

70376

SECRETARIA
DE
PÓS-GRADUAÇÃO
I. B.

UNIVERSIDADE ESTADUAL DE CAMPINAS - UNICAMP
INSTITUTO DE BIOLOGIA
DEPARTAMENTO DE GENÉTICA MÉDICA

ESTUDO MOLECULAR DE UMA SEQUÊNCIA DE DNA ISOLADA DO CROMOSSOMO Y HUMANO

EDI LÚCIA SARTORATO

Orientador: Profa. Dra. SOLANGE BENTO FARAH

Dissertação apresentada ao Instituto de Biologia da
Universidade Estadual de Campinas (UNICAMP)
para a obtenção do grau de Mestre em Ciências
Biológicas, área de concentração Genética Humana.

Este exemplar corresponde à redação final
da tese defendida pelo(a) candidato(a)

Edi Lúcia Sartorato

e aprovada pela Comissão Julgadora.

1994

Solange Bento Farah

UNICAMP
BIBLIOTECA CENTRAL

UNIDADE	IB
N.º CHAMADA:	T/UNICAMP
	Pa 772
V.	Ex.
TOMBO BC	23962
PROC.	433195
C	☐
D	☒
PREÇO	R\$ 11,50
DATA	14/55
N.º CPD	cm00067644-4

Sartorato, Edi Lucia
 Sa77e Estudo molecular de uma sequência de DNA isolada do cromossomo
 Y humano / Edi Lucia Sartorato. - - Campinas, SP : [s.n.], 1994.

Orientador : Solange Bento Farah.
 Dissertação (mestrado) - Universidade Estadual de Campinas.
 Instituto de Biologia.

1. Cromossomo Y. 2. Sondas DNA. 3. Regiões promotoras
 (genética). I - Farah, Solange Bento. II. Universidade Estadual
 de Campinas. Instituto de Biologia III. Título.
 20. CDD -574.873 22 - 574.873 28
 -575.1

Índices para catálogo sistemático:

1. Cromossomo Y 574.873 22
2. Sondas DNA 574.873 28
3. Regiões promotoras (genética) 575.1

*ao meu avô,
Valentim,
in memoriam*

AGRADECIMENTOS

À Profa. Dra. Solange Bento Farah, Departamento de Genética Médica, FCM, Unicamp.

À Profa. Dra. Christine Hackel, Departamento de Genética Médica, FCM, Unicamp.

Ao Prof. Dr. Walter Pinto Jr., Departamento de Genética Médica, FCM, Unicamp.

À Profa. Dra. Anete Pereira de Souza, Departamento de Genética e Evolução, IB, Unicamp.

Ao Prof. Dr. Bernardo Beiguelman, Departamento de Genética Médica, FCM, Unicamp.

Ao Prof. Dr. Paulo Arruda, Coordenador do Centro de Biologia Molecular e Engenharia Genética, Unicamp.

À Profa. Dra. Andréa Trevas Maciel Guerra, Departamento de Genética Médica, FCM, Unicamp.

À Profa. Dra. Laura Maria Mariscal Ottoboni, Centro de Biologia Molecular e Engenharia Genética, Unicamp.

À Profa. Dra. Carmem Silvia Bertuzzo, Departamento de Genética Médica, FCM, Unicamp.

Ao Prof. Dr. Adilson Leite, Centro de Biologia Molecular e Engenharia Genética, Unicamp.

Ao Prof. Dr. Horário Schneider da Universidade Federal do Pará.

Aos professores do Departamento de Genética Médica, FCM, Unicamp.

Aos professores do Centro de Biologia Molecular e Engenharia Genética, Unicamp.

Aos funcionários, alunos, estagiários e residentes do Departamento de Genética Médica, FCM, Unicamp.

Aos funcionários, alunos e estagiários do Centro de Biologia Molecular e Engenharia Genética, Unicamp.

À Cláudia de Lourdes Soraggi, Departamento de Bioquímica, IB, Unicamp.

Ao Emilton da Diretoria de Recursos Didáticos, FCM, Unicamp.

À diretoria da Faculdade de Ciências Médicas, Unicamp.

Às secretárias da Sub-comissão da Pós-graduação em Genética, IB, Unicamp.

Às minhas amigas de sempre, Ercília e Mirna.

Ao Sr. Tercílio Ventura e Sra. Rosária Corrêa Ventura.

Aos meus pais e meus irmãos.

AGRADECIMENTOS ESPECIAIS

Ao meu marido **Alejandro**, que com carinho, amor e paciência fez possível a realização desse trabalho.

À minha filha **Luna**, que com sorrisos, beijos, choros, brincadeiras... deu mais cor à minha vida.

ABREVIATURAS

°C - graus Celsius

µg - micrograma

µl - microlitro

12E7 - anticorpo monoclonal

ACTGP2 - pseudogene relacionado com o gene gama-actina do citoesqueleto

AI -amelogênese imperfeita

Amp - ampicilina

AMG - gene relacionado com desenvolvimento dentário

AMGL - pseudogene localizado no cromossomo Y relacionado com o gene AMG

ASSP6 - pseudogene relacionado com a enzima arginosuccinato sintetase

AZF - fator de azoospermia

Bkm - *banded krait minor satellite DNA*

bp - pares de base

CAT - gene que codifica a enzima cloranfenicol acetil transferase

cDNA - ácido desoxiribonucleico complementar

cen - centrômero

cm - centímetros

CSF2R - gene para receptor GM-CSF

ddNTPs - didesoxinucleotídeos trifosfatos

DNA - ácido desoxiribonucleico

dNTPs - desoxiribonucleotídeos trifosfatos

DO - densidade ótica

DTT - ditioteirol

EDTA - ácido etilenodiaminotetracético

clF-3 - fator de iniciação

EMBL - European Molecular Biology Laboratory

g - grama

GBY - locus gonadoblastoma do cromossomo Y

GCY - gene relacionado com fator de crescimento

GenBank - Genetic Sequence Data Bank

GM-CSF - fator de estimulação de colônia granulócito-macrófago

HMG - *high mobility group*

hUBF - fator de transcrição

H-Y - antígeno de histocompatibilidade

kb - quilobase

l - litro

LB - meio de cultura Luria-Bertani

LINEs - *long interspersed elements*

M - molaridade

Mb - megabase

Mc - proteína requerida na fissão de leveduras

MES - ácido 2-N[-morfolino] etano sulfônico

mg/ml - miligramas por mililitro

MIC2 - gene que codifica um antígeno de superfície celular

ml - mililitro

N - normalidade

ng - nanogramas

nm - nanômetros

ORF - cadeia aberta de leitura de aminoácidos

PAR - região pseudoautossômica

q.s.p. - quantidade suficiente para

RNA - ácido ribonucleico

rpm - rotações por minuto

RPS4 - gene que codifica proteína ribossomal

SDS - dodecil sulfato de sódio

SINEs - *short interspersed elements*

SRY - região de determinação sexual do cromossomo Y

Sry - região de determinação sexual do cromossomo Y em camundongo

STE - sódio/tris/EDTA

STS - gene relacionado com a enzima esteróide sulfatase

STS - sequence-target sites

TAE - tris/ácido acético/EDTA

TDF - fator de determinação testicular

Tdy - fator de determinação testicular em camundongos

TBE - tris/ácido bórico/EDTA

TCF-1 - fator de transcrição

TE - tris-EDTA

TEMED - N, N, N', N' - tetrametiletilenodiamina

tris - tris-(hidroximetil)-aminometano

UV - ultra -violeta

XE7 - gene relacionado com controle da estatura

X-gal - 5-bromo-4-cloro-3-indolil-β-D-galactopiranosídeo

XIC - centro de inativação do cromossomo X

YAC - cromossomo artificial de levedura

Yp - braço curto do cromossomo Y

Yq - braço longo do cromossomo Y

ÍNDICE

Introdução.....	1
1. Estrutura do cromossomo Y	1
2. Papel do cromossomo Y na diferenciação sexual.....	5
3. Genes e pseudogenes no cromossomo Y.....	9
4. Sequências no cromossomo Y	15
5. A sequência Y-1.....	17
Objetivos.....	20
Material e métodos	21
1. Extração em larga escala do plasmídio pIBI-76 contendo o inserto Y-1.....	21
2. Isolamento dos subfragmentos Y-1.a, b1, b2 e c.....	22
3. Análise dos subfragmentos quanto à localização e especificidade.....	23
a. Extração, digestão e eletroforese do DNA genômico	24
b. Transferência pelo método de Southern.....	26
c. Marcação das sondas	25
d. Hibridização.....	26
4. Sequenciamento dos subfragmentos	27
a. Clonagem dos subfragmentos no plasmídio pUC19	27
b. Preparação das células competentes.....	28
c. Transformação bacteriana.....	28
d. Seleção das colônias e mini-preparação.....	29
e. Reação de sequenciamento	29
e.1. Desnaturação	29
e.2. Anelamento	30
e.3. Reação.....	30
e.4. Preparação do gel	31
5. Análise das sequências	32
6. Estimativa do número de cópias	32
7. Técnica de PCR.....	33
8. Conservação da sequência Y-1	34
a. Espécies utilizadas.....	34
b. Extração de DNA de sangue total de macacos	35
c. Extração de DNA de amostras de tecidos de coelho e camundongo	36

Resultados e discussão.....	38
1. Análise quanto à especificidade.....	38
2. Localização da sequência no cromossomo Y	42
3. Estimativa do número de cópias	45
4. Homologias e conservação da sequência	47
5. Sequenciamento do subfragmento Y-1.c, seleção dos <i>primers</i> e padronização da PCR	50
6. Pesquisa de regiões promotoras	53
Conclusões	58
Resumo.....	60
Abstract	61
Bibliografia	62
Apêndice I.....	74
Apêndice II	78

ÍNDICE DAS FIGURAS E TABELAS

Figura 1. Cromossomo Y humano.....	1
Figura 2. Evolução da região pseudoautossômica.....	3
Figura 3. Subdivisão do intervalo 1 do mapa de Vergnaud	7
Figura 4. Principais genes e pseudogenes do cromossomo Y e seus homólogos no cromossomo X.	10
Figura 5. Mapa de restrição do fragmento Y-1.	18
Figura 6. Subfragmentos Y-1.a, Y-1.b1, Y-1.b2 e Y-1.c hibridizados com DNA genômicos de homem (46,XY) e mulher (46,XX) digeridos com as enzimas <i>Hind</i> III, <i>Eco</i> RI, <i>Bgl</i> II e <i>Hae</i> III.	39
Figura 7. Sonda Y-1.b1 hibridizada com DNA genômico de homem (46,XY) e mulher (46,XX) digeridos com <i>Hind</i> III em condições de baixa estringência.	40
Figura 8. Mapa de restrição do subfragmento Y-1.b2.....	41
Figura 9. Subfragmento Y-1.b2B hibridizado com DNA genômico de homem (46,XY) e mulher (46,XX) digeridos com as enzimas <i>Hind</i> III e <i>Bgl</i> II.....	42
Figura 10. Esquema representativo mostrando a porção do cromossomo Y presente em cada linhagem celular e a localização das sondas utilizadas.....	43
Figura 11. Sonda Y-1.c hibridizada com DNA genômico de homem (46,XY), mulher (46,XX), do paciente JPSN e do irmão do paciente JPSN, digeridos com a enzima <i>Eco</i> RI.	45
Figura 12. Hibridização da sonda Y-1.c com filtro contendo 100, 50, 20 e 1 cópia da sequência Y-1, DNA genômico de homem (46,XY) e mulher (46,XX) digeridos com a enzima <i>Eco</i> RI.	46
Figura 13. Sonda Y-1.c hibridizada com DNA 46,XYYY, 46,XYY, 46,XY e 46,XX digeridos com a enzima <i>Eco</i> RI.....	47
Figura 14. Hibridização da sonda Y-1c com (a) espécies animais (b) espécies de macacos.	48

Figura 15. Resultado da reação de PCR	51
Figura 16. Sequência Y-1.c mostrando os <i>primers</i> E1 e E2.	52
Figura 17. Plasmídio pKK232-8 contendo o fragmento Y-1.b2 como inserto.....	54
Figura 18. Esquema das regiões promotoras encontradas na sequência Y-1.b2.....	55
Figura 19. Sequência Y-1.b2 contendo um trecho da sequência do plasmídio pKK232-8.	57
Tabela 1. Cariótipo das linhagens celulares dos pacientes.	23
Tabela 2. Sondas Y-específicas utilizadas na análise das linhagens e dos pacientes...	24
Tabela 3. Resultado das hibridizações do subfragmento Y-1.c com diferentes espécies.....	49

1,

INTRODUÇÃO

INTRODUÇÃO

1. Estrutura do cromossomo Y:

O cromossomo Y é um dos menores cromossomos humanos, possui um tamanho estimado de 60 Mb compondo 1% do genoma de uma célula diplóide. Apresenta uma região heterocromática de tamanho variável, aparentemente sem atividade genética, que representa 50 a 70% do seu total de DNA. Localizada na parte distal do braço longo (Yq), esta região apresenta intensa fluorescência quando corada com quinacrina e observada sob luz UV (COOKE *et al.*, 1979; SCHEMPP & MÜLLER, 1982; CASPERSSON, *et al.*, 1970) (figura 1a).

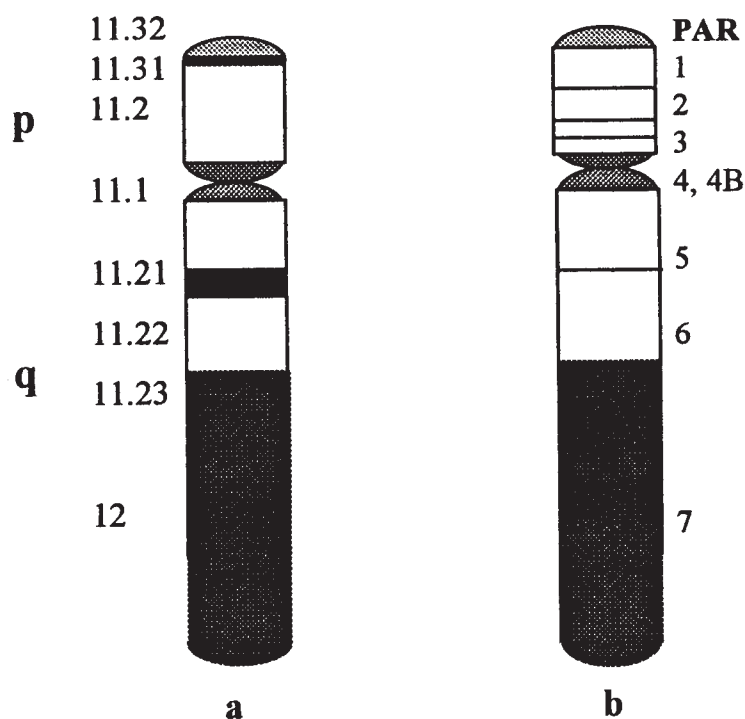


Figura 1 - Cromossomo Y humano; a parte hachurada representa a região heterocromática; (a) bandas G; (b) intervalos do mapa de Vergnaud (VERGNAUD, *et al.*, 1986), PAR - Região pseudoautossômica.

A parte remanescente do cromossomo Y, ou seja, o braço curto (Yp), centrômero e eucromatina distal do braço longo, possui um tamanho mais constante nos indivíduos. Apresenta também, blocos de seqüências homólogas ao cromossomo X, famílias de seqüências Y-específicas repetidas e todos os seus outros genes.

Embora observações morfológicas no cromossomo Y suscitaram interessantes questões, o significado dessas observações só pode ser entendido após estudos moleculares do DNA contido neste cromossomo. Era intrigante o fato de que em fêmeas humanas, todos os cromossomos se emparelham com um homólogo de tamanho e conteúdo genético equivalente mas, os cromossomos X e Y formam um par heteromórfico com padrão de bandas diferentes onde uma pequena porção do cromossomo Y se emparelha com o cromossomo X, além do cromossomo X ser 2 a 3 vezes mais longo que o cromossomo Y. Esta porção de emparelhamento, situada no braço curto dos cromossomos X e Y, sofre recombinação obrigatória na meiose masculina (ROUYER, *et al.*, 1986). Os genes situados nesta região não são estritamente ligados ao sexo e determinam uma herança autossômica, desta forma o adjetivo pseudoautossômico foi introduzido para descrever esse comportamento genético.

A existência de uma região pseudoautossômica (PAR) foi observada pela primeira vez na meiose masculina pela formação de quiasmas na região de emparelhamento dos cromossomos sexuais (KOLLER & DARLINGTON, 1934). Este emparelhamento ocorre no paquíteno e envolve mais da metade do cromossomo Y, sendo a região de emparelhamento muito maior que a região de seqüências homólogas entre os dois cromossomos.

A PAR possui 2,6 Mb e através do seqüenciamento de 1 kb do DNA localizado no ponto de transição da PAR para seqüências cromossomo específicas, quatro regiões distintas foram descritas pela clonagem e estudo deste intervalo e de seqüências próximas:

1. a região pseudoautossômica, apresentando alta homologia entre os dois cromossomos X e Y com somente 2 bases diferentes em cada 245;
2. a inserção de um elemento *Alu* somente no cromossomo Y, sendo que a primeira base da seqüência *Alu* repetida no cromossomo Y é a primeira base que começa a diferir no cromossomo X;

3. uma região de 225 pares de base com homologia parcial (77%) entre os cromossomos X e Y;

4. finalmente, uma região sexo-específica que difere abruptamente entre os cromossomos X e Y (figura 2).

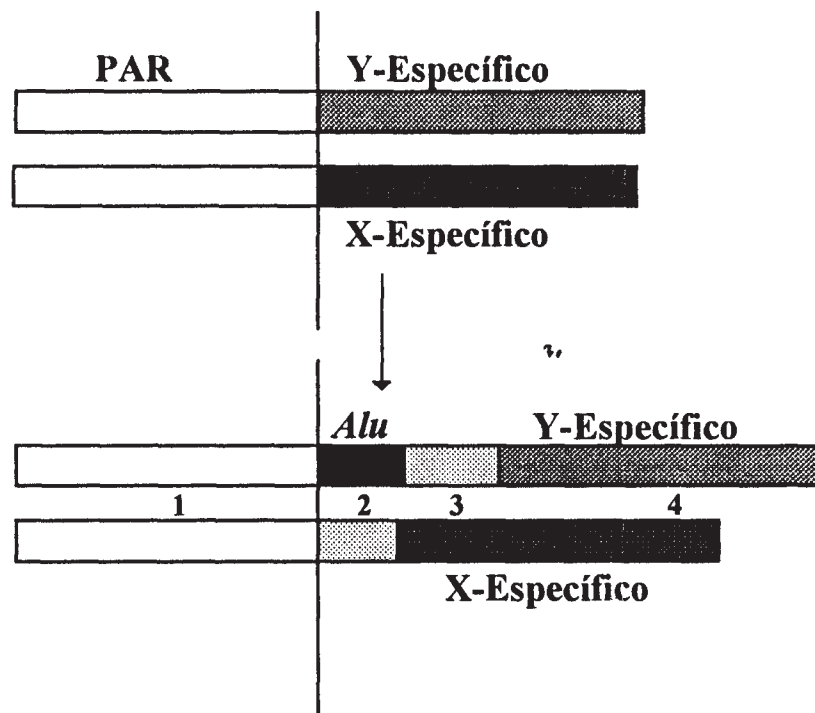


Figura 2- Evolução da região pseudoautossômica e sua estrutura atual.
(ELLIS, *et al.*, 1989).

Uma outra região de homologia foi observada, envolvendo as regiões Xq/Yq. Esta região possui 0,4 Mb mas, nenhuma informação é disponível ainda a respeito de sua estrutura (FREIJE *et al.*, 1992).

Segundo ELLIS *et al.*, 1989, o elemento *Alu*, cujas evidências indicam ser derivado de retro-transposons, foi inserido no cromossomo Y há 25 milhões de anos. Estes dados combinam com a hipótese de que os cromossomos X e Y foram uma vez homólogos como autossomos, com um simples alelo para a determinação sexual (OHNO, 1967).

Por recombinação e acúmulo de mutações no proto-cromossomo Y, outros *loci* se envolveram na determinação sexual, restringindo a recombinação dos cromossomos X e Y. Resultou que a maioria das seqüências que estavam no

cromossomo Y tornaram-se extintas exceto os genes envolvidos na diferenciação sexual.

De acordo com a teoria da evolução, a seleção atua quantitativamente e qualitativamente sob os vários níveis de organização, seja à nível de genoma, indivíduo, espécie, etc. O fenômeno evolutivo observado é resultado de interações positivas e negativas entre os diferentes níveis. Segundo Müller, 1983, existe uma interação positiva em todos os níveis, no que diz respeito à diferenciação em ambos os sexos, favorecendo esta em relação ao hermafroditismo.

Especula-se muito, mas pouco é conhecido, acerca das pressões evolutivas que levaram à divergência do par autossômico que gerou os cromossomos sexuais. É vantajoso para qualquer animal ter cromossomos sexuais distintos, pois evita a recombinação entre eles. A presença dos cromossomos sexuais traz consigo, também, as vantagens da reprodução sexuada, ou seja, a variabilidade genética aumenta devido à recombinação entre os patrimônios genéticos dos progenitores. O aumento da variabilidade genética torna o indivíduo mais flexível às mudanças ambientais sendo a heterozigosidade em muitos casos vantajosa.

A PAR tende, entretanto, a se expandir por recombinação não homóloga, sendo necessários outros rearranjos durante a evolução, como por exemplo, um evento de inversão, para conter essa expansão. Esta região, portanto, poderá se expandir e (ou) contrair durante o processo evolutivo, por isso se apresenta em grande variabilidade entre as espécies (YEN *et al.*, 1988).

O cromossomo Y tem sido estudado ostensivamente pela sua importância, tanto à nível evolutivo quanto em relação ao seu papel na diferenciação sexual. Atualmente ele é dividido em 7 intervalos (VERGNAUD *et al.*, 1986), definidos arbitrariamente segundo padrões de hibridização, obtidos através de 23 fragmentos de restrição Y-específicos com o genoma de 27 indivíduos portadores de alterações estruturais nesse cromossomo.

Deste modo o braço curto (Yp) é dividido em quatro intervalos, excluindo a região pseudoautossômica. A parte eucromática do braço longo (Yq) contém os intervalos 5 e 6 e a parte heterocromática, o intervalo 7. Esta nomenclatura tem

sido utilizada sistematicamente pela maioria dos autores no estudo desse cromossomo (figura 1b).

A viabilidade de indivíduos com deleções no cromossomo Y, relativamente freqüente nas populações humanas, permitiu a construção de vários mapas físicos de deleção. Entretanto, um mapa físico completo deste cromossomo foi construído recentemente, baseado na detecção de STS (*sequence-target sites*), ou seja, pequenas seqüências alvo que são detectadas pela técnica de PCR e mapeadas em um ponto particular do genoma, onde o próprio STS funciona como referência (VOLLRATH *et al.*, 1992).

A ordem das seqüências geradas foi feita pela sobreposição de clones construídos em YAC (cromossomo artificial de levedura) como vetor. Os 160 STS foram obtidos através de centenas de seqüências do cromossomo Y não caracterizadas. Este mapa permite a comparação de mapas feitos por diferentes investigadores, além de sugerir uma interessante estratégia de mapeamento simplificando a análise dos *contigs* de YAC (FOOTE *et al.*, 1992).

2. Papel do cromossomo Y na diferenciação sexual:

Há muito tempo se conhece o papel do cromossomo Y na diferenciação sexual mas, o mecanismo pelo qual o sexo dos indivíduos é determinado, tem sido objeto de especulação desde os tempos de Aristóteles. Até 1900 pensava-se que o sexo dos indivíduos era decidido por fatores ambientais como a nutrição materna. Após a redescoberta das leis de Mendel ficou mais claro que em muitas, mas não em todas as espécies de vertebrados e invertebrados, o sexo era determinado pela constituição cromossômica (McCLUNG, 1902; STEVENS, 1905; WILSON, 1905; MORGAN, 1910; PAGE *et al.*; 1987). Mais tarde, ao ser observado que em insetos com dimorfismo sexual como a *Drosophila*, o sexo era determinado pelo número de cromossomos X independente do número de cromossomos Y, o sexo dos mamíferos, por analogia, também ficou definido desta forma (BRIDGES, 1916).

A existência de cromossomos X e Y em humanos foi demonstrada citologicamente em 1923 (PAINTER, 1923). Em 1959 contudo, estudos de cromossomos sexuais anormais em ratos e humanos (XXY e XO) revelaram a

influência do cromossomo Y na função gonadal masculina independente do número de cromossomos X, já que, embriões carregando pelo menos um cromossomo Y (XY ou XXY) desenvolvem-se como machos e embriões sem cromossomo Y (XX ou X0) desenvolvem-se como fêmeas (WELSHONS & RUSSELL, 1959; FORD *et al.*, 1959; JACOBS & STRONG, 1959).

Até a sétima semana o embrião ainda é sexualmente neutro. A gônada primitiva, formada por células germinativas que migraram do saco vitelino para a superfície das pregas urogenitais, é constituída por células pequenas, em grande número, e por poucas células grandes, as células sexuais primitivas. Esse epitélio germinativo prolifera produzindo cordões sexuais que penetram no mesêquima. Diferenciam-se então, duas regiões, uma cortical e outra medular e nessa fase a gônada é dita bissexual. Na presença do cromossomo Y, haverá desenvolvimento da medula e a gônada primitiva dará origem a um testículo. Caso contrário, na ausência do cromossomo Y, o desenvolvimento do córtex resultará num ovário (WILSON *et al.*, 1981). Desse instante em diante o processo de diferenciação sexual parece ser fundamentalmente hormonal.

Assim, o cromossomo Y, através da sua presença ou ausência constitui o ponto inicial da diferenciação sexual. Entretanto, mesmo nas gônada bissexuais, são requeridas algumas funções controladas por autossomos. Por exemplo, a mutação no cromossomo 17 de embriões de rato XY, resulta no desenvolvimento de ovários (WASBURN & EICHER, 1983). Dessa forma, muitos genes podem atuar em conjunção com o cromossomo Y na diferenciação sexual, incluindo também alguns genes do cromossomo X.

