

**UNIVERSIDADE FEDERAL DE SANTA MARIA  
CENTRO DE CIÊNCIAS NATURAIS E EXATAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA**

**DINÂMICA DA TRANSIÇÃO PRÉ-CÂNCER PARA  
CÂNCER: ESTUDO DA EXPRESSÃO DE VIAS DE  
MANUTENÇÃO DO GENOMA**

**TESE DE DOUTORADO**

**ÉDER MAIQUEL SIMÃO**

**Santa Maria, RS, Brasil**

**2012**

# **DINÂMICA DA TRANSIÇÃO PRÉ-CÂNCER PARA CÂNCER: ESTUDO DA EXPRESSÃO DE VIAS DE MANUTENÇÃO DO GENOMA**

por

**Éder Maiquel Simão**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Física, área de concentração em Física, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para a obtenção do grau de  
**Doutor em Física**

**Orientador: José Carlos Merino Mombach**

**Santa Maria, RS, Brasil**

**2012**

**Universidade Federal de Santa Maria  
Centro de Ciências Naturais e Exatas  
Programa de Pós-graduação em Física**

A Comissão Examinadora, abaixo assinada  
aprova a Tese de Doutorado

**DINÂMICA DA TRANSIÇÃO PRÉ-CÂNCER PARA CÂNCER: ESTUDO  
DA EXPRESSÃO DE VIAS DE MANUTENÇÃO DO GENOMA**

elaborado por  
**Éder Maiquel Simão**

como requisito parcial para obtenção do grau de  
**Doutor em Física**

**COMISSÃO EXAMINADORA:**

**José Carlos Merino Mombach, Dr.**  
(Presidente/Orientador)

**Fábio Passetti, Dr. (INCA)**

**Fábio Mallmann Zimmer, Dr. (UFSM)**

**Eleonir João Calegari, Dr. (UFSM)**

**Juliana Kaizer Vizzotto, Dr. (UFSM)**

**Santa Maria, 09 de julho de 2012**

**aos meus pais, irmã e namorada**

## AGRADECIMENTOS

Gostaria de agradecer a todas as pessoas que de uma forma ou de outra contribuíram para que este trabalho fosse realizado, durante toda a minha jornada acadêmica em que desenvolvi os trabalhos referentes ao meu doutorado:

Uma menção ao professor José Carlos Merino Mombach, pela paciência, amizade e confiança não havendo palavras para agradecê-lo.

A minha mãe Maria Gledes, ao meu pai Aldery e a minha irmã Lariane que sempre me deram forças com o suporte necessário para continuar a minha caminhada.

A minha namorada Daiane Marconato, pelo amor, carinho, força e apoio sempre me motivando e prestando sua solidariedade e atenção necessárias em compreensão ao meu empenho durante essa jornada.

A Cristhian Augusto Bugs, meu compadre, amigo e grande profissional. Os cafezinhos e trocas de ideias valeram-nos mais um capítulo em nossas vidas.

Aos Doutores, Marialva, Mauro, Giovani, Christian Probst e Ronnie pelo empenho e colaboração.

Ao nosso grupo de pesquisas em Sistemas Complexos: Karlise, Bruna, Cecília, Giovana e Daisiane e aos colegas, Carlos Alberto, Giuliano Demarco, Diego, Rafael, Cintia e Isabela pelo companheirismo.

A secretária Saionara Dutra e ao Paulino pelo apoio, dedicação e responsabilidade nas suas atividades.

A CAPES e ao CNPQ pelo apoio financeiro.

Aos meus amigos de São Pedro do Sul, companheiros de trilhas e aos meus familiares pela confiança e apoio depositado.

Por fim agradecer a Deus por me ajudar a enfrentar com coragem os desafios até aqui conquistados.

## **RESUMO**

**Tese de doutorado**

**Programa de Pós-Graduação em Física**

**Universidade Federal de Santa Maria**

### **DINÂMICA DA TRANSIÇÃO PRÉ-CÂNCER PARA CÂNCER: ESTUDO DA EXPRESSÃO DE VIAS DE MANUTENÇÃO DO GENOMA**

**Autor: Éder Maiquel Simão**

**Orientador: Prof. Dr. José Carlos Merino Mombach**

**Local e data da defesa: Santa Maria, 09 de Julho de 2012**

A perda da estabilidade genômica associada com a deterioração genética é um dos muitos aspectos importantes na evolução dos carcinomas. São os mecanismos de manutenção do genoma (GMM) que garantem a integridade do DNA e a sobrevivência da célula. Eles são compostos por vias genéticas que incluem os pontos de controle do ciclo celular, o reparo e a recombinação do DNA, a morte celular programada (apoptose) e a senescência. O objetivo deste trabalho é investigar a atividade dessas vias e genes de manutenção do genoma em doenças relacionadas à evolução do câncer. Para isso foram usados dados públicos de microarranjos obtidos do banco de dados *Gene Expression Omnibus* (GEO). Analisamos a expressão das proteínas e vias de manutenção do genoma que estão alteradas em tecidos pré-cancerosos, cancerosos e tecidos relacionados com a instabilidade genômica e que poderão levar a uma progressão tumoral. Os resultados permitiram caracterizar quantitativamente uma barreira anticâncer na evolução tumoral proposta na literatura.

**Palavras-chave:** Vias de manutenção do genoma, câncer, microarranjos.

## **ABSTRACT**

**Tese de doutorado**  
**Programa de Pós-Graduação em Física**  
**Universidade Federal de Santa Maria**

### **DYNAMICS OF PRE-CANCER TO CANCER TRANSITION: A STUDY OF EXPRESSION PATHWAYS OF GENOME MAINTENANCE**

**Autor: Éder Maiquel Simão**

**Orientador: Prof. Dr. José Carlos Merino Mombach**

**Local e data da defesa: Santa Maria, 09 de Julho de 2012**

The loss of genomic stability associated with genetic deterioration is one of the many important aspects for the development of carcinomas. The genome maintenance mechanisms (GMM) ensure the integrity of DNA and cell survival. They are composed of genetic pathways that include the cell cycle checkpoints, DNA repair and recombination, programmed cell death (apoptosis) and senescence. The purpose of the work is to investigate the activity of these pathways in genome maintenance in diseases related to the evolution of cancer. For this public data from microarrays obtained from the Gene Expression Omnibus (GEO) database were used. We analyzed the expression of proteins and genome maintenance pathways that are altered in pre-cancerous and cancerous tissues related to genomic instability and may lead to tumor progression. The results allowed a quantitative characterization of an anti-cancer barrier in the tumor evolution proposed in the literature.

**Keywords:** Pathways in genome maintenance, cancer, microarrays

## LISTA DE FIGURAS

- Figura 1 - Ciclo da divisão celular, o ciclo inicia na fase  $G_1$  ou  $G_0$  com a síntese de moléculas, passa pelas fases S, onde há duplicação do DNA segue pela fase  $G_2$  e é concluído na fase M, quando ocorre a segregação cromossômica. Imagem obtida de WIKIBOOKS, 2011. ....18
- Figura 2 - O RNA polimerase percorre toda a cadeia de DNA e ajuda a romper as ligações de hidrogênio entre as bases nitrogenadas. Durante o seu movimento ela adiciona a cadeia, uma sequência de ribonucleotídeos de acordo com o pareamento das bases nitrogenadas presentes no DNA. Imagem obtida de MUN, 2011.....20
- Figura 3 - Passos realizados para produzir proteínas através da tradução. a) processo de iniciação, b) processo de prolongação e c) terminação. Imagem obtida de UNM, 2002.....21
- Figura 4 - Três caminhos englobam as redes de proteínas: nós com alta conectividade, módulos de sinalização e vias com formação linear que conectam elementos de entrada e saída (PAPIN *et al.*, 2005). ....23
- Figura 5 - (a) Barreira de evolução do câncer em lesões pré-cancerosas (preCA). O índice de proliferação (P.I.), o índice de resposta ao dano do DNA (DDR.I.) e o índice de apoptose e senescência (A./S.I.) são mostrados para os tecidos normais, pré-câncer (preCA) e câncer de pulmão, melanoma e colonrretal. (b) Modelo de danos do DNA provocado por oncogenes na progressão e no desenvolvimento do câncer. A instabilidade genômica é um resultado direto dos oncogenes que são ativados por estresse na replicação ou por DSB no início do desenvolvimento do câncer, a instabilidade genômica é seguida pela instabilidade cromossômica (CIN) em câncer provocando uma adição de genes mutantes o que facilita o desenvolvimento do câncer (adaptada de HALAZONETIS *et al.*, 2008). ....31
- Figura 6 - Imagem de um microarranjo da *Affymetrix* com 1,3 cm x 1,3 cm de superfície (GOHLMANN *et al.*, 2009). ....36
- Figura 7 - Visão geral simplificada do protocolo *Affymetrix* de análise do nível de expressão. A partir do mRNA é gerado um cDNA usado na transcrição *in vitro*, resultando em uma fita simples de cRNA biotinizado que será fragmentado e hibridizado com os oligonucleotídeos das sondas. As intensidades fluorescentes são adquiridas com um scanner a laser. Imagem adaptada de BECKER *et al.*, 2003. ....38
- Figura 8 - Exemplo mostrando os níveis de expressão de 10 genes quaisquer, com diminuição da diversidade do experimento em relação ao controle ( $H_\alpha^e < H_\alpha^y$ ). Para este caso a diversidade dos genes de controle é maior do que o experimento, desta forma a diversidade relativa é menor que 0,5 deste modo  $h_\alpha = 0,405$ . ....42
- Figura 9 - Grafo de interações físicas entre as 564 proteínas pertencentes as Vias Ontocancro geradas com o banco de dados *String* com opções de interação: experimentos, bancos de dados e coexpressão e com intervalo de confiança de 0,7 (SIMÃO *et al.*, 2010). ....47
- Figura 10 - Projeção da atividade gênica relativa entre os microarranjos de adenoma do colorretal sobre a rede da figura 9. O gradiente de cores é definido pela atividade relativa ( $n_\alpha$ ), se  $n_\alpha > 0,55$  o gradiente aumenta de amarelo até vermelho caracterizando aumento de expressão; se  $n_\alpha < 0,45$  o gradiente

- diminui de azul ciano até azul escuro caracterizando diminuição de expressão e quando  $0,45 < n_{\alpha} > 0,55$  o gradiente é verde e não caracteriza alteração (SIMÃO *et al.*, 2010).....49
- Figura 11- Sequência de tempo de alterações de todas as vias em 0, 4, 8, 16 e 32 horas após a exposição de 2,5  $\mu$ M de Cádmio. O "#" representa a diminuição significativa no instante de tempo, APO - Apoptose, NER - Reparo por excisão de nucleotídeos, BER - Reparo por excisão de bases, MMR - Reparo de bases mal pareadas, HR - Recombinação homóloga, NHJ - Reparo por terminação não homóloga, CS - Estabilidade cromossômica e CC - Ciclo celular (SIMÃO *et al.*, 2010).....53
- Figura 12 - Perfil de significância das vias de Resposta ao dano no DNA (DDR), Apoptose e Ciclo celular para os estudos de colorretal, glândula adrenal, pâncreas e tireóide. O aumento significativo é identificado nos quadros vermelhos, a diminuição significativa nos quadros azuis e os quadros verdes indicam que não há alteração.  $n_{\alpha}$  é a atividade relativa e  $h_{\alpha}$  é a diversidade relativa. ....60
- Figura 13 - Diagramas de Venn usados para avaliar o consenso entre os genes alterados em adenomas e cânceres do colorretal, glândula adrenal, pâncreas e tireóide entre o consenso dos 5 genes com *fold change* mais alterados em cada subvia. Em vermelho estão os genes com evidências e ligados a barreira anticâncer. A razão de alteração entre o experimento (adenoma e câncer) *versus* o controle (tecido normal) foi calculado pela análise de expressão diferencial e os resultados são mostrados nos diagramas de Venn para: (a) e (b) o consenso para os genes da via DDR em adenoma e câncer, respectivamente. (c) e (d) o consenso para os genes da via de Apoptose em adenoma e câncer, respectivamente. (e) e (f) o consenso para os genes da via de Ciclo celular em adenoma e câncer, respectivamente.....65
- Figura 14 - O diagrama ilustra os mecanismos envolvidos na barreira contra o desenvolvimento tumoral com base nos resultados das figuras 12 e 13. Em vermelho estão os genes analisados no consenso. (a) Adenoma: A barreira é ativada em resposta aos danos do DNA, pela proteína *MRE11A* que atua sobre as quinases *ATM* e *ATR* que regulam as múltiplas cascatas de sinalização em resposta aos danos. Elas são reguladas pela *53bp1* e ativam os pontos de verificação, DDR e apoptose pela subvia de Ativação do *NFKB*. As quinases regulam o *CHEK2* e *CHEK1* que estão envolvidas na sinalização dos pontos de verificação. Estas proteínas ativam em conjunto as proteínas *p53* e *cdk6* para promover os pontos de verificação nas fases G1/S e G2/S. O *TP53* promove a parada do ciclo celular, reparo do DNA, ativação da senescência replicativa e apoptose. O reparo é ativado por duas subvias (NHEJ e HR) e pelos genes alterados *BRCA2*, *XRCC4* e *LIG4*. A apoptose é sinalizada pelo *NFKB* que ativa as subvias de sinalização *TNFF1/2* e que ativam o gene *CASP8*. A apoptose é também regulada pela subvia de Regulação da Apoptose e pelo gene alterado *CFLAR*. (b) Câncer: A figura mostra em cinza claro vários genes envolvidos no diagrama em adenoma que estão inibidos em câncer incapazes de ativar a barreira anticâncer. Muitos genes participam da subvia *RB/E2F* que regula o ciclo celular na transição da fase G1 para S. Um dos genes responsáveis pela regulação é o *TP53*. ....69

## LISTA DE ABREVIATURAS E SIGLAS

APO	Apoptose
BER	Reparo por excisão de bases
CC	Ciclo celular
DC	Doença de Crohn
CS	Estabilidade cromossômica
DDR	Reparação ao dano do DNA
DISC	Complexo de sinalização de morte induzida
DSB	Quebras em fitas duplas no DNA
DSS	Quebras em fitas simples no DNA
ENSP	Identificador da sequência de proteínas
GCRMA	Média robusta de múltiplos chips para <i>Affymetrix GeneChip</i>
GEO	<i>Gene Expression Omnibus</i>
GMM	Mecanismos de manutenção do genoma
HR	Reparo por recombinação homóloga
MAS 5.0	Microarranjo suíte versão 5.0
MM	Par incompatível ( <i>Mismatch</i> )
MMR	Reparo de bases mal pareadas
NER	Reparo por excisão de nucleotídeos
NHJ	Reparo por terminação não homóloga
PM	Par perfeito ( <i>Perfect Match</i> )
RMA	Média robusta de múltiplos chips
TNF	Fator de necrose tumoral
UC	Colite Ulcerosa

# SUMÁRIO

1. INTRODUÇÃO .....	12
2. FUNDAMENTAÇÃO TEÓRICA.....	17
2.1 A informação genética .....	17
2.2 Redes biológicas.....	22
2.3 Mecanismos de manutenção do genoma na evolução do câncer .....	25
2.4 Bancos de dados em biologia sistêmica .....	32
2.4.1 Ontologia Ontocancro 1.0 e 2.0 .....	34
2.5 Microarranjos.....	35
2.5.1 Normalizações de microarranjos .....	39
2.5.2 Análises estatísticas de expressões de vias genéticas.....	40
2.5.3 Análise de genes diferencialmente expressos - <i>fold change</i> .....	43
3. MODELAGEM DA REDE DE MANUTENÇÃO DO GENOMA HUMANO .....	46
3.1 Rede de interação de proteínas envolvidas em GMM.....	46
3.2 Análises das vias de manutenção do genoma em estudos de adenoma do colorretal .....	48
3.3 Resposta temporal das vias GMM em células da próstata expostas ao Cádmio .....	52
4. BARREIRA ANTICÂNCER NA EVOLUÇÃO DE TUMORES.....	54
4.1 Descrição das vias e subvias .....	55
4.2 Dados de microarranjos.....	57
4.3 Ferramentas de pesquisa usadas para construir o modelo de barreira anticâncer .....	58
4.3.1 Alterações em subvias de manutenção do Genoma.....	58
4.3.2 Consenso dos genes diferencialmente expressos .....	64
4.3.3 Genes alterados em tecidos do colorretal inflamados .....	67
4.4 Diagrama da barreira anticâncer .....	68
4.4.1 Ativação da barreira em tecidos de adenoma .....	68
4.4.2 Perda de subvias e genes em tecidos cancerosos .....	71
4.5 Conclusões.....	73
5. CONCLUSÕES.....	74
6. REFERÊNCIAS .....	76
ANEXOS .....	86

# 1. INTRODUÇÃO

A análise científica dos seres vivos tem sido focada em subdividir os componentes da vida em partes menores. O método científico denominado reducionismo científico, introduzido pelo filósofo e matemático René Descartes (1596-1650) foi usado por muitos anos para descrever as partes constituintes mais simples dos fenômenos e teorias que dão sentido a vida (AHN *et al.*, 2006). Ao longo dos anos as teorias introduzidas por Descartes foram expandidas por Isaac Newton que considerou a mecânica clássica como a base para a pesquisa científica (MAZZOCHI, 2008).

No início do século XX as pesquisas tomaram novos rumos com os estudos vinculados à exploração específica dos reinos atômicos e subatômicos. As novas descobertas passaram a mudar os paradigmas e uma nova análise intelectual fez surgir novas ideias vinculadas aos estudos do universo, o que acabou refletindo na física quântica (MAZZOCHI, 2008).

Alguns biólogos moleculares da década de 1950, que migraram da física, expandiram a abordagem clássica para o estudo de organismos vivos. A partir daí os pesquisadores adotaram uma abordagem reducionista para entender os sistemas biológicos de acordo com as propriedades químicas e físicas (WESTERHOFF *et al.*, 2004). Com a descoberta do código genético e da estrutura de DNA por James Watson e Francis Crick em 1953, teve início uma revolução científica embasada na descoberta de novas teorias envolvendo o genoma (KEEDWELL *et al.*, 2005). Com o desenvolvimento tecnológico foram criadas ferramentas que possibilitaram a análise de grandes quantidades de dados genéticos, que foram produzidos ao longo do tempo.

A interpretação dos dados levou os pesquisadores a uma mudança de paradigma e a descobrir como as ciências da vida se relacionam. A integração das informações induziu à execução dinâmica de instruções associadas com o desenvolvimento de organismos e suas respostas fisiológicas aos seus ambientes. Com isso, foi necessário quantificar todos os elementos moleculares de um sistema biológico para avaliar suas interações e integrar essa informação a modelos que serviram como hipóteses de previsão para explicar as propriedades emergentes (HOOD *et al.*, 2004).

A biologia sistêmica faz a integração entre os fenômenos e as teorias que envolvem os sistemas biológicos, e faz ainda, emergir as funções e os comportamentos sistêmicos que definem os organismos vivos. Ela trata a genética como uma ciência relacionada à informação, levando em conta uma abordagem sistêmica, assim como, a codificação e a regulação da expressão genética no espaço e no tempo. A genômica, a proteômica e a transcriptômica, são usadas para descrever a hierarquia das ÔMICAS. Através desta hierarquia pode-se verificar o conjunto de alterações dos genes e expressões protéicas ou metabólicas em um experimento. Dentro da biologia sistêmica a genômica descreve como os genes funcionam, interagem entre si e respondem aos estímulos ambientais, através da atividade ou expressão do gene. A proteômica descreve a maneira como as proteínas se transformam e interagem como resultado da atividade dos genes. A análise do nível de expressão dos genes é descrita pela transcriptômica. Essas e outras ÔMICAS interagem como uma rede ou interatoma. Os interatomos são formados por macromoléculas (genes ou proteínas) e as suas interações são propriedades emergentes (KEEDWELL *et al.*, 2005).

A evolução da informação tornou possível armazenar e processar tanto os interatomos quanto os transcriptomas explorados pela hierarquia ÔMICA. Um destes métodos, denominado como microarranjo de DNA, é usado para analisar os transcriptomas e medir os níveis de expressão de grandes números de genes transcritos simultaneamente. Devido ao grande número de genes existem inúmeros bancos de dados usados para armazenar os transcritos produzidos através de microarranjos. O banco de dados *Gene Expression Omnibus* (GEO) desenvolvido por Ron Edgar e colaboradores em 2001, é um repositório público de microarranjos (EDGAR *et al.*, 2002). O primeiro microarranjo foi produzido em 1982 quando Leonard Augenlicht e colaboradores mediram a expressão de 378 colônias de bactérias (*Escherichia coli*). Em humanos a técnica evoluiu com a análise de uma colônia de células de câncer que continha 4000 DNAs complementares (AUGENLICHT *et al.*, 1987). Desde então, a tecnologia formada por um conjunto de todas as transcrições de genes de uma célula, evoluiu ainda mais e muitos outros microarranjos foram produzidos para elucidar o genoma de vários organismos. Mark Shena, em 1995, aprimorou a técnica tornando os microarranjos mais compactos (SCHENA *et al.*, 1995).

## **Câncer**

Desde o primeiro estudo envolvendo microarranjos de câncer em 1987, o aprimoramento e a busca pela cura de algumas doenças estimularam a utilização desta técnica. Em uma das pesquisas mais precursoras, envolvendo o estudo do câncer com microarranjos, DeRisi e colaboradores em 1996, usaram 1161 DNAs complementares para procurar diferenças na expressão gênica associada à supressão do câncer (DERISI *et al.*, 1996).

O câncer sempre foi um elemento relevante dentro das pesquisas relacionadas às doenças humanas. Existem vários tipos de câncer emergentes de várias maneiras que evoluem de acordo com os fatores ambientais e genéticos. Cerca de 1% do genoma humano apresenta implicações na tumorigênese e no desenvolvimento tumoral. A perda da estabilidade do genoma associada com a deterioração genética é um dos muitos aspectos importantes dos carcinomas (JEFFORD *et al.*, 2006). A displasia é caracterizada como uma lesão celular reversível e pelo crescimento anormal de células que poderão evoluir ao câncer passando por três estágios pré-cancerosos, que são definidos como: adenoma primário, intermediário e tardio (LENHARD *et al.*, 2001). Em células normais a instabilidade genômica pode levar a quebras em fitas simples do DNA (DSS) ou a defeitos mais complexos que envolvem quebras em fitas duplas no DNA (DSB) (GORGOLIS *et al.*, 2005). São os mecanismos de manutenção do genoma (GMM) que corrigem estas quebras e garantem a integridade do DNA e a sobrevivência da célula. Os GMM envolvem um conjunto de vias e proteínas que atuam na resposta aos danos (DDR). Na presença de qualquer alteração que envolva instabilidade genômica são ativados os pontos de verificação do ciclo celular, que irão identificar os defeitos e ativar proteínas supressoras de tumor ligadas aos mecanismos de resposta aos danos tais como: a morte celular programada, o reparo e a recombinação do DNA em resposta aos danos (CASTRO *et al.*, 2007; SIMÃO *et al.*, 2010).

## **Pré-câncer**

Os tecidos pré-cancerosos formam um tipo de neoplasia benigna que poderá evoluir para o câncer. Segundo o modelo publicado em 2008 por Halazonetis e

colaboradores há evidências de que tecidos pré-cancerosos apresentam diminuição da expressão em alguns mecanismos de manutenção do genoma relacionados ao índice de proliferação celular, seguidas por um aumento da atividade das vias de reparação ao dano do DNA (DDR) e morte celular programada (HALAZONETIS *et al.*, 2008). Com o aumento das atividades das vias de Apoptose e Reparo do DNA é formada uma barreira contra a progressão do câncer. Uma vez rompida à barreira, há o acúmulo de mutações em células geneticamente alteradas que irão favorecer o desenvolvimento do câncer (LOEB, 2001). As alterações em tecidos pré-cancerosos iniciam pela ativação de oncogenes que irão desencadear uma proliferação aberrante. Em resposta ao estresse replicativo e as DSB do DNA o tecido inicia a resposta aos danos pela ativação de vias especializadas na resposta aos danos (HALAZONETIS *et al.*, 2008).

## Objetivos

Em 2008, Halazonetis e colaboradores propuseram um modelo qualitativo que descreve uma barreira de evolução tumoral, em tecidos pré-cancerosos, responsável pela ativação das vias de Apoptose e de Resposta aos danos, mas quando rompida da origem a um tecido maligno (HALAZONETIS *et al.*, 2008).

O objetivo deste trabalho foi investigar a ativação da barreira anticâncer, em tecidos de adenoma, usando os métodos quantitativos de expressão de genes e vias produzidos pela biologia sistêmica incluindo os métodos de atividade relativa, diversidade relativa e a mudança de expressão (*fold change*) entre os genes. Para descrever a barreira foi proposta a modelagem das vias envolvidas na manutenção do genoma, e em seguida foi desenvolvido um diagrama que descreve as interações entre as vias e entre os genes envolvidos na manutenção do genoma pela ativação da barreira anticâncer. O diagrama complementa os estudos qualitativos relacionados a evolução tumoral através das análises estatísticas que unem em um único modelo as vias, subvias e genes responsáveis por ativar um conjunto de mecanismos que impede a evolução para o câncer.

## Organização do trabalho

A tese está estruturada na forma de capítulos temáticos, que inclui uma revisão bibliográfica da fundamentação teórica apresentada no capítulo 2 e os resultados de dois artigos científicos:

Capítulo 3: contém a descrição dos resultados do artigo publicado na revista *Physica A* (SIMÃO *et al.*, 2010). Neste capítulo são descritos os resultados referentes a modelagem da rede de manutenção do genoma, análise de vias em transcriptomas de adenoma e tempo de resposta das células da próstata a baixas doses de Cádmiio.

Capítulo 4: contém a descrição dos resultados de um artigo submetido para publicação (SIMÃO *et al.*, 2012). Neste capítulo há uma descrição do diagrama de proteínas que atua na barreira anticâncer. Este diagrama foi desenvolvido a partir de análises quantitativas da expressão de vias e dos genes envolvidos na manutenção do genoma.

Capítulo 5: apresenta uma conclusão geral para os artigos abordados nesta pesquisa.

## 2. FUNDAMENTAÇÃO TEÓRICA

Este capítulo descreve alguns dos principais tópicos relacionados aos fundamentos biológicos da área. Estão apresentados os conceitos básicos sobre o processo de transcrição e tradução da informação genética, redes de proteínas, vias de interação, mecanismos de manutenção do genoma, bancos de dados de biologia sistêmica e uma abordagem sobre os microarranjos de DNA.

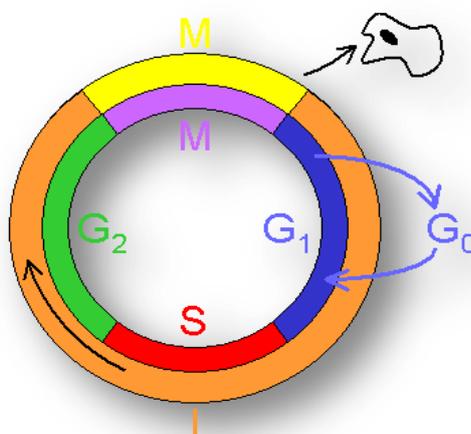
### 2.1 A informação genética

A característica mais importante de uma célula viva é a sua capacidade de transferir propriedades hereditárias de uma geração para a outra. O dogma central é desvendar como essa informação presente no genoma é transferida as novas gerações de células e como elas reagem às adversidades do meio.

A maioria das células em um mesmo organismo possui o mesmo genoma, independente do estágio de desenvolvimento e das condições ambientais (KEEDWELL *et al.*, 2005). Dependendo destes e de outros fatores funcionais como, por exemplo, a cor dos olhos, a célula codifica diferentes partes do genoma em proteínas. No genoma humano o número de genes que contêm informações de codificação de proteínas varia entre 25.000 a 30.000 (GOHLMANN *et al.*, 2009).

Os genomas são codificados em sequências formadas de uma série de cadeias de ácido desoxirribonucléico (DNA). Este, por sua vez, é um composto orgânico que contém todas as instruções genéticas que coordenam o desenvolvimento e o funcionamento de todos os seres vivos. Sua estrutura química é formada por dois filamentos compostos de ácido fosfórico e açúcar (pentose) que se entrelaçam formando uma dupla hélice. Os filamentos são conectados através de ligações de hidrogênio que ligam duas bases (WATSON *et al.*, 2004). A informação genética está codificada em uma sequência específica formada a partir de quatro diferentes bases de nucleotídeos: adenina (A) e guanina (G) (purinas), citosina (C) e timina (T) (pirimidinas). Uma composição diferenciada está presente na fita simples do mRNA, que é produzida na transcrição quando a base da timina é transferida do DNA para o mRNA, na forma de uracila (U).

As informações contidas no genoma de uma célula serão transferidas a células filhas através da divisão celular conforme mostrado na figura 1. Os eventos necessários para a divisão celular são chamados de ciclo celular. O ciclo celular pode ser dividido em quatro fases:  $G_1$ , S,  $G_2$  e M. Na fase  $G_1$  também conhecida como fase de crescimento, ocorre uma síntese intensa de várias moléculas importantes para a sobrevivência da célula, como proteínas estruturais, enzimas e RNA. Na fase S acontece a duplicação do DNA seguida pela fase  $G_2$  onde acontece a síntese de moléculas e organelas relacionadas ao processo de divisão celular. O ciclo celular é concluído, na fase M, quando ocorre a segregação cromossômica (WATSON *et al.*, 2004).



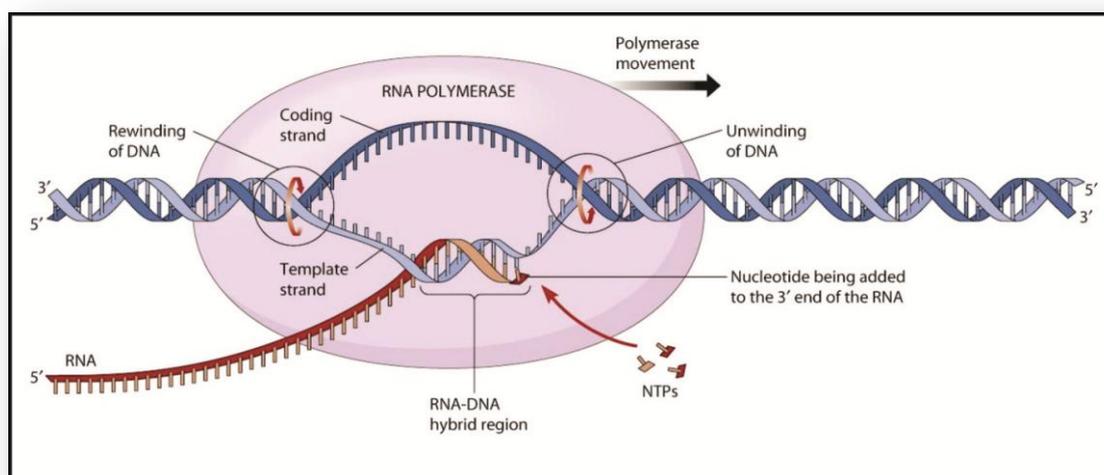
**Figura 1** - Ciclo da divisão celular: o ciclo inicia na fase  $G_1$  ou  $G_0$  com a síntese de moléculas, passa pelas fases S, onde há duplicação do DNA segue pela fase  $G_2$  e é concluído na fase M, quando ocorre a segregação cromossômica. Imagem obtida de WIKIBOOKS, 2011.

