

UNIVERSIDADE DE SÃO PAULO

**CONTRIBUIÇÕES PARA PROBLEMAS ATUAIS EM GENÉTICA HUMANA:
FUNÇÃO GÊNICA E HERANÇA EPIGENÉTICA.**

Lygia da Veiga Pereira Carramaschi

2004

Tese submetida ao Departamento de
Biologia do Instituto de Biociências, USP
para obtenção de título de Livre Docente.

Ao Fabio

AGRADECIMENTOS

Agradeço a todos aqueles que ao longo da minha vida, de alguma forma, sabendo ou não, contribuíram para a minha felicidade. Em especial, à minha tão querida família; à Professora Eliana Abdelhay, por ter me iniciado na carreira de pesquisadora; aos membros do Departamento de Biologia do Instituto de Biociências da USP; à Constância, que me acolheu com tanto carinho na difícil mudança para São Paulo; à Professora Ângela Morgante, grande mestra, pela amizade pelo exemplo profissional; aos meus alunos, que tomaram as bancadas do laboratório com muita competência me deixando confinada ao meu computador; aos meus queridos Professores Russos, Marina, Irina e Alexandre, que me ajudaram a realizar muitos sonhos e, espero, continuarão ajudando a realizar muitos outros; aos Professores José Antônio Visintin, José Eduardo Krieger e todos os parceiros nas aventuras da pesquisa biomédica; às agências de fomento à pesquisa, especialmente à FAPESP, que me fez trocar Ipanema pela Marginal Pinheiros; ao Fabio e à Gabriela, maiores prêmios que recebi por aquela troca; à vida social e cultural de São Paulo, que me enriquecem e ajudam a administrar as tensões da vida profissional.

Agradeço muito especialmente a todos os professores que generosamente se dedicam ao ensino e que, mais do que transmitir informações, despertam a curiosidade pela genética.

ÍNDICE

INTRODUÇÃO: Desafios da pesquisa biomédica na Era Pós-Genoma Humano.....	4
PARTE I - ESTUDO DE FUNÇÃO GÊNICA: FIBRILINA E A SÍNDROME DE MARFAN	8
INTRODUÇÃO.....	8
OBJETIVOS.....	17
CAPÍTULO 1 – <i>FBN1</i> : caracterização do mRNA e da organização genômica.....	18
CAPÍTULO 2 – Triagem de mutações no gene <i>FBN1</i>	20
CAPÍTULO 3 – Diagnóstico molecular da síndrome de Marfan.	23
CAPÍTULO 4 – Caracterização do gene da fibrilina murino (<i>FBN1</i>).	24
CAPÍTULO 5 – Modelos animais para a síndrome de Marfan: estudo <i>in vivo</i> da função da fibrilina-1.	26
CAPÍTULO 6 – O estabelecimento da metodologia de geração de camundongos <i>knockout</i> no Brasil.	29
PARTE II - HERANÇA EPIGENÉTICA: A INATIVAÇÃO DO CROMOSSOMO X.	30
INTRODUÇÃO.....	30
OBJETIVOS.....	40
CAPÍTULO 1 – Uma revisão e um modelo para a iniciação do processo de inativação do cromossomo X.	41
CAPÍTULO 2 – Análise de uma mutação no promotor do gene <i>XIST</i> relacionada a desvio de inativação do cromossomo X.	42
CAPÍTULO 3 – Desenvolvimento de um sistema de monitoração da atividade gênica alelo específica no cromossomo X.	43
CAPÍTULO 4 – A inativação do cromossomo X em clones bovinos.	44
CAPÍTULO 5 – Inativação do cromossomo X em modelos animais e em humanos....	46
CAPÍTULO 6 – O papel da DNMT1 no controle da expressão do gene <i>XIST</i>	47
REFERÊNCIAS.....	48

INTRODUÇÃO: Desafios da pesquisa biomédica na Era Pós-Genoma Humano.

Em 1985, Robert Sinshmeier organizou um encontro de pesquisadores da então insípida área de genômica no campus da University of California, Santa Cruz, para discutir a proposta de sequenciamento do genoma humano (revisto por Roberts, 2001). Na época, o maior genoma seqüenciado era o do vírus Epstein-Barr, 20.000 vezes menor do que o genoma humano, e que havia levado vários anos para ser completado.

Apesar do ceticismo inicial em relação à possibilidade de realizar esse feito; à utilidade de sequenciar o genoma inteiro, incluindo o DNA “lixo”; e ao temor do desvio de verbas de projetos já em andamento, a idéia foi encampada por cientistas com força política. Dentre eles estão Walter Gilbert, que havia desenvolvido a metodologia de sequenciamento de DNA 10 anos antes, trabalho que lhe rendeu o Prêmio Nobel em 1980; James Watson, laureado com o Prêmio Nobel pela descoberta da estrutura do DNA; e Charles DeLisi, que trabalhava com câncer e liderava o “Office of Health and Environmental Research” no Department of Energy (DOE) nos Estados Unidos.

O Projeto Genoma Humano, inicialmente lançado como uma pequena iniciativa dentro do DOE em 1986, ganhou força e simpatizantes ao longo dos anos. Ciência, política e competição (revisto em Roberts, 2001a) levaram à publicação da seqüência do genoma humano por dois grupos simultaneamente em Fevereiro de 2001 (Lander et al., 2001; Venter et al., 2001). Três anos depois, em Abril de 2004 e em comemoração aos 50 anos da elucidação da estrutura do DNA, foi anunciada a versão mais completa do genoma humano, cobrindo 99% da seqüência com uma precisão de 99,99% (<http://www.ncbi.nlm.nih.gov/genome/seq/>).

O sequenciamento do genoma humano foi um marco na história da ciência, e sua disponibilidade como ferramenta de pesquisa mudou uma série de paradigmas na pesquisa biomédica (Tabela 1). De uma forma geral, passamos da era do genoma

estrutural para o genoma funcional. Ou seja, uma vez elucidada a sequência do genoma humano (informação), o novo desafio passa a ser a determinação da função de cada elemento desse genoma e as interações entre esses elementos nos processos fisiológicos normais e patológicos.

Tabela 1: Mudanças de paradigmas na pesquisa biomédica após o sequenciamento do genoma humano.

PRÉ-GENOMA	PÓS-GENOMA
Genômica ESTRUTURAL	Genômica FUNCIONAL
Identificação de genes por POSIÇÃO	Identificação de genes por SEQUÊNCIA
GENÔMICA	PROTEÔMICA
Doenças MONOGÊNICAS	Doenças POLIGÊNICAS
Diagnóstico de DOENÇA	Diagnóstico de SUSCETIBILIDADES
Análise de UM gene	Análise de VÁRIOS genes – flias., vias, sist.
AÇÃO gênica	REGULAÇÃO gênica
ETIOLOGIA (mutação específica)	PATOGÊNESE (mecanismo)
UMA espécie	VÁRIAS espécies

Adaptado de Peltonen e McKusick, 2001.

Dentre as diferentes estratégias de estudo de função gênica, o uso de modelos animais permite a realização de análises *in vivo*. O estudo comparativo de função gênica em modelos mais simples como bactéria, levedura, drosófila, e nematodo permitiu a identificação de função gênica de uma série de genes humanos, principalmente daqueles envolvidos em processos básicos como divisão celular, replicação e reparo de DNA. A importância desses modelos experimentais é tanta, que uma das metas do Projeto Genoma Humano foi o sequenciamento de seus genomas.

Tradicionalmente, no modelo humano a maior parte das informações sobre função gênica veio do estudo de doenças genéticas e a correlação entre genótipos e fenótipos. Apesar dessa estratégia ter permitido a identificação de dezenas de genes e uma primeira caracterização da função dos mesmos, ela depende da existência de pacientes informativos, contendo um genótipo interessante que possa responder alguma pergunta biológica.

Na década de 1980, o desenvolvimento de uma série de técnicas tornou possível a manipulação do genoma do camundongo de forma a se criar um animal com o genótipo desejado de acordo com a pergunta científica a ser respondida (Gordon et al., 1980; Capecchi, 1989). Assim, esse se tornou o modelo animal mamífero mais adequado para o estudo de função gênica *in vivo*. No camundongo transgênico pode-se estudar, entre outros, regulação gênica espacial e temporal, genes envolvidos no desenvolvimento embrionário, assim como mecanismos de patogênese de doenças genéticas humanas.

Que outras informações estão contidas em nosso genoma? Além da informação que determina o produto final produzido, seja ele uma proteína ou um RNA, os genes possuem também informação que determina em que condições ele será expresso. Assim, dentro do grande desafio do genoma funcional, está o estudo da regulação gênica. A capacidade de ligar e desligar diferentes conjuntos de genes possibilita que células com o mesmo genoma possuam fenótipos diferentes, como aqueles presentes nos mais de 200 tipos celulares distintos que compõem o ser-humano. A regulação gênica também permite que células respondam de forma eficiente a estímulos ambientais, adaptando seu perfil de expressão gênica de acordo com esses estímulos.

Em particular, o fenômeno epigenético é aquele que envolve qualquer forma de regulação de expressão gênica que não inclua alterações na seqüência de DNA e que seja herdada por uma ou mais gerações. Assim, a herança genética diz respeito à

presença de genes e a herança epigenética à atividade dos genes. Enquanto conhecemos relativamente bem os mecanismos associados à herança genética, a natureza das marcas epigenéticas, os fatores que as estabelecem e aqueles que as reconhecem ainda são pouco conhecidos.

Para compreendermos como a seqüência de nucleotídeos que compõem o genoma humano funciona na formação do indivíduo, é fundamental que conheçamos não só os fatores codificados por essa seqüência e suas intrerações, mas também a sua regulação. Nesta tese serão descritos trabalhos visando elucidar alguns aspectos desses dois desafios da pesquisa biomédica contemporânea: o estudo da função gênica – especificamente, do gene *FBN1*, envolvido na síndrome de Marfan; e os mecanismos de herança epigenética envolvidos no processo de inativação do cromossomo X em mamíferos.

PARTE I

ESTUDO DE FUNÇÃO GÊNICA:

FIBRILINA E A SÍNDROME DE MARFAN

INTRODUÇÃO:

I – FIBRAS ELÁSTICAS, MICROFIBRAS E FIBRILINA:

1 – Fibras Elásticas:

As funções fisiológicas de tecidos como pulmão, pele e vasos sanguíneos requer que os mesmos tenham a capacidade de esticar e encolher. Esta propriedade elástica é dada pelas fibras elásticas presentes na matriz extra-celular (MEC) nesses tecidos. Análises ultraestruturais e bioquímicas mostraram que essas fibras são compostas de um componente microfibrilar envolvendo um componente amorfo abundante composto da proteína elastina (Fahrenbach et al., 1966; Greenlee et al., 1966; Karrer and Cox, 1961). É essa proteína altamente insolúvel que confere a propriedade elástica às fibras elásticas (Starcher and Galione, 1976).

Durante a elastogênese, a elastina é secretada na MEC na forma de seu precursor, a tropoelastina (Serafini-Fracassini, 1984). Lá, a enzima lisil oxidase é responsável por ligar moléculas de tropoelastinas no que já foi descrito como uma rede de borracha. O componente microfibrilar previamente formado parece dirigir o depósito da elastina na fibra elástica.

2 - Microfibrils:

Entre as várias estruturas fibrilares presentes na MEC, encontra-se uma variedade de fibras não estriadas com pequenos diâmetros. As microfibras foram inicialmente definidas como estruturas fibrilares da matriz com um diâmetro de menos de 20 nm e sem a periodicidade de bandas das fibras de colágeno (Low, 1962). Mais recentemente, as microfibras foram divididas nos grupos de 3-5 nm e de 5-10 nm de diâmetro médio (Mecham and Hauser, 1992).

Na análise por microscopia eletrônica, as microfibras de 10 nm aparecem em seções longitudinais como filamentos com estruturas globulares espaçadas uniformemente e em seções transversais como uma estrutura tubular (Keene et al., 1991; Ren et al., 1991). Nos tecidos elásticos, as microfibras são primeiro depositadas na matriz. Lá, a elastina é posteriormente associada com essas fibras. Conjuntos de microfibras formam fibras que envolvem o centro de elastina, se estendendo para fora por distâncias variáveis (Mecham and Hauser, 1992). Além de normalmente estarem associadas com a elastina, as microfibras também existem não associadas com nenhum elemento elástico, como por exemplo, nos ligamentos das lentes (Mecham and Hauser, 1992).

