UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

INSTITUTO DE CIÊNCIAS BÁSICAS DA SAÚDE

CURSO DE GRADUAÇÃO EM BIOMEDICINA

IGOR ARAUJO VIEIRA

FREQUÊNCIA DA VARIANTE GERMINATIVA rs78378222 EM PACIENTES COM CÂNCER DE MAMA HER2-POSITIVO E POTENCIAIS MECANISMOS MOLECULARES DE PATOGENICIDADE

Porto Alegre

Junho/2014

IGOR ARAUJO VIEIRA

FREQUÊNCIA DA VARIANTE GERMINATIVA rs78378222 EM PACIENTES COM CÂNCER DE MAMA HER2-POSITIVO E POTENCIAIS MECANISMOS MOLECULARES DE PATOGENICIDADE

Trabalho de conclusão de curso de graduação apresentado ao Instituto de Ciências Básicas da Saúde da Universidade Federal do Rio Grande do Sul, como requisito parcial para obtenção do título de Bacharel em Biomedicina.

Orientadora: Profª Drª Patricia Ashton-Prolla
Co-orientador: MSc. Gabriel de Souza Macedo

Porto Alegre

Junho/2014

"Não são os salões de mármore que dão a grandeza intelectual, mas sim a alma e o cérebro de um investigador"

Alexander Fleming

AGRADECIMENTOS

À minha orientadora Patricia Ashton-Prolla, agradeço pela confiança e orientação conduzida sempre de maneira respeitosa e agradável; e por me apresentar o caminho da ciência baseado na dedicação e atualização contínua e, assim, se tornar um exemplo de pesquisadora e pessoa no qual me espelho.

Ao meu co-orientador Gabriel de Souza Macedo, um profissional com nítida vocação para docência e pesquisa, agradeço pela oportunidade de desenvolver esse trabalho que integra o seu projeto de doutorado acadêmico e pela ajuda e incentivo perante todas as dificuldades enfrentadas. A vocês minha gratidão pelos valiosos ensinamentos e pela orientação que vem proporcionando as maiores conquistas da minha vida acadêmica.

Às doutorandas Mariana Kiehl e Bárbara Alemar, agradeço pela incansável ajuda com materiais e explicações sobre câncer de mama HER2-positivo e microRNAs, respectivamente. À doutoranda Clévia, agradeço pela grande amizade e pelo auxílio nas análises de Bioinformática.

A todos os demais colegas e amigos do Laboratório de Medicina Genômica (LMG): Ju, Ana, Pati Silva, Nayê, Lili Todeschini, Gustavo, Fê Vianna, Clê, Nessa e Rudi. Agradeço pelas ajudas, discussões científicas e as inúmeras horas de terapia gratuita. Agradeço também aos grandes irmãos que a vida me concedeu: Ethi, Mel e Marcus. Obrigado pela força durante a graduação e, principalmente, por me fazerem descobrir a amizade verdadeira.

Por fim, agradeço de modo especial ao apoio incondicional da minha família: mãe Lia e pai Rubinho, guerreiros incansáveis que lutaram diariamente comigo nesses 5 anos e sempre acreditaram no meu potencial, fazendo quase o impossível para que esse sonho se concretizasse; o meu irmão William, que certamente alerta meu anjo de plantão nas situações de aperto, me direciona para os melhores caminhos e manda só energias positivas da morada de paz em que se encontra; as primas Vivi e Nieli e a Tia Marisa,

que me incentivaram e me ajudaram desde o início dessa jornada. Essa conquista é NOSSA!

SUMÁRIO

LISTA DE ABREVIATURAS

3'UTR – Região 3' não-traduzida (*3' Untranslated Region*)

BRCA1 – *Breast Cancer Susceptibility Gene 1*

BRCA2 – *Breast Cancer Susceptibility Gene 2*

CHECK2 – *Checkpoint Kinase 2 gene*

CM – **C**âncer de **M**ama

CNV – *Copy Number Variation* (Variação no Número de Cópias)

ERO – **E**spécies **R**eativas de **O**xigênio

EUA – **E**stados **U**nidos da **A**mérica

GWAS – Estudo de associação por varredura genômica (*Genome Wide Association Study*)

HER2 – *Human Epidermal Growth Factor Receptor 2*

IARC – *International Agency for Research on Cancer*

IHQ – **I**muno**h**isto**q**uímica

INCa – **I**nstituto **N**acional de **Câ**ncer

KRAS – *Kirsten rat sarcoma viral oncogene homolog*

LFL – Síndrome de **L**i-**F**raumeni-**l**ike (*Li-Fraumeni-like Syndrome*)

miRNA – **mi**cro**RNA**

OMIM – *On-line Mendelian Inheritance in Man*

OR – *Odds Ratio* (razão de chances)

PTEN – *Phosphatase and tensin homolog gene*

RE – **R**eceptor de **E**strógeno

RISC – Complexo de Silenciamento Induzido por RNA *(RNA-Induced Silencing Complex)*

RT-PCR – *Reverse Transcription Polymerase Chain Reaction*

SLF – **S**índrome de **Li-F**raumeni (*Li-Fraumeni Syndrome*)

SNP – Polimorfismo de nucleotídeo único (*Single Nucleotide Polymorphism*)

SSP – **S**equência **S**inal de **P**oliadenilação

LISTA DE FIGURAS

LISTA DE TABELAS

RESUMO

Mutações germinativas no gene *TP53* estão associadas com a Síndrome de Li-Fraumeni (SLF) e sua variante, a Síndrome de Li-Fraumeni-like (LFL), ambas doenças autossômicas dominantes caracterizadas pela predisposição a múltiplos tumores em idade precoce. O câncer de mama (CM) é o tipo de tumor mais frequente em mulheres com SLF/LFL, apresentando predominantemente o fenótipo molecular HER2-positivo neste grupo de pacientes. Em um estudo recente do nosso grupo de pesquisa, o SNP (*Single Nucleotide Polymorphism*) rs78378222 (A>C), uma variante germinativa rara localizada na região 3' não-traduzida (3'UTR) de *TP53* e, mais especificamente, na sua sequência sinal de poliadenilação, foi identificado em uma alta frequência (5.4%) em um grupo de pacientes SLF/LFL sem mutações germinativas identificadas em regiões codificadoras do gene *TP53*. Sendo assim, os objetivos deste trabalho foram determinar a frequência do SNP funcional rs78378222 em uma amostra de mulheres com CM HER2-positivo do sul do Brasil, bem como investigar os mecanismos moleculares de patogenicidade do alelo variante rs78378222[C]. Inicialmente, através da genotipagem do SNP em nossa série de 105 casos de CM HER2-positivo, não foi identificada nenhuma paciente portadora do alelo rs78378222[C]. A avaliação dos níveis de expressão da proteína p53, tanto em tecido tumoral quanto em células normais (fibroblastos) de pacientes LFL portadores da variante, mostrou uma redução na expressão de p53. Além disso, utilizando-se ferramentas de predição *in silico*, foi possível observar que o SNP rs78378222 está situado em uma sequência-alvo da região 3'UTR do mRNA *TP53* predita para ligação de miR-545-3p. A presença de rs78378222[C] pode enfraquecer a ligação de miR-545-3p ao sítio-alvo predito, mas não parece criar um novo sítio de ligação para miRNAs na região 3'UTR de *TP53*. Nossos resultados sugerem que a haploinsuficiência pode ser um novo mecanismo pelo qual *TP53* promove predisposição ao câncer em famílias SLF/LFL. Entretanto, estudos em novas séries de casos devem ser realizados para confirmar estes achados.

1. INTRODUÇÃO

1.1. Síndrome de Li-Fraumeni (SLF)

No ano de 1969, os pesquisadores Frederick Li e Joseph Fraumeni, ao revisar registros médicos e atestados de óbito de 648 crianças norte-americanas diagnosticadas com rabdomiossarcomas (neoplasia maligna da musculatura esquelética), identificaram uma alta incidência de diferentes tipos de câncer em idade precoce entre os seus familiares. A partir dessa análise, observaram quatro famílias que apresentavam sarcomas de partes moles, câncer de mama (CM) e outros tumores acometendo crianças ou jovens adultos. A suscetibilidade aumentada ao câncer nestas famílias não foi caracterizada apenas pelo grande número de familiares afetados, mas também pela frequência elevada de múltiplos tumores primários em jovens. A distribuição dos casos de câncer nesses agrupamentos familiares sugeria um padrão de transmissão autossômico dominante e levou à descrição de uma nova síndrome de câncer familial, denominada Síndrome de Li-Fraumeni (SLF) (Li e Fraumeni, 1969a, b).

A SLF (OMIM #151623) consiste em uma síndrome monogênica de predisposição hereditária ao câncer, caracterizada por manifestações clínicas heterogêneas. Indivíduos com SLF apresentam risco aumentado para o desenvolvimento de um amplo espectro de tumores em idade precoce. Os tumores mais comuns em indivíduos com a doença, também chamados tumores centrais da síndrome ("*core tumors*"), são os sarcomas de partes moles e osteossarcomas, CM, tumores do sistema nervoso central (em especial meduloblastoma e carcinoma de plexo coróide), leucemias e carcinoma adrenocortical (Li *et al.*, 1988). Outros tumores, tais como melanoma, câncer gástrico e colorretal, câncer de pulmão e de próstata, tumores de células germinativas e tumor de Wilms já foram descritos em várias famílias com SLF (Frebourg *et al.*, 2001; Hartley *et al.*, 1989; Masciari *et al.*, 2011).

A definição clássica e mais bem aceita da SLF é a presença de um probando com diagnóstico de sarcoma antes dos 45 anos de idade, acompanhado de um familiar em primeiro grau com algum câncer antes dos

45 anos de idade, e outro familiar em primeiro ou segundo grau também com câncer antes dos 45 anos de idade ou um sarcoma em qualquer idade (Li *et al.*, 1988). Sempre que presentes, estes critérios indicam claramente o teste genético para diagnóstico da síndrome.

O mecanismo molecular associado com a SLF foi esclarecido apenas em 1990. Anteriormente, análises de segregação já mostravam que a doença tinha uma etiologia genética (Williams e Strong, 1987). Mutações inativadoras somáticas no gene *TP53* já haviam sido detectadas nas formas esporádicas da maioria dos tipos de câncer associados com a SLF, incluindo osteossarcomas, sarcomas de partes moles, leucemias e carcinomas de mama (Nigro *et al.*, 1989), e também em modelos animais da doença. Baseados nessas observações, Malkin *et al.* (1990) sequenciaram o gene *TP53* em leucócitos de indivíduos com suspeita clínica de SLF. Estas análises resultaram na detecção de alterações entre os éxons 5-8, confirmando a associação entre mutações germinativas em *TP53* e o fenótipo SLF.

Posteriormente, foram identificadas inúmeras famílias com mutações germinativas no gene *TP53* que apresentavam características clínicas incompletas quando considerados os critérios clássicos da síndrome. Dessa forma, outros critérios menos restritivos, incluindo os de Birch, Chompret e Eeles (Birch, 1994; Eeles, 1995; Frebourg *et al.*, 2001), foram propostos para indicar o teste genético e investigação da SLF e de suas variantes, definidas como Síndrome de Li-Fraumeni-like (LFL). A prevalência de mutações germinativas em *TP53* nestas famílias tem apresentado variação entre os diferentes estudos. No estudo de Varley *et al.* (2003) foi relatada uma prevalência de mutação de 77% e 40% em famílias com SLF (forma clássica) e LFL (forma variante), respectivamente.

As mutações observadas na SLF apresentam uma penetrância estimada de 90-95% para ocorrência de câncer ao longo da vida (Brosh e Rotter, 2009). Indivíduos portadores de mutações germinativas localizadas no domínio de ligação ao DNA de *TP53* têm, em média, um risco de 50% em desenvolver algum tipo de câncer até os 40 anos de idade e 90% até os 60 anos de idade, comparado com 1% observado na população em geral (Birch *et al.*, 1998; Royds e Iacopetta, 2006). Estudos na Europa e Estados Unidos

(EUA) apontam para uma taxa de portadores de mutação germinativa em *TP53* entre 1:5.000 a 1:20.000 nascidos-vivos (Gonzalez *et al.*, 2009; Lalloo *et al.*, 2006). No Brasil, a prevalência de mutações germinativas que ocorrem entre os éxons 5-8 do gene *TP53* na população geral ainda não foi determinada.

1.2. A proteína p53: funções e domínios

A proteína p53 é codificada pelo gene *TP53* (OMIM #191170) (Levine, 1989), que está localizado no braço curto do cromossomo 17 (17p13.1). Esse gene compreende 20 kb de DNA genômico e está dividido em 11 éxons (Figura 1a), sendo o primeiro não-codificante (Bourdon *et al.*, 2005; Isobe *et al.*, 1986). p53 é uma fosfoproteína tetramérica de cerca de 53 kDa. Cada monômero de p53 é composto por 393 aminoácidos organizados em cinco domínios estruturais e funcionais bem definidos (Figura 1b): a) um domínio de transativação amino-terminal (resíduos 20-60), responsável pela transativação de genes-alvo; b) um domínio rico em prolinas (resíduos 63-97), necessário para interação com proteínas envolvidas na indução de apoptose; c) um domínio central, altamente conservado evolutivamente, responsável pela ligação de p53 a sequências consenso de DNA situadas em regiões promotoras (resíduos 100-300); d) um domínio de oligomerização, fundamental para dimerização e posterior formação de homotetrâmeros de p53 (resíduos 324-355); e) um domínio carbóxi-terminal, envolvido na regulação da ligação ao DNA (resíduos 360-393) (Beckerman e Prives, 2010; Dornan *et al.*, 2003; Levine, 1997).
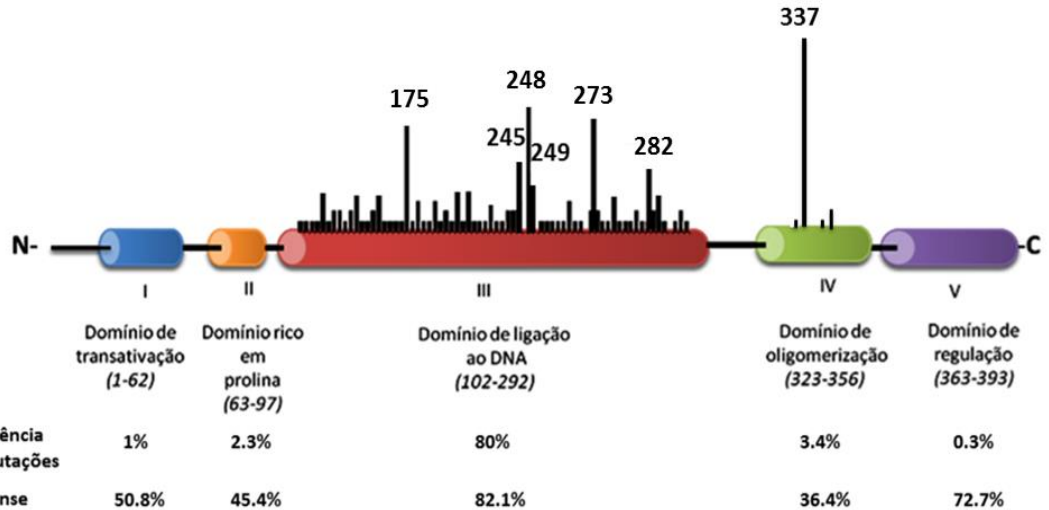
Figura 1. Representação esquemática dos éxons 1-11 do gene *TP53* (a) e da proteína p53 com seus diferentes domínios funcionais (b). Em (a) e (b) as diferentes cores correspondem aos domínios da proteína e os respectivos éxons codificantes (éxon 1 não é codificante). Em (b) são indicados os números dos resíduos de aminoácidos correspondentes a cada domínio, a frequência de mutações germinativas e os percentuais de mutações de sentido trocado em cada um destes domínios. Também são mostrados os resíduos onde ocorrem as sete mutações mais frequentes do gene, considerados códons "*hotspot*s", os quais estão situados no domínio de ligação ao DNA, com exceção do resíduo 337 no domínio de oligomerização, no qual ocorre a mutação p.R337H. Essa mutação apresenta uma frequência populacional elevada (aproximadamente 0,3%) no sul do Brasil devido a um efeito fundador. **Modificado de Bourdon *et al.*, 2005 (a) e Brosh e Rotter, 2009 (b).**

Conhecida como "guardiã do genoma", p53 consiste num fator de transcrição de vida curta, cuja capacidade de mediar a supressão tumoral tem sido amplamente estudada. p53 exerce suas múltiplas funções antiproliferativas através do controle transcricional de diversos genes-alvo e através de interações proteína-proteína (Olivier *et al.*, 2009). Em células normais não expostas a estresse, p53 é mantida em níveis muito baixos, devido a sua rápida degradação proteossomal mediada pela ubiquitina ligase MDM2, principal regulador negativo de p53. As diferentes formas de estresse celular, especialmente eventos genotóxicos, promovem a estabilização e ativação de p53 através de modificações pós-traducionais,

permitindo que a proteína escape da degradação, seja translocada para o núcleo e tenha capacidade de se ligar a sequências específicas de DNA. Esses sítios de ligação reconhecidos por p53, geralmente estão localizados em regiões promotoras de centenas de genes, cuja expressão pode ser induzida ou reprimida dependendo do tipo celular, da natureza do estresse e da extensão do dano (Beckerman e Prives, 2010; Lacroix *et al.*, 2006; Olivier *et al.*, 2009; Vousden e Prives, 2009).

Conforme mostra a figura 2, p53 pode ser ativada por várias formas de estresse celular, entre elas dano ao DNA, sinalização oncogênica, hipóxia, estresse oxidativo e encurtamento telomérico (Oren, 2003). Esta ativação induz a expressão de genes envolvidos na parada de ciclo celular, apoptose, senescência, diferenciação celular, modulação da migração celular, reparo do DNA, inibição de angiogênese e metabolismo (Beckerman e Prives, 2010; Lane e Levine, 2010; Vousden e Lane, 2007). Apoptose e parada do ciclo celular são as respostas celulares induzidas por p53 melhor caracterizadas até o momento (Olivier *et al.*, 2009).

Estudos recentes têm identificado novas funções para p53, incluindo a regulação do metabolismo energético e oxidativo (Feng *et al.*, 2011). p53 selvagem pode reduzir a glicólise e promover a fosforilação oxidativa mitocondrial. Perda de função da proteína resulta em diminuição do consumo de oxigênio e respiração mitocondrial prejudicada, bem como promove uma alta utilização de glicose na glicólise aeróbica (Bensaad *et al.*, 2006; Bensaad e Vousden, 2007; Hu *et al.*, 2010b; Matoba *et al.*, 2006). Sob condições de ausência ou baixa exposição a estresse, p53 estimula a expressão de proteínas que atuam como defesas antioxidantes e, assim, reduz os níveis de espécies reativas de oxigênio (ERO), protegendo a célula do dano ao DNA (Budanov *et al.*, 2004; Sablina *et al.*, 2005).
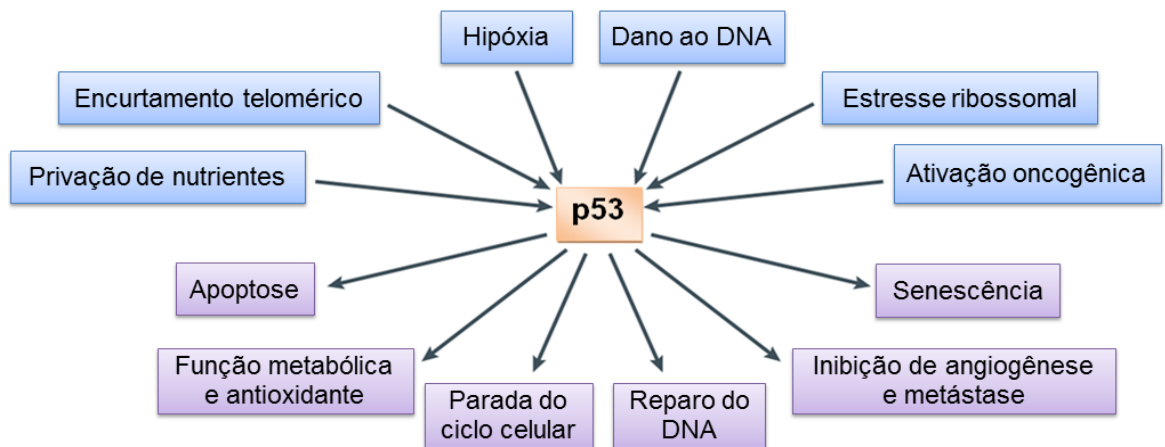
Figura 2. Ativação de p53 por diferentes sinais de estresse celular e suas funções supressoras de tumor. p53 desempenha o papel de integrar respostas celulares (em roxo) a uma variedade de tipos de estresse (em azul). **Modificado de Vousden e Lane, 2007.**

1.3. Variabilidade genética do gene *TP53*

Dois tipos de alterações genéticas podem ser detectadas no gene *TP53*: mutações somáticas encontradas em tumores esporádicos, e variações na sequência (mutações e polimorfismos) em células germinativas (Malkin *et al.*, 1990; Whibley *et al.*, 2009).

Alterações somáticas em *TP53* ocorrem em praticamente todos os tipos de câncer, com frequências de mutação variando entre 38-50% em câncer de ovário, colorretal e pulmão, a cerca de 5% em leucemias, melanoma metastático, entre outros. Considerando todas as mutações somáticas e germinativas já identificadas em *TP53* e compiladas no banco de dados do IARC (*International Agency for Research on Cancer*) (www-p53.iarc.fr, *release R16*), observa-se que a maior parcela destas são do tipo sentido trocado e encontram-se predominantemente na região correspondente ao domínio de ligação ao DNA (éxons 5-8) (Hainaut e Hollstein, 2000; Petitjean *et al.*, 2007). Aproximadamente 30% das mutações ocorrem em seis códons preferenciais ou *"hotspots"* (175, 245, 248, 249, 273 e 282), todos eles localizados neste domínio (Figura 1b). A distribuição de mutações germinativas (Figura 3), associadas com a SLF/LFL, é similar a de

mutações somáticas, com a maioria das mutações do tipo sentido trocado (77%) situadas nos mesmos resíduos *"hotspots"* (Olivier *et al.*, 2010).

Grande parte das mutações de sentido trocado em *TP53* ocasiona a síntese de uma proteína inativa, mas com estabilidade aumentada. Estas mutações levam a altos níveis de p53 no núcleo de células tumorais que são facilmente visualizados através da técnica de imunohistoquímica (IHQ) (Bártek *et al.*, 1991; Dowell *et al.*, 1994). Acredita-se que isso seja uma consequência da proteína p53 mutante não conduzir a expressão de MDM2 necessária para favorecer a sua própria degradação (Lacroix *et al.*, 2006). Entretanto, muitos tumores com mutações em *TP53* não apresentam acúmulo de p53 mutante, especialmente em casos de mutações *frameshift*, *nonsense* e em sítios de *splicing* (Olivier *et al.*, 2005; Soussi e Béroud, 2001).