Ainda em 1959, com a demonstração de que a herança do cromossomo Y resulta na determinação do sexo masculino, aventou-se a existência de um ou mais genes nesse cromossomo que induz à diferenciação testicular. Desde então, a natureza bioquímica desse gene, referido como TDF (fator de determinação testicular), tem sido objeto de grande interesse e especulação. Trabalhos correlacionando o fenótipo com alterações estruturais do cromossomo Y permitiram restringir a localização do TDF ao Yp. Entretanto, foi somente após o desenvolvimento de técnicas de clonagem e isolamento de seqüências Y-específicas, que foram alcançados avanços significativos nessa área.

Através de estudos com homens de cariótipo 46,XX e mulheres com cariótipo 46,XY, os quais são resultado de uma transferência accidental de material do cromossomo Y para o cromossomo X na meiose paterna, PAGE *et al.* em 1987, definiram a posição para o TDF dentro de 140 kb, a região 1A2 (figura 3). Nesta região foi localizado o gene ZFY que codifica uma proteína do tipo *zinc finger*, na época, um sério candidato ao TDF, já que uma das mulheres XY estudada não possuía seqüências derivadas deste gene.

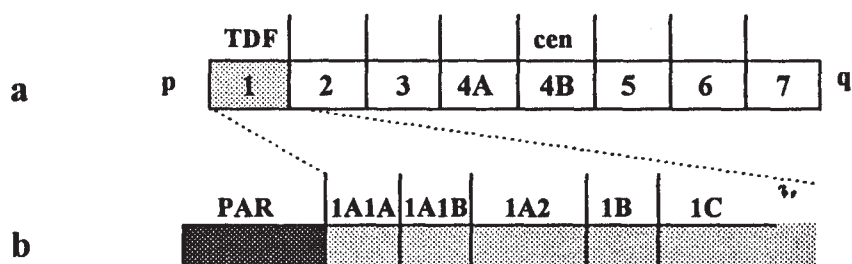


Figura 3 - (a) Mapa de Vergnaud, (b) Subdivisão do intervalo 1 do mapa de Vergnaud, (cen) centrômero, (p) braço curto, (q) braço longo.

Entretanto, posteriormente PALMER *et al.*, (1989), descreveram quatro indivíduos com inversão sexual XX que possuíam seqüências derivadas do cromossomo Y, mas entre as quais não estava incluído o gene ZFY. Analisando o genoma destes indivíduos foi possível definir uma região de 60 kb adjacente ao ZFY, região 1A1 (figura 3), que foi depois reduzida à 35 kb. Nessa região observaram-se muitas seqüências repetidas e dentre elas foi isolado um gene de cópia simples.

Este gene, chamado SRY (região de determinação sexual do Y), codifica uma seqüência de aminoácidos altamente conservada, homóloga à um tipo de proteína Mc requerida na fissão de leveduras de *Schizosaccharomyces pombe*, com a proteína nuclear não histônica HMG (*high mobility group*) e também com dois fatores de transcrição humanos, hUBF e TCF-1 (SINCLAIR, *et al.*, 1990).

Dessa forma, fatos importantes puderam ser evidenciados: o ZFY não é o TDF e o gene SRY, altamente específico, parece até o momento ser consistente com a regra de determinação sexual. Este gene foi também isolado em camundongos, chamado *Sry* e possui uma homologia de 80% com o SRY humano.

Mutações identificadas no gene *Sry* provocaram a incapacidade desse gene para determinar testículos em camundongos, produzindo fêmeas XY. O DNA de camundongos machos normais XY, quando hibridizados com o gene *Sry*, mostra uma banda de 3,5 kb. Esta banda, em DNA de fêmeas mutantes XY, não é observada, uma vez que, esta sequência está alterada (GUBBAY *et al.*, 1990).

Tem sido sugerido que o produto do gene *SRY/Sry* seja um fator de transcrição, ou seja, uma proteína regulatória. Genes regulatórios podem controlar a transcrição por diversos mecanismos. A proteína codificada pela sequência *SRY/Sry* clonada, a qual na verdade acredita-se ser o produto codificado pelo último exon do gene, contém 80 resíduos de aminoácidos, os quais são similares aos encontrados em diversas proteínas conhecidas como fator de transcrição pela função de ligação ao DNA. As primeiras proteínas conhecidas neste grupo foram as chamadas *HMG box* (*high mobility group*). Recentes estudos de homologia do *SRY/Sry* e proteínas com funções de ligação no DNA, mostraram que a proteína codificada pelo gene *SRY/Sry* se liga eficientemente à sequência AACAAAG (HARLEY *et al.*, 1992). Ainda são necessários mais estudos para definir os domínios funcionais de tal proteína, como por exemplo, análise de mutantes criados por mutagênese sítio dirigida. Além disso, seria interessante saber também que genes esta proteína regularia.

Alguns autores partiram por fim, para experimentos que consagrariam o *SRY/Sry* como candidato ao TDF, estudando, por exemplo, o padrão de expressão desse gene em testículo de camundongo. SINCLAIR *et al.*, (1990), mostraram um transcrito de 1,1 kb expresso em testículo de camundongo adulto e GUBBAY *et al.*, (1990), mostraram a expressão do *Sry* nas gônadas masculinas 11,5 dias após o coito, mas não nas gônadas femininas do embrião. Estes resultados são consistentes com o padrão esperado para o gene da determinação testicular, cuja ação induziria à formação de células de Sertoli, e que, na sua ausência, induziria ao aparecimento das células foliculares no ovário.

Um fragmento de DNA de 14 kb contendo o gene *Sry* foi também suficiente para induzir à formação de testículos em embriões transgênicos de camundongos XX, resultando em um fenótipo de inversão sexual. Embora outros genes sejam requeridos no processo de desenvolvimento masculino, os genes *SRY/Sry* têm

mostrado propriedades genéticas e biológicas suficientes para afirmar que este é o suposto TDF (KOOPMAN *et al.*, 1991).

3. Genes e pseudogenes no cromossomo Y:

Na região pseudoautossômica foi isolado e clonado o primeiro gene do cromossomo Y, o **MIC2**, codificando um antígeno de superfície celular de função desconhecida, reconhecido através de anticorpos monoclonais 12E7. O antígeno determinante detectado por este anticorpo é uma proteína expressa em todos os tecidos humanos com possível exceção nos espermatozóides. Este antígeno não é encontrado na superfície celular de ratos, hamster ou camundongos, por isso híbridos celulares humano-roedores têm sido utilizados no estudo da sua expressão. Em híbridos humano-roedores o antígeno é expresso na superfície celular somente quando está presente o cromossomo X humano.

Por análise de deleção, o gene **MIC2** foi mapeado no braço curto do cromossomo X (Xp22.3-pter) e pelo mesmo tipo de estudo, isto é, com híbridos humano-roedor contendo translocações X-Y, o gene foi mapeado também no cromossomo Y (Yq1:1-Ypter), figura 4.

A descoberta de que o antígeno 12E7, que é restrito ao homem, chimpanzé e gorila, é codificado por seqüências em ambos os cromossomos X e Y, levou alguns autores a concluir que o antígeno 12E7 tenha surgido recentemente, ou seja, que tenha havido pelo menos uma troca entre os cromossomos X-Y subsequente à divergência do *Homo sapiens* (GOODFELLOW *et al.*, 1983).

O produto primário do gene **MIC2** foi isolado através de um clone de cDNA pelo sistema de expressão do fago λ gt11 em *E.coli* (DARLING, 1986). Este gene tem sido, uma ferramenta importante no estudo da base molecular da inativação do cromossomo X, associada à uma modificação covalente na estrutura do DNA, já que este gene, e todos os genes da região pseudoautossômica, escapam à inativação do cromossomo X (GOODFELLOW *et al.*, 1988).

O gene para receptor GM-CST (fator de estimulação de colônia granulócito-macrófago) que estimula a proliferação, diferenciação e ativação funcional de granulócitos e macrófagos, também foi localizado na região

pseudoautossômica. Acredita-se que este gene, denominado **CSF2R**, possa estar envolvido na geração de um tipo de leucemia mielóide aguda do subtipo *M2* (GOUGH *et al.*, 1990).

Além dos genes **MIC2** e **CSF2R**, um terceiro *locus* na região pseudoautossômica foi sugerido, que controlaria a estatura (ELLISON *et al.*, 1992). Tem sido proposto que este *locus* estaria próximo à região telomérica mas, HENKE *et al.*, 1991, sugere que ele seja mais proximal. O mais recente gene clonado na PAR é o **XE7**. Embora o gene **XE7**, ainda não esteja precisamente localizado dentro da região pseudoautossômica, é possível que ele seja o gene que controla a estatura. Ele codifica para duas isoformas de proteínas, determinadas por processamento alternativo do RNA. A localização do gene **XE7** concomitante com estudos das proteínas por ele codificadas são ainda necessários para revelar qual seria realmente a função deste gene (ELLISON *et al.*, 1992).

Seguindo-se os três genes presentes na **PAR**, aparecem os genes **SRY** e **ZFY** discutidos no item anterior (figura 4).

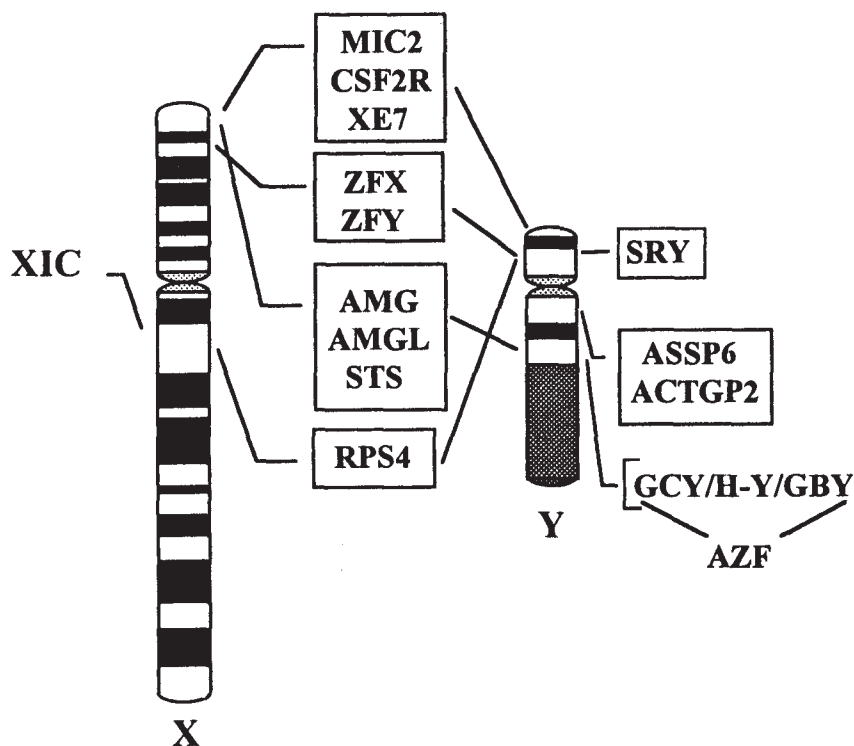


Figura 4 - Principais genes e pseudogenes do cromossomo Y e seus homólogos do cromossomo X.

Ainda no braço curto do Y encontra-se o gene **RPS4Y** com sua cópia correspondente (**RPS4X**), mapeado no braço longo do cromossomo X (Xq13.1), muito próximo ao centro de inativação (**XIC**), figura 4. Ambos os genes, **RPS4X** e **RPS4Y**, codificam isoformas de uma das 33 proteínas existentes na subunidade menor dos ribossomos. A proteína **RPS4** no ribossomo dos mamíferos está localizada na interface das subunidades 40S e 60S (NYGARD & NIKA, 1982; UCHIUMI *et al.*, 1986) onde aparece associada ao fator de iniciação eIF-3. Em mamíferos não se conhece nenhuma outra proteína do ribossomo que possua isoformas ou que seja codificada por mais de um gene.

Desde que, uma das isoformas é codificada pelo gene **RPS4Y**, no cromossomo Y, o qual está presente em machos mas, não em fêmeas, aventou-se a possibilidade de que poderia haver uma diferença estrutural entre os ribossomos de machos e fêmeas. Fêmeas normais, não têm o cromossomo Y e, portanto, não têm o gene **RPS4Y** e sua proteína. Como existem duas isoformas da proteína de **RPS4** por ribossomo, então ribossomos de machos poderiam ser de dois tipos: aqueles contendo o produto do gene **RPS4Y** e aqueles contendo o produto do gene **RPS4X**. Todos os ribossomos das fêmeas, entretanto, conteriam o produto do gene **RPS4X**. É possível que as isoformas codificadas pelos cromossomos X e Y sejam funcionalmente distintas em algum aspecto (FISCHER, *et al.*; 1990).

Acredita-se que a Síndrome de Turner seja resultado da monossomia para um ou mais genes comuns aos cromossomos X e Y. No cromossomo Y este gene tinha sido mapeado nos intervalos 1 e 2 (ROSENFELD *et al.*, 1979; MAGENIS *et al.*, 1984; DITECHE *et al.*, 1986; BLAGOWIDOW *et al.*, 1989; LEVILLIERS *et al.*, 1989). No entanto, em 1990 PAGE *et al.*, através de estudos com fêmeas 46,XY e uma fêmea X,t(Y;22), que não exibem fenótipo de Turner, restringiu o intervalo para a região 1A1B (figura 3). Isto torna os genes **RPS4**, sérios candidatos ao fenótipo da Síndrome de Turner, pois o gene **RPS4Y**, está mapeado justamente na região 1A1B e não há evidências da existência de outros genes neste intervalo, que compreende 90 kb. Além disso, o gene responsável por este fenótipo teria que escapar da inativação senão não haveria diferença entre os fenótipos 45,X e 46,XX e, embora o gene **RPS4X** seja muito próximo ao **XIC**, ele certamente escapa da inativação (FISCHER *et al.*, 1990).

O antígeno de histocompatibilidade H-Y, foi por muitos anos considerado como o indutor de testículos atribuído ao cromossomo Y, já que fêmeas da linhagem isogênica de camundongos C57BL/6, quando enxertadas com pele de macho da mesma linhagem, rejeitavam o transplante (GOLDBERG *et al.*, 1971; WACHTEL *et al.*, 1975b). Esta hipótese foi abandonada desde que demonstrou-se que camundongos possuíam desenvolvimento de testículos mesmo na ausência do antígeno H-Y. LAU, *et al.*, em 1986, demonstraram o envolvimento de um *locus* autossômico na determinação gonadal. Usando anticorpo monoclonal anti-H-Y, uma sequência muito conservada foi isolada de um banco construído a partir de testículos de camundongo. Esta sequência, sem expressão em células ovarianas, foi mapeada no cromossomo 6 humano sugerindo que no cromossomo Y encontra-se somente o gene regulador para o antígeno H-Y, sendo que o gene estrutural estaria localizado no cromossomo 6.

Após a descoberta do gene **SRY**, responsável pela determinação testicular, uma especial importância tem sido dada ao gene que no homem seria responsável pela espermatogênese normal. Tanto dados citogenéticos (TIEPOLO & ZUFFARDI, 1976) como os resultados obtidos com a utilização de sondas de DNA em pacientes azoospermicos (ANDERSON *et al.*, 1988), sugerem a região Yq11, intervalos 5-6 do mapa de Vergnaud, como a localização mais provável do fator de controle da espermatogênese, o **AZF** como foi denominado em humanos.

Atualmente, por estudos moleculares associados à dados citogenéticos, a região mais provável do **AZF** tem se restringido ao intervalo 6 do mapa de Vergnaud, o qual foi dividido em 14 subintervalos designados I a XIV (MA *et al.*, 1992). A busca de microdeleções em pacientes azoospermicos ou com azoospermia severa leva a inferir a existência de mais de um gene no controle da espermatogênese. Microdeleções afetando as regiões proximal ou distal do intervalo 6, normalmente causam uma parada na espermatogênese antes da fase de proliferação celular (VOGHT *et al.*, 1991b, 1992), enquanto que microdeleções centrais no intervalo 6 normalmente estão associadas à uma parada da espermatogênese durante o estágio de paquíteno (RAMOS *et al.*, 1993). Estas observações indicam que o cromossomo Y humano é portador de uma série de genes envolvidos no controle da espermatogênese, cada um com expressão em diferentes e específicos estágios do processo de formação de gametas.

Ainda nos intervalos 5-6, dois genes têm sido propostos, um que codifica para um fator de crescimento (GCY), proximal ao AZF, e outro envolvido com o crescimento dentário, podendo inclusive se tratarem os dois de um mesmo gene (ALVESALO & DE LA CHAPELLE, 1981).

Por outro lado, a constatação de que pacientes com disgenesia gonadal e presença de cromossomo Y apresentam um risco aumentado para desenvolver gonadoblastoma, levou PAGE (1987) a sugerir que nestes pacientes o gene de controle da espermatogênese poderia atuar como um oncogene, causando desenvolvimento do tumor. O autor nominou tal gene hipotético como GBY significando *locus* gonadoblastoma do cromossomo Y, situado no intervalo 5-6 do mapa de Vergnaud. Tendo em vista que os genes GCY, GBY, H-Y e AZF não foram ainda isolados e clonados na espécie humana, e que a existência destes genes é inferida por suas funções, persiste a dúvida se realmente tratam-se de vários genes ou de um único com várias funções.

Além dos genes aqui citados, alguns pseudogenes também são encontrados no cromossomo Y. O primeiro mapeado foi o STS que codifica a enzima esteróide sulfatase, cuja cópia funcional está localizada no cromossomo X. Nesse cromossomo o gene está situado na região pseudoautosômica, mas no cromossomo Y, ele se encontra no braço longo, no intervalo 6 do mapa de Vergnaud (FRASER *et al.*, 1987).

O índice de similaridade entre o STS-X e STS-Y, está entre 85-94%, não havendo diferença deste índice para introns ou exons. Isto indica que o *locus* STS-Y se tornou não funcional simultaneamente com a divergência evolutiva entre os dois genes.

O pseudogene STS-Y, em processo de degeneração, não possui nenhuma atividade de transcrição, já que a divergência ocorrida gerou numerosos códons de terminação na cadeia de leitura da sequência. Além de possuir deleções, com relação à cópia funcional (STS-X), também exibe inserções. Essas inserções constituem-se principalmente de sequências LINEs (*long interspersed elements*), descritas no próximo item. Essas sequências contêm características estruturais de retro-transposons (YEN *et al.*, 1988).

A deficiência da enzima esteróide sulfatase no homem é uma doença que afeta 1 em 5000-6000 homens, em diversos grupos étnicos e localizações geográficas. A alta prevalência desta doença pode ser explicada não por mecanismos clássicos de mutações, mas pelo evento de recombinação desigual X/Y resultando a perda do *locus* STS ligado ao cromossomo X ou na troca pela cópia não funcional do cromossomo Y (YEN *et al.*, 1988).

Uma sequência de cDNA, que foi inicialmente clonada em bovino, codifica uma proteína extracelular responsável pelo desenvolvimento dentário, a qual se expressa exclusivamente nas células germinativas do dente (SNEAD *et al.*, 1985 & SHIMOKAWA *et al.*, 1987). Posteriormente, o gene responsável por esta proteína foi mapeado no braço curto do cromossomo X (AMG; Xp22.1-22.3) e o pseudogene correspondente, próximo ao centrômero do cromossomo Y (AMGL) (NAKAHORI *et al.*, 1991). A deficiência de tal proteína, denominada amelogenina, causa a doença genética amelogenese imperfeita (AI) cujo padrão de herança é geneticamente heterogêneo (WITKOP & RAO, 1971; McKUSICK, 1988).

Além da hipótese de que as sequências dos cromossomos X e Y divergiram há 25 milhões de anos, YEN *et al.* (1988), sugerem que um evento de inversão ocorreu no cromossomo Y há 40 milhões de anos, baseado na localização do gene STS-X e do pseudogene STS-Y. O gene STS-Y estaria inicialmente na região pseudoautossômica e com o evento de inversão, uma parte teria sido perdida, gerando a cópia não funcional no cromossomo Y. Os genes AMG e AMGL foram descritos posteriormente e, estão próximos do STS em suas cópias nos cromossomos X e Y. Isto confirma a hipótese de que ambos os genes AMGL e STS-Y em humanos, pertenceram um dia à região pseudoautossômica e que realmente ocorreu tal inversão no cromossomo Y.

Em camundongo o gene análogo, o *Sts*, ainda se encontra na região pseudoautossômica, já que não ocorreu a inversão. O evento de inversão é reafirmado quanto se observa a localização do gene *TDF*, que em humanos está próximo à região pseudoautossômica, mas em camundongo o *Tdy* ainda está no braço longo.

Outros pseudogenes são encontrados no intervalo 5 do cromossomo Y, como o gene que codifica a enzima arginosuccinato sintetase (ASSP6), o qual participa do ciclo da uréia cujo gene funcional é expresso ativamente no braço longo do cromossomo 9 (9q34-qter) (DAIGER *et al.*, 1982). Nesta mesma região do Y, encontra-se também o pseudogene relacionado ao gene gama-actina do citoesqueleto (ACTGP2) (GOMEZ-PEDROZO *et al.*, 1987).

Uma outra sequência de DNA, com restrição *Eco* RI, Y-específica com 2,8 kb, denominada 49f, que poderia ser um pseudogene, foi mapeada no intervalo 6 (BISHOP *et al.*, 1984). Esse fragmento foi identificado por hibridização com um transcrito de mRNA de 4 kb de testículo humano codificado por um gene localizado no cromossomo 3 (SEBOUN *et al.*, 1986). A sequência 49f se encontra repetida no cromossomo Y cerca de 20-30 cópias e é altamente polimórfica (NGO *et al.*, 1986).

4. Sequências no cromossomo Y:

Muitas sequências no cromossomo Y humano possuem homologia com sequências do cromossomo X e autossomos humanos mas, a maioria das sequências não possui homologia com o cromossomo Y de outros primatas. Isto se deve talvez ao fato de que estas sequências foram adquiridas em humanos quando ocorreu a divergência entre humanos e os outros primatas há milhões de anos.

A evolução de organismos eucarióticos resultou no aumento de sequências no genoma, sequências estas que não codificam para nenhuma proteína e, para as quais, têm-se especulado a função de regulação da expressão gênica. Desta forma, encontram-se espalhados no genoma eucariótico várias famílias de sequências de DNA repetido.

O cromossomo Y é o cromossomo que possui sequências repetidas em maior número e variedade, é também altamente heterocromático e, pelo menos 50% desse cromossomo, é composto por sequências repetidas. Em termos gerais, são três as categorias de sequências de DNA repetidas:

- sequências SINEs (*short interspersed elements*), da qual a maior família é representada por sequências *Alu*, que estão repetidas 300.000 vezes em todo genoma (RINEHART *et al.*, 1981). No cromossomo Y ela está presente numa

seqüência repetida em tandem, isolada por sítios de restrição *Msp* I, possuindo na verdade somente 60% de homologia com as seqüências *Alu* espalhadas nos outros cromossomos (SMITH *et al.*, 1987);

- seqüências LINEs (*long interspersed elements*), também designadas de família *Kpn*, estão vastamente espalhadas no genoma podendo atingir até 6 kb, e no cromossomo Y são indistinguíveis dos outros cromossomos (SCHMECKPEPER *et al.*, 1981);

- seqüências repetidas em tandem, também chamadas de DNA satélites, dentre as quais existem as seqüências Alfóides, que são seqüências cromossomo-específicas com uma periodicidade particular para cada cromossomo (MITCHELL *et al.*, 1985; WOLFE *et al.*, 1985). No cromossomo Y o bloco destas seqüências, situado na região centromérica, possui aproximadamente 400 kb, consistindo de unidades repetidas em tandem em intervalos de 5 - 7 kb que, por sua vez, é formado por subunidades de 170 bp também repetidas em tandem (TYLER-SMITH & BROWN, 1987).

Entretanto, as principais seqüências Y-específicas repetidas são fragmentos espaçados por sítios *Hae* III de 3,4 e 2,1 kb (COOKE, 1976). Estas seqüências não são homólogas entre si e constituem 70% das seqüências repetidas nesse cromossomo. Elas se encontram também espalhadas por todo o genoma mas, devido ao seu grande número de cópias, são muito mais específicas para o Y do que para outros cromossomos.

O DNA satélite *Bkm* (*Banded krait minor satellite DNA*), associado ao sexo heterogamético em cobras, está presente em muitos genomas eucarióticos, consistindo de dinucleotídeos e tetranucleotídeos repetidos de bases GACA e GATA (ARNEMANN, *et al.*, 1986). No cromossomo Y as seqüências *Bkm* são encontradas na região pericentromérica (SINGH & JONES, 1986)

Seqüências que são repetidas do cromossomo Y humano têm sido encontradas, em autossomos de machos e fêmeas de gorila mas, estas mesmas seqüências não têm sido relatadas em chimpanzés. Se o homem e o chimpanzé são parentes mais próximos do que o homem e o gorila, como tem sido mostrado em estudos de reassociação de DNA (BENVENISTE *et al.*, 1976) e estudos de bandas cromossômicas (YUNIS *et al.*, 1982), a evolução de tais seqüências deve ter envolvido um processo de redistribuição e perda.

Se tais seqüências estavam presentes no genoma do ancestral comum ao homem, chimpanzé e gorila, e em alguns cromossomos autossômicos dos gorilas, pode-se supor que no homem essas cópias presentes no cromossomo Y têm sido multiplicadas e mantidas como um bloco homogêneo de seqüências por *crossing-over* desigual. No chimpanzé, entretanto, tais blocos não são mais encontrados e essas cópias acham-se dispersas com outras seqüências.

A hipótese de *crossing-over* desigual pode explicar a manutenção da homogeneidade das seqüências repetidas em um simples cromossomo, mas é difícil imaginar como este mecanismo poderia operar na distribuição de diferentes cromossomos. Deve ter havido algum mecanismo de transferência submetendo tais seqüências à conversão gênica (COOKE *et al.*, 1982).

A aquisição de técnicas de DNA-recombinante para isolar e caracterizar segmentos de DNA de cópia simples no cromossomo Y, tem permitido definir melhor a natureza das seqüências nesse cromossomo, que tem tão poucos genes mapeados quando comparado com qualquer outro cromossomo humano.

5. A seqüência Y-1:

A determinação sexual em mamíferos é sem dúvida o principal papel do cromossomo Y, mas entender o seu atraente mecanismo de regulação tem sido um desafio aos pesquisadores. Caracterizar segmentos de DNA significa uma minúscula, mas importante contribuição para o entendimento do genoma humano.