As microfibras foram imunolocalizadas em diferentes tecidos, incluindo o periósteo, pericôndrio, pleura, aorta, cartilagem, tendão e músculo (Sakai et al., 1986). Acredita-se que elas tenham três funções básicas na MEC (Mecham and Hauser, 1992). Microfibras formam uma matriz onde a elastina é posteriormente depositada, como por exemplo na túnica média da aorta. Nessa estrutura, anéis de microfibras são observados durante a embriogênese antes do depósito da elastina. Microfibras também formam um agregado fibroso que liga a elastina a outros componentes da matriz – por exemplo, na junção entre a derme e a epiderme, microfibras atuam como uma ponte ligando agregados de elastina à membrana basal. Finalmente, em tecidos não elásticos, como a zona ciliar, microfibras parecem ligar as lentes ao corpo ciliar do olho.

Por causa da alta insolubilidade das microfibras, apenas alguns de seus componentes foram caracterizados, dentre eles a glicoproteína associada à microfibrila (microfibril-associated glycoprotein (MAGP) (Gibson et al., 1991), uma glicoproteína de 34kDa com atividade amino oxidase (Serafini-Fracassini et al., 1981), a enzima lisil oxidase (Kagan et al., 1986), a proteína de 58kDa associada à microfibrila (58kDa associated microfibril protein (AMP) (Horrigan et al., 1992)), emilina (Bressan et al., 1993), e fibrilina (Sakai et al., 1986).

3 - Fibrilina:

A fibrilina foi purificada pela primeira vez do meio de cultura de fibroblastos humanos por Sakai et al. em 1986 (Sakai et al., 1986). Diferentes anticorpos monoclonais preparados contra um extrato de microfibras reconheceram a mesma proteína de 350kDa que foi chamada de fibrilina. Em condições não redutoras, a fibrilina tem um peso molecular mais baixo, indicando a presença de pontes di-sulfídicas intra-moleculares. Além disso, através de marcação com glucosamina, demonstrou-se que essa proteína era glicosilada (Sakai et al., 1986).

Further characterization of fibrillin came in subsequent studies in which pepsin digests of this protein were utilized (Maddox et al., 1989). EM rotary shadowing analysis of these peptides suggested that most of the fibrillin molecule is rod-shaped. In addition, another substructure of this protein was identified as a central dense core with flexible arms extending from it, possibly an aggregate of several fibrillin fragments. It was speculated that this region may be a cross-linking domain where fibrillin molecules interact with each other through disulfide bonds (Maddox et al., 1989).

Anticorpos contra produtos de digestão com pepsina da fibrilina foram utilizados para estudar o arranjo das fibrilinas nas microfibras (Sakai et al., 1991). Esses experimentos sugeriram que os filamentos das microfibras são compostos de moléculas

de fibrilina organizadas em paralelo, separadas pelas estruturas globulares (Sakai et al., 1991). Essas seriam o local de interação entre as extremidades amino e carboxi de sucessivas moléculas de fibrilina.

II - SÍNDROME DE MARFAN (SMF):

1 - Aspectos clínicos:

A síndrome de Marfan (SMF) é a mais comum das doenças genéticas do tecido conjuntivo. Herdada de forma autossômica dominante, a SMF tem uma incidência de aproximadamente 1 em 10.000 indivíduos, aonde 30% destes são devidos a novas mutações (McKusick, 1956). O fenótipo da SMF é 100% penetrante, e suas manifestações clínicas afetam primariamente três áreas:

esquelética: crescimento excessivo dos membros, hiperextensibilidade articular, escoliose e deformidades na região anterior do tórax);

ocular: *ectopia lentis* (com deslocamento das lentes para cima), miopia e deslocamento da retina;

cardiovascular: dilatação da raiz da aorta, dissecção da aorta, prolapso e regurgitação da válvula mitral.

A sobrevida de pacientes com a SMF é de aproximadamente dois terços do normal - 40 anos na média. Em 90% dos casos a morte é causada por falha cardiovascular (McKusick, 1956). O diagnóstico precoce seguido de intervenção médica leva a um prolongamento da sobrevida.

A SMF é geneticamente homogênea e compartilha algumas de suas manifestações com outras doenças. Uma destas chamadas doenças "Marfan-like" é *ectopia lentis* herdada de forma dominante (EL), que apresenta somente as

manifestações oculares da SMF. É interessante ressaltar que recentemente foi determinado que ambas doenças são causadas por mutações no mesmo gene (Kainulainen et al., 1994).

2 - Biologia molecular:

Dadas as manifestações da SMF, pesquisadores sempre assumiram que seu defeito básico seria em uma proteína da matriz extra-celular. Em particular, em alguma proteína componente das fibras elásticas, estrutura mais afetada em exames histológicos de aortas de pacientes com esta síndrome (Byers, 1983). Várias destas moléculas foram testadas para o seu envolvimento com a SMF sem nenhum resultado positivo.

Em 1991, a combinação de estudos genéticos e bioquímicos levou à descoberta do gene envolvido na SMF (Lee et al., 1991; Maslen et al., 1991). Este gene, denominado *FBN1*, se localiza na região cromossômica 15q15 -q21.3 (Tsipouras et al., 1991; Dietz et al., 1991). Ele codifica fibrilina, uma proteína de 350 kDa, principal componente das microfibras, que junto com elastina formam as fibras elásticas (Sakai et al., 1986). De fato, a localização da fibrilina nos diferentes tecidos (periosteum, aorta, músculo, ligamentos das lentes, entre outros) se correlaciona bem com as manifestações da SMF.

O mecanismo proposto para a forma pela qual um alelo mutado do gene *FBN1* causa o fenótipo Marfan é o do modelo dominante-negativo (Bonadio et al., 1990; Dietz et al., 1993). Segundo este modelo para a patogênese de proteínas multiméricas, moléculas mutantes interagem com as normais, sendo incorporadas na ultra-estrutura correspondente (no caso da fibrilina, nas microfibras). A presença destas moléculas mutadas altera a organização e a integridade da ultra-estrutura, gerando então o fenótipo. É interessante notar que este modelo prevê então que um alelo nulo do gene *FBN1* não seria deletério. Por outro lado, existe a possibilidade de que a formação das microfibras seja suscetível a um efeito de dosagem, necessitando a expressão dos dois alelos do

FBN1. O estudo de mutações em pacientes com SMF, assim como a geração de camundongos com mutações no gene homólogo, ajudarão a esclarecer esta questão.

III - MODELO ANIMAL:

Modelos animais de doenças genéticas humanas são importantes ferramentas de pesquisa. Estes modelos ajudam a compreensão das vias envolvidas em patologias específicas, e permitem o teste de novas terapias. Modelos *in vitro*, como cultura de tecidos, possuem limitações, uma vez que não reproduzem as complexas interações físicas/bioquímicas do organismo completo. Além disso, quando o modelo animal não reproduz o fenótipo encontrado em humanos, novas vias metabólicas podem ser descobertas tanto no camundongo quanto em humanos, o que auxilia na compreensão do processo patogênico e no desenvolvimento de novos tratamentos.

Outra questão que pode ser abordada nos modelos animais é a influência do *background* genético nas manifestações de uma doença, ou na resposta a diferentes drogas. Através de cruzamentos específicos, mutações podem ser introduzidas em diferentes *backgrounds* genéticos no camundongo, e fenótipos podem ser comparados. Isto pode eventualmente levar à identificação de genes modificadores.

O recente desenvolvimento de técnicas de manipulação do genoma de camundongos faz do camundongo transgênico o melhor sistema disponível para o estudo de doenças genéticas humanas (Capecchi, 1989). A propagação pela linhagem germinativa de DNA microinjetado no pronúcleo de ovos fertilizados de camundongos foi descrita pela primeira vez em 1980 (Gordon et al., 1980). Várias cópias em *tandem* do gene injetado se integram em um sítio aleatório no genoma e são transmitidas de forma Mendeliana. Desde então microinjeção pronuclear tem sido usada para a geração de

modelos animais para várias doenças genéticas dominantes (Kivirikko, 1993) através da inserção de um alelo mutado no genoma do camundongo.

No entanto este método apresenta alguns problemas. Por causa do sítio aleatório de integração do transgene, este poderá não estar sob o controle de todos os elementos em *cis* que controlam a expressão do gene endógeno. Assim, a expressão temporal e espacial do transgene poderá não seguir o padrão de expressão do gene endógeno. Além disto, a introdução de um terceiro alelo, o transgene mutante, cria uma situação artificial no que diz respeito à proporção entre os transcritos normais e mutantes. Esta proporção pode ser crítica em doenças susceptíveis a efeitos de dosagem gênica.

Mais recentemente, a combinação do cultivo de células-tronco embrionárias (células ES) e da recombinação homóloga resultou na criação de um método mais poderoso para modificar o genoma do camundongo. Linhagens de células ES são derivadas da massa celular de blastocistos de camundongos (Evans e Kaufman, 1981). Estas células podem ser mantidas em cultura sem perder sua pluripotência. Quando injetadas em blastocistos de camundongos, estas células são capazes de popular diferentes tecidos, inclusive a linhagem germinativa (Bradley et al., 1984).

Assim, a capacidade de modificar regiões específicas do genoma do camundongo por recombinação homóloga permite a criação em potencial de camundongos com qualquer genótipo desejado. As células ES podem ser modificadas em cultura através da recombinação homóloga, e posteriormente injetadas em blastocistos. Os camundongos resultantes serão quimeras formadas de células do blastocisto recipiente e das células ES recombinantes. Se estas popularem a linhagem germinativa dos animais, a mutação será então passada às novas gerações de camundongos.

O vetor mais comumente usado para recombinação homóloga é o chamado vetor de substituição de sequência, contendo o gene de resistência a neomicina (*neo*) flanqueado por sequências homólogas ao gene alvo (Capecchi, 1989) (Figura 1). Assim,

através de dois eventos de recombinação recíprocos, este tipo de vetor media a substituição de sequências endógenas de DNA por sequência exógenas, permitindo a criação de deleções e inserções no gene de interesse.

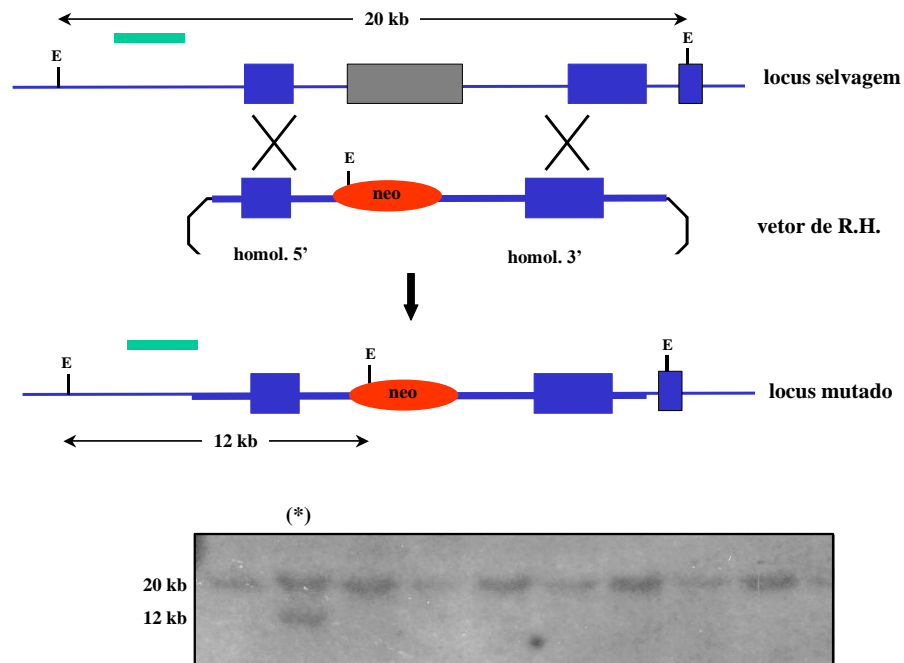


Figura 1: Recombinação homóloga. Exons estão representados por retângulos, íntrons por linhas azuis. O vetor de recombinação homóloga possui duas regiões de homologia ao gene alvo (homol. 5' e 3'). O pareamento dessas regiões com as regiões homólogas do locus endógeno permite que duas recombinações recíprocas levem à substituição de um exon (cinza) pela cassete de expressão *neo*. A detecção do evento de recombinação homóloga é feita por Southern-blot de DNA genômico das colônias de células ES. Uma sonda externa à região de homologia (barra verde) é utilizada em DNA digerido com a enzima de restrição EcoRI (E). A substituição do exon por *neo* introduz um sítio de EcoRI no locus, gerando uma banda menor referente ao alelo mutado. Uma colônia positiva para recombinação homóloga é detectada (*).

