B I B L I O T E C A
INSTITUTO DE QUÍMICA
Universidade de São Paulo

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE QUÍMICA

EXPRESSÃO DIFERENCIAL DE GENES EM CEPAS
DE *TRYPANOSOMA CRUZI* PELA TÉCNICA
DE MICROARRANJOS DE DNA

CASSIO DA SILVA BAPTISTA

TESE DE DOUTORADO

ORIENTADORA: PROFA. DRA. BIANCA ZINGALES

SÃO PAULO

2004

DEDALUS - Acervo - CQ

30100010400

Ficha Catalográfica

Elaborada pela Divisão de Biblioteca e
Documentação do Conjunto das Químicas da USP.

Baptista, Cassio da Silva
B222e Expressão diferencial de genes em cepas de *Trypanosoma*
cruzi pela técnica de microarranjos de DNA / Cassio da
Silva Baptista. -- São Paulo, 2004.
141p

Tese (doutorado) - Instituto de Química da Universidade
de São Paulo. Departamento de Bioquímica.
Orientador : Zingales, Bianca

I. Biologia molecular 2. Biologia celular : *Trypanosoma*
I. T. II. Zingales, Bianca, orientador.

574.88 CDD

“Expressão diferencial de genes em cepas de *Trypanosoma cruzi* pela técnica de microarranjos de DNA”

CASSIO DA SILVA BAPTISTA

Tese de Doutorado submetida ao Instituto de Química da Universidade de São Paulo como parte dos requisitos necessários à obtenção do grau de Doutor em Ciências – Área: Bioquímica.

Aprovado(a) por:

Profa. Dra. BIANCA SILVANA ZINGALES
IQ – USP
(Orientadora e Presidente)

Prof. Dr. WALTER COLLI
IQ – USP

Profa. Dra. ALINE MARIA DA SILVA
IQ – USP

Prof. Dr. JOSÉ FRANCO DA SILVEIRA FILHO
UNIFESP

Profa. Dra. MARIA ISABEL NOGUEIRA CANO
FCM – UNICAMP

SÃO PAULO
10 DE SETEMBRO 2004

AGRADECIMENTOS

A Bianca Zingales pela oportunidade de realizar esse trabalho e por sua orientação, essencial para minha formação científica e para o meu crescimento pessoal.

À minha esposa Mara pela companhia, carinho, amor e por estar sempre ao meu lado, me apoiando e incentivando. Além disto, pela ajuda constante na parte gráfica do meu trabalho, seja da tese, dos artigos, pôster e das apresentações.

Aos meus pais por serem exemplo de vida para mim.

Aos meus irmãos, Mauricio e Murilo, por todos os conselhos e conversas importantes para minha vida, e também pelo companheirismo e amizade.

À toda minha família mais extensa, pois é muito bom tê-los.

Ao Ricardo Vêncio (o físico) não só pela sua atuação marcante no trabalho, desde o planejamento dos experimentos até a submissão do artigo, mas também pelas horas de conversas inflamadas sobre todo tipo de assunto. Sua contribuição na análise estatística dos dados foi imprescindível para o trabalho.

À Sarah pela ótima convivência e por sua competência.

À Camila, uma grande amiga, sempre disposta a colaborar.

Aos colegas do laboratório Marcelo, Paula, Aurélio, Nancy, Margoth por toda ajuda no trabalho e pelo companheirismo.

Ao Aurélio por ceder membranas de bandas cromossômicas de várias cepas e de DNA genômico digerido com *Pst*I.

A Adriana Matsukuma por sua companhia e por sua colaboração durante a realização do trabalho.

Aos amigos Adriana, Paulo, Tié, Denise e Leandro e todos do CAGE, pioneiros nos trabalhos com microarranjo e pela ajuda nesse meu trabalho.

A meus antigos amigos do colégio, da faculdade e de outros lugares, que fazem minha vida mais feliz.

Aos professores Hugo Armelin e Aline M. da Silva pela oportunidade de trabalhar no CAGE e por estarem sempre disponíveis.

Ao Prof. Dr. Crodowaldo Pavan pelos grandes conselhos.

Ao Dr. Edson Rondinelli e Dr. Turan Urmenyi (UFRJ) por terem criado as bibliotecas de cDNA de *T. cruzi* utilizadas no trabalho e pela análise crítica de nossos resultados nas reuniões de Caxambu.

Ao Dr. Wim Degraeve e Dr. Adeilton Brandão (FIOCRUZ-RJ) por disponibilizarem cerca de 700 ESTs para que pudéssemos realizar o trabalho.

A Dra. Santuza Teixeira por disponibilizar as 45 ESTs da cepa Tulahuen.

Ao Prof. Dr. Francisco Nóbrega pelas sugestões, além das ótimas conversas.

Aos professores Fernando Reinach, Sérgio Verjovski, Mari Cleide Sogayar, Roberto Santelli e Hamza El Dorry do Instituto de Química e Arthur Gruber da Faculdade de Medicina Veterinária por disponibilizarem os seqüenciadores automáticos. Esse último também pelos conselhos e por disponibilizar o programa Phred de sua página.

Ao Dr. David Munroe pela oportunidade de trabalhar em seu laboratório nos Estados Unidos e pelas boas horas de conversas.

A Dra. Lisa Gangi pelos conselhos sobre microarranjo e pelo companheirismo nos Estados Unidos.

A todos os amigos do Laboratory of Molecular Technology (NIH-USA).

À FAPESP pelo suporte financeiro.

ÍNDICE

LISTA DE FIGURAS	i
LISTA DE TABELAS	iii
ABREVIACÕES	iv
RESUMO	vi
SUMMARY	viii
1. INTRODUÇÃO	1
1.1. <i>Trypanosoma cruzi</i> e a doença de Chagas	1
1.2. Controle da transmissão	4
1.3. <i>Trypanosoma cruzi</i> e a interação com a célula hospedeira	5
1.4. Expressão diferencial de genes de <i>Trypanosoma cruzi</i>	8
1.5. Projetos Genoma de Parasitas	12
1.6. Tecnologia de microarranjos de DNA	14
1.7. Microarranjos de DNA no estudo de protozoários patogênicos e sua interação com a célula hospedeira	16
1.8. O Consórcio CAGE	18
2. OBJETIVO	19
3. JUSTIFICATIVA	19
4. MATERIAL E MÉTODOS	20
4.1. Soluções	20
4.2. Reagentes e enzimas	20

4.3. Oligonucleotídeos	21
4.4. Isolados de <i>T. cruzi</i> e seu cultivo	21
4.5. Cultivo e purificação de formas tripomastigotas metacíclicas	22
4.5.1. Preparação de DEAE-celulose	23
4.5.2. Purificação de formas tripomastigotas metacíclicas	23
4.6. Bactérias e células competentes	24
4.6.1. Bactérias	24
4.6.2. Meios para cultivo de bactérias	24
4.6.3. Células competentes	24
4.6.4. Transformação	25
4.7. Preparo de DNA plasmidial	25
4.7.1. Digestão com enzimas de restrição	26
4.7.2. Eletroforese convencional	26
4.7.3. Recuperação e purificação de fragmentos de DNA e ligação ao vetor	26
4.8. Obtenção de DNA para seqüenciamento	27
4.8.1. Minipreparação por lise alcalina	27
4.8.2. DNA obtido por PCR de colônia	27
4.8.3. PCR direto de DNA diluído	28
4.9. Seqüenciamento (Tubos individuais e Placas de 96 poços)	28
4.10. Anotação - Busca de homologia em bancos de dados, limpeza de seqüências e classificação por função	29
4.11. Agrupamento de seqüências pelo programa Cap 3	30
4.12. Amplificação, análise e purificação de ESTs	30

4.12.1. Amplificação das ESTs (em Placas de 96 poços)	30
4.12.2. Análise dos produtos de amplificação	31
4.12.2.1. Eletroforese em gel de agarose (96 poços)	31
4.12.3. Purificação dos produtos de PCR (placas de 96 poços)	31
4.13. Construção das lâminas de microarranjos de DNA de <i>T. cruzi</i>	32
4.13.1. Sondas de DNA	32
4.13.2. Amplificação das seqüências e aplicação na lâmina de microarranjos	34
4.13.3. Coloração da lâmina com VISTRA GREEN	37
4.14. Extração de DNA total de <i>T. cruzi</i>	37
4.15. Extração de RNA das cepas	38
4.15.1. Tratamento de RNA com DNase	38
4.16. Marcação de DNA total de CL Brener e Silvio com Cy3 ou Cy5	39
4.16.1. Purificação dos DNAs marcados e quantificação	39
4.17. Hibridização da lâmina com DNA total de <i>T. cruzi</i> marcado com corantes Cy3 ou Cy5	40
4.18. Lavagens após hibridização	40
4.19. Marcação do RNA total das cepas com Cy3 ou Cy5	40
4.20. Obtenção de imagens e análise dos dados	41
4.20.1. Experimentos para validação do microarranjo de DNA	41
4.20.2. Análise da expressão diferencial de genes em cepas isoladas de pacientes chagásicos	42
4.20.3. Análise estatística	42
4.21. Confirmação das hibridizações diferenciais	43

4.21.1. Southern blot	44
4.21.2. Northern Blot	44
4.22. Obtenção de sondas radioativas e condições de hibridização	45
4.22.1. Marcação por “random primer extension”	45
4.22.2. Purificação da sonda em coluna de Sephadex	45
4.22.3. Hibridização de sondas radioativas com ácidos nucleicos imobilizados	45
4.22.4. Lavagem das membranas	46
4.22.5. Desibridização das membranas	46
4.23. Análise da expressão gênica de fibroblastos humanos infectados com duas cepas de <i>T. cruzi</i>	47
4.23.1. Cultivo de parasitas para infecção de fibroblastos humanos	47
4.23.2. Cultivo de Fibroblastos Humanos (FHN) e infecção com cepas de <i>T. cruzi</i>	47
4.23.3. Meio condicionado de parasitas (MCP)	47
4.24. Extração de RNA de fibroblastos humanos não infectados, infectados, e cultivados em meio condicionado	48
4.25. Preparação das sondas, marcação e hibridização	48
4.26. Obtenção de imagem e análise dos dados	49
4.27. Confirmação das hibridizações diferenciais por “real time PCR” RT-PCR	50
5. RESULTADOS	51
5.1. Construção e validação dos microarranjos de DNA para o estudo da expressão diferencial de genes e de genômica comparativa em cepas de <i>T. cruzi</i>	51

5.1.1. Caracterização dos alvos de DNA para o conjunto A de lâminas de microarranjo	51
5.1.1.1. Sequenciamento de ESTs não depositados em bancos de dados	52
5.1.1.2. Genes de diversas cepas de <i>T. cruzi</i> aplicados no microarranjo	54
5.1.2. Amplificação e purificação das ESTs	56
5.1.3. Topologia do conjunto A de microarrajios	57
5.1.4. Análise de genômica comparativa de isolados de CL Brener e Silvio X10 cl1	59
5.1.5. Expressão diferencial de genes em isolados de CL Brener e Silvio	63
5.1.6. Expressão diferencial de genes em isolados de pacientes chagásicos	67
5.2 <i>Transcrição diferencial de DNA satélite em cepas de T. cruzi</i>	74
5.3. Expressão diferencial de genes em seis cepas de <i>T. cruzi</i> isoladas de indivíduos com distintas manifestações da doença de Chagas	78
5.3.1 Caracterização dos alvos de DNA para o conjunto B de lâminas de microarranjos	78
5.3.2. Construção dos microarranjos e obtenção de imagens	79
5.3.3. Expressão diferencial de genes em seis isolados de pacientes chagásicos	80
5.3.4. Análise dos Resultados	81
5.4. Análise da expressão gênica de fibroblastos humanos infectados com duas cepas de <i>T. cruzi</i> (Estágio no Exterior)	93
5.4.1. Obtenção de formas infectantes e infecção de monocamadas	93

5.4.2. Isolamento de RNA	94
5.4.3. Hibridização com microarranjos	95
5.4.4. Validação dos dados do microarranjo	97
6. DISCUSSÃO	99
6.1. Construção e validação dos microarranjos de DNA para o estudo da expressão diferencial de genes e de genômica comparativa em cepas de <i>T. cruzi</i>	99
6.1.1. Caracterização das ESTs	99
6.1.2. Anotação das ESTs	100
6.1.3. Outros genes utilizados na construção do microarranjo	102
6.1.4. Agrupamento das seqüências depositadas nas lâminas	106
6.1.5. Estudo de genômica comparativa e expressão diferencial de genes em Silvio e CL Brener	106
6.1.6. Expressão diferencial de genes em dois isolados de pacientes chagásicos	110
6.2. Transcrição diferencial de DNA satélite em cepas de <i>T. cruzi</i>	113
6.3. Expressão diferencial de genes em seis isolados de pacientes chagásicos	115
6.4. Expressão gênica de fibroblastos humanos infectados com cepas de <i>T. cruzi</i>	123
6.5. Comentários e conclusões finais	128
7. REFERENCIAS BIBLIOGRÁFICAS	130

LISTA DE FIGURAS

Figura 1	O esquema da lâmina de microarranjo com 12 subarranjos	36
Figura 2	PCR de colônia de bactérias transformadas com 7 ESTs	53
Figura 3	Análise em gel de agarose dos produtos de amplificação de 7 ESTs com os iniciadores T3 e T7	53
Figura 4	Análise em gel de agarose dos produtos de amplificação de 16 ESTs com os iniciadores T3 e T7	54
Figura 5	Hibridização de DNA de bandas cromossômicas com sonda do gene de rRNA 24Sa. Cepas.....	55
Figura 6	Southern blot de DNA genômico digerido com <i>Pst</i> I e hibridizado com sonda da RNA helicase.....	56
Figura 7	Bandas cromossômicas hibridizadas com a sonda da RNA helicase.....	56
Figura 8	Produtos de amplificação por PCR dos ESTs de duas placas de 96 poços.....	57
Figura 9	Gráfico de dispersão e gráfico MA da hibridização do microarranjo com cDNA de CL Brener da forma epimastigota marcado com Cy3 e Cy5.....	58
Figura 10	Lâmina de microarranjo de <i>T. cruzi</i> hibridizado com DNA de CL Brener e Silvio marcados com Cy3 e Cy5 respectivamente.....	61
Figura 11	Southern blot do DNA genômico de CL Brener e Silvio digerido com <i>Pst</i> I e hibridizado com sondas marcadas com P ³²	62
Figura 12	Northern blot de RNA total de epimastigotas de CL Brener e Silvio hibridizado com sondas marcadas com P ³²	66
Figura 13	Northern blot de RNA total de isolados de epimastigotas e tripomastigotas metacíclicos de Famema e Hem 179 hibridizados com sondas marcadas com P ³²	72
Figura 14	Alinhamento das ESTs TEUF0045, TEUF0146, TENF0857 e da região ND7 da cepa Silvio x10 com o programa CLUSTAL W.....	73
Figura 15	Northern blot de RNA total de epimastigotas de CL Brener e Silvio hibridizado com com sonda do DNA satélite da cepa Y.....	75
Figura 16	Northern blot de RNA total de CL Brener submetido a distintos tratamentos e hibridizado com sondas de DNA satélite da cepa Y e da cepa Silvio.....	75
Figura 17	Northern blot de RNA total de CL Brener submetido a distintos tratamentos e hibridizado com sonda de DNA satélite da cepa Y.....	77
Figura 18	Northern blot de RNA de várias cepas e estágios, hibridizado com sonda de DNA satélite da cepa Y.....	77
Figura 19	Normalização dos dados de hibridização da lamina com VL 10 versos B13.....	80
Figura 20	Gráficos MA e a determinação do valor de E pelo método do Erro de Classificação de Bayes para seis seqüências.....	85
Figura 21	Gráfico MA com os resultados normalizados de todas as réplicas da EST TENF0857 e legenda de cores do gráfico.....	86

Figura 22	Gráfico MA com os resultados normalizados de todas as réplicas da sonda TENF0103.....	86
Figura 23	Northern blot hibridizado com as sondas com expressão induzida nas cepas de cardíacos comparadas com as cepas de assintomáticos.....	89
Figura 24	Northern blot hibridizado com as sondas com expressão induzida nas cepas de cardíacos comparadas com as cepas de assintomáticos.....	90
Figura 25	Northern blot hibridizado com as sondas com expressão reduzida nas cepas de cardíacos comparadas com as cepas de assintomáticos.....	91
Figura 26	Gel de agarose das amostras de RNA isoladas de fibroblastos Infectados.....	95
Figura 27	PCR em tempo real com oligonucleotídios correspondendo aos genes protease ubiquitina específica, heme oxigenase e gene proteína 5 ligante do fator de crescimento do tipo insulina.....	98
Figura 28	Gráfico MA das ESTs TENU0045 (A) e TENU0146 (B).....	118

LISTA DE TABELAS

Tabela 1	Características de isolados de <i>T. cruzi</i>	22
Tabela 2	Características das seqüências de <i>T. cruzi</i> e DNAs controles imobilizados na lâmina de microarranjos.....	33
Tabela 3	Características das seqüências de <i>T. cruzi</i> imobilizadas no conjunto B de lâminas de microarranjos.....	34
Tabela 4	Agrupamento das 730 seqüências por CAP3.....	52
Tabela 5	Seqüências com maior hibridização com DNA de CL Brener ou Silvio X10 e mais expressas em CL Brener ou SilvioX10.....	59
Tabela 6	Seqüências com maior hibridização com CL Brener ou SilvioX10 DNA e igualmente expressas.....	60
Tabela 7	Seqüências diferencialmente expressas nas cepas Silvio e CL Brener	64
Tabela 8	Seqüências diferencialmente expressas nas formas epimastigotas das cepas Famema Hem 179	67
Tabela 9	Seqüências diferencialmente expressas em formas tripomastigota metacíclicas das cepas Famema e Hem 179	68
Tabela 10	Seqüências coincidentes diferencialmente expressas em epimastigotas e tripomastigotas metacíclicos das cepas Famema e Hem 179	71
Tabela 11	Agrupamento de 787 seqüências de <i>T. cruzi</i> por Cap 3.....	79
Tabela 12	Genes diferencialmente expressos em três cepas de cardíacos em relação às três cepas de assintomáticos.....	82
Tabela 13	Análise dos transcritos das seqüências identificadas pelos microarranjos como diferencialmente expressas em cepas de cardíacos em relação às cepas de assintomáticos.....	92
Tabela 14	Sondas diferencialmente expressas em fibroblastos humanos infectados com a cepa 147 por 24 horas em relação a fibroblastos não infectados	96
Tabela 15	Sondas diferencialmente expressas em fibroblastos humanos infectados com a cepa VL10 por 24 horas em relação a fibroblastos não infectados	97
Tabela A	Resumo dos dados do seqüenciamento de 134 ESTs.....	CD anexo
Tabela B	Valores de correlação de Spearman	CD anexo
Tabela C	Valores de correlação de Pearson	CD anexo

ABREVIACÕES

- ARE**- elemento rico em AU
- ATP**- trifosfato de adenosina
- BSA**- albumina bovina sérica
- Cy3**- corante de cianina 3
- Cy5**- corante de cianina 5
- dATP**- trifosfato de 2'-desoxiadenosina
- dCTP**- trifosfato de 2'-desoxicitidina
- ddATP**- trifosfato de 2', 3'-didesoxiadenosina
- ddCTP**- trifosfato de 2', 3'-didesoxicitidina
- ddGTP**- trifosfato de 2', 3'-didesoxiguanosina
- ddTTP**- trifosfato de 2', 3'-didesoxitimidina
- dGTP**- trifosfato de 2'-desoxiguanosina
- dTTP**- trifosfato de 2'-desoxitimidina
- DEP**- dietilpirocarbonato
- DMEM**- Dulbecco's modified Eagle's medium
- DMSO**- dimetilsufóxido
- DNA**- ácido desoxirribonucléico
- DNase**- desoxirribonuclease
- DNTP**- mistura de 2'-desoxinucleosídeos trifosfato
- DTT**- ditioneitol
- EDTA**- ácido etilenodiaminotetracético
- GRE**- elemento rico em G
- HEPES**- N-[2-hidroxietil] piperazina-N'-[ácido-2-etanossulfônico]
- Kb**- milhares de pares de base (quilobases)
- kDNA**- DNA de cinetoplasto
- LIT**- infusão de fígado e triptose
- mRNA**- RNA mensageiro
- μCi**- microcuries
- μg**- microgramas

- μl**- microlitros
- nm**- nanômetro
- nt**- nucleotídeo
- pb**- pares de bases
- PBS**- solução salina tamponada com fosfato
- PCR**- reação de polimerização em cadeia
- pmol**- picomoles
- RAPD**- DNA polimórfico amplificado aleatoriamente
- rDNA**- genes de RNA ribossômico
- RNA**- ácido ribonucleico
- RNase**- ribonuclease
- rpm**- rotações por minuto
- rRNA**- RNA ribossômico
- RTPCR**- PCR com etapa prévia de transcrição reversa
- SDS**- lauril sulfato de sódio
- SMUG**- gene de mucina
- SSC**- solução salina com citrato de sódio
- TBE**- solução Tris-borato-EDTA
- TcUBP-1**- proteína que se liga à região ARE
- TE**- solução de Tris-EDTA
- Tris**- tris(hidroximetil)aminometano
- UTR**- regiões não traduzidas
- UV**- radiação ultravioleta
- x-Gal**- 5-bromo-4-cloro-3-indolil-β-D-galactopiranosídeo

RESUMO

Com o objetivo de analisar a expressão diferencial de genes em cepas de *Trypanosoma cruzi* pela técnica de microarranjos de DNA, construímos um protótipo de lâminas contendo 710 etiquetas de seqüências transcritas (ESTs) de epimastigotas de CL Brener e 20 genes de várias cepas do parasita. A seqüência nucleotídica estava disponível para 576 ESTs. Desta forma, seqüenciamos as 134 ESTs restantes e sua seqüência foi depositada no dbEST do GenBank. O agrupamento das sondas indicou que a lâmina continha 665 seqüências únicas.

Inicialmente, avaliamos o microarranjo para a análise de genômica comparativa entre as cepas CL Brener (*T. cruzi* II) e Silvio (*T. cruzi* I). Um total de 4 hibridizações foram realizadas. Verificamos, com uma significância estatística ($P \leq 0,01$ no teste-*t*) que 31 seqüências do microarranjo (4,3%) apresentaram maior sinal de hibridização com o DNA de CL Brener, enquanto 37 sondas (5,0%) apresentaram situação inversa. Esse resultado sugere diferenças no número de cópias gênicas e/ou variação na similaridade das seqüências. Várias sondas com hibridização diferencial foram confirmadas em experimentos de Southern blot com DNA genômico dos dois isolados.

Em seguida, o microarranjo foi avaliado para a análise da expressão diferencial de genes. Para isto, a lâmina foi hibridizada com cDNA de formas epimastigotas (fase mid-log de crescimento) de CL Brener e Silvio. Um total de 8 experimentos de hibridização foi realizado com preparações de RNA obtidas de três cultivos independentes. Identificamos, com uma significância estatística ($P \leq 0,01$), 84 sondas (84/730; 11,5%) diferencialmente expressas: 35 sondas mais expressas em Silvio e 49 sondas mais expressas em CL Brener. Algumas destas sondas foram confirmadas por Northern blot. A comparação dos dados dos experimentos de hibridização do microarranjo com DNA e cDNA das duas cepas permitiu concluir que não há uma correlação direta entre a abundância de determinado gene e seu nível de expressão. Em alguns casos, observamos que seqüências com maior hibridização com o DNA de uma cepa são mais expressas na outra.

Visando iniciar os estudos de expressão diferencial de genes em cepas isoladas de pacientes chagásicos, escolhemos as cepas Famema e Hem 179 (*T. cruzi* II), isoladas, respectivamente, de um paciente com a forma indeterminada da doença de Chagas e de um paciente com alterações cardíacas e digestivas. Observamos que aproximadamente 2,5% (18/730) das sondas apresentaram expressão diferencial nas formas epimastigotas das duas cepas e que 9,0% (65/730) das sondas mostraram razões de hibridização diferente nas formas infectantes (tripomastigotas metacíclicas). Algumas destas sondas foram confirmadas por Northern blot. Entre as sondas mais expressas em Hem 179, está a seqüência que codifica a subunidade 7 da NADH desidrogenase (ND7).

O conjunto de dados permitiu concluir que microarranjos contendo predominantemente ESTs de CL Brener são uma ferramenta útil para estudos de genômica comparativa, análise de expressão gênica em isolados de *T. cruzi*, bem como para a descoberta de novos genes. Além disto, este estudo indicou uma regulação pós-transcricional intensa dos níveis de RNA em *T. cruzi*.

Ampliamos os estudos para um maior número de cepas e construímos outro microarranjo contendo 710 ESTs de CL Brener (forma epimastigota), 45 ESTs da cepa Tulahuen (forma amastigota) e 32 genes de várias cepas do parasita. As 787 seqüências representam 714 seqüências únicas.

O RNA total foi isolado de formas epimastigotas de cepas de três pacientes com a forma cardíaca (cepas 115, B13 e 147), e de três pacientes assintomáticos (cepas VL10, Famema e Berenice 62). Cinco das seis cepas são de regiões endêmicas de Minas Gerais (exceto Famema que é do estado de São Paulo). Após hibridização competitiva, a análise estatística dos dados indicou expressão diferencial de 14 sondas (14/787; 17,7%) entre os dois grupos de cepas. Destas, 9 sondas estão induzidas em todas as cepas isoladas de pacientes cardíacos e 4 sondas estão reprimidas em todas as cepas de cardíacos quando comparadas com as cepas de assintomáticos. Dentre as sondas mais expressas em cepas de cardíacos estão: as ESTs que codificam a subunidade 7 da NADH desidrogenase (ND7), a aspartato aminotransferase, enzima importante no processo de regeneração de metionina, e a triparedoxina peroxidase de *T. cruzi*, envolvida na resistência a peróxidos exógenos. Também foram reveladas 8 ESTs que não apresentam similaridade em bancos de dados. Parte destas sondas foi confirmada por Northern blot com RNA total de todas as cepas. Em alguns casos, identificamos diferenças do tamanho de transcritos entre as cepas. Em resumo, a análise da expressão diferencial de genes em cepas de pacientes com a forma cardíaca e indeterminada da doença de Chagas permitiu identificar alguns genes que podem estar relacionados à patogênese e/ou serem usados como alvo para o prognóstico da evolução da doença.

Posteriormente, realizamos um experimento-piloto para investigar a resposta transcricional de fibroblastos humanos em cultura à infecção com duas cepas de *T. cruzi*: VL10 (isolada de indivíduo assintomático) e 147 (isolada de indivíduo com cardiopatia severa). Após 24 horas de infecção, os RNAs das monocamadas foram extraídos. No mesmo período, RNA foi extraído de fibroblastos humanos não infectados e usado como referência. Experimentos de hibridização competitiva foram realizados em lâminas de microarranjo contendo oligonucleotídeos derivados de 24.000 genes humanos, construídas no Laboratório de Tecnologia Molecular do Instituto Nacional do Câncer do NIH, em Frederick, EUA, chefiado pelo Dr. David Munroe. Identificamos 28 sondas diferencialmente transcritas nas células infectadas com a cepa 147 (paciente cardíaco). Destas, 27 sondas (27/28; 96,4%) estão com expressão reduzida nas células infectadas. Por outro lado, apenas uma sonda, correspondente ao gene da “heme oxigenase (decycling) 1”, apresentou indução nas células infectadas com 147. Na infecção de células com a cepa VL10, encontramos 17 sondas diferencialmente transcritas. Destas, 16 sondas (16/17; 94,1%) estão com expressão aumentada nas células infectadas (dentre elas, a heme oxigenase), enquanto que apenas uma sonda (proteína 5 ligante do fator de crescimento do tipo insulina) mostra situação inversa. Concluímos que as diferenças de resposta da mesma célula hospedeira à infecção pelas duas cepas são marcantes, o que determina um estudo temporal ao longo da infecção para a confirmação das observações acima. Comprovamos por PCR em tempo real o aumento da expressão da heme oxigenase na infecção pelas duas cepas e repressão da expressão da proteína 5 ligante do fator de crescimento do tipo insulina apenas em células infectadas por VL10.

SUMMARY

To investigate differential gene expression in *Trypanosoma cruzi* strains by microarray DNA technology, we have constructed a prototype slide that contains 710 expression sequence tags (ESTs) of CL Brener epimastigotes and 20 genes of various parasite strains. The nucleotide sequence was available for 576 ESTs. Therefore, we have sequenced 134 additional ESTs and the sequences have been deposited in dbEST of GenBank. Clustering of these probes indicated that slide contains 665 unique sequences.

Initially, we have evaluated the microarray for comparative genomic analysis between CL Brener (*T. cruzi* II) and Silvio (*T. cruzi* I) strains. Four independent experiments were performed. We verified that 31 probes (4.3%) showed statistical significant ($P \leq 0.01$ in *t*-test) higher hybridization with CL Brener DNA, whereas 37 probes (5.0%) showed the inverse situation. This suggests differences in the abundance of the genes in the genomes of the strains and/or variation in sequence similarity. Some of these probes were confirmed by Southern blot with genomic DNA of both strains.

Next, we evaluated the microarray for the analysis of differential gene expression. For this purpose, the slide with hybridized with cDNA obtained from epimastigote forms in mid-log growth phase of CL Brener and Silvio strains. A total of eight hybridization experiments were performed with RNA preparations obtained from three independent parasite harvests. We identified 84 probes (84/730; 11.5%) that showed statistically significant changes ($P \leq 0.01$ in *t*-test) in expression: 35 probes up regulated in Silvio and 49 probes up regulated in CL Brener. Some of these probes were confirmed by Northern blot. When we compared the data of the hybridization of the microarray with genomic DNA and cDNA of CL Brener and Silvio, we verified no correlation between the putative abundance of a particular gene and its expression level. In a few cases, we noticed that sequences with higher hybridization with genomic DNA of one strain were more expressed in the other.

In an initial attempt of investigate differential gene expression in human *T. cruzi* isolates, we selected the strains Famema and Hem 179 (both *T. cruzi* II), obtained, respectively, from an asymptomatic individual and from a patient with severe cardiopathy and digestive disorders. Differential expression of 2.5% (18/730) and 9.0% (65/730) of the probes was observed, respectively, in epimastigotes and metacyclic trypomastigotes of the two strains. Some of the probes were confirmed by Northern blot. Among the probes more expressed in Hem 179 is the gene encoding subunit 7 of the NADH dehydrogenase (ND7).

We concluded that microarrays, containing predominantly ESTs of CL Brener, are a powerful tool for comparative genomics and gene expression analyses in *T. cruzi* as well as for discovery of new genes. In addition, this study has provided further evidence for a high level of post-transcriptional regulation of RNA abundance in *T. cruzi*.

We have extended the analysis to a larger number of strains and constructed a new set of microarray slides, which contain 710 ESTs of CL Brener epimastigotes, 45 ESTs Tulahuen strain amastigotes and 32 characterized genes of various strains. These probes represent 714 unique sequences. Total RNA was extracted from mid-log phase epimastigotes of three strains isolated from patients with cardiac disorders (115, B13 and 147) and three strains from asymptomatic patients (VL10, Famema and Berenice 62). Five of these strains are from endemic areas of Minas Gerais (except Famema that is from the state of São Paulo). After competitive hybridization with the microarray, statistical

analysis of the data indicated differential expression of 14 probes (14/787, 17.7%) between the two groups of strains. Among these sequences, 9 probes are up regulated in all the strains from cardiac patients and 4 probes are down regulated in all these strains as compared with all the strains isolated from asymptomatic individuals. Among the probes up regulated in the strains from cardiac patients there are ESTs encoding subunit 7 of the NADH dehydrogenase (ND7); aspartate aminotransferase, key enzyme in the methionine regeneration process; and *T. cruzi* trypanothione peroxidase, an important enzyme for resistance to exogenous peroxides. Eight ESTs with no similarity matches in databases were also disclosed. Some of these probes were confirmed by Northern blot assays with RNA of all the strains. In some cases, we noticed differences of the molecular size of the transcripts among the strains. In summary, the analysis of differential gene expression in strains isolated from human chagasic patients with the cardiac and indeterminate forms allowed the identification of some potential genes that may be related to pathogenesis and/or may be used as targets for prognosis of the evolution of Chagas disease.

Subsequently, we performed a pilot experiment to investigate the transcriptional response of human fibroblasts in culture to the infection by two *T. cruzi* strains: VL10 (isolated from an asymptomatic patient) and 147 (isolated from a patient with severe cardiomyopathy). After 24 hours of infection, total RNA was extracted from the infected monolayers. In the same period, RNA was extracted from not infected human fibroblasts, and used as a control. Competitive hybridization experiments were performed in microarray slides containing oligonucleotides representing 24,000 human genes, constructed in the Laboratory of Molecular Technology of the National Institute of Cancer of NIH, in Frederick, EUA headed by Dr. David Munroe. We identified 28 probes differentially transcribed in cells infected with 147 strain. Of these, 27 probes (27/28; 96.4%) were down regulated in the infected cells, and only one probe, corresponding to the heme oxygenase (decycling) 1 gene, showed up regulation. For the fibroblast infected with VL 10 strain, we found 17 probes differentially transcribed. Of these, 16 probes (16/17; 94.1%) were up regulated (among these probes there is the heme oxygenase sequence), and only one probe (insulin-like growth factor binding protein 5) showed the inverse situation. We concluded that there are drastic differences in the response of the same cell line to the infection by two parasite strains. This determines that a temporal analysis of gene expression during the infection process should be performed to corroborate the above observations. The increase of heme oxygenase transcripts in fibroblasts infected by the two strains was confirmed by real time-PCR. This methodology also confirmed down-regulation of the insulin-like growth factor binding protein 5 mRNA in fibroblasts infected with VL 10 strain.

1. INTRODUÇÃO

1.1 - *Trypanosoma cruzi* e a doença de Chagas

O *Trypanosoma cruzi* é um protozoário flagelado, agente etiológico da doença de Chagas, que afeta cerca de 16 a 18 milhões de pessoas em países da América Latina e coloca em risco de infecção cerca de 100 milhões de pessoas. No Brasil existem 5 milhões de pessoas infectadas com o protozoário (www.who.int/ctd/chagas/burdens.html).

T. cruzi pertence à ordem Kinetoplastida, família Trypanosomatidae, gênero *Trypanosoma*. O nome da ordem Kinetoplastida é decorrente da presença de uma estrutura chamada cinetoplasto que é constituída por uma rede de moléculas de DNA que equivale a 16-30% do DNA do organismo, e está localizada na mitocôndria única do protozoário.

T. cruzi é transmitido por insetos da subfamília Triatominae tais como *Triatoma infestans*, *Panstrongylus megistus* e *Rhodnius prolixus*. Os triatomídeos têm hábitos noturnos e alimentam-se de sangue. Assim, para obter seu alimento, picam o futuro hospedeiro em regiões descobertas do corpo como rosto, pescoço (fato que dá ao inseto o nome popular de “barbeiro”), mãos e pés. Após a alimentação, o inseto defeca próximo ao local da picada e, ao coçar o local irritado, a vítima contamina-se com o parasita. Outras vias de contaminação são: por transfusão sangüínea, sendo esta via responsável pelo maior número de casos de Chagas depois da transmissão vetorial e por transmissão congênita, que depende de lesão placentária. Também podem ocorrer contaminações por via oral pela ingestão do parasita com a alimentação; por transplantes de órgão contaminado e por acidentes laboratoriais ocasionados por manipulação inadequada do parasita.

O ciclo de vida do *T. cruzi* alterna-se entre vertebrados e insetos triatomíneos com estágios de desenvolvimento diferentes nos hospedeiros: epimastigotas e tripomastigotas metacíclicos no inseto vetor e amastigotas e tripomastigotas sangüíneos no hospedeiro mamífero. *T. cruzi* é diplóide e sua replicação é predominantemente assexuada e, portanto, as cepas representam linhagens clonais independentes (Tibayrenc & Ayala, 1988). Nos últimos anos, evidências obtidas por marcadores de DNA tais como genes de RNA ribossômico (rRNA) e de mini-exon (Souto *et al.*, 1996), perfis de isoenzimas (Tibayrenc, 1995) e análises filogenéticas do gene de rRNA 18S (Kawashita *et al.*, 2001) indicaram que

os genótipos clonais de *T. cruzi* podiam ser agrupados em duas grandes linhagens filogenéticas. A comunidade decidiu padronizar a nomenclatura dos dois grupos de cepas que passaram a ser denominados *T. cruzi* I e *T. cruzi* II. (Anonymous, 1999). As características epidemiológicas dos dois grupos foram definidas, concluindo-se que *T. cruzi* I predomina no ambiente silvestre, ao passo que *T. cruzi* II prevalece no ambiente doméstico, tendo sido isolado principalmente de infecções crônicas humanas (Zingales *et al.*, 1998). No ciclo silvestre, *T. cruzi* interage com insetos triatomíneos e com mamíferos como marsupiais, roedores, tatus, carnívoros e primatas, que atuam como reservatórios naturais. No ciclo silvestre, há um balanço ecológico entre o parasita e seus vetores selvagens e animais silvestres. Nesses animais, aparentemente, o parasitismo não é nocivo. No ciclo doméstico, o parasita interage com o homem, mamíferos como gatos e cachorros e vetores domiciliados. Esse ciclo estabelece-se como resultado da invasão do homem no ambiente silvestre (desmatamento) e domiciliação de triatomíneos e de mamíferos que trazem o parasita em contato com o homem. Os reservatórios de *T. cruzi* desempenham um papel fundamental nos ciclos silvestre e doméstico da doença de Chagas, assim como na interação altamente dinâmica entre eles. Mamíferos infectados, como gambás e roedores silvestres, invadem áreas domésticas e peridomiciliares aumentando o risco de infecção de humanos. Entre os reservatórios domésticos, dependendo da região analisada e dos hábitos populacionais, estão pequenos roedores como preás e porquinhos-da-índia; carnívoros, como gatos e cachorros e marsupiais (gambás) (Zingales *et al.*, 1998; Breniere *et al.*, 1998; Solari *et al.*, 2001).

Recentemente, a análise de diversos marcadores moleculares mostrou a presença de genomas híbridos em *T. cruzi* (Machado & Ayala 2001; Sturm *et al.*, 2003; Stolf *et al.*, 2003; Pedroso *et al.*, 2003; Augusto-Pinto *et al.*, 2003). CL Brener, o organismo referência do projeto genoma de *T. cruzi* (Zingales *et al.*, 1997a), é um exemplo de cepa híbrida e algumas evidências indicam ser um mosaico de genes de *T. cruzi* I e *T. cruzi* II (Pedroso *et al.*, 2003; Sturm *et al.*, 2003). A escolha dessa cepa para o projeto genoma é vantajosa já que seqüenciando-se o genoma de CL Brener, em princípio, poderia ser obtida informação sobre o genoma dos dois maiores grupos de *T. cruzi*. Por outro lado, isto acarreta um maior esforço de seqüenciamento.

As cepas de *T. cruzi* são extremamente heterogêneas em relação a várias características como: morfologia, virulência, sensibilidade a drogas, infectividade para mamíferos, tropismo tissular, associação com vetores triatomíneos e indução de manifestações clínicas da doença de Chagas (WHO – Expert Committee, 2002).

A doença de Chagas, descoberta em 1909 por Carlos Chagas, possui uma fase aguda e uma fase crônica. A fase crônica caracteriza-se por diferentes manifestações clínicas. Dependendo da área geográfica, 50-70% dos indivíduos infectados são assintomáticos (forma indeterminada), 10-30% apresentam danos cardíacos, com graves alterações eletrocardiográficas, 8-10% desenvolvem anomalias no trato digestivo (megaesôfago e megacólon) e 10% mostram alterações no sistema nervoso periférico (WHO-Expert Committee, 2002). As manifestações cardíacas e digestivas podem aparecer 20 a 30 anos após a fase aguda (Brener *et al.*, 2000). É importante salientar que, por ano, 2-4% dos indivíduos assintomáticos desenvolvem problemas cardíacos e/ou digestivos. Os fatores que influenciam a variabilidade clínica e epidemiológica da doença de Chagas não estão claros, mas, muito provavelmente, envolvem características genéticas do parasita e do hospedeiro (Andrade *et al.*, 1999 e 2002).

Apesar dos avanços na caracterização genotípica dos isolados, na definição de duas linhagens filogenéticas principais e de sua distribuição epidemiológica, até o momento não foram identificados marcadores capazes de correlacionar determinado genótipo com a manifestação clínica da doença de Chagas. Um grande desafio da comunidade científica tem sido a identificação de marcadores moleculares de isolados de *T. cruzi* que possam estar relacionados com o desenvolvimento das formas cardíacas e digestivas da doença. Esses marcadores poderiam ser usados para o prognóstico das manifestações da doença, permitindo intervenções clínicas.

A tecnologia de microarranjos de DNA representa uma das abordagens potenciais para identificar esses genes, seja por genômica comparativa seja por diferenças da expressão gênica.

1.2- Controle da transmissão

Existem quatro estratégias principais para o controle da doença de Chagas que são: a interrupção da transmissão vetorial; interrupção da transmissão nos bancos de sangue; quimioterapia e vacinas.

A interrupção da transmissão vetorial dá-se principalmente pelo tratamento químico com inseticidas da área domiciliar na qual o parasita se aloja. A transmissão domiciliar é resultante da qualidade precária das moradias que favorece a domiciliação do inseto vetor. Em 1980 foi lançado um projeto de combate à doença de Chagas, denominado "Iniciativa do Cone Sul", do qual participaram Argentina, Bolívia, Brasil, Chile, Paraguai e Uruguai. A iniciativa, que tinha como objetivo interromper a transmissão da doença a partir da eliminação de vetores domiciliados e controle nos bancos de sangue, teve alto nível de sucesso no Uruguai, Chile, Brasil e Argentina, onde muitos estados foram considerados livres da transmissão vetorial por *T. infestans*. No entanto, para que este resultado seja mantido é fundamental que haja uma vigilância epidemiológica periódica para monitorar possíveis reinfestações de moradias pelo vetor. Outras iniciativas de interrupção da transmissão vetorial estão em andamento em Países Andinos e da América Central, com alguns resultados promissores (WHO –Expert Committee, 2002).

A triagem sorológica de doadores de sangue e a adição de substâncias tripanocidas ao sangue a ser transfundido são abordagens importantes para a eliminação da transmissão da doença nos bancos de sangue. Deve ser lembrado que na fase crônica, o diagnóstico da doença de Chagas é feito pela monitoração da resposta imune do hospedeiro infectado, enquanto na fase aguda, o diagnóstico é feito principalmente pela identificação do parasita circulante e pela monitoração de anticorpos da classe IgM. Na fase crônica, o diagnóstico sorológico pode ser seguido de análise do eletrocardiograma, raio-X do tórax e do aparelho digestivo do paciente. Entre os métodos sorológicos mais comuns citam-se a hemaglutinação indireta, o teste imunoenzimático de ELISA e a imunofluorescência indireta. Cada teste apresenta um limiar de especificidade e sensibilidade. A especificidade indica a capacidade do teste de reconhecer os indivíduos não-infectados, enquanto que a sensibilidade indica a capacidade do teste de reconhecer os indivíduos infectados. Muitas vezes o emprego de mais de um teste sorológico, com princípios diferentes, é indicado para obter-se maior confiabilidade dos resultados (Brener *et al.*, 2000). Os métodos de

diagnóstico parasitológico mais utilizados são: na fase aguda, o exame direto microscópico do sangue a fresco, ao passo que na fase crônica utilizam-se métodos indiretos, como a hemocultura, o xenodiagnóstico ou a amplificação de seqüências abundantes por PCR. Entre as seqüências do parasita utilizadas como alvo para a PCR temos a seqüência de DNA satélite de 195pb e os minicírculos do kDNA (Brener *et al.*, 2000, Gomes, 1997).

Quanto à quimioterapia, as duas drogas atualmente utilizadas no tratamento da doença de Chagas são Nifurtimox e Benzonidazol. Nifurtimox gera um aumento na taxa de radicais livres intracelular (ânions superóxido e peróxido de hidrogênio). Estes radicais atacam o parasita que possui um sistema de detoxificação para metabólitos de oxigênio pouco eficiente, o que torna o parasita mais sensível a estresse oxidativo do que células do vertebrado (Docampo, 1990). Nifurtimox também está relacionado com a inibição da enzima tripanotona redutase muito importante para a resposta a estresse oxidativo do parasita (Chung *et al.*, 2003). Benzonidazol age, aparentemente, por um mecanismo diferente (estresse redutivo) que envolve modificações covalentes de macromoléculas por intermediários nitroreductores (Urbina & Docampo, 2003). Ambas as drogas são eficazes na fase aguda da doença com até 80% de cura, mas não o são na fase crônica da doença quando a maior parte dos parasitas está no interior das células. Além disto, essas drogas apresentam muitos efeitos colaterais nos pacientes como perda de peso, tonturas, anorexia, e dores abdominais e sua eficiência varia de acordo com a área geográfica, provavelmente devido a diferenças das populações de *T. cruzi* na susceptibilidade às drogas. Estes dados mostram a necessidade de desenvolvimento de novos quimioterápicos para o combate da doença.

No caso de vacinas para profilaxia da doença de Chagas, apesar de terem sido feitos inúmeros estudos, até o momento não foram identificados antígenos capazes de induzir uma proteção eficiente em animais experimentais (WHO-TDR, 2002).

