UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

TESE DE DOUTORADO

ALTERAÇÕES MOLECULARES DO GENE *TP53* E DE GENES QUE REGULAM A ATIVIDADE DA P53 NA INFERTILIDADE E NO CÂNCER

DIEGO D'AVILA PASKULIN

Porto Alegre - Junho

2013

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

TESE DE DOUTORADO

ALTERAÇÕES MOLECULARES DO GENE *TP53* E DE GENES QUE REGULAM A ATIVIDADE DA P53 NA INFERTILIDADE E NO CÂNCER

DIEGO D'AVILA PASKULIN

Tese submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da UFRGS como requisito parcial para a obtenção do grau de Doutor em Ciências (Genética e Biologia Molecular).

Orientadora: Prof. Dra. Patricia Ashton-Prolla

Co-Orientadora: Prof. Dra. Marileila Varella-Garcia

Porto Alegre - Junho

2013

APRESENTAÇÃO

O presente estudo compreende duas abordagens paralelas de caracterização do papel de p53 na reprodução humana e no desenvolvimento do câncer. Conforme formato requerido pelo Programa de Pós-Graduação em Genética e Biologia Molecular da Universidade Federal do Rio Grande do Sul, esta tese está dividida em: *Introdução*; *Objetivos*; *Capítulos* (substituirão as seções de Material e Métodos e Resultados descrevendo os resultados obtidos no período sob forma de artigos científicos publicados, artigos submetidos ou ainda resultados de investigações laboratoriais em andamento ou recentemente concluídas); *Discussão e Perspectivas*; *Conclusões*. Os estudos foram orientados pela Dra. Patricia Ashton-Prolla no Brasil e Dra. Marileila Varella-Garcia nos Estados Unidos da América. O aluno recebeu bolsa de estudos concedida pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) durante o período de doutoramento no Brasil e bolsa de estudos concedida pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) para seu doutorado "sandwich".

INSTITUIÇÕES EXECUTORAS: Este trabalho foi desenvolvido no Laboratório de Medicina Genômica do Centro de Pesquisa Experimental do Hospital de Clínicas de Porto Alegre e no *CytoCore Laboratory – Department of Medicine - University of Colorado* em Aurora, Colorado, Estados Unidos da América.

FONTES FINANCIADORAS: CAPES, CNPq (Edital Universal 475471/2009-1), FAPERGS (Edital PPSUS 2009 e Edital PRONEX), Fundo de Incentiivo a Pesquisa e Eventos – HCPA (Projeto 09-430) e GlaxoSmithKline (Ethnic Research Initiative – ERI – Grant 2009).

DEDICATÓRIA

À Jéssica, pelo amor, por ser parte da minha vida e representar o que nela há de melhor.

Aos meus pais Giorgio e Katya, e a minha irmã Livia pelo amor, união, exemplo, respeito e pelo constante apoio na busca dos meus objetivos.

AGRADECIMENTOS

À professora Dra. Patricia Ashton-Prolla, por ter me proporcionado esses quatro anos de muito estudo, pesquisa, crescimento pessoal e profissional, e pelas oportunidades de colaboração.

À professora Dra. Marileila Varella-Garcia, pela incrível experiência que eu vivenciei em seu laboratório na Universidade do Colorado, pelo constante incentivo, cobrança e pelo investimento para o crescimento profissional e pessoal de seus alunos.

Aos colegas do Hospital de Clínicas de Porto Alegre MSc. Filippo Vairo, MSc. Mariana Kiehl, Dra. Juliana Giacomazzi, MSc. Bárbara Alemar, Dra. Ana Paula Brandalize, MSc. Gabriel Macedo, Patricia Silva, Patricia Izetti, MSc. Liliane Todeschini, Dra. Liliana Cossio, Dra. Ingrid Ewald e Dra. Patricia Koehler pela amizade e convivência diária.

Aos professores Dr. João Sabino Lahorgue Cunha-Filho e Dr. Carlos Augusto Bastos de Souza pela coleta de dados e discussões sobre infertilidade.

Às professoras Dra. Lavínia Faccini, Dra. Úrsula Matte, Dra. Maria Cátira Bortolini e aos colegas Dra. Vanessa Paixão-Cortes, MSc. Alice Tagliani e MSc. Lucas Rosa Fraga, pelas eternas discussões envolvendo p53 e fertilidade.

Ao professor Dr. Sidney Emanuel Batista dos Santos e ao Dr. Pierre Hainaut, pelas discussões envolvendo evolução molecular e a mutação *TP53* p.R337H.

Aos colegas do *CytoCore Lab* da Universidade do Colorado Dra. Margareth Skokan, MSc. Sakshi Mahale, MSc. Severine Kako, Cecilia Nguyen, Teresa Nguyen, Nathan Schulte e Dr. Liang-Guo Xu, pelo auxílio durante os experimentos de FISH e cultura celular.

Ao colegas do *ArrayCore Lab* da Universidade do Colorado professor Dr. Bifeng Gao, Dra. Okyong Cho, Jessica Rice, Steven Glidewell e em especial ao meu

amigo Dr. Ichiro Nakachi pelo auxílio durante os experimentos de SNP Array e pela análise de bioinformática.

Ao amigo Elmo Cardoso pelos inenarráveis auxílios durante todo o meu período de doutoramento.

Ao CNPq e a CAPES, por terem me concedido as bolsas de estudo e os *grants* necessários para a realização de todos meus estudos.

Ao Programa de Pós-Graduação em Genética e Biologia Molecular e seus professores que me proporcionaram um ambiente propício para meu crescimento e amadurecimento profissional.

SUMÁRIO

LISTA DE ABREVIATURAS 10

LISTA DE FIGURAS...13

RESUMO 15

ABSTRACT 16

1. INTRODUÇÃO..17

1.1 Carcinogênese e o gene TP53 17

1.2 Estrutura e Regulação do gene TP53..19

1.3 *TP53* e Câncer 21

1.3.1 TP53 e câncer hereditário...21

1.3.2 Mutação TP53 p.R337H...23

1.4 *TP53* e Fertilidade 27

1.4.1 Função e regulação de p53 relacionadas à fertilidade 27

1.4.2 Família p53 e fertilidade...29

1.4.3 Polimorfismos nos genes TP53, MDM2, MDM4, USP7, LIF, TP63, TP73 e ESR1...30

1.4.4 Infertilidade, endometriose e polimorfismos em genes da via de sinalização de p53 33

2. OBJETIVOS..35

2.1 Objetivo Geral 35

2.2 Objetivos Específicos...35

3. CAPÍTULO 1 37

3.1 *TP53 PIN3 and PEX4 polymorphisms and infertility associated with endometriosis or post-in vitro fertilization failure* 37

4. CAPÍTULO 2 50

4.1 *The TP53 fertility network* 50

5. CAPÍTULO 3 60

5.1 *ESR1 rs9340799 is associated with endometriosis-related infertility and in vitro fertilization failure* 60

6. CAPÍTULO 4 84

6.1 Análise genômica de tumores de pacientes com mutações germinativas no gene *TP53*.....84

6.2 Validação de regiões genômicas alteradas em pacientes portadores de mutações germinativas no gene *TP53* utilizando hibridização *in situ* fluorescente (FISH).....103

7. DISCUSSÃO e PERSPECTIVAS 109

7.1 *TP53* e Fertilidade.....109

7.2 *TP53* e Câncer 111

8. CONCLUSÕES.....118

8.1 *TP53* e Fertilidade.....118

8.2 *TP53* e Câncer 119

9. REFERÊNCIAS BIBLIOGRÁFICAS 120

10. PUBLICAÇÕES COMPLEMENTARES NO PERÍODO DE DOUTORADO 128

10.1 *High twinning rate in Cândido Godói: a new role for p53 in human fertility* 128

10.2 *Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer*.....135

10.3 *p53 signaling pathway polymorphisms associated to recurrent pregnancy loss*.....154

11. ANEXOS 177

11.1 ANEXO A – Descrição do Human Omni 2.5-8 Bead Chip 177

11.2 ANEXO B – Região de perda de heterozigosidade em 11p15.5-15.4 do paciente ACC-3.....178

LISTA DE ABREVIATURAS

A – Adenina

ABL – Abelson Murine Leukemia Viral Oncogene Homolog 1

ABL2 – v-abl Abelson Murine Leukemia Viral Oncogene Homolog 2

ACC – Adrenocortical Carcinoma

ALK – Anaplastic Lymphoma Receptor

Amp – Amplificação

Arg – Arginina

ATM – Ataxia Telangiectasia Mutated

BAF – B Allele Frequency

BC – *Breast Cancer*

C – Citosina

G – Guanina

CDKN1C – Cyclin Dependent Kinase Inhibitor 1C

CHORI – *Children's Hospital Oakland Research Institute*

CN-LOH – *Copy Neutral Loss of Heterozygosity*

CNV – *Copy Number Variation*

DAPI – *4',6-Diamidino-2-Fenilindol*

DBD – *DNA Binding Domain*

Del – Deleção

DNA – *Deoxyribonucleic Acid*

ESR1 – Estrogen Receptor 1

FISH – *Fluorescent in situ Hybridization*

FIV – Fertilização *in vitro*

H – Histidina

H19 – Adult Skeletal Muscle

HMMR – Hyaluronan-Mediated Motility Receptor

HRAS – v-Ha-ras Harvey Rat Sarcoma Viral Oncogene Homolog

IARC – *International Agency Cancer Research*

KRAS – v-Ki-ras2 Kirsten Rat Sarcoma Viral Oncogene Homolog

LFL – *Li-Fraumeni Like Syndrome*

LFS – *Li-Fraumeni Syndrome*

LIF – Leukemia Inhibitory Factor

LKB1 – Serine/Threonine Kinase 11

LOH – *Loss of Heterozygosity*

LRR – *Log_2R Ratio*

MB – Meduloblastoma

MDM2 – Mouse Double Minute 2 Homolog

MDM4 – Mouse Double Minute 4 Homolog

mRNA – *Messenger RNA*

MT1 – Melatonin Receptor 1

OD – *Oligomerization Domain*

OMIM – *Online Mendelian Inheritance in Man*

P – Prolina

PEX4 – *Polymorphism in exon 4*

pH – Potencial de Hidrogênio

PIN2 – *Polymorphism in intron 2*

PIN3 – *Polymorphism in intron 3*

Pro – Prolina

PRR – *Proline Rich Region*

PTEN – Phosphatase and Tensin Homolog

R – Arginina

RD – Regulatory Domain

SG – *Spectrum Green*

SLC22A1L – Solute Carrier Family 22 A1L

SNP – *Single-Nucleotide Polymorphism*

Sp1 – Sp1 Transcription Factor

SR – *Spectrum Red*

SRC – V-Src Sarcoma (Schmidt-Ruppin A-2) Viral Oncogene Homolog

T – Timina

TAD1 – *Transactivation Domain 1*

TAD2 – *Transactivation Domain 2*

TP53 – Tumor Protein p53

TP63 – Tumor Protein p63

TP73 – Tumor Protein p73

TSC1 – Tuberous Sclerosis 1

TSC2 – Tuberous Sclerosis 2

USP7 – Ubiquitin Specific Protease 7

XP – Xeroderma pigmentosum

LISTA DE FIGURAS

Figura 1 - A ampla diversidade de funções do gene *TP53*. 18

Figura 2 - Distribuição dos onze éxons do gene *TP53* e dos seis domínios da proteína p53. 19

Figura 3 - A alça de *feedback* p53-MDM2. 20

Figura 4 - Perfil de desnaturação das proteínas p53 mutante (p.R337H em homozigose) (a) e p53 selvagem (b). 24

Figura 5 - Polimorfismos em genes da via de sinalização de p53 escolhidos para o estudo. 29

Figura 6 – Resultado do SNP Array por todos os cromossomos do Paciente ACC-1. 87

Figura 7 – Resultado do SNP Array por todos os cromossomos do Paciente ACC-2. 89

Figura 8 – Resultado do SNP Array por todos os cromossomos do Paciente ACC-3. 91

Figura 9 – Resultado do SNP Array por todos os cromossomos do Paciente ACC-4. 93

Figura 10 – Resultado do SNP Array por todos os cromossomos do Paciente ACC-1. 95

Figura 11 – Resultado do SNP Array por todos os cromossomos do Paciente ACC-6. 97

Figura 12 – Resultado do SNP Array por todos os cromossomos do Paciente BC-1. 99

Figura 13 – Resultado do SNP Array por todos os cromossomos do Paciente BC-2. 101

Figura 14 – Validação da sonda ABL2(SR)/CEN1(SG) em linhagem celular AG09391 mapeando corretamente em 1q25.2..105

Figura 15 – Paciente ACC-1. KRAS(SR)/CEP12(SG). 106

Figura 16 – Paciente ACC-6. ABL2(SR)/CEP1(SG). 107

Figura 17 – Paciente BC-1. CDKN1C(SR)/CEP11(SG)...................................108

Figura 18 – Paciente BC-2. CDKN1C(SR)/CEP11(SG). FISH escore: Perda. Escala 10µm 108

Figura 19 - Resultado do SNP Array para o cromossomo 17 do paciente ACC-5 demonstrando a perda de heterozigosidade sem alteração do número de cópias na região 17p13.2-p13.1. 115

Figura 20 - Resultado do SNP Array para parte do cromossomo 17 do paciente ACC-5 demonstrando a perda de heterozigosidade sem alteração do número de cópias na região 17p13.2-p13.1 (~4.2Mb)..............................116

Figura 21 - Resultado do SNP Array para o cromossomo 17 do paciente ACC-5 demonstrando as BAFs do paciente ACC-5 em relação aos pais portadores da mutação *TP53* p.R337H. 117

Figura 22 – HumanOmni2.5-8 BeadChip. 177

Figura 23 - Resultado do SNP Array para o cromossomo 11 do paciente ACC-3 demonstrando a perda de heterozigosidade sem alteração do número de cópias na região 11p15.5 - 11p15.4. 178

RESUMO

O gene *TP53* atua como um fator de transcrição regulando a expressão de genes que controlam a progressão do ciclo celular, angiogênese e apoptose. *TP53* é reconhecido como o gene mais frequentemente mutado em câncer, e mutações germinativas estão associadas à síndrome de predisposição hereditária ao câncer denominada Li-Fraumeni (SLF). A mutação germinativa p.R337H em *TP53* está presente em cerca de 1 em cada 300 indivíduos da população geral do Brasil. Em nosso estudo utilizamos uma plataforma de varredura genômica para buscar marcadores moleculares associados ao processo carcinogênico em portadores da mutação *TP53* p.R337H com câncer de mama e carcinoma adrenocortical e utilizamos a técnica de FISH para validá-las em tecido tumoral. Nossos resultados demonstraram presença de variações do número de cópias (CNVs): aumento do número de cópias em 12p13.3-12p11.23, 16p13.3-16q24.3, amplificações da região 1q24.2-1q25.3, deleções em 2p25.3-2q37.3 e 17p13.1 e por fim, perda de heterozigosidade sem alteração do número de cópias em 11p15.5-p11.2 e 17p13.1 em pacientes portadores da mutação p.R337H. Nossos resultados apontam novas regiões genômicas associadas com a SLF em portadores da mutação *TP53* p.R337H. Além de seu posto de supressor tumoral, *TP53* tem importante função na reprodução humana pela regulação do gene *LIF*, citocina essencial no processo de implantação do blastocisto. Desta forma, avaliamos a frequência de polimorfismos de genes da via de sinalização de *TP53* em mulheres inférteis com endometriose e em mulheres inférteis com falhas recorrentes de implantação durante fertilização *in vitro.* Os resultados evidenciam associação entre o polimorfismos *TP53* rs17878362, *TP53* rs1042522 e *ESR1* rs9340799 e os desfechos de infertilidade estudados. A caracterização de polimorfismos da via de sinalização de *TP53* poderá ser importante no entendimento da etiopatogenia da endometriose e da infertilidade associada a anormalidades neste período gestacional, e poderão ser utilizados como biomarcadores com impacto na decisão sobre estratégias de tratamento para estas condições.

ABSTRACT

The p53 tumor suppressor gene is a pivotal regulator of different cellular pathways including apoptosis, DNA damage, oncogene activation, or hypoxia. Germline mutations in *TP53* are the underlying defect of Li-Fraumeni Syndrome (LFS), an autosomal dominant disorder characterized by predisposition to multiple early-onset cancers. A variant form of LFS is commonly detected due to the high prevalence of the *TP53* p.R337H mutation, which is present in about 0.3% of the population of Southern Brazil. Our goal is to use high density SNP genotyping arrays to discover new imbalanced regions associated with p.R337H and to use *Fluorescent in situ Hybridization* to validate these regions in adrenocortical carcinoma and breast cancer patients. Our results show gains in 12p13.3-12p11.23, 16p13.3-16q24.3, amplifications in 1q24.2-1q25.3, deletions in 2p25.3-2q37.3 and 17p13.1 and copy neutral loss of heterozygosity in 11p15.5-p11.2 and 17p13.1 in patients with *TP53* p.R337H mutation. Our findings point out to new chromosomal regions associated with LFS patients carrying the *TP53* p.R337H mutation. Besides its major position as a tumor suppressor gene, *TP53* plays a crucial role in human fertility as it regulates leukemia inhibitory factor (LIF) expression. Although the important interaction between p53 and LIF is crucial for embryo implantation, we believe that not only *LIF* and *TP53*, but also other genes in the *TP53* signaling pathway may be important in the implantation process. To test this hypothesis we decided to examine whether *TP53* signaling pathway genes (*MDM2, MDM4, USP7, LIF, TP63, TP73* e *ESR1*) polymorphisms may be involved with infertile women with endometriosis and with recurrent *in vitro* fertilization (IVF) failure. Our results demonstrate that *TP53* rs17878362, *TP53* rs1042522 and *ESR1* rs9340799 are associated with endometriosis-related infertility and with failure of implantation after IVF. *TP53* signaling pathway genes polymorphisms may have a role as biomarkers and could add to the development of a clinically relevant genetic profile that would be of great help for clinicians to identify patients at higher risk for IVF failure.

1. INTRODUÇÃO

1.1 CARCINOGÊNESE E O GENE TP53

Carcinogênese (do grego *karkinos*, câncer; *genesis*, geração) é definida como um processo de múltiplas alterações genéticas que modificam a homeostasia normal de uma célula determinando um desenvolvimento celular descontrolado que origina câncer. Os mecanismos moleculares que regem a carcinogênese foram definidos inicialmente por Hanahan e Weinberg em 2000 como "*Hallmarks of Cancers*" e incluem seis importantes mecanismos de controle do ciclo celular: apoptose, diferenciação celular, senescência, angiogênese, motilidade e capacidade migratória (Hanahan *et al.*, 2000). Em 2011, o controle do metabolismo energético e do sistema imune, o aumento da instabilidade genômica e a promoção de mecanismos de inflamação foram incluídos como a nova geração dos "*Hallmarks of Cancers*" (Hanahan *et al.*, 2011). Recentes avanços no entendimento dos processos que regulam a carcinogênese têm sugerido uma complexa rede de interação entre células normais e tumorais gerando o conceito de "microambiente tumoral" (Hanahan *et al.*, 2011).

O gene *TP53*, codificador da fosfoproteina p53, ocupa uma posição de destaque no entendimento do complexo processo carcinogênico no microambiente tumoral. Inicialmente identificado com importante função no processo de reparo ao DNA, o gene *TP53* foi nomeado "Guardião do Genoma" em 1992, um termo figurativo que reflete sua função de supressor tumoral ao bloquear o processo de proliferação em células com DNA alterado por fatores carcinogênicos, levando-as à parada do ciclo celular ou à apoptose (Lane, 1992). O papel da proteína p53 como supressora tumoral está amplamente caracterizado experimentalmente. Ela induz parada no ciclo celular, senescência e apoptose em resposta a uma grande variedade de estresses celulares como hipóxia, dano ao DNA, perda de adesão celular, quebra de fuso mitótico, frio e calor, encurtamento telomérico e ativação de oncogenes

(Vousden *et al.*, 2009). A proteína p53 atua diretamente em diversos efetores como fator de transcrição sequência-específico, mas também regula proteínas da família Bcl-2 (Moll *et al.*, 2005), atua no processamento de microRNAs (Blount *et al.*, 1994) e lncRNAs (Huarte *et al.*, 2010) e também possui função essencial no processo de implantação regulando LIF (Hu *et al.*, 2007b) (Figura 1).

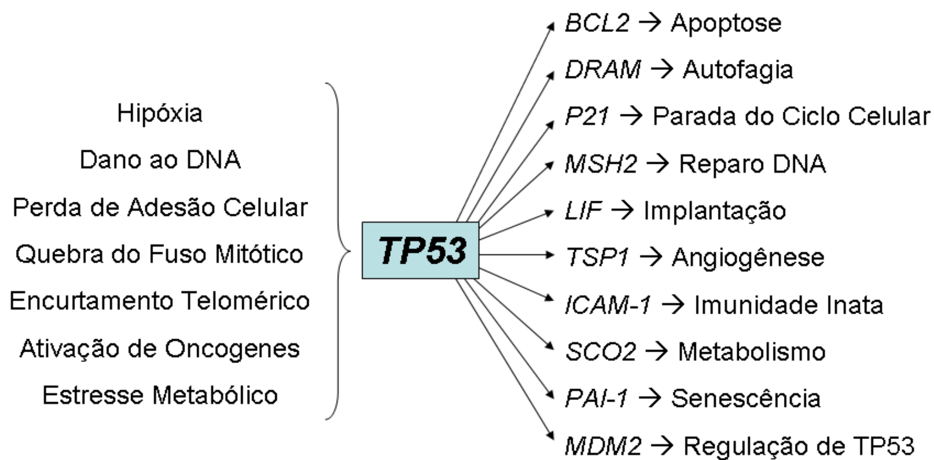


Figura 1 - A ampla diversidade de funções do gene *TP53*. Exemplos de estresses celulares que resultam na ativação de uma cascata de eventos relacionada à sinalização de *TP53* e exemplos de genes-alvo que atuam em diferentes processos celulares. Modificado de Vousden *et al.,* 2009.

Diversos estudos e dados do IARC (*International Agency for Research in Cancer*) apontam o gene *TP53* como o mais frequentemente mutado em câncer independente do tipo tumoral, localização, histologia ou história natural do tumor (Petitjean *et al.*, 2007). Desde sua descrição em 1979, o gene *TP53* e sua complexa via de sinalização têm sido importantes alvos de pesquisa em câncer devido à sua ampla gama de funções em diferentes processos celulares (Lane *et al.*, 2010). Recentemente estudos que avaliam padrões de alterações genéticas ao longo de todo o genoma confirmam que *TP53* é o gene mais frequentemente alterado no câncer em humanos (Pfeifer *et al.*, 2011), sendo demonstrado que aproximadamente 50% de todos os tipos de câncer apresentam alterações que inativam a função supressora tumoral de p53 (Lane *et al.*, 2010; Marcel *et al.*, 2010).

1.2 ESTRUTURA E REGULAÇÃO DO GENE TP53

O gene supressor tumoral *TP53* (OMIM 191170) está localizado no braço curto do cromossomo 17 (17p13.1), possui onze éxons, sendo o primeiro não codificante, 19.198 nucleotídeos e codifica a fosfoproteína nuclear p53 com 393 aminoácidos (Linzer *et al.*, 1979). A proteína p53 apresenta seis domínios e sua expressão é controlada por modificações transcricionais, pós-transcricionais, traducionais e pós-traducionais (Hollstein *et al.*, 2010). A região N terminal da proteína possui dois domínios iniciais de transativação (*Transactivation Domain* 1, TDA1 e *Transactivation Domain* 2, TDA2) no qual se encontra o sítio preferencial de ligação da fosfoproteína MDM2 entre os aminoácidos 18 e 26, e um domínio rico em prolina (*Proline Rich Region,* PRR) envolvido na interação proteína-proteína e na função pró-apoptótica de p53. A região alvo de 84% das mutações somáticas no gene *TP53* que inclui os éxons 5 ao 8 é chamada de domínio de ligação ao DNA (*DNA binding domain,* DBD), sendo responsável pela ligação da proteína p53 a seus genes alvos. A região C terminal é formada pelo domínio de oligomerização (*Oligomerization Domain,* OD) que possui papel na conformação quartenária da proteína e por fim um domínio de regulação (*Regulatory Domain,* RD) importante local de fosforilação de p53 (Figura 2) (Hollstein *et al.*, 2010).

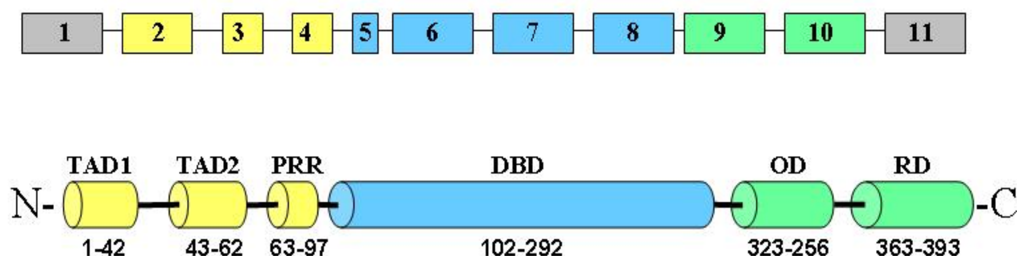


Figura 2 - Distribuição dos onze éxons do gene *TP53* e dos seis domínios da proteína p53. TAD1: Domínio de Transativação 1; TAD2: Domínio de Transativação 2; PRR: Região Rica em Prolina; DBD: Domínio de Ligação ao DNA; OD: Domínio de Oligomerização; RD: Domínio de Regulação.

Em resposta a um fator de estresse, os níveis da proteína p53 são rapidamente elevados pela combinação do aumento dos níveis de transcrição do gene *TP53 (Fu et al., 1996)* e pela diminuição das taxas de degradação da proteína p53, o que ocorre com participação de várias proteínas, em especial, MDM2 (Maltzman *et al.*, 1984). A proteína MDM2 é uma ubiquitina-ligase que regula os níveis e a atividade da proteína p53, ligando-se à região N-terminal desta última e adicionando covalentemente ubiquitinas (resíduos de 76 aminoácidos) que "marcam" a proteína p53 para rápida proteólise pelo proteossomo 26S. Por outro lado, a transcrição do gene *MDM2* (OMIM 164785) é regulada pela proteína p53 que se liga a regiões reguladoras e estimula a transcrição de *MDM2*. Ao ligar-se à p53, MDM2 cria uma "alça de *feedback":* os níveis de p53 são controlados por MDM2 e vice-versa (Figura 3).

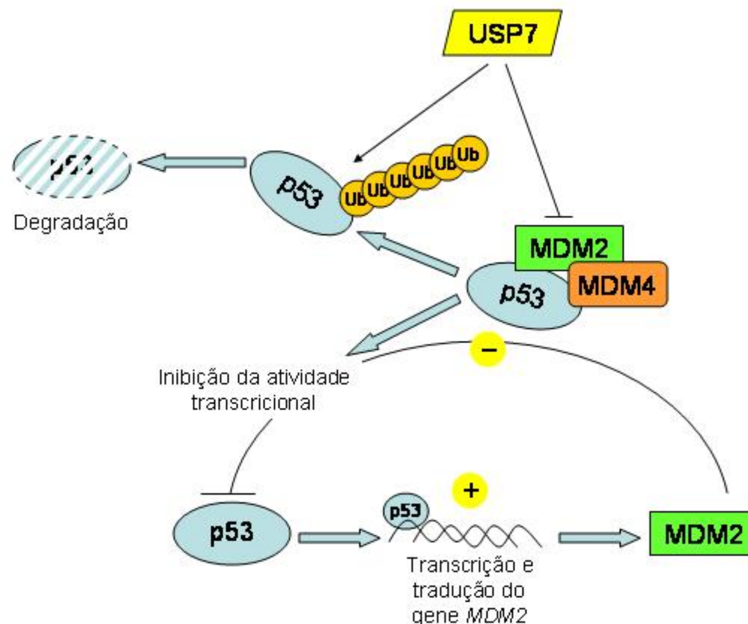


Figura 3 - A alça de *feedback* p53-MDM2. As proteínas MDM2 e MDM4 se ligam e inativam p53 pela inibição das atividades transcricionais de p53 e pela promoção da ubiquitinação e subsequente degradação proteossômica de p53. A proteína USP7 desubiquitina p53 e inativa complexo MDM2 e MDM4. p53 liga-se à região promotora de *MDM2* regulando sua expressão. Adaptado de Hu *et al.*, 2007.

Além de MDM2, a proteína MDM4 (OMIM 602704) também possui função de ubiquitina-ligase. Através de sua hetero-oligomerização com MDM2, MDM4 também exerce função reguladora de p53. Atwal *et al.* demonstraram que a haploinsuficiência de *MDM4* leva à diminuição da atividade de p53, resultando em aumento do dano ao DNA e susceptibilidade a tumores (Atwal *et al.*, 2009). Ao contrário de MDM2 e MDM4 que ubiquitinam a proteína p53, a proteína USP7 (OMIM 602519) que pertencente à família das proteases ubiquitina-específicas promove a desubiquitinação de p53 (Brooks *et al.*, 2004; Everett *et al.*, 1997), possuindo importante papel na regulação dos níveis de p53.

1.3 *TP53* E CÂNCER

1.3.1 *TP53* E CÂNCER HEREDITÁRIO

Genes ligados às vias de sinalização de *TP53* que apresentam mutações altamente penetrantes, porém de baixa frequência, são responsáveis por uma gama de síndromes de predisposição ao câncer, como por exemplo Ataxia Telangiectasia (gene *ATM*), Síndrome de Cowden (gene *PTEN*), Xeroderma Pigmentoso (gene *XP*), Complexo Esclerose Tuberosa 1 e 2 (genes *TSC1* e *TSC2*) e a Síndrome de Peutz–Jeghers (gene *LKB1*) (Kraemer *et al.*, 1994; Liaw *et al.*, 1997; McKinnon, 2004; Miyoshi *et al.*, 2002; Orlova *et al.*, 2010).

O gene *TP53* é altamente polimórfico em regiões codificadoras e não-codificadoras. Polimorfismos em *TP53* são associados a uma maior susceptibilidade ao câncer e também foram demonstrados como modificadores do fenótipo de câncer em pacientes portadores de mutações germinativas do próprio gene *TP53* e também de outros genes (Whibley *et al.*, 2009). A maioria das mutações somáticas (82%) de *TP53* são polimorfismos de um único nucleotídeo (SNP) e 18% são inserções ou deleções que resultam em uma verdadeira perda de função do alelo mutado. As alterações mais frequentes (90%) são mutações de sentido trocado (*missense mutations*), onde a mudança

de um par de nucleotídeos acarreta na modificação de um códon sendo a maioria destas localizadas na região do domínio de ligação ao DNA (entre os éxons 5-8 do gene) e associadas à mudança da atividade transcricional do gene (Petitjean *et al.*, 2007). Aproximadamente 30% das mutações de sentido trocado germinativas ou somáticas ocorrem em seis códons preferenciais (175, 245, 248, 249, 273 e 282), também chamados de "*hotspots*" mutacionais de *TP53*. Outras alterações incluem mutações sem sentido (6%) e mutações silenciosas (4%) (Olivier *et al.*, 2002).

O gene *TP53* foi associado à carcinogênese frente à detecção de mutações germinativas em pacientes que apresentavam uma síndrome hereditária de predisposição a múltiplos tumores de um espectro comum, descrita por Frederick Pei Li e Joseph Fraumeni Jr. em 1969 (Li *et al.*, 1969; Malkin *et al.*, 1990). As famílias originalmente descritas por Li e Fraumeni apresentavam em sua maioria uma alta frequência de câncer de mama em adultos jovens, sarcomas, tumores do sistema nervoso central, leucemias e carcinoma adrenocortical na infância ou adolescência. Realizando sequenciamento dos éxons 5-8 do gene *TP53* de indivíduos provenientes de 5 famílias com critérios clínicos da hoje conhecida como Síndrome de Li-Fraumeni (SLF OMIM151623), Malkin *et al.* associaram pela primeira vez uma alteração genética germinativa com a ocorrência da SLF, ao encontrarem mutações germinativas no gene *TP53* em todas as famílias estudadas (Malkin *et al.*, 1990). A SLF possui prevalência estimada entre 1:5.000 e 1:20.000 nos Estados Unidos e no Reino Unido (Gonzalez *et al.*, 2009; Lalloo *et al.*, 2003) e é caracterizada por agregação familiar de casos de sarcomas, tumores cerebrais, tumores de mama, leucemia e carcinoma adrenocortical (ACC) (Li *et al.*, 1969; Malkin *et al.*, 1990; Varley *et al.*, 1997). Famílias apresentando quadros clínicos similares aos da SLF são classificadas como pertencentes à Síndrome de Li-Fraumeni-*Like* (LFL) com espectro de tumores maior que o originalmente descrito e penetrância reduzida. Outras neoplasias que parecem estar associadas com a síndrome incluem melanomas, leucemias, carcinomas colorretais, cânceres de pulmão e gástricos, tumores phyllodes da mama, tumores de células germinativas, papilomas de plexo coróide e tumores de

Wilms (Birch *et al.*, 2001; Birch *et al.*, 1994; Chompret *et al.*, 2001).

1.3.2 MUTAÇÃO *TP53* p.R337H

Em 2001, Ribeiro *et al.* detectaram a troca de uma arginina (R) por uma histidina (H) no códon 337 (A→G, coordenada genômica GRCCh37.p5 17:7572929) do éxon 10 do gene *TP53* em 35 crianças com carcinoma adrenocortical, provenientes de 25 famílias do estado do Paraná, Brasil, mas que não apresentavam outros tumores associados, nem mesmo em seus familiares, sugerindo que esta nova mutação, p.R337H, poderia estar associada a um fenótipo tecido-específico e predispor exclusivamente a esse tipo de câncer (Ribeiro *et al.*, 2001). A proteína p53 liga-se ao DNA na forma de um tetrâmero formado por dímeros de dímeros sendo a estabilidade do tetrâmero essencial para a correta função da proteína. Resultados de um único estudo experimental *in vitro* realizado com células p.R337H mutantes indicam que a proteína mutante p.R337H tem propriedades funcionais especiais dependentes de pH. O que se depreende a partir dos resultados deste estudo é que a proteína mutante forma oligômeros e retém atividade normal em pH fisiológico (próximo de 7), mas quando o pH é elevado a 8 a proteína perde estabilidade, por desnaturação do tetrâmero, ou seja, o fenótipo mutante somente se evidencia em pH alcalino (DiGiammarino *et al.*, 2002). Em condições fisiológicas de pH (em torno de 7) e temperatura (cerca de 37°C) aproximadamente 20% da proteína p53 mutante encontrava-se desnaturada, ao passo que com a mudança do pH para 8, 70% da p53 encontrava-se desnaturada. O mesmo não ocorreu com a proteína p53 selvagem, mesmo com drásticas mudanças de pH a proteína permanecia inalterada e aparentemente funcional (Figura 4).

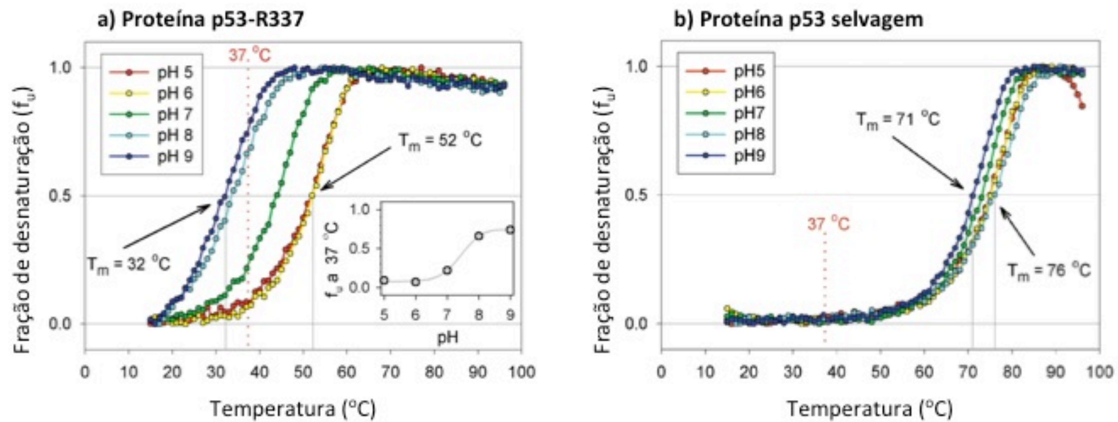


Figura 4 - Perfil de desnaturação das proteínas p53 mutante (p.R337H em homozigose) (a) e p53 selvagem (b). Adaptado de Giammarino *et al.*, 2002.

Do ponto de vista clínico, no entanto, a detecção da mutação p.R337H em heterozigose em indivíduos de famílias com uma ampla gama de tumores hereditários, compatível com as definições SLF/LFL, indica que esta mutação tem o potencial de predispor a tumores em diversos tecidos, diferente de uma predisposição tecido-específica para carcinoma adrenocortical como sugerido inicialmente (Achatz *et al.*, 2007; Hainaut, 2002; Ribeiro *et al.*, 2001). Estudos posteriores envolvendo famílias brasileiras demonstraram que outros tumores além de carcinoma adrenocortical ocorriam com alta frequência em portadores da mutação, tais como: carcinoma de plexo coróide, cânceres gástrico, colônico e de tireóide, osteossarcoma e câncer de mama (Assumpcao *et al.*, 2008; Custodio *et al.*, 2011; Gomes *et al.*, 2012; Latronico *et al.*, 2001; Seidinger *et al.*, 2011).

Em um estudo caso-controle envolvendo 123 mulheres com câncer de mama no sudeste do Brasil, Assumpção *et al.* identificaram três portadoras da mutação p.R337H, duas destas apresentando histórico familiar de LFL (Assumpcao *et al.*, 2008). Estes resultados corroboraram os achados de Achatz *et al.,* que identificaram a mutação germinativa p.R337H em indivíduos com diferentes tumores, inclusive câncer de mama, sendo demonstrada perda de heterozigosidade para a mutação em amostra de câncer de mama de uma portadora (Achatz *et al.*, 2007). Adicionalmente, Palmero *et al.* verificaram a

prevalência da mutação em um grupo de 750 mulheres não afetadas por câncer, participantes de um programa de rastreamento mamográfico no Rio Grande do Sul. Os achados do estudo demonstram que a mutação foi encontrada em dois indivíduos assintomáticos cujas histórias familiares não preenchiam critérios de SLF/LFL. Apesar do pequeno tamanho amostral, a proporção de portadores (~0,3%), era a mesma daquela observada em recém-nascidos participantes do programa de triagem neonatal do Estado do Paraná (Palmero *et al.*, 2008)(Palmero *et al.*, 2008; Piovezan, 2006).

Em recente estudo de prevalência da mutação p.R337H no Brasil Giacomazzi *et al.* identificaram portadores da mutação em diferentes grupos: mulheres com câncer de mama sem história de SLF/LFL com menos de 45 anos e com mais de 45 anos; mulheres com câncer de mama não selecionadas para história familiar de câncer com menos de 45 anos ou mais de 55 anos; mulheres com tumores phyllodes da mama; crianças com carcinoma adrenocortical e de plexo coroide, concluindo novamente que a mutação está presente em pacientes com e sem história de SLF/LFL (Giacomazzi *et al.,* 2013 – dados não publicados). Neste estudo, foi também identificado um indivíduo potencialmente homozigoto para a mutação p.R337H (genótipo GG) cujos pais foram caracterizados como heterozigotos. Embora o genótipo fosse aparentemente homozigoto mutante, os autores não descartaram a possibilidade de uma região com perda de heterozigosidade por dissomia uniparental em 17p13.1 (Giacomazzi *et al.*, 2013).

Em busca de um melhor entendimento do processo carcinogênico associado a mutações germinativas de *TP53* e possíveis marcadores moleculares relacionados especificamente ao ACC infantil, Letouze *et al.* avaliaram 25 pacientes com ACC (13 deles portadores da mutação *TP53* p.R337H) e identificaram alterações recorrentes em 4q34, 9q33 e 19p além de perda de heterozigosidose nas regiões 17p13 e 11p15 e concluíram que outras vias celulares, além da via de *TP53,* também contribuem para o desenvolvimento de ACC infantil (Letouze *et al.*, 2012). Em pacientes com SLF/LFL, Silva *et al.* demonstraram uma grande variação de número de cópias (CNVs) em portadores de mutações somáticas localizadas na região gênica

correspondente ao domínio de ligação ao DNA, DBD (n=9) quando comparados aos portadores da mutação p.R337H (n=12) e também quando comparados a sujeitos controles. Os autores concluíram que o reduzido número de raras CNVs em portadores da mutação p.R337H pode ter contribuído para a manutenção da mutação na população, acarretando em sua elevada frequência, sendo hoje 300 vezes maior do que mutações somáticas localizadas no DBD do gene *TP53* (Silva *et al.*, 2012).

Até o momento, diferentes estudos demonstram em diferentes coortes que a mutação germinativa *TP53* p.R337H ocorre em indivíduos com vários tipos de câncer (câncer de mama, ACC, phyllodes e plexo coróide, entre outros), porém, nenhum estudo utilizou plataformas de análise genômica na linhagem constitutiva tanto em pacientes com carcinoma adrenocortical como com câncer de mama para identificação de alterações genéticas associadas à mutação p.R337H. Portadores da mutação p.R337H apresentam diferentes desfechos (câncer na infância, câncer na vida adulta, um ou múltiplos tumores, nenhum tumor), sendo desta forma, um grande desafio clínico realizar uma previsão do fenótipo de portadores da mutação p.R337H. Portanto, nossa abordagem utilizando uma ferramenta de análise genômica é muito importante para identificação de novos modificadores que podem estar associados à grande variabilidade fenotípica encontrada. Nesse sentido a nossa hipótese é que outros fatores genéticos ainda não identificados possam agir como modificadores da penetrância e do fenótipo associados com p.R337H tanto em carcinoma adrenocortical como em câncer de mama, e uma análise genômica poderia servir como ferramenta inicial de rastreamento para posterior investigação de alterações potencialmente relacionadas e modificadoras do fenótipo de câncer.

1.4 *TP53* E FERTILIDADE

1.4.1 FUNÇÃO E REGULAÇÃO DE p53 RELACIONADAS À FERTILIDADE

Embora a riqueza de publicações acerca das vias de sinalização de p53 e sua relação com a função de gene supressor tumoral seja muito grande, poucos estudos versam sobre as funções da proteína em um ambiente celular normal. O gene *TP53* é conservado desde invertebrados até vertebrados, e ortólogos do gene humano *TP53* já foram identificados em *Caenorhabditis elegans, Danio rerio, Drosophila melanogaster* e *Xenopus laevis* (Cheng *et al.*, 1997; Jin *et al.*, 2000; Soussi *et al.*, 1987). A existência de p53 em animais de vida curta sem ocorrência de câncer (*Drosophila melanogaster*) sugere que a atuação da proteína p53 como supressora tumoral não seja a sua função original, e certamente não a única. Diversos estudos demonstrando que alelos de *TP53* e de seus reguladores negativos *MDM2 e MDM4* estão sofrendo processo de seleção positiva (Atwal *et al.*, 2007; Murphy, 2006) e o fato de cruzamentos de camundongos $p53^{+/+}$, $p53^{+/-}$ e $p53^{-/-}$ indicarem um padrão não usual de efeitos maternais na taxa de gravidez e de tamanho de prole corroboram a hipótese de uma função de p53 no processo reprodutivo.

Nos últimos anos, uma ação anteriormente desconhecida de p53 foi descrita: a proteína possui importante papel no processo gestacional através da regulação do fator inibidor de leucemia (*LIF, leukemia inhibitory factor*) (Hu *et al.*, 2007b). LIF, uma proteína codificada pelo gene *LIF* (OMIM 159540), localizado no cromossomo 22q12, é uma das citocinas mais importantes no processo de implantação e o aumento de sua expressão coincide com o momento de implantação do blastocisto em mamíferos (Chen *et al.*, 2000). O processo de implantação é um estágio crítico no desenvolvimento embrionário quando o blastocisto estabelece uma grande interação com tecidos uterinos, fato que leva à formação da placenta que sustentará o crescimento e o desenvolvimento do feto. A falha de implantação é a causa mais frequente de

perda gestacional após transferência embrionária e fertilização *in vitro* (FIV) (Kay *et al.*, 2006). No útero, níveis normais da proteína LIF são essenciais para o processo de implantação e baixos níveis da mesma foram demonstrados em mulheres inférteis (Hambartsoumian, 1998). Níveis reduzidos de LIF no tecido endometrial estão associados com diminuição da taxa de implantação em humanos, e variações genéticas no próprio gene assim como alterações nos níveis de p53, resultantes de variações polimórficas de *TP53* podem influenciar a expressão de LIF (Mikolajczyk *et al.*, 2003). Em acordo com esses resultados, um polimorfismo (SNP) localizado na região 3' não traduzida do gene *LIF* (rs929271, T→G, coordenada genômica GRCCh37.p5 chr22:30638226) foi recentemente associado à infertilidade em mulheres com menos de 35 anos (Kang *et al.*, 2009). Além de ser regulado pela expressão de p53, o gene *LIF* também possui regulação hormonal através do hormônio estrógeno, cujo efeito é mediado por seu receptor nuclear alfa (*ESR1*, OMIM 133430). Feng *et al.* identificaram um aumento significativo da proteína ESR1 nas glândulas do endométrio de camundongos no momento da implantação e concluíram que a expressão de LIF no momento da implantação do blastocisto (quarto dia após a fecundação em camundongos e décimo segundo dia em humanos) depende da ativação de p53, de elevados níveis de estrógeno, e da ativação de *ESR1* (Feng *et al.*, 2011).

Considerando a estrita regulação de LIF por p53 e ESR1, a modulação da atividade da mesma por polimorfismos nestes dois genes e em outros genes da via de sinalização de p53 poderia afetar o processo de implantação embrionária (Figura 5) (Bond *et al.*, 2004; Murphy, 2006).

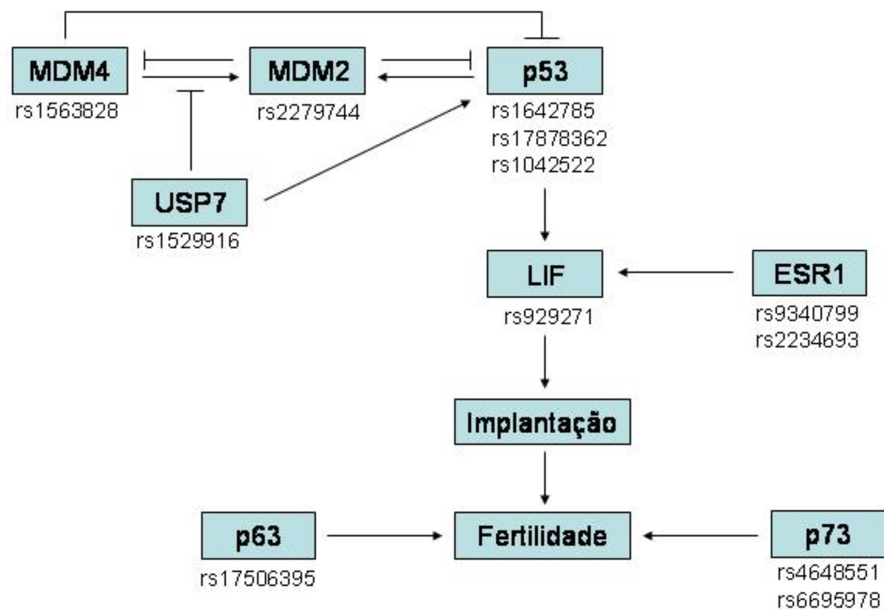


Figura 5 - Polimorfismos em genes da via de sinalização de p53 escolhidos para o estudo. Variantes de sequência nos genes *TP53, MDM2*, *MDM4*, *USP7*, *LIF*, *ESR1*, *TP63* e *TP73* que podem potencialmente modular a fertilidade em humanos (Modificado de Kang *et al.*, 2009).

1.4.2 FAMÍLIA p53 E FERTILIDADE

A família p53 inclui dois homólogos do *TP53*: genes *TP63* (OMIM 603273) e *TP73* (OMIM 601990). Estes três membros da família p53 compartilham de grande homologia, sendo os genes *TP63* e *TP73* mais relacionados entre si do que com *TP53*. Enquanto *TP53* exerce suas funções de fator de transcrição em resposta a diversos estímulos de estresse (em relação à fertilidade atua estimulando LIF), p63 atua como fator de transcrição de uma série de genes relacionados com desenvolvimento epitelial (essencial para o desenvolvimento dos membros) e p73 atua no desenvolvimento de sistema nervoso central e imune (Belyi *et al.*, 2010).

As funções dos genes da família p53 foram descritas a partir de experimentos em camundongos *knockout*. Embora camundongos $p53^{-/-}$ sejam viáveis e apresentem apenas pequenas anormalidades de desenvolvimento, praticamente todos desenvolvem câncer antes dos seis meses de vida,

principalmente sarcomas e linfomas (Donehower *et al.*, 1992). Estudos com camundongos p63$^{-/-}$ demonstraram que a deleção do gene *TP63* gera defeitos no desenvolvimento epitelial e ausência dos membros, dentes e glândulas mamárias (Gonfloni *et al.*, 2009; Mills *et al.*, 1999). Camundongos p73$^{-/-}$ apresentam defeitos de desenvolvimento como hidrocefalia e disgenesia hipocampal, dificuldade de controle de infecções e anormalidades nos receptores de ferormônio, levando a uma falta de interesse por animais do sexo oposto (Yang *et al.*, 2000). Esses resultados sugerem que os defeitos nas vias sensoriais e hormonais podem contribuir para ausência de comportamento sexual em camundongos p73$^{-/-}$. Além do efeito fenotípico da ausência dos genes *TP63* e *TP73*, estudos com camundongos *knockout* demonstram o envolvimento dos genes no processo gestacional em mamíferos em dois estágios não associados com o processo de implantação. O gene *TP63* atua no desenvolvimento, formação, maturação e apoptose de ovócitos (Kurita *et al.*, 2005; Suh *et al.*, 2006) e o gene *TP73* tem papel importante na regulação da divisão celular modulando a formação do fuso mitótico e a correta separação de cromátides irmãs durante a anáfase (Tomasini *et al.*, 2009; Tomasini *et al.*, 2008).

1.4.3 POLIMORFISMOS NOS GENES *TP53, MDM2, MDM4, USP7, LIF, TP63, TP73 E ESR1*

O gene *TP53* é altamente polimórfico e possui 85 SNPs identificados pelo banco de dados da Agência Internacional para Pesquisa em Câncer (IARC - *International Agency for Research on Cancer*) (Petitjean *et al.*, 2007). Três polimorfismos no gene *TP53* (rs1642785, rs17878362 e rs1042522), assim como os SNPs rs2279744 do gene *MDM2* e rs929271 do gene *LIF* têm sido estudados e relacionados com diminuição da função das vias de sinalização de p53, com repercussões clínicas relacionadas ao papel de p53 no desenvolvimento do câncer (associação de diversos SNPs com idade ao diagnóstico de câncer em portadores de mutações germinativas de *TP53*).

O polimorfismo rs1042522 do gene *TP53* (também conhecido como Arg72Pro, códon 72, arg72pro, P72 ou GP, coordenada genômica GRCCh37.p5 chr17:7579472), é um SNP no éxon 3 que resulta em uma alteração de aminoácido (C→G; prolina→arginina; P→R) comprovadamente associada a efeitos fisiológicos (Matlashewski *et al.*, 1987; Thomas *et al.*, 1999). Embora o alelo mutante G resulte em uma proteína com maior capacidade de indução da apoptose, o alelo selvagem C resulta em uma proteína que induz com maior eficiência a senescência e interrompe com maior eficácia o ciclo celular (Dumont *et al.*, 2003; Pim *et al.*, 2004). Kay *et al.* demonstraram que pacientes que apresentaram falha recorrente de implantação em uma clínica de FIV apresentavam, com maior frequência, o alelo C (Kay *et al.*, 2006). Corroborando com este achado, Pietrowski *et al.* e Firouzabadi *et al.* evidenciaram que o alelo C estaria associado a perdas gestacionais recorrentes (Firouzabadi *et al.*, 2009; Pietrowski *et al.*, 2005).

Os SNPs intragênicos de *TP53* rs1642785 (*TP53* PIN2, G→C, coordenada genômica GRCCh37.p5 chr17:7579801) e rs17878362 (*TP53* PIN3, duplicação de 16 pares de base no intron 3; alelo N: não-duplicado e alelo D: duplicado, coordenada genômica GRCCh37.p5 chr17:7579644) têm sido recentemente estudados em pacientes com câncer e famílias com a síndrome de Li-Fraumeni e suas variantes, que demonstram um efeito clínico de antecipação na idade ao diagnóstico de diversos tumores na presença de determinados alelos (Marcel *et al.*, 2010; Whibley *et al.*, 2009). No entanto, não há relatos acerca de uma possível influência em relação à fertilidade, mais especificamente, às consequências relativas à implantação de embrião. Embora a exata relação destes polimorfismos com o efeito fenotípico não esteja ainda totalmente elucidada, Gemignani *et al.* demonstraram que os mesmos têm influência sobre os níveis de mRNA da proteína p53 em células linfoblastóides (Gemignani *et al.*, 2004). Corroborando esta hipótese, Marcel *et al.* demonstraram que o intron 3 do gene *TP53* está envolvido no processamento (*splicing*) do intron 2, regulando a produção de diferentes isoformas de p53 (Marcel *et al.*, 2011).

O polimorfismo *MDM2* rs2279744 (substituição T→G, coordenada genômica GRCCh37.p5 chr12:69202580, região promotora) do gene *MDM2* é um polimorfismo funcional que aumenta os níveis da proteína MDM2 atenuando a via de sinalização de p53 mediante interferência com o sítio de ligação do fator de transcrição Sp1 (Bond *et al.*, 2004). A atividade da proteína Sp1 é ótima na presença do alelo mutante G do SNP MDM2 rs2279744, pois garante um aumento da expressão do gene, atenuação da via de p53, e consequente aumento do risco para tumorigênese (Bond *et al.*, 2004; Park *et al.*, 2006). Para verificar uma possível associação dos genes da via de sinalização de p53 e infertilidade, Kang *et al.* realizaram um estudo caso-controle comparando pacientes de uma clínica de fertilização *in vitro* (FIV) com 100 mulheres férteis. Estudando SNPs nos genes *TP53, LIF, MDM2, MDM4 e USP7*, eles demonstraram que o alelo *TP53* rs1042522 C era mais frequente nas pacientes submetidas à FIV, e consideraram o SNP rs1042522 como fator de risco para falhas de implantação embrionária. Além disso, demonstraram que os níveis de LIF encontravam-se significativamente mais baixos na presença do alelo *TP53* rs1042522 C, fato que pode contribuir para a diminuição da implantação e infertilidade (Kang *et al.*, 2009). O mesmo grupo de pesquisadores recentemente demonstrou que alelos de polimorfismos nos genes *TP63* (rs17506395, T→G, coordenada genômica GRCCh37.p5 chr3:189521319) e *TP73* (rs4648551, G→A, coordenada genômica GRCCh37.p5 chr1:3632730 e rs6695978, G→A, coordenada genômica GRCCh37.p5 chr1:3648345) são mais frequentes em mulheres inférteis com mais de 35 anos (Feng *et al.*, 2011).

Poucos estudos investigaram o papel de polimorfismos no gene *ESR1* em infertilidade em mulheres. Lamp *et al.* demonstraram que o SNP rs2234693 (C→T, coordenada genômica GRCCh37.p5 chr6:152163335) no gene *ESR1* pode estar relacionado a infertilidade e endometriose, embora não tenham encontrado associação estatisticamente significativa entre os grupos estudados (Lamp *et al.*, 2011).

1.4.4 INFERTILIDADE, ENDOMETRIOSE E POLIMORFISMOS EM GENES DA VIA DE SINALIZAÇÃO DE P53

A endometriose, definida como presença de tecido glandular e estroma endometrial fora da cavidade uterina, foi descrita pela primeira vez em 1860 pelo patologista austríaco Karl Freiherr von Rokitansky e é considerada uma importante causa de infertilidade em mulheres, pela sua morbidade e alta frequência. A etiopatogenia da endometriose ainda não está totalmente elucidada podendo estar relacionada a metaplasia celômica, implantes ectópicos e indução de células multipotenciais (Halis *et al.*, 2004). Recentemente, foram descritos diversos SNPs que poderiam relacionar-se com endometriose (Tempfer *et al.*, 2009), entretanto essa associação foi refutada. Existe especial interesse no estudo dos sistemas imunológico e endócrino e em estudos que envolvem SNPs relacionados com transformação maligna e apoptose. Como há clara evidência de risco maior para endometriose em mulheres que tenham irmãs ou mãe com a doença, o componente genético é importante para seu desenvolvimento (Vigano *et al.*, 2007).