A combinação dessas estratégias de manipulação do genoma do camundongo permite potencialmente a criação de linhagens de animais mutantes com o genótipo desejado, de acordo com a pergunta científica a ser respondida. Assim, a geração de animais mutantes é grande importância para pesquisa biomédica, tanto para o estudo de doenças genéticas quanto para o estudo de função gênica *in vivo*. O final do projeto genoma humano nos apresentou 30.000 a 40.000 genes, a grande maioria com função desconhecida. Ao gerarmos camundongos mutantes para esses genes, podemos aprender sobre sua função *in vivo*.

O estudo de correlações entre genótipo e fenótipo na SMF poderá indicar regiões funcionalmente importantes da fibrilina-1, assim como sua função nos diferentes sistemas fisiológicos nos quais ela está presente. Porém, essa estratégia depende da identificação de mutações informativas em pacientes, o que escapa ao controle dos pesquisadores. Por outro lado, a estratégia de recombinação homóloga em células ES permitirá a geração de camundongos com mutações específicas no gene *Fbn1*, desenhadas de forma a permitir a comparação das consequências de mudanças qualitativas e quantitativas na fibrilina *in vivo*.

OBJETIVOS:

A elucidação do defeito molecular na SMF levantou uma série de questões em relação ao papel da fibrilina na biogênese das microfibras e o envolvimento destas nas patologias associadas com a doença. No entanto, uma investigação profunda desses processos depende do conhecimento da estrutura completa do gene da fibrilina e do desenvolvimento de um modelo experimental adequado. Nesse contexto, o objetivo principal deste trabalho foi o de gerar essas ferramentas básicas. Especificamente:

- (i) Clonar o transcrito completo da fibrilina e obter sua estrutura primária;
- (ii) Determinar a estrutura do gene *FBN1*;
- (iii) Identificar marcadores polimórficos no gene *FBN1*;
- (iv) Desenvolver modelos animais para o estudo *in vivo* de correlações entre estrutura e função da proteína fibrilina.

CAPÍTULO 1 – *FBN1*: caracterização do mRNA e da organização genômica.

A partir de sequências parciais de amino ácidos de peptídeos da fibrilina, em 1991 dois grupos geraram oligonucleotídeos degenerados e os utilizaram para isolar clones de cDNA da fibrilina de bibliotecas de cDNA humano (Lee et al., 1991; Maddox et al., 1989; Maslen et al., 1991). Consistente com o peso molecular da proteína, a análise por Northern-blot mostrou que os clones de cDNA da fibrilina detectavam um mRNA de aproximadamente 10kb (Lee et al., 1991). Em conjunto, 70% da sequência completa do mRNA da fibrilina foram isolados nesses estudos (Lee et al., 1991; Maslen et al., 1991).

Uma característica marcante notada nesses primeiros trabalhos foi que a fibrilina parecia ser composta principalmente de vários módulos repetidos ricos em cisteínas, homólogos a módulos encontrados no precursor do fator de crescimento epidérmico (EGF). Esses módulos possuem seis resíduos de cisteínas espaçados de forma específica (Davis, 1990). A estrutura terciária do EGF humano foi elucidada por ressonância magnética: pontes di-sulfídicas entre os resíduos de cisteína 1- 3, 2 - 4 e 5-6 estabilizam a molécula em uma estrutura com uma β -sheet tripla no terminal amino (Figura 2) (Cooke et al., 1987). Os módulos EGF encontrados na fibrilina possuem uma sequência consenso adicional: D, (D/N), N e Y/F, os identificando com uma sub-classe de módulos EGF envolvidos na ligação com cálcio (EGF-CB) (Handford et al., 1990) (Figura 2). O trabalho de Lee et al. (1991) também apresentou dados preliminares sobre a organização genômica do gene *FBN1*. A caracterização de três exons mostrou que cada um codificava um módulo EGF, sugerindo que o gene *FBN1* segue a mesma estrutura modular da proteína fibrilina (Lee et al., 1991).

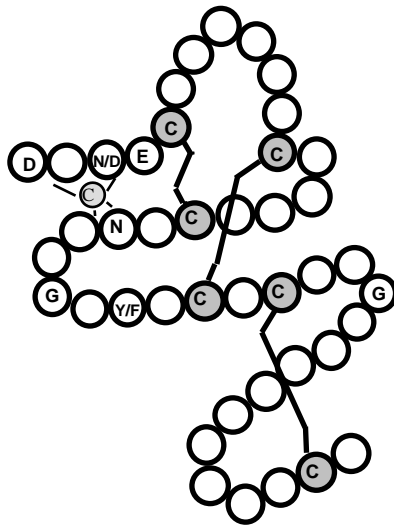


Figura 2: Estrutura terciária do módulo EGF. Aminoácidos estão representados por círculos. Aminoácidos específicos estão indicados no centro dos círculos. Pontes di-sulfídicas entre cisteínas estão representadas por linhas pretas. Resíduos do consenso de ligação ao cálcio e dos módulos EGF da fibrilina (Pereira et al., 1993) estão indicados, assim como o sítio de ligação ao cálcio.

Este trabalho teve dois objetivos principais:

(i) Determinar a sequência completa do mRNA da fibrilina:

A obtenção da sequência completa do mRNA da fibrilina permitiu a dedução da sequência primária da proteína. A partir daí podemos inferir algumas características estruturais e funcionais da fibrilina. Além disso, essa informação tornou possível a triagem de mutações em pacientes com SFM em todo o mRNA da fibrilina.

(ii) Determinar a estrutura genômica do gene *FBN1*:

A caracterização da organização genômica do gene *FBN1*, com a determinação das sequências das junções entre introns e exons, facilitou a triagem de mutações em pacientes com a SMF. Esses dados também forneceram indicações das origens evolucionárias da família gênica das fibrilinas.

CAPÍTULO 2 – Triagem de mutações no gene *FBN1*.

Uma vez identificado o gene envolvido em uma doença genética, a análise de mutações em pacientes pode fornecer dados sobre regiões estrutural e funcionalmente importantes da proteína. Especificamente, a comparação entre a natureza da mutação e seu efeito fenotípico, tanto celular quanto clínico, pode ajudar na elucidação dos mecanismos moleculares envolvidos tanto na doença quanto no sistema biológico normal. Além disso, o estabelecimento de uma correlação genótipo/fenótipo é importante para a predição da gravidade do quadro clínico da doença, como observado na doença de Gaucher (Germain et al, 1998).

Como resultado da clonagem do gene *FBN1* e sua associação com a SMF, iniciaram-se intensos estudos de triagem de mutações em pacientes. Até 1995, somente 38 mutações no gene *FBN1* haviam sido descritas (Dietz et al., 1991a, 1992a, 1992b, 1993a, 1993b; Kainulainen et al., 1992, 1994; Hewett et al., 1993, 1994b; Tynan et al., 1993; Hayward et al., 1994a, 1994b; Kartunnen et al., 1994; Piersall et al., 1994; Stahl-Hallengren et al., 1994). Essas podem ser divididas em três grandes grupos. O primeiro consiste de mutações *missense* que alteram resíduos da seqüência consenso EGF-CB. Acredita-se que essas mutações afetam a estrutura secundária ou a capacidade dos módulos EGF-CB de se ligarem ao cálcio. O segundo grupo contém aquelas mutações que criam códons de parada prematuros. Finalmente, o terceiro grupo de mutações no gene *FBN1* criam moléculas de fibrilina com deleções internas.

O mecanismo preciso pelo qual moléculas mutantes de fibrilina perturbam a formação das microfibras ainda não é conhecido. Por analogia com as collagenopatias, acredita-se que as fibrilinas anormais agem de forma dominante-negativa durante a formação das microfibras (Bonadio et al., 1990; Dietz et al., 1993a). De acordo com esse modelo, as fibrilinas mutantes interagem com as moléculas normais e perturbam a

organização geral dessas fibras ao se incorporarem nas mesmas (Bonadio et al., 1990; Dietz et al., 1993a). Isso pode tornar as microfibras menos resistentes à forças mecânicas ou mais suscetíveis à ação de proteases. Essas hipóteses podem ser confirmadas através do estudo de correlações entre genótipos e fenótipos em pacientes com a SMF. Alternativamente, elas também podem ser testadas em modelos animais, como será discutido mais tarde nesta tese.

Levando-se em conta que a SMF é uma doença monogênica, a taxa de detecção de mutações no gene *FBN1* em pacientes com a doença é baixa. Em média, mutações foram identificadas em somente 10% dos pacientes estudados (Byers, 1993; Pyeritz and Francke 1993; Sykes 1993; Tynan et al., 1993). Uma das possíveis razões dessa baixa taxa de detecção de mutações no gene *FBN1* é a técnica utilizada: análise por SSCP de amplicons derivados de cDNA. Além das limitações inerentes do método de SSCP, a triagem de mutações no gene *FBN1* a partir de seu mRNA não identificaria aquelas que levam a baixos níveis do transcrito mutado, como as mutações que geram códons de parada prematuros (Urlaub et al., 1989; Cheng et al., 1990; Dietz et al., 1993a, 1993b). Outro fator complicador nessa estratégia é a necessidade de se estabelecer culturas de fibroblastos dos pacientes para a extração de mRNA.

Com a elucidação da estrutura genômica do gene *FBN1* e a determinação das sequências das junções entre introns e exons (Pereira et al., 1993), foi possível desenhar estratégias de triagens de mutações nesse gene a partir de DNA genômico. Nesses trabalhos, um conjunto de 65 pares de primers de PCR, que em conjunto amplificam toda a região codificante do gene *FBN1* e os sítios de processamento entre introns e exons a partir de DNA genômico, é utilizado para a triagem de mutações (Nijbroek et al., 1995; Perez et al., 1998). No primeiro trabalho, a análise dos amplicons por heteroduplex foi capaz de identificar mutações em sete dos nove pacientes estudados, indicando uma taxa de detecção de mutações de 78% (Nijbroek et al., 1995). No segundo estudo, realizado

com pacientes brasileiros com a SMF, a combinação dos métodos de SSCP e heteroduplex para a análise de aplicons do gene *FBN1* foi mais eficiente do que somente o uso do heteroduplex (Perez et al., 1998).

CAPÍTULO 3 – Diagnóstico molecular da síndrome de Marfan.

Como já mencionado, a SMF apresenta grande variabilidade clínica mesmo dentro de uma mesma família, muitas vezes dificultando seu diagnóstico (Pyeritz and McKusick, 1979). A elucidação do defeito molecular na SMF trouxe esperanças que o diagnóstico clínico dessa doença pudesse ser substituído, ou pelo menos complementado, pelo diagnóstico molecular. No entanto, análises de mutações no gene *FBN1* indicaram que não existe nenhuma mutação comum levando à SMF (Sykes, 1993). Além disso, dado o tamanho do gene *FBN1*, a taxa de detecção de mutações em pacientes com SMF é baixa (Sykes, 1993). Esses dois fatores dificultaram o desenvolvimento dos diagnósticos pré-natal e pré-sintomático da SMF. Esses seriam de particular importância para o controle das manifestações cardíacas da doença. Foi proposto que quanto mais precoce o tratamento de indivíduos em risco com β -bloqueadores adrenérgicos, que diminuem a taxa de dilatação da aorta, maior a eficácia desse tratamento (Shores et al., 1994). Assim, a identificação precoce de pacientes com a SMF melhoraria seu tratamento clínico. Além disso, a exclusão do diagnóstico da doença evitaria acompanhamentos desnecessários e testes diagnósticos em casos ambíguos.