1.3- *Trypanosoma cruzi* e a interação com a célula hospedeira

A forma infectante de *T. cruzi* encontrada no inseto vetor é denominada tripomastigota metacíclico. Os metacíclicos ao atingirem a corrente sanguínea do hospedeiro mamífero penetram em várias células, dentre as quais macrófagos e células

epiteliais, e transformam-se em amastigotas. Os amastigotas sofrem vários ciclos de divisão binária intracelular e diferenciam-se em tripomastigotas. O ciclo intracelular completo dura em média de 3 a 5 dias. Os tripomastigotas, após ruptura das células, vão para a circulação, podendo ser ingeridos por um inseto triatomíneo, ou invadir novas células do hospedeiro. Desta forma, a habilidade de infectar e replicar-se em uma grande variedade de tipos celulares é uma característica essencial do ciclo de vida de *T. cruzi* em hospedeiros mamíferos. No inseto, as formas tripomastigotas sangüíneas transformam-se em epimastigotas, que se dividem e, na porção terminal do tubo digestivo, diferenciam-se em tripomastigotas metacíclicos.

Até o momento, foram identificados alguns componentes moleculares da superfície do parasita e da célula hospedeira envolvidos no processo de invasão celular (Zingales & Colli, 1985; Schenkman *et al.*, 1991, 1992, 1993, 1994; Burleigh & Andrews, 1995; Burleigh e Woolsey, 2002; Giordano *et al.*, 1994, 1999; Magdesian *et al.*, 2001). No entanto, os fatores que determinam a multiplicação intracelular do parasita ainda são pouco entendidos. Neste processo, a resposta imune da célula à infecção por *T. cruzi* deve ter um papel fundamental (Burleigh & Woolsey, 2002).

Os eventos observados durante a interação de *T. cruzi* com células de mamífero em cultura foram divididos em seis fases: (1) adesão do tripomastigota à superfície da célula hospedeira; (2) interiorização via vacúolo parasitóforo; (3) reorganização de tripomastigotas em amastigotas e escape para o citoplasma; (4) multiplicação de amastigotas; (5) diferenciação de amastigotas em tripomastigotas; (6) ruptura da célula e liberação de tripomastigotas para o meio. Deve ser salientado que nenhuma linhagem celular revelou-se resistente à invasão por tripomastigotas *in vitro* e que epimastigotas não infectam células em cultura, exceto macrófagos, que fagocitam essas formas e as destroem nos fagolisossomos (Dvorak & Hyde, 1973; Zingales & Colli, 1985; de Araujo-Jorge, 1989; Burleigh & Andrews, 1995).

Após adesão do tripomastigota à superfície da célula hospedeira, ocorre sua internalização via formação do pseudópodos ou de uma depressão da superfície celular. Esse processo é lento e pode levar de 5 a 10 minutos. A internalização do parasita coincide com a formação do vacúolo parasitóforo constituído pela membrana plasmática da célula que envolve o parasita. Em seguida, ocorre a fusão dos lisossomos com a membrana

plasmática. Glicoproteínas e glicolipídios localizados na superfície celular do parasita e da célula hospedeira estão envolvidos no processo. Após a interiorização, ocorre a desorganização do vacúolo e os parasitas são liberados no citoplasma, diferenciando-se em amastigotas. A forma amastigota fica em contato direto com o citoplasma da célula e começa um processo de multiplicação em grande escala, em geral 24 horas após infecção. Após aproximadamente 5 a 7 dias, quando a célula hospedeira contém cerca de 500 amastigotas, essas formas começam a diferenciar-se em tripomastigotas. Após ruptura da célula, os tripomastigotas são liberados e, *in vivo*, atingem a corrente circulatória disseminando-se no organismo (Zingales & Colli, 1985; Brener *et al.*, 2000; Vaena de Avalos *et al.*, 2002).

A infecção da célula de mamífero é um evento crítico para a sobrevivência do parasita no hospedeiro. Neste sentido, o parasita parece ter desenvolvido uma série de estratégias adaptativas que envolvem a utilização de diversos ligantes da superfície da célula hospedeira para aderir e penetrar e a modulação de vias de transdução de sinal para garantir sua sobrevivência intracelular. A manipulação ativa de múltiplas vias de transdução de sinal das células hospedeiras é mediada pela interação direta parasita-hospedeiro e por fatores solúveis secretados (Burleigh & Woolsey, 2002). Um dos eventos de sinalização melhor caracterizado e relacionado com a invasão de células hospedeiras é a fusão de lisossomos com a membrana plasmática, processo sinalizado por Ca^{2+} (Rodriguez *et al.*, 1997; revisto em Burleigh & Woolsey, 2002). Embora *T. cruzi* seja capaz de manter o ciclo de infecção em células fagocíticas não-profissionais, muito pouco se sabe sobre as alterações induzidas pelo parasita na célula. Foi relatado que em macrófagos infectados com *T. cruzi*, citocinas pró-inflamatórias são induzidas por mucinas do parasita (Machado *et al.*, 2000) após ativação das vias NF- κ B e MAPK1 da célula hospedeira (Ropert *et al.*, 2001). Coincidentemente, foi verificado que citocinas pró-inflamatórias produzidas precocemente na infecção experimental são moduladores importantes da resposta imune inata e adaptativa do hospedeiro (Ropert & Gazzinelli, 2000). É previsível que a identificação de vias de transdução de sinal induzidas por *T. cruzi* e que determinam a invasão coordenada, manutenção da replicação intracelular e influência no tropismo tissular será uma etapa importante na compreensão da patogênese da doença de Chagas. Nesse

âmbito, a tecnologia de microarranjos de DNA poderá dissecar os eventos moleculares da resposta transcricional temporal da célula hospedeira à invasão pelo *T. cruzi*.

Por outro lado, a influência da diversidade genética do parasita no processo de invasão celular deve ser levada em conta nos estudos desta natureza. De fato, foram relatadas diferenças na infectividade *in vitro* de distintos isolados de *T. cruzi* e na capacidade de diferentes linhagens celulares em suportar a infecção. Dvorak e Howe (1976) reportaram que células BESM (bovine embryo skeletal muscle) interiorizam de 2 a 5 vezes mais tripomastigotas de cepas miotrópicas que células HeLa e Vero, respectivamente. Bertelli e Brener (1980) mostraram que tripomastigotas sanguíneos da cepa Y são mais infectantes que tripomastigotas de cepa CL para células BESM e Vero. Essas observações sugerem que a interação entre parasita e célula hospedeira é dependente de características genéticas particulares de ambas as células.

Na infecção experimental de camundongos, Andrade *et al.* (1999) demonstraram que dois clones de *T. cruzi* com características genéticas próprias apresentam tropismo tissular diferente. Posteriormente, os autores verificaram que a distribuição tissular dos dois clones varia em diferentes linhagens de camundongo (Andrade *et al.*, 2002). Estes resultados demonstram mais uma vez que fatores genéticos do parasita e do hospedeiro interagem no processo de infecção.

1.4- Expressão diferencial de genes de *Trypanosoma cruzi*

A regulação da expressão gênica de tripanossomatídeos ainda é pouco conhecida, apesar de terem sido definidos vários parâmetros que indicam ser este processo bastante complexo e peculiar (revisto em Teixeira, 1998; Teixeira & Darocha 2003; D'Orso *et al.*, 2003). Os dados existentes indicam: (a) ausência de introns em genes codificadores de proteína; (b) ocorrência de transcrição policistrônica; (c) processamento do pré-mRNA em unidades gênicas independentes por um processo de "*trans-splicing*" onde o mini-exon – sequência de 39pb - é adicionado na extremidade 5' de cada mRNA; (d) processamento pós-transcricional que envolve degradação ou estabilização de mRNAs; (e) transcrição de genes codificadores de proteína pela RNA polimerase I (ou enzima com propriedades análogas); (f) processamento pós-transcricional de mRNAs do genoma mitocondrial por um

mecanismo de editoração ("RNA editing") com inserção ou deleção de uridinas nos pré-mRNAs; (g) presença de regiões ricas em pirimidinas que antecedem o dinucleotídeo AG onde ocorre o "trans-splicing" do mini-exon; (h) ausência de seqüências consenso de poliadenilação nos mRNAs, sugerindo-se que os processos de adição de mini-exon e poliadenilação são eventos interdependentes e concomitantes sendo que o primeiro evento precede a poliadenilação; (i) ausência de regiões promotoras em genes que codificam proteína (revisto em Teixeira, 1998; Teixeira & Darocha 2003; D'Orso *et al.*, 2003).

Apesar de a maior parte dos genes codificadores de proteína não possuir introns, Mair e colaboradores (2000) a partir do isolamento e seqüenciamento de genes codificadores da poli A polimerase de *T. brucei* e *T. cruzi* mostraram claramente a presença de introns de 653 e 302 pb, respectivamente. Estes introns obedecem à regra geral de GT/AG para o processamento por "cis-splicing", mostrando que além de "trans-splicing" esses organismos também são capazes de fazer "cis-splicing".

Pouco se conhece a respeito dos promotores de tripanossomatídeos. Os dados existentes sugerem que existem seqüências pouco específicas reconhecidas pela RNA polimerase I (ou enzimas com propriedades similares) (revisto em Teixeira, 1998). Esses dados estão em contraposição com o que se observa em eucariontes, nos quais os genes codificantes de proteínas são transcritos pela RNA polimerase II, enquanto a RNA polimerase I transcreve genes de RNA ribossômico. O único gene identificado com promotor para RNA polimerase II é o gene de mini-exon. Por outro lado, regiões promotoras foram claramente identificadas para o cistron de RNA ribossômico (revisto em Teixeira, 1998).

Para a caracterização de regiões promotoras de genes de tripanossomatídios, plasmídios contendo genes repórter foram utilizados em experimentos de transfecção (estável ou transitória). Apesar de existirem muitos estudos, ainda não foi possível uma identificação precisa de regiões promotoras de genes que codificam proteínas, ao contrário do que ocorre com os cistrons ribossômicos nos quais uma região promotora foi identificada. Desta forma, postula-se que a RNA polimerase teria a capacidade de ligar-se aleatoriamente em pontos do DNA e iniciar a transcrição. O pré-mRNA gerado seria, posteriormente, clivado em pontos corretos a partir dos processos de "*trans-splicing*" e poliadenilação.

Os dados existentes sugerem que o controle da expressão gênica de tripanossomatídeos ocorre no nível pós-transcricional e que esses processos regulatórios são específicos para diferentes estágios de desenvolvimento do parasita, o que mostra a relevância desses sistemas para a adaptação de *T. cruzi* a diferentes condições ambientais. As interações entre elementos *cis* e fatores *trans* têm sido consideradas responsáveis por esse tipo de atividade regulatória (revisto em Teixeira, 1998; Teixeira & Darocha 2003 e D'Orso *et al.*, 2003).

Experimentos realizados em *T. brucei* e *Leishmania* spp. mostram que certas espécies de mRNA apresentam diferenças de até 100 vezes nos seus níveis de expressão em estágios do parasita. Algumas possibilidades foram consideradas para explicar estas observações, incluindo diferenças na eficiência do processo de "*trans-splicing*" e a presença de seqüências regulatórias nas regiões 3' não traduzidas (UTR) dos mRNAs, que afetariam a estabilidade e o "turn-over" dos mensageiros. Utilizando transfecções transitórias, Hehl *et al.* (1994) demonstraram que uma seqüência de 16 nucleotídios, que formaria uma alça no mRNA do antígeno PARP (procyclic acid repetitive protein) do estágio pró-cíclico de *T. brucei*, regularia a tradução desta proteína. Posteriormente, Furger e colaboradores (1997) a partir da análise das 297 pb da região 3'UTR do mRNA de PARP demonstraram a presença de três elementos regulatórios adicionais envolvidos na taxa do "turn-over" deste mensageiro. Os autores mostraram que a deleção de uma dessas regiões gerava uma discreta queda na quantidade do mRNA (2 vezes) mas causava uma redução em até 25 vezes na quantidade da proteína correspondente na célula. Os autores também mostraram que a deleção de outra região da 3'UTR desse mRNA causava a diminuição de até 3 vezes tanto no nível do mRNA quanto da proteína. Esses dados mostraram que as regiões 3'UTR são importantes tanto para a regulação da transcrição gênica como para a regulação da tradução.

A expressão diferencial dos genes de amastina e tuzina de *T. cruzi* foi bastante investigada. Esses genes estão agrupados em um mesmo "cluster" no genoma do parasita, no entanto, os genes de amastina são expressos preferencialmente no estágio amastigota ao passo que tuzina não é regulada no ciclo (Teixeira *et al.*, 1995). Teixeira *et al.* (1996) sugerem que a expressão preferencial da amastina nos estágios intracelulares seja resultado de um processamento mais eficiente deste mRNA, envolvendo uma proteína deste estágio

que se ligaria a uma sequência específica na 3'UTR. Nozaki e Cross (1995) também chegam a conclusões semelhantes a respeito da importância da região 3'UTR para a expressão estágio-específica de outros genes. Por outro lado, as características estruturais encontradas nas regiões 3'UTR dos mRNAs da tuzina não parecem ser uma regra geral para os mRNAs de *T. cruzi*, uma vez que a região 3'UTR de outros genes de expressão estágio-específica (gp72 e gp85) apresenta características diferentes.

Brandão (1999) analisou as características estruturais das regiões 5'UTR e 3'UTR de 92 ESTs geradas no Projeto Genoma de *T. cruzi*. Nesse estudo, o autor concluiu que as regiões não traduzidas seguem o perfil de distribuição de tamanho observado para eucariotos superiores, com uma média de 50 a 100 pb. Também observou que na 5'UTR predominam purinas, ao passo que na 3'UTR há maior abundância de pirimidinas. Em ambas as UTRs são encontradas repetições contínuas de di e trinucleotídeos, com um máximo de repetição de 16 - 17 vezes. A 3'UTR apresenta maior conteúdo de repetições. Nesse estudo não foi observada uma correlação entre o nível de expressão de determinada classe gênica e características das UTRs, tais como, tamanho, conteúdo de GC e frequência de repetições.

Di Noia *et al.* (2000) identificaram pela primeira vez em *T. cruzi* um elemento em *cis* rico em AU (ARE "AU rich element"), localizado na região 3'UTR de um gene da família das mucinas (SMUG-"small mucin gene"). Essa região confere instabilidade ao mRNA de maneira estágio-específica, modula a eficiência de transcrição e é reconhecida por um fator em *trans*. D'Orso e Frasch (2001a) identificaram um elemento rico em G (GRE "G rich element") também presente na região 3'UTR do gene da família das mucinas que não possuía nenhuma identidade estrutural e de sequência com outros elementos conhecidos. Essa sequência confere estabilidade ao mRNA de mucina de maneira estágio-específica, mostrando funcionalidade diferente das regiões ARE (ver acima) (D'Orso & Frasch, 2001a). Posteriormente, os autores detectaram uma proteína denominada TcUBP-1 ("U rich RNA binding protein"), que se liga a regiões ricas em AU e interage com outros fatores de maneira estágio-específica (TcUBP-1 e TcPABP1), e que está envolvida na desestabilização dos mRNAs de mucinas (D'Orso *et al.*, 2003; D'Orso & Frasch 2001b). Esses resultados mostram a importância das regiões não traduzidas e de proteínas que se

ligam ao RNA no mecanismo de estabilização/desestabilização de mRNAs em diferentes estágios evolutivos e para a regulação de expressão gênica em *T. cruzi*.

A expressão diferencial de genes nos estágios de *T. cruzi* tem sido inferida a partir da análise de bibliotecas de cDNA estágio-específicas, produtos de transcrição de núcleos isolados e, mais recentemente, utilizando-se uma técnica denominada RDE ("Representation of Differential Expression"), que envolve a hibridização subtrativa de cDNAs obtidos a partir de duas populações de parasitas (Krieger & Goldenberg, 1998). Estas abordagens permitiram identificar genes expressos preferencialmente na forma amastigota (Teixeira *et al.*, 1994) e na forma tripomastigota metacíclica (Krieger & Goldenberg, 1998). Pela da técnica de "differential display" foi possível identificar um transcrito do gene da RNA helicase da família de DEAD-box que é expresso cerca de oito vezes mais em tripomastigotas metacíclicas de CL Brener em relação a epimastigotas (Diaz-Anel *et al.*, 2000). Utilizando a técnica convencional de Northern blot, também foi possível identificar níveis de expressão diferencial de espécies de mRNA em formas epimastigotas de diferentes cepas (dados de nosso laboratório não publicados). O conjunto de dados indica a existência de processos ativos na regulação da expressão diferencial de genes em *T. cruzi*.

1.5- Projetos Genoma de Parasitas

O projeto genoma de *T. cruzi* foi lançado em 1994 pela Organização Mundial da Saúde, juntamente com os projetos genoma de *Filaria*, *Schistosoma*, *Leishmania* e *Trypanosoma brucei*. Esses projetos têm como objetivo final aumentar os conhecimentos sobre a biologia dos parasitas visando gerar novas estratégias para o seu combate. Neste sentido, buscavam-se genes cujos produtos poderiam ser utilizados como alvos para novos quimioterápicos e vacinas, ou ainda, em métodos mais adequados para seu diagnóstico e/ou prognóstico (<http://www.dbbm.fiocruz.br/TcruziDB/>).

O projeto genoma de *T. cruzi* foi desenvolvido inicialmente por vários laboratórios do Hemisfério Norte e, no Hemisfério Sul, principalmente no Brasil e Argentina. O organismo escolhido para o projeto foi um clone derivado da cepa CL, denominado clone CL Brener em homenagem a Zigman Brener que o isolou em seu laboratório. Este clone apresenta todas as características importantes de *T. cruzi*: derivou de uma cepa isolada de

Triatoma infestans; sofre diferenciação em meio líquido; infecta monocamadas de células; tem parasitismo preferencial para células de músculo e de coração; promove mortalidade em ratos e apresenta fase aguda em infecção humana accidental (Zingales, *et al.*, 1997a, 1997b). O cariótipo molecular de CL Brener foi definido indicando haver 20 bandas cromossômicas com tamanho entre 3,5 e 0,45 Mpb e um tamanho do genoma de 87Mpb (Cano *et al.*, 1995). O clone CL Brener pertence a *T. cruzi* II (Zingales *et al.*, 1997a).

No projeto foram construídas duas bibliotecas de EST (Expressed Sequence Tags, ou Etiquetas de Sequências Transcritas) normalizada e não normalizada, a partir de mRNA da forma epimastigota de CL Brener (Urmenyi *et al.*, 1999). Com a biblioteca normalizada pretendia-se obter o maior numero de ESTs sem que houvesse muita redundância nos clones, já que por essa abordagem é possível eliminar os mRNAs que são expressos em grande quantidade na célula como mRNAs de histonas, de tubulinas e RNA ribossômicos. Foram também construídas bibliotecas genômicas do clone CL Brener em YACS, BACS e cosmídios (ver em Zingales *et al.*, 1997a). Laboratórios integrantes da rede seqüenciaram fragmentos de cerca de 11.000 ESTs, principalmente a partir da extremidade 5' e as seqüências obtidas estão depositadas no banco de dados do NCBI (National Center for Biotechnology).

Um estudo realizado em 1.949 ESTs de CL Brener (Verdun *et al.*, 1998) demonstrou que 14,6% dos clones possuíam similaridade com genes previamente caracterizados de *T. cruzi*, enquanto 18,4% tinham similaridade com genes de outros organismos. Por outro lado, 67% das ESTs não possuíam similaridade com seqüências depositadas em bancos de dados sugerindo serem genes específicos de *T. cruzi* (Verdun *et al.*, 1998).

No início de 2000, foi anunciada a criação de um consórcio entre o NIAID/NIH (National Institute of Allergy and Infectious Diseases do NIH), WHO, Wellcome Trust, Burroughs Wellcome Fund, TIGR (The Institute for Genomic Research), Karolinska Institute e SBRI (Seattle Biomedical Research Institute) para o seqüenciamento do genoma de *T. cruzi* num período de quatro anos. Nesse mesmo ano, um número limitado de seqüências de “shotgun” foi liberado (Aguero *et al.*, 2000). Em 2004, o seqüenciamento genômico completo por “shotgun” de *T. cruzi* foi liberado e está disponível na página <http://Tcruzi.org> (Luchtan *et al.*, 2004). As estratégias de seqüenciamento, resumo dos

dados gerados e outras informações podem ser obtidas também nas páginas:

<http://www.tigr.org/tdb/e2k1/tba/genome.shtml/>;

<http://www.genedb.org/>; <http://www.apps.sbri.org/genome/>; <http://web.cgb.ki.se/tcg/>

Até o momento existem 730 milhões de pares de bases de *T. cruzi* disponíveis e grupo de seqüenciamento pretende publicar o genoma finalizado até o final de 2004.

1.6-Tecnologia de microarranjos de DNA

Até pouco tempo atrás, a análise da expressão gênica só era possível com um número limitado de genes. Com o aparecimento da tecnologia de microarranjos de DNA ("DNA microarrays"), desenvolvida pioneiramente no laboratório do Dr. P.D. Brown na Stanford University, cientistas podem agora examinar, simultaneamente, a expressão de milhares de genes, sendo possível determinar a relação complexa entre genes individuais em diferentes organismos e tecidos e também agrupar os genes pelo seu padrão de expressão (White *et al.*, 1999, Behr *et al.*, 1999, DeRisi *et al.*, 1997, Duncan, 2004).

O princípio da tecnologia de microarranjos de DNA é a hibridização de ácidos nucléicos, cuja elevada especificidade se baseia no pareamento entre as bases destas moléculas. O procedimento envolve imobilizar milhares de seqüências de DNA (cDNAs, DNAs ou oligonucleotídeos) em lâminas de vidro recobertas com um material especial (3-aminopropyltrimethoxysilane por exemplo), usando um sistema robotizado para criar um arranjo com a maior densidade possível.

A análise da expressão diferencial de genes é obtida a partir da hibridização dos DNAs imobilizados com alvos de cDNAs obtidas por transcrição reversa a partir de mRNA total de duas populações do organismo ou tecido em estudo. Cada cDNA é sintetizado com nucleotídeo precursor contendo marcador de fluorescência diferente (Cy3 e Cy5). Após hibridização, a intensidade da fluorescência é quantificada em microscópio adequado ("confocal laser scanning") que reconhece pelo menos dois espectros de fluorescência. Após a normalização da intensidade das sondas, a abundância relativa de cada RNA é calculada. Os dados obtidos a partir desta análise descrevem o amplo espectro da expressão de genes de uma célula, tecido ou organismo.

Uma aplicação dos microarranjos de DNA tem sido o estudo do padrão transcricional de tumores, com o objetivo de entender o processo que acompanha a

transformação maligna e com isso tentar correlacioná-la com os fenótipos clínicos. Também nesse campo pretende-se obter ferramentas para o prognóstico do câncer.

A elucidação das vias metabólicas utilizadas por determinado organismo ou tecido é outro campo de pesquisa que utiliza bastante os microarranjos. Em neurobiologia, esta tecnologia tenta determinar a composição molecular de diferentes regiões do cérebro, associando-a a suas propriedades funcionais. O estudo do "splicing" alternativo, a identificação de pequenos RNAs de função regulatória e a identificação de sítios de ligação RNA-DNA são outros temas abordados com os microarranjos de DNA (vide revisão de Shoemaker & Linsley, 2002).

Um estudo muito interessante realizado por Enard *et al.* (2002) compara a expressão diferencial de genes em cérebros de humanos e chimpanzés, utilizando uma lâmina de microarranjos com 18.000 cDNAs humanos. Apesar de haver 98,7% de similaridade nas seqüências de DNA genômico entre humanos e chimpanzés, é evidente que os dois primatas diferem em vários aspectos morfológicos, comportamentais e cognitivos. No estudo, foi observada uma diferença marcante de expressão gênica no cérebro de humanos e chimpanzés que poderia explicar as diferenças comportamentais. Os autores sugerem que a razão básica da diferença de expressão pode ser: duplicação ou deleção de genes; variação de promotores; variação do nível de fatores de transcrição e variação da composição celular dos tecidos (Enard *et al.*, 2002).

Em relação a estudos de genômica comparativa existe uma extensa bibliografia tanto para procariontes como para eucariontes. Como exemplo temos o trabalho de Behr *et al.* (1999) que utilizando a tecnologia de microarranjo, realizaram uma análise de genômica comparativa em duas espécies próximas de *Mycobacterium*: *M. tuberculosis* e *M. bovis* e de várias cepas de BCG. Nestes estudos, utilizando-se microarranjos contendo fases abertas de leitura ("ORFs") de *M. tuberculosis* hibridizados com DNA genômico dos outros organismos marcados por fluorescência (por "random primer") foi possível comparar a presença e ausência de ORFs nas cepas e assim relacionar ORFs com o grau de virulência dos organismos.

1.7- Microarranjos de DNA no estudo de protozoários patogênicos e sua interação com a célula hospedeira

Microarranjos de DNA contendo 3.548 insertos randômicos foram utilizados para analisar a expressão diferencial de genes no estágio tropofozoito sanguíneo e no estágio de gametócito de *Plasmodium falciparum*. Grandes diferenças foram encontradas e os resultados foram confirmados por Northern blot (Hayward *et al.*, 2000). Posteriormente, Bozdech *et al.* (2003) e Ben Mamoun *et al.* (2001) utilizaram microarranjos de *P. falciparum* mostrando padrões de regulação de expressão coordenada entre genes no estágio tropofozoito sanguíneo e durante o desenvolvimento assexual intraeritrocítico do parasita.

Apesar de os estudos com parasitas do filo Apicomplexa terem demonstrado que a técnica de microarranjos de DNA é uma ferramenta muito útil para o estudo da regulação da expressão gênica e a identificação de novos genes, poucos laboratórios têm utilizado esta tecnologia em tripanossomatídeos (revisado em Duncan, 2004). Nesse contexto, a expressão de genes estágio-específico tem sido observada em formas sanguíneas e procíclicas de *T. brucei*, parasita causador da doença do sono em gado e humanos no Continente Africano (El-Sayed *et al.*, 2000; Diehl *et al.*, 2002), assim como a indução de vários genes entre diferentes estágios de *Leishmania major*, causador de um grande espectro de manifestações clínicas (Almeida *et al.*, 2002 e Saxena *et al.*, 2003).

Um microarranjo foi construído com cerca de 400 seqüências de *T. brucei*, das quais 193 representavam DNA genômico fragmentado aleatoriamente (1 a 3 Kb) e 207 eram clones de cDNA de uma biblioteca da forma sanguínea do parasita. As populações de mRNA foram obtidas de formas procíclicas e sanguíneas. Dos 400 clones analisados, cerca de 100 mostraram padrões reprodutíveis de expressão diferencial (variação em até duas vezes). Por outro lado, quatro clones apresentaram expressão maior que 10 vezes na forma sanguínea em relação à forma procíclica (El-Sayed *et al.*, 2000).

Que seja de nosso conhecimento, apenas um trabalho foi publicado até o momento em que a tecnologia de microarranjos foi utilizada em *T. cruzi*. Este arranjo de DNA foi construído com 4.400 sondas, derivadas majoritariamente de GSS (“genome sequence survey”) de CL Brener (Minning *et al.*, 2003). Com este microarranjo, os autores

analisaram a expressão diferencial de genes durante a transformação de formas tripomastigotas em amastigotas na cepa Brazil. Os autores encontraram 60 genes diferentemente expressos durante a diferenciação.

Apesar de os dados não terem sido publicados em revistas especializadas, resultados apresentados em reuniões científicas indicam que microarranjos de DNA estão sendo usados para o estudo da expressão diferencial de genes durante a metaciclogênese da cepa Dm28c de *T. cruzi* (Nunes *et al.*, 2001).

A expressão diferencial de genes de células HFF (human foreskin fibroblasts) infectadas por *Toxoplasma gondii* foi investigada por Blader *et al.* (2001) utilizando lâminas de microarranjo contendo 22.000 genes e ESTs de cDNAs humanos. Foi verificado que logo após a infecção (1-2 h), menos de 1% dos genes imobilizados na lâmina apresentou alteração de expressão. Dos 63 genes com função conhecida, 27 genes estavam relacionados com a resposta imune. Também concluiu-se que essa resposta não é dependente necessariamente da presença do parasita, já que estes genes também podem ser regulados positivamente pela presença de fatores extracelulares do parasita. Para verificar se esses genes eram modulados independentemente da natureza do parasita, as células HFF foram infectadas com tripomastigotas de *T. cruzi*. Surpreendentemente, observou-se que apenas três genes são modulados após 2 horas de infecção, indicando que a resposta de células HFF varia com a natureza do patógeno. Os genes modulados por ambos os parasitas são *MCPI*, *pré-B-cell-colony-enhancing factor* e *bradykinin receptor B2*. A análise da resposta das células 24 horas pós-infecção com *T. gondii* mostrou a modulação de genes envolvidos com numerosos processos metabólicos da célula hospedeira, dentre os quais, enzimas da via glicolítica e da biossíntese de colesterol. Os genes modulados durante infecção por *T. gondii* foram divididos em três classes: genes necessários para a defesa celular, genes necessários para o crescimento do parasita e genes que indiretamente são regulados como consequência da modulação das duas primeiras classes (Blader *et al.*, 2001).

Dando continuidade aos estudos, Vaena de Avalos *et al.* (2002) infectaram células HFF com *T. cruzi* (cepa Y, isolada de humano na fase aguda da doença de Chagas) e analisaram a expressão diferencial de genes da célula hospedeira ao longo da infecção (2 a 24 h). Nesse estudo, utilizando-se lâminas com 27.000 cDNAs humanos, foi confirmado

que *T. cruzi* não induz expressão de genes nas primeiras horas de infecção (2, 4 ou 6 h). Esse resultado foi considerado surpreendente pelos autores, tendo em vista que dados da literatura obtidos com outras metodologias indicavam a ativação de múltiplas vias de transdução de sinal no início da infecção por *T. cruzi*, incluindo a reorganização do citoesqueleto e a mobilização dos lisossomos da célula hospedeira. Após 24 horas de infecção, foi observada uma alteração significativa na abundância de transcritos de HFF, com indução de cerca de 106 genes. Alguns destes genes foram confirmados por Northern blot (Vaena de Avalos *et al.*, 2002).

1.8 – O Consórcio CAGE

Em 1999 foi criado um Consórcio entre laboratórios do Departamento de Bioquímica do Instituto de Química (IQ-USP) e do Instituto de Matemática e Estatística (IME-USP) num projeto denominado CAGE (Cooperation for Analysis of Gene Expression). Este projeto, apoiado pela FAPESP, envolvia duas grandes atividades: (a) uma operacional, representada pela criação de um laboratório para o preparo dos microarranjos (IQ-USP) e de um grupo de Modelagem Matemática e Computação (IME-USP), visando geração de "softwares" e análise de dados; e (b) uma atividade de pesquisa de bioquímica e biologia molecular, representada por sub-projetos em que esta tecnologia seria utilizada no estudo do Genoma *Xylella*; Genoma Câncer; Ciclo Celular de Mamíferos e Genoma de Microorganismos. Neste último sub-projeto insere-se o Genoma de *T. cruzi*, coordenado pela Profª. Bianca Zingales.

O laboratório (CAGE lab), inaugurado em dezembro de 2000, está instalado no IQ-USP e conta com uma sala limpa com classificação 10.000, com temperatura controlada ($22\pm 2^{\circ}\text{C}$) e umidade (45-53%). Nessa sala encontram-se o robô para confecção dos microarranjos de DNA, a estação de hibridização e o "scanner" para a análise dos resultados. Maiores detalhes podem ser encontrados na página do CAGE lab <http://bioinfo.iq.usp.br/cagelab>. O coordenador do projeto CAGE é o Prof. Dr. Hugo Armelin e a responsável pelo CAGE lab é a Profª. Dra. Aline M. da Silva. Esta infraestrutura foi utilizada para a construção dos microarranjos de *T. cruzi* utilizados no presente trabalho.

2. OBJETIVOS

O projeto contempla quatro objetivos principais:

2.1 - Construção de microarranjos de DNA contendo cerca de 800 ESTs de CL Brener, organismo de referência do projeto genoma de *T. cruzi*, e genes de outras cepas do parasita.

2.2 - Validação dos microarranjos de DNA para o estudo da expressão diferencial de genes e de genômica comparativa em cepas de *T. cruzi*.

2.3 – Análise da expressão diferencial de genes em cepas isoladas de pacientes chagásicos apresentando a forma indeterminada e cardíaca da doença de Chagas.

2.4 – Realização de experimentos-piloto visando analisar a resposta de células humanas em cultura à infecção com duas cepas de *T. cruzi*, utilizando microarranjos de DNA com seqüências de humanos.

3. JUSTIFICATIVA

Tendo em vista que as cepas de *T. cruzi* apresentam elevada diversidade genética, é fundamental verificar se lâminas de microarranjos contendo predominantemente ESTs de CL Brener podem ser usadas em análises de genômica comparativa e para investigar diferenças de expressão gênica em isolados de *T. cruzi*. No caso afirmativo, nos propusemos investigar a expressão diferencial de genes em cepas isoladas de pacientes chagásicos que apresentam a forma indeterminada e a forma cardíaca da doença de Chagas. Esta avaliação teve por objetivo identificar possíveis marcadores moleculares que poderiam ser utilizados para o prognóstico da evolução da Doença, ditando procedimentos terapêuticos.

O experimento-piloto visando a análise da expressão diferencial de genes de células humanas frente à infecção por duas cepas do parasita isoladas, respectivamente, de um indivíduo assintomático e com cardiopatia chagásica, teve como objetivo padronizar as condições para que estudos mais abrangentes pudessem ser planejados em projetos futuros.

A tecnologia de microarranjos de DNA amplia muito a escala de análise e quantificação da expressão diferencial de genes em cepas do parasita e, certamente,

fornece informações relevantes para a compreensão de questões relacionadas com a epidemiologia de *T. cruzi* e da doença de Chagas.

4. MATERIAL E MÉTODOS

4.1. Soluções

- PBS**- Tampão fosfato de sódio 10 mM pH 7,2; NaCl 150 mM
- PS**- Na₂HPO₄ Anidro 0,1 M; NaH₂PO₄ 0,005 M pH 8,0
- SDS**- lauril sulfato de sódio
- SSC 1x**- NaCl 0,15 M; citrato de sódio 0,015 M pH 7,0
- SSPE- 20X**-NaCl 3 M; NaH₂PO₄ H₂O 0,2 M; EDTA 0,02 M pH 7,4
- TAE 1x**- Tris acetato 4 mM; EDTA 1mM
- TBE 1x**- Tris.ácido bórico 90 mM; EDTA 2 mM pH 8,0
- TE**- Tris-HCl 10 mM pH 8,0; EDTA 1 mM pH 8,0.

4.2. Reagentes e enzimas

- Sais e solventes diversos - **Merck**
- RNase, Meio de cultura para diferenciação Grace, DEAE-Celulose (D3764), Dietil pirocarbonato; Formamida deionizada- **Sigma**
- Enzimas de restrição- **New England Biolabs**
- Kit para clonagem (SureClone™ Ligation Kit)- **Pharmacia Biotech**
- Matriz de nylon Hybond-N e Kit para extração de DNA de gel (GPX), [α -P³²]dATP (3.000 Ci/mmol)- **Amersham**
- Kit de sequenciamento de DNA (BigDye™ Terminator Cycle Sequencing Ready Reaction), e dNTPs - **Applied Biosystems (Perking Elmer)**
- Inibidor de RNase (RNasin) (Recombinant Rnasin R, Inhibitor 10,000 U (400 U/ μ L)); Ribonuclease H (RNase H); RQ1 RNase-Free DNase 1,000 U (concentração: 1U/ μ L)- **Promega**
- Fragmento Klenow da DNA polimerase e DNase- **Boehringer Mannheim**

-Trizol reagent, Random Primer DNA Labeling System- **Life Technologies**

-*Taq* DNA Polimerase- **Gibco**

-Cy3-dCTP; Cy5-dCTP; dNTP set (Ultrapure); “Microarray hybridization buffer”; “Nonamer primers”; “Random primers” (hexamer); Nucleotide mix; Oligo dT primers (Anchored dT25); “Microarray crosslinking reagent”; Vistra green nucleic acid gel-

Amersham Pharmacia Biotech

-Formamida super pure- **Fisher Scientific:**

-Guanidine Hydrochloride; Superscript TM II RNase H (10.000unid/epp); Bioprime DNA labeling System (Klenow fragment 40unidades/ul)- **Invitrogen:**

- SDS 20% solution ultrapure; SSC 20X solution pH 7,0; Água RNase free- **USB:**

4.3. Oligonucleotídeos

Iniciadores utilizados para amplificação e sequenciamento de ESTs: T7 (5'-GTA ATA CGA CTC ATA GGG C-3'), T3 (5'- AAC CCT CAC TAA AGG G-3'), SP6 (5'-TAT TTA GGT GCA ACT ATA G-3'), M13 Forward (5'-CCAGTCACGACGTTGTAAAACG -3'), M13Reverse (5'-AGCGGATAACAATTTCACACAGG-3').

4.4. Isolados de *T. cruzi* e seu cultivo

Algumas características dos isolados de *T. cruzi* usados neste estudo são resumidas na Tabela 1. Formas epimastigotas foram cultivadas em meio LIT (Liver Infusion Tryptose) contendo 10% de soro fetal bovino, a 28° C (Castellani *et al.*, 1967). Para estimar o número de células, a cultura foi diluída com LIT e os parasitas foram contados em câmara de Neubauer.

Tabela 1 – Características de isolados de *T. cruzi*

Isolado	Estado/País	Hospedeiro/manifestação	Grupos de <i>T. cruzi</i> ¹	Híbridos Putativos ²
CL Brener	Rio Grande do Sul/ Brasil	<i>Triatoma infestans</i>	II	+
SílvioX10 cl1	Pará/ Brasil	Humano/fase aguda	I	
Esmeraldo cl3	Bahia/ Brasil	Humano/fase aguda	II	
VL10	Minas Gerais/ Brasil	Humano/fase crônica/forma indeterminada	II	
Berenice 62	Minas Gerais/ Brasil	Humano/fase crônica /forma indeterminada	II	
Famema	São Paulo/ Brasil	Humano/fase crônica /forma indeterminada	II	
B13	Minas Gerais/ Brasil	Humano/fase crônica /forma cardíaca	NA	+
147	Minas Gerais/ Brasil	Humano/fase crônica /forma cardíaca	NA	+
115	Minas Gerais/ Brasil	Humano/fase crônica /forma cardíaca	NA	+
Hem 179	Minas Gerais/ Brasil	Humano/fase crônica /forma cardíaca e digestiva	II	
Y	São Paulo/ Brasil	Humano/fase aguda	II	
YuYu	Bahia/ Brasil	<i>Triatoma infestans</i>	I	
NR cl3	Salvador/ Chile	Humano/fase crônica/forma indeterminada	NA	+
SO3 cl5	Potosi/ Bolívia	<i>Triatoma infestans</i>	NA	+
SC43 cl1	Santa Cruz/ Bolívia	<i>Triatoma infestans</i>	NA	+
Dm 28c	---/ Venezuela	<i>Didelphis marsupialis</i>	I	

NA = não se aplica

¹Classificados de acordo com Anonymous (1999)²Classificado de acordo com Souto *et al.* (1996), Machado & Ayala (2001); Sturm *et al.* (2003); Stolf *et al.* (2003); Pedroso *et al.* (2003).

4.5. Cultivo e purificação de formas tripomastigotas metacíclicas

A forma tripomastigota metacíclica foi obtida a partir da diferenciação de 2×10^7 epimastigotas/ml em meio LIT-10% de soro fetal bovino e meio de inseto Grace 20% (v/v) a 28 °C por 7–10 dias (Zingales *et al.* 1997 a). As formas metacíclicas foram purificadas das formas epimastigotas contaminantes por cromatografia em coluna de DEAE-celulose (DE-52, Whatman) de acordo com De Souza, (1983).

4.5.1. Preparação de DEAE-celulose

Para a preparação de DEAE-celulose foram adicionados 5g de resina (D3764, Microgranular Anion Exchanger - Sigma) a um béquer com 75ml de HCl 0,5N. Essa solução foi agitada com um imã por 30 minutos. Após esse tempo, a resina foi filtrada em Büchner com papel de filtro e lavada com água destilada deionizada até pH 5,0-7,0 no filtrado. A resina foi transferida para béquer contendo 75ml de NaOH 0,5N. A solução foi agitada com imã por 30 minutos e filtrada em Büchner com papel de filtro. A resina foi lavada com água destilada até pH 7,0 no filtrado e transferida para um béquer contendo 75ml de água. O pH foi abaixado para 4,0 com solução H₃PO₄ 5% (p/v). Em quitassato pequeno, a solução foi deaerada com trompa de água por 10 minutos e após esse procedimento foi adicionado NaOH 0,5N para ajustar o pH para 8,0. A resina foi armazenada a 4°C com 25% (v/v) de etanol.

4.5.2. Purificação de formas tripomastigotas metacíclicas

Para cada $1,5 \times 10^9$ células (epimastigotas e tripomastigotas metacíclicos) foram utilizados 3ml de DEAE-celulose. A resina da solução estoque foi centrifugada em tubo Corning de 15ml. O sobrenadante foi descartado e a resina ressuspensa em 5ml de PBSG (Na₂HPO₄ Anidro 0,06M; NaH₂PO₄ 0,003M; glicose 0,3M, pH 8,0). Após centrifugação, a resina foi novamente ressuspensa em 10ml de PBSG. A solução foi deixada agitando por 30 minutos. Após esse tempo, a resina foi empacotada em uma seringa de 5 ml com uma camada de lã vidro na ponta. Os parasitas presentes no meio de cultura foram centrifugados (7,000 rpm por 7 minutos) e lavados com 10ml de PBSG. Após nova centrifugação, as células foram ressuspensas em 0,5ml de PBSG e adicionadas no topo da resina. Nestas condições, as formas epimastigotas se ligam à DEAE-celulose e os tripomastigotas metacíclicos passam livremente. A resina foi eluída com 8ml de PBSG. Aliquotas de 1ml foram coletadas e a presença de parasitas foi monitorada ao microscópio. Em geral, os tripomastigotas são eluídos com 3 a 4ml de PBSG. A eluição da coluna é interrompida

quando começam a ser eluídas as formas epimastigotas. Os tripomastigotas são concentrados por centrifugação e utilizados para extração de RNA ou armazenados a -70°C.

4.6. Bactérias e células competentes

4.6.1. Bactérias

E. coli -DH5 α : hsdR17(rk⁻mk⁺)supE44, thi-1, recA1, gyrA (NaI^r), relA1, Δ (lacZYA-argF)u169(m80lacZ Δ M15) (Ausubel *et al.*, 1995). O estoque de bactérias foi mantido em meio de cultura líquido contendo 50% (v/v) de glicerol a -70°C (indefinidamente), a -20°C (6 meses), ou por curto período em placas de meio de cultura a 4°C.

4.6.2. Meios para cultivo de bactérias.

LB: triptona 10g/L; extrato de levedo 5g/L; NaCl 5g/L; glicose 1g/L.

Terrific Broth: Triptona 12g/L; extrato de levedo 24g/L; glicerol 4g/L, KH₂PO₄ 0,017M; K₂HPO₄ 0,072 M (Birren *et al.*, 1997).

SOB: Bacto-triptona 20g/L; extrato de levedura 5g/L; NaCl 0,5g/L; pH 7,4.

SOC: SOB; glicose 20mM; MgSO₄ 10mM.

Para meio sólido adicionou-se 2% (p/v) de agar. Dependendo da seleção desejada, no meio foi adicionado ampicilina (100 μ g/ml concentração final). Os meios foram esterilizados por autoclavagem por 20 minutos a 1 atmosfera de pressão (Sambrook *et al.*, 1989).

4.6.3. Células competentes

Uma colônia de *E. coli* DH5 α de uma placa de LB foi inoculada em 10mL de SOB e incubada a 37°C sob agitação moderada (250 rpm) durante a noite. Em 50mL de meio SOB contendo MgSO₄ (10mM) foi diluído 0,5mL da cultura de crescimento. A cultura foi inoculada sob agitação a 37°C por 2-4 horas, até atingir densidade de aproximadamente 5x10⁷ células/mL (DO₆₀₀ entre 0,4 a 0,5). As células foram centrifugadas a 3000 rpm por 15 min a 4°C e o sobrenadante descartado. As células foram ressuspensas em 1/3 do

volume inicial da cultura em solução TFB (KCl 100mM, MnCl₂.4H₂O 45mM, CaCl₂.2H₂O 10mM, HAcOCl₃ 3mM, K-MES 10mM pH 6.3) e transferidas para um único tubo de centrífuga. O tubo foi incubado em gelo por 10-15 minutos e em seguida centrifugado a 3000 rpm por 15 minutos a 4°C. O sobrenadante foi descartado e as células ressuspensas em 1/3 do volume inicial da cultura, em solução RF1 (RbCl 100mM, MnCl₂.4H₂O 50mM, CaCl₂.2H₂O 10mM, acetato de potássio 3 mM, pH 7,5, Glicerol 15% (v/v)). Em seguida, as células foram centrifugadas a 3000 rpm por 15 minutos a 4°C e o sobrenadante descartado. Por fim, as células foram ressuspensas em 1/12,5 do volume inicial da cultura, em solução RF2 (MOPS 0,5M pH 6,8, RbCl 10mM, CaCl₂.H₂O 75mM, Glicerol 15% (v/v)). Os estoques foram conservados em tubos eppendorf (200µl/tubo) a -80°C.

4.6.4. Transformação

Para transformação, foi adicionado 20ng de DNA a cada eppendorf contendo células competentes. Esses eppendorfs foram incubados em gelo por 30-40 minutos. Os eppendorfs foram incubados a 42°C por 90 segundos e transferidos imediatamente para gelo. Em cada tubo foi adicionado 400µL de SOC recém preparado e a mistura foi incubada por 40 minutos a 37°C. Uma alíquota das células foi espalhada em placa de LB contendo ampicilina (100µg/mL final) e 2% (p/v) de X-gal (5 Bromo- 4 Cloro-3 Indolil- β- D- Galactopiranosido 0,05M em dimetil formamida). As colônias brancas foram selecionadas e inoculadas para purificação de plasmídeo.