Estima-se que 10% a 15% da população feminina apresentem endometriose e os sintomas mais comuns desta condição são infertilidade (30-60% dos casos) e dor pélvica (Olive *et al.*, 1993). Embora a associação entre endometriose e infertilidade seja bem estabelecida (mulheres com endometriose têm 20 vezes mais chance de serem inférteis), a causa exata da infertilidade em portadoras de endometriose não é conhecida (Koninckx, 1994). Podem estar implicadas alterações hormonais (anormalidades na secreção de progesterona, estradiol, prolactina e na função lútea, como insuficiência), ciclos anovulatórios (Cunha-Filho *et al.*, 2001) e foliculogênese anormal (Cunha-Filho *et al.*, 2002). Recentemente, foi demonstrado que mulheres com endometriose mínima e leve apresentam uma alteração da coorte folicular com uma diminuição da reserva ovariana medida pelos níveis séricos do hormônio anti-Mulleriano no 3º dia do ciclo menstrual (Lemos *et al.*, 2008). Estudos sobre polimorfismos genéticos e endometriose têm sido realizados no sentido de melhor entender a etiopatogenia da doença e também na busca de biomarcadores de prognóstico e resposta a diferentes estratégias de tratamento. Esses marcadores poderiam

explicar dois aspectos fundamentais em relação a essa doença: i) a proliferação e crescimento extra-endometrial de tecido e ii) a sub-fertilidade que essas mulheres possuem, mesmo quando submetidas à fertilização *in vitro*, por exemplo.

Em relação a polimorfismos de *TP53,* o estudo do SNP rs1042522 em vários grupos populacionais teve resultados controversos até o momento. Em 2006, Hsieh e Lin descreveram uma associação negativa do genótipo G/G com o diagnóstico de endometriose em mulheres chinesas (genótipo protetor) (Hsieh *et al.*, 2006), entretanto, este efeito não foi observado em mulheres italianas ou brasileiras (Ammendola *et al.*, 2008; Ribeiro Junior *et al.*, 2009; Vietri *et al.*, 2007). Em relação a *MDM2*, apenas um estudo avaliou a expressão deste gene (bem como de *TP53*) em endometriomas, sendo observada uma expressão anormal em relação ao tecido endometrial normal, sugerindo assim um papel das proteínas codificadas por estes genes na regulação do crescimento dos endometriomas (Goumenou *et al.*, 2005). Por fim, em relação a polimorfismos nos genes *MDM4*, *USP7*, *LIF*, *TP63* e *TP73* não há relatos específicos em pacientes com endometriose. No entanto, a observação por vários autores de expressão significativamente reduzida da proteína LIF em mulheres inférteis com endometriose em relação a controles sem endometriose, especialmente no período de implantação embrionária, sugere que LIF (ou seus reguladores) tenha(m) um papel importante na infertilidade associada à endometriose. Os níveis de LIF na cavidade endometrial de mulheres com endometriose e infertilidade poderiam ser utilizados como um biomarcador preditivo do sucesso reprodutivo em procedimentos de FIV (Dimitriadis *et al.*, 2005; Mikolajczyk *et al.*, 2003; Novotny *et al.*, 2009). Considerando evidências previamente publicadas, em especial os resultados de Hu *et al.* (2007), Hu *et al.* (2008), Kang *et al.* (2009) e Firouzabadi *et al.* (2009), hipotetizamos que mulheres inférteis, com ou sem endometriose, apresentam frequências alélicas e genotípicas de polimorfismos nos genes *TP53*, *MDM2, MDM4, USP7, LIF*, *TP63, TP73* e *ESR1* distintas de mulheres férteis (Firouzabadi *et al.*, 2009; Hu *et al.*, 2008; Hu *et al.*, 2007a; Kang *et al.*, 2009).

2. OBJETIVOS

2.1 OBJETIVO GERAL

O objetivo geral deste estudo é a identificação de alterações moleculares associadas à via de sinalização do gene *TP53* em câncer e infertilidade. Os objetivos específicos são apresentados separadamente, para cada uma destas duas abordagens empregadas na tese.

2.2 OBJETIVOS ESPECÍFICOS

2.2.1 Identificação de alterações moleculares associadas com a via e sinalização do gene *TP53* e fertilidade

2.2.1.1 Determinar as frequências alélicas e genotípicas dos polimorfismos PIN2 (rs1642785), PIN3 (rs17878362) e PEX4 (rs1042522) do gene *TP53* em mulheres inférteis com falha recorrente de fertilização *in vitro* e com endometriose, comparando-os com os de mulheres férteis da mesma região geográfica de recrutamento.

2.2.1.2 Calcular o desequilíbrio de ligação entre os três polimorfismos do gene *TP53* estudados e comparar os haplótipos encontrados com os desfechos de infertilidade e endometriose.

2.2.1.3 Avaliar polimorfismos em outros genes da via de sinalização de *TP53* (genes *MDM2, MDM4, USP7, LIF*, *TP63, TP73* e *ESR1)* em mulheres inférteis com falha recorrente de fertilização *in vitro* e endometriose comparando-os com os de mulheres férteis da mesma região geográfica de recrutamento.

2.2.1.4 Construir uma árvore de proteínas relacionadas à via de sinalização de *TP53* com desfechos de infertilidade utilizando análises *in silico*.

2.2.2 Identificação de alterações moleculares associadas com a via e sinalização do gene *TP53* e câncer em pacientes portadores da mutação *TP53* p.R337H

2.2.2.1 Identificar e validar regiões cromossômicas com variação do número de cópias em dois pacientes com câncer de mama portadores da mutação *TP53* p.R337H.

2.2.2.2 Identificar e validar regiões cromossômicas com variação do número de cópias em seis pacientes com carcinoma adrenocortical, sendo quatro portadores da mutação *TP53* p.R337H.

3. CAPÍTULO 1

3.1 *TP53 PIN3 AND PEX4 POLYMORPHISMS AND INFERTILITY ASSOCIATED WITH ENDOMETRIOSIS OR POST-IN VITRO FERTILIZATION FAILURE*

Título do manuscrito: *TP53 PIN3 and PEX4 polymorphisms and infertility associated with endometriosis or post-in vitro fertilization failure*

Autores: Paskulin DD, Cunha-Filho JS, Souza CA, Bortolini MC, Hainaut P, Ashton-Prolla P

Situação: Publicado

Revista: *Cell Death and Disease*

Referência: *Cell Death Dis*. 2012 Sep 27;3:e392. doi: 10.1038/cddis.2012.116

Website: http://www.ncbi.nlm.nih.gov/pubmed/23013791

Citation: Cell Death and Disease (2012) 3, e392; doi:10.1038/cddis.2012.116
© 2012 Macmillan Publishers Limited All rights reserved 2041-4889/12

www.nature.com/cddis

TP53 PIN3 and PEX4 polymorphisms and infertility associated with endometriosis or with post-*in vitro* fertilization implantation failure

DD Paskulin[1,3,6], JSL Cunha-Filho[2], CAB Souza[2], MC Bortolini[1,3], P Hainaut[4] and P Ashton-Prolla*[,1,3,5,6]

p53 has a crucial role in human fertility by regulating the expression of leukemia inhibitory factor (LIF), a secreted cytokine critical for blastocyst implantation. To examine whether *TP53* polymorphisms may be involved with *in vitro* fertilization (IVF) failure and endometriosis (END), we have assessed the associations between *TP53* polymorphism in intron 2 (PIN2; G/C, intron 2), PIN3 (one (N, non-duplicated) or two (D, duplicated) repeats of a 16-bp motif, intron 3) and polymorphism in exon 4 (PEX4; C/G, p.P72R, exon 4) in 98 women with END and 115 women with post-IVF failure. In addition, 134 fertile women and 300 women unselected with respect to fertility-related features were assessed. *TP53* polymorphisms and haplotypes were identified by amplification refractory mutation system polymerase chain reaction. *TP53* PIN3 and PEX4 were associated with both END ($P=0.042$ and $P=0.007$, respectively) and IVF ($P=0.004$ and $P=0.009$, respectively) when compared with women both selected and unselected for fertility-related features. Haplotypes D-C and N-C were related to higher risk for END ($P=0.002$, $P=0.001$, respectively) and failure of IVF ($P=0.018$ and $P=0.002$, respectively) when compared with the Fertile group. These results support that specific *TP53* haplotypes are associated with an increased risk of END-associated infertility and with post-IVF failure.
Cell Death and Disease (2012) **3,** e392; doi:10.1038/cddis.2012.116; published online 27 September 2012
Subject Category: Cancer

TP53 encodes the multi-functional tumor suppressor transcription factor p53 which has a crucial role in maintaining genomic stability in somatic cells exposed to oncogenic or genotoxic stress, thus preventing tumor formation.[1] In response to a wide range of stress signals, p53 accumulates in the nucleus and regulates the expression of a large panel of genes involved in the control of cell cycle arrest, apoptosis, cell senescence, DNA repair and energy metabolism. One of the transcriptional targets of p53 is leukemia inhibitory factor (*LIF*), the gene encoding LIF. LIF is a secreted cytokine with broad roles in the control of lymphocyte proliferation and differentiation. It has also been identified as a critical factor for blastocyst implantation.[2] Control of p53 over LIF expression is operated through a p53-response element located in intron 1 and conserved in both mouse and human *LIF* genes.[3]

Recent studies have demonstrated that p53 regulates female reproduction and blastocyst implantation through LIF. Implantation is a critical step in mammalian embryonic development during which the blastocyst establishes close interactions with the uterus, leading to the formation of the placenta supporting fetal development.[4] Hu *et al.*[4] have demonstrated that p53 regulates LIF expression in the uterus of female mice. p53-deficient mice express lower levels of LIF than their p53-competent counterparts and show impaired blastocyst implantation and consequently, impaired fertility.

There is strong evidence that genes at critical regulatory nodes in the p53 pathway are under evolutionary selection[5,6] and that SNPs in the p53 pathway influence human fertility.[7] Of these, one of the most studied is *TP53* polymorphism in exon 4 (PEX4 of the *TP53* gene), widely known as p.P72R (C/G, rs1042522). This single-nucleotide polymorphism (SNP) located at the second position of the codon 72 consist in either an ancestral C allele whose frequency in African populations is around 0.70 or a derived G allele whose frequency in European and Asian populations varies from around 0.50 to 0.80. Presence of the C allele results in a proline in codon 72, and presence of the G allele, in an arginine. These polymorphic protein variants significantly differ in their biological properties and there is evidence that R72p53 has higher transcriptional activity toward a particular subset of p53 target genes, including *LIF*, than P72p53.[8] Previous studies have identified an association between *TP53* PEX4 and infertility[7] or endometriosis (END).[9–11] It has been suggested that the effect of PEX4 on *LIF* expression and fertility may account for population differences in the distribution of PEX4 alleles in different parts of the world.

[1]Department of Genetics and Molecular Biology, Federal University of Rio Grande do Sul, Porto Alegre, Brazil; [2]Department of Obstetrics and Gynecology, Federal University of Rio Grande do Sul, Porto Alegre, Brazil; [3]Post-Graduate Program in Genetics and Molecular Biology, Federal University of Rio Grande do Sul, Porto Alegre, Brazil; [4]International Prevention Research Institute, Lyon, France; [5]Medical Genetics Service, Hospital de Clínicas de Porto Alegre, and National Institute of Science and Technology in Populational Medical Genetics (INAGEMP), Porto Alegre, Brazil and [6]Genomic Medicine Laboratory, Experimental Research Center, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil
*Corresponding author: P Ashton-Prolla, Departamento de Genética, UFRGS e Serviço de Genética Médica e Centro de Pesquisa Experimental, Hospital de Clínicas de Porto Alegre, Rua Ramiro Barcelos, 2350, Porto Alegre 90035-903, Brazil. Tel: +55 51 3359 8010; Fax: +55 51 3359 8760; E-mail: pprolla@hcpa.ufrgs.br
Keywords: *TP53*; single-nucleotide polymorphisms; IVF; endometriosis
Abbreviations: TP53, tumor protein p53; LIF, leukemia inhibitory factor; PIN2, polymorphism in intron 2; PIN3, polymorphism in intron 3; PEX4, polymorphism in exon 4; SNP, single-nucleotide polymorphism; END, endometriosis; IVF, *in vitro* fertilization; D, duplicated allele; N, non-duplicated allele
Received 30.4.12; revised 21.6.12; accepted 02.7.12; Edited by Y Shi

These differences may reflect subtle adaptation to environmental constraints affecting fertility. However, the magnitude of the PEX4 effect on infertility associated with different pathological causes remains controversial.[12,13]

PEX4 is in strong linkage disequilibrium with another common polymorphism located in its close vicinity, PIN2 (polymorphism in intron 2; rs1642785; G/C). The PIN2 G allele has been associated with human papillomavirus persistence[14] and individuals with two copies of the PIN2 G allele have been reported as having an increased risk of osteosarcoma.[15] Recently, it has been shown that another polymorphism in intron 3 of the *TP53* gene, PIN3 (Polymorphism in Intron 3, rs17878362, 16 bp duplication, N = non-duplicated, D = duplicated) overlaps with a G-quadruplex motif, which regulates p53 mRNA splicing generating an alternatively spliced form, which supports the synthesis of an isoform of p53 lacking the N-terminal transactivation domain (Delta40p53).[16] PIN3 D allele is associated with increased risk of colorectal,[17] lung[18] and breast cancer,[19] whereas the N allele has been reported in association with an average acceleration of 19 years in the mean age at first cancer diagnosis in a Brazilian cohort of *TP53* germline mutation carriers.[20] The effects of this polymorphism in END or infertility have not been investigated so far.

Although the association between END and infertility is well known (END affects up to 50% of women with infertility),[21] the cause of infertility in the disease is not fully understood but is thought to involve hormonal,[22] immunological,[23] genetic,[24] proliferative (endometrial) and uterine alterations.[25] We hypothesized that *TP53* polymorphisms that alter p53 function may be associated with *in vitro* fertilization (IVF) failure and with END-associated infertility.

Results

Patients and healthy study subjects did not differ significantly regarding self-attributed skin color (Supplementary Table S1). Overall, a self-denomination of 'white' color predominated in all study subgroups (END, FIV, Unselected and Fertile). In terms of reproductive history, the mean number of pregnancies in women of the fertile and unselected for fertility groups was 3.62 ± 1.9 and 3.22 ± 2.1, respectively. In the later, nulliparity was observed in 2.6%.

Women in the fertile and unselected for fertility groups presented higher mean age at recruitment (42.68 ± 12.8 years and 43.2 ± 12.7 years, respectively) as compared with END (32.87 ± 4.7 years) and IVF (31.65 ± 3.2 years) groups. Hardy–Weinberg equilibrium was achieved in all study groups for PIN2, PIN3 and PEX4 (all $P > 0.05$, Supplemental Materials, Table S2. Genotypic and allelic frequencies of the *TP53* polymorphisms are shown in Table 1. In all four study subgroups, PIN3 and PEX4 allele frequencies did not differ significantly from those previously described in European populations (Supplementary Table S3).

Single marker analysis (Table 1) revealed a significant association between PIN2 (rs1642785) genotypes and IVF ($P = 0.016$), and a borderline association with the END group ($P = 0.052$) when compared with the Fertile group. There was an increased frequency of the PIN2 C allele in both the END and IVF groups. When analyzing *TP53* PIN3 (rs17878362) polymorphism, a clear difference between IVF and END groups was observed when compared with the Fertile group. Allele D (the duplicated allele) was enriched in patients in both groups as compared with Fertile ($P = 0.042$ and $P < 0.0004$ for the END and IVF groups, respectively). For *TP53* PEX4 (rs1042522), a statistically significant difference between both the END and IVF groups and the Fertile group was also demonstrated, with enrichment of the PEX4 C allele in both groups ($P = 0.007$ and $P = 0.009$, respectively).

When the Fertile and Unselected groups were compared, we observed that the allelic frequencies of PIN2 G and PEX4 G were significantly higher in the Fertile group, whereas PIN2 and PEX4 genotype distribution did not differ between groups.

Table 1 Genotypic and allelic frequencies of selected *TP53* polymorphisms between Fertility Unselected, Fertile, END and IVF groups

TP53		Unselected, *n* (%)	Fertile, *n* (%)	*P*-value[a]	END, *n* (%)	*P*-value[b]	*P*-value[c]	IVF, *n* (%)	*P*-value[d]	*P*-value[e]
PIN2	GG	166 (55.3)	88 (65.7)	0.114	53 (54.1)	0.304	0.052	63 (54.8)	0.049	0.016
rs1642785	GC	112 (37.3)	40 (29.9)		33 (33.7)			35 (30.4)		
	CC	22 (7.3)	6 (4.5)		12 (12.2)			17 (14.8)		
	G	444 (74.0)	216 (80.6)	0.007	139 (70.9)	0.397	0.015	161 (70.0)	0.245	0.005
	C	156 (26.0)	52 (19.4)		57 (29.1)			69 (30.0)		
PIN3	NN	222 (74.0)	94 (70.1)	0.658	49 (50.0)	<0.001	0.042	72 (62.6)	<0.001	0.004
rs17878362	ND	70 (23.3)	35 (26.1)		32 (32.7)			29 (25.2)		
	DD	8 (2.7)	5 (3.7)		17 (17.3)			14 (12.2)		
	N	514 (85.7)	223 (83.2)	0.350	130 (66.3)	<0.001	0.005	173 (75.2)	<0.001	0.027
	D	86 (14.3)	45 (16.8)		66 (33.7)			57 (24.8)		
PEX4	GG	158 (52.7)	89 (66.4)	0.013	50 (51.0)	0.535	0.007	63 (54.8)	0.159	0.009
rs1042522	GC	114 (38.0)	40 (29.9)		35 (35.7)			35 (30.4)		
	CC	28 (9.3)	5 (3.7)		13 (13.3)			17 (14.8)		
	G	430 (71.7)	218 (81.3)	0.002	135 (68.9)	0.455	0.001	161 (0.70)	0.635	0.003
	C	170 (28.3)	50 (18.7)		61 (31.1)			69 (0.30)		

Abbreviations: D, duplicated; END, endometriosis; IVF, *in vitro* fertilization; N, non-duplicated.
[a] χ^2-test, significant difference observed between women unselected for fertility and Fertile women.
[b] χ^2-test, significant difference observed between END patients and women unselected for fertility.
[c] χ^2-test, significant difference observed between END patients and Fertile group.
[d] χ^2-test, significant difference observed between IVF patients and women unselected for fertility.
[e] χ^2-test, significant difference observed between IVF patients and Fertile group.
IVF group: women with recurrent failure of IVF; END group: infertile women with minimal or mild endometriosis; Fertile: Fertile women; Unselected: women unselected with respect to fertility or infertility-related symptoms.

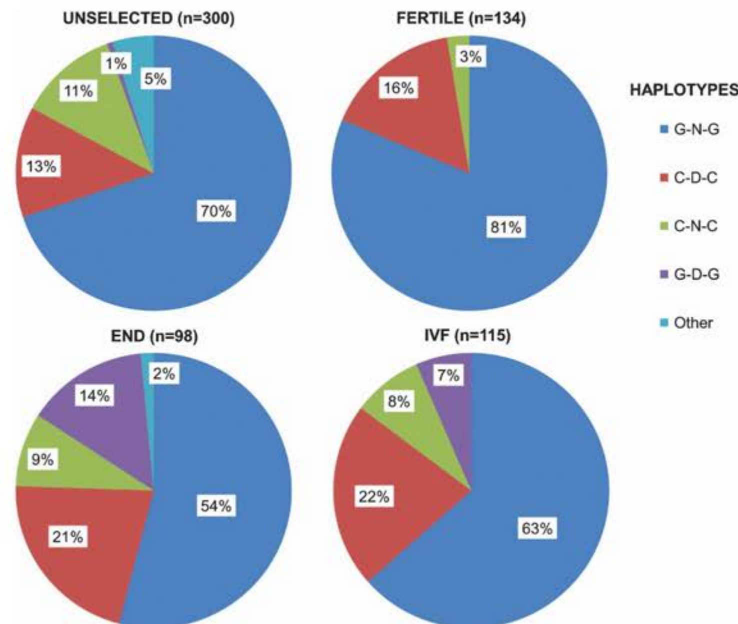


Figure 1 Distribution of the most frequent haplotypes among Unselected, Fertile, END and IVF groups. Haplotypes frequencies are shown as '%'. Haplotypes were constructed as PIN2 (G/C) – PIN3 (N/D) – PEX4 (C/G)

Similarly, PIN3 genotypic or allelic frequencies did not differ between groups (Table 1). For both the END and IVF groups, the allelic frequencies of PIN2, PIN3 and PEX4 differed significantly from those observed in the Fertile group. The allelic frequencies of PIN3 in infertile women (either END and IVF groups) differed significantly from both the Fertile and Unselected groups.

Haplotype analysis showed strong linkage disequilibrium between *TP53* PIN2 and PEX4 ($D' = 1$; $r^2 = 0.94$ in all studies groups, Supplementary Table S4) as previously described.[20] Therefore, we have only considered *TP53* PIN3 and *TP53* PEX4 in further analyses and in our discussion. We carried out a binary logistic regression analysis to evaluate the effect of *TP53* haplotypes with regard to END and IVF. Figure 1 shows the distribution of the most frequent haplotypes encountered (see Figure 1). Table 2 shows the odds ratios for the END and IVF groups of the most frequent haplotypes when compared with the reference N-G haplotype. Haplotypes D-C and N-C were related to higher risk for END ($P = 0.002$, $P = 0.001$, respectively) and failure of IVF ($P = 0.018$ and $P = 0.002$, respectively) when compared with the Fertile group. However, when the Unselected group (unselected for fertility) was used as the comparison group in the logistic regression model, the risk association with haplotypes D-C and N-C was not observed (data not shown).

Discussion

In this study, we have analyzed the distribution of three common polymorphisms in the *TP53* gene (PIN2, PIN3 and PEX4) in infertile women with failure of IVF treatment or with END-associated infertility. Our results demonstrate an association between these two forms of infertility and *TP53* alleles PIN3 D and PEX4 C, suggesting that variations in p53 activity specified by these polymorphisms may be involved in the pathogenesis of both conditions. These results support previously reported observations on associations between PEX4 and infertility, in particular IVF failure. Furthermore, these results provide clear evidence in favor of an association between *TP53* polymorphism and infertility-related END. Regarding *TP53* PIN3, several studies have evaluated the association between this polymorphism and lung[18] or breast cancer,[19] but to our knowledge, no previous study has analyzed its association with infertility or END.

Table 2 Binary logistic regression model for *TP53* haplotypes regarding PIN3 (N/D) and PEX4 (C/G) using the Fertile group as reference

	END		IVF	
Haplotype	*P*-value	OR (95% CI)	*P*-value	OR (95% CI)
N-G				
D-C	0.002	2.1 (1.3–3.5)	0.018	1.7 (1.1–2.7)
N-C	0.001	4.9 (2.1–12.4)	0.002	4.1 (1.6–9.8)

Abbreviations: CI, confidence interval; D, duplicated; END, endometriosis; IVF, *in vitro* fertilization; N, non-duplicated; OR, odds ratio.
OR (95% CI) was calculated by binary logistic regression analysis. IVF group: women with recurrent failure of IVF; END group: infertile women with minimal or mild endometriosis.

Previous studies have shown associations between PEX4 C allele and END,[10,26,27] whereas others fail to demonstrate this association.[11,12] These controversies may be due to the environmental and genetic background of the studied populations but also because of differences in illness classifications (END is sometimes asymptomatic and often can only be diagnosed by laparoscopy). In our study, the END group was carefully diagnosed according to the American Society for Reproductive Medicine (ASRM) and women with END were excluded from both IVF and Fertile groups after laparoscopic examination. In addition, all four study groups described here were quite homogeneous in terms of self-reported skin color (a feature used as proxy for 'race' or ancestry background in Brazil) corroborating with previous population-based studies that demonstrate predominance of European genomes in this specific region.[28–32] This observation was further confirmed by comparative analysis of PIN3 and PEX4 allele frequencies encountered here and those previously described in European/European-derived and African/African-derived populations, showing that in all four study groups, allelic distribution was not statistically different from the observed in Europeans/Europeans-derived.

Kay *et al.*[33] were the first to associate *TP53* PEX4 C allele with women experiencing recurrent implantation failure. Other studies also associated the PEX4 C allele with the occurrence of idiopathic recurrent miscarriages[34] and implantation failure,[35] and Kang *et al.*[7] demonstrated that PEX4 C was significantly enriched among IVF patients, serving as risk factor for implantation failure. Our results are in agreement with these previous findings regarding the *TP53* PEX4 C allele and confirm this allele as a risk factor for both END-associated infertility and IVF failure in a different sample set.

These results and the findings of our study suggest that PIN3 and PEX4 polymorphisms present specific functional differences in p53 protein variants, having an impact on events that are critical for embryo implantation and/or early development. In the case of PEX4, there is experimental evidence from cell and animal studies that the p53 protein encoded by the PEX4 C allele (P72p53) is more efficient in initiating senescence that the product of the PEX4 G allele (R72p53), which in turn appears to have a stronger effect on p53-mediated apoptosis and suppression of cell transformation.[36] In the case of PIN3, presence of the *TP53* PIN3 D allele has been associated with reduced levels of *TP53* mRNA in lymphoblastoid cell lines.[17] Whether this effect also occurs *in vivo* remains to be determined. Marcel *et al.*[16] demonstrated that *TP53* PIN3 is located within a GC-rich region of intron 3 that form G-Quaduplex structures, which modulate splicing of intron 2. *In silico* models predict that PIN3 may alter the topology of these *G-quadruplex* structures, thus modifying the patterns of p53 mRNA isoform expression. The p53 isoform encoded by alternatively spliced p53 retaining intron 2 lacks the N-terminal domain containing the main transactivation activity of p53, thus resulting in an N-terminally truncated protein, which binds DNA but does not activate transcription through p53-response elements. It is important to emphasize that in our study *TP53* PIN3 presented an allelic distribution that was significantly different in infertile women (either END or IVF groups) when compared with women selected for fertility but also when compared with women from a community sample and unselected for fertility, suggesting that genetic variations in PIN3 may have a critical effect on infertility. Further experimental studies are needed to evaluate the possible impact of p53 isoforms in regulating these biological events, especially their impact on transactivation of key genes involved in the early stages of gestation, such as *LIF*. Haplotypes D-C and N-C were related to higher risk for END and IVF only when a group of women selected in favor of normal fertility (the Fertile group) was used as comparison group; this was not observed when the comparison group included women unselected for reproductive history. This observation suggests that specific haplotypes of *TP53* may be associated with high fertility features. Given the associations between specific SNPs and infertility, it is reasonable to assume that particular combinations of SNPs might provide a genetic marker for women with high fertility features.

Our results support that *TP53* polymorphisms have a role in both END-associated infertility and IVF failure; although current evidence points to a strong effect of the PEX4 polymorphism in embryo implantation and fertility, other SNPs in *TP53*, especially PIN3, may have a key role in the modulation of this process and in other biological processes related to early embryonic development. PIN3 was the only SNP that showed differential frequencies in infertile women (either END and IVF groups) when compared with either fertility-selected or unselected groups, whereas PIN2 and PEX4 only showed a differential distribution in END and IVF patients when compared with a group of patients at the other extreme of the phenotype (fertile group).

In conclusion, the data presented here add to the current evidence that variations in expression and activity of p53 may have an effect on the expression of key genes related to the control of cellular growth and invasion, which have been associated with END (*BAX, FAS, PIG11, PTEN*), as well as on genes associated with embryo implantation (*LIF*). Infertility associated to END could be related, at least in part, to embryo implantation failure in a mechanism similar to that seen in other infertile women without END. It may also involve other mechanisms affecting early embryonic development as well as cell–cell communications during the pre-implantation and implantation phases. In agreement with this hypothesis, previous studies have demonstrated lower implantation and pregnancy rates in endometriotic patients.[37] *TP53* polymorphisms, especially PIN3 and PEX4 may have an interest as biomarkers and could add to the development of a clinically relevant genetic profile that would be of great help for clinicians to identify patients at higher risk for IVF failure. The results of this study should be confirmed in larger cohorts with well defined phenotypes of END and infertility and long-term follow-up data. They also emphasize the importance of a clear definition of clinical phenotypes and of study design when analyzing the effects of specific polymorphisms on fertility.

Materials and Methods

Patients and subjects. All patients and subjects were informed about the procedures of the study when invited to participate and signed a consent form at inclusion. The research project was approved by the Institutional Ethics Committee (Hospital de Clinicas de Porto Alegre – GPPG 05-182; GPPG 09-430).

At inclusion, patients and subjects were also asked to provide a description of their perceived skin color. In Brazil, skin color is normally used to define an equivalent to 'race' or ancestry background.[32,38] We used the words 'White' and 'non-White'

to identify women who defined themselves with some term that suggests only European ancestry and with other terms that suggest some level of African ancestry (such as *mulato* or *pardo*), respectively. No term that reports some level of Amerindian ancestry was used by volunteers.

Patients and subjects were divided into four study groups. The IVF Group consisted of 115 women (<35 years) with at least one IVF failure, defined as a failure after IVF cycle treatment with transfer of two or more top quality embryos (8 cell embryos with $<20\%$ fragmentation). Briefly, inclusion criteria of this group were: age <35 years, exclusion of END by laparoscopy and the main factor was of mild masculine (oligospermia) or tubal origin. All patients in this group were submitted to conventional IVF. Patients with previous thyroid disease, positive anti-lupus or anti-cardiolipin antibodies and trombophilias were also excluded from our sample. Controlled ovarian hyperstimulation was performed with the use of recombinant human FSH and pituitary suppression with GnRh antagonist (fixed day-6 protocol). Ovulation was induced by 6500 IU recombinant hCG when at least three follicles had reached a diameter of >17 mm, and transvaginal follicle aspiration was performed 36 h later under ultrasound guidance. Embryos were classified according to the cumulative embryo classification, taking into account cleavage speed, blastomere symmetry, extent of fragmentation and the presence or absence of multinucleated blastomeres.

The END group comprised 98 infertile women with minimal or mild END as diagnosed by laparoscopy recruited at the Gynecology Service of Hospital de Clinicas de Porto Alegre (HCPA), in Southern Brazil. Infertility was defined as the inability of a couple to achieve pregnancy after 1 year of regular unprotected sexual intercourse.[39] Other causes of infertility were excluded by hysterosalpingography, sperm evaluation and hormonal measurements whenever necessary. END diagnosed during laparoscopy was categorized according to the classification proposed by the ASRM.[39]

The Fertile group consisted of 134 women with no history of infertility, who already had children without any difficulties or assisted reproduction and underwent laparoscopy for tubal ligation at HCPA. END was excluded in women from IVF and Fertile groups. In addition, we studied a group of 300 asymptomatic women, who volunteered for a community-based breast cancer screening program in Southern Brazil (from the same geographic recruitment area of the patients included in the IVF and Fertile groups). This group ('Unselected') was unselected with respect to fertility or infertility-related symptoms, as described elsewhere.[40]

Genotyping. Genomic DNA was extracted from peripheral blood using the Ilustra blood genomic Prep Mini spin Kit (GE Healthcare, Piscataway, NJ, USA) as described by the manufacturer. Genotypes and haplotypes defined by the three *TP53* gene polymorphisms (PIN2 rs1642785 G/C, PIN3 rs17878362 16 pb duplication and PEX4 rs1042522 C/G) were determined by Amplification Refractory Mutation System as previously described.[20]

Statistical analysis. The clinical characteristics of the women in all study groups were compared by one-way analysis of variance. Differences in genotype/allele distribution between IVF, END, Fertile and Unselected groups were evaluated using χ^2-analysis, also used to test for Hardy–Weinberg equilibrium. Linkage disequilibrium was assessed calculating D' value (the relative magnitude of D as compared with its theoretical maximum, calculated as D/D^{max}) as described by Lewontin.[41]

Binary logistic regression analysis was carried out to estimate the odds ratios with 95% confidence intervals in order to assess the influence of *TP53* haplotypes for END and IVF using the Fertile group as reference. Haplotype frequencies were calculated by direct count. Statistical analysis was performed using the SPSS 18.0 statistical package. All reported *P*-values are two-tailed and considered statistically significant when $0.05 \geqslant$.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements. The work of DDP was supported by fellowships from CAPES and CNPQ (Brazil). The study was supported in part by grants from GlaxoSmithKline Oncology (Ethnic Research Initiative Grant Award 2009), UK; CNPq to PA-P (Grant 307779 2009-2), Brazil; FAPERGS-PPSUS (Grant no. 09/0103-0), FAPERGS PRONEX (Grant no. 10/0051-9) and Fundo de Incentivo a Pesquisa e Eventos, Hospital de Clínicas de Porto Alegre (GPPG no. 09430), Brazil.

1. Levine AJ, Hu W, Feng Z. The P53 pathway: what questions remain to be explored? *Cell Death Differ* 2006; **13**: 1027–1036.
2. Stewart CL, Kaspar P, Brunet LJ, Bhatt H, Gadi I, Kontgen F *et al*. Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. *Nature* 1992; **359**: 76–79.
3. Hu W, Feng Z, Teresky AK, Levine AJ. p53 regulates maternal reproduction through LIF. *Nature* 2007; **450**: 721–724.
4. Hu W, Feng Z, Atwal GS, Levine AJ. p53: a new player in reproduction. *Cell Cycle* 2008; **7**: 848–852.
5. Atwal GS, Bond GL, Metsuyanim S, Papa M, Friedman E, Distelman-Menachem T *et al*. Haplotype structure and selection of the MDM2 oncogene in humans. *Proc Natl Acad Sci USA* 2007; **104**: 4524–4529.
6. Atwal GS, Kirchhoff T, Bond EE, Montagna M, Menin C, Bertorelle R *et al*. Altered tumor formation and evolutionary selection of genetic variants in the human MDM4 oncogene. *Proc Natl Acad Sci USA* 2009; **106**: 10236–10241.
7. Kang HJ, Feng Z, Sun Y, Atwal G, Murphy ME, Rebbeck TR *et al*. Single-nucleotide polymorphisms in the p53 pathway regulate fertility in humans. *Proc Natl Acad Sci USA* 2009; **106**: 9761–9766.
8. Feng Z, Zhang C, Kang HJ, Sun Y, Wang H, Naqvi A *et al*. Regulation of female reproduction by p53 and its family members. *Faseb J* 2011; **25**: 2245–2255.
9. Ammendola M, Gloria-Bottini F, Sesti F, Piccione E, Bottini E. Association of p53 codon 72 polymorphism with endometriosis. *Fertil Steril* 2008; **90**: 406–408.
10. Chang CC, Hsieh YY, Tsai FJ, Tsai CH, Tsai HD, Lin CC. The proline form of p53 codon 72 polymorphism is associated with endometriosis. *Fertil Steril* 2002; **77**: 43–45.
11. Vietri MT, Molinari AM, Iannella I, Cioffi M, Bontempo P, Ardovino M *et al*. Arg72Pro p53 polymorphism in Italian women: no association with endometriosis. *Fertil Steril* 2007; **88**: 1468–1469.
12. Lattuada D, Vigano P, Somigliana E, Abbiati A, Candiani M, Di Blasio AM. Analysis of the codon 72 polymorphism of the TP53 gene in patients with endometriosis. *Mol Hum Reprod* 2004; **10**: 651–654.
13. Patounakis G, Treff N, Tao X, Lonczak A, Scott RT Jr, Frattarelli JL. The p53 codon 72 single nucleotide polymorphism lacks a significant effect on implantation rate in fresh in vitro fertilization cycles: an analysis of 1,056 patients. *Fertil Steril* 2009; **92**: 1290–1296.
14. Koshiol J, Hildesheim A, Gonzalez P, Bratti MC, Porras C, Schiffman M *et al*. Common genetic variation in TP53 and risk of human papillomavirus persistence and progression to CIN3/cancer revisited. *Cancer Epidemiol Biomarkers Prev* 2009; **18**: 1631–1637.
15. Savage SA, Burdett L, Troisi R, Douglass C, Hoover RN, Chanock SJ. Germ-line genetic variation of TP53 in osteosarcoma. *Pediatr Blood Cancer* 2007; **49**: 28–33.
16. Marcel V, Tran PL, Sagne C, Martel-Planche G, Vaslin L, Teulade-Fichou MP *et al*. G-quadruplex structures in TP53 intron 3: role in alternative splicing and in production of p53 mRNA isoforms. *Carcinogenesis* 2011; **32**: 271–278.
17. Gemignani F, Moreno V, Landi S, Moullan N, Chabrier A, Gutierrez-Enriquez S *et al*. A TP53 polymorphism is associated with increased risk of colorectal cancer and with reduced levels of TP53 mRNA. *Oncogene* 2004; **23**: 1954–1956.
18. Wu X, Zhao H, Amos CI, Shete S, Makan N, Hong WK *et al*. p53 Genotypes and Haplotypes Associated With Lung Cancer Susceptibility and Ethnicity. *J Natl Cancer Inst* 2002; **94**: 681–690.
19. Wang-Gohrke S, Rebbeck TR, Besenfelder W, Kreienberg R, Runnebaum IB. p53 germline polymorphisms are associated with an increased risk for breast cancer in German women. *Anticancer Res* 1998; **18**: 2095–2099.
20. Marcel V, Palmero EI, Falagan-Lotsch P, Martel-Planche G, Ashton-Prolla P, Olivier M *et al*. TP53 PIN3 and MDM2 SNP309 polymorphisms as genetic modifiers in the Li-Fraumeni syndrome: impact on age at first diagnosis. *J Med Genet* 2009; **46**: 766–772.
21. Eskenazi B, Warner ML. Epidemiology of endometriosis. *Obstet Gynecol Clin North Am* 1997; **24**: 235–258.
22. Cunha-Filho JS, Gross JL, Lemos NA, Brandelli A, Castillos M, Passos EP. Hyperprolactinemia and luteal insufficiency in infertile patients with mild and minimal endometriosis. *Horm Metab Res* 2001; **33**: 216–220.
23. Podgaec S, Abrao MS, Dias JA Jr., Rizzo LV, de Oliveira RM, Baracat EC. Endometriosis: an inflammatory disease with a Th2 immune response component. *Hum Reprod* 2007; **22**: 1373–1379.
24. Bischoff F, Simpson JL. Genetic basis of endometriosis. *Ann N Y Acad Sci* 2004; **1034**: 284–299.
25. Giudice LC, Kao LC. Endometriosis. *Lancet* 2004; **364**: 1789–1799.
26. Hsieh YY, Lin CS. P53 codon 11, 72, and 248 gene polymorphisms in endometriosis. *Int J Biol Sci* 2006; **2**: 188–193.
27. Ribeiro CL Jr, Arruda JT, Silva CT, Moura KK. Analysis of p53 codon 72 gene polymorphism in Brazilian patients with endometriosis. *Genet Mol Res* 2009; **8**: 494–499.
28. Salzano FM, Bortolini MC. *The Evolution and Genetics of Latin American Populations*. Cambridge University Press: Cambridge, 2002, p 509.
29. Parra FC, Amado RC, Lambertucci JR, Rocha J, Antunes CM, Pena SD. Color and genomic ancestry in Brazilians. *Proc Natl Acad Sci USA* 2003; **100**: 177–182.
30. Marrero AR, Das Neves Leite FP, De Almeida Carvalho B, Peres LM, Kommers TC, Da Cruz IM *et al*. Heterogeneity of the genome ancestry of individuals classified as White in the state of Rio Grande do Sul, Brazil. *Am J Hum Biol* 2005; **17**: 496–506.

31. Zembrzuski VM, Callegari-Jacques SM, Hutz MH. Application of an African Ancestry Index as a genomic control approach in a Brazilian population. *Ann Hum Genet* 2006; **70**(Pt 6): 822–828.
32. Pena SD, Di Pietro G, Fuchshuber-Moraes M, Genro JP, Hutz MH, Kehdy Fde S *et al.* The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. *PLoS One* 2011; **6**: e17063.
33. Kay C, Jeyendran RS, Coulam CB. p53 tumour suppressor gene polymorphism is associated with recurrent implantation failure. *Reprod Biomed Online* 2006; **13**: 492–496.
34. Pietrowski D, Bettendorf H, Riener EK, Keck C, Hefler LA, Huber JC *et al.* Recurrent pregnancy failure is associated with a polymorphism in the p53 tumour suppressor gene. *Hum Reprod* 2005; **20**: 848–851.
35. Goodman C, Jeyendran RS, Coulam CB. P53 tumor suppressor factor, plasminogen activator inhibitor, and vascular endothelial growth factor gene polymorphisms and recurrent implantation failure. *Fertil Steril* 2009; **92**: 494–498.
36. Vousden KH, Prives C. Blinded by the Light: The Growing Complexity of p53. *Cell* 2009; **137**: 413–431.
37. Pellicer A, Oliveira N, Ruiz A, Remohi J, Simon C. Exploring the mechanism(s) of endometriosis-related infertility: an analysis of embryo development and implantation in assisted reproduction. *Hum Reprod* 1995; **10**(Suppl 2): 91–97.
38. Guerreiro-Junior V, Bisso-Machado R, Marrero A, Hunemeier T, Salzano FM, Bortolini MC. Genetic signatures of parental contribution in black and white populations in Brazil. *Genet Mol Biol* 2009; **32**: 1–11.
39. ASRM. Revised American Society for Reproductive Medicine classification of endometriosis: 1996. *Fertil Steril* 1997; **67**: 817–821.
40. Palmero EI, Schuler-Faccini L, Caleffi M, Achatz MI, Olivier M, Martel-Planche G *et al.* Detection of R337H, a germline TP53 mutation predisposing to multiple cancers, in asymptomatic women participating in a breast cancer screening program in Southern Brazil. *Cancer Lett* 2008; **261**: 21–25.
41. Lewontin RC. On measures of gametic disequilibrium. *Genetics* 1988; **120**: 849–852.

Cell Death and Disease **is an open-access journal published by *Nature Publishing Group*. This work is licensed under the Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/3.0/**

Supplementary Information accompanies the paper on Cell Death and Disease website (http://www.nature.com/cddis)

TABLE S1 – Self-reported skin color distribution in the four subgroups of individuals included in this study.

		Self-reported skin color							
		White	**Non-white**	**Total**	*P*-value[a]	*P*-value[b]	*P*-value[c]	*P*-value[d]	*P*-value[e]
		N (%)	**N (%)**	**N (%)**					
Study groups	**Fertile**	116 (86.6)	18 (13.4)	134 (100)					
	Unselected	259 (86.3)	41 (13.7)	300 (100)					
	END	87 (88.8)	11 (11.2)	98 (100)	0.948	0.615	0.156	0.533	0.102
	IVF	106 (92.2)	9 (7.8)	115 (100)					
Total		568 (87.8)	79 (12.2)	647 (100)					

([a]) χ^2 test between Fertile and Unselected; ([b]) χ^2 test between Fertile and END; ([c]) χ^2 test between Fertile and IVF; ([d]) χ^2 test between Unselected and END; ([e]) χ^2 test between Unselected and IVF;

Table S2: Hardy-Weinberg equilibrium calculation

TP53	Unselected	Fertile	END	IVF
	P-value*	*P*-value*	*P*-value*	*P*-value*
PIN2 rs1642785	0.915	0.929	0.533	0.154
PIN3 rs17878362	0.816	0.917	0.205	0.066
PEX4 rs1042522	0.729	0.940	0.482	0.154

(*) p-value for the Hardy-Weinberg equilibrium; IVF Group: women with recurrent failure of IVF; END Group: infertile women with minimal or mild endometriosis; Fertile: Fertile women; Unselected: women unselected with respect to fertility or infertility–related symptoms.

TABLE S3 – *TP53* PIN3 and PEX4 allele frequencies in control populations from different ethnic backgrounds.

	Ethnic group								
	Caucasian		African						
Allele frequencies	**Weston *et al.*1997***	**Marcel *et al.* 2010****	**Weston *et al.* 1997***	**Marcel *et al.* 2010****	**Present study*****	*P*-value[a]	*P*-value[b]	*P*-value[c]	*P*-value[d]
***TP53* PIN3**									
N	0.89	0.90	0.68	0.67	0.86	0.521	0.384	<0.01	<0.01
D	0.11	0.10	0.32	0.33	0.14				
***TP53* PEX4**									
C	0.21	0.22	0.63	0.67	0.28	0.249	0.327	<0.01	<0.01
G	0.79	0.78	0.37	0.33	0.72				

PIN3 N = non-duplicated; D = duplicated.
* Controls (breast cancer unaffected women, n = 117)
** HapMap sample set (n = 90)
*** Women of the general population, unselected for fertility (n=300)
([a]) χ^2 test between Caucasians from Weston *et al.* and the present study; ([b]) χ^2 test between Caucasians from Marcel *et al.* and the present study; ([c]) χ^2 test between Africans from Weston *et al.* and the present study; ([d]) χ^2 test between Africans from Marcel *et al.* and the present study.

Table S4: Linkage disequilibrium calculation

		IVF					END					Fertile		
		D'					D'					D'		
		PIN2	PIN3	PEX4			PIN2	PIN3	PEX4			PIN2	PIN3	PEX4
r^2	PIN2	-	0.8	1	r^2	PIN2	-	0.55	1	r^2	PIN2	-	0.948	1
	PIN3	0.5	-	0.76		PIN3	0.25	-	0.63		PIN3	0.74	-	0.01
	PEX4	1	0.45	-		PEX4	0.9	0.34	-		PEX4	0.94	1E-04	-

D': measure of linkage disequilibrium; r^2: correlation coefficient between pairs of loci; IVF Group: women with recurrent failure of IVF; END Group: infertile women with minimal or mild endometriosis; Fertile: Fertile women.

Supplementary information: Ethnic variations in the Brazilian population

It is widely recognized that population stratification may be a confounding factor in genetic association studies and the impact of stratification in population-based association studies has been discussed extensively (1-3). With respect to the polymorphisms analyzed here, especially PIN3 and PEX4, previous studies demonstrate that allelic and genotypic frequencies are distributed differentially according to ethnic background. For instance, in a case-control study performed by Weston et al. (1997) on the association of *TP53* polymorphisms and breast cancer risk in individuals from the ethnically heterogeneous population of New York city, PEX4 and PIN3 allele frequencies were significantly different between Caucasian, Hispanic and African-american individuals (4). These differences were also shown by Marcel et al. (2010) who analyzed 90 Caucasian, Asian and African individuals from the Hapmap dataset (5). Thus, it is be reasonable to consider that population structure interferes with association studies that involve these polymorphisms, in particular in the case of populations reputed as ethnically diverse and heterogeneous such as the population of Brazil.

In the context of the present study, however, both the information available on the population structure in Southern Brazil and our data on phenotypic and genotypic variations among the subjects we have recruited, support that the impact of population structure on allele distributions and associations are, at the most, limited, and do not significantly bias our results. This interpretation is based on the following considerations:

1. Landmark studies conducted with a large number of Brazilian subjects have concluded that the current Brazilians have contributions of different continental ancestries (European, Amerindian and African). However the proportion of the contribution of these parental stocks varies significantly among different Brazilian regions. In one of such recent molecular analyses, Pena and collaborators (2011) demonstrated that European ancestry was predominant in all Brazilian regions, with proportions ranging from around 60% in the Northeast to around 80% in the South (6), corroborating previous studies that show that populations from Southern Brazil have a major European contribution (7-10). Thus, Southern Brazil (including the State of Rio Grande do Sul and the city of Porto Alegre, where the study was conducted), has the largest proportion of European contribution in the country. This has been demonstrated in several studies, where populations from this region have been described as having "transplanted European genomes" (9). Furthermore, even in individuals who self-declare themselves as "Black", the genomic European contribution is highly significant (6, 7). This first point argues in favor of a less heterogeneous study population with a predominance of European genomes and ancestry, as defined by population-based studies

2. When analyzing self-reported skin color, which in Brazil is commonly used as surrogate to define an equivalent of "race" or ethnic background (6, 11), patients were separated in two groups: "white" (women who defined themselves with Caucasian features) and "non-white" (women that identified themselves with features suggestive of some level of African ancestry, such as "mulato" or "pardo"). No significant difference in the distribution of self-denominated skin color was observed among the four study subgroups. As expected from previous population-based studies in the region of Porto Alegre, Rio Grande do Sul, the majority of individuals in all four groups self-denominated them as "white", corroborating with the previous assumptions described in argument 1 (see Table S1).

3. When analyzing allele frequencies of *TP53* polymorphisms PIN3 and PEX4, no significant differences were observed between any of the 4 subgroups of the present study (IVF, END, Fertile and Unselected) and allele frequencies reported in other studies on European populations (see Table S3).

References

1. Thomas DC, Witte JS. Point: population stratification: a problem for case-control studies of candidate-gene associations? *Cancer Epidemiol Biomarkers Prev* 2002 Jun; **11**(6)**:** 505-512.

2. Wacholder S, Rothman N, Caporaso N. Counterpoint: bias from population stratification is not a major threat to the validity of conclusions from epidemiological studies of common polymorphisms and cancer. *Cancer Epidemiol Biomarkers Prev* 2002 Jun; **11**(6)**:** 513-520.

3. Wang H, Haiman CA, Kolonel LN, Henderson BE, Wilkens LR, Le Marchand L*, et al.* Self-reported ethnicity, genetic structure and the impact of population stratification in a multiethnic study. *Hum Genet* 2010 Aug; **128**(2)**:** 165-177.

4. Weston A, Pan CF, Ksieski HB, Wallenstein S, Berkowitz GS, Tartter PI*, et al.* p53 haplotype determination in breast cancer. *Cancer Epidemiol Biomarkers Prev* 1997 Feb; **6**(2)**:** 105-112.

5. Marcel V, Palmero EI, Falagan-Lotsch P, Martel-Planche G, Ashton-Prolla P, Olivier M*, et al.* TP53 PIN3 and MDM2 SNP309 polymorphisms as genetic modifiers in the Li-Fraumeni syndrome: impact on age at first diagnosis. *J Med Genet* 2010 Nov; **46**(11)**:** 766-772.

6. Pena SD, Di Pietro G, Fuchshuber-Moraes M, Genro JP, Hutz MH, Kehdy Fde S*, et al.* The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. *PLoS One* 2011; **6**(2)**:** e17063.

7. Salzano FM BM. *The Evolution and Genetics of Latin American Populations*. Cambridge University Press: Cambridge, 2002, 509 pppp.

8. Parra FC, Amado RC, Lambertucci JR, Rocha J, Antunes CM, Pena SD. Color and genomic ancestry in Brazilians. *Proc Natl Acad Sci U S A* 2003 Jan 7; **100**(1)**:** 177-182.

9. Marrero AR, Das Neves Leite FP, De Almeida Carvalho B, Peres LM, Kommers TC, Da Cruz IM*, et al.* Heterogeneity of the genome ancestry of individuals classified as White in the state of Rio Grande do Sul, Brazil. *Am J Hum Biol* 2005 Jul-Aug; **17**(4)**:** 496-506.

10. Zembrzuski VM, Callegari-Jacques SM, Hutz MH. Application of an African Ancestry Index as a genomic control approach in a Brazilian population. *Ann Hum Genet* 2006 Nov; **70**(Pt 6)**:** 822-828.

11. Guerreiro-Junior V, Bisso-Machado R, Marrero A, Hunemeier T, Salzano FM, Bortolini MC. Genetic signatures of parental contribution in black and white populations in Brazil. *Genet Mol Biol* 2009 Jan; **32**(1)**:** 1-11.

4. CAPÍTULO 2

4.1 *THE TP53 FERTILITY NETWORK*

Título do manuscrito: *The TP53 fertility network*

Autores: Paskulin DD, Paixão-Côrtes VR, Hainaut P, Bortolini MC, Ashton-Prolla P

Situação: Publicado

Revista: *Genetics and Molecular Biology*

Referência: Genet Mol Biol. 2012 Dec;35(4 (suppl)):939-46. Epub 2012 Dec 18. http://dx.doi.org/10.1590/S1415-47572012000600008

Website: http://www.ncbi.nlm.nih.gov/pubmed/23412905

Genetics and Molecular Biology, 35, 4 (suppl), 939-946 (2012)
Copyright © 2012, Sociedade Brasileira de Genética. Printed in Brazil
www.sbg.org.br

Research Article

The *TP53* fertility network

Diego d'Avila Paskulin[1,2], Vanessa Rodrigues Paixão-Côrtes[1,3], Pierre Hainaut[4], Maria Cátira Bortolini[1,3] and Patricia Ashton-Prolla[1,2,5*]

[1]*Programa de Pós-Graduação em Genética e Biologia Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil.*
[2]*Laboratório de Medicina Genômica, Hospital de Clínicas de Porto Alegre, Porto Alegre, RS, Brazil.*
[3]*Laboratório de Evolução Humana e Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil.*
[4]*International Prevention Research Institute, Lyon, France.*
[5]*Serviço de Genética Médica, Hospital de Clínicas de Porto Alegre, Porto Alegre, RS, Brazil.*

Abstract

The *TP53* gene, first described in 1979, was identified as a tumor suppressor gene in 1989, when it became clear that its product, the p53 nuclear phosphoprotein, was frequently inactivated in many different forms of cancers. Nick-named "guardian of the genome", *TP53* occupies a central node in stress response networks. The p53 protein has a key role as transcription factor in limiting oncogenesis through several growth suppressive functions, such as initiating apoptosis, senescence, or cell cycle arrest. The p53 protein is directly inactivated in about 50% of all tumors as a result of somatic gene mutations or deletions, and over 80% of tumors demonstrate dysfunctional p53 signaling. Beyond the undeniable importance of p53 as a tumor suppressor, an increasing number of new functions for p53 have been reported, including its ability to regulate energy metabolism, to control autophagy, and to participate in various aspects of differentiation and development. Recently, studies on genetic variations in *TP53* among different populations have led to the notion that the p53 protein might play an important role in regulating fertility. This review summarizes current knowledge on the basic functions of different genes of the *TP53* family and *TP53* pathway with respect to fertility. We also provide original analyses based on genomic and genotype databases, providing further insights into the possible roles of the *TP53* pathway in human reproduction.

Keywords: *TP53*, fertility, p53 network.

The *TP53* Gene, its Products and Regulation

The transcription factor p53 is encoded by the Tumor Protein p53 gene (*TP53*, OMIM 191170), which in humans is located on the short arm of chromosome 17 (17p13.1). *TP53* is composed of 19,198 nucleotides, spanning 11 exons and encoding a 393 amino acid protein that functions primarily as a transcription factor and is biologically active as a homotetramer. The p53 protein has six major domains and its expression is subject to multiple regulation, at transcriptional, post-transcriptional, translational and post-translational levels (Hollstein and Hainaut, 2010). A further level in complexity is generated by the expression of p53 as up to 10 distinct isoforms produced by alternative splicing, alternative promoter usage, and alternative translation initiation (Bourdon *et al.*, 2005).

Send correspondence to Patricia Ashton-Prolla. Serviço de Genética Médica, Hospital de Clínicas de Porto Alegre, Rua Ramiro Barcelos 2350, 90035-903 Porto Alegre, RS, Brazil. E-mail: pprolla@hcpa.ufrgs.br.

The p53 protein functions as a multitarget transcription factor. Upon cellular stress signals (including DNA damage, oncogene activation, hypoxia, nutrient deprivation, telomere erosion and ribosomal stress) p53 is activated through protein stabilization and post-translational modifications, allowing full p53 transactivation potential (Kruse and Gu, 2009). Induction of growth arrest or cell death upon activation of p53 prevents the replication of damaged DNA and the division of genetically altered cells, therefore, playing an important role in maintaining the integrity of the genome (Lane, 1992). The importance of p53 as a tumor suppressor is illustrated by the observation that many individuals affected by Li-Fraumeni syndrome (a high penetrance hereditary cancer syndrome that predisposes to multiple early-onset cancers) are carriers of loss of function germline mutations in *TP53* (Malkin *et al.*, 1990).