Dezenas de polimorfismos germinativos no lócus *TP53* já foram identificados, no entanto, para a maioria deles as alterações funcionais associadas ainda não foram testadas. Cerca de 90% destas variantes estão localizadas em íntrons e parecem não estar associadas a características relacionadas à suscetibilidade ao câncer (Whibley *et al.*, 2009). p.R72P é o SNP mais estudado no gene *TP53* e já foi caracterizado como uma variante funcional, sendo observado em estudos *in vitro* que a proteína p53-R72 é um indutor mais potente da apoptose quando comparado ao alelo P (Dumont *et al.*, 2003). Muitos estudos apontam associações entre p.R72P e o risco para diversas neoplasias, mas os resultados são contraditórios. O polimorfismo p.R72P está em desequilíbrio de ligação com PIN3, um polimorfismo intrônico que consiste na duplicação de 16 pares de base no íntron 3 de *TP53* (Matakidou *et al.*, 2003; Olivier *et al.*, 2010; Schmidt *et al.*, 2007). PIN3 tem sido associado a um risco aumentado para diversos tipos de câncer. A presença desse alelo foi associada com menores níveis de transcrito *TP53*, sugerindo que o mesmo cause uma alteração no processamento do mRNA (Gemignani *et al.*, 2004; Wang-Gohrke *et al.*, 1998, 1999).
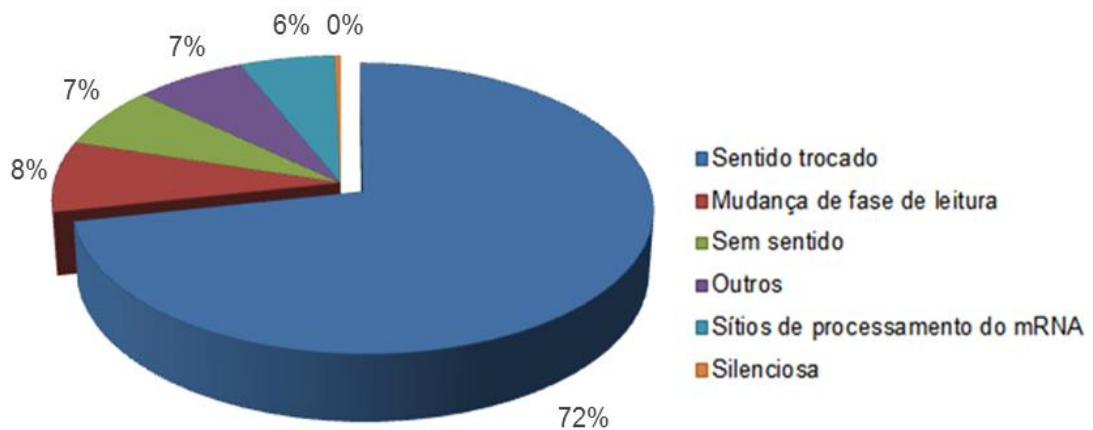
Figura 3. Distribuição dos tipos de mutações germinativas identificadas no gene *TP53*. Conforme indicado, a maioria das mutações germinativas em *TP53* é do tipo sentido trocado (*missense*). **Modificado de *IARC TP53 Mutation Database release R16*.** Disponível em: <http://www-p53.iarc.fr>. (Acessado em 16 de abril de 2014).

1.4. SNP rs78378222

1.4.1. Dados publicados

O SNP rs78378222 (A>C) consiste em uma variante germinativa rara localizada na região 3' não-traduzida (3'UTR) do gene *TP53*. Esta variante foi recentemente identificada em um estudo de associação por varredura genômica (GWAS) como um alelo de risco para carcinoma basocelular (OR = 2.36, *P*=5.2 X 10^{-17}), apresentando uma frequência estimada de 1,92% na população da Islândia. Nesse mesmo estudo, também foi demonstrado que rs78378222 confere suscetibilidade a outros tipos de tumores, como glioma (OR=2.35) e câncer de próstata (OR=1.44) em populações da Europa e Estados Unidos (EUA), com sua frequência nos casos variando de 0,7 a 4,9%. Por fim, foi encontrada uma associação significativa com adenoma colorretal (OR=1.39), mas não com câncer colorretal (OR=1.06), CM (OR=1.07) e melanoma (OR=1.06). Embora raros, foram identificados indivíduos homozigotos para a variante (CC), o que exclui a possibilidade de a mesma ser recessiva letal. Além disso, foi observado que essa variante gênica altera a sequência sinal de poliadenilação (SSP) do gene *TP53* (de AATA**A**A para AATA**C**A) e também diminui os níveis de transcritos de *TP53*

em indivíduos heterozigotos rs78378222[AC], em comparação aos homozigotos selvagens rs78378222[AA] (P = 0.041). Adicionalmente, estratégias de RT-PCR e sequenciamento indicaram que os mRNA *TP53* poli-A (+) eram originados predominantemente do alelo selvagem, com 73% dos mRNAs contendo o alelo selvagem A (P=1.6 X 10^{-6}). Esses achados sugeriram que rs78378222 leva a um processamento 3' anormal do mRNA *TP53*, cujo efeito seria a degradação mais rápida desse transcrito e, por sua vez, uma diminuição da expressão proteica, constituindo um novo mecanismo pelo qual *TP53* promoveria oncogênese (Stacey *et al.*, 2011).

Estudo funcional também testou o papel dessa SSP variante na regulação de p53 usando uma linhagem de células p53-nula. Através da transfecção de um construto contendo o alelo rs78378222[C] se observou diminuição dos níveis de mRNA *TP53* e de proteína p53, e, consequentemente, capacidade reduzida de indução de apoptose celular (Li *et al.*, 2013).

Em relação ao hexanucleotídeo AA[T/U]AAA, sabe-se que essa foi a primeira sequência regulatória de mRNA eucariótico descoberta, mostrando-se altamente conservada em eucariotos, conforme esquematizado na figura 4a (Proudfoot, 1991; Proudfoot e Brownlee, 1976). Desenvolvimento de sistemas *in vitro* para avaliar o processo de poliadenilação em mamíferos (Manley, 1983; Moore e Sharp, 1984, 1985) mostraram a importância da SSP AAUAAA na terminação 3' apropriada do pré-mRNA (Manley *et al.*, 1985), embora outras sequências contendo SNPs em AAUAAA permitam a poliadenilação em níveis menores, mas mensuráveis *in vivo* (Wickens e Stephenson, 1984) e *in vitro* (Conway e Wickens, 1987; Sheets *et al.*, 1990; Wilusz *et al.*, 1989). Vários estudos sugeriram que a maioria dos sítios de poliadenilação em genes humanos contém a sequência consenso AAUAAA (cerca de 70%), porém uma parcela menor de genes apresenta variabilidade nessa sequência (Beaudoing *et al.*, 2000; Graber *et al.*, 1999; MacDonald e Redondo, 2002). Esse tipo de alteração genética em humanos foi primeiramente identificado no sinal de poliadenilação do gene *HBA2* (AATAAA>AATAA**G**), que codifica a alfa-2-globulina, causando a alfa-talassemia (Higgs *et al.*, 1983).

Evidências experimentais indicam que o motivo AAUAAA é reconhecido pelo complexo proteico de clivagem do transcrito primário, o qual cliva a fita simples do pré-mRNA em um sítio localizado 10-30 nucleotídeos a montante dessa SSP. Após essa clivagem específica, ocorre a adição da cauda poli-A pela poli-A polimerase, aumentando a estabilidade e tradução do mRNA. Uma região rica em U/GU, pouco conservada e possivelmente redundante, situada 14-70 nucleotídeos a jusante de AAUAAA, também estaria envolvida na poliadenilação (Ryan *et al.*, 2004; Tian *et al.*, 2005; Zhao *et al.*, 1999). Essa região e AAUAAA constituem os chamados elementos centrais da poliadenilação (Lutz, 2008), destacados na figura 4b.
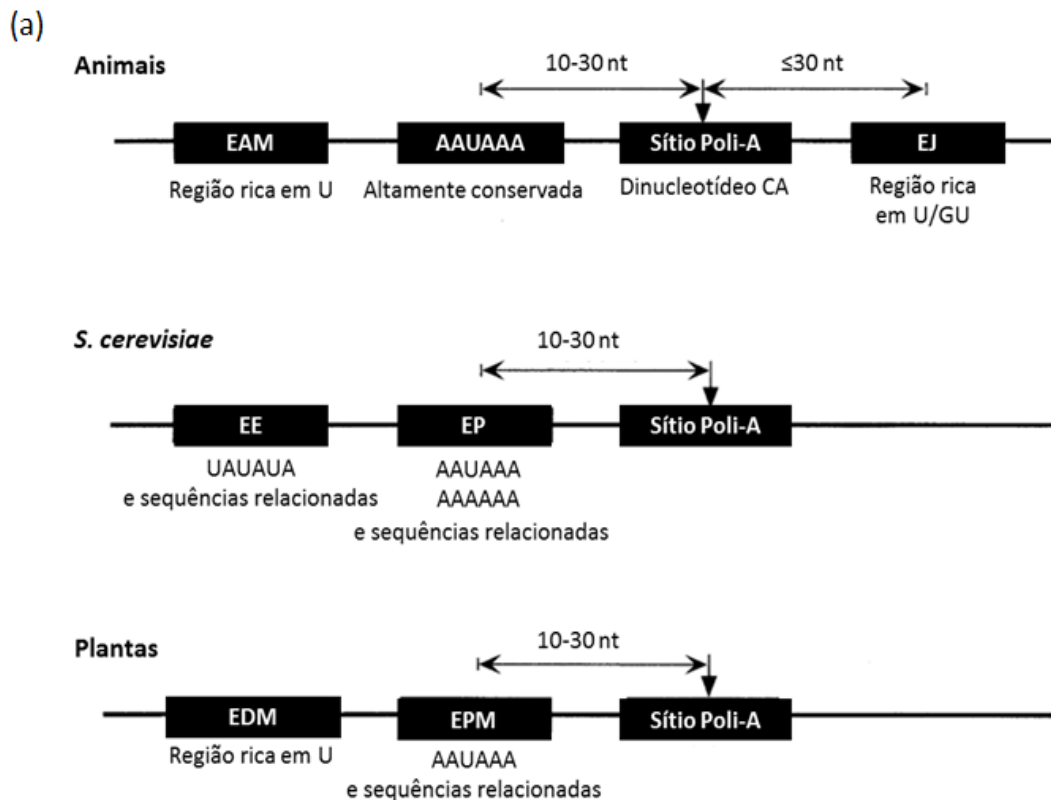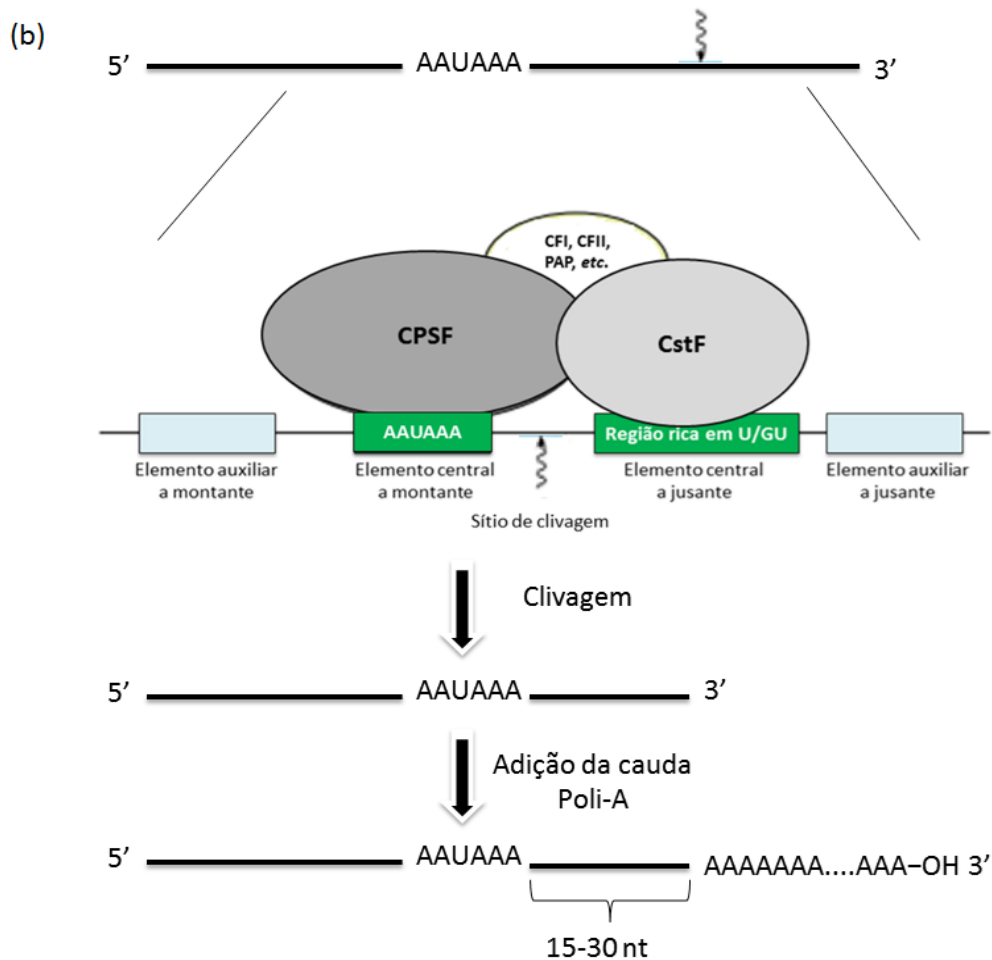
Figura 4. Modificações da extremidade 3' no processamento de mRNA em eucariotos (b), destacando a alta conservação da SSP AAUAAA no transcrito primário (a) e seu reconhecimento pelo complexo proteico de clivagem do mRNA (b). nt, nucleotídeos; EAM, *enhancer* auxiliar a montante; EJ, elemento a jusante; EE, elemento de eficiência; EP, elemento de posicionamento; EDM, elemento distante a montante; EPM, elemento próximo a montante (a). Proteínas que fazem parte da maquinaria central de clivagem e poliadenilação em mamíferos: CPSF, *cleavage and polyadenylation specificity factor*; CstF, *cleavage and polyadenylation stimulation factor;* CFI e CFII (fatores de clivagem I e II); PAP, poli-A polimerase (b). **Modificado de Tian *et al.*, 2005; Lutz, 2008 (a); e Zhao *et al.*, 1999 (b).**

Mais recentemente, dois estudos confirmaram a associação de rs78378222[C] com glioma: um deles utilizando um grande número de casos do Norte da Europa e EUA (OR=3.74) e o outro realizado em uma coorte menor dos EUA (OR=3.5) (Egan *et al.*, 2012; Enciso-Mora *et al.*, 2013). Esse SNP parece ser responsável por cerca de 6% do risco familiar para desenvolvimento de glioma, uma porcentagem maior do que o impacto

conjunto de todas variantes de risco comuns relacionadas a esse tipo de câncer (Enciso-Mora *et al.*, 2013). Egan *et al.* (2012) examinaram pela primeira vez o impacto desse SNP no prognóstico de 413 pacientes com glioma e, curiosamente, constataram uma sobrevida significativamente aumentada no grupo de pacientes portadores do alelo C quando comparados ao grupo de homozigotos selvagens. Esse aumento de sobrevida entre portadores de uma variante genética que aumentaria a incidência do tipo de tumor avaliado foi um achado inesperado, sendo levantada a hipótese de que os níveis reduzidos da proteína p53 selvagem, associados com a presença de rs78378222[C], forneceriam uma atividade residual de p53 suficiente para manter um fenótipo menos agressivo. Porém, em 2013, foi verificada uma ausência de relação entre o genótipo para rs78378222 e a sobrevida global em 1699 casos de glioma, indicando que este pode não ser um marcador de prognóstico clinicamente útil (Enciso-Mora *et al.*, 2013).

Em outros dois estudos também foi demonstrada uma associação de rs78378222 [C] com carcinoma esofágico de células escamosas (OR=3.22) na população chinesa, e com neuroblastoma (OR=2.3) em coortes de ancestralidade europeia e africana (Diskin *et al.*, 2014; Zhou *et al.*, 2012). No entanto, a mesma variante foi associada com uma proteção contra carcinoma de células escamosas de cabeça e pescoço em uma coorte de caucasianos dos EUA (Guan *et al.*, 2013). A tabela 1 resume os principais achados dos estudos de associação entre rs78378222[C] e diversos tipos de tumores, publicados até o presente momento.

Apesar de alguns tipos de tumores associados com rs78378222[C] fazerem parte do espectro da SLF/LFL, como os tumores cerebrais e câncer de próstata, os trabalhos publicados até o momento não relatam a investigação da presença de história familiar de câncer entre os indivíduos portadores da variante. Apenas no estudo de Stacey *et al.* (2011) a presença de história familiar foi considerada, sendo que não foi identificado nenhum paciente SNP-positivo com critérios para síndrome de SLF ou LFL.

Tabela 1. Amostragem e resultados dos estudos de associação entre a variante rs78378222 (A>C) e o risco para diversos tipos de neoplasias.

Tipo de tumor	Populações estudadas	n (casos/controles)	Frequência em casos**	*P*	OR (IC 95%)	Referência
Carcinoma basocelular	Islândia, Dinamarca, Espanha, Bulgária, Eslováquia e Romênia	4.319/>51.810*	0.004 – 0.0442	**2.2X10[-20]**	2.16 (1.83-2.54)	Stacey *et al.* 2011
Câncer de próstata	Islândia, Espanha, EUA, Holanda, Reino Unido e Romênia	7.790/>50.629*	0.007 – 0.0258	**2.4X10[-6]**	1.44 (1.24-1.68)	
Adenoma colorretal	Islândia	4.095/>43.022*	0.0252 – 0.0294	**1.6X10[-4]**	1.39 (1.17-1.65)	
Glioma	Islândia e EUA	1.395/>45.937*	0.0228 – 0.0486	**1.0X10[-5]**	2.35 (1.61-3.44)	
	EUA	566/603	0.037	**1.0X10[-3]**	3.54 (1.87-6.71)	Egan *et al.* 2012
	EUA, França, Reino Unido e Alemanha	4.147/7.435	0.023 – 0.032	**6.86X10[-24]**	3.74 (2.89-4.83)	Enciso-Mora *et al.* 2013
Carcinoma esofágico de células escamosas	China (etnia Han)	405/810	0.031	**1.34X10[-4]**	3.22 (1.71-6.33)	Zhou *et al.* 2012
Neuroblastoma	Coortes de ancestralidade europeia, africana e italiana	2.801/7.446	0.009 – 0.027	**2.03X10[-11]**	2.3 (1.8-2.9)	Diskin *et al.* 2014
Câncer colorretal	Islândia, Espanha, EUA, Holanda e Suécia	5.796/>47.749*	0.0084 – 0.0218	0.51	1.06 (0.89-1.27)	Stacey *et al.* 2011
Câncer de mama	Islândia, Espanha e Holanda	4.985/>42.995*	0.005 – 0.0191	0.57	1.06 (0.88-1.27)	
Melanoma	Islândia, Espanha e Holanda	2.520/>46.644*	0.0017 – 0.0234	0.64	1.07 (0.81-1.42)	
	Coortes dos EUA de caucasianos não-hispânicos	1.329/3.000	0.0139	0.682	1.14 (0.77-1.71)	Guan *et al.* 2013
Câncer de pulmão	Coortes dos EUA de caucasianos não-hispânicos	1.013/3.000	0.01037	0.382	0.84 (0.51-1.37)	
Carcinoma de células escamosas de cabeça e pescoço	Coortes dos EUA de caucasianos não-hispânicos	1.096/3.000	**0.00593**	**0.008**	0.44 (0.24-0.8)	

IC: intervalo de confiança. * Tamanho amostral efetivo leva em conta a eficiência da genotipagem *in silico* empregada no estudo. ** Valores citados como (n - n) representam a amplitude da frequência do alelo rs78378222[C] entre as populações analisadas.

1.4.2. Achados preliminares do grupo de pesquisa

Em um estudo ainda não publicado por nosso grupo de pesquisa, a presença de rs78378222[C] foi investigada em um estudo caso-controle envolvendo casos de CM hereditário e esporádico (n=213) do sul do Brasil. O alelo não foi identificado em nenhum dos grupos de casos, mas esteve presente em 3 dos 299 indivíduos controles, corroborando a ausência de associação com CM (*P*=1) encontrada por Stacey *et al.* (2011). Adicionalmente, buscamos determinar a frequência desse SNP funcional em pacientes com fenótipo SLF/LFL, com e sem mutações germinativas previamente identificadas no gene *TP53*, já que uma parcela significativa das famílias com critérios clínicos para essa síndrome não apresentam mutações em regiões codificantes do gene (Varley et al., 2003). Como resultado preliminar, a variante não foi detectada no grupo de pacientes SLF/LFL com mutação patogênica em *TP53* (n=81), mas foi detectada em 7 dos 130 pacientes com fenótipo clínico da síndrome e sem mutações patogênicas, uma frequência significativamente maior do que aquela identificada na população controle (5,4% *vs.* 1%, *P*=0,01). Os 7 pacientes portadores do SNP preenchiam critérios clínicos para a síndrome de LFL e haviam sido diagnosticados com os tumores "core", sendo que entre eles temos: 3 pacientes com CM, 2 delas diagnosticadas em idade jovem e outra com CM bilateral diagnosticado em idade avançada; e 4 pacientes com diferentes tipos de sarcoma: 1 caso de osteossarcoma e 1 de rabdomiossarcoma, ambos desenvolvidos na infância, 1 caso de sarcoma de tecidos moles diagnosticado aos 63 anos e 1 caso de sarcoma de Ewing aos 25 anos. Até o momento esse trabalho consiste na primeira descrição de rs78378222[C] em pacientes com SLF/LFL.

Estratégias adicionais estão sendo empregadas a fim de se confirmar a patogenicidade da variante. Recentemente, utilizamos ferramentas disponíveis no *USCS Genome Browser* (Fujita *et al.*, 2011; http://genome.ucsc.edu) com o intuito de analisar a conservação da SSP do gene *TP53*. Foi observado que o motivo AATAAA não apresenta nenhum tipo de variação em sua sequência dentre 63 espécies de mamíferos

analisadas. Além disso, para realização de estudos funcionais, foi realizada uma biópsia de pele e obteve-se uma cultura primária de fibroblastos de uma paciente portadora da variante. Foi observada uma produção elevada de ERO nas células dessa paciente em comparação com aquelas de pacientes que apresentam mutações germinativas em *TP53* e controles. Esses achados indicam uma perda da função antioxidante de p53 em portadores do alelo, representando mais uma evidência sobre o impacto funcional de rs78378222[C] (dados não publicados da tese de doutorado do MSc. Gabriel Macedo).

1.5. Câncer de mama (CM)

O CM é a segunda neoplasia mais frequente no mundo e a primeira entre as mulheres, sendo também responsável pela maior taxa de mortalidade por neoplasias em mulheres. Segundo dados do INCa (Instituto Nacional de Câncer), para o ano de 2014 estima-se o surgimento de 57.120 casos novos da doença no Brasil, o que indica um risco estimado de 56 casos a cada 100 mil mulheres. O sul do Brasil apresenta uma das maiores taxas de incidência de CM no país: cerca de 88 casos a cada 100 mil mulheres em certas regiões (INCa, 2014).

Os determinantes genéticos do CM ainda não estão totalmente caracterizados. Estima-se que 5 a 10% dos casos de CM estejam relacionados à predisposição hereditária associada a alterações em genes supressores de tumor, entre eles *BRCA1*, *BRCA2*, *CHECK2, TP53* e *PTEN* (Easton *et al.*, 1993; Ford *et al.*, 1998; Walsh *et al.*, 2006). Porém, estima-se que todas as mutações e SNPs conhecidos explicam menos de 30% do risco total ao CM (Brewster *et al.*, 2012). O CM também se destaca por ser o tipo de tumor mais frequente em famílias com SLF, correspondendo a 24-31,2% dos tumores diagnosticados neste grupo de pacientes (Kleihues *et al.*, 1997; Olivier *et al.*, 2003).