O isolamento de seqüências Y-específicas tem despertado grande interesse devido às suas aplicações em estudos clínicos, na determinação da evolução deste cromossomo, na compreensão da sua estrutura e na detecção de genes. Até o momento, foram mapeadas mais de 200 seqüências de DNA nesse cromossomo (GOODFELLOW & WEISSENBAACH, 1991).

Na expectativa de se isolar novas seqüências Y-específicas e possíveis genes, procedeu-se a triagem de um banco do cromossomo Y, construído pelo método de citometria de fluxo na Divisão de Ciências Biomédicas do Laboratório Nacional Lawrence Livermore, Califórnia. A partir desse banco, o fragmento Y-1

foi isolado, e clonado no plasmídio pBI-76 para posterior caracterização (FARAH *et al.*, 1992).

A sequência Y-1 de aproximadamente 3,2 kb, quando digerida com as enzimas de restrição *Bam* HI e *Eco* RI, exibe o mapa de restrição ilustrado na figura 5, o qual foi determinado por NOVAIS (1990).

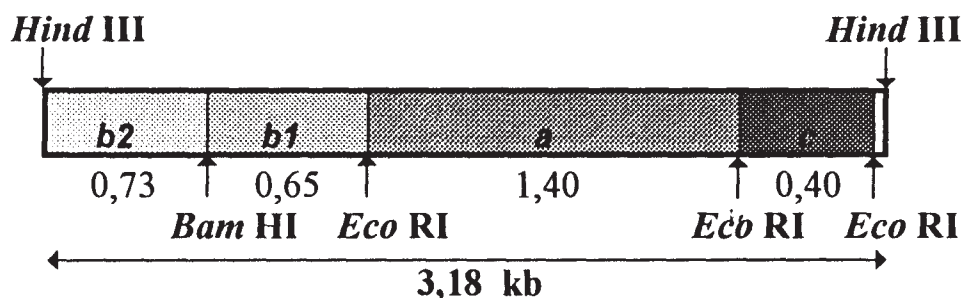


Figura 5 - Mapa de restrição do fragmento Y-1.

Embora a sequência Y-1 hibridize com DNA humano de origem masculina e feminina, algumas bandas Y-específicas são observadas após a digestão do DNA genômico com *Bgl* II, *Xmn* I, *Hinc* II e *Taq* I. O fragmento Y-1, portanto, representa uma sequência repetida, com características semelhantes à muitas outras encontradas no cromossomo Y, ou seja, apresenta bandas Y-específicas e bandas que mostram homologia com autossomos e cromossomo X. Além disso, a sequência Y-1 quando hibridizada com DNA genômico de 15 homens e 15 mulheres normais não aparentados, digeridos com diversas enzimas de restrição, revelou um padrão consistente de bandas não tendo sido detectado portanto, qualquer polimorfismo. Com o objetivo de aumentar a especificidade para o cromossomo Y, NOVAIS (1990) hibridizou separadamente cada um dos subfragmentos determinados no mapa de restrição e observou que o subfragmento Y-1.b1 era o mais específico em condições de alta estrigência.

O fragmento Y-1, foi também subclonado no plasmídio pKK232-8. Este vetor é chamado "caça-promotor", pois, além do gene que confere à bactéria transformada resistência à ampicilina, ele possui o gene para resistência ao cloranfenicol (CAT). Entretanto, este gene não apresenta sua região promotora. Quando um fragmento de DNA é clonado neste plasmídio anteriormente ao gene

CAT, se existir alguma região promotora neste inserto, o gene poderá ser transcrito, conferindo à bactéria transformada resistência também ao cloranfenicol (BROSIUS, 1984).

O fragmento Y-1, ao ser clonado no plasmídio pKK232-8, conferiu à bactéria resistência ao cloranfenicol, sugerindo a possibilidade de conter uma região promotora. Pela posição de clonagem foi demonstrado que o subfragmento Y-1.b2 está localizado imediatamente anterior ao gene CAT sendo, portanto, o mais provável candidato a conter a região promotora (NOVAIS, 1990).

Pelas características aqui apresentadas, a seqüência Y-1 merece um estudo mais aprofundado, constituindo o objetivo principal deste trabalho. Além da sua potencial importância em diagnósticos clínicos, guarda grandes expectativas na localização de um gene devido a possível presença de regiões promotoras. Por outro lado, o sequenciamento do fragmento Y-1 permitiria a determinação de *primers* para a amplificação dessa seqüência pela técnica de PCR (reação em cadeia da polimerase), bem como a identificação de seqüências repetidas e específicas, abrindo novos caminhos para estudos evolutivos, de expressão e dos produtos gênicos.

OBJETIVOS

OBJETIVOS:

1. Localizar o fragmento Y-1 no cromossomo Y para estudos clínicos em indivíduos portadores de aberrações estruturais nesse cromossomo.
2. Estimar o número de cópias do subfragmento mais específico bem como sua distribuição no cromossomo Y.
3. Estudar a homologia e conservação desta sequência com DNA genômico de algumas espécies animais.
4. Sequenciar um subfragmento com vistas à construção de *primers* que permitam a amplificação Y-específica pela técnica de PCR para aplicações em diagnóstico clínico.
5. Otimizar as condições da reação de PCR de forma a amplificar especificamente o DNA genômico de origem masculina.
6. Pesquisar a presença de possíveis regiões promotoras no subfragmento Y-1.b2.

MATERIAL E MÉTODOS

MATERIAL E MÉTODOS

1 -Extração em larga escala do plasmídio pIBI-76 contendo o inserto Y-1:

Inoculou-se 10ml de meio LB (10g bacto-triptone; 5g extrato de levedura; 10g NaCl/litro de solução, pH7,0) contendo 50µg/ml de ampicilina a partir de uma única colônia da bactéria transformada com o plasmídio pIBI-76 (APÊNDICE I) que contém como inserto o fragmento Y-1. A cultura foi mantida durante à noite a 37°C sob agitação a 250rpm.

Foram então inoculados em 500ml do mesmo meio LB com ampicilina, 500µl da cultura obtida na etapa anterior. Incubou-se à noite a 37°C sob agitação a 300rpm.

Cada 250ml do cultivo foi centrifugado a 7000rpm em rotor Beckman JA-14 por 10 minutos a 4°C. Para provocar a lise das células o sedimento obtido foi ressuspenso em 2ml de solução I (tris-HCl 1M; EDTA 0,5M pH8,0; glicose 2M). A suspensão foi transferida para tubo de polipropileno e acrescentou-se mais 1ml de solução I contendo 25mg/ml de lisozima. Manteve-se à temperatura ambiente por 10 minutos.

Adicionou-se 10ml de solução II (NaOH 2N; SDS 10%) recentemente preparada, o conteúdo foi misturado repetidas vezes lentamente por inversão e mantido no gelo por 10 minutos.

A cada tubo adicionou-se 7,5ml de solução gelada de acetato de potássio 3M. Manteve-se no gelo por mais 10 minutos. O precipitado floculoso formado foi separado por centrifugação em rotor Beckman JA-20 a 12.000rpm por 20 minutos a 4°C. O sobrenadante foi filtrado com 4 camadas de gaze e o filtrado distribuído em dois tubos novos de centrífuga.

Adicionou-se 0,6 volumes de isopropanol em cada tubo. Após misturar bem, manteve-se à temperatura ambiente por 10 minutos. Centrifugou-se então, a 11.000rpm em rotor Beckman JA-20 por 10 minutos à temperatura ambiente. O sedimento foi lavado uma vez com etanol 70% e depois seco.

O material obtido foi ressuspensionado com 1ml de TE (tris-HCl 10mM, pH7,5; EDTA 1mM) para cada tubo. A concentração de DNA foi lida em espectrofotômetro em 260nm.

2 - Isolamento dos subfragmentos Y-1.a, Y-1.b1, Y-1.b2 e Y-1.c:

Após a extração em larga escala do plasmídeo pIBI-76 contendo o fragmento Y-1, o plasmídeo foi digerido com as enzimas de restrição *Hind* III, *Bam* HI e *Eco* RI, para se obter cada subfragmento de interesse, segundo o mapa de restrição do fragmento Y-1 (figura 5).

Em cada reação, utilizou-se um excesso de 5 vezes de enzima e foram digeridos 50µg do plasmídeo. Os produtos de digestão foram separados em gel de agarose e os subfragmentos foram purificados com auxílio do kit *Gene clean* (laboratório Bio 101 Inc., cat. 3105), da seguinte forma:

A banda correspondente foi excisada do gel de agarose com o auxílio de um bisturi estéril, colocada em um tubo tipo Falcon e pesada. Foram adicionados 2,5 volumes de solução de NaI (90,8g NaI; 2g Na₂SO₃/100ml de solução) em relação ao peso obtido e incubou-se 5 minutos a 55°C.

Adicionou-se a esse tubo uma resina composta por uma matriz silicosa denominada *Glassmilk*. Para cada 5µg de DNA a ser recuperado teoricamente, acrescentaram-se 5µl da resina e para cada 0,5µg de DNA adicional, mais 1µl de resina. O material foi deixado durante à noite a 4°C.

Na manhã seguinte, o material foi centrifugado e o sedimento foi lavado por 3 vezes com a solução *New wash* que acompanha o kit. Após a última lavagem com *New wash* o DNA foi eluído da resina em quantidade apropriada de tampão TE (para cada 10µl de *Glassmilk* colocou-se 10µl de TE) e incubado a 60°C por 15 minutos. Essa suspensão foi centrifugada a 12000 rpm por 10 minutos e o DNA foi recuperado no sobrenadante. Parte desse DNA foi colocado em um mini-gel para estimar a concentração.

3 - Análise dos subfragmentos quanto à localização e especificidade:

Os subfragmentos isolados foram analisados quanto à sua especificidade para o cromossomo Y através da hibridização com DNA genômico de homem e mulher normais. Para a localização do fragmento Y-1 no cromossomo, foram hibridizados com os subfragmentos isolados, filtros de DNA genômico extraído de leucócitos periféricos de indivíduos normais do sexo masculino e feminino, assim como de indivíduos com aberrações estruturais no cromossomo Y.

Algumas linhagens celulares de indivíduos com aberrações estruturais no cromossomo Y foram obtidas no Human Genetic Repository, Copewood St., Camden NY 08103 (GM6967, GM1709, GM9403). A amostra do paciente J.Reed foi cedida pela Unidade de Genética Reprodutiva da Universidade da Califórnia, os pacientes DIC(Y) e PDR foram cedidos pelo Departamento de Imunologia do ICB da USP, o homem XX, o paciente BR e JPSN, pelos Ambulatórios do Hospital de Clínicas da UNICAMP.

A tabela 1 apresenta os cariótipos descritos em cada linhagem ou paciente e a diferenciação sexual observada.

LINHAGEM	CARIÓTIPO	GENITÁLIA EXTERNA
GM6967	45,X/46,X,dic(Y)(p11)/47,X,dic(Y),+dic(Y)	Ambígua
GM1709	46,X,dic(Y)(p11.2)	Feminina
GM9403	46,X,t(X;Y)(p22.3;q11.2)	Feminina
DIC(Y)	46,X,dic(Y)(q11.2101)	Masculina
J.REED	46,X,t(Y;D?)(p12;p12)	Masculina
PDR	45,X/46,XYnf	Feminina
HOMEM XX	46,XX	Masculina
JPSN	46,XY	Masculina
BR	45,Xt,(Y;15)(?;Yq11.23)	Masculina

Tabela 1 - Cariótipo das linhagens celulares e pacientes.

Para melhor caracterizar as alterações estruturais do cromossomo Y, o DNA dos pacientes e das linhagens celulares foram também analisados com

sondas já bem caracterizadas. A tabela 2 mostra o intervalo detectado por cada uma dessas sondas com a enzima usada para a digestão do DNA genômico.

SONDA	ENZIMA	LOCALIZAÇÃO	REFERÊNCIA
pDP1007	<i>Eco</i> RI	Yp(1A)	PAGE <i>et al.</i> , 1987
pDP105	<i>Eco</i> RI	YP(3)	PAGE (comunicação pessoal)
pDP97	<i>Eco</i> RI	Ycen(4B)	WOLFE <i>et al.</i> , 1985
pY-5	<i>Eco</i> RI	Yq(6)	FARAH <i>et al.</i> , 1992

Tabela 2 - Sondas Y-específicas utilizadas na análise das linhagens e dos pacientes.

a) Extração, digestão e eletroforese do DNA genômico:

Utilizando-se como anticoagulante o EDTA (1,2mg/ml), foram coletados 15ml de sangue periférico dos indivíduos de interesse. O plasma foi descartado após centrifugação por 10 minutos a 2500 rpm a 4°C. O sedimento foi ressuspenso em 9 volumes de tampão A (10ml triton-X 100; 5ml tris-HCl 2M pH7,5; 109,55g sacarose; 1,02g MgCl₂/litro de solução) e mantido 5 minutos no gelo. Centrifugou-se novamente a 3000 rpm por 15 minutos a 4°C. Ressuspendeu-se em PBS (NaCl 137mM; KCl 2,7mM; Na₂HPO₄.7H₂O 4,3mM; KH₂PO₄ 1,4mM) resultando um volume final de 1,5ml. O material foi então colocado no Extrator de Ácidos Nucleicos modelo 340A da *Applied Biosystems*.

A extração de DNA de cultura de células, no caso das linhagens celulares, foi realizada a partir de culturas com aproximadamente 90% de confluência em frascos com 850cm². As células foram lavadas com 30ml de tampão PBS e foram removidas gentilmente do fundo do tubo. A suspensão foi transferida para um tubo tipo Falcon e centrifugada a 2500rpm a 4°C por 15 minutos.

O sobrenadante foi descartado e as células foram ressuspendidas para extração automática em um volume final de 1,5ml de tampão de lise que acompanha o Extrator de Ácidos Nucleicos modelo 340A da *Applied Biosystems*.

O DNA obtido foi então ressuspenso em quantidade apropriada de TE e a concentração foi lida em espectrofotômetro. A seguir, 10µg de DNA genômico foram digeridos durante à noite com a enzima de restrição apropriada de modo a manter um excesso de 5 vezes de enzima e 200µl de volume final da reação. Os fragmentos digeridos foram então precipitados com 0,1 volume de acetato de sódio 2M e 2,5 volumes de etanol absoluto e incubados durante à noite a -20°C. O sedimento é recuperado por centrifugação e lavado 2 vezes com etanol 70%.

As amostras foram ressuspendidas em 15µl de TE mais 3µl de azul de bromofenol e aplicadas em um gel de agarose 0,8% preparado com TBE (tris-base 0,089M; ácido bórico 0,089M; 0,75g EDTA). A eletroforese foi feita à uma voltagem média de 40 volts, em tampão TBE. Como marcador de peso molecular utilizou-se DNA de fago λ digerido com enzima de restrição *Hind* III.

b) Transferência pelo método de Southern:

Antes da transferência pelo método de Southern (Southern, 1975) dos fragmentos de DNA contidos no gel para membrana de nylon (S&S Nytran lote nº04199, Keene nit03431), procedeu-se a desnaturação do gel de acordo com o seguinte tratamento:

1. lavagem com solução de HCl 0,25N por 15 minutos;
2. duas lavagens com solução alcalina (NaOH 2N, NaCl 1,5N) por mais 15 minutos;
3. duas lavagens com solução neutralizante (NaCl 3M, tris-HCl 0,5M) por 40 minutos.

Após este tratamento o gel foi colocado em uma cuba contendo solução de 20xSSC (NaCl 3M, citrato de sódio 0,3M) e procedeu-se a transferência dos fragmentos de restrição nele contido com o auxílio de pontes de papel Whatmann 3MM previamente embebidas em uma solução de 20xSSC. Sobre o gel colocou-se uma membrana de nylon umidecida com água estéril, mais duas folhas de papel Whatmann e mais 10cm de toalhas de papel absorvente. Sobre todo o conjunto acrescentou-se um peso homogeneamente distribuído. Após um período de 12 a 18 horas o sistema foi desmontado e a membrana foi secada por duas horas a 80°C.

c) Marcação das sondas:

A marcação dos subfragmentos foi feita pelo método de *random primer* (FEINBERG & VOGELSTEIN, 1983) com nucleotídeos radioativos adquiridos pela Amersham Internacional, Amersham UK PB10204. A reação foi realizada da seguinte forma:

Foram diluídos em água estéril para um volume final de 9µl, 100ng de DNA a ser marcado, e a sonda foi desnaturada a 95°C por 10 minutos. Acrescentou-se 1µl de cada dNTP exceto aquele que foi usado marcado, 2µl da mistura de reação que consiste numa mistura de oligonucleotídeos de 6bp, 5µl do nucleotídeo marcado com ^{32}P e, finalmente, 1µl da enzima de polimerização *Klenow*. O volume final da reação foi de 20µl e esta foi incubada por no mínimo 30 minutos a 37°C. Para parar a reação, utilizou-se 130µl de STE (tris-HCl 10mM, pH7,5; NaCl 10mM; EDTA 1mM) e passou-se por uma coluna de sílica G-50 saturada com TE para retirar os nucleotídeos não incorporados.

d) Hibridização:

Após a transferência pelo método de Southern, o filtro obtido foi colocado em um saco plástico onde foi adicionado 10ml de uma solução de pré-hibridização que tem a seguinte composição:

formamida desionizada	50ml
20xSSC	25ml
tris-HCl 2M pH 7,5	1ml
100x <i>Denhart's</i>	10ml
SDS 10%	10ml
q.s.p. H ₂ O	100ml

100x *Denhart's*
 10g Ficoll 400
 10g polivinilpirrolidona
 10g albumina bovina
 q.s.p. H₂O 500ml

A pré-hibridização foi feita por no mínimo 4 horas a 42°C, após este período a solução foi descartada. A solução foi trocada por uma solução de hibridização semelhante à solução de pré-hibridização, isto é, o mesmo tampão onde a concentração da solução de *Denhart's* é diminuída de 1 para 0,2 vezes. Além disso nesta solução são adicionados sulfato de dextrana numa concentração final de 25% e a sonda, previamente desnaturada a 95°C por 10 minutos (contida num volume total de 150µl). O material foi incubado então por 16 horas à 42°C sob agitação.

A lavagem dos filtros foi realizada em condições de alta estringência. Após retirar a sonda os filtros foram lavados por duas vezes por 15 minutos à temperatura ambiente, com uma solução contendo 2xSSC e 0,1% de SDS.

Uma terceira lavagem foi feita por 30 minutos a 55°C com uma solução contendo 0,1% de SSC e 0,1% de SDS. Finalmente, o filtro foi lavado rapidamente com uma solução de 0,1% de SSC e envolvido em vitafilme. Alguns filtros foram lavados em condições de baixa estringência onde a terceira lavagem foi feita com a solução 2xSSC e 0,1% de SDS à temperatura de 50°C.

A membrana foi a seguir exposta à um filme X-OMAT AR (Kodak diagnostic film cat 1651454) na presença de placas intensificadoras (Quanta III - Sigma) por períodos variáveis de 3 a 15 dias a -70°C.

4. Sequenciamento dos subfragmentos:

a) Clonagem dos subfragmentos no plasmídeo pUC19:

Para a determinação da sequência, os subfragmentos foram clonados no plasmídeo pUC19 (APÊNDICE I) de dupla fita o que permite o sequenciamento nos dois sentidos. O vetor pUC19 (2686bp) foi preparado através da digestão com a enzima de restrição correspondente ao subfragmento sequenciado.

Para a reação de ligação utilizou-se 100ng de DNA do vetor e uma quantidade de subfragmento correspondente a 3 vezes o número de moles de DNA do vetor. Utilizou-se 1 unidade da enzima T4 DNA *ligase* e um volume final de reação de 10µl. A reação foi incubada durante à noite a 14°C.

b) Preparação das células competentes:

Os produtos das reações de ligação do plasmídeo pUC19 com cada um dos subfragmentos foram transformados pelo método modificado de MANDEL, M. & HIGA, A. (1970) na linhagem DH5 α da bactéria *E.coli* (APÊNDICE I).

Para preparação das células competentes uma única colônia de *E.coli* DH5 α foi crescida em 5ml de meio LB por 16 horas a 37°C. Inoculou-se 500ml de meio LB com 1ml do pré-inóculo obtido durante à noite. Incubou-se sob agitação moderada (250rpm) até atingir a densidade aproximada de 5×10^7 células/ml ($DO_{550}=0,2-0,4$).

A cultura foi transferida para tubos de polipropileno e resfriada por 10 minutos no gelo, em seguida centrifugada a 3000 rpm por 10 minutos a 4°C. O sobrenadante foi descartado e o sedimento ressuspensionado em 1/3 do volume inicial da cultura (33ml) em solução SB (1,85g KCl; 2,22g $MnCl_2 \cdot 4H_2O$; 0,38g $CaCl_2 \cdot 2H_2O$; 5ml K-MES 0,5M pH 6,3/250ml solução), todas as suspensões foram transferidas para um mesmo tubo e incubadas por uma hora no gelo. A suspensão celular foi centrifugada novamente à 3000 rpm por 10 minutos a 4°C e o sobrenadante foi descartado. As células foram ressuspensionadas em um 1/12,5 do volume da cultura original (8ml) em solução SB e separadas em alíquotas de 200 μ l em tubos de microcentrífuga.

c) Transformação bacteriana:

Às células competentes, adicionou-se 2 μ l da reação de ligação. Misturou-se gentilmente e incubou-se no gelo por uma hora. Incubou-se a 42°C por 90 segundos e transferiu-se imediatamente para o gelo. Adicionou-se a seguir 0,8ml de meio LB e incubou-se a 37°C.

Uma alíquota de 200 μ l das células foi espalhada em placas de meio LB contendo ampicilina (50 μ g/ml) e X-gal previamente preparadas, e incubou-se durante a noite a 37°C. As colônias transformadas são reconhecidas pela cor branca, uma vez que as não transformadas desenvolvem coloração azul.

d) Seleção das colônias e mini-preparação:

Inoculou-se 5ml de meio LB contendo ampicilina (50µg/ml) com uma única colônia branca. Incubou-se durante a noite a 37°C sob agitação. Transferiu-se para um tubo de microcentrifuga 1,5ml da cultura e centrifugou-se por 30 segundos.

O sobrenadante foi descartado e o sedimento ressuspendido em 200µl de solução I (tris-HCl 25mM; glicose 50mM; EDTA 10mM). Adicionou-se 400µl de solução II (NaOH 0,2N; SDS 1%) recentemente preparada. Após misturar por inversão manteve-se por 5 minutos no gelo. Adicionou-se 300µl de solução III gelada (60ml acetato de potássio 5M; 11,5ml ácido acético glacial; 28,5ml de água), misturou-se bem e manteve-se no gelo por mais 10 minutos.

Centrifugou-se por 5 minutos à temperatura ambiente e transferiu-se 800µl do sobrenadante para outro tubo de microcentrifuga sem remover o precipitado floculoso. Adicionou-se 0,6 volumes (480µl) de isopropanol e após misturar bem, manteve-se 10 minutos à temperatura ambiente. Centrifugou-se 10 minutos à temperatura ambiente e o sedimento foi lavado com etanol 70%.

O sedimento foi ressuspendido em 50µl de TE contendo 20µg/ml de RNase. A análise dos insertos foi feita digerindo os plasmídios obtidos das mini-preparações com a devida enzima de restrição e observado em gel de agarose.

e) Reação de sequenciamento:

As reações de sequenciamento foram realizadas segundo o método de Sanger (SANGER *et al.*, 1977) utilizando-se o kit de sequenciamento *Sequenase*, versão 2.0, fabricado pela United States Biochemical. Para cada série de 8 reações, quatro diretas e quatro inversas, relativas a cada fita do DNA, utilizou-se 10µg de plasmídio ressuspendido num volume de 40µl de água. Os plasmídios obtidos nas mini-preparações foram previamente purificados por uma extração com fenol/clorofórmio/álcool isoamílico (5:24:1) e uma extração clorofórmio/álcool isoamílico (24:1) seguido de precipitação com 0,1 volume de acetato de sódio 2M e 2 volumes de etanol absoluto gelado. O precipitado foi finalmente lavado com etanol 70% e ressuspendido em água de forma a obter uma concentração final ajustada para 250µg/ml.

e.1) Desnaturação:

Adicionou-se 40µl de tampão de desnaturação (NaOH 0,4M; EDTA 0,4mM) em 40µl do plasmídeo e incubou-se 5 minutos à temperatura ambiente. A seguir, adicionou-se 0,1 volumes de acetato de amônio 2M pH4,5 (8µl) e 2 vols. de etanol absoluto gelado (160µl), sendo os dois adicionados ao mesmo tempo.

Deixou-se precipitando à -20°C por uma hora. Centrifugou-se por 30 minutos a 4°C. O sedimento foi lavado com etanol 70% e ressuspendido em 15µl de água.

e.2) Anelamento:

Foram adicionados 2µl de tampão de reação e 1µl de *primer* (soluções que acompanham o kit) em 7µl do DNA e incubados a 37°C por 15 minutos. A seguir, manteve-se 10 minutos à temperatura ambiente. Para a reação direta adicionou-se *primer* universal, para a inversa o *primer* inverso.

e.3) Reação:

Preliminares: 2,5µl de cada um dos dideoxinucleotídeos (ddNTP) foram mantidos em banho-maria à 37°C. O *labelling mix* foi diluído em água na proporção de 2:8 e a enzima em um tampão específico (que acompanha o kit) na proporção de 1:7.

Foram colocados no tubo contendo o DNA anelado, 1µl de DTT 0,1M, 2µl do *labelling mix* diluído, 0,5µl de dATP marcado com ³⁵S (correspondente à 5µCi) e 2µl da enzima *Sequenase* diluída. Manteve-se 5 minutos à temperatura ambiente; 3,5µl da reação foram então colocados em cada um dos tubos contendo os ddNTPs e foram mantidos à 37°C por mais 5 minutos. Adicionou-se por fim 4µl da solução *stop* e levou-se ao gelo imediatamente. As reações foram mantidas a -20°C até o momento da aplicação no gel de poliacrilamida.

e.4) Preparação do gel:

Uma solução de acrilamida 40% foi preparada diluindo-se em água 38g acrilamida e 2g de N,N'-metilacrilamida num volume final de 100ml de solução. Essa solução foi desionizada colocando-se 1g de resina AG 501-X8(D) e mantendo-se sob agitação por uma hora. Preparou-se um gel contendo 5% de poliacrilamida para a corrida longa, e um gel a 6%, para uma corrida curta. A preparação do gel é feita colocando-se numa proveta de 100ml as seguintes substâncias:

33,6g de uréia

12ml de acrilamida 40% (para o gel a 6%) ou

10ml de acrilamida 40% (para o gel a 5%)

8ml de TBE 10X.