The p53 protein interacts with a large number of partner proteins, but special attention has been given to the p53-Mdm2 interaction. Among other biochemical func-

tions, the Mdm2 protein (encoded by *MDM2*, the human homolog of the Murine Double Minute 2 gene, OMIM 164785) operates as E3 ubiquitin ligase to induce p53 polyubiquitination, mediating its nuclear export and targeting it to the proteasome for degradation (Lain and Lane, 2003). Interestingly, Mdm2 forms a negative-feedback loop with p53, as *MDM2* transcription is positively and directly regulated by p53 (Michael and Oren, 2003). This autoregulatory loop maintains a delicate equilibrium in the precise regulation of protein levels and activities of both p53 and Mdm2. Like *MDM2*, its structural homolog, *MDM4* (OMIM 602704), can also bind directly to p53, inhibiting its ability to function as a transcriptional activator, as well as regulating its stability, most likely through interactions with Mdm2 (Toledo *et al.*, 2006). However, unlike Mdm2, Mdm4 (also known as MdmX), is devoid of autonomous E3 ligase activity. Another important regulator of p53 function is the Ubiquitin-Specific Protease 7, *USP7* (OMIM 602519), which de-ubiquitylates p53 and protects it from proteasome-mediated degradation (Shan *et al.*, 2008). The pivotal role of Mdm2 and Mdm4 in the control of p53 function supports the notion that polymorphisms at these *loci* might be potential modifiers of p53 function (Atwal *et al.*, 2009).

The *TP53* Family

The *TP53* family of genes includes *TP53* and two structural and functional p53 homologs: *TP63* (OMIM 603273) and *TP73* (OMIM 601990). These three members share a very high homology in the DNA binding domain, as well as in overall protein architecture, with p63 and p73 being more closely related to each other than to p53 (Belyi *et al.*, 2010). Although *TP63* and *TP73* have been discovered well after p53, they seem to have appeared earlier than *TP53* during evolution. The current view is that the three genes derive from a single ancestor which was most likely a p63/p73-like gene (Nedelcu and Tan, 2007). The overall protein architecture is highly conserved from *Drosophila* to man, and consists of a central sequence specific DNA binding domain (DBD), an N-terminal transactivation domain (TA) and a C-terminal oligomerization domain (OD). p63 and p73 contain a sterile α-motif (SAM) domain at their C-terminus that plays a role in protein-protein interactions and which has no structural equivalent in p53, and a transcription inhibition domain (TID) that decreases their transcriptional activity by enforcing a closed conformation through interaction with the amino-terminal TA domain (Straub *et al.*, 2010). The competition for binding to DNA through their structurally similar DBDs suggests that p53, p63 or p73 may cooperate or compete for the regulation of common transcriptional targets.

Given that the main forms of p63 and p73 are the so-called delta-N isoforms lacking the main, N-terminal TA domain, interferences between the three proteins can result in either synergistic effects or dominant-negative effects, in which p63 or p73 isoforms may down-regulate p53 (Levine *et al.*, 2011). While p53 acts mainly in response to different stresses, p63 is one of the major transcription factors required for the development of stratified epithelia, making it essential for the limb and for the formation of a functional skin (Gonfloni *et al.*, 2009; Mills *et al.*, 1999). Accordingly, ΔNp63 isoforms play distinct roles in regulating epithelium-mesenchyme interactions through the regulation of TGFβ, resulting in a more invasive phenotype in the presence of ΔNp63γ (Lindsay *et al.*, 2011; Oh *et al.*, 2011). p73 is involved in the development of the immune and central nervous system (Belyi *et al.*, 2010) and has important functions in the regulation of the spindle assembly checkpoint (SAC) during meiosis and mitosis (Tomasini *et al.*, 2008), as it prevents aneuploidy through sensing the improper attachment of sister chromatids to the mitotic or meiotic spindle and delays anaphase until chromosomes are correctly oriented for segregation (Gardner and Burke, 2000). Considering these observations, a functional divergence among p53, p63 and p73 clearly emerges. While p53 behaves as a canonical tumor suppressor gene which is mostly active after induction by various forms of stress, both p63 and p73 play major roles in normal ectodermal differentiation and neurogenesis and only a secondary one in response to the types of stress that activate p53.

TP53 Regulates Reproduction Through *LIF*

The current knowledge on p53 regulation and functions has been the subject of a detailed recent review (Vousden and Prives, 2009). However, most studies have concentrated on p53 as a stress-induced tumor suppressor gene, and little is known about its function in normal cellular processes. The p53 protein accomplishes its function by transcriptionally regulating target genes. In 2002, several genomic DNA sequences were detected where the p53 protein was most likely able to bind and activate transcription (Hoh *et al.*, 2002).

Among these, a potential candidate was the gene encoding the leukemia inhibitory factor (*LIF*, OMIM 159540), a secreted cytokine that is critical for blastocyst implantation (Stewart *et al.*, 1992). This gene contains a putative p53-binding consensus DNA sequence in intron 1, which is conserved in both mouse and human gene sequences (Hu *et al.*, 2007a). In fact, implantation cannot occur unless epithelial cells lining the uterus are exposed to LIF (Stewart *et al.*, 1992), most likely expressed at the onset of implantation, which occurs at day 4 of pregnancy in mice (day 12 in humans). LIF null mice have a defect in maternal reproduction caused by the complete lack of uterine decidualization at the implantation stage, with consequent failure of blastocyst implantation, which can be rescued by LIF injection at the implantation stage (the 4th day of pregnancy in mice) (Chen *et al.*, 2000). Hu *et al.*

(2007b) demonstrated that p53 plays a significant role in fertility, since p53-null female mice present reduced uterine expression of LIF and, as expected, reduced maternal reproduction due to impaired implantation functions. Administering LIF to p53 deficient female mice at day 4 of pregnancy significantly increased the pregnancy rate and litter size with improved blastocyst implantation. These findings demonstrate that inactivation of p53 decreases the levels and function of uterine LIF, thus indicating a function for p53 in maternal reproduction through the regulation of LIF.

In addition, estrogen is also involved in the regulation of transient LIF expression at the implantation stage (Chen *et al.*, 2000), mediated through its nuclear receptor alpha (ERα, encoded by *ESR1*, OMIM 133430). Feng *et al.* (2011) demonstrated a significant increase in nuclear ERα levels in endometrial glands at the implantation stage in mice, and concluded that the increased expression of LIF at this stage requires the activation of p53, increased estrogen levels, and activated ERα.

Single Nucleotide Polymorphisms and Infertility

Considering the strict regulation of LIF by p53, it is reasonable to expect that modulation of p53 function by single nucleotide polymorphisms (SNPs) in *TP53* and *TP53*-related genes may affect fertility. In humans there are many naturally occurring SNPs in genes at critical nodes in the *TP53* pathway, including *TP53*, *MDM2*, *MDM4*, and *USP7*, all of which have known functional variants that can modify the levels or activity of the p53 protein (Atwal *et al.*, 2009; Bond *et al.*, 2004). One of the most commonly studied *TP53* variants, the non-silent polymorphism Pro72Arg (rs1042522; C/G), is associated with biochemical and functional differences in protein functions, since the protein carrying the Pro72 allele is more efficient in initiating senescence and cell cycle arrest, while the one with the Arg72 allele is more active in inducing apoptosis and suppressing cellular transformation (Dumont *et al.*, 2003; Thomas *et al.*, 1999). The Pro72 isoform is also observed in other primates, including the chimpanzee, while the Arg72 one is only present in humans, thus suggesting that the C (Pro72) allele may correspond to the ancestral allele.

Compared with *TP53* Pro72, the Arg72 allele presents higher transcriptional activity toward a subset of p53 target genes, including *LIF*. The induction of LIF is over 2-fold higher in cells with the Arg72 allele than in cells with the Pro72 allele (Kang *et al.*, 2009), leading to decreased implantation success. Kay *et al.* (2006) associated the Pro72 allele with recurrent implantation failure and demonstrated that the Pro72 allele is enriched in women with unexplained infertility from an *in vitro* fertilization clinic, compared with a fertile control population (Kang *et al.*, 2009). In Brazil, Ribeiro Junior *et al.* (2009) associated the Pro72 allele with intense pain in a cohort of endometriotic patients and Bianco *et al.* (2011) considered that the Pro72Arg polymorphism was not a risk factor for infertility or endometriosis in Brazilian infertile patients. Interestingly, we found that both *TP53* Pro72Arg and *MDM4* rs1563828 are associated with twinning in Cândido Godói (Tagliani-Ribeiro *et al.*, 2012), a small town in Brazil remarkable for showing a high frequency of both dizygotic and monozygotic twins (Tagliani-Ribeiro *et al.*, 2011).

An additional remarkable fact regarding Pro72Arg is that the allele frequencies for this SNP vary widely across human populations. For instance, Arg72 frequencies range from ~20% in some Sub-Saharan populations to ~80% in northern Europeans, while in Asians the values are intermediate (HapMap and Alfred database, respectively). These distinct allele frequencies promote a level of differentiation (F_{ST}) of 19% between Caucasians and Yoruba from Nigeria (Table S1). Recently, the complete nuclear genomes of two extinct hominids belonging the genus *Homo*, *Homo neanderthalensis* and Denisova specimen were published (Green *et al.*, 2010; Reich *et al.*, 2010). Based on these genomic data sets, compiled from UCSC Genome Browser, only Pro72 is present in both archaic human sequences. Inference from this observation is that the C → G mutation may have a relatively recent origin, *i.e.* the Arg72 variant may be *Homo sapiens*-specific. Additional studies on archaic human species will be needed to confirm this hypothesis.

The *MDM2* SNP309 (rs2279744; T/G) is another commonly described variant that attenuates the *TP53* pathway. It is a gain of function variant that increases the affinity of a sequence in *MDM2* for the Sp1 transcription factor leading to increased transcription of the Mdm2 protein, and consequent inhibition or attenuation of the *TP53* pathway-mediated tumor suppression functions (Bond *et al.*, 2004). Interestingly, SNP309 is located in a transcriptional enhancer region of *MDM2* regulated by estrogen signaling (Phelps *et al.*, 2003). Because SNP309 increases the binding affinity for Sp1, a co-activator of multiple hormone receptors, it could potentially affect the hormone-dependent regulation of *MDM2* transcription and result in further elevation of the Mdm2 protein levels, as estrogen preferentially stimulates transcription of the 309G allele (Hu *et al.*, 2007b). In addition to *TP53* Pro72 and *MDM2* 309G, other variants in *TP53*-related genes (*MDM4*, rs1563828: T/C; *USP7*, rs1529916: T/C; and *LIF*, rs929271: G/T) have been proposed as functional variants with a role in reproduction, showing differential allele frequencies in young infertile women submitted to *in vitro* fertilization when compared to fertile women (Kang *et al.*, 2009).

From an evolutionary perspective, *TP53* Arg72 and *MDM2* 309G seem to have been positively selected in European and Asian populations, which can be interpreted as a result of adaptive pressures during the dispersion of *Homo sapiens* from Africa to other continents (Atwal *et al.*, 2007; Shi *et al.*, 2009; Belyi *et al.*, 2010). Several studies indicate

that p53 has evolutionarily conserved functions other than acting as a tumor suppressor, and the existence of p53-like proteins in short-lived organisms that do not exhibit adult cancer incidence, such as flies and worms, adds to the argument that tumor suppression was not the original function for p53 and its pathway (Lu *et al.*, 2009). In addition, the major impact of p53 in cancer prevention or longevity in humans likely occurs in post-reproductive years, which would exclude a major evolutionary role associated to these functions. Like *TP53* and *MDM2*, *MDM4* and *USP7* also appear to have alleles or haplotypes that are under selection pressure and show geographic variations in allele distribution (Atwal *et al.*, 2007; Shi *et al.*, 2009). In a recent study, Feng *et al.* (2007) reported that SNPs in the *TP63* (rs17506395; T/G) and *TP73* (rs4648551 G/A and rs6695978 G/A) genes are associated with infertility in women, independently of age for *TP63* and specifically in women aged over 35 years for *TP73*. The authors proposed that the possible mechanisms of infertility associated with variations in *TP53* might be impaired implantation, whereas variations in *TP63* and *TP73* may affect the quality of oocytes and induce chromosomal aneuploidy (Feng *et al.*, 2007).

Based on these findings, it is reasonable to assume that alleles in genes of the *TP53* family and *TP53* pathway with reproductive implications may have been important targets for selection pressure during the human evolutionary history.

The *TP53* Fertility Network

The p53 protein and its signal transduction pathway are composed of a set of genes and their protein products that are designed to respond to a wide variety of intrinsic and extrinsic stress signals. Although the important interaction between p53 and *LIF* is crucial for embryo implantation, current evidence suggests that not only *LIF*, but other genes may be important in the reproductive stages of decidualization and implantation. To further test this hypothesis, we constructed a network with 18 *TP53* related genes involved in decidualization and implantation processes, including *LIF*, *MDM2* and others (Figure 1). Genes related to decidualization and implantation were compiled from the Gene Ontology website database using the AmiGO browser. Association among these genes was tested using the STRING 9 software which tests available known and predicted gene/protein interactions (Szklarczyk *et al.*, 2011). This "two-step" approach was chosen to minimize the possibility of false associations during the STRING analysis due to co-existence of words. Table S2 summarizes the 18 genes of this network, as well as their interconnections and wide range of functions. It is important to note that neither p63 nor p73 are on this list, since their main reproduction-related functions refer to the control of ovulation and female germ cell integrity in humans, and they apparently are not involved in decidualization or embryo implantation stages.

This "*TP53* Fertility Network" illustrates the importance of multiple genes in these specific stages of human fertility and opens a wide range of possibilities for genetic variation studies in genes not yet being investigated with regard to fertility. For example, *IGFBP7* (insulin-like growth factor binding protein 7) is predominantly expressed in the vasculature of developing embryos and regulates vascular endothelial growth factor-A-dependent neoangiogenesis (Hooper *et al.*, 2009). *ESR1* (estrogen receptor 1) gene is critical for LIF expression (Feng *et al.*, 2011) and *IL1B* (interleukin 1, beta) is involved in a variety of cellular activities that are essential for decidualization, including cell proliferation, differentiation, and apoptosis (Ben-Sasson *et al.*, 2009).

An additional analysis to verify if these 18 genes belong to a specific functional cluster was performed using GeneDecks V3 software. Thirteen of them were functionally clustered as having an involvement in the reproductive system (*CYP27B1, ESR1, LIF, MEN1, PLA2G4A, PLAU, PPARD, PTGS2, SOD1, SPP1, TP53, UBE2A* and *VDR*). More specifically, seven genes (*CALCA, IL1B, LIF, PPARD, PTGS2, SOD1* and *SPP1*) were associated with embryo implantation ($p = 1x10^{-16}$) and seven (*CYP27B1, LIF, PLA2G4A, PPARD, PTGS2, SPP1* and *VDR*) with decidualization ($p = 1x10^{-16}$). It is noteworthy that some genes are present in all functional clusters cited above (*e.g. PPARD;* peroxisome proliferator-activated receptor delta). This analysis brought additional evidence of the role of these genes in key stages of fertility.

Evolutionary Pattern of the *TP53* Fertility Network

In order to explore certain evolutionary aspects of the network we expanded the analysis on its suitability as a model using two different approaches and taking into consideration inter- (vertebrate) and intra- (human) species variations of the 18 genes that comprise the network.

The first approach was to assess the level of conservation of the genes included in the network along evolutionary lineages using comparative analysis between humans and other 21 vertebrate species. Data were compiled in the STRING 9.0 database whereas the level of identity of the amino acid sequences between humans and the others species was obtained using the LALIGN software. The 18 genes presented variable levels of amino acid sequence conservation (Table S3). Protein preservation among the species belonging to the primate order (human, chimpanzee, orangutan, and rhesus monkey) was on average 97%, while among placental mammals (human, chimpanzee, orangutan, and rhesus monkey, mouse, rat, guinea pig, rabbit, cow, cat, dog, horse, pig, and armadillo) it was 85%. In contrast, the degree of protein identity decreased signifi-

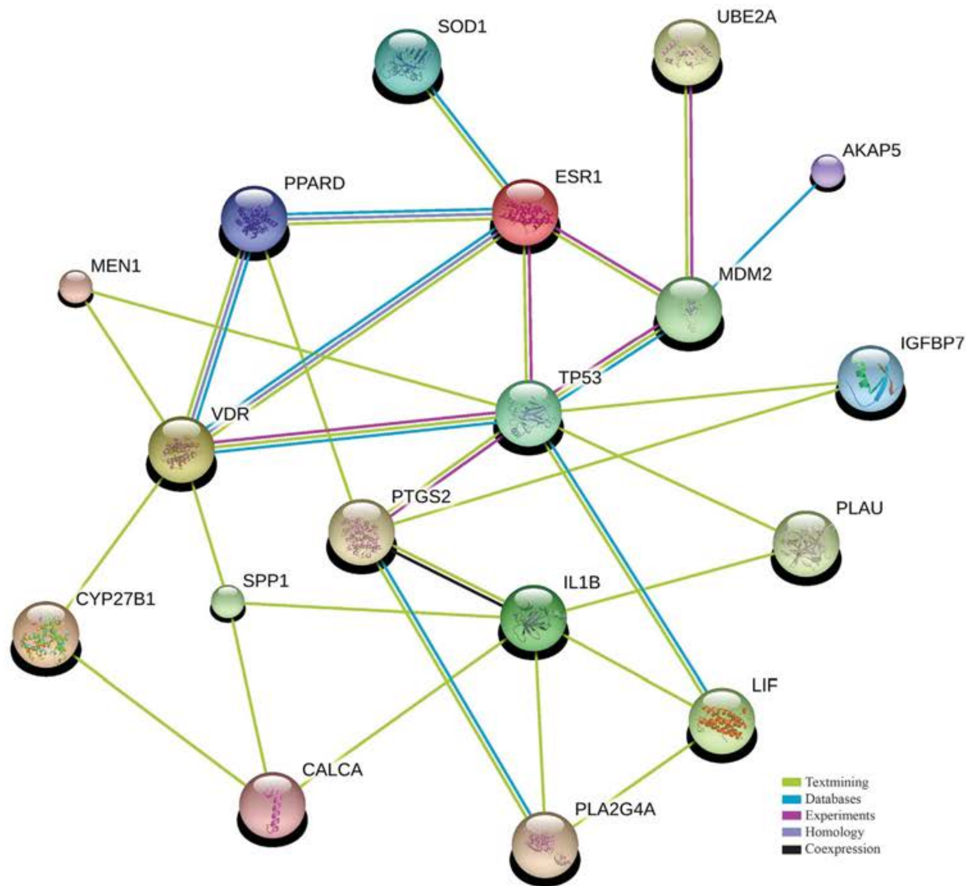


Figure 1 - Eighteen genes of the *TP53* Fertility Network. Image created by STRING Software 9.0 with high confidence score (0.7). Different types of lines represent the kind of evidence for the association. *AKAP5*: A-Kinase anchor protein 5; *CALCA*: Calcitonin-related polypeptide alpha; *CYP27B1*: Cytochrome P450; *IGFBP7*: Insulin-like growth factor binding protein 7; *IL1B*: Interleukin 1 Beta; *ESR1*: Estrogen Receptor alpha; *LIF*: Leukemia inhibitory factor; *MDM2*; Mouse Double Minute 2 Homolog; *MEN1*: Multiple Endocrine Neoplasia I; *PLA2G4A*: Phospholipase A2, group IVA; *PLAU*: Plasminogen activator urokinase; *PPARD*: Peroxisome proliferator-activated receptor delta; *PTGS2*: Prostaglandin-endoperoxide synthase 2; *SOD1*: Superoxide Dismutase 1, *SPP1*: Secreted Phosphoprotin 1; *TP53*: Tumor Protein p53; *UBE2A*: Ubiquitin-Conjugating Enzyme E2A; *VDR*: Vitamin D receptor.

cantly (average of 41%) when the comparison involved only humans and fish species. The results generated from STRING 9.0 show that overall 90% of the network's connections (edges) were retrieved in primates, while for placental mammals the value was reduced to 80%. However, when all vertebrates were considered, only 42% of the network is recovered. These results suggest that some of the 18 network genes may have acquired novel functions in different taxa, throughout vertebrate evolution, in addition to ancestral functions, a similar situation to that reported previously for the *HOX* family genes (Chen *et al.*, 2010).

The second approach was to study variation within the 18 genes between human populations. The data were compiled from HapMap and ENSEMBL databases. Using this strategy we identified 10,918 polymorphisms, only 1.4% of which being non-synonymous changes (Table S4). Of these, 82 (*e.g.* Pro72Arg) are predicted to be deleterious (Table S5). For most other potentially deleterious polymorphisms no striking difference was found in allele fre-

quencies among continental populations. However, some notable exceptions can be highlighted. Reminiscent of the Pro72Arg polymorphism, the rs5241 SNP located in the *CALCA* gene shows a frequency of 17% in Africans, whereas it is absent in European-descendents. On the other hand, for another SNP (rs2227564) located in the *PLAU* gene, the rare allele is only present in Euro-Asian populations (23%-33%), whereas it is absent in Africans (Table S1). Overall, these examples suggest that selection pressure for specific alleles in defined populations has affected only a small subset of the genes involved in the proposed network.

Since reproduction is central to the evolutionary process, in all vertebrates, as well as in other organisms, the genome is expected to be optimized for reproductive success. However, even among vertebrates there is an immense diversity in how reproduction occurs, including care and rearing of the offspring (de Magalhaes and Church, 2005; Plunkett *et al.*, 2011). There are many reasons in the evolutionary history of each species, including of our own and of other phylogenetically close species, such as Neanderthal, Denisova and chimpanzee that can explain shared and unique reproductive traits. Our results show for instance, that the human fertility network is not identical in all vertebrate species investigated here. An additional complicating factor is that for humans reproductive strategies can have changed drastically due to cultural practices, as well as in response to environmental pressures (*e.g.* climate change). Thus, it is expected that part of this diversity is the result of variable selection pressures encountered by human populations as they progressively expanded over the world.

Finally, this analysis adds further support to the idea that there is not a unique major effect gene involved in fertility and that an approach based on a wider gene network and taken under an evolutionary perspective may lead to the delineation of a more comprehensive view of the impact of p53 on the complex biology of fertility. Additional investigations at population level, as well as functional studies are needed to clarify the exact implications of the inter- (vertebrate) and intra- (human) differences highlighted in the present study.

Acknowledgments

This study was supported by grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq grant 475471/2009-1) and Fundo de Incentivo à Pesquisa do Hospital de Clínicas de Porto Alegre FIPE/HCPA (grant 09–430). DDP and VRPC are supported by grants from CNPq.

References

Atwal GS, Bond GL, Metsuyanim S, Papa M, Friedman E, Distelman-Menachem T, Ben Asher E, Lancet D, Ross DA, Sninsky J, *et al.* (2007). Haplotype structure and selection of the MDM2 oncogene in humans. Proc Natl Acad Sci USA 104:4524-4529.

Atwal GS, Kirchhoff T, Bond EE, Montagna M, Menin C, Bertorelle R, Scaini MC, Bartel F, Bohnke A, Pempe C, *et al.* (2009) Altered tumor formation and evolutionary selection of genetic variants in the human MDM4 oncogene. Proc Natl Acad Sci USA 106:10236-10241.

Belyi VA, Ak P, Markert E, Wang H, Hu W, Puzio-Kuter A and Levine AJ (2010) The origins and evolution of the p53 family of genes. Cold Spring Harb Perspect Biol 2:a001198.

Ben-Sasson SZ, Hu-Li J, Quiel J, Cauchetaux S, Ratner M, Shapira I, Dinarello CA and Paul WE (2009) IL-1 acts directly on CD4 T cells to enhance their antigen-driven expansion and differentiation. Proc Natl Acad Sci USA 106:7119-7124.

Bianco B, Christofolini DM, Brandes A, Lerner TG, Gonçalves-Filho RP, Souza AM and Barbosa CP (2011) Analysis of codon 72 polymorphism of the TP53 gene in infertile women with and without endometriosis. Rev Bras Ginecol Obstet 33:37-42 (in portuguese).

Bond GL, Hu W, Bond EE, Robins H, Lutzker SG, Arva NC, Bargonetti J, Bartel F, Taubert H, Wuerl P, *et al.* (2004) A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. Cell 119:591-602.

Bourdon JC, Fernandes K, Murray-Zmijewski F, Liu G, Diot A, Xirodimas DP, Saville MK and Lane DP (2005) p53 isoforms can regulate p53 transcriptional activity. Genes Dev 19:2122-2137.

Chen JR, Cheng JG, Shatzer T, Sewell L, Hernandez L and Stewart CL (2000) Leukemia inhibitory factor can substitute for nidatory estrogen and is essential to inducing a receptive uterus for implantation but is not essential for subsequent embryogenesis. Endocrinology 141:4365-4372.

Chen L, Zhao P, Wells L, Amemiya CT, Condie BG and Manley NR (2010) Mouse and zebrafish Hoxa3 orthologues have nonequivalent *in vivo* protein function. Proc Natl Acad Sci USA 107:10555-10560.

de Magalhaes JP and Church GM (2005) Genomes optimize reproduction: Aging as a consequence of the developmental program. Physiology (Bethesda) 20:252-259.

Dumon P, Leu JI, Della Pietra 3rd AC, George DL and Murphy M (2003) The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. Nat Genet 33:357-365.

Feng Z, Hu W, de Stanchina E, Teresky AK, Jin S, Lowe S and Levine A J (2007) The regulation of AMPK beta1, TSC2, and PTEN expression by p53: Stress, cell and tissue specificity, and the role of these gene products in modulating the IGF-1-AKT-mTOR pathways. Cancer Res 67:3043-3053.

Feng Z, Zhang C, Kang HJ, Sun Y, Wang H, Naqvi A, Frank AK, Rosenwaks Z, Murphy ME, Levine AJ, *et al.* (2011) Regulation of female reproduction by p53 and its family members. FASEB J 25:2245-2255.

Gardner RD and Burke DJ (2000) The spindle checkpoint: Two transitions, two pathways. Trends Cell Biol 10:154-158.

Gonfloni S, Di Tella L, Caldarola S, Cannata SM, Klinger FG, Di Bartolomeo C, Mattei M, Candi E, De Felici M, Melino G, *et al.* (2009) Inhibition of the c-Abl-TAp63 pathway protects mouse oocytes from chemotherapy-induced death. Nat Med 15:1179-1185.

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH, *et al.*(2010) A draft sequence of the Neandertal genome. Science 328:710-722.

Hoh J, Jin S, Parrado T, Edington J, Levine AJ and Ott J (2002) The p53MH algorithm and its application in detecting p53-responsive genes. Proc Natl Acad Sci USA 99:8467-8472.

Hollstein M and Hainaut P (2010) Massively regulated genes: The example of TP53. J Pathol 220:164-173.

Hooper AT, Shmelkov SV, Gupta S, Milde T, Bambino K, Gillen K, Goetz M, Chavala S, Baljevic M, Murphy AJ, *et al.* (2009) Angiomodulin is a specific marker of vasculature and regulates vascular endothelial growth factor-A-dependent neoangiogenesis. Circ Res 105:201-208.

Hu W, Feng Z, Ma L, Wagner J, Rice JJ, Stolovitzky G and Levine AJ (2007a) A single nucleotide polymorphism in the MDM2 gene disrupts the oscillation of p53 and MDM2 levels in cells. Cancer Res 67:2757-2765.

Hu W, Feng Z, Teresky AK and Levine AJ (2007b) p53 regulates maternal reproduction through LIF. Nature 450:721-724.

Kang HJ, Feng Z, Sun Y, Atwal G, Murphy ME, Rebbeck TR, Rosenwaks Z, Levine AJ and Hu W (2009) Single-nucleotide polymorphisms in the p53 pathway regulate fertility in humans. Proc Natl Acad Sci USA 106:9761-9766.

Kay C, Jeyendran RS and Coulam CB (2006) p53 tumour suppressor gene polymorphism is associated with recurrent implantation failure. Reprod Biomed Online 13:492-496.

Kruse JP and Gu W (2009) Modes of p53 regulation. Cell 137:609-622.

Lain S and Lane D (2003) Improving cancer therapy by non-genotoxic activation of p53. Eur J Cancer 39:1053-1060.

Lane DP (1992) Cancer. p53, guardian of the genome. Nature 358:15-16.

Levine AJ, Tomasini R, McKeon FD, Mak TW and Melino G (2011) The p53 family: Guardians of maternal reproduction. Nat Rev Mol Cell Biol 12:259-265.

Lindsay J, McDade SS, Pickard A, McCloskey KD and McCance DJ (2011) Role of DeltaNp63gamma in epithelial to mesenchymal transition. J Biol Chem 286:3915-3924.

Lu WJ, Amatruda JF and Abrams JM (2009) p53 ancestry: Gazing through an evolutionary lens. Nat Rev Cancer 9:758-762.

Malkin D, Li FP, Strong LC, Fraumeni Jr JF, Nelson CE, Kim DH, Kassel J, Gryka MA, Bischoff FZ, Tainsky MA, *et al.* (1990) Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science 250:1233-1238.

Michael D and Oren M (2003) The p53-Mdm2 module and the ubiquitin system. Semin Cancer Biol 13:49-58.

Mills AA, Zheng B, Wang XJ, Vogel H, Roop DR and Bradley A (1999) p63 is a p53 homologue required for limb and epidermal morphogenesis. Nature 398:708-713.

Nedelcu AM and Tan C (2007) Early diversification and complex evolutionary history of the p53 tumor suppressor gene family. Dev Genes Evol 217:801-806.

Oh JE, Kim RH, Shin KH, Park NH and Kang MK (2011) DeltaNp63alpha protein triggers epithelial-mesenchymal transition and confers stem cell properties in normal human keratinocytes. J Biol Chem 286:38757-38767.

Phelps M, Darley M, Primrose JN and Blaydes JP (2003) p53-independent activation of the hdm2-P2 promoter through multiple transcription factor response elements results in elevated hdm2 expression in estrogen receptor alpha-positive breast cancer cells. Cancer Res 63, 2616-2623.

Plunkett J, Doniger S, Orabona G, Morgan T, Haataja R, Hallman M, Puttonen H, Menon R, Kuczynski E, Norwitz E, *et al.* (2011) An evolutionary genomic approach to identify genes involved in human birth timing. PLoS Genetics 7:e1001365.

Reich D, Green RE, Kircher M, Krause J, Patterson N, Durand EY, Viola B, Briggs AW, Stenzel U, Johnson PL, *et al.* (2010) Genetic history of an archaic hominin group from Denisova Cave in Siberia. Nature 468:1053-1060.

Ribeiro Junior CL, Arruda JT, Silva CT and Moura KK (2009) Analysis of p53 codon 72 gene polymorphism in Brazilian patients with endometriosis. Genet Mol Res 8:494-499.

Shan J, Brooks C, Kon N, Li M and Gu W (2008) Dissecting roles of ubiquitination in the p53 pathway. Ernst Schering Found Symp Proc 2008-1:127-136.

Shi H, Tan SJ, Zhong H, Hu W, Levine A, Xiao CJ, Peng Y, Qi XB, Shou WH, Ma RL, *et al.* (2009) Winter temperature and UV are tightly linked to genetic changes in the p53 tumor suppressor pathway in eastern Asia. Am J Hum Genet 84:534-541.

Stewart CL, Kaspar P, Brunet LJ, Bhatt H, Gadi I, Kontgen F and Abbondanzo SJ (1992) Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. Nature 359:76-79.

Straub WE, Weber TA, Schafer B, Candi E, Durst F, Ou HD, Rajalingam K, Melino G and Dotsch V (2010) The C-terminus of p63 contains multiple regulatory elements with different functions. Cell Death Dis 1:e5.

Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, Doerks T, Stark M, Muller J, Bork P, *et al.* (2011) The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res 39:D561-D568.

Tagliani-Ribeiro A, Oliveira M, Sassi AK, Rodrigues MR, Zagonel-Oliveira M, Steinman G, Matte U, Fagundes NJ and Schuler-Faccini L (2011) Twin Town in South Brazil: A Nazi's experiment or a genetic founder effect? PLoS One 6:e20328.

Tagliani-Ribeiro A, Paskulin DD, Oliveira M, Zagonel-Oliveira M, Longo D, Ramallo V, Ashton-Prolla P, Saraiva-Pereira ML, Fagundes NJ, Schuler-Faccini L, *et al.* (2012) High twinning rate in Cândido Godói: A new role for p53 in human fertility. Hum Reprod 27:2866-71.

Thomas M, Kalita A, Labrecque S, Pim D, Banks L and Matlashewski G (1999) Two polymorphic variants of wild-type p53 differ biochemically and biologically. Mol Cell Biol 19:1092-1100.

Toledo F, Krummel KA, Lee CJ, Liu CW, Rodewald LW, Tang M and Wahl GM (2006) A mouse p53 mutant lacking the proline-rich domain rescues Mdm4 deficiency and provides insight into the Mdm2-Mdm4-p53 regulatory network. Cancer Cell 9:273-285.

Tomasini R, Tsuchihara K, Wilhelm M, Fujitani M, Rufini A, Cheung CC, Khan F, Itie-Youten A, Wakeham A, Tsao MS, *et al.* (2008) TAp73 knockout shows genomic instability with infertility and tumor suppressor functions. Genes Dev 22:2677-2691.

Vousden KH and Prives C (2009) Blinded by the light: The growing complexity of p53. Cell 137:413-431.

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH, *et al.*(2010) A draft sequence of the Neandertal genome. Science 328:710-722.

Hoh J, Jin S, Parrado T, Edington J, Levine AJ and Ott J (2002) The p53MH algorithm and its application in detecting p53-responsive genes. Proc Natl Acad Sci USA 99:8467-8472.

Hollstein M and Hainaut P (2010) Massively regulated genes: The example of TP53. J Pathol 220:164-173.

Hooper AT, Shmelkov SV, Gupta S, Milde T, Bambino K, Gillen K, Goetz M, Chavala S, Baljevic M, Murphy AJ, *et al.* (2009) Angiomodulin is a specific marker of vasculature and regulates vascular endothelial growth factor-A-dependent neoangiogenesis. Circ Res 105:201-208.

Hu W, Feng Z, Ma L, Wagner J, Rice JJ, Stolovitzky G and Levine AJ (2007a) A single nucleotide polymorphism in the MDM2 gene disrupts the oscillation of p53 and MDM2 levels in cells. Cancer Res 67:2757-2765.

Hu W, Feng Z, Teresky AK and Levine AJ (2007b) p53 regulates maternal reproduction through LIF. Nature 450:721-724.

Kang HJ, Feng Z, Sun Y, Atwal G, Murphy ME, Rebbeck TR, Rosenwaks Z, Levine AJ and Hu W (2009) Single-nucleotide polymorphisms in the p53 pathway regulate fertility in humans. Proc Natl Acad Sci USA 106:9761-9766.

Kay C, Jeyendran RS and Coulam CB (2006) p53 tumour suppressor gene polymorphism is associated with recurrent implantation failure. Reprod Biomed Online 13:492-496.

Kruse JP and Gu W (2009) Modes of p53 regulation. Cell 137:609-622.

Lain S and Lane D (2003) Improving cancer therapy by non-genotoxic activation of p53. Eur J Cancer 39:1053-1060.

Lane DP (1992) Cancer. p53, guardian of the genome. Nature 358:15-16.

Levine AJ, Tomasini R, McKeon FD, Mak TW and Melino G (2011) The p53 family: Guardians of maternal reproduction. Nat Rev Mol Cell Biol 12:259-265.

Lindsay J, McDade SS, Pickard A, McCloskey KD and McCance DJ (2011) Role of DeltaNp63gamma in epithelial to mesenchymal transition. J Biol Chem 286:3915-3924.

Lu WJ, Amatruda JF and Abrams JM (2009) p53 ancestry: Gazing through an evolutionary lens. Nat Rev Cancer 9:758-762.

Malkin D, Li FP, Strong LC, Fraumeni Jr JF, Nelson CE, Kim DH, Kassel J, Gryka MA, Bischoff FZ, Tainsky MA, *et al.* (1990) Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science 250:1233-1238.

Michael D and Oren M (2003) The p53-Mdm2 module and the ubiquitin system. Semin Cancer Biol 13:49-58.

Mills AA, Zheng B, Wang XJ, Vogel H, Roop DR and Bradley A (1999) p63 is a p53 homologue required for limb and epidermal morphogenesis. Nature 398:708-713.

Nedelcu AM and Tan C (2007) Early diversification and complex evolutionary history of the p53 tumor suppressor gene family. Dev Genes Evol 217:801-806.

Oh JE, Kim RH, Shin KH, Park NH and Kang MK (2011) DeltaNp63alpha protein triggers epithelial-mesenchymal transition and confers stem cell properties in normal human keratinocytes. J Biol Chem 286:38757-38767.

Phelps M, Darley M, Primrose JN and Blaydes JP (2003) p53-independent activation of the hdm2-P2 promoter through multiple transcription factor response elements results in elevated hdm2 expression in estrogen receptor alpha-positive breast cancer cells. Cancer Res 63, 2616-2623.

Plunkett J, Doniger S, Orabona G, Morgan T, Haataja R, Hallman M, Puttonen H, Menon R, Kuczynski E, Norwitz E, *et al.* (2011) An evolutionary genomic approach to identify genes involved in human birth timing. PLoS Genetics 7:e1001365.

Reich D, Green RE, Kircher M, Krause J, Patterson N, Durand EY, Viola B, Briggs AW, Stenzel U, Johnson PL, *et al.* (2010) Genetic history of an archaic hominin group from Denisova Cave in Siberia. Nature 468:1053-1060.

Ribeiro Junior CL, Arruda JT, Silva CT and Moura KK (2009) Analysis of p53 codon 72 gene polymorphism in Brazilian patients with endometriosis. Genet Mol Res 8:494-499.

Shan J, Brooks C, Kon N, Li M and Gu W (2008) Dissecting roles of ubiquitination in the p53 pathway. Ernst Schering Found Symp Proc 2008-1:127-136.

Shi H, Tan SJ, Zhong H, Hu W, Levine A, Xiao CJ, Peng Y, Qi XB, Shou WH, Ma RL, *et al.* (2009) Winter temperature and UV are tightly linked to genetic changes in the p53 tumor suppressor pathway in eastern Asia. Am J Hum Genet 84:534-541.

Stewart CL, Kaspar P, Brunet LJ, Bhatt H, Gadi I, Kontgen F and Abbondanzo SJ (1992) Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. Nature 359:76-79.

Straub WE, Weber TA, Schafer B, Candi E, Durst F, Ou HD, Rajalingam K, Melino G and Dotsch V (2010) The C-terminus of p63 contains multiple regulatory elements with different functions. Cell Death Dis 1:e5.

Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, Doerks T, Stark M, Muller J, Bork P, *et al.* (2011) The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res 39:D561-D568.

Tagliani-Ribeiro A, Oliveira M, Sassi AK, Rodrigues MR, Zagonel-Oliveira M, Steinman G, Matte U, Fagundes NJ and Schuler-Faccini L (2011) Twin Town in South Brazil: A Nazi's experiment or a genetic founder effect? PLoS One 6:e20328.

Tagliani-Ribeiro A, Paskulin DD, Oliveira M, Zagonel-Oliveira M, Longo D, Ramallo V, Ashton-Prolla P, Saraiva-Pereira ML, Fagundes NJ, Schuler-Faccini L, *et al.* (2012) High twinning rate in Cândido Godói: A new role for p53 in human fertility. Hum Reprod 27:2866-71.

Thomas M, Kalita A, Labrecque S, Pim D, Banks L and Matlashewski G (1999) Two polymorphic variants of wild-type p53 differ biochemically and biologically. Mol Cell Biol 19:1092-1100.

Toledo F, Krummel KA, Lee CJ, Liu CW, Rodewald LW, Tang M and Wahl GM (2006) A mouse p53 mutant lacking the proline-rich domain rescues Mdm4 deficiency and provides insight into the Mdm2-Mdm4-p53 regulatory network. Cancer Cell 9:273-285.

Tomasini R, Tsuchihara K, Wilhelm M, Fujitani M, Rufini A, Cheung CC, Khan F, Itie-Youten A, Wakeham A, Tsao MS, *et al.* (2008) TAp73 knockout shows genomic instability with infertility and tumor suppressor functions. Genes Dev 22:2677-2691.

Vousden KH and Prives C (2009) Blinded by the light: The growing complexity of p53. Cell 137:413-431.

Internet Resources

Online Mendelian Inheritance in Man (OMIM). http://www.ncbi.nlm.nih.gov/OMIM (May 30, 2011).

International Agency for Research on Cancer (IARC) *TP53* Mutation Database. http://www-p53.iarc.fr/index.html (May 30, 2011).

The International *HapMap* Project. http://hapmap.ncbi.nlm.nih.gov/ (May 30, 2011).

The ALlele FREquency *Database* (*ALFRED).* http://alfred.med.yale.edu/ (May 30, 2011).

Gene Ontology AmiGO browser. http://amigo.geneontology.org/cgi-bin/amigo/go.cgi (March 28, 2011).

STRING 9.0 software. http://string-db.org/ (March 28, 2011).

LALIGN software. http://www.ch.embnet.org/software/LALIGN_form.html (March 28, 2011).

GeneDecks V3 software. http://www.genecards.org/ (June 30/2011).

Supplementary Material

The following online material is available for this article:

- Table S1 - Frequencies of deleterious polymorphisms available in the HapMap database.
- Table S2 - Eighteen genes which influence the TP53 Fertility Network.
- Table S3 - Amino acid conservation of the Fertility Network genes.
- Table S4 - The variation and conservation of the Fertility Network genes.
- Table S5 - Non-coding variation and damaging prediction of Fertility Network genes.

This material is available as part of the online article from http://www.scielo.br/gmb.

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

5. CAPÍTULO 3

5.1 *ESR1 RS9340799 IS ASSOCIATED WITH ENDOMETRIOSIS-RELATED INFERTILITY AND IN VITRO FERTILIZATION FAILURE*

Título do manuscrito: *ESR1* rs9340799 is associated with endometriosis-related infertility and *in vitro* fertilization failure

Autores: Paskulin DD, Cunha-Filho JS, Souza CA, Ashton-Prolla P

Situação: Submetido

Revista: *Human Reproduction*

Website: humrep.oxfordjournals.org/

Draft Manuscript For Review. Reviewers should submit their review at
http://mc.manuscriptcentral.com/humrep

ESR1 rs9340799 is associated with endometriosis-related infertility and in vitro fertilization failure

Journal:	*Human Reproduction*
Manuscript ID:	HUMREP-13-0520
Manuscript Type:	Original Article
Date Submitted by the Author:	29-Apr-2013
Complete List of Authors:	Paskulin, Diego; Universidade Federal do Rio Grande do Sul, Programa de Pos-Graduação em Genética e Biologia Molecular; Hospital de Clínicas de Porto Alegre, Centro de Pesquisa Experimental Cunha Filho, Joao Sabino; Universidade Federal do Rio Grande do Sul, Ob/Gyn; INSEMINE Human Reproduction Centre, Souza, Carlos; Universidade Federal do Rio Grande do Sul, Ob/Gyn Ashton-Prolla, Patricia; Universidade Federal do Rio Grande do Sul, Programa de Pos-Graduação em Genética e Biologia Molecular; INAGEMP, Instituto Nacional de Genetica Medica Populacional; Universidade Federal do Rio Grande do Sul, Ob/Gyn; Hospital de Clínicas de Porto Alegre, Centro de Pesquisa Experimental
Keywords:	ENDOMETRIOSIS, ESTROGEN, INFERTILITY
Specialty:	Infertility

SCHOLARONE™
Manuscripts

1 **TITLE**

2 ***ESR1* rs9340799 is associated with endometriosis-related infertility and *in vitro***
3 **fertilization failure**

4

5 **Running Title**

6 ESR1 rs9340799 in endometriosis and IVF failure

7

8 **Authors**

9 Paskulin DD[1,4] , Cunha-Filho JS[2], Souza CA[2], Ashton-Prolla P[1,3,4]

10

11 [1]Department of Genetics and Molecular Biology, Federal University of Rio Grande do
12 Sul, Porto Alegre, Brazil; Av. Bento Gonçalves, 9500 - Prédio 43323 M -
13 91501-970, Caixa Postal 15053, Porto Alegre, RS, Brazil.

14 [2]Department of Obstetrics and Gynecology, Federal University of Rio Grande do Sul,
15 Porto Alegre, Brazil; . Rua Ramiro Barcelos, 2350, Porto Alegre 90035-903, Brazil.

16 [3]Medical Genetics Service and National Institute of Science and Technology in
17 Populational Medical Genetics (INAGEMP), Hospital de Clinicas de Porto Alegre, Porto
18 Alegre, Brazil. Rua Ramiro Barcelos, 2350, Porto Alegre 90035-903, Brazil.

19 [4]Genomic Medicine Laboratory, Experimental Research Center, Hospital de Clinicas de
20 Porto Alegre, Porto Alegre, Brazil. Rua Ramiro Barcelos, 2350, Porto Alegre 90035-
21 903, Brazil. Tel: +55 51 3359 8010; Fax: +55 51 3359 8760.

22

23

24 **ABSTRACT**

25 *TP53* has a central role in human fertility by regulating, together with estrogen
26 receptors, the expression of leukemia inhibitory factor (LIF), a cytokine critical for
27 blastocyst implantation. Here we hypothesize that *TP53* signaling pathway gene
28 polymorphisms may be associated with endometriosis-related infertility and also with
29 recurrent *in vitro* fertilization (IVF) failure. We included 98 infertile women with
30 endometriosis, 115 infertile women with at least one IVF failure and also 134 fertile
31 women as controls. Taqman SNP assays were used for genotyping *LIF* (rs929271),
32 *MDM2* (rs2279744), *MDM4* (rs1563828), *USP7* (rs1529916), *TP63* (rs17506395), *TP73*
33 (rs4648551 and rs6695978) and *ESR1* (rs9340799 and rs2234693) polymorphisms.
34 The SNP *ESR1* rs9340799 was associated with endometriosis related infertility
35 (P<0.001) and also with recurrent IVF failure (P=0.018). After controlling for age,
36 infertile women with *ESR1* rs9340799 GG genotype presented 4 fold-increased risk of
37 endometriosis (OR 4.67, 95% CI 1.84-11.83, P=0.001) and 3 fold-increased risk of
38 recurrent IVF failure (OR 3.33, 95% CI 1.38-8.03, P=0.007). Our results demonstrate an
39 association between *ESR1* rs9340799 polymorphism and infertile women with
40 endometriosis and also with women who were submitted to IVF procedures and had no
41 blastocyst implantation.

42

43 **KEY WORDS**

44 Endometriosis, estrogen, infertility

45

46 **INTRODUCTION**

47 Tumor suppressor protein p53 plays many roles as a master coordinator of
48 cellular stress response and tumor suppression pathways (Vousden and Prives, 2009).
49 *TP53* interacts with a large number of proteins as hundreds of genes are up or down-
50 regulated by *TP53* as shown in recent genome-wide analyses (Smeenk *et al.*, 2008).
51 The expression of *TP53* is very complex, as it is kept at very low levels under normal
52 conditions and finely modulated by positive and negative regulation involving several
53 genes (Giaccia and Kastan, 1998; Haupt and Haupt, 2004). Special attention has been
54 given to the classic feedback loop of *TP53* and *MDM2-MDM4* complex (Haupt and
55 Haupt, 2004; Sarkari *et al.*, 2010; Shvarts *et al.*, 1996) and also to its family members
56 *TP63* and *TP73* (Belyi and Levine, 2009; Levrero *et al.*, 2000). Many studies have
57 associated alterations in *TP53* network genes with different types of cancers as well as
58 with embryogenesis, metabolism, innate immunity, and recently to fertility (Hu *et al.*,
59 2009; Vousden and Ryan, 2009).

60 Besides its tumor suppression activities, appears to have *TP53* has a central role
61 in human fertility, in part by regulating the expression of leukemia inhibitory factor (*LIF*)
62 (Hu *et al.*, 2007a), a secreted cytokine critical for blastocyst implantation (Stewart,
63 2007). *LIF* has been described as an important gene in differentiation, proliferation, and
64 cell survival pathways (Metcalf, 1992) and its expression is reduced in endometrium
65 from women with unexplained infertility (Laird *et al.*, 1997). *LIF* expression is also
66 controlled by estrogen and progesterone at the implantation stage (Feng *et al.*, 2011),
67 supporting the idea hypothesis of multifactorial control of *LIF* expression.

68 Functional single nucleotide polymorphisms (SNPs) that modulate the activity of
69 *TP53* and its target genes have been associated with different reproductive phenotypes,
70 including infertility, endometriosis and even high twinning rates (Kang *et al.*, 2009;
71 Paskulin *et al.*, 2012; Shi *et al.*, 2009; Tagliani-Ribeiro *et al.*, 2011). Endometriosis is a
72 common gynecological disorder responsible for pelvic pain and infertility and affects up
73 to 10% of premenopausal women and up to 50% of infertile women (Eskenazi and
74 Warner, 1997; Goldman and Cramer, 1990). Although the association between
75 endometriosis and infertility is well known, a biological or mechanistic link between the
76 two conditions remains elusive (Vercellini *et al.*, 2009).
77 A few studies have shown associations of estrogen receptor α gene (*ESR1*)
78 polymorphisms in women with endometriosis-related infertility (Hsieh *et al.*, 2007; Wang
79 *et al.*, 2012), but to our knowledge, no study has focused on infertile women who were
80 submitted to conventional *in-vitro* fertilization (IVF) procedures with unsuccessful
81 blastocyst implantations. Therefore, the hypothesis we explored in the present study
82 was that SNPs in genes of the *TP53* signaling pathway might be associated with
83 endometriosis-related infertility and recurrent *in vitro* fertilization failure.
84
85 **MATERIAL AND METHODS**
86 **Subjects**
87 Patients and subjects were invited to participate and signed a consent form at inclusion.
88 The research project was approved by the Hospital de Clinicas de Porto Alegre (HCPA)
89 Ethics Committee (GPPG 05-182; GPPG 09-430). Infertile patients with and without
90 endometriosis and controls were divided into three study groups as previously

91 described (Paskulin *et al.*, 2012). Infertility was defined as the inability of a couple to
92 achieve pregnancy after 1 year of regular unprotected sexual intercourse (ASRM,
93 1997). The IVF Failure Group consisted of 115 infertile women with at least one IVF
94 failure, submitted to conventional IVF with 35 years or less. Patients with endometriosis,
95 previous thyroid disease, positive anti- lupus or anti-cardiolipin antibodies and
96 trombophilias were excluded from our sample. Controlled ovarian hyperstimulation was
97 performed with the use of recombinant human FSH and pituitary suppression with
98 GnRh antagonist (fixed day-6 protocol). Ovulation was induced by 6500 IU recombinant
99 hCG when at least three follicles had reached a diameter of 17mm, and transvaginal
100 follicle aspiration was performed 36 h later under ultrasound guidance. Embryos were
101 classified according to the cumulative embryo classification, taking into account
102 cleavage speed, blastomere symmetry, extent of fragmentation and the presence or
103 absence of multinucleated blastomeres. The Endometriosis Group comprised 98
104 infertile women with minimal or mild endometriosis as diagnosed by laparoscopy
105 according to the classification proposed by the ASRM recruited at the Gynecology
106 Service of HCPA, in Southern Brazil. Other causes of infertility were excluded by
107 hysterosalpingography, sperm evaluation and hormonal measurements whenever
108 necessary. The Fertile Group consisted of 134 women with no history of infertility, who
109 already had two or more children without any difficulties or assisted reproduction and
110 underwent laparoscopy for tubal ligation at HCPA.

111 **Genotyping**

112 Genomic DNA was extracted from peripheral blood leukocytes using the Ilustra blood
113 genomic Prep Mini spin Kit (GE Healthcare, Piscataway, NJ, USA) as described by the

114 manufacturer. Genotyping was performed using Taqman assays (Applied Biosystems,
115 USA) (Table S1).

116 **Statistical Analysis**

117 Clinical features of women in all study groups were compared by t-test. Differences in
118 genotype distribution were assessed by chi-square analysis, which was also used to
119 test for Hardy–Weinberg equilibrium. Logistic regression analysis was carried out to
120 estimate the odds ratios with 95% confidence intervals in order to assess the influence
121 of *ESR1* rs9340799 genotypes in endometriosis-related infertility and IVF failure.
122 Statistical analyses were performed using the SPSS 20.0 statistical package. All
123 reported P-values are two-tailed and were considered statistically significant when 0.05
124 or less.

125

126 **RESULTS**

127 The clinical and demographic characteristics of the women enrolled in the study
128 are shown in Table 1. Mean age at recruitment was higher in the Fertile Group (42.6 ±
129 12.88 years) than in both the Endometriosis (32.87 ± 4.7 years) and IVF Failure
130 (31.65±3.24 years) groups since only women with 35 years or less were included in
131 these two latter groups. The population based fertile control women presented a mean
132 of 3.62 ± 1.94 pregnancies reflecting the average number of pregnancies in the normal
133 population from Southern Brazil. Both Endometriosis and IVF Failure presented low
134 frequencies of pregnancy, abortion and caesarean due to their infertility status. Patients
135 and healthy study subjects did not differ significantly regarding self-attributed skin color

136 as a self-denomination of 'white' color predominated in all study subgroups (Paskulin *et*
137 *al.*, 2012).
138 Hardy–Weinberg equilibrium was achieved for all SNPs in the three study groups
139 (P>0.05) (Table S2). Table 2 presents genotype frequencies of the SNPs included in the
140 study. No association was found between *LIF*, *MDM2*, *MDM4*, USP7, *TP73* and *TP63*
141 SNPs and endometriosis-related infertility or *in vitro* fertilization failure. However, a
142 strong association was found between the *ESR1* rs9340799 SNP and clinical
143 phenotype in both case groups (Endometriosis P<0.001 and IVF Failure P=0.018) when
144 compared with the Fertile Group. Interestingly, no association was found between *ESR1*
145 rs2234693 and the outcomes.
146 To evaluate the effects of the *ESR1* rs9340799 SNPs we carried out a logistic
147 regression analysis, controlled by age, with endometriosis-related infertility and IVF
148 failure as outcomes. Results are summarized in Table 3 and show a statistically
149 significant effect of AG (OR 2.67, 95% CI 1.49-4.78, P=0.001) and GG (OR 4.67, 95%
150 CI 1.84-11.83, P=0.001) genotypes with endometriosis related infertility. Regarding IVF
151 failure, genotype GG contributed significantly to the outcome as women with genotype
152 GG had 3 fold-increased risk of recurrent IVF failure (OR 3.33, 95% CI 1.38-8.03,
153 P=0.007).
154
155 **DISCUSSION**
156 In the present study we have analyzed nine common SNPs in *TP53* network
157 genes that might modulate fertility status in women with endometriosis or recurrent
158 failure of *in vitro* fertilization procedures. Our results demonstrate an association

159 between *ESR1* rs9340799 polymorphism with infertile women with endometriosis and
160 also with women who were submitted to IVF procedures and had no embryo
161 implantation.

162 *TP53* regulates maternal reproduction through the expression of *LIF* (Hu *et al.*,
163 2007b). At 12 days of pregnancy, *LIF* is expressed at high levels making the uterus
164 receptive to the blastocyst (Hu *et al.*, 2007b). Both *TP53* and estrogen are essential for
165 *LIF* expression in the endometrial glands, and impaired function of these proteins is
166 clearly associated with failure of blastocyst implantation. Different studies have
167 demonstrated that SNPs modulating the activity of *TP53*, and also in its regulators
168 *MDM2*, *MDM4* and *USP7*, are more frequent in IVF patients (Kang *et al.*, 2009; Paskulin
169 *et al.*, 2012). We have previously shown that *TP53* polymorphisms are associated with
170 both endometriosis related infertility and IVF failure in patients from Southern Brazil
171 (Paskulin *et al.*, 2012). Using the same cohort, we expanded the analysis to other genes
172 of the *TP53* signaling network, and in contrast with previous findings, our results
173 demonstrated no association of *MDM2*, *MDM4*, *USP7* and *LIF* polymorphisms with
174 endometriosis related infertility or IVF failure patients. Interestingly *TP63* and *TP73*
175 SNPs are associated with infertility only when women older than 35 years of age are
176 investigated (Feng *et al.*, 2011). *TP63* and *TP73* act maintaining cell integrity and
177 preventing aneuploidy which are classical mechanisms of infertility in older women. Our
178 results corroborate these findings, since only patients with 35 years or less were
179 included in our study, and we also point out that in our cohort no association was found
180 between *TP63* and *TP73* SNPs and endometriosis related infertility.

