UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE QUÍMICA

**Seqüenciamento e Análise de um Banco de cDNA de Glândulas
Salivares de *Rhynchosciara americana* e Caracterização do gene
RaDup.**

Fábio Siviero

Tese de Doutorado

ORIENTADORA

Profa. Dra. Gláucia Maria Machado Santelli

SÃO PAULO

São Paulo, 27 de fevereiro de 2004.

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE QUÍMICA

**Seqüenciamento e Análise de um Banco de cDNA de Glândulas
Salivares de *Rhynchosciara americana* e Caracterização do gene
RaDup.**

Fábio Siviero
Tese de Doutorado

ORIENTADORES

Prof. Dr. Roberto Vicente Santelli (*in memoriam*)

Profa. Dra. Gláucia Maria Machado Santelli

SÃO PAULO

São Paulo, 27 de fevereiro de 2004.

Para minha Mãe Elisabete e meu Pai Marcus,
e para o Santo (*in memoriam*).

Prof. Dr. Roberto Vicente Santelli

O primeiro Roberto Santelli que conheci foi o Professor, seu jeito descontraído e suas avaliações com questões que exigiam interpretação de dados experimentais, geralmente envolvendo algum aspecto da futura profissão dos alunos, sejam químicos, farmacêuticos ou médicos, criaram uma fama que o precedia. Impossível esquecer de sua sacola “de feira” colorida, contendo dezenas de transparências e diversos livros de Bioquímica, até hoje não sei se o uso desta sacola era proposital, com a intenção de prender a atenção do aluno. Não se envergonhava de preparar suas aulas, de rever o conteúdo, de estudar e de perguntar nossa (alunos de pós) opinião sobre suas provas e aulas; era comum vê-lo remexendo seu material didático dias antes de suas aulas, fazendo suas transparências em casa, criando DNAs e enzimas com isopor e papelão para ilustrar suas aulas.

O segundo Roberto que conheci foi o Pesquisador, apaixonado pelo trabalho, dedicou-se por três décadas ao estudo do díptero *Rhynchosciara americana*, introduziu técnicas modernas de biologia molecular neste sistema, que até então era abordado com bioquímica clássica, sempre manteve seus objetivos e o ânimo em relação a *Rhynchosciara*, mesmo quando todos desistiram deste modelo alegando dificuldades no seu estudo. Recentemente ainda participou dos projetos genoma da cana-de-açúcar e *Xylella fastidiosa* com o mesmo afincamento que tinha pela sua linha principal de pesquisa. Sempre conseguia um pretexto para vir ao laboratório nos finais de semana, nas férias, e quando ia “descansar”, estava no litoral procurando larvas do seu objeto de estudo.

O terceiro Roberto que conheci foi o amigo, amizade que não era restrita aos que trabalham em seu laboratório, mas era com todo o Instituto, todos eram tratados com respeito e carinho, alunos, funcionários e professores. Roberto era acessível a todos, nunca deixou de atender um aluno em sua sala, nunca agendou hora para conversar com qualquer um que o procurasse. Este era o mesmo Roberto que conheci, e ele o era com todos.

A amizade do Santo, como era chamado, nunca permitiu que ele fosse somente um orientador, em primeiro lugar era um amigo, sempre preocupado conosco, não somente com nossos projetos, sua amizade era presente dentro e fora da Universidade. No laboratório cheguei a chamá-lo de “Pai Acadêmico” diversas vezes, e ele realmente era um segundo Pai para todos nós. Nos ensinava pessoalmente as suas técnicas já padronizadas para o sistema, dividiu sua própria bancada comigo, e à medida que ficamos numerosos ele se mudou para outra.

Agradeço ao Santo por mostrar-me o universo acadêmico, por ensinar-me a observar através de diferentes prismas, por apontar-me o caminho e incentivar-me a segui-lo sozinho e, mais do que isso, por contagiar-me com sua paixão pela pesquisa e pelo ensino.

A lembrança que sempre terei do Santo é da pessoa Dedicada e Bem Humorada, do Amigo e principalmente do Pai Acadêmico.

Obrigado Santelli.
Fábio

Profa. Dra. Gláucia Maria Machado Santelli

Devo um agradecimento especial à Profa. Dra. Gláucia Maria Machado Santelli, que sempre foi a principal colaboradora de nossa equipe, e que assumiu com muita coragem nossas orientações.

Por tanta energia e dedicação dispensadas conosco, mesmo nos momentos mais tristes. Por incentivar e apoiar nossos projetos e ambições, por nos dar direção e equilíbrio.

Sou breve nestas palavras por não pretender reavivar uma série de lembranças dolorosas, porém esteja certa que independentemente dos acontecimentos, sempre foi merecedora de grande reconhecimento.

Agradeço por sempre ter acreditado (e ainda acreditar) em meu potencial, pelo apoio em todos os momentos e por manter acesa a Ordem da Mosca Negra.

Agradecimentos

Muitos foram os que tiveram importância fundamental para minha formação acadêmica e pessoal nestes últimos anos, estas simples palavras nesta lista de agradecimentos não são o suficiente para expressar minha gratidão, mas é um começo...

Agradeço aos meus pais por todo o amor, e por serem forças da natureza, capazes de remover rios e montanhas de meu caminho.

Agradeço à Miriam por todo o amor e apoio.

À vó Iracema, por sempre acreditar.

Aos meus irmãos na Ordem da Mosca Negra, Alexandre, Paula e Marcelo, pelo companheirismo, amizade, parentesco, pelos bolos e brigadeiros, e também pelas críticas, todas elas.

Ao Dr. Ricardo de Marco, pela sua amizade, pontualidade, e por ser um cientista, contribuindo com valiosas discussões e sugestões.

Aos meus mais recém parentes, Queridos Mestres Sheila e Herbert, pela amizade incondicional, a qual tenho tido a oportunidade de desfrutar por todos estes anos e espero que por muitos outros.

Ao indefenestrável Dr. Guilherme Marson, pelo seu humor constante, surtos de hiper-criatividade, pelos momentos “Discovery Channel” no Shigueo’s e por me apresentar ao lendário Boturoca.

Ao meu primo Paulo Ponte, por sempre ter sido um amigo, mesmo em seus momentos de patifaria; por ter sido um anti-herói, e por tentar corromper minha alma e meus ideais em todos os meus períodos de “Workaholism” e jejum social.

Ao casal Remo e Raphaela, por serem amigos sinceros, por compartilharem comigo tantos bons momentos no bandex, e por se preocuparem tanto, mas tanto, com minha saúde.

Ao Prof. Dr. Bayardo Baptista Torres, pela amizade e incentivo acadêmico.

Aos amigos César, Jú, Gláucia, Rodrigo Luiz, pela constância e aos amigos Aurélio, Camila, Pedro Kawamura, Bartira, Cássio, Marcelo Nunes, Juliana, Daniela, Celso, Myriam, Roberto, Eric, Wilton, Ari e Augusto por sempre serem vocês mesmos.

À Profa. Dra. Maria Júlia Manso Alves e à Melissa, por todo apoio nos diferentes momentos deste projeto, pela amizade e pelas aulas e protocolos de imuno.

À Profa. Carla Columbano pela amizade, ricas discussões acadêmicas e chocolates.

À Egle e à Giselle pelas normas ABNT, que doravante regem minha vida acadêmica.

Ao Christian e à Ariadne por me ensinarem a domar vetores amotinados.

À Profa. Dra. Bianca Zingalles pelas discussões sobre purificação de proteínas e por me ensinar a lecionar para batalhões em laboratórios didáticos.

À Luci, pela amizade e pelo apoio.

Aos amigos Reinaldo e Fernanda, pela amizade.

À Luíze e ao Igor, pela companhia e por entenderem minha necessidade de mais experimentos.

Ao Departamento de Bioquímica pela infra-estrutura.

À CAPES pelo auxílio financeiro concedido em forma de bolsa.

À Fapesp pelo suporte aos diferentes projetos sobre *Rhynchosciara*.

ABREVIATÖES	2
RESUMO.....	3
SUMMARY	4
1. INTRODUÇÃO	5
1.1. <i>RHYNCHOSCIARA AMERICANA</i>	5
1.2. ENDOREPLICAÇÃO	8
1.3. COMPLEXOS CDKS/CICLINAS ORQUESTRAM O CICLO CELULAR	9
1.4. ESTADOS DAS ORIGENS DE REPLICAÇÃO	11
1.5. FATORES DE LICENCIAMENTO	13
1.5.1. <i>ORC</i>	13
1.5.2. <i>Cdc6/18</i>	13
1.5.3. <i>Cdt1/Dup</i>	14
1.5.4. <i>MCMs</i>	14
1.5.5. <i>CDC45</i>	15
1.6. MODELO ATUAL DE LICENCIAMENTO DAS ORIGENS DE REPLICAÇÃO.....	15
1.7. INIBINDO A RE-REPLICAÇÃO	18
1.8. E OS ENDOCICLOS?	20
1.8.1. <i>Elementos Atuando nos Endociclos</i>	21
1.8.2. <i>Sciarídeos e outros sistemas</i>	24
2. OBJETIVOS.....	26
3. RESULTADOS E DISCUSSÖES	27
3.1. GERAÇÃO E ANÁLISE DE ESTs	27
3.1.1. – <i>Anotação</i>	32
3.1.2. – <i>Categorização</i>	36
3.1.3. – <i>Polimorfismos</i>	46
3.1.4. – <i>Códon Usage</i>	48
3.2. - RADUP	51
3.2.1. – <i>Estrutura Gênica e Localização</i>	51
3.2.2. – <i>Proteína e Expressão</i>	60
3.2.3. – <i>RaDup no gênero Rhynchosciara</i>	67
4. CONCLUSÖES	69
5. PERSPECTIVAS	70
6. MATERIAIS E MÉTODOS	71
6.1. - ANIMAIS	71
6.2. – CONSTRUÇÃO DA BIBLIOTECA DE CDNA	71
6.3. – SEQUENCIAMENTO DE DNA	72
6.4. – ANÁLISE E MONTAGEM DOS ESTs E SEQUÊNCIAS GENÔMICAS	73
6.5. – CATEGORIZAÇÃO SEGUNDO O GENE ONTOLOGY CONSORTIUM	74
6.6. – PRODUÇÃO DE SUB-BIBLIOTECAS GENÔMICAS	76
6.7. – HIBRIDIZAÇÕES: SOUTHERNS E NORTHERNS.....	78
6.8. – REAÇÕES DE PCR.....	79
6.9. – VARREDURA DO BANCO GENÔMICO DE <i>RHYNCHOSCIARA</i>	81
6.10. – PURIFICAÇÃO DE DNA DE FAGOS EM CACL	82
6.11. – CONSTRUÇÃO DO PLASMÍDEO PGEX-4T2-RADUP.....	82
6.12.-INDUÇÃO DE PROTEÍNA DE FUSÃO EM PGEX	84
6.13. – PURIFICAÇÃO DE PROTEÍNA DE FUSÃO EM COLUNA DE GLUTATIONA IMOBILIZADA EM AGAROSE.....	85
6.14.-PURIFICAÇÃO DE CORPOS DE INCLUSÃO	85
6.15. – PRODUÇÃO DE SOROS ANTI-RADUP	87
6.16. – MARCAÇÃO DE WESTERNS.....	87
6.17. – HIBRIDIZAÇÕES <i>IN SITU</i>	88
7. REFERÊNCIAS BIBLIOGRÁFICAS	89
8. SÚMULA CURRICULAR.....	106

Abreviações

2P	Segundo Período
AA	Aminoácidos
ACS	ARS Consensus Sequence
APC/C	Anaphase Promoting Complex/Cyclosome
ARS	Autonomously Replicating Sequence
B	Pufe B2 expandido
C	Pufe C3 expandido
Cab	Cabeças de moscas adultas
CAF1	Chromatin Assembly Factor 1
CAK	Quinase Ativadora de CDKs
CDK	Quinase Dependente de Ciclina
Cdt1	Cdc10 Dependent Transcript 1
DDK	Quinase Dependente de Dbf4
DmDup	<i>Drosophila melanogaster</i> Dup
DNA	Ácido Desoxirribonucleico
Dup	Double Parked
EST	Expressed Sequence Tag
GO	Gene Ontology
Int	Intestinos de indivíduos adultos
IR	Início de Rede
MCM	Mini-Chromosome Maintenance
NBT	Nitro Blue Tetrazolium
ORF	Open Reading Frame
Ova	Ovários de moscas adultas
Pós-RC	Complexo Pós-Replicativo
PP2A	Protein Phosphatase 2A
Pré-RC	Complexo Pré-Replicativo
RaDup	<i>Rhynchosciara americana</i> Dup
RNA	Ácido Ribonucleico
SNP	Single Nucleotide Polymorphism
SU	Sequências Únicas
TC	Consenso Tentativo
TIGR	The Institute for Genome Research
UTR	Untranslated Region

Resumo

Durante o desenvolvimento deste projeto adotou-se como estratégia o sequenciamento de ESTs, com a finalidade de encontrar mensagens relacionadas com desenvolvimento, metabolismo e principalmente amplificação/politenização em glândulas salivares de *Rhynchosciara americana*, um díptero (Sciarídeo) que apresenta cromossomos politênicos e amplificação gênica rigidamente regulada ao longo do desenvolvimento larval, tanto neste tecido quanto em outros. Um total de 8193 ESTs foi gerado, estas foram anotadas e categorizadas segundo os termos do Gene Ontology Consortium, proporcionando uma visão geral do *status* metabólico, como em um *Northern* eletrônico, de um ponto importante no desenvolvimento desta espécie, quando surgem amplificações gênicas específicas e a glândula salivar necessita secretar as proteínas do casulo. Outros frutos deste sequenciamento foram a determinação de 91 polimorfismos e a criação de uma tabela de códon usage.

Diversos ESTs foram identificados com potencial envolvimento com os endociclos observados neste tecido, destes, RaDup e RaMCM5 foram selecionados para estudo. Suas regiões genômicas foram isoladas e suas localizações cromossômicas foram identificadas, em relação a RaDup, toda a porção codificante de seu mensageiro e 12kb de DNA genômico contendo seu gene foram sequenciados, revelando sua estrutura gênica. Anticorpos foram produzidos para detectar esta proteína, gerando evidências de sua participação tanto na replicação mitótica como nos endociclos presentes nas glândulas salivares. A localização cromossômica de RaDup é um dado muito interessante, pois pela primeira vez um pufe amplificado é relacionado com um gene regulatório.

Summary

In this work EST sequencing was used as strategy to find messages related to development, metabolism and polyteny/amplification in salivary glands of *Rhynchosciara americana*, a dipteran (Sciaridae) that shows in this tissue giant polytene chromosomes and gene amplification tightly regulated throughout development of the larvae. A total of 8193 EST sequences were generated, annotated and categorized using Gene Ontology Consortium terms, providing a general view of the metabolic status, like an electronic Northern, of an important point in development of the larvae, that shows where specific genes are amplified and the salivary gland needs to secrete the proteins to form the cocoon. Other data include determination of 91 SNPs and a statistic of codon usage.

Several ESTs were identified with potential connection to endocycles, from these RaDup and RaMCM5 were selected for further studies. Both chromosomal loci were identified and genomic regions isolated, for RaDup the coding region of its mRNA and 12kb of genomic region were completely sequenced, revealing its gene structure, and antibodies were raised against this protein, making evident data about its involvement in replication in mitotic cells and in endocycles in salivary glands. About the chromosomal locus of RaDup, it becomes very interesting, because for the first time one amplified puff can be related to a regulatory gene.

1. Introdução

1.1. *Rhynchosciara americana*

O díptero (Sciaridae) *Rhynchosciara americana* tem sido objeto de estudo desde os anos 50 descrito por Nonato e Pavan (1951) como *Rhynchosciara angelae*. Posteriormente, teve sua taxonomia revista por Breuer (1969), sendo constatado que na realidade esta espécie era *R. americana* já descrita previamente por Wiedemann (1821). Suas características o tornaram um interessante sistema paradigma para o estudo da biologia dos cromossomos politênicos (Dreyfus, Nonato, Breuer e Pavan, 1951), já que estes estão presentes nas glândulas salivares e outros órgãos deste gênero. Apresenta ainda amplificações gênicas em diferentes regiões dos cromossomos das células das glândulas salivares rigidamente reguladas ao longo do desenvolvimento, comprovadas inicialmente por experimentos auto-radiográficos de incorporação de timidina tritiada (Ficq e Pavan, 1957) e por citofotometria (Rudkin e Corlette, 1957). Na época destacava-se o fato de ser possível a visualização de um padrão na atividade gênica durante a vida do indivíduo, e pela primeira vez evidenciou-se que o conteúdo genômico de uma célula poderia naturalmente ser alterado, ou seja, o número de cópias de um gene/região poderiam ser modificadas diferencialmente.

Em relação ao seu estudo em laboratório podem-se citar como vantagens em relação a outros insetos as grandes proporções de suas larvas, o que facilita a sua manipulação e permite a obtenção de ácidos nucléicos com o sacrifício de poucos animais; o fato de apresentar um ciclo de vida longo (aproximadamente 70 dias), o que permite maior controle sobre experimentos envolvendo diferentes pontos do desenvolvimento; facilidade de obtenção de ótimas preparações citológicas e, finalmente, o que talvez seja sua principal característica, o desenvolvimento sincronizado de indivíduos irmãos, desde a postura dos ovos até a emergência dos imagos, passando pela metamorfose, o que nos permite construir séries de dados com espécimes dos mesmos grupos.

A politenização dos cromossomos das glândulas salivares tem início desde o começo da vida larval, porém no segundo período do quarto estágio de desenvolvimento este processo passa a ser sincrônico, é nesta ocasião que as larvas param de se alimentar. O conteúdo de DNA destas células continua a aumentar até o

sexto período do mesmo estágio de desenvolvimento, quando as larvas estão terminando o casulo. A politenia na glândula salivar de *Drosophila* é obtida através de sucessivas replicações a partir de um par (2C) até 9 vezes produzindo um número máximo de 1024C (Swift, 1962), sendo que o número de replicações pode variar em diferentes órgãos deste organismo. Em *Rhynchosciara*, medidas do aumento do volume nuclear ao longo do desenvolvimento da glândula salivar a partir do dia da eclosão dos ovos, indicam um mínimo de 8 duplicações (512C) (Machado-Santelli, resultados não publicados).

O momento em que as larvas começam a secretar o casulo coincide com o último ciclo de replicação, que corresponde à fase de grande atividade gênica na glândula salivar como pode ser observado pela ocorrência de inúmeros pufes nos seus cromossomos (Breuer e Pavan, 1955; Machado-Santelli e Basile, 1973; Guevara e Basile, 1973; Glover *et al* 1981). Além dos pufes de RNA, ou de expressão, que ocorrem normalmente em cromossomos politênicos, observa-se nesta espécie os chamados pufes de DNA em que, além do RNA há a síntese localizada de DNA. Os produtos de transcrição dos pufes amplificados podem ser diretamente relacionados com peptídeos secretados para a construção do casulo (Winter *et al*, 1977).

Em *Rhynchosciara* tanto a politenização quanto a amplificação constituem diferenciações terminais nas glândulas salivares, sendo que estas são as últimas modificações celulares antes deste tecido perder sua função e regredir por apoptose durante a metamorfose.

Segundo dados obtidos em experimentos envolvendo extirpação de glândulas endócrinas e cérebros, implantes de glândulas salivares na cavidade abdominal de larvas e injeção de hormônio ecdisona (Amabis e Cabral, 1970; Amabis *et al*, 1977; Stocker *et al*, 1984), o “timing” de abertura e fechamento dos pufes seria regulado por via hormonal.

Um modelo de amplificação gênica muito bem estudado está presente nas células foliculares de *Drosophila* durante a oogênese, quando dois clusters de genes coriônicos presentes nos cromossomos 3 e X, que codificam proteínas constituintes do córion (membrana de revestimento dos ovos), são amplificados (Calvi *et al*, 1998). Acredita-se que a amplificação destes genes justifica-se pela necessidade de grandes quantidades de mRNA para suprir uma maciça maquinaria de síntese protéica em um curto período de tempo.

O cluster do cromossomo 3 de *Drosophila* é o mais bem detalhado, possui quatro genes coriônicos (s18, s15, s19, s16 nesta ordem) dispostos com o mesmo sentido de transcrição, precedidos por um elemento de controle da amplificação (ACE3) (de Cicco e Spradling, 1984) e intercalados por regiões intensificadoras de amplificação (AERs) (Delidakis and Kafatos 1989; Orr-Weaver et al. 1989); estes elementos reguladores em *cis* foram somente identificados em *Drosophila*. Em sciarídeos algumas regiões de origem foram identificadas através de análises de mapeamento de replicons em géis bidimensionais neutro/neutro (N/N) e neutro/alcalino (N/A), com tamanhos próximos de 1kb em pufe II/9 de *Sciara* (Liang e Gerbi, 1994; Liang *et al*, 1993) e 6kb em pufe C3 de *Rhynchosciara* (Yokosawa *et al*, 1999; Soares et al, 2003). Dentro desta região-origem em *Rhynchosciara* são utilizados múltiplos sítios de início de replicação, em *Sciara* foram determinados a posição onde é disparada a replicação e o sítio de ligação de ORC (complexo de reconhecimento de origens, descrito a seguir), porém não foi confirmada a presença de nenhum elemento de controle agindo em *cis* nestes sistemas.

Até o momento foi encontrada apenas uma unidade de transcrição para cada um dos grandes pufes de DNA estudados em *Rhynchosciara* (C3, C8 e B2) (Penalva et al, 1997; Frydman et al, 1993; Santelli et al, no prelo), ao contrário do que ocorre nos clusters coriônicos de *Drosophila* nos cromossomos X e 3, que possuem 6 e 4 genes respectivamente.

Tanto no caso de células foliculares de *Drosophila* quanto nas glândulas salivares de sciarídeos postula-se um modelo para a amplificação, chamado de casca de cebola (Onion Skin Model, Spradling, 1981) (Figura 1), onde a origem de replicação (ou região-origem) seria reutilizada mais de uma vez dentro de um mesmo ciclo celular gerando um gradiente de amplificação. As regiões próximas da origem ou região-origem seriam mais amplificadas que as regiões mais distantes, este modelo implicaria que os mecanismos normais de controle do ciclo celular estariam sendo modificados, permitindo os endociclos.

Outro aspecto interessante é que como não existem terminadores específicos para a amplificação, a extensão da região amplificada seria determinada pela primeira forquilha de replicação, os pontos onde ela cessasse o movimento seriam os limites do gradiente, estes pontos poderiam ocorrer por algum tipo de impedimento estérico. A baixa velocidade de síntese de DNA nos amplicons, observada tanto em *Rhynchosciara* e *Drosophila* como em outros organismos (Spradling e Leys, 1988; Cordeiro e

Meneghini, 1973) também poderia ser causada por motivos espaciais, conseqüentes das sucessivas ondas de replicação, produzindo novas fitas de DNA que continuam agregadas ao cromossomo.

Figura 1 – Modelo Onion Skin, origens de replicação são disparadas sucessivas vezes, acarretando em um gradiente de amplificação.

1.2. Endoreplicação

O processo de replicação de DNA nos eucariotos é extremamente conservado, tanto na duplicação em si quanto nos mecanismos de controle. Homólogos de proteínas envolvidas nestes processos são encontrados em organismos desde leveduras a eucariotos superiores, muitas vezes podemos perceber um certo aprimoramento ao longo da escala evolutiva, tais como, um maior número de proteínas envolvidas ou redundância em certos pontos de controle do ciclo celular, mas a semelhança na sequência de eventos necessária a estes processos se mantém.

Esta semelhança se mantém também na iniciação da replicação, um processo com múltiplas etapas que restringe o disparo da replicação a apenas uma vez por ciclo e por origem. Quando este passo sofre um desvio que permite o re-disparo da replicação nas origens mais de uma vez por ciclo celular, este passa a ser chamado de endociclo ou ciclo endoreplicativo, caracterizado pelo aumento do conteúdo de DNA genômico sem divisão celular subsequente.

Podem-se encontrar diversos exemplos de endociclos em protistas, plantas e animais (incluindo insetos, moluscos e mamíferos) (Figura 2), onde os endociclos geram células poliplóides, politênicas, células multinucleadas (sincícios), e pufes de DNA amplificado em casos específicos, que é o principal interesse dos estudos em *Rhynchosciara* das últimas décadas. Em todos os exemplos que encontrarmos podemos perceber que o conteúdo genômico aumentado em tecidos destes organismos representa

uma vantagem, tornando o endociclo fundamental na estratégia de desenvolvimento destas espécies.

Uma célula poliplóide, por exemplo, pode ter como vantagem sua massa aumentada, ou uma maior capacidade metabólica e uma proteção natural contra danos no DNA.

Figura 2 – Diferentes exemplos de endoreplicação (adaptado de Edgar e Orr-Weaver, 2001); a – células nurse do ovário de *Drosophila* após o 6º ciclo; b – células nurse do ovário de *Drosophila* até o 4º ciclo; c – trofoblastos de mamíferos; d – endomitose de megacariócitos de mamíferos.

1.3. Complexos CDKs/Ciclinas orquestram o ciclo celular

Em eucariotos o ciclo celular é regido por diferentes quinases dependentes de ciclina (CDK) que se associam a ciclinas específicas para cada fase da vida da célula, fosforilando determinados alvos que vão ativar os mecanismos necessários ao avanço do ciclo celular, respeitando o término dos passos anteriores através do uso de pontos de checagem, a atividade de CDKs está diretamente ligada a modulação da atividade das origens bem como mudanças nos “footprints” destas.

Experimentos de fusão de células de mamíferos foram os primeiros a demonstrar este tipo de controle por pontos de checagem (Rao e Johnson, 1970). Células em G1 fundidas com células em S iniciam prematuramente a síntese de DNA, enquanto que células em G2 fundidas com células em S são incapazes de reiniciar a

replicação, e também são impedidas de entrar na fase M até que finalize a fase S do núcleo em replicação. Estas observações revelaram algumas características do ciclo celular: células em S produzem um sinal (atualmente compreendido como ponto de checagem) que impede o início precoce da mitose; células na fase S possuem um fator positivo em favor da replicação agindo em *trans*; somente células em G1 são competentes para iniciar a replicação.

No modelo atual adotado para explicar o ciclo celular acredita-se que a replicação do DNA é controlada pela ação sequencial de CDKs acopladas a ciclinas da fase S e depois à ciclinas mitóticas. Em leveduras uma única CDK, Cdc2 em *S. pombe* e Cdc28 em *S. cerevisiae* controla o início das fases S e M através de sua associação com ciclinas diferentes (Riabowol *et al*, 1989; Vas *et al*, 2001). Em metazoários a fase S é iniciada e progride sob o controle de complexos de ciclina E/Cdk2 e ciclina A/Cdk2, e a mitose é conduzida pelos complexos Ciclina A/Cdk1 e Ciclina B/Cdk1, outros eventos do ciclo celular, inclusive a transição G0 → G1, envolvem ciclina D complexada com Cdk4 ou com Cdk6 (Mihaylov *et al*, 2002; Nishitani e Lygerou, 2002).

Células com DNA danificado ou com a replicação incompleta são impedidas de iniciar a fase M, através da inativação da Cdk1 (Cdc2 em leveduras) até o reparo ou o término da replicação.

Os mecanismos de regulação parecem ser simplificados em células que empregam o endociclo, havendo a eliminação da expressão de componentes de outras fases do ciclo que não são mais necessários, por exemplo, alguns tipos de células em endoreplicação não expressam Cdk1 nem seus ativadores Cdc25, Ciclina A e Ciclina B, conseqüentemente desviando-se da mitose.

Outro ponto de controle crucial é o impedimento do reinício da replicação na fase S ou G2. Blow e Laskey (1988) postularam a existência de um fator positivo para a replicação que só fosse capaz de acessar a cromatina no término da fase M, e que seria destruído ou inativado após o disparo da replicação. Esta molécula hipotética foi chamada de fator de licenciamento. Nos últimos anos diversos novos elementos relacionados ao disparo da replicação foram descobertos, refinando este modelo e modificando o conceito de fatores de licenciamento, estendendo este *status* a diferentes proteínas com a identificação de novos fatores positivos e negativos que agiriam na regulação da competência das origens de replicação.

Como o conteúdo das células na fase M induz as células em S ou G1 a entrarem em M prematuramente nos experimentos de fusão, mesmo que a replicação não tenha sido finalizada, e esses dados associados ao fato de a superexpressão ectópica de Cdk1 (Cdc2) levar a célula à mitose a partir de qualquer fase do ciclo celular, indica que a atividade de CDKs mitótica é dominante sobre os controles das outras fases do ciclo.

Uma pergunta a ser respondida seria sobre o que ocorreria se as CDKs mitóticas fossem inativadas. A resposta começaria a surgir com as evidências encontradas com o uso de um mutante termo-sensível de *S. pombe* (*cdc2^{ts}*) (Broek *et al*, 1991), neste o ciclo celular parava em G2 em temperaturas restritivas, e quando exposto a temperaturas mais altas apresentava Cdc2 completamente destruída e reiniciava a replicação quando retornava a temperaturas permissivas, em vez de entrarem em mitose. A deleção de Cdc13, uma ciclina mitótica, leva a levedura a reiniciar a replicação do DNA por múltiplas vezes, gerando células poliplóides (Hayles *et al*, 1994). Resultados semelhantes também foram obtidos com a expressão ectópica de inibidores de CDKs tanto em *S. pombe* quanto em *S. cerevisiae*, em células de mamíferos o “knock out” condicional de Cdk1 (Itzhaki *et al*, 1997) e o uso de drogas inibidoras da sua atividade (Usui *et al*, 1991) induzem a re-replicação.