Adicionou-se água até 75ml e dissolveu-se em banho-maria a 60°C, completando-se com água até 80ml. Filtrou-se com membrana marca Millipore de 0,45µm.

Procedeu-se então a preparação e montagem das placas. Estas foram mantidas durante à noite em uma solução de NaOH 0,4N e na manhã seguinte, lavadas com água e detergente. Para desengordurar, passou-se etanol seguido de acetona.

A placa que aderiu o gel, foi tratada com uma mistura de 5ml de etanol 95%, 175µl de ácido acético e 17µl de *bind-silane* (fabricado pela Pharmacia cod. N°80 1129-41) e em seguida com mais 5ml de etanol 95%. A placa que repeliu o gel foi tratada com 5ml de *repel-silane* (fabricado pela Pharmacia cod. N°80 1129-42).

As placas foram montadas com a parte tratada para dentro e utilizou-se espaçadores com gradiente de espessura. Para ocorrer a polimerização do gel, na hora do uso adicionou-se à solução de poliacrilamida 640µl de persulfato de amônio (100mg/ml) e 40µl de TEMED (Sigma cat. n°T-7024), vertendo-se o gel entre as placas. As placas foram mantidas em geladeira durante à noite para completa polimerização do gel.

Foram aplicados 2µl da reação, previamente desnaturada, em cada canaleta de ambos os géis 5% e 6%. A aplicação foi feita na ordem G, A, T e C e cada gel foi corrido à uma potência constante de 50 watts e à 2000 volts. Para finalizar a corrida do gel, a reação aplicada apresenta uma mistura de corantes contidos na solução *stop*. O gel de corrida curta (6%) foi interrompido após a saída do primeiro corante (aproximadamente 3 horas de corrida) e o gel de corrida longa (5%), manteve-se, após a saída do segundo corante, por mais 2 horas.

Ao serem desmontadas as placas, o gel foi fixado com uma solução de metanol 10% e ácido acético 12%, por meia hora sob agitação. As placas contendo o gel fixo foram secas durante a noite a 45°C e a seguir expostas em um filme X-OMAT AR (Kodak diagnostic film cat 1651454) por 2 dias.

5 -Análise das seqüências:

A seqüência dos subfragmentos foram analisadas, utilizando-se o programa DNASIS versão 6.0 (fabricado pela Hitachi Software Engineering CO., LTD.), quanto ao mapa de restrição, tradução e análise de possíveis regiões promotoras (APÊNDICE II). As seqüências foram também enviadas via Bitnet para o banco de dados EMBL e GenBank à fim de serem comparadas com outras seqüências já descritas (ALTSCHUL, *et al.*, 1990).

Os *primers* foram desenhados com o auxílio do programa OLIGO versão 3.4 (fabricado pela Wojciech Rychlik), e foram confeccionados pela GIBCO, BRL.

6 - Estimativa do número de cópias:

A estimativa de quantas vezes uma dada seqüência se repete no genoma foi feita, comparando-se as intensidades das bandas obtidas na hibridização dessa seqüência, com DNA genômico e quantidades relativas conhecidas dela mesma clonada em plasmídio (GLICK, *et al.*, 1993).

Estas quantidades relativas conhecidas, foram calculadas comparando-se o tamanho do plasmídio contendo esta sequência como inserto e o tamanho do genoma haplóide. Utilizando-se, a seguinte fórmula:

$$M_p = \frac{T_p + T_i \times M_g}{T_{gh}}, \text{ onde}$$

M_p = massa relativa do plasmídio mais inserto (correspondente à 1 cópia)

T_p = tamanho do plasmídio em bp

T_i = tamanho do inserto em bp

M_g = massa em μg de DNA genômico ($10\mu\text{g}$)

T_{gh} = tamanho do genoma humano haplóide em bp (3×10^9 bp)

Preparou-se um filtro com quantidades de plasmídio mais inserto equivalentes à 100, 50, 20 e 1 cópia. O plasmídio utilizado foi o pIBI-76, contendo o fragmento Y-1 como inserto. A massa para os respectivos números de cópias foi calculada, a partir de uma digestão deste plasmídio com a enzima *Hind* III, para que o fragmento Y-1 fosse liberado. No mesmo filtro foram colocados $10\mu\text{g}$ de DNA genômico de homem (46,XY) e $10\mu\text{g}$ de DNA de mulher (46,XX). A hibridização foi feita com o subfragmento Y-1.c. A intensidade das bandas obtidas foi lida utilizando-se um densitômetro modelo ULTROSCAN XL (Pharmacia - LKB). A curva padrão foi feita comparando-se a absorbância das bandas obtidas com o número de cópias colocadas no filtro. A absorbância das bandas obtidas no DNA genômico foi relacionada no gráfico da curva padrão e, por regressão linear, foi calculado o número correspondente de cópias.

7 - Técnica de PCR:

Foi realizada a amplificação do subfragmento Y-1.c por PCR utilizando-se os *primers* E1 (GTT ATC TCC TCC ACT GTT CC) e E2 (AAT CCC AAC AAG ACC CAC TG), selecionados pelo programa OLIGO, que fornecem um produto de amplificação de 294 bp.

Foram testadas várias condições de reação, sendo que a que apresentou o resultado esperado foi a seguinte: 100ng de DNA genômico; 2mM de $MgCl_2$; 50mM de KCl; 20mM de Tris-HCl pH 8,4; 3,5% de formamida; 15% de glicerol; 0,2mM de cada um dos dNTPs; 0,5 μ M de cada *primer* e 1,25U de *Taq* polimerase para um volume final de reação de 25 μ l.

Para a amplificação foram realizados 30 ciclos de 94°C, por 45 segundos, 55°C, por 30 segundos e 72°C, por 30 segundos. Finalmente um último ciclo foi realizado de 94°C, por 45 segundos, 55°C, por 30 segundos e 72°C, por 5 minutos.

Em todas as reações de amplificação, o material foi submetido a um ciclo inicial de 94°C, por 5 minutos, após este tempo a enzima era então acrescentada à reação.

Os produtos de amplificação foram separados por eletroforese em gel de agarose *Nusieve* 2% em tampão de corrida TBE. Foram colocados 5 μ l de cada reação no gel.

Foram realizadas reações de controle negativo (branco) e controle positivo (20pg de plasmídeo contendo como inserto a sequência Y-1.c).

8 - Conservação da sequência Y-1:

A sequência Y-1.c foi hibridizada pelo método de Southern com amostras de DNA de várias espécies de mamíferos para determinar a presença de sequências homólogas e, caso existissem, se elas eram específicas de um sexo em particular.

a) Espécies utilizadas:

As amostras de sangue ou de DNA já extraído dos macacos analisados foram enviadas pelo Dr. Horário Schneider da Universidade Federal do Pará, que detém tradição no estudo citogenético de espécies de primatas brasileiros. Dessa forma, as amostras com a classificação sistemática definida e as espécies estudadas foram as seguintes:

Cebus nigrivittatus (macaco prego)

Saimiri sciureus (macaco-de-cheiro)

Ateles belsebuth (macaco aranha)

Lagothrix lagotricha (macaco barrigudo)

Aotus azarae (macaco-da-noite)

Quanto às outras espécies de mamíferos, as amostras foram obtidas em laboratórios do Instituto de Biologia, UNICAMP, e referem-se à animais de uso corrente em pesquisa e são: boi, coelho e camundongo.

b) Extração de DNA de sangue total de macacos:

Foram coletados quantidades variáveis de sangue de diversas espécies de macacos com anticoagulante EDTA (1,2mg/ml). Centrifugou-se o sangue total a 3000rpm, à temperatura ambiente por 20 minutos. Transferiu-se cuidadosamente a camada de células brancas para um outro tubo de 15ml com tampa .

Adicionou-se a esse tubo, 4,75ml de STE, 0,125ml de SDS 20% e 0,1ml de ribonuclease a 10mg/ml . Incubou-se em banho-maria a 37°C por 1 hora, agitando suavemente os tubos de 15 em 15 minutos. A seguir, foram adicionados 100µl de proteinase K (10mg/ml) e incubou-se a 55°C durante a noite sob agitação suave.

Procedeu-se uma extração com 5ml de fenol. Agitou-se manualmente por inversão delicadamente durante 10 minutos. Incubou-se em gelo por 10 minutos e a seguir, centrifugou-se a 4000rpm na temperatura de 4°C , por 30 minutos.

A camada superior (sobrenadante) foi transferida para outro tubo tipo Falcon. Em seguida, uma outra extração foi feita com 5 ml de uma mistura 1:1 de fenol e clorofórmio/álcool isoamílico (24:1). Agitou-se por 10 minutos e centrifugou-se a 4000rpm por 30 minutos a 4°C.

A última extração foi feita com 5ml de clorofórmio/álcool isoamílico. O sobrenadante foi transferido para um tubo de polipropileno e o DNA foi precipitado adicionando-se igual volume de etanol absoluto gelado e 1/20 volumes de acetato de sódio 3M, pH5,2 e mantendo-se a -20°C durante à noite.

Centrifugou-se a 7500rpm, a 4°C, por 30 minutos. O sedimento foi lavado com etanol 70%. Depois de seco o DNA foi ressuspendido com quantidade apropriada de TE e a concentração foi lida em espectrofotômetro.

c) Extração de DNA de amostras de tecidos de coelho e camundongo:

Coletou-se os órgãos (fígado, rins ou baço) de coelho e camundongo. A amostra foi colocada em tubo tipo Falcon e congelada no freezer a -70°C até o momento de sua extração.

Retirou-se os órgãos do freezer para descongelar em cima de uma placa de Petri esterilizada, incubada sobre gelo. Retirou-se 500mg de material que foi macerado até se tornar homogêneo.

Ao tecido macerado foi acrescentado 50ml de STE com 0,5ml de proteinase K (10mg/ml) e 2,5ml de SDS a 10%. Colocou-se num tubo de centrífuga e deixou-se sob agitação a 51°C durante a noite.

Adicionou-se igual volume de fenol/clorofórmio/álcool isoamílico (25:24:1) e misturou-se por inversão durante 10 minutos. Incubou-se em gelo por 10 minutos e em seguida centrifugou-se a 3000rpm por 20 minutos na temperatura de 4°C.

Outras duas extrações foram repetidas com fenol. A seguir, foram feitas duas extrações com igual volume de clorofórmio/álcool isoamílico (24:1) nas mesmas condições das extrações anteriores.

O sobrenadante foi transferido para um tubo de polipropileno. Adicionou-se igual volume de etanol absoluto gelado e 1/20 do volume total de acetato de sódio 3M, pH 5,2 para precipitar o DNA. O tubo foi mantido em freezer a -20°C durante a noite.

Na manhã seguinte, centrifugou-se por 20 minutos, a 7500rpm na temperatura de 4°C. O sedimento foi lavado com 10ml de etanol 70% e, depois de seco, ressuspendido com volume apropriado de TE. A concentração do DNA foi aferida em espectrofotômetro.

A digestão, eletroforese, transferência para o filtro e hibridização foram feitas como descrito para o DNA genômico humano (item 3).

4.

RESULTADOS E DISCUSSÃO

RESULTADOS E DISCUSSÃO

1 -Análise quanto à especificidade:

O estudo da especificidade da sequência Y-1 foi iniciado repetindo-se experimentos realizados por NOVAIS (1990). Cada subfragmento, obtido através do mapa de restrição da sequência Y-1 (figura 5), foi hibridizado com DNA genômico de origem masculina (46,XY) e feminina (46,XX), digeridos com as enzimas *Hind* III, *Eco* RI, *Bgl* II e *Hae* III. Estas enzimas foram escolhidas ao acaso e a hibridização foi feita em condições de alta estringência. As bandas detectadas por estes subfragmentos são mostradas na figura 6.

Quando se observa o resultado da hibridização com o subfragmento Y-1.b1 (figura 6B) o qual foi determinado como o mais Y-específico por NOVAIS (1990), são realmente visualizadas bandas Y-específicas para as diferentes enzimas usadas. Entretanto, posteriormente observou-se que este resultado é muito sensível à variação nas condições do experimento, ou seja, uma pequena diminuição da estringência resulta em um padrão muito menos específico. Na figura 7, pode-se observar a hibridização deste subfragmento com DNA de homem 46,XY e mulher 46,XX, digeridos com *Hind* III em condições de baixa estringência. Apenas uma banda, supostamente Y-específica é visualizada. Entretanto, esta banda deve ser atribuída à uma homologia parcial, já que ela nem aparece nas condições de alta estringência.

O subfragmento Y-1.c de 400 bp (figura 6C), outro lado, embora exiba diversas bandas homólogas em homem e mulher evidencia uma banda Y-específica de 400 bp, quando o DNA genômico é digerido com a enzima *Eco* RI. Este resultado é muito interessante porque esta banda por ser do mesmo tamanho do subfragmento, representa o seu próprio *locus*.

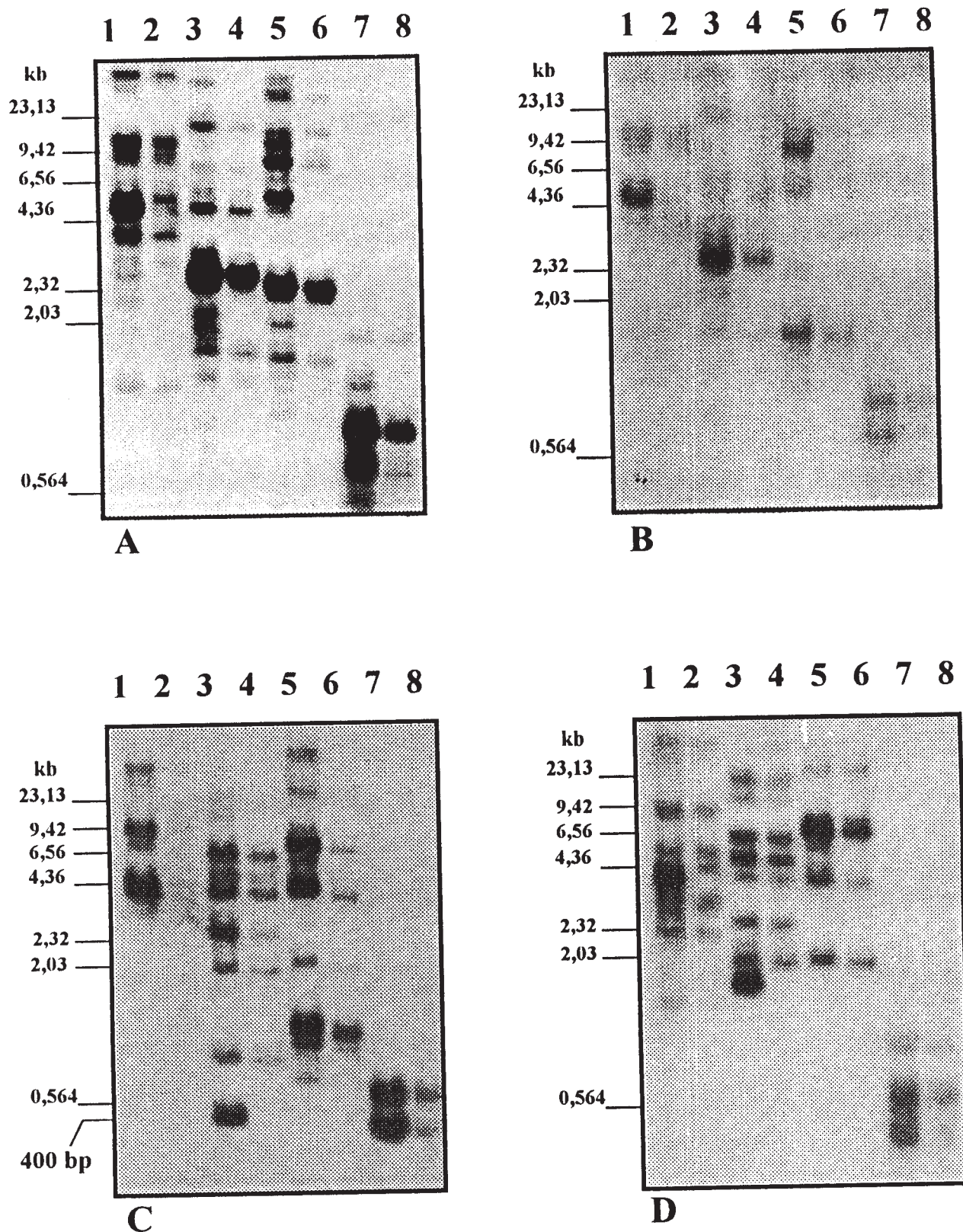


Figura 5 - Subfragmentos (A) Y-1.b2, (B) Y-1.b1, (C) Y-1.c e (D) Y-1.a, hibridizados com DNAs de homem (46,XY - 1, 3, 5 e 7) e mulher (46,XX - 2, 4, 6 e 8) digeridos com *Hind* III (1 e 2), *Eco* RI (3 e 4), *Bgl* II (5 e 6) e *Hae* III (7 e 8).

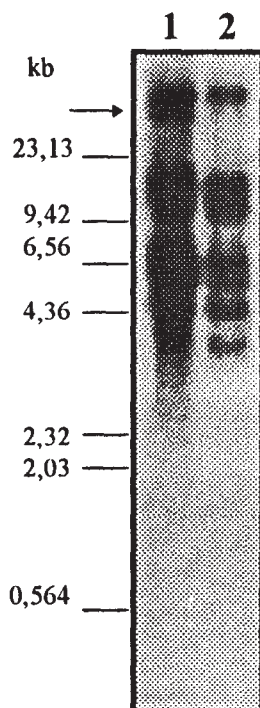


Figura 7 - Sonda Y-1.b1 hibridizada com DNA de homem (46,XY - 1) e mulher (46,XX - 2) digeridos com *Hind* III, em condições de baixa estringência. A flecha indica a banda Y-específica.

Esta banda, que é, portanto, *locus*-específica, foi utilizada em outras abordagens, tais como, estudos de homologias e conservação nas espécies, estimativa do número de cópias, determinação de *primers* para utilização da técnica de PCR e localização do fragmento Y-1 no cromossomo Y.

O subfragmento Y-1.b2, por outro lado, quando foi sequenciado para a pesquisa de regiões promotoras (resultados discutidos no item 5) demonstrou ser uma sequência de 730 bp que apresenta uma cadeia aberta de leitura de 71 aminoácidos (ORF - *open reading frame*) com início na base 517 da sequência e no sentido de restrição *Bam* HI → *Hind* III.

A descoberta deste ORF no subfragmento Y-1.b2, sugere a presença de uma sequência única nesta região, que talvez estivesse somente presente no cromossomo Y.

À fim de se analisar a especificidade deste ORF, o subfragmento Y-1.b2 foi digerido com a enzima de restrição *Ava* I, a qual corta na base 511 e gera dois outros subfragmentos como mostra a figura 8.

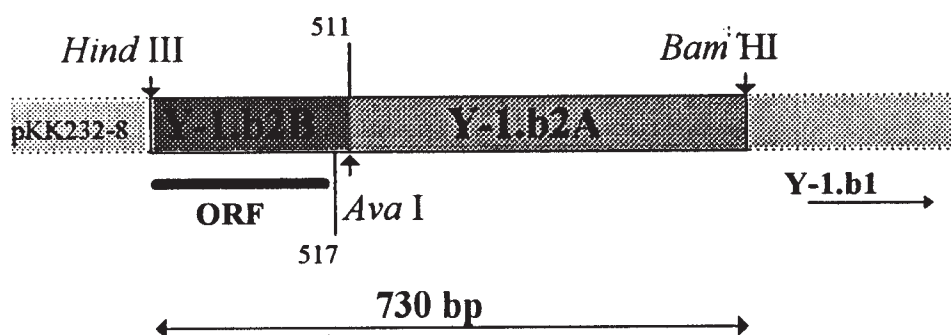


Figura 8 - Mapa de restrição do subfragmento Y-1.b2.

Destes novos subfragmentos gerados (Y-1.b2A e Y-1.b2B) o subfragmento Y-1.b2B, que contém praticamente só o ORF, foi hibridizado com DNA genômico de origem masculina e feminina digeridos com as enzimas *Hind* III e *Bgl* II. Um padrão específico é observado (figura 9), evidenciando a importância da sequência pela presença de regiões promotoras (resultados discutidos no item 5).

Desta forma, tanto o subfragmento Y-1.c quanto o subfragmento Y-1.b2B se prestam para a utilização como sondas quando a intenção é detectar o cromossomo Y, uma vez que ambos apresentam um padrão diferencial que permite distinguir o DNA de origem masculina daquele de origem feminina.

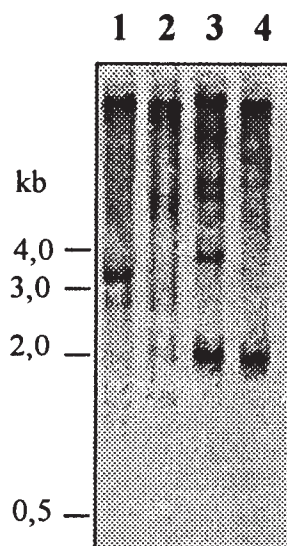


Figura 9 - Subfragmento Y-1.b2B hibridizado com DNA genômico de origem masculina (46,XY - 1 e 3) e feminina (46,XX - 2 e 4), digeridos com as enzimas *Hind* III (1 e 2) e *Bgl* II (3 e 4).

2 - Localização da seqüência no cromossomo Y:

O mapeamento do fragmento Y-1 foi feito utilizando um conjunto de linhagens celulares que apresentavam diferentes aberrações estruturais no cromossomo Y. O subfragmento Y-1.c foi utilizado como referência para a localização, mais precisamente, observou-se a presença ou ausência da banda *locus*-específica de 400 bp.

O esquema da figura 10, mostra a porção do cromossomo Y presente em cada linhagem celular, bem como, os resultados obtidos nas hibridizações e a região do cromossomo Y detectada por cada sonda previamente descrita por outros pesquisadores. Deve ser salientado que a porção do cromossomo Y descrita estar presente em cada linhagem celular baseou-se, inicialmente, em dados citogenéticos (cariótipos na tabela 1) e, portanto, está sujeita à imprecisões. Por isso, utilizamos outras sondas para incrementar o nível de resolução.

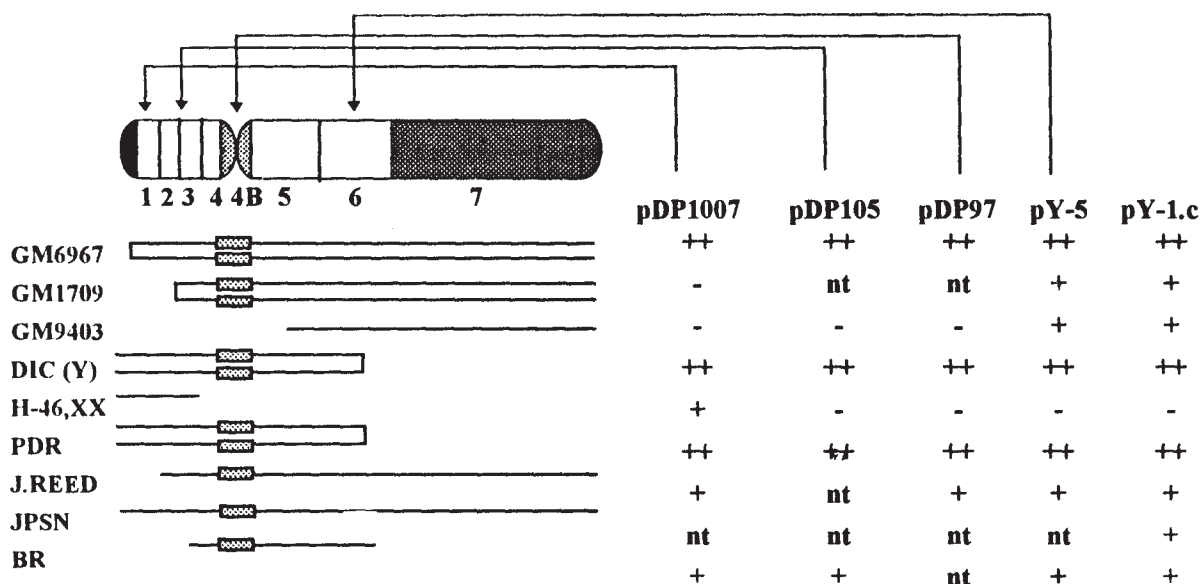


Figura 10 - Esquema representativo mostrando a porção do cromossomo Y presente em cada linhagem celular e a localização das sondas utilizadas (nt = não testado).

Analisando os resultados, observa-se que a banda de 400 bp do subfragmento Y-1.c, somente não apareceu na hibridização com a linhagem de um homem 46,XX (H-46,XX) que contém a região pseudoautosômica (PAR), os intervalos 1, 2 e parte do intervalo 3 do cromossomo Y (FARAH *et al.*, 1991). A partir destes resultados, duas hipóteses podem ser formuladas à respeito deste mapeamento: 1) o fragmento Y-1 se encontra nos intervalos 5-6, pois esta é a região comum à todas as linhagens que apresentaram resultado positivo, ou 2) esta sequência se encontra dispersa por todo o cromossomo, exceto na região presente no homem XX.

A hibridização com DNA do paciente JPSN foi realizada, na tentativa de se obter um resultado mais preciso por exclusão. Este paciente foi selecionado para estudo em um projeto em desenvolvimento no laboratório de Genética Humana do

CBMEG (Centro de Biologia Molecular e Engenharia Genética), o qual pretende relacionar pacientes que apresentam azoospermia idiopática com a presença de microdeleções no intervalo 6 do cromossomo Y. Uma vez que o gene **AZF**, que controla a espermatogênese, foi mapeado no intervalo 6, acredita-se que a presença de microdeleções nesta região poderia interromper a função deste gene levando ao desenvolvimento de homens estéreis.

Realmente usando-se um painel de 15 sondas demonstrou-se no paciente JPSN uma microdeleção no intervalo 6, a qual se estende dos intervalos VII à VIII e que poderia justificar, sua azoospermia (RAMOS *et al.*, 1993) Desta forma, a hibridização da sequência Y-1.c com DNA deste paciente poderia trazer informações interessantes sobre a localização da sequência Y-1 no cromossomo Y. Se o resultado fosse negativo, poderíamos afirmar que o subfragmento Y-1.c se encontra na região deletada no cromossomo Y do paciente. Entretanto, o resultado da hibridização foi positivo (figura 11), não sendo possível então concluir se este subfragmento se encontra realmente no intervalo 6. Nesta hibridização o irmão normal do paciente foi colocado como controle para afastar a possibilidade de polimorfismos. Por outro lado, deve-se considerar que a microdeleção presente no paciente não envolve toda a região de interseção das linhagens utilizadas na figura 10, não sendo possível também descartar a hipótese de que o subfragmento se encontra nas regiões 5-6.