181 *LIF* is regulated by both *TP53* and estrogen. Estrogen signaling is mediated

182 through its nuclear receptor alpha (*ESR1*). Studies have demonstrated an association
183 between *ESR1* polymorphisms and endometriotic women with and without infertility
184 (M'Rabet *et al.*, 2012; Wang *et al.*, 2013), but to our knowledge, no study has evaluated
185 *ESR1* polymorphisms in recurrent IVF failure. Our results demonstrate an association
186 between *ESR1* rs9340799 polymorphism (also known as *ER*-α *XbaI*) and endometriosis
187 related infertility. The classification of endometriosis is changing from a local disorder to
188 a complex disease as new molecular mechanisms are being discovered (Bulun, 2009).
189 The endometriotic process is classified as an estrogen-dependent inflammatory disease
190 similar to cancer due to its capability to invade surrounding tissues, to promote
191 angiogenesis, inflammation and apoptosis in favor of the new endometriotic tissue
192 survival (Abrao *et al.*, 2006; Beliard *et al.*, 2004; Dmowski *et al.*, 2001; Dmowski *et al.*,
193 1989; Osteen and Sierra-Rivera, 1997; Taylor *et al.*, 2002). Estrogen production plays a
194 central role in the pathology of endometriosis enhancing the survival of the
195 endometriotic tissue, and together with prostaglandins and cytokines, mediating pelvic
196 pain and infertility (Bruner *et al.*, 1997; Ryan and Taylor, 1997). The fact that estrogen
197 inhibitors such as GnRh analogues, oral conceptives, and aromatase inhibitors are used
198 to reduce pelvic disease and pain also corroborate to the fact that estrogen signaling is
199 critical for endometriosis (Olive and Pritts, 2001). Estrogen receptor polymorphisms
200 have been associated with ovarian response to follicle stimulating hormone in IVF
201 patients (Boudjenah *et al.*, 2012), with poor responders to IVF (Anagnostou *et al.*,
202 2012), with IVF parameters such as the number of follicles and collected oocytes,
203 maturation, pregnancy rates and embryo quality in women with unexplained infertility
204 (Ayvaz *et al.*, 2009) and with the outcome of ovarian stimulation in IVF (Altmae *et al.*,

205 2012). This is the first time that an association between *ESR1* rs9340799 SNP and
206 recurrent failure of IVF is demonstrated. Remarkably, we did not find any association
207 between *ESR1* rs2234693 polymorphism (also known as *PvuII*) and endometriosis
208 related infertility or failure of IVF. Both rs9340799 (A-351G) and rs2234693 (C-397T)
209 SNPs are localized in intron 1 of the *ESR1* gene in chromosome 6q25 and are in
210 linkage disequilibrium (Ayvaz *et al.*, 2009).

211 In conclusion, our results reveal a potential novel candidate biomarker for the
212 diagnostic and prognostic assessment of endometriosis-related infertility and recurrent
213 IVF failure. Our results demonstrate a 4-fold increased risk of endometriosis and a 3-
214 fold increased risk of recurrent IVF failure in infertile women with *ESR1* rs9340799 GG
215 genotype.

216

217 **SUPPLEMENTARY DATA**

218 Supplementary data are available at http://humrep.oxfordjournal.org/.

219

220 **AUTHORS' ROLES**

221 DDP, PA-P, JSC-F designed the study. DDP performed the experiments and
222 drafted the manuscript. JSC-F and CABS contributed to data collection. All authors
223 participated in data analysis, manuscript revision and approved the final version.

224

225 **FUNDING**

226 The work of DDP was supported by fellowships from CAPES and CNPQ (Brazil).
227 The study was supported in part by grants from GlaxoSmithKline Oncology (Ethnic

228 Research Initiative Grant Award 2009), UK; CNPq to PA-P (Grant 307779 2009-2),
229 Brazil; FAPERGS-PPSUS (Grant 09/0103-0), FAPERGS PRONEX (Grant 10/0051-9)
230 and Fundo de Incentivo a Pesquisa e Eventos, Hospital de Clinicas de Porto Alegre
231 (GPPG 09-430), Brazil.

232

233 **CONFLICT OF INTEREST**

234 None declared

235

236 **REFERENCES**

237 Abrao MS, Podgaec S, Dias JA, Jr., Averbach M, Garry R, Ferraz Silva LF,Carvalho
238 FM. Deeply infiltrating endometriosis affecting the rectum and lymph nodes. *Fertil*
239 *Steril* 2006;**86**,543-7.
240 Altmae S, Reimand J, Hovatta O, Zhang P, Kere J, Laisk T, Saare M, Peters M, Vilo J,
241 Stavreus-Evers A*et al*. Research resource: interactome of human embryo
242 implantation: identification of gene expression pathways, regulation, and
243 integrated regulatory networks. *Mol Endocrinol* 2012;**26**,203-17.
244 Anagnostou E, Mavrogianni D, Theofanakis C, Drakakis P, Bletsa R, Demirol A, Gurgan
245 T, Antsaklis A,Loutradis D. ESR1, ESR2 and FSH receptor gene polymorphisms
246 in combination: a useful genetic tool for the prediction of poor responders. *Curr*
247 *Pharm Biotechnol* 2012;**13**,426-34.
248 ASRM. Revised American Society for Reproductive Medicine classification of
249 endometriosis: 1996. *Fertil Steril* 1997;**67**,817-21.

250 Ayvaz OU, Ekmekci A, Baltaci V, Onen HI,Unsal E. Evaluation of in vitro fertilization
251 parameters and estrogen receptor alpha gene polymorphisms for women with
252 unexplained infertility. *J Assist Reprod Genet* 2009;**26**,503-10.
253 Beliard A, Noel A,Foidart JM. Reduction of apoptosis and proliferation in endometriosis.
254 *Fertil Steril* 2004;**82**,80-5.
255 Belyi VA,Levine AJ. One billion years of p53/p63/p73 evolution. *Proc Natl Acad Sci U S*
256 *A* 2009;**106**,17609-10.
257 Boudjenah R, Molina-Gomes D, Torre A, Bergere M, Bailly M, Boitrelle F, Taieb S,
258 Wainer R, Benahmed M, de Mazancourt P*et al*. Genetic polymorphisms influence
259 the ovarian response to rFSH stimulation in patients undergoing in vitro
260 fertilization programs with ICSI. *PLoS One* 2012;**7**,e38700.
261 Bruner KL, Matrisian LM, Rodgers WH, Gorstein F,Osteen KG. Suppression of matrix
262 metalloproteinases inhibits establishment of ectopic lesions by human
263 endometrium in nude mice. *J Clin Invest* 1997;**99**,2851-7.
264 Bulun SE. Endometriosis. *N Engl J Med* 2009;**360**,268-79.
265 Dmowski WP, Ding J, Shen J, Rana N, Fernandez BB,Braun DP. Apoptosis in
266 endometrial glandular and stromal cells in women with and without
267 endometriosis. *Hum Reprod* 2001;**16**,1802-8.
268 Dmowski WP, Gebel HM,Rawlins RG. Immunologic aspects of endometriosis. *Obstet*
269 *Gynecol Clin North Am* 1989;**16**,93-103.
270 Eskenazi B,Warner ML. Epidemiology of endometriosis. *Obstet Gynecol Clin North Am*
271 1997;**24**,235-58.

272 Feng Z, Zhang C, Kang HJ, Sun Y, Wang H, Naqvi A, Frank AK, Rosenwaks Z, Murphy
273 ME, Levine AJ*et al.* Regulation of female reproduction by p53 and its family
274 members. *Faseb J* 2011.
275 Giaccia AJ,Kastan MB. The complexity of p53 modulation: emerging patterns from
276 divergent signals. *Genes Dev* 1998;**12**,2973-83.
277 Goldman MB,Cramer DW. The epidemiology of endometriosis. *Prog Clin Biol Res*
278 1990;**323**,15-31.
279 Haupt S,Haupt Y. Manipulation of the tumor suppressor p53 for potentiating cancer
280 therapy. *Semin Cancer Biol* 2004;**14**,244-52.
281 Hsieh YY, Wang YK, Chang CC,Lin CS. Estrogen receptor alpha-351 XbaI*G and -397
282 PvuII*C-related genotypes and alleles are associated with higher susceptibilities
283 of endometriosis and leiomyoma. *Mol Hum Reprod* 2007;**13**,117-22.
284 Hu W, Feng Z,Levine AJ. The regulation of human reproduction by p53 and its pathway.
285 *Cell Cycle* 2009;**8**,3621-2.
286 Hu W, Feng Z, Ma L, Wagner J, Rice JJ, Stolovitzky G,Levine AJ. A single nucleotide
287 polymorphism in the MDM2 gene disrupts the oscillation of p53 and MDM2 levels
288 in cells. *Cancer Res* 2007a;**67**,2757-65.
289 Hu W, Feng Z, Teresky AK,Levine AJ. p53 regulates maternal reproduction through LIF.
290 *Nature* 2007b;**450**,721-4.
291 Kang HJ, Feng Z, Sun Y, Atwal G, Murphy ME, Rebbeck TR, Rosenwaks Z, Levine
292 AJ,Hu W. Single-nucleotide polymorphisms in the p53 pathway regulate fertility in
293 humans. *Proc Natl Acad Sci U S A* 2009;**106**,9761-6.

294 Laird SM, Tuckerman EM, Dalton CF, Dunphy BC, Li TC,Zhang X. The production of
295 leukaemia inhibitory factor by human endometrium: presence in uterine flushings
296 and production by cells in culture. *Hum Reprod* 1997;**12**,569-74.
297 Levrero M, De Laurenzi V, Costanzo A, Gong J, Wang JY,Melino G. The p53/p63/p73
298 family of transcription factors: overlapping and distinct functions. *J Cell Sci*
299 2000;**113 (Pt 10)**,1661-70.
300 M'Rabet N, Moffat R, Helbling S, Kaech A, Zhang H,de Geyter C. The CC-allele of the
301 PvuII polymorphic variant in intron 1 of the alpha-estrogen receptor gene is
302 significantly more prevalent among infertile women at risk of premature ovarian
303 aging. *Fertil Steril* 2012;**98**,965-72 e1-5.
304 Metcalf D. Leukemia inhibitory factor--a puzzling polyfunctional regulator. *Growth*
305 *Factors* 1992;**7**,169-73.
306 Olive DL,Pritts EA. Treatment of endometriosis. *N Engl J Med* 2001;**345**,266-75.
307 Osteen KG,Sierra-Rivera E. Does disruption of immune and endocrine systems by
308 environmental toxins contribute to development of endometriosis? *Semin Reprod*
309 *Endocrinol* 1997;**15**,301-8.
310 Paskulin DD, Cunha-Filho JS, Souza CA, Bortolini MC, Hainaut P,Ashton-Prolla P.
311 TP53 PIN3 and PEX4 polymorphisms and infertility associated with
312 endometriosis or with post-in vitro fertilization implantation failure. *Cell Death Dis*
313 2012;**3**,e392.
314 Ryan IP,Taylor RN. Endometriosis and infertility: new concepts. *Obstet Gynecol Surv*
315 1997;**52**,365-71.

316 Sarkari F, Sheng Y,Frappier L. USP7/HAUSP promotes the sequence-specific DNA
317 binding activity of p53. *PLoS One* 2010;**5**,e13040.
318 Shi H, Tan SJ, Zhong H, Hu W, Levine A, Xiao CJ, Peng Y, Qi XB, Shou WH, Ma RL*et*
319 *al.* Winter temperature and UV are tightly linked to genetic changes in the p53
320 tumor suppressor pathway in Eastern Asia. *Am J Hum Genet* 2009;**84**,534-41.
321 Shvarts A, Steegenga WT, Riteco N, van Laar T, Dekker P, Bazuine M, van Ham RC,
322 van der Houven van Oordt W, Hateboer G, van der Eb AJ*et al.* MDMX: a novel
323 p53-binding protein with some functional properties of MDM2. *EMBO J*
324 1996;**15**,5349-57.
325 Smeenk L, van Heeringen SJ, Koeppel M, van Driel MA, Bartels SJ, Akkers RC,
326 Denissov S, Stunnenberg HG,Lohrum M. Characterization of genome-wide p53-
327 binding sites upon stress response. *Nucleic Acids Res* 2008;**36**,3639-54.
328 Stewart CL. Reproduction: the unusual suspect. *Nature* 2007;**450**,619.
329 Tagliani-Ribeiro A, Oliveira M, Sassi AK, Rodrigues MR, Zagonel-Oliveira M, Steinman
330 G, Matte U, Fagundes NJ,Schuler-Faccini L. Twin Town in South Brazil: a Nazi's
331 experiment or a genetic founder effect? *PLoS One* 2011;**6**,e20328.
332 Taylor RN, Lebovic DI,Mueller MD. Angiogenic factors in endometriosis. *Ann N Y Acad*
333 *Sci* 2002;**955**,89-100; discussion 118, 396-406.
334 Vercellini P, Somigliana E, Vigano P, Abbiati A, Barbara G,Crosignani PG. Surgery for
335 endometriosis-associated infertility: a pragmatic approach. *Hum Reprod*
336 2009;**24**,254-69.
337 Vousden KH,Prives C. Blinded by the Light: The Growing Complexity of p53. *Cell*
338 2009;**137**,413-31.

339 Vousden KH,Ryan KM. p53 and metabolism. *Nat Rev Cancer* 2009;**9**,691-700.

340 Wang HS, Wu HM, Cheng BH, Yen CF, Chang PY, Chao A, Lee YS, Huang HD,Wang
341 TH. Functional analyses of endometriosis-related polymorphisms in the estrogen
342 synthesis and metabolism-related genes. *PLoS One* 2012;**7**,e47374.

343 Wang W, Li Y, Maitituoheti M, Yang R, Wu Z, Wang T, Ma D,Wang S. Association of an
344 oestrogen receptor gene polymorphism in Chinese Han women with
345 endometriosis and endometriosis-related infertility. *Reprod Biomed Online*
346 2013;**26**,93-8.

347

348

Table 1 – Characteristic of controls and patients.

Characteristic*	Fertile n=134	Endometriosis n=98	P-value	IVF Failure n=115	P-value
Age	42.68 ± 12.88	32.87 ± 4.70	P<0.001	31.65 ± 3.24	P<0.001
Pregnancies	3.62 ± 1.94	0.34 ± 0.92	P<0.001	0.17 ± 0.51	P<0.001
Spontaneous Abortions	0.45 ± 1.04	0.16 ± 0.63	P=0.015	0.13 ± 0.39	P<0.001
Caesarean sections	0.65 ± 0.95	0.07 ± 0.33	P<0.001	0 ± 0	P=0.005

* (mean ±SD)

Table 2 – Genotype and allele frequencies of *TP53* signaling pathway gene polymorphisms.

		Fertile n (%)	Endometriosis n (%)	P-value*	IVF Failure n (%)	P- value**
***MDM2* rs2279744**	TT	57 (42.5)	41 (41.8)	0.824	48 (41.7)	0.918
	TG	64 (47.8)	45 (45.9)		54 (47)	
	GG	13 (9.7)	12 (12.2)		13 (11.3)	
	G	0.67	0.65	0.765	0.65	0.765
	T	0.33	0.35		0.35	
***MDM4* rs1563828**	CC	34 (25.4)	34 (34.7)	0.268	40 (34.8)	0.141
	CT	71 (53)	43 (42.9)		59 (51.3)	
	TT	29 (21.6)	21 (21.4)		16 (13.9)	
	C	0.52	0.57	0.477	0.6	0.254
	T	0.48	0.43		0.4	
***HAUSP* rs1529916**	CC	73 (54.5)	53 (54.1)	0.977	53 (46.1)	0.224
	CT	52 (38.8)	39 (39.8)		48 (41.7)	
	TT	9 (6.7)	6 (6.1)		14 (12.2)	
	C	0.74	0.74	1	0.67	0.277
	T	0.26	0.26		0.33	
***LIF* rs929271**	TT	57 (42.5)	47 (48)	0.702	46 (40)	0.784
	TG	60 (44.8)	39 (39.8)		51 (44.3)	
	GG	17 (12.7)	12 (12.2)		18 (15.7)	
	T	0.65	0.68	0.653	0.62	0.659
	G	0.35	0.32		0.38	
***TP63* rs17506395**	GG	7 (5.2)	7 (7.1)	0.655	8 (7)	0.096
	GT	39 (29.1)	32 (32.7)		47 (40.9)	
	TT	88 (65.7)	59 (60.2)		60 (52.2)	
	G	0.2	0.23	0.605	0.27	0.243
	T	0.8	0.77		0.73	
TP73	AA	25 (18.7)	21 (23.5)	0.359	17 (14.8)	0.066

rs4648551	AG	56 (41.8)	40 (44.9)		65 (56.5)	
	GG	53 (39.6)	37 (41.5)		33 (28.7)	
	A	0.4	0.42	0.773	0.43	0.666
	G	0.6	0.58		0.57	
***TP73* rs6695978**	AA	0 (0)	0 (0)	0.894	1 (0.9)	0.419
	AG	13 (9.7)	9 (9.2)		8 (7)	
	GG	121 (90.3)	89 (90.8)		106 (92.2)	
	A	0.05	0.04	0.733	0.04	1
	G	0.95	0.96		0.96	
***ESR1* rs9340799**	AA	71 (53)	27 (27.6)	<0.001	45 (39.1)	0.018
	AG	54 (40.3)	55 (56.1)		51 (44.3)	
	GG	9 (6.7)	16 (16.3)		19 (16.5)	
	A	0.73	0.55	0.008	0.61	0.071
	G	0.27	0.45		0.39	
***ESR1* rs2234693**	CC	27 (20.1)	18 (18.4)	0.861	17 (14.8)	0.105
	CT	69 (51.5)	54 (55.1)		51 (44.3)	
	TT	38 (28.4)	26 (26.5)		47 (40.9)	
	C	0.46	0.46	1	0.37	0.196
	T	0.54	0.54		0.63	

*** = chi-square analsys for the difference between FERTILE and ENDOMETRIOSIS.**

**** = chi-square analsys for the difference between FERTILE and IVF FAILURE.**

Table 3 – Logistic Regression model for *ESR1* rs9340799 using the Fertile group as reference.

	***ESR1* rs9340799**	**OR (95% CI)**	**P-value**
Endometriosis	AA	-	-
	AG	2.67 (1.49-4.78)	0.001
	GG	4.67 (1.84-11.83)	0.001
IVF Failure	AA	-	-
	AG	1.49 (0.87-2.54)	0.114
	GG	3.33 (1.38-8.03)	0.007

OR (95% CI) was calculated by binary logistic regression analysis. IVF FAILURE: women with recurrent failure of IVF; ENDOMETRIOSIS group: infertile women with minimal or mild endometriosis

Table S1 – Taqman SNP genotyping assays description.

Gene	rs #	SNP	Taqman Assay
MDM2	rs2279744	T/G	FAM-TCCCGCGCCGCAG
			VIC-CTCCCGCGCCGAAG
			F 5'-CGGGAGTTCAGGGTAAAGGT-3'
			R 5'-ACAGGCACCTGCGATCATC-3'
MDM4	rs1563828	C/T	C_9493064_10
USP7	rs1529916	C/T	C_9688119_1
LIF	rs929271	T/G	C_7545901_10
TP63	rs17506395	G/T	C_32460279_10
TP73	rs4648551	A/G	C_26892242_10
TP73	rs6695978	A/G	C_26892266_10
ESR1	rs9340799	A/G	C_3163591_10
ESR1	rs2234693	C/T	C_3163590_10

Table S2 – Hardy-Weinberg equilibrium calculations.

Gene	rs #	Fertile*	Endometriosis*	IVF Failure*
MDM2	rs2279744	0.980	1	0.932
MDM4	rs1563828	0.778	0.571	0.737
USP7	rs1529916	1	0.946	0.831
LIF	rs929271	0.989	0.687	0.826
TP63	rs17506395	0.631	0.666	0.955
TP73	rs4648551	0.344	0.279	0.261
TP73	rs6695978	0.839	0.891	0.214
ESR1	rs9340799	0.955	0.400	0.782
ESR1	rs2234693	0.913	0.557	0.837

(*) p-value for the Hardy-Weinberg equilibrium; IVF Failure: women with recurrent failure of IVF; Endopmetriosis: infertile women with minimal or mild endometriosis; Fertile: Fertile women

6. CAPÍTULO 4

Este capítulo compreende os resultados do trabalho realizado durante o período de estágio “sandwich” na Universidade do Colorado, EUA. Utilizando amostras pareadas de pacientes portadores da mutação *TP53* p.R337H (DNA extraído de tecido tumoral e sangue periférico), propusemos utilizar uma técnica de varredura genômica com 2,5 milhões de marcadores para identificação de regiões com variação do número de cópias para posterior validação das amplificações e deleções encontradas utilizando FISH (*fluorescent in situ hybridization)*.

6.1 ANÁLISE GENÔMICA DE TUMORES DE PACIENTES COM MUTAÇÕES GERMINATIVAS NO GENE *TP53*

Com o intuito de identificar novos marcadores moleculares associados a mutação *TP53* p.R337H, empregamos uma técnica de varredura de todo o genoma com a plataforma Illumina (Illumina Inc, CA, EUA). A utilização do *Human Omni 2.5-8 BeadChip,* que analisa aproximadamente 2,5 milhões de SNPs (detalhes no Anexo A – Figura 11.1), possibilita a identificação de regiões com variação do número de cópias (deleções e amplificações), a genotipagem dos SNPs analisados, a identificação de regiões de perda de heterozigosidade em alta resolução por todo o genoma. Incluímos neste estudo quatro grupos de pacientes e sujeitos com ou sem a mutação *TP53* p.R337H, conforme descrito abaixo (tabela 6.1.1).

Tabela 6.1.1 – Descrição das amostras incluídas para SNP Array.

ID	Amostra		*TP53* p.R337H
ACC-1	Sangue	Carcinoma Adrenocortical	GA
	Tumor		GA
ACC-2	Sangue	Carcinoma Adrenocortical	GA
	Tumor		GA
ACC-3	Sangue	Carcinoma Adrenocortical	GG
	Tumor		GG
ACC-4	Sangue	Carcinoma Adrenocortical	GG
	Tumor		GG
ACC-5	Sangue*	Carcinoma Adrenocortical	AA
ACC-6	Sangue*	Carcinoma Adrenocortical	GA
BC-1	Sangue	Câncer de Mama	GA
	Tumor		GA
BC-2	Sangue	Câncer de Mama	GA
	Tumor		GA

*: Devido à indisponibilidade de tecido tumoral, apenas DNA extraído de sangue periférico foi utilizado para análise nestes dois casos. *TP53* p.R337H (G→A).

Após extração do DNA com o kit *Illustra blood genomic Prep Spin* (*GE HealthCare*), as amostras de DNA foram quantificadas, diluídas a 50ng/μl, e hibridizadas aos Bead Chips conforme protocolo padrão do fabricante. Após a hibridização, os sinais de fluorescência foram importados para o programa GenomeBeadStudio da Illumina (Illumina Inc, CA, EUA) onde os genótipos foram gerados e transformados em intensidade normalizada (R) e razão de alelos (θ) utilizando as fórmulas de cálculo descritas por Pfeiffer *et al*. (Peiffer *et al*., 2006):

$$R = X_A + Y_B$$

$$\theta = (2/\pi) \times \tan^{-1} (Y_B/X_A),$$

onde X_A e Y_B são as intensidades dos sinais de fluorescência transformados a partir de um SNP específico. Nas amostras pareadas (sangue/tumor) estes

parâmetros foram transformados em duas variáveis: Log2($R_{\text{tumor ou teste}}$/$R_{\text{sangue ou referência}}$) chamado de *Log₂R Ratio* (LRR) e *B Allele Frequency* (BAF) demonstrando a $BAF_{\text{tumor ou teste}}$ e $BAF_{\text{sangue ou referência}}$. Analisando LRR, qualquer desvio do número zero em determinada região evidencia uma mudança no número de cópias. BAF é uma medida normalizada das intensidades dos sinais de dois alelos alternativos, A e B. Desvios dos valores esperados (0, 0,5 e 1 representando os genótipos AA, AB e BB respectivamente) é um indício de alteração cromossômica. A variável Delta-θ é gerada através da diferença entre $\theta_{\text{tumor ou teste}}$ e $\theta_{\text{sangue ou referência}}$ conforme cálculo a seguir:

$$\text{Delta } \theta = |\theta_{\text{tumor ou teste}} - \theta_{\text{sangue ou referência}}|.$$

O valor de Delta-θ não sofre alteração em uma região normal (Delta-θ = 0), porém no caso de uma alteração somática em um *locus* específico, Delta-θ apresentará valores positivos (0-0,5). Delta-θ é analisado primeiramente, pois assinala a existência de uma alteração, e subsequentemente, LRR e BAF são utilizados para a compreensão da natureza da alteração previamente identificada pelo Delta-θ. A visualização dos dados foi realizada utilizando o programa *Partek Genomic Studio 6.6* (*Partek Incorporated*, MO, EUA) que também foi utilizado para delinear estatisticamente as regiões alteradas através de análise de segmentação utilizando o algoritmo *Circular Binary Segmentation* (Olshen *et al.*, 2004) para LRR e também para Delta-θ.

A seguir estão apresentados diagramas gráficos com a representação das regiões identificadas como tendo uma variação do número de cópias (deleções e amplificações) e perda de heterozigosidade, identificadas pela análise dos resultados do SNP Array em cada paciente. Utilizando o programa de análise de dados em grande escala Partek Genomic Studio 6.6 delimitamos cada região afetada, identificando a coordenada genômica de início e fim da alteração. Possíveis genes de interesse localizados nas respectivas regiões foram identificados utilizando o software GeneDecks situado no website *GeneCards.org* (http://www.genecards.org/GeneDecks).

Paciente ACC-1

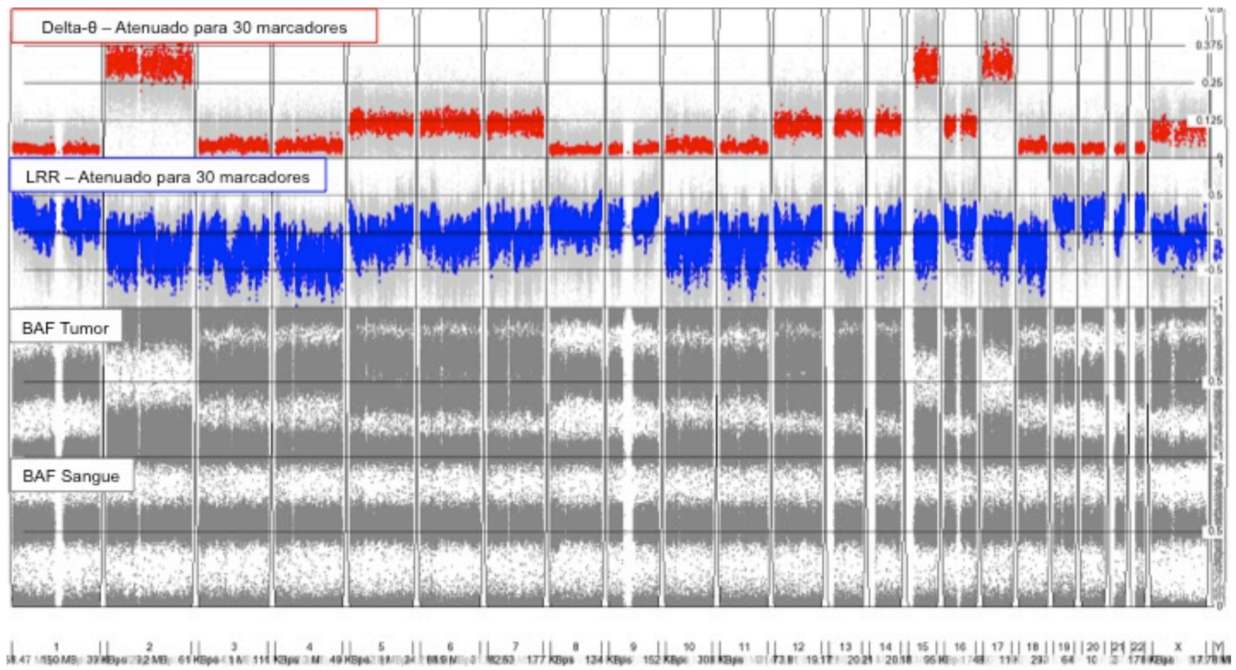


Figura 6 – Resultado do SNP Array por todos os cromossomos do Paciente ACC-1. Delta-θ (vermelho); LRR (azul); BAF Tumor; BAF Sangue.

Tabela 6.1.2 – Descrição de regiões com variação no número de cópias no paciente ACC-1 na análise comparativa entre DNA extraído de tecido tumoral e sangue periférico.

Crom	**Início**	**Fim**	**Banda**	**Alteração**	**P**	**Associações Destacadas***
2	38938	96057070	2p25.3-2q11.1	Del	4.55E-24	Câncer Coloretal: *BUB1.*
2	188804140	218695320	2q32.1-2q35	Del	7.98E-13	Câncer Gástrico: *CASP10*; Apoptose: *CASP8*; Câncer de Mama: *BARD1.*
12	190052	27415505	12p13.3-12p11.23	Amp	1.42E-20	Câncer de Mama e Câncer de Pulmão: KRAS.
16	94036	90141356	16p13.3-16q24.3	Amp	4.81E-22	Crescimento Celular: *MT1A, MT1B, MT1E, MT1F.*
17	66283	56308416	17p13.3-17q22	Del	7.48E-16	Câncer de Mama e Ovário: *ERBB2, BRAC1;* Li-Fraumeni: *TP53*.

*: Associações destacadas pelo GeneDecks (http://www.genecards.org/?path=/GeneDecks); P: Valor de P calculado utilizando o algoritmo de Segmentação Binária Circular; Del: Deleção; Amp: Amplificação.

Paciente ACC-2

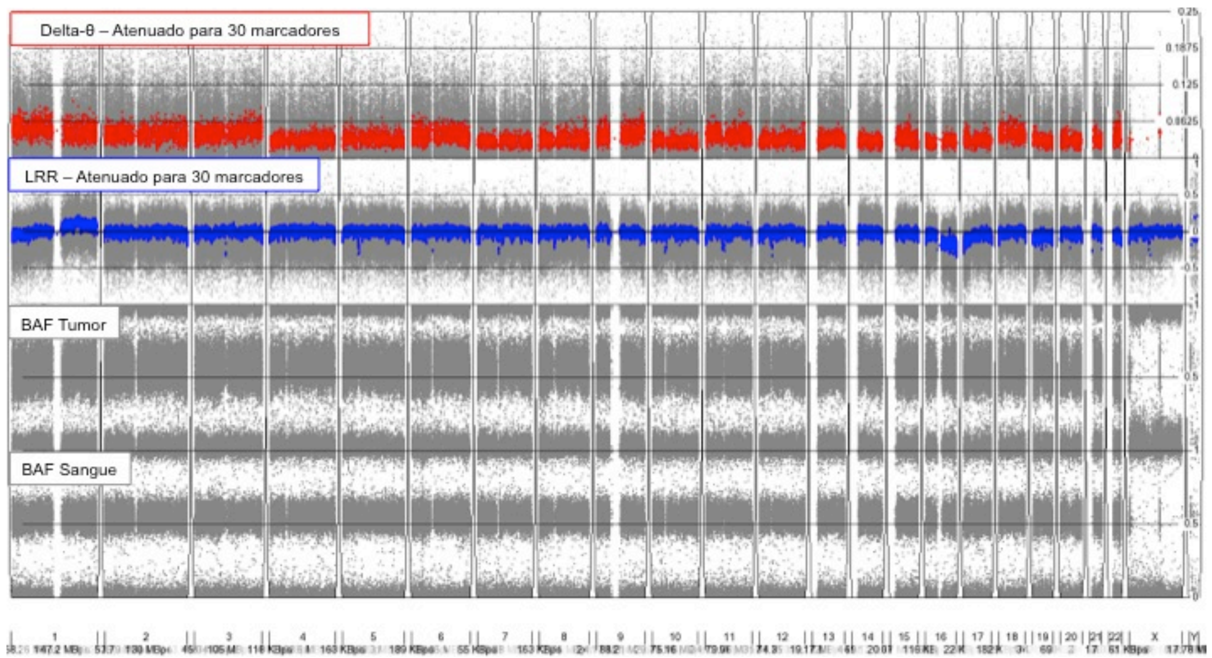


Figura 7 – Resultado do SNP Array por todos os cromossomos do Paciente ACC-2. Delta-θ (vermelho); LRR (azul); BAF Tumor; BAF Sangue.

Tabela 6.1.3 – Descrição de regiões com variação no número de cópias no paciente ACC-2 na análise comparativa entre DNA extraído de tecido tumoral e sangue periférico.

Crom	Início	Fim	Banda	Alteração	P	Associações Destacadas*
2	69431	242947986	2p25.3-2q37.3	R/A	7.98E-13	Câncer de Pulmão: *ALK.*
9	88406	141077353	9p24.3-9q34.3	R/A	8.69E-12	Leucemia: *ABL1.*
11	199153	97408529	11p15.5-11q22.1	R/A	7.01E-45	Tumor Adrenocortical e Tumor de Wilms: *CDKN1C.*

*: Associações destacadas pelo GeneDecks (http://www.genecards.org/?path=/GeneDecks); P: Valor de P calculado utilizando o algoritmo de Segmentação Binária Circular; R/A: Ruído ou Região Anormal; Amp: Amplificação.

Paciente ACC-3

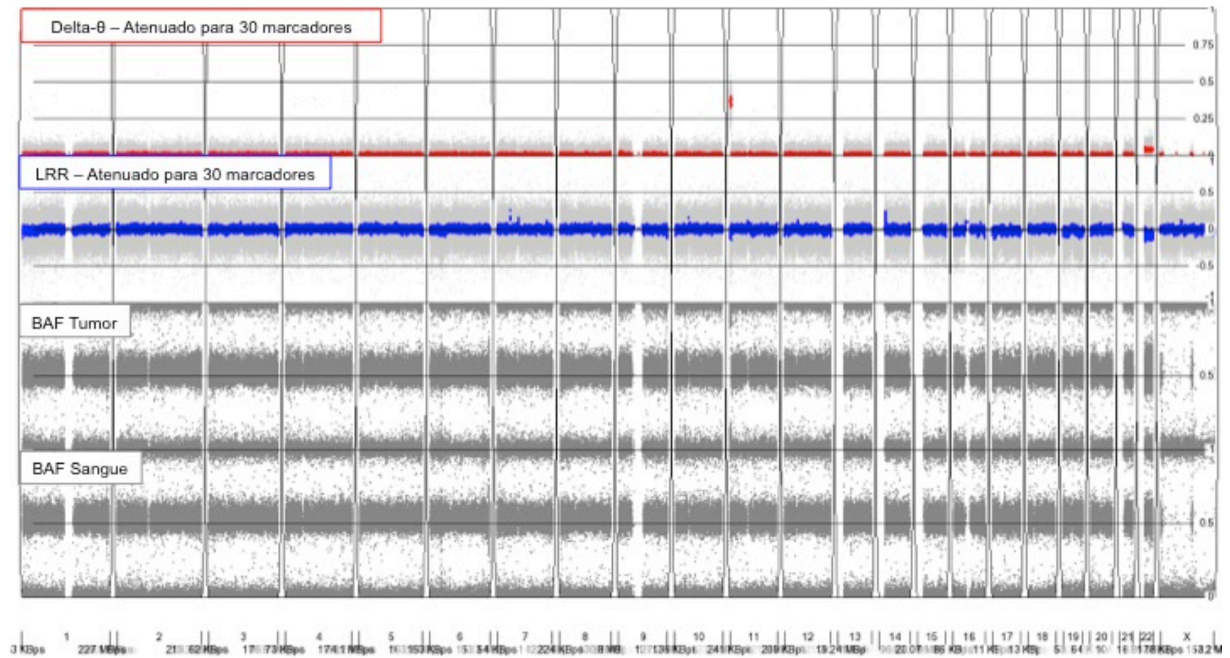


Figura 8 – Resultado do SNP Array por todos os cromossomos do Paciente ACC-3. Delta-θ (vermelho); LRR (azul); BAF Tumor; BAF Sangue.

Tabela 6.1.4 – Descrição de regiões com variação no número de cópias no paciente ACC-3 na análise comparativa entre DNA extraído de tecido tumoral e sangue periférico.

Crom	Início	Fim	Banda	Alteração	P	Associações Destacadas*
11	241432	4059198	11p15.5 - 11p15.4	CN-LOH	7.48E-16	Carcinoma Adrecortical: *ACCS, ADM, CALCA, CARS, IGF2, NAP1L4, NUP98, PHLDA2, TSPAN32, TSSC4*. Tumor de Wilms: *BDNF, CARS, CDKN1C, H19, IGF2, TSSC4.*
22	16464422	51208538	22q11.1 - 22q13.33	LOH	7.35E-09	Li-Fraumeni, Câncer de Mama e Câncer Coloretal: *CHEK2*; Sarcoma de Ewing: *EWS*.

*: Associações destacadas pelo GeneDecks (http://www.genecards.org/?path=/GeneDecks); P: Valor de P calculado utilizando o algoritmo de Segmentação Binária Circular; LOH: *Loss of Heterozygosity;* CN-LOH: *Copy Neutral LOH.*

Paciente ACC-4

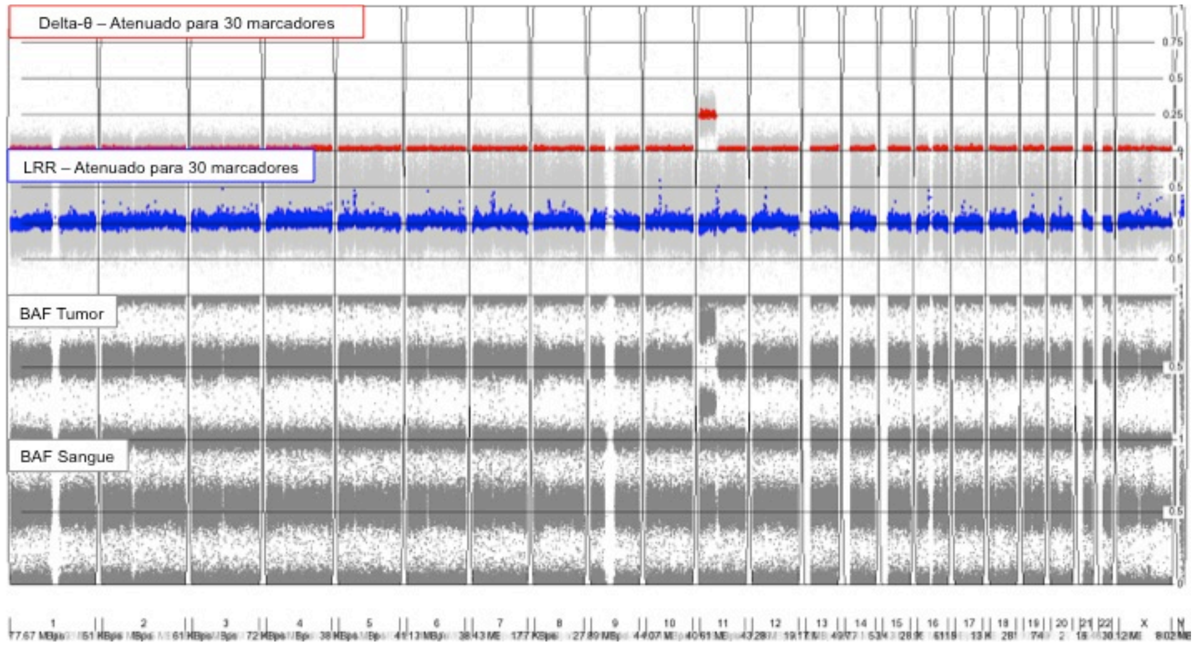


Figura 9 – Resultado do SNP Array por todos os cromossomos do Paciente ACC-4. Delta-θ (vermelho); LRR (azul); BAF Tumor; BAF Sangue.

Tabela 6.1.5 – Descrição de regiões com variação no número de cópias no paciente ACC-4 na análise comparativa entre DNA extraído de tecido tumoral e sangue periférico.

Crom	Início	Fim	Banda	Alteração	P	Associações Destacadas*
11	202547	47995984	11p15.5 - 11p11.2	CN-LOH	3.76E-13	Carcinoma Adrenocortical: *ACCS, ADM, CALCA, CARS, IGF2, NAP1L4, NUP98, PHLDA2, SLC22A18, TSPAN32, TSSC4*. Tumor de Wilms: *BDNF, CARS, CDKN1C, H19, IGF2, TSSC4*.

*: Associações destacadas pelo GeneDecks (http://www.genecards.org/?path=/GeneDecks); P: Valor de P calculado utilizando o algoritmo de Segmentação Binária Circular; CN-LOH: *Copy Neutral Loss of Heterozygosity*.

Paciente ACC-5

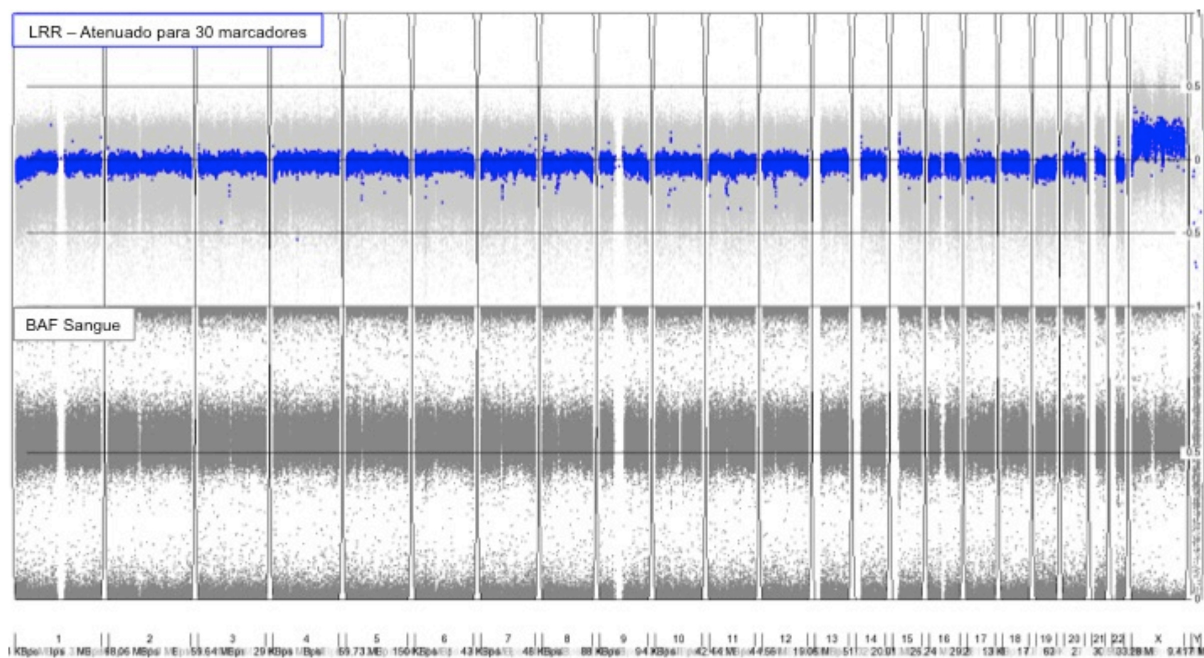


Figura 10 – Resultado do SNP Array por todos os cromossomos do Paciente ACC-1. LRR (azul); BAF Sangue.

Tabela 6.1.6 – Descrição de regiões com variação no número de cópias no paciente ACC-5 na análise comparativa entre DNA extraído de sangue periférico e o banco de dados do programa *GenomeBeadStudio* (*Illumina* Inc, CA, EUA).

Crom	Início	Fim	Banda	Alteração	P	Associações Destacadas*
10	67489804	76861491	10q21.3-q22.1	LOH	7.22E-07	Controle do Ciclo Celular e Apoptose: *TET1, CCAR1, LZP.*
15	22632789	23272533	15q11.2	Amp	2.16E-14	Síndrome Prader-Willi: *NIPA2.*
17	3816737	8103307	17p13.2-p13.1	CN-LOH	5.87E-06	Li-Fraumeni: *TP53.*
20	25567497	26018000	20q11.2	CN-LOH	1.40E-38	Câncer de Mama: *BASE.*

*: Associações destacadas pelo GeneDecks (http://www.genecards.org/?path=/GeneDecks); P: Valor de P calculado utilizando o algoritmo de Segmentação Binária Circular; LOH: *Loss of Heterozygosity*; CN-LOH: *Copy Neutral LOH* ;Amp: Amplificação.

Paciente ACC-6

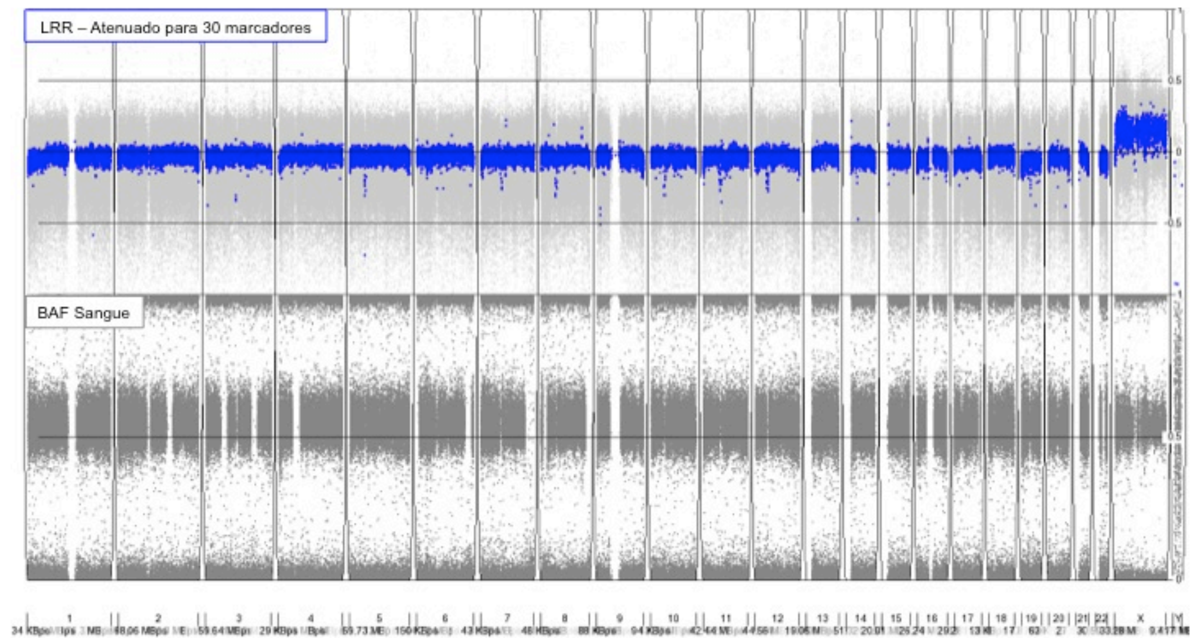


Figura 11 – Resultado do SNP Array por todos os cromossomos do Paciente ACC-6. LRR (azul); BAF Sangue.

Tabela 6.1.7 – Descrição de regiões com variação no número de cópias no paciente ACC-6 na análise comparativa entre DNA extraído de sangue periférico e o banco de dados do programa GenomeBeadStudio (Illumina Inc, CA, EUA).

Crom	Início	Fim	Banda	Alteração	P	Associações Destacadas*
11	199153	6245803	11p15.5-11p15.4	Amp	7.35E-09	Carcinoma Adrenocortical: *ACCS, ADM, CALCA, CARS, IGF2, NAP1L4, NUP98, PHLDA2, SLC22A18, TSPAN32, TSSC4*. Tumor de Wilms: BDNF, *CARS, CDKN1C, H19, IGF2, TSSC4.*
12	25019237	27450854	12p12.1-12p11.23	Amp	2.36E-22	Câncer de Pulmão: *BHLHE41, CASC1, KRAS, LRMP, SSPN*.

*: Associações destacadas pelo GeneDecks (http://www.genecards.org/?path=/GeneDecks); P: Valor de P calculado utilizando o algoritmo de Segmentação Binária Circular; Amp: Amplificação.

Paciente BC-1

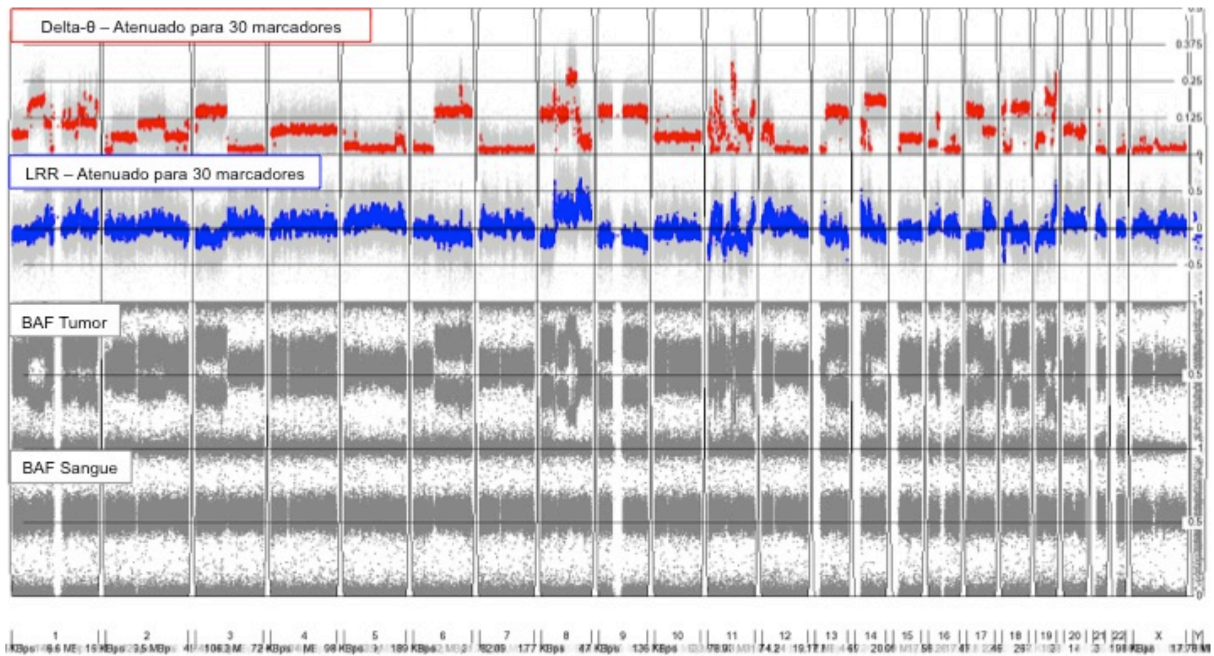


Figura 12 – Resultado do SNP Array por todos os cromossomos do Paciente BC-1. Delta-θ (vermelho); LRR (azul); BAF Tumor; BAF Sangue.

Tabela 6.1.8 – Descrição de regiões com variação no número de cópias no paciente BC-1 na análise comparativa entre DNA extraído de tecido tumoral e sangue periférico.

Crom	Início	Fim	Banda	Alteração	P	Associações Destacadas*
1	169294972	185712088	1q24.2-1q25.3	Amp	2.80E-08	Leucemia: *ABL2*
4	72048	10380803	4p16.3-4p16.1	Amp	5.47E-06	Síndrome de Wolf-Hirschhorn: *FGFRL1, LETM1, MFSD10, MSX1, NOP14, SLBP, TNIP2, WFS1, WHSC1, ZNF141.*
5	151418702	16887264	5q33.1-5q34	Amp	3.87E-07	Câncer de Mama: *HMMR*
9	46587	141077353	9p24.3-9q34.3	Del	2.19E-36	Leucemia: *ABL1*
11	199153	6245803	11p15.5-11p15.4	Del	1.51E-38	Carcinoma Adrenocortical: *ACCS, ADM, CALCA, CARS, IGF2, NAP1L4, NUP98, PHLDA2, SLC22A18, TSPAN32, TSSC4.* Tumor de Wilms: *BDNF, CARS, CDKN1C, H19, IGF2, TSSC4.*
12	25019854	27330939	12p12.1-12p11.23	Amp	7.92E-35	Câncer de Pulmão: *BHLHE41, CASC1, KRAS, LRMP, RASSF8, SSPN*.
16	72091118	76319482	16q22.2-16q23.1	Del	2.25E-14	Metabolismo do Colesterol: *DAT1, BCAR1, CFDP1, CHST5, CTRB1, FA2H, GLG1, KARS, LDHD, MLKL, PSMD7, TERF2IP, WDR59, ZFP1, ZNRF1*.
17	51088	46614695	17p13.3-17q21.32	Del	3.04E-09	Li-Fraumeni: *TP53*. Câncer de Mama e Ovário: *AARSD1, ARL4D, BRCA1, CBX1, CSF3, DUSP3, ERBB2, HSD17B1, IGFBP4, ITGB3, JUP, KRT9, STAT3, THRA, TP53, WNT3.*

*: Associações destacadas pelo GeneDecks (http://www.genecards.org/?path=/GeneDecks); P: Valor de P calculado utilizando o algoritmo de Segmentação Binária Circular; Amp: Amplificação; Del: Deleção.

Paciente BC-2

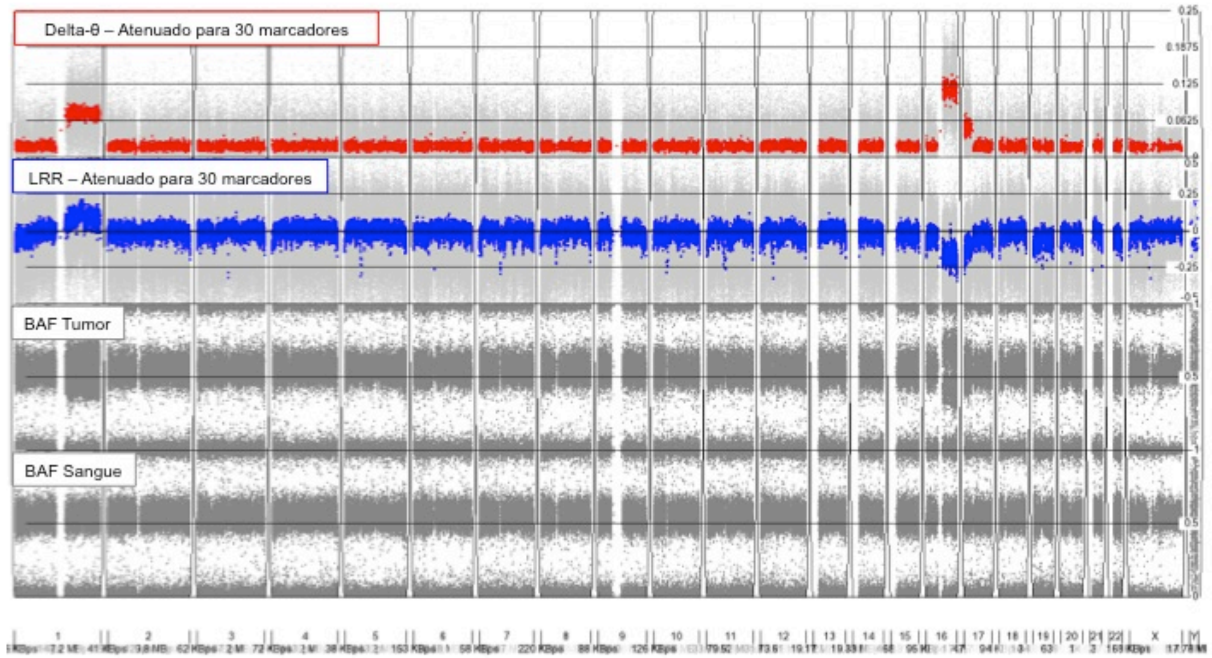


Figura 13 – Resultado do SNP Array por todos os cromossomos do Paciente BC-2. Delta-θ (vermelho); LRR (azul); BAF Tumor; BAF Sangue.

Tabela 6.1.9 – Descrição de regiões com variação no número de cópias no paciente BC-2 na análise comparativa entre DNA extraído de tecido tumoral e sangue periférico.

Crom	Início	Fim	Banda	Alteração	P	Associações Destacadas*
1	145249242	249174683	1q21.1-1q44	Amp	1.82E-44	Leucemia: *ABL2.*
16	40822575	90163276	16q11.2-16q24.3	Del	2.08E-16	Crescimento Celular: *PMT1A*, *MT1B*; Wnt.
17	3587	21791564	17p13.3-17p11.2	Del	3.24E-20	Síndrome de Li-Fraumeni: *TP53*.

*: Associações destacadas pelo GeneDecks (http://www.genecards.org/?path=/GeneDecks); P: Valor de P calculado utilizando o algoritmo de Segmentação Binária Circular; Amp: Amplificação; Del: Deleção.

6.2 VALIDAÇÃO DE REGIÕES GENÔMICAS ALTERADAS EM PACIENTES PORTADORES DE MUTAÇÕES GERMINATIVAS NO GENE *TP53* UTILIZANDO HIBRIDIZAÇÃO *IN SITU* FLUORESCENTE (FISH)

Após a identificação de regiões alteradas utilizando o *Human Omni 2.5-8 BeadChip* da Illumina procuramos validar as regiões comumente alteradas utilizado *Enumeration Fluorescent in situ Hybridization (FISH) Probes* em amostras de tecidos fixados em formalina e embebidos em parafina. Utilizando o algoritmo de análise de segmentação previamente descrito, identificamos três regiões comumente alteradas em pelo menos dois pacientes para construção, validação e aplicação de sondas de FISH conforme tabela 6.2.1.