Atualmente, *microarrays* têm sido empregados na classificação molecular dos tumores de mama. O uso desta ferramenta, associada a achados clínicos, permitiu a identificação de 6 subgrupos de CM: (a)

luminais A, B e C, que apresentam positividade para receptores de estrogênio (RE) e expressão elevada (luminal A) ou moderada a baixa (luminal B e C) de genes tipicamente expressos pelas células epiteliais luminais da glândula mamária; (b) basal-*like*, também chamado de triplo-negativo (ausência de expressão de RE, receptores de progesterona e HER2) e com um padrão de expressão gênica semelhante ao do componente de células epiteliais basais/mioepiteliais; (c) normal *breast-like,* identificado pelo aumento da expressão de genes comumente expressos pelas células epiteliais normais da mama, células adiposas e outras células do estroma; e (d) tumores com superexpressão de HER2 (mais detalhes no tópico 1.5.2) (Perou *et al.*, 2000; Sørlie, 2004; Sørlie *et al.*, 2001). Devido ao seu menor custo, painéis de IHQ têm sido propostos para a identificação desses subtipos, buscando reproduzir com certa aproximação os perfis de expressão gênica (Bhargava *et al.*, 2008; Nielsen *et al.*, 2004).

1.5.1. Mutações no gene *TP53* e câncer de mama

Mutações somáticas em *TP53* são detectadas em aproximadamente 35% de todos os tumores de mama (*IARC TP53 Database*, *release R16*; 2012), sendo estas mais frequentes em subtipos moleculares com comportamento mais agressivo: tumores com superexpressão de HER2 (70%) e triplo-negativos (83%) (Langerød *et al.*, 2007). Por outro lado, a contribuição de mutações germinativas em *TP53* para o risco de CM não é completamente conhecida. Alguns estudos mostram que a frequência destas mutações em mulheres diagnosticadas com CM antes dos 30 anos de idade, não selecionadas pela história familiar, varia de 1 a 7% (Gonzalez *et al.*, 2009; Lalloo *et al.*, 2006; Mouchawar *et al.*, 2010; Walsh *et al.*, 2006).

Nos últimos quatro anos, resultados de três trabalhos indicaram que tumores de mama em portadores de mutação germinativa em *TP53* são predominantemente (63-83%), mas não exclusivamente, HER2-positivos (Wilson *et al.* 2010; Melhem-Bertrandt *et al.* 2012; Masciari *et al.* 2012). O estudo de Wilson *et al.* (2010) identificou que 83% (10 de 12) dos tumores de mulheres portadoras de mutação germinativa apresentavam

superexpressão de HER2, enquanto que este fenótipo foi identificado em apenas 19% (44 de 229) das pacientes sem mutação germinativa em *TP53*. Os resultados deste estudo, confirmados pelos outros dois trabalhos posteriores, sugerem uma forte associação entre a presença de mutações germinativas nesse gene e a superexpressão de HER2 em CM. Em contraste, apenas uma fração pequena dos casos de CM HER2-positivos estão associados a mutações germinativas em *TP53*. Um estudo publicado recentemente analisou 213 casos de CM HER2-positivos antes dos 50 anos de idade e observou-se uma prevalência de mutações germinativas em *TP53* de apenas 1,4% (Rath *et al.*, 2013). Já em um estudo conduzido por nosso grupo de pesquisa, de 61 pacientes com CM HER2 2+ ou 3+ diagnosticados antes dos 60 anos de idade, foram identificadas mutações germinativas (patogênicas) de *TP53* em 3,27% (dados não publicados da tese de doutorado da MSc. Mariana Fitarelli-Kiehl).

1.5.2. Câncer de mama com superexpressão de HER2

A proteína HER2 (*Human Epidermal Growth Factor Receptor 2*) é codificada pelo oncogene *ERBB2* (também conhecido como c-*erbB*-2), localizado no cromossomo 17q21-q22 (Coussens *et al.*, 1985). Esta proteína atua como um receptor transmembrana, com atividade tirosina quinase, em uma variedade de tipos celulares epiteliais, incluindo células epiteliais dos sistemas gastrointestinal, respiratório, reprodutivo e urinário, bem como na pele, mama e placenta (Press *et al.*, 1990). HER2 pertence a uma família de quatro receptores para fatores de crescimento epidermal (HER1 ou *Epidermal Growth Factor Receptor* - EGFR, HER2, HER3 e HER4) que apresentam homologia estrutural, com um domínio extracelular de ligação a ligante (com exceção de HER2), um domínio transmembrana e um domínio intracelular com função tirosina quinase (Figura 5) (Hynes e MacDonald, 2009; Moasser, 2007).

Quando superexpressa na superfície das células, HER2 é capaz de ativar vias de sinalização intracelular com funções oncogênicas (Gutierrez e Schiff, 2011). Além disso, tumores com superexpressão de HER2 são

particularmente agressivos, com aumento do potencial proliferativo e metastático (Ross *et al.*, 2003). A proteína HER2 apresenta-se superexpressa em aproximadamente 15 a 30% dos casos de CM (Collins *et al.*, 2012; Ross e Fletcher, 1998; Slamon *et al.*, 1989). Em cerca de 90% dos casos de CM HER2-positivos, a superexpressão da proteína HER2 está associada à amplificação (aumento do número de cópias) do oncogene *ERBB2*. Para os demais casos, os mecanismos de superexpressão ainda não estão totalmente esclarecidos, mas parecem estar relacionados com alterações na transcrição gênica, na meia-vida ou no processo de reciclagem da proteína HER2 (Magnifico *et al.*, 2007).

A análise dos níveis de expressão de HER2 por IHQ consiste num procedimento de rotina para todos os pacientes com diagnóstico de CM invasivo. Assim, os carcinomas de mama podem ser classificados em uma escala de 0 a 3+, sendo HER2 0 ou 1+ (negativos para expressão de HER2), HER2 2+ (expressão intermediária, ou seja, resultado indeterminado para superexpressão de HER2) e HER2 3+ (positivos para superexpressão de HER2). Esse nível de expressão de HER2 tem se tornado um importante biomarcador para indicação do uso de terapia anti-HER2 específica. Duas drogas estão aprovadas pela FDA (*Food and Drug Administration*) para tratamento de CM HER2-positivo: o trastuzumab (Herceptin®, Genentech) e o lapatinib (Tykerb®, GlaxoSmithKline) (Gutierrez e Schiff, 2011; Wolff *et al.*, 2013).
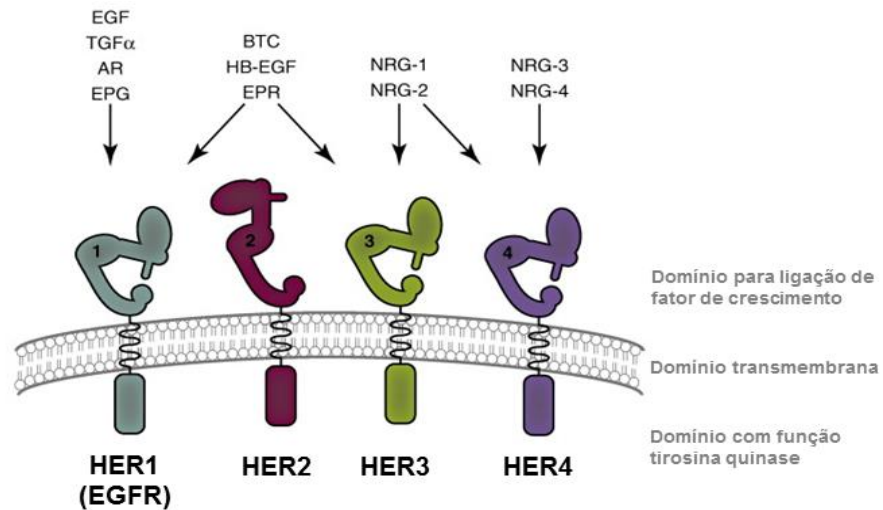
Figura 5. Representação comparativa da família de receptores para fatores de crescimento HER e seus ligantes externos. Modificado de Hynes e MacDonald, 2009.

1.6. microRNAs

Os microRNAs (também chamados de miRNAs ou miRs) consistem em uma classe de pequenos RNAs não-codificantes de fita simples (18 a 25 nucleotídeos de tamanho), envolvidos na regulação pós-transcricional da expressão gênica em todos eucariotos multicelulares (Ambros, 2004; Bartel, 2004). Estima-se que os miRNAs regulem aproximadamente 30% de todos os transcritos gênicos (Rachagani *et al*., 2010) e, atualmente, a principal base de dados de miRNAs contabiliza 2578 diferentes miRNAs humanos (miRBase versão 20.0; 2013).

Em geral, a transcrição dos genes codificadores de miRNAs é mediada pela RNA polimerase II. O transcrito primário, ou pri-miRNA, que possui 5'*cap* e cauda poli-A, é formado por algumas centenas a milhares de nucleotídeos. O pri-miRNA é clivado no núcleo pelo complexo enzimático formado pela RNAse III *Drosha* e proteínas associadas do Microprocessador, produzindo um pre-miRNA de aproximadamente 70 nucleotídeos. Este pre-miRNA é posteriormente clivado no citoplasma pelo complexo enzimático RNAse III *Dicer*/Argonauta, a fim de gerar o miRNA maduro, o qual é então incorporado ao complexo de silenciamento induzido por RNA (RISC), que atua no reconhecimento da região 3'UTR dos mRNAs-

alvo através de pareamento de bases imperfeito com o miRNA. O complexo miRNA e RISC (miRISC) regula negativamente a expressão gênica através de dois mecanismos pós-transcricionais: a degradação do mRNA, quando a complementariedade entre as bases do mRNA e do miRNA for total (mais comum em plantas); ou a inibição da tradução do mRNA, quando essa complementariedade for parcial (Bartel, 2004; Esquela-Kerscher e Slack, 2006). Neste último caso, evidências sugerem que a repressão pode ocorrer durante o início da tradução ou mesmo após este evento. Exemplos de mecanismos de repressão inicial são a deadenilação da cauda poli-A e o bloqueio da associação entre as subunidades ribossomais 60S e 40S promovidos pelo complexo miRISC. Os mecanismos posteriores ao início da tradução incluem: dissociação prematura dos ribossomos, diminuição da taxa ou parada no alongamento proteico e degradação do polipeptídeo nascente (Nilsen, 2007).

1.6.1. miRNAs em câncer

Cerca de 50% dos genes de miRNAs descritos em humanos estão localizados em regiões genômicas associadas ao câncer, pontos de quebra cromossômica e regiões de frequente perda de heterozigosidade, sugerindo que os miRNAs estão envolvidos na iniciação e progressão de diversos tipos de tumores (Calin *et al.*, 2004; Rachagani *et al.*, 2010). Adicionalmente, sabe-se que os miRNAs estão envolvidos na regulação de genes com papel crucial em todas etapas da tumorigênese, como apoptose, proliferação celular, angiogênese, invasão, metástase, entre outras (Ruan *et al.*, 2009).

Evidências recentes indicam que os miRNAs podem atuar como supressores de tumor ou oncogenes, sendo estes últimos chamados de "*oncomirs*". Em tecidos normais, o funcionamento adequado da transcrição, processamento e ligação dos miRNAs às sequências complementares resulta em taxas balanceadas de crescimento, proliferação, diferenciação e morte celular. Já em tecidos tumorais, duas alterações podem ser observadas em relação ao nível de expressão de miRNAs, ambas relacionadas com falhas em algum ponto da biogênese desses miRNAs: (a)

a redução ou eliminação da expressão de um miRNA que atua como um supressor de tumor, processo este que favoreceria a formação de tumores devido a não-repressão da expressão de oncoproteínas; e (b) o aumento nas taxas de expressão de miRNAs que atuam como oncogenes (Esquela-Kerscher e Slack, 2006).

1.6.2. Predição de sítios-alvo para ligação de miRNAs

Já que comumente os miRNAs ligam-se aos seus alvos através de complementariedade imperfeita de bases, a identificação de mRNAs-alvo tem sido cada vez mais complexa. Muitos estudos indicam que a porção 5' (sequência *seed*) do miRNA é crucial para a estabilidade e o carregamento da molécula no complexo RISC, e representa uma região de pareamento perfeito com o mRNA-alvo. Por este motivo, ferramentas de bioinformática para predição de alvos levam em consideração essencialmente esta sequência, que corresponde aos nucleotídeos 2 a 8 da sequência do miRNA maduro. Estudos resultantes dessas análises mostram que um único miRNA pode regular negativamente múltiplos alvos (até 200 mRNAs), com funções completamente distintas (Esquela-Kerscher e Slack, 2006; Ruan *et al.*, 2009). Também tem sido descrito múltiplos sítios de ligação para um mesmo miRNA em um determinado mRNA, o que pode aumentar o nível de regulação da expressão (Grimson *et al.*, 2007).

Como os sítios para ligação de miRNAs estão predominantemente localizados na região 3'UTR dos mRNAs, espera-se que variantes somáticas ou germinativas nessas sequências-alvo possam alterar a expressão gênica por abolir, enfraquecer ou criar complementariedade com miRNAs, possivelmente levando à variação fenotípica (Endale Ahanda *et al.*, 2012). Já é bem conhecido que alterações na expressão gênica representam um dos mecanismos-chave a partir dos quais células normais progridem a um estado tumoral, portanto, polimorfismos germinativos nessas sequências-alvo podem influenciar no risco individual para o desenvolvimento de câncer (López de Silanes *et al.*, 2007). Um exemplo é o caso do SNP rs61764370 na região 3'UTR do proto-oncogene *KRAS*, cuja presença rompe a ligação

do miRNA *let-7*, resultando em uma superexpressão de *KRAS* e risco aumentado para o desenvolvimento de câncer de pulmão (Chin *et al.*, 2008).

Um número crescente de estratégias de bioinformática tem sido proposto para avaliar as interações entre miRNAs e regiões 3'UTR de diferentes transcritos. A maioria delas exige uma complementariedade de bases perfeita entre a sequência *seed* do miRNA e o seu sítio de ligação na região 3'UTR. Algumas utilizam critérios adicionais, como a conservação filogenética desses sítios-alvo (Friedman *et al.*, 2009; Gaidatzis *et al.*, 2007), a estabilidade termodinâmica do pareamento miRNA:mRNA (Rehmsmeier *et al.*, 2004) e características estruturais da região 3'UTR, como a acessibilidade ao sítio-alvo ou contexto local (Grimson *et al.*, 2007; Long *et al.*, 2007; Marín e Vaníček, 2011). No estudo de Endale Ahanda *et al.* (2012), predições combinadas das ferramentas *TargetScan* (Grimson *et al.*, 2007), *PACMIT* (Marín e Vaníček, 2011) e *TargetSpy* (Sturm *et al.*, 2010) considerando diferentes parâmetros biológicos, permitiram a identificação de sequências-alvo para ligação de miRNAs e a descoberta de SNPs na região 3'UTR de genes que codificam o MHC suíno.

Estudo recente de Li *et al.* (2013) avaliou o impacto funcional de 3 variantes anteriormente não descritas na região 3'UTR de *TP53*, detectadas a partir de amostras de tumores de pacientes com linfoma de grandes células B difuso. Conforme predito por *TargetScan* e confirmado por análises de expressão, essas 3 variantes aboliam a complementariedade entre o transcrito *TP53* e a sequência *seed* de miR-125b, impedindo, dessa maneira, a repressão da expressão de p53 por esse miRNA, o que resultou em um aumento na expressão de p53 *in vitro* em comparação à sequência 3'UTR selvagem. No entanto, devido o tipo de amostra analisada, essas variantes novas podem ser de origem somática (Li *et al.*, 2013).

1.6.3. Regulação da expressão de p53 por miRNAs

Vários miRNAs contribuem para um controle refinado da expressão de p53 ao interagir diretamente com sítios-alvo no mRNA *TP53* e devem ser oncogenes clinicamente relevantes. Até o momento, já foram validados

experimentalmente 12 miRNAs que, dessa maneira, regulam a expressão do gene *TP53* em humanos: miR-25, miR-30d, miR-33, miR-125b, miR-504, miR-380-5p, miR-92a, miR-141, miR-200a, miR-1285, miR-15a e miR-16; sendo que apenas miR-33, miR-504 e miR-1285 apresentam 2 sítios-alvo no mRNA *TP53* (Hermeking, 2012; Li *et al.*, 2013) Estudo funcional e em modelo animal observou que a superexpressão de miR-504 reduz os níveis da proteína p53 e prejudica suas funções, promovendo o crescimento de tumores *in vivo* (Hu *et al.*, 2010a). Além disso, conforme indica a figura 6, existem vários exemplos de mecanismos de regulação indireta de p53 através de uma rede de miRNAs que diminuem a expressão de reguladores a montante da sua via de ativação (Hermeking, 2012; Li *et al.*, 2013).

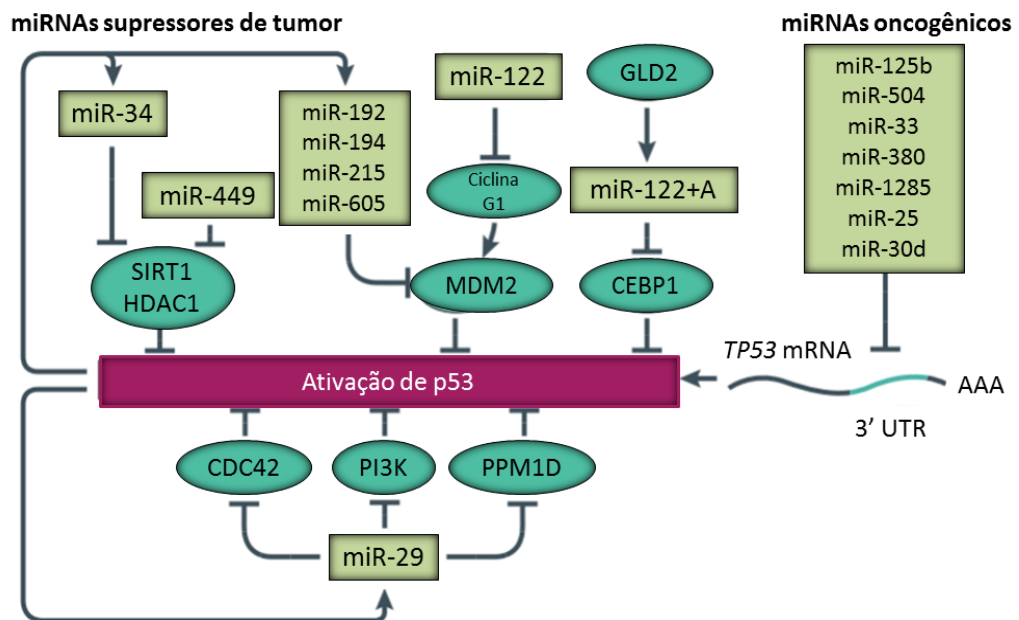**Figura 6. Regulação direta e indireta da atividade de p53 por miRNAs.** Os níveis de expressão de p53 podem ser regulados diretamente por miRNAs oncogênicos que apresentam sítios-alvo na região 3'UTR do mRNA *TP53* (à direita), bem como indiretamente pela atuação de miRNAs supressores de tumor que inibem a expressão de proteínas reguladoras negativas da ativação de p53 (à esquerda). **Modificado de Hermeking, 2012.**

2. JUSTIFICATIVA

Considerando o comportamento agressivo dos tumores de mama HER2-positivo, um melhor entendimento dos mecanismos genéticos que promovem a superexpressão deste receptor, particularmente em relação à amplificação do oncogene *ERBB2*, é crucial para aprimorar as estratégias de tratamento já existentes. Por outro lado, vários estudos apontaram a instabilidade genômica, evidenciada por um número aumentado de CNVs (*Copy Number Variations*), como uma característica fenotípica presente em pacientes com SLF ou LFL com mutações germinativas em *TP53* (Shlien *et al.*, 2008; Silva *et al.*, 2012). Tumores de mama nesse grupo de pacientes são, predominantemente, HER2-positivos. Nesse contexto, e sabendo-se que o SNP rs78378222 tem sido associado a uma redução na expressão de p53 e, possivelmente, a uma perda parcial nas suas funções supressoras de tumor, nossa hipótese é de que essa variante germinativa possa contribuir para o estabelecimento de instabilidade genômica e, por sua vez, estar relacionada com o mecanismo de amplificação (aumento do número de cópias) de *ERBB2*. Sendo assim, seria relevante avaliar a sua frequência em casos de CM HER2-positivos.

Paralelamente, mostra-se relevante confirmar e adicionar evidências acerca da patogenicidade desse SNP, bem como elucidar seu papel como uma determinante causal do fenótipo LFL nos pacientes portadores do alelo rs78378222[C] identificados previamente pelo grupo. Até o momento, a expressão proteica de p53 não foi avaliada em amostras biológicas de pacientes portadores da variante. A confirmação do papel causal dessa variante teria uma implicação direta no diagnóstico molecular, apontando a necessidade de uma análise mais abrangente do gene *TP53*, incluindo essa e demais sequências regulatórias. Por fim, sendo o SNP localizado em uma região 3'UTR, torna-se pertinente utilizar ferramentas de predição *in silico* descritas na literatura para investigar se este altera algum(s) sítio(s) de ligação para miRNA(s), o que poderia apontar um novo mecanismo de doença associado a essa variante.

3. OBJETIVOS

3.1. Objetivos gerais

Determinar a frequência da variante germinativa rs78378222 em pacientes com CM HER2-positivo e estudar os mecanismos moleculares de patogenicidade dessa variante.

3.2. Objetivos específicos

3.2.1. Determinar as frequências alélicas e genotípicas do SNP rs78378222 em uma amostra de mulheres com CM HER2-positivo do sul do Brasil;

3.2.2. Investigar potenciais mecanismos de patogenicidade do alelo mutante p53 contendo o SNP rs78378222 mediante:

(a) Avaliação do nível de expressão da proteína p53 em tecido tumoral conservado em blocos de parafina de pacientes SLF/LFL portadores do SNP rs78378222, comparado ao de portadores de diferentes mutações patogênicas em regiões codificadoras do gene *TP53*;

(b) Análise do nível de expressão de p53 após indução de estresse celular (por exposição à radiação ionizante), a partir de cultura primária de fibroblastos, nos seguintes grupos experimentais: pacientes SLF/LFL portadores do alelo rs78378222[C], pacientes SLF/LFL portadores de mutações patogênicas no gene *TP53* e controles (sem mutações identificadas em *TP53*);

(c) Investigação da possibilidade da variante rs78378222 romper e/ou criar sítio(s) de ligação para miRNAs na região 3'UTR do gene *TP53*.

4. TRABALHO EXPERIMENTAL NA FORMA DE ARTIGO CIENTÍFICO

Formatado para submissão no periódico *Cancer genetics*

Rare germline variant (rs78378222) in the *TP53* 3'UTR: additional evidence for a new mechanism of cancer predisposition

Igor Araujo Vieira[a], Gabriel de Souza Macedo[a,b], Mariana Fitarelli-Kiehl[a,b], Juliana Giacomazzi[a,b], Vanessa Beck Sempé[a], Patrícia Santos-Silva[a,c], Cristina Brinkmann Oliveira Netto[d], Patricia Ashton-Prolla[a,b,c,d,e*]

[a] Laboratório de Medicina Genômica, Centro de Pesquisa Experimental, Hospital de Clínicas de Porto Alegre (HCPA), Porto Alegre, Rio Grande do Sul (RS), Brazil;

[b] Programa de Pós-Graduação em Genética e Biologia Molecular, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil;

[c] Programa de Pós-Graduação em Ciências Médicas: Medicina, UFRGS, Porto Alegre, RS, Brazil;

[d] Serviço de Genética Médica, HCPA, Porto Alegre, RS, Brazil;

[e] Departamento de Genética, Instituto de Biociências, UFRGS, 90040-060, and Centro de Pesquisa Experimental, HCPA, Porto Alegre, RS, Brazil.