As fases  $G_1$  e  $G_2$  também são conhecidas como fases de intervalo do ciclo celular. Elas possuem dois propósitos: fornecem tempo para a célula se preparar para a próxima fase do ciclo e para verificar se a fase anterior foi completada corretamente. Nestas duas fases ocorrem pontos de verificação do DNA. As células com DNA danificados param o ciclo celular em  $G_1$ , antes da síntese, ou em  $G_2$  antes da mitose (fase M), para impedir que esses eventos ocorram com cromossomos danificados. O atraso dá tempo para que a lesão seja corrigida, antes que a célula prossiga no ciclo celular.

Entre as fases  $G_1$  e  $G_2$  a fase de síntese ou fase S engloba o processo de duplicação do DNA. O modelo estrutural do DNA proposto por Watson e Crick explica que na duplicação dos genes as duas cadeias do DNA se separam e cada uma delas orienta a fabricação de uma metade complementar. A duplicação do DNA

é semiconservativa, cada metade da molécula original se conserva íntegra em cada uma das duas moléculas filhas (WATSON *et al.*, 2004). Os procedimentos de formação de novas réplicas do DNA iniciam com a ativação das enzimas topoisomerasas que são responsáveis por desenrolar a hélice de DNA, diminuindo a tensão na medida em que a enzima helicase avança para quebrar as pontes de hidrogênio entre as bases. Uma vez separadas as duas cadeias de DNA, a enzima DNA polimerase promove a polimerização das novas fitas de DNA. A enzima sintetiza uma nova cadeia de DNA modelada a partir das bases nitrogenadas presentes no DNA original. A polimerização ocorre em sentidos contrários nas cadeias separadas pela helicase. Uma de 5' para 3' (cadeia líder) e na outra cadeia da 3' para 5' (cadeia atrasada). Na cadeia atrasada os fragmentos de Okasaki (pequeno fragmento de DNA com um *primer* de RNA no terminal 5') são ligados pela enzima DNA ligase que atua antes do processo de polimerização da enzima DNA polimerase. O *primer* de RNA é um fio curto de RNA que será orientado da posição 5' pela DNA ligase para se ligar a cadeia de DNA pela enzima DNA polimerase. Ao final da replicação do DNA têm-se duas fitas duplicadas idênticas a original (WATSON *et al.*, 2004).

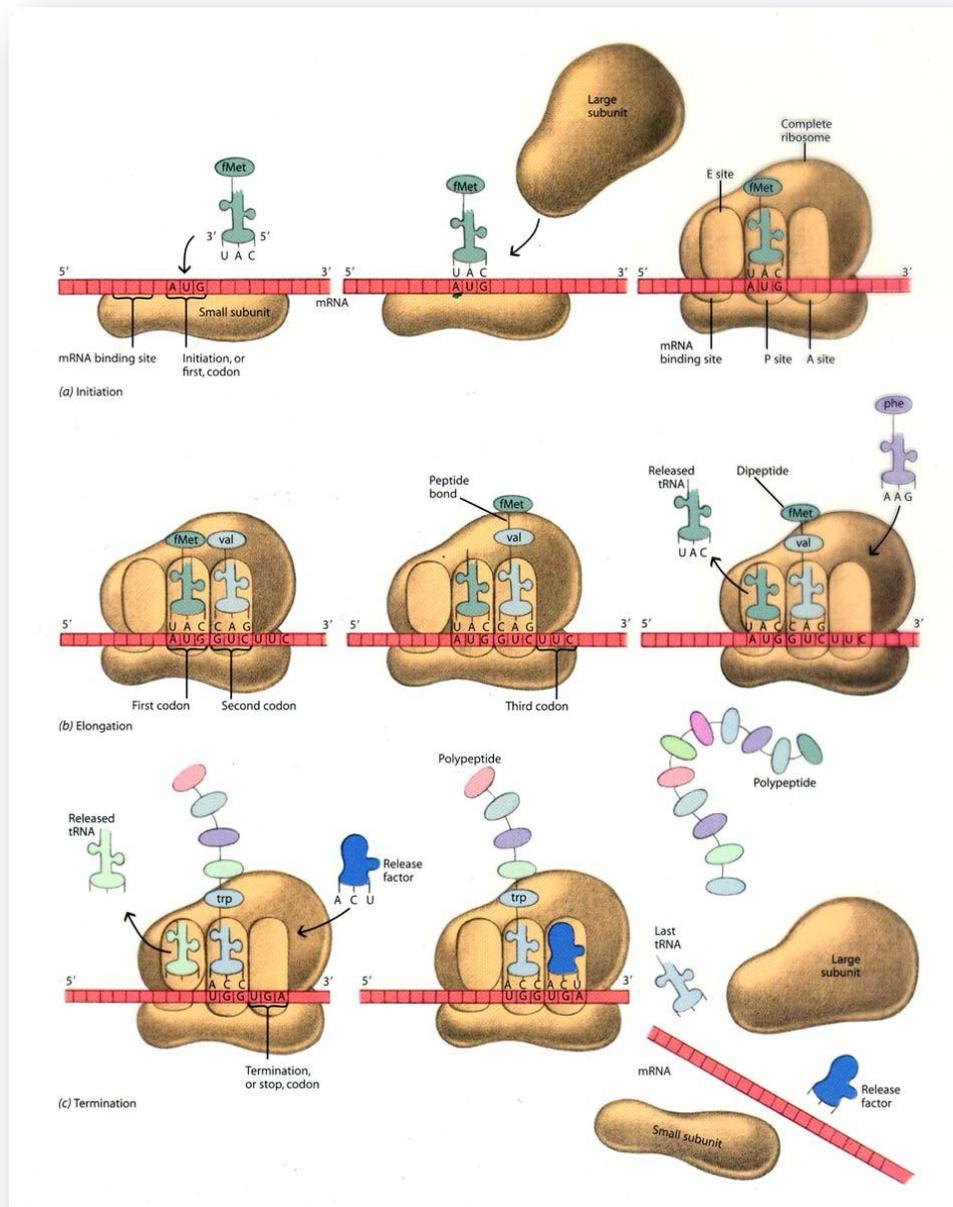
A transcrição é um processo de síntese de uma molécula de mRNA a partir da informação biológica codificada na molécula de DNA. O mRNA contém a informação que determina a sequência de aminoácidos de uma proteína. Sendo assim, o mRNA direciona a síntese protéica. Em um fragmento de uma molécula de DNA, mais precisamente em um gene delimitado por uma região promotora (local onde se inicia a síntese, que é processada da extremidade 5' para a 3' na cadeia de DNA). A RNA polimerase mostrada na figura 2 sintetiza o RNA percorrendo a cadeia de DNA até a outra extremidade do gene onde se encontra uma região terminal da síntese do RNA. Os fatores de transcrição são proteínas que se ligam ao DNA e auxiliam a enzima RNA polimerase a romper as ligações de hidrogênio entre as bases nitrogenadas dos dois filamentos de DNA. A RNA polimerase percorre o gene e adiciona a cadeia uma sequência de ribonucleotídios de acordo com o pareamento das bases nitrogenadas presentes no DNA (KEEDWELL *et al.*, 2005). Ao final do processo de transcrição, a RNA polimerase se dissocia e é sintetizada uma nova molécula de mRNA que será utilizada no processo de tradução de proteínas (GOHLMANN *et al.*, 2009).



**Figura 2** - O RNA polimerase percorre toda a cadeia de DNA e ajuda a romper as ligações de hidrogênio entre as bases nitrogenadas. Durante o seu movimento ela adiciona a cadeia, uma sequência de ribonucleotídeos de acordo com o pareamento das bases nitrogenadas presentes no DNA. Imagem obtida de MUN, 2011.

A tradução é o processo de síntese de proteínas a partir de um mRNA, esse processo é dividido em três partes: iniciação, prolongação e finalização (figura 3). O mRNA possui na sua extremidade 3' a cauda poli A composta por uma sequência de adeninas (A), as quais servem como proteção contra a degradação enzimática do mRNA. Logo após a cauda poli A o mRNA é composto por uma sequência de bases e ao final da sequência estão três códons de finalização (UGA, UAA ou UAG) que irão delimitar o término da tradução. Na extremidade 5' do mRNA há uma cápsula de metionina que indica o início de onde a subunidade menor do ribossomo se ligará para iniciar o processo de tradução (LODISH *et al.*, 2003).

O tRNA é uma molécula fundamental que contém um anticódon que é complementar ao códon do mRNA o qual irá ligar-se. Os códons são sequências de três nucleotídeos do mRNA que na síntese de proteínas representam a codificação da informação genética para um determinado aminoácido ou para a terminação da cadeia polipeptídica. O anticódon é também composto por uma sequência de três nucleotídeos de ácido ribonucléico transportador que se liga ao códon de uma molécula de mRNA (KEEDWELL *et al.*, 2005).



**Figura 3** - Passos realizados para produzir proteínas através da tradução. a) processo de iniciação, b) processo de prolongação e c) terminação. Imagem obtida de UNM, 2002.

O processo de codificação inicia quando a subunidade maior do ribossomo se liga a subunidade menor para criar o sítio P e o sítio A. O primeiro tRNA ocupa o sítio P, e é composto pela sequência AUG que codificará o aminoácido metionina (Met), em seguida, o segundo tRNA entra no sítio A e é complementar ao segundo códon, então a metionina é transferida para o aminoácido do sítio A. O primeiro tRNA é descartado deixando ligado a cadeia o aminoácido metionina. O ribossomo move-se ao longo do mRNA e o próximo tRNA entra no sítio A para prolongar a

cadeia de aminoácidos. Na medida em que o processo continua, a cadeia polipeptídica é continuamente transferida para o tRNA do sítio A. Quando o códon de finalização é encontrado no sítio A um fator de liberação entra no sítio e a tradução termina dissociando o ribossomo e dando origem a uma proteína (PALSSON, 2006).

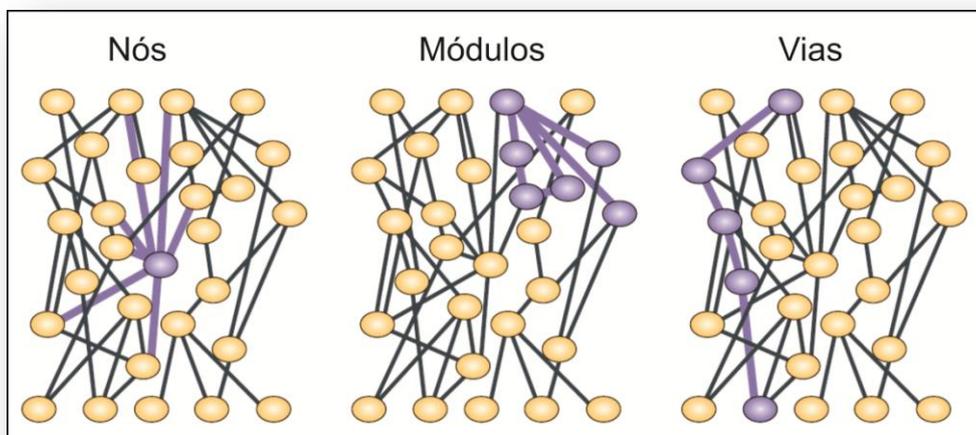
As proteínas compõem cerca de 50 a 80% do peso seco da célula, sendo, portanto, o composto orgânico mais abundante de matéria viva. Elas são polímeros unidos por ligações peptídicas (macromoléculas formadas pela união de várias moléculas menores: aminoácidos). Tem funções estruturais, ou seja, fornecem boa parte da estrutura do citoesqueleto das células e funções enzimáticas, tais como a produção de enzimas que realizam várias reações químicas dentro das células. São responsáveis por carregar sinais de uma parte do organismo para outra, ou de uma célula para outra. Atuam como um sistema de transporte, carregando moléculas como o oxigênio na circulação sistêmica. No sistema imunológico as proteínas detectam patógenos no organismo humano e ajudam a montar um sistema eficaz de defesa imunológico (KEEDWELL *et al.*, 2005).

## 2.2 Redes biológicas

As proteínas estão organizadas em redes funcionais, responsáveis por disseminar funções específicas dentro do organismo onde os nós que representam as proteínas dentro de uma rede de interações estão fortemente associados com propriedades biológicas distintas (CARLSON *et al.*, 2006). A descrição completa da função de uma proteína requer o conhecimento de todas as proteínas com quem ela interage. A partir de uma perspectiva funcional em grupo, a associação forma uma ligação física direta ou indireta, tal como a participação da proteína na mesma via metabólica ou processo celular (MERING *et al.*, 2005; EISENBERG *et al.*, 2000).

Na matemática uma rede de interações é chamada de grafo. A construção de grafos de proteínas é definida por três elementos principais: nós, módulos e vias (figura 4). A primeira abordagem consiste nos nós que possuem alta conectividade e abrangem os compostos e as reações que estão associadas com uma determinada proteína. A sinalização de nós altamente conectados indicam ativação de proteínas

que exercem funções importantes dentro do complexo funcional de uma célula (PAPIN *et al.*, 2005).



**Figura 4** - Três caminhos englobam as redes de proteínas: nós com alta conectividade, módulos de sinalização e vias com formação linear que conectam elementos de entrada e saída (PAPIN *et al.*, 2005).

A segunda abordagem engloba os módulos que são constituídos por grupos de compostos e proteínas que funcionam juntos sob certas condições. Eles representam as proteínas que desempenham funções relacionadas, ou seja, são proteínas muito conectadas que atuam em vias diferentes (PAPIN *et al.*, 2005).

Como terceiro elemento as vias conectam as entradas de sinalizações as sinalizações de saídas. Por exemplo: tal via pode ser a delimitação de todos os passos da ligação de um fator de crescimento ao seu receptor até a posterior ativação de um fator de transcrição que induz a expressão de genes alvos (PAPIN *et al.*, 2005). Em outras palavras, as vias determinam uma ligação funcional entre proteínas que apresentam funções semelhantes. Se a função de uma das proteínas é conhecida, então pode-se inferir que as proteínas ligadas agem na mesma via ou em complexos (EISENBERG *et al.*, 2000). Dentro do contexto de vias, existem as subvias que representam um conjunto de genes com funções bem específicas. Por exemplo, pode-se definir a via de apoptose (morte celular) e uma subvia como um caminho dentro da via de apoptose, ou seja, subvia extrínseca de apoptose.

Na teoria de grafos dirigidos, com interação em um único sentido, um caminho é uma cadeia de nós conectados por arestas dirigidas, sem ramificações ou ciclos. As vias nestas redes podem representar, por exemplo, um caminho de

transformação de um nutriente em um produto em uma rede metabólica, ou seja, dentro de um caminho de transformação a via relaciona-se com a interação entre as proteínas necessárias para desencadear uma função específica (AITTOKALLIO *et al.*, 2006).

O conjunto de interações que ocorrem em uma via dependem de funções individuais, onde, cada um dos nós irá segregar funções que irão ativar ou inibir produtos que atuarão em outros elementos da via, ou, em outras proteínas que irão desempenhar outras funções relacionadas. Uma alteração em um elemento da via poderá interromper a conexão entre genes informativos que agem antes e depois do elemento alterado. Isso é similar a uma ruptura em uma linha telefônica entre duas cidades. Muitas vezes é possível que os dois grupos de proteínas isolados ainda terão alguns elos de comunicação, apesar de terem perdido a capacidade de ter uma atividade completa na via específica (AITTOKALLIO *et al.*, 2006).

Considerando o grande número de organismos e as informações relevantes associadas às vias, alguns bancos de dados visam simplificar o acesso a informações e a conjuntos de genes relacionados a funções específicas em um determinado organismo. O *National Cancer Institute - Pathways Interaction Database* é um banco de dados público que armazena vias relacionadas ao desenvolvimento do câncer. Além das vias curadas do *NCI - Nature* o banco de dados apresenta vias importadas dos bancos de dados *BioCarta* e *Reactome*. As vias relacionadas ao câncer poderão ser encontradas em ambiente virtual<sup>1</sup> (mais informações sobre os bancos de dados de vias e de genes estão apresentadas na seção 2.4).

As analogias entre as vias e as redes de proteínas estão intimamente relacionadas quando se pretende integrar as funções isoladas em cada conjunto de genes. Alguns bancos de dados visam simplificar o acesso a essas informações, fornecendo um conjunto abrangente, mas com controle das interações entre as proteínas (MERING *et al.*, 2005). O banco de dados *String* integra várias informações de associação entre proteínas a partir de mineração de bases de dados e da literatura, além de previsões com base em análise do contexto genômico. Na versão 8.0, publicada em 2009, são encontradas mais de 2,5 milhões de proteínas em 630 organismos (JENSEN *et al.*, 2009). As interações entre as proteínas de cada

---

<sup>1</sup> Disponível em: < <http://www.ncbi.nlm.nih.gov> >

organismo são associadas por seis comparações genômicas sistemáticas. As interações por conservação de genes vizinhos englobam proteínas que possuem um vínculo funcional. Muitas vezes um dos dois genes é um regulador de transcrição visando o gene vizinho (JENSEN *et al.*, 2009). A associação de genes por eventos de fusão explora o fato de que as famílias de determinadas proteínas em uma dada espécie consistem em domínios fundidos que normalmente correspondem ao comprimento total de proteínas em outras espécies. Essas proteínas foram definidas como proteínas de fusão e suas famílias como componentes de proteínas (ENRIGHT *et al.*, 1999). As interações de co-ocorrência ocorrem quando há transferência de interações entre organismos, duas proteínas que interagem em um organismo servirão de evidência para a interação em outro organismo (JENSEN *et al.*, 2009). Coexpressão exhibe interações entre as proteínas que apresentam expressões semelhantes a partir da expressão de um primeiro gene (JENSEN *et al.*, 2009). As interações experimentais são extraídas de experimentos "*in vitro*" onde ocorre a ligação de duas proteínas. As interações por bancos de dados incluem interações procedentes de outros bancos de dados. E, por fim, as interações por mineração de texto são extraídas através de interações de proteínas publicadas em resumos e em outros textos científicos (MERING *et al.*, 2007). Toda pesquisa de interações entre proteínas processadas pelo banco de dados *String* são ainda complementadas pela análise de confiança que inclui uma normalização pelo número de espécies abrangidas e pelos genes envolvidos (MERING *et al.*, 2003).

### **2.3 Mecanismos de manutenção do genoma na evolução do câncer**

Os tumores são formados por células que perderam a capacidade de montar e criar tecidos com formas e funções normais. Desta forma, o câncer é visto como uma doença do mau funcionamento celular (WEINBERG, 2007). Através de uma sequência de mutações aleatórias uma célula perde a capacidade de regular seu crescimento. Com o desenvolvimento acelerado e o rompimento de barreiras que contribuem para a evolução das mutações, as células continuam a se dividir acumulando mais alterações e contornando proteções que impedem a evolução de danos (JEFFORD *et al.*, 2006). O resultado do acúmulo de mutações será um tumor

formado exclusivamente por uma massa de células malignas. Na medida em que se multiplicam, as células malignas podem invadir a corrente sanguínea e entrar em metástase espalhando-se para outros órgãos dando origem a um tumor mais agressivo (WEINBERG, 2007). Com a evolução das alterações, o câncer aumenta o seu índice de proliferação seguido da perda da adesão celular, reparação do DNA e morte celular programada (CASTRO *et al.*, 2007).

Vincent e Gatenby em 2008 descrevem três etapas importantes na formação do câncer. No processo conhecido como iniciação, ocorre a exposição das células a agentes mutagênicos, tais como, vírus, agentes químicos, irradiações ou por fatores hereditários. No passo de promoção do desenvolvimento, o tumor pré-maligno apresenta poucas alterações no tecido e pode ainda regredir a um tecido normal. Na progressão do tumor pode ocorrer uma transformação maligna, o tumor toma características próprias e se dissocia do tecido normal (VINCENT *et al.*, 2008).

O pré-câncer possui potencial para causar alterações em tecidos saudáveis. Embora exista uma pré-disposição maligna, o pré-câncer poderá regredir a um tecido normal ou ter a proliferação tumoral estagnada. O adenoma é um tipo de pré-câncer com origem glandular onde são observadas inúmeras alterações (WEINBERG, 2007). Vassilis Gorgoulis e colaboradores (2005) propuseram que em adenomas há estresse na replicação do DNA, o que leva às quebras em fitas duplas no DNA (DSB) e a mutações no gene *TP53* (Proteína 53 ou *tumor protein 53, p53*) (GORGOULIS *et al.*, 2005). As DSBs são danos agressivos que podem ser reparados por dois diferentes tipos de mecanismos. O primeiro deles conhecido como reparo por recombinação homóloga (HR) obtém informações de cromossomos homólogos para reparar o DNA duplamente quebrado (ALBERTS *et al.*, 2002). O segundo mecanismo envolve o reparo por terminação não homóloga (NHJ), onde o reparo é realizado pela união aleatória de quaisquer extremidades do DNA (FERGUSON *et al.*, 2000). Além dos reparos por terminação não homóloga e recombinação, existem outros mecanismos de reparo que irão corrigir as quebras nas fitas simples no DNA (DSS) gerados pela exposição a agentes cancerígenos, tais como a luz ultravioleta. Segundo Wood e colaboradores existem três tipos de reparos que requerem funções diferentes e irão corrigir as alterações provocadas aos nucleotídeos (WOOD *et al.*, 2001):

- Reparo por excisão de nucleotídeos (NER): este reparo inclui a distorção da dupla hélice do DNA e a danos causados a fita simples. O NER retira os

nucleotídeos danificados em blocos e adiciona um novo nucleotídeo que será alocado no ponto onde ocorreu o dano ao DNA.

- Reparo por excisão de bases (BER): ocorre quando os agentes mutagênicos atingem as bases nitrogenadas. O BER remove a base defeituosa por um processo de clivagem da ligação base nitrogenada (desoxirribose), seguida pelo preenchimento da base correta por ação da DNA polimerase.

- Reparo de bases mal pareadas (MMR): este reparo é responsável pela correção de erros realizados na replicação do DNA e na recombinação de genes que resulta em um mau pareamento dos nucleotídeos.

A proteína *p53*, conhecido como supressor de tumor, age como um fator de transcrição, que se liga a enzima RNA polimerase permitindo a transcrição do RNA. O estresse na replicação do DNA provoca mutações do *TP53* que por sua vez apresenta abrangente importância na prevenção de tumores. Na fase  $G_1$  do ciclo celular a proteína *p53* atua como uma barreira que impede a propagação de danos ao DNA. Ela sinaliza outras proteínas que ativarão a morte celular programada ou a senescência, impedindo a progressão do dano. Se as alterações atingirem o *TP53* é possível que o dano seja transferido às novas células e uma sucessão de mutações ocorra dando origem a um tumor maligno (WEINBERG, 2007). Para garantir a transmissão de alta fidelidade da informação genética, as células desenvolveram mecanismos para monitorar a integridade do genoma (pontos de verificação) que irão responder ao dano no DNA, ativando vias complexas de resposta. São os mecanismos de manutenção do genoma (GMM) que garantem a integridade do genoma e sobrevivência da célula. Estas vias incluem a parada do ciclo celular, a morte celular programada, o reparo e a recombinação do DNA (HOEIJMAKERS *et al.*, 2007; CASTRO *et al.*, 2007; SIMÃO *et al.*, 2010).

Os pontos de verificação do ciclo celular garantem a fidelidade da proliferação celular em cada fase celular. A função mais importante dos pontos de verificação é de avaliar os danos ao DNA e tentar corrigi-los através de sinais que inibem a divisão celular para que o reparo seja realizado. Caso o reparo não seja atingido os pontos de verificação poderão ativar mecanismos de destruição da célula através da morte celular programada (apoptose) (KASTAN, *et al.*, 2004). Dependendo da fase da divisão celular em que ocorram as alterações há ativação de diferentes pontos de verificação: Na Fase  $G_1$  do ciclo celular ocorre a decisão de continuar a divisão celular ou não. Nesta fase a presença de alterações no DNA poderá levar a célula a

uma fase de parada temporária chamado de  $G_0$  ou interfase indefinida (quiescente). Dependendo do tipo de dano e da célula, poderá ocorrer uma estimulação e o retorno da célula ao ciclo celular, caso contrário, a célula poderá atingir o estágio de senescência. No final da fase  $G_2$ , ocorre o segundo ponto de verificação do ciclo celular. Se a célula ultrapassar essa barreira de controle ocorrem processos moleculares que sinalizam o início da mitose (WEINBERG, 2007).

A senescência é conhecida como um processo natural de envelhecimento celular que envolve a deterioração dos telômeros e ativação de genes supressores de tumor que impedem a transformação de uma célula senescente em um câncer. (BARTKOVA *et al.*, 2006). Os telômeros são localizados nas extremidades dos cromossomos e ficam mais curtos na medida em que os filamentos do DNA telomérico se dividem. Algumas proteínas apresentam dificuldade em rearranjar o DNA para que ocorra a divisão celular, quando essas proteínas não conseguem encontrar o início do DNA para que ocorra a cópia, o telômero diminui. Quando os telômeros atingem um tamanho crítico podem ocorrer fusões entre os cromossomos e estes poderão se prender uns nos outros e provocar alterações celulares, por esse motivo a célula é programada para cessar as divisões e entrar em senescência (HOLDBROOK *et al.*, 1996).

A senescência pode ser desencadeada também por danos ao DNA e tem como objetivo proteger o organismo contra o desenvolvimento do câncer. Em tecidos epiteliais a incidência de câncer ocorre exponencialmente com a idade e associado a esse processo, há uma maior participação dos processos senescentes. A partir dos 50 anos de idade, os seres humanos adotam uma capacidade exponencial de transformar tecidos normais em câncer. Esse fator exponencial varia de 1 para cada 2 homens e 1 para cada 3 mulheres. Dentro deste contexto estão associados alguns fatores que contribuem para a carcinogênese, tais como, a ativação de oncogenes e a perda de genes supressores de tumor. Em alguns estudos recentes os autores sugerem que as alterações nos estromas, que fazem parte de um grupo de células compostas por fibroblastos, vasos sanguíneos e células inflamatórias, crescem em conjunto com as células cancerosas e colaboram para uma transformação maligna. A presença dos fibroblastos em alguns tipos de cânceres epiteliais pode promover o crescimento do epitélio e do câncer. Além disso, a rede inflamatória tem sido implicada na formação tumoral proporcionando novas pesquisas sobre a influência dos estromas epiteliais no câncer. Desta forma,

tem-se analisado a influência da senescência na formação da carcinogênese epitelial, com ênfase na mobilização das redes inflamatórias presentes nos estromas e que tem efeito sobre a formação do câncer epitelial (SHAN *et al.*, 2010).

Para criar uma barreira que impede a evolução do câncer, em células pré-cancerosas, ocorre um encurtamento prematuro dos telômeros que acionam a senescência. Assim a senescência torna-se um mecanismo eficiente para suprimir o desenvolvimento do câncer (COLLADO *et al.*, 2007). Quando algumas proteínas não conseguem corrigir determinados erros em seu DNA ocorre a eliminação completa da célula, chamada de apoptose ou suicídio celular programado.

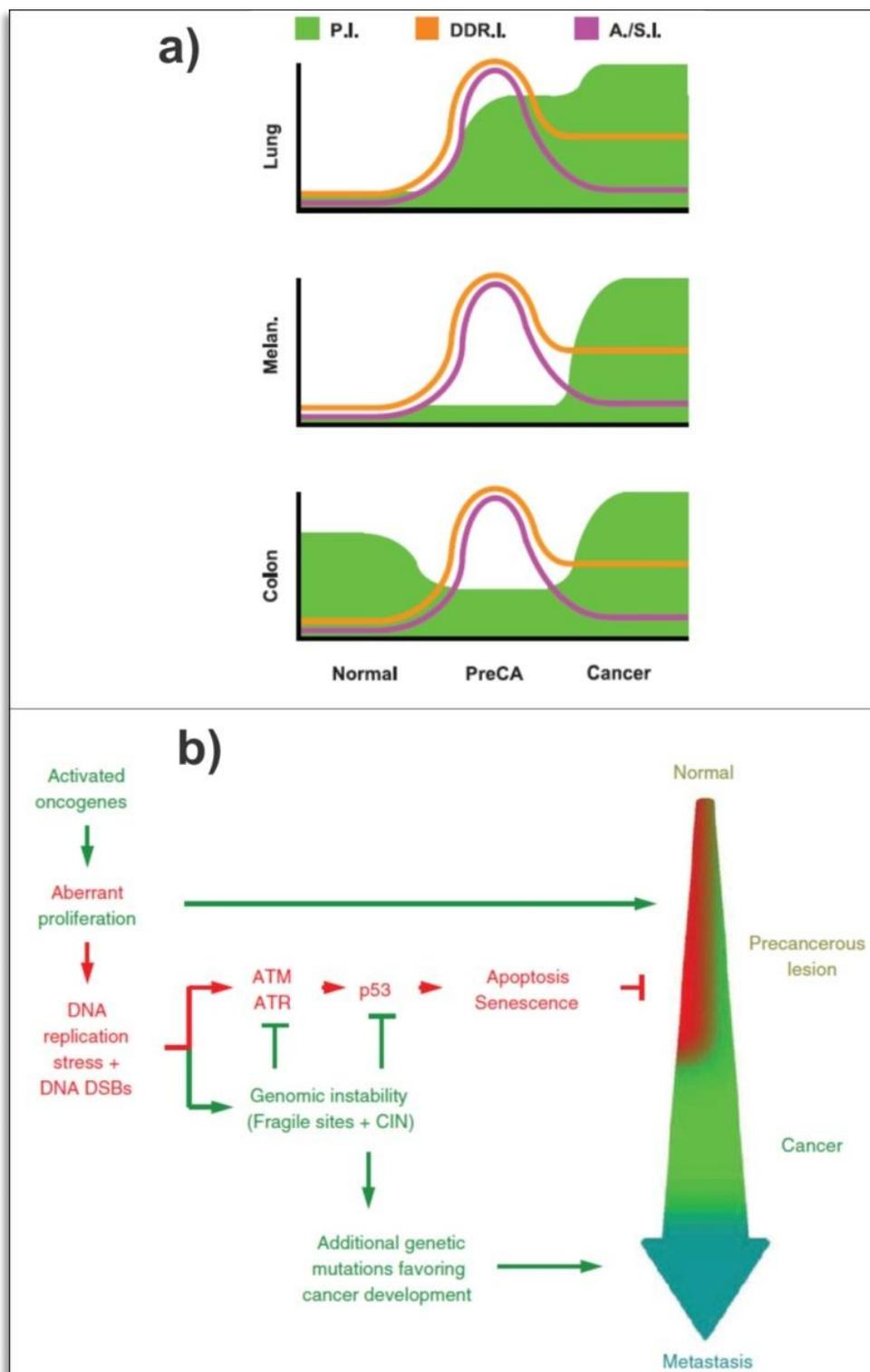
Células que apresentam um DNA muito danificado eliminam a si mesmas para não prejudicar o DNA como um todo. A autodestruição celular causada pela apoptose é um fenômeno bastante rápido que tem início pela retração da célula causando perda de aderência com a matriz extracelular e células vizinhas (GRIVICICH *et al.*, 2007). A morte celular pode ocorrer nas mais diversas situações, como por exemplo, na resposta inflamatória causada pela liberação do conteúdo celular ou por autofagia que ocorre em resposta a um estresse metabólico que resulta na degradação de componentes celulares (RANGANATH *et al.*, 2003).