Embora não tenhamos uma linhagem celular capaz de excluir ou de provar uma das duas hipóteses, a possibilidade da sequência Y-1 estar restrita ao intervalo 5-6 nos parece a mais provável. Isto porque seria difícil se justificar o fato de uma sequência, que aparece dispersa no cromossomo Y, não estar presente na **PAR** e nos intervalos 1, 2 e parte do 3, como ficou demonstrado no caso do homem 46,XX.

Evidentemente, seriam necessárias hibridizações com outras linhagens celulares para refinar o mapeamento da sequência Y-1 no cromossomo Y. Mas, a simples possibilidade deste fragmento se encontrar no intervalo 6, já o torna especialmente importante em estudos de mapeamento do gene **AZF**. Por outro lado, a utilização da sonda Y-1.c em outros pacientes que apresentam microdeleções no intervalo 6, eventualmente nos fornecerá a precisa localização desta sequência no cromossomo Y.

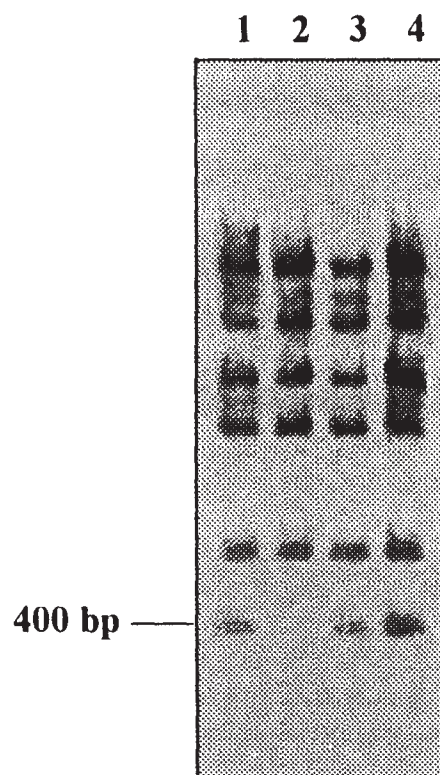


Figura 11 - Sonda Y-1.c hibridizada com homem (46,XY - 1), mulher (46,XX - 2), paciente JPSN (3) e irmão do paciente JPSN (4), digeridos com a enzima *Eco* RI.

3 - Estimativa do número de cópias:

Um outro experimento para determinar o número de cópias do fragmento Y-1.c foi realizado, com a finalidade de se definir o quanto esta sequência se encontra repetida no cromossomo Y.

Mais uma vez, sequência Y-1.c foi utilizada como referência por ser *locus*-específica. Como já foi dito, ela apresenta a banda 400 bp que aparece quando o subfragmento é hibridizado com DNA de homem 46,XY, digerido com a enzima *Eco* RI. Esta banda representa o próprio subfragmento Y-1.c e foi usada para uma análise indireta da organização do resto da sequência Y-1.

A figura 12-A mostra a autorradiografia obtida pela hibridização da sonda Y-1.c com quantidades conhecidas de cópias da sequência Y-1 e DNA genômicos masculino e feminino digeridos com *Eco* RI (material e métodos página 32), a curva da figura 12-C mostra a absorbância de cada banda lida em densitômetro referente ao DNA genômico masculino.

Pelos resultados obtidos (figura 12-B e C), conclui-se que as bandas homólogas em DNA masculino e feminino apresentam aproximadamente o mesmo número de cópias e, considerando-se imprecisões inerentes ao método, estimou-se que a banda Y-específica de 400 bp se encontra repetida no cromossomo Y aproximadamente 70 vezes.

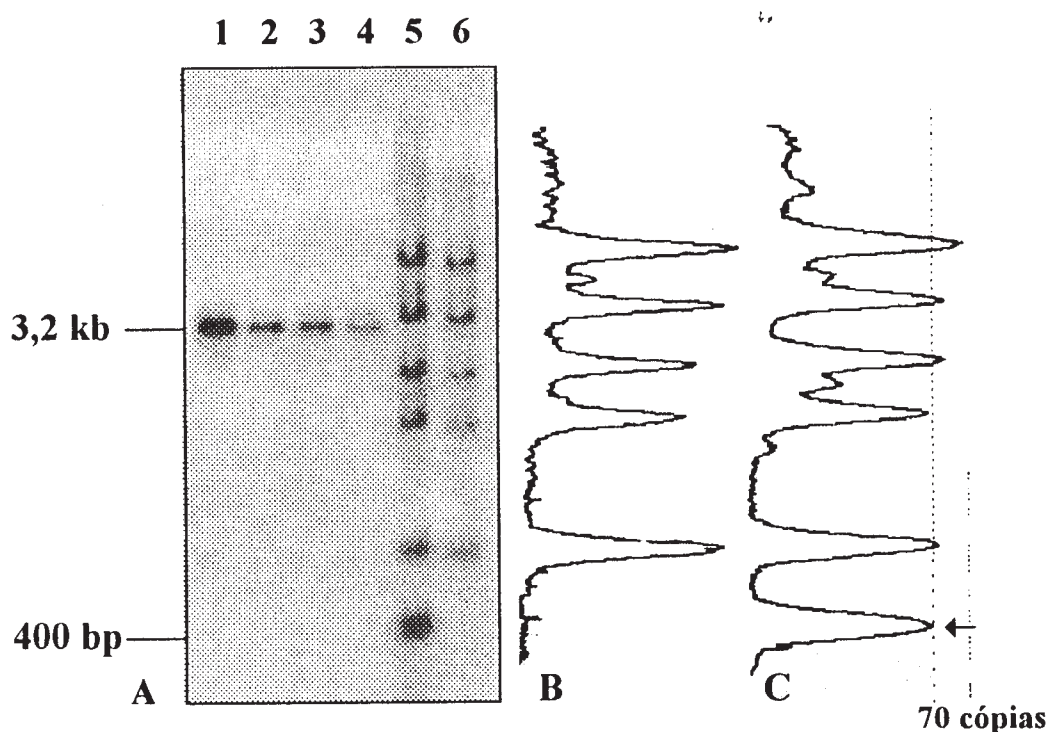


Figura 12 - (A) - Hibridização da sonda Y-1.c com filtro contendo: (1) - 100 cópias, (2) - 50 cópias; (3) - 20 cópias e (4) - 1 cópia da sequência Y-1; (5) - DNA genômico masculino, 46,XY e (6) - DNA genômico feminino, 46,XX, digeridos com *Eco* RI - **(B)** - Absorbância lida em densitômetro referente à canaleta 6 (DNA genômico feminino) - **(C)** - absorbância lida em densitômetro referente à canaleta 5 (DNA genômico masculino). A flecha indica o pico de absorbância da banda Y-específica.

4 - Homologias e conservação da seqüência:

Dentre as diversas bandas detectadas pelo fragmento Y-1 quando é hibridizado com DNA genômico, algumas são Y-específicas, outras são comuns à DNAs de origem masculina e feminina.

O subfragmento Y-1.c também mostra este comportamento, isto é, exibe uma banda diferencial de 400 bp e bandas comuns à homem e mulher. Este padrão de bandas deve ser resultado da homologia entre o cromossomo Y e autossomos. Isto pode ser comprovado observando-se a figura 12, onde a intensidades das bandas comuns nos DNAs de origem masculina e feminina é praticamente a mesma. O mesmo pode ser concluído quando este subfragmento é hibridizado com linhagens celulares onde o número de cromossomos X é aumentado, neste caso as bandas comuns não variam de intensidade. Por outro lado, quando o número de cromossomos Y é aumentado, a banda de 400 bp aumenta de intensidade. Resultados mostrados na figura 13.

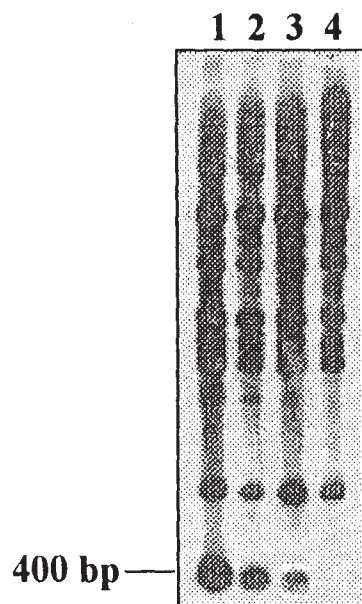


Figura 13 - Sonda Y-1.c hibridizada com DNA 46,XYYY (1), 46,XYY (2), 46,XY (3) e 46,XX (4), digeridos com a enzima *Eco* RI.

A sequência Y-1.c também foi hibridizada com DNA de origem masculina e feminina de várias espécies de macacos, boi, coelho e camundongo. Quanto à especificidade da banda de 400 bp, em macacos ela é conservada em todas as espécies estudadas mas, não é Y-específica, aparecendo em igual intensidade em machos e fêmeas. Em bovinos observa-se uma banda Y-específica mas, de tamanho maior que 400 bp. No coelho uma mesma banda de tamanho maior que 400 bp é visualizada em macho e fêmea. Entretanto, nenhuma banda é observada na hibridização com DNA genômico de camundongos macho ou fêmea. A figura 14, mostra o resultado das hibridizações com camundongo, coelho e boi e com algumas espécies de macacos. A tabela 3 mostra resumidamente o resultado obtido nessas hibridizações.

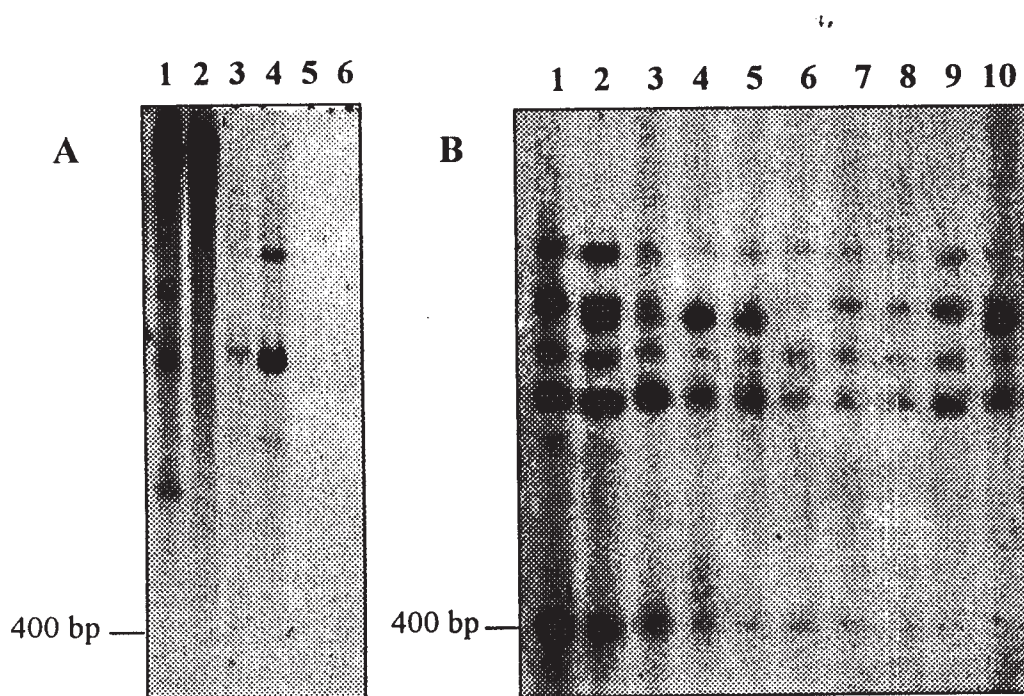


Figura 14 - Hibridização da sonda Y-1.c com (A) boi (1), vaca (2), coelho (3), coelha (4), camundongo macho (5), camundongo fêmea (6); (B) macaco prego macho (1, 2) macaco-de-cheiro macho (3, 4), macaco aranha (5, 6), macaco barrigudo (7, 8) e macaco-da-noite (9, 10); (1, 3, 5, 7 e 9) machos e (2, 4, 6, 8 e 10) fêmeas.

ESPÉCIE	MACHO	FÊMEA	TAMANHO (bp)
camundongo	-	-	-
coelho	+	+	≠ 400
boi	+	-	≠ 400
macaco	+	+	400
homem	+	-	400

Tabela 3 - Resultado das hibridizações do subfragmento Y-1.c com diferentes espécies. Hibridização positiva (+) e negativa (-).

Como não foi observado qualquer homologia desta sequência com camundongo, outras espécies de aves e anfíbios, por exemplo, nem foram testadas, por estarem ainda mais distantes na escala evolutiva.

A sequência Y-1.c, portanto, de ausente em camundongos à Y-específica em humano, sofreu uma clara expansão no genoma em geral e uma tendência de particularização ao cromossomo Y. É possível que essa sequência pertença ao grupo daquelas que se duplicaram e se tornaram dispersas no genoma, como ocorreu, por exemplo, com diversas famílias de multigenes.

Sequências que sofreram expansão e contração durante o processo evolutivo, apresentam grande variabilidade como tem sido demonstrado com relação as sequências da região pseudoautosômica (YEN et al., 1988). Isto poderia justificar a não conservação da sequência Y-1.c, como foi verificado pela sua ausência já em camundongos.

Ao longo da evolução, sequências homólogas à Y-1.c devem ter acumulado inúmeras mutações de forma que os sítios de restrição sofreram alterações gerando bandas de tamanhos diferentes de 400 bp. Esse fenômeno é bem descrito com relação ao cromossomo Y e resulta em genes não-funcionais, o que justifica a alta frequências de pseudogenes nesse cromossomo em relação aos outros (NAKAHORI et al., 1991).

Por outro lado, a sequência Y-1.c presente atualmente no cromossomo Y humano deve ter sofrido pressões evolutivas levando à sua particularização neste cromossomo.

Assim como a região pseudoautossômica indica os caminhos na evolução humana, estudos de homologias de seqüências do cromossomo Y podem francamente mostrar outros processos evolutivos ocorridos.

Estudos evolutivos com o DNA mitocondrial, com estrita herança materna têm indicado que a origem de todas as mitocôndrias contemporâneas humanas são descendentes de uma única mulher que viveu há 200.000 anos, da mesma forma, estudos nas regiões polimórficas do cromossomo Y têm buscado o "pai" de todos nós (GIBBONS, 1991).

5- Sequenciamento do subfragmento Y-1.c, seleção dos *primers* e padronização da PCR:

O subfragmento Y-1.c foi sequenciado por ser *locus*-específico e portanto, o mais favorável para ser amplificado em técnica de PCR para identificação desta seqüência em estudos clínicos. A sequência Y-1.c possui 394 bp (APÊNDICE II). A análise da seqüência no programa DNASIS não revelou nenhuma região codificante, assim como a comparação com outras seqüências, quando enviada para os bancos EMBL e GenBank, não mostrou nenhuma homologia importante.

Para o desenho dos *primers* foram considerados diversos critérios. Procurou-se obter *primers* que não formassem dímeros entre si, ou pelo menos, que a formação dos dímeros não interferisse na reação de polimerização, por exemplo, obstruindo a extremidade 3' do *primer*. Foi levado também em conta o conteúdo de G+C, o qual influi diretamente na temperatura de anelamento garantindo-se que este conteúdo fosse no mínimo de 50%. Na escolha dos *primers* procurou-se alcançar a maior especificidade possível, isto é, que não apresentassem nenhuma homologia com o resto da seqüência e (ou) com a outra fita de DNA. A seleção dos *primers* foi feita utilizando-se o programa OLIGO, o qual considera também outros critérios, como entropia e energia livre, além de fornecer os dados teóricos para as condições de reação de PCR, como temperatura de anelamento e salinidade.

Os *primers* selecionados, E1 (GTT ATC TCC TCC ACT GTT CC) e E2 (AAT CCC AAC AAG ACC CAC TG), fornecem um produto de amplificação de 294 bp e estão localizados na seqüência na figura 16. As condições padronizadas para a reação de PCR a partir de DNA genômico utilizando os *primers* E1 e E2, visou a amplificação específica do DNA de origem masculina. Após algumas tentativas, estas condições foram estabelecidas como descrito em material e métodos na página 33.

O resultado da reação de PCR é mostrado na figura 15, onde pode-se observar a presença da banda de 294 bp somente no controle positivo e no DNA genômico de homem 46,XY. A amplificação específica para DNA de origem masculina alcançada a partir destes *primers* torna a 'seqüência Y-1.c uma importante ferramenta na identificação do cromossomo Y em qualquer linhagem com alteração neste cromossomo.

Uma vez obtida a localização precisa da seqüência Y-1 no cromossomo Y, os *primers* E1 e E2 terão também grande aplicação no emprego da técnica de PCR para a identificação de partes do cromossomo Y como acontece, por exemplo, em cromossomos marcadores

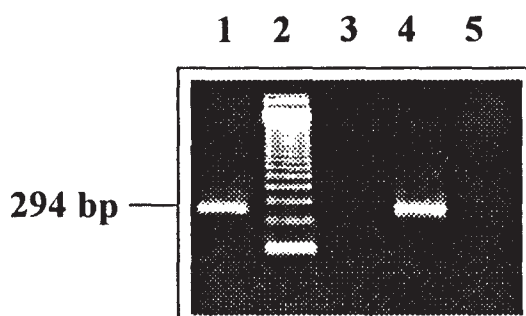


Figura 15 - Resultado da reação de PCR utilizando os *primers* E1 e E2 da sonda Y-1.c. (1) Controle positivo - plasmídio pUC(19) + Y-1.c, (2) marcador 123 *ladder*, (3) controle negativo, (4) homem 46,XY e (5) mulher 46,XX.

10 *primer superior* → 40 50
 5' GAATTCAAAT TTTATCTCC TCCACTTTC CCTGGAGTGC TAACAGGTGA
 3' CTTAAGTTTA ACAATAGAGG AGGTGACAAG GGACCTCACG ATTGTCCACT
 60 70 80 90 100
 CAGAATAGCT CAACTGTTGC TTTTACCATA TGTTAAGTTA GGAGAAATCT
 GTCTTATCGA GTTGACAACG AAAATGGTAT ACAATTCAAT CCTCTTTAGA
 110 120 130 140 150
 CAGAAAAAA GGATATGGAG CACAAATCAG CAGGCAAGGC ACCCTATTGG
 GTCTTTTTTT CCTATACCTC GTGTTTAGTC GTCCGTTCCG TGGGATAACC
 160 170 180 190 200
 TAAATCAAGC TCTGACAATA GTCCTATTGT AGGTCACTAT TCAAGGAAAA
 ATTTAGTTCG AGACTGTTAT CAGGATAACA TCCAGTGATA AGTTCCTTTT
 210 220 230 240 250
 CAATTTGAGG CTCTGGTCGA TACAGGAGCA GACGTGTCAT CATAGCTTTT
 GTTAAACTCC GAGACCAGCT ATGTCCTCGT CTGCACAGTA GTATCGAAAA
 260 270 280 290 300
 ATCAACGGCC CAAAACTGA CCCAAACAAA AGGCCAGTG GGTCTTGTTG
 TAGTTGCCGG GTTTTGGACT GGGTTTGTTC TCCGGGTCAC CCAGAACAAAC
 310 320 330 340 350
 GGATTGAGGA TTTGCTTATA TCTCACCAGG AGAGAATCAA CTTCTTGTCT
 CCTAACTCCT AAACGAATAT AGAGTGGTCC TCTCTTAGTT GAAGAACAGA
 360 370 380 390
 GGGTACCCAC AAGACATCTT AAGCTGTGCC AGGAGCCAGA ATTC 3'
 CCCATGGGTG TTCTGTAGAA TTCGACACGG TCCTCGGTCT TAAG 5'

← *primer inferior*

Figura 16 - Sequência Y-1.c mostrando os *primers* E1 (superior) e E2 (inferior) selecionados para a amplificação da sequência pela técnica de PCR.

6 - Pesquisa de regiões promotoras:

O plasmídio pKK232-8 possui o gene que codifica a enzima cloranfenicol acetil-transferase (CAT) mas com ausência da sua região promotora. Os sítios de clonagem são situados antes do códon de iniciação deste gene, na construção do plasmídio. Desta forma, quando um inserto contendo uma região promotora é clonado neste plasmídio, o gene CAT poderá ser transcrito, conferindo à bactéria transformada resistência ao cloranfenicol. O plasmídio pKK232-8 contendo a sequência Y-1 como inserto, tornou a bactéria transformada resistente ao cloranfenicol numa concentração de 30µg/ml, sugerindo a presença de sequências promotoras (NOVAIS, 1990).

Pela forma como o fragmento Y-1 foi clonado no plasmídio, o subfragmento que se encontra mais próximo ao gene CAT é o Y-1.b2. Isto pôde ser demonstrado por análise de restrição (NOVAIS, 1990). O mapa do plasmídio contendo o inserto é mostrado na figura 17. Este subfragmento foi então sequenciado à fim de se pesquisar regiões promotoras.

O subfragmento Y-1.b2, de restrição *Hind* III/*Bam* HI, possui 730 bp (APÊNDICE II). Sua sequência foi enviada para os bancos de dados EMBL e GenBank via Bitnet, para ver se havia alguma homologia com outras já descritas. Nenhuma homologia significativa foi encontrada. Por outro lado, a análise no programa DNASIS mostra uma cadeia aberta de aminoácidos de 220 bp, com códon de iniciação metionina. Na posição -35 bp em relação a esta metionina, aparece uma TATA *box* e à -91 bp uma CAAT *box*, as quais são regiões promotoras consenso para organismos eucarióticos. Este esquema é mostrado na figura 18.

É difícil afirmar se foram as regiões promotoras detectadas no sequenciamento do fragmento Y-1.b2 que realmente provocaram a expressão do gene CAT. Outras análises seriam necessárias para esta confirmação como, por exemplo, experimentos que demonstram exatamente os sítios em que se ligam as subunidades da polimerase (*footprint*) ou a quantificação da enzima cloranfenicol acetil transferase. Entretanto, considerando-se os experimentos a que nos propomos realizar neste trabalho, não poderíamos esperar esclarecimentos maiores do que os obtidos com relação à regiões promotoras.

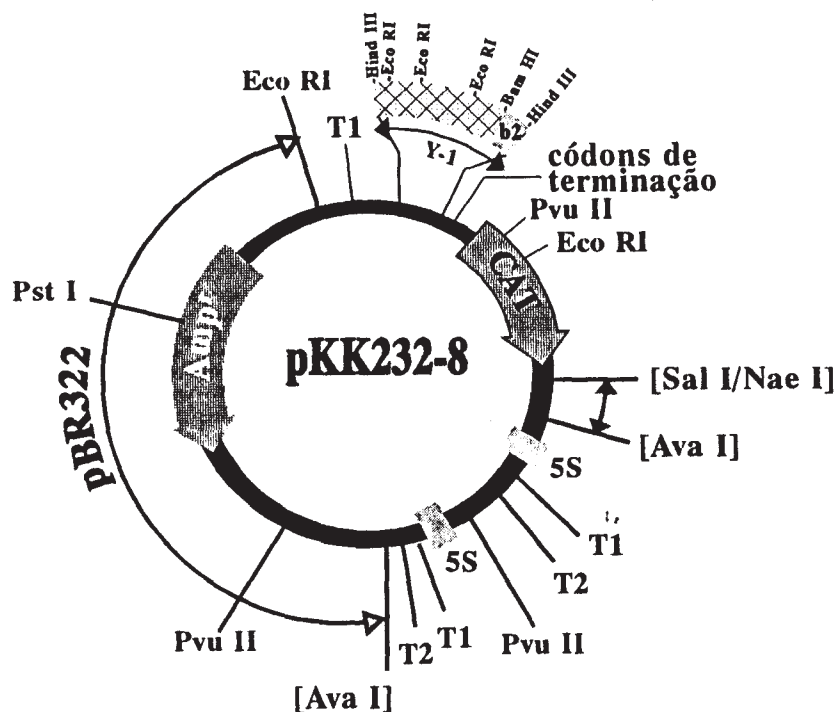


Figura 17 - Plasmídio pKK232-8 contendo o fragmento Y-1 como inserto.

Um fato importante a ser salientado, diz respeito às características do vetor pKK232-8, que foi construído exatamente com o objetivo de rastrear regiões promotoras. Este plasmídio possui códons de terminação nas três cadeias de leituras de aminoácidos anteriores à metionina de iniciação do gene CAT. Estes códons de terminação não permitem que haja a transcrição contínua do inserto clonado com o gene CAT, evitando a formação de um produto gênico híbrido, o que poderia impedir um estudo quantitativo da enzima cloranfenicol acetil transferase ou mesmo provocar a desestabilização da bactéria hospedeira.

A sequência completa do subfragmento Y-1.b2 com uma parte do plasmídeo pKK232-8 incluindo o códon de iniciação do gene CAT é mostrada na figura 19. Nesta sequência são mostrados os códons de terminação, os sítios de clonagem e a cadeia aberta de aminoácidos (ORF - *open reading frame*) encontrada. As regiões promotoras consenso encontradas podem ter permitido a transcrição do gene CAT.

Embora, a expressão do ORF encontrado dentro do subfragmento Y-1.b2, não possa ser inferida nem na bactéria nem no genoma humano pelos resultados obtidos até então, a presença deste ORF sugere a existência de um gene. Deve-se considerar que, em organismos procariotos que não possuem introns, a probabilidade de se encontrar um ORF com 50 códons é menor que 10%, com 100 códons é menor que 1%, portanto, esta probabilidade seria ainda menor já que se trata de uma sequência em humano (BISHOP & RAWLINGS, 1987).