Tabela 6.2.1 – Genes e Clones de BAC selecionados para construção de sondas.

Região	*SNP Array*	Clone de BAC	Gene	Associações
1q21.1-q44	Amplificação	RP11-689C20	*ABL2*	Leucemia, Câncer de Mama, Câncer de Pulmão (Gil-Henn *et al.*, 2012; Zhou *et al.*, 2012).
11p15.5-p11.2	Deleção	RP11-494F4	*CDKN1C*	Tumor de Willms, Tumor Adrenocortical (Kavanagh *et al.*, 2011).
12p13.33-p11.12	Amplificação	RP11-1119I8	*KRAS*	Câncer de Pulmão, Câncer Coloretal (Eklof *et al.*, 2013; Stella *et al.*, 2013).

Para criação das sondas selecionamos genes representativos para as regiões selecionadas (genes *ABL2, CDKN1C* e *KRAS*) obedecendo aos seguintes critérios: 1- Genes descritos no *Online Mendelian Inheritance in Man (OMIM)* (http://www.ncbi.nlm.nih.gov/omim/); 2- Genes Supressores Tumorais e Oncogenes; 3- Genes com literatura sugerindo associação com tumores diversos; 4- Genes com clones de BAC disponíveis no *BACPAC Resources Center* do *Children's Hospital Oakland Research Institute (CHORI)* em Oakland, California, EUA. Clones de BAC foram selecionados especificamente para as regiões de interesse conforme a tabela 6.2.1, foram cultivados em mini

culturas, e colônias únicas foram utilizadas para extração de DNA usando o *QIAamp DNA Mini Kit* (Qiagen, Hilden, Alemanha). O DNA genômico foi então submetido à amplificação genômica utilizando-se *REPLI-g Midi Kit* (Qiagen, Hilden, Alemanha) e 1 µg de cada clone de BAC extraído foi marcado em vermelho com sulforodamina 101 ácido clorídrico (*Texas-Red* ou *SpectrumRed* ou SR) utilizando-se o kit *Vysis Nick Translation* (*Abbott Laboratories*, Illinois, EUA). Adicionalmente, sondas centroméricas marcadas em verde com isotiocianato de fluoresceína (FITC ou *SpectrumGreen* ou SG) comercialmente disponíveis para os cromossomos 1, 11 e 12 foram adquiridas (Kreatech, Amsterdam, Holanda). Para validação e mapeamento de cada sonda criada, a linhagem celular linfoblastóide humana (2n) AG09391 (*Coriell Institute*) foi utilizada. As sondas alvo criadas (marcadas em vermelho) foram hibridizadas em conjunto com as sondas centroméricas (marcadas em verde) em lâminas preparadas com suspensão de células com a linhagem celular AG09391. Por final o corante fluorescente DAPI (4',6-diamidino-2-fenilindol) foi utilizado para contraste. As análises foram realizadas em microscópio de epifluorescência utilizando filtros designados para azul (DAPI), verde (FITC) e vermelho (Texas-Red). Imagens foram capturadas para cada filtro e depois sobrepostas utilizando o programa *CytoVision* (Leica Microsystems, Buffalo Grove, IL, EUA). Para mapeamento, 20 metáfases cariotipicamente normais foram avaliadas, e para validação qualitativa da sonda, 100 núcleos interfásicos (2n) foram utilizados. Os clones RP11-689C20, RP11-494F4, RP11-1119I8 marcaram respectivamente nos cromossomos 1q25.2 (Figura 14), 11p15.4 e 12p12.1 conforme o esperado.

Para análise em amostras de tecidos fixados em formalina e embebidos em parafina pelo menos 50 núcleos foram identificados, analisados, pontuados e fotografados. Para cada núcleo analisado, o número de sinais verdes e vermelhos foram anotados, e a razão entre as médias de sinais vermelho e verdes (SR/SG) foi utilizada para avaliação do número de cópias da região conforme a tabela 6.2.2.

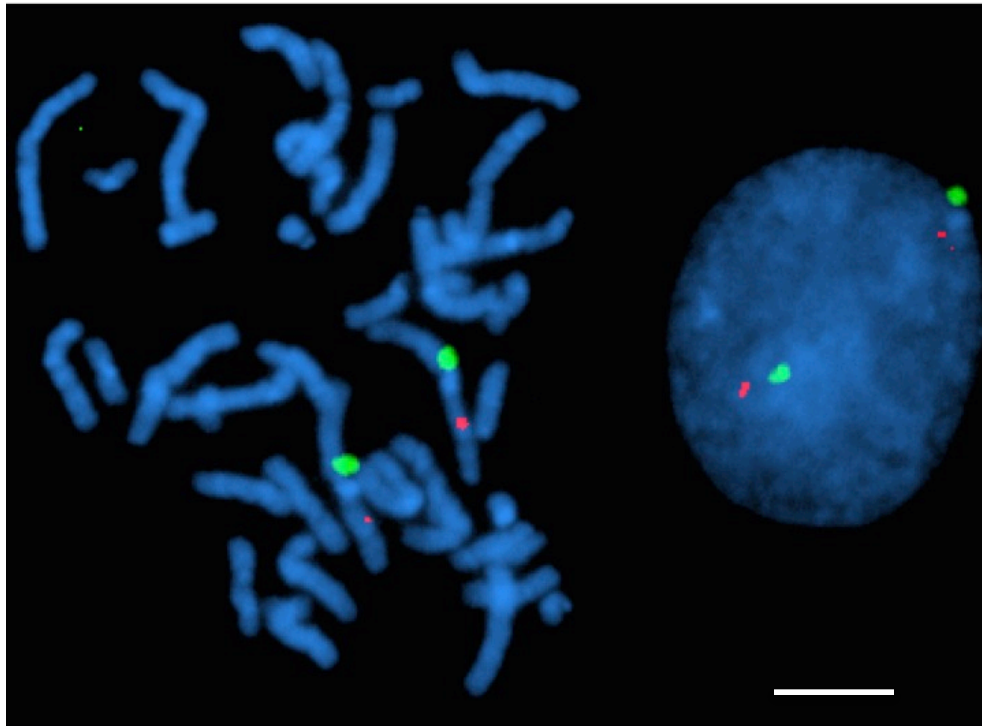


Figura 14 – Validação da sonda ABL2(SR)/CEN1(SG) em linhagem celular AG09391 mapeando corretamente em 1q25.2. Escala 10µm.

Tabela 6.2.2 - Classificação para avaliação de número de cópias em *Enumeration Fluorescent in situ Hybridization.*

Classificação	**SR/SG***
Perda	razão < 0,8
Normal	0,8 < razão <1,2
Ganho	1,2 < razão < 2,0
Amplificação	razão > 2

*: Razão obtida pela divisão entre a média de sinais vermelhos (SR) e de sinais verdes (SG).

Primeiramente, com o intuito de validar os achados encontrados com a varredura genômica, as análises foram efetuadas nos oito pacientes com resultados previamente obtidos a partir do SNP Array, representados na tabela 6.2.3.

Tabela 6.2.3 – Comparação entre resultados identificados por SNP Array e os resultados encontrados utilizando FISH.

		SNP Array			FISH		
ID	***TP53* R337H**	**1q21.1-q44**	**11p15.5-p11.2**	**12p13.33-p11.12**	***ABL2*/CEP1**	***CDKN1C*/CEP11**	***KRAS*/CEP12**
ACC-1	GA	N	N	Amp	N	N	Ganho
ACC-2	GA	Amp	R/A	R/A	N	N	N
ACC-3	GG	N	CN-LOH	N	N	N	N
ACC-4	GG	N	CN-LOH	N	N	N	N
ACC-5	AA	N	N	N	-	-	-
ACC-6	GA	R/A	Amp	Amp	Amp	Amp	Ganho
BC-1	GA	Amp	Del	Amp	Ganho	Perda	Ganho
BC-2	GA	N	Del	Amp	N	Perda	Ganho

N: Normal; Amp: Amplificação; P: Perda; Del: Deleção; -: falha da técnica; R/A: ruído/anormal devido a dificuldades durante o processamento das amostras e da baixa qualidade do DNA utilizado, algumas regiões apresentaram resultados inconsistentes; CN-LOH: *Copy Neutral Loss of Heterozygosity*.

Com esta análise quantitativa, embora utilizando apenas oito pacientes, conseguimos corroborar a maioria das alterações de número de cópias demonstradas pela técnica de SNP Array. Utilizando a amostra ACC-1 como exemplo, identificamos uma amplificação da região 12p12.33-p11.12 por SNP Array e, após a construção de sonda pra o gene *KRAS*, demostramos um ganho na região contendo esse gene utilizando FISH (Figura 15).

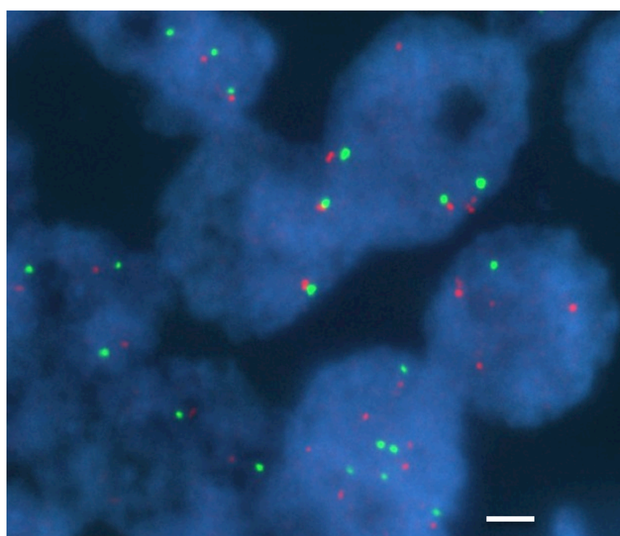


Figura 15 – Paciente ACC-1. KRAS(SR)/CEP12(SG). FISH Escore: Ganho. Escala 10µm.

Embora com resultados ambíguos relativos à análise de SNP Array para as regiões escolhidas 11p15.5-p11.2 e 12p13.33-p11.12, o pacientes ACC-2 apresentou resultados normais para os 3 genes avaliados por FISH. Os pacientes ACC-3 e ACC-4 apresentaram perda de heterozigosidade por SNP Array na região 11p15.5-p11.2, porém, resultados normais para FISH pois apresentaram um número de cópias normal para a região 11p15.5-p11.2. Não foi possível realizar hibridização no tecido do paciente ACC-5 pois a amostra disponível era muito pequena, de baixa qualidade e não suportava as variações de temperatura necessárias para a hibridização das sondas. Interessantemente, a amostra ACC-6 que também apresentou resultados ambíguos para 1q21.1-q44, amplificação de 11p15.5-p11.2 e 12p13.33-p11.12 por SNP Array, apresentou importantes resultados nas análises de FISH, possuindo uma grande amplificação da região 1q21.1-q44 como demonstrado pela figura 16 com amplificação do gene *ABL2*, amplificação de *CDKN1C* assim como ganho de *KRAS*.

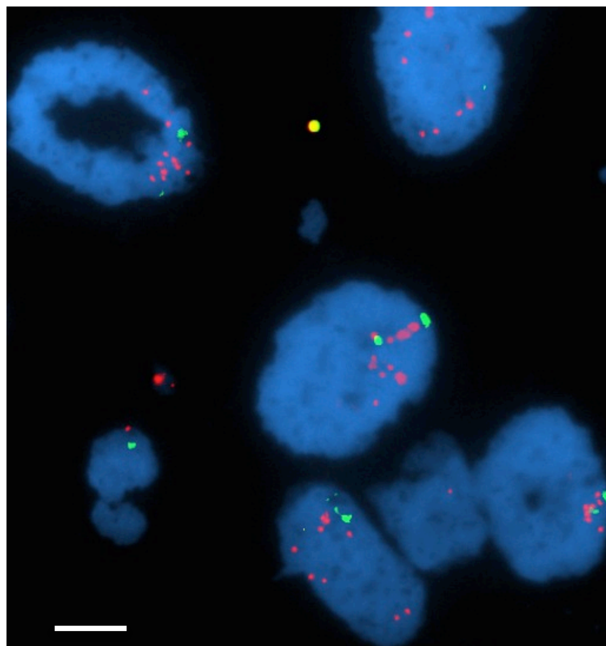


Figura 16 – Paciente ACC-6. ABL2(SR)/CEP1(SG). FISH escore: Amplificação. Escala 10μm.

Por último, analisamos dois pacientes com câncer de mama que apresentaram claros resultados na análise do SNP Array para as regiões

1q21.1-q44, 11p15.5-p11.2 e 12p13.33-p11.12. Ambos os pacientes apresentaram deleções na região 11p15.5-p11.2 definida pelo SNP Array e apresentaram perda pela análise de FISH conforme as figuras 17 e 18. Além disso, os mesmos apresentaram amplificação da região 12p13.33-p11.12 por SNP Array e ganho quando analisados por FISH.

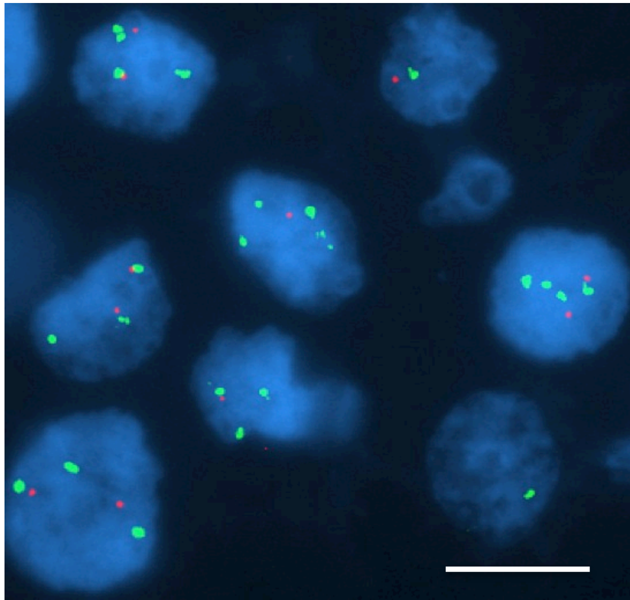


Figura 17 – Paciente BC-1. CDKN1C(SR)/CEP11(SG). FISH escore: Perda. Escala 10μm.

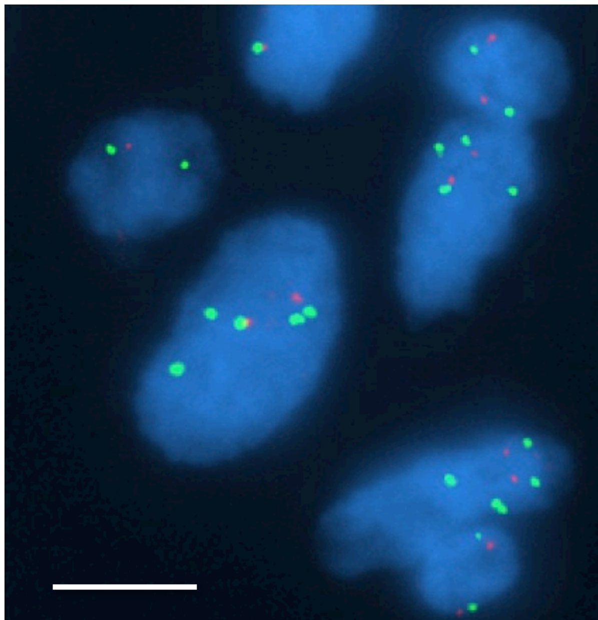


Figura 18 – Paciente BC-2. CDKN1C(SR)/CEP11(SG). FISH escore: Perda. Escala 10μm

7. DISCUSSÃO E PERSPECTIVAS

Os resultados específicos desta tese, cujo objetivo geral é a identificação de alterações moleculares associadas à via de sinalização do gene *TP53* em câncer e infertilidade, serão discutidos separadamente, para cada uma das abordagens empregadas: *TP53* e Câncer; *TP53* e Fertilidade.

7.1 *TP53* E FERTILIDADE

Neste estudo, nossos resultados corroboram os de estudos prévios que sugerem que o polimorfismo *TP53* rs1042522 (PEX4) tem uma associação com desfechos relacionados ao processo de implantação do embrião em mulheres com menos de 35 anos e também demonstramos sua associação ao fenótipo de endometriose com fertilidade. A maior frequência do alelo *TP53* rs1042522 C, associado a níveis inferiores de LIF, foi significativamente menor (em comparação com a do alelo *TP53* rs1042522 G) em mulheres inférteis com endometriose e também em mulheres com falha de implantação, sugerindo um papel importante deste polimorfismo na infertilidade.

Além de corroborar estudos prévios envolvendo *TP53* rs1042522 (PEX4), nosso estudo avaliou pela primeira vez possíveis associações entre o polimorfismo *TP53* rs17878362 (PIN3) e desfechos reprodutivos. PIN3 está envolvido no processamento do intron 2, modulando ativamente os níveis de mRNA do *TP53*. A associação do alelo D (duplicação de 16pb) com os desfechos estudados também indica uma função importante deste polimorfismo na infertilidade. Na mesma linha, a demonstração de que os polimorfismos rs17878362 (PIN2) e rs1042522 (PEX4) estão em desequilíbrio de ligação e que a combinação dos alelos *TP53* rs17878362 D e *TP53* rs1042522 C aumenta 2,1 vezes o risco para endometriose e 1,7 vezes o risco para falha de implantação após fertilização *in vitro,* sugerem que a análise mais detalhada do efeito funcional da proteína p53 polimórfica nos tecidos-alvo (por exemplo, endométrio) nos poderá ser de grande auxílio no entendimento

da etiopatogenia da endometriose e da infertilidade. Identificamos ainda a associação de dois polimorfismos, *TP53* rs17878362 (PIN3) e *ESR1* rs9340799, aos desfechos estudados de infertilidade. Nosso estudo evidencia pela primeira vez a associação de um SNP no gene *ESR1* com falha recorrente de implantação e endometriose, sugerindo que não apenas *TP53*, mas outros genes ligados à via de sinalização de *TP53* influenciam a fertilidade. Por fim, utilizando ferramentas de bioinformática, identificamos novos genes da via de sinalização de *TP53* que parecem ter funções essenciais nos processos de implantação e decidualização como *AKAP5, CALCA, CYP27B1*, *IGFBP7, IL1B*, *MEN1, PLA2G4A*, *PLAU*, *PPARD*, *PTGS2*, *SOD1*, *SPP1, UBE2A e VDR* abrindo novos caminhos para projetos de pesquisa nessa área e contribuindo no entendimento da rota de *TP53* no contexto da fertilidade.

Em relação ao polimorfismo *TP53* rs1042522 (PEX4), além dos nossos resultados em relação à infertilidade e endometriose, evidências recentes sugerem também uma função pós-implantacional do gene *TP53*. Estudando uma população da Cidade de Cândido Godói, caracterizada por elevadas taxas de gemelaridade, Tagliani-Ribeiro e colaboradores, em estudo do qual participamos, associaram a presença do alelo *TP53* rs1042522 C como um fator de risco para gemelaridade e demonstraram que os portadores do alelo T do polimorfismo *MDM4* rs1563828 possuem menores taxas de gravidez em comparação a não- portadores (Tagliani-Ribeiro *et al.*, 2012). Desta forma, podemos inferir uma função pós-implantacional para o gene *TP53*, onde uma proteína p53 e mecanismos que são menos efetivos na indução de apoptose (na presença do alelo rs1042522 C) podem aumentar a probabilidade de sobrevivência de embriões, aumentando a chance de gestação gemelares, tanto mono- quanto dizigóticas, como se observa na cidade de Cândido Godói. Corroborando a hipótese que o gene *TP53* pode apresentar função pós-implantacional, em outro estudo de Fraga e colaboradores do qual participamos, mulheres portadoras dos genótipos *TP53* rs1042522 GG e *MDM2* rs2279744 TT apresentaram maiores taxas de perdas gestacionais recorrentes (Fraga *et al.* 2013, artigo submetido, publicação complementar 10.3).

As evidências de nosso estudo e de outros indicam claramente que o polimorfismo *TP53* rs17878362, também conhecido como PIN3, é um regulador da expressão de *TP53*, com potencial alvo terapêutico. Esta duplicação de 16 pares de base no intron 3 está inserida em uma região que forma um motivo G-quadruplex e está associada ao processamento do intron 2, gerando desta forma, variantes da proteína p53 com ausência do domínio de transativação N-terminal. Corroborando com nosso achados envolvendo desfechos de infertilidade, em recente meta-análise para avaliar o risco associado a câncer do polimorfismo *TP53* rs17878362, incluindo 10.786 pacientes com diversos tipos de câncer e 11.760 controles, Sagne *et al.* identificaram que homozigotos portadores do genótipo *TP53* rs17878362 DD apresentaram maior risco ao desenvolvimento de câncer, especialmente câncer de mama e colorretal (Sagne *et al.*, 2013).

A busca por novas variantes funcionais que possam estar relacionadas à fertilidade abre novas e instigantes perspectivas para inferências sobre o papel de genes da rota p53 na evolução do processo de reprodução em humanos. Os próximos passos para elucidar o papel do polimorfismo *TP53* rs17878362 na expressão e capacidade de transativação de p53 em mulheres inférteis com e sem endometriose será a análise da proteína mutante em tecido endometrial de portadoras com endometriose e infertilidade, e naquelas com falhas recorrentes de implantação.

7.2 *TP53* E CÂNCER

Embora a mutação *TP53* p.R337H tenha sido inicialmente descrita como uma mutação tumor específica associada a carcinoma adrenocortical, posteriormente, portadores com câncer de mama, tumores phyllodes, carcinoma de plexo coróide, sarcomas, e vários outros tumores foram identificados. A penetrância desta mutação é incompleta e é difícil predizer o fenótipo de um portador, que tem risco aumentado para diversos tumores, em diversas faixas etárias. Por muitos anos, a maioria dos estudos envolvendo

identificação de mutações germinativas no gene *TP53* analisava alterações pontuais nos éxons 5-8 que codificam o domínio de ligação ao DNA da proteína, que são as mais comuns, e não incluíam análise de outras regiões do gene, especialmente o éxon 10, onde se encontra a mutação p.R337H. Embora muito conhecimento tenha sido produzido acerca desta mutação desde 2001, quando da sua descrição inicial, o mecanismo de carcinogênese a ela associado, e fatores genéticos que influenciam a sua penetrância não estão completamente elucidados.

Utilizando uma plataforma de alta resolução de varredura genômica, o presente estudo procurou identificar potenciais alterações genéticas potenciais associadas aos tumores mais prevalentes em pacientes portadores da mutação *TP53* p.R337H. Analisando pacientes portadores com ACC (n=4; ACC-1, ACC-2, ACC-5, ACC-6) identificamos variações de número de cópias de DNA não identificadas anteriormente em pacientes com este diagnóstico, inclusive não em pacientes com esta mutação e ACC. Ganhos em 12p13.3-12p11.23 (KRAS), 16p13.3-16q24.3 (Família *MT1*), e deleções em 2p25.3-2q37.3 (*ALK*), 17p13.1 (*TP53*) foram também identificados.

Supreendentemente, o paciente ACC-6 apresentou amplificação das regiões 1q21.1-1q44 (confirmada por de FISH para o gene *ABL2 que se encontra nesta região*) e 11p15.5-p11.2 (confirmada por FISH para o gene *CDKN1C*). Nenhum estudo na literatura atual versa sobre amplificações no gene *ABL2* em carcinoma adrenocortical, entretanto, há descrição de tumores com aumento do número de cópias deste gene que são muito sensíveis ao tratamento com desatinib (inibidor de SRC/ABL) tanto *in vitro* como em camundongos apresentando câncer de pulmão de não pequenas células (Sos *et al.*, 2009). Por outro lado, rearranjos no gene *ABL2* já foram identificados em modelos animais de câncer de mama (Gil-Henn *et al.*, 2012) e em leucemia linfoblástica aguda (Zhou *et al.*, 2012). Além das alterações aqui identificadas, Letouze *et al.* avaliaram pacientes com ACC (13 portadores da mutação *TP53* p.R337H) e demonstraram alterações em 4q34, 9q33 e 19p e perda de heterozigosidade em 17p13 e 11p15 e concluíram que outras vias celulares além de *TP53* contribuem para o desenvolvimento de ACC infantil, sugerindo que alterações importantes possam estar relacionadas com genes da região do cromossomo 11p15 (Letouze *et al.*, 2012). Nosso estudo corrobora com

este achado, pois, embora não houvesse alteração do número de cópias em cinco pacientes analisados, o paciente ACC-6 apresentou uma grande amplificação da região 11p15, conforme citado anteriormente, que envolve entre muitos outros o gene *CDKN1C*. *CDKN1C* tem um importante papel em diversos processos carcinogênicos como apoptose, invasão celular, diferenciação e angiogênese. Sua expressão é regulada através de mudanças epigenéticas como metilação e compressão de histonas na região promotora do gene (Gonzalez *et al.*, 2005; Kavanagh *et al.*, 2011; Vlachos *et al.*, 2007), sendo desta forma, alvo para possíveis intervenções terapêuticas.

Curiosamente, quando analisamos os pacientes ACC-3 e ACC-4, os quais não apresentam mutações germinativas no gene *TP53*, identificamos uma clara região com perda de heterozigosidade sem alteração no número de cópias na região 11p15.5 - 11p15.4 (Anexo B – Figura 23). Esta região comum aos dois pacientes se estende por 3.817.766 nucleotídeos e inclui, entre outros, os genes *SLC22A1L*, associado com câncer de mama e com baixa sobrevida de pacientes com câncer de mama após cirurgia (Gallagher *et al.*, 2006; He *et al.*, 2011), *CDKN1C*, *H19* que foi associado a câncer de mama e próstata (Couch *et al.*, 2009; Odefrey *et al.*, 2010) e *HRAS* associado a melanomas e câncer de bexiga (Kunz *et al.*, 2013; Ouerhani *et al.*, 2011).

A interpretação dos resultados do paciente ACC-5 gerou muita discussão devido ao seu status de homozigose para *TP53* p.R337H (genótipo AA). Analisando os gráficos de BAF e LRR do paciente ACC-5 (Figuras 19 e 20) e de seus pais (Figura 21) constatamos uma grande região de perda de heterozigosidade sem alteração no número de cópias no paciente (região de 4.2Mb). Esta região encontra-se normal em ambos os pais portadores da mutação (Figura 7.1.3). Nesta região situam-se 79 genes descritos no banco de dados OMIM incluindo os genes *TP53*, *WRAP53*, *TNFSF12*, *BCL6B* e *USP6*. Esta região provavelmente corresponde ao haplótipo fundador identificado por Garritano *et al.* (2009) em portadores da mutação *TP53* p.R337H.

Distinto do padrão encontrado em pacientes portadores da mutação *TP53* p.R337H com ACC, os pacientes portadores da mutação *TP53* p.R337H que desenvolveram câncer de mama (BC-1 e BC-2) apresentaram deleções na região 11p15. Nosso estudo é o primeiro a analisar uma estratégia de

varredura genômica em pacientes portadores da mutação *TP53* p.R337H que desenvolveram câncer de mama para estudar alterações genômicas no tumor, e, embora tenham sido utilizados apenas dois casos nesse momento inicial, a análise conjunta de SNP Array e FISH, demonstrou amplificações da região 1q24.2-1q25.3 (identificadas pelo SNP Array e validadas por FISH) nos dois pacientes e novas amplificações não anteriormente descritas em tumores de mama incluindo 4p16.3-4p16.1 e 5q33.1-5q34, que inclui *HMMR*, cuja amplificação está fortemente associada a maior risco para câncer de mama (Maxwell *et al.*, 2011).

A identificação de novos marcadores associados à mutação *TP53* p.R337H abre novas perspectivas de pesquisa em portadores da mutação. Inicialmente, pretendemos aumentar a análise para incluir um maior número de pacientes portadores da mutação *TP53* p.R337H. Devido ao alto custo da metodologia de SNP-array, uma possibilidade é utilizar a técnica de sequenciamento de nova geração para identificação, em larga escala, de genes associados a esta mutação em pacientes com diferentes tumores do espectro da SLF. Adicionalmente, tendo identificado com maior precisão a extensão e limites prováveis do bloco haplotípico associado à mutação (região de perda de heterozigosidade sem alteração no número de cópias na paciente ACC-5) nosso objetivo agora é identificar, com auxílio de marcadores STRs ao redor desta região, a origem e a idade da mutação *TP53* p.R337H.

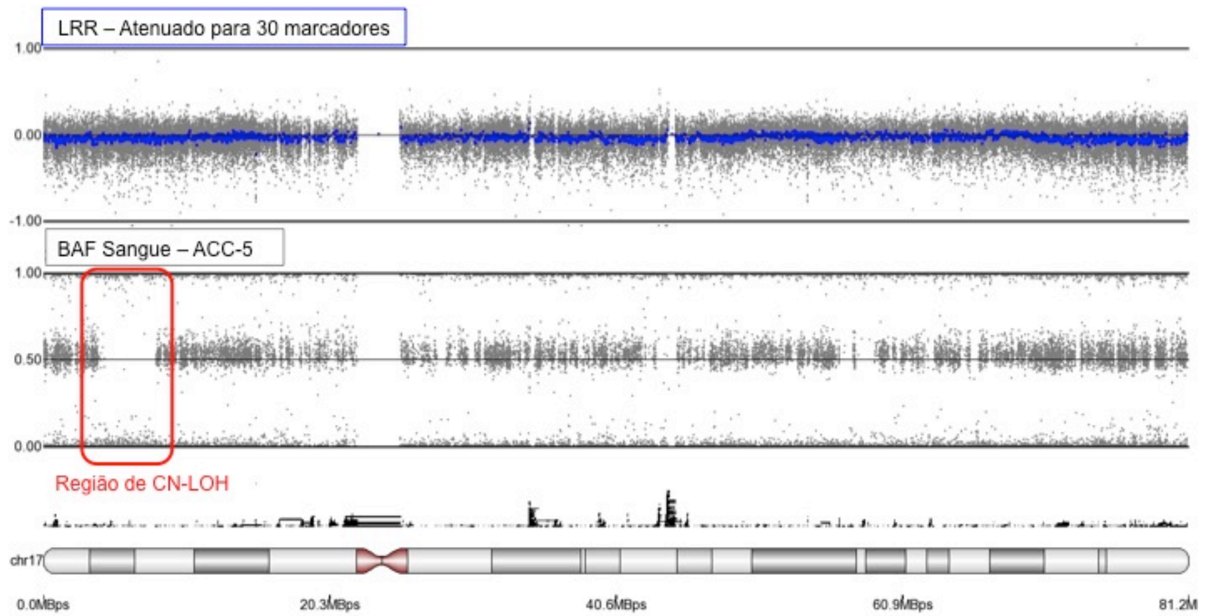


Figura 19 - Resultado do SNP Array para o cromossomo 17 do paciente ACC-5 demonstrando a perda de heterozigosidade sem alteração do número de cópias na região 17p13.2-p13.1. LRR (azul); BAF Sangue; CN-LOH: *Copy Neutral Loss of Heterozygosity.*

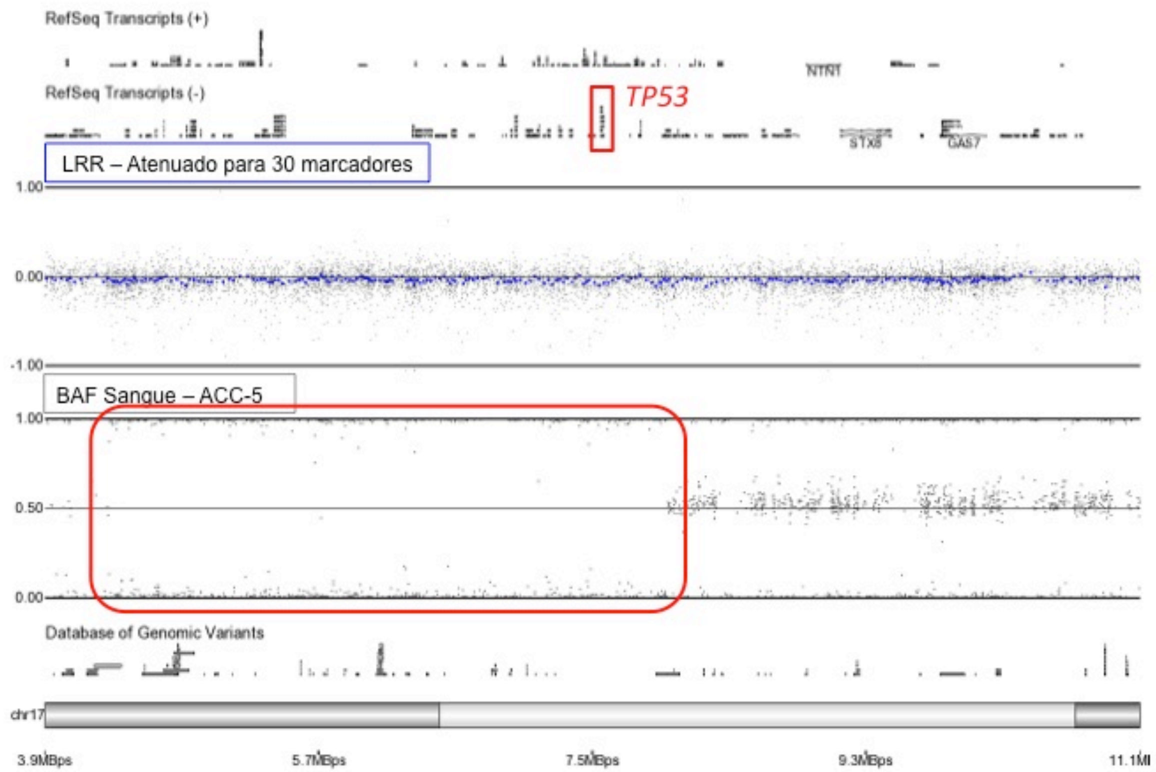


Figura 20 - Resultado do SNP Array para parte do cromossomo 17 do paciente ACC-5 demonstrando a perda de heterozigosidade sem alteração do número de cópias na região 17p13.2-p13.1 (~4.2Mb) LRR (azul); BAF Sangue.

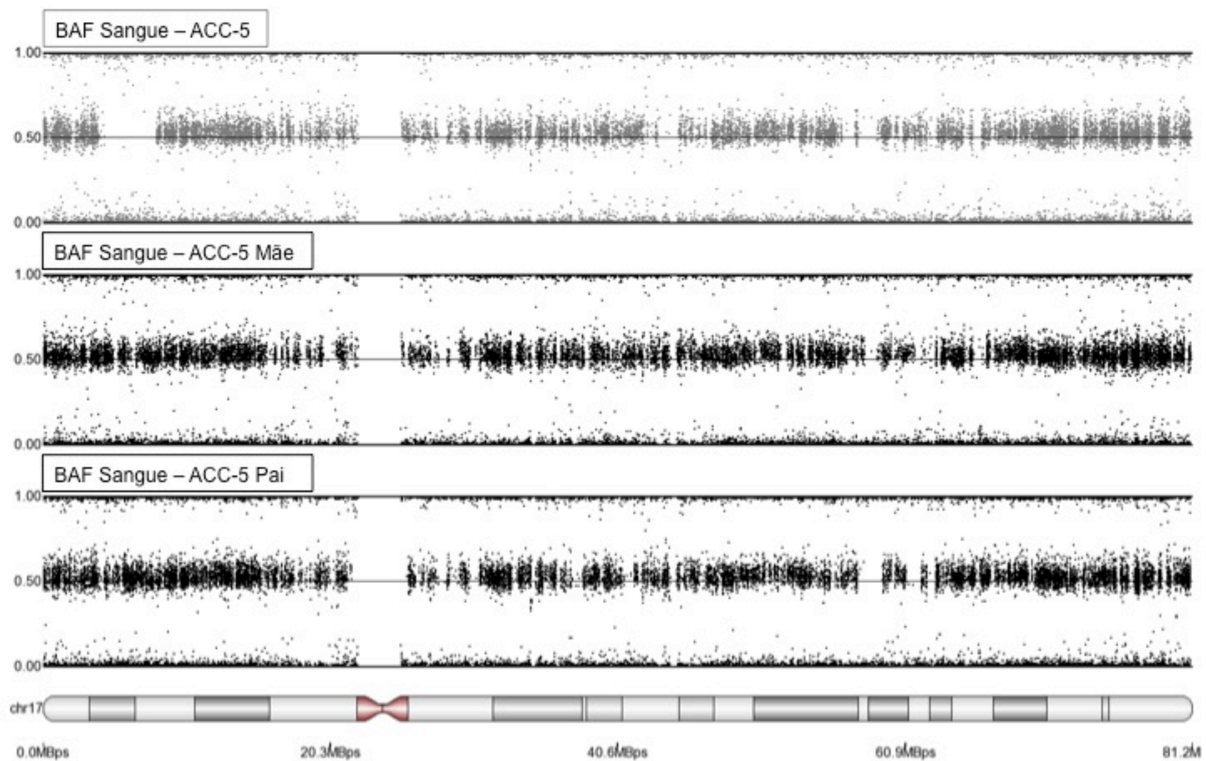


Figura 21 - Resultado do SNP Array para o cromossomo 17 do paciente ACC-5 demonstrando as BAFs do paciente ACC-5 em relação aos pais portadores da mutação *TP53* p.R337H.

8. CONCLUSÕES

As conclusões desta tese serão apresentadas separadamente, para cada abordagem empregada na tese: *TP53* e Câncer; *TP53* e Fertilidade.

8.1 *TP53* E FERTILIDADE

8.1.1 Comparando os polimorfismos *TP53* rs1642785 (PIN2), rs17878362 (PIN3) e rs1042522 (PEX4) do gene *TP53* em mulheres inférteis com falha recorrente de fertilização *in vitro* e com endometriose com mulheres férteis, demonstramos uma associação entre os polimorfismos rs17878362 e rs1042522 com os desfechos de infertilidade analisados.

8.1.2 Analisando três polimorfismos de *TP53* em mulheres inférteis com falha recorrente de fertilização *in vitro* e com endometriose com mulheres férteis, identificamos forte desequilíbrio de ligação entre *TP53* rs17878362 (PIN3) e rs1042522 (PEX4) e demonstramos a associação dos haplótipo PIN3-D/PIN4-C e PIN3-N/PEX4-C com os desfechos de infertilidade analisados.

8.1.3 Com uma abordagem *in silico*, identificamos um conjunto de genes não anteriormente relacionados com os processos de implantação e decidualização os quais se associam também a *TP53: AKAP5, CALCA, CYP27B1*, *IGFBP7, IL1B*, *MEN1, PLA2G4A*, *PLAU*, *PPARD*, *PTGS2*, *SOD1*, *SPP1, UBE2A e VDR*.

8.1.4 Analisando polimorfismos em outros genes da via de sinalização de *TP53* (genes *MDM2, MDM4, USP7, LIF*, *TP63, TP73* e *ESR1)* em mulheres inférteis com falha recorrente de fertilização *in vitro* e endometriose, identificamos o polimorfismo *ESR1* rs9340799 como fator de risco para os desfechos estudados.

8.2 *TP53* E CÂNCER

8.2.1 Utilizando a técnica de SNP Array, validada por FISH, foram identificadas diversas regiões cromossômicas com variação do número de cópias em pacientes com ACC portadores da mutação *TP53* p.R337H, destacando-se: ganhos em 12p13.3-12p11.23, 16p13.3-16q24.3, e deleções em 2p25.3-2q37.3, 17p13.1.

8.2.2 Utilizando a técnica de SNP Array, validada por FISH, foram identificadas diversas regiões cromossômicas com variação do número de cópias em pacientes com câncer de mama portadores da mutação *TP53* p.R337H, destacando-se amplificações de 1q21.1-1q44 e deleções de 11p15.5-p11.2.

8.2.3 Foi identificada uma região de perda de heterozigosidade sem alteração no número de cópias com extensão de 4,2 Mb em 17p13.1 que provavelmente corresponde ao bloco haplotípico no qual incide a mutação *TP53* p.R337H.

8.1.4 Os achados de regiões com variação no numero de cópias deverão ser confirmados em uma série independente de casos de pacientes portadores da mutação *TP53* p.R337H com câncer de mama e com carcinoma adrenocortical.

9. REFERÊNCIAS BIBLIOGRÁFICAS

Achatz MI, Olivier M, Le Calvez F, Martel-Planche G, Lopes A, Rossi BM, Ashton-Prolla P, Giugliani R, Palmero EI, Vargas FR *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.

Ammendola M, Gloria-Bottini F, Sesti F, Piccione E, Bottini E. Association of p53 codon 72 polymorphism with endometriosis. *Fertil Steril* 2008;**90**,406-8.

Assumpcao JG, Seidinger AL, Mastellaro MJ, Ribeiro RC, Zambetti GP, Ganti R, Srivastava K, Shurtleff S, Pei D, Zeferino LC *et al.* Association of the germline TP53 R337H mutation with breast cancer in southern Brazil. *BMC Cancer* 2008;**8**,357.

Atwal GS, Bond GL, Metsuyanim S, Papa M, Friedman E, Distelman-Menachem T, Ben Asher E, Lancet D, Ross DA, Sninsky J *et al.* Haplotype structure and selection of the MDM2 oncogene in humans. *Proc Natl Acad Sci U S A* 2007;**104**,4524-9.

Atwal GS, Kirchhoff T, Bond EE, Montagna M, Menin C, Bertorelle R, Scaini MC, Bartel F, Bohnke A, Pempe C *et al.* Altered tumor formation and evolutionary selection of genetic variants in the human MDM4 oncogene. *Proc Natl Acad Sci U S A* 2009;**106**,10236-41.

Belyi VA, Ak P, Markert E, Wang H, Hu W, Puzio-Kuter A, Levine AJ. The origins and evolution of the p53 family of genes. *Cold Spring Harb Perspect Biol* 2010;**2**,a001198.

Birch JM, Alston RD, McNally RJ, Evans DG, Kelsey AM, Harris M, Eden OB, Varley JM. Relative frequency and morphology of cancers in carriers of germline TP53 mutations. *Oncogene* 2001;**20**,4621-8.

Birch JM, Hartley AL, Tricker KJ, Prosser J, Condie A, Kelsey AM, Harris M, Jones PH, Binchy A, Crowther D *et al.* Prevalence and diversity of constitutional mutations in the p53 gene among 21 Li-Fraumeni families. *Cancer Res* 1994;**54**,1298-304.

Blount PL, Galipeau PC, Sanchez CA, Neshat K, Levine DS, Yin J, Suzuki H, Abraham JM, Meltzer SJ, Reid BJ. 17p allelic losses in diploid cells of patients with Barrett's esophagus who develop aneuploidy. *Cancer Res* 1994;**54**,2292-5.

Bond GL, Hu W, Bond EE, Robins H, Lutzker SG, Arva NC, Bargonetti J, Bartel F, Taubert H, Wuerl P *et al.* A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. *Cell* 2004;**119**,591-602.

Brooks CL, Gu W. Dynamics in the p53-Mdm2 ubiquitination pathway. *Cell Cycle* 2004;**3**,895-9.

Chen JR, Cheng JG, Shatzer T, Sewell L, Hernandez L, Stewart CL. Leukemia inhibitory factor can substitute for nidatory estrogen and is essential to inducing a receptive uterus for implantation but is not essential for subsequent embryogenesis. *Endocrinology* 2000;**141**,4365-72.

Cheng R, Ford BL, O'Neal PE, Mathews CZ, Bradford CS, Thongtan T, Barnes DW, Hendricks JD, Bailey GS. Zebrafish (Danio rerio) p53 tumor suppressor gene: cDNA sequence and expression during embryogenesis. *Mol Mar Biol Biotechnol* 1997;**6**,88-97.

Chompret A, Abel A, Stoppa-Lyonnet D, Brugieres L, Pages S, Feunteun J, Bonaiti-Pellie C. Sensitivity and predictive value of criteria for p53 germline mutation screening. *J Med Genet* 2001;**38**,43-7.

Couch FJ, Wang X, McWilliams RR, Bamlet WR, de Andrade M, Petersen GM. Association of breast cancer susceptibility variants with risk of pancreatic cancer. *Cancer Epidemiol Biomarkers Prev* 2009;**18**,3044-8.

Cunha-Filho JS, Gross JL, Lemos NA, Brandelli A, Castillos M, Passos EP. Hyperprolactinemia and luteal insufficiency in infertile patients with mild and minimal endometriosis. *Horm Metab Res* 2001;**33**,216-20.

Cunha-Filho JS, Gross JL, Lemos NA, Dias EC, Vettori D, Souza CA, Passos EP. Prolactin and growth hormone secretion after thyrotrophin-releasing hormone infusion and dopaminergic (DA2) blockade in infertile patients with minimal/mild endometriosis. *Hum Reprod* 2002;**17**,960-5.

Custodio G, Komechen H, Figueiredo FR, Fachin ND, Pianovski MA, Figueiredo BC. Molecular epidemiology of adrenocortical tumors in southern Brazil. *Mol Cell Endocrinol* 2011;**351**,44-51.

DiGiammarino EL, Lee AS, Cadwell C, Zhang W, Bothner B, Ribeiro RC, Zambetti G, Kriwacki RW. A novel mechanism of tumorigenesis involving pH-dependent destabilization of a mutant p53 tetramer. *Nat Struct Biol* 2002;**9**,12-6.

Dimitriadis E, White CA, Jones RL, Salamonsen LA. Cytokines, chemokines and growth factors in endometrium related to implantation. *Hum Reprod Update* 2005;**11**,613-30.

Donehower LA, Harvey M, Slagle BL, McArthur MJ, Montgomery CA, Jr., Butel JS, Bradley A. Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* 1992;**356**,215-21.

Dumont P, Leu JI, Della Pietra AC, 3rd, George DL, Murphy M. The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. *Nat Genet* 2003;**33**,357-65.

Eklof V, Wikberg ML, Edin S, Dahlin AM, Jonsson BA, Oberg A, Rutegard J, Palmqvist R. The prognostic role of KRAS, BRAF, PIK3CA and PTEN in colorectal cancer. *Br J Cancer* 2013.

Everett RD, Meredith M, Orr A, Cross A, Kathoria M, Parkinson J. A novel ubiquitin-specific protease is dynamically associated with the PML nuclear domain and binds to a herpesvirus regulatory protein. *Embo J* 1997;**16**,1519-30.

Feng Z, Zhang C, Kang HJ, Sun Y, Wang H, Naqvi A, Frank AK, Rosenwaks Z, Murphy ME, Levine AJ *et al.* Regulation of female reproduction by p53 and its family members. *Faseb J* 2011.

Firouzabadi RD, Ghasemi N, Rozbahani MA, Tabibnejad N. Association of p53 polymorphism with ICSI/IVF failure and recurrent pregnancy loss. *Aust N Z J Obstet Gynaecol* 2009;**49**,216-9.

Fu L, Minden MD, Benchimol S. Translational regulation of human p53 gene expression. *EMBO J* 1996;**15**,4392-401.

Gallagher E, Mc Goldrick A, Chung WY, Mc Cormack O, Harrison M, Kerin M, Dervan PA, Mc Cann A. Gain of imprinting of SLC22A18 sense and antisense transcripts in human breast cancer. *Genomics* 2006;**88**,12-7.

Gemignani F, Moreno V, Landi S, Moullan N, Chabrier A, Gutierrez-Enriquez S, Hall J, Guino E, Peinado MA, Capella G *et al.* A TP53 polymorphism is associated with increased risk of colorectal cancer and with reduced levels of TP53 mRNA. *Oncogene* 2004;**23**,1954-6.

Giacomazzi J, Selistre S, Duarte J, Ribeiro JP, Vieira PJ, de Souza Macedo G, Rossi C, Czepielewski M, Netto CB, Hainaut P *et al.* TP53 p.R337H is a conditional cancer-predisposing mutation: further evidence from a homozygous patient. *BMC Cancer* 2013;**13**,187.

Gil-Henn H, Patsialou A, Wang Y, Warren MS, Condeelis JS, Koleske AJ. Arg/Abl2 promotes invasion and attenuates proliferation of breast cancer in vivo. *Oncogene* 2012.

Gomes MC, Kotsopoulos J, de Almeida GL, Costa MM, Vieira R, Filho Fde A, Pitombo MB, PR FL, Royer R, Zhang P *et al.* The R337H mutation in TP53 and breast cancer in Brazil. *Hered Cancer Clin Pract* 2012;**10**,3.

Gonfloni S, Di Tella L, Caldarola S, Cannata SM, Klinger FG, Di Bartolomeo C, Mattei M, Candi E, De Felici M, Melino G *et al.* Inhibition of the c-Abl-TAp63 pathway protects mouse oocytes from chemotherapy-induced death. *Nat Med* 2009;**15**,1179-85.

Gonzalez KD, Noltner KA, Buzin CH, Gu D, Wen-Fong CY, Nguyen VQ, Han JH, Lowstuter K, Longmate J, Sommer SS *et al.* Beyond Li Fraumeni Syndrome: clinical characteristics of families with p53 germline mutations. *J Clin Oncol* 2009;**27**,1250-6.

Gonzalez S, Perez-Perez MM, Hernando E, Serrano M, Cordon-Cardo C. p73beta-Mediated apoptosis requires p57kip2 induction and IEX-1 inhibition. *Cancer Res* 2005;**65**,2186-92.

Goumenou A, Panayiotides I, Mahutte NG, Matalliotakis I, Fragouli Y, Arici A. Immunohistochemical expression of p53, MDM2, and p21Waf1 oncoproteins in endometriomas but not adenomyosis. *J Soc Gynecol Investig* 2005;**12**,263-6.

Hainaut P. Tumor-specific mutations in p53: the acid test. *Nat Med* 2002;**8**,21-3.

Halis G, Arici A. Endometriosis and inflammation in infertility. *Ann N Y Acad Sci* 2004;**1034**,300-15.

Hambartsoumian E. Endometrial leukemia inhibitory factor (LIF) as a possible cause of unexplained infertility and multiple failures of implantation. *Am J Reprod Immunol* 1998;**39**,137-43.

Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000;**100**,57-70.

Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011;**144**,646-74.

He H, Xu C, Zhao Z, Qin X, Xu H, Zhang H. Low expression of SLC22A18 predicts poor survival outcome in patients with breast cancer after surgery. *Cancer Epidemiol* 2011;**35**,279-85.

Hollstein M, Hainaut P. Massively regulated genes: the example of TP53. *J Pathol* 2010;**220**,164-73.

Hsieh YY, Lin CS. P53 codon 11, 72, and 248 gene polymorphisms in endometriosis. *Int J Biol Sci* 2006;**2**,188-93.

Hu W, Feng Z, Atwal GS, Levine AJ. p53: a new player in reproduction. *Cell Cycle* 2008;**7**,848-52.
Hu W, Feng Z, Ma L, Wagner J, Rice JJ, Stolovitzky G, Levine AJ. A single nucleotide polymorphism in the MDM2 gene disrupts the oscillation of p53 and MDM2 levels in cells. *Cancer Res* 2007a;**67**,2757-65.
Hu W, Feng Z, Teresky AK, Levine AJ. p53 regulates maternal reproduction through LIF. *Nature* 2007b;**450**,721-4.
Huarte M, Guttman M, Feldser D, Garber M, Koziol MJ, Kenzelmann-Broz D, Khalil AM, Zuk O, Amit I, Rabani M *et al.* A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* 2010;**142**,409-19.
Jin S, Martinek S, Joo WS, Wortman JR, Mirkovic N, Sali A, Yandell MD, Pavletich NP, Young MW, Levine AJ. Identification and characterization of a p53 homologue in Drosophila melanogaster. *Proc Natl Acad Sci U S A* 2000;**97**,7301-6.
Kang HJ, Feng Z, Sun Y, Atwal G, Murphy ME, Rebbeck TR, Rosenwaks Z, Levine AJ, Hu W. Single-nucleotide polymorphisms in the p53 pathway regulate fertility in humans. *Proc Natl Acad Sci U S A* 2009;**106**,9761-6.
Kavanagh E, Joseph B. The hallmarks of CDKN1C (p57, KIP2) in cancer. *Biochim Biophys Acta* 2011;**1816**,50-6.
Kay C, Jeyendran RS, Coulam CB. p53 tumour suppressor gene polymorphism is associated with recurrent implantation failure. *Reprod Biomed Online* 2006;**13**,492-6.
Koninckx PR. Is mild endometriosis a condition occurring intermittently in all women? *Hum Reprod* 1994;**9**,2202-5.
Kraemer KH, Lee MM, Andrews AD, Lambert WC. The role of sunlight and DNA repair in melanoma and nonmelanoma skin cancer. The xeroderma pigmentosum paradigm. *Arch Dermatol* 1994;**130**,1018-21.
Kunz M, Dannemann M, Kelso J. High-throughput sequencing of the melanoma genome. *Exp Dermatol* 2013;**22**,10-7.
Kurita T, Cunha GR, Robboy SJ, Mills AA, Medina RT. Differential expression of p63 isoforms in female reproductive organs. *Mech Dev* 2005;**122**,1043-55.
Lalloo F, Varley J, Ellis D, Moran A, O'Dair L, Pharoah P, Evans DG. Prediction of pathogenic mutations in patients with early-onset breast cancer by family history. *Lancet* 2003;**361**,1101-2.
Lamp M, Peters M, Reinmaa E, Haller-Kikkatalo K, Kaart T, Kadastik U, Karro H, Metspalu A, Salumets A. Polymorphisms in ESR1, ESR2 and HSD17B1 genes are associated with fertility status in endometriosis. *Gynecol Endocrinol* 2011;**27**,425-33.
Lane D, Levine A. p53 Research: the past thirty years and the next thirty years. *Cold Spring Harb Perspect Biol* 2010;**2**,a000893.
Lane DP. Cancer. p53, guardian of the genome. *Nature* 1992;**358**,15-6.
Latronico AC, Pinto EM, Domenice S, Fragoso MC, Martin RM, Zerbini MC, Lucon AM, Mendonca BB. 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.

Lemos NA, Arbo E, Scalco R, Weiler E, Rosa V, Cunha-Filho JS. Decreased anti-Mullerian hormone and altered ovarian follicular cohort in infertile patients with mild/minimal endometriosis. *Fertil Steril* 2008;**89**,1064-8.

Letouze E, Rosati R, Komechen H, Doghman M, Marisa L, Fluck C, de Krijger RR, van Noesel MM, Mas JC, Pianovski MA *et al.* SNP array profiling of childhood adrenocortical tumors reveals distinct pathways of tumorigenesis and highlights candidate driver genes. *J Clin Endocrinol Metab* 2012;**97**,E1284-93.

Li FP, Fraumeni JF, Jr. Soft-tissue sarcomas, breast cancer, and other neoplasms. A familial syndrome? *Ann Intern Med* 1969;**71**,747-52.

Liaw D, Marsh DJ, Li J, Dahia PL, Wang SI, Zheng Z, Bose S, Call KM, Tsou HC, Peacocke M *et al.* Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. *Nat Genet* 1997;**16**,64-7.

Linzer DI, Levine AJ. Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. *Cell* 1979;**17**,43-52.

Malkin D, Li FP, Strong LC, Fraumeni JF, Jr., Nelson CE, Kim DH, Kassel J, Gryka MA, Bischoff FZ, Tainsky MA *et al.* Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science* 1990;**250**,1233-8.

Maltzman W, Czyzyk L. UV irradiation stimulates levels of p53 cellular tumor antigen in nontransformed mouse cells. *Mol Cell Biol* 1984;**4**,1689-94.

Marcel V, Palmero EI, Falagan-Lotsch P, Martel-Planche G, Ashton-Prolla P, Olivier M, Brentani RR, Hainaut P, Achatz MI. TP53 PIN3 and MDM2 SNP309 polymorphisms as genetic modifiers in the Li-Fraumeni syndrome: impact on age at first diagnosis. *J Med Genet* 2010;**46**,766-72.

Marcel V, Tran PL, Sagne C, Martel-Planche G, Vaslin L, Teulade-Fichou MP, Hall J, Mergny JL, Hainaut P, Van Dyck E. G-quadruplex structures in TP53 intron 3: role in alternative splicing and in production of p53 mRNA isoforms. *Carcinogenesis* 2011;**32**,271-8.

Matlashewski GJ, Tuck S, Pim D, Lamb P, Schneider J, Crawford LV. Primary structure polymorphism at amino acid residue 72 of human p53. *Mol Cell Biol* 1987;**7**,961-3.

Maxwell CA, Benitez J, Gomez-Baldo L, Osorio A, Bonifaci N, Fernandez-Ramires R, Costes SV, Guino E, Chen H, Evans GJ *et al.* Interplay between BRCA1 and RHAMM regulates epithelial apicobasal polarization and may influence risk of breast cancer. *PLoS Biol* 2011;**9**,e1001199.