4.7. Preparo de DNA plamidial

A extração de DNA foi feita utilizado-se mini preparação manual baseada no método de lise alcalina (Birnboim & Doly, 1979) com adaptações realizadas pelo Dr. Marcelo Bento Soares - UIowa - Iowa City - USA. Por esse método as bactérias foram inoculadas por 3 horas a 37°C em 3ml de LA (LB mais Ampicilina) sob agitação durante a noite e centrifugadas a 4000rpm. As células foram ressuspensas em 400µl de GET/RNase (Glicose 0,5M; Tris 0,025M, EDTA 0,01M pH 8, adicionar 1,5ml de RNase (10mg/ml) para cada 100ml de solução) e em seguida foram adicionados 400µl de solução

de lise (SDS 0,035M e NaOH 0,2M). Após 5 minutos a solução de neutralização (acetato de potássio 2,65M, pH 4,8) foi adicionada e a preparação foi deixada por 30 minutos no gelo. Os componentes celulares foram sedimentados por 15 minutos a 13,000rpm. Uma parcela do sobrenadante (800µl) foi retirada e adicionada a 600µl de isopropanol. A preparação foi centrifugada por 10 minutos na mesma velocidade anterior e o DNA lavado com etanol duas vezes. O DNA foi dissolvido em água ou TE. Os clones recombinantes foram confirmados após digestões dos DNAs plasmidiais com uma ou duas enzimas de restrição.

4.7.1. Digestão com enzimas de restrição

As condições de reação, a quantidade de enzima, temperatura e tempo seguiram as recomendações do fabricante. As reações foram interrompidas por adição de 2µl de Mistura de Parada (0,25% (p/v) de azul de Bromofenol, 20% (v/v) glicerol ultra puro, 1/10 (v/v) 1x TBE).

4.7.2. Eletroforese convencional

Os fragmentos de DNA foram separados por eletroforese em gel de agarose 1,0% ou 1,2%. As soluções utilizadas foram TBE 1x ou TAE 1x e as voltagens utilizadas foram de 100V a 150V por cerca de 30 a 50 minutos, dependendo do tamanho do fragmento a ser analisado. Os géis foram corados com 0,4µg/ml de brometo de etídio para visualização das bandas de DNA por iluminação com luz ultravioleta (~300nm) e fotografados com filme Polaroid 667 através de filtro laranja (Wratten 21- Kodak). Como marcador de tamanho molecular foi usado DNA de fago λcI857 digerido com *EcoRI* e *HindIII*. Também foram utilizados padrões de massa de DNA da GIBCO.

4.7.3. Recuperação e purificação de fragmentos de DNA e ligação ao vetor

A recuperação de fragmentos de DNA de géis de agarose foi feita por membranas DE-81 (Wathman) ou o Kit GPX da Amersham. As ligações foram feitas utilizando-se a

enzima ligase do "kit" da SureClone (Pharmacia): 50-200ng de DNA tratado na coluna (GPX), 50ng de vetor pBluescript, tampão de ligação e 1µl de DNA ligase (T4-7000U/ml), em uma reação com volume final de 20µl. A ligação foi efetuada por 1 a 3 horas a 16°C. Após a ligação, os plamídeos foram transformados em bactéria pelo método do cálcio e selecionados através da expressão de lacZ em placas de meio contendo Xgal.

4.8. Obtenção de DNA para sequenciamento

4.8.1. Minipreparação por lise alcalina

O método de lise alcalina está descrito no item 4.7.

4.8.2. DNA obtido por PCR de colônia

Um volume de 100µl de meio Terrific Broth (TB) contendo ampicilina na concentração final de 100 µg/ml, foi distribuído nos 96 poços de microplacas de cultura. As colônias brancas contendo plasmídeos recombinantes foram picadas isoladamente com palitos estéreis e inoculadas nos poços com meio. A placa foi incubada em estufa a 37°C por aproximadamente 14 horas. Em seguida, foi adicionado 100µL de glicerol 30% estéril por poço. Foi preparado um "MIX" para 100 reações de PCR em gelo (Primers T7 e T3 20 pmol/µL, dNTP 25mM, MgCl₂ 50mM, 10x tampão *Taq* DNA polimerase 5U/µL, solução final para 20µl). O "MIX" foi distribuído em microplacas de PCR de 96 poços onde cada poço recebeu 15µL do mix. O replicador de 96 agulhas (Boekel) foi flambado com álcool e colocado na microplaca de cultura. Após misturar bem, o replicador foi transferido para a placa de PCR. Após inoculação do DNA, a microplaca de PCR foi centrifugada por 1 pulso até 700g e transferida para termociclador (PTC-200 MJ). As condições do PCR foram: desnaturação 95°C/4min; 40 ciclos de: 95°C/45seg, 55°C/45seg, 72°C/1min. A reação foi terminada por incubação a 72°C/5min e 4°C/infinity. Após a PCR, 3µl da reação foram aplicados em gel de agarose 1,2% e os DNAs foram submetidos a eletroforese.

4.8.3. PCR direto de DNA diluído

Para fazer PCR diretamente dos DNA das ESTs, sem usar o PCR de colônia, pois a maior parte das ESTs já haviam sido purificadas por lise alcalina, foi preparado um “MIX” igual ao do PCR de colônia (item 4.8.2) e nele foi adicionado 1µl do DNA (1- 5ng) das ESTs estoques diluído 1:50. As temperaturas usadas para a PCR foram: desnaturação 94°C /4 minutos; 40 ciclos de três temperaturas: 94°C/1 min, 55°C/1 min, 72°C/1 min. A reação foi terminada por incubação a 72 °C/7 minutos e 4 °C infinito. Após a PCR, 3µl da reação foram aplicados em gel de agarose 1,2% e os DNAs foram submetidos a eletroforese.

4.9. Sequenciamento (Tubos individuais e Placas de 96 poços)

O sequenciamento de DNA foi feito pelo método de terminação de cadeia (Sanger *et al.*, 1977) modificado para uso de polimerase termoestável e com reações de amplificação terminadas com dNTPs marcados com substâncias fluorescentes.

O sequenciamento de DNA foi realizado com kits da PERKIN-ELMER. O protocolo original do kit utiliza 8µl de mistura de reação (dNTPs, ddNTPs fluorescentes, AmpliTaq polimerase, MgCl₂, tampão Tris-HCl, pH 9.0), 500ng de DNA, 3,2 pmoles de iniciador, completando-se o volume para 20µl. Para as ESTs foi utilizado um protocolo com volume de mistura de reação de 2µl. Também foi adicionado à reação 6µl de tampão de diluição (“save money” - Tris-HCl 200nM pH 9,0 e MgCl₂ 5mM) e 30 pmoles de iniciador (iniciadores T3 ou T7) para volume final de reação de 20µl. A reação de amplificação linear (“cycle sequencing”) ocorreu em termociclador com posterior precipitação do DNA com isopropanol. Condições de amplificação para placas: 96°C/10segundos, e 40 ciclos de 52°C/20seg, 60°C/4min. Condições para tubos individuais no termociclador Minicycler da MJ: 96°C/30 segundos, e 34 ciclos de: 96°C/10 seg., 50°C/5 seg, 60°C/4min.

Para a purificação do DNA, 20 µL de água e 60 µl de isopropanol 100% foram adicionados por poço na microplaca. A placa foi incubada por 15 minutos à temperatura ambiente e centrifugada a 3000 g por 30 minutos. O sobrenadante foi descartado e 150µL de isopropanol 70% foi adicionado por poço. A microplaca foi centrifugada a 3000 g por 10

minutos e o sobrenadante descartado. A microplaca foi deixada por 10 minutos em estufa a 37°C para evaporar o isopropanol residual. Após a precipitação, as amostras foram ressuspensas em “loading buffer” (para corrida em seqüenciador automático ABI Prism 377 DNA Sequencer, 96 poços, Perkin Elmer) e todo o conteúdo dos poços foi adicionado no gel de poliacrilamida (LONG RANGER GEL-FMC). Normalmente a eletroforese foi feita por 7 horas. Para seqüenciador de capilar ABI3700, as amostras foram ressuspensas em formamida.

4.10. Anotação - Busca de homologia em bancos de dados, limpeza de seqüências e classificação por função

O formato FASTA (Pearson & Lipman, 1988) das seqüências geradas foi utilizado para busca de similaridade através do programa BLAST (Basic Local Alignment Search Tool) do NCBI (Altschul *et al.*, 1998). O BLAST pode ser acessado através do site <http://www.ncbi.nlm.nih.gov/BLAST/BLAST>. Esse programa é utilizado para a comparação de uma seqüência nucleotídica contra um banco de dados de seqüências nucleotídicas (BLASTN). Para se comparar seqüências protéicas, utiliza-se o BLASTP. Finalmente, se o objetivo for a comparação de uma seqüência nucleotídica contra um banco de seqüências protéicas, usa-se o BLASTX. Nesse caso, a seqüência nucleotídica é traduzida nas seis fases de leitura possíveis. Cada uma das seqüências aminoácidos geradas é então comparada às seqüências de proteínas do banco de dados. Outros tipos de BLAST podem ser encontrados na página do NCBI citada acima.

O programa executa uma busca de homologia em vários bancos de dados. Entre eles estão: sp (Swiss Protein database), Pir (Protein Information Resource), PDB (Protein DataBase), GenBank, *E.coli* (*E. coli* database), EMBL (European Molecular Biology Laboratory), DBJ (Data Base of Japan), GI (GenPept). Também é possível fazer uma busca através de vários bancos de dados por “nr” (“non redundant protein or nucleotide database”) e DBEST que é um banco de dados de ESTs (Baxevanis & Oullette, 1998). Esses bancos de dados foram utilizados no trabalho. A resposta da similaridade é dada após segundos, quando aparece uma janela onde são listadas todas as proteínas ou as seqüências nucleotídicas com as quais a seqüência tem similaridade. Nessa janela também é possível

ver o valor de E (expected) que seria a chance desse alinhamento ter ocorrido por acaso (Altschul & Koonin, 1998). Um $E=0$ é o valor máximo de confiança na homologia. Sendo assim quanto menor o valor de E maior a chance desta homologia ser confiável. Para as seqüências geradas em nosso laboratório foram utilizados valores confiáveis de $E \leq 10^{-5}$.

O próprio BLAST disponibiliza um software (BLASTCl3) que permite buscar uma grande quantidade de seqüências através do computador do laboratório, mas, utilizando os computadores do NCBI para fazer os alinhamentos. Para a “limpeza” das seqüências (plasmídeos e primers) foi criado um banco de seqüências de vetor e primers e foi feito um BLAST local contra ele. A limpeza manual foi realizada a partir desse resultado. A classificação da função das seqüências foi feita de acordo com Verdun *et al.* (1998).

4.11. Agrupamento de seqüências pelo programa Cap3

Para realizar a análise de agrupamentos, as seqüências das ESTs e de genes de *T. cruzi* foram gravadas em arquivos. O agrupamento (“clustering”) das ESTs foi executado com o programa Cap3 (Huang & Madan 1999). O Cap3 inicialmente mostra todos os “contigs” formados com os nomes de seus representantes (quando uma ESTs não tem similaridade com nenhuma outra EST, o programa coloca a EST em um “contig” que contém uma única EST). Quando o programa termina de mostrar todos os “contigs”, aparecem todos os alinhamentos feitos para achar esses agrupamentos. A seqüência consenso é mostrada abaixo do alinhamento. Foi considerado o valor de 80%, como valor de similaridade mínimo para aceitar a localização das seqüências no mesmo agrupamento.

4.12. Amplificação, análise e purificação de ESTs

4.12.1. Amplificação das ESTs (em Placas de 96 poços)

Na reação de amplificação foram utilizados de 1 a 5 ng de DNA (para isso o DNA das ESTs estoques foi diluído em 1:50 e somente 5 μ L foram utilizados para amplificar). As amplificações foram feitas na presença de 0,5 μ M de cada oligonucleotídeo iniciador, 0,2mM de cada desoxinucleotídeo, tampão de DNA polimerase 10X (pH 8,8; (NH₄)₂SO₄ 160mM; Tris-HCl 670mM, Tween 20 0,1%), MgCl₂ 1,5mM, 1U de DNA polimerase

(BIOLASE) em um volume final de 100µL, em placas de 96 poços. O termociclador utilizado foi PTC200 da MJ de 96 poços. As temperaturas utilizadas foram: desnaturação a 94°C por 4 minutos mais 40 ciclos de três temperaturas: 1 min a 94°C para desnaturação, 1 min a 55°C para associação e 1 min a 72°C para extensão, terminando-se com incubação a 72°C por 7 minutos. (<http://info.med.yale.edu/genetics/ward/tavi/PCR.html>).

4.12.2. Análise dos produtos de amplificação

4.12.2.1. Eletroforese em gel de agarose (96 poços)

Inicialmente foi adicionado aos DNAs da reação de amplificação tampão de amostra 6x (Glicerol 50%; Tampão Fosfato de Sódio 0,1M pH 7,0; azul de bromofenol 0,25%; Xilenocianol 0,25%). As amostras foram aplicadas em gel de agarose 1,2% com 0,5µg/mL de brometo de etídio em TBE 1x. Foram adicionados 4µL de reação de PCR por poço. Após purificação, com Multiscreen (ver item 4.12.3) foram adicionados 2µL de reação de PCR no gel. Como marcador de massa de DNA foi utilizado o marcador de DNA (“Mass ladder”) da Pharmacia. O gel foi submetido à eletroforese com voltagem de 100V por 30 minutos. O gel foi fotografado em transiluminador com luz UV.

4.12.3. Purificação dos produtos de PCR (placas de 96 poços)

Para a purificação dos produtos de PCR foram utilizadas as placas da Millipore (Multiscreen para placas de 96 poços). Inicialmente foram adicionados 10µL de tampão de ligação do PCR (“PCR Binding Buffer”-Guanidina HCl 7M; Tampão MÊS 200mM pH 5,6) em cada poço da placa Millipore que foi centrifugada a 2800 rpm por 1 minuto. Em seguida, foram adicionados 100µL tampão de ligação da PCR em cada poço da placa contendo os produtos amplificados. Após misturar bem, todo líquido foi transferido para a placa Millipore que novamente foi centrifugada a 2800 rpm por 1 minuto, a 25°C. O líquido foi descartado e a placa foi lavada duas vezes: uma vez com 200µL de etanol 80% com posterior centrifugação a 2800 rpm por 1 minuto e uma segunda vez com etanol 80% com centrifugação a 2800 rpm por 12 minutos). O DNA foi eluído com 50µL de TE em placas de

96 poços após centrifugação a 3000 rpm por 5 minutos. A análise da purificação foi feita em gel de agarose como descrito acima.

*4.13. Construção das lâminas de microarranjos de DNA de *T. cruzi**

No decorrer deste trabalho, contruímos dois conjuntos de lâminas de microarranjos. O **conjunto A** foi utilizado na validação desta tecnologia para o estudo da expressão diferencial de genes e de genômica comparativa em cepas de *T. cruzi* (ver item 5.1 de Resultados). O **conjunto B** foi empregado na análise da expressão diferencial de genes em cepas isoladas de indivíduos assintomáticos e com cardiopatia chagásica (ver item 5.2 de Resultados).

4.13.1. Sondas de DNA

O **conjunto A** de lâminas de microarranjos foi construído com 710 ESTs (tamanho médio de 800pb) obtidas de bibliotecas normalizadas e não normalizadas de cDNA da forma epimastigota de CL Brener (Urmenyi *et al.*, 1999). As ESTs são identificadas pelas abreviaturas TEUF e TEUQ (obtidas de uma biblioteca não normalizada) e TENF e TENQ (obtidas de uma biblioteca normalizada). As seqüências das TEUF e TENF foram determinadas na FIOCRUZ (Rio de Janeiro) e seqüências das TEUQ e TENQ no Instituto de Química (USP). A descrição e caracterização destas seqüências podem ser encontradas na página <http://www.vision.ime.usp.br/~rvencio/Tcruzi/>. Neste conjunto de lâminas foram também aplicados insertos representando 20 seqüências de *T. cruzi*, isoladas de diferentes cepas e representantes de regiões codificadoras de proteínas ou RNA, assim como de regiões não codificadoras (Tabela 2). Nove DNAs fragmentados também foram imobilizados na lâmina para servir como controle (Tabela 2).

Tabela 2 – Características das seqüências de *T. cruzi* e DNAs controles imobilizados na lâmina de microarranjos

Identificação	Características	Isolado – Grupo de <i>T. cruzi</i>	Número de acesso
G3	DNA satélite de 195-bp	Y – Tc II	K01771
S-56	DNA satélite de 195-bp	Silvio – Tc I	AY188964
18SG	18S rDNA	G – Tc I	AF239981
18SY	18S rDNA	Y – Tc II	AF301912
24S α	24S α rDNA	Y – Tc II	M28885
B11	Retrotransposon	Y – Tc II	U15615
B12	Antígeno de 230 kDa	Y – Tc II	L07759
B13	Antígeno de 140/116 kDa	Y – Tc II	U15616
C6	Elemento de DNA repetitivo	G – Tc I	U16295
CL65	DNA satélite de 195-bp	Silvio – Tc I	AY327543
GP82	Glicoproteína de superfície estágio específica gp82	CL – Tc II	AF128843
GP85	pseudogene de glicoproteína gp85/sialidase	Y – Tc II	M91469
IF8	Proteína flagelar que se liga a Cálcio	Y – Tc II	U70035
Mini-exon	gene de mini-exon e região intergênica	Silvio – Tc I	X62674
ACT	actina 2	Dm28c – Tc I	
EF	Fator de elongação subunidade α	Y – Tc II	L76077
RNA helicase	RNA helicase da família DEAD-Box	CL Brener – Tc II	AF067953
SLA	Pequenos RNAs nucleares	Tulahuen – Tc II	AF016400
Tb	Seqüência telomérica	<i>T. brucei</i>	K01633
<i>T. cruzi</i>	Seqüência sub-telomérica de <i>T. cruzi</i>	CL Brener – Tc II	AF100651
DNAs controle			
Camundongo			NA
<i>Rhynchosciara americana</i>			NA
<i>T. cruzi</i> – CL Brener			NA
Esperma de Salmão			NA
Iniciadores T7 e T3			NA
M13-Iniciadores forward e reverse			NA
pGEX			U78872
pUC18			U07649
pBluescript			X52329

NA – Não se aplica

O **conjunto B** de lâminas de microarranjos continha as mesmas seqüências do **conjunto A** (710 ESTs, 20 seqüências caracterizadas e 9 DNAs controle) e 45 ESTs da forma amastigota da cepa Tulahuen (DaRocha *et al.*, 2000). Também foram imobilizados insertos de 7 seqüências codificadoras de proteína de diversas cepas de *T. cruzi* e 5 ESTs (Tabela 3). A descrição e caracterização destas seqüências podem ser encontradas no CD anexo.

Tabela 3 – Características das seqüências de *T. cruzi* imobilizadas no conjunto B de lâminas de microarranjos

Identificação	Características	Isolado – Grupo de <i>T. cruzi</i>	Número de acesso
TR	Tripanotiona redutase	CL Brener – Tc II	M38051
DHFR	Diidrofolato redutase-timidilato sintetase	Y – Tc II	L22484
TcMSH2	Proteína reparadora de "mismatch" de DNA	167 - TcII	AY092825
T. hexose	Transportador de hexose	CL Brener – Tc II	U05588
TEUF0001	Histona	CL Brener – Tc II	AA399704
TEUF0062	Sem Similaridade	CL Brener – Tc II	AA441768
TEUF0008	Sem Similaridade	CL Brener – Tc II	AA441772
TEUF0058	Sem Similaridade	CL Brener – Tc II	AA441764
TEUF0075	ND	CL Brener – Tc II	ND
HSP70	Proteína de choque térmico 70	Maracai - TcII	X07083
H49	Antígeno imunodominante associado a citoesqueleto e flagelo	G – Tc I	U16294
JL8	Antígeno	Tulahuen 2 -TcII	AF147956

ND: Não disponível

4.13.2. Amplificação das seqüências e aplicação na lâmina de microarranjos

As ESTs (tamanho médio de 800 pb) foram amplificadas por PCR, utilizando como iniciadores os pares de oligonucleotídios T3 e T7 ou M13 Forward e Reverse para as TENF e SP6 e T7, para TEUF, usando DNA polimerase (BIOLASE). Os produtos de amplificação foram purificados com placas Multiscreen (Millipore). Insertos representando

as seqüências de *T. cruzi* foram liberados dos plasmídios correspondentes com enzimas de restrição adequadas, separados por eletroforese em gel de agarose e purificados. Os DNAs genômicos controle foram fragmentados por sonicação, analisados em gel de agarose e fragmentos genômicos entre 600 e 1500pb foram selecionados para imobilização na lâmina.

As seqüências amplificadas e purificadas foram imobilizadas em lâminas a partir de uma diluição de 1:1 com Reagente D Microarranjo crosslinking reagent (DMSO). Esses DNAs foram colocados em microplacas de 384 poços (Amersham Pharmacia Biotech) que é a placa que o aplicador ("spotter") de 12 agulhas da Molecular Dynamics utiliza. Essa diluição é necessária, pois o reagente D permite uma maior ligação do DNA com a lâmina. Utilizamos lâminas de microarranjo tipo 7 e algumas do tipo 5 (Tipo 7 é recoberta com 3-aminopropiltrimetoxisilano e DMSO e entre o vidro e essa solução existe uma película de metal. A lâmina de tipo 5 não tem a película). Utilizamos as condições de imobilização especificadas pelo fabricante do Generation III Microarray System (Molecular Dynamics). A lâmina de microarranjo é dividida em duas áreas (lado esquerdo e lado direito). Cada lado é dividido em 12 subarranjos e cada subarranjo é preenchido por cada uma das agulhas. Em cada subarranjo podem ser aplicadas 32 x 12 amostras ("spots"). Na **Figura 1** apresenta-se o esquema da lâmina. Após a imobilização, o DNA foi irradiado com luz UV a 254 nm (Stratagene lamp) com energia total de 50 mJ.

Figura 1. No painel à esquerda mostra-se o esquema da lâmina de microarranjo com 12 subarranjos nos lados esquerdo (E) e direito (D). No painel à direita, mostra-se a distribuição de 384 amostras em um subarranjo (amostras 12 x 32).

No **conjunto A** de lâminas, uma vez que aproximadamente 740 amostras foram aplicadas, foi possível incluir várias réplicas da mesma sonda em cada subarranjo. Nestas lâminas utilizamos 10 subarranjos com 288 amostras em cada um. Cada subarranjo foi imobilizado em duplicada (lado esquerdo e direito) resultando em no mínimo 6 e no máximo 60 réplicas de cada seqüência. O mapa detalhado deste conjunto de lâminas pode ser encontrada na página <http://www.vision.ime.usp.br/~rvencio/Tcruzi/>. No primeiro e último subarranjo não foi colocado DNA, mas uma mistura de Tris 10mM, pH 8,0 e Reagente D.

No **conjunto B** de lâminas, as 800 amostras (item 4.13.1) foram imobilizadas seguindo o mesmo padrão utilizado para o conjunto A (ver acima), mas o número de réplicas de cada amostra foi um pouco menor. O mapa detalhado deste conjunto de lâminas pode ser encontrado no CD anexo.

4.13.3. Coloração da lâmina com VISTRA GREEN

Para visualizar os DNAs imobilizados, a lâmina foi corada com Vistra Green. Para isso diluímos o corante 1:10000 em tampão TAE. Adicionamos 15ml do corante diluído na câmara de hibridização e incubamos 10 minutos à temperatura ambiente. Após esse período, retiramos a lâmina da câmara e deixamos secar à temperatura ambiente ou com jatos de nitrogênio. Quando a lâmina estava completamente seca, foi analisada utilizando o laser verde (550 nanômetros-PMT 700) no "confocal laser scanning" da Molecular Dynamics. Vistra Green liga-se não covalentemente aos grupos fosfato da molécula de DNA permitindo a visualização dos DNAs imobilizados na lâmina. Após coloração, as lâminas ficam inutilizadas. As lâminas a serem coradas são colocadas em locais "estratégicos" no processo de aplicação do DNA: no começo, meio e fim para poder inferir a qualidade de todas as lâminas.

4.14. Extração de DNA total de *T. cruzi*

Formas epimastigotas cultivadas em meio LIT foram lavadas por centrifugação três vezes a 4°C, a 5900xg por 10 minutos em PBS. Para a extração de DNA total, 10⁹ células foram ressuspensas em 1ml de tampão de lise (NaCl 0,1M; EDTA 0,25mM; proteinase K 100µg/ml; lauril sarcosinato de sódio 0,5% (p/v); RNase A (livre de DNase) 10µg/ml) e incubadas a 50°C por 30 minutos. A extração foi feita com volume igual de fenol, agitação leve por uma hora à temperatura ambiente e centrifugação por 5 minutos a 1500xg. A fase aquosa foi retirada e extraída por mais duas vezes com igual volume de fenol. O fenol residual foi removido com igual volume de clorofórmio/álcool isoamílico (24:1) e centrifugação por 30 seg. A fase aquosa foi dialisada contra 1 litro de tampão (Tris-HCl 50mM, pH 8,0; EDTA 1mM, pH 8,0; NaCl 100mM) durante a noite, com agitação constante à temperatura ambiente. Acrescentou-se RNase A (livre de DNases) para concentração final de 100µg /mL, seguida de incubação a 37°C por 3 horas. O DNA foi purificado por uma extração com igual volume de fenol e dialisado com TE (Tris-HCl

10mM, pH 8,0, EDTA 1mM, pH 8,0) a 4°C durante a noite. A concentração foi determinada pela leitura da absorbância a 260 nm.

4.15. Extração de RNA das cepas

Para a extração de RNA, as células foram contadas em câmara de Neubauer. Cerca de 5×10^8 formas epimastigotas (fase logarítmica de crescimento) ou tripomastigotas metacíclicos (purificados conforme item 4.5) foram adicionados a tubos plásticos autoclavados que foram centrifugados a 7.000rpm por 7 minutos a 4°C. O sobrenadante foi descartado e o sedimento foi ressuspenso em 15ml de PBS. Outra centrifugação foi efetuada como acima. As células foram ressuspendidas em 1ml de PBS e transferidas para tubo eppendorf, sendo centrifugadas por 15 segundos. Em seguida, foi adicionado 1ml de TRIZOL (Life Technologies) e os eppendorfs foram deixados por 5 minutos à temperatura ambiente. Após esse tempo, foi adicionado 200µL de clorofórmio e os tubos foram incubados por 3 minutos. Após centrifugação a 12.000rpm por 15 minutos a 4°C, a fase aquosa foi transferida para outro tubo onde foi adicionado 0,5ml de isopropanol. O tubo foi deixado à temperatura ambiente por 10 minutos e depois centrifugado 12.000rpm por 10 minutos a 4°C. Em seguida, o sobrenadante foi removido. O RNA foi ressuspendido com 1ml de etanol 75% e centrifugado por 5 minutos a 9.500rpm a 4°C. O tubo contendo o RNA foi seco a vácuo e dissolvido em 52µL de água. O RNA foi transferido para banho a 65°C por 10 minutos. Uma alíquota (2µL) foi analisada por eletroforese em gel de agarose 0,8%. O restante do RNA foi armazenado a -70°C.

4.15.1. Tratamento de RNA com DNase

A um tubo eppendorf foi adicionado 50µg de RNA (10µL), 10µL de tampão 10x (Tris-HCl 400mM, pH8,0; MgSO₄ 100mM, CaCl₂ 10mM), 6,5µL de DNase 1U/ul, 2,75µL de RNAsin 40U/ul, 1µL de DTT 1mM, e água tratada com DEPC para volume final de 100µl. O tubo de reação foi incubado a 37°C por 30 minutos e em seguida foi adicionado 100µL de água para volume final 200µL. Em seguida o RNA foi purificado com 200µL de fenol pH 5,2 (equilibrado com NaOAc 20mM, EDTA 1mM) e com 200µl de clorofórmio. RNA foi precipitado com 1/10 do volume de NaOAc 3M (pH 5,3), EDTA 1mM e 2,0

volumes de etanol 100%, deixado a -70°C por 30 minutos. O tubo foi centrifugado a 12.000rpm por 25 minutos a 4°C e o RNA lavado com etanol 70%. Após centrifugação a 12.000rpm por 25 minutos a 4°C , o RNA foi seco e ressuspendido no mesmo volume inicial do RNA adicionado na reação (10 μL). 2 μL da preparação foram analisados em gel de agarose em cuba de eletroforese tratada por duas horas com DEPC.

4.16. Marcação de DNA total de CL Brener e Silvio com Cy3 ou Cy5 por síntese com random primer.

Nesta etapa é muito importante a utilização de luvas e a não exposição do corante à luz. Como molde ("template") para reação, utilizou-se 4 μg de DNA dividido em dois tubos. Em cada tubo foi adicionado 2 μg de DNA, Cy3- ou Cy5-dCTP 60 μM (Amersham Biosciences), Tris-HCl 50mM pH 6,8, MgCl_2 5mM, 2-mercaptoetanol 10mM, dATP, dGTP, dTTP 0,12mM e dCTP 0,06mM, 1 μL de "random nonâmeros" (Amersham) e 80U de Klenow DNA polimerase (GIBCO-BRL). O volume final da reação de marcação foi de 50 μL . A incubação foi a 37°C por 2 h. A reação foi interrompida pela adição de 5 μL de EDTA 0,5M.

4.16.1. Purificação dos DNAs marcados e quantificação

Imediatamente após a reação de marcação, fez-se uma purificação do DNA com placas Multiscreen da Millipore de 96 poços. Inicialmente foram adicionados 80 μL de tampão de ligação do DNA ("filter Binding Buffer"- Guanidina HCl 5,3M; KOAc 150mM pH 4,8) a cada amostra de DNA marcado, que foi transferida para a placa Millipore. Em seguida, centrifugou-se a 1000g por 1 minuto. A placa foi lavada 4 vezes com 200 μL de etanol 80%, com posterior centrifugação a 1000g por 1 minuto. Fez-se uma última centrifugação de 5 minutos e eluiu-se por 3 vezes com 30 μL de Tris-HCl 10mM pH 8,0.

Após a purificação, fez-se uma quantificação das sondas por medidas de absorbância (550nm para Cy3 e 650nm para Cy5), utilizando-se Cubeta - Ultra microcell de 50 μL ; 8,5 mm (Ficher Scientific) e espectrofotômetro da Hitachi (U2001). Ressalta-se que a eficiência de incorporação dos dois nucleotídeos modificados é diferente.

4.17. Hibridização da lâmina com DNA total de T. cruzi marcado com corantes Cy3 ou Cy5.

As amostras foram liofilizadas em vácuo e dissolvidas em H₂O para atingir 25% da solução mais 25% de tampão de hibridização (Amersham Pharmacia) e 50% de formamida. O DNA foi desnaturado a 98°C por 2 minutos. A hibridização com o microarranjo foi realizada sob lamínulas (24 x 60 mm). As lâminas foram colocadas na câmara de hibridização e incubadas por 12 a 18 horas a 42°C.

4.18. Lavagens após hibridização

Inicialmente, as lâminas foram mantidas numa câmara de hibridização à prova de água num banho à 42°C por no mínimo 16 horas. Após a hibridização, as lâminas foram lavadas com 1x SSC, 0,2% SDS, seguido por duas lavagens com 0,1x SSC, 0,2% SDS. Todas as lavagens foram realizadas a 55°C por 10 minutos. Uma lavagem adicional foi realizada à temperatura ambiente com 0,1x SSC por 1 minuto. Logo em seguida, as lâminas foram lavadas com água destilada por 10 segundos, secas com jatos de nitrogênio e submetidas à detecção de fluorescência com o “confocal laser scanning” da Molecular Dynamics. Utilizamos um PMT de 750 para Cy3 e para Cy5.

4.19. Marcação do RNA total das cepas com Cy3 ou Cy5.

Para a reação de síntese da primeira fita de cDNA aproximadamente 20µg de RNA foram marcados com fluoróforo. A mistura de reação continha Cy3- ou Cy5-dCTP 0,05mM, 1 µl do primer oligo-dT₍₁₅₎ (Amersham), 4 µl de random nonâmeros (Amersham), dATP, dCTP, dTTP 0,1mM e dCTP 0,05mM, ditiotretitol 10mM e 400 unidades de Superscript II Reverse Transcriptase (Invitrogen) no tampão fornecido pelo fabricante. A incubação foi feita 42°C por 2,5 h. Em seguida, o RNA foi hidrolisado pela adição de 2µl de NaOH 2,5M e incubação a 37°C por 15 min. A solução foi neutralizada pela adição de 10µl de Hepes 2M. As amostras foram purificadas com placas Multiscreen (Millipore). A

medida de incorporação, hibridização com as lâminas e condições de lavagem foram as mesmas descritas acima para o DNA.

4.20. Obtenção de imagens e análise dos dados

4.20.1. Experimentos para validação do microarranjo de DNA

As lâminas foram analisadas com um scanner a laser (Molecular Dynamics) resultando em imagens de 16-bit. As intensidades foram extraídas usando o software ArrayVision 6.0 (Image Research). Para verificar a influência do método de extração de imagem, os dados foram também analisados utilizando o software Scanalyse 2.5 (<http://rana.lbl.gov>) resultando em nenhuma diferença significativa. Para a análise de genômica comparativa, foram realizados 4 experimentos de hibridização. Isso significa que cada sonda foi medida no mínimo 24 e no máximo 240 vezes dependendo do número de réplicas de cada sequência. Para a análise da expressão diferencial de genes, a preparação de RNA foi obtida de 3 culturas independentes de cada cepa (réplicas biológicas). Combinações diferentes do cDNA de cada par de cepas (CL Brener x Silvio e Famema x Hem 179) foram usadas para hibridizar com o microarranjo, resultando em oito experimentos para cada par de isolados. A normalização dos dados foi feita por Lowess (incluído no pacote R disponível em <http://www.r-project.org>) utilizando-se o gráfico de dispersão MA, onde $M = \log_2(\text{Cy5/Cy3 normalizado})$ e $A = 1/2 \log_2(\text{Cy5} \times \text{Cy3})$, como descrito em Yang *et al.* (2002). O ruído de fundo local (“background”) foi subtraído dos valores de intensidade para cada sonda da lâmina. Réplicas com baixa intensidade ($A \leq 2$) e com intensidades saturadas ($A \geq 9,5$) foram excluídas.

Para cada sonda, a média ($\bar{M}$) dos valores de M de todas as réplicas técnicas e biológicas foi calculada. O resultado final apresentado para cada sonda é a razão de expressão ($R = 2^{\bar{M}}$). Para a identificação das sondas diferencialmente expressas utilizamos o teste-*t*. A hipótese nula testada foi “ $\bar{M} \geq -0,58$ ” e “ $\bar{M} \leq 0,58$ ” para sondas com expressão reduzida e aumentada, respectivamente. Sondas com valores de $P \leq 0,01$ no teste-*t* foram consideradas diferencialmente expressas. Detalhes adicionais referentes a este tipo de análise podem ser encontrados em Koide *et al.* (2004). O critério de corte de $\pm 0,58$ para a definição

de expressão diferencial (ou seja, 1,5 vezes de expressão aumentada ou reduzida) foi definido num experimento conhecido como “self-self setup” onde hibridizamos a lâmina com cDNA de CL Brener marcado alternativamente com os corantes Cy3 e Cy5. Neste experimento, descrito em Resultados, verificamos que 99% das sondas apresentavam valores de M dentro do critério de corte ($\pm 0,58$). Detalhes adicionais podem ser encontrados na página <http://www.vision.ime.usp.br/~rvencio/Tcruzi/>.

4.20.2. *Análise da expressão diferencial de genes em cepas isoladas de pacientes chagásicos*

Nesse experimento investigamos a expressão diferencial de genes em três cepas isoladas de pacientes chagásicos que apresentam a forma indeterminada (VL10, Famema e Berenice 62) e três cepas de pacientes com a forma cardíaca da doença de Chagas (115, B13 e 147). Após hibridização as imagens foram analisadas com um scanner a laser (Molecular Dynamics). As intensidades foram extraídas usando o software ArrayVision 6.0 (Image Research). Para todas as hibridizações competitivas, a cepa VL10 (derivada de um paciente assintomático) foi usada como referência. Desta forma, foram feitas cinco combinações: VL10 x 115; VL10 x 147; VL10 x B13; VL10 x Famema e VL10 x Berenice 62. O RNA foi isolado de duas culturas independentes de cada cepa, o que representam duas réplicas biológicas. O RNA de cada réplica biológica foi hibridizado em uma lâmina gerando no mínimo 12 e no máximo 120 réplicas de cada gene (réplicas técnicas) dependendo da sonda. A normalização dos dados foi feita por Lowess, pelo gráfico de dispersão MA (Yang *et al.*, 2002). Como descrito acima, o ruído de fundo (“background”) foi subtraído dos valores de intensidade de cada sonda da lâmina. Réplicas com baixa intensidade ($A \leq 2$) e com intensidades saturadas ($A \geq 9,5$) foram excluídas.

4.20.3. *Análise Estatística*

Para a seleção dos genes cuja expressão varia nas cepas isoladas de cardíacos e de assintomáticos, usamos como premissa escolher genes que apresentam aumento ou diminuição de expressão em **todas** as cepas de cardíacos em relação a **todas** cepas de assintomáticos. Selecionamos estes genes aplicando três métodos. Baseamo-nos na

abordagem usada por Singh *et al.* (2002) e Reis *et al.* (2004) para a seleção de genes cuja expressão está correlacionada com classes de tumor. Neste caso, usamos as análises de correlação de Pearson e de Spearman. Esta análise foi aplicada para cada seqüência, sendo que as duas variáveis foram o valor M de todas as réplicas da seqüência e a classificação das cepas em cardíacas ou assintomáticas. Em ambas análises de correlação, utilizamos o valor de corte de $|\rho|=0,75$ para aceitar genes diferencialmente expressos nos dois grupos de cepas. A fórmula para o cálculo de $|\rho|$ pode ser encontrada em Petrie & Sabin (2000) e na página http://mathworld.wolfram.com/_SpearmanRankCorrelationCoefficient.html. Após definição das seqüências que apresentam diferenças no perfil de expressão pelos dois métodos de correlação, aplicamos o método “Bayes Error Rate” (Duda *et al.*, 2000) ou Erro de Classificação de Bayes. Esta abordagem tem sido utilizada em diferentes situações, por exemplo, na descoberta de padrões de seqüências (Chudova & Smith, 2003) e na busca de genes diferencialmente expressos em dados de SAGE (Serial Analysis of Gene Expression) de tumor de cérebro (ver página www.vision.ime.usp.br/~rvencio/SAGEbetaBin/bayes_error_rate.htm) e (Vêncio *et al.*, 2004). No nosso caso, este método verificou se as réplicas de cada gene estão separadas nos dois grupos de cepas (isoladas de assintomáticos e de cardíacos). A separação ou sobreposição dos dois grupos é indicada pelo valor (E). Ou seja, quanto mais próximo de zero for o valor de E, maior será a separação entre os dois grupos. Conseqüentemente, quanto mais próximo de 1 for o valor de E, maior será a sobreposição entre os grupos, indicando não haver variação de expressão gênica entre eles. Definimos o valor de corte para o Erro de Classificação de Bayes de $E=0,05$. Para cada sonda diferencialmente expressa apresentamos a razão de expressão R ($R = 2^{\bar{M}}$) em relação à cepa de referência VL10 e o valor de E (definido pelo “Bayes Error Rate”).

4.21. Confirmação das hibridizações diferenciais

Algumas ESTs que apresentaram hibridização diferencial nos microarranjos foram analisadas em experimentos de Southern blot e/ou Northern blot com preparações de DNAs e RNAs das cepas, respectivamente.

4.21.1. Southern blot

Após digestão de 10µg de DNA com a enzima de restrição *Pst* I (New England Biolabs) por 14 horas a 37 °C, o DNA foi submetido a eletroforese em gel de agarose 0,8%. Após a corrida, o DNA foi despurinado com HCl 0,2M por 10 minutos. Em seguida, o DNA foi desnaturado duas vezes com NaOH 0,5M por 15 minutos, e neutralizado com NaCl 1,5M, Tris-HCl 0,5M pH 7,2; EDTA 1mM por 30 minutos (duas vezes). O DNA foi transferido para membranas de Nylon (Hybond-N, Amersham Pharmacia) em 2xSSC por 16 horas e foi fixado por exposição a UV por 5 minutos e aquecimento a 80°C por 2 horas. Em seguida, as membranas foram hibridizadas com sondas derivadas das sondas de interesse, marcadas com [α -P³²]dATP.

4.21.2. Northern blot

Agarose suficiente para 0,8% foi dissolvida em água tratada com DEPC (Dietilpirocarbonato) e resfriada até cerca de 60°C. Acrescentou-se tampão MOPS 10X para concentração final 1X (MOPS 1X= ácido 3-(N-morfolino) propanossulfônico 20mM, acetato de sódio 5mM, EDTA 1mM) e formaldeído 0,67M. As amostras de RNA foram liofilizadas (10µg de RNA total), ressuspensas em 20µl de formamida 50%, formaldeído 6% e MOPS 1X e aquecidas a 65°C por 15 min. Adicionou-se 2µl de tampão de amostra para RNA (glicerol 50%, Azul de bromofenol 0,4%, xileno cianol 0,4%, EDTA 1mM) e o material foi aplicado no gel. A eletroforese foi realizada em MOPS 1X a 70V, à temperatura ambiente por 3 horas. Após a eletroforese, o gel foi submetido a tratamento alcalino suave (NaOH 50mM, NaCl 10mM por 45 min), neutralização (Tris-HCl 0,1M pH 7,5 por 45 min) e, em seguida, equilibrado com 20x SSC por uma hora. O RNA foi transferido para membrana de nitocelulose (Hybond-C, Amersham Pharmacia) usando como tampão de transferência 20x SSC. A transferência ocorreu à temperatura ambiente por 16 horas. O RNA foi fixado por aquecimento a 80°C por 2 horas.

4.22. Obtenção de sondas radioativas e condições de hibridização

4.22.1. Marcação por “random primer extension”

Para a incorporação de nucleotídeos radioativos sondas de DNA foram marcadas com [α -P³²]dCTP com o kit de marcação de DNA utilizando random primers (Invitrogen). O equivalente a 200ng de DNA foram diluídos para um volume final de 23 μ L em água e desnaturados em banho fervente por 5 minutos. O eppendorf contendo o DNA foi transferido para gelo e foi adicionado ao tubo 20 μ M de dTTP, dCTP, dGTP e [α -P³²]dATP (10 μ Ci), tampão de mistura do “Random primer” (Hepes 0,67mM, Tris-HCl 0,17mM, MgCl₂ 17mM, β -mercaptoetanol 33mM, BSA 1,33mg/ml, hexanucleotídeos oligodeoxiribonucleotídeo iniciadores 18 DO₂₆₀ unidades/ml (pH 6,8)). Também foi adicionado 3U de Klenow em volume final de reação de 50 μ L. A reação foi deixada à temperatura ambiente por 3 horas e interrompida pela adição de 5 μ L de tampão de parada (EDTA 0,5M, pH 8,0).

4.22.2. Purificação da sonda em coluna de Sephadex

O DNA marcado foi purificado de nucleotídeos não incorporados por passagem em coluna de Sephadex G50 (Sambrook *et al.*, 1989). Para a medida de incorporação de P³² foi feita contagem em contador de cintilação líquida.

4.22.3. Hibridização de sondas radioativas com ácidos nucléicos imobilizados

As membranas de Southern foram pré-hibridizadas em forno HyBaid a 60°C por 90 min com 10ml de tampão de hibridização (Ficoll 0,1%, PVP 0,05%, EDTA 1mM, SSC 3x, SDS 0,1 % e DNA de esperma de salmão 100 μ g/mL). Essa solução foi substituída por 10ml de tampão de hibridização contendo a sonda desnaturada por aquecimento. A hibridização ocorreu por 12 a 16 horas a 60°C.

As membranas de Northern foram pré-hibridizadas em forno HyBaid a 42°C por 90 min com 10ml de tampão de hibridização (SSPE 5x, formamida 50%, Denhardt 5x, SDS 0,5%). Essa solução foi substituída por 10ml de tampão de hibridização contendo a sonda desnaturada por aquecimento. A hibridização ocorreu por 12 a 16 horas a 42°C.

4.22.4. Lavagem das membranas.

Após a hibridização, as membranas de Southern blot foram lavadas sequencialmente a 60°C nas seguintes condições: 2 lavagens em solução 3x SSC, SDS 0,1% por 15 minutos; 1 lavagem 1x SSC, SDS 0,1% por 15 minutos; 1 lavagem 0,1x SSC, SDS 0,1% por 30 minutos. As imagens radioativas foram coletadas em Phosphor Screens (Kodak) expostos por diferentes períodos e escaneados com o aparelho Storm (Molecular Dynamics). A análise densitométrica do sinal radioativo foi obtida com o programa ImageQuant Molecular Dynamics.

As membranas de Northern blot foram lavadas sequencialmente a 42°C nas seguintes condições: 2 lavagens em solução 2x SSPE, SDS 0,1% a 42°C por 15 minutos; 1 lavagem em 1x SSPE, SDS 0,1% a 42°C por 15 minutos; 1 lavagem 0,1x SSC, SDS 0,1% por 15 minutos. As imagens radioativas foram coletadas como acima.

4.22.5. Desibridização das membranas

Para reutilização das membranas de Southern blot, as sondas foram retiradas por lavagem com 0,1x SSC, SDS 0,5% a 90°C por 60 minutos. As membranas foram expostas a Phosphor Screens para verificar a eficiência do processo. Para as membranas de Northern blot, as sondas foram retiradas por lavagem com 0,1x SSPE, SDS 0,1% a 90°C por 30 minutos.