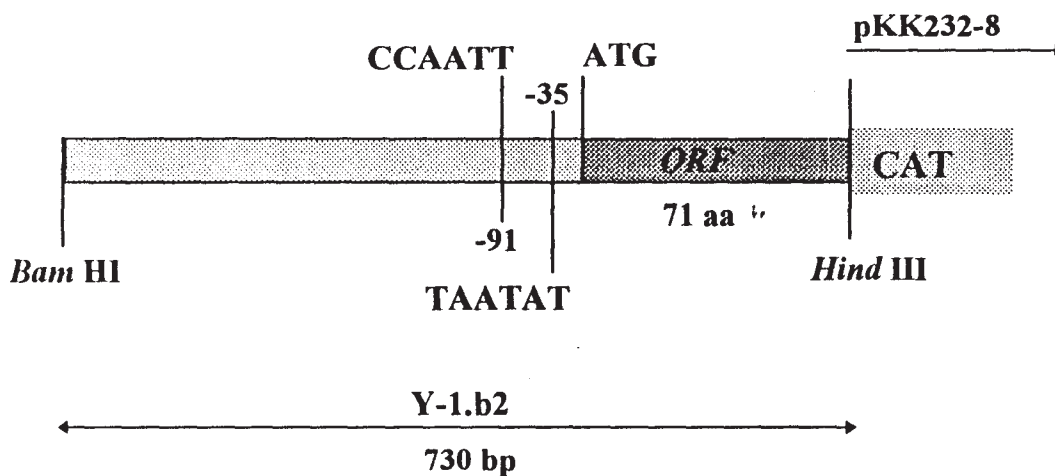


Figura 18 - Esquema das regiões promotoras encontradas na sequência Y-1.b2.

Apesar de alguns genes terem sido mapeados no cromossomo Y, são poucas as informações à respeito dos produtos por eles codificados. Embora o número de genes funcionais no cromossomo Y seja muito baixo em relação aos outros cromossomos, a possibilidade de se ter um gene ou parte dele já é argumento suficiente para maiores estudos. O ORF aqui encontrado poderia ser o início de um gene que foi restringido quando o fragmento Y-1 foi isolado, já que este segmento se encontra em um dos seus extremos. Esta sequência assume uma especial importância quando consideramos a possibilidade de ser parte de um gene ainda não isolado, uma vez que não apresenta homologia com genes já descritos e sequenciados. Além disso, provavelmente encontra-se no cromossomo Y humano dentro de uma região (intervalo 6) onde pelo menos três genes já foram mapeados (H-Y, GCY e AZF).

Embora seja discutível o fato de um promotor originário de um eucarioto estar se expressando num organismo procarioto, este sistema, já foi observado anteriormente por NEVE *et al.*, 1979. Entretanto, permanecem ainda incertas quais as regiões exatamente dos fragmentos eucarióticos que teriam sido reconhecidas pela RNA polimerase de *E.coli*.

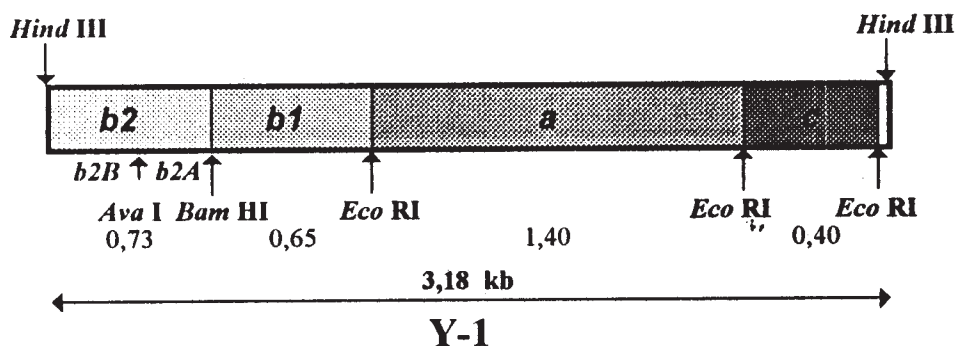


Figura 19 - Sequência Y-1.b2 contendo um trecho de sequência do plasmídeo pKK232-8.

CONCLUSÕES

CONCLUSÕES

A partir dos estudos realizados com a sequência Y-1 isolada de um banco do cromossomo Y, a qual é composta pelos subfragmentos ilustrados abaixo, seguintes conclusões podem ser delineadas:



1. O subfragmento Y-1.b1, inicialmente considerado Y-específico, apresenta um padrão de hibridização extremamente variável com as condições de estringência, não sendo, portanto, ideal para ser utilizado como sonda na detecção do cromossomo Y humano.
2. O subfragmento Y-1.c é, por outro lado, a melhor região da sequência Y-1 para reconhecer o cromossomo Y, pois produz uma banda *locus*-específica constante de 400 bp, somente quando hibridizado com DNA genômico humano de origem masculina. Por este motivo foi utilizado como sonda nos estudos de localização da sequência no cromossomo Y, da estimativa do número de cópias, de homologia e conservação nas espécies e seleção de *primers* para posterior emprego da técnica de PCR.
3. Embora não tenha sido excluída a possibilidade da sequência Y-1 estar distribuída ao longo de todo o cromossomo Y, com exceção da PAR e dos intervalos 1, 2 e parte do 3, os resultados das hibridizações da sonda Y-1.c com um painel de linhagens celulares apresentando alterações estruturais neste cromossomo, apontam fortemente que esta sequência estaria restrita aos intervalos 5-6 do Yq.

4. A sequência Y-1.c se repete no cromossomo Y humano aproximadamente 70 vezes e apresenta homologia com outras sequências presentes nos autossomos.
5. A sequência Y-1.c, durante a evolução, sofreu um processo de expansão e de particularização no cromossomo Y, passando de ausente em camundongos a Y-específica em humanos.
6. O sequenciamento do subfragmento Y-1.c revelou nas 394 bp, ausência de região codificante aparente e nenhuma homologia importante com outras sequências já descritas.
7. Os *primers* selecionados a partir da sequência Y-1.c, amplificam de maneira reproduzível em reações de PCR, um produto de 294 bp especificamente em DNA genômico de origem masculina.
8. O sequenciamento das 730 bp do subfragmento Y-1.b2, demonstrou a presença de um ORF de 71 aminoácidos com códon de iniciação para metionina, que se inicia na base 517 da sequência sendo interrompido pelo sítio de restrição *Hind* III o qual encerra o fragmento Y-1. A presença de regiões promotoras consenso para eucariotos, com um TATA *box* à -35 bp e um CAAT *box* à -91 bp do códon de iniciação do ORF, sugere a existência de parte de um gene interrompido pela restrição do fragmento Y-1.
9. A porção do subfragmento Y-1.b2 que inclui o ORF (Y-1.b2B), apresentou um padrão de hibridização diferencial entre DNA genômico de origem masculina e feminina, se prestando também como sonda na identificação do cromossomo Y.
10. A caracterização da sequência Y-1 permitiu identificar duas novas sondas para detecção do cromossomo Y (Y-1.b2B e Y-1.c), a produção de *primers* que amplificam uma sequência Y-específica e a indicação de um gene em potencial.

RESUMO

RESUMO

O cromossomo Y humano desempenha principalmente as funções de determinação do sexo masculino, relacionada com a presença do gene **SRY** no braço curto (Yp), e de controle da espermatogênese, regulada por gene(s) presente(s) no braço longo (Yq).

A relação entre a estrutura e função do cromossomo Y tem sido estudada através da análise molecular de seqüências de DNA derivadas deste cromossomo. Além disso, seqüências isoladas do cromossomo Y, quando bem caracterizadas, representam importantes ferramentas em estudos evolutivos e clínicos.

A seqüência de DNA Y-1 de aproximadamente 3,2 kb, foi isolada de um banco do cromossomo Y. Estudos prévios definiram quatro subfragmentos dentro desta seqüência (Y-1.a, Y-1.b1, Y-1.b2 e Y-1.c) a partir do mapa de restrição e sugeriram a presença de regiões promotoras.

Na expectativa de se isolar novas seqüências Y-específicas e possíveis genes, procedeu-se a caracterização detalhada do fragmento Y-1. Nossos resultados sugerem que o subfragmento Y-1.c sofreu uma expansão em mamíferos durante o processo evolutivo, tendendo a se restringir praticamente ao cromossomo Y, e que na espécie humana aparece repetida cerca de 70 vezes, provavelmente dentro dos intervalos 5-6 no braço longo deste cromossomo (Yq).

O subfragmento Y-1.c e parte do subfragmento Y-1.b2 (Y-1.b2B) apresentam padrão de hibridização Y-específico, tornando estas seqüências ideais para a utilização como sondas em estudos de identificação do cromossomo Y. O sequenciamento do subfragmento Y-1.c possibilitou a construção de *primers* que permitem a amplificação específica de DNA de origem masculina. A seqüência obtida a partir do subfragmento Y-1.b2 mostra regiões promotoras consenso para eucariotos e a presença de uma cadeia aberta de leitura (*open reading frame* - ORF) de 71 aminoácidos, sugerindo a existência de um gene.

ABSTRACT

ABSTRACT

The main functions of the human Y chromosome are male sex determination, related to the presence of **SRY** on the short arm (Yp), and the regulation of spermatogenesis which is controlled by a gene(s) on the long arm (Yq).

The relationship between the structure and function of the Y chromosome has been studied by molecular analysis of DNA sequences derived from this chromosome. Isolated Y chromosome sequences, when well-characterized, represent important tools in evolutionary and clinical studies.

The DNA Y-1 sequence of approximately 3.2 kb, was isolated from a Y chromosome library. Previous studies defined four fragments in this sequence (Y-1.a, Y-1.b1, Y-1.b2 and Y-1.c) based on restriction mapping and suggested the presence of promoter regions.

With the expectation of isolating new Y-specific sequences and possibly genes, we performed a detailed characterization of the Y-1 fragment. Our results suggest that the Y-1.c fragment has been expanded in mammals during their evolution, and is restricted almost exclusively to the Y-chromosome. The sequence is repeated about 70 times in humans, probably in the 5-6 intervals on the long arm of this chromosome (Yq).

The Y-1.c fragment and part of the Y-1.b2 fragment (Y-1.b2B) show a Y-specific hybridization pattern which makes these sequences ideal for use as probes in Y chromosome identification. The sequencing of the Y-1.c fragment allowed the design of primers for specific amplification of male DNA. The sequence obtained from the Y-1.b2 fragment reveals consensus promoter regions for eukaryotes as well as the presence of an open reading frame with 71 aminoacids, which suggest the existence of one gene.

BIBLIOGRAFIA

BIBLIOGRAFIA

- ALTSCHUL, F. S.; GISH, W.; MILLER, W.; MEYERS, E. W.; LIPMAN, D.
Basic local alignment search tool. *J. Mol. Biol.* **215**: 403-410, 1990.
- ALVESALO, L. & DE LA CHAPELLE, A. Tooth size in two males with deletions of long arm of the Y chromosome. *Am. Hum. Genet.* **54**: 49-54, 1981
- ANDERSON, M.; PAGE, D. C.; PETTAY, D.; SUBRT, I.; TURLEAU, C.; DE GROUCHY, J.; DE LA CHAPELLE, A. Y: autosome translocations and mosaicism in the aetiology of 45,X maleness: Assignment of fertility factor to distal Yq11. *Hum. Genet.* **79**: 2-7, 1988
- ARNEMANN, J.; JAKUBICZKA, S.; SCHÄFER, R.; EPPLEN, J. T. Clustered GATA repeats (*Bkm* sequences) on the human Y chromosomes. *Hum. Genet.* **73**: 301-303, 1986.
- BENVENISTE, R. S.; TODASO, C. J. Evolution of C type viral genes: evidence for any Asian origin of man. *Nature* **261**: 101-107, 1976.
- BISHOP, C. E.; GUELLAEN, G.; GELDWERTH, D.; FELLOUS, M.; WEISSENBACH, J. Extensive sequence homology between Y and other human chromosomes. *J. Mol. Biol.* **173**: 403-417, 1984.
- BISHOP, M. J.; RAWLINGS, J. *Nucleic acid and protein sequence analysis: a practical approach*. IRL Press Limited, Oxford, 1987.
- BOLIVAR, F.; BACKMAN, K. Plasmids of *Escherichia coli* as cloning vector. *Methods Enzymol.* **68**: 245, 1979.
- BOYER, H. W.; ROULLAND-DUSSOIX, D. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. *J. Mol. Biol.* **41**: 469, 1969.

- BLAGOWIDOW, N.; PAGE, D. C.; HUFF, D.; MENNUTI, M.T. Ullrich-Turner syndrome in an XY female fetus with deletion of the sex-determining portion of the Y chromosome. *Am. J. Med. Genet.* **34**: 159-162, 1989
- BRIDGES, C. B. Non disjunction as proof of the chromosome theory of heredity. *Genetics* **1**: 1-51, 107-163, 1916.
- BROSIUS, J. Plasmid vectors for the selection of promoters. *Gene* **936**:151-160, 1984.
- CASPERSSON, T.; ZECH, L.; JOHANSSON, C.; MODEST, E. J. Identification of human chromosomes by DNA - binding fluorescent agents. *Chromosoma*, **30**: 215-227, 1970.
- COOKE, H. Repeated sequence specific to human males. *Nature* **262**: 182-186, 1976.
- COOKE, H. J.; HINDLEY, J. Cloning of human satellite III: different components are on different chromosomes. *Nucl. Acids Res.* **6**: 3177-3198, 1979.
- COOKE, H. J.; SCHMIDTKE, J.; GOSDEN, J. R. Characterization of a human Y chromosome repeated sequence and related sequences in higher primates. *Chromosoma* **87**: 491-505, 1982.
- DAIGER, S. P. ; WILDIN, R. S.; SU T-S. Sequences on the Y-chromosome homologous to the autosomal gene for arginosuccinate syntetase. *Nature* **298**: 682-684, 1982
- DARLING, S. M.; BANTING, G. S. ; PYM, B.; WOLFE, J.; GOODFELLOW, P. N. Cloning and expressed gene shared by the human sex chromosome. *Proc. Natl. Acad. Sci. USA* **83**: 135-135, 1986.
- DITECHE, C. M.; CASANOVA, M.; SAAL, H.; FRIEDMAN, C.; SYBERT, V.; GRAHAM, J.; THULINE, H.; PAGE, D. C., FELLOUS, M. Small deletions of the short arm of the Y chromosome in 46,XY females. *Proc Natl. Acad. Sci. USA* **83**: 7841-7844, 1986.

- ELLIS, N. A.; GOODFELLOW, P. J.; PYM, B.; SMITH, M.; PALMER, M.; FRISCHAUF, A-M.; GOODFELLOW, P. N. The pseudoautosomal boundary in man is defined by an *Alu* repeat sequence inserted on the Y chromosome. *Nature* **337**: 81-84, 1989
- ELLISON, J.; PASSAGE, M.; YU, L-C.; YEN, P.; MOHANDAS, T. K.; SHAPIRO, L. Directed isolation of human genes that escape X inactivation. *Somat. Cell Mol. Genet.* **18**: 259-268, 1992
- ERLICH, H. A. GELFAND, D. H.; SAIKI, R. K. Specific DNA amplification. *Nature* **331**: 461-462, 1988.
- FARAH, S.B.; GARMES H.M.; CAVALCANTI D.P.; MELLO M.P.; PORRELLI R.N.; RAMOS, C.F.; SARTORATO E.L. Use of Y chromosome-specific DNA probes to evaluate an XX male. *Brazilian J. Med. Biol. Res.* **24**: 149-156, 1991.
- FARAH, S. B.; GRUENERT, D. C.; LEPERCQ, J. A.; CHUEH, J.; METZGER, A.; SARTORATO, E. L.; LEBO, R. V. GOLBUS, M. S. Isolation and characterization of Y chromosome DNA probes. *Biochem. Biophys. Res. Comm.* **189**: 581-589, 1992.
- FARAH, S. B.; RAMOS, C. F.; MELLO, M. P.; SARTORATO, E. L.; HORELLI-KUITUNEN, N.; LOPES, V. L. G. S.; CAVALCANTI, D. P.; HACKEL, C. Two cases of Y; autosome translocations: a 45,X male and a clinically trisomy 18 patient. *Am. J. Hum. Genet.* **49**:388-392, 1994.
- FEINBERG, A. P. & VOGELSTEIN, B. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**: 6, 1983.
- FISCHER, E. M. C.; BEER-ROMERO, P.; BROWN, L. G.; RIDLEY, A.; McNEIL, J. A.; LAWRENCE, J. B.; WILLARD, H. F.; BIEBER, F. R. PAGE, D. C. Homologous ribosomal protein genes on the human X and Y chromosomes: Escapes from X inactivation and possible implications for Turner syndrome. *Cell* **63**: 1205-1218, 1990

- FOOTE, S.; VOLLRATH, D.; HILTON, A. PAGE, D. C. The human Y chromosome: overlapping DNA clones spanning the euchromatic region. *Science* **258**: 60-66, 1992.
- FORD, C. E.; MILLER, O. J.; POLANI, P. E.; DE ALMEIDA, J. C.; BRIGGS, J. H. A sex-chromosome anomaly in a case of gonadal dysgenesis (Turner's syndrome). *Lancet* **I**: 711-713, 1959
- FRASER, N.; BALLABIO, A.; ZOLLO, M.; PERSICO, G.; CRAIG, I. Identification of complete coding sequences for steroid sulphatase on the human Y chromosome: evidence for an ancestral pseudoautosomal gene? *Development* **1015**: 127-132, 1987.
- FREIJE, D.; HELMES, C.; MICHAEL, S. W.; DONIS-KELLER, H. Identification of a second pseudoautosomal region near Xq and Yq telomeres. *Science* **258**: 1784-1787, 1992.
- GIBBONS, A. Looking for the father of us all. *Science* **251**: 378-379, 1991.
- GLICK, R. B.; THOMPSON, J. E. *Methods in plant molecular biology and biotechnology*. CRC Press, Florida, 1993.
- GOLDBERG, E. H.; BOYSE, E. A.; BENNETT, D.; SCHEID, M.; CARSWELL, E. A. Serological demonstration of H-Y (male) antigen on mouse sperm. *Nature* **302**: 346-349, 1983.
- GOMEZ-PEDROZO, M.; MOHANDAS, T.; SPARKES, R.; SHAW, J-P.; HESS, J. F.; AYRES, T. M.; SHEN, C-KJ. Evolution of human cytoplasmic actin gene sequences: chromosome mapping and structural characterization of three cytoplasmic actin-like pseudogenes including one the Y chromosome. *J. Hum. Evol.* **16**: 215-230, 1987
- GOODFELLOW, P.; BANTING, G.; SHEER, D.; ROPERS, H. H.; CAINE, A.; FERGUSON-SMITH, M. A.; POVEY, S.; VOSS, R. Genetic evidence that a Y-linked gene in man is homologous to a gene on the X chromosome. *Nature* **302**: 346-349, 1983.

- GOODFELLOW, P. N.; DARLING, S. M. Genetics of sex determination. *Development* **102**: 251-258, 1988.
- GOODFELLOW, P. N.; WEISSENBAACH, J. Report of the committee on the genetic constitution of the Y chromosome. Human Gene Mapping. *Cytogenet. Cell. Genet.* **58**: 967-985, 1991.
- GOUCH, N. M.; GEARING, D. P.; NICOLA, N. A.; BAKER, E.; PRITCHARD, M.; CALLEN, D. F.; SUTHERLAND, G. R.; Localization of the human GM-CSF receptor gene to the X-Y pseudoautosomal region. *Nature* **345**: 734-736, 1990.
- GUBBAY, Y. J.; COLLIGNON, J.; KOOPMAN, P.; CAPEL, B.; ECONOMOU, A.; MÜNSTERBERG, A.; VIVIAN, N.; GOODFELLOW, P.; LOVELL-BADGE, R. A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes, *Nature* **346**: 245-250, 1990.
- HANAHAN, D. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**: 557.
- HARLEY, V. R.; JACKSON, D. I.; HEXTALL, P. J.; HAWKINS, J. R.; BERKOVITZ, G. D.; SOCKANATHAN, S.; LOVELL-BADGE, R.; GOODFELLOW, P. N. DNA binding activity of recombinant SRY from normal males and XY females. *Science* **255**: 453-456, 1992.
- HENKE, A.; WAPENAAR, M.; VAN OMMEN, G. J.; MARASCHIO, P.; CAMERINO, G.; RAPPOLD, G. Deletions within the pseudoautosomal region help map three new markers and indicate a possible role of this region in linear growth. *Am. J. Hum. Genet.* **49**: 811-819, 1991.
- JACOBS, P. A. & STRONG, J. A. A case of human intersexuality having a possible XXY sex-determining mechanism. *Nature* **183**: 302-302, 1959.

- JAKUBICZKA, S.; ARNEMANN, J.; COOKE, H.K.; KRAWCZAK, M.; SCHMIDTKE, J. A search for restriction fragment length polymorphism on the human Y chromosome. *Hum. Genet.* **84**: 86-88, 1989.
- KOLLER, P. C. & DARLINGTON, C. D. The genetical and mechanical properties of the sex chromosomes. *J. Genet.* **29**: 159-173, 1934.
- KOOPMAN, P.; GUBBAY, J.; VIVIAN, N.; GOODFELLOW, P.; LOVELL-BADGE, R. Male development of chromosomally female mice transgenic for Sry. *Nature* **351**: 117-121, 1991.
- LAU, E. C.; MOHANDAS, T. K.; SHAPIRO, L. J.; SLAVKIN, H. C.; SNEAD, M. L. Human and mouse amelogenin gene loci are on the sex chromosome. *Genomics* **4**: 162-168, 1989.
- LEVILLIERS, J.; QUACK, B.; WEISSENBAACH, J. PETIT, C. Exchange of terminal portions of X- and Y-chromosomal short arms in human X-Y females. *Proc. Natl. Acad. Sci. USA* **86**: 2296-2300, 1989.
- MA, K.; SHARKEY, A.; KIRSCH, S.; VOGT, P.; KEIL, R.; HARGREAVE, T. B.; McBEATH, S.; CHANDLEY, A. C. Towards the molecular localization of the AZF locus: mapping of microdeletions in azoospermic men within 14 subintervals of interval 6 of the human Y chromosome. *Hum. Mol. Gen.* **1**: 29-33, 1992.
- MANDEL, M. & HIGA, A. Calcium dependent bacteriophage DNA infection. *J. Mol. Biol.* **53**: 159, 1970.
- MAGENIS, R. E.; TOMAR, D.; SHEEHY, R.; FELLOUS, M.; BISHOP, C.; CASANOVA, M. Y short arm material translocated to distal X short arm in XX males: evidence from in situ hybridization of a Y specific single copy DNA probe. *Am. J. Hum. Genet.* **36**: 1025, 1984.
- MANIATIS, T.; FRITISCH, E. F. & SAMBROCK, J. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, New York, USA. 1990.

- McCLUNG, C. E. The accessory chromosome-sex determinant? *Biol. Bull.* **3**: 43-84, 1902.
- McKUSICK, V. A. "*Mendelian Inheritance in Man*". Johns Hopkins Univ. Press, Baltimore, 1988.
- MESSING, J. New M13 vector for cloning. *Methods Enzimol.* **101**:20, 1983.
- MITCHELL, A. R.; GOSDEN, J. R.; MILLER, D. A. A cloned sequence p82H, of the alphoid repeated DNA family found at the centromeres of all human chromosomes. *Chromosoma* **92**: 369-377, 1985.
- MORGAN, T. H. Sex-limited inheritance in *Drosophila*. "*Science* **32**: 120-122, 1910.
- MÜLLER, U. Mechanisms of gonadal differentiation in vertebrates. *Differentiation* **23** (suppl.): S1-S2, 1983.
- NAKAHORI, Y.; TAKENAKA, O.; NAKAGOME, Y.; A human X-Y homologous region encodes "Amelogenin". *Genomics* **9**: 264-269, 1991.
- NEVE, R. L.; WEST, R. W.; RODRIGUEZ, R. L. Eukaryotic DNA fragments which act as promoters for a plasmid gene. *Nature* **277**: 324-325, 1979.
- NGO, K. Y.; VERGNAUD, G.; JOHNSON, C.; LUCOTTE, G.; WEISSENBAACH, J. A DNA probe detecting multiple haplotypes of the human Y chromosome. *Am. J. Hum. Genet.* **38**: 407-418, 1986.
- NYGARD, O & NIKA, H. Identification by RNA-protein cross-linking of ribosomal proteins located at the interface between the small and large subunits of mammalian ribosomes. *EMBO J.* **1**: 357-362, 1982.
- NORRANDER, J.; KEMPE, T.; MESSING, J. Construction of improved M13 vectors using oligodeoxynucleotide direct mutagenesis. *Gene* **26**:101, 1983.

- NOVAIS, R.C. *Caracterização molecular de uma sequência do cromossomo Y humano*. Campinas, 1990. 88 páginas. Tese de Mestrado, Universidade Estadual de Campinas.
- OHNO, S. Sex chromosomes and sex-linked genes. New York: Springer-Verlag, New York, 1967.
- PAINTER, T. Studies in mammalian spermatogenesis II. The spermatogenesis of man. *J. Exp. Zoology* **37**: 291-335, 1923.
- PAGE, D. C. Hypothesis: a Y-chromosomal gene causes gonadoblastoma in disgenetic gonads. *Development* **101 (suppl.)**: 151-155, 1987.
- PAGE, D. C.; FISCHER, E. M. C.; MCGILLIVRAY, B.; BROWN, L. G. additional deletion in sex-determining region of human Y chromosome resolves paradox of X, t(Y;22) female. *Nature* **346**: 279-281, 1990
- PALMER, M. S.; SINCLAIR, A. H.; BERTA, P.; ELLIS, N. A.; GOODFELLOW, P. N.; ABBAS, N. E.; FELLOUS, M. Genetic evidence that ZFY is not the testis-determining factor. *Nature* **342**: 937-939, 1989.
- RAMOS, C. F.; BORTOLETO, R. K.; CARDINALI, I. A.; VOGT, P.; FARAH, S. B. Estudo molecular da região cromossômica Yq11 em pacientes com azoospermia idiopática. *Brazil. J. Genetics* **16 (suppl. 3)**: 114, 1993.
- RAPPOLD, G. A. The pseudoautosomal regions of human sex chromosomes. *Hum. Genet.* **92**:315-324, 1993.
- RINEHART, F. P.; RITCH, T. G.; DEININGER, P. L.; SCHMID, C. W. Renaturation rate of a single family of interspersed repeated sequences in human deoxyribonucleic acid. *Byochemistry* **20**: 3003-3010, 1981.
- ROSENFELD, R. G.; LUZZATTI, L.; HINTZ, R. L.; MILLER, O. J.; KOO, G. C.; WACHTEL, S. S. Sexual and somatic determinants of the human Y chromosomes: studies in a 46,XYp⁻ phenotypic female. *Am. J. Hum. Genet.* **31**: 458-468, 1979.