McKinnon PJ. ATM and ataxia telangiectasia. *EMBO Rep* 2004;**5**,772-6.

Mikolajczyk M, Skrzypczak J, Szymanowski K, Wirstlein P. The assessment of LIF in uterine flushing--a possible new diagnostic tool in states of impaired fertility. *Reprod Biol* 2003;**3**,259-70.

Mills AA, Zheng B, Wang XJ, Vogel H, Roop DR, Bradley A. p63 is a p53 homologue required for limb and epidermal morphogenesis. *Nature* 1999;**398**,708-13.

Miyoshi H, Nakau M, Ishikawa TO, Seldin MF, Oshima M, Taketo MM. Gastrointestinal hamartomatous polyposis in Lkb1 heterozygous knockout mice. *Cancer Res* 2002;**62**,2261-6.

Moll UM, Wolff S, Speidel D, Deppert W. Transcription-independent pro-apoptotic functions of p53. *Curr Opin Cell Biol* 2005;**17**,631-6.

Murphy ME. Polymorphic variants in the p53 pathway. *Cell Death Differ* 2006;**13**,916-20.

Novotny Z, Krizan J, Sima R, Sima P, Uher P, Zech N, Hutelova R, Baborova P, Ulcova-Gallova Z, Subrt I *et al.* Leukaemia inhibitory factor (LIF) gene mutations in women diagnosed with unexplained infertility and endometriosis have a negative impact on the IVF outcome. A pilot study. *Folia Biol (Praha)* 2009;**55**,92-7.

Odefrey F, Stone J, Gurrin LC, Byrnes GB, Apicella C, Dite GS, Cawson JN, Giles GG, Treloar SA, English DR *et al.* Common genetic variants associated with breast cancer and mammographic density measures that predict disease. *Cancer Res* 2010;**70**,1449-58.

Olive DL, Schwartz LB. Endometriosis. *N Engl J Med* 1993;**328**,1759-69.

Olivier M, Eeles R, Hollstein M, Khan MA, Harris CC, Hainaut P. The IARC TP53 database: new online mutation analysis and recommendations to users. *Hum Mutat* 2002;**19**,607-14.

Olshen AB, Venkatraman ES, Lucito R, Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* 2004;**5**,557-72.

Orlova KA, Crino PB. The tuberous sclerosis complex. *Ann N Y Acad Sci* 2010;**1184**,87-105.

Ouerhani S, Elgaaied AB. The mutational spectrum of HRAS, KRAS, NRAS and FGFR3 genes in bladder cancer. *Cancer Biomark* 2011;**10**,259-66.

Palmero EI, Schuler-Faccini L, Caleffi M, Achatz MI, Olivier M, Martel-Planche G, Marcel V, Aguiar E, Giacomazzi J, Ewald IP *et al.* Detection of R337H, a germline TP53 mutation predisposing to multiple cancers, in asymptomatic women participating in a breast cancer screening program in Southern Brazil. *Cancer Lett* 2008;**261**,21-5.

Park SH, Choi JE, Kim EJ, Jang JS, Han HS, Lee WK, Kang YM, Park JY. MDM2 309T>G polymorphism and risk of lung cancer in a Korean population. *Lung Cancer* 2006;**54**,19-24.

Peiffer DA, Le JM, Steemers FJ, Chang W, Jenniges T, Garcia F, Haden K, Li J, Shaw CA, Belmont J *et al.* High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. *Genome Res* 2006;**16**,1136-48.

Petitjean A, Achatz MI, Borresen-Dale AL, Hainaut P, Olivier M. TP53 mutations in human cancers: functional selection and impact on cancer prognosis and outcomes. *Oncogene* 2007;**26**,2157-65.

Pfeifer GP, Hainaut P. Next-generation sequencing: emerging lessons on the origins of human cancer. *Curr Opin Oncol* 2011;**23**,62-8.

Pietrowski D, Bettendorf H, Riener EK, Keck C, Hefler LA, Huber JC, Tempfer C. Recurrent pregnancy failure is associated with a polymorphism in the p53 tumour suppressor gene. *Hum Reprod* 2005;**20**,848-51.

Pim D, Banks L. p53 polymorphic variants at codon 72 exert different effects on cell cycle progression. *Int J Cancer* 2004;**108**,196-9.

Piovezan G (2006). Prevalência do alelo TP53 R337H no Estado do Paraná [dissertação] Curitiba (Paraná): Universidade Federal do Paraná (UFPR);.

Ribeiro Junior CL, Arruda JT, Silva CT, Moura KK. Analysis of p53 codon 72 gene polymorphism in Brazilian patients with endometriosis. *Genet Mol Res* 2009;**8**,494-9.

Ribeiro RC, Sandrini F, Figueiredo B, Zambetti GP, Michalkiewicz E, Lafferty AR, DeLacerda L, Rabin M, Cadwell C, Sampaio G *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.

Sagne C, Marcel V, Amadou A, Hainaut P, Olivier M, Hall J. A meta-analysis of cancer risk associated with the TP53 intron 3 duplication polymorphism (rs17878362): geographic and tumor-specific effects. *Cell Death Dis* 2013;**4**,e492.

Seidinger AL, Mastellaro MJ, Paschoal Fortes F, Godoy Assumpcao J, Aparecida Cardinalli I, Aparecida Ganazza M, Correa Ribeiro R, Brandalise SR, Dos Santos Aguiar S, Yunes JA. Association of the highly prevalent TP53 R337H mutation with pediatric choroid plexus carcinoma and osteosarcoma in southeast Brazil. *Cancer* 2011;**117**,2228-35.

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.

Sos ML, Michel K, Zander T, Weiss J, Frommolt P, Peifer M, Li D, Ullrich R, Koker M, Fischer F *et al.* Predicting drug susceptibility of non-small cell lung cancers based on genetic lesions. *J Clin Invest* 2009;**119**,1727-40.

Soussi T, Caron de Fromentel C, Mechali M, May P, Kress M. Cloning and characterization of a cDNA from Xenopus laevis coding for a protein homologous to human and murine p53. *Oncogene* 1987;**1**,71-8.

Stella GM, Scabini R, Inghilleri S, Cemmi F, Corso S, Pozzi E, Morbini P, Valentini A, Dore R, Ferrari S *et al.* EGFR and KRAS mutational profiling in fresh non-small cell lung cancer (NSCLC) cells. *J Cancer Res Clin Oncol* 2013.

Suh EK, Yang A, Kettenbach A, Bamberger C, Michaelis AH, Zhu Z, Elvin JA, Bronson RT, Crum CP, McKeon F. p63 protects the female germ line during meiotic arrest. *Nature* 2006;**444**,624-8.

Tagliani-Ribeiro A, Paskulin DD, Oliveira M, Zagonel-Oliveira M, Longo D, Ramallo V, Ashton-Prolla P, Saraiva-Pereira ML, Fagundes NJ, Schuler-Faccini L *et al.* High twinning rate in Candido Godoi: a new role for p53 in human fertility. *Hum Reprod* 2012;**27**,2866-71.

Tempfer CB, Simoni M, Destenaves B, Fauser BC. Functional genetic polymorphisms and female reproductive disorders: part II--endometriosis. *Hum Reprod Update* 2009;**15**,97-118.

Thomas M, Kalita A, Labrecque S, Pim D, Banks L, Matlashewski G. Two polymorphic variants of wild-type p53 differ biochemically and biologically. *Mol Cell Biol* 1999;**19**,1092-100.

Tomasini R, Tsuchihara K, Tsuda C, Lau SK, Wilhelm M, Ruffini A, Tsao MS, Iovanna JL, Jurisicova A, Melino G *et al.* TAp73 regulates the spindle assembly checkpoint by modulating BubR1 activity. *Proc Natl Acad Sci U S A* 2009;**106**,797-802.

Tomasini R, Tsuchihara K, Wilhelm M, Fujitani M, Rufini A, Cheung CC, Khan F, Itie-Youten A, Wakeham A, Tsao MS *et al.* TAp73 knockout shows genomic instability with infertility and tumor suppressor functions. *Genes Dev* 2008;**22**,2677-91.

Varley JM, McGown G, Thorncroft M, Santibanez-Koref MF, Kelsey AM, Tricker KJ, Evans DG, Birch JM. Germ-line mutations of TP53 in Li-Fraumeni families: an extended study of 39 families. *Cancer Res* 1997;**57**,3245-52.

Vietri MT, Molinari AM, Iannella I, Cioffi M, Bontempo P, Ardovino M, Scaffa C, Colacurci N, Cobellis L. Arg72Pro p53 polymorphism in Italian women: no association with endometriosis. *Fertil Steril* 2007;**88**,1468-9.

Vigano P, Somigliana E, Vignali M, Busacca M, Blasio AM. Genetics of endometriosis: current status and prospects. *Front Biosci* 2007;**12**,3247-55.

Vlachos P, Nyman U, Hajji N, Joseph B. The cell cycle inhibitor p57(Kip2) promotes cell death via the mitochondrial apoptotic pathway. *Cell Death Differ* 2007;**14**,1497-507.

Vousden KH, Prives C. Blinded by the Light: The Growing Complexity of p53. *Cell* 2009;**137**,413-31.

Whibley C, Pharoah PD, Hollstein M. p53 polymorphisms: cancer implications. *Nat Rev Cancer* 2009;**9**,95-107.

Yang A, Walker N, Bronson R, Kaghad M, Oosterwegel M, Bonnin J, Vagner C, Bonnet H, Dikkes P, Sharpe A *et al.* p73-deficient mice have neurological, pheromonal and inflammatory defects but lack spontaneous tumours. *Nature* 2000;**404**,99-103.

Zhou MH, Gao L, Jing Y, Xu YY, Ding Y, Wang N, Wang W, Li MY, Han XP, Sun JZ *et al.* Detection of ETV6 gene rearrangements in adult acute lymphoblastic leukemia. *Ann Hematol* 2012;**91**,1235-43.

10. PUBLICAÇÕES COMPLEMENTARES NO PERÍODO DE DOUTORADO

10.1 *HIGH TWINNING RATE IN CÂNDIDO GODÓI: A NEW ROLE FOR P53 IN HUMAN FERTILITY*

Título do manuscrito: *High twinning rate in Cândido Godói: a new role for p53 in human fertility*

Autores: Tagliani-Ribeiro A, Paskulin DD, Oliveira M, Zagonel-Oliveira M, Longo D, Ramallo V, Ashton-Prolla P, Saraiva-Pereira ML, Fagundes NJ, Schuler-Faccini L, Matte U

Situação: Publicado

Revista: *Human Reproduction*

Referência: *Hum Reprod*. 2012 Sep;27(9):2866-71. doi: 10.1093/humrep/des217. Epub 2012 Jun 26

Website: http://www.ncbi.nlm.nih.gov/pubmed/22736329

Human Reproduction, Vol.0, No.0 pp. 1–6, 2012
doi:10.1093/humrep/des217

human reproduction

ORIGINAL ARTICLE *Reproductive genetics*

High twinning rate in Cândido Godói: a new role for p53 in human fertility

A. Tagliani-Ribeiro[1,2], D.D. Paskulin[2,3], M. Oliveira[1,2], M. Zagonel-Oliveira[1,2], D. Longo[1,2], V. Ramallo[1,2], P. Ashton-Prolla[1,2,3,4], M.L. Saraiva-Pereira[1,2,3], N.J.R. Fagundes[2,4], L. Schuler-Faccini[1,2,3,4], and U. Matte[1,2,3,*]

[1]INAGEMP – Instituto Nacional de Genética Médica Populacional, Porto Alegre, RS, Brazil [2]Post-Graduation Program in Genetics and Molecular Biology, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil [3]Experimental Research Center, Hospital de Clínicas de Porto Alegre, Rua Ramiro Barcelos 2350, Porto Alegre, RS 90035-903, Brazil [4]Departament of Genetics, Biosciences Institute, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil

*Correspondence address. Tel: +55-51-3359-8841/8838; Fax: +55-51-3359-8010; E-mail: umatte@hcpa.ufrgs.br

Submitted on November 28, 2011; resubmitted on April 19, 2012; accepted on May 3, 2012

BACKGROUND: Cândido Godói (CG) is a small town in South Brazil, which has the highest prevalence of twin births in Brazil. Recently, a number of studies have shown that p53 plays an important role in reproduction through blastocyst implantation and *intra utero* embryo survival. Thus, gene polymorphisms in the p53 pathway were investigated in this population.

METHODS: Single nucleotide polymorphisms from five genes in the p53 pathway were investigated, as well as background characteristics of 42 mothers of twins (cases) and 101 mothers of singletons (controls), all residents from CG.

RESULTS: Mothers of twins have higher number of pregnancies and higher frequencies of P72 allele at *TP53* and T allele at *MDM4* genes compared with controls. Logistic regression shows that both *TP53* and number of pregnancies maintained their association with twinning ($P = 0.004$ and $P = 0.002$, respectively), with *TP53* having a higher odds ratio than number of pregnancies (2.73 versus 1.70, respectively). No interactive effect between *TP53* and *MDM4* ($P = 0.966$) is observed. As expected, mothers of twins have three times more cases of cancer in their first-degree relatives than control mothers ($P = 0.011$).

CONCLUSIONS: Our results suggest that the P72 allele of *TP53* is a strong risk factor for twinning in CG, while the number of pregnancies and the T allele at *MDM4* may represent weaker risk factors. These two alleles are associated with infertility, but the anti-apoptotic effect of low levels of p53 in general, and of the P72 allele in particular, may play a role after implantation, enhancing the chance for a double pregnancy to succeed to term.

Key words: twinning / p53 / fertility / Cândido Godói

Introduction

Cândido Godói (CG) is a small town in South Brazil (27°57′07″S, 54°45′07″W) with approximately 6700 inhabitants, which has the highest prevalence of twin births in Brazil (1.5 versus 1.0%, respectively), probably due to a genetic founder effect (Tagliani-Ribeiro *et al.*, 2011). The town was founded by a few families of German ancestry at the beginning of 20th century (Roche, 1969), and the highest prevalence of twin births occurs in an isolated district called Linha São Pedro (Matte *et al.*, 1996; Tagliani-Ribeiro *et al.*, 2011). However, the genes or alleles which may influence the high frequency of twin births in this population are still puzzling. Several studies have been done in different populations based either on a candidate gene approach, considering genes possibly involved in poly-ovulation (Al-Hendy *et al.*, 2000; Derom *et al.*, 2001; Montgomery *et al.*, 2001, 2004; Palmer *et al.*, 2006) or in the maintenance of a successful pregnancy (Hasbargen *et al.*, 2000; Montgomery *et al.*, 2003), or on genome-wide linkage scans (Busjahn *et al.*, 2000; Derom *et al.*, 2006; Painter *et al.*, 2010). However, to date, no common pattern has emerged from these studies.

The p53 protein, encoded by the *TP53* gene, was first identified as a virus-associated tumor antigen (Kress *et al.*, 1979; Linzer and Levine, 1979). Its function as a tumor-suppressor influencing cycle cell progression and apoptosis was only discovered one decade latter (Eliyahu *et al.*, 1984; Finlay *et al.*, 1988). It is part of a gene family of transcription factors that regulates a variety of biological functions, including development,

© The Author 2012. Published by Oxford University Press on behalf of the European Society of Human Reproduction and Embryology. All rights reserved.
For Permissions, please email: journals.permissions@oup.com

Downloaded from http://humrep.oxfordjournals.org/ at UCHSC Denison Memorial Library Box A-003 on July 12, 2012

stemness, natural immunity, oxidative stress, ageing and cancer (Levine et al., 2011). p53 protein levels and activities are maintained by three regulators, HAUSP, Mdm2 and its homolog Mdm4 (Mandinova and Lee, 2011; Levine et al., 2006). Mdm4 forms heterodimers with Mdm2, which enhances the ability of Mdm2 to induce p53 degradation (Manfredi, 2010; Wade et al., 2010). HAUSP can stabilize Mdm2, Mdm4 and p53 as a specific deubiquitinase and is an important regulator of the p53 pathway (Brooks et al., 2007).

More recently, a number of studies have shown single nucleotide polymorphisms (SNPs) in genes in the p53 pathway associated with 'unexplained infertility' in women (Kay et al., 2006; Kang et al., 2009; Feng et al., 2011). The *TP53* SNP p.P72R was considered a risk factor for implantation failure through its impact on *LIF* expression levels (Hu et al., 2007). In addition, p53 may act after implantation increasing the *intra utero* embryo survival probability. Abnormal embryos induced by ionizing radiation were maintained at higher percentage in p53-deficient mice, while p53-proficient embryos die, most likely through apoptosis (Norimura et al., 1996). Moreover, human embryo survival may be influenced by genotypes at *TP53*, *LIF* and *MDM2* in a sex-dependent fashion (Ucisik-Akkaya et al., 2010).

Considering that the births of both dizygotic (DZ) and monozygotic (MZ) twins have increased prevalence in CG (Tagliani-Ribeiro et al., 2011), the aim of this study is to investigate gene polymorphisms in the p53 pathway related to blastocyst implantation and *intra utero* embryo survival to test if any of these polymorphisms affect twinning in this population.

Materials and Methods

Study design

This was a population-based case–control study. The study was presented to the population in two community events—a talk in the community center of Candido Godoi and an announcement at the end of the Sunday mess in the Catholic Church of Linha São Pedro (the district with a highest twinning rate). After that, families willing to participate in the study contacted the authors and the interviews were scheduled with the families by telephone or personal contact. Families were then visited in their houses, for filling the questionnaire and blood collection. Families without twins, invited to participate in the same occasions, were also recruited at the community's health center.

Study population

Forty-two women from independent households who gave birth to twins (cases) answered a structured questionnaire and donated a blood sample. These 42 women were divided in 10 mothers of MZ and 20 mothers of DZ twins. One woman was mother of triplets, two of them MZ. The remaining 11 women were mothers of same sex twins that were not investigated for zygosity. A control group comprised 101 women from independent households who only gave birth to singletons and had no first-degree relative who was a twin. All 143 women are residents from CG, and the cases sampled represented 40% of all mothers of twins born in this town after 1955. Background characteristics such as age, age at first and last pregnancy, number of pregnancies, occurrence of gestational losses, height, body mass index (BMI), fertility treatments, as well as the incidence of cancer in their first-degree relatives, were collected with a standardized questionnaire. Both women in case and control groups signed a written consent form before collection of information and blood samples. This research project was approved by the Hospital de Clínicas de Porto Alegre Ethics Research Committee under protocol number 09-359.

Molecular analysis

DNA was extracted according to Lahiri and Nurnberger (1991) and quantified in NanoDrop 2000 (Thermo Scientific, USA). SNPs from five genes in the p53 signaling pathway were genotyped including the *TP53* P72R, (rs1042522), *LIF* (rs929271), *MDM2* (SNP309, rs2279744), *MDM4* (rs1563828) and *HAUSP* (rs1529916), thus representing genes both downstream and upstream of the p53 regulatory cascade. Genotypes were determined by allelic discrimination using TaqMan SNP Genotyping Assay (Applied Biosystems, USA), with the following on demand assay numbers: C___2403545_10 (*TP53* P72R), C___7545904_10 (*LIF*, rs929271), C___9493064_10 (*MDM4*, rs1563828) and C___9688119_1 (*HAUSP*, rs1529916). *MDM2* rs2279744 SNP was assessed using an assay-by-design with Probes: FAM 5′-cccgcgccgcagc-3′ and HEX 5′-cccgcgccgaagc-3′, and primers 5′-ttcagggtaaaggtcacggg-3′ and 5′-tcaacc tgcccactgaacc-3′.

Statistical analysis

For numerical variables Student's *t*-test was used, while for categorical variables the χ^2 test was used. For molecular data, Hardy–Weinberg equilibrium was tested for cases, controls and for the total sample. The association between SNPs and twin pregnancies were evaluated using the χ^2 test, with Bonferroni correction for multiple testing. Odds ratio (OR) and the 95% confidence interval (CI) were calculated using the lower risk genotype as the reference. Genetic and non-genetic factors possibly involved with twinning, including interaction among genes, were subjected to a logistic regression analysis to evaluate the robustness of these associations. For the logistic regression, genotypes were encoded assuming either a co-dominant or recessive model according to the value of OR estimated in a univariate analysis. All analyses were performed using the SPSS statistical package (version 19). Pedigrees information as reported by each individual sampled as case or control was used to assess relatedness (*r*) within and between groups.

Results

Women in the case and control groups were similar in the mean age at first pregnancy, BMI and mean reproductive years at the time of data collection. However, the number of pregnancies was significantly higher in cases compared with that in controls ($P = 0.007$; Table I), with cases having a mean parity of 3.1 live births (95% CI: 2.52–3.63) compared with a mean of 2.1 live births (95% CI: 1.89–2.21) in controls. The frequencies of women with a positive history for spontaneous abortions did not differ between cases and controls (Table I). All women denied the use of assisted conception or ovarian stimulation therapies.

For the five genes studied, univariate analysis in the p53 pathway reveals statistically significant results for *TP53* and *MDM4* genes (Table II). However, after Bonferroni correction for multiple testing, only *TP53* remained significant. For *TP53* there are both genotypic and allelic differences between cases and controls, with the P72 allele (cytosine) enriched in twin's mothers. The OR analyses suggests an almost co-dominant effect, with heterozygous individuals having intermediary risk compared with individuals homozygous for either allele. We did not observe any difference in allelic or genotypic frequencies between cases and controls for the other polymorphisms tested in *HAUSP*, *LIF* and *MDM2* genes.

Downloaded from http://humrep.oxfordjournals.org/ at UCHSC Denison Memorial Library Box A-003 on July 12, 2012

Table I Characteristics of the study population.

Characteristic	Cases	Controls	P-value
Number of pregnancies, mean	3.1 ($n = 36$)	2.1 ($n = 101$)	0.007*
Age at first pregnancy, mean	24.5 ($n = 36$)	24.2 ($n = 96$)	0.717*
Age at last pregnancy, mean	30.4 ($n = 34$)	28.7 ($n = 94$)	0.108*
Reproductive years (difference between first and last pregnancy), mean	5.9 ($n = 34$)	4.4 ($n = 94$)	0.123*
Women with positive history for spontaneous abortions	9 ($n = 36$)	12 ($n = 100$)	0.113#
Cancer incidence in first-degree relatives	47% (17/36)	23% (23/100)	0.011#^
Height, mean	1.63 m	1.64 m	0.923¥
BMI, mean	21.4	22.3	0.336¥

BMI (body mass index).
*Independent samples T-test.
#Chi-square.
^OR (95% CI) = 3.00 (1.34–6.68).
¥Mann–Whitney.

Table II Frequency of SNPs on genes in the p53 pathway analyzed in mothers of twins (cases) and singletons (controls).

SNP	Cases % (n)	Controls % (n)	OR (95% CI)	P-value
TP53 (rs1042522)	$n = 42$	$n = 101$		
GG	45.2 (19)	72.3 (73)	1.00	
CG	42.9 (18)	24.8 (25)	2.76 (1.2–6.08)	0.017
CC	11.9 (5)	3.0 (3)	6.40 (1.4–29.2)	0.026
G	66.7 (56)	84.7 (171)		0.0006#
C	33.3 (28)	15.3 (31)		
MDM4 (rs1563828)	$n = 42$	$n = 101$		
CC	28.6 (12)	43.6 (44)	1.00	
TC	45.2 (19)	43.6 (44)	1.58 (0.68–3.64)	0.382
TT	26.2 (11)	12.9 (13)	3.10 (1.11–8.6)	0.052
C	51.2 (43)	65.3 (132)		0.025#
T	48.8 (41)	34.7 (70)		
MDM2 (rs2279744)	$n = 40$	$n = 95$		
TT	35.0 (14)	44.2 (42)	1.00	
TG	45.0 (18)	42.1 (40)	1.35 (0.59–3.07)	0.611
GG	20.0 (8)	13.7 (13)	1.84 (0.63–5.37)	0.395
G	42.5 (34)	34.7 (66)		0.227#
T	57.5 (46)	65.3 (124)		
HAUSP (rs1529916)	$n = 42$	$n = 101$		
CC	47.6 (20)	57.4 (58)	1.00	
CT	45.2 (19)	39.6 (40)	1.37 (0.65–2.9)	0.514
TT	7.1 (3)	3.0 (3)	2.90 (0.54–15.5)	0.415
C	70.2 (59)	77.2 (156)		0.212#
T	29.8 (25)	22.8 (46)		
LIF (rs929271)	$n = 42$	$n = 101$		
TT	40.5 (17)	45.5 (46)	1.00	
TG	47.6(20)	48.5 (49)	1.10 (0.51–2.36)	0.951
GG	11.9 (5)	5.9 (6)	2.25 (0.60–8.36)	0.379
G	35.7 (30)	30.2 (61)		0.361#
T	64.3 (54)	69.8 (141)		

SNP, single nucleotide polymorphism.
#Chi-square, Bonferroni-corrected P-value 0.01.

Based on the results presented above we performed a logistic regression analysis considering the number of pregnancies, genotypes at *TP53* (coded as co-dominant) and *MDM4* (coded as recessive; Table III). When these three variables were analyzed together, both *TP53* and the number of pregnancies maintained their association with twinning ($P = 0.004$ and $P = 0.002$, respectively), with *TP53* having a higher OR than number of pregnancies (2.73 versus 1.70, respectively). Also, the logistic regression analysis suggested that there is no association between genotypes at the *MDM4* locus and twinning when considering the other variables. Because *MDM4* and *TP53* are in the same pathway, and both exhibited different frequencies between cases and controls in the univariate analysis, an interactive effect of these two SNPs was investigated, but was non-significant ($P = 0.966$). This suggests that these polymorphisms have independent effects on twinning in this community, even though the risk associated with the *TP53* gene seems to be much stronger than the risk associated with the *MDM4* gene as suggested by the logistic regression (Table III). We also compared allelic frequencies between mothers of dizygotic and monozygotic twins for SNPs at *TP53* and *MDM4*. However, no difference between these two groups was observed ($P = 0.286$ and $P = 0.827$, respectively).

As the P72 allele is known for its association with development or progression of tumors (Wu *et al.*, 2002; Harris and Levine, 2005; Fang *et al.*, 2010), we compared the incidence of cancer in first-degree relatives of cases and controls. The mothers of twins have three times more cases of cancer in their first-degree relatives than the control mothers ($P = 0.011$). To evaluate the effects of population structure on our results, we checked the relatedness index *r* for cases, controls and between groups. The mean *r* value in our sample was always <1%, which means that >99% of individuals in any group have $r = 0$.

Discussion

CG is a small town in Brazil, remarkable for showing a high frequency of both DZ and MZ twins (Tagliani-Ribeiro *et al.*, 2011). Based on that, we decided to investigate candidate genes related to blastocyst implantation and/or *intra utero* embryo survival. Our set of candidate genes included not only *TP53* itself but also its upstream regulators *MDM2*, *MDM4* and *HAUSP*, and one of its targets, *LIF*. We found an association between two polymorphisms and twinning in this

Downloaded from http://humrep.oxfordjournals.org/ at UCHSC Denison Memorial Library Box A-003 on July 12, 2012

Table III Logistic regression analyses considering the number of pregnancies, genotypes at TP53 and MDM4.

	B	SE	Wald	OR (95% CI)	P-value
TP53	1.004	0.350	8.253	2.730 (1.7–5.41)	0.004
MDM4	−0.589	0.556	1.124	0.555 (0.18–1.65)	0.289
Npreg	0.531	0.169	9.906	1.701 (1.2–2.37)	0.002
Constant	−2.721	1.263	4.647	0.066	0.031

Npreg, number of pregnancies; B, coefficient of logistic regression; SE, standard error.

population, indicating that twinning in CG is somehow linked to reduced levels of p53-induced apoptosis.

Considering the univariate and the logistic regression analyses together, our results suggest that the P72 allele of *TP53* is the main risk factor for twinning. The P72 allele is weaker than the R72 allele in inducing apoptosis and suppressing cellular transformation but appears to be better at initiating senescence and cell cycle arrest (Pim and Banks, 2004; Bergamaschi *et al*., 2006; Jeong *et al*., 2010). Mdm4 is a negative regulator of p53 and co-operates with Mdm2 to inhibit the p53 activity in cellular response to DNA damage (Atwal *et al*., 2009; Wade and Wahl, 2009). *MDM4T* allele (*MDM4*-T) alters Mdm4 levels, resulting in decreased p53 activity and a less effective apoptosis (Atwal *et al*., 2009). Mdm4 also interacts with the other members of the p53 family, p63 and p73 (Kadakia *et al*., 2001; Zdzalik *et al*., 2010), which may also regulate maternal reproduction. p63 is important for maturation of the egg and p73 ensures normal mitosis in the developing blastocyst (Feng *et al*., 2011; Levine *et al*., 2011). Therefore, our results consistently suggest that twinning in CG is somehow linked to reduced levels of p53-induced apoptosis. Nonetheless, recent data associated these two SNPs with infertility (Kang *et al*., 2009). This could be expected from the fact that p53 directly regulates *LIF*, and therefore a less active form of p53 (P72) or reduced levels of p53 (induced by *MDM4*-T) negatively impact *LIF* expression levels leading to implantation failure. Actually, *TP53* P72 allele and *MDM4T* allele seems to be more frequent in women requesting IVF than in women with no reproduction difficulties (Kay *et al*., 2006; Kang *et al*., 2009; Feng *et al*., 2011). Recently, Sohr and Engeland (2011) reported that p53 selectively induces the expression of human chorionic gonadotrophin acting as a transcription factor for the hCGβ7 (*CGB7*) gene.

These apparently contradictory findings could be reconciled considering that *LIF* is not induced by p53 alone. Prokineticin 1 also induces its expression (Evans *et al*., 2009). Interleukin-1-β, tumor necrosis factor-α, transforming growth factor-β (TGF-β) and estrogen have also been shown to stimulate *LIF* (Stewart, 1995, 2007; Sawai *et al*., 1997). Moreover, the expression of *LIF* is critical for implantation (Chen *et al*., 2000; Lass *et al*., 2001), but it is not the only factor involved. Different studies have shown that there is a wide range of genes necessary for blastocyst implantation, including cyclooxygenase-2, heparin-binding epidermal growth factor, Dickkopf-1 and interleukin 11 (Evans *et al*., 2008). Thus, the impact of p53 over LIF levels may be variable across populations provided that other factors maintain normal LIF expression in spite of low p53. Second, the anti-apoptotic behavior of low levels of p53 in general and of the P72 allele in particular may indicate an effect after implantation, enhancing the chance for a double pregnancy to succeed to term. Twin conceptions in humans are not rare. However, the frequency of multiple pregnancies that continue to birth is considerably rarer. Multiple pregnancies may constitute >12% of all natural conceptions, of which ~2% survive to term as twins and ~12% result in single births (Boklage, 1990). Thus, it is reasonable to suppose that mechanisms which are less effective in inducing apoptosis could increase the probability of an embryo's survival. Such general mechanisms would be compatible with the increased prevalence of both DZ and MZ twin births, as observed in CG.

Furthermore, it is interesting to observe that P72 and MDM4-T allele are more prevalent in African populations (around 70%) than in European (around 25%) and Brazilian ones (around 30%; Beckman *et al*., 1994; Tornesello *et al*., 2005; Marcel and Hainaut, 2009). Coincidentally, African populations show the highest rate of twin births (Nylander, 1969; Little, 1988). However, recent data about ethnic differences in infertility indicates that women with African ancestry have significant reductions (25–38%) in live birth rates after IVF when compared with cohorts of women of European ancestry (Seifer *et al*., 2008, 2010; Fujimoto *et al*., 2010; Huddleston *et al*., 2010). This apparent contradiction provides further possible evidence that the role of p53 on human fertility is not related only to implantation period, but that this protein may be crucial after implantation, increasing the chance of embryonic survival.

Previous studies suggest that increased parity and a higher BMI are factors influencing multiple pregnancy (Hoekstra *et al*., 2008). However, in our study, only the number of pregnancies shows an association with twinning. In spite of this significant association, *TP53* seems to be the variable conferring the highest increase in the relative risk for twinning in this population as suggested by the logistic regression analysis. It would be very interesting to evaluate data on FSH levels in this population. However, as this is a population-based study and many women included in our sample are no longer at reproductive age, we have considered that measuring FSH levels at the present time would not be informative.

Despite the fact that twins show a decreased or unchanged cancer risk for different types of childhood tumors (Inskip *et al*., 1991; Puumala *et al*., 2009), the incidence of cancer in first-degree relatives of twin's mothers (47%) is almost twice that of cases (23%). This, in turn, is similar to that reported in other studies assessing cancer history based on relative's reports (Hall *et al*., 2001; Ramsey *et al*., 2006). In another study from the same Brazilian region studying 8880 women reported that cancer incidence in first-degree relatives was found to be 25% (Roth *et al*., 2009). Even though cancer history was not among the primary goals of this study, an overview of reported types of cancer suggests that among twin's families there seems to be a prevalence in environmental-related tumors, such as lung, laryngeal and gastric cancer, whereas in the controls there is a prevalence in skin, breast and colorectal carcinoma. To the best of our knowledge, there are no studies on this subject or in the incidence of p53-related tumors in twin's families, but our data suggests that this needs further investigation.

Although our sample size was limited by the total number of twins' mothers resident in such a small town as CG, there is also the advantage of investigating a population with low levels of gene flow and with a known history of a founder effect (Charow, 2004; Zlotogora, 2007). The control group was composed of mothers of singletons from the

Downloaded from http://humrep.oxfordjournals.org/ at UCHSC Denison Memorial Library Box A-003 on July 12, 2012

same location to minimize the interference of environmental factors, as both groups were exposed to the same kind of hazards and experiences. Moreover, the founder effect which happened during the colonization of CG may act as a confounding variable since all polymorphisms whose frequency changed due to this founder effect would result in false-positive association signals if a control group composed of women from other localities was used. In addition to that, the mean *r* value was always $<1\%$, therefore suggesting that relatedness alone cannot explain the difference between cases and controls. This is also corroborated by the fact that other polymorphisms presented a similar distribution between groups.

It must be emphasized that our findings were observed in a small population with a very specific history and ethnic background. Twinning is a complex phenomenon and therefore we could expect a wide variation in the relative risks of different genetic polymorphisms across populations, as variants conferring high risk in one population may only confer a minor risk in another. Nevertheless, the results presented in this study suggest a genetic explanation for the high increase in the twinning rate in CG and also suggest a new, testable, role for p53 on reproduction.

Acknowledgements

The authors wish to thank the community of CG for their hospitality and Dr Anthony Fensom for kindly reviewing this manuscript.

Authors' roles

A.T-R., D.D.P., P.A-P., N.J.R.F., L.S-F. and U.M. designed the study. M.O., M.Z-O., D.L., M.L.S.P. and V.R. participated in the acquisition of the data. A.T-R., N.J.R.F. and L.S-F. did the statistical analyses. All authors participated in the analysis and interpretation of data. A.T-R., N.J.R.F., L.S-F. and U.M. drafted the report. All authors participated in the critical revision of the intellectual content of the report.

Funding

This work was supported by a grant from National Institute of Science and Technology in Populational Medical Genetics (INAGEMP) and Fundo de Incentivo a Pesquisa e Eventos (FIPE-HCPA).

Conflict of interest

None declared.

References

Al-Hendy A, Moshynska O, Saxena A, Feyles V. Association between mutations of the follicle-stimulating-hormone receptor and repeated twinning. *Lancet* 2000;**356**:914.

Atwal GS, Kirchhoff T, Bond EE. Altered tumor formation and evolutionary selection of genetic variants in the human MDM4 oncogene. *Proc Natl Acad Sci USA* 2009;**106**:10236–10241.

Beckman G, Birgander R, Själander A, Saha N, Holmberg PA, Kivelä A, Beckman L. Is p53 polymorphism maintained by natural selection? *Hum Hered* 1994;**44**:266–270.

Bergamaschi D, Samuels Y, Sullivan A, Zvelebil M, Breyssens H, Bisso A, Del Sal G, Syed N, Smith P, Gasco M *et al.* iASPP preferentially binds p53 proline-rich region and modulates apoptotic function of codon 72-polymorphic p53. *Nat Genet* 2006;**38**:1133–1141.

Boklage CE. Survival probability of human conceptions from fertilization to term. *Int J Fertil* 1990;**35**:75, 79–80, 81–94.

Brooks CL, Li M, Hu M, Shi Y, Gu W. The p53-Mdm2-HAUSP complex is involved in p53 stabilization by HAUSP. *Oncogene* 2007; **26**:7262–7266.

Busjahn A, Knoblauch H, Faulhaber HD, Aydin A, Uhlmann R, Tumilehto J, Kaprio J, Jedruski P, Januszewics A, Strelau J *et al.* A region on chromosome 3 is linked to dizygotic twinning. *Nat Genet* 2000; **26**:398–399.

Charow J. Ashkenazi Jewish genetic disorders. *Fam Can* 2004;**3**:201–206.

Chen JR, Cheng JG, Shatzer T, Sewell L, Hernandez L, Stewart CL. Leukemia inhibitory factor can substitute for nidatory estrogen and is essential to inducing a receptive uterus for implantation but is not essential for subsequent embryogenesis. *Endocrinology* 2000;**141**:4365–4372.

Derom C, Groenen P, Vlietinck R. Follicle-stimulating-hormone receptor and twinning. *Lancet* 2001;**357**:230–231.

Derom C, Jawaheer D, Chen WV, McBride KL, Xiao X, Amos C, Gregersen PK, Vlietinck R. Genome-wide linkage scan for spontaneous DZ twinning. *Eur J Hum Genet* 2006;**14**:117–122.

Eliyahu D, Raz A, Gruss P, Givol D, Oren M. Participation of p53 cellular tumour antigen in transformation of normal embryonic cells. *Nature* 1984;**312**:646–649.

Evans J, Catalano RD, Morgan K, Critchley HO, Millar RP, Jabbour HN. Prokineticin 1 signaling and gene regulation in early human pregnancy. *Endocrinology* 2008;**149**:2877–2887.

Evans J, Catalano RD, Brown P, Sherwin R, Critchley HO, Fazleabas AT, Jabbour HN. Prokineticin 1 mediates fetal-maternal dialogue regulating endometrial leukemia inhibitory factor. *FASEB J* 2009; **23**:2165–2175.

Fang S, Krahe R, Lozano G, Han Y, Chen W, Post SM, Zhang B, Wilson CD, Bachinski LL, Strong LC *et al.* Effects of MDM2, MDM4 and TP53 codon 72 polymorphisms on cancer risk in a cohort study of carriers of TP53 germline mutations. *PLos One* 2010;**5**:e10813.

Feng Z, Zhang C, Kang HJ, Sun Y, Wang H, Naqsi A, Frank AK, Rosenwaks Z, Murphy ME, Levive AJ *et al.* Regulation of female reproduction by p53 and its family members. *FASEB J* 2011; **25**:2245–2255.

Finlay CA, Hinds PW, Tan TH, Eliyahu D, Oren M, Levine AJ. Activating mutations for transformation by p53 produce a gene product that forms an hsc70-p53 complex with an altered half-life. *Mol Cell Biol* 1988;**8**:531–539.

Fujimoto VY, Luke B, Brown MB, Jain T, Armstrong A, Grainger DA, Hornstein MD. Racial and ethnic disparities in assisted reproductive technology outcomes in the United States. *Fertil Steril* 2010; **93**:382–390.

Hall IJ, Burke W, Coughlin S, Lee NC. Population-based estimates of the prevalence of family history of cancer among women. *Community Genet* 2001;**4**:134–142.

Harris SL, Levine AJ. The p53 pathway: positive and negative feedback loops. *Oncogene* 2005;**24**:2899–2908.

Hasbargen U, Lohse P, Thaler CJ. The number of dichorionic twin pregnancies is reduced by the common MTHFR 677C–>T mutation. *Hum Reprod* 2000;**15**:2659–2662.

Hoekstra C, Zhao ZZ, Lambalk CB, Willemsen G, Martin NG, Boomsma DI, Montgomery GW. Dizygotic twinning. *Hum Reprod Update* 2008;**14**:37–47.

Hu W, Feng Z, Teresky AK, Levine AJ. p53 regulates maternal reproduction through LIF. *Nature* 2007;**450**:721–724.

Huddleston HG, Cedars MI, Sohn SH, Giudice LC, Fujimoto VY. Racial and ethnic disparities in reproductive endocrinology and infertility. *Am J Obstet Gynecol* 2010;**202**:413–419.

Downloaded from http://humrep.oxfordjournals.org/ at UCHSC Denison Memorial Library Box A-003 on July 12, 2012

Inskip PD, Harvey EB, Boice JD Jr, Stone BJ, Matanoski G, Flannery JT, Fraumeni JF Jr. Incidence of childhood cancer in twins. *Cancer Causes Control* 1991;**2**:315–324.

Jeong BS, Hu W, Belyi V, Rabadan R, Levine AJ. Differential levels of transcription of p53-regulated genes by the arginine/proline polymorphism: p53 with arginine at codon 72 favors apoptosis. *FASEB J* 2010;**24**:1347–1353.

Kadakia M, Slader C, Berberich SJ. Regulation of p63 function by Mdm2 and MdmX. *DNA Cell Biol* 2001;**20**:321–330.

Kang HJ, Feng Z, Sun Y, Atwal G, Murphy ME, Rebbeck TR, Rosenwaks Z, Levine AJ, Hu W. Single-nucleotide polymorphisms in the p53 pathway regulate fertility in humans. *Proc Natl Acad Sci USA* 2009;**106**:1–6.

Kay C, Jeyendran RS, Coulam CB. p53 tumour suppressor gene polymorphism is associated with recurrent implantation failure. *Reprod Biomed Online* 2006;**13**:492–496.

Kress M, May E, Cassingena R, May P. Simian virus 40-transformed cells express new species of proteins precipitable by anti-simian virus 40 tumor serum. *J Virol* 1979;**31**:472–483.

Lahiri DK, Nurnberger JI Jr. A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. *Nucleic Acids Res* 1991;**19**:5444.

Lass A, Weiser W, Munafo A, Loumaye E. Leukemia inhibitory factor in human reproduction. *Fertil Steril* 2001;**76**:1091–1096.

Levine AJ, Hu W, Feng Z. The p53 pathway: what questions remain to be explored? *Cell Death Differ* 2006;**13**:1027–1236.

Levine AJ, Tomasini R, McKeon FD, Mak TW, Melino G. The p53 family: guardians of maternal reproduction. *Nat Rev Mol Cell Biol* 2011; **12**:259–265.

Linzer DI, Levine AJ. Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. *Cell* 1979;**17**:43–52.

Little J. (1988) Descriptive epidemiology. In MacGillivray I, Campbell DM, Thompson B (eds), *Twinning and Twins*. John Wiley & Sons Ltd, New York, NY.

Mandinova A, Lee SW. The p53 pathway as a target in cancer therapeutics: obstacles and promise. *Sci Transl Med* 2011;**3**:64rv1.

Manfredi JJ. The Mdm2–p53 relationship evolves: Mdm2 swings both ways as an oncogene and a tumor suppressor. *Genes Dev* 2010;**24**:1580–1589.

Marcel V, Hainaut P. p53 isoforms—a conspiracy to kidnap p53 tumor suppressor activity? *Cell Mol Life Sci* 2009;**66**:391–406.

Matte U, Le Roux MG, Bénichou B, Moisan JP, Giugliani R. Study on possible increase in twinning rate at a small village in south Brazil. *Acta Genet Med Gemellol* 1996;**45**:431–437.

Montgomery GW, Duffy DL, Hall J, Kudo M, Martin NG, Hsueh AJ. Mutations in the follicle-stimulating hormone receptor and familial dizygotic twinning. *Lancet* 2001;**357**:773–774.

Montgomery GW, Zhao ZZ, Morley KI, Marsh AJ, Boomsma DI, Martin NG, Duffy DL. Dizygotic twinning is not associated with methylenetetrahydrofolate reductase haplotypes. *Hum Reprod* 2003; **18**:2460–2464.

Montgomery GW, Zhao ZZ, Marsh AJ, Mayne R, Treloar SA, James M, Martin NG, Boomsma DI, Duffy DL. A deletion mutation in GDF9 in sisters with spontaneous DZ twins. *Twin Res* 2004;**7**:548–555.

Norimura T, Nomoto S, Katsuki M, Gondo Y, Kondo S. p53-dependent apoptosis suppresses radiation-induced teratogenesis. *Nat Med* 1996; **2**:577–580.

Nylander PP. The frequency of twinning in a rural community in Western Nigeria. *Ann Hum Genet* 1969;**33**:41–44.

Painter JN, Willemsen G, Nyholt D, Hoekstra C, Duffy DL, Henders AK, Wallace L, Healey S, Cannon-Albright LA, Skolnick M *et al.* A genome wide linkage scan for dizygotic twinning in 525 families of mothers of dizygotic twins. *Hum Reprod* 2010;**25**:1569–1580.

Palmer JS, Zhao ZZ, Hoekstra C, Hayward NK, Webb PM, Whitman DC, Martin NG, Boomsma Di, Duffy DL, Montgomery GW. Novel variants in growth differentiation factor 9 in mothers of dizygotic twins. *J Clin Endocrinol Metab* 2006;**91**:4713–4716.

Pim D, Banks L. p53 polymorphic variants at codon 72 exert different effects on cell cycle progression. *Int J Cancer* 2004;**108**:196–199.

Puumala SE, Spector LG, Robison LL, Bunin GR, Olshan AF, Linabery AM, Roesler MA, Blair CK, Ross JA. Comparability and representativeness of control groups in a case-control study of infant leukemia: a report from the Children's Oncology Group. *Am J Epidemiol* 2009;**1**:379–387.

Ramsey SD, Yoon P, Moonesinghe R, Khoury MJ. Population-based study of the prevalence of family history of cancer: implications for cancer screening and prevention. *Genet Med* 2006;**8**:571–575.

Roche Jean. *A colonização alemã e o Rio Grande do Sul. Brasil*, 1st edn. Porto Alegre, Brasil: Globo, 1969.

Roth FL, Camey SA, Caleffi M, Schuler-Faccini L, Palmero EI, Bochi C, Moreira SM, Kalakun L, Giugliani R, Ashton-Prolla P. Consistency of self-reported first-degree family history of cancer in a population-based study. *Fam Cancer* 2009;**8**:195–202.

Sawai K, Matsuzaki N, Okada T, Shimoya K, Koyama M, Azuma C, Saji F, Murata Y. Human decidual cell biosynthesis of leukemia inhibitory factor: regulation by decidual cytokines and steroid hormones. *Biol Reprod* 1997;**56**:1274–1280.

Seifer DB, Frazier LM, Grainger DA. Disparity in assisted reproductive technologies outcomes in black women compared with white women. *Fertil Steril* 2008;**90**:1701–1710.

Seifer DB, Zackula R, Grainger DA. Trends of racial disparities in assisted reproductive technology outcomes in black women compared with white women: Society for Assisted Reproductive Technology 1999 and 2000 vs. 2004–2006. *Fertil Steril* 2010;**93**:626–635.

Sohr S, Engeland K. The tumor suppressor p53 induces expression of the pregnancy-supporting human chorionic gonadotropin (hCG) *CGB7* gene. *Cell Cycle* 2011;**10**:3758–3767.

Stewart CL. Leukaemia inhibitory factor and the regulation of pre-implantation development of the mammalian embryo. *Mol Reprod Dev* 1995;**39**:233–238.

Stewart CL. Reproduction: the unusual suspect. *Nature* 2007;**450**:619.

Tagliani-Ribeiro A, Oliveira M, Sassi AK, Rodrigues MR, Zagonel-Oliveira M, Steinman G, Matte U, Fagundes NJR, Schuler-Faccini L. Twin town in South Brazil: a Nazi's experiment or a genetic founder effect? *PLoS One* 2011;**6**:e20328.

Tornesello ML, Waddell KM, Duraturo ML, Biryahwaho B, Downing R, Lucas SB, Giani U, Buonaguro L, Buonaguro FM. TP53 codon 72 polymorphism and risk of conjunctival squamous cell carcinoma in Uganda. *Cancer Detect Prev* 2005;**29**:501–508.

Ucisik-Akkaya E, Davis CF, Do TN, Morrison BA, Stemmer SM, Amadio Wj, Dorak MT. Examination of genetic polymorphisms in newborns for signatures of sex-specific prenatal selection. *Mol Hum Reprod* 2010;**16**:770–777.

Wade M, Wahl GM. Targeting Mdm2 and Mdmx in cancer therapy: better living through medicinal chemistry? *Mol Cancer Res* 2009;**7**:1–11.

Wade M, Wang YV, Wahl GM. The p53 orchestra: Mdm2 and Mdmx set the tone. *Trends Cell Biol* 2010;**20**:299–309.

Wu X, Zhao H, Amos CI, Shete S, Makan N, Hong WK, Kadlubar FF, Spitz MR. p53 genotypes and haplotypes associated with lung cancer susceptibility and ethnicity. *J Natl Cancer Inst* 2002;**94**:681–690.

Zdzalik M, Pustelny K, Kedracka-Krok S, Huben K, Pecak A, Wladyka B, Jankowski S, Dubin A, Potempa J, Dubin G. Interaction of regulators Mdm2 and Mdmx with transcription factors p53, p63 and p73. *Cell Cycle* 2010;**15**:4584–4591.

Zlotogora J. Multiple mutations responsible for frequent genetic diseases in isolated populations. *Eur J Hum Gen* 2007;**15**:272–278.

Downloaded from http://humrep.oxfordjournals.org/ at UCHSC Denison Memorial Library Box A-003 on July 12, 2012

10.2 *COMPREHENSIVE GENOMIC ANALYSIS IDENTIFIES SOX2 AS A FREQUENTLY AMPLIFIED GENE IN SMALL-CELL LUNG CANCER*

Título do manuscrito: *Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer*

Autores: Rudin CM, Durinck S, Stawiski EW, Poirier JT, Modrusan Z, Shames DS, Bergbower EA, Guan Y, Shin J, Guillory J, Rivers CS, Foo CK, Bhatt D, Stinson J, Gnad F, Haverty PM, Gentleman R, Chaudhuri S, Janakiraman V, Jaiswal BS, Parikh C, Yuan W, Zhang Z, Koeppen H, Wu TD, Stern HM, Yauch RL, Huffman KE, Paskulin DD, Illei PB, Varella-Garcia M, Gazdar AF, de Sauvage FJ, Bourgon R, Minna JD, Brock MV, Seshagiri S.

Situação: Publicado

Revista: *Nature Genetics*

Referência: Nat Genet. 2012 Oct;44(10):1111-6. doi: 10.1038/ng.2405. Epub 2012 Sep 2.

Website: http://www.ncbi.nlm.nih.gov/pubmed/22941189

nature genetics

Comprehensive genomic analysis identifies *SOX2* as a frequently amplified gene in small-cell lung cancer

Charles M Rudin[1,8], Steffen Durinck[2,3,8], Eric W Stawiski[2,3,8], John T Poirier[1,8], Zora Modrusan[2,8], David S Shames[4,8], Emily A Bergbower[1], Yinghui Guan[2], James Shin[1], Joseph Guillory[2], Celina Sanchez Rivers[2], Catherine K Foo[2], Deepali Bhatt[2], Jeremy Stinson[2], Florian Gnad[3], Peter M Haverty[3], Robert Gentleman[3], Subhra Chaudhuri[2], Vasantharajan Janakiraman[2], Bijay S Jaiswal[2], Chaitali Parikh[2], Wenlin Yuan[2], Zemin Zhang[3], Hartmut Koeppen[5], Thomas D Wu[3], Howard M Stern[5], Robert L Yauch[4], Kenneth E Huffman[6], Diego D Paskulin[7], Peter B Illei[1], Marileila Varella-Garcia[7], Adi F Gazdar[6], Frederic J de Sauvage[2], Richard Bourgon[3], John D Minna[6], Malcolm V Brock[1] & Somasekar Seshagiri[2]

© 2012 Nature America, Inc. All rights reserved.

Small-cell lung cancer (SCLC) is an exceptionally aggressive disease with poor prognosis. Here, we obtained exome, transcriptome and copy-number alteration data from approximately 53 samples consisting of 36 primary human SCLC and normal tissue pairs and 17 matched SCLC and lymphoblastoid cell lines. We also obtained data for 4 primary tumors and 23 SCLC cell lines. We identified 22 significantly mutated genes in SCLC, including genes encoding kinases, G protein–coupled receptors and chromatin-modifying proteins. We found that several members of the SOX family of genes were mutated in SCLC. We also found *SOX2* amplification in ~27% of the samples. Suppression of *SOX2* using shRNAs blocked proliferation of *SOX2*-amplified SCLC lines. RNA sequencing identified multiple fusion transcripts and a recurrent *RLF-MYCL1* fusion. Silencing of *MYCL1* in SCLC cell lines that had the *RLF-MYCL1* fusion decreased cell proliferation. These data provide an in-depth view of the spectrum of genomic alterations in SCLC and identify several potential targets for therapeutic intervention.

Lung cancer is the leading cause of cancer mortality in the United States, where it is responsible for over 160,000 deaths annually[1]. Approximately 10–15% of the new lung cancer cases diagnosed each year are SCLC[2]. The genomic landscape of SCLC is of particular interest compared to those of other solid tumors, given the unique biological characteristics of this tumor type[3]. SCLC is an exceptionally aggressive malignancy with a high proliferative index and an unusually strong predilection for early metastasis.

Previous efforts to characterize the genetic alterations present in SCLC tumors identified high prevalence of inactivating mutations in *TP53* (75–90%)[4], *RB1* (60–90%)[5,6] and *PTEN* (2–4%)[7], rare activating mutations in *PIK3CA*, *EGFR* and *KRAS*[8–10], amplification of *MYC* family members, *EGFR* and *BCL2*, and loss of *RASSF1A*, *PTEN* and *FHIT*[6,11].

A better understanding of the genomic changes in this cancer will be essential to developing new therapeutics. To this end, we have applied next-generation sequencing technologies to characterize multiple exomes and a single genome of primary SCLC, as well as exomes of SCLC cell lines, together with genome-wide copy-number analysis and whole-transcriptome sequencing.

Specifically, we characterized 80 human SCLCs, including 36 primary SCLC human tumor and adjacent normal sample pairs and 17 paired SCLC cell lines and their patient-matched lymphoblastoid cell lines, as well as 4 primary SCLC tumors and 23 SCLC cell lines without matched normal controls (**Supplementary Table 1**). We sequenced and analyzed the complete genome of one SCLC tumor–normal tissue pair.

Exome capture, sequencing and analysis of 42 SCLC tumor–normal tissue pairs identified 26,406 somatic mutations. Approximately 30% (7,977) of these mutations were protein altering (**Fig. 1a** and **Supplementary Table 2**). The somatic mutations identified included 7,154 missense, 536 nonsense, 12 stop loss, 243 essential splice site, 32 protein-altering insertion and/or deletion (indel), 2,674 synonymous, 11,460 intronic and 4,295 other types (**Fig. 1a** and **Supplementary Tables 3** and **4**). Comparison of the protein-altering changes identified in this study with those reported in the Catalogue of Somatic Mutations in Cancer (COSMIC)[12] showed that 98% (7,824/7,977) of these variations are newly identified somatic changes. Nineteen percent of the protein-altering somatic mutations reported were validated using RNA sequencing (RNA-seq) data or mass spectrometry genotyping, with a validation rate of 91% (**Supplementary Table 3**). We confirmed the effect of several splice-site mutations using

[1]The Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University, Baltimore, Maryland, USA. [2]Department of Molecular Biology, Genentech, South San Francisco, California, USA. [3]Department of Bioinformatics and Computational Biology, Genentech, South San Francisco, California, USA. [4]Department of Oncology Biomarker Development, Genentech, South San Francisco, California, USA. [5]Department of Pathology, Genentech, South San Francisco, California, USA. [6]Hamon Center for Therapeutic Oncology Research, University of Texas Southwestern Medical Center, Dallas, Texas, USA. [7]Division of Medical Oncology, University of Colorado Cancer Center, Aurora, Colorado, USA. [8]These authors contributed equally to this work. Correspondence should be addressed to C.M.R. (rudin@jhmi.edu) or S.S. (sekar@gene.com).