*Corresponding author: Patricia Ashton-Prolla, Centro de Pesquisa Experimental, Hospital de Clínicas de Porto Alegre, Rua Ramiro Barcelos 2350, 90035-903, Porto Alegre, RS, BRAZIL, Tel: +55 51 3359 7473. Email address: pprolla@hcpa.ufrgs.br

Abstract

Germline mutations in *TP53* are the underlying defect of Li-Fraumeni Syndrome (LFS) and its variant, Li-Fraumeni-like (LFL) Syndrome, both autosomal dominant disorders characterized by predisposition to multiple early onset cancers. Recently, we identified a germline 3'UTR *TP53* variant, rs78378222 (A>C), at a high frequency in a cohort of LFS/LFL patients without coding germline *TP53* mutations, but a direct causative genotype-phenotype effect has not been demonstrated. Herein, we assessed rs78378222[C] frequency in HER2+ breast cancers (BC), an immunohistochemical phenotype frequently associated with LFS-related breast tumors. To further evaluate a potential association of the variant with the LFS/LFL phenotype, we analysed p53 expression in tumor specimens and fibroblasts from rs78378222[C] carriers. Additionally, we investigated whether rs78378222[C] affects microRNA (miRNA) binding sites in the 3'UTR of *TP53* mRNA using *in silico* tools. The rs78378222[C] allele was not identified among 105 HER2+ BC-affected patients. However, lower p53 protein levels were observed in tumors and fibroblasts from carriers when compared to non-carriers. Finally, *in silico* analyses indicated that rs78378222[C] could alter a putative target site of miR-545-3p, a novel miRNA predicted to directly target *TP53*. Taken together, these findings suggest a new mechanism by which a germline *TP53* alteration may result in increased cancer predisposition.

37

Introduction

The p53 tumor suppressor protein, encoded by the *TP53* gene, plays a critical role in the maintenance of genomic stability as well as in control of cell growth and apoptosis [1, 2]. *TP53* germline mutations are the underlying defect in Li-Fraumeni Syndrome (LFS) and its variant, Li-Fraumeni-like (LFL) Syndrome, both rare autosomal dominant disorders characterized by increased predisposition to a wide spectrum of early-onset cancers [3, 4]. *TP53* germline mutations are detected in about 77% of LFS and 40% of LFL families [5], and most of these are missense mutations located between exons 5-8, corresponding to the DNA-binding domain (DBD) of p53 [6,7].

Until recently, no genome-wide association study reported a significant association between any germline *TP53* polymorphism and any cancer with a *P*-value ≤ 10^{-7} (a threshold for genome-wide significance) [8]. This changed with the description of rs78378222 (A>C), a single nucleotide polymorphism (SNP) in the 3´UTR of *TP53* which was found to be strongly associated with skin basal cell carcinoma (OR=2.16, *P*= 2.2 x 10^{-20}). At a frequency of 0.0192 in the Icelandic population, this rare variant changes the highly conserved polyadenylation signal (PAS) sequence (AATA**A**A to AATA**C**A), leading to impaired 3′-end processing of the *TP53* transcript [9]. Further studies revealed its association with other types of cancer, including some described in LFS/LFL families, such as brain tumors and prostate cancer [9-14]. In addition to direct modification of a highly conserved region of the gene, a functional study showed that the presence of the rs78378222[C] allele dramatically lowers *TP53* mRNA levels and results in reduced p53 expression and cellular apoptosis [15].

In a recent study from our group, the rs78378222[C] allele was not identified among 81 probands with the LFS/LFL phenotype who were carriers of coding germline *TP53* mutations (exons 2-11). However, the heterozygous genotype, rs78378222[AC], was identified in 7/130 probands with LFS/LFL phenotypes and no identifiable germline *TP53* mutation, a significantly higher frequency than observed in a control group from the same geographic region of the cases (*P*=0.01). All probands met LFL criteria and had been diagnosed

with at least one of the "core" tumors of the syndrome, including breast cancers (BC) and soft tissue and bone sarcomas [16].

Molecular phenotype of breast tumors in carriers of germline *TP53* mutations is predominantly (63-83%) HER2-positive (HER2+) with features of an aggressive subtype [17-19]. The HER2 protein, encoded by the oncogene *ERBB2*, normally acts as a transmembrane receptor with tyrosine kinase activity in various epithelial cell types, and is overexpressed (HER2+) in approximately 15-30% of breast tumors [20-24]. HER2 overexpression is associated with gene amplification in about 90% of HER2+ BC cases [25]. On the other hand, recent studies identified an abnormally increased number of DNA copy number variations in LFS/LFL patients carrying germline *TP53* mutations [26, 27]. Given that rs78378222[C] leads to a reduction in p53 expression and possibly reduces its tumor suppressor functions, we hypothesized that it could be related to the mechanism of *ERBB2* amplification in breast tumors.

In the present study, we aimed to determine rs78378222[C] frequency in HER2+ BC cases from Southern Brazil. In order to further verify its pathogenicity and strengthen a previous association with the LFS/LFL phenotype, we characterized p53 expression in tumor specimens and normal fibroblasts from rs78378222[C] carriers in comparison to controls and carriers of different coding germline *TP53* mutations. Since miRNA target sites (miRTS) are mainly located in the 3'UTR of mRNAs, and germline variants in these target sequences may have a significant impact on gene expression by abolishing, weakening or creating miRTS [28,29], we also investigated potential effects of rs78378222[C] on miRTS in the *TP53* 3'UTR.

Materials and methods

Study groups and ethical aspects

The current study included 105 women with HER2+ BC, diagnosed at Hospital de Clínicas de Porto Alegre (HCPA) between years 2007 and 2012. The following inclusion criteria were applied: women with invasive breast

tumors diagnosed before age 60 years, and immunohistochemical analysis of BC showing HER2 overexpression (2+ and 3+). Patients were unselected for family history of cancer. An additional set of formalin-fixed, paraffin-embedded (FFPE) tumor/metastatic specimens from cancer-affected probands of LFS/LFL families was included in this study. Samples were divided into derived from (a) non-rs78378222[C] carriers LFS/LFL probands carrying coding germline *TP53* mutations (n=5), and (b) rs78378222[C] carriers LFL patients without identifiable germline *TP53* mutations (n=3). Primary fibroblast cultures derived from LFS/LFL patients carrying known pathogenic germline *TP53* mutations (n=3), rs78378222[C] SNP-positive patient with LFL phenotype (n=1) and a control sample from an individual without pathogenic *TP53* mutations (n=1) were obtained by skin biopsies after informed consent.

Clinical data on all recruited patients and families were obtained by interviews, specific questionnaires and/or review of medical records. All individuals/families recruited for the study provided written informed consent and authorized the publication of their pedigrees. The study was approved by the institutional research and ethics committees of the coordinating center, HCPA, under protocol #11-0427 for HER2+ BC recruitment and investigation, #08-080/08-023 and #100405 for tissue and fibroblast studies in LFS/LFL patients, respectively.

Molecular analysis

Peripheral blood samples were collected in EDTA vials and DNA isolated using the GE extraction kit (GE Healthcare Lifesciences, BR). Germline *TP53* mutation screening in HER2+ BC cases was previously performed by High Resolution Melting analysis followed by Sanger sequencing of the entire coding sequence and flanking intronic regions [Fitarelli-Kiehl *et al.*, unpublished results]. Mutation screening in LFS/LFL patients had been conducted as described in two previous studies by our research group [30, 31]. *TP53* rs7837822 genotyping was performed in duplicates by real-time PCR using allele-specific TaqMan® probes (Life Tech.

Inc., USA, reference number AHHS2U2, PN4331349), including positive and negative internal controls in all experiments. Briefly, each reaction contained 20ng of genomic DNA, 6.25ul of 2x MasterMix Genotyping TaqMan (Applied Biosystems Inc.), 0.31ul of probes specific for each allele (40x) and 4.94ul of DNase-free water. A StepOne PCR Real-Time System was used for all experiments, with an initial cycle of 10 minutes at 95°C, followed by 45 cycles of 15 seconds at 92°C and 1 minute at 60°C.

p53 immunohistochemistry

Immunohistochemistry (IHC) staining of p53 was performed on 4µm sections from FFPE tumor/metastatic tissue samples according to conventional methods. Blockade of nonspecific background staining using Protein Block reagent (Spring Bioscience Inc., USA) was followed by incubation overnight at 4°C with a monoclonal anti-p53 antibody (1:200 dilution, DO-7; Dako Corp., Denmark). Primary antibody was detected by application of the streptavidin-biotin-peroxidase complex (Polyvalent HRP Plus Kit, Spring Bioscience) and visualized with 3,3'-diaminobenzidine (DAB Substrate System, Spring Bioscience) according to manufacturer's protocol. Counter staining was done with Harris hematoxylin. Breast tumor sections with known high expression of p53 were used as positive control, and a tumor section stained without primary antibody was used as a negative control.

Cell culture and western blot

Primary fibroblast cultures were grown in a humidified incubator at 5% CO_2 and 37°C, and maintained in HAM F-10 medium, prepared with 1% penicillin/streptomycin and supplemented with 20% fetal bovine serum (all reagents from Gibco Laboratories., USA). Cultured cell flasks from each patient with distinct *TP53* genotype were either exposed to ionizing radiation (treated cells) at a dose of 1Gy, or not irradiated (untreated controls). Treated and untreated fibroblasts were collected and lysed, 20h after DNA damage

induction, and prepared for western blotting as previously described with minor modifications [32]. After 15 µg of protein was separated by SDS-PAGE and electroblotted, PVDF membranes were incubated overnight at 4 °C with primary antibody against p53 (1:500; DO-1; Santa Cruz Biotechnology, CA), followed by incubation with anti-mouse horseradish peroxidase-conjugated secondary antibody (1:4000; GE Healthcare Lifesciences) for 2h at 4 °C. Chemoluminescence was detected by ImageQuant LAS 500 (GE Healthcare Lifesciences). β-actin was used as a endogenous control and band density was analyzed using the ImageJ software.

Bioinformatic analysis

In silico prediction tools were used to investigate whether the rs78378222 SNP alters and/or creates any miRTS in the *TP53* 3'UTR, such as Target Scan Human (release 6.2), MicroCosm Targets (version 5), Target Spy, RNA22 microRNA target detection (version 2.0), RNAhybrid and PITA algorithm [33-39]. All of them, except RNA22 and PITA algorithm, require a stringent base complementarity between the miRNA seed sequence (nucleotides 2-8 at the animal miRNA 5' end) and its miRTS in the mRNA 3'UTR which is often sufficient for functional binding specificity [33, 40, 41]. Supplementary table A.1 shows the key criteria for prediction used by these algorithms. miRNA sequences were obtained from miRBase database (http://mirbase.org/) and aligned with the *TP53* 3'UTR reference sequence corresponding to the NM_000546 using MEGA 5.04 software. A *P*-value<0.05 was considered statistically significant in all analyses.

Results

rs78378222 frequency in HER2+ BC cases

Clinical information and IHC analysis for HER2, estrogen receptor (ER) and progesterone receptor (PR) are summarized in Table 1. 59 (56.19%) of the 105 breast tumors were classified as HER2 3+ and the

remaining, HER2 2+. About half (51.43%) of the cases were also positive for both ER and PR. Germline *TP53* mutations with known pathogenic/damaging effect had been previously detected in 2 cases (1.9%) and variants with uncertain significance were identified in 4 patients (3.8%). None of the 105 HER2+ BC-affected patients were carriers of the rs78378222[C] allele (Table 1).

p53 expression in rs78378222[C] carriers fulfilling LFL criteria

We assessed p53 expression by IHC in primary tumor/metastatic specimens from (a) LFS/LFL probands with known pathogenic germline *TP53* mutations, and (b) rs78378222[C] carriers LFL patients. Clinical features and sample origin of these patients are shown in Table 2. In the first group, specimens of FFPE primary tumors were encountered for all patients, and 3 of the 5 probands were carriers of a common Brazilian founder germline mutation, p.R337H, located in the tetramerization domain (exon 10) of the *TP53* gene [42-45]. All patients of this group met criteria for LFL, except one carrier of the p.R273H DBD mutation who had the classic LFS phenotype. In the second group, primary tumor specimen of only one rs78378222[C] carrier was available (Figure 1-A1); for the remaining patients metastatic tissue samples were obtained (Figure 1-A2, A3). In the group of rs78378222[C] carriers, we observed weak and focal p53 immunostaining regardless of the specimen origin (Figure 1-A), whereas in tumors from patients with coding germline *TP53* mutations, intense nuclear accumulation of p53 was seen (Figure 1-B1, B2).

In addition, we also found lower p53 protein levels in fibroblasts harboring the rs78378222[C] allele than in cells from the control (*TP53* wild-type, WT/WT) both in the non-irradiated condition and 20 hours after DNA damage induction (Figure 2). When not exposed to ionizing radiation, fibroblasts from the rs78378222[C] carrier showed a similar reduced p53 expression compared to those carrying p.R337H and R273H pathogenic mutations (Figure 2A). However, 20h after induction of DNA damage, we verified predominantly a lower p53 protein level in SNP-positive fibroblasts

than in cells derived from pathogenic mutation carriers (Figure 2B). These data indicate that the rs78378222 variant hinders p53 expression at baseline and after exposure to cellular stress. Pedigrees of all rs78378222[C] carriers whose biological materials were included in the current study are depicted in Figure 3.

Prediction of miRTS around rs78378222 in the *TP53* 3'UTR

Target Scan, TargetSpy, MicroCosm Targets and RNA 22 analysis identified that the rs78378222 variant is located in a putative miRTS. Though it is not exactly located in a seed match, this variant lies within a region where there is a perfect base complementarity between the human *TP53* 3'UTR and hsa-miR-545-3p 3' end (Figure 4). miR-545-3p is not included among the twelve miRNAs reported so far to target the human *TP53* gene [15, 46, 47]. Based on this initial finding, scores assigned by each algorithm to functionally validated miRNAs which directly regulate mRNA *TP53* ("reference" miRNAs), as well as to experimentally tested miRNAs that demonstrated not to target this transcript, were observed for comparison to the candidate miRNA data. Scores provided by these computational tools, with the exception of RNA22, indicated a more favorable miR-545-3p:WT 3'UTR interaction than that predicted to *TP53*-repressive reference miRNAs (Supplementary Table A.2).

PITA algorithm and RNAhybrid allowed to compare energetic scores for the heteroduplex formed between miR-545-3p and WT *TP53* 3'UTR as well as "variant 3'UTR" containing rs78378222[C]. Since they are energetic scores, the lower (more negative) its value, the stronger the binding of the miRNA to the given site is expected to be [38, 39]. According to RNAhybrid, WT sequence seemed more favorable for binding of miR-545-3p (figure 5-A1), because rs78378222[C] creates a larger internal loop in the heteroduplex secondary structure evidenced by an increased minimum free energy for hybridization to variant 3'UTR (Figure 5-A2). Finally, we inputted full-length and the latter 70bp of *TP53* 3'UTR encompassing rs78378222 position in PITA algorithm. In the first analysis, miR-545-3p arose as the

unique miRNA presenting a remarkable ΔΔG score variation with the rs78378222[C] presence (Figure 5-B). In the second, all miRNAs predicted to target the WT and variant 3'UTR were the same, including miR-545-3p and indicating that rs78378222[C] does not create a new miRTS in the *TP53* 3'UTR (Supplementary Table A.3). As PITA does not supply a representation of the miRNA-target interaction, we performed an alignment between the predicted miRNA sequences and variant 3'UTR that surprisingly showed miR-545-3p as the unique candidate miRNA presenting perfect seed pairing (Supplementary Figure A.1). Supporting RNAhybrid results, ΔΔG score assigned to miR-545-3p:variant 3'UTR interaction was higher than that for WT 3'UTR in both analyses, suggesting that the binding of miR-545-3p in the putative miRTS could be weakened by rs78378222[C]. However, scores provided by these tools did not clearly distinguish *TP53*-repressive candidate miRNAs from reference miRNAs (Supplementary Table A.2).

Discussion

Currently, germline mutations in the *TP53* gene are the only known genetic defect underlying LFS/LFL, and molecular diagnosis mainly focuses on the screening for mutations in coding (exons 2-8) and flanking intronic regions. Missense mutations, that cluster in highly conserved regions of the DBD, are the most common *TP53* germline mutations associated with these syndromes [48]. The known functional effects of such mutations include "loss" of the p53 suppressor functions; "gain of function", conferring new oncogenic activities to the protein; and dominant-negative activity over the WT allele, when the mutation is present in heterozygosis [49]. Recently, a germline variant occurring in the sole PAS of *TP53*, rs78378222, has been associated with certain tumor types that belong to the LFS/LFL spectrum [9, 11, 12, 14]. Unpublished results of our group have also identified that this variant is present at a significantly higher frequency (5.4%) in a cohort of LFS/LFL patients without identifiable coding germline *TP53* mutations when compared to controls (1%) (P=0.01) [16]. Previously, Stacey *et al.* reported that rs78378222[AC] heterozygotes expressed "somewhat less" *TP53*

transcript levels than WT homozygotes (*P*=0.041), being that correctly terminated polyA(+) mRNAs were produced predominantly (73%) from the WT allele (*P*=1.6 X 10−6) [9]. This latter finding was also reported by Disk *et al.* in samples from two rs78378222 heterozygous primary neuroblastomas [12].

In the current study, we reported for the first time the p53 expression analysis in tumoral and non-tumoral biological samples from rs78378222[C] carriers fulfilling LFL criteria. We identified lower p53 protein levels in tumor/metastatic specimens and fibroblasts from rs78378222[C] carriers, in agreement with previous results from *in vitro* studies performed by Li and colleagues involving transfection of exogenous p53 constructs [15]. Since the 3'UTR SNP occurs in the PAS signal of *TP53*, we expect that none of the downstream p53 protein modifications are significantly affected by the variant. Thus, it presents a unique instance of *TP53* alleles associated with reduced mRNA stability, in contrast with the phenotypic effects of those missense mutations in *TP53* coding sequence which produce a mutant protein.

To further explore protein expression we assessed the p53 immunostaining profile in tumors from coding germline *TP53* mutations carriers, p.R337H and p.R273H, in addition to the samples from SNP-positive patients. Comparing to p.R337H and p.R273H, which showed nuclear accumulation of p53, specimens from rs78378222[C] carriers presented weak and focal p53 immunostaining. Thus, this finding reinforces a phenomenon identified more than 20 years ago that the majority of *TP53* missense mutations lead to abnormal accumulation of p53 in the nucleus of tumor cells [50,51]. Given that rs78378222[C] is a rare variant, a limitation of our study was the lack of primary tumor samples for most rs78378222[C] carriers. Other rs78378222[AC] samples retrieved were from metastatic origin, which could create an important bias by the fact that this tissue usually accumulates numerous somatic alterations that may trigger the metastasis process. Although no significant difference in protein expression patterns between metastatic tumor and the single primary tumor samples was observed, confirmation of this finding should be sought with a larger number

of primary tumor samples from rs78378222[C] carriers. Interestingly, reinforcing our findings in tumors, western blot analysis in fibroblasts from the rs78378222[C] carrier also showed lower p53 protein levels compared to coding germline *TP53* mutation carriers and WT control after DNA damage induction. This is probably due to a faster degradation of the *TP53* transcript harboring rs78378222[C], displaying impaired polyadenylation.

To our knowledge, the present study is the first evaluating the frequency of rs78378222 SNP in HER2+ BC-affected patients. None of the participants in this cohort carried the rs78378222[C] allele. Consistent with this result, recently we did not identify any SNP-positive patient among 213 BC cases of non-specific molecular subtype (excluding LFS/LFL phenotype) [Macedo *et al.*, unpublished results]. Furthermore, Stacey *et al.* did not see any convincing evidence that the variant 'C' allele confers risk for BC (OR=1.06, *P*=0.57), neither specific associations with ER-negative BC (OR=1.15, *P*=0.61) or high-risk BC (OR=0.87, *P*=0.36) [9]. Taken together, these data suggest that rs78378222[C] may not be associated with BC, or at least not with HER2+ BC predominantly related to the *ERBB2* amplification. Nevertheless, further studies in larger cohorts of BC-affected patients are warranted in order to confirm these findings.

Based on outputs from four different online computational algorithms, other important results of the present study were the prediction of miR-545-3p (also referred simply as miR-545) as a possible novel miRNA to directly target the 3'UTR of mRNA *TP53*, and the finding that its putative miRTS in this 3'UTR sequence encompasses the rs78378222 position. In addition, energetic scores computed by RNAhybrid and PITA algorithm suggested that rs78378222[C] could weaken the binding of this candidate miRNA to its putative miRTS in the *TP53* 3'UTR. Recent *in vitro* study reported the first target gene regulated by miR-545-3p, showing that it binds to two target sites in the *BRCA1* 3'UTR. miR-545-3p was found to be expressed through much of the cell cycle and, when overexpressed, significantly down-regulated expression of endogenous *BRCA1* p220 (main gene product that performs its tumor supression function) [52]. Within this context, if miR-545-3p is validated

by functional assays as a *TP53*-repressive miRNA, a broader oncogenic function could be attributed for it.

Ultimately, several different lines of evidences contribute positively to the hypothesis that rs78378222[C] may be the causative molecular alteration leading to the LFL phenotype in those previously identified SNP-positive patients. First, many of the tumors identified in the 7 variant-positive families are characteristic of the LFS/LFL (i.e. soft tissue and bone sarcomas, brain tumors and BC) [16]. Second, previously published data [9, 12, 15] along with those presented here demonstrate that rs78378222 is a functional SNP that impairs proper termination and polyadenylation of the *TP53* transcript, resulting in reduced p53 expression. Third, although the numbers of our previous study are relatively small, the rs78378222[C] allele does not seem to co-segregate with other germline pathogenic *TP53* mutation [16]. Fourth, the polyadenylation signal sequence within which the variant occurs, the hexanucleotide AAUAAA, is located 10-30 bases upstream of the cleavage/polyadenylation site and is present in 90% of all sequenced polyadenylation elements known to date, being recognized as one of the most highly conserved sequence elements [53-55]. Unlike those high-penetrance germline mutations which cause LFS, polymorphic variants are classically expected to have more modest effects on phenotype. Although the frequency of this rare variant reaches about 3% in most reports to date and be classified as a polymorphism, we must not exclude the fact to mean a variant with low or intermediate effect. Supporting it, we saw in most studies conducting until now a higher frequency in cases than in controls.

In summary, our findings suggest a new mechanism of carcinogenesis associated to reduction in the *TP53* dosage (haploinsufficiency) in rs78378222[C] carriers. This highlights the need for a more comprehensive analysis of the *TP53* gene in the molecular diagnosis routine, including this one and remaining regulatory gene regions. Additionally, further functional and *in vivo* studies as well as analysis of larger series of families with the LFS/LFL phenotype should be undertaken to confirm the rs78378222[C] as a pathogenic variant associated with this disorder.

Acknowledgments

Referenccs

[1] Levine AJ. p53, the cellular gatekeeper for growth and division. Cell 1997;88:323-31.

[2] Vogelstein B, Lane D, Levine AJ. Surfing the p53 network. Nature 2000;408:307-10.