A ativação da via de Apoptose evita a propagação do fenótipo mutador e a sua inibição provoca a incapacidade de responder adequadamente ao dano do DNA o que leva à instabilidade genética e ao aumento da taxa de desenvolvimento de câncer (ZHIVOTOVSKY *et al.*, 2004). O acúmulo de mutações que ocorrem ao longo da evolução da doença, elimina genes supressores de tumor e ativa outros genes de promoção do crescimento aumentando a instabilidade genômica (MARX, 2002).

Algumas causas da instabilidade genética estão associadas com mutações em vias de Reparo por excisão de nucleotídeos (NER), Reparo por excisão de bases (BER), Reparo de bases mal pareadas (MMR) ou por Recombinação de genes. Em mutações germinativas que estão associadas a síndromes e a hereditariedade podem ser encontrados genes alterados relacionados aos três sistemas de reparo citados acima. Por exemplo: Na *Polipose Adenomatosa* há genes alterados relacionados a MMR e BER e na *Xeroderma Pigmentosa* as alterações são relacionadas a NER (MOMBACH *et al.*, 2008). Desta forma, a instabilidade genômica torna as células pré-cancerosas suscetíveis a novas mutações e ao desenvolvimento de novas aberrações (HOEIJMAKERS *et al.*, 2001, MOMBACH *et al.*, 2008).

No processo de transformação tumoral a ativação de algumas vias forma uma barreira de progressão do câncer (HALAZONETIS *et al.*, 2008). No modelo de Thanos Halazonetis mostrado na figura 5 há evidências de que na evolução tumoral em tecidos de pré-câncer exista uma barreira envolvendo a ativação das vias de apoptose e DDR que resiste à transformação maligna, mas que é rompida mais tarde dando origem a um tecido maligno. Segundo o modelo de Halazonetis os oncogenes induzem um estresse na replicação do DNA levando à formação de DSB's e conseqüentemente a ativação do gene *TP53* que induz a apoptose e a senescência para barrar a evolução tumoral. A origem do estresse replicativo ocorre provavelmente pela desregulação do ciclo celular através de oncogenes que aumentam a atividade das quinases dependentes de ciclinas (*cdk's*) e que atuam nas fases G1 e S. A constante formação de quebras duplas do DNA associada a instabilidade genômica leva ao acúmulo de mutações e a progressão ao câncer. Na figura 5a é mostrada a barreira anticâncer em tecidos pré-cancerosos, onde percebe-se uma diminuição da proliferação celular e aumento da morte celular programada e da resposta ao dano do DNA (DDR) (HALAZONETIS *et al.*, 2008). A diminuição da proliferação em tecidos pré-cancerosos é notória quando há ativação dos pontos de verificação do ciclo celular, que irão ativar mecanismos responsáveis pela reparação do DNA (HALAZONETIS *et al.*, 2008, BARTKOVA *et al.*, 2006).

Na figura 5b é mostrado o modelo de danos ao DNA provocado por oncogenes. O modelo pode ajudar a explicar muitas características do câncer através da ativação de oncogenes que levam os tecidos saudáveis a uma proliferação aberrante. Com a ativação dos promotores tumorais há um estresse replicativo seguido de DSB. Segundo o modelo as principais proteínas que irão ativar a resposta aos danos serão a *atm* e *atr* que irão acionar a *p53*. Em pré-câncer a *p53* ativa a apoptose e a senescência para conter os defeitos inerentes e com isso barrar um possível estresse na replicação que poderá levar a célula a uma proliferação aberrante. Na figura 5b é possível verificar ainda que a instabilidade genômica é seguida pela instabilidade cromossômica (CIN) em câncer provocando uma adição de genes mutantes o que facilita o desenvolvimento do câncer e a formação de metástase (HALAZONETIS *et al.*, 2008).



**Figura 5 -** (a) Barreira de evolução do câncer em lesões pré-cancerosas (preCA). O índice de proliferação (P.I.), o índice de resposta ao dano do DNA (DDR.I.) e o índice de apoptose e senescência (A./S.I.) são mostrados para os tecidos normais, pré-câncer (preCA) e câncer de pulmão, melanoma e colorretal. (b) Modelo de danos do DNA provocado por oncogenes na progressão e no desenvolvimento do câncer. A instabilidade genômica é um resultado direto dos oncogenes que são ativados por estresse na replicação ou por DSB no início do desenvolvimento do câncer, a instabilidade genômica é seguida pela instabilidade cromossômica (CIN) em câncer provocando uma adição de genes mutantes o que facilita o desenvolvimento do câncer (adaptada de HALAZONETIS *et al.*, 2008).

A transformação do fenótipo de uma célula normal em um fenótipo alterado é observada em tumores que surgem de duas formas: a primeira delas é baseada no acúmulo de anomalias genéticas observadas ao longo do período de vida, chamadas de mutações somáticas e que aparecem na maioria dos tumores sólidos, e em segundo lugar aparecem às mutações germinativas que são passadas aos descendentes de forma hereditária (CASTRO *et al.*, 2007).

A probabilidade da ocorrência do fenótipo mutador é maior em tecidos renováveis, como as glândulas, pele, próstata e colorretal. A alta taxa de proliferação celular destes tecidos prejudica a integridade do genoma, onde se observa uma maior probabilidade de aquisição e propagação das mutações deletérias que contribuem para o aumento de transformações malignas ao longo do tempo de latência da doença (LOEB, 2001).

## **2.4 Bancos de dados em biologia sistêmica**

Os bancos de dados em biologia sistêmica são essenciais para o armazenamento e gerenciamento da grande demanda de informações. A maior dificuldade dos cientistas é desenvolver um padrão de armazenamento de dados, o que causa hoje uma enorme dificuldade em integrar diferentes bancos e informações. Essa dificuldade foi solucionada com a criação de alguns bancos de dados usados para padronizar as descobertas provenientes do projeto genoma humano (WATSON,1990). O *Gene Nomenclature Committee* (HUGO) é um banco de dados de genes, utilizado para atribuir um único símbolo e nome a mais de 32 mil genes, dos quais mais de 19 mil codificam proteínas (POVEY *et al.*, 2001). Além do HUGO outros bancos de dados se destacam como repositórios de dados: *Biocarta*, *Reactome*, *Kegg*, *Gene Ontology* (GO) e *National Center for Biotechnology Information* (NCBI) que são grandes repositórios de informações moleculares provenientes de vários organismos. Uma descrição dos seus endereços virtuais é apresentada na tabela 1.

Os genes encontrados no banco de dados *Gene Ontology* apresentam evidências físicas inferidas por curadores que extraem informações relevantes que apóiam a participação dos genes em determinadas funções. Os códigos de

evidências do GO se dividem em: experimental, análise computacional, declaração de autores e declaração de curadores.

Entre as evidências do GO cabe salientar as três mais específicas:

- Inferida a partir de fenótipo mutador (IMP): são evidências que implicam em mutações que resultam em deficiência parcial ou completa que possam perturbar o funcionamento normal de outros genes;

- Inferida a partir de Interação genética (IGI): apresentam evidências como complementação funcional ou até mesmo de genes retirados do fenótipo de uma mutação em um gene diferente;

- Inferida a partir da declaração de autores (NAR): são evidências de genes que foram declarados por autores em resumos, introduções ou discussões onde não é possível rastrear a publicação original.

**Tabela 1** - Lista de bancos de dados de genes e proteínas com os seus respectivos sítios.

<b>Nome</b>	<b>Sítio</b>
<i>Biocarta</i>	<a href="http://www.biocarta.com/">http://www.biocarta.com/</a>
<i>Reactome</i>	<a href="http://www.reactome.org">http://www.reactome.org</a>
<i>Kegg</i>	<a href="http://www.genome.jp/kegg/">http://www.genome.jp/kegg/</a>
<i>Gene Ontology (GO)</i>	<a href="http://www.geneontology.org/">http://www.geneontology.org/</a>
<i>National Cancer Institute</i>	<a href="http://www.pid.nci.nih.gov/">http://www.pid.nci.nih.gov/</a>
<i>National Center for Biotechnology Information (NCBI)</i>	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>

Existem também bancos de dados que são usados para armazenar uma grande quantidade de estudos gerados a partir de microarranjos. O banco de dados *Gene Expression Omnibus* (GEO) é um repositório público funcional de dados usado para armazenar transcriptomas (EDGAR *et al.*, 2002). O banco de dados é composto por 777.727 amostras e 9.993 plataformas divididas em mais de 29 mil estudos (microarranjos) diferentes (GEO, 2012). O GEO é considerado o maior banco de dados de microarranjos e nele estão depositados transcriptomas fabricados por diversas empresas em destaque a *Affymetrix Gene Chip*. A *Affymetrix* é uma empresa fundada em 1992 por Stephen Fodor e tem como objetivo fabricar microarranjos de DNA (GOHLMANN *et al.*, 2009).

### 2.4.1 Ontologia Ontocancro 1.0 e 2.0

A falta de um vocabulário unificado para genes, subvias e vias de manutenção do genoma e estabilidade genômica fez com que o grupo<sup>2</sup> desenvolvesse a ontologia *Ontocancro* (LIBRELOTTO *et al.*, 2009). A ontologia é um modelo de dados que representa um conjunto de conceitos dentro de um determinado domínio e os relaciona entre si. O banco de dados é baseado em uma ontologia que busca organizar e integrar as informações de interatomas e transcriptomas disponíveis a partir dos bancos de dados da tabela 1 (LIBRELOTTO *et al.*, 2009, SIMÃO *et al.*, 2010). A ontologia é formada por vias chamadas de vias *Ontocancro*, disponíveis na versão 1.0 da ontologia *Ontocancro*, e por subvias que apresentam funções GMM mais específicas que as vias. As vias *Ontocancro*, são formadas por genes citados na literatura e bases de dados com evidências de envolvimento em GMM e que não estão incluídas na definição de vias específicas em outros bancos de dados. As vias *Ontocancro* abrangem as seguintes vias (LIBRELOTTO *et al.*, 2009):

- Apoptose (APO)
- Reparo por excisão de nucleotídeos (NER);
- Reparo por excisão de bases (BER);
- Reparo por recombinação homóloga (HR);
- Reparo de bases mal pareadas (MMR);
- Reparo por terminação não homóloga (NHJ)
- Ciclo celular (CC)
- Estabilidade cromossômica (CS)

As subvias disponíveis na versão 2.0 da ontologia foram extraídas do banco de dados *National Cancer Institute*, *Reactome* e *Biocarta*, além de algumas outras subvias que foram incluídas a partir da literatura. Todas as vias implementadas na ontologia *Ontocancro* são curadas e suas informações mais relevantes são extraídas de diferentes repositórios. Entre estas informações buscadas nos bancos destaca-se a descrição do gene extraída do Banco de dados *String*, o símbolo aprovado pelo *HUGO* e o *Entrez Gene*, que é um identificador numérico do gene.

---

<sup>2</sup> Projeto desenvolvido pelo grupo de pesquisa, referenciado por: Giovani H. Librelotto, José C. M. Mombach, Marialva Sinigaglia, Éder M. Simão, Heleno B. Cabral, Mauro A. A. Castro, Karlise Soares e Cristhian Augusto Bugs.

A *Ontocancro 2.0*<sup>2</sup>, é um banco de dados em desenvolvimento que além de integrar informações de vias e subvias de manutenção do genoma envolvidas no ciclo celular, resposta ao dano do DNA, apoptose e senescência integra também estudos de microarranjos envolvidos em câncer extraídos do banco de dados *Gene Expression Omnibus* (GEO). Em sua última atualização o banco disponibiliza quatro estudos:

1- Série GSE4183 de colorretal (côlon): com 8 estudos de tecido normal, 15 estudos de processo inflamatório, 15 estudos de adenoma e 15 estudos de carcinomas (GSE4183, 2007);

2- Série GSE10927 do córtex adrenal: com 10 estudos de tecidos normais, 22 estudos de adenomas e 33 estudos de carcinomas (GSE10927, 2009);

3- Série GSE19650 de pâncreas: com 7 estudos de tecidos normais dos dutos pancreáticos, 6 estudos de adenoma mucinoso papilar e 6 estudos de carcinoma mucinoso papilar do pâncreas (GSE19650, 2010);

4- Série GSE27155 de tireóide: com 4 estudos de tecidos normais da tireóide, 17 estudos de adenoma foliculares da tireóide e 13 estudos de carcinoma folicular da tireóide (GSE27155, 2011).

Os quatro estudos encontrados na *Ontocancro 2.0* foram normalizados com MAS 5.0 com chamadas de P/M/A (presente, marginal e ausente). Mais informações sobre normalizações estão disponíveis na seção 2.5.1 e informações adicionais sobre a ontologia *Ontocancro 1.0* e *2.0* podem ser encontrados em ambiente virtual<sup>3</sup>.

## 2.5 Microarranjos

A expressão gênica corresponde a vários eventos e inicia pela transcrição do gene no núcleo até a tradução do mRNA no ribossomo. Assim a expressão gênica corresponde ao processo em que a informação codificada por um determinado gene é compilada em uma molécula. Com algumas exceções as células que constituem o organismo humano possuem a mesma carga genética, ou seja, o mesmo DNA. O que diferencia dois grupos celulares morfológicamente distintos são os genes

---

<sup>3</sup> Disponível em <[ontocancro.inf.ufsm.br](http://ontocancro.inf.ufsm.br)>

expressos e os níveis de expressão desses genes (WATSON *et al.*, 2004). Para medir os níveis de expressão dos genes são usadas várias técnicas experimentais e, uma delas, os microarranjos se destacam na biologia sistêmica desde a sua descoberta em 1982 (AUGENLICHT *et al.*, 1982).

Os microarranjos são usados para medir grandes quantidades simultâneas de níveis de expressão de transcritos e consistem em conjuntos ordenados de milhares de moléculas de DNA organizadas em sondas. A *Affymetrix GeneChip* produz sondas que são fixadas quimicamente a uma superfície sólida (*chip*) com aproximadamente 1,3 cm x 1,3 cm de área, composta por vidro, nylon ou *chip* de silício (figura 6). Em cada uma das sondas estão imobilizados oligonucleotídeos. Os oligonucleotídeos são fragmentos curtos com aproximadamente 25 bases de uma cadeia simples de DNA ou RNA. Nos microarranjos fabricados pela *Affymetrix* são encontradas sondas que identificam a expressão de genes e hibridizações cruzadas, (tabela 2).



**Figura 6** - Imagem de um microarranjo da *Affymetrix* com 1,3 cm x 1,3 cm de superfície (GOHLMANN *et al.*, 2009).

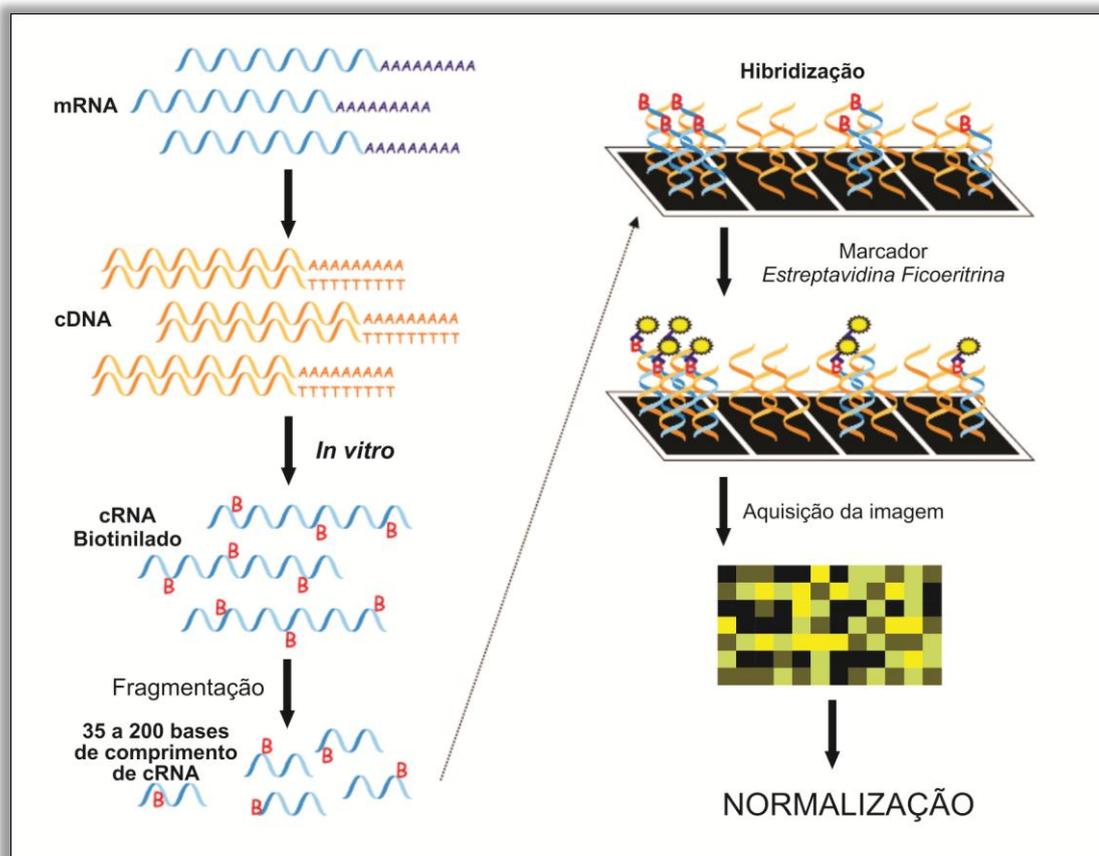
Cada conjunto de aproximadamente 8 sondas impressas em um *chip* é chamado de *probset*. O conjunto de *probsets* irá definir a expressão de um determinado gene. Em geral as *probsets* variam entre 20 a 55 mil dependendo do tipo de análise desejada (GOHLMANN *et al.*, 2009). Para verificar o nível de expressão de uma amostra celular são extraídas amostras do mRNA (*in vivo*) e convertidas em DNA complementar (cDNA) (por transcriptase reversa) em dupla fita para que ocorra uma amplificação linear da amostra. Esse processo inicial é

importante também para garantir a estabilidade e criar um modelo de DNA que sirva para a síntese de RNA complementar (cRNA) biotilado (*in vitro*) em fita simples. No processo seguinte o cRNA é purificado com partículas magnéticas e quantificado espectrofotometricamente. O cRNA purificado é fragmentado, a fim de facilitar a hibridização com os oligonucleotídeos presos nas sondas do *chip* (BECKER *et al.*, 2003).

**Tabela 2** - Interpretação das sondas em um *chip Affymetrix*. A primeira coluna indica a terminação da sonda. A segunda coluna traz informações do conjunto de sondas. Na terceira coluna são destacadas as dificuldades de interpretar o sinal obtido pelos conjuntos de sondas. A última coluna refere-se ao tipo de arranjo *Affymetrix* em que eles são usados (GOHLMANN *et al.*, 2009).

Sonda	Descrição	Questões	Affymetrix
_at	Conjunto de sondas únicas. Contém sondas que medem sequências específicas de uma transcrição única		HG-U133
_a_at	Transcrições alternativas. Reconhecem transcritos cruzados para uma determinada família de genes	O sinal da expressão não pode ser atribuído a uma transcrição simples, portanto é difícil de interpretar	"Plus Array"
_s_at	Conjunto de sondas com hibridização cruzadas para um determinado conjunto de sequências	O sinal da expressão não pode ser atribuído a uma transcrição simples, portanto é difícil de interpretar	HG-U133
_x_at	Pelo menos algumas sondas hibridizam cruzadas com sequências de outros alvos para o <i>chip</i>	Além da transcrição alvo, há hibridizações imprevisíveis, resultando em um sinal que pode ser difícil de interpretar	HG-U133

Com a ajuda de uma solução de hibridização e de um marcador fluorescente (*Streptavidina Ficoeritrina*) o cRNA se ligará ao microarranjo. Após algum tempo a solução excedente será removida por um processo de lavagem que eliminará os cRNA que não se hibridizaram ao microarranjo. Cada cRNA ligado a um oligonucleotídeo é excitado usando um scanner a laser e as posições e intensidades das emissões fluorescentes são capturadas e convertidas em dados que são transferidos para um sistema de arquivos (BECKER *et al.*, 2003). Uma síntese do processo de formação de imagens em microarranjos está descrita na figura 7.



**Figura 7** - Visão geral simplificada do protocolo *Affymetrix* de análise do nível de expressão. A partir do mRNA é gerado um cDNA usado na transcrição *in vitro*, resultando em uma fita simples de cRNA biotilado que será fragmentado e hibridizado com os oligonucleotídeos das sondas. As intensidades fluorescentes são adquiridas com um scanner a laser. Imagem adaptada de BECKER *et al.*, 2003.

Para cada mRNA analisado há um conjunto de aproximadamente 8 sondas (*probset*) diferentes que hibridizam os fragmentos do cRNA em diferentes posições ao longo da sequência de mRNA. Os microarranjos produzidos pela *Affymetrix* contêm milhares de dezenas de conjuntos de sondas em seus *chips*. As sondas são selecionadas de tal forma que a sequência dos fragmentos representa uma parte de uma sequência dos genes que serão detectadas. Estas sondas são chamadas de "par perfeito" (PM - *Perfect Match*). Uma sonda PM deveria ser única para um determinado mRNA eliminando os falsos sinais de transcrições que poderiam se ligar a sondas devido à similaridade da sequência de bases (BECKER *et al.*, 2003). Os microarranjos da *Affymetrix* usam outro tipo de sonda chamada de "par incompatível" (MM - *Mismatch*). As sondas MM também possuem 25 bases de oligonucleotídeos com a mesma sequência da sonda PM, mas eles contêm uma única base modificada na posição 13. Estas sondas têm a intenção de detectar a

quantidade de sinal inespecífico e também os fundos de escala detectados pelas sondas PM. Nos últimos anos têm sido demonstrado que as sondas MM não apenas detectam sinais inespecíficos, mas também detectam mesmas transcrições que os seus homólogos em PM (GOHLMANN *et al.*, 2009).

### 2.5.1 Normalizações de microarranjos

Devido as impurezas geradas durante o processo de hibridização e detecção dos níveis de expressão das amostras de mRNA, os dados gerados a partir do protocolo *Affymetrix* devem passar por um processo de normalização. Os dados armazenados em arquivos *.CEL* apresentam todas as informações dos valores de expressão e localização para todas as sondas na imagem produzida pela hibridização do microarranjo (GOHLMANN *et al.*, 2009).

Existem inúmeros métodos de normalização e correção do ruído de fundo para os microarranjos, porém alguns são os mais importantes:

- Microarranjo suíte versão 5 (MAS 5.0 - *Microarray Suite Version 5.0*),
- Média robusta de múltiplos chips (RMA - *Robust Multichip Average*) (IRIZARRY *et al.*, 2003),
- RMA para *Affymetrix GeneChip* (GCRMA - *GeneChip Affymetrix RMA*) (GOHLMANN *et al.*, 2009, WU *et al.*, 2004).
- MAS 5.0 com chamadas P/M/A (Presente/Marginal/Ausente)

A aplicação dos métodos de normalização, MAS 5.0, RMA, GCRMA e MAS 5.0 com chamadas P/M/A em arquivos no formato *.CEL* podem ser aplicadas pelos pacotes de dados de normalização disponíveis pela base de dados do *Bioconductor*. A base de dados é o resultado de um projeto iniciado em 2001, que tem por objetivo proporcionar o acesso a análises estatísticas e gráficas de dados genômicos e, ainda, facilitar a integração de bancos de dados na análise de dados experimentais. Os pacotes do *Bioconductor* são aplicados pela linguagem estatística de programação R e apresenta 460 pacotes de dados desenvolvidos para a análise de microarranjos. Cada pacote é desenvolvido em estruturas de códigos que desempenham funções diferenciadas para a interface do *Bioconductor*. Para efetuar os cálculos de normalização, MAS 5.0, RMA, GCRMA e MAS 5.0 com chamadas

P/M/A o software utiliza pacotes específicos que poderão ser adquiridos em ambiente virtual<sup>4</sup> (GENTLEMAN *et al.*, 2004).

### 2.5.2 Análises estatísticas de expressões de vias genéticas

A análise de vias e subvias em microarranjos é alimentada por métodos que envolvem raciocínios e suposições diferentes. Devido à grande dimensão dos dados de microarranjos e também pela grande variabilidade na expressão de genes isolados, podem ser usadas metodologias de análise de microarranjos destinadas especificamente a estudos isolados. Desta forma, dependendo da aplicação existem aproximações estatísticas que variam de acordo com o tipo de microarranjos. Por exemplo, na análise por representação excedente há uma procura por vias enriquecidas com genes significativos. A partir de uma lista de genes é aplicado um teste hipergeométrico que avaliará a super-expressão do conjunto. O teste hipergeométrico que é descrito através de uma distribuição de probabilidades discreta é usado para calcular a probabilidade de conter genes diferencialmente expressos do que seria esperado por acaso (GOHLMANN *et al.*, 2009). Outro método que se destaca é a contagem funcional de classes que é usado em microarranjos que apresentam baixos valores  $p$  ( $p$ -value) em suas amostras. O valor de  $p$  é a probabilidade de obter-se uma estatística de teste no mínimo tão extrema como o que foi realmente observado, muitas vezes o valor de  $p$  é menor que o nível de significância que geralmente varia entre 0,05 e 0,1. Esse método é usado para calcular uma pontuação que define as alterações significativas de todos os genes incluídos na análise. Um terceiro método analisa os conjuntos de genes diretamente ligados a conjuntos pré-definidos com o objetivo de identificar as mudanças de expressão entre essas vias (GOHLMANN *et al.*, 2009).

Castro e colaboradores (2007) introduziram outro método de análise direto de expressão de vias e subvias. Os cálculos da atividade relativa e da diversidade relativa são usados para avaliar os níveis de expressão de conjuntos de genes e definir os padrões de expressão (reprogramação) entre os genes das vias (CASTRO *et al.*, 2007).

---

<sup>4</sup> Disponível em: <[www.bioconductor.org](http://www.bioconductor.org)>

Para calcular a atividade relativa de uma dada via  $\alpha$  com um número de genes  $M_\alpha$ , deve-se somar a expressão dos genes em dois grupos de vias: o primeiro grupo representa as amostras de tecidos alterados ou experimentais  $N_\alpha^e$  e o segundo grupo é composto pelas amostras de tecidos normais ou controle  $N_\alpha^y$ . Então, a atividade relativa  $n_\alpha$  da via  $\alpha$  será dada por:

$$n_\alpha = \frac{N_\alpha^e}{N_\alpha^e + N_\alpha^y} . \quad (1)$$

O valor de  $n_\alpha$  varia entre ,  $0 \leq n_\alpha \leq 1$ , se  $n_\alpha < 0,5$  implica que  $N_\alpha^e < N_\alpha^y$ , isto é, a atividade da via com amostra alterada é menor que a atividade do controle, enquanto que  $n_\alpha > 0,5$  representa o caso inverso (CASTRO *et al.*, 2007).