Received 4 June; accepted 10 August; published online 2 September 2012; doi:10.1038/ng.2405

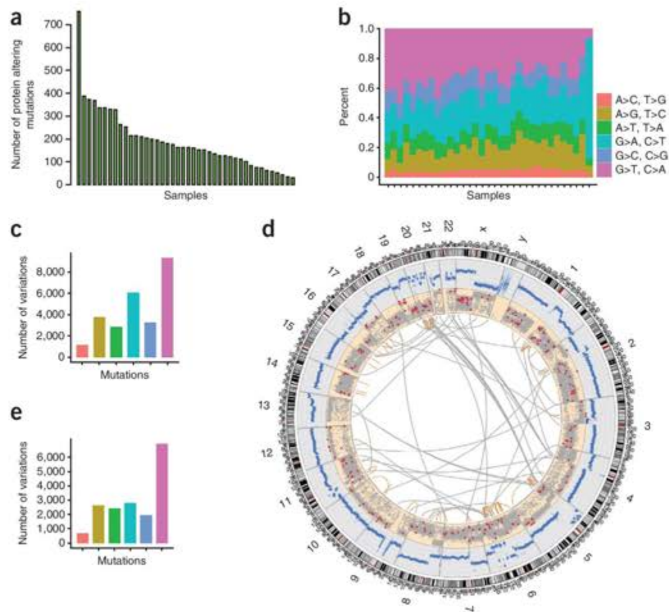


Figure 1 SCLC somatic mutations. (**a**) Histogram of the number of mutations in each primary tumor sample. (**b**) Base-level transitions and transversions in each SCLC sample shown in **a**. (**c**) Average number of transitions and transversions in the SCLC samples based on the exome sequencing data. (**d**) Whole genome of an SCLC sample shown as a Circos plot. Copy-number changes measured using sequencing reads are shown in blue. Somatic nonsynonymous, splice-site and stop-gain mutations are shown as red dots. Other somatic mutations are depicted as gray dots. Intra- (orange lines) and interchromosomal (gray lines) rearrangements are also shown. (**e**) Average number of transitions and transversions in the whole-genome sequence of an SCLC sample. Colors in **c** and **e** correspond to those defined in **b**.

RNA-seq data (**Supplementary Table 3**). We validated all of the indels reported using Sanger sequencing (**Supplementary Table 4**). One sample represented a distinct profile, with 2,953 mutations (757 validated protein-altering variants; **Fig. 1a** and **Supplementary Table 3**). Given the exceptionally high number of mutations in this sample, we excluded it from our calculations of the background mutation rate. Excluding the hypermutated sample, the SCLC tumors had an average of 175 protein-altering single-nucleotide variants (range 31–388) with a mean nonsynonymous mutation rate of 5.5 mutations per megabase (**Fig. 1a**). This is comparable to the 92 protein-altering variants observed in the previously sequenced genome of a single SCLC cell line[13].

Analysis of the base-level transitions and transversions showed that G-to-T transversions were predominant, followed in prevalence by G-to-A and A-to-G transitions (**Fig. 1b**), both at the exome (**Fig. 1c**) and whole-genome (**Fig. 1d,e** and **Supplementary Fig. 1**) levels. This pattern is consistent with demonstrated effects of tobacco smoke carcinogens on DNA[13].

In assessing the whole-genome data for an SCLC tumor–normal tissue pair, we found 59,784 somatic mutations, of which 286 were protein-altering changes (256 missense, 19 nonsense, 11 essential splice site, 77 synonymous, 13,924 intronic and 45,497 others). The average whole-genome mutation rate was 21.34 mutations per megabase (**Fig. 1d**). Previously, 22,910 somatic variants were reported for the NCI-H209 SCLC cell line[13].

Our mutation analysis identified protein-altering somatic single-nucleotide variants in 5,179 genes, including 4,775 genes that were mutated in the non-hypermutated SCLC sample set. Frequently mutated classes included genes encoding kinases, G protein–coupled receptors and chromatin-modifying proteins. To further understand the impact of the mutations on gene function, we applied SIFT[14], Polyphen[15] and Condel[16] and found that ~53% of the somatic mutations identified are likely to have functional consequences according to at least two of the three methods (**Supplementary Table 3**). In contrast, only approximately 17% of germline variants identified in the normal samples are predicted by these methods to have a functional impact (**Supplementary Fig. 2**).

To further assess the relevance of mutated genes, we applied a q-score metric[17] to rank significantly mutated cancer-associated genes. We identified 22 significantly mutated genes in SCLC (q score $\geq$ 1; false discovery rate $\leq$ 10%; **Supplementary Table 5**). These genes included *TP53* and *RB1* and several genes that have not previously been reported as mutated in SCLC (**Fig. 2a** and **Supplementary Table 5**).

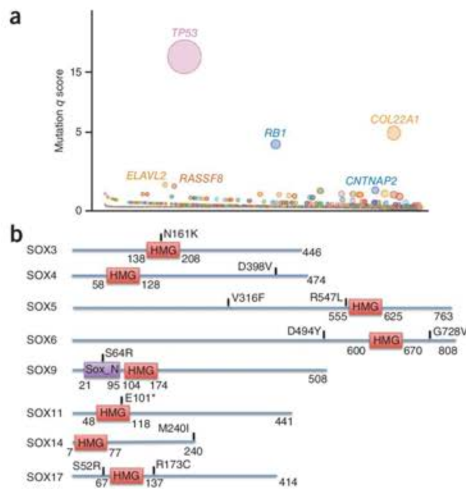


Figure 2 Significantly mutated genes in SCLC. (**a**) Genes evaluated for significance on the basis of q score are shown. Each gene is represented as a circle, where the size of the circle is proportional to the observed frequency of mutation in that gene. Genes are arranged on the x axis in order of increasing number of expected mutations from left to right. Genes with significant q scores are labeled. (**b**) Alterations affecting the SOX family. *, nonsense change; HMG, high-mobility group; Sox_N, Sox developmental protein N terminal.

© 2012 Nature America, Inc. All rights reserved.

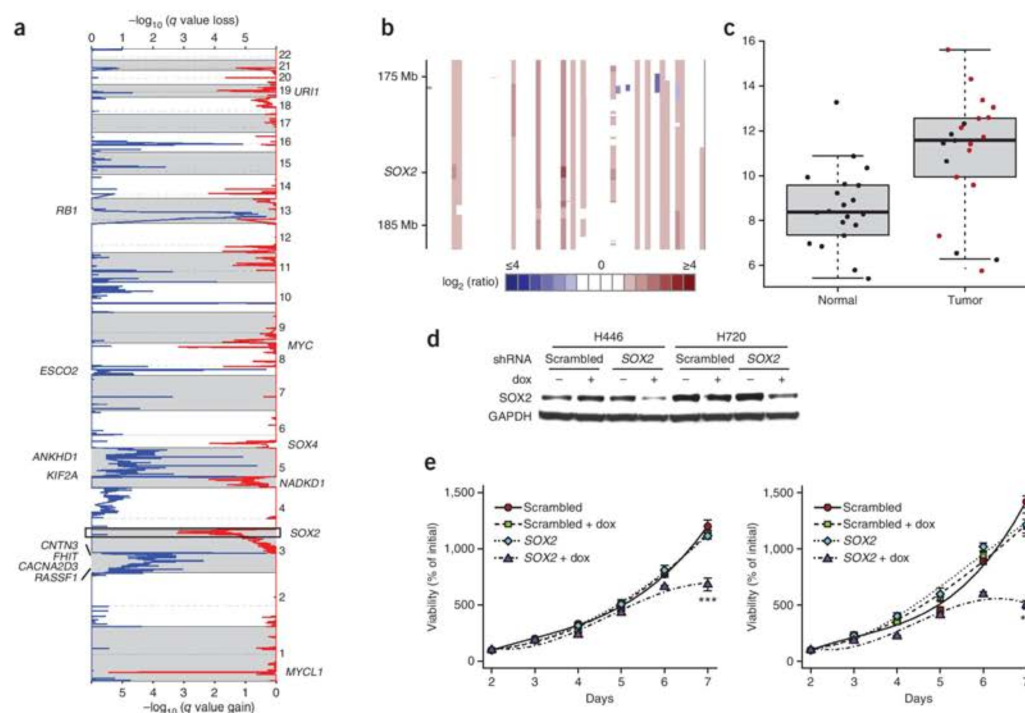


Figure 3 *SOX2* is amplified in SCLC and drives proliferation. (**a**) GISTIC plot depicting recurrent amplifications in SCLC samples ($n = 56$) with copy-number data. (**b**) Heatmap of segmented copy-number $\log_2$ (ratio) values from the 3q chromosomal region containing the *SOX2* locus. (**c**) Box plot of *SOX2* expression in SCLC and adjacent normal samples measured by RNA-seq. Samples with *SOX2* amplification are highlighted in red. Error bars at the top indicate the maximum values excluding outliers, and error bars at the bottom indicate the minimum values excluding outliers. Outliers are defined as values more than the third quartile +1.5 × IQR or less than the first quartile −1.5 × IQR, where IQR is the innerquartile range. (**d**,**e**) Doxycycline-inducible shRNA targeting of *SOX2* suppresses SOX2 protein levels (**d**) and inhibits cell proliferation (**e**) in H460 and H720 SCLC lines compared to scrambled control shRNA. Error bars in **e**, s.e.m. $^{**}P < 0.01$; $^{***}P < 0.001$.

© 2012 Nature America, Inc. All rights reserved.

To further confirm the relevance of the 22 genes, we assessed the mutation frequency for these genes using exome data from a set of 21 additional samples (**Supplementary Table 6**). We found a significant correlation between the mutation frequencies of the 22 genes in the initial sample set and the validation cohort ($P = 1.16 \times 10^{-5}$, $r^2 = 0.63$; **Supplementary Table 7**). In addition, we found that 42 genes that were mutated in our primary tumor samples (**Supplementary Table 8**) were also previously reported to be mutated in the genome of the NCI-H209 SCLC cell line[13].

Mutational hotspots are indicative of genes that are relevant to cancer. In this study, we have identified 17 genes with 18 hotspot mutations (**Supplementary Table 9**). By comparing our mutations with those reported in COSMIC[12] and a large-scale colon cancer mutation screen[18], we identified an additional 150 hotspot mutations in 116 genes (**Supplementary Table 9**). Besides known hotspots in *TP53*, *RB1*, *PIK3CA*, *CDKN2A* and *PTEN*, several new hotspot mutations were identified. These included genes encoding Ras family regulators (*RAB37*, *RASGRF1* and *RASGRF2*), chromatin-modifying enzymes or transcriptional regulators (*EP300*, *DMBX1*, *MLL2*, *MED12L*, *TRRAP* and *RUNX1T1*), ionotropic glutamate receptor (*GRID1*), kinases (*STK38*, *LRRK2*, *PRKD3* and *CDK14*), protein phosphatases (*PTPRD* and *PPEF2*) and G protein–coupled receptors (*GPR55*, *GPR113* and *GPR133*). Further, three of the genes with the top *q* scores—*RUNX1T1*, *CDYL* and *RIMS2*—contained a hotspot mutation.

In addition to the hotspots, we found mutations clustering in particular gene families and pathways (**Supplementary Table 10**). Evidence of clustering was found in genes in the phosphatidylinositol 3-kinase (PI3K) pathway (*PIK3CA*, *AKT1–3*, *MTOR*, *RPS6KA2* and *RPS6KA6*), the mediator complex (*MED12*, *MED12L*, *MED13*, *MED13L*, *MED15*, *MED24*, *MED25*, *MED27* and *MED29*), Notch and Hedgehog family members (*NOTCH1*, *NOTCH2*, *NOTCH3* and *SMO*), glutamate receptor family members (*GRIA1*, *GRIA2*, *GRIA3*, *GRIA4*, *GRIND1*, *GRID2* and *GRM1–3*, *GRM5*, *GRM7* and *GRM8*), SOX family members (*SOX3*, *SOX4*, *SOX5*, *SOX6*, *SOX9*, *SOX11*, *SOX14* and *SOX17*; **Fig. 2b**) and DNA repair and/or checkpoint pathway genes (*ATM*, *ATR*, *CHEK1* and *CHEK2*). The mutations in SOX family members were mutually exclusive (**Supplementary Fig. 3**). In contrast to non–small-cell lung cancer (NSCLC)[12], we did not observe any SCLC samples with a *KRAS* mutation. Among the receptor tyrosine kinase genes, we identified mutations in *FLT1*, *FLT4*, *KDR* and *KIT* and members of the Ephrin family (*EPHA1–7* and *EPHB4*). Notably,

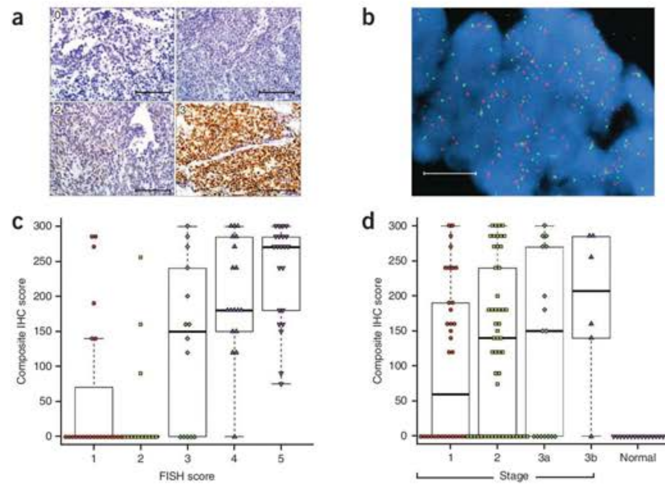


Figure 4 SOX2 gene amplification and protein expression in SCLC. SOX2 protein expression was assessed by IHC, and *SOX2* gene copy number was assessed by FISH on a set of 110 SCLC cases and 15 normal lung controls. (**a**) Representative images showing variability of staining intensity by IHC, from 0 to 3. Scale bars, 100 μm. (**b**) Representative image showing very high *SOX2* copy number by FISH. Red, *SOX2* probe; green, centromeric probe. Scale bar, 10 μm. (**c**) Correlation between SOX2 IHC score (staining intensity × percent with positively stained nuclei) and *SOX2* FISH score (1–6). (**d**) Composite SOX2 IHC score of SCLC samples by stage and normal lung controls. Plots in **c** and **d** are box plots where the box encloses the first to third quartiles, the bar inside the box represents the median, the whisker at the top indicates the maximum value excluding outliers and the whisker at the bottom indicates the minimum value excluding outliers. Outliers are defined as values more than the third quartile +1.5 × IQR or less than the first quartile −1.5 × IQR, where IQR is the interquartile range.

© 2012 Nature America, Inc. All rights reserved.

the *KIT* mutation affecting codon 761 has previously been reported in mast cell activation disorder and is likely an activating change[19] (**Supplementary Fig. 4**).

Chromosomal copy-number analysis of 56 SCLC samples identified recurrent copy gains and losses (**Supplementary Tables 11** and **12**). Genes with copy-number loss included the previously reported *RB1*, *RASSF1* and *FHIT* (**Fig. 3**) and several genes not previously known to be altered in SCLC, including *KIF2A* and *CNTN3* (refs. 6,20). Among the genes with recurrent copy-number gain, we confirmed previously reported amplifications involving *MYC*, *SOX4* and *KIT* (**Fig. 3a**, **Supplementary Fig. 4b** and **Supplementary Table 11**)[6,20,21].

In addition, we identified high levels of amplification (copy number of ≥4) of *SOX2* in ~27% (15/56) of the SCLC samples (**Fig. 3b**). RNA-seq data showed that the majority of the SCLC samples, including those with *SOX2* amplification, had higher *SOX2* expression compared to adjacent normal samples (**Fig. 3c**). We further examined the expression of SOX2 by immunohistochemistry (IHC) and copy-number change by FISH in an independent cohort of 110 primary SCLC tumor samples (**Fig. 4a,b**). Expression of SOX2 was strongly correlated with increased gene copy number and with clinical stage (**Fig. 4c,d**).

To further assess the relevance of *SOX2* in SCLC, we analyzed a panel of SCLC cell lines for SOX2 protein expression and gene copy number (**Supplementary Fig. 5**). Among these cell lines, H446 and H720 both had strong SOX2 protein expression, and H720 was found to have elevated gene copy number. SOX2 has previously been implicated in the maintenance of proliferative potential and stem cell function[22–25]. To test whether H446 and H720 were dependent on SOX2 for continued growth and proliferation, we stably transduced them with lentiviruses carrying either a doxycycline-inducible *SOX2*-targeting short hairpin RNA (shRNA) or a scrambled control shRNA. Induction of *SOX2* shRNA in both H446 and H720 resulted in lower amounts of SOX2 protein and reduced cell proliferation (**Fig. 3d,e**). Previously, amplification of *SOX2* and its role as an oncogene have been reported in lung and esophageal squamous cell carcinoma[26]. Our findings further support the idea of *SOX2* as a genuine SCLC driver gene.

Analysis of RNA-seq data obtained from SCLC samples for fusion transcripts

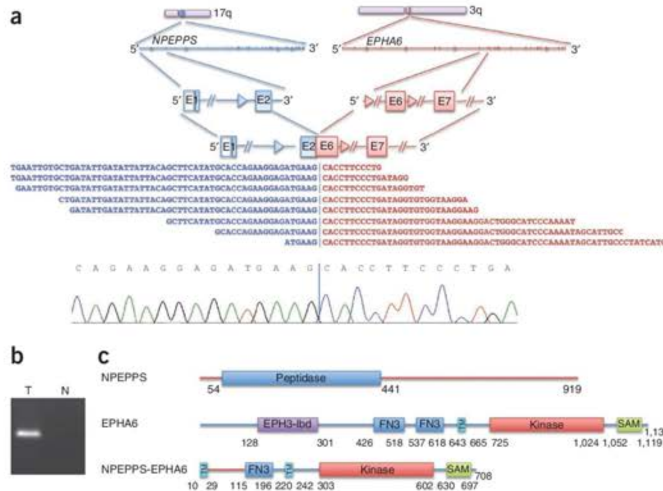


Figure 5 Kinase fusions. (**a**) *NPEPPS-EPHA6* fusion identified using RNA-seq (top) along with a representative Sanger sequencing chromatogram derived from this fusion product (bottom). E, exon. (**b**) Independent product derived by RT-PCR confirming the *NPEPPS-EPHA6* somatic fusion resolved on an agarose gel. RT-PCR was performed on a tumor (T) and normal (N) sample. (**c**) Schematic of the NPEPPS-EPHA6 fusion protein. EPH3-lbd, Ephrin receptor ligand–binding domain; FN3, fibronectin type 3 domain; TM, transmembrane domain; SAM, sterile α motif.

identified 41 gene fusions, including 4 recurrent fusions (**Supplementary Table 13**). A majority of the predicted gene fusions were intrachromosomal (83%, 34/41). All of the gene fusions reported were verified and confirmed to be somatic by RT-PCR **Supplementary Table 13**). A fusion involving *RLF* and *MYCL1* (**Supplementary Fig. 6a**) was found in one primary SCLC tumor and four SCLC cell lines (H889, HCC33, H1092 and COR-L47). *RLF* and *MYCL1* are ~259 kb apart and are encoded by opposing strands. The observed fusion requires an inversion event that brings exon 1 of *RLF* in frame with *MYCL1*, leading to the expression of a fusion protein composed of the first 79 amino acids of RLF and a MYCL1 protein lacking its first 27 amino acids, thereby generating a 446-residue fusion protein. The clinical sample that had the *RLF-MYCL1* fusion also overexpressed *MYCL1*. This fusion has previously been noted[27], but its role as an oncogene in SCLC has not been established. We found that small interfering RNA (siRNA)-mediated targeting of *MYCL1* in H1092 and CORL47 fusion-positive cells effectively reduced the proliferation of these cells, strongly suggesting a functional role for *MYCL1* in SCLC (**Supplementary Fig. 6**).

Multiple gene fusions involving kinase genes have recently been shown to be activating[28]. We identified four such fusions—*NPEPPS-EPHA6*, *SKP1-CDKL3*, *NEK4-SFMBT1* and *ZAK-RAPGEF4*—that are predicted by sequence to result in functional fusion proteins (**Fig. 5** and **Supplementary Figs. 7–9**). The roles of these fusion products in cancer remain to be elucidated.

In this study, we have identified multiple new recurrent somatic mutations in SCLC, including multiple mutations and copy-number alterations in SOX gene family members. The potential role of SOX family members in SCLC is further emphasized here by the identification of *SOX2* amplification and overexpression in approximately a quarter of the SCLC samples analyzed. SOX proteins have an important role in diverse biological processes, including cell type specification. Among the SOX family members, SOX2 in particular is a key factor in the maintenance of pluripotency and self-renewal of stem cells[23]. Aberrant *SOX2* expression has also been implicated in reprogramming mature cells to acquired pluripotency[24]. Its expression in mouse fibroblasts, together with *FoxG1*, has been shown to generate self-renewing neural precursor cells[25]. Conditional deletion of *Sox2* in mice indicates its critical role in lung development[22]. Conversely, overexpression of *SOX2* in lung epithelial cells has been shown to promote tumorigenesis[29].

Notably, conditional induction of *SOX2* in lung epithelial cells is also known to increase the number of neural progenitor cells[30]. SCLCs are tumors with neuroendocrine features. SOX2 protein overexpression has previously been noted in high-grade SCLC[31], and immunoreactive antibodies against SOX2 have been detected in sera from SCLC patients[32]. These observations, together with the frequent amplifications identified here, imply that *SOX2* has an important role as a putative lineage-survival oncogene in SCLC. This suggestion is further supported by the correlation of SOX2 expression with SCLC stage and the role of *SOX2* expression in maintaining SCLC proliferation.

The recurrent nature of the *RLF-MYCL1* fusion and its functional relevance provide additional opportunities for therapeutic intervention in SCLC. Recently, oncogenic kinase gene fusions have become a major focus of interest in the therapeutic targeting of NSCLC[33–35]. Understanding the role of tumor-specific in-frame kinase fusion transcripts identified in SCLC in this study may provide promising opportunities for targeted therapy development.

URLs. EGA, http://www.ebi.ac.uk/ega/.

© 2012 Nature America, Inc. All rights reserved.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Sequencing and genotype data have been deposited at the European Genome-phenome Archive, which is hosted by the European Bioinformatics Institute (EBI), under accession EGAS00001000334.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

The authors would like to thank Genentech DNA Sequencing and Oligo groups for their help with the project. We thank M.A. Huntley and J. Degenhardt for bioinformatics support and the Pathology Core Labs for providing histology, IHC and tissue management support. This work was also supported by grants from the Burroughs Wellcome Fund, the Flight Attendant Medical Research Institute, the Johns Hopkins Specialized Programs of Research Excellence (SPORE) NCI P50CA058184 (M.V.B. and C.M.R.), the Colorado SPORE NCI P50 CA058187 (M.V.-G) and the University of Texas SPORE NCI P50CA70907 (J.D.M., A.F.G. and K.E.H.). D.D.P. is supported by the Coordenacao de Aperfeicoamento de Passoal de Nivel Superior (CAPES) Foundation and the Ministry of Education of Brazil.

AUTHOR CONTRIBUTIONS

C.M.R. and S.S. conceived the study and designed the experiments. E.W.S. and S.D. performed the exome and whole-genome sequencing, RNA-seq and copy-number analysis. Z.M. and Y.G. performed validation of the fusions. Z.M. managed exome capture. J.T.P., E.A.B., S.C., V.J., B.S.J., W.Y. and C.P. performed biological validated studies. J. Shin, D.D.P., P.B.I. and M.V.-G. performed SOX2 IHC and FISH studies. K.E.H., A.F.G. and J.D.M. provided reagents and analysis support. J. Stinson, C.K.F., D.B., C.S.R. and J.G. collected sequencing data and performed mutation validation. F.G. and Z.Z. predicted the functional effects of mutations. E.W.S., P.M.H., R.B., T.D.W. and R.G. provided bioinformatics support, including the algorithm for variant calling, fusion detection and copy-number calling. R.B. and P.M.H. analyzed SNP array data. H.K., H.M.S., P.B.I., M.V.B. and A.F.G. provided pathology support. F.J.d.S., D.S.S., R.L.Y. and J.D.M. provided critical analysis and organizational support. D.S.S., E.W.S., S.D., Z.M., C.M.R. and J.T.P. wrote the manuscript, which was reviewed and edited by the other coauthors.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

Published online at http://www.nature.com/doifinder/10.1038/ng.2405.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Siegel, R., Ward, E., Brawley, O. & Jemal, A. Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *CA Cancer J. Clin.* **61**, 212–236 (2011).
2. Siegel, R., Naishadham, D. & Jemal, A. Cancer statistics, 2012. *CA Cancer J. Clin.* **62**, 10–29 (2012).
3. Hann, C.L. & Rudin, C.M. Fast, hungry and unstable: finding the Achilles' heel of small-cell lung cancer. *Trends Mol. Med.* **13**, 150–157 (2007).
4. Wistuba, I.I., Gazdar, A.F. & Minna, J.D. Molecular genetics of small cell lung carcinoma. *Semin. Oncol.* **28**, 3–13 (2001).
5. Mori, N. *et al.* Variable mutations of the RB gene in small-cell lung carcinoma. *Oncogene* **5**, 1713–1717 (1990).
6. Arriola, E. *et al.* Genetic changes in small cell lung carcinoma. *Clin. Transl. Oncol.* **10**, 189–197 (2008).
7. Yokomizo, A. *et al.* *PTEN/MMAC1* mutations identified in small cell, but not in non-small cell lung cancers. *Oncogene* **17**, 475–479 (1998).
8. Tatematsu, A. *et al.* Epidermal growth factor receptor mutations in small cell lung cancer. *Clin. Cancer Res.* **14**, 6092–6096 (2008).
9. Shibata, T., Kokubu, A., Tsuta, K. & Hirohashi, S. Oncogenic mutation of *PIK3CA* in small cell lung carcinoma: a potential therapeutic target pathway for chemotherapy-resistant lung cancer. *Cancer Lett.* **283**, 203–211 (2009).
10. Onuki, N. *et al.* Genetic changes in the spectrum of neuroendocrine lung tumors. *Cancer* **85**, 600–607 (1999).
11. Sher, T., Dy, G.K. & Adjei, A.A. Small cell lung cancer. *Mayo Clin. Proc.* **83**, 355–367 (2008).
12. Forbes, S.A. *et al.* COSMIC (the Catalogue of Somatic Mutations in Cancer): a resource to investigate acquired mutations in human cancer. *Nucleic Acids Res.* **38**, D652–D657 (2010).

13. Pleasance, E.D. *et al.* A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* **463**, 184–190 (2010).
14. Ng, P.C. & Henikoff, S. Accounting for human polymorphisms predicted to affect protein function. *Genome Res.* **12**, 436–446 (2002).
15. Ramensky, V., Bork, P. & Sunyaev, S. Human non-synonymous SNPs: server and survey. *Nucleic Acids Res.* **30**, 3894–3900 (2002).
16. González-Pérez, A. & Lopez-Bigas, N. Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. *Am. J. Hum. Genet.* **88**, 440–449 (2011).
17. Kan, Z. *et al.* Diverse somatic mutation patterns and pathway alterations in human cancers. *Nature* **466**, 869–873 (2010).
18. Seshagiri, S. *et al.* Recurrent R-spondin fusions in colon cancer. *Nature* published online, doi:10.1038/nature11282 (15 August 2012).
19. Molderings, G.J. *et al.* Multiple novel alterations in *Kit* tyrosine kinase in patients with gastrointestinally pronounced systemic mast cell activation disorder. *Scand. J. Gastroenterol.* **42**, 1045–1053 (2007).
20. D'Angelo, S.P. & Pietanza, M.C. The molecular pathogenesis of small cell lung cancer. *Cancer Biol. Ther.* **10**, 1–10 (2010).
21. Medina, P.P. *et al.* The SRY-HMG box gene, *SOX4*, is a target of gene amplification at chromosome 6p in lung cancer. *Hum. Mol. Genet.* **18**, 1343–1352 (2009).
22. Tompkins, D.H. *et al.* Sox2 is required for maintenance and differentiation of bronchiolar Clara, ciliated, and goblet cells. *PLoS ONE* **4**, e8248 (2009).
23. Wegner, M. SOX after SOX: SOXession regulates neurogenesis. *Genes Dev.* **25**, 2423–2428 (2011).
24. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
25. Lujan, E., Chanda, S., Ahlenius, H., Sudhof, T.C. & Wernig, M. Direct conversion of mouse fibroblasts to self-renewing, tripotent neural precursor cells. *Proc. Natl. Acad. Sci. USA* **109**, 2527–2532 (2012).
26. Bass, A.J. *et al.* *SOX2* is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. *Nat. Genet.* **41**, 1238–1242 (2009).
27. Mäkelä, T.P., Saksela, K., Evan, G. & Alitalo, K. A fusion protein formed by L-myc and a novel gene in SCLC. *EMBO J.* **10**, 1331–1335 (1991).
28. Robinson, D.R. *et al.* Functionally recurrent rearrangements of the MAST kinase and Notch gene families in breast cancer. *Nat. Med.* **17**, 1646–1651 (2011).
29. Lu, Y. *et al.* Evidence that SOX2 overexpression is oncogenic in the lung. *PLoS ONE* **5**, e11022 (2010).
30. Gontan, C. *et al.* Sox2 is important for two crucial processes in lung development: branching morphogenesis and epithelial cell differentiation. *Dev. Biol.* **317**, 296–309 (2008).
31. Sholl, L.M., Long, K.B. & Hornick, J.L. Sox2 expression in pulmonary non-small cell and neuroendocrine carcinomas. *Appl. Immunohistochem. Mol. Morphol.* **18**, 55–61 (2010).
32. Güre, A.O. *et al.* Serological identification of embryonic neural proteins as highly immunogenic tumor antigens in small cell lung cancer. *Proc. Natl. Acad. Sci. USA* **97**, 4198–4203 (2000).
33. Kwak, E.L. *et al.* Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N. Engl. J. Med.* **363**, 1693–1703 (2010).
34. Bergethon, K. *et al.* *ROS1* rearrangements define a unique molecular class of lung cancers. *J. Clin. Oncol.* **30**, 863–870 (2012).
35. Takeuchi, K. *et al.* *RET*, *ROS1* and *ALK* fusions in lung cancer. *Nat. Med.* **18**, 378–381 (2012).

© 2012 Nature America, Inc. All rights reserved.

ONLINE METHODS

Samples, DNA and RNA preparations. In this study, we have characterized 80 human SCLCs, including 36 primary SCLC human tumor and adjacent normal sample pairs and 17 paired SCLC cell lines and their patient-matched lymphoblastoid lines, as well as 4 primary SCLC tumors and 23 SCLC cell lines without matched normal controls (**Supplementary Table 1**).

Patient-matched fresh-frozen primary SCLC tumors and normal tissue samples were obtained from commercial sources or the Johns Hopkins tissue repository (**Supplementary Table 1**). All samples used in the study had appropriate IRB approval and informed consent from study participants. All tumor and normal tissues were subjected to review by a pathologist to confirm diagnosis and tumor content. The Qiagen AllPrep DNA/RNA kit was used to prepare DNA and RNA.

Exome capture and sequencing. We analyzed the exomes of 42 SCLC samples (30 primary tumor–normal tissue pairs and 12 paired cell lines) and their patient-matched normal samples to assess their mutational burden. We also obtained exome data for an additional 21 SCLC samples that included 5 primary SCLC tumors and 16 SCLC cell lines (**Supplementary Table 1**). Exome capture was performed using the Agilent SureSelect Human All Exome kit (38 Mb or 50 Mb). The SureSelect 50 Mb kit includes all of the capture probes from the 38 Mb kit plus some additional content derived from CCDS, GENCODE and RefSeq. Exome capture libraries were sequenced by HiSeq 2000 (Illumina) to generate 2 × 75-bp paired-end data (**Supplementary Table 1**). Targeted mean coverage of 80× and 162× with 96% and 92% of bases covered at ≥10× was achieved for 38 Mb and 50 Mb exome libraries, respectively (**Supplementary Table 2**).

RNA-seq. We obtained RNA-seq data for 55 samples (24 primary tumor–normal tissue pairs, 7 primary tumors, 2 adjacent normal samples and 22 SCLC cell lines) using the TruSeq RNA Sample Preparation kit (Illumina). Libraries were multiplexed two per lane and sequenced on HiSeq 2000 to obtain at least ~30 million paired-end (2 × 75-bp) reads per sample.

Sequence data processing. All sequencing reads were evaluated for quality using the Bioconductor ShortRead package[36]. Sample identity was confirmed by comparing data derived from exome sequencing and RNA-seq against Illumina 2.5 M array data as described[18].

Variant calling and validation. Sequencing reads were mapped to the UCSC human reference genome (GRCh37/hg19) using Burrows-Wheeler Aligner (BWA) software[37] set to default parameters. Local realignment, duplicate marking and raw variant calling were performed as described previously[38]. Known germline variations represented in dbSNP Build 131 (ref. 39) but not represented in COSMIC v56 (ref. 12) were filtered out. Variations present in the tumor sample but absent in matched normal tissue were predicted to be somatic. Predicted somatic variations were additionally filtered to include only positions with a minimum of 10× coverage in both the tumor and matched normal tissue, as well as an observed variant allele frequency of <3% in the matched normal tissue and a significant difference in variant allele counts, as determined using Fisher's exact test. To control for possible low-level tumor contamination in adjacent normal tissue, the allele frequency cutoff was expanded to 5% if a gene was significantly mutated, allowing for an additional 11 variants to be included. We performed whole-genome sequencing of the 1 hypermutated sample and only report the 755 protein-altering variants that were found in both the exome and whole-genome data for this sample. This sample was excluded from background mutation rate calculations. For unpaired samples, in addition to dbSNP, variants were filtered against normal variants from this data set, as well as normal variants from a published colon data set[18]. In addition, data from 2,500 normal exomes in the National Heart, Lung, and Blood Institute (NHLBI) Exome Sequencing Project was used to filter out variants and hotspot mutations. To evaluate the performance of the variant calling algorithm, we randomly selected 594 protein-altering variants and validated them using Sequenom, as described previously[17]. Of these variants, 91% (539) were validated as somatic. All variants that were invalidated were removed from the final set. Variants that were also validated by RNA-seq are labeled as VALIDATED: RNA-Seq to show confirmed expression of the variant (**Supplementary Table 3**). Indels were called using the GATK Indel Genotyper Version 2 (ref. 28). Indel validation was performed as described in a recent study[18]. The effects of all nonsynonymous somatic mutations on gene function were predicted using SIFT[14], PolyPhen[15] and Condel[16]. All variants were annotated using Ensembl (release 59).

Mutational significance. We evaluated the mutational significance of genes using a previously described method[17], with the addition of an expression filter, as mutation rates are known to vary with expression level[13,40]. The hypermutated sample was excluded from analysis so that it did not affect the background mutation rate. Because of the variability in background mutation, the uniform background mutation rate used to assess the significance of mutation in cancer-associated genes is at times lower than the actual mutation rate in some regions, resulting in false positive candidates, such as the olfactory genes, seeming to be significantly mutated cancer-associated genes. To address this, a recent study used an RNA-seq–based expression filter to focus on expressed genes, thereby potentially filtering out genes that are expressed at very low levels or are not expressed at all[41]. In this study, we classified average gene expression on the basis of RNA-seq data into tertiles (high, medium and low) and used this information to remove low expressors that would otherwise be identified as significantly mutated cancer-associated genes.

Whole-genome, RNA-seq and pathway analysis. Whole-genome analysis, RNA-seq–based expression assessment and pathway-level analysis were performed as described previously[18].

SNP array data generation and analysis. Illumina HumanOmni2.5_4v1 arrays were used to assay 56 samples (36 primary tumor–normal pairs, 15 SCLC cell line–normal pairs, 1 SCLC cell line and 4 unpaired primary tumors) for genotype, DNA copy number and loss of heterozygosity (LOH) at ~2.5 million SNP positions. These samples all passed our quality control metrics for sample identity and data quality. A subset of 2,295,239 high-quality SNPs was selected for all analyses.

After making modifications to permit use with Illumina array data, we applied the PICNIC[42] algorithm to estimate total copy number, allele-specific copy number and LOH, as described recently[18]. Recurrent genomic regions with DNA copy gain and loss were identified using GISTIC, version 2.0 (ref. 43).

Fusion detection and validation. Fusion identification and validation were performed as has been recently described[18].

Cell lines and culture conditions. All cell lines used in the study, except where noted, were cultured in RPMI 1640 supplemented with 10% FBS. H446 and H720 were cultured in RPMI 1640 with 10% tetracycline-free FBS (Hyclone, R10). Cell line identity for lines used to assess *SOX2* copy number was confirmed by short tandem repeat (STR) profiling using the StemElite ID System (Promega). HCC33, HCC2433, H289, H2141, H2107, H209, H1963, H1672, H1607, H1450, H1339, H1184, H2171, HCC1772, HCC970, H128 and H2195 SCLC cell lines, their patient-matched lymphoblastoid lines and their culture conditions have been described previously[44–46] (**Supplementary Table 1**). Additional SCLC cell lines were obtained from the American Type Culture Collection (ATCC).

Doxycycline-inducible shRNA-expressing cell lines and protein blotting. Scrambled or *SOX2*-targeting (TRC Clone TRCN0000003253) shRNAs were cloned as annealed oligonucleotides (Sigma) into Tet-pLKO-puro (Addgene plasmid 21915) digested with AgeI and EcoRI according to published protocols[47,48]. Sequence-verified clones were used to produce lentiviral particles according to TRC protocols. Lentiviral supernatants were used to infect cultured H446 or H720 cells in R10 medium at low multiplicity of infection in the presence of 8 μg/ml polybrene for 16 h. After incubation, medium was replaced with fresh R10, and cells were cultured for an additional 24 h before being selected and maintained in 500 ng/ml puromycin. The optimal doxycycline dose for inducible knockdown was determined to be 2 μg/ml, which was the minimum dose that resulted in maximal knockdown of *SOX2* after 96 h. The effect of

© 2012 Nature America, Inc. All rights reserved.

SOX2 knockdown on the amount of SOX2 protein was assessed by protein blot using antibody to SOX2 (Cell Signaling Technology 27485) or GAPDH (Santa Cruz Biotechnology, sc-25778) horseradish peroxidase (HRP)-conjugated secondary antibodies, followed by signal detection with chemiluminescence (GE Healthcare Life Sciences).

Cell viability and proliferation assays. Stable cell lines were plated in quadruplicate at a density of 1×10^3 cells per well in opaque 96-well plates in the presence or absence of 2 μg/ml doxycycline. Cells were plated in replicate plates for each time point tested. ATP content was measured as an indicator of metabolically active cells using the CellTiter-Glo Luminescent Cell Viability Assay (Promega) read on a SpectraMax M2e plate reader in luminescence mode (Molecular Devices). Viability was normalized between cell lines at 48 h to correct for differences in the initial number of cells plated in each group. All experiments were repeated a minimum of three times with similar results, and one representative experiment is shown.

Analysis of copy-number variation in SCLC cell lines. *SOX2* copy number was assessed by quantitative RT-PCR using TaqMan Copy Number Assays (Hs02719379_cn) on a StepOnePlus Real-Time PCR System (Applied Biosystems). *RPPH1* served as the reference gene (Applied Biosystems). Copy-number calls relative to normal human genomic DNA (Promega) were made with CopyCaller v2.0 (Applied Biosystems).

Tissue microarrays. SCLC tissue microarrays were obtained from US Biomax (LC703, LC802a, LC1009 and LC10010a) for IHC and FISH as fresh-cut slides. The four tissue microarrays contain replicate cores and a small set of overlapping cases. For analysis, missing or inconclusive cores were removed, and the replicate or overlapping case core with the highest percentage of tumor area was used for analysis, yielding 110 unique SCLC cases and 15 normal lung cases. Histological diagnosis with SCLC was confirmed by an attending pathologist.

Immunohistochemistry. IHC for SOX2 was performed on the tissue microarrays using a Leica Bond-III automated slide stainer (Leica Microsystems). The 4-μm sections were deparaffinized and subjected to antigen retrieval with Cell Conditioning Solution (high pH CC1 standard, Ventana Medical Systems) for 60 min. Sections were then incubated for 44 min with rabbit monoclonal antibody to SOX2 (1:100 dilution; clone SP76, Cellmarque). Reactions were developed through biotin-free, polymer detection (Ultra-view, Ventana Medical Systems) according to the manufacturer's instructions.

Scoring was performed on each sample. Nuclear labeling was scored by intensity (no (0), weak (1), moderate (2) or strong (3)) and for extent (expressed as the percentage of nuclei that were positive). Results were expressed by assigning a composite IHC score that was calculated by multiplying the intensity score by the percentage of nuclei with positive staining, with a maximum value of 300.

FISH analysis. FISH was performed on the tissue microarrays. The BAC clone RP11-459K6 containing a human DNA insert from the genomic region of *SOX2* (previously validated by PCR) was used for preparation of the *SOX2* FISH probe. The *SOX2* probe was validated for chromosome mapping and quality of hybridization in the human lymphoblastoid cell line AG09391 (Coriell Institute).

One slide of each tissue microarray was subjected to a two-color FISH assay using a mixture of the *SOX2* probe (red) and a commercially available probe for the chromosome 3 centromere (Kreatech) (green). The steps before hybridization were performed using the Zymed Spot-Light Tissue Pretreatment kit (Invitrogen) according to the manufacturer's instructions.

Analysis was performed on an epifluorescence microscope using single interference filter sets for blue (DAPI), green (FITC) and red (Texas red). For each interference filter, monochromatic images were acquired and merged using CytoVision (Leica Microsystems). Tumor cells were scored for copy-number signals of *SOX2* in 30–50 cells. In this analysis, a scoring system was proposed to identify increased levels of copy number per cell. Scores were assigned on a scale from 1–6 (according to pattern of copy-number gain, median per-cell change): 1 (no, 1–2), 2 (low, 2–3), 3 (moderate, 3–4), 4 (high, 4–5), 5 (very high, >5), 6 (gene amplification, gene clusters).

***MYCL1* knockdown studies.** The SCLC cell lines, NCI-H1092, CORL47 and NCI-H2171 were transfected with siRNA pools targeting *MYCL1* (Dharmacon) or with a non-targeting control siRNA (Dharmacon) following a reverse transfection protocol. The cells were incubated at 37 °C for 5 d after transfection and were subjected to a cell viability assay using the CellTiter-Glo kit (Promega).

MYCL1 (Hs00420495_m1) and *GAPDH* (Hs00266705_g1) TaqMan probes and primers were obtained from Life Technologies and were used to assess knockdown according to the manufacturer's instructions. Data were analyzed using the $\Delta\Delta C_T$ method by normalizing to *GAPDH* and mock-transfected controls. TaqMan reactions were performed in duplicate to obtain a mean value and s.d. *P* values were calculated by *t* test.

36. Morgan, M. *et al.* ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics* **25**, 2607–2608 (2009).
37. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
38. DePristo, M.A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
39. Sherry, S.T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).
40. Martincorena, I., Seshasayee, A.S. & Luscombe, N.M. Evidence of non-random mutation rates suggests an evolutionary risk management strategy. *Nature* **485**, 95–98 (2012).
41. Barbieri, C.E. *et al.* Exome sequencing identifies recurrent *SPOP*, *FOXA1* and *MED12* mutations in prostate cancer. *Nat. Genet.* **44**, 685–689 (2012).
42. Greenman, C.D. *et al.* PICNIC: an algorithm to predict absolute allelic copy number variation with microarray cancer data. *Biostatistics* **11**, 164–175 (2010).
43. Mermel, C.H. *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* **12**, R41 (2011).
44. Simms, E., Gazdar, A.F., Abrams, P.G. & Minna, J.D. Growth of human small cell (oat cell) carcinoma of the lung in serum-free growth factor–supplemented medium. *Cancer Res.* **40**, 4356–4363 (1980).
45. Phelps, R.M. *et al.* NCI–Navy Medical Oncology Branch cell line data base. *J. Cell. Biochem. Suppl.* **24**, 32–91 (1996).
46. Carney, D.N., Bepler, G. & Gazdar, A.F. The serum-free establishment and *in vitro* growth properties of classic and variant small cell lung cancer cell lines. *Recent Results Cancer Res.* **99**, 157–166 (1985).
47. Sarbassov, D.D., Guertin, D.A., Ali, S.M. & Sabatini, D.M. Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science* **307**, 1098–1101 (2005).
48. Wiederschain, D. *et al.* Single-vector inducible lentiviral RNAi system for oncology target validation. *Cell Cycle* **8**, 498–504 (2009).

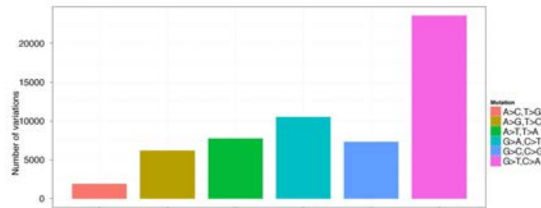
© 2012 Nature America, Inc. All rights reserved.

Comprehensive genomic analysis identifies *SOX2* as a frequently amplified gene in small cell lung cancer

Charles M. Rudin[1§*], Steffen Durinck[2,3§], Eric W. Stawiski[2,3§], John T. Poirier[1§], Zora Modrusan[2§], David S. Shames[5§], Emily A. Bergbower[1], Yinghui Guan[2], James Shin[1], Joseph Guillory[2], Celina Sanchez Rivers[2], Catherine K. Foo[2], Deepali Bhatt[2], Jeremy Stinson[2], Florian Gnad[3], Peter M. Haverty[3], Robert Gentleman[3], Subhra Chaudhuri[2], Vasantharajan Janakiraman[2], Bijay S. Jaiswal[2], Chaitali Parikh[2], Wenlin Yuan[2], Zemin Zhang[3], Hartmut Koeppen[4], Thomas D. Wu[3], Howard M Stern[4], Robert L. Yauch[5], Kenneth E. Huffman[6], Diego D. Paskulin[7], Peter B. Illei[1], Marileila Varella-Garcia[7], Adi F. Gazdar[6], Frederic J. de Sauvage[2], Richard Bourgon[3], John D. Minna[6], Malcolm V. Brock[1] and Somasekar Seshagiri[2*]

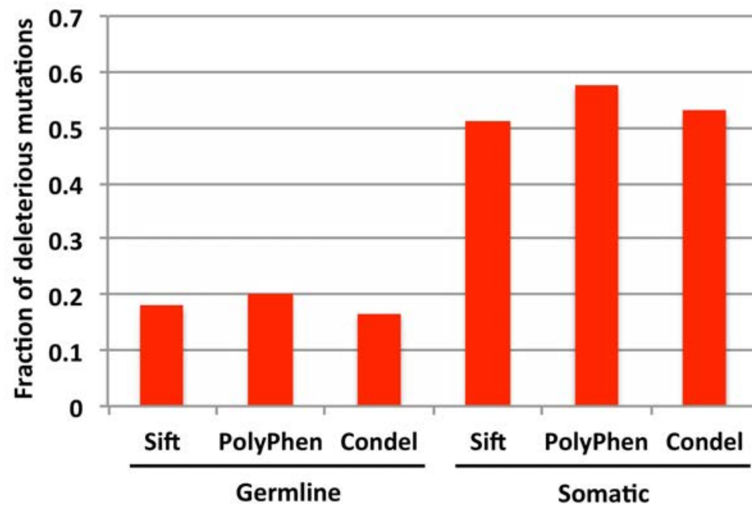
[1]The Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University, Baltimore, MD 21287; [2]Department of Molecular Biology, Genentech Inc., 1 DNA WAY, South San Francisco, CA 94080; [3]Department of Bioinformatics and Computational Biology, Genentech Inc., 1 DNA WAY, South San Francisco, CA 94080; [4]Department of Pathology Genentech Inc., 1 DNA WAY, South San Francisco, CA 94080; [5]Department of Oncology Biomarker Development. Genentech Inc., 1 DNA WAY, South San Francisco, CA 94080; [6] Hamon Center for Therapeutic Oncology Research, University of Texas Southwestern Medical Center, Dallas TX 75390; [7]University of Colorado Cancer Center, Aurora, CO 80045

Nature Genetics: doi:10.1038/ng.2405



Supplementary Figure 1. Average number of transitions and transversions in the whole genome sequence of the hypermutated SCLC sample.

Nature Genetics: doi:10.1038/ng.2405



Supplementary Figure 2. Plot of percentage of germline and somatic mutations scored as deleterious by Polyphen, SIFT and Condel.

Nature Genetics: doi:10.1038/ng.2405

TP53
RB1
SOX2
SOX3
SOX4
SOX5
SOX6
SOX9
SOX11
SOX14
SOX17

Mutation
WT
het
hom/multiple

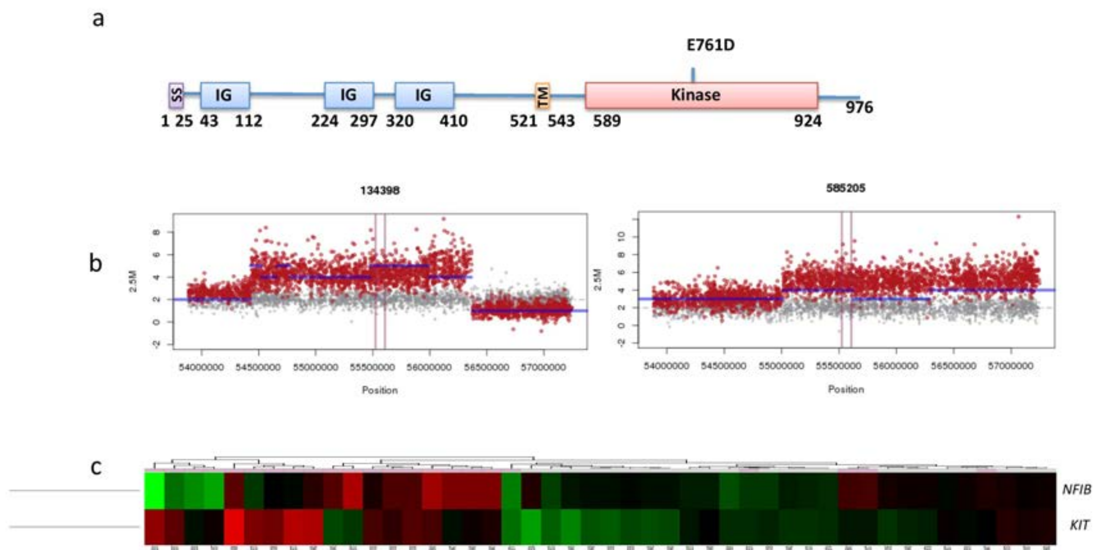
Expression
Increased
Decreased

Copy Number
Gain
Normal
Loss

* No RNA-seq data
^ No RNA-seq and 2.5M data

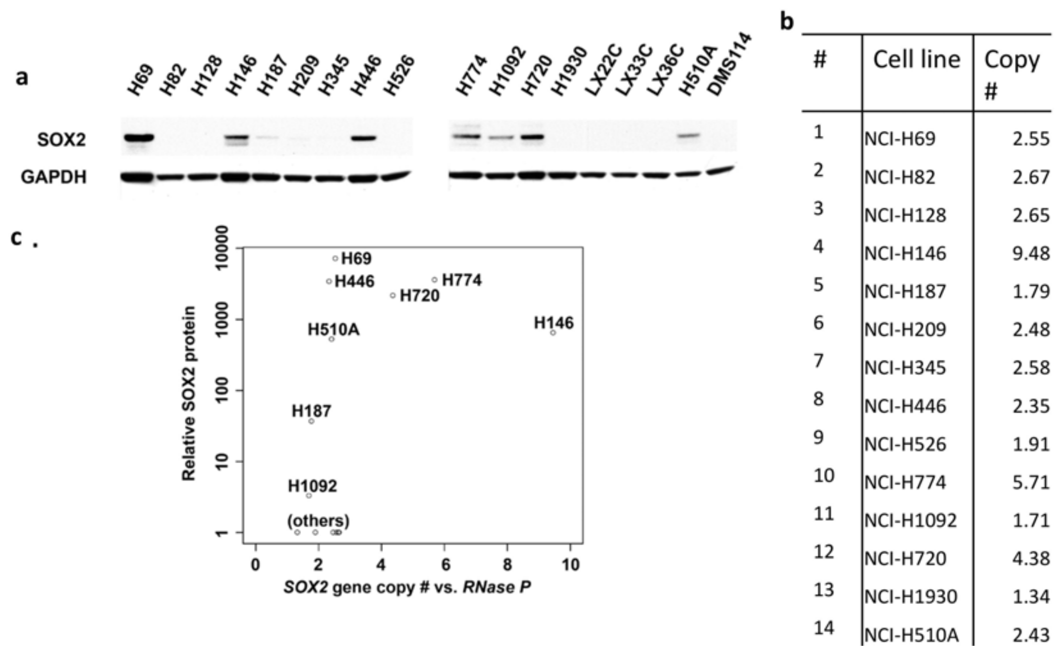
Supplementary Figure 3. Quilt plot depicting the co-occurrence of alteration in samples in *SOX*-family genes, *TP53* and *RB1*. Each column represent is data from a single tumor sample.

Nature Genetics: doi:10.1038/ng.2405



Supplementary Figure 4. (**a**) *KIT* mutation occurs at a codon previously know to be mutated in systemic mast cell disorder (Scandinavian Journal of Gastroenterology 2007; 42:1045). (**b**) Gene amplification of the *KIT* locus in two representative samples. (**c**) Heat map showing the overexpression of *KIT* in SCLC.

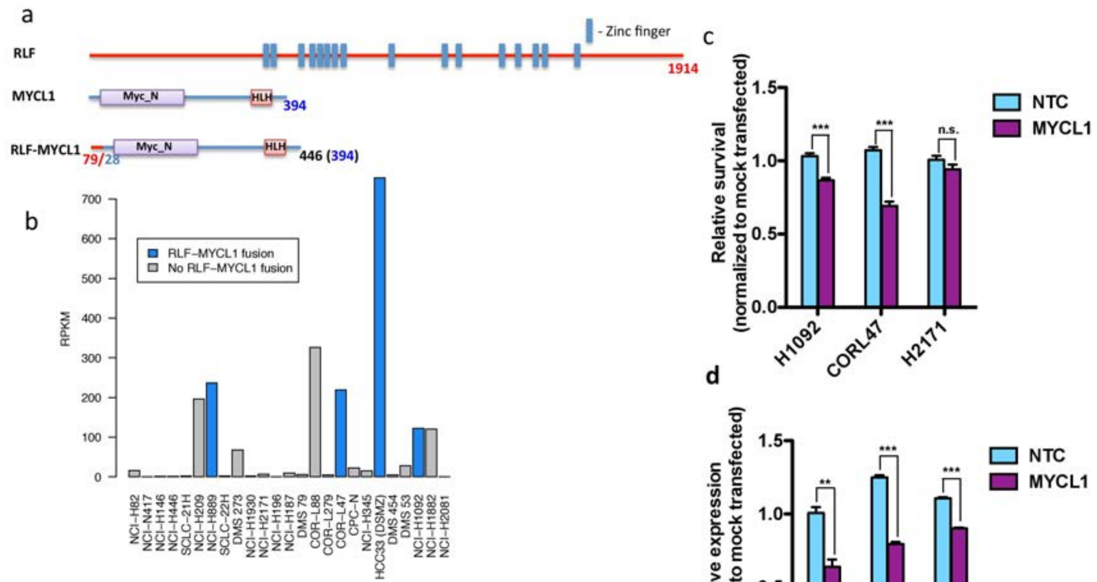
Nature Genetics: doi:10.1038/ng.2405



#	Cell line	Copy #
1	NCI-H69	2.55
2	NCI-H82	2.67
3	NCI-H128	2.65
4	NCI-H146	9.48
5	NCI-H187	1.79
6	NCI-H209	2.48
7	NCI-H345	2.58
8	NCI-H446	2.35
9	NCI-H526	1.91
10	NCI-H774	5.71
11	NCI-H1092	1.71
12	NCI-H720	4.38
13	NCI-H1930	1.34
14	NCI-H510A	2.43

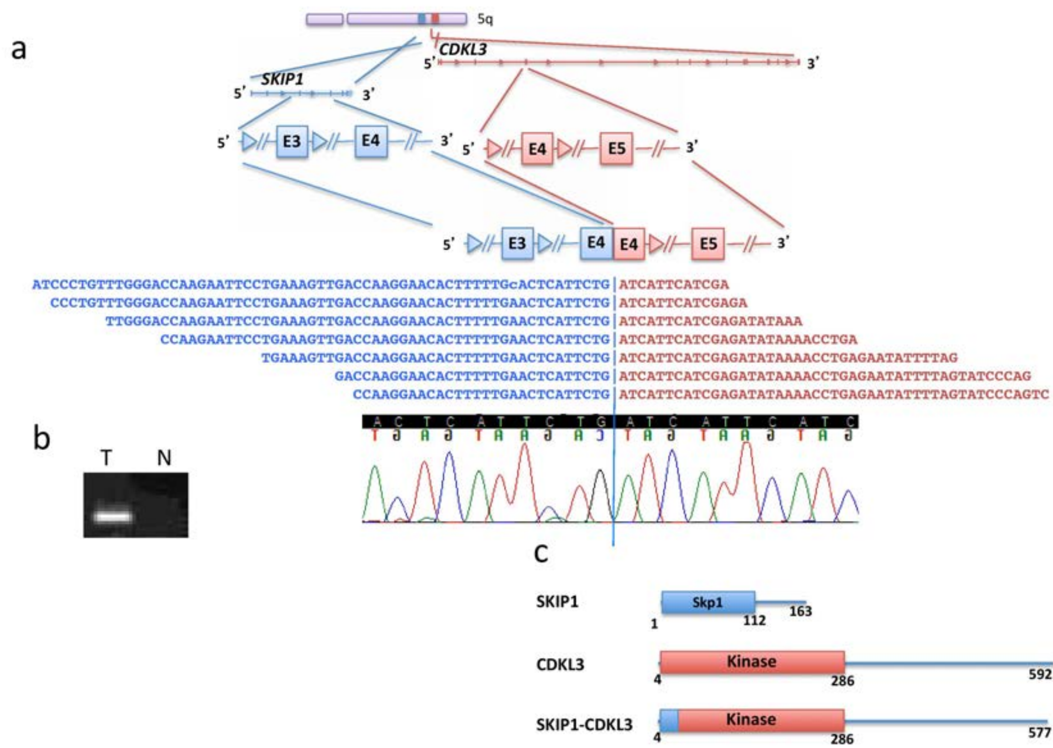
Supplementary Fig 5. (**a**) Expression of *SOX2* in cell lines as assessed by Western blot (**b**) *SOX2* copy number in cell lines assessed by TaqMan quantitative real-time PCR. Data presented is normalized against *RPPH1* copy #. (**c**) Plot of the copy number in panel **b** vs the relative expression levels of SOX2 protein in panel **a**.