Teoricamente, os diagnósticos pré-natal e pré-sintomático da SMF poderiam ser feitos através de estudos de ligação em casos familiares utilizando-se marcadores polimórficos intragênicos. Porém, os três marcadores inicialmente descritos no gene *FBN1* provaram ser pouco informativos para serem utilizados no diagnóstico molecular (Clark et al., 1991; Hewett et al., 1991; Lee et al., 1991).

Nesse trabalho foram identificados quatro novos marcadores polimórficos no gene *FBN1*. A heterozigocidade combinada de 86% fez desse um painel de marcadores polimórficos altamente informativo para o diagnóstico molecular da SMF em casos familiares através do estudo de segregação.

CAPÍTULO 4 – Caracterização do gene da fibrilina murino (*FBN1*).

O gene *FBN1* humano foi localizado em 15q15-q21.3 (Dietz et al., 1991). Valendo-se da homologia prevista entre os genes humano e murino, o gene murino da fibrilina-1, *Fbn1*, foi mapeado através da análise de híbridos de células somáticas camundongo/hamster por PCR e os resultados foram confirmados por hibridização in situ fluorescente (FISH) (Li et al., 1993). O homólogo murino foi mapeado no cromossomo 2, bandas F. Esse resultado foi consistente com estudos prévios que demonstraram que essa região cromossômica possui sintonia conservada com o braço longo do cromossomo 15.

A caracterização do gene da fibrilina humana *FBN1* indicou que seu mRNA possui aproximadamente 10 kb (Pereira et al., 1993; Corson et al., 1993). Porém, esses mesmos estudos apresentaram resultados conflitantes quanto ao sítio de início de transcrição desse gene. Utilizando mapeamento por extensão de *primers* (*primer extension*) e por ensaios de proteção de RNase, Pereira et al. (1993) demonstraram em células de osteosarcoma que a região 5' não traduzida do transcrito da fibrilina possuía 134 nucleotídeos. Além disso, uma análise superficial da região a montante deste sítio indicou a ausência de caixas TATA e CCAAT, sugerindo que o promotor da fibrilina-1 humana teria um padrão amplo de expressão.

Já Carson et al. (1993) encontraram evidências de um transcrito de fibrilina-1 mais longo, possuindo três exons a montante do exon 1 descrito por Pereira et al. (1993) que seriam processados alternativamente. Esses resultados levantaram questões importantes sobre a estrutura do mRNA do gene *FBN1*, envolvido na função normal do tecido conjuntivo assim como na patogênese da síndrome de Marfan.

Uma estratégia para se resolver esta questão é a análise da conservação desse transcrito e da proteína em outras espécies, em particular no camundongo. No trabalho

de Win et al. (1995), a seqüência completa do mRNA do gene murino de fibrilina-1 (*Fbn1*) foi descrita, assim como a análise da aproximadamente 2kb de seqüência genômica murina a montante do ATG iniciador de tradução. Esses resultados demonstraram um alto grau de conservação entre as seqüências dos transcritos humano e murino da fibrilina-1. Além disso, a baixa homologia encontrada entre sequências genômicas a montante do primeiro ATG determinado por Pereira et al. (1993) indicou ser aquele de fato o sítio de início de tradução da fibrilina-1.

CAPÍTULO 5 – Modelos animais para a síndrome de Marfan: estudo *in vivo* da função da fibrilina-1.

A fibrilina-1 é o componente estrutural mais abundante das microfibras, que associadas ou não à elastina dão origem a diversas estruturas extra-celulares em tecidos elásticos e não elásticos. Como durante o desenvolvimento embrionário as microfibras são formadas antes do depósito da elastina, acreditava-se que a sua função seria justamente a de dirigir a formação da fibra elástica nos tecidos elásticos. Assim, imaginava-se que a SMF resultava de distúrbios na formação das microfibras e, logo, das fibras elásticas devido à presença de moléculas de fibrilina-1 mutantes.

Durante a clonagem do cDNA da fibrilina um segundo cDNA, que codifica uma proteína semelhante à fibrilina, foi isolado (Lee et al. 1991). O gene correspondente, *FBN2*, foi mapeado no cromossomo 5, e foi demonstrada ligação entre este gene e uma condição semelhante à SMF: síndrome de Beals (SB). Indivíduos afetados por esta rara anomalia do tecido conjuntivo compartilham algumas das características ósseas da SMF, porém apresentam contraturas em vez de hiperestensibilidade articular (McKusick, 1956). Além de não apresentarem manifestações oculares e cardiovasculares da SMF, estes pacientes apresentam orelhas com deformidades características. A identificação de duas proteínas homólogas ligadas a condições semelhantes levou à classificação destas como uma nova família de proteínas extracelulares: fibrilina-1, produto do gene *FBN1*, e fibrilina-2, produto do gene *FBN2* (Zhang et al., 1994).

A fibrilina-2 é altamente semelhante à fibrilina-1, sendo também composta das mesmas unidades repetitivas ricas em cisteínas. É interessante notar que a única grande diferença entre as duas proteínas é um domínio na região amino da fibrilina-1 rico em prolinas que na fibrilina-2 é rico em glicinas (Zhang et al., 1994).

Estudos dos padrões de expressão destes dois genes durante a embriogênese mostram que estes são expressos de maneiras diferentes, tanto em termos de distribuição tecidual quanto de estágios de desenvolvimento (Zhang et al., 1995). Em geral, mRNAs da fibrilina-2 aparecem mais cedo e se acumulam por um período mais curto de tempo do que os da fibrilina-1. A síntese de fibrilina-1 está relacionada com o período mais tardio da morfogênese e o aparecimento de estruturas mais definidas.

Por outro lado, a síntese da fibrilina-2 coincide com etapas iniciais da morfogênese, em particular com o começo da elastogênese. Logo, as microfibras morfológicamente homogêneas são na verdade heterogêneas no que diz respeito à sua composição das diferentes fibrilinas. Baseado nestes resultados, foi proposto que as fibrilinas têm funções distintas, porém relacionadas, na fisiologia das microfibras. Segundo esta teoria, fibrilina-1 provê principalmente suporte e força, enquanto fibrilina-2 tem uma função principal na regulação do processo inicial da formação da fibra elástica (Zhang et al., 1995). Para abordarmos essas questões *in vivo*, através da recombinação homóloga em células-tronco embrionárias foram geradas duas linhagens de camundongos contendo diferentes alelos mutados do gene *Fbn1* (Pereira et al., 1997; Pereira et al., 1999).

Experimentos com a primeira linhagem de camundongos *knockout* contendo o alelo $mg\Delta$ do gene *Fbn1* (Pereira et al., 1997) indicaram que as microfibras compostas de fibrilina-1 são responsáveis principalmente pela homeostase de tecidos em vez de serem necessárias para a formação das fibras elásticas. Isso por sua vez sugere que as manifestações cardiovasculares da SMF, principalmente a dilatação da aorta, são um resultado primário da incapacidade do sistema microfibrilar de suportar o estresse hemodinâmico e que as alterações da rede elástica são um efeito secundário.

Apesar da lesão molecular responsável pela síndrome de Marfan ter sido identificada, o mecanismo preciso através do qual mutações no gene *FBN1* levam ao

desenvolvimento das patologias associadas à doença permanece desconhecido. A geração da linhagem de camundongos contendo o alelo mutante do gene *Fbn1* mg Δ (Pereira et al., 1997) indicou que o aneurisma dessecante da aorta, característico da doença, é o resultado da presença de microfibrilas defeituosas que não suportam o estresse hemodinâmico ao qual aquela estrutura é submetida. No entanto, a seqüência de eventos patogênicos responsáveis pelo colapso da lamela elástica da parede da aorta não pode ser estudada nesses animais porque os mesmos morriam logo após o nascimento.

Durante a geração dos animais mg Δ (Pereira et al., 1997), um evento anormal de recombinação homóloga criou uma segunda linhagem de animais contendo uma mutação diferente no gene *Fbn1* (Pereira et al., 1999). Esse alelo mutante, mgR, expressa a fibrilina-1 normal, porém em níveis baixos. Animais homozigotos morrem no terceiro mês de vida de alterações vasculares. A análise da progressão dessas alterações nesses animais indicou um papel importante da resposta inflamatória no processo de dissecação da aorta.

CAPÍTULO 6 – O estabelecimento da metodologia de geração de camundongos

***knockout* no Brasil.**

A iniciativa da FAPESP de criar a rede ONSA trouxe grande desenvolvimento em pesquisa “genômica” no Brasil (Simpson e Perez, 1998; Simpson et al., 2000). Porém, como em todo projeto genoma, depois de terminadas as etapas de sequenciamento, anotação de genes e determinação do *transcritoma* correspondente, o próximo desafio é a determinação de função gênica. Uma estratégia poderosa para a análise de função gênica *in vivo* é a geração de mutantes. No caso de genes humanos, modelos animais em camundongos são mais utilizados devido à nossa vasta capacidade de manipular o genoma deste animal através da recombinação homóloga em células tronco embrionárias (ES) – os animais mutantes gerados dessa forma são chamados de animais *knockout*. A derivação de linhagens pluripotentes de células ES não é trivial. Atualmente, grupos interessados em gerar animais *knockout* dependem de linhagens estabelecidas por alguns grupos ou de células comerciais. No entanto, geralmente essas células já se encontram em passagens avançadas e podem ter sua pluripotência comprometida.

Com o objetivo de estabelecer a metodologia de criação de camundongos *knockout* no Brasil de forma a atender a crescente demanda de geração desses modelos de pesquisa por diferentes grupos nacionais, foram estabelecidas novas linhagens de células ES a partir de blastocistos de camundongos (Sukoyan et al., 2002). Essas linhagens foram testadas quanto à sua pluripotência, sendo capazes de popular a linhagem germinativa de animais quiméricos e assim se mostrando adequadas para experimentos de geração de camundongos *knockout*. A linhagem de células ES USP-1 foi utilizada em experimentos de recombinação homóloga para a geração de animais com novas mutações no gene *Fbn1* (dados não publicados).

PARTE II

HERANÇA EPIGENÉTICA: A INATIVAÇÃO DO CROMOSSOMO X.

INTRODUÇÃO:

I - HERANÇA EPIGENÉTICA:

A partir dos estudos de Mendel, no final do século XIX, os mecanismos de herança genética foram progressivamente compreendidos nos mais diferentes seres vivos, desde a reprodução assexuada até a reprodução sexuada com sua característica mistura de material genético dos progenitores. Este processo de aprendizado teve alguns pontos altos, como a descoberta da natureza bioquímica da informação genética (DNA) (Avery et al., 1944) e a elucidação da estrutura da molécula do DNA por Watson e Crick (1953).

No entanto, nos últimos 60 anos a observação de fenômenos como expressão gênica dependente da origem parental em mamíferos e o silenciamento gênico em plantas deixou claro que as sequências de nucleotídeos não são a única forma de informação genética na célula. Outros fatores modulam a expressão ou silenciamento de seqüências idênticas em células diferentes ou até mesmo na mesma célula em cromossomos diferentes. Essa atividade de regulação gênica que não envolve mudanças na seqüência de nucleotídeos e que pode persistir por uma ou mais gerações é chamada de *fenômeno epigenético*.

Mais recentemente, tem-se demonstrado que a atividade gênica é influenciada pelas proteínas que compactam o DNA em cromatina, por enzimas que modificam tanto

essas proteínas quanto o próprio DNA, e até por moléculas de RNA (revisto por Feil e Khosla, 1999; Reik e Walter, 2001). Essas proteínas e RNAs controlam padrões de expressão gênica que são passados para gerações sucessivas. Assim, mudanças epigenéticas permitem que células respondam a mudanças do meio ambiente sem ter que alterar a seqüência do DNA, também tornando possível que durante o desenvolvimento células adquiram características próprias (se diferenciem) sem perder ou ganhar material genético. No entanto, os mecanismos precisos da herança epigenética ainda não foram elucidados.

