



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE BIOCÊNCIAS
BACHARELADO EM CIÊNCIAS BIOLÓGICAS

Trabalho de Conclusão de Curso

**BUSCANDO OS RASTROS DE EVOLUÇÃO ADAPTATIVA NO
GENOMA DE *ORDOSPORA COLLIGATA***

Jennifer Stein de Lima

Orientadora: Profa. Dra. Karen Luisa Haag

Porto Alegre, junho de 2015.

Buscando os rastros de evolução adaptativa no genoma de *Ordospora colligata*

Orientadora: Profa. Dra. Karen Luisa Haag

Banca Examinadora:

Prof.Dr.Nelson Jurandi Rosa Fagundes

Dr. Rodrigo Ligabue Braun

Porto Alegre, junho de 2015.

*“A vida é apenas uma visão momentânea
das maravilhas deste assombroso
universo.”*

(Carl Segan)

Agradecimentos

Agradeço primeiramente aos meus pais (Alice e Helio) não só pelos genes por mim herdados, mas principalmente pelo apoio incondicional que recebi durante todos esses anos de estudo. Obrigada por acreditarem em mim desde o começo!

Agradeço a minha orientadora, Karen Haag, pela paciência e pelos ensinamentos durante esse ano de pesquisa.

Agradeço também aos meus colegas de laboratório (Sarah, Sebastian e Juliano) por me aguentarem durante o período de finalização deste trabalho.

Agradeço ao professor Nelson e ao Rodrigo por aceitarem fazer parte da minha banca examinadora.

E por último, mas não menos importante, aos amigos incríveis que a Biologia me deu (Amanda, Bruna, Débora, Jeniffer, Lilith, Lucas, Malu, Nati, Pâmela, Paty, Pietro, Samuel e Verônica). Obrigada por alegrarem todos meus dias durante esses anos de graduação! Vocês são incríveis!

O filo *Microspora* é formado por fungos unicelulares, endobióticos, parasitas intracelulares obrigatórios e produtores de esporos. Para explicar a evolução adaptativa destes organismos, que possuem como característica análoga a dos procariotos a redução extrema do genoma, tem sido estudados principalmente os eventos de transferência gênica horizontal. Entretanto, novas abordagens para identificar possíveis genes candidatos envolvidos com a adaptação destes parasitos aos seus hospedeiros são necessárias. O grupo mais derivado nos microsporídeos é o gênero *Encephalitozoon*, que infecta principalmente hospedeiros mamíferos. Quatro espécies do gênero já tiveram seus genomas totalmente sequenciados – *Encephalitozoon cuniculi*, *E. intestinalis*, *E. romaleae* e *E. hellem*. *Ordospora colligata* é a espécie mais ancestral descrita neste grupo sendo a única que infecta invertebrados (microcrustáceos do gênero *Daphnia*). Tanto *O. colligata* como as espécies do gênero *Encephalitozoon* possuem genomas muito similares, mas hábitos de vida e modos de desenvolvimento muito diferentes. O excesso de divergência (K) de *O. colligata* em relação ao polimorfismo genético (Pi) dentro do gênero *Encephalitozoon* indica os genes que sofreram mais mudanças neste salto evolutivo. O objetivo deste trabalho é investigar os possíveis genes que sofreram evolução adaptativa na passagem do hospedeiro invertebrado para o vertebrado. Para tanto, utilizamos programas de bioinformática e banco de dados genômicos a fim de calcular a razão entre a divergência e o polimorfismo genético utilizando alinhamentos derivados dos genomas completos das cinco espécies analisadas. Consideramos um gene com alta divergência genética aquele que teve razão $K/Pi > 2,0$. Encontramos 233 genes considerados altamente divergentes em *O. colligata*. Quando categorizamos estes genes de acordo com sua hierarquia funcional, é possível visualizar um enriquecimento nas funções de transporte de controle do ciclo celular. Esse fato pode nos indicar que, mesmo que as espécies de encephalitozoonídeos e *O. colligata* possuam genomas parecidos, sua forma de interação com as células do hospedeiro é muito diferente.

Palavras-chaves: *Ordospora colligata*; *Encephalitozoon*; Microsporídeos; Divergência.

The phylum *Microspora* consists of unicellular endobiotic fungi, which are obligate intracellular parasites and that produce spores. The adaptive evolution of these organisms, which are analogous to prokaryotes in having suffered an extreme genome reduction, has been mainly studied by looking at horizontal gene transfer events. However, new approaches to identify possible candidate genes involved in the adaptation of these parasites to their hosts are necessary. The most derived group in microsporidia belongs to the genus *Encephalitozoon*, which infects mainly mammalian hosts. Four species of the genus have had their full genomes sequenced - *Encephalitozoon cuniculi*, *E. intestinalis*, *E. hellem* and *E. romaleae*. *Ordospora colligata* is the most ancestral species described in this group being the only one that infects microcrustaceans invertebrates from the genus *Daphnia*. Both *O. colligata* and the species of the genus *Encephalitozoon* have very similar genomes, but distinct lifestyles and different development modes. The excess of nucleotide divergence (K) of *O. colligata* in relation to the genetic polymorphism (Pi) within the genus *Encephalitozoon* indicates genes that have undergone major changes over this evolutionary jump. The objective of this study is to investigate the possible candidate genes that have undergone adaptive evolution in the passage of an invertebrate to the vertebrate host. Therefore, we used bioinformatics softwares and genome databases in order to calculate the ratio of divergence to genetic polymorphism using nucleotide sequence alignments derived from complete genomes of the five species analyzed. We consider a gene showing $K/Pi > 2.0$ highly divergent. We found 233 genes considered highly divergent. Categorizing them according to a functional hierarchy, an enrichment in genes involved in the control of cell cycle and transport was found. This may indicate that, although encephalitozoonids and *O. colligata* have similar genomes, their way of interaction with host cells is very different.