[3] Li FP, Fraumeni JF, Mulvihill JJ, Blattner WA, Dreyfus MG, Tucker MA, et al. A cancer family syndrome in twenty-four kindreds. Cancer Res 1988;48:5358-62.

[4] Malkin D, Li FP, Strong LC, Fraumeni JF, Nelson CE, Kim DH, et al. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science 1990;250:1233-8.

[5] Varley JM. Germline TP53 mutations and Li-Fraumeni syndrome. Hum Mutat 2003;21:313-20.

[6] Hainaut P, Hollstein M. p53 and human cancer: the first ten thousand mutations. Adv Cancer Res 2000;77:81-137.

[7] Petitjean A, Mathe E, Kato S, Ishioka C, Tavtigian SV, Hainaut P, et al. Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. Hum Mutat 2007;28:622-9.

[8] Deng Q, Becker L, Ma X, Zhong X, Young K, Ramos K, et al. The dichotomy of p53 regulation by noncoding RNAs. J Mol Cell Biol 2014.

[9] Stacey SN, Sulem P, Jonasdottir A, Masson G, Gudmundsson J, Gudbjartsson DF, et al. A germline variant in the TP53 polyadenylation signal confers cancer susceptibility. Nat Genet 2011;43:1098-103.

[10] Zhou L, Yuan Q, Yang M. A functional germline variant in the P53 polyadenylation signal and risk of esophageal squamous cell carcinoma. Gene 2012;506:295-7.

[11] Egan KM, Nabors LB, Olson JJ, Monteiro AN, Browning JE, Madden MH, et al. Rare TP53 genetic variant associated with glioma risk and outcome. J Med Genet 2012;49:420-1.

[12] Diskin SJ, Capasso M, Diamond M, Oldridge DA, Conkrite K, Bosse KR, et al. Rare Variants in TP53 and Susceptibility to Neuroblastoma. J Natl Cancer Inst 2014;106:dju047.

[13] Guan X, Wang LE, Liu Z, Sturgis EM, Wei Q. Association between a rare novel TP53 variant (rs78378222) and melanoma, squamous cell carcinoma of head and neck and lung cancer susceptibility in non-Hispanic Whites. J Cell Mol Med 2013;17:873-8.

[14] Enciso-Mora V, Hosking FJ, Di Stefano AL, Zelenika D, Shete S, Broderick P, et al. Low penetrance susceptibility to glioma is caused by the TP53 variant rs78378222. Br J Cancer 2013;108:2178-85.

[15] Li Y, Gordon MW, Xu-Monette ZY, Visco C, Tzankov A, Zou D, et al. Single nucleotide variation in the TP53 3' untranslated region in diffuse large B-cell lymphoma treated with rituximab-CHOP: a report from the International DLBCL Rituximab-CHOP Consortium Program. Blood 2013;121:4529-40.

[16] Macedo GS, Vieira IA, Cortez-Paixão VR, Giacomazzi J, Santiago KM, Achatz MI, et al. Identification of a rare germ-line variant in the TP53 3'UTR in individuals with the Li-Fraumeni-like phenotype: A new mechanism of cancer predisposition? (Unpublished results).

[17] Wilson JR, Bateman AC, Hanson H, An Q, Evans G, Rahman N, et al. A novel HER2-positive breast cancer phenotype arising from germline TP53 mutations. J Med Genet 2010;47:771-4.

[18] Melhem-Bertrandt A, Bojadzieva J, Ready KJ, Obeid E, Liu DD, Gutierrez-Barrera AM, Litton JK, et al. Early onset HER2-positive breast cancer is associated with germline TP53 mutations. Cancer 2012;118:908-13.

[19] Masciari S, Dillon DA, Rath M, Robson M, Weitzel JN, Balmana J, et al. Breast cancer phenotype in women with TP53 germline mutations: a Li-Fraumeni syndrome consortium effort. Breast Cancer Res Treat 2012;133:1125-30.

[20] Coussens L, Yang-Feng TL, Liao YC, Chen E, Gray A, McGrath J, et al. Tyrosine kinase receptor with extensive homology to EGF receptor

shares chromosomal location with neu oncogene. Science 1985;230:1132-9.

[21] Slamon DJ, Godolphin W, Jones LA, Holt JA, Wong SG, Keith DE, et al. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. Science 1989;244:707-12.

[22] Press MF, Cordon-Cardo C, Slamon DJ. Expression of the HER-2/neu proto-oncogene in normal human adult and fetal tissues. Oncogene 1990;5:953-62.

[23] Ross JS, Fletcher JA, Linette GP, Stec J, Clark E, Ayers M, et al. The Her-2/neu gene and protein in breast cancer 2003: biomarker and target of therapy. Oncologist 2003;8:307-25.

[24] Collins LC, Marotti JD, Gelber S, Cole K, Ruddy K, Kereakoglow S, et al. Pathologic features and molecular phenotype by patient age in a large cohort of young women with breast cancer. Breast Cancer Res Treat 2012;131:1061-6.

[25] Magnifico A, Albano L, Campaner S, Campiglio M, Pilotti S, Ménard S, et al. Protein kinase Calpha determines HER2 fate in breast carcinoma cells with HER2 protein overexpression without gene amplification. Cancer Res 2007;67:5308-17.

[26] Shlien A, Tabori U, Marshall CR, Pienkowska M, Feuk L, Novokmet A, et al. Excessive genomic DNA copy number variation in the Li-Fraumeni cancer predisposition syndrome. Proc Natl Acad Sci U S A 2008;105:11264-9.

[27] Silva AG, Achatz IM, Krepischi AC, Pearson PL, Rosenberg C. Number of rare germline CNVs and TP53 mutation types. Orphanet J Rare Dis 2012;7:101.

[28] Endale Ahanda ML, Fritz ER, Estellé J, Hu ZL, Madsen O, Groenen MA, et al. Prediction of altered 3'- UTR miRNA-binding sites from RNA-Seq data: the swine leukocyte antigen complex (SLA) as a model region. PLoS One 2012;7:e48607.

[29] López de Silanes I, Quesada MP, Esteller M. Aberrant regulation of messenger RNA 3'-untranslated region in human cancer. Cell Oncol 2007;29:1-17.

[30] Giacomazzi J, Selistre SG, Rossi C, Alemar B, Santos-Silva P, Pereira FS, et al. Li-Fraumeni and Li-Fraumeni-like syndrome among children diagnosed with pediatric cancer in Southern Brazil. Cancer 2013;119:4341-9.

[31] Giacomazzi J, Graudenz MS, Osorio CABT, Koehler-Santos P, Palmero EI, Zagonel-Oliveira M et al. Prevalence of the TP53 p.R337H mutation in breast cancer patients in Brazil. (Unpublished results).

[32] Zamin LL, Dillenburg-Pilla P, Argenta-Comiran R, Horn AP, Simão F, Nassif M, et al. Protective effect of resveratrol against oxygen-glucose deprivation in organotypic hippocampal slice cultures: Involvement of PI3-K pathway. Neurobiol Dis 2006;24:170-82.

[33] Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 2009;19:92-105.

[34] Garcia DM, Baek D, Shin C, Bell GW, Grimson A, Bartel DP. Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. Nat Struct Mol Biol 2011;18:1139-46.

[35] Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ. miRBase: tools for microRNA genomics. Nucleic Acids Res 2008;36:D154-8.

[36] Sturm M, Hackenberg M, Langenberger D, Frishman D. TargetSpy: a supervised machine learning approach for microRNA target prediction. BMC Bioinformatics 2010;11:292.

[37] Miranda KC, Huynh T, Tay Y, Ang YS, Tam WL, Thomson AM, et al. A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. Cell 2006;126:1203-17.

[38] Krüger J, Rehmsmeier M. RNAhybrid: microRNA target prediction easy, fast and flexible. Nucleic Acids Res 2006;34:W451-4.

[39] Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E. The role of site accessibility in microRNA target recognition. Nat Genet 2007;39:1278-84.

[40] Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004;116:281-97.

[41] Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell 2009;136:215-33.

[42] Ribeiro RC, Sandrini F, Figueiredo B, Zambetti GP, Michalkiewicz E, Lafferty AR, et al. An inherited p53 mutation that contributes in a tissue-specific manner to pediatric adrenal cortical carcinoma. Proc Natl Acad Sci U S A 2001;98:9330-5.

[43] Latronico AC, Pinto EM, Domenice S, Fragoso MC, Martin RM, Zerbini MC, et al. An inherited mutation outside the highly conserved DNA-binding domain of the p53 tumor suppressor protein in children and adults with sporadic adrenocortical tumors. J Clin Endocrinol Metab 2001;86:4970-3.

[44] Achatz MI, Olivier M, Le Calvez F, Martel-Planche G, Lopes A, Rossi BM, et al. The TP53 mutation, R337H, is associated with Li-Fraumeni and Li-Fraumeni-like syndromes in Brazilian families. Cancer Lett 2007;245:96-102.

[45] Garritano S, Gemignani F, Palmero EI, Olivier M, Martel-Planche G, Le Calvez-Kelm F, et al. Detailed haplotype analysis at the TP53 locus in p.R337H mutation carriers in the population of Southern Brazil: evidence for a founder effect. Hum Mutat 2010;31:143-50.

[46] Hu W, Chan CS, Wu R, Zhang C, Sun Y, Song JS, et al. Negative regulation of tumor suppressor p53 by microRNA miR-504. Mol Cell 2010;38:689-99.

[47] Hermeking H. MicroRNAs in the p53 network: micromanagement of tumour suppression. Nat Rev Cancer 2012;12:613-26.

[48] Malkin D. Li-fraumeni syndrome. Genes Cancer 2011;2:475-84.

[49] Brosh R, Rotter V. When mutants gain new powers: news from the mutant p53 field. Nat Rev Cancer 2009;9:701-13.

[50] Bártek J, Bártková J, Vojtěsek B, Stasková Z, Lukás J, Rejthar A, et al. Aberrant expression of the p53 oncoprotein is a common feature of a wide spectrum of human malignancies. Oncogene 1991;6:1699-703.

[51] Dowell SP, Wilson PO, Derias NW, Lane DP, Hall PA. Clinical utility of the immunocytochemical detection of p53 protein in cytological specimens. Cancer Res 1994;54:2914-8.

[52] Dimitrov SD, Lu D, Naetar N, Hu Y, Pathania S, Kanellopoulou C, et al. Physiological modulation of endogenous BRCA1 p220 abundance suppresses DNA damage during the cell cycle. Genes Dev 2013;27:2274-91.

[53] Tian B, Hu J, Zhang H, Lutz CS. A large-scale analysis of mRNA polyadenylation of human and mouse genes. Nucleic Acids Res 2005;33:201-12.

[54] Zhao J, Hyman L, Moore C. Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. Microbiol Mol Biol Rev 1999;63:405-45.

[55] Lutz CS. Alternative polyadenylation: a twist on mRNA 3' end formation. ACS Chem Biol 2008;3:609-17.

Figure legends

Figure 1. Immunohistochemical analysis of p53 expression in primary tumor and metastatic tissue sections from LFS/LFL patients with different *TP53* germline alterations. (A1, A2, A3) Specimens of rs78378222[C] carriers without coding region *TP53* germline mutations: rhabdomyosarcoma section with arrows showing cells positive for p53 (A1); lung section (metastasis of osteosarcoma) (A2); and pleura biopsy (metastasis of breast cancer) (A3). (B1, B2) Samples from carriers of coding germline *TP53* mutations: sarcoma section from a p.R273H mutation carrier (B1); and section of adrenocortical carcinoma from a p.R337H mutation carrier (B2). Positive control (breast tumor section) is also shown (B3). p53 is seen as brown particles and is located both in cytoplasm and nucleus. All magnifications x 400.

Figure 2. Western blot of p53 expression in fibroblasts non-irradiated (A) and 20 hours after irradiation (B) from rs78378222[C] carrier with LFL phenotype compared to a control individual (WT/WT) and LFS/LFL patients carrying coding germline *TP53* mutations (p.R337H and p.R273H). β-actin was used as a control. Genotypes for the *TP53* coding region and band intensity of p53 in relation to control individual are indicated. WT, wild-type allele.

Figure 3. Pedigrees of the rs78378222[C] carriers LFL probands without identifiable germline *TP53* mutations for whom p53 expression was analyzed in tumor/metastatic specimens (1-3) and primary cultured fibroblasts (4). Cancer-affected relatives are represented by blackened symbols. Arrows indicate the proband. Dx indicates age at diagnosis; (n) indicates current age; d, death age; CNS, central nervous system.

Figure 4. Outputs provided by Target Scan (A), TargetSpy (B), MicroCosm Targets (C) and RNA 22 v.2 (D) algorithms suggesting that rs78378222 variant is located in a perfect pairing region (no seed sequence) between hsa-miR-545-3p and the *TP53* 3'UTR. Although the predicted miRNA is designated by miR-545 generic identifier in databases from outputs A-C,

sequence comparison (considering T=U) indicates that it refers to the same mature miRNA mentioned in output D, miR-545-3p. miRNA seed region is underlined, while rs78378222 position is indicated by black rectangles in the miRNA-mRNA alignment. ID, number identifiers; 7mer-m8, Watson-Crick base-pairing from positions 2 to 8 of the miRNA 5'end; P_{CT}, probability of preferentially conserved targeting; NA, not available (A); Base, Poisson and Org P= different P-values calculated for target sites described in Website URL of MicroCosm Targets (C).

Figure 5. Predictions by RNAhybrid (A) and PITA algorithm (B) showing that the miR-545-3p:*TP53* 3'UTR interaction could be weakened by the rs78378222[C] variant allele. Since they provide energetic scores, the lower (more negative) score, the stronger the binding of the miRNA to the given site is expected to be.

(A) RNAhybrid outputs. Minimum free energy (mfe) is more favorable for binding of miR-545-3p with the wild-type (WT) 3'UTR sequence (A1) than with "variant 3'UTR" containing rs78378222[C], because this SNP creates a larger internal loop in the heteroduplex secondary structure (A2). Latter 200bp of *TP53* 3'UTR encompassing rs78378222 position was entered as input sequence. rs78378222 position is highlighted by black rectangles and miRNA seed region is underlined.

(B) Plot of ΔΔG scores assigned to selected candidate *TP53*-repressive miRNAs, including miR-545-3p, and some *TP53*-repressive miRNAs experimentally validated predicted by PITA algorithm to target both WT and variant *TP53* 3'UTR. It indicates miR-545-3p as the unique miRNA presenting considerable ΔΔG score variation with the presence of the rs78378222 SNP. Besides miR-545-3p, we selected other candidate miRNAs that showed any ΔΔG score variation. All *TP53*-repressive validated miRNAs shown no ΔΔG score variation. Full-length *TP53* 3'UTR was inputted in this analysis. ΔΔG score variation = (ΔΔG score for WT 3'UTR - ΔΔG score for variant 3'UTR).

*ΔΔG score variation for miR-545-3p= - 2.89. All other candidate miRNAs with ΔΔG score variation ≥ -0.15.

Figure 1
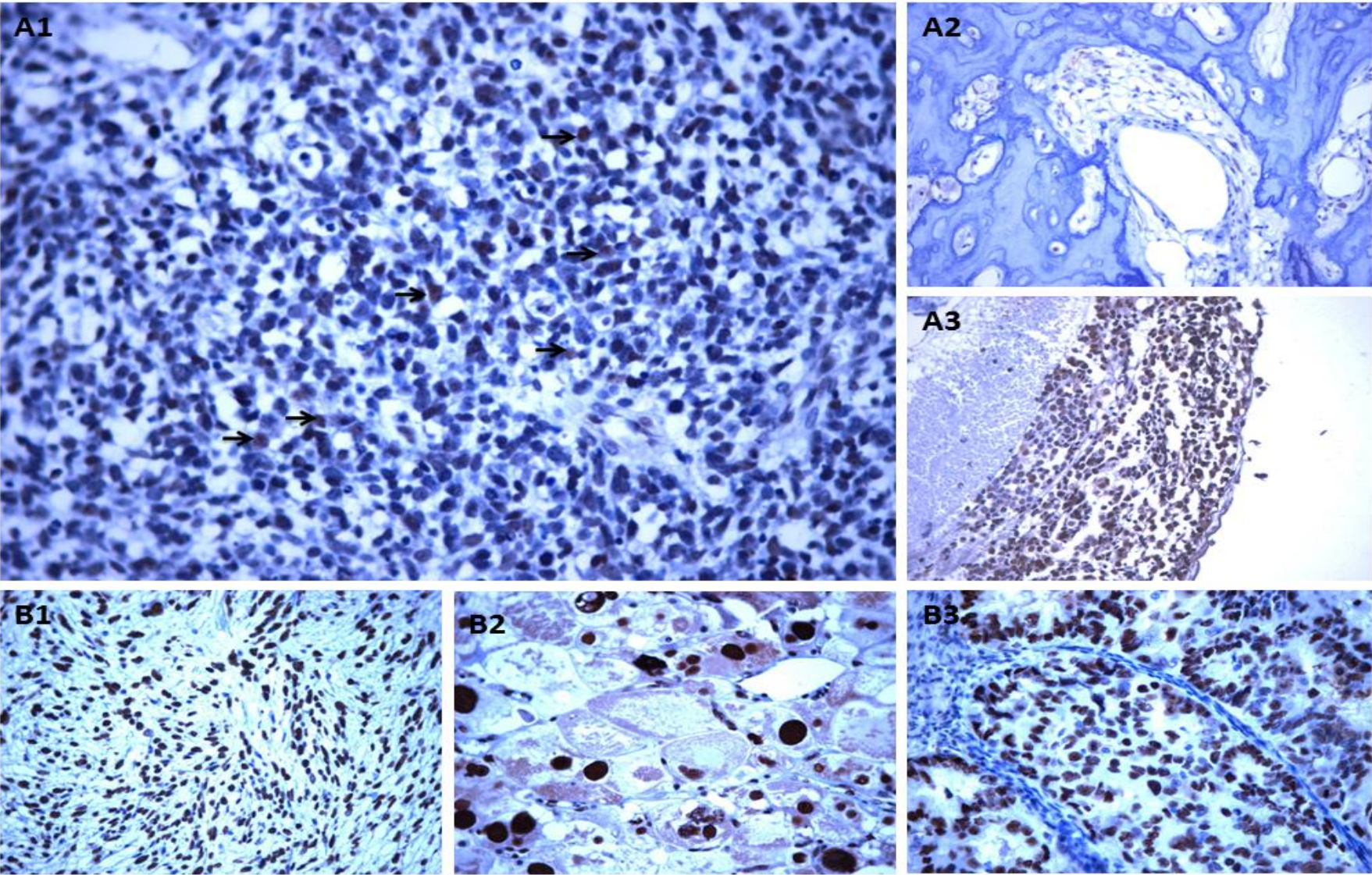
Figure 2

A Non-irradiated fibroblasts:
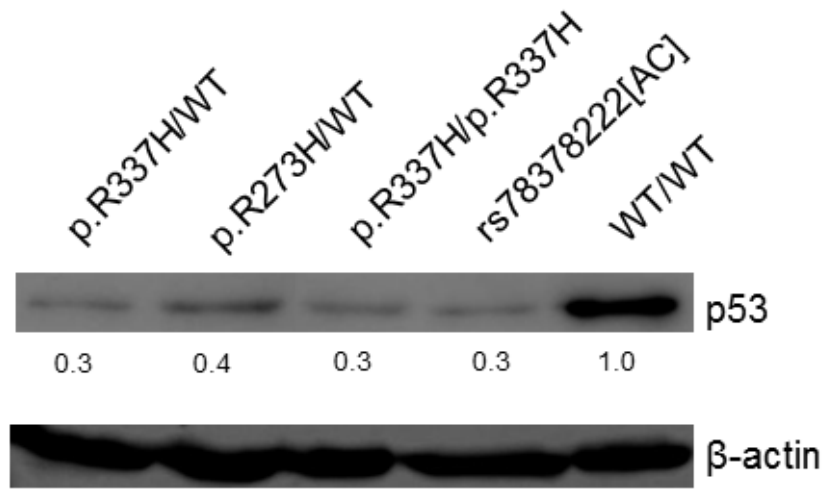
B 20h after irradiation:
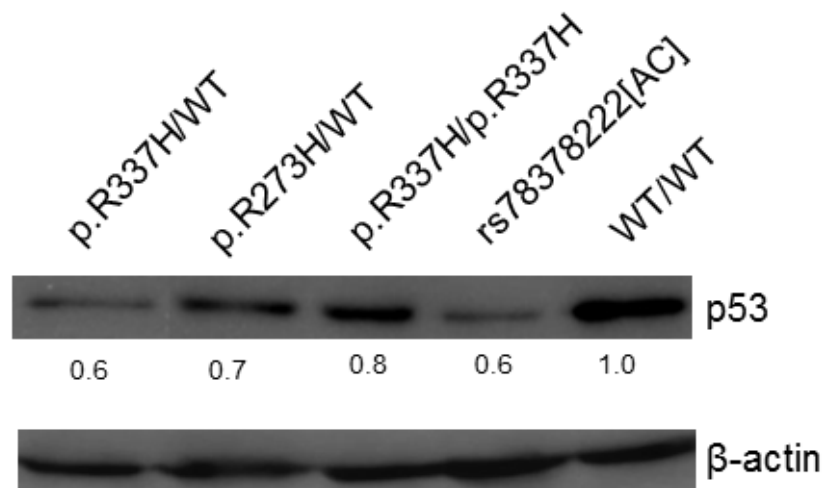
Figure 3

Pedigree 1
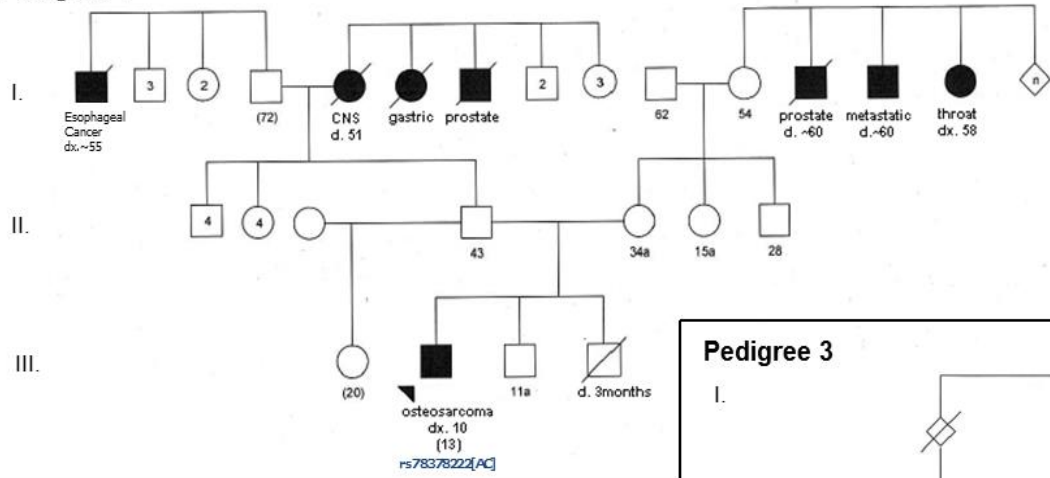
Pedigree 2
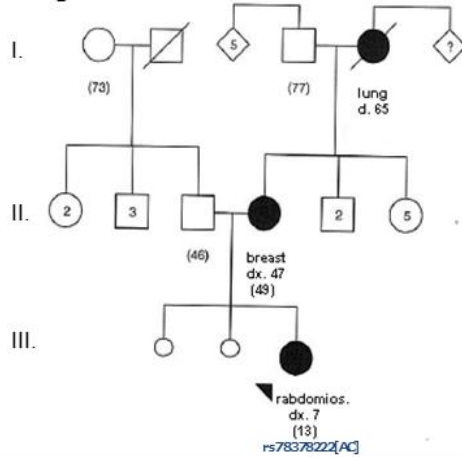
Pedigree 3