Todos estes dados sugerem que as CDKs mitóticas são as responsáveis pela indução da mitose e pela prevenção do reinício da replicação durante S e G2, assim, com base nestas observações, um novo aspecto seria acrescentado ao modelo de ciclo celular, a separação temporal das fases S e M seria consequência da atividade das diferentes CDKs. Este novo conceito é importante para explicar o modelo atual de como é o mecanismo que impede o reinício da replicação.

1.4. Estados das Origens de Replicação

Origens de replicação são os pontos da cromatina onde as forquilhas de replicação começam seu movimento, em procariotos e leveduras são muito bem caracterizadas. Em *E. coli* o *locus* *oriC* é a única origem de replicação de seu genoma, possuindo 245 bp e regiões ricas em AT. Em leveduras, as origens distribuídas na cromatina, abrangem aproximadamente 100 bp, e são chamadas ARS (Autonomously Replicating Sequence) e possuem uma região consenso de 11 bp rica em AT, chamada de ACS (ARS Consensus Sequence), que é fundamental para sua função. Nos

metazoários são conhecidas diferentes origens de replicação, porém não há nenhum consenso definido, e muitas vezes as origens são definidas por regiões grandes, que têm a abertura da bolha replicativa em diferentes posições dentro destas regiões ou várias pequenas bolhas abertas simultaneamente. Sem dúvida eucariotos superiores possuem muito a ser caracterizado sobre suas origens de replicação e sua regulação.

Experimentos de “footprinting” *in vivo* mostraram que as origens de replicação são protegidas por complexos protéicos durante todo o ciclo celular, mas entre a mitose até o início da fase S a extensão desta proteção é maior, o que sugere que as origens possuam dois estados, um apto para a replicação ocorrendo em G1, e outro estado onde as origens não são competentes para a replicação, ocorrendo em S e G2 e caracterizado por um complexo menor protegendo estas regiões da cromatina.

O complexo presente nas origens antes do evento da replicação é chamado de complexo pré-replicativo (Pré-RC), momento quando a origem é dita licenciada, após o disparo da replicação parte deste complexo se desmonta e a estrutura remanescente chama-se complexo pós-replicativo (Pós-RC), este complexo permanecerá desta forma até o fim da mitose.

O período compreendido entre o final da metáfase até meados da fase G1 é caracterizado por uma baixa atividade de CDKs, proporcionada pela destruição das ciclinas de G2 mediada através do complexo promotor da anáfase (Anaphase Promoting Complex – APC), que se encontra ativo nesta ocasião. Todas as observações em diferentes organismos sugerem que neste período ocorra a montagem do complexo pré-replicativo nas origens, tornando-as aptas para replicar na fase S.

Este mecanismo que somente permitiria a instalação do Pré-RC em um período de baixa atividade de CDKs seria um primeiro nível de controle que impediria o reinício da replicação, oferecendo uma repressão ampla sobre diversos fatores simultaneamente. Ao longo desta introdução serão descritos ainda outros mecanismos atuando para garantir que o DNA seja duplicado uma vez por ciclo, muitas vezes transmitindo uma idéia de redundância.

O complexo pré-replicativo é formado basicamente por ORC (Origin Recognition Complex), Cdc6/18, Cdt1/Dup, seis MCMs (Mini-Chromosome Maintenance) e Cdc45, sendo que estes elementos são encaixados nas origens nesta ordem. Estas proteínas também são chamadas de fatores de licenciamento, devido à sua função. Dados recentes sugerem que outras proteínas também podem fazer parte deste complexo, tais como MCM10. Não causaria estranheza a descoberta de novos

participantes no processo de licenciamento das origens, dada a sua complexidade e importância.

1.5. Fatores de Licenciamento

Os componentes do Pré-RC são extensivamente estudados em praticamente todos os grandes organismos paradigma: leveduras, insetos, anfíbios e mamíferos. A seguir são apresentadas as principais características conhecidas de cada um.

1.5.1. ORC

O complexo de reconhecimento de origens é composto por seis proteínas, Orc1 a Orc6, que são as primeiras a ocupar as origens de replicação durante o ciclo celular, ligam-se a estas de forma dependente de ATP, processo provavelmente mediado por Orc1 e Orc5, que são subunidades que possuem sítios de ligação de ATP. As ORCs foram primeiramente purificadas e caracterizadas em levedura (*S. cerevisiae*) por Bell e Stillman (1992) devido sua habilidade de se ligar ao ACS. Experimentos de deleção e mutação mostraram que todas as Orcs são proteínas essenciais.

Mesmo tendo sido encontrados homólogos destas proteínas em diversos eucariotos, apenas em *S. cerevisiae* foi identificada uma sequência consenso para o seu sítio de ligação nas origens.

1.5.2. Cdc6/18

A proteína Cdc6, inicialmente descrita em *S. cerevisiae*, possui um homólogo em *S. pombe* chamado Cdc18, sua sequência de aminoácidos possui similaridade com Orc1 e possui um domínio de ligação de ATP que é fundamental para sua função (Bell *et al*, 1995).

A depleção desta proteína essencial impede o início da replicação do DNA, mas em leveduras não impede o avanço da mitose, gerando um fenótipo letal, o que foi a primeira evidência de que fatores de licenciamento poderiam participar de um ponto de controle entre a fase S e o início da mitose.

1.5.3. Cdt1/Dup

Cdt1/Dup é um fator essencial para a replicação, visto que linhagens de levedura que não possuem o gene de Cdt1 apresentam defeitos similares aos causados pela perda de Cdc6/18.

Inicialmente caracterizado em *S. pombe*, Cdt1 (cdc10 dependent transcript 1) foi descrito como um gene expresso após o estímulo de cdc10, um fator de transcrição específico para a transição G1/S (Hofmann e Beach, 1994). Foram identificados homólogos em *Drosophila*, *Xenopus*, humanos, ratos e camundongos, todos envolvidos com o disparo da replicação. Em *Drosophila* é chamado de Dup (Whittaker *et al*, 2000), devido ao mutante *double parked*, e foi co-localizado com complexos MCMs nas origens de replicação dos clusters coriônicos no momento da amplificação (Claycomb *et al*, 2002).

Postula-se que Cdt1 em conjunto com Cdc6/18 são os responsáveis pela inclusão do complexo MCM nas origens. Em Cdt1 de camundongos foi reportada a capacidade de se ligar a DNA simples e dupla fita de forma independente de sequência e conformação (Yanagi *et al*, 2002). Dados obtidos com células NIH3T3 expressando Cdt1 de camundongo sugerem que o gene Cdt1/Dup possui um forte potencial oncogênico.

1.5.4. MCMs

O complexo de proteínas MCM (Mini-Chromosome Maintenance) é composto por seis proteínas (MCM2 a MCM7), e atua como helicase, função necessária à maquinaria de replicação, viajando com as forquilhas durante a replicação. Foram descobertas em *S. cerevisiae* durante varreduras para descobrir mutações que causassem a perda de mini-cromossomos (Maine *et al*, 1984).

As seis proteínas possuem uma alta similaridade entre si, principalmente em uma região de 240 aminoácidos contendo um domínio de ligação de ATP, formam um complexo hexamérico de aproximadamente 600 kDa com composição estequiométrica de cada uma das seis proteínas. Todas se mostraram essenciais e são conservadas de leveduras a mamíferos.

As subunidades MCM4-6-7 são chamadas de catalíticas, possuindo atividade de ATPase e fraca atividade de helicase quando dois destes trímeros se

associam, enquanto que as subunidades MCM2-3-5 são regulatórias. Acredita-se que a conformação necessária à atividade correta de helicase dependente de ATP seja adquirida apenas quando as seis proteínas formam o complexo. Interessantemente em arqueobactérias uma única proteína forma um homohexâmero com função de helicase replicativa.

1.5.5. CDC45

A proteína Cdc45 foi identificada em *S. cerevisiae* em mutantes de ciclo celular sensíveis ao frio que quando na temperatura restritiva as células permaneciam na interfase indefinidamente sem replicar o DNA. Caracterizada como um fator essencial para a replicação e envolvida na iniciação do processo. De fato todos os seus homólogos encontrados em outros organismos apresentam um comportamento muito similar e um envolvimento com o início da replicação, experimentos em extratos celulares de *Xenopus* indicam que Cdk2 controla a ligação de Cdc45 na cromatina e que sua depleção reduz a síntese de DNA em 90%.

Em *Drosophila*, Cdc45 é muito expressa nos primeiros estágios de desenvolvimento embrionário, e seu padrão de marcação imunológica sugere uma participação na amplificação dos sítios coriônicos, consistentes com os estágios característicos em que esta ocorre. Seu gene também parece estar sob a ação do fator de transcrição E2F, o que indica uma expressão regulada de forma dependente do ciclo celular (Loebel *et al*, 2000).

Dados obtidos por imunoprecipitação indicam que Cdc45 interage com Orc2, MCM2, MCM3, MCM4, MCM5 e MCM7 (Loebel *et al*, 2000), liga-se tardiamente ao Pré-RC, mas antes do disparo da replicação (Zou e Stillman, 1998), acredita-se que sua principal atuação nas origens seja o recrutamento da DNA Polimerase α .

1.6. Modelo Atual de Licenciamento das Origens de Replicação

Licenciamento das origens como a entendemos hoje se inicia com a ligação de ORC à origem de replicação de maneira dependente de ATP, em leveduras este complexo permanece ligado à cromatina por todo o ciclo, mas há indícios de que em

eucariotos superiores a remoção de ORC possa fazer parte da estratégia para evitar o disparo indevido da replicação.

Ao final da mitose, um momento de baixa atividade de ciclinas-CDKs, Cdc6/18 e Cdt1/Dup são adicionados às origens (Figura 3), o posicionamento destas proteínas é dependente de ORC, mas aparentemente a ligação destas à cromatina independe da presença prévia de qualquer uma das duas. Cdc6/18 e Cdt1/Dup interagem fisicamente e agem sinergicamente no recrutamento dos dois complexos formados pelas proteínas MCM2-7. Estudos *in vitro* envolvendo a montagem do Pré-RC com extratos de ovos de *Xenopus* indicam ainda a participação de Cdc6/18 na etapa de ligação de ORC à origem (Sun *et al*, 2002).

Figura 3 – Primeiras etapas da montagem do Pré-RC

Após a montagem dos complexos MCMs a estrutura presente na origem é considerada licenciada (Figura 4), mas outro fator ainda participa no Pré-RC antes do início da replicação. MCM10 é uma proteína identificada na mesma varredura responsável por encontrar as outras proteínas MCMs, aparentemente seu papel é estabilizar o complexo MCM2-7 na cromatina.

Figura 4 – Estado licenciado das origens de replicação.

A atividade de CDKs associadas a ciclinas específicas da transição G1/S, Cdk2-Ciclina E em eucariotos superiores, começa a surgir pouco antes da fase S, neste momento o Pré-RC sofre modificações que levarão a inclusão de Cdc45 no complexo, a abertura do DNA na origem e ao recrutamento de DNA polimerases α e ϵ , RPA (Replication Protein A) e PCNA (Proliferating Cell Nuclear Antigen) (Figura 5). As supostas funções de RPA durante a replicação são estabilizar fitas simples de DNA no momento de sua síntese (Wold, M.S., 1997) e regular a polimerização através da interação com outras proteínas presentes na forquilha. PCNA, referido por alguns autores como um grampo móvel (“sliding clamp”), é um fator envolvido na processabilidade da maquinaria de replicação e no recrutamento do complexo CAF1 (Chromatin Assembly Factor 1), que por sua vez traz as “Core Histones”, proporcionando a formação dos nucleossomos durante a replicação (Krude & Keller, 2001; Haracska *et al*, 2002).

Além da ação de CDKs para a ativação das origens licenciadas também é necessária a atuação de DDKs (Dbf4 Dependent Kinases), assim como as ciclinas regulam a atividade das CDKs, Dbf4 regula a atividade das DDKs, seu acúmulo na transição G1/S somente permite a atividade dessas quinases neste período de tempo (Nishitani e Lygerou, 2002).

Diversos autores sugerem que DDKs agem localmente nas origens de replicação, fosforilando proteínas MCMs (MCM2 e MCM4), enquanto as CDKs atuam globalmente regendo a formação de diferentes complexos e modulando as atividades das proteínas envolvidas nos passos iniciadores da replicação. No entanto os alvos de fosforilação de CDKs não são muito definidos, em alguns casos (MCM4 e Cdc45) há dúvida se a fosforilação foi mediada por DDKs ou CDKs. A maioria das proteínas do

Pré-RC possuem possíveis sítios de fosforilação, mas não são sempre utilizados em todos os eucariotos.

Figura 5 – Ativação do Pré-RC e disparo da replicação.

1.7. Inibindo a re-replicação

Uma vez as origens estando ativadas a função de helicase dos complexos MCMs é adquirida, a partir do momento que a replicação se inicia Cdc45 e os complexos MCMs movem-se junto com as forquilha de replicação. O Pré-RC sem estes componentes é convertido ao Pós-RC, esta conversão é fundamental para garantir apenas um disparo de replicação por fase S, outras modificações ainda ocorrem nas origens para evitar que novos fatores de licenciamento sejam atraídos.

Em leveduras, algumas subunidades de ORC sofrem fosforilação com o intuito de evitar sua associação com Cdc6/18 após o início da síntese de DNA, mas permanecem ligadas à cromatina durante o ciclo celular. Em *Xenopus* foi observada a migração de ORC para o citoplasma durante a mitose e em *Drosophila* e mamíferos Orc1 é desligada das origens e pelo menos em mamíferos dados sugerem sua

degradação via ubiquitinação (Li e DePamphilis, 2002). De modo geral a estratégia aparenta ser evitar que ORC associe-se a outros fatores de licenciamento e não à origem.

Cdc6/18 é fosforilado e degradado em leveduras, em *Xenopus* e mamíferos é fosforilado e exportado do núcleo, para depois ser degradado. Sua fosforilação parece ser via CDKs, e impede sua interação com ORC, fosforilado se torna alvo para ubiquitinação e posterior degradação pelo proteassomo 26S.

Em todos os eucariotos a proteína Cdt1/Dup surge e desaparece em um período bem definido na transição G1/S, mesmo que seu mRNA esteja presente. Mais especificamente esta proteína é degradada rapidamente após o início da replicação, seus níveis são regulados não só por controle da sua expressão mas também por proteólise. Em leveduras e mamíferos Cdt1 é excluída do núcleo dependentemente de Cdc28 e destinada à degradação por intermédio de SCF (Skp1-Cul1/Cdc53; Li *et al*, 2003) que proporcionará sua ubiquitinação.

Eucariotos superiores desenvolveram ainda um refinado método de regular o licenciamento através de um regulador negativo que inativa Cdt1/Dup, este novo fator chamado Geminina (Geminin) foi descoberto em *Xenopus* como sendo um substrato para o complexo APC/ciclossomo, sua precipitação a partir de extratos celulares revelava uma segunda proteína fortemente ligada a Geminina, tanto em *Xenopus* quanto em humanos foi constatado que esta proteína se tratava de Cdt1/Dup.

Geminina possui um domínio de destruição que é reconhecido pelo sistema APC/C e que vai direcioná-la para degradação no período de licenciamento das origens. Sua destruição é essencial para o prosseguimento do ciclo celular, células com uma mutação neste domínio de destruição são incapazes de iniciar a fase S.

A alta afinidade de Geminina por Cdt1/Dup inibe sua ligação ao Pré-RC em formação e sua atividade de recrutamento dos MCMs. Mesmo sendo antagonistas estas duas proteínas não são co-expressas, Geminina durante G2 e S parece atuar de forma redundante com a finalidade de inibir qualquer quantidade de Cdt1/Dup que tenha escapado da proteólise.

O controle negativo exercido por Geminina pode ser sobrepujado com quantidades extras de Cdt1/Dup, mas sua depleção não leva necessariamente ao licenciamento, o que reforça a idéia de redundância no controle negativo sobre as origens.

Segundo Lygerou e Nishitani (2002), Geminina pode fazer parte de um ponto de controle, ligando as CDKs/Ciclinas presentes na mitose à próxima rodada de síntese de DNA, sua presença coincide com a atividade de CDKs durante o ciclo celular e sua degradação se inicia com a ativação do APC que também é responsável pela redução da atividade das CDKs (Figura 6).

Figura 6 – Período de atividade de complexos envolvidos na regulação da replicação.

1.8. E os Endociclos?

Não há dúvidas que ainda existe muito a ser elucidado em relação aos endociclos. Muitos dos elementos envolvidos no processo já são conhecidos, novos participantes e novas funções desempenhadas por proteínas já conhecidas estão surgindo a medida que pesquisas se aprofundam no assunto.

Evidências em diferentes sistemas paradigma apontam os endociclos utilizando a mesma maquinaria de replicação e grande parte dos mesmos reguladores que a replicação de uma célula normal com o ciclo celular completo. Ou seja, os metazoários não desenvolveram novas proteínas para efetuar um trabalho que a célula já possui larga experiência em fazê-lo, que é replicar o DNA, em vez disso parecem ter

derivado novos mecanismos utilizando os recursos previamente disponíveis para iniciar os endociclos.

Por outro lado, esta nova habilidade de aumentar o conteúdo genômico não é fruto de um simples relaxamento nos rígidos e redundantes dispositivos de controle que impedem o re-disparo da replicação, certamente possui um controle tão exato e rígido quanto os existentes no ciclo celular canônico, visto que em todos os exemplos de politenização e amplificação não é permitido a re-utilização das origens dentro de uma mesma “fase S”. Outro aspecto que reforça a idéia de complexidade da regulação dos endociclos é o fato destes estarem ligados ao desenvolvimento do organismo, o que sugere que estejam sob influência de sinais exógenos. Um exemplo bastante extremo de controle sobre os endociclos ocorre em células epidérmicas de *Manduca sexta* enquanto lagarta, estas tem o conteúdo genômico aumentado até 32C e a partir do momento em que aumenta os níveis de ecdisona elas voltam aos ciclos mitóticos e reduzem sua ploidia a 2C (Kato *et al*, 1987).

1.8.1. Elementos Atuando nos Endociclos

Os endociclos também estão sob controle de complexos Ciclinas/CDKs, em *Drosophila* foi constatada participação de Ciclina E/Cdk2 na amplificação dos genes coriônicos (Lilly e Spradling, 1996; Calvi *et al*, 1998), e a expressão de *Dacapo* (p27), um inibidor de Ciclina E, se mostrou capaz de impedir a amplificação coriônica, porém a superexpressão de Ciclina E não induz a amplificação ou a politenização, mostrando que Ciclina E é necessária, mas não suficiente, para este fenômeno.

O fator de transcrição formado pela associação das proteínas E2F e DP, descrito anteriormente participando da transição G1/S, também se mostrou necessário para o progresso dos endociclos. Relacionadas à transcrição do gene Ciclina E, que ocorre sempre pouco antes da fase S do endociclo, estas proteínas mostram-se como fundamentais para o estabelecimento e manutenção de endociclos. Os níveis de Ciclina E/Cdk2 oscilam ao longo do endociclo, o que concorda com a necessidade de um período de baixa atividade de CDKs para a renovação das proteínas do Pré-RC, gerando uma periodicidade que lembra uma fase S e uma G dentro do endociclo. Acredita-se que o complexo Ciclina E/Cdk2 atua diretamente nos Pré-RCs fosforilando alguns de seus componentes, sendo que um provável candidato é Orc2 (Calvi *et al*, 1998; Duronio *et al*, 1998).

Assim como as ciclinas mitóticas são marcadas para degradação antes da transição G1/S, acredita-se que o oscilador Ciclina E/Cdk2 seja mantido por via proteolítica, além do controle transcricional, já que se observou a presença do mensageiro de Ciclina E mesmo quando não há a presença desta proteína no núcleo.

Outro participante importante para os endociclos é o APC (Anaphase Promoting Complex), também chamado de ciclossomo, composto por onze subunidades conhecidas, possui a atividade de ubiquitina ligase E3 e é responsável por adicionar “etiquetas de poli-ubiquitina” de forma específica em pelo menos doze proteínas envolvidas com a regulação da mitose e com o disparo da replicação, entre estas estão as Ciclinas A e B, Cdc6/18 e Geminina. Estas proteínas marcadas são degradadas pelo complexo proteassomo 26S, desta forma APC é capaz de desligar a atividade dos complexos formados por ciclinas mitóticas e Cdk1, destruir ou inativar os promotores da mitose, disparar os eventos responsáveis pela separação das cromátides e permitir a montagem do Pré-RC.

Os diferentes substratos do APC/C são destruídos sequencialmente em diferentes períodos do ciclo celular, seguindo uma rígida programação, Ciclina A por exemplo é degradada na pro-metáfase, Cdc6 em G1 enquanto Geminina na transição G1/S. A regulação do APC/C fica a cargo de FIZZY e FIZZY-RELATED (Cdc20 e Cdh1 em *S. cerevisiae* respectivamente), conferindo-lhe especificidade de substrato e “timing” apropriado.

Estas proteínas são capazes de reconhecer dois domínios de destruição diferentes (d-box), FIZZY possui a propriedade de ligar-se a substratos que possuam o domínio RXXL e os alvos de FIZZY-RELATED podem possuir tanto RXXL como o domínio KEN. Basicamente estas proteínas são ativadas sequencialmente permitindo a atividade do APC sobre dois grupos de substratos em diferentes momentos (metáfase e G1). Interessantemente a atividade do APC parece ser regulada tanto positivamente quanto negativamente por Ciclina B/Cdk1. Para ser ativado e poder ubiquitinar os polipeptídeos reconhecidos por FIZZY, o APC deve ser fosforilado por este complexo CDK ao mesmo tempo em que FIZZY-RELATED também sofre fosforilação e possui sua atividade inibida, à medida que a atividade de Ciclina B/Cdk1 diminui, FIZZY-RELATED deixa de ser inibido e os seus substratos passam a sofrer degradação (Zur e Brandeis, 2002).

Em relação aos endociclos, APC parece desempenhar um papel importante nos primeiros passos da transição entre o ciclo mitótico e o endociclo, degradando os

reguladores mitóticos. Após esta etapa dados indicam que este complexo está envolvido com a diferenciação entre politenia e poliploidia (Kashevsky *et al*, 2002), sendo que em células poliplóides participa da degradação da proteína Securina, que resultará na ativação da Separase e conseqüente separação das cromátides, já em células em politenização o APC estaria inativo, mantendo a atividade constitutiva da Securina e a integridade do complexo Coesina, que mantém as cromátides-irmãs unidas.

A hipótese da utilização dos componentes usuais do Pré-RC e da maquinaria de replicação durante os endociclos é sustentada por diversas evidências, no entanto, os mecanismos que regem a passagem de um ciclo mitótico para um endociclo com objetivo de politenização, poliploidização ou amplificação de diferentes *loci* não são claros e são poucos os elementos conhecidos que possam estar envolvidos nesta transição.

Em leveduras surgiram as primeiras evidências da possível participação de Cdt1/Dup e Cdc6/18 nos endociclos: a superexpressão de Cdc18 leva à replicação contínua em *S. pombe*, enquanto que na presença de Cdt1 as quantidades requeridas de Cdc18 para induzir a re-replicação são muito menores, o que sugere que Cdt1/Dup pode ser um participante importante durante os endociclos.

Em *Drosophila* foram co-localizados nos amplicons coriônicos proteínas constituintes do ORC, bem como Cdt1/Dup e Cdc45 (Royzman *et al*, 1999; Whittaker *et al*, 2000), e também foram identificados mutantes de Orc2, Dbf4, Mcm6 e Cdt1/Dup incapazes de amplificar corretamente os clusters coriônicos e produzindo ovos envoltos em cascas finas e defeituosas (Whittaker *et al*, 2000), estas observações sugerem a participação destes elementos nos endociclos desempenhando os mesmos papéis que nos ciclos mitóticos, porém recentemente surgiram indícios de um comportamento diferenciado para Cdt1/Dup durante a amplificação conforme descrito a seguir.

Tanto em extratos de ovos de *Xenopus*, como em *S. pombe* ou *S. cerevisiae*, Cdt1/Dup parece ser necessário apenas durante a montagem do Pré-RC, sendo desligado da cromatina após o disparo da replicação e exportado no núcleo depois do início da fase S (Maiorano *et al*, 2000).

Porém nos clusters amplificados das células foliculares de *Drosophila* surge um comportamento inusitado, observa-se que a proteína Cdt1/Dup apresenta marcação nas origens de replicação, e desempenha sua função recrutando o complexo MCM, no entanto após o disparo da síntese de DNA a marcação de Cdt1/Dup na origem desvanece e começa a surgir um padrão de marcação na forquilha de replicação,

evidenciado através de técnicas de microscopia confocal de alta resolução e de deconvolução (Claycomb *et al*, 2002), tais dados indicam que neste caso Cdt1/Dup está viajando com a maquinaria de síntese de DNA.

A observação de Cdt1/Dup participando da fase de elongação contradiz as observações anteriores em outros sistemas, vale lembrar que os dados anteriores foram obtidos em sistemas “normais” de replicação, e que a amplificação gênica é um sistema peculiar. Segundo os próprios autores sugerem, a presença de Cdt1/Dup na elongação pode ser necessária para a manutenção do complexo MCM, mantendo-o ligado à cromatina ou modulando sua atividade de helicase, visto que as condições estéricas nestas bolhas de replicação são bastante drásticas, como pode ser evidenciado pela sua baixa velocidade de síntese de DNA, aproximadamente 28 vezes mais lenta que em uma replicação normal.

1.8.2. Sciarídeos e outros sistemas

Faltam muitas peças para a elucidação dos mecanismos de controle dos diferentes tipos de endociclos, *Drosophila* demonstrou ser um modelo formidável para o avanço das pesquisas sobre este tema e certamente ainda proporcionará muitos resultados. Em relação aos endociclos o alto grau de detalhamento já conhecido sobre seu desenvolvimento neste sistema nos faz notar grandes diferenças em relação a outros organismos modelo que utilizam este artifício, tais como sciarídeos e plantas.

É difícil a comparação da amplificação dos sciarídeos com a observada em *Drosophila*, dado que os tecidos são diferentes e possuem objetivos diferentes. A amplificação coriônica em *Drosophila* se mostra necessária para a ovogênese que ocorre em uma hora, um período de tempo curto para uma síntese protéica de grandes proporções. Esta relação tempo/massa de proteína tem sido utilizada como justificativa para a amplificação gênica de um modo geral. Porém em Sciarídeos a amplificação que ocorre em glândulas salivares e está relacionada à síntese das proteínas do casulo, fundamental para a sua metamorfose, a relação tempo/massa de proteína é bem diferente. O tempo para realização do processo de secreção do casulo em Sciarídeos é muito mais longo do que o da *Drosophila* para proteger os ovos.

Em Sciarídeos há por se elucidar a existência de seqüências reguladoras atuando em cis nos *loci* amplificados, sendo que nestes sistemas apenas algumas regiões-origem foram determinadas. Um experimento envolvendo a inserção de

grandes cassetes de DNA contendo o *locus* C3 de *Rhynchosciara* em *Drosophila* não produziu pufes em nenhum tecido, apesar de ter sido detectada a expressão do gene C3 (Soares *et al*, 2003), além das dificuldades técnicas há a possibilidade de os possíveis elementos reguladores deste pufe não serem reconhecidos pelas células de *Drosophila*. Dadas as diferenças entre os sistemas torna-se claro o potencial dos Sciarídeos como modelo de amplificação.

2. Objetivos

O projeto descrito neste trabalho teve por metas identificar elementos envolvidos com a regulação e disparo da amplificação gênica em *Rhynchosciara americana*, e sua posterior caracterização.

Dentre as alternativas de metodologia o seqüenciamento aleatório (“shotgun”) de uma biblioteca de cDNA mostrou produzir mais resultados. Esta biblioteca fora construída em um momento do desenvolvimento das larvas em que a probabilidade de se encontrar mensagens relacionadas à amplificação era maior. Os ESTs gerados serviriam como ponto de partida para um detalhamento posterior, sendo usados como sondas para varreduras (“screening”) de um banco genômico de *Rhynchosciara* construído em bacteriófago lambda DASH II, e suas seqüências ainda permitiriam o desenho de oligonucleotídeos para reações de seqüenciamento e RACE, permitindo a obtenção da seqüência completa dos mensageiros de nosso interesse. Tanto os mensageiros como os insertos genômicos seriam utilizados como sondas para hibridizações *in situ*, proporcionando a determinação da localização cromossômica de nossas seqüências alvo.

Estas informações são valiosas para a geração de anticorpos contra uma série de proteínas interessantes, o que poderia trazer um aspecto mais citológico aos estudos deste gênero. Este seqüenciamento gerando novas informações a respeito da biologia molecular deste modelo teria ainda outro fruto, que seria a possibilidade de abertura de diversas frentes de trabalho, baseadas em aspectos do próprio gênero, ou seja, o uso de sondas heterólogas e tentativas de reações cruzadas com anticorpos contra outros organismos deixariam de ser a única opção e passariam a ser uma alternativa a ser analisada em cada caso.

3. Resultados e Discussões

3.1. Geração e Análise de ESTs

Durante a execução do projeto foram geradas ESTs de um banco de cDNA construído a partir de glândulas salivares de *Rhynchosciara* no período em que se inicia a síntese do casulo, fase referida doravante de Início de Rede (IR). Neste ponto do desenvolvimento as larvas estão começando o último ciclo de replicação politênica nas glândulas salivares e a amplificação gênica de uma parte dos pufes de DNA característicos deste gênero. A escolha deste estágio de desenvolvimento foi feita visando maximizar a probabilidade de detecção de mensageiros relacionados aos ciclos endoreplicativos, tanto politenização como amplificação.