4.23. Análise da expressão gênica de fibroblastos humanos infectados com duas cepas de T. cruzi.

4.23.1. Cultivo de parasitas para infecção de fibroblastos humanos.

Formas epimastigotas das cepas VL 10 e 147 (Tabela 1) foram cultivadas em meio LIT (Liver Infusion Tryptose) contendo 10% de soro fetal bovino, a 28° C. Tripomastigotas metacíclicos foram obtidos pela diferenciação de 2×10^7 epimastigotas/ml em meio LIT contendo 20% (v/v) de meio de inseto Grace (Sigma) por 7–10 dias. Formas tripomastigotas foram obtidas do sobrenadante dos fibroblastos infectados com 10^7 tripomastigotas metacíclicos totais em meio DMEM contendo 2% soro bovino por 10 dias.

4.23.2. Cultivo de Fibroblastos Humanos (FHN) e infecção com cepas de T. cruzi

Fibroblastos Humanos (gentilmente cedidos pela Profa. Dra. Mari Sogayar, IQUSP) foram cultivados em frascos de 75cm² de área em 15ml de meio DMEM-10% soro fetal bovino até semi-confluência. Tripomastigotas ($2,5 \times 10^7$ parasitas) da cepa 147 foram adicionados às monocamadas e incubados por 24 horas a 37°C em atmosfera de 5% CO₂ com DMEM-2% de soro fetal bovino. Para a cepa VL10, as monocamadas foram infectadas com uma mistura de tripomastigotas e amastigotas ($1,5 \times 10^7$ parasitas) nas condições acima. Após 24 horas, o meio foi removido e substituído por DMEM-2% de soro fetal bovino. A cada dois dias, o sobrenadante das monocamadas foi observado ao microscópio para monitorar a saída de tripomastigotas. No caso da cepa 147, tripomastigotas eclodiram nos dias 8 a 10 pós-infecção. No caso da cepa VL10, amastigotas e tripomastigotas foram observados no sobrenadante a partir do dia 8 pós-infecção. Estes parasitas foram utilizados para infectar novas monocamadas de fibroblastos.

4.23.3. Meio condicionado de parasitas (MCP)

Tripomastigotas da cepa 147 ($2,5 \times 10^7$ parasitas) e tripomastigotas e amastigotas da cepa VL10 ($1,5 \times 10^7$ parasitas) recém eclodidos foram incubados em meio DMEM

contendo 2% de soro fetal bovino (FCS) por 24 horas a 37°C em atmosfera de 5% CO₂. Após incubação, os parasitas foram removidos por centrifugação e o meio foi filtrado em membranas Millipore de 0,2 µm. Fibroblastos humanos foram cultivados por 24 horas com esse meio condicionado a 37°C em atmosfera de 5% CO₂.

4.24. Extração de RNA de fibroblastos humanos não infectados, infectados e cultivados em meio condicionado

Após lavagens das células com PBS 1X, o RNA total foi purificado pelo método de Trizol a partir de monocamadas infectadas por 24 horas. RNA também foi extraído de células não infectadas ou cultivadas em meio DMEM condicionado (item 4.23.3). A integridade do RNA foi verificada por eletroforese em gel de agarose. As preparações de RNA foram tratadas com DNase de acordo com o protocolo do fabricante (Invitrogen). Brevemente, 10µl de reação contendo 1ug de RNA, 1X tampão de reação da DNase, 1U de DNase livre de RNase foram incubados à temperatura ambiente por 15 minutos. A reação foi interrompida pela adição de EDTA 2,5mM e incubação por 10 minutos a 65°C. Os RNAs tratados foram purificados em colunas de purificação da Qiagen (RNeasy mini column). Brevemente, o volume da reação foi ajustado para 100µl com água livre de RNase. Foram adicionados 350µl de tampão RLT e 250µl de etanol 100% à mistura de reação para promover a ligação do RNA com a coluna. Toda a mistura foi aplicada na coluna RNeasy que foi centrifugada a 10.000rpm por 15 segundos. Foram adicionados 500µl de solução de lavagem RPE e novamente a coluna foi centrifugada a 10.000rpm por 15 segundos. O RNA purificado foi eluído adicionando-se 30µl de água. A quantificação do RNA foi determinada pela medida da absorbância a 260 nm e 280 nm.

Nota: RLT e RPE soluções fornecidas pela Qiagen.

4.25. Preparação das sondas, marcação e hibridização.

Aproximadamente 10µg de RNA foram marcados com dUTP ligado a fluoróforo em uma reação de síntese da primeira fita de cDNA. A mistura de reação continha Cy3- ou Cy5-dUTP 0,02mM, 2µM do primer oligo-dT₍₁₅₎, 2µg de random nonâmeros (Amersham),

0,5mM de dATP, dCTP, dGTP e 0,08mM de dUTP, 20 unidades de inibidor de RNase e 2,5µl de LabelStar Reverse Transcriptase em 1X tampão da transcriptase reversa fornecido pelo fabricante. O volume final de reação foi 50µl. A incubação foi feita a 37°C por 2,0 h. Em seguida, o RNA foi hidrolisado pela adição de 10µl de NaOH 1 M e 5µl de EDTA 0,5M e incubação a 65°C por 20 min. A solução foi neutralizada pela adição de 25µl de Tris-Cl 1M pH 8,5. As amostras foram purificadas com colunas de purificação da Qiagen. Brevemente, foi adicionado 260µl de tampão PB à reação de marcação (tampão que possibilita a ligação do DNA à coluna). Essa solução foi adicionada a coluna MinElute e a coluna foi centrifugada por 1 minuto. Após duas lavagens consecutivas com tampões de lavagens (LS e PE) o cDNA marcado foi eluído com 10µl tampão de eluição EB (Tris-Cl 10mM, pH8,5).

Em cada experimento de hibridização, as sondas foram diluídas numa solução contendo 4µg de Poly A₍₄₀₎, 4µg de DNA humano Cot-1, 6,2µl de 20X SSC e 1µl de SDS 10%.

A solução de DNA foi desnaturada a 99°C por 2 minutos antes da hibridização com a lâmina. A hibridização com os microarranjos de DNA contendo oligonucleotídeos representando 24.000 genes humanos foi realizada sob lâminulas (24 x 60 mm). A lâmina foi preparada no NCI-SAIC (Frederick-USA). As lâminas foram mantidas numa câmara de hibridização à prova de água num banho a 57°C por no mínimo 16 horas. Após a hibridização, as lâminas foram lavadas com 2x SSC, SDS 0,1%, seguido por uma lavagem com 1x SSC por 5 segundos à temperatura ambiente. Uma lavagem adicional foi realizada à temperatura ambiente com 0,1x SSC por 30 segundos. Dependendo da imagem obtida da lâmina, uma nova lavagem com 0,1x SSC foi realizada. As lâminas foram secas por centrifugação e submetidas à detecção de fluorescência.

4.26. Obtenção de imagem e análise dos dados

As lâminas foram analisadas com um scanner a laser GenePix (Axon instruments) do NCI-NCC e as intensidades foram extraídas usando o programa GenePix Pro 5.0 Analysis Software. A normalização foi feita por Lowess (Dudoit *et al.*, 2002) realizada através do sistema automatizado mAdb do NCI-NCC. Sondas com intensidade menor que

duas vezes o “background” em ambos os canais foram excluídas. Somente sondas que apresentaram razão maior que 1,5 ou 0,66 (experimento reverso) foram consideradas como diferencialmente expressas.

4.27. Confirmação das hibridizações diferenciais por “real time PCR” RT-PCR

Para validar alguns resultados obtidos no microarranjo, a expressão diferencial de alguns genes foi analisada por RT-PCR nas facilidades do NCI-NCC. A primeira fita de cDNA foi sintetizada incubando-se 0,5µg de RNA das amostras com a mistura de reação da Omniscript kit (Qiagen). Na mistura de reação foi colocado tampão de reação 1X, 5µl dNTP (Omniscript kit-Qiagen), 10µM do primer oligo-dT₍₁₅₎, 50U de inibidor de RNase e 5µl de transcriptase reversa (Omniscript kit-Qiagen), em um volume de reação final de 100µl. A reação foi incubada por 2 horas a 37°C.

PCR em tempo real de alguns genes foi realizada no sistema ABI 7900 Sequence Detection System (Applied Biosystems) adicionando-se a uma mistura de reação 2µl de cDNA molde. A mistura de reação consiste de 1X Assay Mix (Applied Biosystem) contendo 5pmoles de primers “forward” e “reverse” (específicos para cada sequência), 1X TaqMan Universal PCR Máster Mix (Applied Biosystems) em um volume de reação de 10µl. A análise dos dados foi feita pelo software ABI 7900 e a normalização foi feita utilizando-se valores encontrados com o gene de rRNA 18S humano. A razão final ($2^{-\Delta\Delta CT}$) de expressão de cada gene é expressa em relação a células humanas não infectadas (Livak & Schmittgen, 2001). Os dados de RT-PCR foram obtidos a partir de duplicatas experimentais.

5. RESULTADOS

5.1. Construção e validação dos microarranjos de DNA para o estudo da expressão diferencial de genes e de genômica comparativa em cepas de T. cruzi.

5.1.1. Caracterização dos alvos de DNA para o conjunto A de lâminas de microarranjo

Neste estudo adotamos a definição “sonda” e “alvo” com base nas recomendações do suplemento da Nature Genetics de Janeiro de 1999 [The Chipping forecast. 21 (suppl.)], onde o termo “sonda” refere-se à sequência de DNA aplicada no microarranjo e o termo “alvo” refere-se à amostra a ser hibridizada na lâmina.

Para a construção do conjunto A de microarranjos de *T. cruzi*, foram aleatoriamente selecionados 710 clones de uma biblioteca normalizada e não normalizada de cDNA da forma epimastigota de CL Brener. As sequências de nucleotídeos de 576 ESTs estavam disponíveis no banco de dados. Estas ESTs são identificadas por TENF (biblioteca normalizada) e TEUF (biblioteca não normalizada). Sequenciamos 134 clones de cDNA que foram depositados no dbEST do GenBank e denominados TENQ (biblioteca normalizada) e TEUQ (biblioteca não normalizada). Os números de acesso destas sequências são: CF134346 a CF134365 e CF243279 a CF243392 (Tabela A no CD anexo). A busca de similaridade com os programas BLASTN e BLASTX indicou que 75% das ESTs (532/710) não apresentavam similaridade no banco de dados.

Vinte sequências de *T. cruzi* clonadas a partir de diversas cepas e caracterizadas em distintos laboratórios e que representam genes que codificam proteína ou RNA, ou sequências não codificantes também foram imobilizados na lâmina juntamente com DNAs controles (Ver Item 5.1.1.2 e Tabela 2). Considerando as sequências com similaridade no banco de dados (178 ESTs e 20 genes clonados), 77,7% (154/198) das sequências apresentaram similaridade com genes de *T. cruzi* ou de outros tripanossomatídeos, enquanto 22,3% (44/198) apresentaram similaridade significativa com genes de outros organismos. As categorias funcionais mais representadas são sequências relacionadas à síntese protéica (24%); metabolismo celular (16%); proteínas de superfície celular (14%) e elementos estruturais e do citoesqueleto (12%). (Gráfico 1).

Gráfico 1: Classificação funcional das seqüências contidas no conjunto A de lâminas de microarranjo

O agrupamento das seqüências foi feito com o programa Cap3, indicando 665 (91%) seqüências únicas (Tabela 4).

Tabela 4- Agrupamento das 730 seqüências por CAP 3

Agrupamentos ("clusters")	Número de agrupamentos ("clusters")
Com 2 a 5 seqüências	51
Singletons	614
Total de seqüências únicas	665

5.1.1.1. Seqüenciamento de ESTs não depositadas em bancos de dados.

Conforme comentado acima, 134 ESTs não tinham sua seqüência depositada em bancos de dados. Desta forma, julgamos importante determinar sua seqüência para poder depositá-la em bancos de dados públicos e também para poder construir nosso banco de dados do microarranjo.

O seqüenciamento das ESTs foi realizado utilizando os iniciadores de T3 e T7 que flanqueiam o inserto. Tendo em vista a clonagem direcional do cDNA, a seqüência obtida com o iniciador T7 corresponde à porção 5' do mRNA e aquela obtida com o iniciador T3, corresponde à porção 3' (Ürmenyi *et al.*, 1999). As ESTs foram seqüenciadas a partir de

DNAs amplificado por PCR de colônia ou pela preparação plasmidial original diluída 1:50 em TE. Os fragmentos amplificados foram usados na reação de sequenciamento. Com as seqüências obtidas a partir de uma ou ambas as extremidades foi realizada uma busca de similaridade com os programas BLASTN e BLASTX. Aceitamos como possível similaridade valores de E menores que 10^{-5} .

Inicialmente transformamos várias ESTs e fizemos PCR de colônia, para posterior sequenciamento. Na Figura 2 mostram-se algumas amplificações de colônias analisadas em gel de agarose 1%. As ESTs foram sequenciadas e o resumo das informações sobre a similaridade em bancos de dados pode ser visto no CD anexo (Tabela A).

Figura 2. PCR de colônia de bactérias transformadas com 7 ESTs. Foram feitas réplicas para a maioria das ESTs para confirmação dos resultados. Em 1 (EST0318), 2 e 3 (EST0546), 4 e 5 (EST0548), 6 e 7 (ESTs0566), 8 e 9 (EST0576), 10 (EST0581), 11 (EST0583).

Para acelerar o processo de sequenciamento, investigamos a possibilidade de fazer o sequenciamento direto das ESTs a partir do DNA plasmidial, ao invés de fazê-lo por PCR de colônia como descrito acima. Desta forma, o DNA da preparação plasmidial original foi diluído 1:50 em TE e 0,5µl foram utilizados na reação de PCR com os iniciadores T3 e T7. Nas Figuras 3 e 4 mostram-se os produtos de amplificação de 23 ESTs. Observa-se que a grande maioria dos produtos de amplificação das ESTs são fragmentos com tamanho molecular em torno de 800 pb.

Figura 3. Análise em gel de agarose dos produtos de amplificação das ESTs com os iniciadores T3 e T7. Foram feitas réplicas para a maioria das ESTs para confirmação dos resultados. Amostras: 1, 2 e 15 (EST0589), 3 e 4 (EST0597), 5 e 6 (EST0602), 7 e 8 (EST0607), 9 e 10 (EST0610), 11 e 12 (EST0614), 13 e 14 (EST0618).

Figura 4. Análise em gel de agarose dos produtos de amplificação das ESTs com os iniciadores T3 e T7. Amostras: 1 (EST0633), 2 (EST0635), 3 (EST0636), 4 (EST0640), 5 (EST0643), 6 (EST0644), 7 (EST0645), 8 (EST0646), 9 (EST0647), 10 (EST0648), 11 (EST0649), 12 (EST0650), 13 (EST 0651), 14 (EST0658), 15 (EST 0659), 16 (EST0659), 17 (Padrão de massa molecular).

Uma vez confirmada a amplificação das ESTs, uma alíquota foi utilizada nas reações de sequenciamento usando os iniciadores T3 ou T7. Devido à grande dificuldade de sequenciar regiões 3' das ESTs que possuem cauda de poli A, 90% das seqüências correspondem à região 5' e somente 10%, à região 3'. A busca de similaridade com os programas BLASTN e BLASTX indicou que 52% das ESTs (70/134) não possuíam similaridade no banco de dados. Considerando as seqüências com similaridade no banco de dados (64 seqüências), 76% (48/134) das seqüências apresentaram similaridade com genes previamente identificados em *T. cruzi* ou em outros tripanossomatídeos, enquanto 24% (16/134) apresentaram similaridade significativa com genes de outros organismos. A Tabela A com a identidade e a similaridade das 134 ESTs seqüenciadas pode ser encontrada no CD anexo

5.1.1.2. Genes de diversas cepas de *T. cruzi* aplicados no microarranjo

Vários genes de *T. cruzi* previamente caracterizados foram colocados na lâmina do microarranjo. A maioria dos genes estava clonada em plasmídios, os quais tiveram que ser transformados e purificados. Os insertos foram purificados com GFX após digestão com enzima de restrição apropriada ou por amplificação com primers específicos. Para confirmação da obtenção do inserto, alguns dos genes foram analisados apenas por eletroforese em gel de agarose 1% e outros, por Southern blot (por exemplo, gene ribossômico 24S α , ver abaixo). Em alguns casos recebemos em doação apenas o inserto do gene. Este teve que ser clonado em plasmídeo, antes de ser amplificado e aplicado na lâmina (por exemplo, RNA helicase, ver abaixo). Na Tabela 2 (em Material e Métodos) são apresentados os 20 genes de *T. cruzi* utilizados nas lâminas do microarranjo.

Gene de RNA ribossômico 24S α : O cistron ribossômico da cepa Y de *T. cruzi* foi clonado em nosso laboratório (Arruda *et al.*, 1990). Para verificar a representatividade do gene de rRNA 24S α no genoma de distintas cepas, o inserto, liberado do plasmídio que o continha com a enzima *Pst* I, foi purificado e marcado por "random primer synthesis" com (α -P³²)dATP. Esta sonda foi hibridizada com uma membrana, gentilmente cedida por Aurélio Pedroso Júnior, contendo bandas cromossômicas de várias cepas separadas por eletroforese em campo pulsado (PFGE) (Cano *et al.*, 1995). Observamos que a sonda hibridiza com bandas de tamanho molecular diferente nas cepas, ilustrando sua representatividade e a variabilidade genética das mesmas. O autoradiograma é mostrado na Figura 5.

Figura 5. Hibridização de DNA de bandas cromossômicas de distintas cepas com sonda do gene de rRNA 24S α . Marcador de massa molecular à esquerda da figura.

Gene da RNA helicase da família DEAD-box: Este gene foi caracterizado no laboratório do Prof. Dr. Hector Torres do INGEBI (Buenos Aires) em trabalho em colaboração com nosso grupo. Foi observado que o gene da RNA helicase apresenta um nível de expressão oito vezes maior em tripomastigotas metacíclicos de CL Brener em relação a epimastigotas (Diaz-Añel *et al.*, 2000). Para obter quantidade desta seqüência, clonamos um fragmento de 614 pb da porção 3' terminal do gene da RNA helicase no vetor pBluescript nos sítios *Sal* I e *Hind* III. O seqüenciamento do inserto confirmou que

háviamos clonado a porção correta (dado não mostrado). Este inserto foi marcado por "random primer" com (α -P³²) dATP e utilizado como sonda em Southern blot de DNA genômico de várias cepas digerido com *Pst* I (Figura 6). Observa-se que a sonda hibridiza com uma banda de 2,7 kb nas cepas Y e Esmeraldo, com uma banda de 1,15 kb nas cepas Silvio, Dm28c e YuYu e com duas bandas de 1,15 e 2,7 kb nas cepas CL Brener, NR, SO3 e Sc43. Este resultado indica que esta sequência está representada no genoma das várias cepas de *T. cruzi*, mas que o arranjo genômico apresenta variabilidade. A sonda da RNA helicase foi também hibridizada com bandas cromossômicas das cepas e o resultado é mostrado na Figura 7.

Figura 6. Southern blot de DNA genômico digerido com *Pst* I e hibridizado com sonda da RNA helicase. Cepas: (1) CL Brener, (2) Silvio, (3) Y, (4) NR, (5) SO3, (6) SC43, (7) Dm28, (8) Esmeraldo, (9) YuYu.

Figura 7. Bandas cromossômicas hibridizadas com a sonda da RNA helicase. Cepa (1) CL Brener, (2) Esmeraldo, (3) Y, (4) NR, (5) SO3, (6) SC43, (7) Dm28, (8) Silvio, (9) YuYu.

5.1.2. Amplificação e purificação das ESTs

A amplificação das ESTs da biblioteca normalizada foi obtida com os primers T3 e T7 ou M13 Forward e Reverse, e T7 e SP6 para as ESTs da biblioteca não normalizada,

usando DNA polimerase (BIOLASE). Os padrões dos produtos amplificados de duas placas de 96 poços, analisados por eletroforese em gel de agarose, são mostrados na Figura 8. Poços vazios significam reações que não apresentam produtos amplificados. Após amplificação, as ESTs foram purificadas em placas da Millipore (Multiscreen) de 96 poços e analisadas novamente em gel de agarose. A concentração de DNA foi cerca de 40ng/ml, o que está nos níveis aceitáveis de acordo com protocolos sugeridos pelo fabricante do Microarranjo (Molecular Dynamics).

Figura 8- Produtos de amplificação por PCR das ESTs de duas placas de 96 poços. No lado direito de cada fileira aplicamos padrão de peso molecular, cujos tamanhos são mostrados do lado direito da figura.

5.1.3. Topologia do conjunto A de microarrajos

As aproximadamente 740 sondas foram imobilizados em 10 subarranjo com 288 amostras cada. Cada subarranjo foi imobilizado em duplicada (lado esquerdo e direito da lâmina) resultando em no mínimo 6 e no máximo 60 réplicas de cada sequência.

Uma fonte conhecida de ruído que ocorre em todos os experimentos de microarranjo é resultante da variação experimental intrínseca. Nesse estudo, a estimativa de ruído foi determinada em um experimento onde duas amostras idênticas de RNA de CL Brener foram marcadas com Cy3 ou Cy5 e hibridizadas em uma mesma lâmina. Este experimento é conhecido como “self-self setup”. As intensidades de fluorescência foram obtidas pelo programa ArrayVision e o gráfico de dispersão comparando a hibridização do cDNA marcado com os corantes Cy3 e Cy5 é mostrado na Figura 9A. A linha no topo da figura indica sinal de saturação. A normalização foi realizada por Lowess (Dudoit *et al.*, 2002). Os dados também podem ser visualizados no gráfico de dispersão MA (Yang *et al.*, 2002), onde M é a razão de hibridização [$M = \log_2(\text{Cy5}/\text{Cy3})$] e A está relacionado com a intensidade de fluorescência [$A = \frac{1}{2} \log_2(\text{Cy5} \times \text{Cy3})$] (Figura 9B). Neste experimento, a razão de hibridização Cy3/Cy5 esperada é 1 e o valor de M esperado é zero. Com base nos dados experimentais, definimos um valor de corte do log da razão de expressão (M) de ± 0.58 , valor onde se encontram 99% das sondas imobilizadas na lâmina. Esse valor corresponde a uma razão de corte de 1,5 ($M = 2^{0.58}$), ou seja, regulação aumentada ou reduzida em 1,5 vezes. Assim toda a sonda que apresentar um valor de M dentro desse intervalo é considerada como não diferencialmente expressa.

Figura 9 – Gráfico de dispersão (A); gráfico MA (B) da hibridização do microarranjo com cDNA da forma epimastigota de CL Brener marcado com Cy3 e Cy5.

5.1.4. Análise de genômica comparativa dos isolados CL Brener e Silvio X10 cll

Visando a validação dos microarranjos de DNA para análise de genômica comparativa, escolhemos dois clones de cepas "polares" de *T. cruzi*: CL Brener, isolado do vetor triatomíneo e considerado um genoma híbrido de *T. cruzi* II, e Silvio X10 cll, isolado de um indivíduo em fase aguda da doença de Chagas e classificado como *T. cruzi* I (Tabela 1). DNA dos parasitas foram marcados com Cy3 e Cy5, misturados e hibridizados com a lâmina (Figura 10). Em quatro experimentos independentes, observamos que 9,3% (68/730) das sondas apresentaram hibridização diferencial com significância estatística ($P \leq 0,01$): 31 sondas apresentaram maior sinal de hibridização com o DNA de CL Brener, enquanto 37 sondas hibridizaram mais intensamente com DNA de Silvio. Esse dado sugere diferenças na abundância dos genes no genoma das cepas e/ou variação na similaridade das seqüências entre das duas cepas (a identidade de algumas destas seqüências é mostrada nas Tabelas 5 e 6). A análise dos dados foi realizada como descrito nos Materiais e Métodos (Item 4.20.1).

Tabela 5 – Seqüências com maior hibridização com DNA de CL Brener ou Silvio X10 e mais expressas em CL Brener ou SilvioX10.

Seqüências (n = 11 de 730; 1,5%) com maior hibridização com DNA de CL Brener e mais expressas em CL Brener			
Identidade	Similaridade ^a	Razão de hibridização com DNA ^b	Razão de hibridização com cDNA ^b
CL65 ^d	DNA repetitivo de 195pb	1,68	4,67
G3 ^d	DNA repetitivo de 195pb	6,07	6,47
TENF0385	Tc mRNA com seqüência SIRE ^h	1,88	1,88
TENF0396 ^e	Tc gene com seqüência repetitiva	2,06	2,95
TENF0421	Região intergênica entre genes de Tc para OMPDCase ⁱ -OPRTase ^j e proteína de superfície	1,61	2,69
TENF0593	Sem similaridade	2,14	3,01
TENF0599	Sem similaridade	2,22	1,61
TENF0693 ^e	Tc gene com seqüência repetitiva	2,00	3,05
TENF0775	Sem similaridade	1,90	2,74
TENF0855	Tc histona H2A	2,00	2,00
TENQ0711	Tc DNA genômico da cepas Y	1,90	2,10
Seqüências (n = 7 de 730; 0,96%) com maior hibridização com DNA de Silvio X10 e mais expressas em Silvio			
Identidade	Similaridade ^a	Razão de hibridização com DNA ^c	Razão de hibridização com cDNA ^c
B11 ^f	Retrotransposon	1,75	4,29
Pact	Actina 2	1,83	2,48

Telomero	Região telomérica	1,61	2,95
TENF0453	<i>T. cruzi</i> “sterile alpha motif”	1,70	1,70
TENQ0659	Sem similaridade	2,00	1,77
TENQ0889	Tc com sequência SIRE ^h	1,73	1,65
TEUF0049 ^f	Retrotransposon	1,72	3,46
<i>Seqüências (n = 3 de 730; 0,41%) com maior hibridização com DNA de Silvio X10 e mais expresso em CL Brener</i>			
TENF0310	Sem similaridade	2,05	2,10
TENQ0584 ^g	Tc Esmeraldo histona H3	1,80	1,90
TENF0516 ^g	Tc Tulahuen histona H3	2,04	1,99

^a A melhor similaridade obtida pelos programas BLASTN e/ou BLASTX ($E \leq 10^{-5}$)

^b CL Brener/Silvio X10 (P valor $\leq 0,01$ no teste t).

^c Silvio X10/CL Brener (P valor $\leq 0,01$ no teste t).

^{d-g} Seqüências pertencentes ao mesmo agrupamento (“cluster”)

^h SIRE = pequeno elemento repetitivo intercalado (“short interspersed repetitive element”)

ⁱ OMPDCase = orotidine-5'-fosfato descarboxilase (síntese de pirimidina)

^j OPRTase = orotate fosforribosil transferase (síntese de pirimidina)

Tabela 6– Seqüências com maior hibridização com DNA de CL Brener ou SilvioX10 e igualmente expressas nas duas cepas

<i>Seqüências (n = 20 de 730; 2,7%) com maior hibridização com DNA de CL Brener</i>		
Identidade	Similaridade ^a	Razão de hibridização ^b
TENF0018	Sem similaridade	1,58
TENF0065	CL Brener cosmídio 1m17 cromossomo 3	2,71
TENF0146	Sem similaridade	1,83
TENF0198	Sem similaridade	1,90
TENF0267	Sem similaridade	2,06
TENF0296	Sem similaridade	2,02
TENF0303	Sem similaridade	1,94
TENF0306	Sem similaridade	1,83
TENF0333	Sem similaridade	1,97
TENF0409	Sem similaridade	1,77
TENF0426	Sem similaridade	2,75
TENF0479	Tc RNA polimerase II subunidade maior	1,97
TENF0588	Sem similaridade	2,27
TENF0598	Tc Histona H1 (H1.M6.1)	1,80
TENF0714	Seqüência SIRE ^g em fator de tradução (eIF6)	1,75
TENF0787	Sem similaridade	1,68
TENF0912	Sem similaridade	2,10
TENQ0900	Sem similaridade	1,63
TENU1969	Tc adenilato ciclase (TczAC)	1,87
TEUF0051	Sem similaridade	1,69
<i>Seqüências (n = 25 de 730; 3,42%) com maior hibridização com DNA de Silvio X10</i>		
Identidade	Similaridade ^a	Razão de hibridização ^c
18SG ^d	Tc 18S rRNA	5,80
18SY ^d	Tc 18S rRNA	5,10
24S	Tc 24S rRNA α	2,86
B12	Antígeno de 230 kDa	1,87
EF ^e	Fator de elongação subunidade α	2,07
Minixon	“Spliced leader” e região intergênica	21,52
TENF0105	<i>L. major</i> cepa BT010 proteína ribossômica S2 mRNA	2,18
TENF0163	Sem similaridade	2,10
TENF0233	<i>T. brucei</i> complexo de proteínas de editoração de RNA	1,94

TENF0388	Sem similaridade	1,85
TENF0452	Tc clone 21 sequência SIRE	2,14
TENF0536	Sem similaridade	2,33
TENF0625	Tc G2RP gene de trans-sialidase	2,16
TENF0702	Sem similaridade	1,59
TENF0718	Tc 24S β rRNA	4,18
TENQ0583	Tc glicoproteína de superfície (TSA-1) de tripomastigota	1,66
TENQ0828	Tc fator de elongação 1 alfa	1,89
TENQ0861	Sem similaridade	2,40
TENQ0901 ^d	Tc cepa Y 18S rRNA	3,91
TEUF0050	Tc gene mitocondrial HSP70 (MTP70)	1,60
TEUF0096 ^f	Tc mRNA de alfa tubulina	1,69
TEUF0102	<i>L. major</i> antígeno	1,67
TEUF0113	Sem similaridade	1,65
TEUF0253 ^f	Tc mRNA de alfa tubulina	2,20
TEUF0259 ^e	Tc fator de elongação 1 alfa	1,88

^a A melhor similaridade obtida pelos programas BLASTN e/ou BLASTX ($E \leq 10^{-5}$)

^b CL Brener/Silvio X10 (P valor $\leq 0,01$ no teste t).

^c Silvio X10/CL Brener (P valor $\leq 0,01$ no teste t).

^{d-f} Sequências pertencentes ao mesmo agrupamento ("cluster")

^g SIRE = pequeno elemento repetitivo intercalado ("short interspersed repetitive element")

Figura 10. Lâmina de microarranjo de *T. cruzi* hibridizado com DNA de CL Brener e Silvio X10 cl1 marcados com Cy3 e Cy5, respectivamente. As duas imagens representam o lado esquerdo e direito da lâmina.

Figura. 11 - (A) Southern blot do DNA genômico de CL Brener (C) e Silvio (S) digerido com *Pst* I e hibridizado com sondas marcadas com P^{32} . O primeiro painel mostra o padrão de coloração do DNA com brometo de etídio. Abaixo dos outros painéis indica-se a denominação da sonda e a razão de hibridização (C/S) na lâmina de microarranjo.

(B) DNA genômico de algumas cepas digerido com *Pst* I. Do lado esquerdo, gel corado com brometo de etídio. Do lado direito, Southern blot hibridizado com sonda TENF0426. Canaletas: CL Brener (1); Silvio X10 cl1 (2); Y (3); NR cl3 (4); SO3 cl5 (5); SC43 cl1 (6); Dm 28c (7); Esmeraldo cl3 (8); YuYu (9).

Algumas destas sondas foram confirmadas por experimentos de Southern blot com DNA genômico das cepas CL Brener e Silvio. Na Figura 11A são apresentados os padrões obtidos para seis sondas que mostraram razões de hibridização CL Brener/Silvio de 2,1 a 2,7 e Silvio/CL Brener de 5,1 (ou CL Brener/Silvio= 0,19). Duas sondas controle com razão de hibridização de 1 foram também analisadas (TENF0066 e TENF0102). Esses experimentos confirmam diferenças no número de cópias no genoma das duas cepas e/ou diferenças de similaridade das seqüências. Em alguns casos, observamos variação no arranjo genômico das sequencias (ver padrões obtidos para TENF0593; 0065; 18S rDNA, Figura 11A).

A distribuição genômica da TENF0426, que mostra razão de hibridização CL Brener/Silvio = 2,7, foi analisada em 9 isolados, representativos de *T. cruzi* I, *T. cruzi* II e genomas híbridos (Tabela 1). O Southern blot mostrado na Figura 11B confirma a grande variação na quantidade desta seqüência no genoma dos parasitas e sugere que esta EST pode ser uma seqüência repetitiva.

5.1.5. Expressão diferencial de genes em isolados de CL Brener e Silvio

A expressão diferencial em CL Brener e Silvio foi investigada a partir da hibridização da lâmina de microarranjos com cDNA obtido do RNA total isolado de formas epimastigotas na fase mid-log de crescimento. O tempo de duplicação das formas epimastigotas de CL Brener e Silvio é 44 h e 45 h, respectivamente. Oito experimentos de hibridização foram realizados com RNAs preparados de três culturas celulares obtidas independentemente. Identificamos 84 sondas diferencialmente transcritas que mostraram significância estatística ($P \leq 0,01$): 35 sondas mais expressas em Silvio e 49 sondas mais expressas em CL Brener. As seqüências são identificadas na Tabela 7. Entre as seqüências que não apresentaram variação de expressão entre as duas cepas estão os ESTs TENF0463, TENF0273, TENF0296. Estes ESTs não apresentam similaridade no banco de dados.

Tabela 7: Sequências diferencialmente expressas nas cepas Silvio e CL Brener

Expressão aumentada em Silvio (n = 35 sondas de 730; 4,79%)				
Identidade	Número de acesso	Similaridade ^a	Razão Silvio/CL Brener	Northern blot ^b
B11 ^c	U15615	Retrotransposon (similar a transcriptase reversa)	4,29	confirmado
pact	U20234	Actina 2	2,48	
telomere	K01633	Região telomérica	2,95	
TENF0102	AA556014	Tc retrotransposon L1Tcg62	2,33	confirmado
TENF0125	AA525729	Sem similaridade	1,64	
TENF0153	AA532208	Sem similaridade	2,28	confirmado
TENF0193	AA532128	Sem similaridade	1,74	
TENF0280	AA532175	Sem similaridade	1,96	
TENF0304	AA556073	Tc região intergênica	1,95	
TENF0317	AA556076	Sem similaridade	1,77	
TENF0335	AA556083	Tc caseína quinase 1.2 (CK1.2) mRNA	1,66	
TENF0453	AA675965	Tc Tcc1i14-2.7	1,70	
TENF0527	AA676013	Sem similaridade	1,78	
TENF0543	AA676026	Sem similaridade	1,79	
TENF0574	AA676054	Tc muc1 pseudogene	2,18	
TENF0587	AA676063	Tc glutationa-S-transferase/glutarredoxina putativa	1,56	
TENF0641	AA676105	<i>T. brucei</i> proteína ribossômica L1	1,93	
TENF0686	AA676135	Sem similaridade	1,61	
TENF0715	AA676154	Tc elemento repetitivo intergênico	2,29	confirmado
TENF0733	AI717863	Sem similaridade	1,60	
TENF0743	AI717872	Sem similaridade	1,65	
TENF0784	AI717749	<i>L. major</i> cromossomo 35 clone L8650 cepa Friedlin	1,84	
TENF0831	AI717778	Sem similaridade	1,63	
TENF0851	AI717796	<i>T. b. rhodesiense</i> similar a proteína ribossômica	1,72	
TENQ0581	CF243334	Sem similaridade	1,76	
TENQ0659	CF243283	Sem similaridade	1,77	
TENQ0668	CF243285	Sem similaridade	1,73	
TENQ0676	CF243286	<i>S. cerevisiae</i> biossíntese de esterol; Erg2p	2,06	
TENQ0698	CF243380	Sem similaridade	2,01	
TENQ0758	CF243302	Sem similaridade	2,10	
TENQ0816	CF243316	Sem similaridade	2,13	
TENQ0823	CF243322	Sem similaridade	1,77	
TENQ0889	CF243348	Tc região repetitiva SIRE	1,65	
TEUF0049 ^c	AA426691	Tc mRNA de não-LTR retrotransposon	3,46	
TEUF0061	AA426693	Sem similaridade	1,77	
Expressão aumentada em CL Brener (n = 49 sondas de 730; 6,71%)				
Identidade	Número de acesso	Similaridade ^a	Razão CL Brener/Silvio	Northern blot ^b
CL56 ^d	AY188964	DNA satélite de 195pb	3,82	
CL65 ^d	AY327543	DNA satélite de 195pb	4,67	
G3 ^d	K01771	DNA satélite de 195pb	6,47	confirmado
TENF0002	AA525711	Sem similaridade	1,92	

TENF0023	AA556024	Sem similaridade	1,72	
TENF0073	AA556144	Sem similaridade	1,98	
TENF0094	AA525746	Sem similaridade	1,63	
TENF0191	AA556046	Tc triparedoxina peroxidase mRNA	1,81	
TENF0209	AA556049	Sem similaridade	1,78	
TENF0228	AA532138	Sem similaridade	1,55	
TENF0235	AA556054	Tb ATPase subunidade alfa	3,63	confirmado
TENF0270	AA532168	Sem similaridade	1,81	
TENF0310	AA532189	Sem similaridade	2,10	
TENF0337	AA532201	Sem similaridade	2,18	
TENF0372	AA532078	Sem similaridade	1,75	
TENF0385	AA556097	Tc mRNA com sequência SIRE ^f	1,88	
TENF0394	AA532086	Sem similaridade	1,83	
TENF0396 ^e	AA556101	Tc gene 35.2 com repetição	2,95	
TENF0399	AA532091	Sem similaridade	1,85	
TENF0421	AA532099	Tc região intergência entre OMPDCase ^g -OPRTase ^h e proteína de superfície	2,69	confirmado
TENF0432	AA532103	Sem similaridade	2,30	
TENF0471	AA532112	Sem similaridade	1,65	
TENF0516	AA676002	Tc Tulahuen histona H3	1,99	
TENF0520	AA676006	Sem similaridade	2,45	
TENF0528	AA676014	Sem similaridade	1,97	
TENF0582	AA676060	Sem similaridade	1,78	
TENF0591	AA676066	Sem similaridade	1,82	
TENF0593	AA676068	Sem similaridade	3,01	
TENF0599	AA676073	Sem similaridade	1,61	
TENF0603	AA676076	Sem similaridade	1,64	
TENF0672	AA676123	Sem similaridade	3,55	confirmado
TENF0691	AA676138	Tc antígeno com sequência repetitiva 101	2,38	
TENF0693 ^e	AA676140	Tc uma repetição do gene 35.2	3,05	confirmado
TENF0736	AI717865	Sem similaridade	2,46	
TENF0775	AI717885	Sem similaridade	2,74	
TENF0777	AI717887	Sem similaridade	1,89	
TENF0802	AI717766	Sem similaridade	1,73	
TENF0844	AI717790	Sem similaridade	2,27	
TENF0846	AI717791	Sem similaridade	2,09	
TENF0855	AI717800	Tc histona H2A	2,00	
TENF0911	ND	ND	1,71	
TENQ0584	CF134352	Tc Esmeraldo histona H3	1,90	
TENQ0711	CF243292	Tc cepa Y (genoma)	2,10	
TENQ0859	CF243328	Sem similaridade	1,80	
TENQ0867	CF243338	Tc proteína de dedo de zinco PDZ5 mRNA	1,71	
TENQ0871	CF243340	Sem similaridade	1,62	
TENU0080	AA866508	Sem similaridade	1,73	
TEUF0008	AA441772	Sem similaridade	2,21	
TEUF0075	ND	ND	2,64	

^a A melhor similaridade obtida pelos programas BLASTN e/ou BLASTX ($E \leq 10^{-5}$)

^b Confirmação por Northern blot

^{c-e} Sequências pertencentes ao mesmo agrupamento ("cluster")

^f SIRE = pequeno elemento repetitivo intercalado ("short interspersed repetitive element")

^g OMPDCase = orotidine-5'-fosfato descarboxilase (síntese de pirimidina)

^h OPRTase = orotate fosforribosil transferase (síntese de pirimidina)

n = 84 sondas de 730; 11,50% ($P \leq 0,01$ no teste t)

ND = não determinado

Para confirmar a expressão diferencial, algumas seqüências foram usadas como sondas em Northern blots de RNA total de CL Brener e Silvio, e os padrões obtidos para sete dessas sondas estão na Figura 12. O sinal de hibridização foi quantificado por análise densitométrica usando-se o programa ImageQuant Molecular Dynamics. A hibridização com a sonda de gliceraldeído 3-fosfato desidrogenase (GAPDH) foi usada na normalização dos valores (dado não mostrado). As razões de CL Brener/Silvio obtidas nos Northern blots foram comparadas com as razões obtidas no experimento de microarranjos (Figura 12). Na maior parte dos casos observamos que a razão de hibridização no Northern Blot é maior que aquela obtida no experimento com microarranjo.

Figura 12 –Northern blot de RNA total de epimastigotas de CL Brener (C) e Silvio (S) hibridizado com sondas marcadas com P³². No primeiro painel mostra-se a região do gel onde migram os rRNAs 18S; 24S α e 24S β corados com brometo de etídio. Abaixo de cada painel, a identificação da sonda, as razões de hibridização no microarranjo (M) e no Northern blot (N) são indicadas.

Decidimos verificar se a expressão diferencial de genes entre CL Brener e Silvio estava relacionada com diferenças no número de cópias gênicas e/ou similaridade de seqüência entre as cepas estudadas. Assim, comparamos os dois conjuntos de experimentos onde os microarranjos foram hibridizados com os alvos de DNA e cDNA. Entre as seqüências mais expressas em Silvio, somente 7 (7/35; 20%) mostraram maior hibridização com o DNA de Silvio. Entre as 49 sondas com expressão induzida em CL Brener, 11 sondas (11/49, 22,4%) apresentaram maior hibridização com o DNA dessa cepa. Um resultado interessante foi encontrado em 3 sondas (3/49; 6,1%) que mostraram maior hibridização com o DNA de Silvio e maior expressão em CL Brener. Os dados referentes à

identidade das sondas e às razões de hibridização com DNA e cDNA são descritos nas Tabelas 5, 6 e 7.

5.1.6. Expressão diferencial de genes em isolados de pacientes chagásicos

Em uma tentativa inicial de identificar genes diferencialmente expressos em cepas de *T. cruzi* isoladas de pacientes chagásicos, RNA total foi purificado da fase mid-log de formas epimastigotas de Famema e Hem 179, ambas cepas pertencentes ao grupo *T. cruzi* II, e obtidas, respectivamente, de um indivíduo assintomático e de um paciente com cardiopatia grave e problemas digestivos (Tabela 1). O cDNA foi obtido a partir de diferentes culturas das cepas (réplicas biológicas) e hibridizado com lâminas de microarranjo (conjunto A). A análise das réplicas indicou uma expressão diferencial de 18 sondas (18/730; 2,36%) com valores de $P \leq 0,01$ (Tabela 8). A expressão diferencial de genes correspondentes a seis sondas foi confirmada por Northern blot. Na Figura 13A mostra-se o padrão de três destas sondas.

Tabela 8: Sequências diferencialmente expressas nas formas epimastigotas das cepas Famema (F) e Hem 179 (H)

Expressão aumentada em Famema (n = 7 sondas de 730; 0,95%)				
Identidade	Número de acesso	Similaridade ^a	Razão F/H	Northern blot ^b
B13	U15616	Tc antígeno de 140/116 kDa	1,90	confirmado
TENF0073	AA556144	Sem similaridade	1,64	
TENF0209	AA556049	Sem similaridade	2,06	
TENF0330	AA556080	Sem similaridade	1,77	
TENF0753 ^c	AI717881	Tc proteína de ligação (RBP) a RNA	2,96	confirmado
TENQ0649	CF243279	Sem similaridade	1,84	
TENQ0689 ^c	CF243287	Tc proteína de ligação (RBP) a RNA	2,84	confirmado
Expressão aumentada em Hem 179 (n = 11 sondas de 730; 1,50%)				
Identidade	Número de acesso	Similaridade ^a	Razão H/F	Northern blot ^b
B12	L07759	Antígeno de 230 kDa	1,89	
CL56 ^d	AY188964	DNA satélite de 195pb	2,61	
CL65 ^d	AY327543	DNA satélite de 195pb	3,31	
G3 ^d	K01771	DNA satélite de 195pb	3,07	
TENF0236	AA532144	Sem similaridade	2,60	confirmado
TENF0317	AA556076	Sem similaridade	2,18	
TENF0857 ^e	AI717801	Tc maxicírculo do cinetoplasto - NADH desidrogenase subunidade 7 (ND7)	51,22	confirmado

TEUF0045 ^e	AA676164	Tc maxicirculo do cinetoplasto - NADH desidrogenase subunidade 7 (ND7)	4,15	confirmado
TEUF0066	AA676168	Sem similaridade	3,26	
TEUF0105	AA399706	<i>L. amazonensis</i> 40S proteina ribossômica S24e mRNA	2,24	
TEUF0146 ^e	AA676184	Tc maxicirculo do cinetoplasto - NADH desidrogenase subunidade 7 (ND7)	3,81	

^a A melhor similaridade obtida pelos programas BLASTN e/ou BLASTX ($E \leq 10^{-5}$)

^b Confirmação por Northern blot

^{c-d} Sequências pertencentes ao mesmo agrupamento ("cluster")

Sequências diferencialmente expressas ($P \leq 0,01$ no teste *t*)

Também investigamos a expressão diferencial de genes em tripomastigotas metacíclicos das duas cepas humanas. Sessenta e seis sondas (66/730; 9,0%) foram identificadas como diferencialmente expressas (Tabela 9) e o transcrito correspondente à sonda TEUF 0045 é mostrado na Figura 13B. Observa-se sua elevada abundância em metacíclicos da cepa Hem 179, assim como observado para epimastigotas da mesma cepa (Figura 13 A).