- ROUYER, F.; SIMMLER, M-C; JOHNSON, C.; VERGNAUD, G.; COOKE, H. J.; WEISSENBAUGH, J. A gradient of sex linkage in the pseudoautosomal region of the human sex chromosomes. *Nature* **319**: 291-295, 1986.
- SANGER, F.; NICKLEN, S; COULSON, A. R. DNA sequencing with chain terminating inhibitors. *Proc. Nat. Acad. Sci. USA* **74**: 5463-5467, 1977.
- SEBOUN, E.; LEROY, P.; CASANOVA, M.; MAGENIS, E.; BOUCEKKINE, C.; DISTECHE, C.; BISHOP, C.; FELLOUS, M. A molecular approach to the study of the human Y chromosome and anomalies of sex determination in man. *Cold Spring Harbor Symp. Quant. Biol.* **51**: 237-248, 1986.
- SCHMECKPEPER, B. J.; WILLARD, H. F.; SMITH, K. D. Isolation and characterization of cloned human DNA fragments carrying reiterated sequences common to both autosomes and the X chromosome. *Nucl. Acids Res.* **8**: 1853-1872, 1981.
- SCHEMP, W. & MULLER, U. High resolution replication patterns of the human Y chromosome. Intra and interindividual variation. *Chromosoma* **86**: 229-237, 1982.
- SHIMOKAWA, H.; SOBEL, M. E.; SASAKI, M.; TERMINE, J. D.; YOUNG, M. F. Heterogeneity of amelogenin mRNA in the bovine tooth germ. *J. Biol. Chem.* **262**: 4042-4047, 1987.
- SINCLAIR, A.H.; BERTA, P.; PALMER, M.S.; HAWKINS, J.R.; GRIFFITHS, B.L.; SMITH, M.J.; FOSTER, J.W.; FRISCHAUF, A.M.; LOVELL-BADGE, R.; GOODFELLOW, P.N. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* **346**: 240-244, 1990.
- SINGH, L & JONES, K. W. *Bkm* sequences are polymorphic in humans and are clustered in pericentric regions of various acrocentric chromosomes including the Y. *Hum. Genet.* **73**: 304-308, 1986.

- SMITH, K. D.; YOUNG, K. E.; TALBOT, C. C.Jr; SCHMECKPEPER, B. J.
Repeated DNA of the human Y chromosome. *Development* **101** (suppl.): 77-92, 1987.
- SNEAD, M. L.; LAU, E. C.; ZEICHNER, D.; FINCHAM, A. G.; WOO, S. L.;
SLAVKIN, H. C. DNA sequence for cloned cDNA for murine amelogenin
reveal the amino acid sequence for enamel-specific protein. *Biochem.
Biophys. Res. Commun.* **129**: 812-818, 1985.
- SOUTHERN, E. Detection of specific sequences among DNA fragments separated
by gel eletrophoresis. *J. Mol. Biol.* **98**: 503, 1975.
- STEVENS, N. M. *Studies in spermatogenesis with especial reference to the
"accessory chromosome"*. Carnegie Institute Report 36 . Washington, DC,
1905.
- TIEPOLO, L. & ZUFFARDI, O. Localization of factors controlling
spermatogenesis in the nonfluorescent portion of the human Y chromosome
long arm. *Hum. Genet.* **34**: 119-124, 1976.
- TYLER-SMITH, C. & BROWN, W. R. A. Structure of the major block of alphoid
satllite DNA on the human Y chromosome. *J. molec. Biol.* **195**: 457-470,
1987.
- UCHIUMI, T.; KIKUCHI, M.; OGATA, K. Cross-linking study on protein
neighborhoods at the subunit interface of rat liver ribosomes with 2-
iminothiolane. *J. Biol. Chem.* **261**: 9663-9667, 1986.
- VERGNAUD, G.; PAGE, D. C.; SIMMLER, M. C.; BROWN, L.; ROYER, F.;
NOEL, B.; BOTSTEIN, D.; DE LA CHAPELLE, A.; WEISSENBAACH, J. A
deletion map of the human Y chromosome based on DNA hybridization. *Am.
J. Hum. Genet.* **38**: 109-124, 1986.

- VOGT, P.; CHANDLEY, A. C.; HARGREAVE, T. B.; KEIL, R.; MA, K.; SHARKEY, A. Microdeletions in interval 6 of the Y chromosome of the males with idiopathic sterility point to disruption of AZF , a human spermatogenesis gene. *Hum. Genet.* **89**: 491-496, 1992.
- VOGHT, P.; KEIL, R.; KÖHLER, M.; LENE, D.; LENE, G. Selection of DNA sequences from interval 6 of the human Y chromosome with homology to a Y chromosomal fertility gene sequence of *Drosophila hydei*. *Hum. Genet.* **86**: 341-349, 1991a.
- VOGHT, P.; CHANDLEY, A. C.; HARGREAVE, T. B.; KEIL, R.; MA, K.; SHARKEY, A. Towards the molecular localization of 3, a male fertility gene on the human Y chromosome by comparative mapping of microdeletions in the Y chromosome of men with idiopathic infertility. *Am. J. Hum. Genet.* **49**: 390, 1991b.
- VOLLRATH, D.; FOOTE, S.; HILTON, A.; BROWN, L. G.; BEER-ROMERO, P.; BOGAN, J. S.; PAGE, D. C. The human Y chromosome a 43-interval based on naturally occurring deletions. *Science* **258**: 52-59, 1992.
- WACHTEL, S.S.; KOO, G. C.; BREG., W. R.; ELIAS, S.; BOYSE, E. A. MILLER, O. J. Expression of H-Y antigen in human males with two Y chromosomes. *New Eng. J. Med.* **293**: 1070-1072, 1975b.
- WASBURN, L. L. & EICHER, E. M. Sex reversal in XY mice caused by dominant mutation on chromosome 17. *Nature* **303**: 338-340, 1983.
- WELSHONS, W. J. & RUSSEL, L. B. The Y chromosome as the bearer of male determining factors in the mouse. *Proc. Natl. Acad. Sci. USA* **45**: 560-566, 1959.
- WILSON, E. B. The chromosomes in relation to the determination of sex in insects. *Science* **22**: 500-502, 1905.

- WILSON, J. D.; GRIFFIN, J. E.; LESHIN, M.; GEORGE, F. W. Role of gonadal hormones in development of the sexual phenotypes. *Hum. Genet.* **58**: 78-84, 1981.
- WITKOP, C. J. & RAO, S. Inherited defects in tooth structure. *Brith defects* **7**: 153- 184, 1971.
- WOLFE, J.; DARLING, S. M.; ERICKSON, R. P.; CRAIG, R. W.; BUCKLE, V. J.; RIGBY, P. W. J.; WILLARD, H. F.; GOODFELLOW, P. N. Isolation and characterization of an alphoid centromeric repeat family from the human Y chromosome. *J. Molec. Biol.* **182**: 477-485, 1985.
- YANISCH-PERRON, C.; VIEIRA, J.; MESSING, J. Improved M13 phage cloning vectors and hosts strain: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**:103, 1985.
- YEN, P. H.; MARSH, B.; ALLEN, E.; TSAI, S. P.; ELLISON, J.; CONNOLLY, L.; NEISWANGER, K. & SHAPIRO, L. J. The human X-linked steroid sulfatase gene and a Y-encoded pseudogene: evidence for an inversion of the Y chromosome during primate evolution. *Cell* **55**: 1123-1135, 1988.
- YUNIS, J. J.; PRAKASH, O. The origin of man: A chromosomal pictorial legacy. *Science* **215**: 1525-1529, 1982.

APÊNDICE

APÊNDICE I

A - Linhagens bacterianas utilizadas:

1 - Bactéria DH5 α : esta linhagem é defectiva em recombinação (*Rec A1*). A deleção de genes do operon *lac* ($\Delta lacU16a$) e a presença do fragmento *lacZ* $\Delta M15$ no profago $\phi 80$ permite a utilização de α -complementação na seleção dos recombinantes. A ausência de superprodução do repressor do operon *lac* (*lacI^q*) possibilita a realização do teste de complementação na ausência de IPTG.

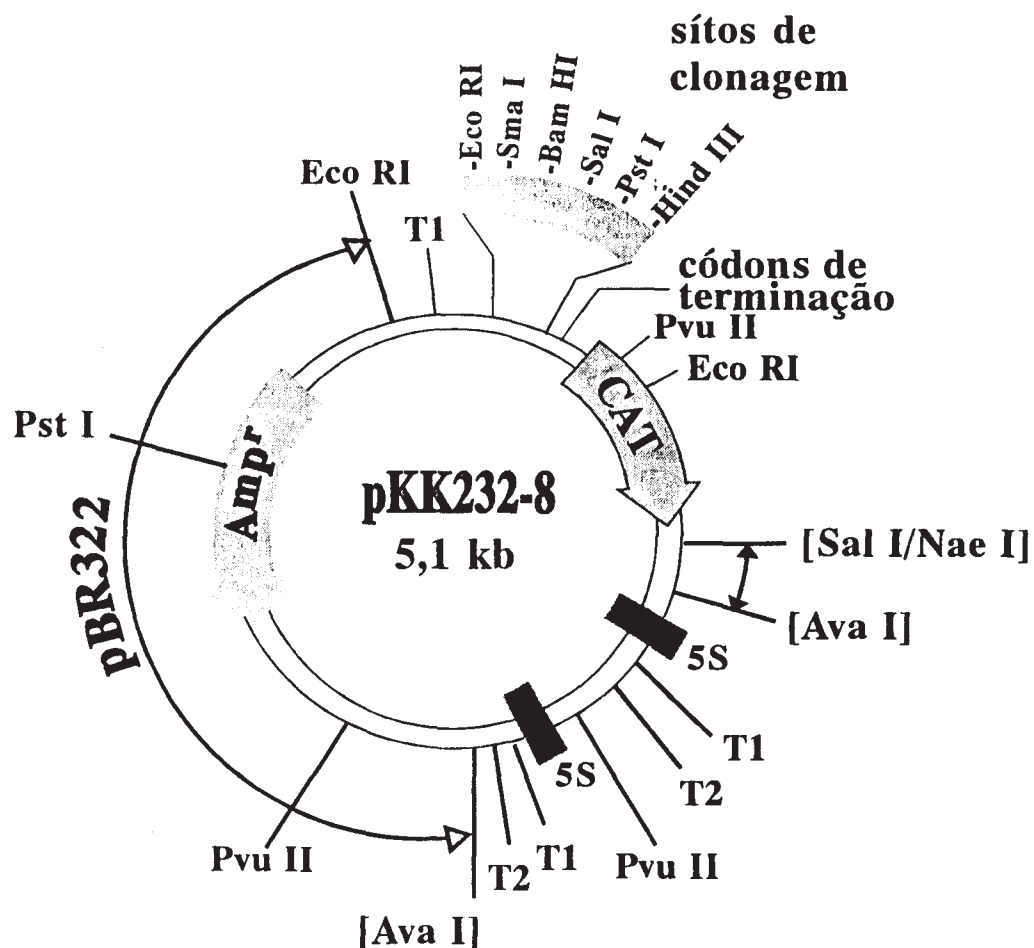
Genótipo relevante: *F- supE44 $\Delta(lacZYA-argF)$ U169 ($\phi 80 lacZ\Delta M15$) hsdR17 rec A1 end A1 gyrA96 thi-relA1* (HANAHAH, 1983).

2 - Bactéria HB101: esta linhagem é um híbrido das linhagens de *E.coli* K-12 e *E.coli* B. É uma hospedeira eficiente para transformação e obtenção de plasmídios em larga escala (BOYER & ROULLAND-DUSSOIX, 1969; BOLIVAR & BACKMAN, 1979).

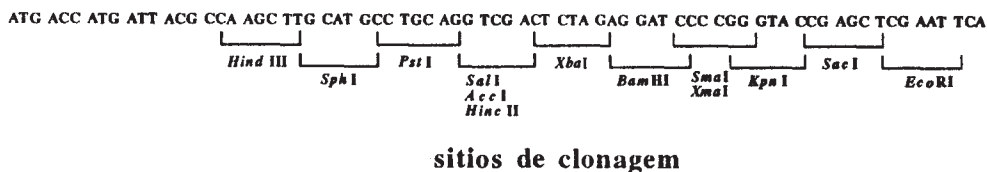
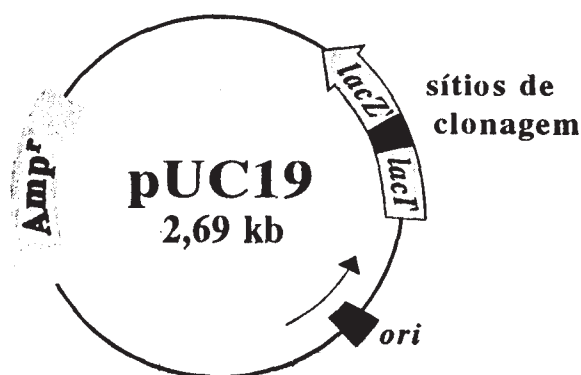
Genótipo relevante: $\Delta(gpt-proA)62 leuB6 thi-1 lacY1 hsdS_B20 recA rpsL20$ (*Str^r*) *ara-14 galK2 xyl-5 mtl-1 supE44 mcrB_B*.

B - Vetores utilizados:

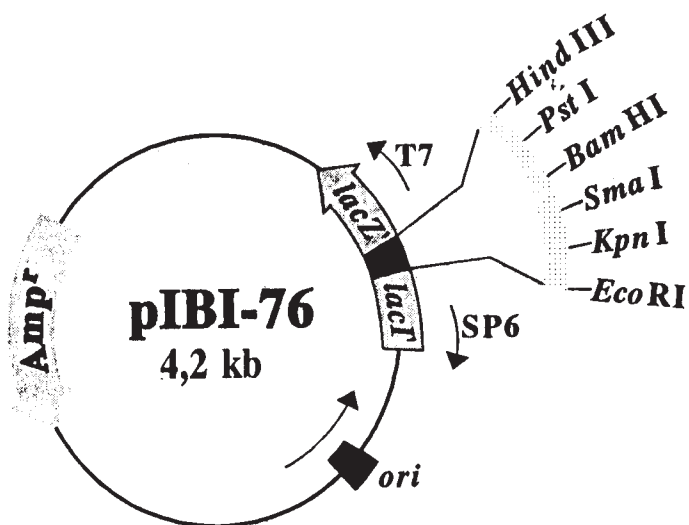
1 - Plasmídio pKK232-8: este vetor é derivado do plasmídio pBR322 e possui o gene que codifica para enzima cloranfenicol acetil-transferase (CAT) com a sua região promotora deletada. Por possuir os sítios de clonagem na região 5'-terminal do gene CAT permite a análise de regiões promotoras nele inseridas (BROSIUS, J., 1984).



2 - Plasmídio pUC19: possui 2690 bp e possibilita a identificação histoquímica das colônias obtidas pelo plaqueamento das células transformadas por um processo conhecido como α -complementação. Este vetor possui os sítios de clonagem inseridos dentro do gene *lacZ* que é a parte N-terminal que codifica para a enzima β -galactosidase do operon *lac* de *E.coli*. Esta parte do operon pode então ser complementada por uma forma defectiva da β -galactosidase codificada pela célula hospedeira. Obtida a α -complementação, a proteína sintetizada tem a capacidade de degradar o substrato X-gal que desenvolverá nas colônias transformadas a coloração azul. Uma vez clonado um fragmento no plasmídio haverá a inativação do gene *lacZ*, impossibilitando a complementação do polipeptídeo, as colônias transformadas recombinantes vão se apresentar com coloração branca. Além disso, este plasmídio possui o gene que confere à bactéria resistência ao antibiótico ampicilina (NORRANDER *et al.*; 1983).



3 - Plasmídio pIBI-76: este plasmídio é derivado do vetor pEMBL8(-) e contém a origem de replicação do bacteriófago F1 permitindo a produção de DNA simples fita. Possui o promotor T7 na fita não codificante e SP6 na fita codificante. Contêm os marcadores de seleção para resistência à ampicilina e identificação histoquímica pela β -galactosidade



APÊNDICE II - SEQUÊNCIAS DOS SUBFRAGMENTOS

A - Subfragmento Y-1.b2:

1 - Sequência de bases:

	11	21	31	41	51
5'	GATCCTGTCA	AACTGAGTTA	AAAGAAGGAT	AAATAGTAAC	AGGGTCATGA
3'	CTAGGACAGT	TTGACTCAAT	TTTCTTCCTA	TTTATCATTG	TCCAGTACT
	61	71	81	91	101
	ATTTTTTTTCC	TGGTATCTTC	AGCATATAGT	TCTGAGCTGA	TCAACAGCCT
	TAAAAAAAGG	ACCATAGAAG	TCGTATATCA	AGACTCGACT	AGTTGTCGGA
	111	121	131	141	151
	CATCACCCAT	TACCATCTGT	TGATTTACAG	TACCCAATGC	CTGTCCAATT
	GTAGTGGGTA	ATGGTAGACA	ACTAAATGTC	ATGGGTTACG	GACAGGTTAA
	161	171	181	191	201
	CCAAGCAATT	GATCAGATGT	GATATTAATA	AGAGGTTGGG	CCTGAGCATT
	GGTTCGTAA	CTAGTCTACA	CTATAATTAT	TCTCCAACCC	GGACTCGTAA
	211	221	231	241	251
	TCTCGCCTGC	CTGACTTGTT	GCCTTCATCC	TGTCCACCAG	GTTTTAAATT
	AGAGCGGACG	GACTGAACAA	CGGAAGTAGG	ACAGGTGGTC	CAAAATTTAA
	261	271	281	291	301
	GGAGAAATTG	AGATGGGGAC	AGGGTGGATT	GAGCTAATGA	CTCCAATCCA
	CCTCTTTAAC	TCTACCCCTG	TCCCACCTAA	CTCGATTACT	GAGGTTAGGT
	311	321	331	341	351
	TAGGTATTAA	GGGCCTGTTA	TAAGCCGCGA	AATTATAATG	AGGAATGAAC
	ATCCATAATT	CCCGGACAAT	ATTCGGCGCT	TTAATATTAC	TCCTTACTTG
	361	371	381	391	401
	CTCAGAAGAA	TTTGGCCTAT	ATTGTCCAAT	TGCTTTCTTA	AATTGCTTGT
	GAGTCTTCTT	AAACCGGATA	TAACAGGTTA	ACGAAAGAAT	TTAACGAACA
	411	421	431	441	451
	CCAATTTGGC	CTATATTGTC	CAATTGCTTT	CTTAAATTGC	TTGTCCAATT
	GGTTAAACCG	GATATAACAG	GTTAACGAAA	GAATTTAACG	AACAGGTTAA
	461	471	481	491	501
	TGGGCCTATA	TTGTCCAATT	GCTTTTTTAAT	ATTTTTTAAA	GAAAATGGCT
	ACCCGGATAT	AACAGGTTAA	CGAAAAATTA	TAAAAATTTT	CTTTTACCGA
	511	521	531	541	551
	CCCAGCAGGC	CCGAGCATGT	TCTCTGGGCT	CCTGAGCTGG	AGAAATAATT
	GGGTCGTCCG	GGCTCGTACA	AGAGACCCGA	GGACTCGACC	TCTTTATTAA
	561	571	581	591	601
	ACCAGAAACT	GCCAAGCATC	CAAATCCCCC	ATTTCTTGTG	CCTGTCAAAT
	TGGTCTTTGA	CGGTTCGTAG	GTTTAGGGGG	TAAAGAACAC	GGACAGTTTA
	611	621	631	641	651
	GGATGCCTGA	ATTGCCCCTG	CACCATAATT	TGTATTTGGA	CTGGCTCACA
	CCTACGGACT	TAACGGGGAC	GTGGTATTAA	ACATAAACCT	GACCGAGTGT
	661	671	681	691	701
	GCATTTTTCT	CCCACAATTA	ACAGGTGGAC	AAGCCTGGAT	CATTCCCTCA
	CGTAAAAAGA	GGGTGTTAAT	TGTCCACCTG	TTCGGACCTA	GTAAGGGAGT
	711	721			
	CCGTAATTGG	CTGTTGGGCT	GTTAAGCTT	3'	
	GGCATTAACC	GACAACCCGA	CAATTCGAA	5'	

2 - Tradução da sequência:

Ile Leu Ser Asn *** Val Lys Arg Arg Ile Asn Ser Asn Arg Val Met Asn Phe
 Asp Pro Val Lys Leu Ser *** Lys Lys Asp Lys *** *** Gln Gly His Glu Phe
 Gly Ser Cys Gln Thr Glu Leu Lys Glu Gly *** Ile Val Thr Gly Ser *** Ile

5' GGA TCC TGT CAA ACT GAG TTA AAA GAA GGA TAA ATA GTA ACA GGG TCA TGA ATT
 9 18 27 36 45 54
 3' CCT AGG ACA GTT TGA CTC AAT TTT CTT CCT ATT TAT CAT TGT CCC AGT ACT TAA

Ser Gly Thr Leu Ser Leu *** Phe Phe Ser Leu Tyr Tyr Cys Pro *** Ser Asn
 Ile Arg Asp Phe Gln Thr Leu Leu Leu Ile Phe Leu Leu Leu Thr Met Phe Lys
 Asp Gln *** Val Ser Asn Phe Ser Pro Tyr Ile Thr Val Pro Asp His Ile Lys

Phe Pro Gly Ile Phe Ser Ile *** Phe *** Ala Asp Gln Gln Pro His His Pro
 Phe Ser Trp Tyr Leu Gln His Ile Val Leu Ser *** Ser Thr Ala Ser Ser Pro
 Phe Phe Leu Val Ser Ser Ala Tyr Ser Ser Glu Leu Ile Asn Ser Leu Ile Thr

TTT TTC CTG GTA TCT TCA GCA TAT AGT TCT GAG CTG ATC AAC AGC CTC ATC ACC
 63 72 81 90 99 108
 AAA AAG GAC CAT AGA AGT CGT ATA TCA AGA CTC GAC TAG TTG TCG GAG TAG TGG

Lys Glu Gln Tyr Arg *** Cys Ile Thr Arg Leu Gln Asp Val Ala Glu Asp Gly
 Lys Gly Pro Ile Lys Leu Met Tyr Asn Gln Ala Ser *** Cys Gly *** *** Gly
 Lys Arg Thr Asp Glu Ala Tyr Leu Glu Ser Ser Ile Leu Leu Arg Met Val Trp

Leu Pro Ser Val Asp Leu Gln Tyr Pro Met Pro Val Gln Phe Gln Ala Ile Asp
 Ile Thr Ile Cys *** Phe Thr Val Pro Asn Ala Cys Pro Ile Pro Ser Asn ***
 His Tyr His Leu Leu Ile Tyr Ser Thr Gln Cys Leu Ser Asn Ser Lys Gln Leu

CAT TAC CAT CTG TTG ATT TAC AGT ACC CAA TGC CTG TCC AAT TCC AAG CAA TTG
 117 126 135 144 153 162
 GTA ATG GTA GAC AAC TAA ATG TCA TGG GTT ACG GAC AGG TTA AGG TTC GTT AAC

Met Val Met Gln Gln Asn Val Thr Gly Leu Ala Gln Gly Ile Gly Leu Leu Gln
 Asn Gly Asp Thr Ser Lys Cys Tyr Gly Ile Gly Thr Trp Asn Trp Ala Ile Ser
 *** Trp Arg Asn Ile *** Leu Val Trp His Arg Asp Leu Glu Leu Cys Asn Ile

Gln Met *** Tyr *** *** Glu Val Gly Pro Glu His Phe Ser Pro Ala *** Leu
 Ser Asp Val Ile Leu Ile Arg Gly Trp Ala *** Ala Phe Leu Ala Cys Leu Thr
 Ile Arg Cys Asp Ile Asn Lys Arg Leu Gly Leu Ser Ile Ser Arg Leu Pro Asp

ATC AGA TGT GAT ATT AAT AAG AGG TTG GGC CTG AGC ATT TCT CGC CTG CCT GAC
 171 180 189 198 207 216
 TAG TCT ACA CTA TAA TTA TTC TCC AAC CCG GAC TCG TAA AGA GCG GAC GGA CTG

Asp Ser Thr Ile Asn Ile Leu Pro Gln Ala Gln Ala Asn Arg Ala Gln Arg Val
 *** Ile His Tyr *** Tyr Ser Thr Pro Gly Ser Cys Lys Glu Gly Ala Gln Ser
 Leu His Ser Ile Leu Leu Leu Asn Pro Arg Leu Met Glu Arg Arg Gly Ser Lys

Val Ala Phe Ile Leu Ser Thr Arg Phe *** Ile Gly Glu Ile Glu Met Gly Thr
 Cys Cys Leu His Pro Val His Gln Val Leu Asn Trp Arg Asn *** Asp Gly Asp
 Leu Leu Pro Ser Ser Cys Pro Pro Gly Phe Lys Leu Glu Lys Leu Arg Trp Gly

TTG TTG CCT TCA TCC TGT CCA CCA GGT TTT AAA TTG GAG AAA TTG AGA TGG GGA
 225 234 243 252 261 270
 AAC AAC GGA AGT AGG ACA GGT GGT CCA AAA TTT AAC CTC TTT AAC TCT ACC CCT

Gln Gln Arg *** Gly Thr Trp Trp Thr Lys Phe Gln Leu Phe Gln Ser Pro Ser
 Thr Ala Lys Met Arg Asp Val Leu Asn *** Ile Pro Ser Ile Ser Ile Pro Val
 Asn Gly Glu Asp Gln Gly Gly Pro Lys Leu Asn Ser Phe Asn Leu His Pro Cys

Gly Trp Ile Glu Leu Met Thr Pro Ile His Arg Tyr *** Gly Pro Val Ile Ser
 Arg Val Asp *** Ala Asn Asp Ser Asn Pro *** Val Leu Arg Ala Cys Tyr Lys
 Gln Gly Gly Leu Ser *** *** Leu Gln Ser Ile Gly Ile Lys Gly Leu Leu ***

CAG GGT GGA TTG AGC TAA TGA CTC CAA TCC ATA GGT ATT AAG GGC CTG TTA TAA
 279 288 297 306 315 324
GTC CCA CCT AAC TCG ATT ACT GAG GTT AGG TAT CCA TAA TTC CCG GAC AAT ATT