Para caracterizar de forma quantitativa a diversidade para uma via  $\alpha$  é utilizada a entropia de Shannon (SHANNON *et al.*, 1951) que é descrita como:

$$H_\alpha = -\frac{1}{\ln(M_\alpha)} \sum_i^{M_\alpha} p(i, \alpha) \ln p(i, \alpha) , \quad (2)$$

onde  $M_\alpha$  é o número de genes na via e  $p(i, \alpha)$  é a frequência da diversidade do gene  $i$ , dada por:

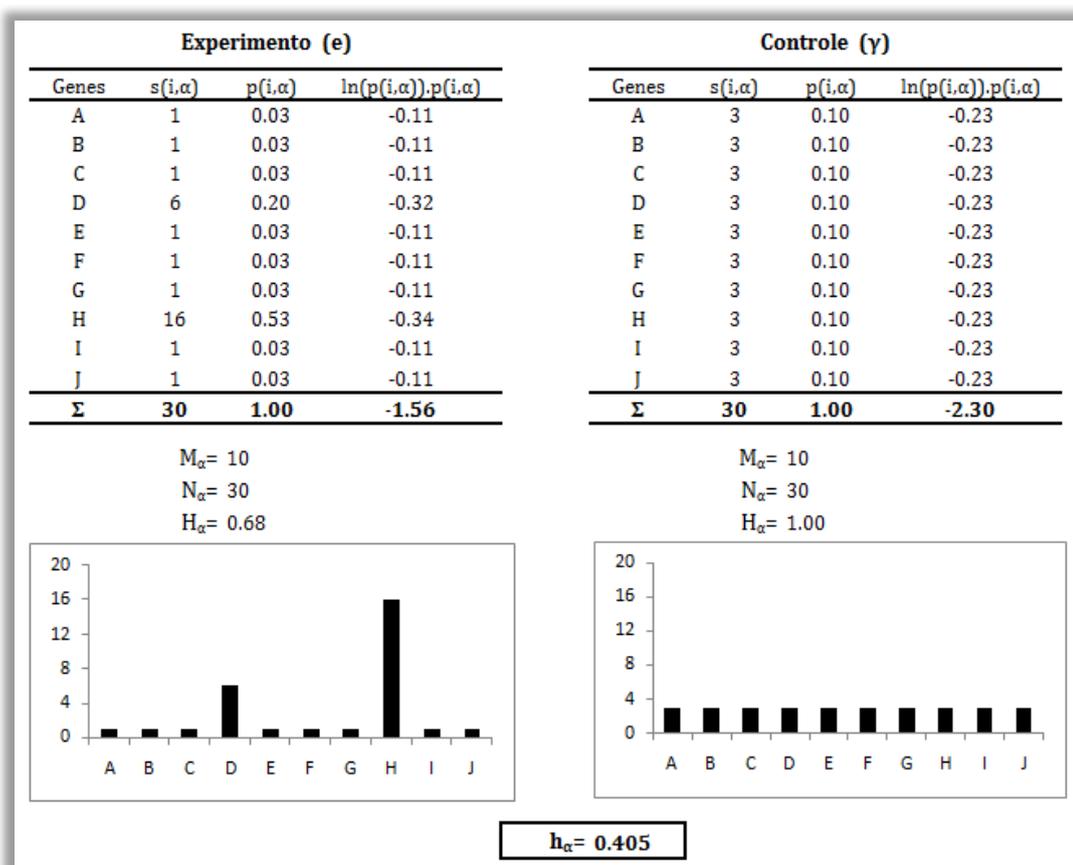
$$p(i, \alpha) = \frac{s(i, \alpha)}{N_\alpha} \quad (3)$$

com  $s(i, \alpha)$  sendo a atividade do gene ( $i$ ) e  $N_\alpha$  a soma da expressão dos genes na via ( $\alpha$ ). O termo  $\ln(M_\alpha)$  é um fator de normalização que garante  $0 \leq H_\alpha \leq 1$ , desta forma, pode-se comparar as vias com diferentes quantidades de genes. Tendo como referência o sinal controle da amostra, pode-se definir a diversidade relativa ( $h_\alpha$ ) como:

$$h_\alpha = \frac{H_\alpha^e}{H_\alpha^e + H_\alpha^y}, \quad (4)$$

onde  $H_\alpha^e$  e  $H_\alpha^y$  são as diversidades das amostras com alteração (experimento) e o controle, respectivamente. O valor de  $h_\alpha$  varia entre  $0 \leq h_\alpha \leq 1$ , se  $h_\alpha < 0,5$  implica que  $H_\alpha^e < H_\alpha^y$ , isto é, a diversidade dos valores de expressão dos genes na via é

menor para a amostra alterada do que para o controle, enquanto que  $h_\alpha > 0,5$  representa o caso inverso (CASTRO *et al.*, 2007). A figura 8 ilustra a diversidade relativa para um conjunto de 10 genes alterados com seus respectivos genes de controle pertencentes a uma via qualquer. Na figura 8 os genes das amostras de controle apresentam níveis de expressão com valores mais próximos (constantes ou sem dispersão) do que os níveis de expressão dos genes do experimento, isso indica que  $H_\alpha^e < H_\alpha^y$  e  $h_\alpha < 0,5$ . Ou seja, há uma diminuição da diversidade nos níveis de expressão dos genes de experimento em relação ao controle. Para  $H_\alpha^e > H_\alpha^y$  tem-se o caso contrário e haverá aumento da diversidade relativa  $h_\alpha > 0,5$  do experimento *versus* o controle. Na reprogramação da expressão dos genes poderão ocorrer níveis de aumento ou diminuição da diversidade relativa entre os genes de uma via.



**Figura 8** - Exemplo mostrando os níveis de expressão de 10 genes quaisquer, com diminuição da diversidade do experimento em relação ao controle ( $H_\alpha^e < H_\alpha^y$ ). Para este caso a diversidade dos genes de controle é maior do que o experimento, desta forma a diversidade relativa é menor que 0,5 deste modo  $h_\alpha = 0,405$ .

Para determinar se uma alteração em uma via (atividade) ou em um conjunto de genes (diversidade) é estatisticamente significativo em um determinado estudo

(microarranjo), aplica-se o método de *bootstrap*. Esse método é usado para calcular a distribuição amostral de  $h_\alpha$  e  $n_\alpha$  através de uma análise de reamostragem aleatória que cobre todos os genes do estudo com repetições que variam de 100 a 100.000 com as mesmas quantidades de genes das vias de interesse para investigar a convergência do valor de  $p$  (*p-value*) da amostra sobre uma distribuição de probabilidades normal.

O nível de significância em uma distribuição de probabilidades é comparado com o valor de  $p$  e serve para delimitar se os resultados da atividade ou diversidade relativa de um determinado conjunto de genes são ou não significativos a um determinado nível. Desta forma, se o nível de significância for fixado em 0,05 os valores de  $p$  menores que 0,05 ( $p < 0,05$ ) apresentaram aumento ou diminuição significativa de expressão da via em relação ao estudo. Dependendo em qual dos lados da distribuição de probabilidade unicaudal cair o valor de  $p$  está poderá ter aumento ou diminuição de expressão.

As análises mostradas acima podem ser calculadas usando o *software ViaComplex*. O *software* é um aplicativo usado para construir mapas funcionais de expressão gênica e utiliza a entropia de Shannon para obter um parâmetro quantitativo usado para caracterizar a atividade e a diversidade relativa de vias (CASTRO *et al.*, 2009). O *software* é utilizado para calcular significâncias de vias pelo método de *bootstrap* e pela correção de falsos positivos. Os falsos positivos são usados como método de controle estatístico quando se faz os testes de múltiplas hipóteses (*bootstrap*) para efetuar uma correção nas múltiplas comparações (entre as vias). O código fonte do *software ViaComplex* contendo todas as equações apresentadas acima está disponível em ambiente virtual<sup>5</sup>.

### 2.5.3 Análise de genes diferencialmente expressos - *fold change*

O objetivo deste tipo de análise é identificar os genes com diferenças extremas de expressão entre as amostras de microarranjos. Em termos estatísticos formais, um gene é expresso diferencialmente se o seu nível de expressão é alterado sistematicamente entre duas condições de tratamento (exemplo: amostras

---

<sup>5</sup> Disponível em: <[www.lief.if.ufrgs.br/pub/biossoftwares/viacomplex](http://www.lief.if.ufrgs.br/pub/biossoftwares/viacomplex)>

de experimento *versus* controle) (McCARTHY *et al.*, 2009). A utilização destes procedimentos de análise diferencial garante que as conclusões sobre os genes identificados sejam válidas para uma possível população amostral de interesse, e não somente para a amostra de mRNA em questão (HUBER *et al.*, 2003).

Em estudos com mais de um tipo de tecido (ex: pré-câncer e tecido normal, no GEO representado por uma série GSE) a análise de genes diferencialmente expressos busca identificar quais genes são expressos de forma distinta entre os tecidos. Existem várias abordagens estatísticas envolvendo a análise de genes diferencialmente expressos (HUBER *et al.*, 2003). O *fold change*, mudança de expressão ou número de vezes que a expressão varia, é usada em conjuntos de microarranjos cujo número de amostras  $h$  é grande e as estimativas da variância por gene possuem muitos graus de liberdade e são muito estáveis. Em geral o *fold change* é aplicado em conjuntos de amostras com mais de cinco repetições e para conjuntos de genes muito pequenos, no caso das vias e subvias (RUIZ *et al.*, 2004).

Considerando um conjunto de amostras  $H$  (amostras de uma série de microarranjos, ex: Série GSE10927, 2009) de um estudo de microarranjos, com  $K$  genes referente ao conjunto onde deseja-se calcular o *fold change* para três tecidos: normal  $n_1$ , adenoma  $n_2$  e câncer  $n_3$ , pode-se construir uma matriz de expressão  $X$ , sendo  $x_{ijh}$  o valor de expressão normalizado, para o  $i$ -ésimo gene no  $j$ -ésimo conjunto de tecidos ( $n_1, n_2, n_3$ ) e na  $h$ -ésima amostra:

$$X(n_1, n_2, n_3) = \begin{bmatrix} \text{Normal} & \text{Adenoma} & \text{Câncer} \\ x_{1n_11} \dots x_{1n_1H} & x_{1n_21} \dots x_{1n_2H} & x_{1n_31} \dots x_{1n_3H} \\ x_{2n_11} \dots x_{2n_1H} & x_{2n_21} \dots x_{2n_2H} & x_{2n_31} \dots x_{2n_3H} \\ \vdots & \vdots & \vdots \\ x_{Kn_11} \dots x_{Kn_1H} & x_{Kn_21} \dots x_{Kn_2H} & x_{Kn_31} \dots x_{Kn_3H} \end{bmatrix} \quad (5)$$

onde cada linha da matriz corresponde a um gene, e cada coluna a uma amostra dentro de um determinado tecido ( $n_1, n_2, n_3$ ).

Para calcular o *fold change* deve-se agrupar as amostras de um tecido através da média aritmética:

$$\bar{x}_{i(n_1, n_2, n_3)} = \frac{1}{H} \sum_{h=1}^H x_{ih} \quad (6)$$

onde  $H$  é o número total de amostras e  $x_{ih}$  é a expressão normalizada do gene  $i$  na amostra  $h$ . Como exemplo tem-se o seguinte resultado:

$$X'(n_1, n_2, n_3) = \begin{bmatrix} \textit{Normal} & \textit{Adenoma} & \textit{Câncer} \\ \bar{x}_{1n_1} & \bar{x}_{1n_2} & \bar{x}_{1n_3} \\ \bar{x}_{2n_1} & \bar{x}_{2n_2} & \bar{x}_{2n_3} \\ \vdots & \vdots & \vdots \\ \bar{x}_{Kn_1} & \bar{x}_{Kn_2} & \bar{x}_{Kn_3} \end{bmatrix} \quad (7)$$

De acordo com as abordagens feitas por Ruiz e colaboradores em 2004, pode-se calcular o *fold change* para uma determinada via extraída da matriz (7) (RUIZ *et al.*, 2004). Desta forma o *fold change* de cada gene da via será a razão entre os valores de expressão experimentais pelo seu controle (McCARTHY *et al.*, 2009):

$$FC_i = \frac{x_{i(n_2, n_3)}}{x_{i(n_1)}} \quad (8)$$

Em outras palavras o *fold change* será a razão entre a expressão de um determinado gene de adenoma *versus* normal ou câncer *versus* normal. Considerando uma via com  $M_\alpha$  genes de microarranjos, pode-se calcular o *fold change* para cada gene da via e classificar os genes  $i$  mais alterados em conjuntos chamados de *TOP-N*, onde  $N$  será o número de genes (arbitrário) mais alterados que irão representar os genes de uma via (GOHLMANN *et al.*, 2009). A análise do *fold change* de genes mais alterados em vias ou em subvias, pode ser calculada pela linguagem estatística de programação R, através dos pacotes do *Bioconductor*: *limma*, *biobase*, *annaffy* e *hgplus2.db* disponíveis em ambiente virtual<sup>6</sup>.

---

<sup>6</sup> Disponível em: < <http://www.bioconductor.org/> >

### 3. MODELAGEM DA REDE DE MANUTENÇÃO DO GENOMA HUMANO

Neste capítulo descrevem-se os resultados de três abordagens publicadas em 2010 por SIMÃO e colaboradores (SIMÃO *et al.*, 2010):

- Rede de interação de proteínas envolvidas em GMM,
- Análise das vias de manutenção do genoma em estudos de adenoma do colorretal,
- Resposta temporal das células da próstata a baixas doses de Cádmio.

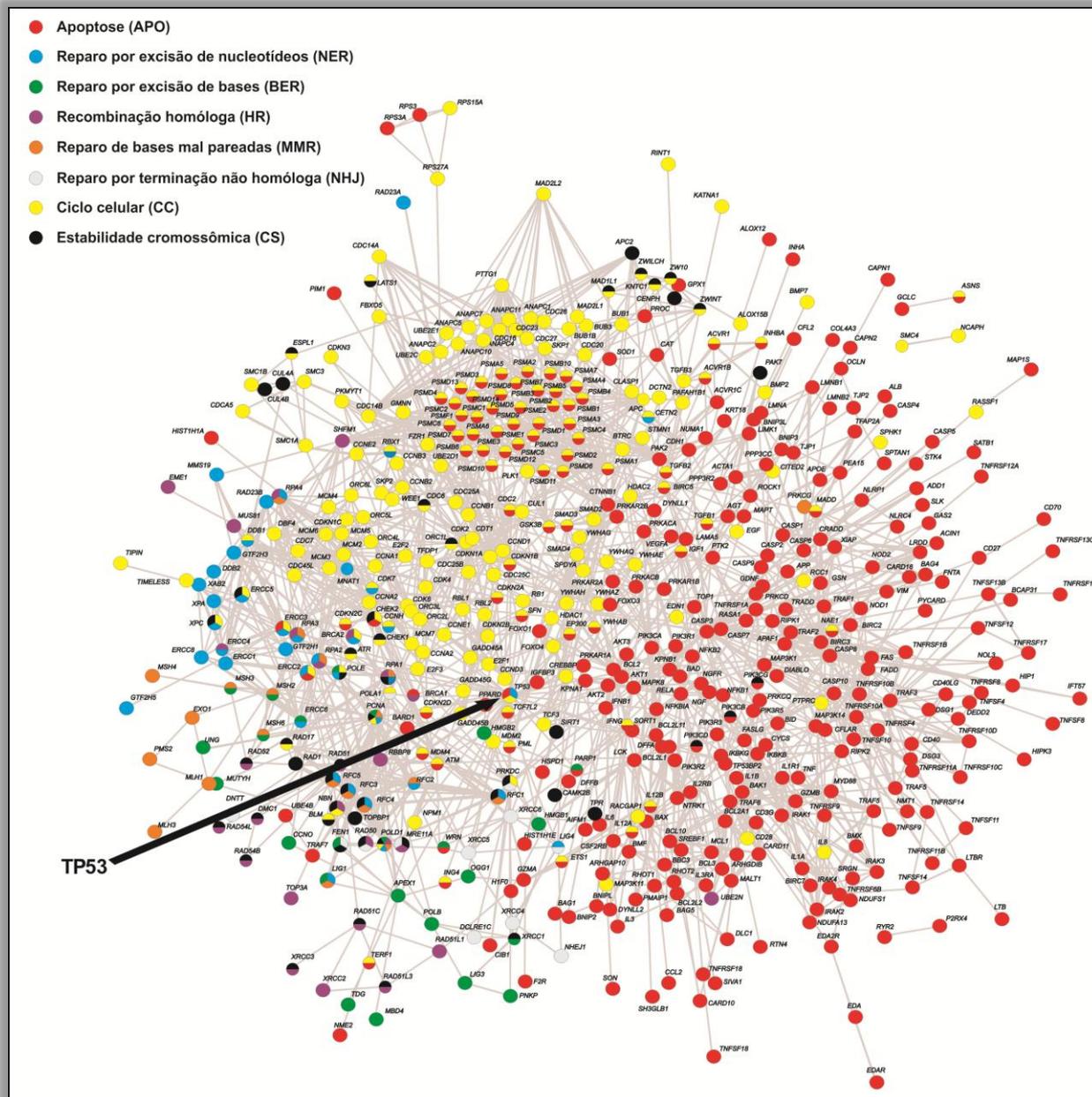
#### 3.1 Rede de interação de proteínas envolvidas em GMM

Em um trabalho desenvolvido pelo grupo de pesquisa em 2009, foi elaborado um banco de dados para integrar informações relevantes de genes e vias envolvidas nos mecanismos de manutenção do genoma (GMM) (LIBRELOTTO, *et al.*, 2009). Na ontologia *Ontocancro 1.0* estão apresentados 834 genes pertencentes às vias Ontocancro com exceção da via de Apoptose expandida.

Com isso foi construída uma rede de proteínas com informações relacionadas as vias de Apoptose (APO), Reparo por excisão de nucleotídeos (NER), Reparo por excisão de bases (BER), Reparo por recombinação homóloga (HR), Reparo de bases mal pareadas (MMR), Reparo por terminação não homóloga (NHJ) e Ciclo celular (CC) além de um grupo de genes relacionados a estabilidade cromossômica (CS) (VOGELSTEIN *et al.*, 2004) (ver seção 2.3). O objetivo da rede foi buscar por interações físicas entre as proteínas pertencentes às vias GMM.

Para cada uma das vias Ontocancro foi extraído o conjunto de identificadores ENSP das proteínas. Esse identificador permitiu localizar as interações físicas no banco de dados *String*. Os identificadores dos 834 genes foram introduzidos no banco de dados *String* que foi usado para construir a rede de proteínas com intervalo de confiança de 0,7 e com três tipos de interações: experimentos, coexpressão e bancos de dados.

O maior aglomerado de interações da rede de proteínas envolveu 564 genes e 8240 interações. A figura 9 mostra todos os nós e as interações entre as proteínas na rede.



**Figura 9** - Grafo de interações físicas entre as 564 proteínas pertencentes as Vias Oncocâncer geradas com o banco de dados *String* com opções de interação: experimentos, bancos de dados e coexpressão e com intervalo de confiança de 0,7 (SIMÃO *et al.*, 2010).

Nesta figura observa-se ainda que há uma sobreposição funcional de alguns genes, que apresentam funções em mais de uma via, como por exemplo, o *TP53* que atua em três vias: Apoptose, Reparo por excisão de nucleotídeos e Ciclo celular. Também é importante salientar que a rede mostra uma clara segregação espacial de

nós pertencentes a três vias: Ciclo celular (nós amarelos), Apoptose (nós vermelhos) e Reparo do DNA (outras cores).

A direita da rede (nós vermelho) estão representados os 331 genes envolvidos na via de Apoptose que estão conectados a outros dois grandes módulos de proteínas envolvendo as vias de Reparo (NER, BER, HR, MMR e NHJ, nós coloridos) com 124 genes e as via de Ciclo celular (nós amarelos) com 223 genes. Estes três grandes módulos irão representar três vias importantes para a manutenção do genoma, que são: Apoptose, Reparo e Ciclo celular. Outra observação importante está relacionada aos genes que representam a estabilidade cromossômica (48 nós pretos) que apresentam proteínas espalhadas por toda a rede. A distribuição destes genes de estabilidade cromossômica ocorre devido a funcionalidade de cada proteína, algumas proteínas além de proporcionar estabilidade cromossômica também exercem funções que irão acionar vias de Apoptose, Reparo e Ciclo celular (VOGELSTEIN *et al.*, 2004).

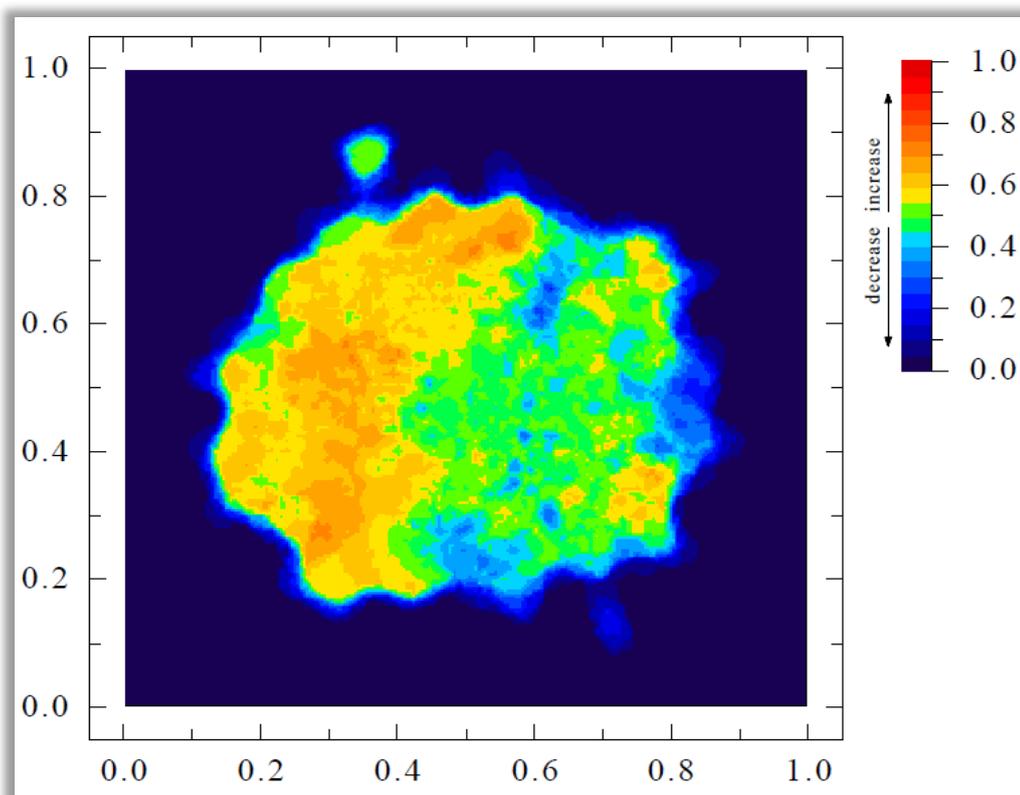
### **3.2 Análises das vias de manutenção do genoma em estudos de adenoma do colorretal**

Para avaliar a sensibilidade da rede de proteínas e detectar as alterações funcionais dos novos genes vinculados as vias GMM armazenadas na ontologia *Ontocancro*, foi selecionado um estudo de microarranjo envolvendo 32 amostras de adenomas de colorretal e 32 amostras de colorretal normal publicados no banco de dados *Gene Expression Omnibus* (GEO), com série GSE8671 (GSE8671, 2010).

As análises foram realizadas em um conjunto de adenomas que foi usado para evidenciar a presença das alterações de vias GMM e relacionar as alterações dos adenomas à evolução do câncer. Todas as amostras foram produzidas com plataformas GPL 570 da *Affymetrix*, normalizadas usando o método microarranjo suíte versão 5 (MAS 5.0) e abrangem 32 biópsias de adenomas e 32 biópsias de mucosa normal do colorretal (SABATES-BELLVER *et al.*, 2007). As 32 amostras de adenoma foram agrupadas por média aritmética (equações 5, 6 e 7 da seção 2.5.3) e as 32 amostras da mucosa normal sofreram o mesmo procedimento. As matrizes de dados de adenoma *versus* mucosa normal do colorretal foram inseridos no

software *ViaComplex* onde foi calculada a funcionalidade dos dados de microarranjos sobre a rede de proteínas da figura 9<sup>7</sup>.

A figura 10 mostra o resultado da análise funcional da rede de proteínas do seção anterior. Nessa figura é observado um aumento de expressão do conjunto de vias de reparo (localizada na parte inferior esquerda da figura, ou observar os nós de reparo localizados na figura 9). Trabalhos anteriores também mostraram que as alterações encontradas em vias de reparo precedem uma transformação maligna de tecidos de adenoma e sugerem uma evolução para o câncer (CASTRO *et al.*, 2007).



**Figura 10** - Projeção da atividade gênica relativa entre os microarranjos de adenoma do colorretal sobre a rede da figura 9. O gradiente de cores é definido pela atividade relativa ( $n_\alpha$ ), se  $n_\alpha > 0,55$  o gradiente aumenta de amarelo até vermelho caracterizando aumento de expressão; se  $n_\alpha < 0,45$  o gradiente diminui de azul ciano até azul escuro caracterizando diminuição de expressão e quando  $0,45 < n_\alpha < 0,55$  o gradiente é verde e não caracteriza alteração (SIMÃO *et al.*, 2010).

Para a via de Apoptose não foi possível verificar um resultado claro de aumento ou diminuição de expressão da morte celular em estudos de adenoma. Em alguns pontos isolados foi notado aumento da atividade (pontos amarelos localizados a direita da figura 9), alguns pontos que não sofreram alteração da

<sup>7</sup> A funcionalidade dos dados de microarranjos sobre a rede de proteínas da figura 9 foi calculada pelo software *ViaComplex* usando os seguintes itens: arranjos (a,b), construção pelos nós e interações e com 50% de resolução, contraste, suavidade e zoom (SIMÃO *et al.*, 2010).

expressão (em verde) e outros com diminuição da expressão (por exemplo: pontos azuis na parte inferior central e na parte superior central da figura 10).

Em uma segunda análise, os microarranjos dos 32 pacientes do colorretal foram usados para comparar os níveis de expressão entre as cinco vias de reparo da ontologia *Ontocancro 1.0* (localizadas na rede de proteínas da figura 9) entre cinco vias de reparação do DNA, publicadas por Castro em 2007 (CASTRO *et al.*, 2007; KRISTOFF, 2011):

- Reparo por excisão de nucleotídeos (NER),
- Reparo por excisão de bases (BER),
- Reparo de bases mal pareadas (MMR),
- Recombinação (REC - fusão das vias de Reparo por recombinação

homóloga e Reparo por terminação não homóloga) (SIMÃO *et al.*, 2010). A fusão destas vias foi necessária, uma vez que, elas apresentam poucos genes que poderiam prejudicar a análise comparativa entre as vias de reparo da *Ontocancro* versus as vias de reparo utilizadas por Castro (CASTRO *et al.*, 2007).

O objetivo do teste foi verificar a sensibilidade das novas vias de reparo publicadas por Librelotto e colaboradores (2009) na ontologia *Ontocancro 1.0* e comparar os resultados encontrados com os resultados da atividade relativa e estatística significativa das vias de reparo publicadas anteriormente por Castro e colaboradores em 2007.

As 32 amostras de biopsias de adenoma foram agrupadas por média aritmética (equações 5, 6 e 7 da seção 2.5.3) e as 32 amostras de tecido controle (normais) sofreram o mesmo processo. As amostras de adenoma e controle foram inseridas juntamente com as vias de reparo na base de cálculo da análise estatística do *software ViaComplex* (CASTRO *et al.*, 2009). O *software* foi usado para medir a atividade relativa e o nível de significância das vias de reparo com *bootstrap* de 10.000 passos aleatórios e falso positivo de 5% (ver capítulo 2.5.2).

Os resultados apresentados na tabela 3 mostram as alterações encontradas para as vias de reparo da ontologia *Ontocancro 1.0* versus as vias de reparo publicadas por Castro em 2007. Na segunda coluna estão representados os números de genes em cada via. As listas de genes estão disponíveis na ontologia *Ontocancro 1.0* e no arquivo suplementar publicado por Castro e colaboradores,

disponível em ambiente virtual<sup>8</sup>. Na terceira coluna está mostrada a atividade relativa seguida da estatística significativa que descreve os níveis mais significativos de expressão das vias em relação ao microarranjo. Para as vias de reparo *Ontocancro*, observa-se que as vias apresentam maiores valores de atividade relativa quando comparadas as vias de reparo utilizadas por Castro (CASTRO *et al.*, 2007). Além disso, é observado que os níveis de significância estatística estão aumentados para as vias *Ontocancro*.

O cálculo da atividade relativa também foi realizado isoladamente para cada uma das 32 amostras de adenoma *versus* cada uma das 32 amostras da mucosa normal. Em 78% dos pacientes foi detectada alteração nas vias *Ontocancro* (SIMÃO *et al.*, 2010).

**Tabela 3** - Comparação da sensibilidade para as alterações entre as vias *Ontocancro* e as vias de reparo publicadas por castro e colaboradores em 2007. A avaliação da atividade relativa ( $n_a$ ), foi calculada para as amostras do colorretal adenoma *versus* normal para as 5 vias (SIMÃO *et al.*, 2010).

	Genes	Atividade Relativa ( $n_a$ )	Estatística Significativa
<b>Vias <i>Ontocancro</i></b>			
<b>(LIBRELOTTO <i>et al.</i>, 2009)</b>			
NER	51	0,596	SIM
BER	44	0,600	SIM
MMR	28	0,620	SIM
REC	46	0,593	SIM
<b>Vias Padrão</b>			
<b>(CASTRO <i>et al.</i>, 2007)</b>			
NER	28	0,585	NÃO
BER	17	0,600	NÃO
MMR	26	0,605	NÃO
REC	25	0,579	NÃO

Com os resultados apresentados nesta seção, as informações da ontologia *Ontocancro 1.0* quando aplicadas a microarranjos de adenoma do colorretal e mucosa normal, podem ajudar a expandir as investigações na progressão do câncer. Isso é possível devido a implementação de novos genes associados a vias de manutenção do genoma ligadas especificamente a reparação do DNA. Um resultado oposto foi observado para os resultados da rede funcional, mais precisamente para a via de Apoptose. Como a via de Apoptose apresenta um número elevado de genes que representam diferentes funções, não é possível verificar uma estatística significativa. Os resultados da expressão das vias de morte celular sugeriram o

<sup>8</sup> Disponível em: < <http://nar.oxfordjournals.org/content/35/6/1859/suppl/DC1> >

desenvolvimento de uma nova análise envolvendo especificamente subvias de apoptose.

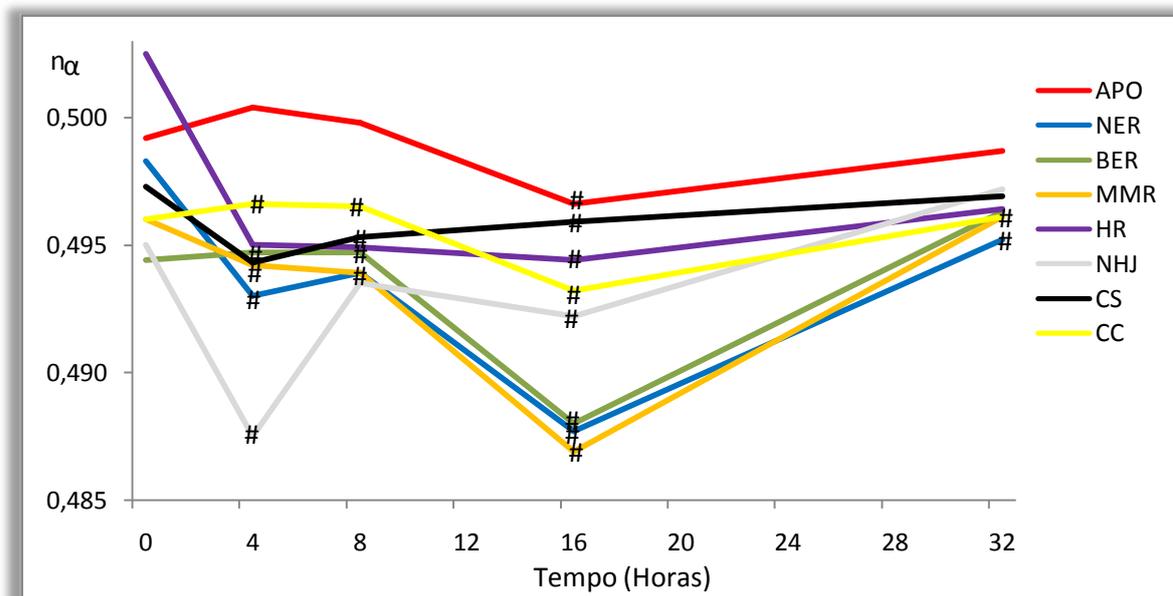
### **3.3 Resposta temporal das vias GMM em células da próstata expostas ao Cádmio**

Nesta aplicação foram analisados os dados de um estudo onde foram investigadas as alterações no crescimento de célula epiteliais normais de linhagens da próstata após a exposição a baixas doses de Cádmio (Cd). O objetivo foi de investigar o efeito do Cádmio sobre as vias GMM, que é considerado como uma substância carcinogênica.