Tabela 9: Sequências diferencialmente expressas em formas tripomastigotas metacíclicas das cepas Famema (F) e Hem 179 (H)

Expressão aumentada em Famema ($n = 33$ sondas de 730; 4,52%)				
Identidade	Número de acesso	Similaridade ^a	Razão F/H	Northern Blot ^b
B13	U15616	Antígeno de 140/116 kDa	1,92	
TENF0102	AA556014	Tc retrotransposon L1Tcg62, sequência completa	2,09	
TENF0191	AA556046	Tc triparedoxina peroxidase mRNA	1,78	
TENF0204	AA532132	Sem similaridade	1,82	
TENF0209	AA556049	Sem similaridade	1,88	
TENF0222	AA532134	Sem similaridade	1,68	
TENF0287	AA532180	Sem similaridade	1,68	
TENF0304	AA556073	Tc retrotransposon codificador dos genes de proteína gag e transcriptase reversa endonuclease	1,81	
TENF0358	AA532070	Sem similaridade	1,77	
TENF0413	AA556106	Tc "shed-acute-phase-antigen" (SAPA) gene	1,72	
TENF0488	AA675977	Sem similaridade	2,09	
TENF0490	AA675979	Sem similaridade	1,70	
TENF0511	AA675997	Sem similaridade	2,16	

TENF0573	AA676053	Sem similaridade	1,93	
TENF0604	AA676077	Sem similaridade	1,78	
TENF0605	AA676078	Sem similaridade	2,00	
TENF0753 °	AI717881	Tc proteína de ligação a RNA	6,24	
TENF0785	AI717750	Sem similaridade	1,64	
TENF0797	AI717761	Sem similaridade	2,16	
TENF0846	AI717791	Sem similaridade	2,81	
TENF0908	AI717803	Sem similaridade	1,67	
TENF0915	AI717809	Sem similaridade	2,33	
TENF1052	AI717894	Sem similaridade	1,72	
TENQ0643	CF134360	Tc cepa CL Brener	2,95	
TENQ0689 °	CF243287	Tc proteína de ligação a mRNA	5,66	
TENQ0826	CF243324	Sem similaridade	1,70	
TENQ0873	CF243342	<i>Crithidia fasciculata</i> citocromo c oxidase subunidade 9 [N-terminal]	1,73	
TENQ0888	CF243373	<i>L. major</i> antígeno	1,61	
TENQ0895	CF243350	Sem similaridade	1,70	
TENU2633	AI069632	<i>L. major</i> genes de proteína ribossômica S8	1,78	
TENU4317	AW325013	<i>T. brucei</i> apocitocromo c1	1,76	
TEUF0049	AA426691	Tc retrotransposon L1Tcg62	2,21	
TEUF0224	AA426674	Tc proteína de membrana do cinetoplasto	2,64	
Expressão aumentada em Hem 179 (n = 33 sondas de 730; 4,52%)				
Identidade	Número de acesso	Similaridade ^a	Razão H/F	Northern blot ^b
CL56 ^d	AY188964	DNA satélite de 195pb	3,11	
CL65 ^d	AY327543	DNA satélite de 195pb	3,31	
G3 ^d	K01771	DNA satélite de 195pb	3,07	
SLA	AI069901	Pequeno RNA nuclear	3,76	
TENF0012	AA525724	Tc OMPDCase ^g -OPRTase ^h	1,83	
TENF0066	AA556141	Sem similaridade	2,83	
TENF0129	AA556019	Sem similaridade	2,21	
TENF0236	AA532144	Sem similaridade	2,34	
TENF0409	AA532094	Sem similaridade	2,49	
TENF0497	AA675986	Sem similaridade	2,49	
TENF0520	AA676006	Sem similaridade	1,89	
TENF0579	AA676058	Sem similaridade	2,19	
TENF0591	AA676066	Sem similaridade	3,00	
TENF0593	AA676068	Sem similaridade	2,46	
TENF0621	AA676089	Tc CL Brener cosmídio 1m17 cromossomo 3	1,67	
TENF0638	AA676103	Sem similaridade	1,81	
TENF0718	AI717740	Tc 24S β rRNA	2,65	
TENF0722	AI717743	Tc triptofanil-tRNA sintetase (TRYtRNA)	2,01	
TENF0749	AI717877	Sem similaridade	1,93	
TENF0775	AI717885	Sem similaridade	4,25	
TENF0852	AI717797	Sem similaridade	1,72	
TENF0857 °	AI717801	Tc maxicirculo do cinetoplasto - NADH desidrogenase subunidade 7 (ND7)	15,70	
TENQ0862	CF243331	Tc FL-160-2 gene, 3' UTR	1,95	

TENQ0877	CF243368	Sem similaridade	1,78	
TEUF0044 ^f	AA676163	Tc DNA de maxicírculo de cinetoplasto, clone TRCKPMAx	2,40	
TEUF0045 ^e	AA676164	Tc maxicírculo do cinetoplasto - NADH desidrogenase subunidade 7 (ND7)	25,29	confirmado
TEUF0050	AA676165	Tc gene mitocondrial HSP70 (MTP70)	2,56	
TEUF0067	AA426695	Tc proteína de choque térmico do retículo endoplasmático (grp78)	1,69	
TEUF0083	AA676173	Sem similaridade	2,26	
TEUF0084 ^f	AA676174	Tc DNA de maxicírculo de cinetoplasto, clone TRCKPMAx	2,40	
TEUF0102	AA676178	Proteína quinase 2	1,92	
TEUF0105	AA399706	<i>L. amazonensis</i> 40S proteína ribossômica S24e	4,12	
TEUF0146 ^e	AA676184	Tc maxicírculo do cinetoplasto - NADH desidrogenase subunidade 7 (ND7)	15,73	

^a A melhor similaridade obtida pelos programas BLASTN e/ou BLASTX ($E \leq 10^{-5}$)

^b Confirmação por Northern blot

^{c-f} Sequências pertencentes ao mesmo agrupamento ("cluster")

^gOMPDCase = orotidine-5'-fosfato descarboxilase (síntese de pirimidina)

^hOPRTase = orotate fosforribosil transferase (síntese de pirimidina)

Sequências diferencialmente expressas ($P \leq 0,01$ no teste *t*)

A análise dos dados da Tabela 8 indica que 7 sondas estão mais expressas no estágio epimastigota de Famema em relação a epimastigotas de Hem 79, enquanto que 11 sondas mostram situação inversa. Além disso, 33 sondas estão mais expressas no estágio tripomastigota metacíclico de Famema em relação ao mesmo estágio de Hem 179. Trinta e três sondas apresentam situação inversa. As listas das sondas diferencialmente expressas em epimastigotas e tripomastigotas metacíclicos dessas cepas são apresentadas nas Tabelas 8 e 9.

A análise conjunta dos dados mostra que 4 sequências coincidentes são diferencialmente expressas em ambos os estágios epimastigota e tripomastigota metacíclico de Famema, em relação aos mesmos estágios de Hem 179 (Tabela 10). Por outro lado, situação análoga é observada para 8 sondas coincidentes que estão mais expressas em ambos os estágios de desenvolvimento de Hem 179 quando comparadas com a cepa Famema. Essas sequências estão especificadas na Tabela 10, onde são apresentadas as razões de sinal obtidas para cada estágio de desenvolvimento.

Tabela 10 – Sequências coincidentes diferencialmente expressas em epimastigotas e tripomastigotas metacíclicas das cepas Famema (F) e Hem 179 (H).

<i>Expressão aumentada em Famema (n = 4 of 730; 0,54%)</i>				
Identidade	Nº de Acesso	Similaridade ^a	Razão epimastigota ^b F/H	Razão metacíclico ^b F/H
B13	U15616	Tc antígeno de 140/116 kDa	1,90	1,92
TENF0209	AA556049	Sem similaridade	2,06	1,88
TENF0753 °	AI717881	Tc proteína de ligação a RNA	2,96	6,24
TENQ0689 °	CF243287	Tc proteína de ligação a RNA	2,84	5,66
<i>Expressão aumentada em Hem 179 (n = 8 de 730; 1,09%)</i>				
Identidade	Nº de Acesso	Similaridade ^a	Razão epimastigota ^b H/F	Razão metacíclico ^b H/F
CL56 ^d	AY188964	DNA repetitivo de 195pb	2,61	3,11
CL65 ^d	Ay327543	DNA repetitivo de 195pb	3,31	3,31
G3 ^d	K01771	DNA repetitivo de 195pb	3,07	3,07
TENF0236	AA532144	Sem similaridade	2,60	2,34
TENF0857 °	AI717801	Tc maxicírculo do cinetoplasto - NADH desidrogenase subunidade 7 (ND7)	51,22	15,70
TEUF0045 °	AA676164	Tc maxicírculo do cinetoplasto - NADH desidrogenase subunidade 7 (ND7)	4,15	25,29
TEUF0105	AA399706	<i>L. amazonensis</i> 40S proteína ribossômica S24e	2,24	4,12
TEUF0146 °	AA676184	Tc maxicírculo do cinetoplasto - NADH desidrogenase subunidade 7 (ND7)	3,81	15,87

^a A melhor similaridade obtida pelos programas BLASTN e/ou BLASTX ($E \leq 10^{-5}$)

^b Estatisticamente significativa ($P \leq 0,01$ no teste *t*).

^{c-e} Sequências pertencentes ao mesmo agrupamento ("cluster")

Dentre as seqüências mais expressas em Hem 179 encontra-se a TENF0857, que apresenta uma razão de hibridização Hem 179/Famema de 51, para epimastigotas e 15,7, para tripomastigotas metacíclicas. Esta ETS apresenta elevada similaridade com o gene da subunidade 7 da NADH desidrogenase, codificado pelo DNA do maxicírculo do cinetoplasto (Tabela 10). É interessante notar que a TENF0857 pertence ao mesmo agrupamento onde encontram-se a TEUF0045 e a TEUF0146 que apresentam similaridade com a mesma seqüência e que também mostram aumento de expressão em Hem 179, principalmente em tripomastigotas metacíclicas (25 e 15 vezes, respectivamente). O alinhamento das seqüências depositadas em bancos de dados das três ESTs é mostrado na Figura 14. Na figura alinha-se também a região da ND7 da cepa Silvio x10.

Conforme comentado acima, a expressão do mRNA de ND7 foi confirmado em Northern blot contendo RNA total de epimastigotas e metacíclicas de Famema e Hem 179,

usando-se como sonda a TEUF0045 (Figura 13A e B). Os dados mostram claramente ausência do transcrito ou transcrição muito baixa deste RNA em ambos os estágios de Famema. Também foi investigada a presença deste mensageiro em diferentes isolados e estágios de CL Brener (Figura 13D). Os dados mostram ausência ou baixa transcrição de TEUF0045 em epimastigotas das cepas Y (canaleta 4), quando comparado com CL Brener (canaletas 1 e 2), Esmeraldo, Silvio e Dm28c (canaletas 5 a 7). Além disto, a abundância desse transcrito é muito baixa em tripomastigotas metacíclicos de CL Brener (canaleta 3) quando comparado com a forma epimastigota (canaletas 1 e 2).

Fig. 13 - Northern blot de RNA total de epimastigotas (A) e tripomastigotas metacíclicos (B) de Famema (F) e Hem 179 (H) hibridizados com sondas marcadas com P^{32} . Abaixo de cada painel a identificação da sonda, a razão de hibridização (F/H) no microarranjo (M) e no Northern Blot (N) são indicados. (C) Eletroforese do RNA total de parasitas corado com brometo de etídio. A região do gel onde os rRNAs 18S; 24S α e 24S β migram é mostrada. (D) Hibridização do Northern blot com sonda TEUF0045 marcada com P^{32} . Canaletas: CL Brener, epimastigotas em fase mid-log (1) e estacionária (2); tripomastigotas metacíclicos (3); epimastigotas de Y (4); Esmeraldo cl3 (5); Silvio X10 cl1 (6) e Dm 28C (7).

TEUF0146	-----	
ND7	AGAAAAGCTATGTGAATATAAAACAATAGAACAGTGTTTGCCGATTTTGATAGGTTTGA	60
TEUF0045	-----	
TENF0857	-----	
TEUF0146	-----	
ND7	ATGATTTGGGGAAGAACAGGCAGGAAAAGGCGGGACGAACGGGCAAGCGGAAGAGAAACC	120
TEUF0045	-----	
TENF0857	-----	
TEUF0146	-----	
ND7	TTTTCGGGAAGGAAGAACAGTTTAGGGGAGGGGACTCCAAAGACAGAGTTGAATTCGTTT	180
TEUF0045	-----	
TENF0857	-----	
TEUF0146	-----	
ND7	GATGGGGGACGGGGTTTTGGGGGCGAGCAAGCTTTTCAGGAGGGACTCGAGAGGGGGAA	240
TEUF0045	-----	
TENF0857	-----	
TEUF0146	-----AATTTCCGAGCGACGGGACGGATTGAG	27
ND7	AATGAAAACGGGGAGGAGATTTTTTGAGGGGGGAATTTCCGAGCGACGGGACGGATTGAG	300
TEUF0045	-----GTTTTCCGAGCGACGGGACGGCTTGAG	28
TENF0857	-----	
TEUF0146	ATT-GTTGATATTTTAGAAGGGGGCCCGCAAGNAGGAGGGATGTT-CGAAGGATTTGATT	85
ND7	ATTTGTTGATATTTTAGAAGGGGGCCCGCAAGA-GGAGGGATGTTTCGAAGGATTTGATT	359
TEUF0045	ATT-TTTAATATTTAAGAAGGGGGCCCGCAAGGAGGAGGGATGTC-CGAAGGATTTAATT	86
TENF0857	-----GATATTTGAGACCGGGGGCCCGCAAGGAGGAGGGATGTT-CGAAGGATTTGATT	52
	***** **	
TEUF0146	ATATATTTTAGGAGAAGGAATGGGAGCGGATTTTCGAACGAGAATTTTTTGCCGAGAGGA	145
ND7	ATATATTTTAGGAGAAGGAATGGGAGCGGATTTTCGAACGAGAATTTTTTGC-GAGAGGA	418
TEUF0045	ATATATTTAAGGAGAAGGAATGGAAGCGGATTTCCGAACGAAAATTTTTTCC-AAGAGGA	145
TENF0857	ATATATTTTAGGAGAAGGAATGGGAGCGGATTTTCGAACGAGAATTTTTTGC-GAGAGGA	111

TEUF0146	ATTTGTAAATTTCAAGGGGGTTTTTGGGGGGAAGAGGATTTTAAACCAGTTTAGACCCA	205
ND7	ATTTGTAAATTTCAAGGGGGTTTTTGGGGGGAAGAGGATTTTAAACC-AGTTTAGACCAA	477
TEUF0045	ATTTTAAATTTCAAGNNGGGTTTTTGGGGGGAAGAGGATTTTAAACCAGTTTAGACCAA	205
TENF0857	ATTTGTAAATTTCAAGGGGGTTTTTGGGGGGAAGAGGATTTTAAACCAGTTTAGACCTT	171
	**** *	
TEUF0146	AAGGC-AAGGGAGGGACCCGAATC-----TAGCTTGTTGGAACACCCCAAGTTTTT	255
ND7	AGG--AAGGGAGGGACCCGAATC-----TAGCTTGTTGGAACACCCCAAAGTTTT	525
TEUF0045	AGG--AAGGGAGGGACCCGGANTC-----TAGCTTGTTGGAAC-CCCAAGTTTTT	252
TENF0857	GTGCCTATGAGTAGTAACCATAGCGATGCCCGAACAGTTGGCACAGTGGCCGGGCAGTCG	231
	* * * * *	
TEUF0146	TTGGGA-----	261
ND7	TTGGATTTTGATTTTT--GAGTTGGT--GAGTTCC-----	557
TEUF0045	TTGGGATNTGAATTTATTGGAGTTGGGTGGAGTTCCCGACGACGGGCGGAGCATTTTAAT	312
TENF0857	GTACCGCAGAAATTTTATTGCTTTGTCAAACCCCTCCCTTCAGAAGGTGAGATCTTAA	291
	*	
TEUF0146	-----	
ND7	-----	
TEUF0045	GGAGAGGACGGG-----	325
TENF0857	AAGAACAACACCAGCTTATTACCCCATTCGCAACTAACT	330

Figura 14. Alinhamento das ESTs TEUF0045, TEUF0146, TENF0857 e da região ND7 da cepa Silvio x10 com o programa CLUSTAL W (Thompson *et al.*, 1994) disponível na página <http://www.ebi.ac.uk/clustalw/>. Asterisco significa identidade de nucleotídeos entre as quatro seqüências.

5.2. Transcrição diferencial de DNA satélite em cepas de *T. cruzi*

Os experimentos de microarranjos de DNA apresentados acima mostraram a expressão diferencial de um DNA satélite de 195 pb nas cepas CL Brener e Silvio x10 cl1 (Tabela 7) e nas cepas Famema e Hem 179 (Tabelas 8 e 9). Esta sequência, caracterizada na cepa Y (Gonzalez *et al.*, 1984; Sloof *et al.*, 1983a), está representada em cerca de 120.000 cópias e encontra-se organizada em agrupamentos de 30 ± 10 kb presentes em alguns cromossomos (Elias *et al.*, 2003). A função desta sequência não é conhecida, embora se acredita desempenhar algum papel estrutural nos cromossomos (Elias *et al.*, 2003).

Na lâmina do microarranjo imobilizamos DNA satélite de 195 pb da cepa Y (Acc AY188964) e da cepa Silvio X10 cl1 (Acc AY188964). A identidade entre as 2 sequências é de 89%. A análise dos resultados de hibridização com o programa ArrayVision indicou que o DNA satélite da cepa Y era expresso 6 vezes mais na cepa CL Brener do que em Silvio (Tabela 7). Para o DNA satélite de Silvio, observamos que sua expressão em CL Brener era 4 vezes maior que em Silvio (Tabela 7). A comparação da expressão em formas epimastigotas e metacíclicas de Hem 179 e Famema mostra razões de hibridização Hem 179/Famema de 2,6 a 3,3 vezes (Tabelas 8 e 9).

Na literatura há dados conflitantes em relação à transcrição da sequência satélite de 195pb. Por um lado, experimentos de Northern blot indicaram que esta sequência não é transcrita (Sloof *et al.*, 1983a), por outro, em experimentos de “run on” nuclear, foi observado que o satélite é ativamente transcrito (Elias *et al.*, 2001). Desta forma, visando confirmar os resultados do microarranjo, fizemos ensaios de Northern blot contendo RNA total de formas epimastigotas de CL Brener e Silvio. Na Figura 15 podemos observar que a sonda do DNA satélite de CL Brener hibridiza com populações de RNA de vários tamanhos desde 1,9 kb até 6,6 kb. Também nota-se que a hibridização da sonda com o RNA total de CL Brener é muito maior do que com o RNA de Silvio, confirmando o resultado do microarranjo (Tabela 7). A sonda de GAPDH foi usada para normalização dos dados de Northern blot.

Figura 15- Northern blot de RNA total de epimastigotas de CL Brener (C) e Silvio (S) hibridizado com com sonda do DNA satélite da cepa Y. Abaixo das amostras indica-se a razão da hibridização no microarranjo e no Northern blot. A hibridização do Northern blot com sonda de GAPDH foi usada para normalização dos dados.

Com o objetivo de verificar se este resultado não era ocasionado por uma contaminação do RNA com DNA genômico, tratamos o RNA de CL Brener, alternativamente, com NaOH, RNase ou DNase e realizamos novos experimentos de Northern blot (Figura 16). Notamos que a preparação de RNA controle e aquela tratada com DNase apresentam hibridização com a sonda de DNA satélite (Figura 16A e B canaletas, 1 e 2). Por outro lado, RNAs tratados com RNase ou NAOH, agentes que degradam RNA, não apresentaram hibridização com a sonda do satélite (Figura 16A e B, canaletas 3 e 4). Estes dados sugerem que a sequência satélite é transcrita.

Figura 16 - Northern blot de RNA total de CL Brener submetido a distintos tratamentos e hibridizado com sondas de DNA satélite da cepa Y (A) e da cepa Silvio (B). Canaletas: (1) RNA controle; (2) tratado com DNase; (3) tratado com RNase; (4) tratado com NaOH.

Para garantir a completa eliminação de uma possível contaminação por DNA, a preparação de RNA de CL Brener foi submetida a dois tratamentos com DNase. Os dados apresentados na Figura 17 indicam que nos RNAs tratados uma ou duas vezes com DNase a sonda de DNA satélite continua apresentando hibridização. Na amostra tratada duas vezes com DNase a hibridização é mais fraca, provavelmente devido à perda de RNA durante o processo (ver coloração com brometo de etídio, Figura 17, canaleta 3). Por outro lado, nas amostras de RNA não tratado com DNase mas tratado com RNase, ou tratado duas vezes com DNase e uma vez com RNase não se observa hibridização com a sonda satélite (Figura 17, canaletas 4 e 5).

Figura 17 - Northern blot de RNA total de CL Brener submetido a distintos tratamentos. (A) Padrão de coloração com brometo de etídio das três espécies de rRNA. (B) Hibridização com sonda de DNA satélite da cepa Y. Canaletas: (1) RNA controle; (2) tratado com DNase uma vez; (3) tratado com DNase duas vezes; (4) tratado com RNase; (5) tratado com DNase duas vezes e RNase uma vez.

Visando verificar se o padrão de transcrição diferencial da sequência de DNA satélite era uma característica estágio-específica, a sonda satélite da cepa Y foi hibridizada com RNA total de CL Brener isolado de formas epimastigotas em fase mid-log e em fase estacionária de crescimento e de formas tripomastigotas metacíclicas. Também foram analisadas preparações de RNA de formas epimastigotas das cepas Esmeraldo, Silvio e

Dm28c (Figura 18). Todos os RNAs foram tratados com DNase. Os dados mostram que os RNAs isolados dos estágios de CL Brener e de epimastigotas das cepas Y e Esmeraldo apresentam forte hibridização com a sequência satélite (Figura 18, canaletas de 1 a 5). Por outro, lado os RNAs isolados das cepas Silvio e Dm28c apresentam sinal muito fraco de hibridização (Figura 18, canaletas 6 e 7).

Figura 18- Northern blot de RNA de várias cepas e estágios evolutivos, hibridizado com sonda de DNA satélite da cepa Y. Todos os RNAs foram tratados com DNase. Canaletas: CL Brener epimastigotas, fase de crescimento mid-log (1); estacionária (2); formas metacíclicas (3). Epimastigotas das cepas Y (4); Esmeraldo (5); Silvio (6) e Dm 28c (7).

Tomados em seu conjunto, os resultados indicam que a sequência satélite é transcrita diferencialmente em cepas de *T. cruzi* II (CL Brener, Y e Esmeraldo) em relação a *T. cruzi* I (Silvio e Dm 28c). Estudos estão em andamento no laboratório para melhor investigar este aspecto.

Decidimos verificar se esta sequência estava representada nas ESTs de CL Brener depositadas nos bancos de dados. Para isto, realizamos uma busca de similaridade no banco de dados de ESTs do NCBI, que contém ESTs de todos os organismos exceto de humanos e rato (www.ncbi.nlm.nih.gov). A busca revelou que apenas a EST de CL Brener (TENU4403) continha a sequência satélite (valor de similaridade de $4e^{-84}$). Verificamos que nesta EST a sequência satélite estava localizada na região 3' do mRNA. A TENU4403 apresenta similaridade com uma sequência de DNA repetitiva do minicromossomo de *T. cruzi* com um valor de $E=0,0$.

5.3 Expressão diferencial de genes em seis cepas de *T. cruzi* isoladas de indivíduos com distintas manifestações da doença de Chagas

Tendo verificado que lâminas de microarranjos contendo predominantemente ESTs de CL Brener podiam ser usadas em análises de genômica comparativa e de expressão diferencial de genes em isolados *T. cruzi*, passamos a investigar a diferença de expressão gênica em três cepas isoladas de indivíduos que apresentam a forma indeterminada (VL10, Famema e Berenice 62) e três cepas isoladas de pacientes com a forma cardíaca da doença de Chagas (115, B13 e 147) (ver Tabela 1). Cinco destas cepas foram isoladas em Minas Gerais e uma em São Paulo (Famema). Para essa análise, utilizamos o conjunto B de lâminas de microarranjos de DNA contendo 755 ESTs e 32 genes de várias cepas de *T. cruzi*.

5.3.1. Caracterização dos alvos de DNA para o conjunto B de lâminas de microarranjos

Para a construção do conjunto B de microarranjos, 755 clones de uma biblioteca normalizada e não normalizada de cDNA da forma epimastigota de CL Brener e de amastigotas da cepa Tulahuen foram aleatoriamente selecionados (710 dessas seqüências tinham sido imobilizadas no conjunto A). A busca de similaridade com os programas BLASTN e BLASTX indicou que 70% das ESTs (532/755) não têm similaridade no banco de dados. Também foram imobilizados na lâmina insertos representando 32 seqüências clonadas de *T. cruzi* e ESTs que fazem parte de grupos sintênicos (Vargas *et al.*, 2004). Entre as seqüências com similaridade no banco de dados, as categorias funcionais mais representadas são as mesmas encontradas no conjunto A de lâminas com uma pequena variação nas porcentagens de algumas categorias: seqüências relacionadas à síntese protéica (24%); metabolismo celular (18%); proteínas de superfície celular (15%) e elementos estruturais e do citoesqueleto (13%). (Gráfico 2).

Gráfico 2: Classificação funcional das seqüências contidas no conjunto B de microarranjos

O agrupamento das 787 seqüências foi feito com o programa Cap 3, indicando 713 (90%) seqüências únicas (Tabela 11). Tendo em vista que na lâmina aplicamos também 10 ESTs que não conseguimos seqüenciar, essa ultima versão da lâmina de microarranjo continha aproximadamente 800 seqüências de *T. cruzi*.

O mapa detalhado da imobilização do conjunto B de lâminas pode ser encontrado no CD anexo.

Tabela 11. Agrupamento de 787 seqüências de *T. cruzi* por Cap 3

Agrupamentos ("clusters")	Número de agrupamentos ("clusters")
Com 2 a 6 seqüências	57
Singletons	657
Total de seqüências únicas	714

5.3.2. Construção dos microarranjos e obtenção de imagens

As 800 sondas, representando 787 seqüências de *T. cruzi* e DNAs controles, foram distribuídas em 10 sub-arranjos. Cada subarranjo foi imobilizado em duplicata (lado direito e lado esquerdo da lâmina) e cada sonda foi imobilizada no mínimo em seis e no máximo

em trinta réplicas por lâmina, dependendo da sonda. Os dados de hibridização foram extraídos com o programa ArrayVision 6.0. Para eliminar viés de marcação, a normalização dos dados foi feita por Lowess (Dudoit *et al.*, 2002) através do gráfico de dispersão MA (Yang *et al.* 2002), conforme especificado acima. Na Figura 19 apresentamos como exemplo de normalização os dados de hibridização obtidos com o par de cepas VL10 e B13. Réplicas com baixa intensidade ($A \leq 2$) e com intensidades saturadas ($A \geq 9.5$) foram excluídas.

Figura 19. Normalização dos dados de hibridização da lâmina com cDNA da cepa VL 10 e da cepa B13. (A) razão de intensidade não normalizada. (B) razão normalizada por Lowess.

5.3.3. Expressão diferencial de genes em seis isolados de pacientes chagásicos

RNA total foi isolado da fase mid-log de crescimento de formas epimastigotas de seis cepas isoladas de pacientes chagásicos. Para reduzir o número de experimentos, as hibridizações competitivas foram feitas usando-se o cDNA da cepa VL10 (derivada de um paciente assintomático) como referência. Desta forma, foram feitas seis combinações: VL10 x VL10; VL10 x 115; VL10 x 147; VL10 x B13; VL10 x Famema e VL10 x Berenice 62.

O RNA foi isolado de duas culturas independentes de cada cepa, o que representa duas réplicas biológicas. O RNA de cada réplica biológica foi hibridizado em uma lâmina gerando no mínimo 12 e no máximo 60 réplicas de cada gene (réplicas técnicas), dependendo da sonda.

5.3.4. *Análise dos Resultados*

A análise dos resultados foi direcionada para identificar genes que apresentam aumento ou diminuição de expressão em **todas** as cepas de cardíacos em relação às cepas de assintomáticos. Nesta análise utilizamos três métodos: as análises de correlação de Pearson e de correlação Spearman e o método “Bayes Error Rate” ou Erro de Classificação de Bayes (Duda *et al.*, 2000). As análises de Pearson e de Spearman foram usadas para medir as correlações entre os padrões de expressão de **cada** sequência utilizando-se o valor M ($M = \log_2 \text{Cy5/Cy3}$) normalizado por Lowess de todas as réplicas da sequência e a classificação das cepas em cardíacas ou assintomáticas. O coeficiente de correlação de Pearson foi usado anteriormente na análise de dados de microarranjos para medir correlações entre o padrão de expressão gênica em câncer de próstata e características dos pacientes (como idade e presença de PSA no soro) (Singh *et al.*, 2002). Mais recentemente, este coeficiente também foi utilizado para medir a correlação entre níveis de transcritos específicos em amostras de tumor de próstata com diferentes valores de Gleason (Reis *et al.*, 2004). Nas análises de correlação, utilizamos o valor de corte de $|\rho|=0,75$ para aceitar genes diferencialmente expressos nos dois grupos de cepas. Nas Tabelas B e C que constam no CD anexo, apresentamos os valores de ρ obtidos nas análises de correlação de Pearson e de Spearman das sequências que foram aceitas como diferencialmente expressas.

Após a definição das sequências que apresentam valores de ρ aceitos pelos dois métodos de correlação, aplicamos o método “Bayes Error Rate” ou Erro de Classificação de Bayes (Duda *et al.*, 2000). Este método permite verificar se **as réplicas** de cada gene estão, de fato, separadas nos dois grupos de cepas. A separação ou sobreposição dos dois grupos é indicada pelo valor (E). Ou seja, quanto mais próximo de zero for o valor de E , maior será a separação entre os dois grupos. Conseqüentemente, quanto mais próximo de 1 for o valor de E , maior será a sobreposição entre os grupos, indicando não haver variação de expressão

gênica entre eles. Neste caso, definimos o valor de corte para o Erro de Classificação de Bayes de $E=0,05$. Salientamos que esta definição é bastante rígida como será discutido adiante. A abordagem “Bayes Error Rate” foi utilizada na análise de resultados de SAGE de tumor de cérebro em trabalho realizado no Instituto Ludwig para Pesquisa sobre o Câncer (Vêncio *et al.*, 2004) ver também na página www.vision.ime.usp.br/~rvencio/SAGEbetaBin/bayes_error_rate.htm).

A utilização conjunta dos três métodos permitiu identificar 14 sondas que são diferencialmente expressas nos dois grupos de cepas (Tabelas 12 e 13). Nas Tabelas, para cada sonda diferencialmente expressa apresentamos o valor de E (definido pelo “Bayes Error Rate”) e a razão de expressão R em relação à cepa de referência VL10 ($R = 2^{\bar{M}}$), onde $\bar{M}$ é a média dos valores de M de todas as réplicas técnicas e biológicas.

Tabela 12: Genes diferencialmente expressos em três cepas de cardíacos em relação às três cepas de assintomáticos.

Seqüências com expressão aumentada nas cepas de cardíacos em relação às cepas de assintomáticos (n=10 sondas de 800; 1,25%)									
Nome	Número de acesso	Similaridade ^a	Valor de E ^b	Razão de expressão em relação à cepa de referência VL10 ^c					
				VL/VL	Fam/VL	Be/VL	115/VL	B13/VL	147/VL
TENF0857	AI717801	Subunidade 7 da NADH desidrogenase.	0	1,2	1,2	1,0	31,5	17,5	30,5
B6-PI08_F10	CB923575	<i>T. brucei</i> aspartato aminotransferase -	0	1,0	1,3	1,1	2,7	2,2	3,3
TENF0305	AA532186	Sem similaridade	0	ND	1,1	1,2	4,7	1,7	2,0
TENF0641	AA676105	<i>T. brucei</i> gene da proteína ribossômica L1	0	ND	1,2	1,0	4,0	1,5	1,5
TENQ0823	CF243322	Sem similaridade	0,02	1,0	1,1	1,2	1,9	1,4	1,6
A4-PI04_C06	CB923896	<i>Tc</i> triparedoxina	0,02	1,3	1,1	1,3	1,7	1,7	2,1
TENF0593	AA676068	Sem similaridade	0,03	1,1	1,1	1,0	2,3	1,5	1,8
TENF0426	AA532101	Sem similaridade	0,04	1,0	1,0	1,1	1,5	1,2	1,8
TENU2842	AI069865	<i>L. major</i> proteína ribossômica 40S S13	0,04	1,7	1,1	1,0	1,8	1,6	1,3
TENF0228	AA532138	Sem similaridade	0,05	1,1	1,4	1,1	2,7	4,3	2,2

Seqüências com expressão reduzida em todas as cepas de cardíacos quando comparadas com as cepas de assintomáticos (n=4 sondas de 800; 0,5%)									
Nome	Número de acesso	Similaridade ^a	Valor de E ^b	Razão de expressão em relação à cepa de referência VL10 ^c					
				VL/VL	VL/Fam	VL/Be	VL/113	VL/B13	VL/147
TENF0103	AA525713	Sem similaridade	0,01	1,0	1,1	1,0	1,8	1,6	1,5
TENF0853	AI717798	Sem similaridade	0,03	1,1	1,0	1,1	1,6	1,4	1,4
D2-PI01_E01	CB923590	<i>L. donovani</i> ciclofilina	0,04	1,0	1,0	1,0	1,6	1,5	1,5
TENF0280	AA532175	Sem similaridade	0,04	1,0	1,1	1,2	1,5	1,5	1,5

^a - A melhor similaridade obtida pelos programas BLASTN e/ou BLASTX ($E \leq 10^{-5}$)

^b E= Erro de Classificação de Bayes (Duda *et al.*, 2000)

^c Razão de expressão R em relação à cepa de referência VL10 ($R = 2^{\bar{M}}$), onde $\bar{M}$ é a média dos valores de M de todas as réplicas técnicas e biológicas.

ND-dado não disponível porque as réplicas apresentaram na hibridização VL10 x VL10 valores de intensidades abaixo do nosso valor de corte ($A \leq 2$) (ver Material e Métodos, Item 4.20).

Para ilustrar a análise realizada, na Figura 20 apresentamos os gráficos MA e a determinação do valor de E pelo método do Erro de Classificação de Bayes para seis seqüências que constam da Tabela 12. Nestes gráficos as réplicas das cepas de assintomáticos e cardíacos são apresentadas com duas cores diferentes.

Na Figura 21, apresentamos o gráfico MA com os resultados normalizados das réplicas da TENF0857 indicando com cores diferentes os valores obtidos para as três cepas de cardíacos e de assintomáticos. Valores de M positivos indicam que o gene está sendo mais expresso nas cepas de cardíacos. Pode-se notar uma dispersão relativamente baixa dos valores de M dentro do mesmo grupo de cepas.

Na Figura 22, apresentamos o mesmo tipo de análise para todas as réplicas da sonda TENF0103 que, neste caso apresenta redução de expressão em todas as cepas de cardíacos em relação às cepas de assintomáticos. Neste caso, a dispersão dos valores das réplicas é maior.

Gráficos MA para as outras seqüências diferencialmente expressas podem ser encontrados no CD anexo.

A análise dos dados da Tabela 12 indica que a expressão de 10 sondas está aumentada nas 3 cepas isoladas de pacientes cardíacos quando comparado com a expressão nas 3 cepas isoladas de pacientes assintomáticos. Por outro lado, observa-se que a

expressão de 4 sondas está reprimida em todas as cepas de pacientes cardíacos quando comparado com as cepas de pacientes assintomáticos. Os valores da razão de expressão R em relação à cepa de referência VL10 das sondas com expressão aumentada nas cepas de cardíacos variam bastante entre as seqüências. Assim, para a sonda TENF0857 obtivemos valores de R de 17 a 31 (nas três cepas de cardíacos); para a sonda TENF0305, valores de 1,7 a 4,7 e para a sonda A4-l04_C06, valores de 1,7 a 2,1. Isto reflete a característica da sonda em estudo e a variação interna no grupo de cepas de cardíacos, esperada em função da diversidade genética das mesmas. É interessante notar que a variação dos valores de R é menor para as sondas com expressão reduzida nas cepas de cardíacos em relação às cepas de assintomáticos (Tabela 12).

Figura 20. Gráficos MA e a determinação do valor de E pelo método do Erro de Classificação de Bayes para seis seqüências que constam da Tabela 12. Resultados normalizados das réplicas das cepas isoladas de pacientes com cardiopatia chagásica crônica estão representados em vermelho e com as cepas isoladas de indivíduos assintomáticos em verdes.

Figura 21. (A) Gráfico MA com os resultados normalizados das réplicas da TENF0857 para cepas isoladas de pacientes com cardiopatia chagásica crônica (quadrados vermelho, marron e rosa) e com as cepas isoladas de indivíduos assintomáticos (quadrados verdes, amarelos e azul claro). Valores de M positivos indicam genes induzidos em relação à cepa VL10. M=0 indica que não houve expressão diferencial.

Figura 22. (A) Gráfico MA com os resultados normalizados das réplicas da sonda TENF0103, para cepas isoladas de pacientes com cardiopatia chagásica crônica (quadrados vermelho, marron e rosa) e com as cepas isoladas de indivíduos assintomáticos (quadrados verdes, amarelos e azul claro).

A expressão diferencial de 11 genes identificados pelo microarranjo foi investigada por hibridização em membranas de Northern blot contendo RNA das cepas de cardíacos e assintomáticos analisadas neste estudo. Também foi colocado na membrana RNA total da cepa Hem 179 isolada de paciente com manifestações cardíacas e digestivas (Tabela 1). As Figuras 23, 24 e 25 mostram o padrão de hibridização dessas sondas e o padrão de migração dos rRNAs 18S, 24S α e 24S β corados com brometo de etídio. Na Figura 23F apresentamos o padrão de hibridização da sonda correspondente ao gene de actina 2, um “house keeping gene”, que não mostra variação de expressão no microarranjo. O sinal de hibridização de todas as sondas foi quantificado por análise densitométrica com o programa ImageQuant Molecular Dynamics. Para a sonda de actina 2 (Figura 23F), atribuímos o valor unitário para o menor valor de hibridização (cepa Berenice 62) e calculamos o valor relativo dos sinais de hibridização nas outras cepas. Estes valores foram utilizados para a normalização dos valores de hibridização de cada sonda em cada cepa, determinados por análise densitométrica. Abaixo de cada canaleta indica-se a razão de hibridização normalizada do RNA de cada cepa em relação à cepa VL10. Isto foi feito para seguir o critério utilizado na análise dos dados do microarranjo, onde os pares de hibridizações sempre incluíram o cDNA de VL10 como referência.

Na Figura 23B observa-se a presença de um transcrito de 1,1 kb, correspondente à TENF0857, apenas nas cepas de cardíacos. Esta observação confirma os dados do microarranjo que indicam que este mRNA é de 17 a 30 vezes mais expresso em cepas de cardíacos em relação às cepas de assintomáticos (Tabela 12). Esta sequência possui alta similaridade com o gene da subunidade 7 da NADH desidrogenase codificado pelo DNA de maxicirculo.

A sonda B6-PI08_F10, que possui alta similaridade com o gene da aspartato aminotransferase de *T. brucei*, apresenta um padrão de hibridização interessante. Transcritos de 1,7 kb são observados nas três cepas de assintomáticos (VL10, Famema, e Berenice 62) ao passo que em duas cepas de cardíacos (115 e 147) observa-se um transcrito de 1,1 kb. Por outro lado, nas cepas de cardíaco B13 e na cepa Hem 179 (forma cardíaca e digestiva) ambos os transcritos são observados (Figura 23C). A análise densitométrica dos sinais de hibridização (computando-se os dois transcritos, quando presentes) mostra pequenas variações nas razões de hibridização em relação à cepa VL 10 (de 0,6 a 1,4). Por

outro lado, pelos microarranjos, a razão de expressão da sonda nas três cepas de cardíacos é em média 2,8 vezes maior em relação às cepas de assintomáticos (Tabela 12).

A expressão da sonda TENF0305, que não possui similaridade no banco de dados, é em média 2,7 vezes maior nas cepas de cardíacos quando comparado com as cepas de assintomáticos. (Tabela 12). O experimento de Northern blot (Figura 23D) indica que o mRNA da sonda TENF0305 está mais expresso em todas as cepas de cardíacos, apesar da diferença de expressão nas cepas 147, B13 e Hem 179 ser pequena.

O resumo dos dados da análise dos resultados de Northern blot para as sondas mais expressas em cardíacos é apresentado na Tabela 13.

Três das quatro sondas que apresentam expressão reduzida em cepas de cardíacos em relação às cepas de assintomáticos foram analisadas por Northern blot (Figura 25 e Tabela 13). O padrão de hibridização mostrado na Figura 25B indica que o nível do transcrito de 2 kb reconhecido pela sonda TENF0103, que não possui similaridade no banco de dados, não varia entre as cepas.

A sonda D2-PI01_E01, que possui similaridade com uma ciclofilina de *Leishmania donovani*, hibridizou com três transcritos (ver Figura 25C e Tabela 13) nas cepas. A integração das áreas de hibridização permitiu concluir que estes transcritos são menos abundantes nas cepas de cardíacos e em Hem 179. Situação análoga é observada para os transcritos da TENF0853, que não possui similaridade no banco de dados. O padrão de hibridização é mostrado na Figura 25E e indica a presença de dois transcritos de 2,4 e 1,3 kb nas cepas de assintomáticos e em Hem 179. A banda de 1,3 kb não é visível nas cepas de cardíacos 115, 147 e B13. Na cepa 115 observa-se além do transcrito de 2,4 kb, um mRNA de 0,9 kb (Figura 25E).

Os dados das análises por Northern blot são resumidos na Tabela 13. Em alguns casos, estes dados confirmam os resultados do microarranjo. Em outros casos, as diferenças das razões de hibridização são pequenas. Por outro lado, os padrões de hibridização para algumas sondas mostram diferenças no tamanho dos transcritos entre as cepas que podem determinar diversidades no processo de expressão gênica. Isto determina que estudos mais aprofundados devam ser conduzidos para elucidar este aspecto.

Figura 23– Northern blot hibridizado com as sondas com expressão induzida nas cepas de cardíacos comparadas com as cepas de assintomáticos. A) O gel corado com brometo de etídio indica a região onde migram os rRNAs 18S; 24S α e 24S β . B) sonda TENF0857, C) sonda B6Pl08F10, D) sonda TENF0305, E) sonda TENF0641, F) sonda de actina 2. A seta do lado direito indica o tamanho dos transcritos em kb. Abaixo das canaletas da hibridização com a sonda actina 1 está indicado o valor relativo de hibridização em relação ao valor de Berenice 62. Essas razões foram usadas para normalizar os valores de hibridização das sondas nas Figuras 23 a 25. Abaixo da canaleta de hibridização de cada sonda está indicada a razão de hibridização em relação ao valor de VL 10.

Figura 25 – Northern blot hibridizado com as sondas com expressão reduzida nas cepas de cardíacos comparadas com as cepas de assintomáticos. A) e D) gel corado com brometo de etídio indica a região do gel onde migram os rRNAs 18S; 24S α e 24S β . B) sonda TENF0103, C) sonda D2-PI01_E01, E) sonda TENF0853. A seta do lado direito indica o tamanho dos transcritos em kb. Abaixo de cada canaleta esta indicada razão de hibridização em relação ao valor de VL10, após a normalização.

Tabela 13: Análise dos transcritos das seqüências identificadas pelos microarranjos como diferencialmente expressas em cepas de cardíacos em relação às cepas de assintomáticos.

Seqüências com expressão aumentada nas cepas de cardíacos em relação às cepas de assintomáticos (n=10 sondas de 800; 1,25%)					
Nome	Número de acesso	Similaridade ^a	Análise por Northern blot ^b	Transcrito (kb) ^c	Observação
TENF0857	AI717801	Subunidade 7 da NADH desidrogenase.	Confirma	1,1	Presente só em cepas de cardíacos
B6-PI08_F10	CB923575	<i>T. brucei</i> aspartato aminotransferase -	Confirma	1,7 e 1,1	Variação do tamanho dos transcritos
TENF0305	AA532186	Sem similaridade	Confirma	1,0	Variação de expressão
TENF0641	AA676105	<i>T. brucei</i> gene da proteína ribossômica L1	Não confirma	1,3	Variação de expressão aleatória
TENQ0823	CF243322	Sem similaridade	Confirma	0,95	Variação de expressão
A4-I04_C06	CB923896	<i>Tc</i> triparedoxina	Confirma	1,4 e 1,0	Variação de expressão
TENF0593	AA676068	Sem similaridade	Confirma	0,9	Variação de expressão
TENF0426	AA532101	Sem similaridade	Não analisado	ND	ND
TENU2842	AI069865	<i>L. major</i> proteína ribossômica 40S S13	Não analisado	ND	ND
TENF0228	AA532138	Sem similaridade	Confirma	1,7 e 0,4	Variação da intensidade dos transcritos
Seqüências com expressão reduzida em todas as cepas de cardíacos quando comparadas com as cepas de assintomáticos (n=4 sondas de 800; 0,5%)					
TENF0103	AA525713	Sem similaridade	Não confirma	2,0	Não varia
TENF0853	AI717798	Sem similaridade	Confirma	2,4; 1,3 e 0,95	Variação de expressão e do tamanho dos transcritos
D2-PI01_E01	CB923590	<i>L. donovani</i> ciclofilina	Confirma	2,6 e 1,3 1,0	Variação de expressão e do tamanho dos transcritos
TENF0280	AA532175	Sem similaridade	Não analisado	ND	ND

^a A melhor similaridade obtida pelos programas BLASTN e/ou BLASTX ($E \leq 10^{-5}$)

^b Confirma variação entre as cepas dos dois grupos

^c Tamanho do transcrito identificado no Northern blot

ND-dado não disponível

5.4. *Análise da expressão gênica de fibroblastos humanos infectados com duas cepas de T. cruzi (Estágio no Exterior)*

Conforme descrito na Introdução, Vaena de Avalos et al. (2002) analisaram a expressão de genes em fibroblastos humanos (HFF) infectados com a cepa Y de *T. cruzi*. Nesse estudo, utilizando lâminas com 27.000 cDNAs humanos, foi observado que *T. cruzi* não induz expressão de genes nas primeiras horas de infecção (2, 4 ou 6 h). Após 24 horas, foi observada uma alteração significativa na abundância de transcritos de HFF, com indução de cerca de 106 genes. Alguns resultados foram confirmados por Northern blot (Vaena de Avalos et al., 2002).