Keywords: *Ordospora colligata*; *Encephalitozoon*; Microsporidia; Divergence.

O Reino *Fungi* compreende organismos eucariontes, heterotróficos que se alimentam de nutrientes absorvidos do meio, com espécies unicelulares e multicelulares, cujas células são revestidas por uma parede composta principalmente de quitina. O grupo é subdividido em cinco filos: *Ascomycetes*, *Phycomycetes*, *Basidiomycetes*, *Deuteromycetes* e *Microspora*. O filo *Microspora* possui aproximadamente 1200 espécies distribuídas por cerca de 150 gêneros (Wittner, 1999). Os microsporídeos são unicelulares, parasitas intracelulares obrigatórios de praticamente todas as linhagens animais nos mais diversos tipos de ambientes (Figura 1). As infecções causadas por microsporídeos são crônicas, levando a uma redução da aptidão dos hospedeiros, isto é, na sua sobrevivência e fecundidade em relação a animais não infectados.

Os microsporídeos são considerados fungos altamente derivados (Keeling & Fast, 2002). São claramente distintos de outros eucariotos, pois apesar de possuírem um núcleo tipicamente eucariótico e um sistema de membranas internas com citoesqueleto, revelam características moleculares e citológicas similares às de procariotos (Méténier & Vivarés, 2001). Estas características incluem 1) tamanho reduzido do genoma – o menor genoma eucariótico sequenciado, com 2,3Mb, pertence ao microsporídeo *Encephalitozoon intestinalis* (Corradi et al., 2010); 2) a ausência de mitocôndrias, as quais foram reduzidas a estruturas denominadas mitossomos, cuja função ainda não está bem determinada; e 3) ausência do típico aparelho de Golgi (Mathis et al., 2005). A análise de genomas de diversos microsporídeos também mostrou um encurtamento das sequências de proteínas e a perda de regiões intergênicas (Katinka et al., 2001). Os microsporídeos possuem ainda uma estrutura exclusiva chamada tubo polar, que é um longo filamento espiralado (50 µm a 100µm) utilizado pelo esporo durante a invasão da célula hospedeira (Xu & Weiss, 2005). O tubo polar contém um disco de ancoragem na porção terminal anterior e na porção posterior está ligado ao interior do esporo. Este disco de ancoragem perfura a membrana da célula hospedeira e o conteúdo celular presente no interior do esporo é ejetado diretamente para o citoplasma da célula hospedeira (Xu & Weiss, 2005).

Análises de genomas podem fornecer um quadro comparativo, auxiliando na compreensão de como esses organismos tornaram-se parasitas tão bem sucedidos. Os genomas dos microsporídeos já sequenciados mostram uma grande variação de tamanho – entre 2,3 a 24 Mb (Corradi et al. 2009; Pombert et al., 2015; Figura 1). Essas grandes diferenças no tamanho do genoma ocorrem principalmente devido a diferenças na densidade gênica, causada pela presença de elementos repetitivos nas regiões intergênicas e são menos influenciadas pela variação no número total ou no tamanho de genes (Heinz et al., 2012; Peyretailade et al., 2012). Quando analisamos o proteoma das espécies de microsporídeos que já tiveram seus genomas sequenciados, temos uma variação que vai de 1.801 proteínas, na espécie de *Ordospora colligata*, a 4.280 proteínas na espécie *Hamiltosporidium tvaerminnensis* (Pombert et al., 2015).

Embora algumas das diferenças no tamanho dos proteomas tenham significado funcional conhecido, como a presença ou ausência de enzimas glicolíticas (Keeling et al., 2010) a maioria das diferenças envolve proteínas hipotéticas, ou seja, sem função determinada (Cuomo et al., 2012; Heinz et al., 2012; Peyretailade et al., 2012). Há

também expansões de famílias de genes específicos de cada linhagem (Nakjang et al., 2013) como, por exemplo, a família de genes denominados *interB* presente em espécies dos gêneros *Encephalitozoon*, *Vittaforma*, e *Anncalia*. A importância funcional desta família proteica ainda é desconhecida.