As amostras utilizadas neste trabalho estão publicadas no banco de dados GEO, com série GSE9951, foram produzidas com a plataforma GPL 570 da *Affymetrix Human Genome U133 Plus* e normalizadas usando a média robusta de múltiplos chips (RMA) (GSE9951, 2010).

Foram analisadas 19 amostras, cada condição experimental com duas réplicas biológicas, exceto para a amostra do Cádmio da quarta hora que tinha apenas uma amostra. As amostras foram coletadas com condições experimentais de zero hora (no ato de aplicação do Cádmio), 4 horas após a aplicação, 8 horas, 16 horas e 32 horas após a aplicação do Cádmio. Cada réplica com aplicação experimental do Cádmio foi agrupada por média aritmética (equações 5, 6 e 7 da seção 2.5.3) e combinada com as réplicas de tecido normal em cada condição experimental. A combinação das réplicas experimentais do Cádmio foram submetidas ao cálculo da atividade relativa e estatística significativa sobre 8 vias Ontocancro: APO, NER, BER, MMR, HR, NHJ, CS e CC (exceto para a via apoptose expandida). Para realizar está análise foi utilizado o *software ViaComplex*, com *bootstrap* de 10.000 passos aleatórios e 5% de falsos positivos (SIMÃO, *et al.*, 2010). A figura 11 mostra a atividade relativa das vias Ontocancro em função do tempo de exposição ao Cádmio. Na figura é verificada uma diminuição significativa ("#") transitória de todas as vias Ontocancro após a exposição ao Cádmio (O símbolo "#" representa a diminuição significativa no instante de tempo). Este resultado mostra claramente uma diminuição da atividade das vias envolvidas na

manutenção do genoma, como maior destaque para as vias de reparo (SIMÃO *et al.*, 2010).



**Figura 11-** Sequência de tempo de alterações de todas as vias em 0, 4, 8, 16 e 32 horas após a exposição de 2,5 µM de Cádmio. O "#" representa a diminuição significativa no instante de tempo, APO - Apoptose, NER - Reparo por excisão de nucleotídeos, BER - Reparo por excisão de bases, MMR - Reparo de bases mal pareadas, HR - Recombinação homóloga, NHJ - Reparo por terminação não homóloga, CS - Estabilidade cromossômica e CC - Ciclo celular (SIMÃO *et al.*, 2010).

Sendo assim é mostrada a sensibilidade das vias Ontocancro na presença do Cádmio durante as 32 horas. Os resultados apontam a diminuição da atividade das vias de manutenção do genoma sobre a exposição da próstata ao Cádmio. Além disso, é mostrada a sensibilidade das vias Ontocancro na presença do Cádmio e após as 16 horas elas começam a retornar ao seu estado normal devido à eliminação do Cádmio pelo tecido da próstata.

## 4. BARREIRA ANTICÂNCER NA EVOLUÇÃO DE TUMORES

Algumas neoplasias malignas ligadas a lesões pré-cancerosas podem evoluir para o câncer. Um exemplo desta progressão tumoral ocorre em tecidos de adenomas que evoluem para carcinomas em tecidos do colorretal, conforme descrito na seção 2.3. No modelo publicado em 2008, Halazonetis e colaboradores propõem que em tecidos pré-cancerosos exista uma barreira envolvendo a ativação das vias de Apoptose e DDR que resiste à transformação maligna, mas que é rompida mais tarde dando origem a um tecido maligno. Segundo o modelo de Halazonetis os oncogenes induzem um stress na replicação do DNA levando à formação de DSB's e conseqüentemente a ativação da proteína *p53* que induz a apoptose e a senescência para barrar a evolução tumoral. A origem do stress replicativo se dá provavelmente pela desregulação do ciclo celular através de oncogenes que aumentam a atividade das quinases dependentes de ciclinas (*cdk's*) que atuam na fase G1 e S. A constante formação de quebras duplas associada a instabilidade genômica leva ao acúmulo de mutações e a progressão ao câncer (HALAZONETIS *et al.*, 2008).

O objetivo foi investigar quantitativamente os principais mecanismos envolvidos na ativação da barreira anticâncer que impede a progressão tumoral em tecidos de adenoma do colorretal (cólon), glândula adrenal, pâncreas e tireóide. Para isso foi construído um diagrama baseado em alterações em subvias envolvidas em GMM, especificamente ligado as vias de Resposta aos danos (DDR), Apoptose e Ciclo celular. Para complementar a análise de subvias foi calculado o *fold change* das subvias com alteração e de outras subvias que não apresentaram alteração em alguns tecidos. As alterações em subvias foram estabelecidas pela análise estatística do *software ViaComplex* e as alterações em genes pela análise do *fold change* dos pacotes do *Bioconductor* (descrição completa dos métodos estão disponíveis na seção 2.5).

#### 4.1 Descrição das vias e subvias

Conforme descrito no seção 2.3, no contexto de vias, existem as subvias que representam um conjunto de genes com funções bem específicas dentro de uma determinada via. Por exemplo, pode-se definir a via de Apoptose (morte celular) e uma subvia como um caminho dentro da via de apoptose, ou seja, subvia extrínseca de apoptose.

Para desenvolver o modelo de barreira anticâncer, foram selecionadas subvias com três funções de manutenção do genoma envolvidas em vias de Resposta ao dano (DDR), Apoptose e Ciclo celular.

Sendo assim, foram selecionadas subvias com funções específicas em cada uma das três funções citadas acima e extraídos dos bancos de dados:

- *National Cancer Institute (NCI) - Nature - Interaction Pathway Database*<sup>9</sup>,
- *Reactome*<sup>10</sup>,
- *BioCarta*<sup>11</sup>,
- *Ontocancro 1.0*<sup>12</sup>.

Outras três subvias utilizadas: Senescência replicativa, Ativação do *NFKB* e subvia *RB/E2F* foram extraídas de publicações recentes (SABATEL *et al.*, 2011; CALZONE *et al.*, 2008). A subvia de senescência replicativa foi desenvolvida com base em um conjunto de evidências de genes envolvidos no encurtamento dos telômeros que participam do processo natural de envelhecimento celular (mais detalhes dos genes pertencentes a esta subvia podem ser encontrados na ontologia *Ontocancro 2.0*)

As subvias de Ativação do *NFKB* e *RB/E2F* foram extraídas dos trabalhos desenvolvidos por Sabatel e colaboradores (2011) e Calzone e colaboradores (2008). Na tabela 4 estão listados os três conjuntos de vias com funções biológicas relevantes para o desenvolvimento do câncer: DDR, Apoptose e Ciclo celular, e suas respectivas subvias com funções específicas incluindo o número de genes presentes e os seus bancos de dados de origem. Todas as informações adicionais sobre estas subvias podem ser obtidas na ontologia *Ontocancro 2.0*<sup>12</sup>.

---

<sup>9</sup> Disponível em: < <http://pid.nci.nih.gov/>>

<sup>10</sup> Disponível em: <<http://www.reactome.org>>

<sup>11</sup> Disponível em: <<http://www.biocarta.com/>>

<sup>12</sup> Disponível em: <<http://ontocancro.inf.ufsm.br/>>

**Tabela 4** - Vias DDR, Apoptose e Ciclo celular, e subvias GMM usadas na análise de expressão de tecidos de adenoma e carcinoma com seus números de genes e suas fontes de dados.

<b>Vias de Resposta ao Dano (DDR)</b>		
<b>SUBVIAS</b>	<b>NÚMERO DE GENES</b>	<b>BANCOS DE DADOS</b>
Reparo por excisão de bases	44	Ontocancro
Via da anemia de Fanconi	24	Reactome
Recombinação homóloga	34	Ontocancro
Reparo de bases mal pareadas	28	Ontocancro
Reparo por excisão de nucleotídeos	51	Ontocancro
Reparo por terminação não homologa	14	Ontocancro
Reparo das quebras em fitas duplas no DNA	22	Reactome
<b>Vias de Apoptose /Senescência Replicativa</b>		
<b>SUBVIAS</b>	<b>NÚMERO DE GENES</b>	<b>BANCOS DE DADOS</b>
Sinalização <i>TNFR1</i>	13	BioCarta
Sinalização <i>TNFR2</i>	10	BioCarta
Sinalização do receptor <i>TNF</i>	40	NCI-Nature
Receptor de sinalização de morte	13	Reactome
Sinalização da apoptose em DDR	12	BioCarta
Indução da apoptose pelas receptores de morte <i>DR3/4/5</i>	20	BioCarta
Apoptose - Homo sapiens	83	Kegg
Cascata de caspases em apoptose	52	Ontocancro
Regulação da apoptose	60	Reactome
Fase de execução da apoptose	51	Reactome
Via extrínseca de apoptose	13	Reactome
Granzima A mediador de apoptose	13	BioCarta
Senescência Replicativa	56	Ontocancro
Ativação do <i>NFKB</i>	20	(SABATEI <i>et al.</i> , 2011)
<b>Vias de Ciclo Celular</b>		
<b>SUBVIAS</b>	<b>NÚMERO DE GENES</b>	<b>BANCOS DE DADOS</b>
Ciclo celular: pontos de controle G1/S	21	BioCarta
Fase S	112	Reactome
Ciclina e regulação do ciclo celular	15	BioCarta
Pontos de controle G1/S do dano ao DNA	60	Reactome
Postos de controle G2/M	43	Reactome
Fase mitótica G2-G2/M	86	Reactome
Fase mitótica M-M/G1	178	Reactome
Pontos de controle do fuso mitótico	19	Reactome
<i>RB/E2F</i>	78	(CALZONE <i>et al.</i> , 2008)

## 4.2 Dados de microarranjos

Para verificar a expressão das subvias e dos genes envolvidos na manutenção do genoma foram escolhidas amostras de microarranjos que continham tecidos normais, adenomas e câncer e que apresentam plataformas GPL 570 (tecidos do colorretal, adrenal e pâncreas) e GPL 96 (tecidos da tireóide) da *Affymetrix Human Genome U133 Plus*. Foram analisadas as seguintes séries do banco de dados *Gene Expression Omnibus* (GEO):

1- Série GSE4183 de colorretal (cólon): com 8 estudos de tecido normal, 15 estudos de processo inflamatório, 15 estudos de adenoma e 15 estudos de carcinomas (GSE4183, 2007);

2- Série GSE10927 do córtex adrenal: com 10 estudos de tecidos normais, 22 estudos de adenomas e 33 estudos de carcinomas (GSE10927, 2009);

3- Série GSE19650 de pâncreas: com 7 estudos de tecidos normais dos dutos pancreáticos, 6 estudos de adenoma mucinoso papilar e 6 estudos de carcinoma mucinoso papilar do pâncreas (GSE19650, 2010);

4- Série GSE27155 de tireóide: com 4 estudos de tecidos normais da tireóide, 17 estudos de adenoma foliculares da tireóide e 13 estudos de carcinoma folicular da tireóide. Para a análise deste estudo não foram usadas as amostras das papilas da tireóide por não apresentarem amostras de adenoma (GSE27155, 2011).

Todas as amostras foram normalizadas usando: microarranjo suíte versão 5 (MAS 5.0 - *Microarray Suite Version 5.0*), média robusta de múltiplos chips (RMA - *Robust Multichip Average*) (IRIZARRY *et al.*, 2003) e RMA para *Affymetrix GeneChip* (GCRMA - *GeneChip Affymetrix RMA*) (GOHLMANN *et al.*, 2009, WU *et al.*, 2004). As três normalizações foram comparadas e de acordo com os resultados apresentados no Anexo 1 deste trabalho foi usada Média robusta de múltiplos chips (RMA - *Robust Multichip Average*) devido a homogeneidade na distribuição de expressão que essa normalização apresentou. Todos os microarranjos estudados também poderão ser consultados no banco de dados *Ontocancro 2.0* (Ver seção 2.4.1).

### 4.3 Ferramentas de pesquisa usadas para construir o modelo de barreira anticâncer

Para construir o diagrama baseado no modelo quantitativo de expressão de subvias e genes de manutenção do genoma foram usados dois métodos de análise que estão descritos nas fundamentações deste trabalho:

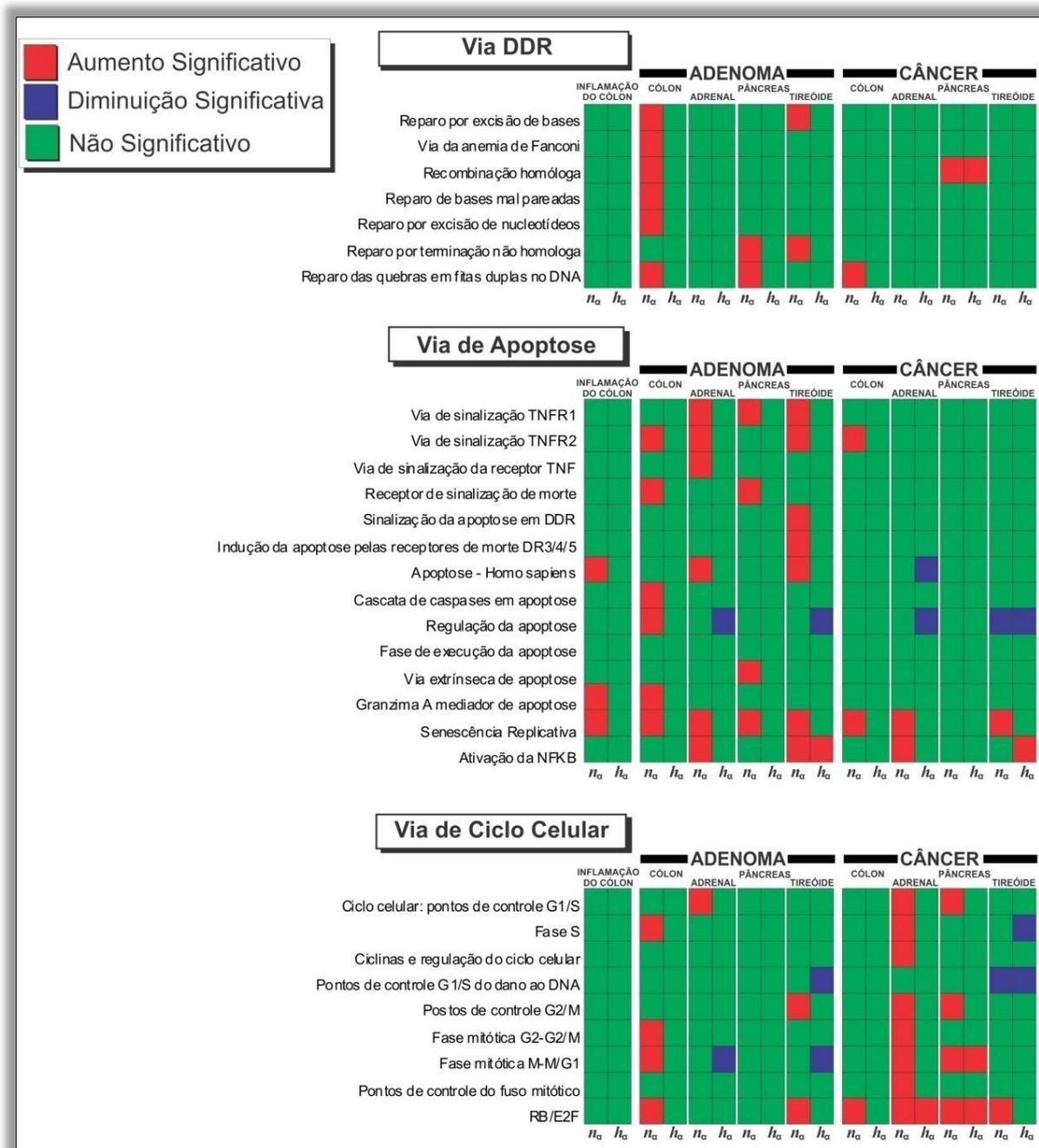
1- Análise significativa da atividade relativa e diversidade relativa de cada uma das subvias envolvidas nas funções (vias) de DDR, Apoptose e Ciclo celular (seção 2.5.2). Para determinar se uma alteração em uma subvia possui estatística significativa, com aumento (vermelho) ou diminuição (azul) foi utilizada uma reamostragem aleatória (*bootstrap*) com 10.000 repetições envolvendo todos os genes da série estudada. O nível de confiança utilizado na identificação de alterações foi de 5%. Para fazer esta análise de alterações da atividade relativa e diversidade relativa das subvias foi utilizado o *software ViaComplex*.

2- Para as alterações observadas na figura 12 foram localizados os genes com maior *fold change*, entre as amostras dos tecidos experimentais (adenoma ou câncer) *versus* o controle (normal). Além das alterações também foram localizados os *TOP-5* genes mais alterados para as doenças que não mostram alteração na figura 12 (Sobre todas as subvias na: via DDR de adenoma adrenal, na via DDR de câncer adrenal e tireóide, na via de Apoptose de câncer do pâncreas e na via de Ciclo celular para o adenoma do pâncreas). A mudança de expressão entre os tecidos foi calculada pelos pacotes estatísticos do *Bioconductor* na interface de linguagem de programação R. Em cada uma das subvias buscou-se os *TOP-5* genes mais alterados e procurou-se pelo consenso destes genes entre pelo menos duas doenças (ver seção 2.5.3).

#### 4.3.1 Alterações em subvias de manutenção do Genoma

A figura 12 descreve as três vias com funções específicas: DDR, Apoptose e Ciclo celular e suas respectivas subvias com as alterações aumentadas (vermelho) e diminuídas (azul), para a atividade relativa e diversidade relativa. Na análise global

da figura 12 observa-se que as vias DDR e Apoptose estão mais expressas nos adenomas do que em cânceres, enquanto que o oposto parece ocorrer com as subvias de Ciclo celular. Em uma análise mais detalhada observam-se alterações da atividade relativa e diversidade relativa das sete subvias de resposta aos danos do DNA (DDR) que estão envolvidas em reparo das quebras em fitas duplas do DNA (DSB) e quebras em fitas simples do DNA (DSS). Nas subvias de DSS foram encontradas 4 alterações: Reparo por excisão de bases (BER), Reparo de bases mal pareadas (MMR) e Reparo de por excisão de nucleotídeos (NER), todas as alterações ocorreram em tecidos de adenoma e nenhuma alteração em câncer. Nas subvias que envolvem a reparação das DSB foram encontradas alterações nas subvias: Anemia de Fanconi, Recombinação Homóloga, Reparo por recombinação não homóloga e em Reparo das quebras em fitas duplas do DNA foram verificadas 2 alterações em adenoma e 1 alteração em câncer. Comparando os resultados para os dois conjuntos de subvias (DSS e DSB) em adenoma e câncer, pode-se concluir que há uma evolução nas alterações de adenoma em relação ao câncer. Observa-se que em adenoma há presença de DSS que podem ser observadas pelas alterações encontradas nas subvias BER, NER e MMR e que diminuem em amostras de câncer devido a presença de alterações mais complexas como o caso das DSB. Em pré-câncer a presença das DSB já foi identificada por Halazonetis e colaboradores (2008), e as alterações encontradas nas subvias envolvidas em DSB indicam alterações em pelo menos três dos adenomas estudados, exceto para adenoma adrenal. O estresse causado por este tipo de anomalia poderá elevar as células a um aumento da morte celular, em adenomas, ocasionado por defeitos inerentes ao DNA (GORGOULIS *et al.*, 2005). Observa-se também que não foram encontradas alterações para os adenomas adrenais e a maior parte das alterações ocorreu em adenomas colorretais.



**Figura 12** - Perfil de significância das vias de Resposta ao dano no DNA (DDR), Apoptose e Ciclo celular para os estudos de colorretal, glândula adrenal, pâncreas e tireóide. O aumento significativo é identificado nos quadros vermelhos, a diminuição significativa nos quadros azuis e os quadros verdes indicam que não há alteração.  $n_{\alpha}$  é a atividade relativa e  $h_{\alpha}$  é a diversidade relativa.

Um padrão similar é observado na via de apoptose, predominando subvias aumentadas em adenomas em relação ao câncer. Particularmente foram encontradas alterações para inflamação apenas na via de Apoptose e uma peculiaridade é a alteração da subvia de Senescência replicativa em quase todas as

amostras e inclusive em inflamação. Das alterações encontradas na via de Apoptose podem ser destacados os aumentos das subvias de Sinalização dos fatores de necrose tumoral (*TNF*), que englobam as subvias de Sinalização do *TNFR1* e Sinalização do *TNFR2* que estão associados com a subvia de Sinalização dos receptores *TNF*. Estas subvias têm papel importante na ativação da morte celular e as alterações encontradas para os adenomas mostram que durante a evolução tumoral há ativação de citosinas envolvidas em processos inflamatórios causando a morte celular programada das células em resposta ao estresse causado pelos danos ao DNA (LONCKSLEY *et al.*, 2001).

A sinalização da morte celular pode ser identificada pela ativação das seguintes subvias em adenoma: Receptor de sinalização de morte, Sinalização da apoptose em DDR e Indução da apoptose pelos receptores de morte *DR3*, *DR4* e *DR5*. Estas três subvias são sensoras de danos e sinalizam a apoptose em resposta aos danos no DNA. A sinalização é induzida pelos genes *FAS*, *TNFR*, *DR3*, *DR4* e *DR5* que recrutam outras proteínas especializadas em ativar a cascata de eventos apoptóticos através da ativação de proteínas efetoras. Na figura 12 observa-se que há ativação de algumas subvias efetoras em adenoma: Apoptose - Homo sapiens, Cascata de caspases em apoptose e Via extrínseca de apoptose. As alterações nestas subvias indicam que há ativação efetiva da apoptose em adenomas e em câncer é evidente a falta de ativação destas subvias. As subvias de Regulação da apoptose e Granzima A mediador de apoptose são formadas por conjuntos de genes responsáveis por regular a ativação da apoptose. A subvia Granzima possui mecanismos usados por células T citotóxicas que ajudam a matar as células defeituosas através da liberação de um conjunto de proteínas granzimas. A presença destas proteínas já foi encontrada em alguns tipos de adenomas do colorretal (BOTS *et al.*, 2006, GOPING *et al.*, 2003). A subvia de Senescência replicativa engloba o conjunto de genes que atua no processo de parada permanente da proliferação celular. Esse processo está envolvido no envelhecimento celular e também possui implicação no encurtamento prematuro dos telômeros provocado por alterações no DNA. Na progressão tumoral a senescência é induzida por estímulos de oncogenes e a sua resposta depende da ativação de duas proteínas essenciais regidas pelo *TP53* e reguladas pelo *PRB* e *P16*. A senescência ocorre em resposta ao estresse intracelular e contribui para a parada do ciclo celular, que impede a propagação de danos para a segunda geração de células evitando uma transição

maligna. Em trabalhos recentes foram identificadas células senescentes em tecidos renováveis e em tecidos que enfrentam uma prolongada inflamação (COPPÉ *et al.*, 2010).

Na análise das subvias alteradas para a inflamação do colorretal foram observadas três subvias alteradas: Apoptose - Homo sapiens, Granzima A mediador de apoptose e em Senescência replicativa (ver figura 12). As alterações nestas subvias de apoptose ocorrem pela dinâmica de renovação das células do colorretal que estão envolvidas em doenças inflamatórias do intestino considerando a apoptose como uma das características encontradas em colite (condição inflamatória do intestino) (CHEN *et al.*, 2010) que está envolvida na regulação e ativação dos processos inflamatórios característicos de Colite Ulcerosa (UC) e na Doença de Crohn (DC) (CHEN *et al.*, 2010; DANESE *et al.*, 2011; KRAUS *et al.*, 2009). Enquanto as duas primeiras subvias analisadas estão diretamente envolvidas com apoptose, a subvia de Senescência replicativa permite a parada permanente do ciclo celular para evitar a progressão de danos no DNA em células com inflamação do colorretal. A importância do estado senescente na inflamação do colorretal aumenta visto que o encurtamento dos telômeros é um dos fatores que contribui para a progressão de UC para o carcinoma do colorretal (AUBERT *et al.*, 2008). Mais detalhes sobre as alterações da apoptose na inflamação do colorretal estão apresentadas na seção 4.3.2 e nas discussões estendidas apresentadas no Anexo 2.

A subvia de Ativação do *NFKB* também está alterada tanto em adenoma quanto em câncer. Ela tem função de controlar a apoptose e a senescência e está associada com resposta ao dano além de coordenar a parada do ciclo celular pela ativação de alguns fatores tais como *TP53* e *NFKB* (SABATEL *et al.*, 2011).

Na via de Ciclo celular não foi observado um padrão claro e esse padrão parece depender de cada tipo específico de tecido. Em uma abordagem mais geral da via percebe-se que a maioria das alterações ocorreu em câncer, enquanto nenhuma alteração é observada nos adenomas do pâncreas. Conforme descrito na seção 2.3, devido ao aumento de mutações, em câncer, há um aumento de proliferação seguido da perda da adesão celular, reparação do DNA e morte celular programada (CASTRO *et al.*, 2007). Esse aumento do índice de proliferação ocorre em diferentes fases conforme mostrado na figura 12. Nas subvias: Ciclo celular: pontos de controle G1/S e Pontos de controle G2/M as alterações são observadas

em cânceres adrenais e pâncreas e nas subvias da Fase S observa-se uma diminuição da diversidade em câncer da tireóide, sugerindo que neste tecido alguns genes possuem expressão muito elevadas assim como as alterações encontradas em câncer da tireóide para a subvia de Pontos de controle G1/S de danos do DNA. As alterações em adenoma ocorrem devido a presença do estresse replicativo causado por quebras em fitas simples e por quebras em fitas duplas, que por sua vez, ativam mecanismos de resposta ao dano do DNA. Desta forma, há uma diminuição do ciclo celular e um aumento dos mecanismos de resposta ao dano e apoptose que formam uma barreira que impede a evolução tumoral (HALAZONETIS *et al.*, 2008). Devido a complexidade dos defeitos em câncer percebe-se que há muitas alterações na via de Ciclo celular. Através da alteração de algumas subvias (ex. Pontos de controle do fuso mitótico) é possível afirmar que existem defeitos estruturais na formação do fuso mitótico em câncer. Além desta alteração pode ser verificado também que há uma desregulação da subvia *RB/E2F* em câncer. Na figura 12 verifica-se que esta subvia está com a atividade e diversidade relativa alterada em câncer adrenal e pâncreas e com a atividade relativa alterada em câncer de colorretal e tireóide, enquanto que em adenoma a subvia tem duas alterações, uma em colorretal e outra em tireóide. A subvia *RB/E2F* está envolvida na regulação da transição da fase G1 para S, sendo assim, a subvia está envolvida na regulação do ciclo celular, replicação do DNA e em outras funções ligadas a ativação das vias de Apoptose. De acordo com Calzone e colaboradores (2008) muitas alterações nesta subvia foram verificadas em retinoblastoma (tumor maligno da retina) (CALZONE *et al.*, 2008).

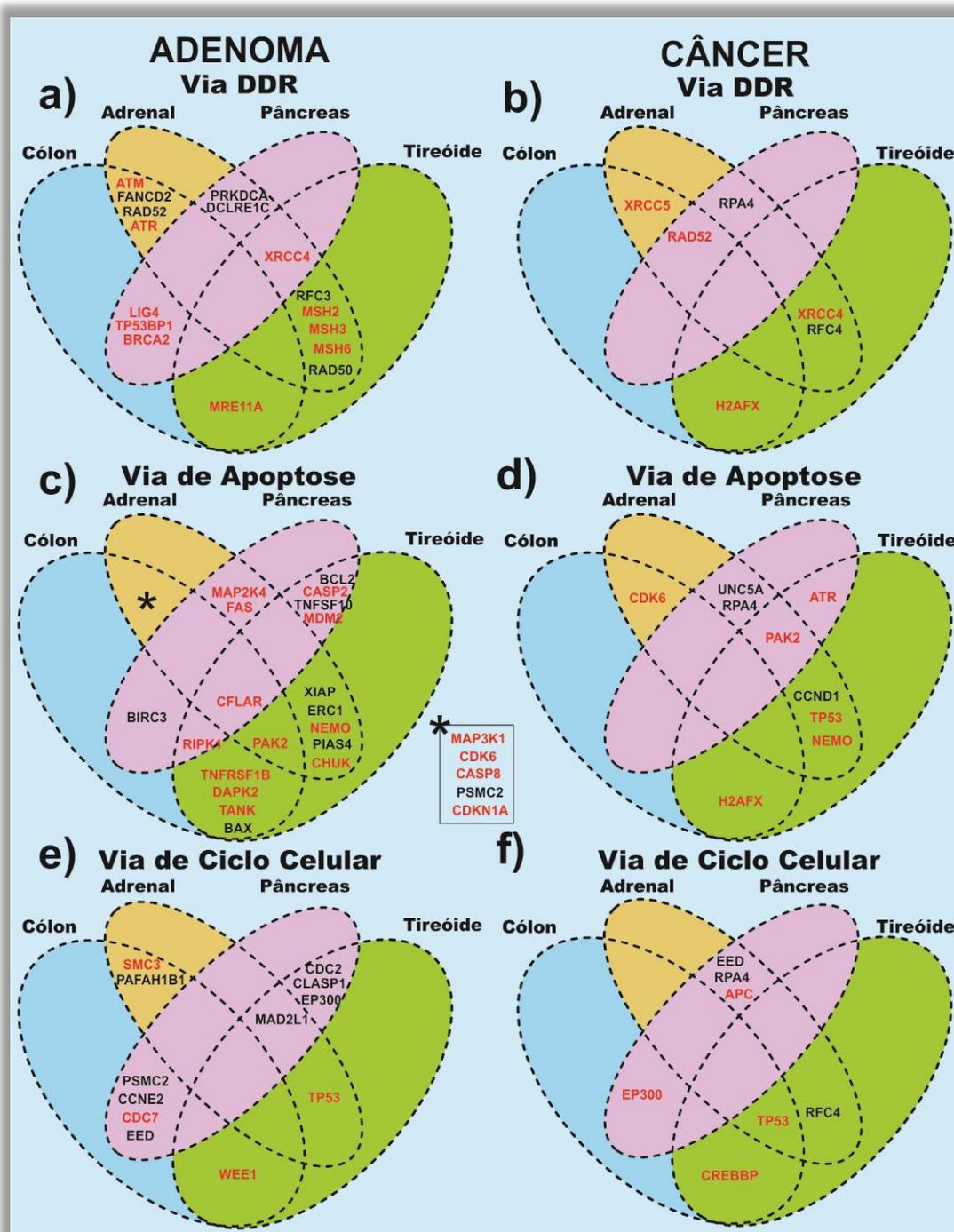
Não foram detectadas alterações em algumas vias como é o caso da via DDR de adenoma adrenal, na via DDR de câncer adrenal e tireóide, na via de Apoptose de câncer do pâncreas e na via de Ciclo celular para o adenoma do pâncreas. Estas doenças sem alteração sugerem uma análise detalhada do *fold change* para todas as subvias (mais detalhes sobre essa análise estão disponíveis na próxima seção).