Pedigree 4

Figure 4

Figure 5

A1
TARGET : TP53_3UTR_WT
length: 200
MIRNA : hsa-miR-545-3p
length: 22

position   160
target 5'   U   AC   C 3'
              UACAAUAAA   UUUGCUG
              GUGUUAUUU   AAACGAC
miRNA 3' CGU   AC   U 5'

mfe: -20.6 kcal/mol

A2
TARGET : TP53_3UTR_variant
length: 200
MIRNA : hsa-miR-545-3p
length: 22

position   160
target 5'   U   CAAC   C 3'
              UACAAUA   UUUGCUG
              GUGUUAU   AAACGAC
miRNA 3' CGU   UUAC   U 5'

mfe: -17.7 kcal/mol

B

ΔΔG score
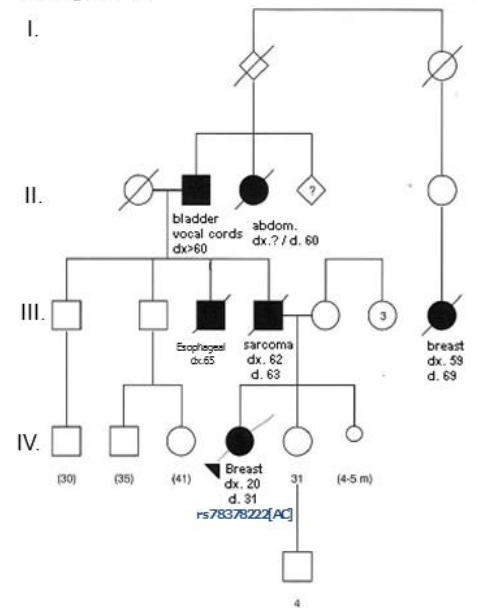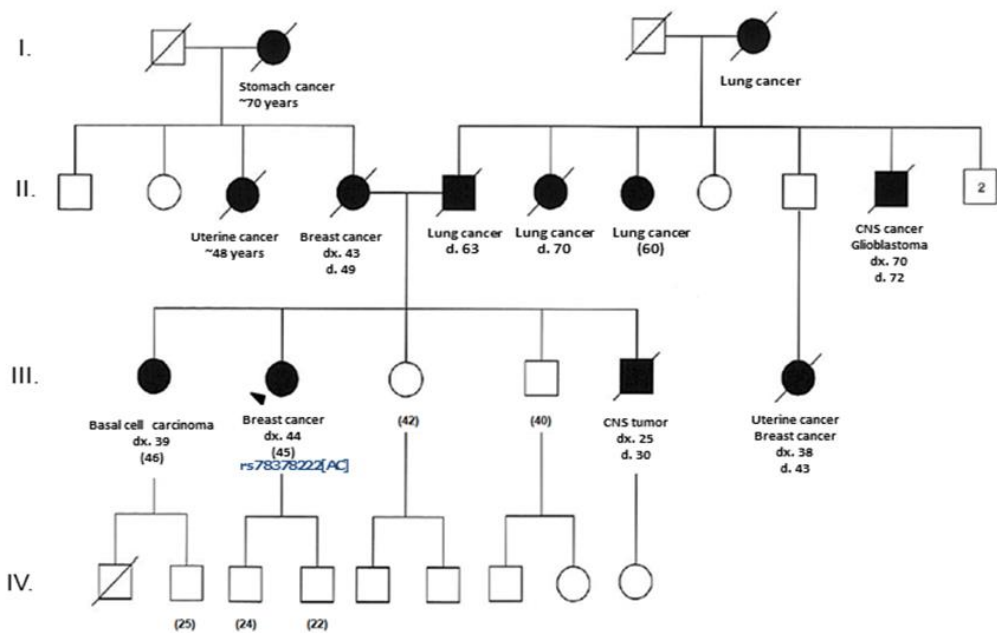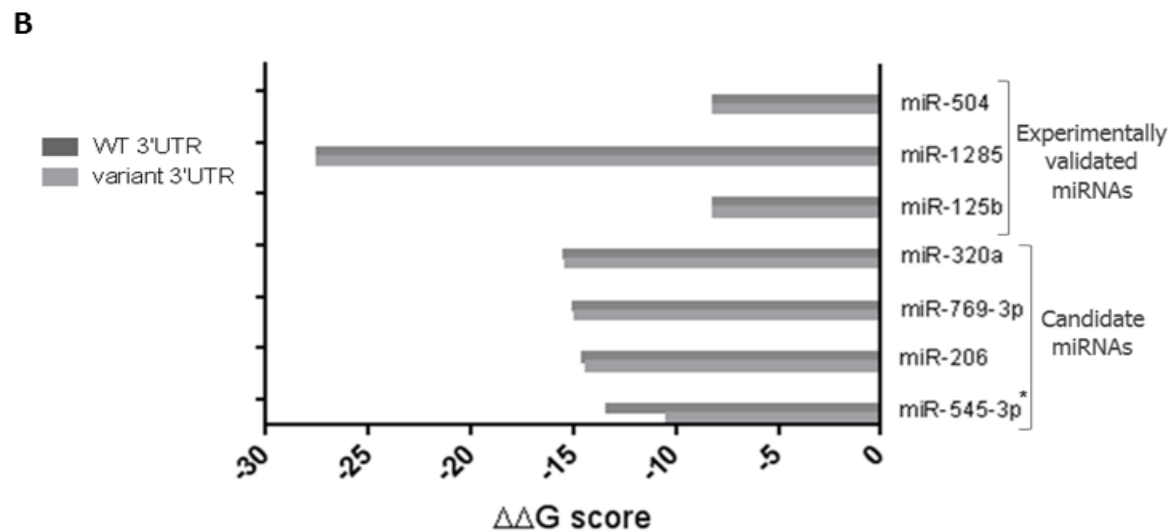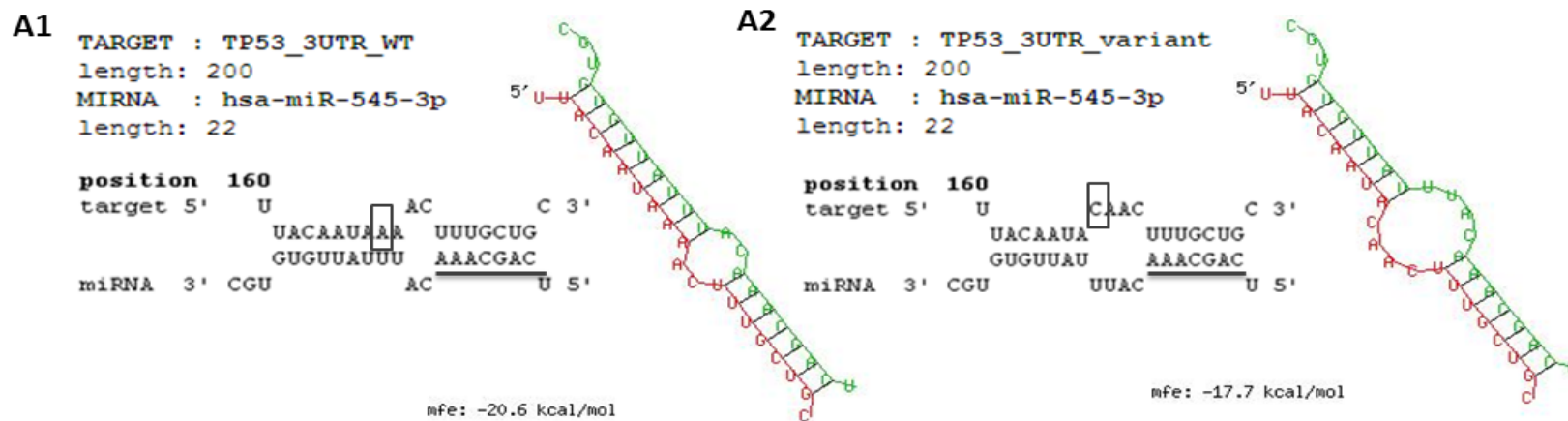
Table 1. Clinical features and rs78378222[C] frequency among HER2-positive BC-affected patients.

Study group description	HER2-positive BC cases (n=105)
rs78378222[C] frequency (%)	0
TP53 germline mutation carriers (%)	2 (1.9)[a]
Age at cancer diagnosis	
Mean (± SD)	46.48 (8.39)
Median	46.08
Cancer diagnosis ≤45y, n (%)	52 (49.52)
IHC analysis	
HER2 2+, n (%)	46 (43.81)
HER2 3+, n (%)	59 (56.19)
ER+, n (%)	66 (62.86)[b]
PR+, n (%)	57 (54.28)[b]
ER/PR+ combined, n (%)	54 (51.43)[b]

ER, estrogen receptor; PR, progesterone receptor; IHC, immunohistochemical.

[a] Germline pathogenic *TP53* mutation p.G244D was previously detected in 1 patient fulfilling classic LFS criteria. A possibly damaging mutation, p.K351E, was identified in other patient without clinical phenotype for SLF/LFL. In addition, 4 patients shown sequence variants with unknown effect on p53 function (all variants with minor allele frequency <0.01): 2 intronic alterations (c.993+12T>C and c.1100+30A>T) and 2 silent alterations (p.P36P and p.R213R).

[b] ER and PR IHC analyses were not available for 1 of the 105 patients.

Table 2. Features of cancer-affected LFS/LFL patients and their tumor/metastatic tissue specimens retrieved for the current study.

rs78378222 genotype	Primary tumor (PT) histology	Origin of sample	Criteria	Sex	Age at diagnosis of PT (years)	*TP53* genotype (coding region)
AC (Carriers)	Osteosarcoma[a]	Metastasis (lung)	LFL Chompret[a]	Male	10	WT
	Rhabdomyosarcoma[b]	Primary tumor	LFL Chompret[b]	Female	7	
	Breast cancer[c]	Metastasis (pleura)	LFL Chompret[c]	Female	26	
AA (Non-carriers)	Soft tissue sarcoma	Primary tumor	Classic LFS	Female	19	p.R237H/ WT
	Adrenocortical carcinoma	Primary tumor	LFL Chompret	Male	43	p.E258K/ WT
	Adrenocortical carcinoma	Primary tumor	LFL Chompret	Male	1,9	p.R337H/ WT
	Adrenocortical carcinoma	Primary tumor	LFL Chompret	Female	0,7	p.R337H/ WT
	Breast cancer	Primary tumor	LFL Chompret	Female	57	p.R337H/ WT

WT, wild-type.

[a] Immunohistochemistry (IHC) result is depicted in Figure 1-A2 and pedigree corresponds to the number 1 in Figure 3.

[b] IHC result in Figure 1-A1 and pedigree corresponds to the number 2 in Figure 3.

[c] IHC result in Figure 1-A3 and pedigree corresponds to the number 3 in Figure 3.

Supplementary data

Supplementary Figure A.1. Alignment between reverse complement sequences of the candidate miRNAs predicted by PITA algorithm and the latter 70bp of *TP53* 3'UTR containing rs78378222[C]. Note that miR-545-3p is the unique candidate miRNA presenting perfect base complementarity with the 3'UTR sequence in its seed region (black rectangle). As shown by the Supplementary Table A.3, in addition to miR-545-3p, miR-1207-5p also presented a remarkable ΔΔG score variation. This energy variation could favor the binding of miR-1207-5p in the target site (rs78378222[C] establishes a new pairing outside the seed region), but the alignment indicates that its seed sequence does not perfectly pairing with the *TP53* 3 'UTR (red circles). Minimum length accepted (70bp) was used as input in order to reduce bias of the computations. Position of the rs78378222 variant is highlighted by blue column. T=U in the miRNA sequences.

Supplementary Figure A.1
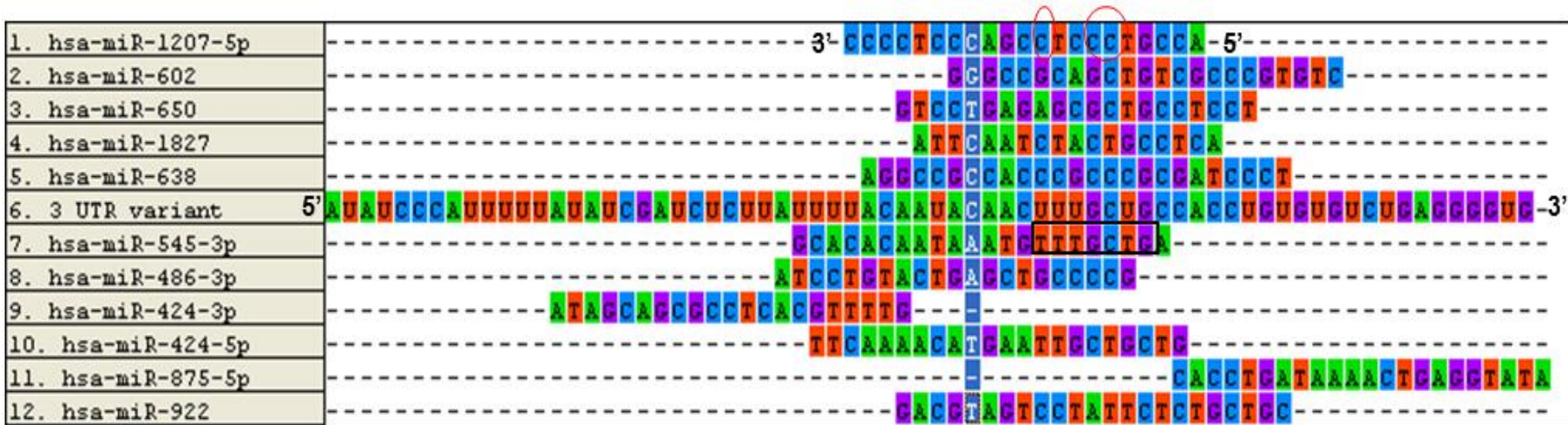
Supplementary Table A.1. Features of the online computational algorithms employed in the current study.

Algorithms	Criteria for prediction an ranking	Score description	Website URL	Last updated, recent reference
Target Scan release 6.2	Stringent seed pairing, site number, site type, site context (which includes factors that influence site accessibility); site conservation, option of ranking by likelihood of preferential conservation rather than site context	Context+ score: sum of the contribution of these six features, calculated as in Garcia *et al.*, 2011: site-type, 3' pairing, local AU, position, target site abundance and seed-pairing stability	http://targetscan.org	June 2012, Friedman *et al.*, 2009
TargetSpy release 1.0	Stringent seed pairing (chosen option), general extent of miRNA-mRNA binding, number of G:U wobble base pairs in the miRNA-mRNA hybrid, bulge-related features of duplexes, position specific features, base composition, accessibility of the target site to RISC, compactness (unpaired nucleotides and pairing length); identify species-specific target sites independent of site conservation	Overall score based in a broad spectrum of pairing requirements. A ranked list of all features regarded is detailed in the respective reference	http://www.targetspy.org/	October 2009, Sturm *et al.*, 2010
MicroCosm Targets version 5[a]	Stringent seed pairing, site conservation, folding energy and thermodynamic stability (using miRanda algorithm)	Overall score that is the sum of the derived scores across the total miRNA *vs.* 3'UTR alignment[b]	http://www.ebi.ac.uk/enright-srv/microcosm	NA, Griffiths-Jones *et al.*, 2008
PITA algorithm	Moderately stringent seed pairing, site number, overall predicted pairing stability, probability of interaction by target accessibility (secondary structure of the mRNA), site conservation algorithm	ΔΔG score: free energy gained by transitioning from the state in which the miRNA and the target are unbound to the state in which the miRNA binds its target[c,d]	http://genie.weizmann.ac.il/pubs/mir07/mir07_prediction.html	NA, Kertesz *et al.*, 2007
RNA 22 version 2.0 microRNA target detection	Moderately stringent seed pairing, matches to sequence patterns generated from miRNA set, overall predicted pairing, folding energy and predicted pairing stability	Folding energy (Kcal/mol) for the heteroduplex formed between miRNA and target mRNA[c]	https://cm.jefferson.edu/rna22v2/	NA, Miranda *et al.*, 2006
RNAhybrid	Minimum free energy hybridization (MFEH) between mRNA and miRNA, hybridization performed in a kind of domain mode (miRNA sequence hybridized to the most favorable site of the mRNA); possibility to disallow G:U base pairs in the seed region, to restrict possible sizes of unpaired regions (loop size), and to define site number per miRNA; threshold for the MFEH and position/length features	Energy score; MFEH for miRNA-target interactions (Kcal/mol)[c]	http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/	February 2014, Krüger *et al.*, 2006

NA, not available.

[a] Formerly named miRBase Targets.

[b] A positive score is given for complementary base pairing, and a negative score is given to mismatches, gap-opening and gap-extension.

[c] For energetic scores, the lower (more negative) its value, the stronger the binding of the microRNA to the given site is expected to be.

[d] As a rough rule of thumb, sites having ΔΔG values below -10 are likely to be functional in endogenous miRNA expression levels.

Supplementary Table A.2. Scores provided by several *in silico* algorithms suggesting miR-545-3p as a novel miRNA that directly target the human mRNA *TP53*.

Algorithms	Score Interpretation	Score for miR-545-3p[a]	Mean score for "reference miRNAs"[b]	Mean score for no *TP53*-repressive miRNAs[c]
TargetSpy	The higher (more positive) its value, the stronger the binding of the microRNA to the given site is expected to be.	0.99748	0.99688 (miR-92a, miR-141)	miR-485-5p was NP and 0.991984 for miR-508
MicroCosm Targets		17.35	16.9658 (miR-92a)	16.26 for miR-485-5p and miR-508 was NP
Target Scan	The lower (more negative) its value, the stronger the binding of the microRNA to the given site is expected to be.	-0.19	-0.16 (miR-125b, miR-504)	- 0.09 (miR-485-5p, miR-508)
PITA algorithm		-13.26	-8.14 (miR-125b, miR-504)	-12.16 (miR-485-5p, miR-508)
RNA 22		-14.2	-16.7 (miR-125b, miR-504)	-16.05 (miR-485-5p, miR-508)
RNAhybrid		-20.6	-26.45 (miR-125b, miR-504)	-24.7 (miR-485-5p, miR-508)

NP, not predicted.

[a] Score for miR-545-3p:mRNA interaction considering wild-type sequence of the *TP53* 3'UTR.

[b] "Reference miRNAs" = miRNAs previously validated by functional assays to regulate human *TP53* expression. miRNAs analyzed are mentioned along with the mean score. We were unable to use the same reference miRNAs in all analyses because miR-504 and miR-125b are not predicted by Target Spy and MicroCosm Targets algorithms.

[c] No *TP53*-repressive miRNAs = miRNAs for which experimentally was demonstrated not negatively regulating p53 protein levels, as described in Hu *et al.*, 2010 [46]. They are mentioned along with the mean score.

Supplementary Table A.3. ΔΔG scores assigned to binding of candidate miRNAs in target sites predicted by PITA algorithm in the latter 70bp of *TP53* 3'UTR.

Candidate miRNA[a]	WT 3'UTR[b]		Variant 3'UTR[b]		ΔΔG score variation[c]
	ΔΔG score	Site number	ΔΔG score	Site number	
miR-602	-13.82	1	-13.57	1	-0.25
miR-1207-5p	-12.95	1	-14.24	1	1.29
miR-545-3p	-12.25	1	-9.59	1	-2.66
miR-650	-11.93	1	-11.57	1	-0.36
miR-638	-11.46	1	-11.26	1	-0.2
miR-1827	-11.45	1	-11.4	1	-0.05
miR-486-3p	-11.16	1	-10.9	1	-0.26
miR-875-5p	-10.68	1	-10.11	1	-0.57
miR-922	-10.25	1	-9.99	1	-0.26
miR-424	-9.91	1	-10.44	1	0.53

[a] We selected only the sites having ΔΔG values below -10 which likely are functional in endogenous miRNA expression levels, as described by Kertesz *et al.*, 2007. miRNAs predicted to target the wild-type (WT) and variant *TP53* 3'UTR (containing rs78378222[C]) were the same using this cutoff, suggesting that rs78378222 variant does not create a new miRNA target site.

[b] Inputted sequence encompassing rs78378222 position.

[c] ΔΔG score variation = (ΔΔG score for WT 3'UTR - ΔΔG score for variant 3'UTR).

5. CONCLUSÕES

Considerando os resultados encontrados no presente estudo, podemos concluir que:

a) O SNP funcional rs78378222(A>C) não foi identificado em nenhuma paciente com CM HER2-positivo, sugerindo que essa variante não está relacionada com o mecanismo de amplificação do oncogene *ERBB2* observado na grande maioria (90%) dos tumores de mama que apresentam esse fenótipo molecular.

b) Amostras de tumores primários e sítios metastáticos de pacientes portadores da variante rs78378222 apresentam uma marcação fraca e focal para a proteína p53 pela técnica de IHQ, um padrão distinto daquele observado em tumores de indivíduos portadores de mutações patogênicas em regiões codificadoras do gene *TP53*, o qual é caracterizado por um intenso acúmulo nuclear de p53.

c) Fibroblastos de pacientes SNP-positivos apresentam menores níveis de expressão da proteína p53 em um *status* celular basal, bem como após indução de dano ao DNA, quando comparado aos de um indivíduo controle (sem mutações identificadas em *TP53*), sugerindo que pelo fato desta variante alterar a SSP do gene *TP53* e, assim, ocasionar um processamento 3' anormal do mRNA *TP53*, menores níveis de expressão de p53 são observados.

d) Fibroblastos de pacientes portadores da variante apresentam menores níveis de p53 após indução de dano ao DNA, quando comparado aos de portadores de mutações em regiões codificadoras de *TP53*, corroborando a hipótese de que, devido sua poliadenilação prejudicada, o transcrito *TP53* contendo rs78378222[C] é degradado mais rapidamente.

e) Pacientes portadores do SNP rs78378222 apresentam uma redução na expressão de p53 nas diferentes amostras biológicas analisadas e esse achado, aliado aos dados previamente publicados acerca do impacto funcional dessa variante, sugere um novo mecanismo pelo qual *TP53* pode promover predisposição ao câncer, provavelmente relacionado com haploinsuficiência.

f) A variante rs78378222 está situada em uma sequência-alvo na região 3'UTR do mRNA *TP53* predita para ligação de miR-545-3p, sendo esse um miRNA ainda não validado experimentalmente como regulador pós-transcricional da expressão do gene *TP53* em humanos.

g) O alelo rs78378222[C] está situado em uma região do mRNA *TP53* que apresenta uma complementariedade de bases perfeita com a extremidade 3' de miR-545-3p, podendo enfraquecer a ligação do miRNA candidato a esse sítio-alvo "putativo", o que sugere um novo mecanismo de doença associado a essa variante.

h) A variante rs78378222 possivelmente não cria um novo sítio de ligação para miRNAs na região 3'UTR do gene *TP53*.

6. PERSPECTIVAS

Como perspectivas da continuidade desse trabalho e sugestões de análises adicionais destacamos as seguintes:

- Avaliar os níveis de mRNA *TP53* em pacientes portadores da variante rs78378222 com fenótipo clínico da síndrome de LFL, em comparação com controles não portadores (sem mutação identificada em *TP53*), a partir do isolamento da fração leucocitária em amostras de sangue periférico.

- Analisar a perda de heterozigosidade no *locus TP53* em amostras de tecido tumoral conservado em blocos de parafina de pacientes LFL portadores da variante estudada.