 Leu Thr Ser Gln Ala Leu Ser Glu Leu Gly Tyr Thr Asn Leu Ala Gln *** Leu
 Pro His Ile Ser Ser Ile Val Gly Ile Trp Leu Tyr *** Pro Gly Thr Ile Leu
 Pro Pro Asn Leu *** His Ser Trp Asp Met Pro Ile Leu Pro Arg Asn Tyr Ala

Arg Glu Ile Ile Met Arg Asn Glu Pro Gln Lys Asn Leu Ala Tyr Ile Val Gln
 Pro Arg Asn Tyr Asn Glu Glu *** Thr Ser Glu Glu Phe Gly Leu Tyr Cys Pro
 Ala Ala Lys Leu *** *** Gly Met Asn Leu Arg Arg Ile Trp Pro Ile Leu Ser

GCC GCG AAA TTA TAA TGA GGA ATG AAC CTC AGA AGA ATT TGG CCT ATA TTG TCC
 333 342 351 360 369 378
CGG CGC TTT AAT ATT ACT CCT TAC TTG GAG TCT TCT TAA ACC GGA TAT AAC AGG

 Gly Arg Phe *** Leu Ser Ser His Val Glu Ser Ser Asn Pro Arg Tyr Gln Gly
 Arg Ser Ile Ile Ile Leu Phe Ser Gly *** Phe Phe Lys Ala *** Ile Thr Trp
 Ala Phe Asn Tyr His Pro Ile Phe Arg Leu Leu Ile Gln Gly Ile Asn Asp Leu

Leu Leu Ser *** Ile Ala Cys Pro Ile Trp Pro Ile Leu Ser Asn Cys Phe Leu
 Ile Ala Phe Leu Asn Cys Leu Ser Asn Leu Ala Tyr Ile Val Gln Leu Leu Ser
 Asn Cys Phe Leu Lys Leu Leu Val Gln Phe Gly Leu Tyr Cys Pro Ile Ala Phe

AAT TGC TTT CTT AAA TTG CTT GTC CAA TTT GGC CTA TAT TGT CCA ATT GCT TTC
 387 396 405 414 423 432
TTA ACG AAA GAA TTT AAC GAA CAG GTT AAA CCG GAT ATA ACA GGT TAA CGA AAG

 Ile Ala Lys Lys Phe Gln Lys Asp Leu Lys Ala *** Ile Thr Trp Asn Ser Glu
 Asn Ser Glu *** Ile Ala Gln Gly Ile Gln Gly Ile Asn Asp Leu Gln Lys Arg
 Gln Lys Arg Leu Asn Ser Thr Trp Asn Pro Arg Tyr Gln Gly Ile Ala Lys Lys

Lys Leu Leu Val Gln Phe Gly Pro Ile Leu Ser Asn Cys Phe Leu Ile Phe Leu
 *** Ile Ala Cys Pro Ile Trp Ala Tyr Ile Val Gln Leu Leu Phe Asn Ile Phe
 Leu Asn Cys Leu Ser Asn Leu Gly Leu Tyr Cys Pro Ile Ala Phe *** Tyr Phe

TTA AAT TGC TTG TCC AAT TTG GGC CTA TAT TGT CCA ATT GCT TTT TAA TAT TTT
 441 450 459 468 477 486
AAT TTA ACG AAC AGG TTA AAC CCG GAT ATA ACA GGT TAA CGA AAA ATT ATA AAA

 *** Ile Ala Gln Gly Ile Gln Ala *** Ile Thr Trp Asn Ser Lys Leu Ile Lys
 Leu Asn Ser Thr Trp Asn Pro Gly Ile Asn Asp Leu Gln Lys Lys Ile Asn Lys
 Phe Gln Lys Asp Leu Lys Pro Arg Tyr Gln Gly Ile Ala Lys *** Tyr Lys ***

Lys Glu Asn Gly Ser Gln Gln Ala Arg Ala Cys Ser Leu Gly Ser *** Ala Gly
 Lys Arg Lys Trp Leu Pro Ala Gly Pro Ser Met Phe Ser Gly Leu Leu Ser Trp
 *** Lys Lys Met Ala Pro Ser Arg Pro Glu His Val Leu Trp Ala Pro Glu Leu

TAA AAG AAA ATG GCT CCC AGC AGG CCC GAG CAT GTT CTC TGG GCT CCT GAG CTG
 495 504 513 522 531 540
ATT TTC TTT TAC CGA GGG TCG TCC GGG CTC GTA CAA GAG ACC CGA GGA CTC GAC

 Leu Leu Phe His Ser Gly Ala Pro Gly Leu Met Asn Glu Pro Ser Arg Leu Gln
 Phe Ser Phe Pro Glu Trp Cys Ala Arg Ala His Glu Arg Pro Glu Gln Ala Pro
 Phe Phe Ile Ala Gly Leu Leu Gly Ser Cys Thr Arg Gln Ala Gly Ser Ser Ser

Glu Ile Ile Thr Arg Asn Cys Gln Ala Ser Lys Ser Pro Ile Ser Cys Ala Cys
 Arg Asn Asn Tyr Gln Lys Leu Pro Ser Ile Gln Ile Pro His Phe Leu Cys Leu
 Glu Lys *** Leu Pro Glu Thr Ala Lys His Pro Asn Pro Pro Phe Leu Val Pro

GAG AAA TAA TTA CCA GAA ACT GCC AAG CAT CCA AAT CCC CCA TTT CTT GTG CCT
 549 558 567 576 585 594
 CTC TTT ATT AAT GGT CTT TGA CGG TTC GTA GGT TTA GGG GGT AAA GAA CAC GGA

Leu Phe Leu *** Trp Phe Ser Gly Leu Met Trp Ile Gly Trp Lys Lys His Arg
 Ser Ile Ile Val Leu Phe Gln Trp Ala Asp Leu Asp Gly Met Glu Gln Ala Gln
 Phe Tyr Asn Gly Ser Val Ala Leu Cys Gly Phe Gly Gly Asn Arg Thr Gly Thr

Gln Met Asp Ala *** Ile Ala Pro Ala Pro *** Phe Val Phe Gly Leu Ala His
 Ser Asn Gly Cys Leu Asn Cys Pro Cys Thr Ile Ile Cys Ile Trp Thr Gly Ser
 Val Lys Trp Met Pro Glu Leu Pro Leu His His Asn Leu Tyr Leu Asp Trp Leu

GTC AAA TGG ATG CCT GAA TTG CCC CTG CAC CAT AAT TTG TAT TTG GAC TGG CTC
 603 612 621 630 639 648
 CAG TTT ACC TAC GGA CTT AAC GGG GAC GTG GTA TTA AAC ATA AAC CTG ACC GAG

Asp Phe Pro His Arg Phe Gln Gly Gln Val Met Ile Gln Ile Gln Val Pro Glu
 *** Ile Ser Ala Gln Ile Ala Gly Ala Gly Tyr Asn Thr Asn Pro Ser Ala ***
 Leu His Ile Gly Ser Asn Gly Arg Cys Trp Leu Lys Tyr Lys Ser Gln Ser Val

Ser Ile Phe Leu Pro Gln Leu Thr Gly Gly Gln Ala Trp Ile Ile Pro Ser Pro
 Gln His Phe Ser Pro Thr Ile Asn Arg Trp Thr Ser Leu Asp His Ser Leu Thr
 Thr Ala Phe Phe Ser His Asn *** Gln Val Asp Lys Pro Gly Ser Phe Pro His

ACA GCA TTT TTC TCC CAC AAT TAA CAG GTG GAC AAG CCT GGA TCA TTC CCT CAC
 657 666 675 684 693 702
 TGT CGT AAA AAG AGG GTG TTA ATT GTC CAC CTG TTC GGA CCT AGT AAG GGA GTG

Cys Cys Lys Glu Gly Val Ile Leu Leu His Val Leu Arg Ser *** Glu Arg Val
 Leu Met Lys Arg Gly Cys Asn Val Pro Pro Cys Ala Gln Ile Met Gly Glu Gly
 Ala Asn Lys Glu Trp Leu *** Cys Thr Ser Leu Gly Pro Asp Asn Gly *** Arg

*** Leu Ala Val Gly Leu Leu Ser
 Val Ile Gly Cys Trp Ala Val Lys Leu
 Arg Asn Trp Leu Leu Gly Cys *** Ala

CGT AAT TGG CTG TTG GGC TGT TAA GCT T 3'
 711 720 729
 GCA TTA ACC GAC AAC CCG ACA ATT CGA A 5'

Thr Ile Pro Gln Gln Ala Thr Leu Ser
 Tyr Asn Ala Thr Pro Ser Asn Leu Lys
 Leu Gln Ser Asn Pro Gln *** Ala

3 - Mapa de restrição da sequência:

```

      10      20      30      40      50      60
****
5'  GGATCCTGTCAAACCTGAGTTAAAAGAAGGATAAATAGTAACAGGGTCATGAATTTTTTTC
    ^  ^      ^      ^      ^      ^      ^
    BamHI  AlwI DdeI MseI      MaeIII  BspHI      EcoRII
    BstI   BinI      NlaIII  MboII
    BstYI
    DpnI
    MboI
    MflI
    NdeII
    NlaIV
    Sau3AI
    XhoII

      70      80      90      100     110     120
****
CTGGTATCTTCAGCATATAGTTCTGAGCTGATCAACAGCCTCATCACCCATTACCATCTG
  ^  ^      ^      ^      ^      ^      ^
  ApyI  Eco57I      DdeIAluIDpnI  CviJI  MnlI
  BstNI      BclI  HphI
  MvaI      CviJI
  ScrFI      MboI
            NdeII
            Sau3AI

      130     140     150     160     170     180
****
TTGATTTACAGTACCCAATGCCTGTCCAATTCCAAGCAATTGATCAGATGTGATATTAAT
  ^  ^      ^      ^      ^      ^      ^
  TaqII      RsaI      TspEI  MfeI BclI      AseI
            TspEI DpnI      MnlI
            MboI      MseI
            NdeII      Tth111II
            Sau3AI      VspI
            ***

      190     200     210     220     230     240
AAGAGGTTGGGCCTGAGCATTCTCGCCTGCCTGACTTGTTGCCTTCATCCTGTCCACCA
    ^  ^      ^      ^      ^      ^
    AsuIDdeI      FokI      ApyI
    Cfr13I      EcoRII
    CviJI      MvaI
    HaeIII      ScrFI
    NspIV
    Pali
    Sau96I

      250     260     270     280     290     300
**
GGTTTTAAATTGGAGAAATTGAGATGGGGACAGGGTGGATTGAGCTAATGACTCCAATCC
  ^  ^      ^      ^      ^      ^
  AhaIII  TspEI  FinI      AluI  HinfI
  DraI      CviJI
  MseI      PleI
  TspEI

      310     320     330     340     350     360
ATAGGTATTAAGGGCCTGTTATAAGCCGCGAAATTATAATGAGGAATGAACCTCAGAAGA
  ^  ^      ^      ^      ^      ^
  MseIASuI      AccIIMnlI      DdeI  TspEI
    Cfr13I      BstUI
    CviJI      CviJI TspEI
    DraII      Fnu4HI
    EcoO109I      ThaI
    HaeIII
    NspIV
    Pali
    Sau96I

```

```

      370      380      390      400      410      420
ATTTGGCCTATATTGTCCAATTGCTTTCTTAAATTGCTTGCCAATTTGGCCTATATTGT
  ^  ^  ^      ^  ^      ^  ^      ^  ^      ^  ^
MnII CviJI      MfeI      MseI      TspEI      CviJI
      HaeI      TspEI      TspEI      HaeI
      HaeIII      Tth111II      HaeIII
      MboII      PalI
      PalI

      430      440      450      460      470      480
CCAATTGCTTTCTTAAATTGCTTGCCAATTTGGGCCTATATTGTCCAATTGCTTTTTAA
  ^  ^  ^      ^  ^      ^  ^      ^  ^      ^  ^
MfeI      MseI      TspEI      AsuI      MfeI      MseI
      TspEI      TspEI      Cfr13I      TspEI
      Tth111II      CviJI
      HaeIII
      NspIV
      PalI
      Sau96I

      490      500      510      520      530      540
TATTTTTTAAAGAAAATGGCTCCCAGCAGGCCCGAGCATGTTCTCTGGGCTCCTGAGCTG
  ^  ^  ^      ^  ^      ^  ^      ^  ^      ^  ^
SspI  AhaIII      CviJI      AquI      NlaIII      AocII  AluI
      DraI      NlaIV      AsuI      Nsp(7524)I  BanII  CviJI
      MseI      AvaI      AvrI      CviJI      GsuI
      Cfr13I      CviJI      DdeI
      HaeIII      EcoT38I
      NspIII      NlaIV
      NspIV      NspII
      PalI      SduI
      Sau96I

      550      560      570      580      590      600
GAGAAATAATTACCAGAACTGCCAAGCATCCAAATCCCCCATTTCTTGTGCCTGTCAAA
  ^  ^  ^      ^  ^      ^  ^      ^  ^      ^  ^
TspEI  FokI  AlwNI      SfaNI      SfaNI
      Tth111II

      610      620      630      640      650      660
TGGATGCCTGAATTGCCCTGCACCATAATTTGTATTGGACTGGCTCACAGCATTTTTC
  ^  ^  ^      ^  ^      ^  ^      ^  ^
TspEI  FokI      TspEI      BsrI
      CviJI

      670      680      690      700      710      720
TCCCACAATTAACAGGTGGACAAGCCTGGATCATTCCTCACCGTAATTGGCTGTTGGGC
  ^  ^  ^      ^  ^      ^  ^      ^  ^      ^  ^
MseI      ApyIDpnI  AlwI      MnII  CviJI      CviJI
TspEI      BstNI  HphI  BinI      TspEI
      CviJI
      EcoRII
      MboI
      MvaI
      NdeII
      Sau3AI
      ScrFI

      730
TGTTAAGCTT 3'
  ^  ^  ^
MseI  AluI
      CviJI
      HindIII

```

B - Subfragmento Y-1.c:

1 - Sequência de bases:

	10	20	30	40	50
5'	GAATTCAAAT	TGTTATCTCC	TCCACTGTTC	CCTGGAGTGC	TAACAGGTGA
3'	CTTAAGTTTA	ACAATAGAGG	AGGTGACAAG	GGACCTCACG	ATTGTCCACT
	60	70	80	90	100
	CAGAATAGCT	CAACTGTTGC	TTTTACCATA	TGTTAAGTTA	GGAGAAATCT
	GTCTTATCGA	GTTGACAACG	AAAATGGTAT	ACAATTCAAT	CCTCTTTAGA
	110	120	130	140	150
	CAGAAAAAAA	GGATATGGAG	CACAAATCAG	CAGGCAAGGC	ACCCTATTGG
	GTCTTTTTTT	CCTATACCTC	GTGTTTAGTC	GTCCGTTCCG	TGGGATAACC
	160	170	180	190	200
	TAAATCAAGC	TCTGACAATA	GTCCTATTGT	AGGTCACTAT	TCAAGGAAAA
	ATTTAGTTTCG	AGACTGTTAT	CAGGATAACA	TCCAGTGATA	AGTTCCTTTT
	210	220	230	240	250
	CAATTTGAGG	CTCTGGTCGA	TACAGGAGCA	GACGTGTCAT	CATAGCTTTT
	GTTAAACTCC	GAGACCAGCT	ATGTCCTCGT	CTGCACAGTA	GTATCGAAAA
	260	270	280	290	300
	ATCAACGGCC	CAAAAACCTGA	CCCAAACAAA	AGGCCCAGTG	GGTCTTGTTG
	TAGTTGCCGG	GTTTTTGACT	GGGTTTGTTT	TCCGGGTCAC	CCAGAACAAC
	310	320	330	340	350
	GGATTGAGGA	TTTGCTTATA	TCTCACCAGG	AGAGAATCAA	CTTCTTGTCT
	CCTAACTCCT	AAACGAATAT	AGAGTGGTCC	TCTCTTAGTT	GAAGAACAGA
	360	370	380	390	
	GGGTACCCAC	AAGACATCTT	AAGCTGTGCC	AGGAGCCAGA	ATTC 3'
	CCCATGGGTG	TTCTGTAGAA	TTCGACACGG	TCCTCGGTCT	TAAG 5'

2 - Tradução da sequência:

Ile Gln Ile Val Ile Ser Ser Thr Val Pro Trp Ser Ala Asn Arg *** Gln Asn
Asn Ser Asn Cys Tyr Leu Leu His Cys Ser Leu Glu Cys *** Gln Val Thr Glu
Glu Phe Lys Leu Leu Ser Pro Pro Leu Phe Pro Gly Val Leu Thr Gly Asp Arg

5' **GAA TTC AAA TTG TTA TCT CCT CCA CTG TTC CCT GGA GTG CTA ACA GGT GAC AGA**
 9 18 27 36 45 54

3' **CTT AAG TTT AAC AAT AGA GGA GGT GAC AAG GGA CCT CAC GAT TGT CCA CTG TCT**

Phe Glu Phe Gln *** Arg Arg Trp Gln Glu Arg Ser His *** Cys Thr Val Ser
Ile *** Ile Thr Ile Glu Glu Val Thr Gly Gln Leu Ala Leu Leu His Cys Phe
Asn Leu Asn Asn Asp Gly Gly Ser Asn Gly Pro Thr Ser Val Pro Ser Leu Ile

Ser Ser Thr Val Ala Phe Thr Ile Cys *** Val Arg Arg Asn Leu Arg Lys Lys
*** Leu Asn Cys Cys Phe Tyr His Met Leu Ser *** Glu Lys Ser Gln Lys Lys
Ile Ala Gln Leu Leu Leu Leu Pro Tyr Val Lys Leu Gly Glu Ile Ser Glu Lys

ATA GCT CAA CTG TTG CTT TTA CCA TAT GTT AAG TTA GGA GAA ATC TCA GAA AAA
 63 72 81 90 99 108

TAT CGA GTT GAC AAC GAA AAT GGT ATA CAA TTC AAT CCT CTT TAG AGT CTT TTT

Tyr Ser Leu Gln Gln Lys *** Trp Ile Asn Leu *** Ser Phe Asp *** Phe Phe
Leu Glu Val Thr Ala Lys Val Met His *** Thr Leu Leu Phe Arg Leu Phe Phe
Ala *** Ser Asn Ser Lys Gly Tyr Thr Leu Asn Pro Ser Ile Glu Ser Phe Phe

Gly Tyr Gly Ala Gln Ile Ser Arg Gln Gly Thr Leu Leu Val Asn Gln Ala Leu
Arg Ile Trp Ser Thr Asn Gln Gln Ala Arg His Pro Ile Gly Lys Ser Ser Ser
Lys Asp Met Glu His Lys Ser Ala Gly Lys Ala Pro Tyr Trp *** Ile Lys Leu

AAG GAT ATG GAG CAC AAA TCA GCA GGC AAG GCA CCC TAT TGG TAA ATC AAG CTC
 117 126 135 144 153 162

TTC CTA TAC CTC GTG TTT AGT CGT CCG TTC CGT GGG ATA ACC ATT TAG TTC GAG

Leu Ile His Leu Val Phe *** Cys Ala Leu Cys Gly Ile Pro Leu Asp Leu Glu
Pro Tyr Pro Ala Cys Ile Leu Leu Cys Pro Val Arg Asn Thr Phe *** Ala Arg
Ser Ile Ser Cys Leu Asp Ala Pro Leu Ala Gly *** Gln Tyr Ile Leu Ser Gln

Thr Ile Val Leu Leu *** Val Thr Ile Gln Gly Lys Gln Phe Glu Ala Leu Val
Asp Asn Ser Pro Ile Val Gly His Tyr Ser Arg Lys Thr Ile *** Gly Ser Gly
*** Gln *** Ser Tyr Cys Arg Ser Leu Phe Lys Glu Asn Asn Leu Arg Leu Trp

TGA CAA TAG TCC TAT TGT AGG TCA CTA TTC AAG GAA AAC AAT TTG AGG CTC TGG
 171 180 189 198 207 216

ACT GTT ATC AGG ATA ACA TCC AGT GAT AAG TTC CTT TTG TTA AAC TCC GAG ACC

Ser Leu Leu Gly Ile Thr Pro *** *** Glu Leu Phe Val Ile Gln Pro Glu Pro
Val Ile Thr Arg Asn Tyr Thr Val Ile *** Pro Phe Cys Asn Ser Ala Arg Thr
Cys Tyr Asp *** Gln Leu Asp Ser Asn Leu Ser Phe Leu Lys Leu Ser Gln Asp

Asp Thr Gly Ala Asp Val Ser Ser *** Leu Leu Ser Thr Ala Gln Lys Leu Thr
Arg Tyr Arg Ser Arg Arg Val Ile Ile Ala Phe Ile Asn Gly Pro Lys Thr Asp
Ser Ile Gln Glu Gln Thr Cys His His Ser Phe Tyr Gln Arg Pro Lys Asn ***

TCG ATA CAG GAG CAG ACG TGT CAT CAT AGC TTT TAT CAA CGG CCC AAA AAC TGA
 225 234 243 252 261 270

AGC TAT GTC CTC GTC TGC ACA GTA GTA TCG AAA ATA GTT GCC GGG TTT TTG ACT

Arg Tyr Leu Leu Leu Arg Thr Met Met Ala Lys Ile Leu Pro Gly Phe Val Ser
Ser Val Pro Ala Ser Thr Asp Asp Tyr Ser Lys Asp Val Ala Trp Phe Ser Val
Ile Cys Ser Cys Val His *** *** Leu Lys *** *** Arg Gly Leu Phe Gln Gly

Gln Thr Lys Gly Pro Val Gly Leu Val Gly Ile Glu Asp Leu Leu Ile Ser His
 Pro Asn Lys Arg Pro Ser Gly Ser Cys Trp Asp *** Gly Phe Ala Tyr Ile Ser
 Pro Lys Gln Lys Ala Gln Trp Val Leu Leu Gly Leu Arg Ile Cys Leu Tyr Leu

CCC AAA CAA AAG GCC CAG TGG GTC TTG TTG GGA TTG AGG ATT TGC TTA TAT CTC
 279 288 297 306 315 324

GGG TTT GTT TTC CGG GTC ACC CAG AAC AAC CCT AAC TCC TAA ACG AAT ATA GAG

Gly Phe Leu Leu Gly Leu Pro Asp Gln Gln Ser Gln Pro Asn Ala *** Ile Glu
 Trp Val Phe Pro Gly Thr Pro Arg Thr Pro Ile Ser Ser Lys Ser Ile Asp ***
 Leu Cys Phe Ala Trp His Thr Lys Asn Pro Asn Leu Ile Gln Lys Tyr Arg Val

Gln Glu Arg Ile Asn Phe Leu Ser Gly Tyr Pro Gln Asp Ile Leu Ser Cys Ala
 Pro Gly Glu Asn Gln Leu Leu Val Trp Val Pro Thr Arg His Leu Lys Leu Cys
 Thr Arg Arg Glu Ser Thr Ser Cys Leu Gly Thr His Lys Thr Ser *** Ala Val

ACC AGG AGA GAA TCA ACT TCT TGT CTG GGT ACC CAC AAG ACA TCT TAA GCT GTG
 333 342 351 360 369 378

TGG TCC TCT CTT AGT TGA AGA ACA GAC CCA TGG GTG TTC TGT AGA ATT CGA CAC

Gly Pro Ser Phe *** Ser Arg Thr Gln Thr Gly Val Leu Cys Arg Leu Ser His
 Trp Ser Leu Ile Leu Lys Lys Asp Pro Tyr Gly Cys Ser Met Lys Leu Gln Ala
 Leu Leu Ser Asp Val Glu Gln Arg Pro Val Trp Leu Val Asp *** Ala Thr Gly

Arg Ser Gln Asn
 Gln Glu Pro Glu Phe
 Pro Gly Ala Arg Ile

CCA GGA GCC AGA ATT C 3'
 387

GGT CCT CGG TCT TAA G 5'

Trp Ser Gly Ser Asn
 Leu Leu Trp Phe Glu
 Pro Ala Leu Ile

3 - Mapa de restrição da sequência:

```

      10      20      30      40      50      60
5' GAATTCAAATTGTTATCTCCTCCACTGTTCCCTGGAGTGCTAACAGGTGACAGAATAGCT
   ^  ^      ^      ^      ^      ^      ^
   EcoRI TspEI      ApyI      MaeIII      AluI
   RsrI      BstNI      Tsp45I      CviJI
   TspEI      EcoRI      HphI
              GsuI
              MnlI
              MvaI
              ScrFI
              SecI
      70      80      90      100      110      120
CAACTGTTGCTTTTACCATATGTTAAGTTAGGAGAAATCTCAGAAAAAAGGATATGGAG
           ^      ^      ^
           NdeI MseI      DdeI
      130      140      150      160      170      180
CACAAATCAGCAGGCAAGGCACCCTATTGGTAAATCAAGCTCTGACAATAGTCCTATTGT
   ^      ^      ^      ^      ^      ^
   AocII      BanI      AluI      EcoNI
   Bsp1286I      NlaIV      CviJI
   HgiAI
   NspII
   SduI
      190      200      210      220      230      240
AGGTCACTATTCAAGGAAAACAATTTGAGGCTCTGGTCGATACAGGAGCAGACGTGTCAT
   ^      ^      ^      ^      ^      ^
   MaeIII      MnlI      CviJI      TaqI      AflIII
   Tsp45I      TspEI      TthHB8I      MaeII
              SstIII
      250      260      270      280      290      300
CATAGCTTTTATCAACGGCCCCAAAACCTGACCCAAACAAAAGGCCAGTGGGTCTTGTG
   ^      ^      ^      ^      ^      ^
   AluI      AsuI      BceFI      AsuI      Tth111III MnlI
   CviJI      Cfr13I      BsrI
              CviJI
              Cfr13I
              CviJI
              HaeIII
              NspIV
              Pali
              Sau96I
              NspIV
              Pali
              Sau96I
      310      320      330      340      350      360
GGATTGAGGATTGCTTATATCTCACCAGGAGAGATCAACTTCTTGTCTGGGTACCCAC
           ^      ^      ^      ^      ^      ^
           HphI      ApyI      HinfI      Asp718I
              BstNI      BanIKpnI
              EcoRI      NlaIV
              MvaI      RsaI
              ScrFI
      370      380      390
AAGACATCTTAAGCTGTGCCAGGAGCCAGAATTC 3'
   ^      ^      ^      ^      ^      ^
   AflIIAluI      ApyI      CviJI
   MseICviJI      BstNI      EcoRI
              EcoRIINlaIVRsrI
              MvaI      TspEI
              ScrFI

```