### 4.3.2 Consenso dos genes diferencialmente expressos

Nesta seção são apresentados os resultados do consenso da análise dos genes diferencialmente expressos. O diagrama de *Venn* da figura 13 descreve o consenso de genes com *fold change* mais alterado entre pelo menos dois tipos de adenomas ou cânceres (colorretal (cólon), glândula adrenal, pâncreas e tireóide) para cada via: DDR, Apoptose ou Ciclo celular. O *fold change* foi calculado sobre os genes das subvias com aumento (vermelho) ou diminuição (azul) e sobre algumas doenças que não apresentaram nenhuma alteração: via DDR de adenoma adrenal, na via DDR de câncer adrenal e tireóide, na via de Apoptose de câncer do pâncreas e na via de Ciclo celular para o adenoma do pâncreas, conforme mostrado na figura 12.

No Anexo 3 encontram-se três tabelas que mostram os *TOP-5* genes do consenso com *fold change* mais elevados entre os adenomas e cânceres do colorretal, adrenal, pâncreas e tireóide. Ainda na figura 13 foram marcados em vermelhos todos os genes relevantes para a ativação da barreira anticâncer e que perderam sua adesão junto aos mecanismos de manutenção do genoma em câncer.

Na figura 13a estão apresentados os genes alterados na via DDR. Visivelmente, muitos dos genes do consenso (*MRE11A*, *TP53BP1*, *BRCA2*, *LIG4*, *ATM* e *XRCC4*) pertencem as subvias de Reparo das quebras em fitas duplas do DNA e Reparo por combinação não homologa (NHEJ). Outros genes do consenso (*MSH2*, *MSH3* e *MSH6*) pertencem a subvia Reparo por excisão de bases (BER). Na via DDR em câncer (figura 13b) o consenso também mostra alguns genes ligados a subvia de Reparo de quebras em fitas duplas do DNA: *H2AFX*, *XRCC4*, *XRCC5* e *RAD52*.



**Figura 13** - Diagramas de Venn usados para avaliar o consenso entre os genes alterados em adenomas e cânceres do colorretal, glândula adrenal, pâncreas e tireóide entre o consenso dos 5 genes com *fold change* mais alterados em cada subvia. Em vermelho estão os genes com evidências e ligados a barreira anticâncer. A razão de alteração entre o experimento (adenoma e câncer) *versus* o controle (tecido normal) foi calculado pela análise de expressão diferencial e os resultados são mostrados nos diagramas de Venn para: (a) e (b) o consenso para os genes da via DDR em adenoma e câncer, respectivamente. (c) e (d) o consenso para os genes da via de Apoptose em adenoma e câncer, respectivamente. (e) e (f) o consenso para os genes da via de Ciclo celular em adenoma e câncer, respectivamente.

No consenso da figura 13c para a via de Apoptose foram encontrados vários genes que estão associados com a regulação e ativação da apoptose através dos fatores de necrose tumoral (*TNF*) e que estão presentes na subvia de Senescência replicativa. Associado com as subvias *TNF* foram encontrados os seguintes genes: Na subvia Sinalização da *TNFR1* os genes *CASP8*, *CASP2* e *RIPK1*, na subvia Sinalização da *TNFR2* os genes *TANK*, *RIPK1* e *TNFRSF1B*. Na subvia de Regulação da apoptose os genes *PAK2* e *DAPK2*. Três genes do consenso pertencem a subvia de Senescência replicativa: *CDK6*, *CDKN1A* e *MDM2*. Outra subvia bastante alterada em adenoma é a de Ativação do NFKB, esta subvia apresenta dois genes importantes: *NEMO* e *CHUK*. O consenso da figura 13c mostra dois genes: o *FAS* alterado nas subvias Receptor de sinalização de morte, Apoptose - Homo sapiens e Via extrínseca de apoptose e o *CFLAR* alterado no consenso entre os quatro adenomas.

Em câncer (figura 13d) há ausência de ativação de algumas subvias de apoptose (HALAZONETIS *et al.*, 2008), com isso o consenso entre os genes é menor e alguns genes têm evidências na regulação e na Subvia de Senescência replicativa, como os genes *DAPK1* e *PAK2* envolvidos na Regulação da apoptose, *CDK6*, *TP53*, *ATR* e *H2AFX* envolvidos na Senescência Replicativa e o *NEMO* envolvido na subvia de Ativação do NFKB.

Na figura 13e o consenso para os genes envolvidos na via de Ciclo Celular em adenoma mostra a subvia Fase mitótica M-M/G1 como a mais alterada, com três genes: *SMC3*, *CDC7* e *WEE1*. Outra proteína que merece destaque neste consenso é a *p53* que está alterada nas subvias: Ciclo celular: pontos de controle G1/S e Pontos de controle G1/S de danos ao DNA. Como a característica mais comum em câncer é a proliferação aberrante (CASTRO *et al.*, 2007), percebe-se a ativação de algumas proteínas importantes para a manutenção do ciclo celular em câncer, em particular percebe-se a ativação da subvia *RB/E2F* com os genes: *EP300*, *CREBBP*, *APC* e *TP53*. Uma discussão estendida da função dos genes do consenso poderá ser verificada no Anexo 4 deste documento.

### 4.3.3 Genes alterados em tecidos do colorretal inflamados

Na análise das subvias alteradas na inflamação do colorretal, foram observadas alterações em três subvias de apoptose: Senescência replicativa, Apoptose - Homo sapiens e para a subvia Granzima A mediador de apoptose (ver figura 12). Considerando que a apoptose e a proliferação celular podem ser encontradas em colites (CHEN *et al.*, 2010) e envolvidas na regulação e ativação dos processos inflamatórios característicos de Colite Ulcerosa (UC) e Doença de Crohn (DC) (CHEN *et al.* 2010, DANESE *et al.* 2011, DI SABATINO *et al.* 2011, KRAUS *et al.* 2009), estas alterações podem ser explicadas como uma consequência da dinâmica de remoção das células do colorretal envolvidas em doenças inflamatórias do intestino.

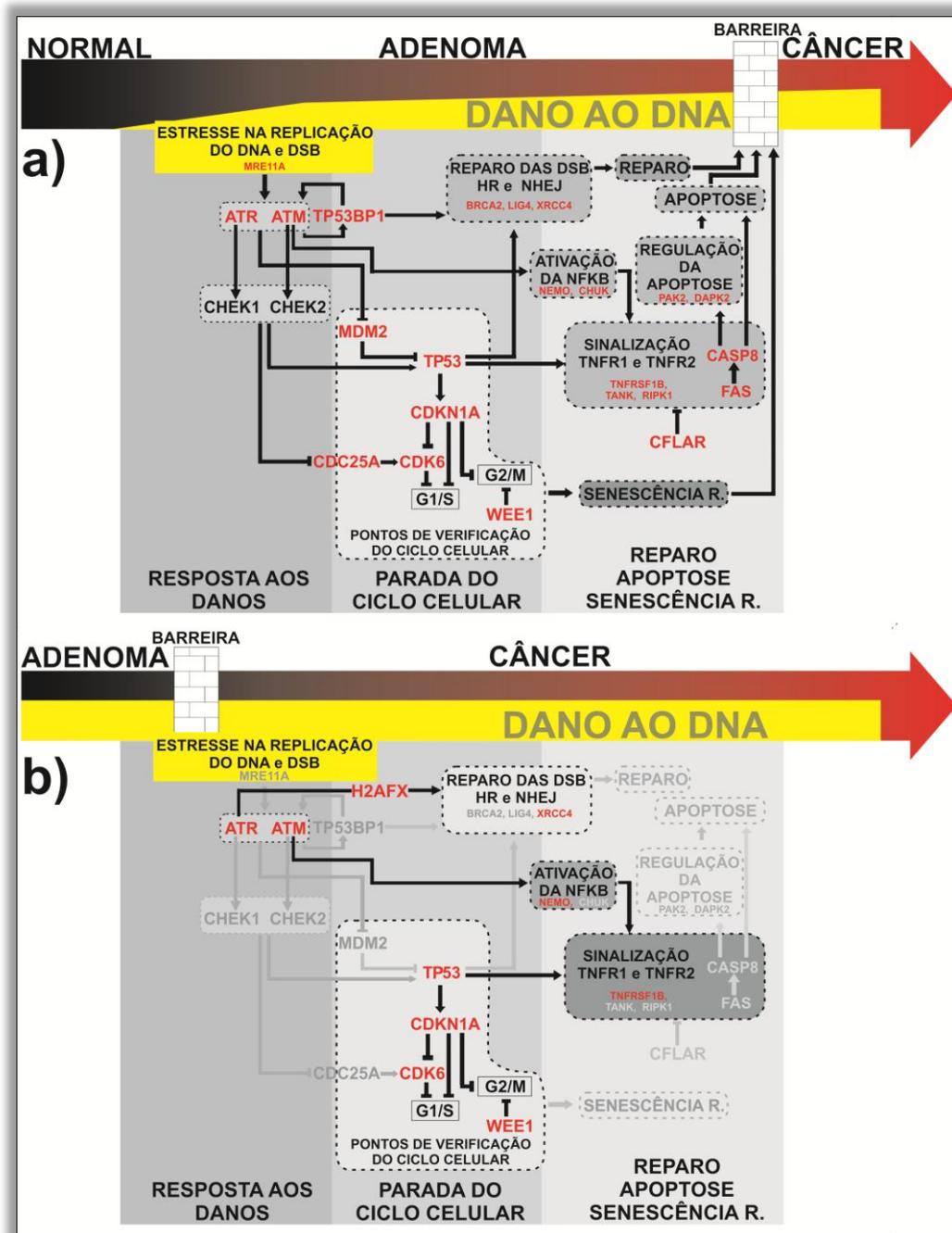
Pela análise do *fold change* das subvias alteradas em apoptose foram localizados os *TOP-5* genes mais alterados. Na subvia de Senescência replicativa os genes mais relevantes são: *CDK6*, *CDC25A* e *CDKN1A* que indicam a regulação do estado senescente no tecido do colorretal inflamado por *CDKN1A*. Na subvia de Apoptose - Homo sapiens os genes *TNFRSF10D/B*, *BIRC3*, *NTRK1* e *BAX* estão diferencialmente expressos. Os genes *TNFRSF10D/B* que codificam os receptores *TRAILR2* e *TRAILR4* possuem uma interação capaz de inibir a apoptose do ligante *TRAIL* que regula a inibição da apoptose através da Ativação do *NFKB* (FALSCHLEHNER *et al.*, 2007). O gene *BIRC3* codifica o *CIPA2* e é também responsável pela inibição da apoptose através da Ativação do *NFKB* e contribui para prevenir a ativação do gene *CASP8* através do *RIPK1* (GYRD-HANSEN *et al.*, 2010). A análise do *fold change* da subvia Granzima A mediador de apoptose localizou os seguintes genes: *SET*, *HMGB2*, *APEX1* e *ANP32A* indicando que a inflamação do colorretal ativa a apoptose e também é coordenada pelas células T. (ELMORE, 2007; BOTS *et al.*, 2006). Mais detalhes e uma discussão estendida das funções biológicas dos genes envolvidos em processos inflamatórios podem ser obtidos no Anexo 2.

#### 4.4 Diagrama da barreira anticâncer

Com os resultados apresentados nos itens anteriores (4.3.1 e 4.3.2), foi construído um diagrama que representa o processo evolutivo de ativação das subvias e genes envolvidos na manutenção do genoma através do reparo, ciclo celular, apoptose e senescência replicativa. O estudo da barreira anticâncer está resumido no diagrama de alterações em adenomas e para o câncer incluindo algumas subvias alteradas e genes (em vermelho) que pertencem ao consenso de genes com maior *fold change*. Nas próximas seções serão ilustradas as cadeias de mecanismos ligados a detecção e ativação de resposta aos danos do DNA, pontos de verificação do ciclo celular, subvias relacionadas a apoptose, senescência replicativa e reparo das DSB que utiliza as subvias de Recombinação homóloga (HR) e Reparo por terminação não-homóloga (NHEJ). A seguir será descrito o cenário para o câncer, incluindo os genes e subvias que possuem expressão diminuída.

##### 4.4.1 Ativação da barreira em tecidos de adenoma

Na figura 14a observa-se que as quebras em fitas duplas do DNA (DSB) desencadeiam uma cascata de eventos responsáveis pela parada do ciclo celular e pela ativação da DDR, apoptose e senescência. A ativação da subvia de Reparo das quebras em fitas duplas em DDR (figura 12) e o gene do consenso *MRE11A* pertencente ao complexo *MRN* e sensor de DSB desencadeiam a ativação dos genes *ATM* e *ATR* e de seus substratos *CHEK1* e *CHEK2* para coordenar uma resposta aos danos do DNA (BOHGAKI *et al.*, 2010; LEE *et al.* 2010). *ATM* e *ATR* são sensores de danos e ativam a sinalização para os pontos de verificação em resposta ao estresse replicativo (HALAZONETIS *et al.*, 2008).



**Figura 14** - O diagrama ilustra os mecanismos envolvidos na barreira contra o desenvolvimento tumoral com base nos resultados das figuras 12 e 13. Em vermelho estão os genes analisados no consenso. (a) Adenoma: A barreira é ativada em resposta aos danos do DNA, pelo gene *MRE11A* que atua sobre as quinases *ATM* e *ATR* que regulam as múltiplas cascatas de sinalização em resposta aos danos. Elas são reguladas pelo *TP53BP1* e ativam os pontos de verificação, DDR e apoptose pela subvia de Ativação do *NFKB*. As quinases ativam o *CHEK2* e *CHEK1* que estão envolvidas na sinalização dos pontos de verificação. Estas proteínas regulam em conjunto as proteínas *p53* e *cdk6* para promover os pontos de verificação nas fases *G1/S* e *G2/S*. O *p53* promove a parada do ciclo celular, reparo do DNA, ativação da senescência replicativa e apoptose. O reparo é ativado por duas subvias (NHEJ e HR) e pelos genes alterados *BRCA2*, *XRCC4* e *LIG4*. A apoptose é sinalizada pelo *NFKB* que ativa as subvias de sinalização *TNFR1/2* e que ativam o gene *CASP8*. A apoptose é também regulada pela subvia de Regulação da Apoptose e pelo gene alterado *CFLAR*. (b) Câncer: A figura mostra em cinza claro vários genes envolvidos no

diagrama em adenoma que estão inibidos em câncer incapazes de ativar a barreira anticâncer. Muitos genes participam da subvia *RB/E2F* que regula o ciclo celular na transição da fase G1 para S. Um dos genes responsáveis pela regulação é o *TP53*.

O *ATM* atua como transdutor de sinal, controla os genes *CHEK2* e *TP53* e ativa a via de Apoptose através dos genes alterados do consenso *NEMO* e *CHUK* pertencentes a subvia de Ativação do *NFKB* enquanto *ATR* regula o gene *CHEK1* (BOXUS *et al.*, 2012; STUDZINSKI, *et al.*, 2010; SALMINEN *et al.*, 2011). No diagrama a proteína *53bp1* ou proteína de ligação *p53* regula as proteínas *atm* e *atr* e desempenha papel na regulação da subvia de Reparo por terminação não homóloga, além de sinalizar os pontos de verificação do ciclo celular através da regulação dos substratos *CHEK1* e *CHEK2* (LEE *et al.*, 2010; SCHULTZ *et al.*, 2000). No caminho de ativação dos genes *ATM* e *ATR*, o gene *TP53* é ativado pelos genes *CHEK1* e *CHEK2* forçando a parada do ciclo celular pela regulação do *MDM2*. Em tecidos normais o *TP53* é mantida em níveis basais de expressão, devido a sua interação com o *MDM2*, no entanto na presença de estresse replicativo o *MDM2* é inibida e *p53* é ativada por *CHEK1* e *CHEK2* em resposta aos danos (DEY *et al.*, 2008; DONHOWER, 2009).

*CHEK2* e *CHEK1* induzem a parada do ciclo celular para permitir o reparo de danos no DNA, bem como a sua participação na regulação de apoptose ou senescência, no caso de danos irreparáveis. A parada do ciclo celular é consequência da inibição do *CDC25A* e na ativação da *p53* resultando na regulação e na inibição das ciclinas quinases dependentes: *cdk6* e *cdkn1a* que atuam diretamente na proliferação celular (STUDZINSKI *et al.*, 2010). As paradas do ciclo celular em G1/S e G2/M serão controladas por *p53* que ativa o gene *CDKN1A* para prevenir a progressão do ciclo celular. O *CDK6* é inibido por *CDKN1A* para induzir a parada nos pontos de verificação da fase G1/S (STUDZINSKI *et al.*, 2010). A inibição do ciclo celular na fase G2/M é regulada pela interação entre o *CDKN1A* e a quinase *wee1* responsável por controlar a transição nos pontos de controle da fase G2/M. *WEE1* inibe a ativação das *cdk's* responsáveis pela parada do ciclo celular e ativa a subvia de Senescência replicativa (SORENSEN *et al.*, 2012).

Através de um caminho independente do gene *TP53* pode ser ativado os mecanismos de reparo das DSB, reconhecidos no modelo pela ativação de 3 subvias de reparo: Reparo das quebras em fitas duplas do DNA (Reparo das DSB), Reparo por terminação não homóloga (NHEJ) e Reparo por recombinação homóloga

(HR) e pelos genes alterados no consenso: *BRCA2*, *LIG4* e *XRCC4*. Na subvia HR o gene *BRCA2* atua como um supressor de tumor que está envolvido no reparo de danos do DNA e tem um papel importante no reparo de erros causados pelas DSB (DUNCAN *et al.*, 1998). Os genes *XRCC4* e *LIG4* formam um complexo responsável pelas etapas de ligação da subvia NHEJ, onde o *XRCC4* aumenta a atividade de adesão do *LIG4* (SENGULTA *et al.*, 2005).

A ativação da apoptose ocorre por dois caminhos diferentes: pela subvia de Ativação do *NFKB* e pela ativação do *TP53*. Pela Ativação do *NFKB*, os genes *NEMO* e *CHUK* interagem com o *RIPK1* e com as subvias de Sinalização *TNFR1* e Sinalização *TNFR2* que ativam o gene *CASP8* e a apoptose (BITON *et al.*, 2011). O gene *CASP8* (regulado pelo *FAS*) participa do complexo de morte celular induzido (DISC) que ativa uma cascata de caspases responsável por desencadear a apoptose (FALSCHLEHMER *et al.*, 2007).

O gene *CFLAR* que aparece no consenso entre todos os adenomas é uma proteína que pode facilitar a apoptose formando um heterodímero com o gene *CASP8* que ajuda na clivagem de proteínas pró-apoptóticas (BITON *et al.*, 2011; JING *et al.*, 2012). A regulação da apoptose ocorre pelos genes alterados do consenso em adenomas: *PAK2* e *DAPK2* presentes na subvia de Regulação da apoptose. Estes genes são ativados por estímulos extracelulares gerados pelos danos ao DNA que ativam o programa de morte celular em adenomas (LI, *et al.*, 2011).

#### **4.4.2 Perda de subvias e genes em tecidos cancerosos**

O diagrama da figura 14b é baseado no consenso entre pelo menos 2 doenças com os *TOP-5* genes menos ativos entre o *fold change* das amostras de câncer (colorretal, adrenal, pâncreas e tireóide) sobre cada uma das alterações de atividade relativa e diversidade relativa (aumento ou diminuição) das subvias de adenoma da figura 12. O objetivo desta análise foi de avaliar a integridade das subvias e dos genes do diagrama da figura 14a em câncer (figura 14b). O resultado completo dos genes menos ativos em câncer é descrita no Anexo 5.

A figura 14b mostra em cinza claro todos os genes que perderam sua atividade na análise do *fold change* e outros genes que permaneceram expressos no consenso da figura 12 e na figura 13 em câncer. Sendo assim, pela análise do *fold change* (consenso da figura 13f) foi possível identificar sete genes do consenso alterados e que pertencem a subvia *RB/E2F* responsável por regular o ciclo celular na transição da fase G1 para S. Um destes genes responsáveis pela regulação é o *TP53* que aparece alterado em cânceres do colorretal, glândula adrenal e tireóide (SENGUPTA *et al.*, 2005). O consenso também mostra o gene *H2AFX* alterado na via de Apoptose, ele é um sensor de DSB e um marcador de câncer (KIM *et al.*, 2010). O *H2AFX* é regulado pela proteína *atr*, e é responsável por identificar as DSB que danificam as células e podem levar o tecido ao câncer. Observando o modelo da figura 14b o *H2AFX* aparece imediatamente alterado após a barreira, em câncer, e o seu objetivo é ativar a subvia NHEJ, monitorar e detectar as DSB durante o processo carcinogênico (BONNER *et al.*, 2008).

Com a análise dos genes com *fold change* diminuídos em câncer foi possível identificar alguns genes que perderam sua capacidade de reparo na via DDR: *LIG4*, *MRE11A* e *TP53BP1*. Como já foi mencionado anteriormente, os genes *XRCC4* e *LIG4* participam da subvia de Reparo das quebras em fitas duplas do DNA, Reparo por terminação não homóloga e Reparo por recombinação homóloga. O *MRE11A* é um sensor de DSB e *TP53BP1* atua como regulador do complexo *MRN* e sobre os genes *ATM* e *ATR*. Como estes genes apresentam inibição em câncer é observado que há uma diminuição da DDR que facilita a proliferação aberrante de células cancerosas. Na via de Apoptose também foram encontrados genes com *fold change* diminuído e que não estão apresentados no modelo da figura 14: *FAS*, *CFLAR*, *RIPK1*, *TANK*, *CASP8* e os genes da subvia de Regulação da Apoptose, *PAK2* e *DAPK2*. Por outro lado verificou-se que os genes envolvidos na via de ciclo celular se mantiveram ativos e os resultados não mostraram nenhum sinal de inibição.

## 4.5 Conclusões

Os estudos identificaram duas características importantes da transformação que complementam o modelo proposto por Halazonetis e colaboradores (2008): (1) A ativação da subvia de apoptose, *NFKB* (independente da *p53*) que está principalmente relacionada com o elevado nível de DSB e (2) a alteração da subvia de Senescência replicativa em quase todas as amostras e inclusive relacionada com a inflamação do colorretal. No primeiro caso observa-se que em adenomas a apoptose é ativada independentemente do gene *TP53*, através das subvias de Ativação do *NFKB* que permanece ativa mesmo em câncer. A atividade do gene *NFKB* pode ter efeitos benéficos ou prejudiciais em relação a homeostase do tecido. Na supressão do tumor esta subvia pode regular a parada do ciclo celular, o estado senescente e a sinalização do sistema imune para remover as células danificadas. O segundo caso sugere o envolvimento da erosão dos telômeros em adenomas e na progressão para o câncer. Particularmente, a inflamação e adenoma do colorretal têm comportamentos diferentes na ativação da barreira anticâncer, enquanto na inflamação há ativação da apoptose e senescência o ciclo celular não se altera, em adenomas existe a ativação de subvias responsáveis por diferentes tipos de reparo do DNA e um equilíbrio entre o ciclo celular e a apoptose característica do colorretal para garantir a integridade dos tecidos e a supressão tumoral.

O modelo mostra ainda o gene *CFLAR* alterado em todos os adenomas e em outros dois estudos que não estão descritos neste trabalho: adenoma pituitário (GSE4488, 2006) e em adenoma do colorretal (GSE8671, 2007)<sup>13</sup>. Todas as alterações encontradas para o *CFLAR* sugerem esse gene como um marcador de adenomas.

---

<sup>13</sup> Estas séries não foram incluídas à tese por que no conjunto de amostras não continham amostras de câncer.

## 5. CONCLUSÕES

Neste trabalho foram abordados métodos de análise estatística da expressão de genes e vias envolvidas nos mecanismos de manutenção do genoma para identificar as alterações em tecidos de adenoma que atuam como uma barreira anticâncer. Baseado no modelo qualitativo apresentado por Halazonetis e colaboradores em 2008, descreveu-se um diagrama quantitativo que define as alterações em vias e genes presentes na barreira que impede a evolução tumoral.

Para desenvolver o diagrama, os genes envolvidos na manutenção do genoma foram caracterizados pela modelagem da rede de manutenção do genoma. Modelagem que foi usada para descrever a expressão dos genes e das vias que atuam como mecanismos de manutenção do genoma em tecidos de adenoma. Para isso, foram utilizadas análises de atividade relativa, diversidade relativa e análise do *fold change* usadas para identificar as alterações dos GMM em amostras de tecidos relacionadas a evolução tumoral, tais como: inflamação, adenoma e câncer do colorretal, amostras de tecidos da próstata, adenomas e cânceres adrenais, pâncreas e tireóide.

Em uma primeira abordagem, uma rede de interação de proteínas envolvidas em GMM foi modelada e foram analisados dados de expressão de genes envolvidos na evolução tumoral através de métodos estatísticos de atividade relativa e diversidade relativa. Na modelagem da rede de manutenção do genoma, foram encontrados dois resultados importantes: (1) As vias GMM apresentaram-se agrupadas em três conjuntos com funções bem específicas de reparo, apoptose e ciclo celular. (2) Alguns genes participam de mais de uma via, como é o caso da proteína *p53* que forma uma ligação importante entre as vias de ciclo celular, apoptose e reparo. O desenvolvimento da rede de interação surgiu como precursor para a análise funcional da rede através de microarranjos de adenomas do colorretal. Os resultados destas análises descreveram o comportamento das vias de Ciclo celular e vias de Reparo em tecidos pré-cancerosos. No entanto, não foi possível explicar o funcionamento da via de Apoptose devido ao grande número de proteínas que estão espalhadas por um grande complexo funcional envolvido na morte celular programada. Em uma segunda proposta verificou-se o comportamento das vias GMM em amostras de tecidos da próstata. Os resultados apontam a diminuição da atividade das vias de manutenção do genoma sobre a exposição da

próstata ao Cádmio. Sendo assim, é mostrada a sensibilidade das vias Ontocancro na presença do Cádmio e após as 16 horas elas começam a retornar ao seu estado normal devido à eliminação do Cádmio pelo tecido da próstata.

Uma vez modelados os mecanismos envolvidos na manutenção do genoma, foi proposto o desenvolvimento do modelo quantitativo que descreve uma barreira anticâncer na evolução de tumores. Nesta fase da pesquisa foram avaliadas as alterações em subvias de manutenção do genoma em diferentes histologias, sugerindo apoio ao modelo qualitativo que descreve a barreira anticâncer (HALAZONETIS *et al.*, 2008). De acordo com os resultados apresentados na análise funcional da rede de proteínas verificou-se que a aplicação da metodologia quantitativa tem maior eficiência em vias de manutenção do genoma que contenham genes com funcionalidades específicas. Ou seja, as vias de Apoptose, Resposta ao dano do DNA (DDR) ou Ciclo celular serão melhor analisadas em desmembramentos de vias em subvias com funções mais específicas. Para complementar a análise estatística aplicada as subvias foram calculados os valores de *fold change* e escolhidos os *TOP-5* genes mais alterados em cada subvia.

O modelo proposto apresenta resultados que descrevem a barreira anticâncer presente em tecidos de adenoma do colorretal, adrenal, pâncreas e tireóide e descreve também como as subvias de manutenção do genoma são perdidas após a ativação da barreira em tumores malignos. Portanto, os resultados apresentados proporcionaram a construção de um modelo que complementa os estudos qualitativos relacionados a evolução tumoral através das análises estatísticas que uniram em um único modelo as vias, subvias e genes responsáveis por ativar um conjunto de mecanismos que impede a evolução para o câncer. O modelo mostra ainda o gene *CFLAR* alterado em todos os adenomas e sua alteração sugere esse gene como um marcador de adenomas.

Sendo assim, o modelo baseado na ativação de uma barreira anticâncer poderá contribuir para a terapia do câncer e ajudar no combate de neoplasias que afetam os tecidos glandulares através do desenvolvimento de drogas que permitem a ativação da barreira em câncer.

Como perspectiva futura é sugerido introduzir amostras de expressão obtidas por RNASeq para determinar os níveis de expressão gênica em vias e genes envolvidos na manutenção do genoma e ainda verificar os mecanismos que estão envolvidos na destruição da barreira anticâncer.