Após o processamento das leituras, onde as seqüências de vetor foram aparadas e as de baixa qualidade foram descartadas, foram geradas 8193 ESTs úteis, 7108 leituras a partir da extremidade 5' e 1085 a partir da extremidade 3', estas não são redundantes entre si e representam 8193 clones diferentes. Uma seqüência foi considerada útil quando apresentava um mínimo de 150 bases contíguas com qualidade inferida pelo Phred ≥ 20 (Ewing *et al*, 1998 a, b), a média obtida das leituras foi de 778 bp e a média de bases com qualidade ≥ 20 foi de 362,4 bp/EST. Todas estas seqüências foram submetidas ao GenBank ("Accession Numbers": CK807874 - CK815989).

Montando estes ESTs por alinhamento de bases (Phrap) (Gordon *et al*, 1998), foram produzidos 846 contigs e 1734 singlets, o que totaliza 2580 Seqüências Únicas (SU). Porém não se pode esperar que cada seqüência única seja um transcrito diferente, pois muitos dos ESTs pertencentes à mesma mensagem não se sobrepõem ou podem estar presentes no banco ESTs representando "splicings" alternativos e pré-mRNAs, além de polimorfismos e regiões de baixa qualidade que dentro das condições de estringência adotadas podem levar a formação de contigs diferentes de um mesmo transcrito.

Analizando manualmente o resultado da comparação das SU com o banco não redundante do GenBank (nr - Blastx), comparando-se os primeiros hits de cada EST que apresentou similaridade com alguma seqüência, foi possível estimar que 8% das Seqüências Únicas sejam redundantes, assim podem estar representados neste banco aproximadamente 2374 transcritos diferentes.

Considerando-se que, filogeneticamente, o genoma da *Rhynchosciara* possui características próximas do genoma da *Drosophila* (13647 genes codificantes de proteínas e 18471 transcritos diferentes, conforme a versão 3.1 do FlyBase¹) pode-se supor que durante este seqüenciamento aproximadamente 13% dos transcritos e 17% dos genes de *R. americana* estão representados nesta coleção de ESTs.

As traduções destas seqüências quando comparadas às seqüências depositadas no GenBank (Blastx – banco de dados não-redundante – nr), revelaram que 83% dos EST e 69,5% das Seqüências Únicas possuem similaridade com alguma proteína depositada no banco público. Destas 34,19% mostraram proteínas de *Drosophila* como maior semelhante (“best hit”), e 32,05% apresentaram maior semelhança com proteínas de *Anopheles*, os “best hits” dos outros 33,76% destes ESTs se distribuem em proporções bem menores entre outros organismos conforme mostra a Figura 7.

Figura 7 – Distribuição das similaridades de proteínas entre *Rhynchosciara* e outros organismos, 83 % dos ESTs apresentaram similaridade com algum organismo presente no GenBank.

¹ www.fruitfly.org/annot/

As similaridades médias entre os ESTs de *Rhynchosciara* e as proteínas de organismos de genoma completo estão detalhadas na figura 8, entendendo-se por similaridade a soma das identidades e das substituições conservativas provenientes das comparações proteína-proteína dentro das janelas de comparação, as diferenças dos valores observados entre ESTs e SU são decorrentes do tamanho das seqüências, os ESTs são leituras das extremidades 5' ou 3' de mensagens e possuem porções significativas de regiões não-traduzidas (5'/3' UTR), já as Seqüências Únicas por possuírem uma região codificante maior apresentam valores mais elevados de similaridade, devido sua janela de comparação ser maior.

Figura 8 – Similaridade média entre *R. americana* e outros organismos de genoma completo, baseada nas traduções dos ESTs, considerou-se como similaridade a soma das identidades e das substituições conservativas.

Comparando nucleotídeos entre os ESTs de *R. americana* e a base de dados do GenBank observa-se que 29,8% destas seqüências possuem identidade significativa com alguma seqüência depositada, 5,58% dos EST apresentam maior identidade com seqüências de *Drosophila* e 3,36% com seqüências de *Anopheles* (Figura 9), com identidade média de 84,4% e 84,3% respectivamente.

Não é incomum encontrar nas respostas de Blastx proteínas de humanos, camundongos, entre outros organismos distantes de *Rhynchosciara* como “best hit”, em nossos resultados estes organismos representam aproximadamente 30% das respostas, pode-se apontar diferentes motivos para encontrarmos organismos muito divergentes de *Rhynchosciara* ocupando a posição de maior semelhante nas comparações entre as traduções de EST e o banco de proteínas do Genbank: proteínas muito conservadas entre diferentes filos (ex: proteínas “housekeeping”, histonas) quando comparadas entre si podem apresentar os mesmos valores de E-values e de Score, e ao listar tais comparações outro critério que não a semelhança seria adotado; outras possibilidades seriam a origem comum destas proteínas; e a convergência ao longo da evolução de proteínas de organismos diferentes a uma estrutura primária semelhante, vale lembrar que mesmo havendo outro organismo como o mais semelhante muito frequentemente *Drosophila* e *Anopheles* ocupam as posições subseqüentes.

Figura 9 – Similaridade de nucleotídeos entre os ESTs de *R. americana* e outros organismos, 29,8% das seqüências mostraram significativa semelhança com algum organismo presente no GenBank.

É possível notar na resposta do Blastn uma presença mais acentuada de insetos, e que os organismos mais distantes de *Rhynchosciara* aparecem em uma proporção bem menor, e novamente a maior parte das seqüências possui maior homologia com *Drosophila*. Tanto nas comparações de proteínas como de nucleotídeos os dados sugerem uma proximidade filogenética maior entre *Rhynchosciara* e *Drosophila*, organismos conhecidamente próximos, por análise de características morfológicas e de desenvolvimento² (Figura 10), apesar de tal análise indicar *Anopheles* com maior proximidade de *Rhynchosciara*.

Figura 10 – Dendrograma gerado no TaxBrowser (Wheeler *et al*, 2000) baseado em aspectos morfológicos e genômicos para alguns dos organismos que surgiram nas comparações das seqüências de *Rhynchosciara*.

² <http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html/>

Porém este quadro se limita aos organismos mais presentes no banco de dados, uma vez que não existem tantas seqüências de outros dípteros depositadas no GenBank, tais como *Ceratitis* e *Mayetiola*, comparáveis aos genomas completos de *Drosophila* e *Anopheles*, assim à medida que mais seqüências de dípteros e de *Rhynchosciara* forem depositadas no GenBank estes gráficos podem ser bastante modificados.

3.1.1. – Anotação

As seqüências geradas foram anotadas manualmente com base nos resultados dos blastx e tblastn, as seqüências foram designadas como “putativas” quando o E-value destas comparações era menor ou igual a 10^{-20} , entre este valor e 10^{-5} foram anotadas como “similar a”, e quando as respostas do blast continham a tradução de um ORF não anotado ou com uma proteína anotada como hipotética estas seqüências eram anotadas como “proteína hipotética conservada”; mensagens com resultados de blast com E-value acima de 10^{-5} ou “no hit found” foram anotadas como “proteína desconhecida”. Estas ESTs anotadas como codificando proteínas desconhecidas podem corresponder a novos peptídeos ou a ESTs “curtos” formados principalmente por regiões não traduzidas dos mensageiros.

Como glândulas salivares são tecidos especializados em secreção e o IR é caracterizado pelo início da construção do casulo, já havia a expectativa de encontrar níveis significativos de mensagens de proteínas ribossômicas e de proteínas conhecidamente constituintes do casulo, de fato, 15,82% dos ESTs foram anotados como mensagens de proteínas ribossômicas e 0,93% correspondiam aos pufes B2 e C8 (Santelli *et al*, 1991) conforme detalhado na Tabela 1. Interessantemente não foram seqüenciadas mensagens do pufe C3, o que corrobora com estudos anteriores onde foi determinado que este pufe se expande e é expresso depois de B2 e de C8 (Santelli *et al*, 1991), o que ocorrerá depois de alguns dias do IR.

Outro aspecto interessante é que 2,78% das mensagens possuem similaridade com pufes de outros dípteros (*Trichosia*, *Bradisia* e *Sciara*) além de pufes já seqüenciados de *Rhynchosciara* (B2 e C8), podendo representar mensagens de pufes ainda não estudados neste modelo. Similaridades entre proteínas de pufes de diferentes espécies já são conhecidas, existindo inclusive sugestões de famílias protéicas,

indicando uma origem comum destes genes (Penalva *et al*, 1997; Santelli *et al*, no prelo). A hipótese de estes mensageiros corresponderem a transcritos de pufes ainda não caracterizados está sendo averiguada por nosso grupo.

Conforme o objetivo inicial de encontrar mensagens ligadas ao ciclo celular, foram identificadas diversas mensagens relacionadas direta e indiretamente a sua progressão/regulação e ainda encontradas outras desconhecidas, que foram selecionadas após a categorização e podem representar candidatos a participar dos endociclos em *Rhynchosciara*, conforme descrito posteriormente neste texto.

Tabela 1 – lista dos 60 ESTs mais representativos e suas proporções relativas (contigs).

Anotação	% ESTs	Anotação	% ESTs
Unknown Protein 1	2,28	Unknown Protein 13	0,49
Putative 60S acidic ribosomal protein P1	1,89	Unknown Protein 14	0,48
Unknown Protein 2	1,53	Unknown Protein 15	0,46
Unknown Protein 3	1,27	Unknown Protein 16	0,45
Putative 40S ribosomal protein S18	1,23	Unknown Protein 17	0,45
Similar to puff C4B protein [T. pubescens]	1,00	Unknown Protein 18	0,44
Unknown Protein 4	0,96	Putative Sec61beta	0,43
Unknown Protein 5	0,95	Unknown Protein 19	0,43
Unknown Protein 6	0,89	Unknown Protein 20	0,41
Putative 40S ribosomal protein S3A	0,88	Unknown Protein 21	0,41
Putative 60S ribosomal protein P2	0,77	Unknown Protein 22	0,40
Unknown Protein 7	0,77	Putative 60S ribosomal protein L17	0,39
Putative 60S ribosomal protein L5	0,74	Unknown Protein 23	0,39
Putative 60S ribosomal protein L7A	0,68	Unknown Protein 24	0,39
Unknown Protein 8	0,63	Putative 40S ribosomal protein S9	0,38
Putative 60S acidic ribosomal protein P0	0,62	Putative abnormal wing disc protein	0,38
Similar to DNA puff protein [T. pubescens]	0,62	Putative 40S ribosomal protein SA	0,38
Putative puff C-8 protein [R. Americana]	0,61	Putative heat shock 70 kDa protein (cog4)	0,34
Unknown Protein 9	0,61	Unknown Protein 25	0,34
Unknown Protein 10	0,60	Unknown Protein 26	0,34
Putative 40S ribosomal protein S7	0,59	Putative cap binding protein 20	0,33
Similar to DNA puff protein [T. pubescens]	0,59	Unknown Protein 27	0,33
Similar to myosin heavy chain	0,59	Unknown Protein 28	0,33
Putative alpha tubulin	0,52	Unknown Protein 29	0,33
Unknown Protein 11	0,52	Putative 40S ribosomal protein S14	0,32
Unknown Protein 12	0,51	Putative 60S ribosomal protein L37A	0,32
Putative 60S ribosomal protein L35A	0,50	Unknown Protein 30	0,32
Putative elong. factor 1-alpha R. americana	0,50	Puff B2 protein [R. Americana]	0,32
Putative 40S ribosomal protein S2	0,49	Putative chitinase	0,31

Foram encontradas 62 proteínas ribossômicas das 80 proteínas que geralmente compõe os ribossomos em eucariotos (Doudna e Rath, 2002), listadas na tabela 2, esta coleção de proteínas ribossômicas torna-se muito interessante para o

modelo *Rhynchosciara*, pois além de participar da síntese protéica, estas têm sido associadas ao desenvolvimento (Warner e Nierras, 1998), ao transporte e endereçamento de peptídeos.

Ainda associados à função secretora da glândula salivar foram identificados transcritos de chaperones, e transportadores de membrana, tais como Sec61 beta, Sec31 e Sec23, e como possíveis candidatos a serem secretados pode-se observar os transcritos dos pufes putativos, transcritos codificando peptídeos antimicrobianos (Diptericinas A e B) e quitinase e seu respectivo precursor, possivelmente associados à proteção do grupo (Diptericinas) e à metamorfose (quitinase), além dos já descritos transcritos dos pufes B2 e C8, codificando proteínas estruturais do casulo (Winter, 1977).

Diferentes proteínas estruturais da célula estão sendo expressas (actina e tubulinas, entre outras) (Tabelas 1 e 6, Figura 11), o que pode implicar que além da expressão constitutiva destas proteínas, que são naturalmente repostas no citoesqueleto, pode estar havendo modificações estruturais nestas células, devido ao volume nuclear que está terminando de aumentar, ou como reflexo da atividade secretora do tecido.

Figura 11 – Alguns dos principais transcritos expressos, excluindo-se os que codificam proteínas ribossômicas e desconhecidas.

Tabela 2 – Lista das 62 proteínas ribossômicas encontradas durante o sequenciamento de EST em *Rhynchosciara* e suas respectivas proporções entre os ESTs.

Proteínas Ribossômicas	% ESTs	Proteínas Ribossômicas	% ESTs
60S acidic ribosomal protein P1	2,32	60S ribosomal protein L18	0,11
40S ribosomal protein S18	1,44	60S ribosomal protein L22	0,11
40S ribosomal protein S3A	0,89	60S ribosomal protein L15	0,11
60S acidic ribosomal protein P2	0,81	60S ribosomal protein L11	0,07
60S ribosomal protein L5	0,81	60S ribosomal protein L35	0,07
60S ribosomal protein L7A	0,68	40S ribosomal protein S21	0,07
60S acidic ribosomal protein P0	0,66	40S ribosomal protein S27a	0,07
40S ribosomal protein S7	0,61	40S ribosomal protein S6	0,07
40S ribosomal protein S2	0,49	60S ribosomal protein L7	0,07
60S ribosomal protein L17	0,45	40S ribosomal protein S9	0,06
Stubarista/40S ribosomal protein SA	0,39	60S ribosomal protein L19	0,06
60S ribosomal protein L37A	0,37	60S ribosomal protein L13A	0,05
60S ribosomal protein L18A	0,34	60S ribosomal protein L14	0,05
40S ribosomal protein S14	0,34	60S ribosomal protein L29	0,05
60S ribosomal protein L27A	0,34	60S ribosomal protein L30	0,05
60S ribosomal protein L35A	0,31	60S ribosomal protein L34	0,05
40S ribosomal protein S4	0,29	60S ribosomal protein L4	0,05
60S ribosomal protein L17A	0,29	60S ribosomal protein L8	0,05
60S ribosomal protein L6	0,26	mitochondrial ribosomal protein S18b	0,05
40S ribosomal protein S26	0,24	60S ribosomal protein L1	0,05
60S ribosomal protein L36	0,24	60S ribosomal protein L21	0,05
40S ribosomal protein S8	0,20	40S ribosomal protein S20	0,05
40S ribosomal protein S16	0,18	40S ribosomal protein S17	0,04
40S ribosomal protein S15	0,17	60S ribosomal protein L27	0,04
60S ribosomal protein L10A	0,16	60S ribosomal protein L28	0,04
40S ribosomal protein S19	0,16	40S ribosomal protein S13	0,02
60S ribosomal protein L13	0,16	40S ribosomal protein S15A	0,02
60S ribosomal protein L3	0,16	40S ribosomal protein S25	0,02
40S ribosomal protein S10	0,12	40S ribosomal protein S27	0,02
60S ribosomal protein L31	0,12	60S Ribosomal protein L12	0,02
40S ribosomal protein S23	0,11	60S ribosomal protein L38	0,02

Possivelmente como consequência das modificações sofridas na cromatina e da necessidade de continuar transcrevendo, foram encontradas mensagens codificando proteínas estruturais da cromatina tais como histona H3.3 e HMG-D. Histona H3.3 é uma variante da histona H3 presente em regiões da cromatina transcricionalmente ativas e que acredita-se ser capaz de contribuir com as características físicas especializadas destas regiões (Ahmad e Henikoff, 2002) e HMG-D (High Mobility Group) é de uma família de proteínas nucleares abundantes capazes de mediar diferentes modificações estruturais na cromatina, interagindo tanto com o DNA, podendo estabilizá-lo, como com os nucleossomos, contribuindo com a regulação da compactação da cromatina (Ner e Travers, 1994; Bustin, 1999; Ner *et al*, 2001).

Inesperadamente também foram encontradas mensagens de transposases e de fatores pró e anti-apoptóticos (Tabela 1, figura 11 e seção Categorização), sendo que este tecido vai regredir somente depois de vários dias.

3.1.2. – Categorização

A fim de gerar uma visão mais ampla do estado metabólico do tecido, os ESTs gerados foram categorizados segundo os critérios do Gene Ontology Consortium (GO) (2001) e com base comparativa no Gene Index de *Drosophila* (TIGR – V 8.0) (conforme descrito em Materiais e Métodos), os termos do Gene Ontology proporcionam um vocabulário para descrever de forma não-unívoca e consistente um produto gênico conforme três aspectos principais: *Função Molecular*, *Processo Biológico* e *Compartimento Celular*; que se ramificam em diferentes níveis hierárquicos de subcategorias inter-relacionadas, ao analisar as seqüências categorizadas é possível ter uma visão mais detalhada dos eventos que estão ocorrendo no tecido como em um “Northern eletrônico”.

O uso do Gene Index de *Drosophila* como banco para categorização permitiu o aproveitamento do extensivo esforço de anotação deste genoma e da alta qualidade de sua caracterização, outros bancos do Gene Index disponíveis poderiam ter sido utilizados, porém não seria obtido o mesmo grau de detalhamento, pois o genoma de *Anopheles* apesar de completo não foi inteiramente anotado, e o uso do Gene Index de organismos inferiores (ex: *S. pombe*, *S. cerevisiae*) ou de organismos com genomas ainda não completamente seqüenciados (ex: *Apis mellifera*) acarretaria na não categorização de grandes porções de ESTs de *Rhynchosciara*.

Outra hipótese seria a soma de todos os Gene Index de interesse em um único banco de dados, porém como a categorização é baseada em comparação proteína-proteína certamente haveria desvios das seqüências favorecendo organismos não tão bem anotados e curados como *Drosophila*, como exemplo, poder-se-ia esperar desvios de aproximadamente 32% das seqüências de *Rhynchosciara* para categorização segundo Gene Index de *Anopheles*, conforme discutido anteriormente. Entretanto a soma de todos os bancos do Gene Index apresentaria como vantagem um ganho na amplitude da categorização, visto que categorizando segundo um único banco perde-se a oportunidade de comparar proteínas exclusivas de diferentes organismos que por ventura possam ser semelhantes com o modelo em estudo. Considerando-se o atual

estágio da pesquisa em *Rhynchosciara*, optou-se para o presente estudo prezar a especificidade e detalhamento em vez da amplitude.

Com o intuito de preservar a integridade do vocabulário do GO e seu sentido original, os seus termos não foram traduzidos para o português.

Os termos do GO relacionados à *Função Molecular* são os mais curados e detalhados, o que resulta em uma proporção maior de ESTs que podem ser classificados dentro desta categoria, no banco de *Rhynchosciara* 59,72% dos ESTs e 52,52% das Sequências Únicas puderam ser categorizados com termos desta classe. As subcategorias mais presentes foram *Moléculas Estruturais*, *Enzimas* e *Ligantes* (Tabela 3).

A subcategoria *Moléculas Estruturais* representa 18,34% dos ESTs e 10,35% das Sequências Únicas, consequência da alta redundância de mensagens codificando proteínas ribossômicas muitas delas categorizadas funcionalmente como *Constituinte Estrutural do Ribossomo*, 15,83% dos ESTs e 8,33% das Sequências Únicas.

Ainda dentro do aspecto *Função Molecular* destacam-se as subcategorias *Enzimas* e *Ligantes*, dada a natureza do tecido em estudo já era esperado que houvesse níveis elevados de mensagens codificantes de enzimas envolvidas em síntese protéica (tais como isomerases, quinases e transferases) e metabolismo energético (ATPases e GTPases), porém a presença de diferentes helicases e de pelo menos duas enzimas quinases dependentes de ciclina (CDC2-like e CDK7) indica uma atividade replicativa.

Também dentro dos transcritos classificados como ligantes encontra-se indícios de uma atividade replicativa além da atividade sintética, a presença de *Ligantes de Nucleotídeos* (3,05%), formados principalmente por ATPases e GTPases, e a abundância de *Ligantes de Ácidos Nucléicos* (15,16%), constituídos por *Ligantes de RNA* (1,28%), *Fatores de Transcrição* (0,75%), *Fatores de Elongação da Tradução* (1,55%) além de fatores de “splicing” e poliadenilação corroboram com a síntese de proteínas, a presença de *Ligantes de DNA* (4,02%), *Ligantes de Fitas Simples de DNA* (0,34%), *Fatores de Replicação de DNA* (0,38%), e uma mensagem categorizada como “*DNA Clamp Loader*” indicam fortemente a atividade replicativa das glândulas salivares no Início de Rede.

Tabela 3 – Distribuição dos ESTs e das Sequências Únicas dentro do aspecto *Função Molecular* do GO.

	% ESTs	% SU
Molecular Function	59,72	52,52
Binding	21,77	26,28
Juvenile hormone binding	0,01	0,04
Odorant binding	0,01	0,04
Oxygen binding	0,04	0,12
Isoprenoid binding	0,06	0,19
Vitamin/Cofactor Transporter	0,06	0,12
Drug Binding	0,07	0,19
Neurotransmitter Binding	0,07	0,16
Receptor binding	0,07	0,16
Carbohydrate Binding	0,09	0,16
Heavy metal binding	0,17	0,16
Lipid Binding	0,20	0,39
Calcium ion binding	0,55	1,01
Protein Binding	1,99	3,76
Nucleotide Binding	3,05	3,53
Nucleic Acid Binding	15,16	16,28
Structural Molecule	18,39	10,35
Structural constituent of chorion (sensu Insecta)	0,01	0,04
Structural constituent of muscle	0,10	0,27
Structural constituent of cuticle	0,11	0,27
Puparial glue (sensu Diptera)	0,23	0,23
Structural constituent of peritrophic membrane	0,23	0,31
Structural constituent of cytoskeleton	0,89	0,74
Extracellular matrix structural constituent	0,99	0,12
Structural constituent of ribosome	15,83	8,33
Enzyme	14,90	26,28
Diphosphoinositol-polyphosphate diphosphatase	0,04	0,08
Phosphatase	0,31	0,81
Small protein conjugating enzyme	0,34	0,66
Helicase	0,35	0,74
Isomerase	0,48	0,89
Ligase	0,50	1,05
Lyase	1,29	1,24
Kinase	1,42	3,02
Oxidoreductase	1,94	3,60
Transferase	2,66	1,09
Hydrolase	5,49	8,72
Transporter	4,28	7,17
Transcription Regulator	2,33	5,50
Chaperone	2,09	1,82
Translation Regulator	1,92	1,36
Signal Transducer	1,66	3,22
Enzyme Regulator	0,43	0,81
Apoptosis Regulator	0,23	0,19
Apoptosis Inhibitor	0,23	0,19
Motor	0,22	0,66
Defense/Immunity Protein	0,21	0,27
Antimicrobial Peptide	0,19	0,62
Protein Tagging	0,11	0,12
Cell Adhesion Molecule	0,09	0,23
Antioxidant	0,01	0,04
Non-categorized	40,28	47,48

No estágio atual de desenvolvimento do GO é natural que uma quantidade menor de transcritos seja classificada nos aspectos *Processo Biológico* e *Compartimento Celular*, reflexo da complexidade da tarefa de associar evidências reais a cada termo do GO e do fato de não ser conhecido o perfil completo de participação em processos biológicos ou a localização em que cada função de uma dada proteína é desempenhada.

Ao continuar a categorização no aspecto *Processo Biológico* também há evidências do funcionamento de um processo replicativo conjuntamente com a produção de proteínas, 31,2% dos ESTs se enquadraram em alguma subcategoria, sendo que as que mais se destacaram foram *Metabolismo de Proteínas* (19,24%) e *Biossíntese* (22,48%) (Tabela 4), 3,11% dos ESTs (correspondendo a 6,28% das SU) foram classificadas como participantes do *Ciclo Celular*, os principais processos que compõem esta categoria são: *Formação e manutenção do complexo pré-replicativo* (0,15% ESTs), *Coesão de cromátides irmãs* (1,27% ESTs), *Replicação do DNA* (0,22% ESTs) e *Chechagem da replicação do DNA* (0,22% ESTs), dentre estes transcritos estão presentes os que foram anotados como RaDUP e RaMCM5, representando componentes do complexo pré-replicativo, e CDC2-like, CDK7, fizzy-related e possíveis componentes do APC/C que podem estar envolvidos com os endociclos e sua regulação neste tecido (Tabela 6).

Conforme relatado anteriormente, estão presentes neste banco fatores pró e anti-apoptóticos, e observando a Tabela 4 nota-se que os transcritos a favor e os contra a morte celular estão praticamente em equilíbrio, como o número de células da glândula não varia até a regressão do tecido que ocorrerá durante a metamorfose, a quebra deste equilíbrio pode representar o gatilho que desencadeará o desaparecimento da glândula.

Tabela 4 – Distribuição dos ESTs e das Sequências Únicas dentro do aspecto *Processo Biológico* do GO.

	% EST	% SU
Biological Process	31,20	28,80
Cell growth and/or maintenance	30,66	45,93
Membrane fusion	0,04	0,12
Cellular morphogenesis	0,07	0,24
Cell proliferation	0,09	0,27
Homeostasis	0,12	0,03
Cell motility	0,18	0,46
Response to stress	1,11	0,97
Transport	2,36	3,41
Cell cycle	3,11	6,28
DNA replication and chromosome cycle	2,10	1,90
DNA replication and chromosome cycle	0,23	0,50
DNA replication	0,22	0,46
Mitotic chromosome condensation	0,08	0,19
Pre-replicative complex formation and maintenance	0,15	0,31
Sister chromatid cohesion	1,27	0,15
G1/S transition of mitotic cell cycle	0,03	0,08
M phase	3,37	1,78
Mitotic cell cycle	1,13	2,44
DNA replication checkpoint	0,22	0,46
Regulation of CDK activity	0,01	0,04
Cell organization and biogenesis	4,81	4,07
Metabolism	25,73	33,80
Amino acid and derivative metabolism	0,23	0,31
Aromatic compound metabolism	0,24	0,35
Energy pathways	0,28	0,39
Carbohydrate metabolism	0,31	0,46
Phosphorus metabolism	0,78	1,90
Catabolism	1,11	1,63
Nucleobase, nucleoside, nucleotide and nucleic acid metabolism	9,42	8,06
Protein metabolism	19,24	13,68
Biosynthesis	22,48	12,52
Development	5,69	12,94
Larval development	0,23	0,58
Morphogenesis	2,37	6,16
Cell communication	4,46	7,13
Behavior	0,65	1,59
Physiological processes	0,38	0,58
Death	0,13	0,27
Anti-apoptosis	0,04	0,08
Apoptosis	0,01	0,04
Caspase activation	0,01	0,04
Caspase activation via cytochrome c	0,01	0,04
Response to stress	0,02	0,08
Non-categorized	68,80	71,20

Componente Celular é a ontologia que fornece menor quantidade de informação, 33,60% dos ESTs (49,19% das SU) puderam ser adequadas em algum dos seus termos, 26,25% dos ESTs (13,26% dos SU) estão associados de alguma forma aos ribossomos e 1,84% dos ESTs (1,20% das SU) estão atribuídos como agindo diretamente nos cromossomos.

Tabela 5 – Distribuição dos ESTs e das Sequências Únicas dentro do aspecto *Compartimento Celular* do GO.

	% EST	% SU
Cellular Component	33,6	49,19
Apical part of cell	0,02	0,08
Cillium	0,05	0,16
Ribonucleoprotein complex	0,07	0,16
Cell cortex	0,16	0,43
Membrane	0,79	1,47
Plasma membrane	1,16	2,67
Endoplasmic reticulum	1,21	0,66
Chromosome	1,84	1,20
Extracellular	1,92	1,43
Mitochondrion	2,62	3,91
Nucleus	9,45	9,92
Cytoplasm	11,28	13,84
Ribosome	26,25	13,26
Non-categorized	66,40	50,81

Construído o banco de dados relacionando os termos do GO aos ESTs de *Rhynchosciara* foi possível identificar os possíveis transcritos envolvidos com o ciclo celular e com a síntese de DNA, a anotação de tais candidatos está listada na tabela 6, bem como sua possível função atribuída por comparação com os homólogos de outros organismos quando possível.

A expectativa de encontrar CDKs e ciclinas já descritas como participantes de endociclos nas células foliculares (CDK2/Ciclina E) (Calvi et al, 1998; Follette et al, 1998) ou de controle geral do ciclo celular não foi concretizada, o que não significa que estas proteínas não estejam participando da amplificação e politenização em *Rhynchosciara*. Por outro lado, algumas destas mensagens selecionadas possuem forte similaridade com algumas quinases, tais como Cdc2, WARTS e CDK7, além de outros transcritos que possuem relativa semelhança com domínios de quinases de serina/treonina mas não foram anotadas, que podem ser quinases (ou CDKs) específicas deste gênero envolvidas de alguma forma na regulação dos endociclos.

Tabela 6 – ESTs selecionados pelo GO por provável envolvimento com o ciclo celular e com a síntese de DNA.