Em mamíferos, a herança epigenética se manifesta em dois fenômenos particularmente interessantes: a expressão gênica dependente de origem parental (ou *imprinting*) e no processo de inativação do cromossomo X em fêmeas. Segundo historiadores, o mecanismo de *imprinting* foi observado pela primeira vez há 3 mil anos atrás por criadores de mulas, que notaram que o cruzamento de uma égua com um burro gerava um mulo, enquanto que o cruzamento de um cavalo com uma burra produzia um jegue, com orelhas menores e pernas mais fortes do que o mulo. Assim, havia fatores parentais específicos influenciando o fenótipo da prole.

Somente em 1991 foram isolados os primeiros genes de expressão dependente da origem parental, ou *imprintados*, *Igf2r*, *Igf2* e *H19* (revisto por Willison, 1991). Desde então, mais de 40 genes *imprintados* já foram identificados (aproximadamente metade expressa a partir do alelo paterno e metade a partir do alelo materno), alguns envolvidos em doenças humanas como as síndromes de Prader-Willy e Angelman (revisto por Nicholls e Knepper, 2001).

O fenômeno de *imprinting* requer que, durante a gametogênese, as marcas de origem parental nos dois alelos dos genes *imprintados* sejam apagadas e re-estabelecidas de acordo com o sexo do indivíduo, de forma que em seus gametas esses alelos tenham a mesma marca, materna ou paterna. A partir daí, essas marcas devem

ser mantidas durante a formação do zigoto e todas as divisões celulares subseqüentes, sendo novamente apagadas e re-estabelecidas nas células germinativas do novo indivíduo. Além disso, ao longo da vida essas marcas devem ser reconhecidas por fatores celulares para que haja então a expressão gênica somente a partir do alelo materno ou paterno.

A natureza dessas marcas e os fatores que as reconhecem vem sendo aos poucos identificados. A identificação de padrões diferentes de metilação de alelos materno e paterno de alguns genes *imprintados* sugeriu um papel importante das DNA-metil-transferases (MTases) na herança epigenética. De fato, em 1993, experimentos em camundongos *knockout* demonstraram o papel da DNA-metil-transferase 1 (Dnmt1) no fenômeno de *imprinting* (Li et al., 1993). A Dnmt1 está envolvida no estabelecimento e na manutenção da metilação e age predominantemente em DNA hemi-metilado (Leonhardt e Bestor, 1993). Animais com mutações no gene *Dnmt1* apresentavam problemas de desenvolvimento e expressão bi-alélica de genes *imprintados* (Li et al., 1993).

Outra forma de controle de expressão gênica alelo-específica envolve transcritos anti-senso, como por exemplo nos genes *Igf2r* e *Ube3a* (Riesewijk et al., 1996; Oudejans et al., 2001; Chamberlain et al., 2001). Nesses loci, a transcrição na orientação anti-senso a partir de um dos alelos gera RNAs que parecem mediar o silenciamento gênico alelo-específico.

Modificações covalentes em histonas, como acetilação, metilação e fosforilação, alteram a afinidade dessas proteínas à cromatina. Essa afinidade, por sua vez, torna a cromatina mais ou menos acessível a fatores de transcrição, modulando assim a expressão gênica. Assim, essas modificações pós-traducionais das histonas representam uma forma importante de controle epigenético em eucariontes, permitindo mudanças herdáveis, porém reversíveis, de expressão gênica (revisto por Jenuwein e Allis, 2001). Diferentes combinações dessas modificações da cromatina expandem o potencial de

informações do código genético, sinalizando regiões transcricionalmente ativas e inativas do genoma.

Modificações em histonas forma associadas ao *imprinting*. Por exemplo, em mamíferos, genes *imprintados* apresentam a di-metilação da Lys-4 da histona H3 (H3K4) restrita a região promotora, enquanto que as H3K4 di-metiladas estão presentes na região promotora e também nos exons de genes não submetidos ao imprinting (Rougeulle et al., 2003). Esse mesmo fenômeno acontece em genes submetidos/que escapam à inativação do cromossomo X, como será discutido a diante.

Finalmente, o primeiro fator envolvido especificamente no estabelecimento das marcas de *imprinting* foi identificado por Aapola et al. (2000): o gene *Dnmt3L*. Homólogo às citosina-5-metiltransferases 3 (*Dnmt3A* e *Dnmt3B*), porém sem atividade enzimática, *Dnmt3L* é essencial para o estabelecimento da metilação e expressão correta de genes imprintados maternos em camundongos (Bourc'his et al., 2002). Além disso, foi demonstrado que a *Dnmt3L* coopera com as *Dnmt3A* e *Dnmt3B* nessa função (Hata et al., 2002).

II - INATIVAÇÃO DO CROMOSSOMO X

1. Compensação de dosagem:

Em mamíferos, a compensação de dose de produtos gênicos produzidos a partir do único cromossomo X dos machos e dos dois das fêmeas é realizada através do silenciamento transcricional de um dos cromossomos X nas fêmeas, o processo conhecido como inativação do cromossomo X (ICX) (Lyon, 1961). A ICX envolve diferentes etapas (Lyon, 1961), incluindo a contagem do número de cromossomos X na célula (n); a escolha aleatória de $(n-1)$ cromossomos X a serem inativos por genoma

diplóide; a iniciação da inativação desses cromossomos; a propagação em *cis* do sinal de inativação por todo o cromossomo; e finalmente, a manutenção do estado de inativação durante as posteriores divisões celulares.

Vários fatores envolvidos em cada uma dessas etapas já foram descritos. Dentre eles destacam-se o elemento de controle do X (ou *X controlling element*, *Xce*), envolvido na escolha do cromossomo X a ser inativo/ativo (Heard et al., 1997); os genes *Xist/XIST* (*X Inactive Specific Transcript*) e *Tsix*, essenciais para o reconhecimento de um cromossomo como X, para a escolha do cromossomo X ativo/inativo e para a iniciação da ICX (Lee et al., 1999); e os elementos LINE-1 (long interspersed elements) que parecem participar da propagação em *cis* do sinal de inativação (Bailey et al., 2000).

Existem na natureza pelo menos dois outros mecanismos de compensação de dose. Em *Drosophila* o único cromossomo X dos machos apresenta nível de transcrição dobrado quando comparado com cada um dos dois cromossomos X das fêmeas (Lucchesi et al., 1987). Já em *Caenorhabditis elegans* (*C. elegans*), os dois cromossomos X dos hermafroditas têm sua expressão gênica reduzida em duas vezes quando comparados com o único X dos machos (Meyer e Casson, 1986; Cline e Meyer, 1996). Nesses dois organismos, os níveis de expressão gênica de cada cromossomo X nas fêmeas são indistinguíveis. Já em mamíferos, o mecanismo de compensação de dose requer que as células da fêmea diferenciem os dois cromossomos X de forma a tornar somente um deles transcricionalmente inativo – essa diferenciação é feita através de marcas epigenéticas.

Na formação do zigoto feminino, os dois cromossomos X estão ativos. No estágio de blastocisto avançado, antes da implantação, em cada célula do embrião é iniciada a ICX. Normalmente, a escolha do cromossomo X a ser inativo é aleatória em cada célula, de forma que após a ICX o embrião feminino é composto de células que inativaram o X

materno (X_m) e células que inativaram o X paterno (X_p), se comportando funcionalmente como um mosaico.

É importante notar que, apesar da maioria dos genes ligados ao X sofrerem inativação no X_i , a ICX não se estende por todo o cromossomo. Um número crescente de genes que escapam da ICX e apresentam expressão bi-alélica vem sendo identificado nos últimos anos. Esses genes devem estar envolvidos nos fenótipos relacionados às síndromes de aneuploidia dos cromossomos sexuais, como Turner e Klinefelter. O estudo mais detalhado sobre esses genes, realizado em híbridos de células somáticas de camundongo/humanas contendo somente o X_i humano analisou aproximadamente 10% dos genes ligados ao X (Carrel et al., 1999). Foi identificada uma distribuição não aleatória dos genes que escapam da ICX com uma maior concentração na região X_p , indicando que os dois braços do cromossomo X são epigeneticamente e/ou evolucionariamente distintos.

2. Herança epigenética na ICX:

Uma vez que um cromossomo X é inativado em uma célula, todas as suas descendentes terão o mesmo cromossomo X inativo (X_i). Ou seja, a ICX estabelece marcas naquele cromossomo herdadas de forma estável por suas descendentes, fazendo com que o estado de inatividade do cromossomo X seja herdado de forma epigenética - o processo chamado de manutenção da ICX (Monk, 1986). A natureza dessas marcas e os fatores celulares que as reconhecem são tema de intenso estudo e nos últimos 10 anos alguns deles foram identificados a partir de diferenças entre os cromossomos X ativo (X_a) e X_i . Em relação ao X_a , o X_i apresenta expressão do gene *XIST/Xist* (revisto em Brockdorff, 2002); maior grau de metilação das ilhas CpG na região 5' de seus genes (revisto em Bestor, 2000); hipoacetilação das histonas H4 (Jeppesen e Turner, 1993);

maior concentração de histonas macroH2A1 (Costanzi & Pehrson, 1998); maior grau de metilação de histonas H3 na posição Lys9 (Heard et al., 2001; Boggs et al., 2002; Peters et al., 2002); e associação com a proteína BRCA1 (Ganesan et al., 2002);.

O gene *XIST/Xist*, localizado no CIX/ciX, é expresso somente a partir do Xi e é transcrito em um mRNA de aproximadamente 19,3/17,8 kb (revisto em Brockdorff, 2002). Esse transcrito não é traduzido e se localiza no núcleo, associado ao cromossomo Xi (Clemson et al., 1996; Sheardown et al., 1997; Panning et al., 1997). Diferentes estudos indicaram que, apesar de ser essencial para a iniciação da ICX, o gene *XIST/Xist* não é necessário para a manutenção da ICX em humanos e em camundongo (Brown e Willard, 1994; Rack et al., 1994; Csankovszki et al., 1999). Mais recentemente, o papel do gene *Xist* na manutenção da ICX foi estudado *in vitro* em células somáticas de camundongo (Csankovszki et al., 2001). Esses autores demonstraram pela primeira vez que a perda da expressão do gene *Xist* promove uma desestabilização do estado inativo do cromossomo Xi. Porém, essa desestabilização é maior quando também são inibidas a hipocetilação de histonas e a metilação de DNA, indicando um sinergismo entre esses três fatores na manutenção do estado inativo do Xi.

Em células não diferenciadas, a expressão do gene *Xist* é controlada pelo gene sobreposto *Tsix*, transcrito na orientação reversa de *Xist*, provavelmente por um mecanismo de inativação gênica por RNA anti-senso (Lee, 1999). Durante a diferenciação, o alelo de *Tsix* no cromossomo X a ser inativado é reprimido, o que por sua vez leva a altos níveis de expressão de *Xist* a partir desse cromossomo e à sua eventual inativação (Lee et al., 1999a). Deleção no promotor de *Tsix* leva à inativação constitutiva do cromossomo X mutado em fêmeas, indicando que esse gene está também envolvido na escolha do cromossomo X a se manter ativo (Lee e Lu, 1999). Atualmente, a repressão da expressão do gene *Tsix* no futuro Xi é o primeiro evento conhecido no processo de ICX.

Como nos genes *imprintados*, a metilação de DNA é uma das marcas epigenéticas que diferenciam os cromossomos Xa e Xi. O cromossomo Xi apresenta um alto grau de metilação de ilhas CpG quando comparado com o Xa (Yen et al., 1984; Wolf et al., 1984), uma característica que pode ser necessária para a manutenção do estado inativado do cromossomo Xi (Lyon, 1991). Por exemplo, o tratamento de células em cultura com agentes desmetiladores aumenta a frequência de reativação dos genes *HPRT* e *G6PDH* no cromossomo Xi (Mohandas et al., 1981; Wolf & Migeon, 1982). Estudos em células tronco embrionárias (ES) e camundongos *knockout* para o gene *Dnmt1* demonstraram o papel dessa MTase no controle da repressão do gene *Xist* no Xa em células diferenciadas (Beard e Jaenisch, 1995; Panning e Jaenisch, 1996). Foi também observada a inativação de genes ligados ao cromossomo X nas células ES *Dnmt1*-negativas diferenciadas, indicando que a ICX pode ser iniciada na ausência de metilação do DNA em camundongos (Panning e Jaenisch, 1996).