## 6. REFERÊNCIAS

- AFFYMETRIX GENECHIP. Statistical Algorithms Description Document. 2002
- AHN, A. C.; TEWARI, M.; POON, C.; PHILLIPS, R. S. The Clinical Applications of a Systems Approach, **Plos Medicine**, v. 3, p. 0956-60, 2006
- AITTOKALLIO, T.; SCHOWIKOWSKI, B. Graph-based Methods for Analysing Networks in Cell Biology, **Briefings in Bioinformatics**, v.7, n.3, p.243-55, 2006
- ALBERTI, L.; CARNITI, C.; MIRANDA, C.; ROCCATO, E.; PIEROTTI, M.A. RET and NTRK1 Proto-Oncogenes in Human Diseases. **Journal of Cellular Physiology**, v.195, p.168–86, 2003
- ALBERTS, B.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. **Molecular Biology of the Cell**, 4 ed, New York, Garland Science. 2002
- AUBERT, G.; LANSDORP, P.M. Telomeres and Aging. **Physiological Reviews**, v.88, p.557–79, 2008
- AUGENLICHT, L. H.; KOBRIN, D. Cloning and Screening of Sequences Expressed in a Mouse Colon Tumor, **Cancer Research**, v.42, p.1066-93, 1982
- AUGENLICHT, L. H.; WAHRMAN, M. Z.; HALSEY, H.; ANDERSON, L.; TAYLOR, J.; LIPKIN, M. Expression of Cloned Sequences in Biopsies of Human Colonic Tissue and in Colonic Carcinoma Cells Induced to Differentiate in Vitro, **Cancer Research**, v.47, p.6017-21, 1987
- BARTKOVA, J.; REZAEI, N.; LIONTOS, M.; KARAKAIDOS, P.; KLETSAS, D.; TAKAOKA, M.; NAKAGAWA, H.; TORT, F.; FUGGER, K.; GORGOULIS, V. G. Oncogene-Induced Senescence is Part of the Tumorigenesis Barrier Imposed by DNA Damage Checkpoints, **Nature**, v.444, p.633-37, 2006
- BECKER, J. D.; FEIJO, J. A. Profiling Genomes With Oligonucleotide Arrays, Affymetrix Core Facility, **Boletim de Biotecnologia**, v.24, p.2-6, 2003
- BEST, S. M. Viral Subversion of Apoptotic Enzymes: Escape from Death Row, **Annual Review of Microbiology**, v.62, p.171-92, 2008
- BITON, S.; ASHKENAZI A. NEMO and RIP1 Control Cell Fate in Response to Extensive DNA Damage via TNF- $\alpha$  Feedforward Signaling. **Cell**. v.145, p.92-103, 2011
- BOHGAKI, T.; BOHGAKI, M.; HAKEM, R. DNA Double-Strand Break Signaling and Human Disorders, **Bio Med Central**, v.1, n.15, p.1-14, 2010
- BONNER, W.M.; REDON, C.E.; DICKEY, J.S.; NAKAMURA, A.J.; SEDELNIKOVA, O.A.; SOLIER, S., POMMIER, Y.  $\gamma$ H2AX and Cancer. **Nature Reviews Cancer**, v.8, p.957-67, 2008.

BOTS, M.; MEDEMA, J.P. Granzymes at a Glance. **Journal of Cell Science**. v.119, p.5011-14, 2006.

BOXUS, M.; WILLEMS, L. How the DNA Damage Response Determines the Fate of HTLV-1 Tax-Expressing cells. **Retrovirology**, v.9, p.2-4, 2012

CALZONE, L.; GELAY, A.; ZINOVYEV, A.; RADVANYI, F.; BARILLOT, E.; A Comprehensive Modular Map of Molecular Interactions in RB/E2F Pathway. **Molecular Systems Biology**, v.173, p.1-12, 2008

CARLSON, M. R. J.; ZHANG, B.; FANG, Z.; MISCHER, P. S.; HORVATH, S.; NELSON, S.F. Gene Connectivity, Function, and Sequence Conservation: Predictions from Modular Yeast Co-Expression Networks, **BMC Genomics**, v.7, n.40, p.1-15, 2006

CASTRO, M. A. A.; MOMBACH, J. C. M.; ALMEIDA, R. M. C.; MOREIRA, J. C. F. Impaired Expression of NER Gene Network in Sporadic Solid Tumors, **Nucleic Acids Research**, v.35, p.1859-67, 2007

CASTRO, M. A.; FILHO, J. L.; DALMOLIN R. J.; SINIGAGLIA M.; MOREIRA J. C.; MOMBACH J. C., de ALMEIRA R. M. ViaComplex: software for landscape analysis of gene expression networks in genomic context, **Bioinformatics** 2009, v.25, n.11, p.1468-69, 2009

CHEN, L.; PARK, S.M.; TURNER, J.R.; PETER, M.E. Cell Death in the Colonic Epithelium During Inflammatory Bowel Diseases: CD95/Fas and Beyond. **Inflammatory Bowel Disease**. v.16, p.1071 - 76, 2010

COLLADO, M.; BLASCO, M. A.; SERRANO, M. Cellular Senescence in Cancer and Aging, **Cell**, v.130, p.223-33, 2007

COPPÉ, J.P.; DESPREZ, P.Y.; KRTOLICA, A.; CAMPISI, J.; The Senescence-Associated Secretory Phenotype: The Dark Side of Tumor Suppression, **Annual Review of Pathology: Mechanisms of Disease**. v.5 p.99-118, 2010

DANESE, S.; FIOCCHI, C. Ulcerative Colitis. **New England Journal of Medicine**. v.365, p.1713-25, 2011

DERISI, J.; PENLAND, L.; BROWN P. O.; BITTNER, M. L.; MELTZER, P. S.; RAY, M.; CHEN, Y.; SU, Y. A.; TRENT, J.M. Use of a cDNA Microarray to Analyse Gene Expression Patterns in Human Cancer, **Nature**, v.14, p. 457-460, 1996

DEY, A.; TERGAONKAR, V.; LANE, D. P. Double-Edged Swords as Cancer Therapeutics: Simultaneously Targeting P53 and NFKB Pathways, **Nature** v.7, p.1031-40, 2008

DI MOLA, F.F.; FRIESS, F.; ZHU, W.Z.; KOLIOPANOS, A.; BLEY, T.; DI SEBASTIANO, P.; INNOCENTI, P.; ZIMMERMANN, A.; BÜCHLER, M.W. Nerve

Growth Factor and TRK High Affinity Receptor (TRKA) Gene Expression in Inflammatory Bowel Disease. **Gut**, v.46, p.670–78. 2000

DONEHOWER, L.A. Using Mice to Examine P53 Functions in Cancer, Aging, and Longevity. **Cold Spring Harb Perspect Biol**, v.4, p.1-17, 2009

DUNCAN, J.A.; REEVES, J.R.; COOKE, T.G. BRCA1 and BRCA2 Proteins: Roles in Health and Disease. **Clinical Molecular Pathology**. v.51. p.237–47, 1998

EDGAR, R.; DOMRACHEV, M.; LASH, A. E. Gene Expression Omnibus: NCBI Gene Expression and Hybridization Array Data Repository, **Nucleic Acids Research**, v.30, n.1, p. 207-210, 2002

EISENBERG, D.; MARCOTTE, E. M.; XENARIOS, I.; YEATES T. O. Protein Function in the Post-Genomic Era, **Nature**, v.405, p.823-26, 2000

ELMORE, S. Apoptosis: A Review of Programmed Cell Death. **Toxicologic Pathology**. v.35, p.495 – 516, 2007

ENRIGHT, A. J.; ILIOPOULOS, I.; KYRPIDES, N. C.; OUZOUNIS, C. A. Protein Interactionmaps for Complete Genomes Basedon Genefusion Events, **Nature**, v.402, p.86-90, 1999

FALSCHLEHMER, C.; EMMERICH, C.H.; GERLACH, B.; WALCZAK, H. TRAIL Signaling: Decisions Between Life and Death. **The International Journal of Biochemistry & Cell Biology**, v.39, p.1462–75, 2007

FAN, Z.; ZHANG, Q. Molecular Mechanisms of Lymphocyte-Mediated Cytotoxicity. **Cellular & Molecular Immunology**, v.2, p.259 – 64, 2005

FANNING, E.; KLIMOVICH, V.; NAGER, A.R. A Dynamic Model for Replication Protein A (RPA) Function in DNA Processing Pathways. **Nucleic Acids Research**, v.15, p.4126-37, 2006

FERGUSON, D. O.; SEKIGUCHI, J. M.; CHANG, S.; FRANK, K. M.; GAO, Y.; DEPINHO, R. A.; ALT, F. W. The Nonhomologous End-Joining Pathway of DNA Repair is Required for Genomic Stability and the Suppression of Translocations, **PNAS**, v.97, n.12, p.6630-33, 2000

FISHEL, R.; WILSON, T.; MutS homologs in Mammalian Cells, **Current Opinion in Genetics & Development**. v.7(1), p.105-13, 1997

GENTLEMAN, R.; CAREY V. J.; G. V.; BATES, D.;BOLSTAD, B.; DETTLING, M.; HOTHORN, T.; HUBER, W.; IACUS, S.; IRIZARRY, R.;LEISCH, F.;LI, C.; MAECHLER, M.; ROSSINI, A. J.; SAWITZKI, G.; SMITH, C.; SMYTH, G.; TIERNEY, L.; YANG, J. Y. H.; ZHANG, J. Bioconductor: Open Software Development for Computational Biology and Bioinformatics, **Open Access**, v.5, p. 1-16, 2004

GEO, GENE EXPRESSION OMNIBUS, 11 de Maio de 2011, Disponível em: <<http://www.ncbi.nlm.nih.gov/geo>> , Acesso em 17 de Abr. de 2012

GYRD-HANSEN, M.; MEIER, P. IAPs: From Caspase Inhibitors to Modulators of NFkB, Inflammation and Cancer. **Nature Reviews Cancer**, v.10, p.561 – 74, 2010

GRIVICICH, I.; REGNER, A.; ROCHA, A. B. Morte Celular por Apoptose, **Revista Brasileira de Cancerologia**, v.53, n.3, p.335-43, 2007

GOHLMANN, H.; TALLOEN, W. **Gene Expression Studies Using Affymetrix Microarrays**, New York, CRC Press, 2009

GOPING, I.S.; BARRY, M.; LISTON, P.; SAWCHUK, T.; CONSTANTINESCU, G.; MICHALAK, K.M.; SHOSTAK, I.; ROBERTS, D.L.; HUNTER, A.M.; KORNELUK, R.; BLEACKLEY, A.M. Granzyme B-Induced Apoptosis Requires Both Direct Caspase Activation and Relief of Caspase Inhibition, **Immunity**, v.18, 355-65, 2003.

GORGOULIS, V.G.; VASSILIOU, L. V. F.; KARAKAIDOS, P.; ZACHARATOS, P.; LEVY, B.; KLETSAS, D.; YONETA, A.; HERLYN, M.; KITTAS, C.; HALAZONETIS, T. D. Activation of the DNA Damage Checkpoint and Genomic Instability in Human Precancerous Lesions, **Nature**, v.434, p.907-913, 2005

GSE10927: Human adrenocortical carcinomas (33), adenomas (22), and normal adrenal cortex (10), on Affymetrix HG\_U133\_plus\_2 arrays, GEO, MI, USA, janeiro de 2009. Disponível em:

<<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE10927>>. Acesso em 04 de mai. 2012

GSE19650: Expression data from epithelial cells during the process of multistep pancreatic carcinogenesis, GEO, Tokyo, dezembro de 2010. Disponível em:

<<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE19650>> Acesso em 30 de abr. 2012

GSE27155: Human thyroid adenomas, carcinomas, and normals, GEO, MI, USA, fevereiro de 2011. Disponível em:

<<http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE27155>> Acesso em 05 de jun. 2012

GSE4183: Inflammation, adenoma and cancer: objective classification of colon biopsy specimens with gene expression signature, GEO, Berlin, outubro de 2007. Disponível em

<<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE4183>>. Acesso em: 19 de set. 2010

GSE8671: Transcriptome profile of human colorectal adenomas. GEO, Zurich, agosto de 2007. Disponível em:

<<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE8671>>. Acesso em: 20 jan. 2010

GSE9951: Transcriptome analyses in normal prostate epithelial cells following exposure to low-dose cadmium, GEO, Cincinnati, março de 2008. Disponível em

<<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE9951>>. Acesso em: 12 fev. 2010

GSE4488: Expression data from whole blood, GEO, Helsinki, março de 2006.  
Disponível em < <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE4488>>  
Acesso em: 06 jun. 2012

HALAZONETIS, T. D.; GORGOULIS, V. G.; BARTEK, J. An Oncogene-Induced DNA Damage Model for Cancer Development, **Science**, v.319 p.1352-55, 2008

HOEIJMAKERS, J. H. J. Genome Maintenance Mechanisms for Preventing Cancer, **Nature**, v.411, p.366-74, 2001

HOEIJMAKERS, J. H. J. Genome Maintenance Mechanisms are Critical for Preventing Cancer as Well as Other Aging-Associated Diseases, **Elsevier**, v.128, p.460-462, 2007

HOLDBROOK, N. J.; MARTIN, G. R.; LOCKSHIN, R. A. **Cellular Aging and Cell Deth**, 16 ed, New York, Modern Cell Biology, 1996

HONG, S. J.; SUNG RIM, J.; IN YANG, H.; SHIK YIN, C.; KOH, H. G.; JANG, M.; KIM, C.; CHOE, B.; CHUNG, J. Bee Venom Induces Apoptosis Through Caspase-3 Activation in Synovial Fibroblasts of Patients with Rheumatoid Arthritis, **Toxicon**, v.46, p.39–45, 2005

HOOD, L.; HEATH, J. R.; PHELPS, M. E.; LIN, B. Systems Biology and New Technologies Enable Predictive and Preventative Medicine, **Science**, v.306, p. 640-43, 2004

HUBER, W.; HEYDEBRECK, A.V.; VINGRON, M. Handbook of Statistical Genetics. 3 ed, John Wiley & Sons, Cap. Analysis of Microarray Gene Expression Data, p. 162-187, 2003

IRIZARRY, R. A. Exploration, Normalization, and Summaries of High Density Oligonucleotide Array Probe Level Data, **Biostatistics**, v.4, p.249-64, 2003

JEFFORD, C. E.; IRMINGER-FINGER, I.; Mechanisms of Chromosome Instability in Cancers, **Elsevier Ireland**, v. 59, p.1-14, 2006

JENSEN, L. J.; KUHN, M.; STARK, M.; CHAFFRON, S.; CREEVEY, C.; MULLER, J.; DOERKS, T.; JULIEN, P.; ROTH, A.; SIMONOVIC, M.; BORK, P.; MERING, C. V. STRING 8 - A Global View on Proteins and Their Functional Interactions in 630 Organisms, **Nucleic Acids Research**, v.37, p.412-16, 2009

JI-HOON; LEE, J. H.; GOODARZI, A. A.; JEGGO, P. A.; PAULL, T. T. 53BP1 Promotes ATM Activity Through Direct Interactions with the MRN Complex, **The Embo Journal**. v.29, p.574-85, 2010

JING, G.; YUAN, K.; LIANG, Q.; SUN, Y.; MAO, X.; MCDONALD, J. M.; CHEAN, Y. Reduced CaM/FLIP Binding by a Single Point Mutation in c-FLIPL Modulates Fas-Mediated Apoptosis and Decreases Tumorigenesis, **Laboratory Investigation** v.92. p.82–90, 2012

KASTAN, M. B.; BARTEK, J. Cell-Cycle Checkpoints And Cancer, **Nature**, v.432, p.316-23, 2004

KAWAI, T.; NOMURA, F.; HOSHINO, K.; COPELAND, N. G.; GILBERT, D. G.; JENKINS, N. A.; AKIRA, S. Death-Associated Protein Kinase 2 is a New Calcium/Calmodulin-Dependent Protein Kinase that Signals Apoptosis Through Its Catalytic Activity, **Oncogene** v.18, p.3471-80, 1998

KAWAMURA, K.; KAWAMURA, N.; KUMAGAI, J.; FUKUDA, J.; TANAKA, T. Tumor Necrosis Factor Regulation of Apoptosis in Mouse Preimplantation Embryos and Its Antagonism by Transforming Growth Factor Alpha/Phosphatidylinositol 3-Kinase Signaling System. **Biology of Reproduction** v.76, p.611-18, 2007

KAYE, K. M.; DEVERGNEI, O.; HARADA, J. N.; IZUMIT, K. M.; YALAMANCHILIT, R.; KIEFFT, E.; MOSIALOST, L. Tumor Necrosis Factor Receptor Associated Factor 2 is a Mediator of NFKCB Activation by Latent Infection Membrane Protein 1, the Epstein-Barr Virus Transforming Protein, **Medical Sciences**, v.93, p.11085-90, 1996

KEEDWELL, E.; NARAYANAN, A. **Intelligent bioinformatics - The Application of Artificial Intelligence Techniques to Bioinformatics Problems**, England, Wiley, 2005

KIM, J.H.; KIM, S.S.; BYUN, S.W.; CHANG, Y.J.; KIM, J.S.; KIM, J.K.; CHO, H.J.; LIM, K.W.; JUNG, E.S. Double Strand Break of DNA in Gastric Adenoma and Adenocarcinoma. **Journal of Gastroenterology**, v.55, p.21-25, 2010

KRAUS, S.; ARBER, N.; Inflammation and Colorectal Câncer. **Current Opinion in Pharmacology**, v.9, p.405410, 2009

KRISTOFF, H. C. Cancer Biomarkers. In: **Bioinformatics analysis of gene networks involved in genomic stability and cancer**, ed. New York: Nova Biomedical, 2011. cap. 4, p.139-64

LABIB, K. How do CDC7 and Cyclin-Dependent Kinases Trigger the Initiation of Chromosome Replication in Eukaryotic Cells?. **Cancer Research**, v.24, p.1208–19, 2010

LEE, J. H.; GOODARZI, A. A.; JEGGO, P. A.; PAULL, T. T. 53BP1 Promotes ATM Activity Through Direct Interactions With the MRN Complex, **The Embo Journal**, v.29, p.574-85, 2010

LENHARD, R. E.; OSTEEN, R. T.; GANSLER, T. **Clinical Oncology**, Atlanta, American Cancer Society, 2001

LI, T.; ZHANG, J., ZHU, F., WEN, W., ZYKOVA, T.; LIU, K.; PENG, C.; MA, W.; SHI, G.; DONG, Z., BODE, A.M. P21-Activated Protein Kinase (PAK2)-Mediated c-Jun Phosphorylation at 5 Threonine Sites Promotes Cell Transformation. **Carcinogenesis**, v.32, p.659-666, 2010.

LIBRELOTTO, G. R.; MOMBACH, J. C. M.; SINIGAGLIA, M.; SIMÃO, E. M.; CABRAL, H. B.; CASTRO, M. A. A. **An Ontology to Integrate Transcriptomics and Interatomics Data involved in Gene Pathways of Genome Stability**. BSB 2009, LNBI 5676, pp. 164-167, 2009

LINDSAY, J.; ESPOSTI, M.D.; GILMORE, A.P. Bcl-2 Proteins and Mitochondria—Specificity in Membrane Targeting for Death. **Biochimica et Biophysica Acta**, v. 1813, p.532–39, 2011.

LOCKSLEY, R.; KILLEEN, N.; LEONARDO, M.J.; The TNF and TNF Receptor Superfamilies: Integrating Mammalian Biology. **Cell**, v.104, p.487-501, 2001

LODISH, H.; BERK, A.; MATSUDAIRA, P.; **Molecular Cell Biology**, 6.ed, Cambridge University Press, 2003

LOEB, L. A. A Mutator Phenotype in Cancer, **Cancer Research**, v.61, p.3230-39, 2001

MANSER, E.; CHONG, C.; ZHAO, Z.S.; LEUNG, T.; MICHAEL, G.; HALL, C.; LIM, L. Molecular Cloning of a New Member of the p21-Cdc42/Rac-activated Kinase (PAK) Family. **The Journal of Biological Chemistry**. v.270, p.5070–78, 1995

MARX, J.; Debate Surges Over the Origins of Genomic Defects in Cancer, **Science**, v.297, p.544-46, 2002

MAZZACCHI, F. Complexity in Biology, **Science & Society**, v.9, n.1, p. 10-14, 2008

MCCARTHY, D.J.; SMYTH, G.K.; Testing Significance Relative to a Fold-Change Threshold is a TREAT, **Bioinformatics**, v.25, n.6, p. 765-771, 2009

MERING, C. V.; HUYNEN, M. A.; JAEGGI, D.; SCHMIDT, S.; BORK, P.; SNEL, B. STRING: A Database of Predicted Functional Associations Between Proteins, **Nucleic Acids Research**, v.31, p.258-61, 2003

MERING, C. V.; JENSEN, L. J.; SNEL, B.; HOOPER, S. D.; KRUPP, M.; FOGLIERINI, M.; JOUFFRE, N.; HUYNEN, M. A.; BORK, P. STRING: known and predicted protein–protein associations, integrated and transferred across organisms, **Nucleic Acids Research**, v.33, p.433-37, 2005

MERING, C. V.; JENSEN, L. J.; KUHN, M.; CHAFFRON, S.; DOERKS, T.; KRUGER, B.; SNEL, B.; BORK, P. STRING 7 - Recent Developments in the Integration and Prediction of Protein Interactions, **Nucleic Acids Research**, v.35, p.358-62, 2007

MEMORIAL UNIVERSITY. MUN, Canadá, 2011. Disponível em: <<http://www.mun.ca/biology/desmid/brian/BIOL2060/BIOL2060-21/2110.jpg>>. Acesso em: 08 de novembro de 2011

MICHEAU, O.; THOME, M.; SCHNEIDER, P.; HOLLER, N.; TSCHOPP, J.; NICHOLSON, D.W.; BRIAND, C.; GRUTTER, M. G. The Long Form of FLIP Is an

Activator of Caspase-8 at the FAS Death-inducing Signaling Complex. **The Journal of Biological Chemistry**, v.277, p.45162-171, 2002

MOMBACH, J. C. M.; CASTRO, M. A. A.; MOREIRA, J. C. F.; ALMEIDA, R. M. C. On the Absence of Mutations in Nucleotide Excision Repair Genes in Sporadic Solid Tumors. **Genetics and Molecular Research**. v.7, n.1, p.152-60, 2008

PALSSON, B. O. **Systems Biology, Properties of Reconstructed Networks**, New York: Cambridge, 2006

PAPIN, J. A.; HUNTER, T.; PALSSON, B. O.; SUBRAMANIAM, S. Reconstruction of cellular signaling networks and analysis of their properties, **Nature Reviews**, v.6, p.99-111, 2005

POVEY, S.; LOVERING, R.; BRUFORD, E.; WRIGHT, M.; LUSH, M.; WAIN, H. The HUGO gene nomenclature committee (HGNC), **Human Genetics**, v.109, n.6, p.678-80, 2001

QIAO, L.; GRIDER, J.R. Colitis Elicits Differential Changes in the Expression Levels of Receptor Tyrosine Kinase TrkA and TrkB in Colonic Afferent Neurons: A Possible Involvement of Axonal Transport. **Pain**, v.151, p.117–27, 2010

RANGANATH, R. M.; NAGASHREE, N. R. Role of programmed cell death in development, **International Review of Cytology**, v.202, p.159-242, 2003

REINHARDT, H.C.; SCHUMACHER, B. The P53 Network: Cellular and Systemic DDR in Aging and Cancer, **Trends in Genetics**, v.28, p.128-36, 2012

REINHARDT, H.C.; YAFFE, M. B. Kinases that Control the Cell Cycle in Response to DNA Damage: CHK1, CHK2, and MK2. **Current Opinion in Cell Biology**, v.21, p.245–55, 2009

RISQUES, J. Ulcerative Colitis-Associated Colorectal Cancer Arises in a Field of Short Telomeres, Senescence, and Inflammation. **Cancer Research**, v.1, p.1669–79

RUIZ, C.; HUANG, W.; HEGI, M.E.; LANG, K.; HAMOU, M.F.; FLURI, E.; OAKELEY, E.J.; CHIQUET-EHRISMANN, R.; OREND, G.; Differential Gene Expression Analysis Reveals Activation of Growth Promoting Signaling Pathways by Tenascin-C, **Cancer Research**, v.64, p. 7377-83, 2004

SABATEL, H.; PIRLOT, C.; PIETTE, J.H.; Importance of PIKKs in NF $\kappa$ B Activation by Genotoxic Stress, **Biochemical Pharmacology**, v.82, p.1371-83, 2011.

SABATES-BELLVER, J.; VAN DER FILER, L.G.; DE PALO, M.; CATTANEO, E.; MAAKE, C.; REHRAURER, H.; LACZKO, E.; KUROWSKI, M.A.; MENIGATTI, M.; RANALLI, T.V.; GOMES, V.; PASTORELLI, A.; FAGGIANI, R.; ANTI, M.; JIRICNY, J.; CLEVERS, H.; MARRA, G. Transcriptome profile of human colorectal adenomas, **Mol Cancer Research**, v.5, p-1263-75, 2007

SALMINEN, A.; KAARNIRANTA, K. Control of P53 and NFkB Signaling by WIP1 and MIF: Role in Cellular Senescence and Organismal Aging. **Cellular Signaling**, v.23, p.747-52, 2011

SCHENA, M.; SHALON, D.; DAVIS, R. W.; BROWN, P. O. Quantitative Monitoring of Gene Expression Patterns With a Complementary DNA Microarray, **Science**, v.270, p.467-70, 1995

SCHULTZ, L.B.; CHEHAB, N.H.; MALIKZAY, A.; HALAZONETIS, T.D. P53 Binding Protein 1 (53BP1) is an Early Participant in the Cellular Response to DNA DSBs **The Journal of Cell Biology**. v.7, p.1381-90, 2000.

SCORRANO, L.; OAKES, S. A.; OPFERMAN, J. T.; CHENG, E. H.; SORCINELLI, M. D.; POZZAN, T.; KORSMEYER, S. J. BAX and BAK Regulation of Endoplasmic Reticulum Ca: A Control Point for Apoptosis, **Science**, v.300, p.135-139, 2003

SENGUPTA, S.; HARRIS, C.C. P53: Traffic Cop at the Crossroads of DNA Repair and Recombination. **Nature**, v.6, p.44-55, 2005

SHAN, W.; YANG, G.; LIU, j.; The Inflammatory Network: Bridging Senescent Stroma and Epithelial Tumorigenesis, **Front Biosci**, v.14, p.4044-57, 2010

SHANNON, C.E. Prediction and Entropy of Printed English. **The Bell System Technical Journal**, v.30 p.50–64, 1951

SHERR, C. J.; ROBERTS, J. M. CDK Inhibitors: Positive and Negative Regulators of G1, **Genes e Development**, v.13, p.1501–1512, 1999

SIMÃO, E. M.; CABRAL, H. C.; CASTRO, M. A. A.; SINIGAGLIA, M.; MOMBACH, J. C. M.; LIBRELOTTO, G. R. Modeling the Human Genome Maintenance Network. **Physica A**, v.389, p. 4188-94, 2010

SIMÃO, E. M.; BUGS, C.A.; CASTRO, M. A. A.; SINIGAGLIA, M.; M.; LIBRELOTTO, G. R.; MOMBACH, J. C. M. Anti-cancer Barrier in Tumor Evolution: Analysis of Differentially Expressed Genome Maintenance Pathways and Genes. **SUBMETIDO A PUBLICAÇÃO**

SORENSEN, C.S.; SYLJUASEN, R.G. Safeguarding Genome Integrity: The Checkpoint Kinases ATR, CHK1 and WEE1 Restrain CDK Activity During Normal DNA Replication. **Nucleic Acids Research**, v.40, p.477–86, 2012

STRACKER, T.H.; PETRINI, J.H.J. The MRE11 Complex: Starting From the Ends. **Nature**, v.12 p.90-103, 2011

STRUNNIKOV, A.V.; JESSBERGER, R. Structural Maintenance of Chromosomes (SMC) Proteins Conserved Molecular Properties for Multiple Biological Functions. **European Journal of Biochemistry**, v.263, p.6-13, 1999

STUDZINSKI, G.P.; WANG, X.; DANILENKO, M. DNA Damage Response: A Barrier or a Path to Tumor Progression? **Cancer Biology & Therapy**, v.9, p.253-55, 2010

UNM Biology Undergraduate, University of New Mexico, 2002. Disponível em:  
<Labshttp://biology.unm.edu/ccouncil/Biology\_124/Images/RNAtranslation.jpeg>,  
Acesso em: 08 de novembro de 2011

VAN GENT, D. C.; HOEIJMAKERS, J. H. J.; KANAAR, R. Chromosomal Stability and the DNA double-Stranded Break Connection. **Nature Reviews Genetics**, v.2, p.196-206, 2001

VICENT, T. L.; GATENBY, R. A. An Evolutionary Model for Initiation, Promotion, and Progression in Carcinogenesis, **International J of Oncology**, v.32, p.729-37, 2008

VOGELSTEIN, B.; KINZLER, K. W.; Cancer Genes and the Pathways they Control. **Nature. Medicine.**, v.10, p.789-99, 2004.