Nature Genetics: doi:10.1038/ng.2405



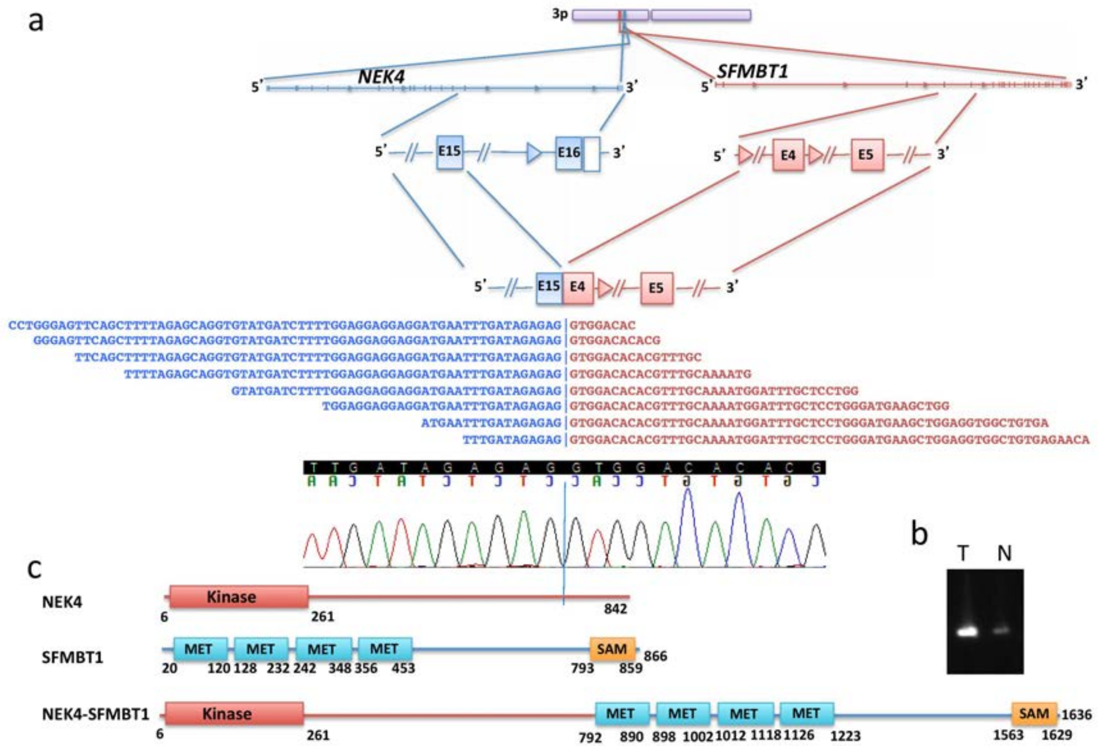
Supplementary Fig 6. siRNA mediated knockdown of *MYCL1* in *RLF-MYCL1* fusion positive cells reduces the viability of fusion positive cells. (**a**) The architecture of RLF-MYCL1 fusion protein. It is a 446 amino acid protein that contains the first 79 amino acids of RLF and 367 amino acids of MYCL1 starting from amino acid 28 to 394 (**b**) Expression of MYCL1 in a panel of SCLC cells measured by RNA-seq. Cells with *RLF1-MYCL1* fusion showed an elevated level of *MYCL1* expression. (**c**) Inhibition of *MYCL1* expression using siRNA in cells lines harboring *RLF-MYC1* fusion (H1092 and CORL47) reduces survival, compared to cells that do not express the fusion (H2171). ($p \leq 0.05$. **, $p \leq 0.01$. *** , n.s. not significant) (**d**) *MYCL1* mRNA levels measured by Taqman at 48h post-transfection in cells transfected with *MYCL1* or non targeting control (NTC) siRNA.

Nature Genetics: doi:10.1038/ng.2405



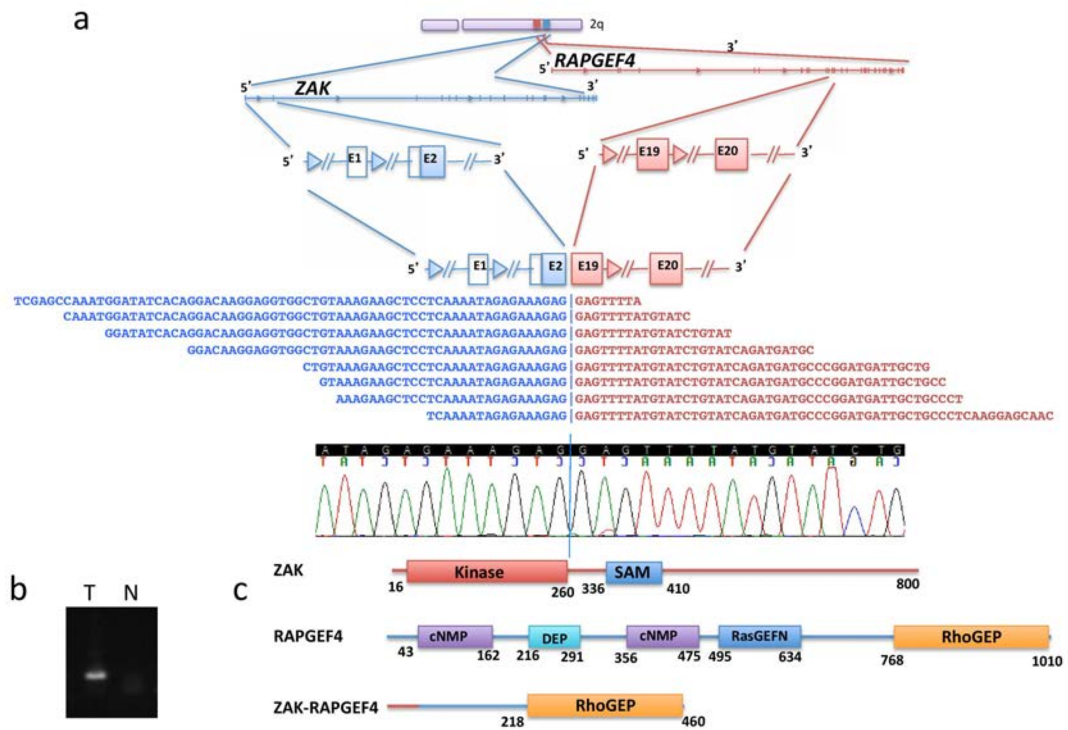
Supplementary Figure 7. (**a**) *SKIP1-CDKL3* fusion identified using RNA-seq along with the read evidence. (**b**) Independent RT-PCR derived product confirming the somatic fusion. A representative chromatogram obtained from this product is shown in **a**. (**c**) Schematic representation of the SKIP1-CDKL3 fusion protein.

Nature Genetics: doi:10.1038/ng.2405



Supplementary Figure 8. (**a**) *NEK4-SFMBT1* fusion identified using RNA-seq along with the read evidence. (**b**) Independent RT-PCR derived product confirming the somatic fusion. A representative chromatogram obtained from this product is shown in (**a**). (**c**) Schematic representation of the NEK4-SFMBT1 fusion protein. MET- Present in Drosophila Scm, l(3)mbt, and vertebrate SCML2.

Nature Genetics: doi:10.1038/ng.2405



Supplementary Figure 9. (**a**) *ZAK-RAPGEF4* fusion identified using RNA-seq along with the read evidence. (**b**) Independent RT-PCR derived product confirming the somatic fusion. A representative chromatogram obtained from this product is shown in (**a**). (**c**) Schematic representation of ZAK-RAPGEF4 fusion protein. cNMP - Cyclic nucleotide-monophosphate binding domain; DEP - Domain found in Dishevelled, Egl-10, and Pleckstrin; RasGEFN - Guanine

Nature Genetics: doi:10.1038/ng.2405

10.3 *P53 SIGNALING PATHWAY POLYMORPHISMS ASSOCIATED TO RECURRENT PREGNANCY LOSS*

Título do manuscrito: *p53 signaling pathway polymorphisms associated to recurrent pregnancy loss*

Autores: Fraga LR, Dutra CG, Boquett JA, Vianna FSL, Gonçalves RO, D.D. Paskulin DD; Costa OL, Ashton-Prolla P, Sanseverino MTV, Schuler-Faccini L.

Situação: Submetido

Revista: *Molecular Biology Reports*

Website: http://www.springer.com/life+sciences/animal+sciences/journal/11033

Molecular Biology Reports

p53 signaling pathway polymorphisms associated to recurrent pregnancy loss

--Manuscript Draft--

Manuscript Number:	MOLE-6873R1
Full Title:	p53 signaling pathway polymorphisms associated to recurrent pregnancy loss
Article Type:	Manuscript
Keywords:	MDM2; p53; Polymorphism; Recurrent pregnancy loss; reproduction
Corresponding Author:	Lucas Fraga, M.D. UFRGS Porto Alegre, BRAZIL
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	UFRGS
Corresponding Author's Secondary Institution:	
First Author:	Lucas Fraga, M.D.
First Author Secondary Information:	
Order of Authors:	Lucas Fraga, M.D.
	Caroline Gross Dutra, M.D.
	Juliano André Boquett
	Fernanda Sales Luiz Vianna
	Rozana Oliveira Gonçalves
	Diego D'Ávila Paskulin
	Olívia Lúcia Costa
	Patrícia Ashton-Prolla
	Maria Teresa Vieira Sanseverino
	Lavínia Schuler-Faccini
Order of Authors Secondary Information:	

Powered by Editorial Manager® and Preprint Manager® from Aries Systems Corporation

To Molecular Biology Reports

May 4, 2013

ANSWERS TO THE REVIEWERS

We wish to thank the careful revision and comments made by the reviewers. Here are our answers to the questions raised.

Reviewer #1: The authors studied the association between polymorphisms in p53 signaling pathway molecules and recurrent pregnancy loss. They concluded that the combination of TP53 Arg/Arg (rs1042522) and MDM2 TT (rs2279744) genotypes may be a risk factor for RPL. It is valuable to estimate the genetic susceptibility to RPL. However, there are some concerns for the present form of the manuscript.

Question 1. The sample size is small.

We agree that the sample is small but it is a highly depurated sample, since we had carefully investigated previously all couples and excluded all with known causes of RPL, such as morphological abnormalities, hormonal or immune disorders as well as chromosomal alterations. That led us to a sample of 120 patients without any known risk for RPL [detailed in the first paragraph of Materials and Methods section (subjects)].

2. Hardy-Weinberg equilibrium was tested using the Pearson chi-square test. The p-value should be shown in table 2.

The p-values from the Hardy-Weinberg equilibrium of both case and control groups were added to Table 2.

3. The interaction of the TP53 Arg/Arg (rs1042522) and MDM2 TT (rs2279744) genotypes was shown to increase the risk to RPL (OR=2.58; 95% CI: 1.31-5.07; p=0.006). Did this effect remained statistically significant in Non-European ancestry subgroup, European ancestry subgroup?

We changed the word "non-European" to "African ancestry", to be clearer. We reformulated the paragraph in the main text as follows

"When the subgroup of women reporting European ethnicity was evaluated independently, the effect of the interaction remained significant (OR = 2.77, 95% CI: 1.31 to 5.85, p = 0.008). This was not the case for the subgroup of African ancestry (OR = 2.05, 95% CI: 0.38 to 11.01, p = 0.401). The interaction remains conferring risk to RPL, nevertheless the CI is high."

4. There was no difference between the two groups for allelic and genotypic distributions of all SNPs, when analyzed separately (Table 2). These were tested using the Pearson chi-square test. Was the logistic regression analysis used to evaluate ORs (odds ratios) with 95% CIs (confidence intervals) for additive genetic model, dominant model, and recessive model respectively? Was subgroup analysis for ethnicity performed?

The analyses performed to evaluate these differences are now explained on Materials and Methods section, as well as in the Results section as follows

a. Materials and Methods

"To test dominant and recessive genetic models we used the Multivariate binary logistic regression analysis, which was performed evaluating these models separately and controlled by clinical risk factors (smoking, alcohol consumption, and number of pregnancies, as well as ethnicity – using European ancestry as the reference category). Multivariate binary logistic regression analysis also was used to evaluate gene-gene interactions between the TP53 Arg/Arg genotype (rs1042522) and other risk genotypes (TT for MDM2 and TT for LIF). To achieve this, manual replacement of these different interaction terms was performed along with the clinical risk factors aforementioned."

a. Results

"There was no difference between the two groups for allelic and genotypic distributions of all SNPs, when analyzed separately, even when controlled by the clinical risk factors or in additive and dominant models (data not shown). All genotypic and allelic frequencies were within the expected by the Hardy-Weinberg distribution (Table 2)."

*Manuscript
Click here to download Manuscript: Fraga et al p53 130504.docx
Click here to view linked References

p53 signaling pathway polymorphisms associated to recurrent pregnancy loss

L.R. Fraga[1]; C.G. Dutra[1]; J.A. Boquett[1]; F.S.L. Vianna[1,2]; R.O. Gonçalves[3]; D.D. Paskulin[1]; O.L. Costa[3]; P. Ashton-Prolla[1,2,4]; M.T.V. Sanseverino[2,4]; L. Schuler-Faccini[1,2,4]

[1]Post-Graduation Program in Genetics and Molecular Biology, Departament of Genetics, Biosciences Institute, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, 91501-970, Brazil.

[2]INAGEMP – Instituto Nacional de Genética Médica Populacional, Universidade Federal do Rio Grande do Sul, Porto Alegre, 91.501-970, Brazil.

[3]Post-Graduation Program in Biotechnology in Health and Investigative Medicine and Department of Obstetrics, Gynecology and Human Reproduction, Medical School, Universidade Federal da Bahia (UFBA), Salvador, 40.110-100, Brazil.

[4]Medical Genetics Service, Hospital de Clínicas de Porto Alegre, Porto Alegre, 90.035-903, Brazil.

*Corresponding author:

Lavinia Schuler-Faccini, MD, PhD

Universidade Federal do Rio Grande do Sul - Departamento de Genética

Caixa Postal 15031 – Agencia Campus UFRGS

CEP 91501-970 / Porto Alegre – RS – Brasil

Phone: (51) 33598008 Fax: (51) 33598011

E-mail: lavinia.faccini@ufrgs.br

Abstract

The p53 protein is known for performing essential functions in the maintenance of genomic stability in somatic cells and prevention of tumor formation. Studies of the p53 signaling pathway have suggested associations between some polymorphisms and infertility, post-in vitro fertilization implantation failure and recurrent abortions. The TP53 Pro72Arg polymorphism has been implicated as a risk factor for recurrent pregnancy loss (RPL); however, the association is controversial. In this study, our objective was to evaluate selected polymorphisms in genes of the p53 signalling pathway [TP53 c.215G>C (Pro72Arg), MDM2 c.14+309T>G (SNP309) and LIF c.1414T>G in the region 3' UTR] and determine their effect as risk factors for RPL. In a case-control study, we investigated 120 women with two or more pregnancy losses and 143 fertile control women reporting at least two live births and no history of pregnancy loss. When analyzed separately, the allele and genotype distributions of the polymorphisms in the two groups were not different. However, in a multivariate analysis adjusted for alcohol consumption, smoking, ethnicity, and number of pregnancies, the interaction between the genotypes TP53 Arg/Arg (rs1042522) and MDM2 TT (rs2279744) showed to be associated to RPL, increasing the risk for this condition (OR = 2.58, 95% CI: 1.31 - 5.07, p = 0.006). In conclusion, our study indicates that the combination of TP53 Arg/Arg (rs1042522) and MDM2 TT (rs2279744) genotypes may be a risk factor for RPL.

Key words: MDM2; p53; polymorphism; recurrent pregnancy loss; reproduction

Introduction

Recurrent pregnancy loss (RPL) is defined as the occurrence of two or more consecutive pregnancy losses before 24 weeks of gestation [1]. This condition affects about 5% of couples in the reproductive period of their lives [2] and can be classified as: (a) primary RPL, in which all pregnancies are lost or (b) secondary RPL, when at least one successful pregnancy to term with a liveborn child is reported, regardless of the number of miscarriages [3, 4].

The etiology of RPL has been widely studied. Among the most common causes are genetic, morphological, hormonal, metabolic, infectious, environmental, and immunologic alterations, as well as thrombophilias and advanced maternal age [5]. A specific cause for RPL is clearly identified in only about 50% of cases [4].

The p53 protein, known for performing essential functions in the maintenance of genomic stability in somatic cells and prevention of tumor formation [6], plays an essential role also in human reproduction [7]. p53 is involved in the protection of female germinative cells and embryos against teratogenic agents, and induces the expression of the leukemia inhibitor factor (Lif) protein, which is an important mediator of embryo implantation [7-9].

The major negative regulator of p53 levels is the protein E3 ubiquitin ligase Mdm2 (Mdm2) [10]. Mdm2 acts in a principal regulatory point, binding to p53 and causing its degradation through polyubiquitination and consequent degradation and attenuation of its activity [11, 12].

Some polymorphisms in these genes have been implicated with adverse pregnancy outcomes like infertility [13], recurrent implantation failure [14], twinning [15], and missed abortions [16]. The Pro72Arg polymorphism of the p53 protein has been considered a risk factor for RPL. However, there is no consistent replication of the observed results [14, 17-19]. The Pro72Arg polymorphism of the TP53 gene results in the alteration c.215G> C in exon 4 of the gene and modifies the apoptosis induction [20, 21]. SNP309 (c.14+309T>G), located in the intronic promoter region of the MDM2 gene affects the level of gene expression and consequently the levels of Mdm2 and p53 [11]. The polymorphism c.1414T>G, located in the 3' UTR region of the LIF gene (rs929271), reduces the stability of the mRNA [13, 22, 23].

Thus, the objective of this study was to evaluate the effect of the aforementioned polymorphisms and their effects on the risk of RPL in a Brazilian population sample.

Material and Methods

Subjects

This is a case-control study in which we assessed women diagnosed with primary RPL at the Prenatal Diagnosis Clinic of the Medical Genetics Service (DPN-SGM) of Hospital de Clinicas in Porto Alegre (HCPA), in Southern Brazil, between 2000 and 2011. All women reporting at least two pregnancy losses before 24 weeks of gestation with the same partner, and with no report of a full-term pregnancy were invited for the study, of which 130 were recruited. We performed a structured interview to obtain information about demographics, gynecological/obstetric history, medical history, current or past use of tobacco (yes-no), alcohol consumption (yes-no), consumption of other drugs or medications, family history of malformations, age at recruitment, weight, height, and occupation. All women were subjected to a preliminary standard diagnostic protocol including hysteroscopy, laparoscopy, ultrasound, and comprehensive determination of hormonal status (gonadotrophins, FSH, LH, prolactin, thyroid hormones, thyroperoxidase) in order to detect known causes for the pregnancy losses, before inclusion in the study. Karyotypic examination of peripheral lymphocytes was conducted to detect chromosomal abnormalities in all couples, and immunological risk factors were investigated through assessment of anticardiolipin, lupus anticoagulant and antinuclear antibodies. Presence of any maternal clinical condition that could prevent full-term pregnancies (e.g. uterine abnormality) was considered an exclusion criterion for this study.

The control group consisted of 143 healthy women who reported at least two live births and no history of pregnancy loss or infertility.

This research study was approved by the Institutional Review Board (Grupo de Pesquisa e Pós-Graduação, Hospital de Clínicas de Porto Alegre) under protocol number #11-242. Patients were recruited after signature of informed consent, in accordance with the Declaration of Helsinki.

Genotyping

Genomic DNA was obtained from saliva using the Oragene® DNA collection kit (DNA Genotek Inc., Canada) in accordance to the manufacturer's instructions. The extracted DNA was quantified with PicoDrop 100 (Picodrop Limited, United Kingdom) equipment. Genotyping TP53 Pro72Arg rs1042522,

MDM2 SNP309 rs2279744 and LIF rs929271 was performed by allelic discrimination using the TaqMan Genotyping Assay (Applied Biosystems, USA); assay numbers: C___2403545_10 (TP53), and C___7545904_10 (LIF). MDM2 SNP309 was assessed with a customized (assay-by-design) assay using probes FAM-TCCCGCGCCGCAG and VIC-CTCCCGCGCCGAAG, forward primer 5'-CGGGAGTTCAGGGTAAAGGT-3'; and reverse primer 5'-ACAGGCACCTGCGATCATC-3'. PCR reactions were conducted in 96-well plates, with each reaction containing: 10ng of genomic DNA, 2 x TaqMan® Genotyping MasterMix (Applied Biosystems, USA), probes specific for each SNP (40x), and water in sufficient quantity for a final reaction volume of 8μl. The PCR conditions were as follows: 10 min at 95°C, followed by 45 cycles of 75 s each (95°C for 15 s and 63°C for 60 s). MDM2 SNP309 genotyping was also conducted in 96-well plates, with each reaction containing: 10ng of genomic DNA, 2 x TaqMan® Genotyping MasterMix (Applied Biosystems, USA), 1μM of each primer and probe, and water to reach a final volume of 25μM. MDM2 PCR conditions were as follows: 2 min at 50°C, 10 min at 95°C, followed by 45 cycles of 75 s (95°C for 15 s and 60°C for 60 s). All the reactions were performed in a StepOnePlus™ Real-Time PCR System (Applied Biosystems, USA) and the results were analyzed using StepOne v.2.2.2 Software (Applied Biosystems, USA).

Statistical Analyses

Sample characterization and comparisons between groups (cases and controls) were done using the Mann-Whitney U-test was used for continuous variables and the Pearson chi-square test for categorical variables. Cases and controls were compared for their genotypic and allelic frequencies, and Hardy-Weinberg equilibrium was tested using the Pearson chi-square test. To test dominant and recessive genetic models we used the Multivariate binary logistic regression analysis, which was performed evaluating these models separately and controlled by clinical risk factors (smoking, alcohol consumption, and number of pregnancies, as well as ethnicity – using European ancestry as the reference category). Multivariate binary logistic regression analysis also was used to evaluate gene-gene interactions between the TP53 Arg/Arg genotype (rs1042522) and other risk genotypes (TT for MDM2 and TT for LIF). To achieve this, manual replacement of these different interaction terms was performed along with the clinical risk factors aforementioned.

Whenever p-value was <0.20, the variable was included as a covariate in a Multivariate Logistic Regression analysis. The odds ratios (OR) with respective confidence intervals (CI) of 95% were calculated to assess the relative risk conferred by each covariate. Sample size for this study was calculated using the EpiInfo software, version 6.0, with the parameters being an alpha value of 0.05 and 80% power. All other analyses cited above were performed using the Statistical Package for Social Sciences software, version 20.0 (SPSS Inc., Chicago, USA). Alpha values less than 0.05 were considered significant. For each logistic regression model, the Bonferroni correction was applied for number of covariates into the model. Considering all analyses, the Bonferroni correction was applied to five tests (two for the multivariate binary logistic regression and three for SNPs evaluated individually), being $\alpha_{Bonf} = 0.01$.

Results

From the original, consecutive sample of 130 women with RPL, 120 were included after initial cytological, hormonal, morphological, immunological and karyotype exams identified no abnormalities. In ten (7.7%) women karyotypic alterations were identified.

In the RPL group, 41 women (34.2%), 41 (34.2%) and 38 (31.7%) reported two, three and four losses, respectively. The number of reported pregnancies was higher in the RPL group than in the control group (Table 1). Average age at recruitment was also higher in the RPL group: 45.2 (± 9.1) versus 33.2 (± 7.5) (p <0.001). However, the average age at first pregnancy did not differ between groups. Women with RPL had higher alcohol consumption, and also a greater proportion of African ancestry when compared to the control group.

There was no difference between the two groups for allelic and genotypic distributions of all SNPs, when analyzed separately, even when controlled by the clinical risk factors or in additive and dominant models (data not shown). All genotypic and allelic frequencies were within the expected by the Hardy-Weinberg distribution (Table 2).

The most appropriate model of multiple associations obtained in this analysis is presented in Table 3. The interaction of the TP53 Arg/Arg (rs1042522) and MDM2 TT (rs2279744) genotypes was shown to increase the risk to RPL (OR=2.58; 95% CI: 1.31-5.07; p=0.006). When the Bonferroni correction was applied for multiple comparisons this effect remained statistically significant. Binary logistic regression

analysis allowed estimating the probability of pregnancy loss for a carrier of these risk genotypes. The estimated probability was 71% for carriers and 48% for non-carriers, considering the overall average number of pregnancies of 3.16. When the subgroup of women reporting European ethnicity was evaluated independently, the effect of the interaction remained significant (OR = 2.77, 95% CI: 1.31 to 5.85, p = 0.008). This was not the case for the subgroup of African ancestry (OR = 2.05, 95% CI: 0.38 to 11.01, p = 0.401). The interaction remains conferring risk to RPL, nevertheless the CI is high.

Discussion

In this study, we focused on selected polymorphisms in the p53 signaling pathway and their interactions as possible risk factors for RPL. No associations were identified between RPL and the three SNPs studied, when individually assessed. However, multivariate logistic regression analysis showed that the interaction between genotypes TP53 Arg/Arg (rs1042522) and MDM2 TT (rs2279744) increased the risk of RPL 2.58 times, when corrected for smoking, alcohol consumption, ethnicity, and number of pregnancies.

Several studies have sought associations between SNPs and RPL. Genes such as tumor necrosis factor alpha (TNF-a) [24-26], vascular endothelial growth factor (VEGF) [27] and nitric oxide syntase 3 (NOS3), which encodes the eNOS protein, [28] are considered to be relevant to the origin of RPL.

Regarding the TP53 gene, to date only four studies have evaluated the Pro72Arg polymorphism in women with RPL (Table 4) [14, 17-19]. In these studies only genotypic data were considered, and neither environmental variables nor interactions between different SNPs were assessed. The allelic and genotypic frequencies of the TP53 Pro72Arg SNP in our study were similar to those described by the International HapMap Project database, and to those observed in previously in the Southeastern Brazilian population [29, 30].

TP53 Pro72Arg has been associated with an increased risk for developing several tumors and with adverse pregnancy outcomes [7]. Functional studies indicate that the Arg allele is more effective in inducing apoptosis and suppression of cellular transformation [21] while the Arg allele seems to be less effective in inducing senescence and interruption of the cell cycle [31]. Arg is also more effective in inducing the expression of LIF, an essential step to ensure appropriate uterine conditions for blastocyst impantation [32,

33]. Thus, the efficiency of p53 is directly reflected in implantation and fertility and a few studies have linked the Pro allele with infertility resulting from implantation failure [34, 35].

MDM2 SNP309 is located in the promoter region (intron 1) of the gene and affects the level of gene expression. The G allele increases binding affinity of the transcription factor of MDM2, resulting in a higher transcriptional level when compared to the T allele and increased degradation of p53 [11]. The interaction between the TP53 Arg/Arg (rs1042522) and MDM2 TT (rs2279744) genotypes observed here, which is associated with an increased risk for RPL, likely results from . reduced transcriptional levels of Mdm2 (and consequent degradation of p53) coupled to the increased effectiveness of the TP53 Arg/Arg genotype in inducing apoptosis. Thus, these genotypes seem to synergistically increase the levels and functionality of p53 in relation to apoptosis, an effect that has been shown to increase the rate of abortion in mice [36].

p53-mediated apoptosis is involved in a number of mechanisms related to human reproduction, such as ovarian cell death for its homeostasis [37] and formation and development of the placenta [38, 39]. p53 is expressed in the trophoblast at all stages of gestation, but with an increase in the first quarter - a critical period for embryo selection [40, 41]. In addition, p53 regulates genes in the trophoblast cell invasion process, which involves degradation and remodeling of the extracellular matrix of the uterus [42, 43]. However, despite being an integral part of many reproductive functions, apoptosis may be pathological and damaging during reproduction in certain scenarios [44]. Increased apoptosis in the maternal-fetal interface has been associated with several reproductive disorders including preeclampsia and intrauterine growth restriction. Furthermore, studies in mice clearly suggest an interaction between increased p53 expression, excessive apoptosis and pregnancy loss [45-47]. Further investigations have shown that the p53 protein is highly expressed in human placentas of abnormal pregnancies [48-50].

Based on our results, we hypothesize that with a TP53 Arg/Arg background there is an enhanced induction of apoptosis, contributing to post-implantation selection, which is directly reflected in intra-uterine survival. This mechanism may be functioning as a post-zygotic selection step, since the majority of RPLs seem to be due to a high rate of embryo aneuploidies [51, 52]. Additional support for this hypothesis is provided by the study of Norimura et al. [36] which shows that p53+/+ pregnant mice exposed do X-rays exhibited a higher rate of apoptosis, reduced survival of abnormal fetuses, and a higher rate of death and

abortion of normal fetuses. In addition, a study in horses (Thoroughbred mares) identified a higher frequency of the TP53 Arg/Arg genotype in animals with pregnancy loss [53].

With respect to established environmental risk factors for RPL, only alcohol consumption was associated to RPL. The adverse effects of alcohol consumption, even in small doses, have been previously described and include fertility and development aspects [1, 54, 55]. Smoking was not associated with RPL, and although data in the literature are quite controversial, several studies implicate smoking as a dose-dependent risk factor for pregnancy loss, an information that we were not able to retrieve [1, 55, 56]. Although we were unable to identify published data on an association of ethnic origin with RPL, one of the SNPs implicated as a risk factor here (the TP53 Arg allele) has suffered strong selective pressure in recent history and its frequency is distributed quite differently among European and African populations [34]. For this reason, this particular genotype was included in the multivariate regression model.

In conclusion, this is the first study implicating the combined inheritance of TP53 Arg/Arg (rs1042522) and MDM2 TT (rs2279744) genotypes as a risk factor for RPL. We hypothesize that the mechanism by which this interaction is acting to contribute to RPL involves post-implantation p-53 mediated embryo selection through apoptosis induction. Although these finding do not currently have a significant clinical impact, they underscore that the investigation of genetic variants of smaller effect can still contribute to our understanding of the mechanisms underlying common complex conditions such as RPL.

References

1. Rai R, Regan L (2006) Recurrent miscarriage. Lancet 368:601-11.

2. Branch DW, Gibson M, Silver RM (2010) Clinical practice. Recurrent miscarriage. N Engl J Med 363:1740-7.

3. Pandey MK, Rani R, Agrawal S (2005) An update in recurrent spontaneous abortion. Arch Gynecol Obstet 272:95-108.

4. Toth B, Jeschke U, Rogenhofer N, Scholz C, Wurfel W, Thaler CJ, Makrigiannakis A (2010) Recurrent miscarriage: current concepts in diagnosis and treatment. J Reprod Immunol 85:25-32.

5. Laurino MY, Bennett RL, Saraiya DS, Baumeister L, Doyle DL, Leppig K, Pettersen B, Resta R, Shields L, Uhrich S et al. (2005) Genetic evaluation and counseling of couples with recurrent miscarriage: recommendations of the National Society of Genetic Counselors. J Genet Couns 14:165-81.

6. Levine AJ, Hu W, Feng Z (2006) The P53 pathway: what questions remain to be explored? Cell Death Differ 13:1027-36.

7. Hu W (2009) The role of p53 gene family in reproduction. Cold Spring Harb Perspect Biol 1:a001073.

8. Brady CA, Attardi LD (2010) p53 at a glance. J Cell Sci 123(Pt 15):2527-32.

9. Hu W, Feng Z, Teresky AK, Levine AJ (2007) p53 regulates maternal reproduction through LIF. Nature 450:721-4.

10. Brooks CL, Gu W (2006) p53 ubiquitination: MDM2 and beyond. Mol Cell 21(3):307-15.

11. Bond GL, Hu W, Bond EE, Robins H, Lutzker SG, Arva NC, Bargonetti J, Bartel F, Taubert H, Wuerl P et al. (2004) A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. Cell 119(5):591-602.

12. Brooks CL, Li M, Hu M, Shi Y, Gu W (2007) The p53--MDM2--HAUSP complex is involved in p53 stabilization by HAUSP. Oncogene 26(51):7262-6.

13. Kang HJ, Feng Z, Sun Y, Atwal G, Murphy ME, Rebbeck TR, Rosenwaks Z, Levine AJ, Hu W (2009) Single-nucleotide polymorphisms in the p53 pathway regulate fertility in humans. Proc Natl Acad Sci USA 106:9761-6.

14. Firouzabadi RD, Ghasemi N, Rozbahani MA, Tabibnejad N (2009) Association of p53 polymorphism with ICSI/IVF failure and recurrent pregnancy loss. Aust N Z J Obstet Gynaecol 49:216-9.

15. Tagliani-Ribeiro A, Paskulin DD, Oliveira M, Zagonel-Oliveira M, Longo D, Ramallo V, Ashton-Prolla P, Saraiva-Pereira ML, Fagundes NJ, Schuler-Faccini L et al. (2012). High twinning rate in Candido Godoi: a new role for p53 in human fertility. Hum Reprod 27:2866-71.

16. Fang Y, Kong B, Yang Q, Ma D, Qu X (2011) The p53-HDM2 gene-gene polymorphism interaction is associated with the development of missed abortion. Hum Reprod 26:1252-8.

17. Coulam CB, Kay C, Jeyendran RS (2006) Role of p53 codon 72 polymorphism in recurrent pregnancy loss. Reprod Biomed Online 12:378-82.

18. Kaare M, Butzow R, Ulander VM, Kaaja R, Aittomaki K, Painter JN (2009) Study of p53 gene mutations and placental expression in recurrent miscarriage cases. Reprod Biomed Online 18:430-5.

19. Pietrowski D, Bettendorf H, Riener EK, Keck C, Hefler LA, Huber JC, Tempfer C (2005) Recurrent pregnancy failure is associated with a polymorphism in the p53 tumour suppressor gene. Hum Reprod 20:848-51.

20. Buchman VL, Chumakov PM, Ninkina NN, Samarina OP, Georgiev GP (1988) A variation in the structure of the protein-coding region of the human p53 gene. Gene 70:245-52.

21. Dumont P, Leu JI, Della Pietra AC, 3rd, George DL, Murphy M (2003) The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. Nat Genet 33(3):357-65.

22. Giess R, Tanasescu I, Steck T, Sendtner M (1999) Leukaemia inhibitory factor gene mutations in infertile women. Mol Hum Reprod 5:581-6.

23. Ishida R, Ezura Y, Iwasaki H, Nakazawa I, Kajita M, Kodaira M, Ito H, Emi M (2001) Linkage disequilibrium and haplotype analysis among four novel single-nucleotide polymorphisms in the human leukemia inhibitory factor (LIF) gene. J Hum Genet 46:557-9.

24. Daher S, Shulzhenko N, Morgun A, Mattar R, Rampim GF, Camano L, DeLima MG (2003) Associations between cytokine gene polymorphisms and recurrent pregnancy loss. J Reprod Immunol 58:69-77.

25. Finan RR, Al-Irhayim Z, Mustafa FE, Al-Zaman I, Mohammed FA, Al-Khateeb GM, Madan S, Issa AA, Almawi WY (2010) Tumor necrosis factor-alpha polymorphisms in women with idiopathic recurrent miscarriage. J Reprod Immunol 84:186-92.

26. Prigoshin N, Tambutti M, Larriba J, Gogorza S, Testa R (2004) Cytokine gene polymorphisms in recurrent pregnancy loss of unknown cause. Am J Reprod Immunol 52:36-41.

27. Su MT, Lin SH, Lee IW, Chen YC, Kuo PL (2011) Association of polymorphisms/haplotypes of the genes encoding vascular. Hum Reprod 26:758-64.

28. Tempfer C, Unfried G, Zeillinger R, Hefler L, Nagele F, Huber JC (2001) Endothelial nitric oxide synthase gene polymorphism in women with idiopathic recurrent miscarriage. Hum Reprod 16:1644-7.

29. Marcel V, Palmero EI, Falagan-Lotsch P, Martel-Planche G, Ashton-Prolla P, Olivier M, Brentani RR, Hainaut P, Achatz MI (2009) TP53 PIN3 and MDM2 SNP309 polymorphisms as genetic modifiers in the Li-Fraumeni. J Med Genet 46:766-72.

30. Thurow HS, Haack R, Hartwig FP, Oliveira IO, Dellagostin OA, Gigante DP, Horta BL, Collares T, Seixas FK (2011) TP53 gene polymorphism: importance to cancer, ethnicity and birth weight in a Brazilian cohort. J Biosci 36:823-31.

31. Salvioli S, Bonafe M, Barbi C, Storci G, Trapassi C, Tocco F, Gravina S, Rossi M, Tiberi L, Mondello C et al. (2005) p53 codon 72 alleles influence the response to anticancer drugs in cells from aged people by regulating the cell cycle inhibitor p21WAF1. Cell Cycle 4:1264-71.

32. Cullinan EB, Abbondanzo SJ, Anderson PS, Pollard JW, Lessey BA, Stewart CL (1996) Leukemia inhibitory factor (LIF) and LIF receptor expression in human endometrium suggests a potential autocrine/paracrine function in regulating embryo implantation. Proc Natl Acad Sci USA 93:3115-20.

33. Stewart CL, Kaspar P, Brunet LJ, Bhatt H, Gadi I, Kontgen F, Abbondanzo SJ (1992) Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. Nature 359:76-9.

34. Feng Z, Zhang C, Kang HJ, Sun Y, Wang H, Naqvi A, Frank AK, Rosenwaks Z, Murphy ME, Levine AJ et al. (2011) Regulation of female reproduction by p53 and its family members. Faseb j 25:2245-55.

35. Kay C, Jeyendran RS, Coulam CB (2006) p53 tumour suppressor gene polymorphism is associated with recurrent implantation. Reprod Biomed Online 13:492-6.

36. Norimura T, Nomoto S, Katsuki M, Gondo Y, Kondo S (1996) p53-dependent apoptosis suppresses radiation-induced teratogenesis. Nat Med 2:577-80.

37. Amsterdam A, Keren-Tal I, Aharoni D, Dantes A, Land-Bracha A, Rimon E, Sasson R, Hirsh L (2003) Steroidogenesis and apoptosis in the mammalian ovary. Steroids 68:861-7.

38. Halperin R, Peller S, Sandbank J, Bukovsky I, Schneider D (2000) Expression of the p53 gene and apoptosis in gestational trophoblastic disease. Placenta 21:58-62.

39. Watson ED, Cross JC (2005) Development of structures and transport functions in the mouse placenta. Physiology (Bethesda) 20:180-93.

40. Cohen M, Meisser A, Haenggeli L, Irminger-Finger I, Bischof P (2007) Status of p53 in first-trimester cytotrophoblastic cells. Mol Hum Reprod 13:111-6.

41. Marzusch K, Ruck P, Horny HP, Dietl J, Kaiserling E (1995) Expression of the p53 tumour suppressor gene in human placenta: an immunohistochemical study. Placenta 16:101-4.

42. Cohen M, Wuillemin C, Irion O, Bischof P (2008) Regulation of MMP-9 by p53 in first trimester cytotrophoblastic cells. Hum Reprod 23:2273-81.

43. Wei P, Jin X, Zhang XS, Hu ZY, Han CS, Liu YX (2005) Expression of Bcl-2 and p53 at the fetal-maternal interface of rhesus monkey. Reprod Biol Endocrinol 3:4.

44. Halperin R, Peller S, Rotschild M, Bukovsky I, Schneider D (2000) Placental apoptosis in normal and abnormal pregnancies. Gynecol Obstet Invest 50:84-7.

45. Minas V, Jeschke U, Kalantaridou SN, Richter DU, Reimer T, Mylonas I, Friese K, Makrigiannakis A (2007) Abortion is associated with increased expression of FasL in decidual leukocytes. Mol Hum Reprod 13:663-73.

46. Veljkovic Vujaklija D, Sucic S, Gulic T, Dominovic M, Rukavina D (2012) Cell death mechanisms at the maternal-fetal interface: insights into the role of granulysin. Clin Dev Immunol. 2012:180272.

47. Savion S, Lepsky E, Orenstein H, Carp H, Shepshelovich J, Torchinsky A, Fein A, Toder V (2002) Apoptosis in the uterus of mice with pregnancy loss. Am J Reprod Immunol 47:118-27.

48. Hu C, Smith SD, Pang L, Sadovsky Y, Nelson DM (2006) Enhanced basal apoptosis in cultured term human cytotrophoblasts is associated with a higher expression and physical interaction of p53 and Bak. Placenta 27:978-83.

49. Levy R, Smith SD, Yusuf K, Huettner PC, Kraus FT, Sadovsky Y, Nelson DM (2002) Trophoblast apoptosis from pregnancies complicated by fetal growth restriction is. Am J Obstet Gynecol 186:1056-61.

50. Yamauchi H, Katayama K, Ueno M, He XJ, Mikami T, Uetsuka K, Doi K, Nakayama H (2007) Essential role of p53 in trophoblastic apoptosis induced in the developing rodent. Apoptosis 12:1743-54.

51. Hodes-Wertz B, Grifo J, Ghadir S, Kaplan B, Laskin CA, Glassner M, Munne S (2012) Idiopathic recurrent miscarriage is caused mostly by aneuploid embryos. Fertil Steril 98:675-80.

52. Quenby S, Vince G, Farquharson R, Aplin J (2002) Recurrent miscarriage: a defect in nature's quality control? Hum Reprod 17:1959-63.

53. Leon PM, Campos VF, Thurow HS, Hartwig FP, Selau LP, Dellagostin OA, Neto JB, Deschamps JC, Seixas FK, Collares T (2012) Association between single nucleotide polymorphisms in p53 and abortion in Thoroughbred mares. Vet J 193:573-5.

54. Kesmodel U, Wisborg K, Olsen SF, Henriksen TB, Secher NJ (2002) Moderate alcohol intake in pregnancy and the risk of spontaneous abortion. Alcohol Alcohol 37:87-92.

55. Kumar S (2011) Occupational, environmental and lifestyle factors associated with spontaneous. Reprod Sci 18:915-30.

56. Lindbohm ML, Sallmen M, Taskinen H (2002) Effects of exposure to environmental tobacco smoke on reproductive health. Scand J Work Environ Health 28 Suppl 2:84-96.

Acknowledgements

The authors would like to thank Dr. Sidia Maria Callegari Jacques for her assistance in the statistical analysis of the data.

Funding

This work was supported by Instituto Nacional de Genética Médica Populacional (INAGEMP), Fundo de Apoio a Pesquisa e Eventos do Hospital de Clínicas de Porto Alegre (FIPE-HCPA), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Table 1
Click here to download Table: Fraga et al. 2013 - Table 1.doc

Table 1: Demographic and Clinical Data of Women reporting RPL and Controls

Characteristic	**RPL group N=120**	**Control group N=143**	**P**
Number of pregnancies [mean (SD)]	3.4 (1.7)	2.9 (1.2)	<0.001 MW
Age at first pregnancy [mean (SD)]	24.4 (7.7)	22.0 (4.7)	0.123 MW
Smoking [N (%)]	17 (14.2)	33 (23.1)	0.067 X2
Alcohol consumption [N (%)]	54 (45.0)	25 (17.6)	<0.001 X2
Consanguinity with partner[N (%)]	3 (2.5)	1 (0.7)	0.228 X2
African ancestry [N (%)]	52 (44.3)	22 (15.4)	<0.001 X2

N: number; SD: Standard Deviation; X2: Pearson chi-square test; MW: Mann-Whitney U-test;

Table 1
Click here to download Table: Fraga et al. 2013 - Table 1.doc

Table 1: Demographic and Clinical Data of Women reporting RPL and Controls

Characteristic	**RPL group N=120**	**Control group N=143**	**P**
Number of pregnancies [mean (SD)]	3.4 (1.7)	2.9 (1.2)	<0.001 ^{MW}
Age at first pregnancy [mean (SD)]	24.4 (7.7)	22.0 (4.7)	0.123 ^{MW}
Smoking [N (%)]	17 (14.2)	33 (23.1)	0.067 ^{X2}
Alcohol consumption [N (%)]	54 (45.0)	25 (17.6)	<0.001 ^{X2}
Consanguinity with partner[N (%)]	3 (2.5)	1 (0.7)	0.228 ^{X2}
African ancestry [N (%)]	52 (44.3)	22 (15.4)	<0.001 ^{X2}

N: number; SD: Standard Deviation; X2: Pearson chi-square test; MW: Mann-Whitney U-test;

Table 2
Click here to download Table: Fraga et al 2013 - Table 2.doc

Table 2: Allelic and genotype frequencies of the polymorphisms from the p53 signaling pathway in women with RPL and controls

Gene	Genotype/Allele	RPL group n (%)	Control group n (%)	P[a]	RPL HWE P[a,b]	Controls HWE P[a,c]
TP53 (rs1042522)	Arg/Arg	57 (47.5)	72 (50.3)			
	Arg/Pro	47 (39.2)	60 (42)			
	Pro/pro	16 (13.3)	11 (7.7)	0.324		
	Arg	161 (67.1)	204 (71.3)			
	Pro	79 (32.9)	82 (28.7)	0.338	0.215	0.757
MDM2 (rs2279744)	TT	60 (50.0)	59 (41.3)			
	TG	45 (37.5)	59 (41.3)			
	GG	15 (12.5)	25 (17.4)	0.301		
	T	165 (68.7)	177 (61.9)			
	G	75 (31.3)	109 (38.1)	0.120	0.163	0.133
LIF (rs929271)	TT	64 (53.3)	66 (46.2)			
	TG	45 (37.5)	63 (44.1)			
	GG	11 (9.2)	14 (9.8)	0.499		
	T	173 (72.1)	195 (68.2)			
	G	67 (27.9)	91 (31.8)	0.380	0.455	0.854

[a]P-values obtained by chi-square test. [b]P-value from Hardy-Weinberg Equilibrium of the RPL group. [c]P-value from Hardy-Weinberg Equilibrium of the Control group.

Table 3
Click here to download Table: Fraga et al. 2013 - Table 3.doc

Table 3: Multivariate Logistic Regression analysis for the effect of the interaction of the TP53 Arg/Arg (rs1042522) and MDM2 TT (rs2279744) genotypes in the risk of recurrent pregnancy losses, corrected for alcohol consumption, smoking, ethnicity, and number of pregnancies.

Variables	Regression coefficient	Standard Error	Wald	OR (CI 95%)	P [a, b, c]
TP53Arg/Arg + MDM2TT	0.948	0.345	7.543	2.58 (1.31 to 5.07)	0.006
Alcohol Consumption	1.343	0.324	20.542	4.33 (2.29 to 8.17)	<0.001
Smoking	-0.598	0.385	5.900	0.39 (0.18 to 0.83)	0.015
Ethnicity	-1.436	0.326	19.936	0.23 (0.12 to 0.44)	<0.001
Number of Pregnancies	0.233	0.103	9.080	1.36 (1.11 to 1.66)	0.003
Constant	-1.508	0.514	8.593	0.221	0.003

[a]P and OR values obtained from the binary logistic regression model, in which RPL patients were coded as 1 and the control subjects as 0 (reference category). [b]Hosmer and Lemeshow test: $p = 0.720$; Omnibus test: $P<0.0001$. [c]Bonferroni-corrected P-value 0.01.

Table 4
Click here to download Table: Fraga et al. 2013 - Table 4.doc

Table 4: Case-control studies with the Pro72Arg polymorphism in the TP53 gene and RPL

Study	Country	Allele/Genotype	RPL	Control group	P
			N (%)	n (%)	
Pietrowski et al., 2005	Austria	Arg	236 (67.4)	216 (75.5)	0.07
		Arg/Arg	83 (47.4)	83 (58.0)	0.03
Coulam et al., 2006	USA	Arg	337 (82.2)	34 (81.0)	0.84
		Arg/Arg	141 (68.8)	13 (61.9)	0.38
Firouzabadi et al., 2009	Iran	Arg	86 (44.3)	29 (45.4)	0.99
		Arg/Arg	23 (23.7)	4 (12.5)	0.06
Kaare et al., 2009	Finland	Arg	64 (69.6)	289 (75.7)	0.29
		Arg/Arg	21 (45.7)	106 (55.5)	0.45
Present Study	Brazil	Arg	161 (67.1)	204 (71.3)	0.32
		Arg/Arg	57 (47.5)	72 (50.3)	0.33

11. ANEXOS

11.1 ANEXO A – DESCRIÇÃO DO HUMAN OMNI 2.5-8 BEAD CHIP

Data Sheet: DNA Analysis

HumanOmni2.5-8 BeadChip

Next-Generation GWAS Content for Genotyping and CNV Analysis

Figure 1: HumanOmni2.5-8 BeadChip

The eight-sample HumanOmni2.5 BeadChip supports rapid cost-effective studies with coverage of the latest common and rare variants from the 1000 Genomes Project, down to 2.5% minor allele frequency (MAF).

Overview

The HumanOmni2.5-8 BeadChip delivers comprehensive coverage of both common and rare SNP content from the 1000 Genomes Project (1kGP; MAF > 2.5%), designed to be maximally informative for diverse world populations. Using the proven HiScan™ or iScan System, this eight-sample BeadChip offers high throughput and optimized tag SNP content, including full support of copy number variation (CNV) applications. The Omni2.5-8 BeadChip is a powerful entry point into the Omni Roadmap, which provides researchers with step-wise, flexible access to five million variants per sample. A convenient kit packaging, streamlined PCR-free protocol, and integrated analysis software are included to provide a comprehensive DNA analysis solution.

HumanOmni2.5-8 BeadChip Kit	Catalog No.
16 samples, single-use kit	WG-311-2511
48 samples, single-use kit	WG-311-2512
96 samples, single-use kit	WG-311-2513
384 samples, single-use kit	WG-311-2514

HumanOmni2.5-8 Product Information

Feature	Description
Number of Markers	2,379,855
Number of Samples	8
DNA Requirement	200 ng
Assay	Infinium® LCG
Instrument Support	HiScanSQ or iScan
Sample Throughput*	~1,067 samples / week
Scan Time / Sample	6.5 min (HiScan) 11.4 min (iScan)

% Variation Captured ($r^2 > 0.8$)	1kGP† MAF > 5%	1kGP† MAF > 1%
CEU	0.83	0.73
CHB + JPT	0.83	0.73
YRI	0.65	0.51

Data Performance	Value** / Product Specification‡
Call Rate	99.65% / > 99% avg
Reproducibility	99.99% / > 99.9%
Log R Dev	0.12 / < 0.30‡

Spacing	Mean / Median / 90th%†
Spacing (kb)	1.19 / 0.64 / 2.76

Marker Categories	Number of Markers†
Number of SNPs with 10kb of RefSeq genes	1,231,382
MHC / ADME	19,238 / 27,335
Non synonymous SNPs	41,900
Sex Chromosome (X / Y / PAR Loci)	55,208 / 2,561 / 418
Mitochondrial	256

* Estimate assumes one HiScan system, one AutoLoader2, one Tecan Robot, and a five-day work week.

† Compared against June 2011 1kGP data release

** Values are derived from genotyping 470 reference samples

‡ Value expected for typical projects, excluding tumor samples or any samples prepared not following standard Illumina protocols.

Illumina, Inc. • 9885 Towne Centre Drive, San Diego, CA 92121 USA • 1.800.809.4566 toll-free • 1.858.202.4566 tel • techsupport@illumina.com • illumina.com

FOR RESEARCH USE ONLY

© 2011 Illumina, Inc. All rights reserved.
Illumina, illuminaDx, BeadArray, BeadXpress, cBot, CSPro, DASL, Eco, Genetic Energy, GAIIx, Genome Analyzer, GenomeStudio, GoldenGate, HiScan, HiSeq, Infinium, iSelect, MiSeq, Nextera, Sentrix, Solexa, TruSeq, VeraCode, the pumpkin orange color, and the Genetic Energy streaming bases design are trademarks or registered trademarks of Illumina, Inc. All other brands and names contained herein are the property of their respective owners. Pub. No. 370-2011-005 Current as of 07 October 2011

Figura 22 – HumanOmni2.5-8 BeadChip.

11.2 ANEXO B – REGIÃO DE PERDA DE HETEROZIGOSIDADE EM 11p15.5-15.4 DO PACIENTE ACC-3

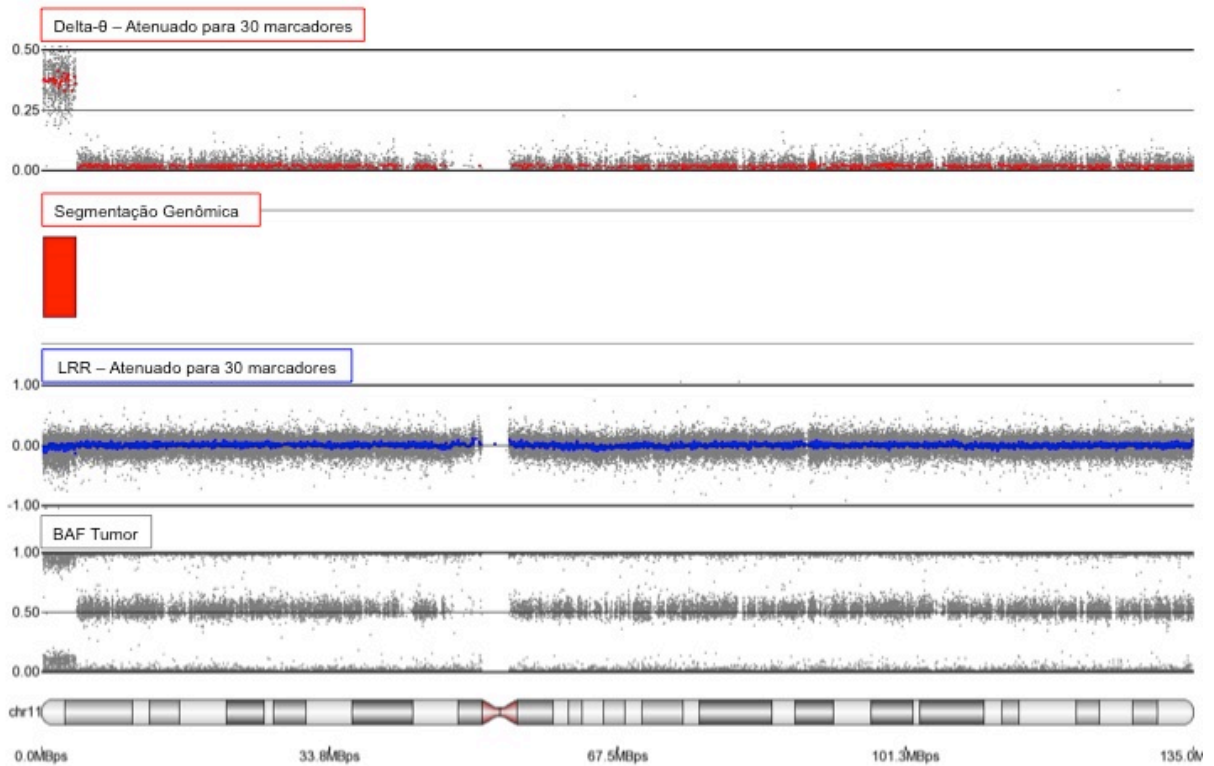


Figura 23 - Resultado do SNP Array para o cromossomo 11 do paciente ACC-3 demonstrando a perda de heterozigosidade sem alteração do número de cópias na região 11p15.5 - 11p15.4. Delta-θ (vermelho); Segmentação Genômica; LRR (azul); BAF Tumor; BAF Sangue.