Uma série de modificações pós-traducionais de histonas estão envolvidas no processo de ICX. A primeira delas identificada foi a desacetilação de histonas H4. O cromossomo Xi apresenta uma grande quantidade de H4 desacetiladas (Jeppesen e Turner, 1993), o que aumenta a afinidade dessa histona com o DNA, alterando a conformação da cromatina e assim modulando a expressão gênica. Este tipo de regulação de expressão gênica através da acetilação/ desacetilação de histonas já foi descrito em levedura (Taunton et al., 1996), e aparentemente está mais envolvida com a manutenção da ICX do que com a iniciação desta (Keohane et al., 1996).

Outra modificação em histonas relacionada à ICX é a metilação da lisina 9 da H3 (H3K9). Peters et al. (2002) e Heard et al. (2001) demonstraram que a metilação da H3K9 ocorre na heterocromatina facultativa do Xi em humanos e camundongo e é mantida durante as mitoses, indicando que essa pode ser uma marca epigenética para a manutenção do estado inativo. Em camundongos, logo após a associação do RNA do

Xist ao cromossomo X a ser inativado, a forma metilada da H3K9 se associa a esse cromossomo, representando assim a primeira modificação de cromatina conhecida durante o processo de ICX (Heard et al., 2001).

Outra modificação de histonas relacionada a ICX é a di-metilação da Lys-4 da histona H3 (H3K4) (Rougeulle et al., 2003). Da mesma forma que genes *imprintados*, genes submetidos à ICX apresentam di-metilação das H3K4 restrita à região promotora, enquanto que nos genes que escapam da ICX, as H3K4 di-metiladas estão presentes na região promotora e também nos exons. Assim, da mesma forma que a metilação do DNA, a di-metilação das H3K4 parece ser uma marca epigenética comum entre o imprinting e a ICX.

O cromossomo Xi em mamíferos também é caracterizado por uma alta concentração da histona macroH2A1.2, uma variante da histona H2A com um grande domínio não-histona (Costanzi e Pehrson, 1998). Em núcleos interfásicos, o Xi pode ser identificado como um grande domínio denso em macroH2A1.2. Essa localização da macroH2A1.2 é dependente da expressão do gene *Xist* (Csankovszki et al., 1999).

Mais recentemente, Ganesan et al. (2002) demonstraram que o supressor de tumor de mama e de ovário BRCA1 também se localiza no cromossomo Xi e parece ser necessário para a associação do RNA do *XIST* a este cromossomo. Em células humanas femininas deficientes em BRCA1 o ambos o RNA do *XIST* e a macroH2A1.2 não estão associados ao Xi, corroborando também os achados de Csankovszki et al. (1999). Estas observações sugerem que a perda de BRCA1 em células femininas podem levar a perturbação e desestabilização do estado inativo do cromossomo Xi, mas sua função no processo de ICX precisa ser melhor estudada.

3. ICX e imprinting:

Enquanto no camundongo a ICX ocorre de forma aleatória nas células do embrião, nos tecidos extra-embrionários o cromossomo X paterno é preferencialmente inativado (Takagi e Sasaki, 1975; revisto em Razin e Shemer, 1995). Nesses tecidos, os dois fenômenos epigenéticos se sobrepõem: a iniciação da ICX é dependente de marcas epigenéticas de origem parental; e a manutenção da ICX, como em todos os tecidos, é dependente de marcas epigenéticas de atividade/inatividade do cromossomo X. E de fato, a inativação dependente de origem parental (ou inativação *imprintada*) está relacionada com a expressão *imprintada* do gene *Tsix* em tecidos extra-embrionários somente a partir do X materno (Lee, 2000; Sado et al., 2001).

Em humanos, a inativação do cromossomo X dependente de origem parental em tecidos extra-embrionários ainda é controversa, onde tanto a inativação aleatória (Harrison, 1989; Goto et al., 1997), como a inativação preferencial (Migeon et al., 1985; Mohandas et al., 1989) já foram descritas. A similaridade de fenótipos associados à síndrome de Turner com X paterno ou X materno é um argumento contra a existência de ICX *imprintada* em humanos.

OBJETIVOS:

A reprodução, o desenvolvimento e o funcionamento corretos dos organismos vivos dependem das heranças genética e epigenética. Apesar de conhecermos relativamente bem os mecanismos de herança genética, os mecanismos moleculares envolvidos na herança epigenética em mamíferos ainda são pouco conhecidos. O processo de manutenção da ICX envolve esse tipo de herança, uma vez que o estado inativo do Xi é herdado durante as mitoses. Assim, o estudo desse processo poderá levar à identificação de fatores envolvidos em herança epigenética. Dentro desse contexto, os objetivos desse trabalho foram:

- (i) O estabelecimento de um modelo *in vitro* para o estudo da manutenção do estado de atividade do cromossomo Xi em humanos;
- (ii) A análise da herança epigenética nos cromossomos X *in vivo* em animais gerados por transferência nuclear.
- (iii) A comparação entre os mecanismos moleculares da ICX em camundongos e em humanos.

CAPÍTULO 1 – Uma revisão e um modelo para a iniciação do processo de inativação do cromossomo X.

Tradicionalmente, a inativação do cromossomo X (ICX) é vista como um processo que envolve a contagem do número de cromossomos X na célula (n); a escolha aleatória dos ($n-1$) cromossomos a serem inativados por genoma diplóide; a propagação do sinal de inativação em *cis* por todo o cromossomo; e a posterior manutenção do estado de inativação durante todas as divisões celulares a partir daí. Apesar desses processos ainda não estarem completamente elucidados, alguns fatores importantes já foram identificados, principalmente os genes *XIST/Xist* e *Tsix* (Borsani et al., 1991; Brown et al., 1991; Lee et al., 1999), envolvidos na iniciação da ICX.

Estudos em camundongos transgênicos e *knockout* para esses genes indicam funções específicas dos mesmos no processo de ICX. Neste trabalho, é feita uma revisão desses experimentos e com base em seus resultados, é proposto um modelo para os eventos iniciais do processo de ICX. De acordo com esse modelo, o mecanismo celular *default* seria a inativação dos cromossomos X, independente do número desses cromossomos na célula. É prevista a existência de uma estrutura celular única por genoma diplóide que protegeria do processo de inativação o cromossomo X a se manter ativo.

CAPÍTULO 2 – Análise de uma mutação no promotor do gene *XIST* relacionada a desvio de inativação do cromossomo X.

Normalmente, a inativação do cromossomo X é um processo aleatório, onde em cada célula 46,XX cada cromossomo X tem 50% de chance de ser inativado. Assim, numa população normal, a proporção relativa de cromossomos X materno (X_m) e X paterno (X_p) inativos tende a seguir uma distribuição normal, onde a maioria das fêmeas possuem o padrão $X_p:X_m$ de 50:50 a 80:20 ou 20:80. Padrões de ICX fora desta faixa são chamados de desvio de ICX.

Em camundongos, desvio de ICX foi associado com o elemento de controle do X (X_{ce}) (Cattanach et al., 1969). Cromossomos X com diferentes alelos do X_{ce} possuem diferentes probabilidades de serem inativados. Em humanos, raras famílias apresentam várias mulheres com acentuado desvio de ICX (95:5 a 100:0) onde o mesmo cromossomo X é preferencialmente inativado, sugerindo a influência de um componente genético ligado ao X.

A mutação C43G no promotor mínimo do gene *XIST* foi associada com desvio de ICX em duas famílias não relacionadas (Plenge et al., 1997). Neste trabalho as autoras estudaram um grupo de 66 mulheres com desvio de inativação do cromossomo X quanto à presença da mutação C43G no promotor do gene *XIST*. A mutação não foi encontrada em nenhuma das mulheres analisadas, sugerindo que, se essa for de fato responsável pelo desvio de inativação do cromossomo X, não é um mecanismo comum.

CAPÍTULO 3 – Desenvolvimento de um sistema de monitoração da atividade gênica alelo específica no cromossomo X.

Enquanto a maioria dos genes no cromossomo X é submetida à inativação transcricional durante o processo de inativação do cromossomo X (ICX), alguns genes escapam à inativação, apresentando expressão bi-alelica (revisto por Disteché, 1995). Os principais estudos de genes que escapam à ICX foram realizados em híbridos de células somáticas roedor/humano contendo somente o cromossomo X humano inativo, o que permite a análise da expressão gênica especificamente a partir desse cromossomo (Carrel and Willard, 1999; Carrel et al., 1999; Anderson e Brown, 1999). No entanto, não se pode garantir que as células híbridas reproduzam os padrões de expressão gênica de células normais.

A recente identificação de polimorfismos de único nucleotídeo (*single nucleotide polymorphisms*, ou SNP) ao longo do genoma humano permite o desenvolvimento de métodos alternativos para o estudo de expressão gênica alelo-específica. Neste trabalho, foi gerado um painel de SNPs em regiões expressas de genes no cromossomo X através da análise de banco de dados de SNPs. Esse painel foi utilizado em linhagens primárias de fibroblastos humanos com desvio completo de ICX, demonstrando o uso desse sistema para a análise de expressão gênica alelo específica em células humanas normais.

CAPÍTULO 4 – A inativação do cromossomo X em clones bovinos.

A clonagem de mamíferos por transferência nuclear está associada com baixas taxas de sucesso e defeitos de desenvolvimento, indicando erros no processo de reprogramação do genoma somático, incluindo os sinais epigenéticos (revisto por Yanagimachi, 2002). Em mamíferos, marcas epigenéticas estão envolvidas nos processos de *imprinting* e de inativação do cromossomo X (ICX), onde elas diferenciam o X ativo (Xa) do X inativo (Xi) (revisto por Goto e Monk, 1998; Brockdorff, 2002; Ganesan et al., 2002).

Na reprodução natural, após a fertilização os dois cromossomos X estão ativos no zigoto feminino (Heard et al., 1997). Já na clonagem por transferência nuclear, o embrião clonado recebe um Xa e um Xi do núcleo somático doador, criando uma situação nova no que diz respeito às marcas epigenéticas presentes nos cromossomos X nesse estágio do desenvolvimento. Logo, como uma forma de inferir como essas marcas são tratadas durante a reprogramação nuclear, é interessante estudar os padrões de ICX em animais clonados. Eggan et al. (1999) estudaram esse processo clones de camundongos. Os autores detectaram ICX aleatória em embriões e completamente desviada em placenta dos embriões clonados, sugerindo que as marcas epigenéticas presentes nos dois cromossomos X somáticos foram apagadas e re-estabelecidas corretamente durante o desenvolvimento.

Neste trabalho, a ICX foi estudada em clones bovinos, tanto viáveis quanto inviáveis (mortos até 24 horas após o nascimento devido a complicações respiratórias) através da análise da expressão alelo específica do gene *MAOA* e dos padrões de expressão de 10 genes ligados ao X em vários órgãos de 9 animais clonados a partir de

diferentes tipos de células somáticas. Os resultados demonstraram pela primeira vez (1) ICX completamente desviada, com inativação preferencial do cromossomo X paterno, na placenta de vacas geradas por reprodução natural, como descrito em camundongo (Takagi e Sasaki, 1975; Harper et al., 1982); (2) padrões anormais de ICX em tecidos somáticos e placenta dos clones inviáveis não compatíveis com padrões descritos em animais gerados por reprodução normal; (3) ICX aleatória e padrões normais de expressão gênica ligada ao X em pele e sangue de clones bovinos viáveis e ICX completamente desviada na placenta desses animais.

Esses resultados indicaram que durante a reprogramação nuclear as marcas epigenéticas presentes nos cromossomos X somáticos não são corretamente apagadas e re-estabelecidas em clones bovinos inviáveis. Isso por sua vez pode estar contribuindo para a alta taxa de insucesso da clonagem como forma de reprodução em mamíferos.

CAPÍTULO 5 – Inativação do cromossomo X em modelos animais e em humanos.