Anotação	Possível função
Putative Dup/Cdt1	Fator de licenciamento
Putative Minichromosome maintenance 5	Fator de licenciamento/Helicase
Putative replication protein A	Estabiliza fitas simples de DNA no momento de sua síntese
Putative Cyclin-dependent kinase 7 (CDK7)	Fosforila outras CDKs (ativação) e RNAPol-II
Putative cdc2-like	Provável CDK envolvida na regulação do ciclo celular ou dos endociclos neste tecido
Putative translation initiation factor eIF-4A	Fator de iniciação da tradução e RNA helicase, sua regulação é essencial para a proliferação celular e crescimento embrionário
Putative eukaryotic initiation factor eIF-4E	Fator de iniciação da tradução, sua regulação é essencial para a proliferação celular
Putative protein phosphatase 2A 55K regulatory chain	Modula o complexo PP2A para regular a atividade de CDK1, participa de um ponto de checagem durante a segregação cromossômica
Putative protein phosphatase 2A	Desfosforila resíduos de serina e treonina em diferentes processos biológicos.
Similar to Serine/threonine protein phosphatase 4	Membro da mesma família de fosfatases de proteínas de PP2A
Putative Serine/threonine protein phosphatase 4	Membro da mesma família de fosfatases de proteínas de PP2A
Similar to Spc110p	Modificações do citoesqueleto associadas ao ciclo celular
Putative Component of the spindle assembly checkpoint (MDF-2)	Participa de um mecanismo que impede erros na segregação cromossômica, inibe FIZZY
Putative lamin precursor	Componente da membrana nuclear que pode ligar-se à cromatina
Putative lamin	Componente da membrana nuclear que pode ligar-se à cromatina
Putative fizzy-related protein	Regula especificidade e “timing” do APC/C
Putative Heat shock protein 83	Assiste a formação dos complexos Ciclinas-CDKs
Putative myosin II regulatory light chain / spaghetti squash	Envolvida na citocinese durante a ovogênese e embriogênese de <i>Drosophila</i> (Jordan e Karess, 1997)
Putative LATS/WARTS	Quinase regulada por Cdc2/ciclina B, regula a montagem dos filamentos de Actina
Putative Ubiquitin Conjugating enzyme E2	Peptídeo que forma um intermediário com ubiquitina, através de uma ligação tiol-éster, antes de esta ser transferida e ligada ao substrato final por uma ubiquitina ligase E3
Putative Ubiquitin Conjugating enzyme E2 / Effete	Peptídeo que forma um intermediário com ubiquitina, através de uma ligação tiol-éster, antes de esta ser transferida e ligada ao substrato final por uma ubiquitina ligase E3
Putative Phosphatidylinositol 4-kinase	Fosforilação de fosfoinosítídeos, participando de respostas celulares à estímulos hormonais
Putative half pint	Regula “splicing” alternativos de mRNAs de eIF4E, outo entre outros
Putative splicing factor	Regula “splicing” alternativos
Outras 13 proteínas putativas anotadas como desconhecidas	Relacionadas ao ciclo celular devido a pequenas regiões com similaridade com os TCs, porém não suficientes para anotação

Cdc2 é uma CDK que regula todo o ciclo celular associando-se a diferentes ciclinas em *S. cerevisiae*, é interessante a presença em nosso banco de uma proteína semelhante com tal CDK, pois dentre todos os ESTs, excetuando-se CDK7, não foram encontradas CDKs descritas como “principais” na regulação do ciclo em outros sistemas (CDK1-2-4-6) tampouco ciclinas, tal transcrito foi identificado recentemente e merece uma caracterização para averiguar se atua na regulação da politenização e/ou amplificação em *Rhynchosciara*.

WARTS corresponde a uma proteína que possui um domínio característico de quinases de serina/treonina, identificada em *Drosophila* pela sua capacidade de suprimir tumores e está relacionada com diferentes eventos mitóticos no citoesqueleto e no centróssomo (Morisaki *et al*, 2002). CDK7 por sua vez, atua na regulação da transcrição e do ciclo celular, fosforilando a porção C-terminal da RNA polimerase II quando atua como uma subunidade do TFIIF, atua também como Quinase Ativadora de CDKs (CAK) ativando outras CDKs através de fosforilação, promovendo o avanço do ciclo (Wallenfang e Seydoux, 2002), esta função é desempenhada em vertebrados quando é formado um complexo ternário composto por CDK7, MAT1 e ciclina H (Hunter e Poon, 1997).

Ainda em relação à regulação de CDKs pode-se notar que também foi identificado pela categorização segundo o GO um homólogo de HSP83 (Heat Shock Protein 83) de *Drosophila*, designado de HSP90 em diferentes organismos. Inicialmente relacionar esta classe de proteínas ao ciclo celular pode causar estranheza, porém diversos autores relatam dados que sugerem que as funções de diversas chaperonas vão além da renaturação e maturação de proteínas, operando em conjunto com receptores em vias de sinalização e assistindo no processo de montagem dos complexos ciclinas-CDKs (Hunter e Poon, 1997), o que pode inferir a participação em um segundo processo neste tecido às chaperonas citadas na Figura 11.

Além de quinases, fosfatases também foram relacionadas ao ciclo celular durante a categorização, duas destas compõem o complexo hetero-trimérico PP2A (Protein Phosphatase 2A), o qual desfosforila resíduos de serina e treonina, este é composto por uma subunidade catalítica, uma estrutural e uma regulatória variável, que lhe confere a capacidade de desfosforilar diferentes substratos, o que envolve o complexo PP2A em diversos processos celulares. Experimentos em *Drosophila* indicam sua importância na segregação cromossômica e descondensação dos cromossomos mitóticos, através da desfosforilação das histonas H3 (Nowak *et al*,

2003), que também ocorre na regulação da transcrição. PP2A atua também em cascatas de sinalização anti-apoptóticas em *Drosophila*, e pró-apoptóticas em mamíferos (Hoof e Goris, 2003).

A subunidade regulatória de 55KDa do complexo PP2A está implicada na regulação da atividade de CDK1 e desfosforilação de seus substratos, além de atuar na separação das cromátides irmãs mediada via APC/C, postula-se que esta cadeia seja substrato do APC/C e que seja degradada (Deak *et al*, 2003).

Três ESTs dentre estes relacionados ao ciclo celular estariam diretamente envolvidos com a origem de replicação, dois são fatores de licenciamento (Dup/Cdt1 e MCM5) e um EST codifica uma proteína ligada diretamente à maquinaria de replicação (RPA – Replication Protein A, subunidade de 32kDa) e está envolvida em sua capacidade de processamento, estabilizando fitas simples de DNA. Recentemente também foi relacionada à ação de telomerasas, agindo como regulador positivo (Schramke *et al*, 2004).

Os fatores de iniciação da tradução eIF4E e eIF4A são subunidades do complexo eIF4F, cuja função é recrutar as subunidades ribossômicas e fatores associados à porção 5' de mRNAs. eIF4E possui a capacidade de ligar-se ao “cap” dos mensageiros, é conhecidamente um oncogene, e sua superexpressão é capaz de induzir a síntese de DNA em células NIH 3T3 quiescentes (Lachance *et al*, 2002). A subunidade eIF4A age como RNA helicase dependente de ATP (Gao *et al*, 2000), e ambos são relacionados a processos replicativos, embriogênese e proliferação celular, além de serem regulados por CDKs.

A categorização também relacionou dois fatores de “splicing” putativos como envolvidos na regulação do ciclo celular, um destes fatores seria homólogo à proteína “Half-Pint” de *Drosophila* (Buskirk e Schüpbach, 2002; Rio, 2002), implicada na regulação do “splicing” alternativo e no endereçamento de eIF4E, do mensageiro de “ovarian tumour” (otu), e de fatores relacionados ao fusoma e microfilamentos mitóticos, durante a ovogênese.

Outros elementos que durante a categorização foram relacionados ao ciclo celular são: enzimas conjugadoras de ubiquitina E2, uma subunidade regulatória de miosina II, uma quinase de fosfatidil inositol, MDF2 e laminas putativas. Laminas são proteínas que compõem a lâmina nuclear disposta internamente à membrana e que se ligam à cromatina, ao longo do ciclo celular esta ligação é desfeita conjuntamente com a membrana nuclear.

Durante a metáfase, Mad1 e Mad2 (MDF2 em *Caenorhabditis*) participam de um mecanismo de ponto de checagem que impede que os cromossomos segreguem até que estejam devidamente ligados aos microtúbulos do fuso mitótico. Um dos alvos deste ponto de checagem seria regular o APC/C, visto que participa da degradação dos elementos envolvidos na união das cromátides conforme já descrito, a tarefa de inibi-lo é de Mad2 (MDF2) através da interação com FIZZY (Kitagawa e Rose, 1999).

Algumas respostas celulares à estímulos hormonais, inclusive no controle do crescimento, são produzidas por intermédio da fosforilação do lipídeo de membrana fosfatidil inositol (Gao *et al*, 2000), que ativa quinases de serina/treonina que são capazes de regular efetores do ciclo celular, da biossíntese e de morte celular programada. Neste contexto a presença da quinase de 4-fosfatidil-inositol nas glândulas salivares durante o IR pode representar um vínculo entre estímulos hormonais e os processos em funcionamento neste tecido (síntese protéica e/ou endocitos).

As enzimas conjugadoras de ubiquitina E2 atuam concomitantemente com as ligases de ubiquitina E3, sendo necessárias em um passo intermediário durante processo de ubiquitinação, formando um conjugado com ubiquitina ativada (pela proteína ativadora de ubiquitina E1) através de uma ligação tiol-éster (Okamoto *et al*, 2003), que será utilizado pela enzima ligase de ubiquitina E3 para transferir a ubiquitina ao substrato (Hershko e Ciechanover, 1998). Logo sendo o APC/C uma ubiquitina ligase E3, pode-se supor que a enzima conjugadora de ubiquitina E2 esteja trabalhando em conjunto com este complexo.

Quanto aos fatores de licenciamento, Dup foi o primeiro EST de *Rhynchosciara* a ser relacionado com replicação e foi o primeiro selecionado para caracterização como será descrito nas próximas seções. Conforme já exposto na introdução Dup é um fator de licenciamento das origens, bem como MCM5, identificado posteriormente.

É interessante notar que estes transcritos relacionados ao ciclo celular pelo GO podem ser divididos em grupos: um grupo ligado à síntese de DNA (RPA e fatores de licenciamento); um grupo relacionado à regulação por adição ou remoção de grupos fosfato (quinases e fosfatases); transcritos envolvidos com o ciclo celular através da regulação da tradução, da transcrição, da proteólise; além de um conjunto de proteínas envolvidas com aspectos estruturais da célula (laminas, Scp110p e até mesmo WARTS).

Diversos destes transcritos aparentemente possuem funções mitóticas (PP2A, WARTS, Scp110p, laminas e cadeia leve regulatória de miosina II), visto que são conhecidas suas participações durante as modificações do citoesqueleto e na segregação cromossômica em outros modelos. Como em glândulas salivares de *Rhynchosciara* não há degradação da membrana nuclear nem movimentação citocinética, não se pode implicar estes transcritos a um tipo de endomitose, porém sua presença pode ser um reflexo das modificações citológicas que estão ocorrendo neste tecido, como o aumento do volume nuclear conseqüente da politenização, ou da dinâmica do citoesqueleto associada à exocitose, ou ainda alguns destes transcritos possam desempenhar outras funções no endociclo (PP2A e WARTS).

O seqüenciamento de ESTs não oferece nenhum dado conclusivo sobre a participação de proteínas em processos biológicos, limita-se a mostrar quais elementos estão sendo expressos em determinado momento, porém a partir desta informação pode-se escolher candidatos para uma caracterização mais aprofundada. Dentro do interesse de estudar os endociclos em nosso modelo, os primeiros genes escolhidos foram RaDup e RaMCM5, interessantes por serem absolutamente necessários tanto para ciclos celulares normais como para ciclos modificados e por servirem de ponto de partida para acessar informações sobre as origens de replicação de *Rhynchosciara*.

Outros candidatos com forte potencial informativo nos endociclos deste modelo são: CDK7, por sua atividade ativadora de CDKs, e os transcritos relacionados com o APC/C, tais como FIZZY, FIZZY-RELATED e MDF2, um possível regulador negativo de FIZZY, que nos induz a imaginar um mecanismo de regulação do APC/C dentro do endociclo. As outras treze mensagens anotadas como proteínas desconhecidas podem realmente representar algum participante do ciclo celular, porém, como foram selecionadas por pequenas regiões de similaridade suficientes para passar pelos critérios de categorização pode-se esperar falsos positivos.

3.1.3. – Polimorfismos

Pela primeira vez em *Rhynchosciara* foi possível identificar polimorfismos em suas seqüências, como detalhado na Tabela 7 foi possível encontrar 91 SNPs (Single Nucleotide Polymorphisms) em 89 Seqüências Únicas, discrepâncias foram consideradas SNPs quando ocorriam entre seqüências de alta qualidade (≥ 45 inferida pelo Phred) e preferencialmente quando presentes em mais de uma seqüência. Com

base nos resultados da comparação dos ESTs de *Rhynchosciara* com as proteínas depositadas no GenBank (blastx) foi possível designar o quadro de leitura de 53 destas Sequências Únicas e assim determinar a posição ocupada nos códons para cada uma destas SNPs além de poder definir se a mutação é silenciosa, conservativa ou modificadora.

Destas 53 seqüências, 30 apresentaram SNPs na terceira base do códon e 44 foram designadas como substituições silenciosas ou conservativas (Tabela 7), quanto ao tipo de substituição no DNA tem-se que de um total de 91 SNPs 31 (34%) correspondem a transversões e 60 (66%) a transições.

Interessante notar que um polimorfismo foi encontrado na seqüência codificante da proteína do pufe B2, sendo uma transição G-A na posição 738 do mensageiro, correspondendo a uma substituição modificadora na primeira posição do códon, os aminoácidos codificados nas duas variantes são isoleucina e valina. No entanto este polimorfismo não altera o mapa de restrição deste mensageiro, pois não cria nem destrói sítios de restrição.

A presença destes SNPs no banco de *Rhynchosciara* pode ser explicada pela necessidade de coletar os espécimes na natureza, como não é possível prever se os espécimes selvagens coletados pertencem à mesma linhagem, e no momento de construir a biblioteca de cDNA foram necessárias glândulas salivares de diferentes grupos, pode-se então esperar observar variações nas seqüências obtidas, outra hipótese é que pelo menos alguns destes genes representados nos ESTs possuam múltiplas cópias no genoma deste modelo.

A presença de alguns destes polimorfismos nas seqüências de proteínas ribossômicas, que de modo geral são conservadas a ponto de serem utilizadas como marcadores filogenéticos por alguns autores (Landais *et al*, 2003; Harris *et al*, 2003), reforça a hipótese de diferentes linhagens terem originado estes polimorfismos.

É interessante constatar que há variabilidade genética na população de *Rhynchosciara* de Ubatuba-SP, local onde foram coletadas as larvas utilizadas para este trabalho. Anteriormente já haviam sido observadas diferenças em Southern de DNA genômico marcados com sondas específicas para pufes, entre grupos de larvas provenientes de Ubatuba-SP e de Mongaguá-SP (Santelli, comunicação pessoal), tais diferenças eram atribuídas a polimorfismos, porém até o momento estes não haviam sido detectados por seqüenciamento.

Tais dados podem indicar que os poucos nichos aonde ainda se encontram *Rhynchosciara* com relativa facilidade, foram povoados por diferentes populações deste gênero que migraram a estes locais, o que explicaria também o fato de ser comum nestes pontos a coexistência de diferentes espécies de *Rhynchosciara*. Obviamente mais estudos comparativos deste gênero são necessários para estabelecer relações entre diferentes populações e diferentes espécies.

Tabela 7 – Dados obtidos sobre os 91 polimorfismos identificados

Polimorfismos (SNPs)	91	
	n°	%
Posicionamento no Códon: 1º Base	14	15,38
Posicionamento no Códon: 2º Base	9	9,89
Posicionamento no Códon: 3º Base	30	32,97
Quadro de leitura não-determinado	38	41,76
Substituição Conservativa	17	18,68
Substituição Silenciosa	27	29,67
Substituição Modificadora	9	9,89
Transições	60	66,00
Transversões	31	34,00

3.1.4. – Códon Usage

Outro fruto deste seqüenciamento foi a criação de uma tabela de “códon usage” para *Rhynchosciara americana* (Tabela 8), dado que até então era baseado em três ORFs conhecidas deste modelo.

Tal tabela fora gerada a partir de 69 seqüências tidas como completas e com a região codificante composta inteiramente por bases de alta qualidade, totalizando 11868 códons, a seleção das seqüências foi manual e o critério consistiu que estas somente seriam consideradas “full lenght” se quando comparadas com outros organismos (principalmente *Drosophila*) codificassem desde a metionina até o códon de parada, sendo que “gaps” e inserções eram permitidas e erros e dúvidas de seqüenciamento não foram permitidos.

A amostragem utilizada para o cálculo destes dados é pequena em relação aos totais de ESTs gerados, menores ainda se considerarmos o genoma envolvido, porém prezou-se a qualidade da informação de entrada para o algoritmo de cálculo³, para evitar tendências e falsos padrões de utilização de códons.

³ <http://www.molbiol.ox.ac.uk/cu/codonW.html>

Tabela 8: de códon usage de *Rhynchosciara americana*: gerada a partir de 11868 códon de 69 ORFs diferentes, os valores à esquerda representam o número de vezes que os códon aparecem dentro da amostragem, e à direita a fração de utilização).

Phe	UUU	172	0.39	Ser	UCU	102	0.14	Tyr	UAU	150	0.47	Cys	UGU	86	0.50
	UUC	261	0.61		UCC	115	0.16		UAC	174	0.53		UGC	86	0.50
Leu	UUA	177	0.17		UCA	173	0.24	TER	UAA	42	0.68	TER	UGA	12	0.19
	UUG	393	0.42		UCG	147	0.20		UAG	8	0.13	Trp	UGG	117	1.00
	CUU	105	0.12	Pro	CCU	65	0.12	His	CAU	148	0.59	Arg	CGU	319	0.42
	CUC	63	0.06		CCC	63	0.11		CAC	103	0.41		CGC	123	0.16
	CUA	100	0.12		CCA	301	0.56	Gln	CAA	311	0.72		CGA	126	0.17
	CUG	106	0.11		CCG	113	0.21		CAG	120	0.28		CGG	52	0.07
Ile	AUU	383	0.52	Thr	ACU	167	0.25	Asn	AAU	260	0.55	Ser	AGU	105	0.15
	AUC	262	0.36		ACC	174	0.26		AAC	211	0.45		AGC	81	0.11
	AUA	91	0.12		ACA	226	0.34	Lys	AAA	629	0.57	Arg	AGA	98	0.13
Met	AUG	281	1.00		ACG	90	0.15		AAG	480	0.43		AGG	35	0.05
Val	GUU	320	0.40	Ala	GCU	401	0.42	Asp	GAU	358	0.65	Gly	GGU	386	0.46
	GUC	199	0.25		GCC	234	0.24		GAC	193	0.35		GGC	157	0.19
	GUA	139	0.17		GCA	243	0.25	Glu	GAA	552	0.77		GGA	254	0.30
	GUG	143	0.18		GCG	81	0.09		GAG	159	0.23		GGG	43	0.05

Tabela 9: códon usage de *Drosophila melanogaster*: baseada em 17.606.080 códon e 33600 CDS, os números em azul representam uma utilização maior do códon correspondente em relação a *Rhynchosciara*, os números em vermelho o inverso.

Phe	UUU	0.38	Ser	UCU	0.08	Tyr	UAU	0.37	Cys	UGU	0.29
	UUC	0.62		UCC	0.24		UAC	0.63		UGC	0.71
Leu	UUA	0.05		UCA	0.09	TER	UAA	0.41	TER	UGA	0.26
	UUG	0.18		UCG	0.20		UAG	0.33	Trp	UGG	1.00
	CUU	0.10	Pro	CCU	0.13	His	CAU	0.40	Arg	CGU	0.16
	CUC	0.15		CCC	0.33		CAC	0.60		CGC	0.33
	CUA	0.09		CCA	0.25	Gln	CAA	0.30		CGA	0.15
	CUG	0.43		CCG	0.29		CAG	0.70		CGG	0.15
Ile	AUU	0.34	Thr	ACU	0.17	Asn	AAU	0.44	Ser	AGU	0.14
	AUC	0.47		ACC	0.38		AAC	0.56		AGC	0.25
	AUA	0.19		ACA	0.19	Lys	AAA	0.30	Arg	AGA	0.09
Met	AUG	1.00		ACG	0.26		AAG	0.70		AGG	0.11
Val	GUU	0.18	Ala	GCU	0.19	Asp	GAU	0.53	Gly	GGU	0.21
	GUC	0.23		GCC	0.45		GAC	0.47		GGC	0.43
	GUA	0.11		GCA	0.17	Glu	GAA	0.33		GGA	0.28
	GUG	0.47		GCG	0.19		GAG	0.67		GGG	0.07

Comparando-se as tabelas de códon usage de *Rhynchosciara* e de *Drosophila*⁴ pode-se notar uma preferência de uso do códon de terminação UAA em *Rhynchosciara* e a inversão dos códon preferenciais para histidina, glutamina, lisina e ácido glutâmico. Quanto aos aminoácidos codificados por 4 ou 6 códon também

⁴ <http://www.kazusa.or.jp/codon/>

observou-se diferenças nas suas frações de utilização, conforme mostra a Tabela 9, sendo que as maiores diferenças estão nos códons que codificam leucina, glicina e alanina. Lembrando que o códon usage gerado para *Rhynchosciara* refere-se a um tecido específico em um determinado estágio do desenvolvimento, uma vez que a amostragem de ORFs aumente e possivelmente passe a representar também outros tecidos, esta tabela pode ser modificada.

Este tipo de informação é uma ferramenta muito útil em organismos com poucas informações moleculares disponíveis sobre sua biologia, auxiliando no desenho de oligos degenerados.

3.2. - RaDup

3.2.1. – Estrutura Gênica e Localização

Dentre as seqüências de ESTs de *Rhynchosciara* a primeira a ser relacionada diretamente com replicação foi RaDup (*Rhynchosciara americana* Double Parked), devido sua homologia com Cdt1/Dup, no entanto “Double Parked” refere-se à um fenótipo específico de uma linhagem mutante de *Drosophila* (Whittaker *et al*, 2000) e como até o presente não existem linhagens mutantes em *Rhynchosciara* o gene representado por este EST deveria chamar-se RaCdt1, porém, como DmDup apresenta maior similaridade com esta mensagem (e como tornou-se um hábito designá-la de RaDup), optou-se por chamar este gene e seu produto de RaDup.

O clone correspondente a esta mensagem foi isolado e seqüenciado por “primer walking”, constatou-se então que este EST codificava um peptídeo que correspondia a aproximadamente 2/3 da proteína DmDUP.

O mesmo cDNA serviu de sonda para uma varredura em uma biblioteca genômica de *Rhynchosciara* construída em bacteriófago λ DASH II (Stratagene), onde um clone contendo 12.674 bp foi isolado e completamente seqüenciado por “shotgun” ou por “primer walking” em regiões de alto conteúdo de GC (Figuras 12 e 13). Este mostrou conter completamente o gene RaDup e a porção final de um gene com alta similaridade com mus210 de *Drosophila melanogaster*, gene homólogo de XPC (Xeroderma Pigmentosum C), envolvido em reparo de DNA por excisão de nucleotídeo. A cobertura média desta seqüência foi de 13,4 vezes e a taxa de erro inferida pelo Consed foi de 2 erros/10kb (Gordon *et al*, 1998).

É interessante notar que tanto no modelo em questão quanto em *Drosophila* os genes RaDup e o putativo XPC estão próximos, em *Rhynchosciara* estes genes possuem a mesma orientação e são separados por aproximadamente mil nucleotídeos, enquanto em *Drosophila* estes genes possuem orientações opostas e são separados por 58 kb (cromossomo 2R posição no mapa citogenético 51F4 para XPC e 51F11 para Dup), havendo outros ORFs neste intervalo, já em *Anopheles* estes genes estão posicionados em braços opostos do cromossomo 2.

Figura 12 – Esquematização do inserto genômico contendo o gene de RaDup

As estruturas dos genes RaDup e do trecho correspondente ao XPC foram preditas através do programa Genscan (Burge e Karlin, 1997), e em relação ao RaDup foi parcialmente confirmada através da clonagem e sequenciamento de outras porções de seu mensageiro. Com base nas seqüências já obtidas no EST e no inserto genômico foram desenhados primers para uso em reações de RT-PCR e RACE visando obter a seqüência completa deste mRNA, infelizmente a região 5' UTR ainda não foi clonada, porém toda a porção codificante do mensageiro foi definida e utilizada para confirmar a estrutura gênica predita.

Figura 13 – Seqüência do inserto genômico contendo o gene RaDup, as porções realçadas em amarelo foram confirmadas pelo sequenciamento do cDNA.

1	CGTGTAAGTGTGAATCGTTTCTCCTACGAAAGTTTGTGTACATCAAAAACGACTTACTTTTCAGTTAGCTAATAAAATGC
81	TCGACCCCTCAACCATTTGGAAATTTTCGGACGATGGCAAACCTGAACCATACAAACCACCTACCGCTGAAAATGGAATCGTA
161	CCGCGAAATGCATACGGAATGTTGAACATTTAAGCCGGAATGTTGCCAAAGAAAACCGTTCATTTGAGACGTAAGTT
241	GAAGCAGGAATCAGTCATGCATGACTGACCTGTAGAACCTGTGTGCGAACTAAACAATTTCCGTGTTTCAGTAAACGGTC
	Íntron Putativo
321	TGAATCGTGTTTGTGCCAAATTAAAAATTGATTGTGCGCAGGCGGTGGTCGGATTTCGACTTTTCATGGAGGCAGTTCACAT
401	CCAACTTTTCGATGGATTTGTTGTATGCGAGGAATTTGTCTGAACCTCGTTACGGAAAAATGGTGTGAGGAACAAGACGAAGC
481	TTATCGCAAAGCGAAAGAGAAGCAGAAACAAAGGGTGTATGATAATTGGAAGAAGCTGGTCAAGGGATTGTTGATCAGGC
561	AACGTCACAACGGAATACAATTTGCTGGCAAAGCATGCCGAAAGTAATTTGTTGAAGTGTAATGATTGTGATTGA
	Putativo mus210 (XPC)
641	GTAACGGGTATTTTATGAGTCTGTAAGACTATGTGTTTGCATTTGTTCTGTGATAAGCCATTTTACTCTTTCTTTTCAAT
721	AAATTCATTTTGAACCGATTTGTTTATTTTCTGCTAAAAGTAAGAAAAGTCTTGAAAAGGTTGTTCAAAAGAACG
	Sinal Putativo de Poliadenilação
801	AAAATCTTCTTGATAAAAAGACTGTTATGACGGGGCGAATTAGAAAAATGACAAATGTGATACACTTCAAAGCAAACA

[illegible]

