–2093, 2013. ISSN 09692126. Disponível em: <http://dx.doi.org/10.1016/j.str.2013.08.026>.

BØRRESEN-DALE, A.-L.; SØRLIE, T.; KRISTENSEN, V. N. On the molecular biology of breast cancer. **Molecular oncology**, v. 4, p. 171–173, 2010. ISSN 15747891.

CAMPBEL, T. M. et al. FGFR2 risk SNPs confer breast cancer risk by augmenting oestrogen responsiveness. **Carcinogenesis**, v. 37, n. 8, p. 741–750, 2016. ISSN 14602180.

CASTRO, M. A. A. et al. Regulators of genetic risk of breast cancer identified by integrative network analysis. **Nature Genetics**, Nature Publishing Group, v. 48, n. November, p. 12–21, 2015. ISSN 1061-4036. Disponível em: <http://www.nature.com/doifinder/10.1038/ng.3458>.

CHEN, W. Y. Exogenous and endogenous hormones and breast cancer. **Best Pract Res Clin Endocrinol Metab**, v. 22, n. 4, p. 573–585, 2008.

CRUK. **Cancer Research UK**. 2018. Disponível em: <http://www.cancerresearchuk.org/about-cancer/breast-cancer/stages-types-grades>.

CUI, F. et al. Variants of FGFR2 and their associations with breast cancer risk: a HUGE systematic review and meta-analysis. **Breast Cancer Research and Treatment**, Springer US, v. 155, n. 2, p. 313–335, 2016. Disponível em: <http://link.springer.com/10.1007/s10549-015-3670-2>.

DRAGHI, F. et al. Ultrasonography of the male breast. **Journal of Ultrasound**, v. 14, n. 3, p. 122–129, 2011. ISSN 19713495.

ELSTON, C. W.; ELLIS, I. O. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow up. **Histopathology**, v. 19, p. 403–410, 1991. ISSN 03090167.

ESWARAKUMAR, V. P.; LAX, I.; SCHLESSINGER, J. **Cellular signaling by fibroblast growth factor receptors**. 2005. 139–149 p.

FLETCHER, M. N. et al. Master regulators of FGFR2 signalling and breast cancer risk. **Nature Communications**, v. 4, 2013. ISSN 20411723.

GIANCOTTI, V. Breast cancer markers. **Cancer Letters**, v. 243, p. 145–159, 2006.

GOWN, A. M.; FULTON, R. S.; KANDALAFT, P. L. **Markers of metastatic carcinoma of breast origin**. 1 2016. 86–95 p. Disponível em: <http://doi.wiley.com/10.1111/his.12877>.

HANAHAN, D.; WEINBERG, R. A. Hallmarks of cancer: The next generation. **Cell**, Elsevier Inc., v. 144, n. 5, p. 646–674, 2011. ISSN 00928674. Disponível em: <http://dx.doi.org/10.1016/j.cell.2011.02.013>.

HEJMADI, M. **Introduction to Cancer Biology**. [s.n.], 2010. 3–27 p. ISSN 1469-185X. ISBN 978-87-7681-478-6. Disponível em: <http://csbl.bmb.uga.edu/mirrors/JLU/DragonStar2017/download/introduction-to-cancer-biology.pdf>.

HU, Z. et al. The molecular portraits of breast tumors are conserved across microarray platforms. **BMC Genomics**, v. 7, p. 1–12, 2006. ISSN 14712164.

INCA. **Instituto Nacional de Câncer José Alencar Gomes da Silva**. 2018. Disponível em: <http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/mama>.

JESINGER, R. A. **Breast anatomy for the interventionalist**. 2014. 3–9 p.

KATOH, M. FGFR inhibitors: Effects on cancer cells, tumor microenvironment and whole-body homeostasis (Review). **International Journal of Molecular Medicine**, v. 38, p. 3–15, 2016. ISSN 1107-3756.

KIM, M. C. et al. Coexistent Loss of the Expressions of BRCA1 and p53 Predicts Poor Prognosis in Triple-Negative Breast Cancer. **Annals of Surgical Oncology**, v. 23, n. 11, p. 3524–3530, 2016. ISSN 15344681.

KIM, S. et al. FGFR2 Promotes Breast Tumorigenicity through Maintenance of Breast Tumor-Initiating Cells. **PLoS ONE**, v. 8, n. 1, 2013. ISSN 19326203.