Grande parte do que se conhece sobre os mecanismos moleculares envolvidos nas diferentes etapas da inativação do cromossomo X em mamíferos vem de experimentos em modelos animais, principalmente no camundongo. Apesar das vantagens práticas de se trabalhar com o modelo animal, nem sempre os resultados obtidos em camundongo podem ser extrapolados para humanos. Nesse trabalho, foi feita uma revisão crítica da literatura, comparando diferentes aspectos da ICX em camundongo e em humanos. Conclui-se que, dadas as diferenças desse processo já identificadas entre as duas espécies, é fundamental o desenvolvimento de modelos experimentais para o estudo da ICX especificamente em humanos.

CAPÍTULO 6 – O papel da DNMT1 no controle da expressão do gene *XIST*.

A inativação do cromossomo X (ICX) em camundongos e humanos envolve a expressão do alelo do gene *Xist/XIST* no X inativo (Xi) e repressão do alelo no X ativo (Xa). Essa repressão está associada à metilação da região 5' do gene *Xist/XIST* (Beard et al., 1995; Hendrich et al., 1997). Em camundongos, a DNA-metil-transferase Dnmt1 é necessária para a metilação e repressão transcricional do alelo do gene *Xist* no Xa (Panning et al., 1996; Beard et al., 1995). Além disso, essa enzima é responsável pela manutenção da metilação global do DNA e de genes *imprintados* (Li et al., 1993).

Em humanos, o papel da DNMT1 na metilação global do DNA parece ser diferente. Dois estudos em cultura de células deficientes para as MTases DNMT1 e DNMT3B indicam que a metilação do DNA é mantida através de uma colaboração entre as duas enzimas (Rhee et al., 2000; Rhee et al., 2002). Enquanto células deficientes em DNMT1 ou DNMT3B mantêm a metilação global do DNA, células deficientes nas duas enzimas simultaneamente apresentam até 95% de desmetilação do DNA, incluindo a desmetilação de alelos inativos de genes *imprintados*.

Neste trabalho, o papel das MTases DNMT1 e DNMT3B na repressão transcricional do alelo do gene *XIST* no Xa é investigada (Vasques et al., submetido). Enquanto desmetilação global do DNA provocada pelo tratamento das células com o agente desmetilador 5-aza-2'-deoxycytidine levou à expressão do *XIST*, ausência de DNMT1, DNMT3B ou de ambas MTases não foi suficiente para a ativação desse gene. Conclui-se que, nessas células humanas, diferentes mecanismos controlam metilação global do DNA e metilação do gene *XIST*.

REFERÊNCIAS

Aapola, U, Kawasaki, K, Scott, HS, Ollila, J, Vihinen, M, Heino, M, Shintani, A, Kawasaki, K, Minoshima, S, Krohn, K, Antonarakis, SE, Shimizu, N, Kudoh, J, e Peterson, P. (2000). Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family. *Genomics*, 65(3):293-8

Anderson, CL e Brown, CJ (1999). Polymorphic X-chromosome inactivation of the human TIMP1 gene. *Am J Hum Genet.* 65(3): 699-708.

Avery OT, MacLeod CM e McCarty M. (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types. *J. Exp. Med.* 79: 137-158.

Bailey JA, Carrel L, Chakravarti A, e Eichler EE. (2000). Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. *Proc Natl Acad Sci U S A.* 97(12):6634-9

Beard, C., Li E., and Jaenisch R. (1995) Loss of methylation activates Xist in somatic but not in embryonic cells, *Genes Dev.* 9:2325-2334.

Bonadio, J, Saunders, TL, Tsai, E, Goldstein, SA, Morris-Wiman, J, Brinkley, L, Dolan, D.F, Altschuler, R.A, Hawkins, J.E, Bateman, J.F, Mascara, T, e Jaenisch, R. (1990). Transgenic mouse model of the mild dominant form of osteogenesis imperfecta. *Proc. Natl. Acad. Sci. U.S.A.*, 87:7145-7149.

Borsani, G, Tonlorenzi, R, Simmler, M.C, Dandolo, L, Arnaud, D, Capra, V, Grompe, M, Pizzuti, A, Muzny, D, Lawrence, C, Willard, H.F, Avner, P, e Ballabio, A (1991). Characterization of a murine gene expressed from the inactive X chromosome. *Nature* 351:325-329.

Bradley, A, Evans, M, Kaufman, M.H, e Robertson, E. (1984). Formation of germ-line chimeras from embryo-derived teratocarcinoma cell lines. *Nature*, 309:255-256.

Bressan, G. M, Daga-Gordini, D, Colombatti, A, Castellani, I, Marigo, V, e Volpin, D. (1993). Emilin, a component of elastic fibers preferentially located at the elastin-microfibrils interface. *J. Cell Biol*, 121:201-212.

Brockdorff N. (2002). X-chromosome inactivation: closing in on proteins that bind Xist RNA. *Trends Genet.* 18(7):352-8.

Brown, CJ, Ballabio, A, Rupert, JL, Lafrenière, RG, Grompe, M, Tonlorenzi, R, e Willard, HF. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* 349:38-44.

Bourc'his, D, Xu, GL, Lin, CS, Bollman, B, e Bestor, TH. (2001). Dnmt3L and the establishment of maternal genomic imprints. *Science.* 294(5551):2536-9.

Boumil RM, e Lee JT. (2001). Forty years of decoding the silence in X-chromosome inactivation. *Hum Mol Genet.* 10(20):2225-32.

Byers, PH (1983). Inherited disorders of collagen biosynthesis: Ehlerst-Danlos syndrome, the Marfan syndrome, and osteogenesis imperfecta. In J. A. Spittel (Eds.), *J. Clin. Med.* (pp. 1). Philadelphia: Harper & Row, publishers.

Capecchi, MP (1989). The new mouse genetics: altering the genome by gene targeting. *Trends Genet*, 5:70-76.

Carrel L, e Willard HF (1999). Heterogeneous gene expression from the inactive X chromosome: an X-linked gene that escapes X inactivation in some human cell lines but is inactivated in others. *Proc Natl Acad Sci USA*. 96(13): 7364-9.

Carrel L, Cottle AA, Goglin KC, e Willard HF (1999), A first-generation X-inactivation profile of the human X chromosome. *Proc Natl Acad Sci USA*. 96(25): 14440-4.

Cattanach BM, Pollard CE, Perez JN (1969). Controlling elements in the mouse X-chromosome. I. Interaction with the X-linked genes. *Genet Res*. 14(3):223-35.

Chamberlain SJ, e Brannan CI. (2001). The Prader-Willi syndrome imprinting center activates the paternally expressed murine Ube3a antisense transcript but represses paternal Ube3a. *Genomics*. 73(3):316-22.

Clark, BA, Maslen, CL, Sakai, LY, Dhalimi, MA, Litt, R, e Litt, M. (1991). A BamHI polymorphism at the fibrillin (FBN) locus. *Nucl. Acids. Res*, 19(15):4309.

Corson, GM, Chalberg, SC, Dietz, HC, Charbonneau, NL, e Sakai, LY. (1993). Fibrillin binds calcium and is coded by cDNAs that reveal a multidomain structure and alternatively spliced exons at the 5'end. *Genomics*, 17:476-484.

De Souza SJ *et al* (2000). Identification of human chromosome 22 transcribed sequences with ORF expressed sequence tags. *Proc Natl Acad Sci U S A*, 97(23):12690-3.

Dias Neto E *et al* (2000). Shotgun sequencing of the human transcriptome with ORF expressed sequence tags. *Proc Natl Acad Sci U S A*, 97(7):3491-6.

Dietz, H. C, Pyeritz, R.E, Hall, B.D, Cadle, R.G, Hamosh, A, Schwartz, J, Meyers, D.A, e Francomano, C.A. (1991). The Marfan syndrome locus: confirmation of assignment to chromosome 15 and identification of tightly linked markers at 15q15-q21.3. *Genomics*, 9:355-361.

Dietz, H. C, Cutting, G.R, Pyeritz, R.E, Maslen, C.L, Sakai, L.Y, Corson, G.M, Puffenberger, E.G, Hamosh, A, Nanthakumar, E.J, Curristin, S.M, Stetten, G, Meyers, D.A, e Francomano, C.A. (1991a). Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene. *Nature*, 352:337-339.

Dietz, H. C, Saraiva, J.M, Pyeritz, R.E, Cutting, G.R, e Francomano, C.A. (1992a). Clustering of fibrillin (FBN1) missense mutations in Marfan syndrome patients at cysteine residues in the EGF-like domains. *Hum. Mut*, 1:366-374.

Dietz, H. C, Pyeritz, R.E, Puffenberger, E.G, Kendzior, R.J, Corson, G.M, Maslen, C.J, Sakai, L.Y, Francomano, C.A, e Cutting, G.R. (1992b). Marfan phenotype variability in a

family segregating a missense mutation in the EGF-like motif of the fibrillin gene. *J. Clin. Invest*, 89:1674-1680.

Dietz, H. C, McIntosh, I, Sakai, L.Y, Corson, G.M, Chalberg, S.C, Pyeritz, R.E, e Francomano, C.A. (1993a). Four novel FBN1 mutations: significance for mutant transcript level and EGF-like domain calcium binding in the pathogenesis of Marfan syndrome. *Genomics*, 17:468-475.

Dietz, H. C, Valle, D, Francomano, C.A, Kendzior, F.J, Pieritz, R.E, e Cutting, G.R. (1993b). The skipping of constitutive exons in vivo induced by nonsense mutation. *Science*, 254:680-683.

Disteche C. (1995) Escape from X inactivation in human and mouse. *Science* 11:17-22.

Eggan K, Akutsu H, Hochedlinger K, Rideout W 3rd, Yanagimachi R, e Jaenisch R. (2000). X-chromosome inactivation in cloned mouse embryos. *Science* 290:1578-1581.

Evans, M. J, e Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature*, 292:154-155.

Feil R, e Khosla S. (1999). Genomic imprinting in mammals: an interplay between chromatin and DNA methylation? *Trends Genet.* 15(11):431-5.

Ganesan S, Silver DP, Greenberg RA, Avni D, Drapkin R, Miron A, Mok SC, Randrianarison V, Brodie S, Salstrom J, Rasmussen TP, Klimke A, Marrese C, Marahrens

Y, Deng CX, Feunteun J e Livingston DM. (2002). BRCA1 supports XIST RNA concentration on the inactive X chromosome. *Cell*, 111: 393-405.

Germain, D.P, Puech, J, Caillaud, C, Kahn, A, e Poenaru, L. (1998) Exhaustive Screening of the Acid -Glucosidase Gene, by Fluorescence-Assisted Mismatch Analysis Using Universal Primers: Mutation Profile and Genotype/Phenotype Correlations in Gaucher Disease. *Am. J. Hum. Genet*, 63:415-427.

Gibson, M. A, Kumaratilake, J.S. e Cleary, E.G. (1989). The protein components of the 12-nanometer microfibrils of elastic and non-elastic tissues. *J. Biol. Chem*, 264:4590-4598.

Gibson, M. A, Sandberg, L.B, Grosso, L.E, e Cleary, E.G. (1991). Complementary DNA cloning establishes microfibril-associated glycoprotein (MAPG) to be a discrete component of the elastin-associated microfibrils. *J. Biol. Chem*, 266:7596-7601.

Gordon, J. W, Scangos, G.A, Plotkin, D.J, Barbosa, J.A, e Ruddle, F.H. (1980). Genetic transformation of mouse embryos by microinjection of purified DNA. *Proc. Natl. Acad. Sci. U.S.A*, 77:7380-7384.

Goto T. e Monk M. (1998). Regulation of X-Chromosome Inactivation in Development in Mice and Human. *Microbiology and Molecular Biology Reviews*, 62: 362 - 378.

Hata, K, Okano, M, Lei, H, e Li E. (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* 129(8):1983-93.

Harper, M. I, Fosten, M, e Monk, M. (1982) Preferential paternal X inactivation in extraembryonic tissues of early mouse embryos. *J. Embryol. Exp. Morphol.* 67, 127-135.

Heard, E, Clerc, P. e Avner, P. (1997) X-chromosome inactivation in mammals. *Annu. Rev. Genet.* 31:571-610.