Julgamos interessante iniciar experimentos-piloto para investigar a resposta transcricional de células humanas à infecção com cepas de *T. cruzi* isoladas de pacientes com distintas manifestações da doença de Chagas. Para este fim, nosso laboratório iniciou uma colaboração com o Dr. David Munroe (Diretor do Laboratório de Tecnologia Molecular do Instituto Nacional do Câncer do NIH, em Frederick, EUA) que disponibilizou a infraestrutura do Centro e microarranjos de genes humanos.

Dessa forma, entre outubro de 2003 e janeiro de 2004, realizei um estágio no NCI-SAIC. Em função da limitação de tempo, decidimos fixar alguns parâmetros: (i) fibroblastos humanos, como célula hospedeira; (ii) a cepa VL10, isolada de indivíduo assintomático e a cepa 147 isolada de indivíduo com cardiopatia severa, como cepas infectantes; e (iii) 24 horas pós-infecção, como tempo de análise. A expressão diferencial de genes nas duas cepas havia sido analisada nos experimentos descritos no item 5.3. Ambas as cepas provém de Minas Gerais.

5.4.1. *Obtenção de formas infectantes e infecção de monocamadas*

Em nosso laboratório no IQ, realizamos todos os experimentos de infecção de monocamadas de fibroblastos (ver Material e Métodos). Formas epimastigotas de ambas as cepas foram diferenciadas em tripomastigotas metacíclicos em meio LIT contendo 10% de meio Grace (Zingales *et al.* 1997a). Os metacíclicos foram utilizados para infectar monocamadas de fibroblastos humanos. A partir do oitavo ao décimo dia pós-infecção,

tripomastigotas e amastigotas começaram a eclodir das células. Estas formas foram utilizadas nas infecções posteriores das monocamadas. Verificamos o fenômeno interessante de que a cepa 147 originava quase exclusivamente tripomastigotas, ao passo que para a cepa VL10 amastigotas eram predominantemente recuperados no sobrenadante das culturas. Este fato ilustra diferenças biológicas entre as cepas.

Os tripomastigotas da cepa 147 (25×10^6 de parasitas) e a mistura de amastigotas e tripomastigotas de VL10 (15×10^6 de parasitas) foram incubados com os fibroblastos por 24h.

5.4.2. Isolamento de RNA

Após 24 horas de infecção, as monocamadas foram lavadas com PBS e o RNA total foi extraído. Ao mesmo tempo, células humanas foram cultivadas com meio condicionado de parasitas (MCP) de ambas as cepas (ver Material e Métodos), também por 24h. Após esse período o RNA dos fibroblastos também foi extraído, assim como o de fibroblastos não infectados. Esta última preparação foi usada como RNA de referência para todas as hibridizações dos microarranjos. A qualidade do RNA foi analisada em gel agarose 0,8%. Dois experimentos independentes para cada cepa, foram realizados caracterizando duas réplicas biológica. A Figura 26 mostra o padrão de migração do RNA das amostras após coloração com brometo de etídio. Observa-se a integridade do rRNA 18S e 28S nas amostras extraídas de fibroblastos incubados com MCP de 147; infectados com VL10 e não infectados (Figura 26 canaletas 2 a 4). Por outro lado, em fibroblastos infectados com a cepa 147 (Figura 26, canaleta 1) verificamos integridade do rRNA 18S e ausência do rRNA 28S. Já na amostra extraída de fibroblastos cultivados com MCP da cepa VL10 (Figura 26, canaleta 5), o RNA está ausente provavelmente por ter havido degradação durante a purificação. Por conseguinte, esta preparação não foi analisada subsequentemente. Também não foi hibridizada com o microarranjo a preparação de RNA obtida dos fibroblastos incubados com MCP da cepa 147.

Figura 26: Gel de agarose das amostras de RNA isoladas de fibroblastos infectados com a cepa 147 (1); e com VL10 (3); incubados com meio condicionado (MCP) da cepa 147 (2) e da cepa VL10 (5) e de fibroblastos não infectados (4). Coloração com brometo de etídio. As setas indicam a posição dos rRNAs 28S e 18S.

5.4.3. Hibridização com microarranjos

As amostras de RNA foram tratadas com DNase e cDNA foi sintetizado, marcado com os fluoróforos Cy3 e Cy5. Experimentos de hibridização competitiva foram realizados, utilizando como referência o cDNA das células não infectadas. Duas réplicas técnicas foram realizadas para cada par de amostras (“dye swap”). A lâmina de microarranjo utilizada continha oligonucleotídeos representando um total de 24.000 genes humanos e foi produzida no NCI-SAIC.

Para a análise dos dados, seguimos o procedimento adotado no Centro, utilizando os programas e “softwares” disponíveis. Após normalização dos dados por Lowess, filtramos os genes que apresentavam redução ou indução de expressão nas células infectadas com a cepa 147, em relação às células não infectadas. Também analisamos os resultados de genes que apresentavam redução ou indução de expressão nas células infectadas com a cepa VL10.

Na Tabela 14 relacionamos 28 sondas diferencialmente expressas nas células não infectadas e infectadas com a cepa 147 (paciente com a forma cardíaca). Destas, 27 sondas estão com expressão reduzida nas células infectadas, ao passo que apenas uma sonda (heme oxigenase) apresentou indução nas células infectadas.

Em estudo análogo realizado com a cepa VL10 (paciente assintomático), identificamos 17 sondas diferencialmente expressas nas células infectadas em relação a células não infectadas (Tabela 15). Destas, 16 sondas estão com expressão induzida nas células infectadas, enquanto que apenas uma sonda (Proteína 5 ligante do fator de crescimento do tipo insulina) mostra situação inversa.

Tabela 14. Sondas^a diferencialmente expressas em fibroblastos humanos infectados (FHNI) com a cepa 147 (paciente com a forma cardíaca) por 24 horas em relação a fibroblastos não infectados (FHN)

Nome do gene	Razão de hibridização FHN/FHNI
<i>Genes com expressão reduzida em fibroblastos infectados</i>	
cytochrome P450, family 7, subfamily B, polypeptide 1	3,2
leucine rich repeat containing 4	1,5
cytochrome c oxidase subunit VIc	1,8
synaptotagmin III	2,1
spectrin, beta, non-erythrocytic 1	1,9
hypothetical protein FLJ13456	1,8
upstream regulatory element binding protein 1	1,8
chromosome 14 open reading frame 127	1,7
Homo sapiens clone FLB9131 PRO2459 mRNA	1,7
Homo sapiens mRNA; cDNA DKFZp586I041 (from clone DKFZp586I041)	1,7
Kruppel-associated box protein	1,7
deiodinase, iodothyronine, type I	1,6
DEAD (Asp-Glu-Ala-Asp) box polypeptide 46	1,7
alanine-glyoxylate aminotransferase 2	1,6
protease, serine, 15	1,6
Chromosome 20 open reading frame 186	1,6
ubiquitin specific protease 34	1,6
a disintegrin and metalloproteinase domain 10 ADAM10	1,5
iduronate 2-sulfatase (Hunter syndrome)	1,6
solute carrier organic anion transporter family, member 1B3	1,5
Putative prostate cancer tumor suppressor	1,8
hypothetical protein FLJ1236	1,6
hypothetical protein LOC116064	1,6
Homo sapiens, clone IMAGE:3916023, mRNA, partial cds	1,5
Homo sapiens mRNA; cDNA DKFZp434O0919	1,5
hypothetical protein FLJ32011	1,5
hypothetical protein MGC2793	1,5
<i>Genes com expressão aumentada em fibroblastos infectados</i>	
heme oxygenase (decycling) 1	0,25

^aNota: preferimos elencar as sondas com sua nomenclatura em inglês

Tabela 15. Sondas^a diferencialmente expressas em fibroblastos humanos infectados (FHNI) com a cepa VL10 (paciente assintomático) por 24 horas em relação a fibroblastos não infectados (FHN).

Nome do gene	Razão de hibridização FHNI/FHN
<i>Genes com expressão aumentada em fibroblastos infectados</i>	
heme oxygenase (decycling) 1	5,0
Homo Homo sapiens similar to hypothetical protein DKFZp434A171 (LOC340012), mRNA	1,5
ESTs, Weakly similar to FRHUL ferritin light chain [H.sapiens]	1,6
chemokine (C-C motif) ligand 26	1,6
meprin A, alpha (PABA peptide hydrolase)	1,6
SH3 domain binding glutamic acid-rich protein like 3	1,6
alkaline phosphatase, intestinal	1,7
prolyl endopeptidase	1,7
ATPase, H ⁺ transporting, lysosomal 9kDa, V0 subunit e	1,7
ribosomal protein S9	1,6
Homo sapiens cDNA: FLJ21436 fis, clone COL04279	1,8
cyclin D1 (PRAD1: parathyroid adenomatosis 1)	2,0
potassium inwardly-rectifying channel, subfamily J, member 1	2,5
FLJ20758 protein	1,5
thioredoxin domain containing 5	1,6
cysteine-rich motor neuron 1	1,8
<i>Genes com expressão reduzida em fibroblastos infectados</i>	
insulin-like growth factor binding protein 5	0,37

^aNota: preferimos elencar as sondas com sua nomenclatura em inglês

5.4.4. Validação dos dados do microarranjo

Para validar alguns resultados obtidos no microarranjo, a expressão diferencial de alguns genes foi analisada por ensaio de RT-PCR (real time-PCR). O padrão obtido para duas dessas sondas é mostrado na Figura 27. A primeira fita de cDNA foi sintetizada utilizando-se o Omniscript kit da Qiagen. O PCR em tempo real foi realizado através do sistema ABI 7900 Sequence Detection System (Applied Biosystems). A análise dos dados foi feita através do ABI 7900 software e a normalização foi feita utilizando-se valores

encontrados com o gene de rRNA 18S humano. A razão final dos corantes para cada gene foi calculada em relação a fibroblastos não infectados.

Na Figura 27A, a análise do RT-PCR indica que o gene da heme oxigenase apresenta indução de expressão nas células infectadas com a cepa 147 de 80 vezes. Valores desta magnitude também foram encontrados no RNA extraído de células incubadas com MCP de 147 (150 vezes) e no RNA extraído de fibroblastos infectados com VL10 (60 vezes). Os resultados obtidos no microarranjo são bem mais baixos. Observa-se uma indução de 4 vezes nas células infectadas com 147 (Tabela 14) e de 5 vezes nas células infectadas com a cepa VL10 (Tabela 15).

Na Figura 27B a análise do RT-PCR indica que o gene da Proteína 5 ligante do fator de crescimento do tipo insulina apresenta repressão de expressão de 6 vezes nas células infectadas com VL10, confirmando o resultado de microarranjo que demonstrou repressão de 2,7 vezes desse gene (Tabela 15).

Figura 27 - PCR em tempo real com oligonucleotídeos correspondendo aos genes da heme oxigenase (A) e gene proteína 5 ligante do fator de crescimento do tipo insulina (B). Em A amostras de RNA isoladas de células não infectadas (1); células infectadas com a cepa 147 (2); células cultivadas com MCP de 147 (3) células infectadas com a cepa VL10 (4). Em B amostras de RNA isoladas de células não infectadas (1); células infectadas com a cepa VL10 (2). Valores de razão positivos indicam que o gene está induzido em relação ao controle. Valores de razão negativos indicam que o gene está reprimido em relação ao controle/referência. Todas as comparações foram feitas utilizando as células não infectadas como referência.

6. DISCUSSÃO

6.1. Construção e validação dos microarranjos de DNA para o estudo da expressão diferencial de genes e de genômica comparativa em cepas de *T. cruzi*.

6.1.1. Caracterização das ESTs

Com o objetivo de validar microarranjos de DNA para a análise da expressão diferencial de genes de genômica comparativa em *T. cruzi*, solicitamos aos Drs. A. Brandão e W. Degraive da FIOCRUZ cerca de 700 ESTs obtidas de bibliotecas normalizadas e não normalizadas de cDNA da forma epimastigota de CL Brener. Ambas bibliotecas foram construídas no laboratório do Dr. E. Rondinelli (UFRJ) e o inserto foi clonado direcionalmente (detalhes em Urmenyi *et al.*, 1999). A biblioteca normalizada foi construída no vetor pT7T3-Pac (Pharmacia) e o inserto é flanqueado pelos promotores T7 e M13 Forward, no lado 5' do cDNA, e promotores T3 e M13 Reverse, no lado 3' (Urmenyi *et al.*, 1999). A biblioteca não normalizada foi construída no vetor pSport 1 (Invitrogen) e o inserto é flanqueado pelo promotor T7, no lado 5' do cDNA, e pelo promotor SP6 no lado 3' (Urmenyi, comunicação pessoal).

As ESTs possuem em média 800 pb e suas seqüências, depositadas em bancos de dados, foram obtidas principalmente a partir da extremidade 5' do cDNA e referem-se, em média, a cerca de 300 pb. Os pesquisadores da FIOCRUZ nos cederam 710 clones de ESTs, sendo que 661 provinham de uma biblioteca de cDNA normalizada (anotadas como TENF) e 49 de uma biblioteca não normalizada (anotadas como TEUF).

A seqüência nucleotídica estava disponível para 576 ESTs. Desta forma, seqüenciamos as 134 ESTs restantes e depositamos as seqüências no dbEST do GenBank. Utilizamos três abordagens para a obtenção do DNA das ESTs para o seqüenciamento: (a) minipreparação do DNA plasmidial por lise alcalina a partir de colônias transformadas; (b) amplificação do inserto por PCR de colônia e (c) amplificação direta a partir dos plasmídios recebidos (diluição 1:50). Este último método é mais vantajoso, por ser mais rápido. Dentre as 134 ESTs, 90% foram seqüenciadas a partir da extremidade 5' do cDNA (iniciador T7) e somente 10% foram seqüenciadas a partir da extremidade 3' do cDNA (iniciador T3). O

seqüenciamento a partir desta extremidade é mais problemático devido à presença da cauda de poli A.

6.1.2. Anotação das ESTs

A anotação das ESTs foi feita inicialmente com seqüências individuais utilizando-se o próprio NCBI para fazer as buscas de similaridade. Implementamos no laboratório dois métodos para fazer um BLAST semi-automático. A primeira abordagem consistiu em fazer “download” do programa BLAST e de alguns bancos como “nr” e “DBEST” em nosso computador e tentar fazer as buscas de similaridade somente utilizando nossos computadores. Essa abordagem é mais difícil pelo fato de os nossos computadores não terem memória suficiente para toda essa operação. De qualquer forma, ter o programa BLAST instalado no nosso computador foi muito útil, pois pudemos analisar nossas seqüências contra pequenos bancos criados por nós. Nesses bancos pudemos colocar seqüências de vetor, sítios de clonagem, etc, e, através do BLAST, “limpar” as seqüências. A outra abordagem foi a utilização de um serviço que o próprio NCBI disponibiliza, onde podemos enviar todas as seqüências que se deseja analisar ao NCBI e a resposta da análise de similaridade é obtida no dia seguinte. Esta abordagem é claramente mais funcional. Outra ferramenta que utilizamos é o programa Phred para a análise da qualidade das seqüências. Utilizamos como padrão de qualidade um índice maior ou igual a 20, de acordo com o programa Phred. Esse programa está disponível na página <http://lbm.fmvz.usp.br/>

O conjunto dos dados de similaridade das 134 ESTs obtido pelo BLASTX e BLASTN indica que 48% das seqüências (n=64) apresenta similaridade em bancos de dados (36% com seqüências de *T. cruzi* e 12%, com seqüências de outros organismos). Entre as seqüências com similaridade com genes de *T. cruzi*, encontramos histonas, mucinas, quinases, desidrogenases, e outras classes de genes. Observamos também que 70 seqüências (52%) não têm similaridade com seqüências depositadas, podendo ser específicas de *T. cruzi*. Porcel *et al.* (2000) analisaram 3.054 seqüências de CL Brener, encontrando, com os programas BLASTN e BLASTX, 33% de similaridade e 67% de não similaridade com seqüências depositadas em bancos de dados. Verdun *et al.* (1998), que analisaram 1.949 ESTs de CL Brener, obtiveram a mesma porcentagem de seqüências sem

similaridade. O aumento de similaridade de nossas seqüências, comparado com os trabalhos citados acima (de 33% para 48%) pode ter ocorrido pelo: a) aumento do número de seqüências depositadas nos bancos; b) análise mais cuidadosa dos nossos dados por ter sido realizada por inspeção visual de cada seqüência e c) o baixo número de seqüências analisadas (n=134), que pode tornar o estudo mais tendencioso. Essa última hipótese é sustentada pelo fato de que quando analisamos todas as ESTs colocadas nas lâminas, 710 EST no conjunto A e 755 no conjunto B, encontramos 75% e 70% de não similaridade respectivamente. Esses valores são muito mais próximos aos encontrados por Verdun *et al.* (1998) e Porcel *et al.* (2000).

As duas categorias funcionais mais representadas nas 730 seqüências do conjunto A de lâminas são: proteínas envolvidas na síntese protéica (24%) e em vias metabólicas (16%). Porcel *et al.* (2000) e Verdun *et al.* (1998) encontraram, respectivamente, 24% e 23,3% de seqüências envolvidas em síntese protéica, concordando com nossos resultados. Em relação a proteínas envolvidas no metabolismo, Porcel *et al.* (2000) observaram 24% de similaridade, enquanto Verdun *et al.* (1998) obtiveram 7,9%. Isso poderia ocorrer devido à particularidade das seqüências analisadas ou diferenças de classificação de determinados genes que podem ser colocados em categorias diferentes.

Outra categoria interessante encontrada em nossas 730 seqüências foi a de genes que codificam antígenos de superfície celular. Verdun *et al.* (1998) encontraram 10,9% de similaridade com essa categoria, enquanto nós encontramos um valor de similaridade para essa categoria de 14%. Vale lembrar que essa categoria de genes é interessante uma vez que os produtos gênicos podem estar relacionados com a infectividade e morfologia do parasita. A título de exemplo: De Jesus *et al.* (1993) reportaram que a deleção do gene que codifica a proteína de superfície GP72 de *T. cruzi*, por integração dirigida, causa uma drástica mudança morfológica do parasita, ocasionando a perda da região de ligação do flagelo à membrana da célula. Ruiz *et al.* (1998) descreveram que em cepas do grupo *T. cruzi* II (Y e CL) não há expressão de gp90. Estas cepas infectam muito bem células de mamíferos em cultura, ao contrário de cepas de *T. cruzi* I (Dm28 e G), que expressam a gp90 e infectam pouco monocamadas de células. Lembramos que cepas de *T. cruzi* II predominam no ciclo doméstico e estão associadas à doença de Chagas, ao passo que o grupo *T. cruzi* I

predomina no ciclo silvestre e, em humanos, determinaria manifestações clínicas mais brandas (Zingales *et al.*, 1998).

6.1.3. Outros genes utilizados na construção do microarranjo

Para a construção do conjunto A e B de lâminas além das 710 ESTs, incluímos 20 seqüências de várias cepas do parasita, que representam genes que codificam proteínas, RNAs e regiões não-codificadoras (ver Tabela 2).

A seguir descreveremos algumas características destas seqüências.

No microarranjo imobilizamos três genes de rRNA: 18S da cepa G e da cepa Y e 24S α da cepa Y. Os genes que codificam os rRNAs 18S, 24S α e 24S β e alguns rRNAs de pequeno tamanho (rRNAs S1-S6) estão agrupados em um mesmo segmento de DNA de aproximadamente 30 kb denominado cistron ribossômico (Hernandez *et al.*, 1988). Existem cerca de 120 cópias do cistron ribossômico por genoma haplóide, dispostas em tandem e separadas por regiões denominadas espaçadores intergênicos (Castro *et al.*, 1981). Em algumas cepas os genes podem estar em uma única banda cromossômica, indicando um único *locus* contendo várias cópias do cistron ou podem estar em dois *loci* situados em dois cromossomos, que podem ser considerados cromossomos “homólogos” de tamanhos diferentes (Pedroso *et al.*, 2003). O inserto contendo a seqüência do rRNA 24S α foi purificado do plasmídio que o continha e utilizado como sonda em membrana contendo o DNA de bandas cromossômicas de várias cepas (Figura 5). Observamos que a sonda hibridiza com bandas de tamanho molecular diferente nas cepas, ilustrando a variabilidade genética das mesmas, mas elevada similaridade de seqüência.

Uma seqüência interessante colocada no microarranjo é derivada do gene da RNA helicase da família das DEAD-box, uma vez que foi anteriormente demonstrado que a transcrição deste gene é 8 a 10 vezes maior em tripomastigotas metacíclicos de CL Brener em relação a epimastigotas (Diaz-Añel *et al.*, 2000). Para obter a seqüência, clonamos e seqüenciamos um fragmento da extremidade 3' terminal do gene da RNA helicase. Um BLAST contra seqüências do banco de dados forneceu uma similaridade de 99% com o próprio gene da RNA helicase depositado no banco. A sonda clonada foi hibridizada com um Southern blot de DNA de várias cepas de *T. cruzi* digerido com *Pst* I (Figura 6). Tendo

em vista que o gene da RNA helicase é um gene de cópia única por genoma haplóide (Díaz-Añel *et al.*, 2000), esperávamos encontrar hibridização com apenas uma banda de 2,7 kb na cepa CL Brener, já que esse é o tamanho do fragmento obtido por digestão com a enzima *Pst* I (Díaz-Añel *et al.*, 2000). No entanto, encontramos duas bandas, a banda esperada de 2,7 kb e uma outra de 1,15 kb. Uma vez que *T. cruzi* é um organismo diplóide, é possível que as duas bandas de hibridização correspondam a cada um dos "alelos" do gene da RNA helicase. Diferenças de sequência em alelos de genes de *T. cruzi* já foram demonstradas (Henriksson *et al.*, 1990). De fato, dados de mapeamento do gene da RNA helicase nos cromossomos de CL Brener indicam hibridização com uma banda de 2,6 Mpb e outra de 0,83 Mpb, que conteriam cada um dos alelos (Díaz-Añel *et al.*, 2000). Nos experimentos realizados por nós confirmamos a hibridização com ambas as bandas cromossômicas de CL Brener, apesar de a banda de 2,6 Mpb ter apresentado um sinal de fraca intensidade e de difícil reprodução na foto (Figura 7). Observamos também hibridização do gene da RNA helicase com bandas de diferentes tamanhos moleculares nas diferentes cepas (Figura 7).

Outra sequência colocada no microarranjo, designada como mini exon (ME), codifica o RNA da sequência líder (exon e intron) e a região intergênica não transcrita. A sequência líder (exon) possui 39 pb e está presente na extremidade 5' de todos os mRNAs de tripanossomatídeos. O gene que codifica o ME apresenta-se em arranjo similar aos RNAs ribossômicos. Existem cerca de 200 cópias do gene, arranjadas em tandem e localizadas em um único *locus* (revisto em Brener *et al.*, 2000).

O gene B11 codifica uma proteína com similaridade com transcriptase reversa e corresponde a um retrotransposon. Há cerca de 2.000 cópias por genoma haplóide (Martin *et al.*, 1995), que estão distribuídas em diferentes cromossomos (Vargas *et al.*, 2004).

Genes que codificam vários antígenos de superfície também foram colocados na lâmina incluindo: o gene que codifica o antígeno imunodominante B13 (140/116 KDa), expresso apenas na superfície da forma tripomastigota (Gruber & Zingales, 1993) e o gene do antígeno B12 (200 kDa) (Gruber & Zingales, 1993); genes que codificam a glicoproteína de superfície GP82 (estágio-específica de formas tripomastigotas metacíclicas) e a GP85 (gp85/sialidase), ambos genes de múltiplas cópias. A GP85 pertence à família II da super-família das sialidasas/gp85. Essa família reúne glicoproteínas de

massas moleculares de 70-90 kDa, presentes na superfície das formas amastigota e tripomastigota e sem atividade enzimática. Tanto a Gp85 quanto a Gp82 estão envolvidas na penetração do parasita na célula hospedeira (Teixeira & Yoshida, 1986). As sequências depositadas na lâmina correspondem às regiões das ORFs que possuem os motivos Asp Box e Fn3 (revisto em Brener *et al.*, 2000). A Gp82, por ser estágio-específica, é um controle interessante para a análise da expressão diferencial de genes entre formas epimastigotas e metacíclicas. No entanto, pode ocorrer hibridização cruzada entre o cDNA da Gp82 e da GP85 por causa das regiões de Asp Box e Fn3 presentes em ambos os genes (Araya *et al.*, 1994).

A proteína flagelar 1F8 que liga cálcio é uma proteína de 24 kDa que pode estar envolvida na mobilidade do parasita. Existem várias cópias deste gene no genoma de *T. cruzi* (Engman *et al.*, 1989).

O gene de actina 2, que foi depositado na lâmina, corresponde a um “house keeping gene” que, em teoria, seria expresso continuamente. O gene do fator de alongação 1 codifica uma proteína que tem um papel central na biossíntese de proteínas em eucariontes. No caso de *T. cruzi* pode também estar envolvido na apoptose (Billaut-Mulot *et al.*, 1996). O gene da SLA (Spliced leader associated RNA) é um membro da classe dos “small nuclear RNAs” que está envolvido na modificação de RNAs ribossômicos. Existem mais de 10 cópias deste gene por genoma haplóide (Roberts *et al.*, 1998).

Uma sequência que também foi colocada na lâmina é a região telomérica de tripanossomatídeos. Esta consiste de repetições em tandem da sequência de 6 pb - 5'ccctaa3'. O DNA telomérico está associado a uma série de proteínas, e esses complexos são importantes na manutenção da integridade cromossômica, evitando sua degradação. Esta sequência não é expressa (Blackburn & Challoner, 1994). Também foi aplicada a sequência sub-telomérica de *T. cruzi* que possui sequências homólogas aos genes das famílias das gp85/sialidase importantes para a patogênese do parasita (Chiurillo *et al.*, 1999).

Imobilizamos sequências repetitivas de DNA, como o elemento C6, representado por 1000 cópias por genoma haplóide (Araya *et al.*, 1997), e a sequência de 195 pb (minisatélite) (Gonzalez *et al.*, 1984), isolada das cepas Silvio e Y. O número de cópias

estimado para os minisatélites é de 120.000. Estas cópias estão organizadas em tandem e estão localizadas em vários cromossomos (Elias *et al.*, 2003).

Os genes comentados acima fazem parte dos conjuntos A e B das lâminas construídas. Os genes discutidos abaixo fazem parte somente do conjunto B.

Tripanotiona redutase é uma enzima chave na proteção de tripanossomatídeos contra os efeitos letais do estresse oxidativo. Seu papel importante no metabolismo de peróxidos foi mostrado através de “knockout” condicional desse gene em células de *T. brucei*. Após o “knockout” as células passaram a apresentar hipersensibilidade a estresse oxidativo (Krieger *et al.*, 2000).

Diidrofolato redutase (DHFR)-timidilato sintetase (TS) é uma proteína bifuncional monomérica envolvida na síntese de 2'-desoxitimidilato. O domínio de DHFR é encontrado na região amino terminal e o domínio TS é encontrado na região carboxila do polipeptídeo (Reche *et al.*, 1994).

A proteína reparadora de "mismatch" de DNA é uma enzima envolvida no reparo de DNA. Esse gene é constitutivamente expresso em todas fases de vida de *T. cruzi* e está representado no genoma desse organismo como um gene de cópia única (Augusto-Pinto *et al.*, 2001).

O transportador de hexose que imobilizamos na lâmina faz parte da superfamília dos transportadores de glicose. Esse gene é similar a transportadores de glicose de *T. brucei* e de *Leishmania* e no genoma de *T. cruzi* é representado em oito seqüências repetitivas em tandem (Tetaud *et al.*, 1994).

A proteína de choque térmico 70 é codificada pelo gene hsp70 que está organizado em pelo menos 10 seqüências repetitivas em tandem no genoma de *T. cruzi*. Essa seqüência possui grande homologia com a seqüência de hsp70 de *T. brucei* e de outros organismos (Requena *et al.*, 1988).

Outros dois antígenos também foram imobilizados no conjunto B: antígeno H49 e JL8. O antígeno H49 é um antígeno imunodominante com repetições seriadas de aminoácidos associado ao citoesqueleto, com peso molecular de 300 kDa e participa da ligação do corpo celular com o flagelo (Cotrim *et al.*, 1995). O antígeno JL8 é outro antígeno repetitivo de *T. cruzi* (Levin *et al.*, 1989).

6.1.4. Agrupamento das seqüências depositadas nas lâminas

Realizamos um agrupamento ("clusterization") das 730 seqüências do conjunto A com o programa Cap 3 e verificamos a presença de 51 agrupamentos com 2 a 5 seqüências cada e 614 "singletons" (agrupamentos com uma única seqüência), o que perfaz um total de 665 (84%) seqüências únicas. Para o conjunto B, o agrupamento das 787 seqüências mostrou 57 agrupamentos com 2 a 6 seqüências e 657 "singletons", o que representa 713 (90%) seqüências únicas. Além disto, no conjunto B também aplicamos 10 ESTs que não conseguimos seqüenciar. Desta forma, este conjunto contém 797 seqüências de *T. cruzi*.

O mapa detalhado dos conjuntos A e B de lâminas pode ser encontrado no CD anexo.

6.1.5. Estudo de genômica comparativa e expressão diferencial de genes em Silvio e CL Brener

Nota: Os estudos relatados nos itens 6.1.5 e 6.1.6 estão descritos no artigo Baptista *et al.*, (2004) que pode ser consultado em formato pdf no CD anexo.

Tendo em vista a grande diversidade biológica das cepas de *T. cruzi*, é esperado que a comparação genômica entre isolados forneça dados importantes, por exemplo, para estudos de epidemiologia molecular. Todavia, a comparação genômica baseada na seqüência nucleotídica de várias cepas, não é viável. Como alternativa, a comparação genômica entre microorganismos tem sido investigada com o auxílio de hibridizações em microarranjo (Koide *et al.*, 2004; Oliveira *et al.*, 2002). Nesse contexto, estudos pioneiros analisaram diferenças entre *Mycobacterium tuberculosis*, *M. bovis* e cepas de BCG e identificaram várias regiões de cromossomos e ORFs que estavam deletadas nas cepas atenuadas (Behr *et al.*, 1999). Mais recentemente, utilizando-se microarranjos de DNA humano, amplificações e deleções de regiões gênicas, resultantes de anomalias cromossômicas, foram identificadas na linhagem HL-60 de tumor mielógeno, quando comparada com leucócitos normais (Ulger *et al.*, 2003). No presente estudo, nós avaliamos a possibilidade de utilizarmos microarranjos de DNA contendo predominantemente ESTs de CL Brener para estudos de genômica comparativa e expressão gênica em isolados de *T. cruzi*.

Para a análise de genômica comparativa, escolhemos cepas dos grupos *T. cruzi* I (Silvio) e *T. cruzi* II (CL Brener). O uso de métodos estatísticos rigorosos ao invés de simplesmente relatar variações na razão de hibridização está se tornando um padrão para a análise de experimentos de microarranjo. Recentemente, o uso do valor *P* do teste-*t* permitiu a identificação, com elevada confiança, de vários genes que apresentam expressão regulada em menos de duas vezes durante a diferenciação de formas procíclicas para metacíclicas em *L. major* (Saxena *et al.*, 2003). Nós usamos o critério de corte $P \leq 0,01$ na análise estatística dos dados do microarranjo para definir hibridizações diferenciais com os alvos de cDNA ou DNA.

A hibridização da lâmina com o DNA genômico de Silvio e CL Brener mostrou que todas as sondas imobilizadas hibridizavam com ambos os DNAs. Entretanto, 37 seqüências (37/730; 5,0%) deram maior sinal com o DNA de Silvio, enquanto 31 sondas (31/730; 4,2%) apresentaram situação inversa. Esse resultado sugere diferenças genômicas entre os isolados, que podem refletir variação no número de cópias gênicas e/ou variação na similaridade da seqüência. A hibridização do DNA das duas cepas com a sonda representada pelo DNA genômico fragmentado de CL Brener mostrou razão CL Brener/Silvio de 1,4. Esse resultado pode ser devido à abundância de seqüências repetitivas em CL Brener e/ou diferenças de similaridade de seqüência. Em favor da primeira hipótese, observamos que a sonda de DNA satélite de 195 pb, a seqüência reiterada mais abundante e que corresponde a 10% do genoma de *T. cruzi*, está com com razão CL Brener/Silvio de 6,5. Esses resultados são consistentes com conclusões anteriores de nosso grupo, via hibridização de cromossomos e quantificação do número de cópias, de que o DNA satélite é 6 vezes mais abundante em CL Brener quando comparado com Silvio (Elias *et al.* 2003).

O cistron de RNA ribossômico dos tripanossomatídeos é constituído pelos genes de rRNA 18S, 24S α e 24S β e genes de RNA pequenos. O cistron é transcrito em precursores de RNA de alto peso molecular pela RNA polimerase I sensível a α -amanitina (revisto por Clayton, 1988). Os dados do microarranjo sugerem aumento de 3 a 5 vezes no número de unidades de cistrons ribossômicos em Silvio quando comparado com CL Brener. De fato, razões de hibridização Silvio/CL Brener de 2,8 a 5,8 foram obtidas para sondas representando os genes ribossômicos 18S (três sondas), 24S α (uma sonda) e 24S β (uma sonda) (Tabela 6). Esse aumento concomitante na razão de hibridização das três espécies de

rRNA é coerente com sua organização na unidade de transcrição do cistron. Esse dado e o fato de as seqüências de rRNA 18S de Silvio e CL Brener apresentarem 99% de similaridade apóiam a conclusão do aumento do número de cópias das unidades de rDNA em Silvio e indicam a confiabilidade dos dados do microarranjo. Apesar disso, não observamos aumento de transcritos de rRNA em Silvio, quando comparado com CL Brener (experimento de hibridização com cDNA). Esse fato pode ser resultante de diferenças na atividade dos promotores de rRNA e eficiência dos fatores de transcrição nas duas cepas. De fato, foram reportadas divergências na atividade das regiões promotoras de rDNA em cepas de *T. cruzi* (Floeter-Winter *et al.*, 1997, Nunes *et al.*, 1997, Stolf *et al.*, 2003). Assim, o aumento no número de cópias do rDNA em Silvio poderia refletir um fenômeno de resposta compensatória. A consistência entre as conclusões derivadas da análise de microarranjos em relação a DNA satélite e genes de rRNA e dados previamente publicados sugerem que microarranjos são adequados para análises de genômica comparativa em *T. cruzi*.

Por outro lado, para a sonda designada como miniexon (Tabela 6), que representa o gene de seqüência líder (exon e intron) e a região intergênica não transcrita, observamos uma razão de hibridização do DNA Silvio/CL Brener de 21,52 ($P \leq 0,01$) (Tabela 6). Nesse caso, a diferença de hibridização muito provavelmente é resultante da diferença de similaridade de seqüência. De fato, o alinhamento da seqüência dos exons de 39 pb e introns de 73 pb do gene de miniexon de quatro cepas indicou aproximadamente 98% de identidade, enquanto somente 57-59% de identidade foi observada entre as regiões intergênicas que variam de 484 a 494pb (Souto *et al.*, 1996). Para Silvio e CL Brener, nós verificamos que a similaridade das regiões é de 57,9%. Não foi verificada variação de expressão na sonda de miniexon entre Silvio e CL Brener.

Várias sondas com hibridização diferencial foram confirmadas por experimentos de Southern blot com DNA genômico dos dois isolados (Figura 11A). Uma das sondas, TENF0426, que no microarranjo apresenta razão de hibridização CL Brener/Silvio = 2,7, foi usada como sonda em Southern blots com DNA de nove isolados (Fig. 11B). Verificamos diferenças na quantidade dessa seqüência no genoma dos isolados, mas não observamos correlação entre a abundância da seqüência TENF0426 e o grupo ao qual os

isolados pertencem (*T. cruzi* I ou *T. cruzi* II). O padrão de hibridização sugere que TENF0426 seja um DNA repetitivo que ainda não foi caracterizado em *T. cruzi*.

Uma vez que a lâmina de microarranjo contém predominantemente sondas de CL Brener, a análise de genômica comparativa é unidirecional e representa somente um subconjunto das diferenças entre os isolados. Ou seja, seqüências somente encontradas no genoma de Silvio não serão evidenciadas em nossos arranjos. Contudo, essa abordagem permite a comparação genômica entre CL Brener e outros isolados, e entre diferentes cepas. Essa informação é valiosa para os projetos de seqüenciamento e para estudos da diversidade genética em *T. cruzi*.

Para validar os microarranjos para estudos de expressão diferencial de genes, a lâmina do conjunto A foi hibridizada com cDNA de CL Brener e Silvio. A hibridização com a sonda contendo DNA de CL Brener fragmentado mostrou razão CL Brener/Silvio de 1,1, sugerindo que o experimento era válido. Observamos que 11,5% das sondas (84 de 730) estavam diferencialmente expressas em epimastigotas de fase mid-log de CL Brener e Silvio e apresentavam uma razão de sinal de 1,6 a 6,5 (Tabela 7). A comparação dos dados de hibridização dos microarranjos com o DNA genômico e com cDNA das cepas CL Brener e Silvio permitiu concluir que não há uma correlação entre a abundância de determinado gene e seu nível de expressão. De fato, somente 20% das sondas com aumento no nível de expressão em Silvio ou CL Brener apresentaram hibridização diferencial com DNA genômico. Em alguns casos, verificamos que seqüências com maior hibridização com o DNA genômico de uma cepa eram mais expressas na outra. Essa observação sugere controle na expressão gênica e/ou diferenças de similaridade de seqüência. Ensaio de “run-on nuclear” e experimentos realizados em epimastigotas permeabilizados com detergente mostraram que a atividade de transcrição não é diretamente proporcional ao número de cópias gênicas (Elias *et al.*, 2001). De fato, a seqüência de satélite de 195pb, que apresenta em torno de 10^5 cópias no genoma da cepa Y (Sloof *et al.*, 1983a; Gonzalez *et al.*, 1984), é transcrita somente 10^3 vezes mais que os genes de cópia única. O mesmo é verificado para genes de múltiplas cópias que codificam mucinas e glicoproteínas de 85 kDa (Elias *et al.*, 2001). Os autores sugerem que algumas cópias desses genes abundantes podem estar localizadas em domínios nucleares silenciosos tais como heterocromatina.

Vários dos genes diferencialmente expressos foram confirmados por Northern blot. Observamos que as razões de hibridização no microarranjo eram quase sempre menores que as razões de hibridização determinadas por análise densitométrica nos experimentos de Southern e Northern blots. Essa discrepância foi verificada anteriormente e pode ser causada pela competição entre os dois alvos fluorescentes, assim como pela não linearidade dos sistemas de detecção (cf Dihel *et al.*, 2002).

6.1.6. Expressão diferencial de genes em dois isolados de pacientes chagásicos

A expressão diferencial de genes foi investigada nas cepas Famema e Hem 179, isoladas de pacientes chagásicos que apresentam a forma indeterminada da doença de Chagas (Famema) e alterações cardíacas e digestivas (Hem 179). Observamos que aproximadamente 2,5% (18/730) das sondas apresentaram expressão diferencial nas formas epimastigotas das duas cepas e que 9,0% (66/730) das sondas mostraram razões de hibridização diferentes nas formas infectantes (tripomastigotas metacíclicas). Não comparamos a expressão gênica entre formas epimastigotas e tripomastigotas metacíclicas de uma mesma cepa, pois 97% das sondas do microarranjo foram originadas de formas epimastigotas. A análise da identidade das sondas diferencialmente expressas indica que seqüências que pertencem a um mesmo agrupamento gênico apresentam mesmo comportamento de expressão (Tabelas 8, 9 e 10). Esse fato também foi observado nos genes com expressão induzida em CL Brener e Silvio. Esses dados mostram uma consistência interna nos dados do microarranjo e que os resultados são confiáveis.

Outra indicação do controle de expressão gênica em *T. cruzi* foi obtida quando analisamos a expressão dos genes B12 e B13. Nosso grupo, com base na análise de seqüência de clones genômicos, havia concluído previamente que o gene B13, que codifica um antígeno de superfície de 140/116 kDa, é contíguo ao gene B12, que codifica um antígeno de 230/200 kDa (Gruber & Zingales, 1993; Gruber, 1994). Os autores também determinaram que esses genes estão separados por 1,5 kb e são transcritos na mesma direção, sugerindo que os genes B13 e B12 podem ser transcritos em um RNA policistrônico (Gruber & Zingales, dados não publicados). Dados do microarranjo indicam que somente o gene B13 é induzido em ambas as formas epimastigota e tripomastigota

metacíclico de Famema com uma razão de hibridização Famema/Hem 179 = 1,90 (Tabela 10), enquanto o gene B12 é induzido somente em epimastigotas de Hem 179 (Hem/Famema = 1,89, Tabela 8). Esse dado sugere diferenças na estabilidade dos dois transcritos e/ou regulação na expressão gênica nos estágios de desenvolvimento. A expressão diferencial dos genes de amastina e tuzina no ciclo celular de *T. cruzi* foi reportada. Os dois genes são contíguos e transcritos em um mesmo pré-RNA policistrônico (Teixeira *et al.*, 1994; 1995). A análise por Northern blot mostrou que o mRNA de amastina é 60 vezes mais abundante em amastigotas quando comparado com epimastigotas e tripomastigotas e que os mRNAs de tuzina são mantidos em um nível constante nos estágios de desenvolvimento (Teixeira *et al.* 1994; 1995).

Entre as 40 sondas mais expressas em epimastigotas e tripomastigotas metacíclicos de Famema, quando comparado com Hem 179, 4 sondas coincidentes foram reconhecidas pelos cDNAs de ambos os estágios de desenvolvimento. Entre elas estão uma EST que não apresenta similaridade no banco de dado, duas ESTs com similaridade com proteínas que se ligam ao RNA e o gene B13 (Tabela 10).

Oito sondas são mais expressas em epimastigotas e tripomastigotas metacíclicos de Hem 179 quando comparado com Famema. Entre essas sondas está a sequência de DNA satélite de 195 pb. Esse resultado muito provavelmente reflete diferenças do número de cópias no genoma dos dois isolados, como observado para CL Brener e Silvio (ver acima e Elias *et al.* 2003) (Tabela 10).

Três ESTs (TENF0045, TENF0857, TEUF0146), que estão mais expressas em Hem 179 quando comparado com Famema, apresentam alta similaridade com o gene que codifica a subunidade 7 da NADH desidrogenase (ND7) da cepa Silvio X10. Estes genes pertencem ao mesmo agrupamento gênico, representando um controle interno da confiabilidade dos resultados. A hibridização de Northern blots com TEUF0045 indica ausência ou nível muito baixo desse transcrito em epimastigotas de cepa Y, quando comparado com epimastigotas de quatro isolados, onde um mRNA de 1,1 kb foi detectado (Fig. 13D). Além disso, o transcrito de RNA de ND7 foi detectado em epimastigotas de CL Brener e está ausente ou muito menos abundante em tripomastigotas metacíclicos (Fig. 13D). O gene ND7 está localizado no DNA de maxicírculo e será discutido em mais detalhes na seção 6.3.

Nesse estudo, observamos que duas ESTs (TENF0753 e TENQ0689), pertencentes ao mesmo agrupamento gênico e que apresentam similaridade com cDNA que codifica proteínas que se associam a rRNA 5S e à família de fosfoproteínas nucleares previamente descritas em *T. brucei* e *T. cruzi* (Li *et al.*, 2003, DaRocha *et al.*, 2002), são mais intensamente expressas em Famema em relação a Hem 179, tanto em epimastigotas (razão de hibridização 2,9) quanto em tripomastigotas metacíclicos (razão de hibridização 5,6 a 6,2) (Tabela 10). Estudos de interferência de RNA (RNAi) mostram que esse gene é essencial para a sobrevivência de *T. brucei* tanto na forma procíclica quanto na forma sanguínea e que o produto codificado deve estar envolvido com a via de importação e/ou montagem do rRNA 5S para a biossíntese de ribossomo (Pitula *et al.*, 2002). Apesar disso, o papel funcional específico dessas proteínas no ciclo de vida de *T. brucei* e nas diferentes cepas de *T. cruzi* ainda precisa ser elucidado.

Regiões SIRE (“short interspersed repetitive element”) de *T. cruzi* são encontradas na região 3’ terminal de vários mRNAs e podem participar em processos de poliadenilação (Vazquez *et al.*, 2000). SIRE também pode estar localizado na região 5’ UTR (“untranslated region”) de alguns transcritos atuando como um doador de sítios de “trans-splicing”. A presença de SIRE em várias ESTs imobilizadas no microarranjo poderia induzir uma hibridização cruzada entre as sondas. Contudo, esse fato não foi observado já que sondas contendo sequências SIRE apresentaram um comportamento independente nos experimentos (ver Tabelas 7, 8, 9, e 10).