LIN, Y.; WANG, F. FGF signalling in prostate development, tissue homoeostasis and tumorigenesis. **Bioscience Reports**, v. 30, n. 5, p. 285–291, 2010. ISSN 0144-8463. Disponível em: <http://bioscirep.org/cgi/doi/10.1042/BSR20100020>.

LU, Y. et al. p21-activated protein kinase 4 (PAK4) interacts with the keratinocyte growth factor receptor and participates in keratinocyte growth factor-mediated inhibition of oxidant-induced cell death. **Journal of Biological Chemistry**, v. 278, n. 12, p. 10374–10380, 2003. ISSN 00219258.

LUO, Y. et al. Novel phosphotyrosine targets of FGFR2IIIb signaling. **Cellular Signalling**, v. 21, n. 9, p. 1370–1378, 2009. ISSN 08986568.

MARSICANO, A. P. et al. Câncer de Mama. **Jornada Científica de Enfermagem 2014 Câncer**, v. 4, p. 1–2, 2014.

MATSUMOTO, A. et al. Biological markers of invasive breast cancer. **Japanese Journal of Clinical Oncology**, v. 46, n. October 2015, p. 99–105, 2016.

NBCF. **National Breast Cancer Foundation**. 2018. Disponível em: <http://www.nationalbreastcancer.org/what-is-breast-cancer>.

ORNITZ, D. M.; ITOH, N. The fibroblast growth factor signaling pathway. **Wiley Interdisciplinary Reviews: Developmental Biology**, v. 4, n. 3, p. 215–266, 2015. ISSN 17597692.

PARKER, J. S. et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. **Journal of Clinical Oncology**, v. 27, n. 8, p. 1160–1167, 2009. ISSN 0732183X.

PENG, L. et al. Association Between BRCA Status and P53 Status in Breast Cancer: A Meta-Analysis. **Medical Science Monitor**, v. 22, p. 1939–1945, 6 2016. ISSN 1643-3750. Disponível em: <http://www.medscimonit.com/abstract/index/idArt/896260>.

PEROU, C.; BORRESEN-DALE, A. Systems Biology and Genomics of. **Cold spring Harb Perspect Biol**, p. 1–18, 2011. ISSN 19430264.

PEROU, C. M. et al. Molecular portraits of human breast tumours. **Nature**, v. 406, n. 6797, p. 747–752, 2000. ISSN 00280836.

RAKHA, E. A. et al. Breast cancer prognostic classification in the molecular era: The role of histological grade. **Breast Cancer Research**, v. 12, n. 4, 2010. ISSN 14655411.

RAMSAY, D. T. et al. Anatomy of the lactating human breast redefined with ultrasound imaging. **Journal of Anatomy**, v. 206, n. 6, p. 525–534, 2005. ISSN 00218782.

SARABIPOUR, S.; HRISTOVA, K. Mechanism of FGF receptor dimerization and activation. **Nature Communications**, Nature Publishing Group, v. 7, p. 1–12, 2016. ISSN 20411723. Disponível em: <http://dx.doi.org/10.1038/ncomms10262>.

TIONG, K. H.; MAH, L. Y.; LEONG, C.-O. Functional roles of fibroblast growth factor receptors (FGFRs) signaling in human cancers. **Apoptosis**, v. 18, n. 12, p. 1447–1468, 2013. ISSN 1360-8185. Disponível em: <http://link.springer.com/10.1007/s10495-013-0886-7>.

TURNER, N.; GROSE, R. **Fibroblast growth factor signalling: From development to cancer**. 2010. 116–129 p.

WEIGELT, B.; GEYER, F. C.; REIS-FILHO, J. S. Histological types of breast cancer: How special are they? **Molecular Oncology**, Elsevier B.V, v. 4, n. 3, p. 192–208, 2010. ISSN 15747891. Disponível em: <http://dx.doi.org/10.1016/j.molonc.2010.04.004>.

YEUNG, S. J.; PAN, J.; LEE, M. H. **Roles of p53, MYC and HIF-1 in regulating glycolysis - The seventh hallmark of cancer**. 2008. 3981–3999 p.

ZHANG, Y. et al. Association between FGFR2 (rs2981582, rs2420946 and rs2981578) polymorphism and breast cancer susceptibility: A meta-analysis. **Oncotarget**, v. 8, n. 2, 2017. ISSN 19492553.

ZHU, X.; ASA, S. L.; EZZAT, S. Genetic and epigenetic mechanisms down-regulate FGF receptor 2 to induce melanoma-associated antigen A in breast cancer. **American Journal of Pathology**, v. 176, n. 5, p. 2333–2343, 2010. ISSN 15252191.