- Realizar análises de segregação do alelo rs78378222[C], a partir do recrutamento de familiares dos pacientes LFL portadores da variante, prevendo uma possível limitação no número amostral imposta pelo fato desta ser uma variante rara.

- Avaliar os níveis de expressão da proteína p53 em amostras de tumores primários conservados em blocos de parafina de um maior número de pacientes portadores da variante.

- Avaliar a possibilidade de miR-545-3p ser um novo miRNA que regula negativamente a expressão do gene *TP53* em humanos através de ensaios funcionais *in vitro*, bem como avaliar se a variante rs78378222 enfraquece a ligação de miR-545-3p ao seu sítio-alvo predito na região 3'UTR do mRNA *TP53*, por meio de estudo funcional envolvendo a transfecção de construtos p53 exógenos contendo o alelo rs78378222[C].

- Estudar a variabilidade genética na SSP (hexanucleotídeo AAUAAA) em outros genes de suscetibilidade ao câncer (*BRCA1*, *BRCA2*, *CHECK2*, *KRAS*, *PTEN*) em pacientes que preenchem critérios clínicos para síndromes de predisposição hereditária ao câncer.

7. REFERÊNCIAS ADICIONAIS

Ambros, V. (2004). The functions of animal microRNAs. Nature *431*, 350-355.

Beaudoing, E., Freier, S., Wyatt, J.R., Claverie, J.M., Gautheret, D. (2000). Patterns of variant polyadenylation signal usage in human genes. Genome Res *10*, 1001-1010.

Beckerman, R., e Prives, C. (2010). Transcriptional regulation by p53. Cold Spring Harb Perspect Biol *2*, a000935.

Bensaad, K., Tsuruta, A., Selak, M.A., Vidal, M.N., Nakano, K., Bartrons, R., Gottlieb, E., Vousden, K.H. (2006). TIGAR, a p53-inducible regulator of glycolysis and apoptosis. Cell *126*, 107-120.

Bensaad, K., e Vousden, K.H. (2007). p53: new roles in metabolism. Trends Cell Biol *17*, 286-291.

Bhargava, R., Beriwal, S., McManus, K., Dabbs, D.J. (2008). CK5 is more sensitive than CK5/6 in identifying the "basal-like" phenotype of breast carcinoma. Am J Clin Pathol *130*, 724-730.

Birch, J.M. (1994). Li-Fraumeni syndrome. Eur J Cancer *30A*, 1935-1941.

Birch, J.M., Blair, V., Kelsey, A.M., Evans, D.G., Harris, M., Tricker, K.J., Varley, J.M. (1998). Cancer phenotype correlates with constitutional TP53 genotype in families with the Li-Fraumeni syndrome. Oncogene *17*, 1061-1068.

Bourdon, J.C., Fernandes, K., Murray-Zmijewski, F., Liu, G., Diot, A., Xirodimas, D.P., Saville, M.K., Lane, D.P. (2005). p53 isoforms can regulate p53 transcriptional activity. Genes & development *19*, 2122-2137.

Brewster, B.L., Rossiello, F., French, J.D., Edwards, S.L., Wong, M., Wronski, A., Whiley, P., Waddell, N., Chen, X., Bove, B.*, et al.* (2012). Identification of fifteen novel germline variants in the BRCA1 3'UTR reveals a variant in a breast cancer case that introduces a functional miR-103 target site. Hum Mutat *33*, 1665-1675.

Budanov, A.V., Sablina, A.A., Feinstein, E., Koonin, E.V., Chumakov, P.M. (2004). Regeneration of peroxiredoxins by p53-regulated sestrins, homologs of bacterial AhpD. Science *304*, 596-600.

Calin, G.A., Sevignani, C., Dumitru, C.D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M.*, et al.* (2004). Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci U S A *101*, 2999-3004.

Chin, L.J., Ratner, E., Leng, S., Zhai, R., Nallur, S., Babar, I., Muller, R.U., Straka, E., Su, L., Burki, E.A.*, et al.* (2008). A SNP in a let-7 microRNA complementary site in the KRAS 3' untranslated region increases non-small cell lung cancer risk. Cancer Res *68*, 8535-8540.

Conway, L., e Wickens, M. (1987). Analysis of mRNA 3' end formation by modification interference: the only modifications which prevent processing lie in AAUAAA and the poly(A) site. EMBO J *6*, 4177-4184.

Dornan, D., Shimizu, H., Burch, L., Smith, A.J., Hupp, T.R. (2003). The proline repeat domain of p53 binds directly to the transcriptional coactivator p300 and allosterically controls DNA-dependent acetylation of p53. Mol Cell Biol *23*, 8846-8861.

Dumont, P., Leu, J.I., Della Pietra, A.C., George, D.L., Murphy, M. (2003). The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. Nat Genet *33*, 357-365.

Easton, D.F., Bishop, D.T., Ford, D., Crockford, G.P. (1993). Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. The Breast Cancer Linkage Consortium. Am J Hum Genet *52*, 678-701.

Eeles, R.A. (1995). Germline mutations in the TP53 gene. Cancer Surv *25*, 101-124.

Esquela-Kerscher, A., e Slack, F.J. (2006). Oncomirs - microRNAs with a role in cancer. Nat Rev Cancer *6*, 259-269.

Feng, Z., Wu, R., Lin, M., Hu, W. (2011). Tumor supressor p53: new functions of an old protein. Front Biol *6*, 58–68.

Ford, D., Easton, D.F., Stratton, M., Narod, S., Goldgar, D., Devilee, P., Bishop, D.T., Weber, B., Lenoir, G., Chang-Claude, J.*, et al.* (1998). Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. Am J Hum Genet *62*, 676-689.

Frebourg, T., Abel, A., Bonaiti-Pellie, C., Brugières, L., Berthet, P., Bressac-de Paillerets, B., Chevrier, A., Chompret, A., Cohen-Haguenauer, O., Delattre, O.*, et al.* (2001). Li-Fraumeni syndrome: update, new data and guidelines for clinical management. Bull Cancer *88*, 581-587.

Fujita, P.A., Rhead, B., Zweig, A.S., Hinrichs, A.S., Karolchik, D., Cline, M.S., Goldman, M., Barber, G.P., Clawson, H., Coelho, A.*, et al.* (2011). The UCSC Genome Browser database: update 2011. Nucleic Acids Res *39*, D876-882.

Gaidatzis, D., van Nimwegen, E., Hausser, J., Zavolan, M. (2007). Inference of miRNA targets using evolutionary conservation and pathway analysis. BMC Bioinformatics *8*, 69.

Gemignani, F., Moreno, V., Landi, S., Moullan, N., Chabrier, A., Gutiérrez-Enríquez, S., Hall, J., Guino, E., Peinado, M.A., Capella, G.*, et al.* (2004). A TP53 polymorphism is associated with increased risk of colorectal cancer and with reduced levels of TP53 mRNA. Oncogene *23*, 1954-1956.

Gonzalez, K.D., Noltner, K.A., Buzin, C.H., Gu, D., Wen-Fong, C.Y., Nguyen, V.Q., Han, J.H., Lowstuter, K., Longmate, J., Sommer, S.S.*, et al.* (2009). Beyond Li Fraumeni Syndrome: clinical characteristics of families with p53 germline mutations. J Clin Oncol *27*, 1250-1256.

Graber, J.H., Cantor, C.R., Mohr, S.C., Smith, T.F. (1999). In silico detection of control signals: mRNA 3'-end-processing sequences in diverse species. Proc Natl Acad Sci U S A *96*, 14055-14060.

Grimson, A., Farh, K.K., Johnston, W.K., Garrett-Engele, P., Lim, L.P., Bartel, D.P. (2007). MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell *27*, 91-105.

Gutierrez, C., e Schiff, R. (2011). HER2: biology, detection, and clinical implications. Arch Pathol Lab Med *135*, 55-62.

Hartley, A.L., Birch, J.M., Kelsey, A.M., Marsden, H.B., Harris, M., Teare, M.D. (1989). Are germ cell tumors part of the Li-Fraumeni cancer family syndrome? Cancer Genet Cytogenet *42*, 221-226.

Higgs, D.R., Goodbourn, S.E., Lamb, J., Clegg, J.B., Weatherall, D.J., Proudfoot, N.J. (1983). Alpha-thalassaemia caused by a polyadenylation signal mutation. Nature *306*, 398-400.

Hu, W., Chan, C.S., Wu, R., Zhang, C., Sun, Y., Song, J.S., Tang, L.H., Levine, A.J., Feng, Z. (2010a). Negative regulation of tumor suppressor p53 by microRNA miR-504. Mol Cell *38*, 689-699.

Hu, W., Zhang, C., Wu, R., Sun, Y., Levine, A., Feng, Z. (2010b). Glutaminase 2, a novel p53 target gene regulating energy metabolism and antioxidant function. Proc Natl Acad Sci U S A *107*, 7455-7460.

Hynes, N.E., e MacDonald, G. (2009). ErbB receptors and signaling pathways in cancer. Curr Opin Cell Biol *21*, 177-184.

IARC TP53 Database. (2012). Database developments release R16. Disponível em: <http://www-p53.iarc.fr>. (Acessado em 16 de abril de 2014).

INCa. (2014). Estimativa de incidência de câncer no Brasil. Disponível em: <http://www.inca.gov.br/estimativa/2014/>. (Acessado em 16 de abril de 2014).

Isobe, M., Emanuel, B.S., Givol, D., Oren, M., Croce, C.M. (1986). Localization of gene for human p53 tumour antigen to band 17p13. Nature *320*, 84-85.

Kleihues, P., Schäuble, B., zur Hausen, A., Estève, J., Ohgaki, H. (1997). Tumors associated with p53 germline mutations: a synopsis of 91 families. Am J Pathol *150*, 1-13.

Lacroix, M., Toillon, R.A., Leclercq, G. (2006). p53 and breast cancer, an update. Endocr Relat Cancer *13*, 293-325.

Lalloo, F., Varley, J., Moran, A., Ellis, D., O'dair, L., Pharoah, P., Antoniou, A., Hartley, R., Shenton, A., Seal, S.*, et al.* (2006). BRCA1, BRCA2 and TP53 mutations in very early-onset breast cancer with associated risks to relatives. Eur J Cancer *42*, 1143-1150.

Lane, D., e Levine, A. (2010). p53 Research: the past thirty years and the next thirty years. Cold Spring Harb Perspect Biol *2*, a000893.

Langerød, A., Zhao, H., Borgan, Ø., Nesland, J.M., Bukholm, I.R., Ikdahl, T., Kåresen, R., Børresen-Dale, A.L., Jeffrey, S.S. (2007). TP53 mutation status and gene expression profiles are powerful prognostic markers of breast cancer. Breast Cancer Res *9*, R30.

Levine, A.J. (1989). The p53 tumor suppressor gene and gene product. Princess Takamatsu symposia *20*, 221-230.

Li, F.P., e Fraumeni, J.F. (1969a). Rhabdomyosarcoma in children: epidemiologic study and identification of a familial cancer syndrome. J Natl Cancer Inst *43*, 1365-1373.

Li, F.P., e Fraumeni, J.F. (1969b). Soft-tissue sarcomas, breast cancer, and other neoplasms. A familial syndrome? Ann Intern Med *71*, 747-752.

Long, D., Lee, R., Williams, P., Chan, C.Y., Ambros, V., Ding, Y. (2007). Potent effect of target structure on microRNA function. Nat Struct Mol Biol *14*, 287-294.

MacDonald, C.C., e Redondo, J.L. (2002). Reexamining the polyadenylation signal: were we wrong about AAUAAA? Mol Cell Endocrinol *190*, 1-8.

Manley, J.L. (1983). Accurate and specific polyadenylation of mRNA precursors in a soluble whole-cell lysate. Cell *33*, 595-605.

Manley, J.L., Yu, H., Ryner, L. (1985). RNA sequence containing hexanucleotide AAUAAA directs efficient mRNA polyadenylation in vitro. Mol Cell Biol *5*, 373-379.

Marín, R.M., e Vanícek, J. (2011). Efficient use of accessibility in microRNA target prediction. Nucleic Acids Res *39*, 19-29.

Masciari, S., Dewanwala, A., Stoffel, E.M., Lauwers, G.Y., Zheng, H., Achatz, M.I., Riegert-Johnson, D., Foretova, L., Silva, E.M., Digianni, L.*, et al.* (2011). Gastric cancer in individuals with Li-Fraumeni syndrome. Genet Med *13*, 651-657.

Matakidou, A., Eisen, T., Houlston, R.S. (2003). TP53 polymorphisms and lung cancer risk: a systematic review and meta-analysis. Mutagenesis *18*, 377-385.

Matoba, S., Kang, J.G., Patino, W.D., Wragg, A., Boehm, M., Gavrilova, O., Hurley, P.J., Bunz, F., Hwang, P.M. (2006). p53 regulates mitochondrial respiration. Science *312*, 1650-1653.

miRBase versão 20.0. (2013). Browse miRBase by species. Disponível em: <http://mirbase.org/cgi-bin/browse.pl>. (Acessado em 22 de abril de 2014).

Moasser, M.M. (2007). The oncogene HER2: its signaling and transforming functions and its role in human cancer pathogenesis. Oncogene *26*, 6469-6487.

Moore, C.L., e Sharp, P.A. (1984). Site-specific polyadenylation in a cell-free reaction. Cell *36*, 581-591.

Moore, C.L., e Sharp, P.A. (1985). Accurate cleavage and polyadenylation of exogenous RNA substrate. Cell *41*, 845-855.

Mouchawar, J., Korch, C., Byers, T., Pitts, T.M., Li, E., McCredie, M.R., Giles, G.G., Hopper, J.L., Southey, M.C. (2010). Population-based estimate of the contribution of TP53 mutations to subgroups of early-onset breast cancer: Australian Breast Cancer Family Study. Cancer Res *70*, 4795-4800.

Nielsen, T.O., Hsu, F.D., Jensen, K., Cheang, M., Karaca, G., Hu, Z., Hernandez-Boussard, T., Livasy, C., Cowan, D., Dressler, L.*, et al.* (2004). Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. Clin Cancer Res *10*, 5367-5374.

Nigro, J.M., Baker, S.J., Preisinger, A.C., Jessup, J.M., Hostetter, R., Cleary, K., Bigner, S.H., Davidson, N., Baylin, S., Devilee, P. (1989). Mutations in the p53 gene occur in diverse human tumour types. Nature *342*, 705-708.

Nilsen, T.W. (2007). Mechanisms of microRNA-mediated gene regulation in animal cells. Trends Genet *23*, 243-249.

Olivier, M., Goldgar, D.E., Sodha, N., Ohgaki, H., Kleihues, P., Hainaut, P., Eeles, R.A. (2003). Li-Fraumeni and related syndromes: correlation between tumor type, family structure, and TP53 genotype. Cancer Res *63*, 6643-6650.

Olivier, M., Hainaut, P., Borresen, A.L. (2005). In: Hainaut P & Wiman KG (eds). 25 Years of p53 Research. Springer, 321-338.

Olivier, M., Hollstein, M., Hainaut, P. (2010). TP53 mutations in human cancers: origins, consequences, and clinical use. Cold Spring Harb Perspect Biol *2*, a001008.

Olivier, M., Petitjean, A., Marcel, V., Pétré, A., Mounawar, M., Plymoth, A., de Fromentel, C.C., Hainaut, P. (2009). Recent advances in p53 research: an interdisciplinary perspective. Cancer Gene Ther *16*, 1-12.

Oren, M. (2003). Decision making by p53: life, death and cancer. Cell Death Differ *10*, 431-442.
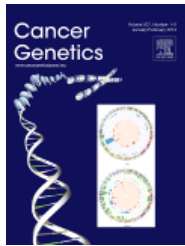
Perou, C.M., Sørlie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., Ross, D.T., Johnsen, H., Akslen, L.A.*, et al.* (2000). Molecular portraits of human breast tumours. Nature *406*, 747-752.

Proudfoot, N. (1991). Poly(A) signals. Cell *64*, 671-674.

Proudfoot, N.J., e Brownlee, G.G. (1976). 3' non-coding region sequences in eukaryotic messenger RNA. Nature *263*, 211-214.

Rachagani, S., Kumar, S., Batra, S.K. (2010). MicroRNA in pancreatic cancer: pathological, diagnostic and therapeutic implications. Cancer Lett *292*, 8-16.

Rath, M.G., Masciari, S., Gelman, R., Miron, A., Miron, P., Foley, K., Richardson, A.L., Krop, I.E., Verselis, S.J., Dillon, D.A.*, et al.* (2013). Prevalence of germline TP53 mutations in HER2+ breast cancer patients. Breast Cancer Res Treat *139*, 193-198.

Rehmsmeier, M., Steffen, P., Hochsmann, M., Giegerich, R. (2004). Fast and effective prediction of microRNA/target duplexes. RNA *10*, 1507-1517.

Ross, J.S., e Fletcher, J.A. (1998). The HER-2/neu Oncogene in Breast Cancer: Prognostic Factor, Predictive Factor, and Target for Therapy. Oncologist *3*, 237-252.

Royds, J.A., e Iacopetta, B. (2006). p53 and disease: when the guardian angel fails. Cell Death Differ *13*, 1017-1026.

Ruan, K., Fang, X., Ouyang, G. (2009). MicroRNAs: novel regulators in the hallmarks of human cancer. Cancer Lett *285*, 116-126.

Ryan, K., Calvo, O., Manley, J.L. (2004). Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing endonuclease. RNA *10*, 565-573.

Sablina, A.A., Budanov, A.V., Ilyinskaya, G.V., Agapova, L.S., Kravchenko, J.E., Chumakov, P.M. (2005). The antioxidant function of the p53 tumor suppressor. Nat Med *11*, 1306-1313.

Schmidt, M.K., Reincke, S., Broeks, A., Braaf, L.M., Hogervorst, F.B., Tollenaar, R.A., Johnson, N., Fletcher, O., Peto, J., Tommiska, J.*, et al.* (2007). Do MDM2 SNP309 and TP53 R72P interact in breast cancer susceptibility? A large pooled series from the breast cancer association consortium. Cancer Res *67*, 9584-9590.

Sheets, M.D., Ogg, S.C., Wickens, M.P. (1990). Point mutations in AAUAAA and the poly (A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation in vitro. Nucleic Acids Res *18*, 5799-5805.

Sørlie, T. (2004). Molecular portraits of breast cancer: tumour subtypes as distinct disease entities. Eur J Cancer *40*, 2667-2675.

Sørlie, T., Perou, C.M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S.*, et al.* (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A *98*, 10869-10874.

Soussi, T., e Béroud, C. (2001). Assessing TP53 status in human tumours to evaluate clinical outcome. Nat Rev Cancer *1*, 233-240.

Vousden, K.H., e Lane, D.P. (2007). p53 in health and disease. Nat Rev Mol Cell Biol *8*, 275-283.

Vousden, K.H., e Prives, C. (2009). Blinded by the Light: The Growing Complexity of p53. Cell *137*, 413-431.

Walsh, T., Casadei, S., Coats, K.H., Swisher, E., Stray, S.M., Higgins, J., Roach, K.C., Mandell, J., Lee, M.K., Ciernikova, S.*, et al.* (2006). Spectrum of mutations in BRCA1, BRCA2, CHEK2, and TP53 in families at high risk of breast cancer. JAMA *295*, 1379-1388.

Wang-Gohrke, S., Rebbeck, T.R., Besenfelder, W., Kreienberg, R., Runnebaum, I.B. (1998). p53 germline polymorphisms are associated with an increased risk for breast cancer in German women. Anticancer Res *18*, 2095-2099.

Wang-Gohrke, S., Weikel, W., Risch, H., Vesprini, D., Abrahamson, J., Lerman, C., Godwin, A., Moslehi, R., Olipade, O., Brunet, J.S.*, et al.* (1999). Intron variants of the p53 gene are associated with increased risk for ovarian cancer but not in carriers of BRCA1 or BRCA2 germline mutations. Br J Cancer *81*, 179-183.

Whibley, C., Pharoah, P.D., Hollstein, M. (2009). p53 polymorphisms: cancer implications. Nat Rev Cancer *9*, 95-107.

Wickens, M., e Stephenson, P. (1984). Role of the conserved AAUAAA sequence: four AAUAAA point mutants prevent messenger RNA 3' end formation. Science *226*, 1045-1051.

William, W.R., e Strong, L.C. (1987). Familial Cancer, 1st International Research Conference (Karger, Basel, 1985). J Natl Cancer Inst *79*, 1213.

Wilusz, J., Pettine, S.M., Shenk, T. (1989). Functional analysis of point mutations in the AAUAAA motif of the SV40 late polyadenylation signal. Nucleic Acids Res *17*, 3899-3908.

Wolff, A.C., Hammond, M.E., Hicks, D.G., Dowsett, M., McShane, L.M., Allison, K.H., Allred, D.C., Bartlett, J.M., Bilous, M., Fitzgibbons, P.*, et al.* (2013). Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. J Clin Oncol *31*, 3997-4013.

ANEXOS DO TRABALHO

Normas de avaliação do periódico *Cancer genetics*

ELSEVIER

Type here to search on Elsevier.com

Advanced search   Follow us: [f] [in] [t] [g+]   Help & Contact

Journals & books   **Online tools**   **Authors, editors & reviewers**   **About Elsevier**   **Community**   **Store**

Browse journals > Cancer Genetics > Guide for authors

Cancer Genetics

View free sample

Guide for authors

Submit your paper

Order journal

View articles

Abstracting and indexing

Editorial board

Guide for Authors

[PDF] Author information pack

BEFORE YOU BEGIN
• Ethics in publishing
• Plagiarism
• Competing Interests
• Conflict of interest
• Submission declaration
• Changes to authorship
• Copyright
• Role of the funding source
• Funding body agreements and policies
• Open access
• Language (usage and editing services)

• Submission
• Referees

PREPARATION
• Use of word processing software
• Article structure
• Essential title page information
• Abstract
• Keywords
• Acknowledgements
• Math formulae
• Footnotes
• Artwork

• Illustration services
• Tables
• References
• Video data
• Supplementary data
• Submission checklist

AFTER ACCEPTANCE
• Use of the Digital Object Identifier
• Online proof correction

AUTHOR INQUIRIES

The aim of *Cancer Genetics* is to publish high quality scientific papers on the cellular, genetic and molecular aspects of cancer, including cancer predisposition and clinical diagnostic applications. Specific areas of interest include descriptions of new chromosomal, molecular or epigenetic alterations in benign and malignant diseases; novel laboratory approaches for identification and characterization of chromosomal rearrangements or genomic alterations in cancer cells; correlation of genetic changes with pathology and clinical presentation; and the molecular genetics of cancer predisposition. To reach a basic science and clinical multidisciplinary audience, we welcome original full-length articles, reviews, meeting summaries, brief reports, and letters to the editor.

Review Articles
Review articles fitting within the Aims and Scope of the journal will be considered for publication. Authors wishing to submit an unsolicited review should send a brief, 250-word synopsis of the planned review to cancergenetics@elsevier.com for presubmission approval by the editor.

Original Research
Original research manuscripts must include the following sections: Abstract (=200 words), Introduction, Materials and Methods, Results, Discussion, and References. Acknowledgments (optional) and figure legends should also be included in the manuscript file but should be on a separate text page. Figure legends, figures, and tables should each be uploaded in separate files. Authors must include a cover letter clearly stating the novel findings described in the paper. All genotype association studies must utilize current molecular methods and include at least 500 cases and 500 controls.

Meeting summaries
Meeting Summaries require pre-submission approval by the Editor. Please email proposals to cancergenetics@elsevier.com.