O conjunto de dados obtidos permite concluir que microarranjos são uma ferramenta poderosa para estudos de genômica comparativa, análise de expressão gênica em isolados de *T. cruzi* e para a descoberta de genes de interesse potencial. Com essa tecnologia identificamos vários genes para os quais a variação no número de cópias e/ou regulação dos níveis de transcrição de RNA foram demonstradas anteriormente por outras abordagens metodológicas. Este estudo também apresenta evidências adicionais de uma regulação pós-transcricional intensa nos níveis de RNA em *T. cruzi*.

6.2. Transcrição diferencial de DNA satélite em cepas de *T. cruzi*

Estudos de cinética de reassociação do DNA indicaram que o genoma de *T. cruzi* contém aproximadamente 9% de DNA reiterado (Castro *et al.*, 1981). A primeira sequência de DNA repetitivo isolada em *T. cruzi* foi denominada DNA satélite. Esta é uma sequência de 195 pb representada por 120.000 cópias no genoma da cepa Y (Sloof *et al.*, 1983a, Gonzalez *et al.*, 1984).

Os dados dos microarranjos indicaram maior hibridização das sondas de DNA satélite de 195 pb com DNA de CL Brener em relação a Silvio. A razão de hibridização CL Brener/Silvio foi de 6,0. Esses resultados são consistentes com conclusões anteriores de nosso grupo em colaboração com o grupo do Dr. S. Schenkman (EPM/UNIFESP) de que o DNA satélite é 6 vezes mais abundante no genoma de CL Brener quando comparado com Silvio (Elias *et al.*, 2003). Por outro lado, uma expressão gênica diferencial das sondas de DNA satélite também foi verificada em microarranjos hibridizados com cDNA de CL Brener e Silvio (Tabela 7). A confirmação dos dados por Northern blot mostrou que o DNA satélite da cepa Y era transcrito 6,5 vezes mais na cepa CL Brener do que em Silvio. Os dados dos microarranjos também mostraram que a sonda do satélite era mais expressa em epimastigotas e tripomastigotas metacíclicos de Hem 179 quando comparado com Famema (Tabela 10).

Foi sugerido que o DNA satélite de 195 pb pode ter um papel estrutural em cromossomos de *T. cruzi*, com base em evidências de que unidades de satélite estão organizadas de modo similar no genoma das cepas e agrupadas em domínios de 30 ± 10 kb (Elias *et al.*, 2003). Todavia, a observação de que essa sequência é transcrita sugere que ela pode estar envolvida na regulação da expressão gênica. Ensaio de “run-on” nuclear e experimentos realizados em epimastigotas permeabilizados com detergente mostraram que o satélite é transcrito e que a intensidade não é diretamente proporcional ao número de cópias gênicas (Elias *et al.*, 2001). Por outro lado, ensaios de Northern blot mostraram que uma sequência satélite de 177 pb de *T. brucei* não é transcrita (Sloof *et al.*, 1983b). Visando verificar se a sequência de 195 pb era de fato transcrita em *T. cruzi* realizamos experimentos de Northern blot com RNA total de formas epimastigotas de CL Brener e Silvio (Figura 15). Verificamos que a sonda hibridiza com espécies de RNA variando de

6,6 a 1,9 kb nas duas cepas e que a hibridização com a cepa Silvio é mais fraca, confirmando o resultado do microarranjo. Vazquez *et al.* (2000) analisando a transcrição de seqüências SIRE verificaram por Northern blot que essa seqüência estava representada em espécies de RNA com tamanhos variando de 7 kb a 800 pb, num padrão semelhante àquele obtido para o satélite de 195 pb.

Com o objetivo de verificar se a transcrição do DNA satélite não estava ocorrendo devido à contaminação do RNA com DNA genômico, submetemos o RNA de CL Brener a diferentes tratamentos e fizemos novos experimentos de Northern blot (Figuras 16 e 17). Verificamos que as preparações de RNA tratadas uma ou duas vezes com DNase continuavam apresentando forte hibridização com as sondas de DNA satélite. Por outro lado, RNAs tratados com RNase ou NaOH, não apresentaram nenhum sinal de hibridização com as sondas satélites, confirmando que as seqüências satélites são transcritas.

Visando verificar se o padrão de transcrição diferencial da seqüência de DNA satélite era uma característica estágio-específica, a sonda da cepa Y foi hibridizada com RNA total de diferentes cepas e estágios de desenvolvimento do parasita (Figura 18). Todos os RNAs foram tratados com DNase. As hibridizações mostraram que o RNA isolado de formas tripomastigotas metacíclicas de CL Brener e de epimastigotas da cepa Y apresentaram maior hibridização com a seqüência satélite. Por outro lado, os RNAs isolados das cepas Silvio e Dm28c apresentaram fraca hibridização. É interessante observar que a abundância e a transcrição da seqüência satélite é maior em cepas do grupo *T. cruzi* II (CL Brener e Y) em relação a *T. cruzi* I (Silvio e Dm 28c).

Foi relatado que a seqüência SIRE está localizada na extremidade 3' de muitos mRNAs de *T. cruzi* (Vazquez *et al.*, 2000). Decidimos então fazer uma busca no banco de dados de ESTs para verificar se a seqüência de DNA satélite era encontrada nestes cDNAs. Surpreendentemente, apenas uma EST foi encontrada, sugerindo que o DNA satélite não faz parte da estrutura de mRNAs.

Dihel *et al.* (2002), utilizando microarranjos contendo fragmentos de uma biblioteca genômica “shotgun” de *T. brucei*, identificaram uma série de seqüências diferencialmente transcritas em formas sanguíneas e procíclicas que não possuíam ORFs. Os autores propuseram que essas seqüências poderiam ser longas regiões de 3'UTR, ou regiões

intergênicas, que são transcritas como parte de transcritos policistrônicos (revisto em Teixeira *et al.*, 1998). Estas regiões seriam processadas durante o evento de “trans splicing”. Em *T. cruzi* a presença de agrupamentos de DNA satélite de 30±10 kb foi descrita (Elias *et al.*, 2003). No entanto, não está afastada a possibilidade de haver pequenos agrupamentos de satélite em regiões intergênicas. As bandas de 6,6 a 1,9 kb observadas nos Northern blots (Figuras 15 a 18) poderiam corresponder a produtos de processamento de agrupamentos que foram transcritos no mensageiro policistrônico. No momento, está sendo investigado se os transcritos do satélite de 195 pb permanecem no núcleo ou vão para o citoplasma (C. A. Martins).

Diaz *et al.* (1981, 1985) demonstraram que DNAs satélites de *Notophthalmus viridescens* (salamandra) formam longos agrupamentos que se intercalam entre os genes de histona. As duas classes de genes são co-transcritas, iniciando-se a transcrição a partir do promotor do gene de histona. Epstein *et al.* (1986) verificaram que esses satélites são encontrados no citoplasma das células de vários tecidos de *N. viridescens*.

Em bactérias do gênero *Neisseria* foi verificado que uma variação no número de repetições de microsátélites na região 5' de alguns genes resultava na regulação do padrão de transcrição (Stern *et al.*, 1986; Martin *et al.*, 2003). Assim, um decréscimo de microsátélites no gene *opa* (responsável pela aderência e infecção de *N. gonorrhoeae* em células epiteliais) produz uma proteína truncada que não permite à bactéria aderir. Um mecanismo de deslizamento da DNA polimerase, durante a replicação, volta a adicionar repetições e, portanto produzir a proteína funcional (Moxon *et al.*, 1994; Martin *et al.*, 2003). Será interessante investigar se o DNA satélite de 195 pb também participa do controle da expressão gênica em *T. cruzi*.

6.3. Expressão diferencial de genes em seis isolados de pacientes chagásicos

Tendo validado os microarranjos para o estudo do perfil de expressão gênica em cepas de *T. cruzi*, iniciamos a construção de um microarranjo com maior número de sondas de DNA visando comparar a expressão gênica em isolados de pacientes humanos com diferentes manifestações clínicas da doença de Chagas. A lâmina utilizada nesse trabalho (conjunto B) possui 710 ESTs de CL Brener (formas epimastigotas), 45 ESTs da cepa Tulahuen (formas amastigotas) e 32 genes de várias cepas do parasita. A anotação das 787

seqüências indica que 70% dessas seqüências não apresentam similaridade no banco de dados e que as classes funcionais mais representadas são proteínas envolvidas na síntese protéica (24%) e em vias metabólicas (18%). Esses valores são muito próximos aos encontrados para as sondas do conjunto A de microarranjos.

Neste estudo, RNA total foi isolado da fase mid-log de formas epimastigotas de cepas de três pacientes com manifestações cardíacas (cepas 115, B13 e 147) e de três pacientes assintomáticos (cepas VL10, Famema e Berenice 62) (Tabela 1). Cinco das seis cepas são de regiões endêmicas de Minas Gerais (exceto Famema que é do estado de São Paulo). Zingales *et al.* (1984) analisaram o padrão de antígenos de superfície de várias cepas de Bambuí (MG) (Chiari *et al.*, 1979) entre as quais as cepas 115, B13 e 147 empregadas no presente estudo. Os autores verificaram que muitos antígenos são conservados em populações distintas do parasita.

Nos experimentos de microarranjos, as hibridizações foram feitas com pares de cDNAs, sendo que o cDNA da cepa VL10 (de indivíduo assintomático) foi escolhida como referência em cada hibridização. Na análise derivada de 2 réplicas biológicas, foram considerados genes diferencialmente transcritos entre as cepas de cardíacos e de assintomáticos, aqueles que apresentavam aumento ou diminuição de expressão em todas as cepas cardíacas em relação a todas as cepas de assintomáticos. A identificação dos genes foi feita a partir de três métodos estatísticos, conforme mostrado em Resultados. Estas abordagens vêm sendo utilizadas na análise de dados de microarranjos (Reis *et al.*, 2004) ou de SAGE (Vêncio *et al.*, 2004). Concluímos haver expressão diferencial de 14 sondas (Tabela 12). Destas, 10 sondas apresentaram indução de expressão em todas as cepas de pacientes cardíacos quando comparadas com as cepas de pacientes assintomáticos. Por outro lado, 4 sondas tiveram expressão reprimida em todas as cepas de cardíacos quando comparadas com as cepas de assintomáticos. Oito destas sondas não apresentaram similaridade no banco de dados e representam possíveis novos genes envolvidos com a patogenicidade de *T. cruzi*.

Para confirmação da expressão diferencial destas sondas, membranas de Northern blot, contendo RNA de todas as cepas de cardíacos e assintomáticos utilizadas no ensaio de microarranjo, bem como RNA da cepa Hem 179, isolada de paciente com as formas cardíaca e digestiva, foram hibridizadas com algumas sondas. As figuras 23, 24 e 25

mostram o padrão de hibridização de 11 sondas analisadas em Northern blot. As sondas TENF0426, TEUN2842 e TENF0280 não foram analisadas por Northern Blot. Três tipos de padrões foram observados: (a) hibridização só com RNA de cepas de cardíacos (TENF0857); (b) hibridização com apenas um transcrito nos dois grupos de cepas (TENF0305; TENF0641; TENQ0823; TENF0593; TENF0103); (c) hibridização com mais de um transcrito e variações no padrão de hibridização (B6-PI08_F10; A4-PI04_C6; TENF0228; D2-PI101_E01; TENF0853). Após normalização dos valores de hibridização e determinação da razão de hibridização do RNA de cada cepa em relação a VL10, concluímos não haver confirmação dos dados do microarranjo para as sondas TENF0641 e TENF0103 (Tabela 13). De um modo geral, observamos que os valores das razões de hibridização no Northern blot são menores que aquelas obtidos no microarranjo (Tabela 12). Isto difere dos dados obtidos na primeira parte do trabalho e não temos explicação para isto.

A sonda TENF0857 mostrou hibridização apenas com cepas de cardíacos. Esta sonda que apresenta similaridade com o gene que codifica a subunidade 7 da NADH desidrogenase (ND7), é a mesma sequência encontrada com expressão aumentada nas formas epimastigotas e tripomastigotas metacíclicas da cepa Hem 179 em relação a Famema (ver acima item 6.1.6). Esse dado ilustra a consistência e reprodutibilidade dos resultados obtidos com os microarranjos e mostra que, apesar da grande variabilidade genética e biológica entre os isolados de *T. cruzi*, algumas características são compartilhadas por grupos de cepas.

Uma vez que a TENF0857 pertence ao mesmo agrupamento das TENU0045 e TENU0146, que também apresentam elevada similaridade com o gene da ND7, esperávamos que estas duas sequências também apresentassem indução de expressão nas cepas de cardíacos. No entanto, os métodos estatísticos empregados apenas haviam revelado expressão diferencial da TENF0857. Desta forma, decidimos analisar os dados do microarranjo para as sondas TENU0045 e TENU0146, com o gráfico MA (Figura 28A e B). Observamos não haver uma nítida separação de todas as réplicas dos dois grupos de cepas, ao contrário do que se verifica para a sonda TENF0857 (Figura 21). Pelo teste de correlação de Spearman encontramos para as sondas TENU0045, TENU0146 e TENF0857 valores de rho de, respectivamente, 0,57; 0,47 e 0,83. Pela correlação de Pearson para estas

mesmas sondas encontramos valores de rho, respectivamente de 0,62; 0,75 e 0,97. Em relação ao valor de E no teste “Bayes Error Rate” verificamos valores de 0,22 (TENU0045); 0,16 (TENF0146) e 0,0 (TENF0857).

Portanto, tendo como premissa a seleção de sondas que apresentam expressão diferencial em todas as cepas de cardíacos em relação a todas as cepas de assintomáticos; valores de $\rho \geq 0,75$ e valor de $E \leq 0,05$ as seqüências de TENU0045 e TENU0146 não foram aceitas pelos testes. Isto mostra o rigor das análises.

Decidimos então analisar a similaridade de seqüência entre as três ESTs, alinhando as seqüências disponíveis nos bancos de dados e que correspondem a cerca 300 pb da extremidade 5' do cDNA. (Figura 14). Observamos certa divergência no lado 3' das seqüências, que poderia justificar as diferenças nos resultados dos microarranjos. Tentamos seqüenciar a porção 3' do cDNA das três ESTs mas não tivemos êxito.

Figura 28. Gráfico MA das ESTs TENU0045 (A) e TENU0146 (B). Quadrados vermelho, marron e rosa representam os resultados normalizados das réplicas nas cepas isoladas de pacientes com cardiopatia chagásica crônica. Quadrados verdes, amarelos e azul claro representam os resultados normalizados das réplicas nas cepas isoladas de indivíduos assintomáticos. Valores de M positivos indicam genes induzidos em relação à cepa VL10. M=0 indica que não houve expressão diferencial.

O gene da ND7 está localizado no DNA dos maxicírculos e foi sequenciado em vários tripanossomatídeos. Este gene é pan-editado em *T. brucei* e *T. cruzi*, envolvendo a deleção e inserção de resíduos de uridina nas regiões codificadoras do mRNA. A edição pode produzir codons de terminação e iniciação em ORFs e também pode afetar a estabilidade e a abundância de transcritos através de edições nas regiões 3'UTR (Maslov *et al.*, 1994; Koslowsky *et al.*, 1990). A sequência da proteína da subunidade ND7 de *T. brucei* é homóloga à da subunidade de 49 kDa do complexo da NADH desidrogenase bovina e, provavelmente, é um componente do complexo ferro-enxofre (Fearnley *et al.*, 1989). A TENF0857, que mostra indução gênica média de 26 vezes nas cepas cardíacas, mostra alta identidade com a região pré-editada do gene de ND7 da cepa Silvio (número de acesso U05815) (Maslov *et al.*, 1994). A hibridização em Northern blots com a TENF0857 indica ausência desse transcrito nas cepas isoladas de pacientes assintomáticos, confirmando os dados do microarranjo.

A expressão de ND7 em *T. brucei* é regulada nos estágios de desenvolvimento e o transcrito completamente editado dessa subunidade só é encontrado nas formas sangüíneas (Koslowsky *et al.*, 1990). Foi reportado que formas sangüíneas de *T. brucei* são totalmente dependentes da glicólise para a obtenção de energia e que citocromos e enzimas do ciclo de Krebs estão ausentes (cf. Koslowsky *et al.*, 1990). Além disso, evidências baseadas em ensaios enzimáticos sugerem que a NADH desidrogenase e o primeiro sítio de fosforilação mitocondrial estão ausentes em *T. cruzi* (revisto por Cazzulo, 1994). Portanto, a presença de transcritos de ND7 em formas sangüíneas de *T. brucei* (Koslowsky *et al.*, 1990), em cepas de *T. cruzi* isoladas de pacientes com a forma cardíaca e com a forma cardíaca e digestiva e em tripanossomatídeos de plantas que apresentam deficiência respiratória (Maslov *et al.*, 1999) é intrigante. Apesar disso, a expressão diferencial do transcrito ND7 em cepas de *T. cruzi* e em estágios de desenvolvimento dos dois tripanossomas sugere que essa sequência tenha um papel importante e merece maior investigação. Além disso, como o transcrito ND7 divide as cepas de assintomáticos e de cardíacos em dois grupos, esse gene poderia ser um marcador potencial para o prognóstico da evolução da doença de Chagas. É claro que a presença deste transcrito deverá ser investigada em um número maior de cepas.

A ausência do transcrito de ND7 nas cepas de assintomáticos poderia ser resultante de diferenças na sequência dos genes e/ou na edição do mRNA. Esta hipótese é sustentada pelo fato de que Maslov *et al.* (1999) demonstraram que em *T. brucei* e *T. cruzi* o mRNA de ND7 é pan-editado em dois grandes domínios, enquanto que em *L. tarentoleae* e em *C. fasciculada* esse mesmo mRNA é editado somente em um domínio, sugerindo variação de edição entre os tripanossomatídeos. A edição em *T. brucei* pode gerar diferenças entre o RNA pré-editado e o RNA maduro de até 450 pb, não permitindo a detecção do gene em experimentos de Southern blot com DNA total de *T. brucei* usando como sonda o RNA completamente editado (Koslowsky *et al.*, 1990). Dessa maneira, é possível que não tenhamos detectado o transcrito de ND7 nas cepas de assintomáticos pelo fato de haver um processo de edição muito extenso nessas cepas ou diferenças de similaridade de sequência. A ausência do transcrito ND7 nas cepas de assintomáticos poderia determinar ausência de atividade de NADH desidrogenase ou de um Complexo I da cadeia respiratória constituído por subunidades diferentes. De fato, Denicola-Seoane *et al.* (1992) e Cazzulo *et al.* (1992; 1994) demonstraram que a atividade de NADH desidrogenase está ausente na cepa Tulahuen de *T. cruzi* (forma epimastigota). Os autores propõem que o metabolismo de carboidratos ocorreria principalmente por um processo denominado “fermentação aeróbica”, uma vez que o nível de citocromos é baixo e a NADH desidrogenase e o primeiro sítio de fosforilação estão ausentes. Isto implicaria numa ineficiência da cadeia de transporte de elétrons e na entrada dos equivalentes reduzidos preferencialmente via Complexo II (revisto por Cazzulo, 1992; 1994). Também foi demonstrada variação na concentração de citocromos entre isolados de *T. cruzi*, o que sugere diferenças no metabolismo oxidativo (Engel *et al.*, 1990). A título de especulação, os dados do microarranjo sugerem que as cepas de cardíacos, que transcrevem a subunidade ND7 poderiam ter NADH desidrogenase e a cadeia de transporte de elétrons mais funcional, em relação às cepas de assintomáticos. O sistema respiratório mais eficiente poderia ser uma resposta adaptativa.

Foi proposto que uma das defesas usadas por macrófagos contra a infecção por *T. cruzi* é a produção de peroxinitritos. Esses compostos agem sobre a fumarato redutase, enzima essencial para a fermentação aeróbica e, conseqüentemente, para o metabolismo energético de *T. cruzi* (Rubbo *et al.*, 1994). Desta forma as cepas de cardíacos que possuem

os dois sistemas respiratórios poderiam apresentar vantagens sobre as cepas de assintomáticos.

Outra possibilidade a ser contemplada é a existência de classes de NADH desidrogenases com diferentes subunidades. Esta hipótese é sustentada pelo fato de que Denicola-Seoane *et al.* (1992) determinaram que a respiração de *T. cruzi* é insensível a rotenona, inibidor específico de NADH desidrogenase (complexo I). Por outro lado, foi relatado que formas sanguíneas de *T. brucei* e formas epimastigotas de *T. cruzi* contêm uma enzima similar a NADH desidrogenase, provavelmente composta por subunidades diferentes das tradicionais do complexo I de mamíferos, o que pode determinar diminuição na sensibilidade da NADH desidrogenase à rotenona. (Tielens & Van Hellemond, 1998). Uma vez que não detectamos transcrito de ND7 nas cepas de assintomáticos, é possível que essas cepas possuam uma NADH desidrogenase com composição diferente de subunidades.

Os dados dos microarranjos indicaram que o mRNA B6-PI08_F10, que possui alta similaridade com o gene da aspartato aminotransferase (AAT) de *T. brucei*, é em média 2,8 vezes mais abundante em cepas de cardíacos em relação às cepas de assintomáticos (Tabela 12). Essa enzima tem um papel importante na regeneração de metionina usada na síntese protéica e de poliaminas. Estes compostos são importantes para o crescimento celular da maioria dos parasitas, bactérias e células tumorais (Sufrin *et al.*, 1995). Nos tripanossomatídeos, uma poliamina importante é a espermidina, utilizada na síntese de tripanotiona, composto essencial para a degradação de peróxidos e, conseqüentemente, para a proteção celular contra estresse oxidativo. Desta forma as cepas de cardíacos poderiam estar produzindo mais espermidina o que conferiria maior proteção contra peróxidos de hidrogênio. Para confirmar o resultado do microarranjo, um experimento de Northern Blot foi realizado. O dado da Figura 23C indica que o transcrito de B6-PI08_F10 apresenta aproximadamente 1,7 kb nas cepas de assintomáticos (VL10, Famema e Berenice 62). Essa banda não é observada nas cepas de cardíacos 115 e 147, onde a sonda hibridiza com uma banda de 1,1 kb. Na cepa de cardíaco B13 e na cepa de paciente com a forma cardíaca e digestiva Hem 179 são observadas ambas as bandas (aproximadamente 1,7 kb e 1,1 kb) (Figura 23C). Esses dados sugerem haver diferenças nas estruturas dos genes de AAT entre as cepas, ou variações no processamento pós-transcricional do mRNA, por exemplo por “splicing” alternativo. Apesar de ser um evento muito raro, a presença de um intron no

gene que codifica a poli A polimerase de *T. cruzi* e *T. brucei* foi descrita (Mair *et al.*, 2000). Berger *et al.* (2001) clonaram e caracterizaram o gene da AAT de alguns protozoários parasitas. No caso de *T. brucei*, os autores verificaram a presença de dois genes que apresentam 1.212 pb e 1.173 pb, e que codificam, respectivamente, polipeptídios de 403 e 390 aminoácidos. Estes produtos correspondem às AATs nuclear e mitocondrial, respectivamente. Para *T. cruzi*, genes que codificam AAT ainda não foram caracterizados. Desta forma não é possível dizer se os transcritos detectados com a sonda B6-PI08_F10 correspondem a enzimas nucleares ou mitocondriais, ou se as duas enzimas apresentam diferenças de tamanho molecular. Em *T. brucei* foi determinado que cada um dos genes da AAT corresponde a um gene de cópia única (Berger *et al.*, 2001). Por outro lado, os dados da literatura apontam que tripanossomatídeos são diplóides e que cada gene apresenta um “alelo”, que está em cromossomos “homólogos”. Para *T. cruzi* foi observado em vários casos que os dois alelos não são idênticos (Henriksson *et al.*, 1990; Pedroso *et al.*, 2003). Desta forma, a presença simultânea de dois transcritos do gene AAT (1,7 kb e 1,1 kb) na cepa B13 e Hem 179 poderia ser explicada por este mesmo fenômeno. Iribar *et al.* (2003) detectaram em *Leishmania* uma sequência que apresentava dois códons de terminação internos e no Northern Blot gerava duas bandas de tamanho diferente (3,8 e 0,4 kb). Os autores verificaram que somente o transcrito de baixo peso molecular era transportado para o citoplasma, sugerindo estabilidade e processamento adequado desse transcrito. Em resumo, as características dos genes AAT e de seus produtos merecem ser estudadas em *T. cruzi*.

Outra sequência interessante identificada nos microarranjos como sendo induzida em cepas de cardíacos –A4-PI04_C06 - apresenta similaridade com o gene que codifica a enzima triparedoxina peroxidase de *T. cruzi* (CB923896). Esta enzima está localizada no citoplasma e no glicossomo e está envolvida em um sistema muito importante de resistência de *T. cruzi* a hidroperóxidos e peróxidos de hidrogênio, já que o parasita não possui catalase ou glutathione redutase (Wilkinson *et al.*, 2002). A título de especulação, poderíamos supor que a indução deste gene em cepas de cardíacos proporcionaria um mecanismo de defesa a peróxidos mais eficiente em relação às cepas de assintomáticos. No entanto, não pode ser descartado que cepas de assintomáticos tenham desenvolvido mecanismos diferentes de proteção, ou que haja diferenças de similaridade de sequência.

A sonda D2-PI01_E01, que possui similaridade com uma ciclofilina de *Leishmania donovani*, revelou redução da abundância do transcrito nas cepas de cardíacos quando comparadas com as cepas de assintomáticos. Esse transcrito apresenta duas bandas no Northern Blot uma de 2,6 kb e outra de 1,3 kb (Fig. 25C). A expressão desse gene pode implicar em resistência a ciclosporina, um potente composto anti-malárico, com amplo espectro antimicrobial (Peterson *et al.*, 2000). Essa droga inibe a replicação de formas epimastigotas de *T. cruzi* *in vitro*, mas não inibe a replicação intracelular de amastigotas em macrófagos e nem a infecção em camundongos (McCabe *et al.*, 1985).

As sondas TENF0853 e TENF0228, cujas seqüências não têm similaridade em bancos de dados, apresentam variação no padrão dos transcritos entre as cepas

Analizando os dados, acreditamos que algumas seqüências apresentam uma diferença de expressão mais significativa entre as cepas de cardíacos e de assintomáticos e merecem estudos mais aprofundados. Entre elas estão as ESTs: TENF0857, B6-PI08_F10, TENF0228, D2-PI01_E01 e TENF0853.

Em resumo, a análise da expressão diferencial de genes em cepas de pacientes com a forma cardíaca e indeterminada da doença de Chagas permitiu a identificação de alguns genes potenciais que podem estar relacionados à patogênese do parasita e que merecem maiores investigações para a elucidação da sua função em *T. cruzi*.

6.4. Expressão gênica de fibroblastos humanos infectados com cepas de *T. cruzi*

Estes experimentos foram realizados num estágio de 3 meses no laboratório do Dr. David Munroe no NCI-SAIC (Frederick, EUA). Salientamos que esses experimentos representam um ensaio-piloto, visando o desenho futuro de um projeto mais abrangente. Neste estudo utilizamos fibroblastos humanos infectados com duas cepas de *T. cruzi* (VL10, isolada de indivíduo assintomático e 147, isolada de indivíduo com cardiopatia severa) e o tempo de infecção, 24 horas.

A primeira diferença observada entre as cepas foi verificar que após 8 a 10 dias de infecção, no sobrenadante da cultura eram obtidos tripomastigotas, para a cepa 147 e predominantemente amastigotas, para a cepa VL10. É possível que, por fatores desconhecidos, os fibroblastos infectados com VL10 litem antes de haver diferenciação de

amastigotas em tripomastigotas. Esta observação mostra que a interação parasita-hospedeiro é dependente de características genéticas de ambas as células.

Após 24h de infecção com tripomastigotas (147) ou amastigotas (VL10), os RNAs das monocamadas foram extraídos. No mesmo período, fibroblastos humanos foram incubados com meio condicionado de parasitas (MCP) de ambas as cepas. O RNA foi extraído destas culturas assim como de fibroblastos humanos não infectados. Este último RNA foi usado como referência nas hibridizações.

A análise da qualidade do RNA em gel de agarose corado com brometo de etídio (Figura 26) mostrou ausência do rRNA 28S na amostra extraída de células infectadas com a cepa 147, mas não com a cepa VL10. Em uma célula normal, em geral, a razão da fluorescência da banda do rRNA 28S em relação à banda de rRNA 18S é 2:1. Em alguns tecidos, entretanto, a razão 28S/18S pode ser 1:2 ou 1:3. Este efeito não é resultante da degradação do RNA durante o processo de purificação. Payao *et al.* (1998) demonstraram que a razão 28S/18S em células do sangue periférico diminui com a idade do indivíduo (envelhecimento) e isto ocorre pela degradação natural do rRNA 28S. Um efeito análogo foi observado em células de cérebro e fígado de camundongo (Mori *et al.*, 1978). Também foi observado que a abundância do rRNA 28S em células de camundongo infectadas com o vírus da hepatite murina decai ao longo da infecção e que, após 24 horas, praticamente não é possível encontrar rRNA 28S (Benerjee *et al.*, 2000). Por outro lado, nenhum dano específico no rRNA 18S é detectado. Os autores concluíram que a indução da degradação do rRNA 28S difere de outros mecanismos de degradação de rRNA descritos na literatura.

Genes de rRNA estão envolvidos em processos de transcrição e tradução e podem modular a expressão gênica (Skrypina *et al.*, 2003). No caso das células infectadas com a cepa 147, podemos especular que o parasita, ao induzir a degradação do rRNA 28S da célula hospedeira (seja por liberação de RNases seja por indução de RNases da célula hospedeira), poderia estar controlando toda a maquinaria de transcrição e tradução. Neste caso, o parasita estaria utilizando os compostos acumulados para sua nutrição e sobrevivência. Salienta-se que após a utilização total da glicose do meio, *T. cruzi* utiliza basicamente aminoácidos para obtenção de energia e fontes de carbono (cf Cazullo, 1994; Urbina, 1994). Coincidentemente, verificamos no experimento de microarranjo que muitos

genes da célula hospedeira estão com a expressão reduzida na infecção com a cepa 147 (Tabela 14). Esse fato pode estar ocorrendo por outros fatores como será discutido abaixo.

Nos experimentos de hibridização competitiva utilizamos lâminas de microarranjo contendo oligonucleotídeos derivados de 24.000 genes humanos. RNA de fibroblastos humanos não infectados foi utilizado como referência em todos os experimentos. Após normalização dos dados por “Lowess”, filtramos os genes que apresentavam redução ou indução de expressão (maior que 1,5 vezes) nas células infectadas, em relação às células não infectadas. Identificamos 28 sondas diferencialmente transcritas nas células infectadas com a cepa 147 (paciente cardíaco) (Tabela 14). Destas, 27 sondas (96,4%) estão com expressão reduzida nas células infectadas. Por outro lado, apenas uma sonda - correspondente ao gene da “heme oxigenase (decycling) 1”, apresentou indução nas células infectadas com 147 (Tabela 14). Na infecção de células com a cepa VL10, encontramos 17 sondas diferencialmente transcritas (Tabela 15). Destas, 16 sondas (94,1%) estão com expressão aumentada nas células infectadas, enquanto que apenas uma sonda (Proteína 5 ligante do fator de crescimento do tipo insulina) mostra situação inversa (Tabela 15). Estas diferenças de comportamento à infecção pelas duas cepas são marcantes. Ainda que seja uma possibilidade remota, as diferenças poderiam ser resultantes do uso de tripomastigotas (cepa 147) e amastigotas (cepa VL 10) para iniciar a infecção. De qualquer forma, isto determina que um estudo temporal ao longo da infecção pelas duas cepas seja realizado para a confirmação das observações acima, ou mesmo que a cepa VL10 seja substituída por outra de indivíduo assintomático que forneça tripomastigotas no sobrenadante da cultura.

Vaena de Avalos *et al.* (2002) observaram uma alteração na abundância de transcritos de HFF (human foreskin fibroblasts) após 24 horas de infecção com a cepa Y, com indução de cerca de 106 genes. Não observamos concordância de similaridade entre os genes relatados por estes autores e aqueles identificados em nossos experimentos.

É importante salientar que o gene da “heme oxigenase (decycling) 1” está induzido em fibroblastos infectados com VL10 e 147. Classicamente a heme oxigenase é conhecida por sua habilidade de catalisar a conversão do grupo heme a biliverdina e CO₂. Esta enzima também pode atuar na proteção celular a danos oxidativos. Essa enzima existe tanto na forma indutiva quanto constitutiva, que são denominadas, respectivamente, heme oxigenase 1 (HO-1) e heme oxigenase 2 (HO-2). A heme-oxigenase é uma enzima endógena,

citoprotetora e seu nível está aumentado em células que estão sofrendo estresse oxidativo (Bulger *et al.*, 2003). A indução deste gene em fibroblastos infectados com ambas as cepas sugere que esta seja uma resposta comum na infecção por *T. cruzi*. Por outro lado, a enzima heme-oxigenase requer citocromo P450 para o seu funcionamento (ambas as enzimas estão localizadas no retículo endoplasmático) (Gong *et al.*, 2003). Desta forma, a repressão do gene do citocromo P450 (Tabela 14) na célula infectada com a cepa 147, pode ser um mecanismo que o parasita usa para controlar a resposta a estresse oxidativo. O fato de o gene da heme-oxigenase estar induzido tanto nas células infectadas com 147 quanto nas células cultivadas com MCP de 147 (Figura 27) sugere que a regulação da expressão desse gene deve ser mediada por algum fator solúvel do parasita que é liberado no meio. Burger *et al.* (2003) sugerem que a indução da proteína heme oxigenase 1, frente a estresse oxidativo e lipopolissacarídeos, poderia estar envolvida na atenuação da resposta proinflamatória de células endoteliais. Dessa forma um aumento da expressão deste gene nos fibroblastos infectados com as duas cepas, suporta a idéia de estresse oxidativo e poderia significar uma resposta da célula humana direcionada para sua proteção.

Entre os genes reprimidos nas células infectadas com a cepa 147 está o gene do citocromo P450 (polipeptídio 1) que havia sido descrito como sendo reprimido em células de coração de murinos infectados com a cepa Brasil (Mukherjee *et al.*, 2003). Garg *et al.* (2003) demonstraram que a infecção de coração murino com a cepa Silvio gera uma redução generalizada na expressão de genes que codificam componentes da via de transporte de elétrons e da fosforilação oxidativa. Os autores propõem que essa redução de expressão gênica pode gerar uma interrupção dessas vias iniciando um aumento de ROS (espécies reativas de oxigênio), que aceleraria a degeneração cardíaca. Em nossa análise, encontramos também redução da expressão do gene que codifica a subunidade VIc da citocromo c oxidase nas células infectadas com a cepa 147 (Tabela 14).

A única sonda que apresenta expressão reduzida nas células infectadas com VL 10 corresponde à Proteína 5 ligante do fator de crescimento do tipo insulina (Tabela 15). Este gene está envolvido na regulação da proliferação celular. A super expressão desse gene foi descrita em vários tipos de câncer e pode estar envolvida com a aceleração do crescimento celular (Stolf *et al.*, 2003). A redução da expressão desse gene nas células infectadas com

VL10 pode representar um mecanismo desenvolvido pelo parasita para modular o crescimento da célula hospedeira.

Entre os genes com expressão aumentada na célula infectada com VL10, estão os genes de ciclina D1, descrito como sendo induzido em células de fibroblastos infectados com *T. gondii* e o gene da quemocina (motivo C-C) (Mukherjee *et al.*, 2003). Esta proteína faz parte de uma família de citocinas que são proteínas secretadas, envolvidas em processos imunoregulatórios e inflamatórios, importantes em vários processos, dentre eles, na atração química de basófilos e eosinófilos. Além disso, Alibert *et al.* (2000) demonstraram que a indução da quemocina CC na invasão de *T. gondii* em células de camundongo é um processo crucial para o recrutamento de células dendríticas, importante para a imunidade célula-dependente a patógenos intracelulares e para a determinação da resposta imunológica subsequente. A ativação das células dendríticas ocorre por ligação da quemocina CC ao receptor CCR5 (receptor 5 de quemocina CC) que aciona a liberação de interleucina-12, promovendo a resposta imunológica específica ao agente infeccioso. A título de especulação, poderíamos supor que o aumento da expressão do gene da quemocina CC nas células infectadas com a cepa VL10, seja um mecanismo de resposta imune da célula hospedeira a esta cepa.

Dois genes que apresentaram diferenças de expressão foram analisados por RT-PCR (real time PCR). O gene da “heme oxigenase (decycling)1”, que apresenta indução de 4 e 5 vezes na infecção com 147 e VL10, respectivamente (Tabelas 14 e 15). O experimento de RT-PCR mostrou uma indução de 80 e 60 vezes, respectivamente (Figura 27A). Esses dados confirmam as conclusões obtidas pelo microarranjo, apesar de os valores serem bem mais altos. Nas células cultivadas com MCP de 147 encontramos indução de 150 vezes deste transcrito. Esses dados indicam que esse gene é importante na resposta celular frente à infecção por *T. cruzi*, e que sua indução pode ser mediada por fatores excretados/secretados pelo parasita.

Outro gene analisado por RT-PCR foi o gene da proteína 5 ligante do fator de crescimento do tipo insulina (Figura 27B). Observamos repressão de expressão de 6 vezes apenas no RNA isolado a partir de células infectadas com a cepa VL10. Esses dados confirmam as conclusões obtidas a partir do microarranjo, onde esta sonda mostrou repressão de 2,7 vezes apenas nas células infectadas com VL10.

Em resumo, o estudo-piloto mostrou que a abordagem experimental é válida e indicou vários genes humanos diferencialmente expressos durante a infecção por cepas de *T. cruzi*. Estes podem ser importante na interação parasita-célula hospedeira.

Estamos cientes de que para um estudo mais abrangente teria sido ideal termos realizado uma análise temporal da expressão gênica ao longo da infecção e termos monitorado a porcentagem de células infectadas pelas cepas. Também teria sido interessante verificar o número de parasitas por célula e saber em que etapa da infecção se encontra o sistema após 24 horas. Neste período, os dados da literatura sugerem que os tripomastigotas devem ter escapado do vacúolo parasitóforo e estariam diferenciando-se em amastigotas para iniciar o processo de divisão. No entanto, tendo em vista a diversidade biológica e genética das cepas, é possível que as duas cepas analisadas estejam em fases diferentes no processo de infecção.

6.5. Comentários e conclusões finais

A tecnologia de microarranjos de DNA é uma ferramenta poderosa na análise simultânea da expressão de milhares de genes. Esta tecnologia está bem estabelecida. A análise das imagens de hibridização é um passo muito importante na identificação de seqüências diferencialmente expressas. Neste caso, uma variedade de métodos e “softwares” estão disponíveis, e são continuamente gerados, cada um com suas vantagens e limitações. Isto também ocorre com os métodos estatísticos empregados para a análise dos resultados e, até o momento, não há um consenso entre os pesquisadores sobre o “melhor” método a ser usado. Desta forma, as seqüências evidenciadas pelos experimentos de microarranjo devem ser consideradas como seqüências “potencialmente” diferencialmente expressas. Estas podem ser validadas em experimentos de Northern blot ou “real-time PCR”. Por outro lado, a caracterização de cada seqüência revelada pelos micorarranjos e validada pelas duas abordagens acima é um longo caminho e pode constituir um projeto de pesquisa independente.

Nosso projeto de doutoramento teve como objetivo inicial construir um protótipo de microarranjo para *T. cruzi* para validar esta tecnologia para um parasita que apresenta grande diversidade genética e para o qual os dados existentes mostram haver um controle

de expressão gênica bastante peculiar. Salientamos que quando iniciamos este projeto, em agosto de 2000, não havia sido descrita na literatura a utilização de microarranjos para a análise de expressão gênica em nenhum tripanossomatídio: *T. cruzi*, *T. brucei* ou *Leishmania*. De fato, o primeiro relato sobre o uso de microarranjos para analisar o perfil de transcrição em *T. brucei* ocorreu em 2000 (El-Sayed *et al.*, 2000) e o primeiro para *T. cruzi*, em 2003 (Minning *et al.*, 2003), seguido pelo nosso artigo em 2004 (Baptista *et al.*, 2004). O protótipo de microarranjo, contendo cerca de 800 sondas de *T. cruzi*, mostrou que esta abordagem é válida para análises genômicas de cepas de *T. cruzi* I e *T. cruzi* II (Silvio e CL Brener) e para a análise da expressão diferencial de genes nos dois grupos de cepas. Também obtivemos evidências de que existem níveis diferentes de transcritos em cepas isoladas de pacientes com a forma indeterminada e cardíaca da doença de Chagas. Para a análise dos dados aplicamos diferentes métodos estatísticos, alguns inéditos na análise de dados de microarranjos como o “Bayes Error Rate”. Muitas das sondas foram validadas por Southern e Northern blot. Desta forma, nossos resultados revelaram uma série de seqüências potenciais que, a partir deste momento, precisam ser caracterizadas em profundidade para verificar seu possível envolvimento com a patogênese ou sua aplicação como alvos de diagnóstico/prognóstico. Infelizmente, em virtude da limitação de tempo para cumprir os prazos determinados para o doutoramento (quatro anos), não tivemos possibilidade de sequer iniciar alguns estudos. Aliás, numa revisão de maio deste ano, Robert Duncan (Duncan, 2004) comenta que “enquanto para parasitas do grupo Apicomplexa - *Plasmodium* e *Toxoplasma*, os microarranjos mostraram ser uma ferramenta importante para a descoberta de novos genes e para a identificação da expressão coordenada de famílias gênicas, no caso de parasitas da ordem Kinetoplastida – *Trypanosoma* e *Leishmania*, os resultados obtidos até o momento não avançaram muito além do estágio de validação e descoberta de novos genes, cuja função deve ainda ser determinada”.

O projeto “piloto” ou “exploratório” que desenvolvemos durante o estágio de três meses no laboratório do Dr. David Munroe no NCI-SAIC nos Estados Unidos teve por objetivo obter dados preliminares para verificar a expressão gênica diferencial de uma linhagem de células humanas infectadas com duas cepas de *T. cruzi* isoladas de pacientes com manifestações da doença de Chagas distintas. Neste caso também, em função da

limitação de tempo, não pudemos fazer uma análise mais abrangente, uma vez que tivemos que fixar vários parâmetros para o estudo. Por outro lado, os resultados obtidos permitiram a delineação de um projeto mais compreensivo, a ser desenvolvido por um pós-doutor em nosso laboratório com a colaboração do grupo do Dr. D. Munroe num futuro próximo.

Em resumo, acreditamos ter cumprido uma etapa importante para a implantação da tecnologia de microarranjos de DNA, para o estudo da expressão gênica em *T. cruzi* e aberto a possibilidade de novas linhas de pesquisa em nosso laboratório.

7. REFERÊNCIAS

- AGUERO, F., VERDUN, R. E., FRASCH, A. C. & SANCHEZ, D. O. (2000). A random sequencing approach for the analysis of the *Trypanosoma cruzi* genome: general structure, large gene and repetitive DNA families, and gene discovery. *Genome Res.* **10**:1996-2005.
- ALIBERTI, J., REIS E SOUSA, C., SCHITO, M., HIENY, S., WELLS, T., HUFFNAGLE, G. B. & SHER, A. (2000). CCR5 provides a signal for microbial induced production of IL-12 by CD8 alpha+ dendritic cells. *Nat Immunol.* **1**: 83-87.
- ALMEIDA, R., NORRISH, A., LEVICK, M., VETRIE, D., FREEMAN, T., VILO, J., IVENS, A., LANGE, U., STOBER, C., McCANN, S. & BLACKWELL, J. M. (2002). From genomes to vaccines: *Leishmania* as a model. *Phyl Tras: Bioll Sciences.* **57**: 5-11.
- ALTSCHUL, S. A. & KOONIN, E. V. (1998). Interactive profile searches with PSI-BLAST – a tool for discovery in protein databases. *TIBS.* **23**: 444-447.
- ANDRADE, L. O., MACHADO, C. R., CHIARI, E., PENA, S. D. & MACEDO, A. M. (2002). *Trypanosoma cruzi*: role of host genetic background in the differential tissue distribution of parasite clonal populations. *Exp Parasitol.* **100**: 269-275.
- ANDRADE, L. O., MACHADO, C. R., CHIARI, E., PENA, S. D. & MACEDO, A. M. (1999). Differential tissue distribution of diverse clones of *Trypanosoma cruzi* in infected mice. *Mol Biochem Parasitol.* **25**:163-172.
- ANONYMOUS (1999). Recommendations from an international Symposium to commemorate the 90th anniversary of Discovery of Chagas disease. *Mem Inst Oswaldo Cruz.* **94**: 429-432.
- ARAYA, J., CANO, M. I., GOMES, H. B., NOVAK, E. M., REQUENA, J. M., ALONSO, C. LEVIN, M. J., GUEVARA, P., RAMIREZ, J. L. & DA SILVEIRA, J. F. (1997). Characterization of an interspersed repetitive DNA element in the genome of *Trypanosoma cruzi*. *Parasitol.* **115**: 563-570.
- ARAYA, J. E., CANO, M. I., YOSHIDA, N. & DA SILVEIRA, J. F. (1994). Cloning and characterization of a gene for the stage-specific 82-kDa surface antigen of metacyclic trypomastigotes of *Trypanosoma cruzi*. *Mol Biochem Parasitol.* **65**:161-169.
- ARRUDA, M. V. REINACH, F. C., COLLI, W. & ZINGALEZ, B. (1990). Sequence of 24Sα ribosomal RNA gene and characterization of a corresponding pseudogene from *Trypanosoma cruzi*. *Mol Bioch Parasitol.* **40**: 35-42.