WATSON, J.; BAKER, T. A.; BELL, S. P., GANN, A.; LEVINE, M.; LOSICK, R.; CSHLP, I. **Molecular Biology of the Gene**, 5 ed, 2004

WATSON, J. The human genome project: past, present, and future, **Science**, v.248, n.4951, p.44-49, 1990

WEINBERG, R. A.; **The biology of cancer**, Garland Science, 2007

WESTERHOFF, H. V.; PALSSON, B. O. The Evolution of Molecular Biology Into Systems Biology, **Nature Biotechnology**, v.22, n.10, p.1249-52, 2004

WIKIBOOKS. **Schematic of the cell cycle**, 2011, Disponível em:  
<http://upload.wikimedia.org/wikipedia/commons/c/c4/Cell\_cycle.png>. Acesso em:  
08 de novembro de 2011

WOOD, R. D.; MITCHELL, M.; SGOUROS, J.; LINDAHL, T.; Human DNA repair genes, **Science**, v.291, p.1284-89, 2001

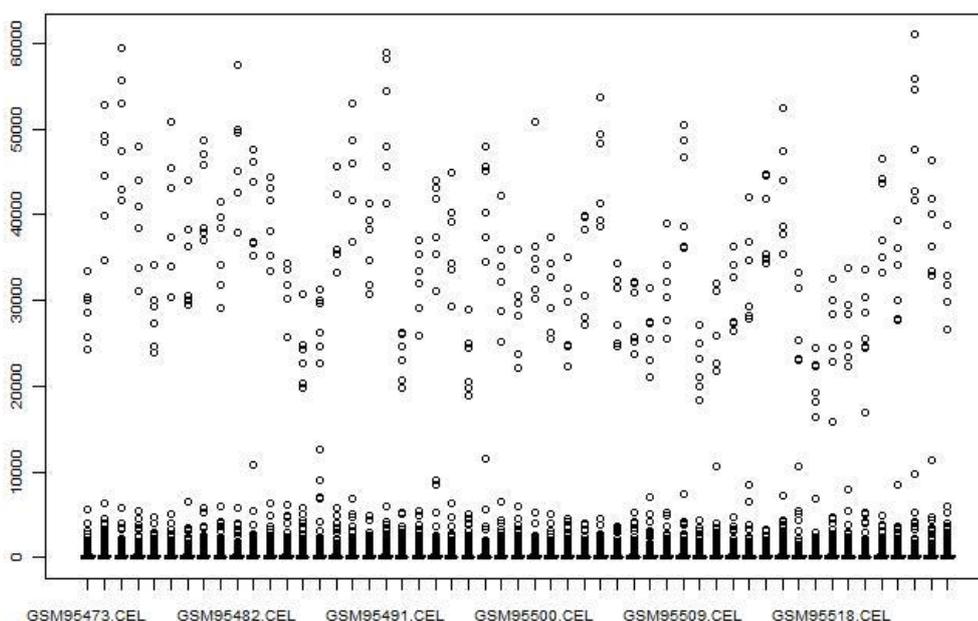
WU, Z.; IRIZARRY, R. A.; GENTLEMAN, R.; MURILLO, F. M.; SPENCER, F. A Model Based Background Adjustment for Oligonucleotide Expression Arrays, **Journal of the American Statistical Association**, v.99, p.909-17, 2004

YEH, W.; ITIE, L.; ELIA, A.J.; NG, M.; SHU, H.B.; WAKEHAM, A.; MIRTOSOS, C.; SUZUKI, N.; BONNARD, M.; GOEDEL, D.V.; MAK, T.W. Requirement for Casper (c-FLIP) in Regulation of Death Receptor–Induced Apoptosis and Embryonic Development. **Immunity**, v.12. p.633-42, 2000

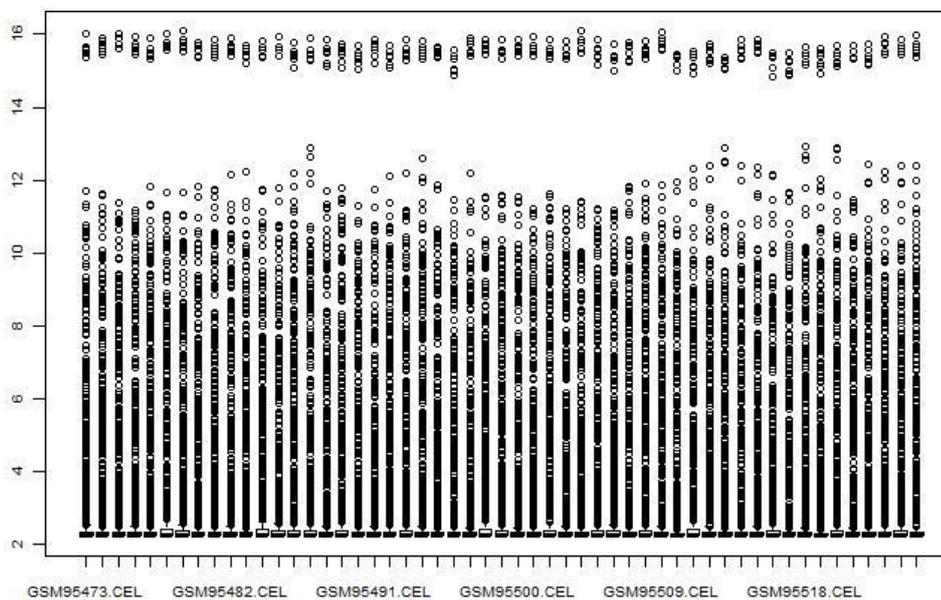
ZHIVOTOVSKY, B., KHOEMER, G., Apoptosis and Genomic Instability, **Molecular and Cellular Biology**, v.5, p.752-62, 2004.

## ANEXO 1 - Resultados das comparações entre as normalizações: MAS 5.0, GCRMA e RMA.

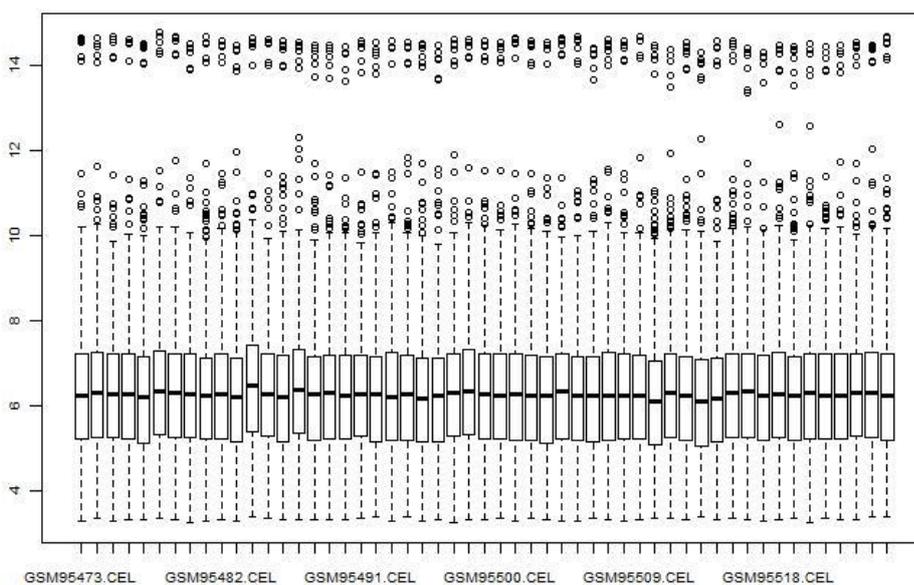
Neste Anexo, estão apresentados através dos gráficos de caixas os resultados das três normalizações (MAS 5.0, GCRMA e RMA) utilizadas nas amostras de colorretal, adrenal, pâncreas e tireóide do capítulo 4. Para definir a melhor normalização os gráficos de caixas que representam a expressão de cada amostra do tecido do colorretal indicam diferentes distribuições de normalização para cada um dos três casos. Enquanto as figuras A e B que representam as normalizações MAS 5.0 e GCRMA mostram uma maior dispersão em relação a média, a figura C que representa a normalização RMA indica uma normalização mais homogênea. O que é característica importante na análise de vias e subvias. As análises de normalização também foi aplicada aos tecidos adrenais, pâncreas e tireóide e os resultados seguiram o mesmo padrão da normalização RMA aplicada ao colorretal.



**Figura A** - Gráfico de caixas aplicado sobre as amostras do tecido do colorretal normalizadas com MAS 5.0 onde cada coluna representa uma das 53 amostra.



**Figura B** - Gráfico de caixas aplicado sobre as amostras do tecido do colorretal normalizadas com GCRMA onde cada coluna representa uma das 53 amostra.



**Figura C** - Gráfico de caixas aplicado sobre as amostras do tecido do colorretal normalizadas com RMA onde cada coluna representa uma das 53 amostra.





		COLON	ADENOMA				CÂNCER			
SUBVIA	R	INFLAM.	COLON	ADRENAL	PÂNCREAS	TIREÓIDE	COLON	ADRENAL	PÂNCREAS	TIREÓIDE
Indução da apoptose pelas receptores de morte DR3/4/5	1									
	2					<i>TNFSF10</i>				
	3					<i>TNFRSF10B</i>				
	4									
	5									
Apoptose - HS	1	<i>BIRC3</i>		<i>FAS</i>		<i>CFLAR</i>				
	2	<i>NTRK1</i>		<i>CFLAR</i>						
	3	<i>TNFRSF10D</i>								
	4	<i>TNFRSF10B</i>								
	5	<i>BAX</i>				<i>CFLAR</i>				
Cascata de caspases em apoptose	1									
	2									
	3		<i>BIRC3</i>							
	4									
	5		<i>BAX</i>							
Regulação da apoptose	1			<i>PAK2</i>		<i>DAPK2</i>		<i>UNC5A</i>		
	2		<i>PSMC2</i>	<i>PSMC2</i>		<i>PAK2</i>		<i>DAPK1</i>		<i>DAPK1</i>
	3									
	4		<i>PAK2</i>						<i>UNC5A</i>	<i>DAPK1</i>
	5		<i>DAPK2</i>					<i>PAK2</i>	<i>PAK2</i>	<i>PAK2</i>
Fase de execução da apoptose	1									
	2									
	3									
	4									
	5									
Via extrínseca de apoptose	1				<i>FAS</i>					
	2				<i>TNFSF10</i>					
	3									
	4				<i>CFLAR</i>					
	5									
Granzima A mediador de apoptose	1	<i>SET</i>								
	2	<i>HMGB2</i>								
	3	<i>APEX1</i>								
	4	<i>ANP32A</i>								
	5	<i>HMGB2</i>								
Senescência Replicativa	1	<i>CDK6</i>	<i>CDK6</i>					<i>TP53</i>		<i>H2AFX</i>
	2	<i>CDC25A</i>				<i>CDKN1A</i>	<i>CDK6</i>		<i>RPA4</i>	
	3	<i>CDK6</i>		<i>CDK6</i>	<i>BCL2</i>	<i>MDM2</i>		<i>RPA4</i>	<i>ATR</i>	<i>ATR</i>
	4	<i>CCNE2</i>	<i>CDK6</i>				<i>H2AFX</i>	<i>CDK6</i>		<i>TP53</i>
	5	<i>CDKN1A</i>	<i>CDKN1A</i>		<i>MDM2</i>			<i>CCND1</i>		<i>CCND1</i>
Ativação da NFκB	1			<i>XIAP</i>		<i>NEMO</i>		<i>NRN</i>		<i>NBN</i>
	2			<i>ERC1</i>		<i>XIAP</i>		<i>NEMO</i>	<i>XIAP</i>	
	3			<i>NEMO</i>		<i>ERC1</i>		<i>PIAS4</i>		<i>PIAS4</i>
	4			<i>CHUK</i>		<i>CHUK</i>		<i>XIAP</i>	<i>NEMO</i>	
	5			<i>PIAS4</i>		<i>PIAS4</i>		<i>XIAP</i>	<i>NBN</i>	<i>NEMO</i>

## CONSENSO DOS GENES DA VIA DE CICLO CELULAR

		COLON	ADENOMA				CÂNCER			
SUBVIA	R	INFLAM.	COLON	ADRENAL	PÂNCREAS	TIREÓIDE	COLON	ADRENAL	PÂNCREAS	TIREÓIDE
Ciclo celular: pontos de controle G1/S	1				CDC2			TP53		
	2			TP53						
	3			TP53						
	4									
	5									
Fase S	1				PSMC2			RFC4		RFC4
	2									
	3									
	4		CCNE2							
	5		WEE1							
Ciclina e regulação do ciclo celular	1				CDC2					
	2									
	3									
	4									
	5									
Pontos de controle G1/S do dano ao DNA	1									TP53
	2				CCNE2					
	3									
	4									
	5					TP53				
Postos de controle G2/M	1					WEE1		RFC4		
	2									
	3				CDC7	WEE1				
	4									
	5					CDC2				
Fase mitótica G2-G2/M	1				CDC2					
	2									
	3									
	4		WEE1		CLASP1					
	5		PAFAH1B1							
Fase mitótica M-M/G1	1		CDC7	MAD2L1	CDC7	CLASP1		RPA4	RPA4	
	2									
	3		SMC3	PAFAH1B1						
	4		PAFAH1B1			MAD2L1				
	5		PSMC2	SMC3	MAD2L1					
Pontos de controle do fuso mitótico	1				MAD2L1					
	2									
	3									
	4									
	5									
RB/E2F	1				CDC2	WEE1	EP300	APC	EP300	CREBBP
	2						CREBBP	TP53	EED	
	3		WEE1		EED			EED	APC	
	4					EP300				
	5		EED		EP300					

### **ANEXO 3 - Discussão estendida da função dos genes alterados no consenso da figura 13.**

A discussão que segue mostra as funções de alguns genes relevantes para a ativação da barreira anticâncer em tecidos de adenoma.

De acordo com o trabalho de Halazonetis e colaboradores (2008), os autores sugerem que os oncogenes induzem um estresse replicativo no DNA que leva a formação de DSB e conseqüentemente ativa o gene *TP53* que induz a ativação das vias de Apoptose e Senescência. Observando o consenso verificou-se a existência de genes alterados responsáveis por identificar e reparar DSB em tecidos de adenoma. Na figura 13a são mostrados os genes alterados no consenso para a via DDR. Muitos dos genes do consenso (*MRE11A*, *TP53BP1*, *BRCA2*, *LIG4*, *ATM* e *XRCC4*) pertencem a subvias de Reparo das quebras em fitas duplas e os genes *XRCC4*, *LIG4* e *MRE11A* também participam da subvia de Reparo por terminação não homologa. Destes 6 genes que estão envolvidos em DSB, o *MRE11A*, *TP53BP1* e *ATM* são proteínas sensoras das alterações em fitas duplas e os outros três genes têm papel como efetores do reparo das DSB realizado pelas subvias de Recombinação homologa e Reparo por terminação não homologa (STRACKER *et al.*, 2011; DUNCAN *et al.*, 1998). Outro gene que merece atenção é o *ATR* que aparece ativo na subvia Anemia de Fanconi e atua como sensor de danos do DNA e ajuda na ativação dos pontos de verificação do DNA (STRACKER *et al.*, 2011). Outros genes do consenso (*MSH2*, *MSH3* e *MSH6*) pertencem a subvia de Reparo por excisão de bases, ou seja, estão envolvidos no reparo das quebras em fitas simples do DNA (DSS). O gene *MSH6* que está alterado na subvia Reparo por excisão (BER) e em Reparo das bases mal pareadas (MMR), o gene *MSH3* que está alterado na subvia Reparo por excisão de bases (BER) e o *MSH2* desempenham papéis importantes na recombinação dos pares de bases trocados e são fundamentais para a manutenção do genoma (FISHEL *et al.*, 1997).

No consenso da via DDR em câncer (figura 13b) não foram identificados genes envolvidos em DSS e 4 genes encontrados estão ligados as subvias de DSB: *H2AFX*, *XRCC4*, *XRCC5* e *RAD52*. Os genes *XRCC4* e *XRCC5* estão envolvidos no Reparo por terminações não homólogas e são ativadas quando há reparo das DSB. O *XRCC4* se liga ao *LIG4* formando um complexo que é responsável por ativar as

etapas de ligação que envolve o reparo pela subvia NHEJ (SENGULTA *et al.*, 2005). A histona *H2AFX* é um sensor de danos e surge muito expressa em câncer. Ele é ativo imediatamente após a barreira e monitora e ajuda na deleção das DSB (BONNER *et al.*, 2008).

Modelos anteriores descrevem a ativação da apoptose e da senescência como uma barreira que impede o desenvolvimento do câncer em tecidos pré-cancerosos. No consenso da figura 13c para a via de apoptose é mostrado que os genes da análise estão associados a regulação e ativação da apoptose através de algumas subvias envolvendo os fatores de necrose tumoral (*TNF*) e pela subvia de Senescência replicativa. Associado as subvias *TNF* foram localizados alterados os seguintes genes: Na subvia de Sinalização *TNFR1* os genes, *CASP8*, *CASP2* e *RIPK1*, na subvia de Sinalização *TNFR2* os genes *TANK*, *RIPK1* e *TNFRSF1B*. A ativação da subvia de Sinalização *TNFR1* é mediada pelo gene do consenso *RIPK1* (BITON *et al.*, 2011) e atua como reguladora da cascata de caspases pela ativação do gene *CASP8* que aparece alterada nos adenomas do colorretal e da glândula adrenal, enquanto que em adenomas do pâncreas e tireóide a ativação da cascata de caspases parece ser ativada a partir do gene *CASP2* (BEST, 2008). No processo de ativação da apoptose o gene *MAP2K4* sinaliza a fragmentação do DNA e o gene *TANK* é um receptor da proteína *tnfr2* que é induzida pelo *RIPK1* que controla a ativação do gene *MAP3K1* (KAYE, 1996). O gene *TNFRSF1B* atua na subvia de Sinalização *TNFR2* e utiliza o *TANK* como supressor apoptótico de *BIRC3* que participa das subvias da Caspase cascade em apoptose e da subvia de Sinalização *TNFR1*. O *FAS* pertence a família Fator de Necrose Tumoral, ela faz parte do complexo de sinalização de morte induzida (DISC) que ativa o gene *CASP8* para mediar a apoptose (MICHEAU *et al.*, 2002). Dois genes do consenso pertencem a subvia de Regulação da apoptose: *DAPK2* e *PAK2*. O *DAPK2* está associada com a morte celular pela quinase 2 e atua como reguladora positiva da apoptose que atuando em conjunto com o *PAK2* na regulação dos eventos apoptóticos (KAWAI *et al.*, 1998]. Na subvia de Senescência replicativa foram encontrados três genes do consenso: *CDK6*, *CDKN1A* e *MDM2*. *CDK6* e *CDKN1A* são ciclinas quinases que atuam com a *p53* na parada do ciclo celular em resposta ao dano do DNA (REINHARDT *et al.*, 2012) e o gene *MDM2* têm função de inibir a *p53* em tecidos normais (DEY *et al.*, 2008). Ainda na figura 13c os genes *NEMO* e *CHUK* pertencem a subvia de Ativação do *NFKB* que aparece muito alterada em tecidos de adenoma

e em câncer como já foi mencionado na seção anterior. Os genes alterados no consenso têm função de regulação e influenciam na ativação da subvia Ativação do *NFKB*, seja através da liberação de neurotransmissores pelo gene *ERC1*, pela modulação do *NFKB* que é dada pelo gene *NEMO* ou pela regulação de outras vias celulares, como o caso da via *p53* que ocorre pela ativação do *PIAS4* (SABATEL *et al.*, 2011). O consenso mostra o gene *CFLAR* alterado entre todos os adenomas. Este gene atua na subvia Extrínseca de apoptose como regulador do gene *CASP8* que é responsável por ativar a morte induzida pelo *TNFRSF1A* (YEH *et al.*, 2000; KAWAMURA *et al.*, 2007). Algumas evidências sugerem que o *CFLAR* pode facilitar a apoptose formando um heterodímero com o *CASP8*, que ajuda na clivagem de pró-caspases 8 (JING *et al.*, 2012).

Seguindo pela analogia de alterações da subvia Ativação do *NFKB* foram encontrados três genes (*NBN*, *PIAS4* e *NEMO*) desta subvia alterados em câncer, conforme mostra a figura 13d. Entre estes três genes destaca-se o *NBN* que participa do complexo *MRN* e tem um papel na resposta aos danos do DNA e na manutenção da integridade cromossômica (JI-HOON *et al.*, 2010). Além destes genes foram encontrados os genes *H2AFX*, *PAK2* e *CDK6* também alterações em câncer. Na figura 13d os genes do consenso *PAK2* e *CDK6* estão envolvidos na subvia de Regulação da Apoptose e o gene *H2AFX* que participa das subvias de Senescência replicativa e Ativação do *NFKB* é um sensor de quebras em fitas duplas do DNA (KIM *et al.*, 2010). Em câncer o *NEMO* controla os mecanismos de resposta a altos níveis de danos do DNA e é ativado pelo *H2AFX* e pelo *ATM* (BITON *et al.*, 2011).

Na figura 13e estão apresentados os resultados do consenso dos genes na via de ciclo celular em adenoma. Neste diagrama a subvia mais alterada entre os adenomas é a Fase mitótica M-M/G1 com três genes alterados: *SMC3*, *CDC7* e *WEE1*. O gene *SMC3* está envolvido na manutenção estrutural dos cromossomos e na montagem dos pólos do fuso durante a mitose (STRUNNIKOV *et al.*, 1999). O *CDC7* ou ciclo de divisão celular, homolog 7 está mutuamente alterado nas subvias da Fase mitótica M-M/G1 e nos Postos de controle G2/M. *CDC7* é uma quinase que ativa proteínas envolvidas no ciclo celular, em tumores essa proteína provoca a morte celular por apoptose independente do *TP53* (LABIB, 2010) e estudos recentes mostram que o *WEE1* inibe as ciclinas quinases na fase G2 (SORENSEN *et al.*, 2012).

Como a característica mais comum do câncer é a proliferação aberrante (CASTRO et al., 2007), observa-se a ativação de algumas proteínas importantes para a manutenção do ciclo celular em câncer, em particular percebe-se a ativação da subvia *RB/E2F* que regula o ciclo celular na transição de G1 para S cuja desregulação já foi observada em diversos tipos de câncer (CALZONE et al., 2008). Nesta subvia foram encontrados 4 genes alterados em consenso: *EP300*, *CREBBP*, *APC* e *TP53*. O gene *EP300* e *CREBBP* formam parte de um caminho de ligação cruzada (*cross talk*) formando um ponto crítico de ligação entre o *NFKB* e o gene *TP53*. Estudos mostram que o *cross talk* entre os genes sugerem ativação do *TP53* e inibição do *NFKB* oferecendo grande potencial para evitar a carcinogênese (DEY et al., 2008). Observando ainda o diagrama da figura 13f foram identificados mais dois genes alterados na subvia *RB/E2F*: *EED* e *APC*. O *EED* codifica proteínas multiméricas que estão envolvidas no estado de regulação de outros genes e o *APC* tem função regulatória e também é um supressor de tumor. Mutações do *APC* são conhecidas em casos de adenocarcinoma do colorretal (CALZONE et al., 2008). Os genes *RPA4* e *RFC4* aparecem expressos nas subvias Fase mitótica M-M/G1 e na Fase S respectivamente e tem função na regulação da replicação (FANNING et al., 2006).

## **ANEXO 4 - Discussão estendida da função dos *TOP-5* genes alterados nos tecidos de inflamação do colorretal.**

Neste Anexo, estão apresentadas as discussões biológicas dos genes analisados pela mudança de expressão (*fold change*) nas subvias: Senescência Replicativa, Apoptose - Homo sapiens e Granzima A mediador de apoptose.

Na análise do *fold change* da subvia de Senescência replicativa, os genes *CDK6*, *CDC25A* e *CDKN1A*, indicam a regulação do estado senescente em inflamação do colorretal a partir de *CDKN1A*. Na subvia Apoptose - Homo sapiens os genes *TNFRSF10D/B* codificam os receptores *TRAILR2* e *TRAILR4* cuja interação entre eles é capaz de inibir a apoptose a partir do ligante *TRAIL*, além de regular a inibição da apoptose através da subvia de Ativação do *NFKB* (FALSCHLEHNER *et al.*, 2007). O *BIRC3* codifica o gene *CIPA2* e também é responsável pela inibição da apoptose através da ativação do *NFKB*, além de contribuir para prevenir a ativação do *CASP8* através do *RIPK1* (GYRD-HANSEN *et al.*, 2010). Enquanto estes 3 genes atuam na inibição da apoptose, a proteína pró-apoptótica *bax* é mantida inativa por proteínas da família *Bcl2*. Contudo esta repressão diminui a partir da ativação o *BH3* (LINDSAY *et al.*, 2011). Como está sendo previsto nessa abordagem há ativação da apoptose em tecidos inflamados do colorretal, com isso ocorre ativação de alguns genes como o *TRAILR2* e *BAX* que indicam a regulação da morte celular programada através da via Intrínseca de apoptose. O gene *NTRK1* foi inicialmente relacionado com carcinomas de colorretal, ele é um oncogene que atua como receptor para a proteína envolvida no fator de crescimento de neurônios (*NGF*), cuja expressão está envolvida na regulação do crescimento, diferenciação e apoptose de neurônios nos sistema nervoso central e periférico (ALBERTI *et al.*, 2003). A expressão gênica do *NGF* e *NTRK1* aparecem aumentadas em pacientes sofrendo UC e DC e com intestinos inflamados, o que sugere a ativação da subvia de Apoptose - Homo sapiens contendo estes elementos. Células imunes presentes na parede intestinal, expressam altos níveis de *NGF* e *NTRK1*, induzindo a ativação de fibras nervosas sensoriais, que por sua vez atuam na manutenção da integridade da mucosa intestinal (DI MOLA *et al.*, 2000, QIAO *et al.*, 2010). Enquanto que a atividade da subvia Apoptose - Homo sapiens indica a ativação das subvias Extrínseca e Intrínseca de apoptose a partir

da expressão de *TNFRSF10D/B* e *BAX* respectivamente, uma terceira subvia indicada pela atividade da subvia Granzima A mediador de apoptose envolve a ativação de células T citotóxicas. A principal função destas células é regular o início da apoptose em células-alvo através da exposição destas células a proteínas tóxicas, incluindo perforina, cuja principal função é formar poros na membrana plasmática com o objetivo de auxiliar as granzimas liberadas pelas células T a entrar no citoplasma das células marcadas para remoção por apoptose de células tumorais (ELMORE, 2007; FAN *et al.*, 2005).

Os linfócitos T expressam de forma abundante as granzimas A e B que coordenam a apoptose por duas vias distintas, enquanto a atividade apoptótica da granzima A independe da ativação das caspases, a via envolvendo a granzima B envolve a mitocôndria e uma rápida ativação das caspases (BOTS *et al.*, 2006). A granzima A regula a morte celular programada através da fragmentação do DNA, de modo que a regulação deste processo envolve um complexo liberado pelo retículo endoplasmático formado pelas proteínas *Set*, *Ape1*, *pp32*, *Hmg2*, *Nm23-h1* e *Trex1*, cuja clivagem dos genes *HMG2*, *APEL* e *SET* pela granzima A, libera a fragmentação do DNA por *Nm23-h1* e *TREX1* (ELMORE, 2007; BOTS *et al.*, 2006). Na pesquisa pelos genes com *fold change* mais alterados foram localizados 4 genes na subvia Granzima mediadora de apoptose: *SET*, *HMGB2*, *APEX1* e *ANP32A*, indicando que na inflamação do colorretal ocorre também a ativação da apoptose coordenada pelas células T a partir da ativação da subvia contendo granzima A, o que reforça a atividade apoptótica de células presentes em inflamação do colorretal.

## ANEXO 5 - Tabelas do consenso entre os genes com *fold Change* menos alterados entre câncer *versus* normal nas subvias alteradas apenas em adenoma.

Neste anexo estão apresentados todos os genes que perderam sua posição no *ranking* e estão pouco ativos em câncer. No Anexo, o R' indicará a posição inversa do gene sendo a posição 1 o gene menos ativo e A é a ausência do gene entre os *TOP-5* genes menos ativos, indica que o gene está presente e não perdeu suas características em câncer.

### CONSENSO DOS GENES DA VIA DDR

		ADENOMA			
SUBVIA	R'	COLORRETAL	ADRENAL	PÂNCREAS	TIREÓIDE
Reparo por excisão de bases	1	<i>MSH6</i>			
	2				
	3				
	4	<i>MSH3</i>			<i>MSH3</i>
	5				<i>MSH6</i>
Via da anemia de Fanconi	1				
	2				
	3				
	4	<i>BRCA2</i>			
	5				
Recombinação homóloga	1				
	2				
	3				
	4	<i>MRE11A</i>			
	5	<i>BRCA2</i>			
Reparo de bases mal pareadas	1	<i>MSH6</i>			
	2				
	3				
	4				
	5				
Reparo por excisão de nucleotídeos	1				
	2				
	3				
	4	<i>LIG4</i>			
	5	<i>BRCA2</i>			
Reparo por terminação não homóloga	1			<i>XRCC4</i>	
	2				
	3				
	4			<i>LIG4</i>	<i>MRE11A</i>
	5				<i>XRCC4</i>
Reparo das quebras em fitas duplas no DNA	1	<i>BRCA2</i>			
	2	<i>LIG4</i>			
	3	<i>MRE11A</i>			
	4			<i>XRCC4</i>	
	5	<i>TP53BP1</i>		<i>TP53BP1</i>	<i>BRCA2</i>

## CONSENSO DOS GENES DA VIA DE APOPTOSE

		ADENOMA			
SUBVIA	R'	COLORRETAL	ADRENAL	PÂNCREAS	TIREÓIDE
Sinalização TNFR1	1				
	2				
	3		CASP8	CASP2 MAP2K4	
	4			BIRC3	
	5		MAP2K4	RIPK1	CASP2
Sinalização TNFR2	1	TNFRSF10B	TNFRSF10B		TNFRSF10B
	2				TANK
	3	RIPK1	TANK		
	4	TANK	RIPK1		
	5	MAP3K1	MAP3K1		
Sinalização do receptor TNF	1				
	2				
	3				
	4		CASP8		
	5				
Receptor de sinalização de morte	1			CFLAR	
	2				
	3	CFLAR			
	4			TNFSF10	
	5	CASP8		FAZ	
	A	TNFRSF10B			
Sinalização da apoptose em DDR	1				
	2				
	3				BCL2
	4				
	5				BAX
Indução da apoptose pelas receptores de morte DR3/4/5	1				
	2				
	3				TNFSF10
	4				
	5				
	A				TNFRSF10B
Apoptose - HS	1				
	2				
	3				
	4				
	5			FAZ	
	A			CFLAR	CFLAR
Cascata de caspases em apoptose	1	BAX			
	2				
	3				
	4				
	5	BIRC3			
Regulação da apoptose	1				
	2	PSMC2			
	3				
	4	PAK2	PSMC2		PAK2
	5	DAPK2	PAK2		DAPK2

		ADENOMA			
SUBVIA	R'	COLORRETAL	ADRENAL	PÂNCREAS	TIREÓIDE
Fase de execução da apoptose	1				
	2				
	3				
	4				
	5				
Via extrínseca de apoptose	1				
	2				
	3				
	4				
	5				
A				<i>CFLAR</i> <i>FAZ</i> <i>TNFSF10</i>	
Granzima A mediador de apoptose	1				
	2				
	3				
	4				
	5				
Senescência Replicativa	1				
	2				
	3				<i>MDM2</i>
	4				<i>BCL2</i>
	5				<i>MDM2</i>
A	<i>CDK6, CDKN1A</i>	<i>CDK6</i>		<i>CDKN1A</i>	
Ativação da NFkB	1				
	2				
	3				
	4				
	5				
A		<i>ERC1</i>			
		<i>PIAS4, NEMO, CHUK</i>		<i>CHUK</i> <i>PIAS4, NEMO, WIAP</i>	

## CONSENSO DOS GENES DA VIA DE CICLO CELULAR

		ADENOMA			
SUBVIA	R'	COLORRETAL	ADRENAL	PÂNCREAS	TIREÓIDE
Ciclo celular: pontos de controle G1/S	1				
	2				
	3				
	4				
	5				
Fase S	1				
	2				
	3				
	4				
	5				
	A	<i>WEE1</i>			
Ciclina e regulação do ciclo celular	1				
	2				
	3				
	4				
	5				
Pontos de controle G1/S do dano ao DNA	1				
	2				
	3				
	4				
	5				
Postos de controle G2/M	1				
	2				
	3				
	4				
	5			<i>CDC7</i>	
Fase mitótica G2-G2/M	1				
	2				
	3				
	4				
	5				
	A	<i>WEE1</i>			
Fase mitótica M-M/G1	1				
	2				
	3				
	4				
	5				
	A	<i>SMC3, CDC7</i>	<i>SMC3</i>		
Pontos de controle do fuso mitótico	1				
	2				
	3				
	4				
	5				
RB/E2F	1				
	2				
	3				
	4				
	5				
	A	<i>WEE1</i>			<i>WEE1</i>