4881 TAGTTCACACACAATAAAGACAGCTGATGTTACCTGTCTCGTCTGTGTAATCATATAAAATTCAAAAACAAGTCAATATTC
4961 AGTACAATTGTTTTTATCAAGGTAGAATGCAAATCAAAGACATTCTGCGTCGACCAAAGTGAAGAAAGTTGTTGTTTTA
5041 TTGTGTCGTTTCAAACAACCTTCTACGATCGATTGAAACAATAGCAATTGATAATAAATTATCGTACATCTCCAAAACAG
5121 GTTTTATAGTAAAAATGGGCTCGATGCCTATGTTAGTGAGCTTGACAAGGTGAATGGAAGAAAATTCATTGAATTTATAT
5201 ACTTTTGTGTTGAGTTCAATTAACCTTGCACAAAATGCAGTAGACGATTTTTTTTTTTCATGTGCTCGTGAAATTTGAATAA
5281 CCGACGTAAAAGTACTCACCTCTGCAGAATTTATTTTTTTTACAAGAAGTATTTTCAAGTCTTAAAATAATTGATGTGCG
5361 TCCAGATGTGGGTAGTTTCATTGTCTTTAAAAATTTATTGTATCCGTAAGAGATCTCTGTAAACAACCTCAAACGCAACTC
5441 AGTAAAAACAATTAGAACGTAGCAGATCGTGCCTACGTTAGGCCTGAACACGCAAGTACACTTTGCGCGGACGTAAACCT
5521 AGTACAATTTGTATTCTTTGTACGAAACACACTACTTTCCGTCGTAGTATTTACTTTATAGCGCAATTACACCTGTGTATG
5601 AAGTGTATAAAAGATGCTCTCATAAAGTGTCTACGTCATGATCCTGTGAACCAATTTTCTTTACAACATTACAGCCTAC
5681 AGCTTTGTTTCAAGGCATAAAAATCGTCTTTTATAGTATGTTGTGGCTCGTTAGCATAGCAGATGTAACCAGTTATTCTG
5761 TTAAATAATATCAAGGGTCCCAACTCATTTTTTACGGGAATTTGTTGGAGGAAGCATTCTGAATCGAAAAATTAATATTT
5841 TAGCAGCTCGGCTCATGTTATAGATTACTTATACTGGCAATGTTTCAAAGTTGTACACTTTACCTGAATTTTTTAGTAC
5921 AAAATTTTCTGATTTTCGTCTGAAATTTTAGCAAAAAACAAAACTTTTCAAAAAATATGGAAGTGCCCGAAAACAGTA
6001 TGACCCCAAGTATGTTGGTACTCCGTACACAATCTTTAATTCAATTTTTGTAAACATAAATTTTGCTATATTTTGCCCTTC
6081 TGTCGCTAATCTTTAACCACATCATTTTTGTTTTGTTGTGCAACCTTTATCTTATCAAAGTTTCATTTACTTTTTAGCCC
S
6161 GAAAAAAGGATTATGCACTCCGACCAAGCCAATGCCTGCATCGCCTTTTGCAGATGGACGAAGCAAAGAAGTAAATGTCAC
P K K G L C T P T K P M P A S P F A D G R S K E L M S
6241 CGTCAACACCGCGACGGCTGTTCAAGTCCAACCAAAGAATCGCCATCAAAGTCGTTGCCAGCTTACCAACGATTCCAGTCG
P S T P R R L F S P T K E S P S K S L P A Y Q R F Q S
6321 CTAGTGGAAGTGGTCGGCCAACGCTTCAGTTGCCATACAAGTATCGGTGCTTAGCAGAAACGTTCAAATGTGTGGACAC
L V E S G R P T L Q L P Y K Y R C L A E T F K C V D
6401 CGTATGCTCGATGTTTTTCAATCGACGCGAAGCTATTACTATGAAAAAAGTGAACCCGGCTGTTCAACGGATGATGCGAA
T V C S M F F N R R E A I T M K K L K P A V Q R M M R
6481 AGAATTTCTACGAAACATCATTGGCCCAAATCAATCATTGTACCCCGAAGCCTATAAATTTTCCCAAAGAAAATGTGCG
K N F Y E T S L A Q I N H L Y P E A Y K F S Q K K M S
6561 AATCCAGGTTCCCTGACGAAATATGATTACTATCAACTGGTAATTACACCAAATGTGACCGAATGTAGCGCAACTGTTAC
N P G S L T K Y D Y Y Q L V I T P N V T E C S A T V
6641 CAAAACCTACAAACACTGACGACGTCATTAAATCAGGTGACGAAAAATCGATGAATCCACAAGTATGATTGAGCGACAAC
T K T T N T D D V I K S G Q Q K S M N P Q V M I E R Q
6721 AACACTTGATATAAATGTTGCTAGATTTGGTAAAGGCAGAACATAACAAATTTCTACAAAACCTTAATCCACCCATGAAC
Q H L H K L L L D L V K A E H N K F L Q N L N P P M N
6801 ATTAATCCGGACAAAAGTAACGCGTTGGCATCCTGAATTTGATTTGCAAAACCTGTCAGACGTCAGTCCAAAAGGAATTGCC
I N P D K V T R W H P E F D L Q N C P D V V Q K E L
6881 GCGACCACCAATATTGAAAAGTTTTCTTCAGCAAAAGACATTTTGTCAACCGCTCGGAACCTGTTGATTGCGCTACAC
P R P P N I E K F S S A A K D I L S T A R N L F D C A T
6961 CAAAGGAACGAATGATGCAACGTTACGAAGCGAAGCAAAAGTTAGAAAAGCAACCGGCACCTTGTGGAACCGAAAGCTTAT
P K E R M M Q R Y E A K Q K L E K Q P A L V E P K A Y
7041 GATCCAGTGTCCAGTGTGTTTGAAGGTGTTCCAAAATCACTGATCGAACGAATTCGTGCACGTGAGCGGCCAAAGCTTT
D P V S S V L K G V P K S L I E R I R A R Q A A K A
7121 AGATTGATGACACGTCGGCCGTCGACGAGCAAGGAGTGTAATAATTATGGAAGACTGCCCGATTGCGCGACATATTC
L D S M T R R P S Q D K E C N N Y G R L P D L A R H I
7201 GAAACATCTTCGTAACGGAGAGTAAGGGAGTACTGGATTTGGATTTCGTTTTAACAAAGCTTGAAAACAGTTACCGTGAA
R N I F V T E S K G V L D L F V L T K L E N S Y R E
7281 CGTGTACCGGTAATGATATGAAAGATCTCTGCAATTACTGGCCAAGGAGACACCAATTAATTGGTTAACAATTTCCACA
R V P R N D M K D L L Q L L A K E T P I N W L T F P
7361 TCTACGGGACAGAGATTTCTGAAAATAAATAAAGTGAAGATTTGACCGAAATTATCAAACAGTTGGAAGCTAAGGCTG
H L R D R D F L K I N K S E D L T E I I K Q L E A K A
7441 AAGAGAAACGCTGAATCGTATAATCTTGTTTAAAGACATTAAGTTTTAATTATCACTTTTAAGTTTGGTTGAGTGAACCTT
E E K R *
7521 TTACTGTTGACTTCCGTTTCTGTTTCGAGAAATAAATGGCACTTTTGACCATATTTTAAACAACACTCGTGTTGAATTTTTTTCA
Sinal de Poliadenilação Sítio de Poliadenilação
7601 CTTTTGCATAAGTTCGTAAGTAAAAATGTCCTCATAAATTTCAAGTGTCCGAAAAATATTTAAATTTTTTTTTTAAAT
Região Rica em U -->
7681 TCATTTAGGCCACAATAACCGATACTGCTGAATGTAGTTCAAATGACCGTTTTTATTTTTTCAAACCTAGCTGTTTATGCCA
7761 TCGAGTGGTAAATTTCTTCGATCAATTGTAAAGCATCTATCAGAATTTTGCTTTACTTTTCTGTAAAATTTTTAAAAA
7841 TCCGCCACCTCAACAAGAACAGCAGAAGAAATTTTTATCGTTTTTTCAGATGGCTGTCAAAAATATCGCCTACAGCGCAAT
7921 GATAGTACCATTGTAACCTAATCAGCTTTTCACTAGCTGTAAAATGATAAATCCGTTGAAGTTTAAACAAAATACCTTTTA
8001 GATACCTTTTCTGAGTGGGCGGTTAGCATATCGACCGTGACGCAAGTCCACTGCCAGCAAAAATTAATAAATTTTCTT
8081 AGGCGTAGATTACAGCAAATCTCCTCTTTGCTCTGAATAATTCCAAGTTAGTCATAAAAGTTATAGGTACGAGTTCAA
8161 CCAGTGTGCAACTATGTGATTAAAGTGGTCACCCAAAATGTATGGAATAAACCTTAGGAAATCAGTTCAACCAAAATGA
8241 GTGGAATAAAGGCTACTTTTCATCTCAACCAAAATAAGGACGCTTTGCAAAAGAGTGTTCTTTGCAACTAATACTATTATT
8321 ATAATAGTTGCTACATCTAGACTGCAATCTAGCTTAAAGTGAATGTGCTCAGATAGTGTGTTTGGTCATCACTTTTACA

8401 AATACTTTAAAAGGCTAAAAGAGTGAGTGGGATGTGATTATTAAAGCTCTCCCTTTGCTATTATAATAATTAATATTAA
8481 TAGAGAGGCATATAAAAAATCTTAGCTCGCTTCGCTCGCGATATGTAAAATAAATTAACGCTCTCAGCGGTCCTGAGAT
8561 AAAACTGAGATAATGACATGTATATGAAAAATGAATATTGCAGCCATCTGTTCAAAACCCACTATTTCCCGATCGATGACC
8641 ATATGTATGATCTCGTTACGACTACTCGATAACCTCACATTAGCATTCCATCTACAACTTAATCCCAAATGTTCTAAATT
8721 CAGGGTATTATTTAAATGTAAATTAATTTGAAATCGACATCTCCGAAAATGACTTGTGCGATGTCATGTTTTGTAAAAAG
8801 TCGAACCAATTCACAAAGTCCCTCAACATAAACATATTTCTCTCGATTACCAACGTTGAGTGGATAAAGTAAACCACACA
8881 ACTCTCAGCCACATTTATTCTTACACTCATTCTTCATGCACACACACAAAATGATACGTCTTACCAATGTTCAATAGCAA
8961 CGGATTGGACATCTCCAAAGATTTATAATAATTTTACAGAGTTTTCAACCATTGTCGTATTGTTAGATAGCAAAAGTTTT
9041 GAAACTAACGCCGATAATTTATTCAAATTTGATTAATGTAAAGTCACACAAAAGGTGTACGTATGGAATTTCTTAAAGTC
9121 TTATGTGCGGCTGGTTTTATTTCGGAATAAGTGGCATGGTATGTGGATTTACCAAGCAGTACATAAGGAAATTTAATTTGC
9201 GGTAAATAGGATTTTGAAGGAAAATTTGATGCAGATTTTCGTTAGGTATATATAGTTGTGCTCTCTAACGACGTTATG
9281 TAATTAACTACAAGAATGAGATAATTTCTTATTGTTGACGAAAATTTGGTTGTAATTAATTTGTTATACTAGCGATACAA
9361 TGAATATAAACACTCGCAGGAGTATGCATTGATGTAAGTGGTAGGAGAGTATCAGAGTTACTCTAGAGTTACTATAATG
9441 TCACCATCGAGTGTCTCTACTCAATAAAAAATGCAATAAAATAATTTTAGTCTCGGTTGCAATGTGTCTGTTCTACTTCTGC
9521 ACTTATTGCAAAATTAGTTTTCATATCACCTAAAGAACACTAGATAGCTACCATACACACTACATCCTCCATAATACAGT
9601 AAAGTCGAAAACAAATTTCCCAAATCTTTTACGATATACGATTTTCAGATTTGTCTAACATCGTAGCATTTTTTATATCGT
9681 CTTCTGGGGTAAATCTTCTTGATATTTATTACATCCACCGTATGTTTCAATGAATGCATGAAGATTTTATGTTAGCTA
9761 TAACGTAAATGTACATCTCCGTCTTAAAGTATTTCTACACAAAACATAATCCACAATAAAATGAAAATGAAAATAAAT
9841 AATCGGAAAACAGACTAAAACCTTTTGAGTCTATAAAAAGTATATCCAGGAAAACATTTCTTTTTGTTTGCCTTTCTCT
9921 GTATTCAGGAATTTTTAGAATATAAGGAGCAAGAATAGAATAATTGAGTTTTACGGTGACTTTGTCTTCTAAAATTTGCGT
10001 AACGCACAGGTACAGAATTTTTGGGGCTATATTGAAATCTTTCTCAAATGAAATAAGGAGTGCAAAGTGCTTTATTAGGA
10081 ACTAGATATGTAATAGACATCGTATGTCTAGTAAAACACTCTTGACTCTCTTTGATTTCCATGAACATAAACTATCTCC
10161 GCAACTTTCTCTACAACATATAAAAAGCCCTGTTTCATTCCCTTCCATTTTCAATGTCACATAAGAACAGTACTGGATCTAA
10241 AAACCCAGCCCAATTCTCTACAACCTTGCCTTCGAAAGTCACTTCCATTTTTTACGATTCTAGAGATACGATCCATCGA
10321 AATTCAGATTCTGGAAGTGGCTGAATCCCGGGGTAGGGACACTTTTCTGAAAAATACTGAAACATATGTTCCGTAAGC
10401 GAAATTTCTCTCGAAGATTTGTATCTAATGATAATGTGTTACATCTTAGCAGTGTGATGTTAACCATGAAAGCACTTTC
10481 AGTGACCTCAACTTTATGAAATAACAGAAGTCACATAAAATGTTGCGAAACGAAAACCTTCTCCTCAACTGTGTATCGGC
10561 ACGCCTATATTACCTTCATCAAACCTAAACTAAACAATCCATCCAAAAGCCCTCCATATCCACCACAGGCATGCATAAAT
10641 ACGAGCATATCCATTAAGCCATGATTAATAGAAATTTTTGATTGACAAACACCCAGTGATAAATATTTGAACAAGATA
10721 ATATAAATTTTGTGTTTGAATAATTAGCGAGACTCGGTTTTTAGAGGGATTCTATTACAACATCAAATCCATTTTCGATG
10801 GAAAATGTTAAGTATGAAAAGAAGCACAAAAATCAATACTTTTCAAGGTAATTTAAACAGTTAATGTTTTTTCGGTTGT
10881 GAAAATATTCGTTTCGTTGGCGCTTATGCAATAGAAAATGATTTTTTCGAAAATGATGTATGAGAATTGATTTTTTGTGG
10961 ATGAGTAATTACTGGCTTGAATATTTTCGGTTAATACAAATGAAAAAGAGATATTTTATTGATGATGACAAACGAAATTCG
11041 TATGGCTTGGTATATTAGATAGAGGCAGTAATTTATTTGCATCATCAAATGTGATTCAAAGCTTTCCCATTCATTTTACTA
11121 TTTATTTTTCGATCCCAAACGTTTGCATGTGTAACGACTAAGGAAAGGAAAAGCGGAAGCTTTTCGACGAAACCGGTTATCG
11201 ATAATAGCGGACGCTGTAATACCGGCTTTACCGTATTTTCCGGTTTTTTCATGCAATATTTTTCGATGAAGAATAAAGATAT
11281 GTCTCACTAATTGCATGTCTTTGAATATAGGTTTTCGGGAGGGCTCTTTCTTCTACCTCTACTTATTAATAACTTATAAGC
11361 GTTATACGAAGGAGTTGTAAAGCTTTTGTGACCTCTATTTGTCCAGGCATGTTTCATACAAATAGTATAGTCCAATGTCCC
11441 TTTTATGAATATCAGATGATTTCGAATATATACAGGGTATCAAAGCATTAGATACATTTTCTGGACCACCATGTGCAATC
11521 ACCCTATATTTATAGTGTGGTTGCATTCGGTGCCTTCGACTCGAGTCACAATCGTAAAAATAAAAAATCTCCAAACAAGTA
11601 CGTAATTAAAGTTTTATATTGCACATCTAAGTGATTCCCTAAAAAATCAGTTTCGTTCTTTTTGTTGATTATTTGTGCAT
11681 GTACCAACTACACTGACCAAGTCGTAATGCACAAGTGATGCAATTTATTCCTACAAAACAAGGTAACATTCCTACTGT
11761 TAATGATTATACATCATAAATTACTTGTGCGATAGTGAACTTGACGACCAACGTTAGCATAGGGCATCGTTTCTTTTAT
11841 TGTTAATTCTTACAGTAGACTAAGTCGCCGTGAAGAGGTTAATATTACATAAATCGTCTTCATTAAATTACCATTAATGT
11921 TTTGAATTAAATTACGTTTCATTTCTTTGATTCATCCTCTGGTCTCGTATTTATTTGATCTTTCTATCGAAATTCATT
12001 TCCCATCACACAAAACACACAATCAAACGAAACGAATCATACTAAACCTTTAATACACACATTTTCGCAAAAAAACGGTC
12081 ATGGAAATACACCTCTTATTTACCATGTTTCATTCAAAAAATAAATTTCCCTTGATTGCTCTCATTTATTTACGTCCTT
12161 GTCATGTGCTTCAACCTTTCAATGGAAAACGAAATGATTTACCATAGGATTTTCCCATGAACATTTTCTTCGTTTCAGCT
12241 TTTATTAGTTTGTCTATTTGCGTTAATGATTCTCGAATTGTCTATATGCTATGTACATTTTAAACAAGAAAGTGTGCAAA
12321 TAACATAAAATACATTGTCTTATGTTTTTCAGAACGATTTTCCGTGCCATTCCATTTTACTATGATGCTGTTTCAACATC
12401 CAACTCAATAAGAAAAACATTTTTCTATTTCGAAGAGTTAAGATTTCCCAATGAAATTATGCAAAATGGAATATTTATGT
12481 CATTTGCTAGAAGTAGCATTTTCGTATACATTATTTTCATTGATGGTATTTTGATATTTTAAATTAATTTCTTTCTCAATG
12561 ATTTTTTGTATATTTCTGAACGTTTCTTTCTTTTGGACATTTTCAGTATAGAAGATGTTAAGAGCTTTAACCATTTTAG
12641 CGTAAAGAGACAAATAAATGCCTCAGCCAAAT

Figura 13 – Sequência do inserto genômico contendo o gene RaDup, as porções realçadas em amarelo foram confirmadas pelo sequenciamento do cDNA.

A estrutura do gene RaDup que foi confirmada pela sequência do mRNA segue as características de um gene usual, possui ATG como códon iniciador, sítios de “splicing” que obedecem ao consenso GT-AG, sinal de poli-adenilação AATAAA e sítio de poliadenilação CA subsequente ao sinal de poliadenilação e precedendo uma região rica em U, possivelmente envolvida no processamento do mRNA (Manley e Takagaki, 1996). Entretanto na porção anterior não foi identificado um consenso TATA e a região promotora foi somente predita, não havendo confirmação do ponto inicial da transcrição.

Existem pequenas diferenças quando se compara o gene RaDup com seu homólogo em *Drosophila* e com seu putativo homólogo em *Anopheles* (agCP6189 – sequência ainda não anotada, identificada por similaridade) (Figura 14), em *Drosophila* a região codificante é dividida em três exons, intercalados por um íntron de 119 bases e outro de 1135, o homólogo de *Anopheles* possui dois exons e um íntron de 103 bases, RaDup por sua vez também é dividido em dois exons, mas possui um íntron bem maior constituído por 3249 nucleotídeos, o que o torna maior do que seus homólogos, englobado em uma região de aproximadamente 7kb.

Quanto à localização cromossômica do gene RaDup, hibridizações *in situ* utilizando o inserto genômico como sonda revelaram que está situado na banda 3 do cromossomo B (Figura 15 e 16), o que é muito interessante dado que esta região é um pufe previamente caracterizado.

Guevara e Basile (1973) observaram através de análises citológicas (câmara lúcida) que a região B3, em glândulas salivares, forma pufes temporariamente aos 20, 30 e 50 dias (2º período) e é amplificada a partir dos 64 dias (4º período). Também em túbulos de Malpighi e em intestinos foram observadas ocorrências de pufe nesta região em diferentes momentos a partir do 2º período até o 4º período. Machado-Santelli e Basile (1975) confirmaram por meio de incorporação de timidina tritiada e autoradiografia que a região B3 sofre amplificação durante o 4º período (Figura 17), atualmente experimentos estão sendo realizados com o objetivo de quantificar a amplificação e precisar o período em que esta ocorre.

Figura 14 – Comparação entre RaDup, DmDup e agCP6189 (homólogo putativo de CDT1 de *Anopheles*), baseada em seqüências genômicas.

Figura 15 – Hibridização *in situ* revelando através de marcação com NBT a localização do gene RaDup na banda 3 do cromossomo B.

Figura 16 – Imagens de microscopia confocal de glândulas salivares de *Rhynchosciara americana* revelando por hibridização *in situ* a localização de RaDup em B3 (cromossomo marcado com iodeto de propídio e banda B3 revelada com anti-digoxigenina).

Figura 17 – Auto-radiografias de glândulas salivares de larvas no final do 4º período, após incorporação de timidina tritiada, setas indicam incorporação em B3 (1: foto modificada com autorização a partir de Machado-Santelli e Basile – 1975; 2: incorporação de 20 minutos; 3: incorporação de 1 hora; 2 e 3 cedidas por Machado-Santelli, arquivo pessoal).

Pela primeira vez em *Rhynchosciara americana* um pufe, que em glândulas salivares é amplificado, pode ser relacionado com um gene regulatório de um processo celular. Não se pode justificar a amplificação do pufe B3 como uma forma de contribuir com a expressão de RaDup, como descrito na introdução a concentração de CDT1/DUP livre para acessar as origens de replicação é rigidamente controlada por diferentes mecanismos, de fato a superexpressão de RaDup não é observada, como descrito na próxima seção.

Possivelmente a amplificação desta região seja necessária à expressão de outro gene presente nesta banda, a presença do homólogo de XPC (caso venha a se confirmar que não se trata de um pseudogene) na adjacência de RaDup pode indicar que esta é uma porção genômica com uma incidência maior de genes, como é a região do cromossomo 2R de *Drosophila* que contém Dup e Mus210 (www.fruitfly.org).

3.2.2. – Proteína e Expressão

A proteína RaDup deduzida através da tradução conceitual de seu mensageiro apresenta um alto grau de homologia com as demais CDT1/DUP já sequenciadas (Tabela 10, Figura 18), em especial com DmDup, que possui 312 resíduos idênticos e outras 189 substituições conservativas em relação a RaDup. Possuindo 679 aminoácidos (78kDa) representa um homólogo de massa intermediária entre os de maior peso molecular, *Drosophila* (743aa) e *Caenorhabditis* (741aa), e os menores conhecidos até o momento, humanos (546aa), camundongos (547aa), *S. pombe* (444aa) e *Anopheles* (629aa).

Tabela 10 – Dados comparativos entre a tradução da mensagem RaDup e diferentes homólogos da proteína CDT1/Dup (Blastx)

Homólogo	Identities (%)	Positivos (%)	Gaps (%)	Score (bits)	E-Value
Double parked [<i>Drosophila melanogaster</i>]	40	54	12	452	10e-125
agCP6189 [<i>Anopheles gambiae</i>]	43	63	9	300	7e-80
Putative CDT1 protein [<i>Xenopus laevis</i>]	38	56	3	253	8e-66
CDT1 protein [<i>Mus musculus</i>]	36	55	2	244	5e-63
DNA replication factor CDT1 [<i>Homo sapiens</i>]	34	54	2	224	4e-57

Ra	1	MAQPSVTRYF	NSRKRTAPDD	VYTRKSK--I	IHLLEEVGCDK	QPNFLQDKED	LFNDLTTVEG
Ag	1	--QPKLVSFV	KKGN-----	-LSPQKR---	--PHTASPSK	RPAVPEKK--	-----
Dm	1	MAQPSVAAFF	TNRKRAALDD	AISIKNRRLV	EPALTVSPAS	APSQLPAG-D	QDADLDTLKA
Ra	59	KSS-PRINTN	-TANKTVTGL	ISDRKKTAT	KRRSCIRRS	QETAESNQPK	IVRFTLAGTL
Ag	34	-----	-IA-----	-SPVKVEPVA	AS-----	-----VVDSK	PVEFSPRNNL
Dm	60	AATGMRTRSG	RTARLIVTAA	QESKKKTAA	AK-----M	EP--HIKQPK	LVQFIKKGTL
Ra	117	SPKKKPKNE-	-----AQFS	AVFENMENV-	-----QKEGD	-----LHTPT	KEQQMEKAKA
Ag	63	N-----	-----	-----NVA	-----	-----KAPT	KASVSRKLQL
Dm	111	SPRKQAQSSK	LDEEELQSSS	AISEHTPKVN	FTITSQONAD	NVQRGLRTP	KQILKDasPI
Ra	159	EQDKQTLLNM	SLPEIKEKII	KSDRIKDLKA	SLNKLRSLEE	SRQKCIKONG	SLSSNFERGP
Ag	81	GKDKDAMS--	-LTDIRNKLL	QHPSLPELKT	KLNAIQSG-L	DKMDRLRRER	LEG----TKP
Dm	171	KADLRRQL--	TFDEVKTKVS	RSAKLQELKA	VIALKAALEQ	KRKEQEERNR	KLR---DAGP
Ra	219	ALPPVAQKMD	GTLGSNLKKF	DHIELEVLSS	PKKGLCTPTK	PMPASPFADG	RSKELMSPST
Ag	133	SVSPAVAAKN	-----LEKF	ATLDVEVMVS	PKKTIASPSK	LLR-TPTKDA	SSSPANNVTP
Dm	226	S--PSKSKMS	----VQLKEF	DTIELEVLIS	PLKTFKTPTK	IPPTTPDKHE	LMSPRHTDVS
Ra	279	PRR--LFSPT	KE----SPSK	----SLPAYQ	RFQSLVESGR	P-TLQLPYKY	RCLAETFKCV
Ag	186	KRLSALMSPI	KEPTAGTPVM	ASPKKLPAYQ	REHNLVEAGV	P-TLQLPEKY	RSLLLEFKCT
Dm	280	KRL--LFSPA	KNG---SPVK	LV--EVPAYK	RYASLVESSR	AGQLPLPYKY	RHLLDVFKGL
Ra	328	DTVCSMFENR	REAITMKKLK	PAVQRMKRKN	FYETSLAQIN	HLYPEAYKFS	QKMSNPGSL
Ag	245	DTVCSMLHNR	KEQITFKKLK	PAVQRMARKN	FFESHLAGIY	SLFPDAFTLS	QEMTKNYGSA
Dm	333	DSVVAMFHN	KETITFKKLK	PAVQRMKRKN	FTETHLAQIK	HIYPDAFIFS	QVKTNRFGSV
Ra	388	TKYDYYQLVI	TPNVTECSAT	VT-----K	TTNTDDVIKS	GOQSMNPQV	MIERQOHLHK
Ag	305	TKHETYQLVI	RPNIADAPEE	PA-----R	KAAEDDEFRT	NTKPAVNSQS	LLERYQHFR
Dm	393	SKADYEQLI	APNVEPLPEQ	QQSEKPOHFT	KINEDDVLAS	AQSTSMNPV	MTARMQRFQS
Ra	441	LLLDLVKAEH	NKFLONLNP	MNINPDKVTR	WHPEFDLQNC	PDVVQKELPR	PPNIEKFSSA
Ag	358	LLLERITKDAH	QAFLQTLDP	LNIDRSKIVR	WHVDFDLENC	PDIEKAELPQ	PPNVERFSSA
Dm	453	LLLDRAHRAH	DQFLRSQDPP	ITIEK-ALTR	WHPOFDLESC	PEVELSPLPQ	PPNVEKYSSA
Ra	501	KDILSTARNL	FDCATPKERM	MORYEAKOKL	EKQ-----	-----	-----
Ag	418	RDVLSTARNL	FSCGTPMERA	LARLEEKKKQ	DAAVATRGG-	-----	TPQPTTTEPE
Dm	512	KDILSTARNL	FNCATPMEAR	MDRYEAKLES	EKQQAESNK	KTEEQQAGEV	TRTSTAIQTS
Ra	533	-----	-----	-----PALV	EPKAYDPVSS	VLKGVPKSLI	ERIRARQAAK
Ag	467	DKKPPPIG---	-----IKT	EKVSTTPSTT	TALLSDPAEQ	MLKNVPKALL	EKIRAKQAAK
Dm	572	QEVPGISGSS	KNPTVPETTN	QTETAKPTVK	DCTVPDASSN	LLKGLPKSLI	EKIRAKQAAK
Ra	568	ALDSMTRRPS	QDKECNNGR	LFDIARHTRN	IFVTESKGVL	DLDFVLTKE	NSYRERVPRN
Ag	517	ALEQMVRRPS	QEKEAMKYSR	LPEIARHLRN	VFVTERKNVL	PLETPVVKIE	NSYRGKLTQ
Dm	632	ALEAMTRRPS	QDQEATKYSR	LPELARHLRN	VFVTERKGVL	TEVVIKKIQ	NSFRANLTPO
Ra	628	DMKDLLQLLA	KETPINWLTF	PHLRDRDFLK	INKSEDLTEI	IKOLEAKAE	KR-
Ag	577	ELEEHIIRMA	QLVPF-WLTL	PEVRKVYAK	IAKCDLAKV	IDVLERKANE	VVQ
Dm	692	EIEAHLKLLA	KELPS-WASF	HEVRKTMYLK	VAKDMMNKI	IEKLESVANA	KSN

Figura 18 – Comparação entre as seqüências protéicas de Dup de *Rhynchosciara americana* (**Ra**), *Anopheles gambiae* (**Ag**) e *Drosophila melanogaster* (**Dm**), os aminoácidos realçados em negro representam identidades, em cinza denotam substituições conservativas.

Na sequência protéica de RaDUP foi possível identificar o domínio KOG4762 (CDART⁵), característico das proteínas CDT1/DUP, entre os resíduos 196 até 672, uma busca no serviço PFAM (www.sanger.ac.uk) também foi capaz de identificar esta assinatura além de um potencial sítio de ligação de DNA (HLH – Helix-Loop-Helix) a partir do resíduo 615 até o 675, o que, se confirmado, seria coerente com as observações em camundongos sobre a capacidade de CDT1 ligar-se com DNA (Yanagi *et al*, 2002).

Baseados nas seqüências protéicas dos homólogos de CDT1/DUP já estabelecidos na literatura e na tradução de nossa mensagem, foi construído um dendrograma relacionando as espécies de forma coerente. A árvore filogenética foi construída pelo método Neighbour Joining, com o programa clustalx (Thompson *et al*, 1997) (Figura 19).

Figura 19 – Dendrograma gerado por “Neighbour Joining” para homólogos de CDT1/Dup, baseado em seqüências protéicas, os valores em preto correspondem às distâncias filogenéticas calculadas, em vermelho são mostrados os valores de “Bootstrap”.

⁵ www.ncbi.nlm.nih.gov

A atividade de CDT1/DUP na montagem do Pré-RC tem sido muito bem caracterizada em organismos tão diversos quanto leveduras, *Xenopus* e humanos, devido o alto grau de similaridade de RaDup com seus homólogos, seria esperado o mesmo comportamento desta proteína em *Rhynchosciara*, constatada a participação de DmDup em um cenário de amplificação gênica nos genes coriônicos de *Drosophila* (Whittaker *et al*, 2000; Claycomb *et al*, 2002) tornou-se uma necessidade caracterizar a participação de RaDup em tecidos em mitose e em endoreplicação neste modelo.

O mensageiro de RaDup no banco de ESTs representa aproximadamente 0,1% das seqüências, um nível de expressão concordante com um gene envolvido na regulação de um processo biológico como a replicação, um sinal razoavelmente nítido em Northern só foi obtido empregando grandes quantidades de RNA (Figura 20). Nestes Northern ficam evidenciadas as presenças deste mRNA no 4º e no 5º período de *Rhynchosciara*, em embriões de 1 dia, e por uma leve marcação no 3º período.