Hendrich, B. D., Brown, C. J., and Willard, H. F. (1993) Evolutionary conservation of possible functional domains of the human and murine XIST genes, *Hum. Mol. Genet.* 2:663-672.

Hewett, DR, Lynch, JR, Godfrey, M, e Sykes, BC. (1991). G/A polymorphism in an intron of the fibrillin gene FBN1. *Nucl. Acids Res.* 19(24): 6975.

Horrigan, SK, Rich, CB, Streeten, BW, Li, Z, e Foster, JA. (1992). Characterization of an associated microfibril protein through recombinant DNA techniques. *J. Biol. Chem.* 267(14): 10087-95.

Jenuwein T, e Allis CD. (2001). Translating the histone code. *Science* 293(5532):1074-80.

Jeppesen P, e Turner BM. (1993). The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell.* 74(2):281-9.

Kainulainen, K, Karttunen, L, Puhakka, L, Sakai, L, e Peltone, L. (1994). Mutations in the fibrillin gene responsible for dominant ectopia lentis and neonatal Marfan syndrome. *Nature Genet.* 6: 64-9.

Kagan, H. M, Vaccaro, CA, Bronson, RE, Tang, SS, e Brody, JS. (1986). Ultrastructural immunolocalization of lysyl oxidase in vascular connective tissue. *J. Cell Biol.* 103: 1121-8.

Keene, DR, Maddox, BK, Kuo, HJ, Sakai, LY, e Glanville, RW. (1991). Extraction of extendable beaded structures and their identification as fibrillin-containing extracellular matrix microfibrils. *J. Histochem. Cytochem.* 39: 441-9.

Karrer, HE, e Cox, J. (1961). Electron microscope study of developing chick embryo aorta. *J. Ultrastruct.Res.* 4: 420-54.

Keene, DR, Maddox, BK, Kuo, HJ, Sakai, LY, e Glanville, RW. (1991). Extraction of extendable beaded structures and their identification as fibrillin-containing extracellular matrix microfibrils. *J. Histochem. Cytochem.* 39: 441-9.

Keohane AM, O'Neill LP, Belyaev ND, Lavender JS, e Turner BM. (1996). X-Inactivation and histone H4 acetylation in embryonic stem cells. *Dev Biol.* 180(2):618-30.

Kivirikko, KI. (1993). Collagens and their abnormalities in a wide spectrum of diseases. *Ann. Med.* 25: 113-26.

Lander ES, Linton LM, Birren M, Nusbaum C et al. (2001). Initial sequencing and analysis of the human genome. *Nature* 409:942 – 943.

Lee JT, Lu N. (1999). Targeted mutagenesis of Tsix leads to nonrandom X inactivation. *Cell* 99(1):47-57.

Lee, JT, Davidow, LS, e Warshawsky, D. (1999). Tsix, a gene antisense to Xist at the X-inactivation centre. *Nature Genetics*. 2: 400-4.

Lee, B, Godfrey, M, Vitale, E, Hori, H, Mattei, MG, Sarfarazi, M, Tsipouras, P, Ramirez, F, e Hollister, DW. (1991). Linkage of Marfan syndrome and a phenotypically related disorder to two fibrillin genes. *Nature*,. 352: 330-4.

Leonhardt H, e Bestor TH. (1993). Structure, function and regulation of mammalian DNA methyltransferase. *EXS*. 64:109-19.

Li, E, Beard, R, e Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. *Nature* 366: 362-5.

Li E, Beard C, Forster AC, Bestor TH, e Jaenisch R. (1993). DNA methylation, genomic imprinting, and mammalian development. *Cold Spring Harb Symp Quant Biol*. 58:297-305.

Low, FN. (1962). Microfibrils: fine filamentous components of the tissue space. *Anat. Rec*. 142: 131-7.

Maddox, BK, Sakai, LY, Keene, DR, e Glanville, RW. (1989). Connective tissue microfibrils; isolation and characterization of three large pepsin-resistant domains of fibrillin. *J. Biol. Chem*. 264: 21381-5.

Maslen, CL, Corson, GM, Maddox, BK, Glanville, RW, e Sakai, LY. (1991). Partial sequence of a candidate gene for the Marfan syndrome. *Nature*. 352: 334-7.

McKusick, V. (1956). *Heritable Disorders of Connective Tissue*. 1st ed. St. Louis: CV Mosby.

Mecham, RP, e Hauser, JE. (1992). The Elastic Fiber. In E. D. Hay (Eds.), *Cell Biology of the Extracellular Matrix*, 2nd ed. (pp. 79). Plenum Publish Co.

Neptune, ER, Frischmeyer, PA, Arking, DE, Myers, L, Bunton, TE, Gayraud, B, Ramirez, F, Sakai, LY, e Dietz, HC. (2003). Dysregulation of TGF-beta activation contributes to pathogenesis in Marfan syndrome. *Nat. Genet.* 33(3):407-11.

Nicholls RD, e Knepper JL. (2001). Genome organization, function, and imprinting in Prader-Willi and Angelman syndromes. *Annu Rev Genomics Hum Genet.* 2:153-75.

Oudejans CB, Westerman B, Wouters D, Gooyer S, Leegwater PA, van Wijk IJ, e Sleutels F. (2001). Allelic IGF2R repression does not correlate with expression of antisense RNA in human extraembryonic tissues. *Genomics.* 73(3):331-7.

Panning B. e Jaenisch R. (1996). DNA hypomethylation can activate Xist expression and silence X-linked genes. *Genes Dev*, 10: 1991-2002.

Panning B, Dausman J, e Jaenisch R. (1997). X chromosome inactivation is mediated by Xist RNA stabilization. *Cell* 90(5):907-16.

Pelton, L, e McKusick, V. (2001). Dissecting human disease in the Postgenomic Era. *Science*. 291: 1224-9.

Pereira, L, D'Alessio, M, Ramirez, F, Lynch, JR, Sykes, B, Pangilinan, T, e Bonadio, J (1993). Genomic organization of the sequence coding for fibrillin, the defective gene product in Marfan syndrome. *Hum. Mol. Genet.* 2(10): 1762.

Pereira, L, Zhang, H, e Ramirez, F. Microfibril pathology and cardiovascular manifestations in Marfan syndrome. Chapter in *Molecular Biology of Cardiovascular Disease*, 1997. Eds. Marks and Taubman. Marcel Dekker, Inc, New York, NY, USA.

Pereira LV, e Vasques LR. (2000). X-chromosome inactivation: lessons from transgenic mice. *Gene* 255(2):363-71.

Plenge RM, Hendrich BD, Schwartz C, Arena JF, Naumova A, Sapienza C, Winter RM, Willard HF(1997). A promoter mutation in the XIST gene in two unrelated families with skewed X-chromosome inactivation. *Nat Genet.* 17(3):353-6.

Pyeritz, RE, e McKusick, VA. (1979). The Marfan syndrome: diagnosis and management. *N. Engl. J. Med.* 300: 772-7.

Reik W. e Walter J. (2001). Genomic imprinting: parental influence on the genome. *Nat Rev Genet.* 2(1):21-32

Ren, ZX Brewton, RG, e Mayne, R. (1991). An analysis by rotary shadowing of the structure of the mammalian vitreous humor and zonular apparatus. *J. Struct. Biol.* 106(1): 57-63.

Rhee I, Jair K-W, Yen R-WC, Lengauer C, Herman JG, Kinzler KW, Vogelstein B, Baylin SB e Schuebel KE. (2000). CpG methylation is maintained in human cancer cells lacking DNMT1. *Nature*, 404: 1003 - 1007.

Rhee I, Bachman KE, Park BH, Jair K-W, Yen R-WC, Schuebel KE, Cui H, Feinberg AP, Lengauer C, Kinzler KW, Baylin SB e Vogelstein B. (2002). DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature*, 416: 552 - 556.

Riesewijk AM, Schepens MT, Welch TR, van den Berg-Loonen EM, Mariman EM, Ropers HH, e Kalscheuer VM. (1996). Maternal-specific methylation of the human IGF2R gene is not accompanied by allele-specific transcription. *Genomics*. 31(2):158-66.

Roberts L .(2001). Controversial form the start. *Science* 291:1182-1188.

Roberts L . (2001b). A history of the human genome project. *Science* 291:1195-1200.

Rougeulle C, Navarro P, e Avner P. (2003). Promoter-restricted H3 Lys 4 di-methylation is an epigenetic mark for monoallelic expression. *Hum Mol Genet.* 12(24):3343-8.

Sakai, LY, Keene, DR, e Engvall, E. (1986). Fibrillin a new 350-kD glycoprotein is a component of extracellular microfibrils. *J. Cell. Biol.* 103(6): 2499-509.

Sakai, LY, Keene, DR, Glanville, RW, e Bachinger, HP. (1991). Purification and partial characterization of fibrillin, a cysteine-rich structural component of connective tissue microfibrils. *J. Biol. Chem.* 266(22): 14763-70.

Serafini-Fracassini, A, Venterella, G, Field, MJ, Hinnie, J, Onyezili, e Griffiths, R. (1981). Characterization of a structural glycoprotein from bovine ligamentum nuchae exhibiting dual amine oxidase activity. *Biochemistry.* 20(19): 5424-9.

Sheardown SA, Newall AE, Norris DP, Rastan S, e Brockdorff N. (1997). Regulatory elements in the minimal promoter region of the mouse Xist gene. *Gene* 203(2):159-68.

Sheardown SA, Duthie SM, Johnston CM, Newall AE, Formstone EJ, Arkell RM, Nesterova TB, Alghisi GC, Rastan S, e Brockdorff N. (1997). Stabilization of Xist RNA mediates initiation of X chromosome inactivation. *Cell* 91(1):99-107.

Shores, J, Berger, KR, Murphy, EA, e Pyeritz, RE. (1994). Progression of aortic dilatation and the benefit of long-term beta-adrenergic blockade in Marfan's syndrome. *N Engl J Med.* 330(19):1335-41.

Simpson, AJ e Perez, JF. (1998). ONSA, the Sao Paulo Virtual Genomics Institute. Organization for Nucleotide Sequencing and Analysis. *Nat Biotechnol.* 16(9):795-6.

Simpson AJ *et al.* (2000). The genome sequence of the plant pathogen *Xylella fastidiosa*. The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing and Analysis. *Nature* 13;406(6792):151-7.

Starcher, BC, e Galione, MJ. (1976). Purification and comparison of elastins from different animal species. *Anal. Biochem.* 74(2): 441-7.

Sykes, B. (1993). Marfan gene dissected. *Nat Genet.* 3(2):99-100.

Tsipouras, P, Sarfarazi, M, Devi, A, Weiffenbach, B, e Boxer, M. (1991). Marfan syndrome is closely linked to a marker on chromosome 15q1.5-q2.1. *Proc. Natl. Acad. Sci.U.S.A.* 88(10):, 4486-8.

Takagi, N. e Sasaki, M. (1975). Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* 256, 640-642.

Vasques LR, e Pereira LV. (2001). Allele-specific X-linked gene activity in normal human cells assayed by expressed single nucleotide polymorphisms (cSNPs). *DNA Res.* 8(4):173-7.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ et al. (2001). The sequence of the human genome. *Science* 291:1304-1351.

Watson JD, e Crick FH. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 25;171(4356):737-8.

Willison, K. (1991). Opposite imprinting of the mouse *Igf2* and *Igf2r* genes. *Trends Genet.* 7(4):107-9.

Yanagimachi, R. (2002). Cloning: experience from the mouse and other animals. *Molecular and Cellular Endocrinology* . 187: 241-248.

Xue F, Tian XC, Du F, Kubota C, Taneja M, Dinnyes A, Dai Y, Levine H, Pereira LV, e Yang X. (2002). Aberrant patterns of X chromosome inactivation in bovine clones. *Nat Genet*. 31(2):216-20.

Zhang, H, Apfelroth, SD, Hu, W, Davis, EC, e Sanguineti, C. (1994). Structure and expression of fibrillin-2, a novel microfibrillar component preferentially located in elastic matrices. *J. Cell. Biol.* 124(5):855-63.

Zhang H, Hu W, e Ramirez F. (1995). Developmental expression of fibrillin genes suggests heterogeneity of extracellular microfibrils. *J Cell Biol.* 129(4):1165-76.