Commentaries
Commentaries on emerging technologies require pre-submission approval. Please email proposals to cancergenetics@elsevier.com.

Brief Communications
Brief Communications must be less than 2,000 words, with a maximum of 2 tables and/or figures. Brief communications must include the following sections: Abstract (=200 words), Introduction, Materials and Methods, Results, Discussion, and References. Authors must include a cover letter clearly stating the novel findings described in the communication.

Letters to the Editor
Letters to the Editor must be <750 words, with no figures or tables.

• Jaclyn A. Biegel, PhD, Editor-in-Chief
• Editorial Office: cancergenetics@elsevier.com

Before You Begin

Ethics in publishing

For information on Ethics in publishing and Ethical guidelines for journal publication see
http://www.elsevier.com/publishingethics and http://www.elsevier.com/journal-authors/ethics.

Plagiarism

Plagiarism and duplicate submission are serious acts of misconduct. Plagiarism is defined as unreferenced use of published or unauthorized use of unpublished ideas, and may occur at any stage of planning, researching, writing, or publication. Plagiarism takes many forms, from 'passing off' another's paper as the author's own paper, to copying or paraphrasing substantial parts of another's paper (without attribution), to claiming results from research conducted by others. Plagiarism in all its forms constitutes unethical publishing behavior and is unacceptable. Duplicate submission/publication occurs when two or more papers, without full cross-reference, share the same hypothesis, data, discussion points, or conclusions.

Competing Interests

Effective February 1, 2012, each author must complete and sign the Research Funding Support, Conflict of Interest Disclosure, and Authorship Agreement form at the time of manuscript submission (http://www.elsevier.com/__data/promis_misc/cg_coi.pdf). All authors must disclose any financial and personal relationships with other people or organizations that could inappropriately influence (bias) their work. Examples of potential conflicts of interest include employment, consultancies, stock ownership, honoraria, paid expert testimony, patent applications/registrations, and grants or other funding.

Conflict of interest

All authors must disclose any financial and personal relationships with other people or organizations that could inappropriately influence (bias) their work. Examples of potential conflicts of interest include employment, consultancies, stock ownership, honoraria, paid expert testimony, patent applications/registrations, and grants or other funding. See also http://www.elsevier.com/conflictsofinterest. Further information and an example of a Conflict of Interest form can be found at: http://help.elsevier.com/app/answers/detail/a_id/286/p/7923.

Submission declaration

Submission of an article implies that the work described has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see http://www.elsevier.com/postingpolicy), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere including electronically in the same form, in English or in any other language, without the written consent of the copyright-holder.

Changes to authorship

This policy concerns the addition, deletion, or rearrangement of author names in the authorship of accepted manuscripts:
Before the accepted manuscript is published in an online issue: Requests to add or remove an author, or to rearrange the author names, must be sent to the Journal Manager from the corresponding author of the accepted manuscript and must include: (a) the reason the name should be added or removed, or the author names rearranged and (b) written confirmation (e-mail, fax, letter) from all authors that they agree with the addition, removal or rearrangement. In the case of addition or removal of authors, this includes confirmation from the author being added or removed. Requests that are not sent by the corresponding author will be forwarded by the Journal Manager to the corresponding author, who must follow the procedure as described above. Note that: (1) Journal Managers will inform the Journal Editors of any such requests and (2) publication of the accepted manuscript in an online issue is suspended until authorship has been agreed.
After the accepted manuscript is published in an online issue: Any requests to add, delete, or rearrange author names in an article published in an online issue will follow the same policies as noted above and result in a corrigendum.

Copyright

This journal offers authors a choice in publishing their research: Open Access and Subscription.

For Subscription articles
Upon acceptance of an article, authors will be asked to complete a 'Journal Publishing Agreement' (for more information on this and copyright, see http://www.elsevier.com/copyright). An e-mail will be sent to the corresponding author confirming receipt of the manuscript together with a 'Journal Publishing Agreement' form or a link to the online version of this agreement.
Subscribers may reproduce tables of contents or prepare lists of articles including abstracts for internal circulation within their institutions. Permission of the Publisher is required for resale or distribution outside the institution and for all other derivative works, including compilations and translations (please consult http://www.elsevier.com/permissions). If excerpts from other copyrighted works are included, the author(s) must obtain written permission from the copyright owners and credit the source(s) in the article. Elsevier has preprinted forms for use by authors in these cases: please consult http://www.elsevier.com/permissions.

For Open Access articles
Upon acceptance of an article, authors will be asked to complete an 'Exclusive License Agreement' (for more information see http://www.elsevier.com/OAauthoragreement). Permitted reuse of open access articles is determined by the author's choice of user license (see http://www.elsevier.com/openaccesslicenses).

Retained author rights
As an author you (or your employer or institution) retain certain rights. For more information on author rights for:
Subscription articles please see http://www.elsevier.com/journal-authors/author-rights-and-responsibilities.
Open access articles please see http://www.elsevier.com/OAauthoragreement.

Role of the funding source

You are requested to identify who provided financial support for the conduct of the research and/or preparation of the article and to briefly describe the role of the sponsor(s), if any, in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. If the funding source(s) had no such involvement then this should be stated.

Funding body agreements and policies

Elsevier has established agreements and developed policies to allow authors whose articles appear in journals published by Elsevier, to comply with potential manuscript archiving requirements as specified as conditions of their grant awards. To learn more about existing agreements and policies please visit http://www.elsevier.com/fundingbodies.

Open access

This journal offers authors a choice in publishing their research:

Open Access
• Articles are freely available to both subscribers and the wider public with permitted reuse
• An Open Access publication fee is payable by authors or their research funder
Subscription
• Articles are made available to subscribers as well as developing countries and patient groups through our access programs (http://www.elsevier.com/access)
• No Open Access publication fee

All articles published Open Access will be immediately and permanently free for everyone to read and download. Permitted reuse is defined by your choice of one of the following Creative Commons user licenses:
Creative Commons Attribution-NonCommercial-ShareAlike (CC BY-NC-SA): for non-commercial purposes, lets others distribute and copy the article, to create extracts, abstracts and other revised versions, adaptations or derivative works of or from an article (such as a translation), to include in a collective work (such as an anthology), to text and data mine the article, as long as they credit the author(s), do not represent the author as endorsing their adaptation of the article, do not modify the article in such a way as to damage the author's honor or reputation, and license their new adaptations or creations under identical terms (CC BY-NC-SA).
Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND): for non-commercial purposes, lets others distribute and copy the article, and to include in a collective work (such as an anthology), as long as they credit the author(s) and provided they do not alter or modify the article.

Elsevier has established agreements with funding bodies, http://www.elsevier.com/fundingbodies. This ensures authors can comply with funding body Open Access requirements, including specific user licenses, such as CC BY. Some authors may also be reimbursed for associated publication fees. If you need to comply with your funding body policy, you can apply for the CC BY license after your manuscript is accepted for publication.

To provide Open Access, this journal has a publication fee which needs to be met by the authors or their research funders for each article published Open Access.
Your publication choice will have no effect on the peer review process or acceptance of submitted articles.

The publication fee for this journal is **$3,000**, excluding taxes. Learn more about Elsevier's pricing policy: http://www.elsevier.com/openaccesspricing.

Language (usage and editing services)

Please write your text in good English (American or British usage is accepted, but not a mixture of these). Authors who feel their English language manuscript may require editing to eliminate possible grammatical or spelling errors and to conform to correct scientific English may wish to use the English Language Editing service available from Elsevier's WebShop (http://webshop.elsevier.com/languageediting/) or visit our customer support site (http://support.elsevier.com) for more information.

Submission

Submission to this journal proceeds totally online and you will be guided stepwise through the creation and uploading of your files. The system automatically converts source files to a single PDF file of the article, which is used in the peer-review process. Please note that even though manuscript source files are converted to PDF files at submission for the review process, these source files are needed for further processing after acceptance. All correspondence, including notification of the Editor's decision and requests for revision, takes place by e-mail removing the need for a paper trail.

Submit your article
Please submit your article via http://ees.elsevier.com/cg.

Referees

Please submit, with the manuscript, the names, addresses and e-mail addresses of three potential referees. Note that the editor retains the sole right to decide whether or not the suggested reviewers are used.

 Preparation

Use of word processing software

It is important that the file be saved in the native format of the word processor used. The text should be in single-column format. Keep the layout of the text as simple as possible. Most formatting codes will be removed and replaced on processing the article. In particular, do not use the word processor's options to justify text or to hyphenate words. However, do use bold face, italics, subscripts, superscripts etc. When preparing tables, if you are using a table grid, use only one grid for each individual table and not a grid for each row. If no grid is used, use tabs, not spaces, to align columns. The electronic text should be prepared in a way very similar to that of conventional manuscripts (see also the Guide to Publishing with Elsevier: http://www.elsevier.com/guidepublication). Note that source files of figures, tables and text graphics will be required whether or not you embed your figures in the text. See also the section on Electronic artwork.
To avoid unnecessary errors you are strongly advised to use the 'spell-check' and 'grammar-check' functions of your word processor.

Article structure

Subdivision - unnumbered sections
Divide your article into clearly defined sections. Each subsection is given a brief heading. Each heading should appear on its own separate line. Subsections should be used as much as possible when cross-referencing text: refer to the subsection by heading as opposed to simply 'the text'.

Introduction
State the objectives of the work and provide an adequate background, avoiding a detailed literature survey or a summary of the results.

Material and methods
Provide sufficient detail to allow the work to be reproduced. Methods already published should be indicated by a reference: only relevant modifications should be described.

Results
Results should be clear and concise.

Discussion
This should explore the significance of the results of the work, not repeat them. A combined Results and Discussion section is often appropriate. Avoid extensive citations and discussion of published literature.

Appendices
If there is more than one appendix, they should be identified as A, B, etc. Formulae and equations in appendices should be given separate numbering: Eq. (A.1), Eq. (A.2), etc.; in a subsequent appendix, Eq. (B.1) and so on. Similarly for tables and figures: Table A.1; Fig. A.1, etc.

Essential title page information

• **Title.** Concise and informative. Titles are often used in information-retrieval systems. Avoid abbreviations and formulae where possible.
• **Author names and affiliations.** Where the family name may be ambiguous (e.g., a double name), please indicate this clearly. Present the authors' affiliation addresses (where the actual work was done) below the names. Indicate all affiliations with a lower-case superscript letter immediately after the author's name and in front of the appropriate address. Provide the full postal address of each affiliation, including the country name and, if available, the e-mail address of each author.
• **Corresponding author.** Clearly indicate who will handle correspondence at all stages of refereeing and publication, also post-publication. **Ensure that phone numbers (with country and area code) are provided in addition to the e-mail address and the complete postal address. Contact details must be kept up to date by the corresponding author.**
• **Present/permanent address.** If an author has moved since the work described in the article was done, or was visiting at the time, a 'Present address' (or 'Permanent address') may be indicated as a footnote to that author's name. The address at which the author actually did the work must be retained as the main, affiliation address. Superscript Arabic numerals are used for such footnotes.

Abstract

A concise and factual abstract is required. The abstract should state briefly the purpose of the research, the principal results and major conclusions. An abstract is often presented separately from the article, so it must be able to stand alone. For this reason, References should be avoided, but if essential, then cite the author(s) and year(s). Also, non-standard or uncommon abbreviations should be avoided, but if essential they must be defined at their first mention in the abstract itself.

Keywords

Immediately after the abstract, provide a maximum of 6 keywords, using American spelling and avoiding general and plural terms and multiple concepts (avoid, for example, 'and', 'of'). Be sparing with abbreviations: only abbreviations firmly established in the field may be eligible. These keywords will be used for indexing purposes.

Acknowledgements

Collate acknowledgements in a separate section at the end of the article before the references and do not, therefore, include them on the title page, as a footnote to the title or otherwise. List here those individuals who provided help during the research (e.g., providing language help, writing assistance or proof reading the article, etc.).

The journal will adhere to the cytogenetics nomenclature of the latest version of the International System of Human Cytogenetic Nomenclature (ISCN 2009). One exception to the ISCN is to use a hyphen (-) instead of a tilde (~) when a break occurs between separate bands. This exception, which is only applicable to the title and abstract, is necessary because some abstracting and indexing databases, including PubMed, do not support the tilde (~). Human genes should be referenced according to HUGO and written in uppercase letters and italicized (e.g., *TP53*). RNA and proteins should be written in uppercase letters but not italicized. Commercial fluorescence in situ hybridization probe names should also be written in uppercase letters, not italicized. When referring to the probe itself, please use the name provided by manufacturer (e.g., TEL/AML1, AML1/ETO); however, the FISH nomenclature should be written using the gene symbols according to HUGO. All abbreviations should be defined when first used. Mutation descriptions must follow the Human Genome Variation Society (HGVS) nomenclature, available at http://www.hgvs.org/mutnomen/.

Math formulae

Present simple formulae in the line of normal text where possible and use the solidus (/) instead of a horizontal line for small fractional terms, e.g., X/Y. In principle, variables are to be presented in italics. Powers of e are often more conveniently denoted by exp. Number consecutively any equations that have to be displayed separately from the text (if referred to explicitly in the text).

Footnotes

Footnotes should be used sparingly. Number them consecutively throughout the article, using superscript Arabic numbers. Many wordprocessors build footnotes into the text, and this feature may be used. Should this not be the case, indicate the position of footnotes in the text and present the footnotes themselves separately at the end of the article. Do not include footnotes in the Reference list.

Table footnotes
Indicate each footnote in a table with a superscript lowercase letter.

Artwork

Electronic artwork
General points

• Make sure you use uniform lettering and sizing of your original artwork.
• Embed the used fonts if the application provides that option.
• Aim to use the following fonts in your illustrations: Arial, Courier, Times New Roman, Symbol, or use fonts that look similar.
• Number the illustrations according to their sequence in the text.
• Use a logical naming convention for your artwork files.
• Provide captions to illustrations separately.
• Size the illustrations close to the desired dimensions of the printed version.
• Submit each illustration as a separate file.
A detailed guide on electronic artwork is available on our website:
http://www.elsevier.com/artworkinstructions
You are urged to visit this site; some excerpts from the detailed information are given here.
Formats
If your electronic artwork is created in a Microsoft Office application (Word, PowerPoint, Excel) then please supply 'as is' in the native document format.
Regardless of the application used other than Microsoft Office, when your electronic artwork is finalized, please 'Save as' or convert the images to one of the following formats (note the resolution requirements for line drawings, halftones, and line/halftone combinations given below):
EPS (or PDF): Vector drawings, embed all used fonts.
TIFF (or JPEG): Color or grayscale photographs (halftones), keep to a minimum of 300 dpi.
TIFF (or JPEG): Bitmapped (pure black & white pixels) line drawings, keep to a minimum of 1000 dpi.
TIFF (or JPEG): Combinations bitmapped line/half-tone (color or grayscale), keep to a minimum of 500 dpi.
Please do not:
• Supply files that are optimized for screen use (e.g., GIF, BMP, PICT, WPG); these typically have a low number of pixels and limited set of colors;
• Supply files that are too low in resolution;
• Submit graphics that are disproportionately large for the content.

Color artwork
Please make sure that artwork files are in an acceptable format (TIFF (or JPEG), EPS (or PDF), or MS Office files) and with the correct resolution. If, together with your accepted article, you submit usable color figures then Elsevier will ensure, at no additional charge, that these figures will appear in color on the Web (e.g., ScienceDirect and other sites) regardless of whether or not these illustrations are reproduced in color in the printed version. **For color reproduction in print, you will receive information regarding the costs from Elsevier after receipt of your accepted article**. Please indicate your preference for color: in print or on the Web only. For further information on the preparation of electronic artwork, please see http://www.elsevier.com/artworkinstructions.
Please note: Because of technical complications which can arise by converting color figures to 'gray scale' (for the printed version should you not opt for color in print) please submit in addition usable black and white versions of all the color illustrations.

Illustration services

Elsevier's WebShop (http://webshop.elsevier.com/illustrationservices) offers Illustration Services to authors preparing to submit a manuscript but concerned about the quality of the images accompanying their article. Elsevier's expert illustrators can produce scientific, technical and medical-style images, as well as a full range of charts, tables and graphs. Image 'polishing' is also available, where our illustrators take your image(s) and improve them to a professional standard. Please visit the website to find out more.

Figure captions
Ensure that each illustration has a caption. Supply captions separately, not attached to the figure. A caption should comprise a brief title (**not** on the figure itself) and a description of the illustration. Keep text in the illustrations themselves to a minimum but explain all symbols and abbreviations used.

Tables

Number tables consecutively in accordance with their appearance in the text. Place footnotes to tables below the table body and indicate them with superscript lowercase letters. Avoid vertical rules. Be sparing in the use of tables and ensure that the data presented in tables do not duplicate results described elsewhere in the article.

References

Citation in text
Please ensure that every reference cited in the text is also present in the reference list (and vice versa). Any references cited in the abstract must be given in full. Unpublished results and personal communications are not recommended in the reference list, but may be mentioned in the text. If these references are included in the reference list they should follow the standard reference style of the journal and should include a substitution of the publication date with either 'Unpublished results' or 'Personal communication'. Citation of a reference as 'in press' implies that the item has been accepted for publication.

Reference links
Increased discoverability of research and high quality peer review are ensured by online links to the sources cited. In order to allow us to create links to abstracting and indexing services, such as Scopus, CrossRef and PubMed, please ensure that data provided in the references are correct. Please note that incorrect surnames, journal/book titles, publication year and pagination may prevent link creation. When copying references, please be careful as they may already contain errors. Use of the DOI is encouraged.

Web references
As a minimum, the full URL should be given and the date when the reference was last accessed. Any further information, if known (DOI, author names, dates, reference to a source publication, etc.), should also be given. Web references can be listed separately (e.g., after the reference list) under a different heading if desired, or can be included in the reference list.

Reference style
Text: Indicate references by number(s) in square brackets in line with the text. The actual authors can be referred to, but the reference number(s) must always be given.
List: Number the references (numbers in square brackets) in the list in the order in which they appear in the text.
Examples:
Reference to a journal publication:
[1] Van der Geer J, Hanraads JAJ, Lupton RA. The art of writing a scientific article. J Sci Commun 2010;163:51–9.
Reference to a book:

[2] Strunk Jr W, White EB. The elements of style. 4th ed. New York: Longman; 2000.
Reference to a chapter in an edited book:
[3] Mettam GR, Adams LB. How to prepare an electronic version of your article. In: Jones BS, Smith RZ, editors. Introduction to the electronic age, New York: E-Publishing Inc; 2009, p. 281–304.
Note shortened form for last page number. e.g., 51–9, and that for more than 6 authors the first 6 should be listed followed by 'et al.' For further details you are referred to 'Uniform Requirements for Manuscripts submitted to Biomedical Journals' (J Am Med Assoc 1997;277:927–34) (see also http://www.nlm.nih.gov/bsd/uniform_requirements.html).

Journal abbreviations source
Journal names should be abbreviated according to the List of Title Word Abbreviations:
http://www.issn.org/services/online-services/access-to-the-ltwa/.

Video data

Elsevier accepts video material and animation sequences to support and enhance your scientific research. Authors who have video or animation files that they wish to submit with their article are strongly encouraged to include links to these within the body of the article. This can be done in the same way as a figure or table by referring to the video or animation content and noting in the body text where it should be placed. All submitted files should be properly labeled so that they directly relate to the video file's content. In order to ensure that your video or animation material is directly usable, please provide the files in one of our recommended file formats with a preferred maximum size of 50 MB. Video and animation files supplied will be published online in the electronic version of your article in Elsevier Web products, including ScienceDirect: http://www.sciencedirect.com. Please supply 'stills' with your files: you can choose any frame from the video or animation or make a separate image. These will be used instead of standard icons and will personalize the link to your video data. For more detailed instructions please visit our video instruction pages at http://www.elsevier.com/artworkinstructions. Note: since video and animation cannot be embedded in the print version of the journal, please provide text for both the electronic and the print version for the portions of the article that refer to this content.

Supplementary data

Elsevier accepts electronic supplementary material to support and enhance your scientific research. Supplementary files offer the author additional possibilities to publish supporting applications, high-resolution images, background datasets, sound clips and more. Supplementary files supplied will be published online alongside the electronic version of your article in Elsevier Web products, including ScienceDirect: http://www.sciencedirect.com. In order to ensure that your submitted material is directly usable, please provide the data in one of our recommended file formats. Authors should submit the material in electronic format together with the article and supply a concise and descriptive caption for each file. For more detailed instructions please visit our artwork instruction pages at http://www.elsevier.com/artworkinstructions.

Submission checklist

The following list will be useful during the final checking of an article prior to sending it to the journal for review. Please consult this Guide for Authors for further details of any item.
Ensure that the following items are present:
One author has been designated as the corresponding author with contact details:
• E-mail address
• Full postal address
• Phone numbers
All necessary files have been uploaded, and contain:
• Keywords
• All figure captions
• All tables (including title, description, footnotes)
Further considerations
• Manuscript has been 'spell-checked' and 'grammar-checked'
• References are in the correct format for this journal
• All references mentioned in the Reference list are cited in the text, and vice versa
• Permission has been obtained for use of copyrighted material from other sources (including the Web)
• Color figures are clearly marked as being intended for color reproduction on the Web (free of charge) and in print, or to be reproduced in color on the Web (free of charge) and in black-and-white in print
• If only color on the Web is required, black-and-white versions of the figures are also supplied for printing purposes
For any further information please visit our customer support site at http://support.elsevier.com.

After Acceptance

Use of the Digital Object Identifier

The Digital Object Identifier (DOI) may be used to cite and link to electronic documents. The DOI consists of a unique alpha-numeric character string which is assigned to a document by the publisher upon the initial electronic publication. The assigned DOI never changes. Therefore, it is an ideal medium for citing a document, particularly 'Articles in press' because they have not yet received their full bibliographic information. Example of a correctly given DOI (in URL format; here an article in the journal *Physics Letters B*):
http://dx.doi.org/10.1016/j.physletb.2010.09.059
When you use a DOI to create links to documents on the web, the DOIs are guaranteed never to change.

Online proof correction

Corresponding authors will receive an e-mail with a link to our online proofing system, allowing annotation and correction of proofs online. The environment is similar to MS Word: in addition to editing text, you can also comment on figures/tables and answer questions from the Copy Editor. Web-based proofing provides a faster and less error-prone process by allowing you to directly type your corrections, eliminating the potential introduction of errors.
If preferred, you can still choose to annotate and upload your edits on the PDF version. All instructions for proofing will be given in the e-mail we send to authors, including alternative methods to the online version and PDF.
We will do everything possible to get your article published quickly and accurately - please upload all of your corrections within 48 hours. It is important to ensure that all corrections are sent back to us in one communication. Please check carefully before replying, as inclusion of any subsequent corrections cannot be guaranteed.

Proofreading is solely your responsibility. Note that Elsevier may proceed with the publication of your article if no response is received.

Author Inquiries

For inquiries relating to the submission of articles (including electronic submission) please visit this journal's homepage. For detailed instructions on the preparation of electronic artwork, please visit http://www.elsevier.com/artworkinstructions. Contact details for questions arising after acceptance of an article, especially those relating to proofs, will be provided by the publisher. You can track accepted articles at http://www.elsevier.com/trackarticle. You can also check our Author FAQs at http://www.elsevier.com/authorFAQ and/or contact Customer Support via http://support.elsevier.com.

Readers	Authors	Librarians	Editors	Reviewers	Advertisers/ Sponsors	Societies
View Articles	Guide for Authors	Ordering Information				
	Submit your paper	Abstracting/ Indexing				
	Author information pack					