Figura 20 – Northern mostrando a presença do mensageiro RaDup em glândulas salivares em diferentes estágios de desenvolvimento e em embriões de 1 dia (1 – 2º período, 2 – 3º período, 3 – 4º período, 4 – 5º período, 5 – 6º período, 6 – Embriões, M – RNA total de *E. coli*)

Os Northern de RaDup não permitiram determinar com clareza sua expressão em embriões e ovários. A partir dos mRNAs purificados destes tecidos, tratados com DNAase, foram feitas reações de RT-PCR que sugerem a expressão de RaDup nestes tecidos (Figura 21). Os primers utilizados flanqueiam o íntron do gene RaDup (Race-1 e DUP-atg – ver mapa de primers em Materiais e Métodos), e os tempos utilizados na fase de extensão da reação de PCR não permitiriam a amplificação de toda a sua extensão, amplificando apenas uma banda de 862 nucleotídeos, além

deste controle interno, as bandas geradas foram clonadas e seqüenciadas, confirmando se tratar do mRNA de RaDup.

Figura 21 – Produto de RT-PCR (862 bp) utilizando os primers Race-1 e DUP-atg, mostrando a presença do mensageiro RaDup em ovários (O) e embriões (E).

A detecção desta mensagem, no 4º período do desenvolvimento de *Rhynchosciara*, em embriões e em ovários, indica uma possível participação desta proteína tanto na replicação normal de um tecido em desenvolvimento como na amplificação gênica das glândulas salivares. Esta idéia é reforçada ao lembrar que os mRNAs presentes nos ovários representam um estoque de mensagens maternas, e a expressão zigótica só ocorrerá depois de diversos ciclos celulares após a fecundação.

Com a finalidade de caracterizar a presença da proteína RaDup em diferentes tecidos foram produzidos dois anti-soros em camundongos (SWISS e Balb-C). Construiu-se um vetor (pGEX4T2-RaDup – Materiais e Métodos) capaz de expressar os últimos 433 aminoácidos de RaDup fusionados com Glutathione S-Transferase (82,5kDa), que quando induzido em *E. Coli* BL21 expressa esta proteína de fusão e forma corpos de inclusão em todas as condições testadas, estes foram purificados e utilizados para imunizar os camundongos. A região de RaDup escolhida para esta imunização corresponde à mesma região utilizada com sucesso para produzir anti-soro contra DmDup (Whittaker *et al*, 2000), considerada pelos autores com maior potencial antigênico.

Optou-se por utilizar o soro produzido em camundongos SWISS, por apresentar um título maior. Apesar de seu pré-inóculo não ter apresentado nenhuma marcação com os extratos de tecidos de *Rhynchosciara*, algumas marcações inespecíficas foram observadas, possivelmente devido à contaminações no inóculo.

Westerns marcados com o anticorpo policlonal RaDup indicam a presença desta proteína durante o desenvolvimento da glândula salivar (Figura 22), a partir do IR começa a ser nítida sua marcação, tornando-se mais clara em extratos de glândulas em B e C, e em extratos de glândulas de pupas sua marca torna a desvanecer, em extratos

de ovários também foi possível observar a presença de RaDup, devido a aplicação de excesso de material nesta marcação houve “inversão” da reação de revelação, em separado pode-se observar a presença de RaDup em extrato de ovário, em um novo Western. As fases de desenvolvimento das glândulas salivares utilizadas foram confirmadas citologicamente.

Como controle negativo escolheu-se utilizar extratos de cabeças e de intestinos, ambos de adultos, por serem órgãos já desenvolvidos. No entanto foi possível observar um sinal claro em extratos de intestino, e outro mais tênue em cabeças. Nestes órgãos pode estar havendo alguma reposição celular em algum tecido (epitélio do intestino por exemplo) ou alguma atividade replicativa, ou estamos observando um remanescente desta proteína (complexado com seu inibidor geminina) que ainda não foi degradado, o que explicaria também a marcação em pupa, onde as glândulas salivares começam a regredir. Como controle positivo a proteína de fusão GST-RaDup foi marcada tanto com anti-RaDup como com anti-DmDup.

Para confirmar estes dados estes Westerns foram repetidos utilizando o anticorpo anti-DmDup (Whittaker *et al*, 2000), gentilmente cedido pela Dra. Terry Orr-Weaver⁶, o mesmo padrão de marcação foi novamente observado, demonstrando a presença de RaDup tanto em glândulas salivares, em um período em que estão ocorrendo endociclos, como em ovários e intestinos de mosca (Figura 22).

Outra forma encontrada de validar estes dados e indicar os tecidos e as fases de desenvolvimento em que está ocorrendo replicação foi marcar Westerns contendo os mesmos extratos com um anticorpo contra uma proteína que faça parte ou trabalhe junto com a maquinaria de replicação, como atualmente RaMCM5 também está sendo caracterizada, utilizou-se um anticorpo anti-DmMCM5 (Loebel *et al*, 2000), enviado pela Dra. Sue Cotterill⁷, como resultado deste protocolo foi observado o mesmo padrão já obtido nos Westerns anteriores.

Apesar de terem sido padronizadas as quantidades de extratos aplicados nos Westerns, ainda não utilizamos uma marcação contra uma proteína constitutiva para servir de controle de quantidade (actina por exemplo), assim apesar de serem perceptíveis mudanças nas intensidades das bandas, ainda não serão tratadas quantificações de RaDup.

⁶ Whitehead Institute - Cambridge

⁷ St. George's Hospital Medical School – Universidade de Londres

Figura 22 – Western Blottings de glândulas salivares em diferentes fases do desenvolvimento, ovários, cabeças e intestinos de adultos, marcados com anticorpos contra RaDup, DmDup e DmMCM5 (2P – 2º período, IR – início de rede, B – pufe B2 expandido, C – Pufe C3 expandido, Pupa, Ova – ovários, Cab – Cabeças, Int – Intestinos de adultos)

A utilização destes anticorpos em extratos de glândulas salivares e de outros tecidos mostra a participação de RaDup em replicação de células mitóticas e em endociclos, conforme esperado.

3.2.3. – RaDup no gênero *Rhynchosciara*

Aproveitando primers utilizados para o seqüenciamento do cDNA de RaDup (primers Dup e Dup-Right), em um momento em que ainda não estava completa a sua seqüência genômica, foram preparadas reações de PCR utilizando o cDNA, DNA genômico de fêmeas adultas e o DNA do fago isolado (até então putativo RaDup) como molde, com a intenção de confirmar a identidade do fago, e talvez obter alguma informação sobre a estrutura deste gene, já que a presença de um íntron poderia ser detectada. O resultado obtido foi a mesma banda de 1100bp como produto de todas as reações, o que confirmou que o fago continha pelo menos grande parte do cDNA conhecido naquele momento e que não havia íntrons naquele trecho codificante, como ficou evidenciado depois.

A mesma reação de PCR feita para DNA de *R. americana* foi feita para um DNA de uma espécie não descrita (*Rhynchosciara sp*), a qual é chamada em nosso grupo de trabalho de *Rhynchosciara* “marrom”, e o resultado foi idêntico ao já obtido, o que incentivou a pedir ao Prof. Dr. Eduardo Gorab⁸ DNAs genômicos de outras espécies de *Rhynchosciara*, gentilmente cedidos. Repetidas as reações para as espécies *R. hollaenderi*, *R. Baschanti* e *R. milleri*, novamente o mesmo resultado anterior fora obtido (Figura 23).

Figura 23 – Produto de PCR utilizando os primers de RaDup e diferentes DNAs (Ra – *R. americana*, Rsp – *R. sp*, Rb – *R. baschanti*, Rm – *R. milleri*, Rh – *R. hollaenderi*)

Um “Zoo-Southern” foi montado e a sonda utilizada foi o DNA genômico do gene RaDup (inserto do fago) (Figura 24), onde é possível notar que as marcações entre as diferentes espécies são muito parecidas, porém algumas diferenças nos tamanhos de bandas (banda de 9kb de *SmaI*) podem indicar polimorfismos ou

⁸ Instituto de Biociências – Universidade de São Paulo

inserções/deleções nesta região que compreende o gene Cdt1/Dup. A marcação positiva de uma sonda heteróloga e o reconhecimento dos primers utilizados em diferentes espécies, indicam que este gene é conservado dentro do gênero.

Figura 24 – Southern de DNAs genômicos de diferentes espécies de *Rhynchosciara*, (Ra – *R. americana*, Rsp – *R. sp*, Rb – *R. baschanti*, Rm – *R. milleri*, Rh – *R. hollaenderi*)

As similaridades cromossômicas entre algumas espécies deste gênero já são descritas (Stocker *et al*, 1993; Stocker e Gorab, 2000), sendo observado entre *R. americana* e *R. hollaenderi* o mesmo perfil de desenvolvimento de pufes, inclusive regiões destes pufes clonadas hibridizam nos dois organismos nas mesmas posições, o que demonstra uma semelhança muito grande nas seqüências de nucleotídeos dos genomas correspondentes. Entre as outras espécies, como *R. milleri* por exemplo, podem haver diferenças mais pronunciadas, visto que possuem características mais distantes de *R. americana*, tais como cores das larvas e formação de casulos individuais, porém como foi observado o mesmo produto de PCR para todas as espécies utilizadas e os Southern mostram muitas semelhanças entre si, espera-se que as diferenças no gene RaDup destes organismos limitem-se a polimorfismos e no tamanho do íntron.

4. Conclusões

O trabalho de seqüenciamento de ESTs de *Rhynchosciara* representa para o modelo um grande passo para iniciar novos estudos, e é neste fato que reside sua importância. Como em todo projeto de seqüenciamento de transcriptomas foram gerados grandes volumes de informação, difíceis até de listá-los em uma tese sem perder o foco ou tornar o volume de dados enfadonho ao leitor. De fato preferiu-se manter o foco sobre o controle do ciclo celular, e evitou-se listar 8193 linhas com resultados de “blast” ou anotações. Estes ESTs foram todos depositados no GenBank, como uma forma de oferecer acesso livre à estas informações.

É verdade que foi possível caracterizar o estado metabólico das glândulas salivares na fase correspondente ao início de rede, identificar possíveis componentes envolvidos em biossíntese, secreção, regulação da replicação, processamento de mRNA e alguns possíveis “secretados” entre outras mensagens interessantes, porém os ESTs por si somente não permitem concluir a respeito de causas e efeitos sobre processos biológicos, desenvolvimento de tecidos ou regulação celular. Constituem sim, um conjunto de dados que deve ser associado a outros para promover uma caracterização mais detalhada do aspecto em questão.

Como exemplo de como estes dados podem formar um alicerce para outros estudos temos o trabalho com RaDup e RaMCM5 (entre outros em andamento em nosso grupo), que tiveram seu início no banco de dados de *Rhynchosciara*. Foi possível identificar em RaDup pela primeira vez uma conexão entre um pufe amplificado e um gene regulatório (pelo menos um), também sua participação em replicação de células mitóticas e em endociclos, o que está em concordância com as funções de seu homólogo mais próximo (DmDup).

Restam muitas perguntas a serem respondidas a respeito da amplificação em *Rhynchosciara*, esta coleção de seqüências pode ajudar a responder algumas, e/ou a levantar mais outras tantas (como levantaram durante a anotação). Lembrando que não foram encontrados correspondentes no GenBank para muitas mensagens seqüenciadas, e que ainda não são definidos elementos de controle nestes amplicons, tão bem detalhados como nos genes coriônicos de *Drosophila*.

De fato, a existência de tantas seqüências conhecidas de *Rhynchosciara* permite não só a continuidade dos estudos sobre amplificação, como o início de

pesquisas em diferentes aspectos de seu desenvolvimento. Sem a necessidade de o primeiro passo desta pesquisa ser “heterólogo”, o que muitas vezes desestimula o início de uma linha neste modelo.

5. Perspectivas

Os próximos passos na pesquisa de RaDup envolvem obter imagens *in situ* de seu envolvimento com a replicação em ovários e embriões em diferentes fases, entre outros tecidos, quantificar sua transcrição em glândulas salivares, túbulos de Malpighi, cabeça e intestino em diferentes fases do desenvolvimento através de RT-PCR quantitativo. Estimar a quantidade de proteína RaDup nestes mesmos tecidos por “Western blottings” utilizando um controle interno (actina ou tubulina). E por fim quantificar a amplificação do pufe B3 e precisar o período em que esta ocorre, através de Southern quantitativo.

Em relação aos ESTs, o seqüenciamento de outras bibliotecas de glândula salivar em fases anteriores, e de outros tecidos, poderia ter seus dados associados aos atuais gerando “mapas” dos transcritos em relação aos tecidos e ao desenvolvimento. Quanto aos possíveis elementos envolvidos em replicação, estes estão sendo isolados e seqüenciados em ambas as extremidades, conhecendo uma porção maior das seqüências destes candidatos talvez seja possível escolher as mensagens com maior probabilidade de participar da regulação dos endociclos em *Rhynchosciara*.

Outros resultados que possivelmente surgirão em breve são identificações de transcritos de puffes ainda não descritos, através de hibridizações *in situ* das mensagens anotadas como “pufes putativos”, possivelmente estes transcritos, ou uma parte deles, venham a ser confirmados como codificando proteínas secretadas constituintes do casulo.

Fica a expectativa de que este conjunto de dados contribua com outras pesquisas em *Rhynchosciara*, e possa servir de estímulo para novas abordagens neste modelo ao mesmo tempo “adorado” e “temido”.

6. Materiais e métodos

6.1. - Animais

Foram utilizadas neste estudo larvas de *Rhynchosciara americana* (Wiedemann, 1821), que foi redescrita como *Rhynchosciara angelae* por Nonato e Pavan (1951), cuja sinonímia foi demonstrada por Breuer (1969). As larvas foram coletadas em sua maioria em Ubatuba-SP e eventualmente em Monguaguá-SP, sendo mantidas em laboratório por algumas gerações segundo método descrito por Lara *et al* (1965).

A maior parte do desenvolvimento larval é compreendida pelo 4º estágio, que se inicia na 3ª muda larval e se estende até a muda pupal (Morgante e Guevara, 1969). Este estágio foi subdividido em períodos, de acordo com modificações fisiológicas e comportamentais da larva, por Terra (1973): 1º Período, a partir da 3ª muda (aproximadamente 21 dias de idade), durante a fase em que as larvas apresentam-se com coloração clara; 2º Período, caracterizado pelo escurecimento das larvas (aproximadamente 32 dias de idade); 3º Período, início de rede, onde ocorre a secreção do casulo (que dura em média 11 dias); 4º, 5º e 6º Períodos são caracterizados pelo surgimento e regressão de diferentes pufes e modificações na textura do casulo.

6.2. – Construção da biblioteca de cDNA

A biblioteca de cDNA foi construída pelo Prof. Dr. Roberto Vicente Santelli com auxílio da Dra. Maria de Fátima Bonaldo no laboratório do Prof. Dr. Marcelo Bento Soares⁹, a partir de glândulas salivares do 3º período (Início de Rede), através do método descrito em Soares *et al* (1994). Foram produzidas duas versões desta biblioteca, uma normalizada e outra não-normalizada, construídas no vetor pT7T3-pac. Após testes preliminares foi decidido prosseguir o seqüenciamento com a biblioteca não-normalizada.

⁹ Universidade de Columbia – Nova York)

6.3. – Sequenciamento de DNA

Tanto para o sequenciamento ao acaso dos clones de cDNA da biblioteca de cDNA como para sub-bibliotecas genômicas utilizou-se a mesma metodologia.

Bactérias competentes *E. Coli* DH10B preparadas em cálcio 50 mM foram transformadas com estes bancos (cDNA ou genômicos), e as bactérias transformadas que continham inserto no plasmídeo foram selecionadas no crescimento em placas com meio LB (10 g/L NaCl, 5 g/L extrato de levedura, 10 g/L triptona) contendo ampicilina 100 mg/L, IPTG e Xgal.

As colônias selecionadas foram crescidas "overnight" em LB com ampicilina 100 mg/L e os plasmídeos foram extraídos com o kit Concert (Gibco) baseado em método de lise alcalina e com uso de coluna de sílica (Birnboim e Doly, 1979; Volgelstein e Gillespie, 1979).

Posteriormente, após a atualização do sequenciador para 96 amostras, utilizamos o crescimento em placa de 96 poços e filtro Multiscreen (Millipore), onde as transformantes eram crescidas em 1 mL de TB (10 g/L NaCl, 24 g/L extrato de levedura, 12 g/L triptona, 0,4% glicerol, ampicilina 100mg/L) durante 18 horas, coletadas na própria placa de crescimento por centrifugação a 4°C (4000 rpm, 5 minutos), ressuspensas em 85 µL de GET (10mM EDTA, 50mM Glicose, 25 mM Tris-HCl) contendo 0,5 mg/mL de RNAase. Eram então lisadas com a adição de 60 µL de solução de lise (0,2M NaOH, 1% SDS) e neutralizadas com 60 µL de acetato de potássio 3M e mantidas a 90°C por 30 minutos. Após centrifugação de 15 minutos (4000 rpm, 4°C) o lisado era filtrado em Multiscreen (Millipore) montado sobre placas de coleta com fundo em V contendo 110µL de isopropanol através de nova centrifugação a 4000 rpm (40 minutos, 4°C), ao final desta os DNAs já precipitados foram lavados com etanol 70% e ressuspensos em 10 mM Tris-HCl (pH 8).

Os DNAs obtidos foram verificados em gel de agarose 0,8 %, a partir destes DNAs preparou-se as reações de PCR de sequenciamento utilizando o modelo de PCR GeneAmp 9700 (Perkin Elmer), terminador BigDye (Perkin Elmer) e as condições sumarizadas na tabela 11.

Os produtos de PCR foram precipitados adicionando em cada amostra 54 µL de solução gelada de precipitação que continha 50 µL de etanol, 2 µL de acetato de sódio (3 M, pH 5) e 2 µL de glicogênio (1 g/L), após uma breve agitação em um vórtex

as amostras permaneceram em gelo por 15 minutos, foram então centrifugadas por 30 minutos a 4000 rpm e a 4°C, após o descarte do sobrenadante adicionava-se 50 µL de etanol gelado e a centrifugação era repetida, as amostras foram secas no escuro a temperatura ambiente por uma hora, eram então ressuspensas em 1,5 µL de tampão de entrada (loading buffer : blue-dextran e formamida).

Tabela 11 – Condições empregadas nos PCRs para seqüenciamento.

Reação	Ciclos	Temperatura (°C)	Tempo
2 µL de BigDye 6 µL de Tampão SM 6,5 µL de água 4 µL de DNA 1,5 µL de Primer	1	96	1 min
	35	96	20 s
		50	1 min
		60	4 min
	1	4	∞

Utilizou-se um seqüenciador modelo ABI-Prism 377 (Perkin Elmer), equipado com placas de 36 cm montadas com gel de acrilamida 5%, as corridas foram de 4 horas de eletroforese quando o aparelho era preparado para 36 amostras, e de 8 horas após seu "upgrade" para 96 amostras. As análises dos géis de seqüenciamento foram feitas em um Machintosh G4, utilizando o programa DNA Sequencing (Perkin Elmer).

6.4. – Análise e Montagem dos ESTs e Seqüências Genômicas

As seqüências foram processadas em um computador Pentium III 933 MHz rodando sobre o sistema operacional Linux (Debian Testing com kernel 2.4.18), com os programas Phred, Phrap, CrossMatch e Consed 11 (Ewing e Green, 1998a e 1998b; Gordon *et al*, 1998) compilados, cujas licenças de uso para fins acadêmicos são gratuitas e obtidas diretamente com os autores. Eletroferogramas gerados no seqüenciador foram convertidos em seqüências, as quais foram aparadas, extirpando-se extremidades de baixa qualidade, e as regiões seqüenciadas de vetor foram mascaradas. Em seguida ao processamento as seqüências foram alinhadas, separando as que se sobrepõem em "contigs" e as isoladas em "singlets". Para o seqüenciamento de regiões genômicas foram mantidas as configurações padrão do phrap, porém para alinhar ESTs foram necessários ajustes a fim de evitar formação de falsas quimeras, foram

modificadas as penalidades de “gaps”, a janela mínima de alinhamento e a nota mínima deste, aplicando a seguinte linha de comando no momento de alinhar as seqüências:

```
:$ phredPhrap -penalty -15 -shatter_greedy -minscore 150 -bandwidth 14 -minmatch 55
```

Através de outras ferramentas de biologia molecular, tais como o pacote Vibrant 6.1, blast2 e o Blast Client 3, softwares distribuídos pelo NCBI (National Center for Biotechnology Information, <http://www.ncbi.nlm.nih.gov/>) foi possível gerar um banco de dados local de *Rhynchosciara* e comparar estas seqüências com bancos internacionais de forma automatizada, facilitando assim o processo de anotação de seqüências.

Inicialmente foi previsto seqüenciar os ESTs a partir do terminal 3' e 5', porém verificamos posteriormente a presença de poli A de tamanhos variados em todos os clones. Isto prejudica o seqüenciamento em larga escala tanto pelo número muito grande de bases A repetidas como pela geração de artefatos durante a fase de extensão. Assim optamos pelo seqüenciamento a partir da extremidade 5'. Os clones mais interessantes foram seqüenciados no sentido contrário na tentativa de sobrepor as seqüências formando um "contig" contendo grande parte do cDNA.

6.5. – Categorização Segundo o Gene Ontology Consortium

A categorização das seqüências de *Rhynchosciara* fez uso dos arquivos disponibilizados pelo Gene Index de *Drosophila* 8.0¹⁰ através de licença acadêmica. Dentre estes arquivos estão uma lista em FASTA de consensos tentativos (TC), baseados em alinhamentos de ESTs e seqüências “full lenght” de *Drosophila*, e uma tabela de anotação contendo os TC com suas respectivas anotações e associados a um código GO (Gene Ontology). Utilizando o comando “formatdb”, fornecido junto com o pacote blast2, foi possível construir um banco local chamado TGI, que foi comparado às seqüências de *Rhynchosciara* por blast local segundo o algoritmo “tBlastx”, onde a tradução das seqüências de entrada são comparadas com as traduções das seqüências do banco de dados. A saída gerada em formato de texto foi processada pelo seguinte script escrito em AWK:

¹⁰ www.tigr.org

Test.Awk

```
BEGIN {nlind=10000000; nlinq=10000000}
/^Database/ {nlind=NR; data1 = $0}
(NR == nlind+1) {data2 = $0}
(NR == nlind+2) {data3 = $0}
/Query=/ {nlinq=NR}
(NR == nlinq+2) {print data1; print data2; print data3}
{print $0}
```

e pelo script em perl descrito abaixo, com a função de “parser” dos dados obtidos com as comparações. Este script faz uso da biblioteca BPlite:
test.perl

```
#!/usr/local/bin/perl
use Bio::Tools::BPlite;
use Bio::Tools::BPlite::HSP;
my $multiple_report = new Bio::Tools::BPlite(\*STDIN);
{
    $query = $multiple_report->query;
    $database = $multiple_report->database;
    $i = 0;
    while(my $sbjct = $multiple_report->nextSbjct) {
        $i = $i + 1;
        $name = $sbjct->name;
        if ($i > 1) {} else {
            while(my $hsp = $sbjct->nextHSP) {
                $score = $hsp->score;
                $bits = $hsp->bits;
                $percent = $hsp->percent;
                $P = $hsp->P;
                $match = $hsp->match;
                $positive = $hsp->positive;
                $length = $hsp->length;
                # $queryBegin = $hsp->qbegin;
                # $queryEnd = $hsp->qend;
                # $sbjctBegin = $hsp->sbjctbegin;
                # $sbjctEnd = $hsp->sbjctend;
                # $queryAlignment = $hsp->qalign;
                # $sbjctAlignment = $hsp->sbjctalign;
                # $alignmentString = $hsp->alignstring;
                # $queryGaps = $hsp->qgaps;
                # $sbjctGaps = $hsp->sbjctgaps;
                print "$query\t$name\t$score\t$P\t$length\t$match\t$positive\n";
            }
        }
    }
    last if ($multiple_report->_parseHeader == -1);
    print "redoing\n";
    redo;
}
```

Através das seguintes linhas de comando:

```
awk -f test.awk Desktop/testes/arquivo.ascii.br >arquivo.new  
./test.pl arquivo.new
```

foi possível gerar tabelas, que foram manualmente editadas, curadas e importadas para o MS Access para incorporar um banco de dados. As outras tabelas deste banco de dados foram geradas manualmente a partir das três ontologias disponíveis no Gene Ontology Consortium¹¹, onde foram desfeitos os indentamentos, que geralmente são interpretados pelos softwares do GO (Amigo), de forma que cada termo GO pudesse ser associado a um código GO.

Os scripts acima descritos também foram úteis nas análises dos blasts das seqüências de *Rhynchosciara* contra o GenBank (nr), que foram efetuados pelo robô Blastcl3, com saída em html, que foram convertidas em ascii com o uso do lynx (lynx -dump -nolist arquivo.html). As tabelas criadas foram curadas manualmente e analisadas com auxílio do software Excel.

6.6. – Produção de sub-bibliotecas genômicas

Os fagos isolados durante este projeto foram digeridos com *NotI* e seus insertos separados em gel de agarose 0,8%, que devidamente recortado permitiu recuperar o inserto por eletroeluição em membrana de diálise (“cutoff” de peso molecular 12000) em 0,5X TAE (0,8mM Tris-Acetato, 0,1mM EDTA).

Este inserto foi fragmentado aleatoriamente por nebulização, a partir de uma quantidade de 50 µg de DNA em 1 mL solução de glicerol (150 µL de DNA (50 µg), 600 µL de glicerol 50 %, 250 µL de água milliQ), que é então colocada em um nebulizador, e fragmentada através de três injeções de nitrogênio a uma pressão de 0,70 kgf/cm² por 45 segundos, sempre batendo em sua parede para forçar a solução para o fundo, o nebulizador é então centrifugado por 4 minutos a 1000 rpm.

Recupera-se aproximadamente 800 µL de solução com DNA o qual é precipitado adicionando 80 µL de acetato de sódio (3 M, pH 7), 2 mL de etanol e resfriando-se a -20 °C por 12 horas, centrifuga-se a 14000 rpm por 30 minutos a

¹¹ www.geneontology.org

temperatura ambiente, após descartar o sobrenadante lava-se o pellet com solução de etanol 80% e repete-se a centrifugação anterior.

O DNA obtido foi ressuspendido em 20 μ L de água milliQ e colocado em uma eletroforese em gel de agarose Low-melting 0,8 % juntamente com marcador de peso molecular.

Como esperado houve a formação de um “smear”, o gel de agarose é cortado em dois trechos, que segundo o marcador de peso molecular contém DNA na faixa de 0,5-1 e 1-2 kb, acondicionados em tubos diferentes adiciona-se 300 μ L de TE em cada um deles e coloca-se em um banho a 95°C por um minuto, afim de derreter a agarose, imediatamente adiciona-se 50 μ L de fenol saturado com TE em ambos os tubos e homogeneíza-se por dez minutos.

Centrifuga-se os tubos por cinco minutos a 12000 rpm a temperatura ambiente, cuidadosamente retira-se a fração superior de cada tubo, recuperando assim a solução contendo DNA e descartando o fenol e a agarose, repete-se este passo de fenolização novamente e precipita-se os 400 μ L de soluções de DNA com 40 μ L de acetato de sódio (ph 7, 3 M) e 1 mL de etanol e congelando a -20°C por 12 horas, após centrifugar a 14000 rpm por 30 minutos, lava-se o pellet com solução de etanol 80 % e centrifuga-se novamente nestas mesmas condições, ressuspende-se o DNA em 5 μ L de água milliQ.

Corre-se então 1 μ L desta solução obtida em gel de agarose 0,8% para confirmar a presença de DNA, em seguida prepara-se a reação de “blunting” (4 μ L de DNA, 1 μ L de Klenow (Amersham), 2 μ L de tampão B/K (Amersham), 1 μ L de PNK (Amersham), 12 μ L de água) totalizando 20 μ L de solução, que foi incubada a 37°C por 30 minutos.

Adiciona-se então 20 μ L de fenol saturado com TE e centrifuga-se 5 minutos a 12000 rpm, recuperar a fase superior e a passar por uma coluna de sephacryl S200 HR previamente empacotada em um “spin cartridge”, purificando assim o DNA de possíveis contaminações de fenol.

A partir destes fragmentos de DNA obtidos monta-se as reações de ligação com o vetor pUC18 ou pBluescript II KS+, previamente aberto e com extremidades “blunt” desfosforiladas (10 μ L de DNA, 4 μ L de Tampão 5x DNA Ligase Reaction (Gibco), 1 μ L T4 DNA Ligase (Gibco), 1 μ L de pUC18/pBluescript II KS+, 4 μ L de

água), esta reação foi mantida a 16°C por 17 horas em um aparelho de PCR modelo Thermocycler PTC-100 (MJ Research).

Em seguida transformou-se por choque térmico (42°C, 2 minutos, solução de CaCl₂ 50mM) duas doses de 40 µL de bactérias DH10B competentes preparadas no dia anterior, com 1µL de cada ligação. Adiciona-se então 1mL de SOB (2% triptona, 0,5% extrato de levedura, 10 mM de NaCl/MgSO₄/MgCl₂ e 20 mM de glicose), transfere-se para eppendorfs e incuba-se as bactérias a 37°C por 1 hora; por fim semea-se as transformantes em placas LB com ampicilina e X-gal/IPTG.