- AUGUSTO-PINTO, L. A., TEIXEIRA, S. M. R., PENA, S. D. J. & MACHADO, C. R. (2003). Single-nucleotide polymorphisms of the *Trypanosoma cruzi* *MSH2* gene support the existence of three phylogenetic lineages presenting differences in mismatch-repair efficiency. *Genetics*. **164**: 117-126.
- AUGUSTO-PINTO, L., BARTHOLOMEU, D. C., TEIXEIRA, S. M., PENA, S. D. & MACHADO, C. R. (2001). Molecular cloning and characterization of the DNA mismatch repair gene class 2 from the *Trypanosoma cruzi*. *Gene*. **272**: 323-333.
- AUSUBEL, F., BRENT, R., KINGSTON, R. E., MOORE, D. D., SEIDMAN, J. G., SMITH, J. A. & STUHL, K. (1995). Short Protocols in Molecular Biology. John Wiley & Sons, Inc. Madison
- BANERJEE, S., AN, S., ZHOU, A., SILVERMAN, R. H. & MAKINO, S. (2000). RNase L-independent specific 28S rRNA cleavage in murine coronavirus-infected cells. *J Virol*. **74**: 8793-8802.
- BAPTISTA, C. S., VENCIO, R., ABDALA, S., VALDARES, M. P., MARTINS, C., BRAGANÇA, C. A. & ZINGALES, B. (2004). DNA microarrays for comparative genomics and analysis of gene expression in *Trypanosoma cruzi*. Manuscrito aceito para publicação na *Molecular and Biochemical Parasitology*.
- BAXEVANIS, A. D. & OULLETTE, B. F. F. (1998). Bioinformatics, a practical guide to the analysis of genes and proteins. Wiley Interscience, New York. 370pp.
- BEHR, M. A., WILSON, M. A., GILL, W. P. SALAMON, H., SCHOOLNIK, G. K., RANE, S. & SMALL, P. M. (1999). Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science*. **284**: 1520-1523.
- BEN MAMOUN, C., GLUZMAN, I. Y., HOTT, C., MACMILLAN, S. K., AMARAKONE, A. S., ANDERSON, D. L., CARLTON, J. M., DAME, J. B., CHAKRABARTI, D., MARTIN, R. K., BROWNSTEIN, B. H. & GOLDBERG, D. E. (2001). Co-ordinated programme of gene expression during asexual intraerythrocytic development of the human malaria parasite *Plasmodium falciparum* revealed by microarray analysis. *Mol Microbiol*. **39**: 26-36.
- BERGER, L. C., WILSON, J., WOOD, P. & BERGER, B. J. (2001). Methionine regeneration and aspartate aminotransferase in parasitic protozoa. *J Bacteriol*. **183**: 4421-4434.
- BERTELLI, M. S. & BRENER, Z. (1980). Infection of tissue culture cells with bloodstream trypomastigotes of *Trypanosoma cruzi*. *J Parasitol*. **66**: 992-997.
- BILLAUT-MULOT, O., FERNANDEZ-GOMEZ, R., LOYENS, M. & OUAISSI, A. (1996). *Trypanosoma cruzi* elongation factor 1- α : nuclear localization in parasites undergoing apoptosis. *Gene*. **174**:19-26.
- BIRNBOIM, H. C. & DOLY, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl Acid Res*. **7**: 1513-1523.
- BIRREN, B., GREEN, E. D., KLAPHOLZ, S., MYERS, R. M. & ROSKAMS, J. (1997). Genome Analysis: a laboratory manual-Analysing DNA, Cold Spring Harbor Laboratory Press.
- BLACKBURN, E. H. & CHALLONER, P. B. (1984). Identification of a telomeric DNA sequence in *Trypanosoma brucei*. *Cell*. **36**: 447-457.
- BLADER, I. J., MANGER, I. D. & BOOTHROYD, J., C. (2001). Microarray analysis reveals previously unknown changes in *Toxoplasma gondii*-infected human cells. *J Biol Chem*. **276**: 24223-24231.

- BOZDECH, Z., ZHU, J., JOACHIMIAK, M. P., COHEN, F. E., PULLIAM, B. & DERISI, J. L. (2003). Expression profiling of the schizont and trophozoite stages of *Plasmodium*. *Genome Biol.* **4**: R9.1-R9.15.
- BRANDÃO, A. A. (1999). Sequenciamento aleatório de cDNAs e análise das regiões não traduzidas em *Trypanosoma cruzi*. Tese de Doutorado. UFRJ. Rio de Janeiro, Brasil.
- BRENER, Z., ANDRADE, Z. A. & BARRAL-NETO, M. (2000). *Trypanosoma cruzi* e doença de Chagas. 2nd ed. Editora Guanabara, Rio de Janeiro.
- BRENIERE, S. F., BOSSENSO, M. F., TELLERIA, J., BASTRENTA, B., YACSIK, N., NOIREAU, F., ALCAZAR, J. L., BARNABE, C., WINCKER, P. & TIBAYRENC, M. (1998). Different behavior of two *Trypanosoma cruzi* major clones: transmission and circulation in young Bolivian patients. *Exp Parasitol.* **89**: 285-295.
- BULGER, E. M., GARCIA, I. & MAIER, R. V. (2003). Induction of heme-oxygenase 1 inhibits endothelial cell activation by endotoxin and oxidant stress. *Surgery.* **134**: 146-152.
- BURLEIGH, B. A. & ANDREWS, N. W. (1995). The mechanisms of *Trypanosoma cruzi* invasion of mammalian cells. *Annu Rev Microbiol.* **49**: 175-200.
- BURLEIGH, B. A. & WOOLSEY, A. M. (2002). Cell signalling and *Trypanosoma cruzi* invasion. *Cell Microbiol.* **11**: 701-11. Review.
- CANO, I., DA SILVEIRA, J. F., RAMIREZ, J. L. (1999). Organization of telomeric and sub-telomeric regions of chromosomes from the protozoan parasite *Trypanosoma cruzi*. *Mol Biochem Parasitol.* **100**: 173-183.
- CANO, M. I., GRUBER, A., VAZQUEZ, M., CORTÉS, A., LEVIN, M. J., GONZÁLES, A., DEGRAVE, W., RONDINELLI, E., ZINGALES, B., RAMIREZ, J. L., ALONSO, C., REQUENA, J. M. & SILVEIRA, J. F. (1995). Molecular karyotype of clone CL Brenner chosen for the *Trypanosoma cruzi* Genome project. *Mol Biochem Parasitol.* **71**: 273-278.
- CASTELLANI, O., RIBEIRO, L. V., & FERREIRA, J. F. (1967). Differentiation of *Trypanosoma cruzi* in culture. *J. Protozool.* **14**: 447-451.
- CASTRO, C., CRAIG, S. P. & CASTANEDA, M. (1981). Genome organization and ploidy number in *Trypanosoma cruzi*. *Mol Biochem Parasitol.* **31**: 273-282.
- CAZZULO, J. J. (1992). Aerobic fermentation of glucose by trypanosomatids. *FASEB J.* **6**: 3153-3161.
- CAZZULO, J. J. (1994). Intermediate metabolism in *Trypanosoma cruzi*. *J Bioenerg and Biomembr.* **26**: 157-164.
- CHIARI, E., DIAS, J. C. P., LANA, M. & CHIARI, C. A. (1979). Hemocultures for parasitological diagnosis of human chagas disease in chronic phase. In: Congresso internacional sobre doença de chagas, Procceding of an internatinal Meeting, Rio de Janeiro, Brazil, Abstr, N-1.
- CHIURILLO, M. A., CANO, I., DA SILVEIRA, J. F. & RAMIREZ, J. L. (1999). Organization of telomeric and sub-telomeric regions of chromosomes from the protozoan parasite *Trypanosoma cruzi*. *Mol Biochem Parasitol.* **100**: 173-83.
- CHUDOVA, D. & SMYTH, P. (2003). Analysis of pattern discovery in sequences using Bayes error rate framework. *Data Mining and Knowledge Discovery J.* **7**: 273-299.
- CHUNG, M. C., GUIDO, R. V., MARTINELLI, T. F., GONÇALVES, M. F., POLLI, M. C., BOTELHO, K. C., VARANDA, E. A., COLLI, W., MIRANDA, M. T. & FERREIRA, E. I. (2003). Synthesis and in vitro evaluation of potential antichagasic hydroxymethylnitrofurazone (NFOH-121): a new nitrofurazone prodrug. *Bioorg Med Chem.* **11**: 4779-4783.
- CLAYTON, C. E. (1988). The molecular biology of the Kinetoplastidae. *Genet Eng.* **7**: 1-56.

- COTRIM, P. C., PARANHOS-BACCALA, G., SANTOS, M. R., MORTENSEN, C., CANO, M. I., JOLIVET, M., CAMARGO, M. E., MORTARA, R. A. & DA SILVEIRA, J. F. (1995). Organization and expression of the gene encoding an immunodominant repetitive antigen associated to the cytoskeleton of *Trypanosoma cruzi*. *Mol Biochem Parasitol.* **71**: 89-98.
- DAROCHA, W. D., BARTHOLOMEU, D. C., MACÊDO, C. D. S., HORTA, M. F., CUNHA-NETO, E., DONELSON, J. E. & TEIXEIRA, S. M. R. (2002). Characterization of cDNA clones encoding ribonucleoprotein antigens expressed in *Trypanosoma cruzi* amastigotes. *Parasitol Res.* **88**: 292-300.
- DE ARAUJO-JORGE, T. C. (1989). The biology of *Trypanosoma cruzi*- macrophage interaction. *Mem Inst Oswaldo Cruz.* **84**: 441-462. Review.
- DE JESUS, A. R., COOPER, R., ESPINOSA, M., GOMES, J. E., GARCIA, E. S., PAUL, S. & CROSS, G. A. (1993). Gene deletion suggests a role for *Trypanosoma cruzi* surface glycoprotein GP72 in the insect and mammalian stages of the life cycle. *J Cell Sci.* **106**: 1023-1033.
- DE SOUZA, W. (1983). A simple method to purify biologically and antigenically preserved bloodstream trypomastigotes of *Trypanosoma cruzi* using DEAE cellulose columns. *Mem Inst Oswaldo Cruz.* **78**: 317-333.
- DENICOLA-SEOANE, A., RUBBO, H., PRODANOV, E. & TURRENS, J. F. (1992). Succinate-dependent metabolism in *Trypanosoma cruzi* epimastigotes. *Mol Biochem Parasitol.* **54**: 43-50.
- DERISI, J. L., LYER, V. R. & BROWN, P. O. (1997). Exploring the metabolic and genetic control of gene expression on a genomic Scale. *Science.* **278**: 680-686.
- Di NOIA, J. M., D'ORSO, I., SANCHEZ, D. O. & FRASCH, A. C. (2000). AU-rich elements in the 3'-untranslated region of a new mucin-type gene family of *Trypanosoma cruzi* confers mRNA instability and modulates translation efficiency. *J Biol Chem.* **275**: 10218-10227.
- DIAZ, M. O. & GALL, J. G. (1985). Giant readthrough transcription units at the histone loci on lampbrush chomosomes of the newt *Notophthalmus*. *Chromo.* **92**: 243-248.
- DIAZ, M. O., BARSACCHI-PILONE, G., MAHON, K. A. & GALL, J. G. (1981). Transcripts from both strands of a satellite DNA occur on lampbrush chomosome loops of the newt *Notophthalmus*. *Cell.* **24**: 649-657.
- DIAZ-ÑEL, A. M., ROSSI, M. S., ESPINOSA, J. M., GUIDA, C., FREITAS, F. A., KORNBLIHTT, A. R., ZINGALES, B., FLAWIA, M. M. & TORRES, H. N. (2000). mRNA encoding a putative RNA Helicase of the DEAD-Box gene family is up-regulated in *Trypanosoma cruzi*. *J. Euk. Microbiol.* **47**: 555-560.
- DIEHL, S., DIEHL, F., EL-SAYED, N. M., CLAYTON, C. & HOHEISEL, J. D. (2002). Analysis of stage-specific gene expression in the bloodstream and the procyclic form of *Trypanosoma brucei* using a genomic DNA-microarrays. *Mol Biochem Parasitol.* **123**: 115-123.
- DOCAMPO, R. (1990). Sensibility of parasites to free radical damage by antiparasitic drugs. *Chem. Biolo. Interact.* **73**: 1-23.
- D'ORSO, I. & FRASCH, A. C. (2001a). Functionally different AU- and G-rich cis-elements confer developmentally regulated mRNA stability in *Trypanosoma cruzi* by interaction with specific RNA-binding proteins. *J Biol Chem.* **276**: 15783-15793.

- D'ORSO, I. & FRASCH, A. C. (2001b). TcUBP-1, a developmentally regulated U-rich RNA-binding protein involved in selective mRNA destabilization in trypanosomes. *J Biol Chem.* **276**: 34801-34809.
- D'ORSO, I., DE GAUDENZI, J. G. & FRASCH, A. C. (2003). RNA-binding proteins and mRNA turnover in trypanosomes. *Trends Parasitol.* **19**: 151-155.
- DUDA, R. O., HART, P. E. & STORK, D. G. (2000). Pattern Classification. 2nd ed. Published by Wiley-Interscience.
- DUDOIT, S., YANG, Y. H., CALLOW, M. J. & SPEED, T. P. (2002). Statistical Methods for identifying genes with differential expression in replicated cDNA microarray experiments. *Stat Sin*; **12**: 111-139.
- DUNCAN, R. (2004). DNA microarray analysis of protozoan parasite gene expression: outcomes correlate with mechanisms of regulation. *TRENDS Parasitol.* **5**: 211-215.
- DVORAK, J. A. & HOWE, C. L. (1976). The attraction of *Trypanosoma cruzi* to vertebrate cells in vitro. *J Protozool.* **23**: 534-537.
- DVORAK, J. A. & HYDE, T. P. (1973). *Trypanosoma cruzi*: interaction with vertebrate cells in vitro. 1. Individual interactions at the cellular and subcellular levels. *Exp Parasitol.* **34**: 268-283.
- ELIAS, C. Q. B. (2001). *Trypanosoma cruzi*- o núcleo em questão. Tese de doutoramento 130p. Universidade Federal de São Paulo-Escola Paulista de Medicina. São Paulo.
- ELIAS, M. C. Q. B., VARGAS, N., ZINGALES, B. & SCHENKMAN, S. (2003). Organization of Satellite DNA in the genome of *Trypanosoma cruzi*. *Mol Biochem Parasitol.* **129**: 1-9.
- ELIAS, M. C., MARQUES-PORTO, R., FREYMULLER, E. & SCHENKMAN, S. (2001). Transcription rate modulation through the *Trypanosoma cruzi* life cycle occurs in parallel with changes in nuclear organization. *Mol Biochem Parasitol.* **112**: 79-90.
- EL-SAYED, N. M., HEGDE, P., QUACKENBUSH, J., MELVILLE, S. E. & DONELSON, J. E. (2000). The African Trypanosome genome. *Int J Parasitol.* **30**: 329-345.
- ENARD, W., KHAITOVICH, P., KLOSE, J., ZOLLNER, S., HEISSIG, F., GIAVALISCO, P., NIESELT-STRUWE, K., MUCHMORE, E., VARKI, A., RAVID, R., DOXIADIS, G. M., BONTROP, R. E. & PAABO, S. (2002). Intra- and interspecific variation in primate gene expression patterns. *Science.* **296**: 340:343.
- ENGEL, J. C., DOYLE, P. S., DVORAK, J. A. (1990). Isolate-dependent differences in the oxidative metabolism of *Trypanosoma cruzi* epimastigotes. *Mol Biochem Parasitol.* **39**: 69-76.
- ENGMAN, D. M., KRAUSE, K. H., BLUMIN, J. H., KIM, K. S., KIRCHHOFF, L. V. & DONELSON, J. E. (1989). A novel flagellar Ca²⁺-binding protein in trypanosomes. *J Biol Chem.* **264**: 18627-18631.
- EPSTEIN, L. M., MAHON, K. A. & GALL, J. G. (1986). Transcription of satellite DNA in the newt. *J Cell Biol.* **103**: 1137-1144.
- FEARNLEY, I. M., RUNSWICK, M. J. & WALKER, J. E. (1989). A homologue of the nuclear coded 49 kd subunit of bovine mitochondrial NADH-ubiquinone reductase is coded in chloroplast DNA. *EMBO J.* **8**: 665-672.
- FLOETER-WINTER, L. M., SOUTO, R. P., STOLF, B. S., ZINGALES, B. & BUCK, G. A. (1997). Can activity of the rRNA gene promoter be used as a marker of speciation? *Exp Parasitol.* **86**: 232-234.
- FURGER, A., SCHURCH, N., KURATH, U. & RODITI, I. (1997). Elements in the 3' untranslated region of procyclin mRNA regulate expression in insect forms of

- Trypanosoma brucei* by modulating RNA stability and translation. *Mol Cell Biol.* **17**:4372-4380.
- GARG, N., POPOV, V. L. & PAPAConstantinou, J. (2003). Profiling gene transcription reveals a deficiency of mitochondrial oxidative phosphorylation in *Trypanosoma cruzi*-infected murine hearts: implications in chagasic myocarditis development. *Biochim Biophys Acta.* **14**: 106-120.
- GIORDANO, R., CHAMMAS, R., VEIGA, S. S., COLLI, W., ALVES, M. J. (1994). An acidic component of the heterogeneous Tc-85 protein family from the surface of *Trypanosoma cruzi* is a laminin binding glycoprotein. *Mol Biochem Parasitol.* **65**: 85-94.
- GIORDANO, R., FOUTS, D. L., TEWARI, D., COLLI, W., MANNING, J. E. & ALVES, M. J. (1999). Cloning of a surface membrane glycoprotein specific for the infective form of *Trypanosoma cruzi* having adesive properties to laminin. *J Biol Chem.* **274**: 3461-3468.
- GOMES, Y. M. (1997). PCR and sero-diagnosis of chronic Chagas's disease. Biotechnological advantages. *Appl Biochem Biotechnolo.* **66**: 107-119.
- GONG, P., CEDERBAUM, A. I. & NIETO, N. (2003). Increased expression of cytochrome P450 2E1 induces heme oxygenase-1 through ERK MAPK pathway. *J. Biol. Chem.* **278**: 29693-29700.
- GONZALEZ, A., PREDIGER, E., HUECAS, M. E., NOGUEIRA, N. & LIZARDI, P. M. (1984). Minichromosomal repetitive DNA in *Trypanosoma cruzi*: its use in a high-sensitivity parasite detection assay. *Proc Natl Acad Sci USA.* **81**: 3356-3360.
- GRUBER, A. & ZINGALES, B. (1993) *Trypanosoma cruzi*: characterization of two recombinant antigens with potential application in the diagnosis of Chagas disease. *Exp Parasitol.* **76**:1-12.
- GRUBER, A. Caracterização de dois genes contíguos de *Trypanosoma cruzi* que codificam antígenos com repetições de epitópos imunodominantes. (1994). 204p. Tese de doutorado-Instituto de Química da Universidade de São Paulo. São Paulo.
- HAYWARD, R. E., DERISI, J. L., ALFADHLI, S., KASLOW, D. C., BROWN, P. O. & RATHOD, P. K. (2000). Shotgun DNA microarray and stage-specific gene expression in *Plasmodium falciparum* malaria. *Mol Microbial.* **35**: 6:14.
- HEHL, A., VASELLA, E., BRAUN, R. & RODITI, I. (1994). A conserved stem-loop structure in 3' untranslated region of procyclic mRNAs regulates expression in *Trypanosoma brucei*. *Proc. Natl Acad Sci USA.* **91**: 370-374.
- HENRIKSSON, J., ASLUND, L., MACINA, R. A., FRANKE, DE CAZZULO, B. M., CAZZULO, J. J., FRASCH, A. C. & PETTERSSON, U. (1990). Chromosomal localization of seven cloned antigen genes provides evidence of diploidy and further demonstration of karyotype variability in *Trypanosoma cruzi*. *Mol Biochem Parasitol.* **42**: 213-223.
- HERNANDEZ, R., DIAZ-DE LEON, F. & CASTANEDA M. (1988). Molecular cloning and partial characterization of ribosomal RNA genes from *Trypanosoma cruzi*. *Mol Biochem Parasitol.* **27**: 275-279.
- HUANG, X. & MADAN, A., (1999). CAP3: A DNA sequence assembly program. *Genome Res.* **9**: 868-877.
- IRIBAR, M. P., TOSI, L. R., CRUZ, A. K. (2003). A processed short transcript of *Leishmania*, ODD1. *Mol Biochem Parasitol.* **127**: 205-208.
- KAWASHITA, S. Y., SANSON, G. F. O., FERNANDES, O., ZINGALES, B. & BRIONES, M. R. S. (2001). Maximum likelihood divergence date estimates based on rRNA gene

- sequences suggest two scenarios of *Trypanosoma cruzi* intraspecific evolution. *Mol Biol Evol.* **18**: 2250-2259.
- KOIDE, T., ZAINI, P. A., MOREIRA, L. M. *et al.* (2004). DNA microarray-based genome comparison of a pathogenic and a non-pathogenic strain of *Xylella fastidiosa* delineates genes important for bacterial virulence. *J Bacteriol.* **186**: 5442-5449.
- KOSLOWSKY, D. J., BHAT, G. J., PERROLLAZ, A. L., FEAGIN, J. E. & STUART, K. (1990). The MURF3 gene of *T. brucei* contains multiple domains of extensive editing and is homologous to a subunit of NADH dehydrogenase. *Cell.* **62**: 901-911.
- KRIEGER, M. A. & GOLDENBERG, S. (1998). Representation of Differential Expression: A new approach to study differential gene Expression in trypanosomatids. *Parasitol Today.* **14**: 165-166.
- KRIEGER, W. S., ARIYANAYAGAM, M. R., FAIRLAMB, A. H., KRAUTH-SIEGEL, R. L. & CLAYTON, C. (2000). Trypanosomes lacking trypanothione reductase are avirulent and show increased sensitivity to oxidative stress. *Mol Microbiol.* **27**: 966-984.
- LEVIN, M. J., MESRI, E., BENAROUS, R., LEVITUS, G., SCHIJMAN, A., LEVY-YEYATI, P., CHIALE, P., RUIZ, A., KAHN, A. & ROSENBAUM, M. (1989). Identification of major *Trypanosoma cruzi* antigenic determinants in chronic Chagas heart disease. *Am. J. Trop. Med. Hyg.* **5**: 530-538.
- LI, J., RUYECHAN, W. T. & WILLIAMS, N. (2003). Stage-specific translational efficiency and protein stability regulate the developmental expression of p37, an RNA binding protein from *Trypanosoma brucei*. *Biochem Biophys Res Commun.* **306**: 918-923.
- LIVAK, K. J. & SCHMITTGEN, D. T. (2001). Analysis of relative gene expression data using Real time quantitative PCR and method. *Methods.* **25**: 402-408.
- LUCHTAN, M., WARADE, C., WEATHERLY, D. B., DEGRAVE, W. M., TARLETON, R. L. & KISSINGER, J. C. (2004). TcruziDB: an integrated *Trypanosoma cruzi* genome resource. *Nucleic Acids Res.* **1**: 32-39.
- MACHADO, C. A. & AYALA, F. J. (2001). Nucleotide sequences provide evidence of genetic exchange among distantly related lineages of *Trypanosoma cruzi*. *Proc Nat Acad Sci USA.* **98**: 7396-7401.
- MACHADO, F. S., MARTINS, G. A., ALIBERTI, J. C., MESTRINER, F. L., CUNHA, F. Q., SILVA, J. S. (2000). *Trypanosoma cruzi*-infected cardiomyocytes produce chemokines and cytokines that trigger potent nitric oxide-dependent trypanocidal activity. *Circulation.* **102**: 3003-3008.
- MAGDESIAN, M. H., GIORDANO, R., ULRICH, H., JULIANO, M. A., JULIANO, L., SCHUMACHER, R. I., COLLI, W. & ALVES, M. J. (2001). Infection by *Trypanosoma cruzi*. Identification of a parasite ligand and its host cell receptor. *J Biol Chem.* **276**: 19382-19389.
- MAIR, G., SHI, H., LI, H., DJOIKENG, A., AVILES, H. O. *et al.* (2000). A new twist in trypanosome RNA metabolism: *cis*-splicing of pre-mRNA. *RNA.* **6**: 163-169.
- MARTIN, F., MARANON, C., OLIVARES, M., ALONSO, C. & LOPEZ M. C. (1995). Characterization of a non-long terminal repeat retrotransposon cDNA (L1Tc) from *Trypanosoma cruzi*: homology of the first ORF with the ape family of DNA repair enzymes. *J Mol Biology.* **247**: 49 - 59.
- MARTIN, P., VEN, V., MOUCHEL, N., JEFFRIES, A. C., HOOD, D. W. & MOXON, E. R. (2003). Experimentally revised repertoire of putative contingency loci in *Neisseria gonorrhoeae* strain MC58: Evidence for novel mechanism of phase variation. *Mol Microb.* **50**: 245-257.

- MASLOV, D. A., NAWATHEAN, P. & SCHEEL, J. (1999). Partial kinetoplast-mitochondrial gene organization and expression in the respiratory deficient plant trypanosomatid *Phytomonas serpens*. *Mol Biochem Parasitol.* **99**: 207-221.
- MASLOV, D., AVILA, H. A., LAKE, J. A. & SIMPSON, L. (1994). Evolution of RNA editing in kinetoplastid protozoa. *Nature.* **368**: 345-348.
- MCCABE, R. E., REMINGTON, J. S. & ARAUJO, F. G. (1985). In vivo and in vitro effects of cyclosporin A on *Trypanosoma cruzi*. *Am J Trop Med Hyg.* **34**: 861-865.
- MINNING, T. A., BUA, J., GARCIA, G. A., MCGRAW, R. A. & TARLETON, R. L. (2003). Microarray profiling of gene expression during trypomastigote to amastigote transition in *Trypanosoma cruzi*. *Mol Biochem Parasitol.* **131**: 55-64.
- MORI, N., MIZUNO, D. & GOTO, S. (1978). Increase in the ratio of 18S RNA to 28S RNA in the cytoplasm of mouse tissues during aging. *Mech Ageing Dev.* **8**: 285-297.
- MOXON, E. R., RAINEY, P. B., NOWAK, M. A. & LENSKE, R. E. (1994). Adaptive evolution of highly mutable loci in pathogenic bacteria. *Current biology.* **4**: 24-33.
- MUKHERJEE, S., BELBIN, T. J., SPRAY, D. C., IACOBAS, D. A., WEISS, L. M., KITSIS, R. N., WITTNER, M., JELICKS, L. A., SCHERER, P. E., DING, A. & TANOWITZ, H. B. (2003). Microarray analysis of changes in gene expression in a murine model of chronic chagasic cardiomyopathy. *Parasitol Res.* **91**: 187-196.
- NOZAKI, T. & CROSS, G. A. M. (1995). Effects of 3' untranslated and intergenic regions on gene expression in *Trypanosoma cruzi*. *Mol Biochem Parasitol.* **75**: 55-67.
- NUNES, L. R., CARVALHO, M. R. C. & BUCK, G. A. (1997). *Trypanosoma cruzi* strains partition into two groups based on the structure and function of the spliced leader and rRNA gene promoters. *Mol Biochem Parasitol.* **86**: 211-224.
- NUNES, L. R., OLIVEIRA, R. C., NUTO, H. N., YAMADA-OGATTA, YABAI, G., ÁVILA, A. R., TOAM, H. K., PROBST, C. M., PAVONI, D. P., MONTEIRO, V. S., SOTOMAIOR, V., FREUD, A., RIBAS, P. D., MIRANDA, A., GOLDENBERG, S. & KRIGER, M. A. (2001). Analysis of the metacyclogenesis process in *Trypanosoma cruzi* through microarray hybridization. Resumos da XXVIII Annual Meeting on Basic Research in Chagas Disease e XVII Annual Meeting of the Brazilian Society of Protozoology.
- OLIVEIRA, R. C., YANAI, G. M., MUTO, N. H. *et al.* (2002) Competitive hybridization on spotted microarrays as a tool to conduct comparative genomic analyses of *Xylella fastidiosa* strains. *FEMS Microbiol Lett.* **10677**: 1-7.
- PAYAO, S. L., SMITH, M. A., WINTER, L. M. & BERTOLUCCI, P. H. (1998). Ribosomal RNA in Alzheimer's disease and aging. *Mech Ageing.* **105**: 265-272.
- PEARSON, W. R. & LIPMAN, D. J. (1988). Improved tools for biological sequence comparison. *Proc Natl Acad Sci USA.* **85**: 2444-2448.
- PEDROSO, A., CUPOLILLO, B. & ZINGALES, B. (2003). Evaluation of *Trypanosoma cruzi* hybrid stocks based on chromosomal size variation. *Mol Biochem Parasitol.* **129**: 79-90.
- PETERSON, M. R., HALL, D. R., BERRIMAN, M., NUNES, J. A., LEONARD, G. A., FAIRLAMB, A. H., & HUNTER, W. N. (2000). The three-dimensional structure of a *Plasmodium falciparum* cyclophilin in complex with the potent anti-malarial cyclosporin A. *J Mol Biol.* **298**: 123-133.
- PETRI, A. & SABIN, C. (2000). Medical statistics at glance. Blackwell Sciences Ltd, Oxford. 138pp.

- PITULA, J., RUYECHAN, W. T. & WILLIAMS, N. (2002). Two novel RNA binding proteins from *Trypanosoma brucei* are associated with 5S rRNA. *Biochem Biophys Res Commun.* **290**: 569-576.
- PORCEL, B. M., TRAN, A.-N., TAMMI, M., NYARADY, Z., RYDAKER, M., URMENYI, T. P., RONDINELLI, E., PETTERSSON, U., ANDERSSON, B. & ASLUND, L. (2000). Gene survey of the Pathogenic Protozoan *Trypanosoma cruzi*. *Genome Res.*, **10**: 1103-1107.
- RECHE, P., ARREBOLA, R., OLMO, A., SANTI, D. V., GONZALEZ-PACANOWSKA, D. & RUIZ-PEREZ, L. M. (1994). Cloning and expression of the dihydrofolate reductase-thymidylate synthase gene. *Mol Biochem Parasitol.* **65**:247-258.
- REIS, E. N., NAKAYA, H. I., LOURO, R., CANAVEZ, F. C., FLATSCHART, A. V. F., ALMEIDA, G. T., EGIDIO, C. M., PAQUOLA, A. C., MACHADO, A. A., FESTA, F., YAMAMOTO, D., ALVARENGA, R., DA SILVA, C. C., BRITO, G. C., SIMON, S. D., MOREIRA-FILHO, C. A., LEITE, K. R., CAMARA-LOPES, L. H., CAMPOS, F. S., GIMBA, E., VIGNAL, G. M., EL-DORRY, H., SOGAYAR, M. C., BARCINSKI, M. A., DA SILVA, A. M. & VERJOVSKI-ALMEIDA, S. (2004). Antisense intronic noncoding RNA levels correlate to the degree of tumor differentiation in prostate cancer. *Oncogene* in press.
- REQUENA, J. M., LOPEZ, M. C., JIMENEZ-RUIZ, A., DE LA TORRE, J. C. & ALONSO, C. (1988) A head-to-tail tandem organization of hsp70 genes in *Trypanosoma cruzi*. *Nucleic Acids Res.* **16**: 1393-1406.
- ROBERTS, T. G., STURM, N. R., YEE, B. K., YU, M. C., HARTSHORNE, T., AGABIAN, N. & CAMPBELL, D. A. (1998). Three small nucleolar RNAs identified from the spliced leader-associated RNA locus in kinetoplastid protozoans. *Mol Cell Biol.* **18**: 4409-17.
- RODRIGUEZ, A., WEBSTER, P., ORTEGO, J. & ANDREWS, N. W. (1997). Lysosomes behave as Ca²⁺-regulated exocytic vesicles in fibroblasts and epithelial cells. *J Cell Biol.* **137**: 93-104.
- ROPERT, C. & GAZZINELLI, R. T. (2000). Signaling of immune system cells by glycosylphosphatidylinositol (GPI) anchor and related structures derived from parasitic protozoa. *Curr Opin Microbiol.* **4**: 395-403. Review.
- ROPERT, C., ALMEIDA, I. C., CLOSEL, M., TRAVASSOS, L. R., FERGUSON, M. A., COHEN, P. & GAZZINELLI, R. T. (2001). Requirement of mitogen-activated protein kinases and I kappa B phosphorylation for induction of proinflammatory cytokines synthesis by macrophages indicates functional similarity of receptors triggered by glycosylphosphatidylinositol anchors from parasitic protozoa and bacterial lipopolysaccharide. *J Immunol.* **166**: 3423-3431.
- RUBBO, H., DENICOLA, A. & RADI, R. (1994). Peroxynitrite inactivates thiol-containing enzymes of *Trypanosoma cruzi* energetic metabolism and inhibits cell respiration. *Arch Biochem Biophys.* **308**: 96-102.
- RUIZ, R. C., FAVORETO, S. J. R., DORTA, M. L., OSHIRO, M. E., FERREIRA, A. T., MANQUE, P. M. & YOSHIDA, N. (1998). Infectivity of *Trypanosoma cruzi* strains is associated with differential expression of surface glycoproteins with differential Ca²⁺ signalling activity. *Biochem J.* **330**: 505-511.
- SAMBROOK, J., FRITSCH, E. & MANIATIS, T. (1989). Molecular cloning: a laboratory manual 2nd ed. Cold Spring Harbor Laboratory Press NY, USA.
- SAMBROOK, J., FRITSCH, E. F. & MANIATIS, T. (1989). Molecular Cloning - a laboratory manual 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

- SANGER, F., NIKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA*. **74**: 5463-5467.
- SAXENA, A., WORTHEY, E. A., YAN, S., LELAND, A., STUART, K. D. & MYLER, P. J. (2003). Evaluation of differential gene expression in *Leishmania major* Friedlinprocyclics and metacyclics using DNA microarray analysis. *Mol Biochem Parasitol*. **129**:103-114.
- SCHENKMAN, R. P., VANDEKERCKHOVE, F. & SCHENKMAN, S. (1993). Mammalian cell sialic acid enhances invasion by *Trypanosoma cruzi*. *Infect Immun*. **3**: 898-902.
- SCHENKMAN, S., EICHINGER, D., PEREIRA, M. E. & NUSSENZWEIG, V. (1994). Structural and functional properties of *Trypanosoma* trans-sialidase. *Annu Rev Microbiol*. **48**: 499-523. Review.
- SCHENKMAN, S., JIANG, M. S., HART, G. W. & NUSSENZWEIG, V. (1991). A novel cell surface trans-sialidase of *Trypanosoma cruzi* generates a stage-specific epitope required for invasion of mammalian cells. *Cell*. **7**: 1117-1125.
- SCHENKMAN, S., KUROSAKI, T., RAVETCH, J. V. & NUSSENZWEIG, V. (1992). Evidence for the participation of the Ssp-3 antigen in the invasion of nonphagocytic mammalian cells by *Trypanosoma cruzi*. *J Exp Med*. **6**: 1635-1641.
- SHOEMAKER, D. D. & LINSLEY, P. S. (2002). Recent developments in microarrays. *Curr. opinion microbiol*. **5**: 334-337.
- SINGH, D., FEBBO, P. G., ROSS, K., JACKSON, D. G., MANOLA, J., LADD, C., TAMAYO, P., RENSHAW, A. A., D'AMICO, A. V., RICHIE, J. P., LANDER, E. S., LODA, M., KANTOFF, P. W., GOLUB, T. R. & SELLERS, W. R. (2002). Gene expression correlates of clinical prostate cancer behavior. *Cancer Cell*. **1**:203-209.
- SKRYPINA, N. A., TIMOFEEVA, A. V., KHASPEKOV, G. L., SAVOCHKINA, L. P. & BEABEALASHVILLI, R. S. H. (2003). Total RNA suitable for molecular biology analysis. *J Biotechnol*. **105**: 1-9.
- SLOOF, P., BOS, J. L., KONINGS, A. F., MENKE, H. H., BORST, P., GUTTERIDGE, W. E. & LEON, W. (1983a). Characterization of satellite DNA in *Trypanosoma brucei* and *Trypanosoma cruzi*. *J Mol Biol*. **167**: 1-21.
- SLOOF, P., MENKE, H. H., CASPERS, M. P. & BORST, P. (1983b). Size fractionation of *Trypanosoma brucei* DNA: localization of the 177-bp repeat satellite DNA and a variant surface glycoprotein gene in a mini-chromosomal DNA fraction. *Nucleic Acids Res*. **11**: 3889-3901.
- SOLARI, A., CAMPILLAY, R., ORTIZ, S. & WALLACE, A. (2001). Identification of *Trypanosoma cruzi* genotypes circulating in Chilean chagasic patients. *Exp Parasitol*. **97**: 226-233.
- SOUTO, R. P., FERNANDES, O., MACEDO, A. M., CAMPBELL, D. A. & ZINGALES, B. (1996). DNA markers define two major phylogenetic lineages of *Trypanosoma cruzi*. *Mol. Biochem. Parasitol*. **83**: 141-152.
- STERN, A., BROWN, M., NICKEL, P. & MEYER, T. F. (1986). Opacity genes in *Neisseria gonorrhoeae*: control of phase and antigenic variation. *Cell*. **47**: 67-71.
- STOLF, B. S., CARVALHO, A. F., MARTINS, W. K., RUNZA, F. B., BRUN, M., HIRATA, R. J. R., JORDÃO NEVES, E., SOARES, F. A., POSTIGO-DIAS, J., KOWALSKI, L. P. & REIS, L. F. (2003). Differential expression of IGFBP-5 and two human ESTs in thyroid glands with goiter, adenoma and papillary or follicular carcinomas. *Cancer Lett*. **191**: 193-202.

- STOLF, B. S., SOUTO, R. P., PEDROSO, A. & ZINGALES, B. (2003). Two types of ribosomal RNA genes in hybrid *Trypanosoma cruzi* strains. *Mol Biochem Parasitol.* **126**: 73-80.
- STURM, N. R., VARGAS, N. S., WESTENBERGER, S. J., ZINGALES, B. & CAMPBELL, D. A. (2003.) Evidence for multiple hybrid groups in *Trypanosoma cruzi*. *Int J Parasitol.* **33**: 269-279.
- SUFRIN, J. R., MESHNICK, S. R., SPIESS, A. J., GAROFALO-HANNAN, J., PAN X. Q. & BACCHI, C. J. (1995). Methionine recycling pathways and antimalarial drug design. *Antimicrob Agents Chemother.* **39**: 2511-2515.
- TEIXEIRA, M. M., & YOSHIDA, N. (1986). Stage-specific surface antigens of metacyclic trypomastigotes of *Trypanosoma cruzi* identified by monoclonal antibodies. *Mol Biochem Parasitol.* **18**: 271-282.
- TEIXEIRA, S. M. & DAROCHA, W. D. (2003). Control of gene expression and genetic manipulation in the Trypanosomatidae. *Genet Mol Res.* **2**: 148-158.
- TEIXEIRA, S. M. R. (1998). Control of gene expression in Trypanosomatidae. *Braz J Med Biol Res.* **31**: 1503-1516.
- TEIXEIRA, S. M. R., COUGHLIN, B. C. KIRCHHOFF, L. V. & DONELSON, J. E. (1996) 5' and 3' regulatory RNA sequences in the developmentally regulated amastin/tuzin gene cluster of *Trypanosoma cruzi*. *Mem Inst Oswaldo Cruz.* **92**: 863-866.
- TEIXEIRA, S. M. R., KIRCHHOFF, L. V. & DONELSON, J. E. (1995). Post-transcriptional elements regulating expression of mRNAs from the amastin/tuzin gene cluster of *Trypanosoma cruzi*. *J Biol Chem.* **270**: 22586-22594.
- TEIXEIRA, S. M. R., RUSSEL, D. G., KIRCHHOFF, L. V. & DONELSON, J. E. (1994). A differentially expressed gene family encoding "amastin," a surface protein of *Trypanosoma cruzi* amastigotes. *J Biol Chem.* **269**: 20509-20516.
- TETAUD, E., BRINGAUD, F., CHABAS, S., BARRETT, M. P. & BALTZ, T. (1994). Characterization of glucose transport and cloning of a hexose transporter gene in *Trypanosoma cruzi*. *Proc Natl Acad Sci USA.* **91**: 8278-8282.
- THE CHIPPING FORECAST. (1999). *Nature Genetics*; 21 (suppl.)
- THOMPSON, J. D., HIGGINS, D. G. & GIBSON, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research.* **22**:4673-4680.
- TIBAYRENC, M. & AYALA, F. (1988). Isoenzyme variability in *Trypanosoma cruzi*, the agent of Chagas' disease: genetical, taxonomical and epidemiological significance. *Evolution.* **42**: 277-292.
- TIBAYRENC, M. (1995). Population genetics of parasitic protozoa and other microorganisms. *Adv Parasitol.* **36**: 47-115.
- TIELENS, A. G. M. & VAN HELLEMOND, J. J. (1998). Differences in energy metabolism between trypanosomatidae. *Parasitol Today.* **14**: 265-271.
- ULGER, C., TORUNER, G. A., ALKAN, M. *et al.* (2003). Comprehensive genome-wide comparison of DNA and RNA level scan using microarray technology for identification of candidate cancer-related genes in HL-60 cell line. *Cancer Genet Cytogenet.* **147**: 28-35.
- URBINA, J. A. & DOCAMPO, R. (2003). Specific chemotherapy of chagas disease: controversies and advances. *TRENDS in parasitology.* **11**: 495-501.
- URBINA, J. A. (1994) Intermediary Metabolism of *Trypanosoma cruzi*. *Parasitol today.* **10**: 107-110

- URMENYI, T. P., BONALDO, M. F., SOARES, M. B. & RONDINELLI, E. (1999). Construction of a normalized cDNA library for the *Trypanosoma cruzi* genome project. *J Eukaryot Microbiol.* **46**: 542-544.
- VAENA DE AVALOS, S., BLADER, I. J., FISHER, M., BOOTHROYD, J. C. & BURLEIGH, B. A. (2002). Immediate/Early response to *Trypanosoma cruzi* infection involves minimal modulation of host cell transcription. *J. Biol Chem.* **277**: 639-644.
- VARGAS, N., PEDROSO, A. & ZINGALES, B. (2004). Chromosomal polymorphism, gene synteny and genome size in *T. cruzi* I and *T. cruzi* II groups. Aceito para publicação.
- VAZQUEZ, M., BEN-DOV, C., LORENZI, H., MOORE, T., SCHIJMAN, A. & LEVIN, M. (2000). The short interspersed repetitive element of *Trypanosoma cruzi*, Sire, is part of Viper, an unusual retroelement related to long terminal repeat retrotransposons. *Proc Natl Acad Sci USA.* **97**: 2128-2133.
- VÊNCIO R. Z. N., BRENTANI H., PATRÃO, D. F. C. & PEREIRA, C. A. B. (2004). Bayesian model accounting for within-class biological variability in SAGE analysis. *BMC Bioinformatics.* Aceito para publicação.
- VERDUN, R. E., PAOLO, N. D., URMENY, T. P., RONDINELLI, E., FRASCH, A. C. C. & SANCHEZ, D. O. (1998). Gene discovery through Expressed Sequence Tag sequencing in *Trypanosoma cruzi*. *Infect Immun.* **66**: 5393-5398.
- WHITE, K. P., RIFTIN, S. A., HURBAN, P., HOGNESS, D. S. (1999). Microarray of *Drosophila* development during metamorphosis. *Science.* **286**: 2179-2184.
- WHO-EXPERT COMMITTEE (2002). Second Report – Control of Chagas Disease - Technical Report Series 905 - World Health Organization (Geneva), 109 pages.
- WILKINSON, S. R., MEYER, D. J., TAYLOR, M. C., BROMLEY, E. V., MILES, M. A., KELLY, J. M., (2002). The *Trypanosoma cruzi* enzyme TcGPXI is a glycosomal peroxidase and can be linked to trypanothione reduction by glutathione or trypanredoxin. *J Biol Chem.* **277**: 17062-17071.
- YANG, Y. H., DUDOIT, S., LUU, P., LIN, D. M., PENG, V., NGAI, J. & SPEED, T. P. (2002). Normalization for cDNA microarrays data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* **15**: 30 e15.
- ZINGALES, B. & COLLI, W. (1985). *Trypanosoma cruzi*: Interaction with Host Cell. *Curr. Topics Microbiol. Immuno.* **117**: 129:152.
- ZINGALES, B., ABUIN, G., ROMANHA, A. J., CHIARI, E. & COLLI, W. (1984) Surface antigens of stocks and clones of *Trypanosoma cruzi* isolated from humans. *Acta Trop.* **41**: 5-16.
- ZINGALES, B., PEREIRA, M. E. S., RIVA, O. P., ALMEIDA, K. A., UMEZAWA, E. S., SOUTO, R. P., VARGAS, N., CANO, M. I., SILVEIRA, J. F., NEHME, N. S., MOREL, C. M., BRENER, Z. & MACEDO, A. (1997a) *Trypanosoma cruzi* genome project: biological characteristics and molecular typing of clone CL Brener. *Acta Tropica.* **68**: 159-173.
- ZINGALES, B., PEREIRA, M. E. S., RIVA, O. P., ALMEIDA, K. A., UMEZAWA, E. S., NEHME, N. S., OLIVEIRA, R. P., MACEDO, A. & SOUTO, R. P. (1997b). Biological parameters and molecular markers of clone CL Brener- The reference organism of the *Trypanosoma cruzi* genome project. *Mem Inst Oswaldo Cruz.* **92**: 811-814.
- ZINGALES, B., SOUTO, R. P., MANGIA, R. H., LISBOA, C. V., CAMPBELL, D. A., COURA, J. R., JASEN, A. & FERNANDES, O. (1998). Molecular epidemiology of American Trypanosomiasis in Brazil based on dimorphism of rRNA and mini-exon gene sequences. *Intern. J. Parasitol.* **28**: 105-112.