6.7. – Hibridizações: Southern e Northern

Tanto os Southern e Northern tiveram suas sondas marcadas por kit random primer (GIBCO/Invitrogen) ou com Nick Translation (GIBCO/Invitrogen).

Os Southern foram hibridizados em solução 6X SSC, 4X Denhardts, 1% SDS, 200µg/mL DNA de esperma de salmão; os filtros foram pré-hibridizados por uma hora à 65°C e hibridizados “overnight” à mesma temperatura. E foram lavados quatro vezes com solução de lavagem 2X SSC, 01% SDS e duas vezes com solução 0,2X SSC, 0,1% SDS.

Northern foram hibridizados em solução 5X SSPE, 1% SDS, 5X Denhardts, 200µg/mL DNA de esperma de salmão e 50% formamida, os filtros foram pré-hibridizados por uma hora a 42°C e hibridizados “overnight” à mesma temperatura. A lavagem foi feita quatro vezes com solução 2XSSC, 0,1%SDS e duas vezes com solução 0,1% SSC, 0,1%SDS.

RNAs extraídos de fases anteriores à amplificação (2º período e início de rede 3º período), e subseqüentes abrangendo a expressão dos pufes de DNA (B2 máximo expresso, B presente e 6º período) foram tratados com glyoxal-DMSO e submetidos a eletroforese em gel de agarose 1,2% sob recirculação do tampão. Após a transferência para membrana de nylon (Zetaprobe BioRad), os filtros foram hibridizados. Tanto os Southern como os Northern foram analisados após exposição em um phosphoimager Storm.

6.8. – Reações de PCR

Para seqüenciar algumas porções internas do cDNA ou do inserto genômico de RaDup, cujo conteúdo CG era alto ou para fechar falhas no seqüenciamento, foram desenhados os seguintes primers listados na tabela 12 e localizados no inserto genômico conforme diagramado na figura 25.

As condições de PCR de seqüenciamento são as mesmas já descritas, os primers DUP e DUP-RIGHT também foram utilizados para amplificar 1166 bp do gene RaDUP, com a finalidade de comprovar a escolha correta do clone genômico (da biblioteca de fagos), verificar presença de íntrons no gene e comprovar sua presença em outras espécies de *Rhynchosciara*, e Dup-atg e Dup-Race-1 foram aproveitados em RT-PCR para averiguar sua expressão em ovos e ovários.

Tabela 12 – Primers utilizados durante a caracterização de RaDup.

Dup-atg	GCGGATCCATGGCGCAACCATCGG
DUP-ex1-left	GGGCTTGATTCTGACCGTA
DUP-ex2-right	TAGGCTTCGGGGTACAAATG
DUP	CGCTTCAGTTGCCATACAAG
DUP RIGHT	TCGAACAGAAACGGAAGTCA
DUP-Left	AGCGACAACAACACTTGCAT
DUP-taa	CGCGAGCTCGATTTCAGCGTTTCTCTTC
DUP-5	CCTTCCACAGTAGTCAAATCATT
Race-2	GCTTGGTCGGAGTGCATAATC
Race-1	AACGACTTTGATGGCGATTC

A condição de PCR para amplificação está sumarizada na tabela 13.

Tabela 13 - condições de PCR para amplificação

Reação	Ciclos	Temperatura (°C)	Tempo
5 µL 10X tampão	1	94	45 s
3 µL MgCl ₂	35	94	30 s
1 µL DNA molde		50	45 s
4 µL mix primers (10pM)		72	1 min 30 s
2 µL dNTP (10mM)	1	72	7 min
0,5 µL Taq (Invitrogen)	1	4	∞
34,5 µL H ₂ O

Figura 25 – Mapa de primers que anelam em RaDup, mostrando suas orientações.

Em RT-PCR a transcrição reversa foi executada a partir de mRNA purificado com o kit OLIGOTEX (QUIAGEN) e tratado com DNAase (GIBCO), utilizando a transcriptase reversa Superscript II (GIBCO), inicialmente desnatura-se o mRNA e o primer (1µL de primer DUP_RIGHT (10 pM), 2µL de mRNA (~10ng), 1µL dNTP Mix 10mM) aquecendo-se a mistura a 65°C por 5 minutos, em seguida adiciona-se o tampão e a transcriptase reversa (4µL 5X First-Strand Buffer, 2µL 01,M DTT, 1µL Superscript II). A reação é então incubada por uma hora à 42°C, neste ponto adiciona-se 1µL de RNAase H e incuba-se por 20 minutos a 37°C. 1µL deste material é o suficiente para a reação de PCR subsequente.

6.9. – Varredura do banco genômico de *Rhynchosciara*

O primeiro passo para um “screening” radioativo do banco genômico de *Rhynchosciara*, que esta inserido em bacteriófago λ DASHII, foi a sua titulação.

Diluiu-se 16 μ L deste banco em 3 mL de λ buffer (50 mM $MgCl_2$ e 50 mM Tris), em seguida incubou-se por 15 minutos a 37°C quantidades diferentes desta solução de fagos (entre 20 e 50 μ L) com *E. coli* DL538 em tubos de ensaio estéreis, os conteúdos destes tubos foram então misturados a 4 mL de top agarose (0,65% agarose, 5 g/L NaCl, 5 g/L extrato de levedura, 10 g/L triptona) previamente fundida e plaqueado em placas de Petri com meio de cultura LB.

A partir do número de placas de lise da placa com 35 μ L de solução diluída de fago (1272 placas de lise), obtém-se o título do banco que é de 7.10^7 pfu/mL (já amplificado).

Prepara-se então 15 placas de Petri com placas de lise quase confluentes, a partir do inóculo de 40 μ L de solução diluída por placa, após o crescimento “overnight” cobre-se as placas com membranas HYBOND N+ (Amersham) por 2 minutos, seca-se as membranas e trata-se sequencialmente em solução 1 (NaCl 1,5M; NaOH 0,5M) por 5 minutos, solução 2 (NaCl 1,5M; tris 0,5M) por 5 minutos e em SSC 2X por 3 minutos, as membranas então são secas sobre papel de filtro e levadas a uma estufa a 80°C por 2 horas.

Deste ponto em diante as membranas são hibridizadas como um Southern comum já descrito acima.

Coleta-se as placas de lise positivas com o uso de uma ponteira de 1 mL de cabeça para baixo, lava-se o plug com 1 mL de λ buffer e então a partir desta suspensão prepara-se 4 placas de fago com placas de lise isoladas, e repete-se os passos referentes à transferência e hibridização.

Identificado(s) o(s) fago(s) positivo(s) retira-se os plugs e prepara-se uma suspensão de fago em 1 mL de λ buffer, plaqueia-se esta suspensão para gerar placas confluentes.

6.10. – Purificação de DNA de fagos em CsCl

Para a purificação de DNA de fagos em larga escala utilizou-se o método descrito em Santelli e Navarro-Cattapan (2000). Uma placa de fago confluenta é coberta com λ buffer (10mM Tris-HCl (pH 7.4), 10mM $MgCl_2$) por duas horas a 4°C com agitação periódica, adiciona-se então duas gotas de clorofórmio e agita-se suavemente, retira-se a suspensão da placa com uma pipeta.

Incuba-se 1 mL de suspensão de fago com 1 mL de bactéria DL538 competente (cultura overnight ressuspensa em $MgCl_2$ 10 mM) por 20 minutos a 37°C. Este inóculo é então transferido para um Erlenmeyer contendo 400 mL de LB, e mantido a 37°C sob forte agitação “overnight”.

Adiciona-se ao Erlenmeyer 2 mL de clorofórmio e agita-se vigorosamente, em seguida adiciona-se 20g de NaCl, após a solubilização do sal centrifuga-se o conteúdo por 10 minutos a 8000 rpm em um rotor GSA. Descarta-se o “pellet” e o sobrenadante é transferido para uma garrafa de centrífuga contendo 40g de PEG 8000, onde é incubado por 2 horas a 4°C.

Centrifuga-se novamente a 8000 rpm por 10 minutos descartando o sobrenadante.

Ressuspende-se o precipitado em 5 mL de λ buffer e transfere-se a solução para um tubo de centrífuga de 12 mL, adiciona-se 6,48 g de CsCl, ajusta-se o volume para 9 mL, cobre-se com óleo mineral e centrifuga-se em rotor SW41 a 35000 rpm por 24 horas.

A banda de DNA é então retirada com auxílio de uma seringa descartável de 3 mL e uma agulha de 22 gauge, perfurando-se a lateral do tubo cuidadosamente.

O DNA é então fenolizado, precipitado adicionando NaCl para uma concentração final de 0,2 M e 2,5 Volume de etanol absoluto gelado. O DNA precipita formando um véu que é então enrolado em um bastão de vidro e lavado em etanol 70%. Finalmente ressuspende-se o DNA em 1 mL de TE.

6.11. – Construção do plasmídeo pGEX-4T2-RaDup

A partir da clivagem SalI/NotI do cDNA RaDup clonado em pT7T3PAC foi construído um vetor de expressão com pGEX-4T2 (Amersham) onde a proteína é expressa fusionada com glutathione S transferase (GST), tal corte em nosso clone e no

vetor preserva o quadro de leitura da proteína RaDup, a unidade de transcrição está sob controle de PLac, o que permite indução com IPTG (Figura 26).

O plasmídeo pGEX-4T2-RaDup foi transformado em *E. coli* DH5 α e BL21, e em ambas as cepas foram gerados corpos de inclusão, foram testadas diferentes condições variando-se sistematicamente o tempo de indução, a temperatura e a concentração de IPTG (Figura 27). Ao final destas baterias de testes o que se visualizou sempre foi que a proteína de fusão estava na fração celular insolúvel, exceto quando a cepa BL21 é induzida com 0,1 mM de IPTG a 25°C por 1 hora, onde pode-se visualizar proteína de fusão na fração insolúvel e muito pouco na fração solúvel, a partir da qual foi purificada pouca proteína utilizando coluna de Glutathione-Agarose.

Figura 26 – Vetor pGEX-4T2-RaDup

Figura 27 – Exemplo de indução em BL21, Alb representa albumina bovina, utilizada como referência (66kDa).

6.12.-Indução de proteína de fusão em pGEX

A indução da proteína se deu em alíquotas de 50 mL para testes de expressão e 500 mL para purificação, em ambos os casos um pré-inóculo de 5 mL foi crescido “overnight”, inocula-se 100 µL de pré-inóculo em 50 mL de LB com ampicilina 100 µg/mL e 1 mL de pré-inóculo em 500 mL de LB com ampicilina 100 µg/mL, incuba-se a 37°C com agitação até atingir DO igual a 0,4-0,8, induz-se então a expressão da proteína adicionando IPTG para uma concentração final entre 0,1 mM e 10 mM, incuba-se então com agitação em temperatura entre 25 e 37 °C.

O meio é então centrifugado por 10 minutos a 5000g e o “pellet” é congelado a -20°C.

Efetua-se então um teste de solubilidade, o “pellet” de 50 mL de cultura é ressuspenso em 1,5 mL de solução de lise (Tris 10 mM (pH 8), NaCl 150 mM, EDTA 1mM, Lisozima 100 µg/mL), após 10 minutos eleva-se o volume para 6 mL com uma solução isotônica contendo inibidores de proteases (Tris 50mM (pH 7,5), MgCl₂ 10 mM, NaCl 100 mM, TLCK 1mM, PMSF 1mM). Em sonificador de TIP (MICROTIP/FISCHER) aplica-se ultrassom a 50% de potência em 5 pulsos de 5 segundos, com intervalos de 2 segundos.

Centrifuga-se 12000g por 10 minutos, ao sobrenadante é adicionado igual volume de 2X SDS loading buffer, e ao precipitado 50 µL de 1X SDS loading buffer, corre-se em gel desnaturante (SDS-PAGE) 8%.

6.13. – Purificação de proteína de fusão em coluna de glutathiona imobilizada em agarose

Ressuspede-se o pellet de 500 mL de cultura em 20 mL de PBS (137mM NaCl, 2,7mM KCl 10mM Na₂HPO₄, 2mM KH₂PO₄), adiciona-se lisozima para uma concentração final de 1mg/mL e 10 µL de PMSF 100mM, incuba-se por 30 minutos.

Adiciona-se 100 µL de Triton X-100 e trata-se em sonificador Microtip com potência 50% em 10 pulsos de 5 segundos.

Após centrifugação de 10 minutos a 10000g adiciona-se ao sobrenadante 500 µL de coluna previamente equilibrada em PBS (“overnight”), incuba-se por 1 hora a 4°C com agitação lenta, a solução é então centrifugada por 30 segundos a 500g e o sobrenadante é separado do precipitado guardando-se alíquotas para SDS-PAGE, lava-se a coluna 4 vezes com PBS (guardando-se alíquotas do PBS para SDS-PAGE); elui-se a coluna por 3 vezes com solução de eluição (glutathiona reduzida 10mM e 50mM Tris, pH8).

Dialisar o eluído “overnight” em 1000 volumes de PBS e correr alíquotas em SDS-PAGE.

Como o rendimento da purificação em coluna de GST foi muito pequeno optou-se por purificar os corpos de inclusão, onde a proteína se concentra, para preparar os inóculos da imunização dos camundongos.

6.14.-Purificação de corpos de inclusão

O pellet de 500 mL de cultura é pesado e para cada grama de bactéria ressuspende-se em 3 mL de Solução de Lise I (50 mM tris pH8, 1 mM EDTA, 100 mM NaCl), mais 100 µL de PMSF 100mM e 80 µL de lisozima 10 mg/mL, agita-se por 20 minutos a temperatura ambiente. Adiciona-se 4 mg de ácido deoxicólico por grama de bactéria, e incuba-se a 37°C por mais 30 minutos.

Aplica-se 10 pulsos de ultrassom de 5 segundos a 50% de potência em sonicador, centrifuga-se 10000g por 10 minutos, ao precipitado adiciona-se 9 volumes da Solução de Lise II (50 mM Tris pH8, 10 mM EDTA, 100 mM NaCl, 0,5% Triton X-100) e incuba-se a temperatura ambiente por 5 minutos.

Centrifuga-se por 15 minutos a 4°C a 12000g, separa-se alíquotas de sobrenadante e precipitado para SDS-PAGE, se a proteína ainda estiver presente no precipitado ressuspende-se o precipitado em 100 µL de Solução de Solubilização de Corpos de Inclusão I (50mM Tris pH8, 1mM EDTA, 100mM NaCl, 8M urea, 100mM PMSF), incuba-se por 1 hora a temperatura ambiente.

Adiciona-se 900 mL de Solução de Solubilização de Corpos de Inclusão II (50mM KH_2PO_4 pH10,7, 1mM EDTA, 50mM NaCl), incubar a temperatura ambiente por 30 minutos reajustando o pH para 10,7 com KOH 10M.

Ajustar o pH para 8 com HCl 12M e incubar por mais 30 minutos a temperatura ambiente, centrifuga-se então por 15 minutos 12000g e corre-se alíquotas do precipitado e do sobrenadante em SDS-PAGE (Figura 28).

Figura 28 – Alíquotas dos passos de purificação do corpo de inclusão. Sobrenadantes e precipitados são indicados por Sob e Ppt respectivamente, Alb indica albumina bovina (66kDa).

6.15. – Produção de Soros Anti-RaDup

A expressão da construção pGEX-4T2-RaDup produziu uma proteína do tamanho esperado, em forma de corpo de inclusão, esta proteína purificada por eletroforese foi utilizada para inocular camundongos (6 Balb-C e 7 Swiss) com a finalidade de produzir um anti-soro contra RaDup.

Em cada animal foram injetadas 5 doses semanais de aproximadamente 10 µg de proteína (intraperitoneal), contendo adjuvante de Freund completo na primeira e incompleto nas subseqüentes. Uma semana após o último inóculo os camundongos foram anestesiados com éter e o sangue foi coletado por punção venosa, os animais foram sacrificados o mais rapidamente possível. As amostras de sangue foram coaguladas a 37°C por 30 minutos, e centrifugados a 4000g por 10 minutos.

Os soros obtidos mostraram atividade contra GST e a própria proteína, os pré-inóculos não mostraram reação com extratos de glândulas salivares, de cabeças e de intestinos de *Rhynchosciara*.

6.16. – Marcação de Westerns

Os extratos de tecidos de *Rhynchosciara* foram desnaturados a 100°C por 5 minutos em tampão de corrida (50mM Tris-HCl, 100mM DTT, 2% SDS, 0,1% azul de bromofenol, 10 % glicerol) e após resfriamento conservados com PMSF.

As amostras foram aplicadas em géis de poliacrilamida 10%, e após eletroforese normal a 100mA em tampão 1X Tris-Glicina (0,1% SDS, 25mM Tris-HCl, 250 mM glicina) foram transferidas para membranas de PVDF, por 2 horas a 250mA em tampão de transferência (25mM Tris-HCl, 192mM glicina, 15% metanol).

As membranas foram bloqueadas “overnight” em albumina 1%, incubadas 2 horas em solução de anticorpo primário diluído em TBS (150mM NaCl, 10 mM Tris-HCl pH 7,5), e após 3 lavagens em TBS-TT (150mM NaCl, 10 mM Tris-HCl pH 7,5, 0,3% Tween, 0,5% Triton X-100), incubadas por mais 2 horas com o anticorpo secundário diluído em TBS. As diluições para os anticorpos utilizados foram: anti-RaDup – 1:3500; anti-DmDup – 1:5000; anti-DmMCM5 – 1:7500; anti-guinea-pig(DmDup) – 1:256; anti-goat(anti-guinea-pig) – 1:7500; anti-mouse – 1:7500. Anti-

mouse e anti-goat eram conjugados com peroxidase e todos os Westerns foram revelados com o kit ECL (Amersham).

6.17. – Hibridizações *in situ*

As glândulas salivares foram fixadas em etanol-Ác. acético (3:1) e esmagadas em Ác. acético 45%, as lamínulas foram retiradas por congelamento em gelo seco. As preparações foram lavadas em PBS (140mM NaCl, 2,7mM KCl, 1,5mM KH_2PO_4 , 6,5mM Na_2HPO_4), desnaturadas em NaOH 0,07M por 3-5 minutos e lavadas em etanol. As sondas foram marcadas com dUTP-digoxigenina por reação de random primer (Roche) e desnaturadas antes da hibridização. As hibridizações com digoxigenina foram reveladas com anti-digoxigenina conjugado a fluoresceína ou com NBT (Nitro Blue Tetrazolium) e os cromossomos marcados com iodeto de propídio, estas amostras foram analisadas em um microscópio confocal de varredura a laser LSM-510 (Zeiss), as amostras marcadas com NBT foram analisadas em microscópio de luz.

7. Referências Bibliográficas

AHMAD, K.; HENIKOFF, S.. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. **Mol. Cell.** v. 9, n. 6, p. 1191-200, 2002.

AMABIS, D.; AMABIS, J.M.; SIMOES, L.C.. Puffing activity in the salivary gland chromosomes of *Rhynchosciara* under experimental conditions. **Chromosoma**, v. 62, n. 2, p. 139-54, 1977.

AMABIS, J.M.; CABRAL, D.. RNA and DNA puffs in polytene chromosomes of *Rhynchosciara*: inhibition by extirpation of prothorax. **Science**, v. 14, n. 169, p. 692-4, 1970.

ARATA, Y.; FUJITA, M.; OHTANI, K.; KIJIMA, S.; KATO, J.. Cdk2-dependent and -independent Pathways in E2F-mediated S Phase Induction. **J. Biol. Chem.**, v. 275, n. 9, p. 6337-6345, 2000.

ARMELIN, H.A.; MENEGHINI, R.; LARA, F.J.S.. Patterns of ribonucleic acid synthesis in salivary glands of *Rhynchosciara angelae* larvae during development. **Genetics**, v. 61, p. 351-360, 1969.

BATEMAN, A.; BIRNEY, E.; DURBIN, R.; EDDY, S.R.; HOWE, K.L.; SONNHAMMER, E.L.. The Pfam protein families database. **Nucleic Acids Res.**, v. 28, n. 1, p. 263-6, 2000.

BELL, S.P.; MITCHELL, J.; LEBER, J.; KOBAYASHI, R.; STILLMAN, B.. The multidomain structure of Orc1p reveals similarity to regulators of DNA replication and transcriptional silencing. **Cell**, v. 83, n. 4, p. 563-8; 1995.

BELL, S.P.; STILLMAN, B.. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. **Nature**, v. 14, n. 357, p. 128-34, 1992.

BENTON, W.D.; DAVIS, R.W.. Screening lambda gt recombinant clones by hybridization to single plaques in situ. **Science**, v. 196, p. 180-182, 1977.

BIRNBOIM, H.C.; DOLY, J.. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. **Nucleic Acids Res.** v. 7, n. 6, p.1513-23, 1979.

BLOW, J.J.; LASKEY, R.A.. A role for the nuclear envelope in controlling DNA replication within the cell cycle. **Nature**, v. 332, n. 6164, p. 546-8, 1988.

BREUER, M.E.. Revision of the genus *Rhynchosciara rübsaamen* (Diptera, Sciaridae) in neotropical region. **Archivo Zool.** (S. Paulo), v. 17, p. 167-198, 1969.

BREUER, M.E.; PAVAN, C. Behavior of polytene chromosomes of *Rhynchosciara angelae* at different stages of larval development. **Chromosoma**, v. 7, p. 371-386, 1955.

BROEK, D.; BARTLETT, R.; CRAWFORD, K.; NURSE, P.. Involvement of p34cdc2 in establishing the dependency of S phase on mitosis. **Nature**, v. 351, n. 349, p. 388-93, 1991.

BURGE, C.; KARLIN, S.. Prediction of complete gene structures in human genomic DNA. **J. Mol. Biol.**, v. 268, p. 78-94; 1997.

BUSTIN, M.. Regulation of DNA-Dependent Activities by the Functional Motifs of the High-Mobility-Group Chromosomal Proteins. **Mol. Cel. Biol.**, v. 19, n. 8, p. 5237–5246, 1999.

CALVI, B.R.; LILLY, M.A.; SPRADLING, A.C. Cell cycle control of chorion gene amplification. **Genes & Development**. v. 12, n. 5, p. 734-744, 1998.

CHIN, M.T.; PELLACANI, A.; WANG, H.; LIN, S.S.; JAIN, M.K.; PERRELLA, M.A.; LEE, M.E.. Enhancement of Serum-response Factor-dependent Transcription and DNA Binding by the Architectural Transcription Factor HMG-I(Y). **J. Biol. Chem.**, v. 273, n. 16, p. 9755-60, 1998.

CLAYCOMB, J.M.; MACALPINE, D.M.; EVANS, J.G.; BELL, S.P.; ORR-WEAVER, T.L.. Visualization of replication initiation and elongation in *Drosophila*. **J. Cell Biol.**, v. 159, n. 2, p. 225-36, 2002.

COATES, P.J.; JAMIESON, D.J.; SMART, K.; PRESCOTT, A.R.; HALL, P.A.. The prohibitin family of mitochondrial proteins regulate replicative life span. **Curr. Biol.**, v. 7, n. 8, p. 607-10, 1997.

CORDEIRO, M.; MENEGHINI, R.. The rate of DNA replication in the polytene chromosomes of *Rhynchosciara angelae*. **J. Mol. Biol.**, v. 78, n. 2, p. 261-74, 1973.

DE CICCO D.V., SPRADLING A.C.. Localization of a cis-acting element responsible for the developmentally regulated amplification of *Drosophila* chorion genes. **Cell**, v. 38, n. 1, p. 45-54, 1984.

DEAK, P.; DONALDSON, M.; GLOVER, D.M.. Mutations in makos, a *Drosophila* gene encoding the Cdc27 subunit of the anaphase promoting complex, enhance centrosomal defects in polo and are suppressed by mutations in twins/aar, which encodes a regulatory subunit of PP2A. **J. Cell. Sci.**, v. 116, n. 20, p. 4147-58, 2003.

DELIDAKIS, C.; KAFATOS, F.C.. Amplification enhancers and replication origins in the autosomal chorion gene cluster of *Drosophila*. **EMBO J.**, v. 8, n. 3, p. 891-901, 1989.

DOUDNA, J.A.; RATH, V.L.. Structure and function of the eukaryotic ribosome: the next frontier. **Cell**, v. 109, n. 2, p. 153-6, 2002.

DREYFUS, A.; NONATO, E.; BREUER, M.E.; PAVAN, C.. Cromossomos politênicos em vários órgãos de *Rhynchosciara angelae*. **Rev. Brasil. Biol.**, v. 11, p. 439-450, 1951.

DURONIO, R.J.; BONNETTE, P.C.; O'FARRELL, P.H.. Mutations of the *Drosophila* dDP, dE2F, and cyclin E genes reveal distinct roles for the E2F-DP transcription factor and cyclin E during the G1-S transition. **Mol. Cell Biol.**, v. 18, n. 1, p. 141-51; 1998.

DURONIO, R.J.; BROOK, A.; DYSON, N.; O'FARRELL, P.H.. E2F induced S phase requires cyclin E. **Genes & Development**, v. 10, p. 2505-2513, 1996.

DURONIO, R.J.; O'FARRELL, P.H.. Developmental control of a G₁-S transcriptional program in *Drosophila*. **Development**, v. 120, p. 1503-1515, 1994.

DURONIO, V.; SCHEID, M.P.; ETTINGER, S.. Downstream signalling events regulated by phosphatidylinositol 3-kinase activity. **Cell Signal.**, v. 10, n. 4, p. 233-9, 1998.

EDGAR, B.A.; ORR-WEAVER, T.L.. Endoreplication cell cycles: more for less. **Cell**, v. 105, n. 3, p. 297-306, 2001.

EWING, B.; GREEN, P.. Base-calling of Automated Sequencer Traces using Phred, 2. Error Probabilities. **Genome Research**, v. 8, p. 186-194, 1998a.

EWING, B.; HILLIER, L.D.; WENDL, M.C.; GREEN, P.. Base-calling of Automated Sequencer Traces using Phred, 1. Accuracy Assessment. **Genome Research**, v. 8, n. 175-185, 1998b.

FICQ, A.; PAVAN, C.. Autoradiography of polytene chromosomes of *Rhynchosciara angelae* at different stages of larval development. **Nature**, v. 180, n. 4593, p. 983-4, 1957.

FOLLETTE, P.J.; DURONIO, R.J.; O'FARRELL, P.H.. Fluctuations in cyclin E levels are required for multiple rounds of endocycle S phase in *Drosophila*. **Curr. Biol.**, v. 8, n. 4, p. 235-8, 1998.

FRYDMAN, H.M.; CADAVID, E.O.;YOKOSAWA, J.; SILVA, F.H.; NAVARRO-CATTAPAN, L.D.; SANTELLI, R.V.; JACOBS, L. M.; GRAESSMANN, A.; GRAESSMANN, M.; STOCKER, A.J.; LARA, F.J.S.. Molecular characterization of the DNA puff C-8 gene of *Rhynchosciara americana*. **J. Mol. Biol.**, v. 233, p. 799-803, 1993.

GAO, M.; FRITZ, D.T.; FORD, L.P.; WILUSZ, J.. Interaction between a poly(A)-specific ribonuclease and the 5' cap influences mRNA deadenylation rates in vitro. **Mol. Cell**, v. 5, n. 3, p. 479-88, 2000.

GERBI, S.A.; STREZOSKA, Z.; WAGGENER, J.M.. Initiation of DNA replication in multicellular eukaryotes. **J. Struct. Biol.**, v. 140, n. 1, p. 17-30, 2002.

GERBI, S.A.; LIANG, C.; WU, N.; DIBARTOLOMEIS, S.M.; BIENZ-TADMOR, B.; SMITH, H.S.; URNOV, F.D.. DNA amplification in DNA puff II/9A of *Sciara coprophila*. **Cold. Spring Harbor Symp. Quant. Biol.**, v. 58, p. 487-494; 1993.

GILLESPIE, P.J.; LI, A.; BLOW, J.J.. Reconstitution of licensed replication origins on *Xenopus* sperm nuclei using purified proteins. **BMC Biochem.**, v. 2, n. 1, p. 15, 2001.

GLOVER, D.M.; ZAHA, A.; STOCKER, A.J.; SANTELLI, R.V.; PUEYO, M.T.; TOLEDO, S.M.; LARA, F.J.S.. Gene amplification in *Rhynchosciara* salivary gland chromosomes. **Proc. Natl. Acad. Sci. USA**, v. 79, p. 2947-2951, 1981.

GORDON, D; ABAJIAN, C; GREEN, P.. Consed: a graphical tool for sequence finishing. **Genome Res.**, v. 8, n. 3, p. 195-202, 1998.

GUEVARA, M.; BASILE, R. DNA and RNA puffs in *Rhynchosciara*. **Caryologia**, v. 26, n. 2, 1973.

Hammond, M.O.; Laird, C.D.. Chromosome structure and DNA replication in nurse and follicle cells of *Drosophila melanogaster*. **Chromosoma**, v. 91, p. 267-278, 1985.

HAYLES, J.; FISHER, D.; WOOLLARD, A.; NURSE, P.. Temporal order of S phase and mitosis in fission yeast is determined by the state of the p34cdc2-mitotic B cyclin complex. **Cell**, v. 78, n. 5, p. 813-22, 1994.

HARACSKA, L.; UNK, I.; JOHNSON, R.E.; PHILLIPS, B.B.; HURWITZ, J.; PRAKASH, L.; PRAKASH, S.. Stimulation of DNA synthesis activity of human DNA polymerase kappa by PCNA. **Mol. Cell Biol.**, v. 22, n. 3, p. 784-91, 2002.

HARRIS, J.K.; KELLEY, S.T.; SPIEGELMAN, G.B.; PACE, N.R. The genetic core of the universal ancestor. **Genome Res.**, v. 13, n. 3, p. 407-12, 2003.

HERSHKO, A.; CIECHANOVER, A.. The ubiquitin system. **Annu. Rev. Biochem.**, v. 67, p. 425-479, 1998.

HODGSON, B.; LI, A.; TADA, S.; BLOW, J.. Geminin Becomes Activated as an Inhibitor of Cdt1/RLF-B Following Nuclear Import. **Curr. Biol.**, v. 12, p. 678-683, 2002.

HOFMANN, J.F.; BEACH, D.. cdt1 is an essential target of the Cdc10/Sct1 transcription factor: requirement for DNA replication and inhibition of mitosis. **EMBO J.**, v. 13, n. 2, p. 425-34, 1994.

HOOFF, C.V.; GORIS, J.. Phosphatases in apoptosis: to be or not to be, PP2A is in the heart of the question. **Biochim. Biophys. Acta.**, v. 1640, n. 2, p. 97-104, 2003.

HUNTER, T.; POON, R.Y.C.. cdc37: a protein kinase chaperone?. **Trends in Cell Biology**, v. 7, p. 157-161, 1997.

JORDAN, P.; KARESS, R.. Myosin light chain-activating phosphorylation sites are required for oogenesis in *Drosophila*. **J. Cell Biol.**, v. 29, n. 139, p. 1805-19, 1997.

ITZHAKI, J.E.; GILBERT, C.S.; PORTER, A.C.. Construction by gene targeting in human cells of a 'conditional' CDC2 mutant that rereplicates its DNA. **Nat. Genet.**, v. 15, n. 3, p. 258-65, 1997.

KASHEVSKY, H.; WALLACE, J.A.; REED, B.H.; LAI, C.; HAYASHI-HAGIHARA, A.; ORR-WEAVER, T.L.. The anaphase promoting complex/cyclosome is required during development for modified cell cycles. **Proc. Natl. Acad. Sci. U S A.**, v. 99, n. 17, p. 11217-22, 2002.

KATO, Y; NAIR, K.K.; DYER, K.A.; RIDDIFORD, L.M.. Changes in ploidy level of epidermal cells during last larval instar of the tobacco hornworm, *Manduca sexta*. **Development**, v. 99, p. 137-143, 1987.

KITAGAWA, R.; ROSE, A.M.. Components of the spindle-assembly checkpoint are essential in *Caenorhabditis elegans*. **Nat. Cell Biol.**, v. 1, n. 8, p. 514-21; 1999.

KNOBLICH, J.A.; SAUER, K.; JONES, L.; RICHARDSON, H.; SAINT, R.; LEHNER, C.F.. Cyclin E controls S phase progression and its down-regulation during *Drosophila* embryogenesis is required for the arrest of cell proliferation. **Cell**, v. 77, p. 107-120, 1994.

KRUDE, T.; KELLER, C.. Chromatin assembly during S phase: contributions from histone deposition, DNA replication and the cell division cycle. **Cell. Mol. Life Sci.**, v. 58, n. 5-6, p. 665-72, 2001.

LACHANCE, P.E.; MIRON, M.; RAUGHT, B.; SONENBERG, N.; LASKO, P.. Phosphorylation of eukaryotic translation initiation factor 4E is critical for growth. **Mol. Cell Biol.**, v. 22, n. 6, p. 1656-63, 2002.

LANDAIS, I.; OGLIASTRO, M.; MITA, K.; NOHATA, J.; LOPEZ-FERBER, M.; DUONOR-CERUTTI, M.; SHIMADA, T.; FOURNIER, P.; DEVAUCHELLE, G.. Annotation pattern of ESTs from *Spodoptera frugiperda* Sf9 cells and analysis of the ribosomal protein genes reveal insect-specific features and unexpectedly low codon usage bias. **Bioinformatics**, v. 19, n. 18, p. 2343-50, 2003.

LARA, F.J.S.; TAMAKI, H.; PAVAN, C.. Laboratory culture of *Rhynchosciara angelae*. **Amer. Natur.**, v. 99, p. 189-191, 1965.

LEI, M.; TYE, B.K.. Initiating DNA synthesis: from recruiting to activating the MCM complex. **J. Cell Sci.**, v. 114, n. 8, p. 1447-1454, 2002.

Li, C.J.; DePAMPHILIS, M.L.. Mammalian Orc1 protein is selectively released from chromatin and ubiquitinated during the S-to-M transition in the cell division cycle. **Mol. Cell Biol.**, v. 22, n. 1, p. 105-16, 2002.

LI, X.; ZHAO, Q.; LIAO, R.; SUN, P.; WU, X.. The SCF(Skp2) ubiquitin ligase complex interacts with the human replication licensing factor Cdt1 and regulates Cdt1 degradation. **J. Biol. Chem.**, v. 278, n. 33, p. 30854-8, 2003.

LIANG, C.; GERBI, S.A.. Analysis of an origin of DNA amplification in *Sciara coprophila* by a novel three-dimensional gel method. **Mol. Cell Biol.**, v. 14, n. 2, p. 1520-9, 1994.

LIANG, C.; SPITZER, J.D.; SMITH, H.S.; GERBI, S.A.. Replication initiates at a confined region during DNA amplification in *Sciara* DNA puff II/9A. **Genes Dev.**, v. 7, n. 6, p. 1072-84, 1993.

LILLY, M.A.; SPRADLING, A.C.. The *Drosophila* endocycle is controlled by cyclin E and lacks a checkpoint ensuring S-phase completion. **Genes & Development**, v. 10, p. 2514-2526, 1996.

LOEBEL, D.; HUIKESHOVEN, H.; COTTERILL, S.. Localization of the DmCdc45 DNA replication factor in the mitotic cycle and during chorion gene amplification. **Nucleic Acids Res.**, v. 28, n. 20, p. 3897-903, 2000.

LYGEROU, Z.; NURSE, P.. License Withheld - Geminin blocks DNA replication. **Science**, v. 290, p. 2271-2273, 2000.

MA, T.; VAN TINE, B.A.V.; WEI, Y.; GARRET, M.D.; NELSON, D.; ADAMS, P.D.; WANG, J.; QIN, J.; CHOW, L.T.; HARPER, J.W.. Cell cycle regulated phosphorylation of p220 NPAT by cyclinE/Cdk2 in Cajal bodies promotes histone gene transcription. **Genes & Dev.**, v. 14, p. 2298-2313, 2000.

MACHADO-SANTELLI, G.M.; BASILE, R.. DNA replication and DNA puffs in salivary chromosomes of *Rhynchosciara*. **Ciência e Cultura**, v. 27, n. 2, 1975.

MACHADO-SANTELLI, G.M.; BASILE, R.. Studies of DNA puff development in *Rhynchosciara*. **Genetics**, v. 74, p. 168, 1973.

MAINE, G.T.; SINHA, P.; TYE, B.K.. Mutants of *S. cerevisiae* defective in the maintenance of minichromosomes. **Genetics**, v. 106, n. 3, p. 365-85, 1984.

MAIORANO, D.; LEMAITRE, J.M.; MECHALI, M.. Stepwise regulated chromatin assembly of MCM2-7 proteins. **J. Biol. Chem.**, v. 275, n. 12, p. 8426-31, 2000.

MAIORANO, D.; MOREAU, J.; MECHALI, M.. XCDT1 is required for the assembly of pre-replicative complexes in *Xenopus laevis*. **Nature**, v. 404, n. 6778, p. 622-5, 2000.

MANLEY, J.L.; TAKAGAKI, Y.. The end of the message--another link between yeast and mammals. **Science**, n. 274, p. 1481-1482, 1996.

MENEGHINI, R.; CORDEIRO, M.. DNA replication in polytene chromosomes of *Rhynchosciara angela*. **Cell. Diff.**, v. 1, n. 167, 1972.

MIHAYLOV, I.S.; KONDO, T.; JONES, L.; RYZHIKOV, S.; TANAKA, J.; ZHENG, J.; HIGA L.A.; MINAMINO, N.; COOLEY, L.; ZHANG, H.. Control of DNA replication and chromosome ploidy by geminin and cyclin A. **Mol. Cell Biol.**, v. 22, n. 6, p. 1868-80, 2002.

MONESI, N.; FERNANDEZ, M.A.; FONTES, A.M.; BASSO, L.R.; NAKANISHI, Y.; BARON, B.; BUTTIN, G.; PAÇÓ-LARSON, M.L.. Molecular characterization of 18 Kb segment of DNA puff C4 of *Bradysia hygida* (Diptera, Sciaridae). **Chromosoma**, v. 103, p. 715-724, 1995.

MORGANTE, J.S.; GUEVARA, M.. Caracterização dos estádios larvais de *Rhynchosciara angelae*. **Ciência e Cultura**, v. 21, p. 288-289, 1969.

MORISAKI, T.; HIROTA, T.; IIDA, S.; MARUMOTO, T.; HARA, T.; NISHIYAMA, Y.; KAWASUZI, M.; HIRAOKA, T.; MIMORI, T.; ARAKI, N.; IZAWA, I.; INAGAKI, M.; SAYA H.. WARTS tumor suppressor is phosphorylated by Cdc2/cyclin B at spindle poles during mitosis. **FEBS Lett.**, v. 529, n. 2-3, p. 319-24, 2002.

NER, S.S.; BLANK, T.; PEREZ-PARALLE, M.L.; GRIGLIATTI, T.A.; BECKER, P.B.; TRAVERS, A.A.. HMG-D and histone H1 interplay during chromatin assembly and early embryogenesis. **J. Biol. Chem.**, v. 276, n. 40, p. 37569-76, 2001.

NER, S.S.; TRAVERS, A.A.. HMG-D, the *Drosophila melanogaster* homologue of HMG 1 protein, is associated with early embryonic chromatin in the absence of histone H1. **EMBO J.**, v. 13, n. 8, p. 1817-22, 1994.

NISHITANI, H.; LYGEROU Z.. Control of DNA replication licensing in a cell cycle. **Genes Cells.**, v. 7, n. 6, p. 523-34, 2002.

NISHITANI, H.; TARAVIRAS, S.; LYGEROU, Z.; NISHIMOTO, T.. The Human Licensing Factor for DNA Replication Cdt1 Accumulates in G1 and Is Destabilized after Initiation of S-Phase. **J. Biol. Chem.**, v. 276, n. 48, p. 44905-44911, 2001.

NONATO, E.; PAVAN, C.. A new species of *Rhynchosciara* Rubsaamen, 1894 (Diptera, Mycetophilidae). **Rev. Brasil. Biol.**, v.11, p. 435-437, 1951.

NOWAK, S.J.; PAI, C.Y.; CORCES, V.G.. Protein phosphatase 2A activity affects histone H3 phosphorylation and transcription in *Drosophila melanogaster*. **Mol. Cell Biol.**, v. 23, n. 17, p. 6129-38, 2003.

ORR-WEAVER, T.L.; JOHNSTON, C.G.; SPRADLING, A.C.. The role of ACE3 in *Drosophila* chorion gene amplification. **EMBO J.**, v. 8, n. 13, p. 4153-62, 1989.

OKAMOTO, Y.; OZAKI, T.; MIYAZAKI, K.; AOYAMA, M.; MIYAZAKI, M.; NAKAGAWARA, A.. UbcH10 is the cancer-related E2 ubiquitin-conjugating enzyme. **Cancer Res.**, v. 63, n. 14, p. 4167-73, 2003.

OKUNO, Y.; MCNAIRN, A.J.; DEN ELZEN, N.; PINES, J.; GILBERT, D.M.. Stability, chromatin association and functional activity of mammalian pre-replication complex proteins during the cell cycle. **EMBO J.**, v. 20, n. 15, p. 4263-4277, 2001.

PENALVA, L.O.F.; YOKOZAWA, J.; STOCKER, A.J.; SOARES, M.A.M.; GRAESSMANN, M.; ORLANDO, T.C.; WINTER, C.E.; BOTELLA, L.M.; GRAESSMANN, A.; LARA, F.J.S.. Molecular characterization of the C-3 DNA puff gene of *Rhinosciara Americana*. **Gene**, v. 193, n. 2, p. 163-172, 1997.

RAO, P.N.; JOHNSON, R.T.. Mammalian cell fusion: studies on the regulation of DNA synthesis and mitosis. **Nature**, v. 225, n. 228, p. 159-64, 1970.

RIABOWOL, K.; DRAETTA, G.; BRIZUELA, L.; VANDRE, D.; BEACH, D.. The cdc2 kinase is a nuclear protein that is essential for mitosis in mammalian cells. **Cell.**, v. 57, n. 3, 393-401, 1989.

RICHARDSON, H.E.; O'KEEFE, L.V.; REED, S.I.; SAINT, R.. A *Drosophila* G₁-specific cyclin E homolog exhibits different modes of expression during embryogenesis. **Development**, v. 119, p. 673-690, 1993.

RIO, D.. Half-pint: alternative splicing in the *Drosophila* ovary. **Mol. Cell.**, v. 9, n. 3, p. 456-7, 2002.

ROYZMAN, I.; AUSTIN, R.J.; BOSCO, G.; BELL, S.P.; ORR-WEAVER, T.L.. ORC localization in *Drosophila* follicle cells and the effects of mutations in dE2F and dDP. **Genes Dev.**, v. 13, n. 7, p. 827-40, 1999.

RUBIN, G.M.. The *Drosophila* genome project: a progress report **TIG**, v. 14, n. 9, p. 340-343, 1998.

RUDKIN, G.T.; CORLETTE, S.L.. Disproportionate synthesis of DNA in polytene chromosome region. **Proc. Nat. Acad. Sci.**, v. 43, p. 964-968, 1957.

SANTELLI, R.V.; NAVARRO-CATTAPAN, L.D.. A simplified CsCl protocol for lambda DNA purification: no enzymatic treatment/one phenol extraction. **Genetics and Molecular Biology**. v. 23, n. 1, p. 65-66, 2000.

SANTELLI, R.V.; SIVIERO, F.; MACHADO-SANTELLI, G.M.; LARA, F.J.S.; STOCKER, A.J.. Molecular characterization of the B-2 DNA puff gene of *Rhynchosciara americana*. **Chromosoma**. no prelo.

SANTELLI, R.V.; MACHADO-SANTELLI, G.M.; PUEYO, M.T.; NAVARRO-CATTAPAN, L.D.; LARA, F.J.S.. Replication and transcription in the course of DNA amplification of the C3 and C8 DNA puffs of *Rhynchosciara americana*. **Mech. Dev.**, v. 36, p. 59-66; 1991.

SCHRAMKE, V.; LUCIANO, P.; BREVET, V.; GUILLOT, S.; CORDA, Y.; LONGHESE, M.P.; GILSON, E.; GELI, V.. RPA regulates telomerase action by providing Est1p access to chromosome ends. **Nat Genet.** v. 36, n. 1, p. 46-54, 2004.

SIMON, C.R.. A morte celular programada na glândula salivar de *Bradysia hygidae* (Díptera, Sciaridae). 1999. Tese (Doutorado em Ciências). Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo.

SIRGRIST, S.J.; LEHNER, C.F.. *Drosophila* fizzy-related down-regulates mitotic cyclins and is required for cell proliferation arrest and entry into endocycles. **Cell**, v. 90, n. 4, p. 671-81, 1997.

SMITH, A.V.; ORR-WEAVER, T.L.. The regulation of the cell cycle during *Drosophila* embryogenesis: The transition to polyteny. **Development**, v. 112, p. 997-1008, 1991.

SOARES, M.B.; BONALDO, M.F.; JELENC, P.; SU, L.; LAWTON, L.; EFSTRATIADIS, A.. Construction and characterization of a normalized cDNA library. **Proc. Natl. Acad. Sci. U S A.**, v. 91, n. 20, p. 9228-32, 1994.

SOARES, M.A.; MONESI, N.; BASSO, L.R. JR; STOCKER, A.J.; PACO-LARSON, M.L.; LARA, F.J.. Analysis of the amplification and transcription of the C3-22 gene of *Rhynchosciara americana* (Diptera: Sciaridae) in transgenic lines of *Drosophila melanogaster*. **Chromosoma**, v. 112, n. 3, p. 144-51, 2003.

SPRADLING, A.C.. ORC binding, gene amplification, and the nature of metazoans replication origins. **Genes & Development**, v. 13, p. 2619-2623, 1999.

SPRADLING, A.C.. The organization and amplification of two chromosomal domains containing *Drosophila* chorion genes. **Cell**, v. 27, p. 193-201, 1981.

SPRADLING, A.C.; LEYS, E.. Slow replication fork movement during *Drosophila* chorion gene amplification. Eukariotic DNA Replication, vol 6. In: KELLY, T.; STILLMAN, B. (Org.). **Cancer Cells**. New York: Cold Spring Harbor Press; 1988.

STOCKER, A.J.; GORAB, E.. The chromosomes of *Rhynchosciara baschanti* (Diptera: Sciaridae): molecular cytogenetic comparisons with taxa in the americana-like group. **Genome**, v. 43, n. 5, p. 786-95, 2000.

STOCKER, A.J.; GORAB, E.; AMABIS, J.M.; LARA, F.J.S.. A molecular cytogenetic comparison between *Rhynchosciara americana* and *Rhynchosciara hollaenderi* (Diptera:Sciaridae). **Genome**, v. 36, p. 831-43, 1993.

STOCKER, A.J.; TROYANO-PUEYO, M.; PEREIRA, S.D.; LARA, F.J.S.. Ecdysteroid titers and changes in chromosomal activity in the salivary glands of *Rhynchosciara Americana*. **Chromosoma**, v. 90, p. 26-38, 1984.

SULLIVAN, E.; SANTIAGO, C.; PARKER, E.D.; DOMINSKI, Z.; YANG, X.; LANZOTTI, D.J.; INGLEDUE, T.; MARZLUFF, W.F.; DURONIO, R.J.. *Drosophila* stem loop binding protein coordinates accumulation of mature histone mRNA with cell cycle progression. **Genes & Dev.**, v. 15, p. 173-187, 2001.

SUN, W.H.; COLEMAN, T.R.; DePAMPHILIS, M.L.. Cell cycle-dependent regulation of the association between origin recognition proteins and somatic cell chromatin. **EMBO J.**, v. 21, p. 1437, 2002.

SWIFT, H.. Nucleic acids and cell morphology in dipteran salivary glands. In: Allen, JM (Org.), **The molecular control of cellular activity**, 1962. p. 75-125.

TERRA, W.R.; DE BIANCHI, A.G.; GAMBARINI, A.G.; LARA, F.J.S.. Haemolymph amino acids and related compounds during cocoon production by the larvae of the fly *Rhynchosciara americana*. **J. Insect Physiol.**, v. 19, p. 2097-2106, 1973.

THE GENE ONTOLOGY CONSORTIUM. Creating the gene ontology resource: design and implementation. **Genome Research**, v. 11, p. 1425-1433, 2001.

THE GENE ONTOLOGY CONSORTIUM. Gene Ontology: tool for the unification of biology. **Nature Genetics**, v. 25, p. 25-29, 2000.

THOMPSON, J.D.; GIBSON, T.J.; PLEWNIAK, F.; JEANMOUGIN, F. AND HIGGINS, D.G.. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. **Nucleic Acids Research**, v. 25, p. 4876-4882, 1997.

THIERINGER, H.A.; SINGH, K.; TRIVEDI, H.; INOUE, M.. Identification and developmental characterization of a novel Y-box protein from *Drosophila*. **Nucleic Acid Research**, v. 25, n. 23, p. 4764-70, 1997.

USUI, T.; YOSHIDA, M.; ABE, K.; OSADA, H.; ISONO, K.; BEPPU, T.. Uncoupled cell cycle without mitosis induced by a protein kinase inhibitor, K-252a. **J. Cell. Biol.**, v. 115, n. 5, p. 1275-82, 1991.

VAN BUSKIRK, C.; SCHÜPBACH, T.. Half pint regulates alternative splice site selection in *Drosophila*. **Dev. Cell.**, v. 2, n. 3, p. 343-53, 2002.

VAS, A.; MOK, W.; LEATHERWOOD, J.. Control of DNA rereplication via Cdc2 phosphorylation sites in the origin recognition complex. **Mol. Cell. Biol.**, v. 21, n. 17, p. 5767-77, 2001.

WALLENFANG, M.R.; SEYDOUX, G.. cdk-7 Is required for mRNA transcription and cell cycle progression in *Caenorhabditis elegans* embryos. **Proc. Natl. Acad. Sci U S A.**, v. 99, n. 8, p. 5527-32, 2002.

WANG, S.; NATH, N.; FUSARO, G.; CHELLAPPAN, S. Rb and prohibitin target distinct regions of E2F1 for repression and respond to different upstream signals. **Mol. Cell. Biol.**, v. 19, n. 11, p. 7447-60, 1999.

WARNER, J.R.; NIERRAS, C.R. Trapping human ribosomal protein genes. **Genome Res.**, v. 8, n. 5, p. 419-21, 1998.

WHEELER, D.L.; CHAPPEY, C.; LASH, A.E.; LEIPE, D.D.; MADDEN, T.L.; SCHULER, G.D.; TATUSOVA, T.A.; RAPP, B.A. Database resources of the National Center for Biotechnology Information. **Nucleic. Acids. Res.**, v. 28, n. 1, p. 10-4, 2000.

WHITE, K.P.; RIFKIN, S.A.; HURBAN, P.; HOGNESS, D.S.. Microarray analysis of *Drosophila* development during metamorphosis. **Science**, v. 286, n. 5547, p. 2179-2184, 1999.

WHITFIELD C.W., BAND M.R., BONALDO M.F., KUMAR C.G., LIU L., PARDINAS J.R., ROBERTSON, H.M.; SOARES, M.B.; ROBINSON, G.E.. Annotated Expressed Sequence Tags and cDNA Microarrays for Studies of Brain and Behavior in the Honey Bee. **Gen. Res.**, v. 12, p. 555-566, 2002.

WHITTAKER, A.J.; ROYZMAN, I.; ORR-WEAVER, T.. *Drosophila double parked*: a conserved, essential replication protein that colocalizes with the origin recognition complex and links DNA replication with mitosis and the down-regulation of S phase transcripts. **Genes & Dev.**, v. 14, p. 1765-1766, 2000.

WINSTON, R.L.; MILLAR, D.P.; GOTTESFELD, J.M.; KENT, S.B.H.. Characterization of the DNA Binding Properties of the bHLH Domain of Deadpan to Single and Tandem Sites. **Biochemistry**, v. 38, p. 5138-5146, 1999.

WINTER, C.E.; BIANCHI, A.G.; TERRA, W.R.; LARA, F.J.S.. The giant DNA puffs of *Rhynchosciara americana* code for polypeptides of the salivary gland secretion. **J. Insect. Physiol.**, v. 23, p. 1455-59, 1977.

WIEDEMANN, C. R.W.. *Aussereuropäische zweiflügelige Insekten* 1: XXXII + p. 608, 7pls., Hamm; 1828.

WIEDEMANN, C. R.W.. *Diptera exotica*, p. 244, 1 fig., 2 pls. Kiliae [=Kiel], 1821.

WOHLSCHLEGEL, J.A.; DWYER, B.T.; DHAR, S.K.; CVETIC, C.; WALTER, J.C.; DUTTA, A.. Inhibition of eukaryotic DNA replication by Geminin binding to Cdt1. **Science**, v. 290, p. 2309-2312, 2000.

WOLD, M.S.. REPLICATION PROTEIN A: A Heterotrimeric, Single-Stranded DNA-Binding Protein Required for Eukaryotic DNA Metabolism. **Annu. Rev. Biochem.**, v. 66, p. 61–92, 1997.

YANAGI, K.; MIZUNO, T.; YOU, Z.; HANAOKA, F.. Mouse geminin inhibits not only Cdt1-MCM6 interactions but also a novel intrinsic Cdt1 DNA binding activity. **J. Biol. Chem.**, v. 277, n. 43, p. 40871-80, 2002.

YANOW, S.K.; LYGEROU, Z.; NURSE, P.. Expression of Cdc18/Cdc6 and Cdt1 during G2 phase induces initiation of DNA replication. **EMBO J.**, v. 20, n. 17, p. 4648-4656, 2001.

YOKOSAWA, J.; SOARES, M.A.; DIJKWEL, P.A.; STOCKER, A.J.; HAMLIN, J.L.; LARA, F.J.. DNA replication during amplification of the C3 puff of *Rhynchosciara americana* initiates at multiple sites in a 6 kb region. **Chromosoma**, v. 108, n. 5, p. 291-301, 1999.

ZHAO, J.; KENNEDY, Y.B.K.; LAWRENCE, B.D.; BARBIE, D.A.; MATERA, A.G.; FLETCHER J.; HARLOW, E.. NPAT links cyclin E-Cdk2 to the regulation of replication-dependent histone gene transcription. **Genes Dev.**, v. 14, p. 18, p. 2283-2297, 2000.

ZOU, L.; STILLMAN, B.. Formation of a preinitiation complex by S-phase cyclin CDK-dependent loading of Cdc45p onto chromatin. **Science**, v. 280, n. 5363, p. 593-6, 1998.

ZUR, A.; BRANDEIS, M.. Securin degradation is mediated by fzy and fzr, and is required for complete chromatid separation but not for cytokinesis. **EMBO J.**, v. 20, n. 4, p. 792-801, 2001.

ZUR A.; BRANDEIS M.. Timing of APC/C substrate degradation is determined by fzy/fzr specificity of destruction boxes. **EMBO J.**, v. 21, n. 17, p. 4500-10, 2002.

8. Súmula Curricular

Dados Pessoais

Nome: Fábio Siviero

Data de Nascimento: 29/03/77

Nacionalidade: Brasileira

N° USP: 0787968

RG:27.536.280-2

Formação 2º Grau

EEPSGES Barão de Piratininga

Início: 01/92

Conclusão: 12/94

Formação Superior

Instituto de Química da Universidade de São Paulo

Curso: Química – Especialização Biotecnológica

Início: 01/95

Conclusão: 12/98

Participações em Congressos

- XXIX Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular – Caxambu - 2000. Com o trabalho “EST Bank from *Rhynchosciara americana* salivary gland cells”.
- XXX Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular – Caxambu - 2001. Com os trabalhos “The complete sequence of the histone gene repeat of *Rhynchosciara*” e “Characterization of the messenger of the B-2 DNA puff of *Rhynchosciara*”.
- XXXII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular – Caxambu – 2003. Com o trabalho “Mobile Elements in *Rhynchosciara*”.
- I Brazilian International Genome Conference – Angra dos Reis – 2001. Com o trabalho “EST Bank from *Rhynchosciara americana* salivary gland cells”.
- I International Meeting of Latin American Society of Developmental Biology – Valle Nevado, Chile – 2003. Com o trabalho “RaDup: a *Rhynchosciara* Homologue of the Replication Origin Licensing Factor DmDup”.

Publicações

1. Possidonio, S.; Siviero, F. and El Seoud, O.A.; “Kinetics of the pH-independent hydrolysis of 4-nitrophenylchloroformate in aqueous micellar solutions: effects of the charge and structure of the surfactant”; *Journal of Physical Organic Chemistry*; 12; 325-332; 1999.
2. Santelli, R. V. and Siviero, F ; “A search for homologues of plant photoreceptor genes and their signaling partners in the sugarcane expressed sequence tag (Sucest) database”; *Genetics and Molecular Biology*, 24 (1-4), 49-53; 2001.
3. Vettore AL, Da Silva FR, Kemper EL, Souza GM, Da Silva AM, Ferro MI, Henrique-Silva F, Giglioti EA, Lemos MV, Coutinho LL, Nobrega MP, Carrer H, Franca SC, Bacci M Jr, Goldman MH, Gomes SL, Nunes LR, Camargo LE, Siqueira WJ, Van Sluys MA, Thiemann OH, Kuramae EE, Santelli RV, Marino CL, Targon ML, Ferro JA, Silveira HC, Marini DC, Lemos EG, Monteiro-Vitorello CB, Tambor JH, Carraro DM, Roberto PG, Martins VG, Goldman GH, De Oliveira RC, Truffi D, Colombo CA, Rossi M, De Araujo PG, Sculaccio SA, Angella A, Lima MM, De Rosa VE Jr, Siviero F, Coscrato VE, Machado MA, Grivet L, Di Mauro SM, Nobrega FG, Menck CF, Braga MD, Telles GP, Cara FA, Pedrosa G, Meidanis J, Arruda P.; “Analysis and Functional Annotation of an Expressed Sequence Tag Collection for Tropical Crop Sugarcane”; *Genome Res.* 2003 Nov 12 (em impressão – versão online disponível).
4. Santelli RV, Siviero F, Machado-Santelli GM, Lara FJS and Stocker AJ; “Molecular characterization of the B-2 DNA puff gene of *Rhynchosciara americana*”; *Chromosoma*; em impressão.

Outras informações

Recebi o Prêmio Lavoisier, concedido pelo CRQ-IV Região, pelo desempenho no curso de Química – Opção Biotecnológica entre 1995 e 1998.

Participei do Projeto Genoma da Cana-de-Açúcar tanto na fase de seqüenciamento como na fase de “Data Mining”.

Em janeiro de 2003 fui selecionado pela “Latin American Society of Developmental Biology” para participar do “Third Practical Course of Developmental Biology”, realizado na Faculdade de Ciências da Universidade do Chile no laboratório do Dr. Roberto Mayor. Curso ministrado pelos Drs. Roberto Mayor, Kathleen E. Whitlock, John Ewer, Miguel Allende Connelly, Ruby Valdivia e Motoe Kato, onde foram enfocados aspectos experimentais da biologia do desenvolvimento de *Drosophila*, *Xenopus*, “Zebrafish” e de camundongos.