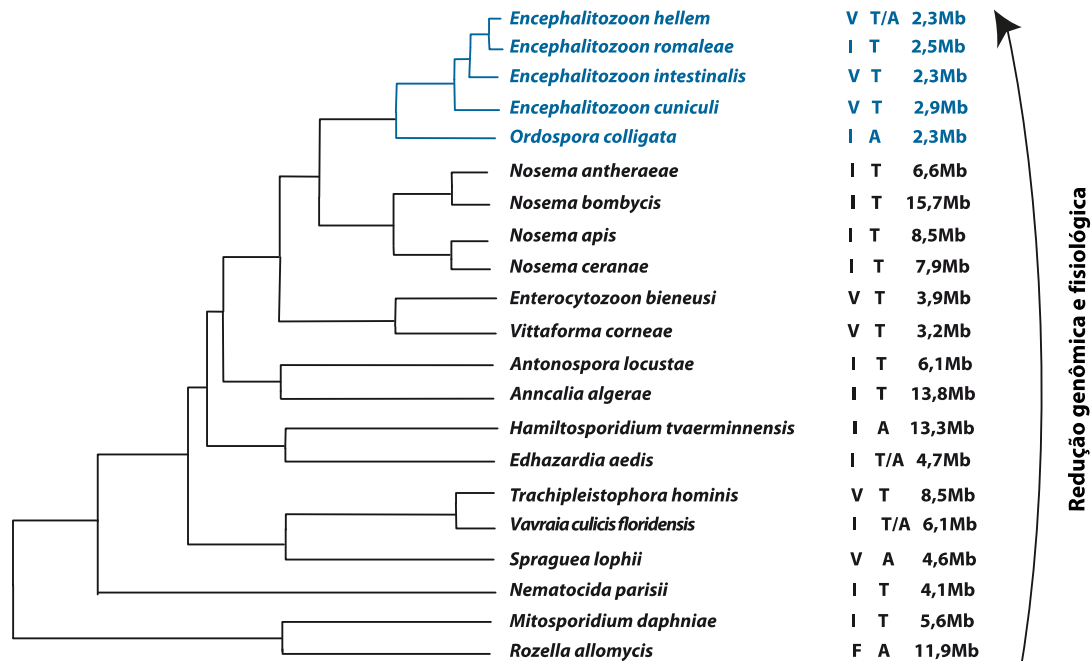


Figura 1: Árvore filogenética simplificada das espécies do filo *Microspora* cujos genomas já foram sequenciados. Os nomes das espécies aparecem seguidos do grupo de hospedeiros principal (Vertebrado –V, Invertebrado – I e Fungo - F), do seu ambiente (Aquático - A e Terrestre - T) e do tamanho dos seus genomas. Algumas espécies de microsporídeos parasitam animais cujo ciclo de vida envolve dois tipos de ambientes, como no caso das espécies *Edhazardia aedis* e *Vavraia culicis* que utilizam tanto o mosquito adulto da espécie *Aedes aegypti* como suas larvas aquáticas. Apesar de haver uma tendência de redução genômica na história evolutiva dos microsporídeos, também encontra-se grupos sofrendo moderadas ou grandes expansões (Adaptado de Pombert et al., 2015).

As mudanças celulares que ocorreram nos microsporídeos revelam uma extrema simplificação fisiológica, incluindo a perda de vias metabólicas completas, como o ciclo de Krebs, por exemplo (Keeling et al., 2010), o que os torna dependentes de seu hospedeiro na obtenção de energia. Este elevado grau de dependência está refletido nos transportadores celulares codificados pelo genoma dos microsporídeos (por exemplo, transportadores de ATP), os quais lhes permitem adquirir energia e nutrientes essenciais de seus hospedeiros. Há evidências de que alguns desses transportadores de ATP possam ter origem na transferência horizontal de genes, adquiridos a partir da coexistência com patógenos bacterianos (Keeling et al., 2010).

Eventos de transferência gênica horizontal, ainda que raros mesmo em organismos que estão intimamente interligados ao seu hospedeiro, são de extrema importância para os microsporídeos, pois são a única evidência de adaptação presentes no grupo (Corradi et al., 2010). Os microsporídeos não adquiriram genes apenas das bactérias, também os obtiveram a partir de seus hospedeiros. *E. romaleae* obteve de seu hospedeiro artrópode

um gene que atua na via das purinas, codificando a purina fosforilase (PNP). Assim, o microsporídeo consegue reciclar suas purinas sem depender diretamente do seu hospedeiro (Corradi et al., 2010).

Estudos recentes descrevem que uma das grandes diferenças entre os genomas de *O. colligata* e das espécies de encephalitozoonídeos é a presença do gene que codifica a proteína septina 7 (Sept7) (Pombert et al., 2015). Este gene provavelmente foi adquirido por meio de um evento de transferência genética horizontal e, embora muito similar ao gene *sept7* encontrado em *Daphnia magna* e *Daphnia pulex*, ainda não é possível determinar se foi obtido diretamente destas espécies ou se é oriundo de alguma espécie ancestral. A função da Sept7 é desconhecida em *O. colligata*, mas em outras espécies de fungos patógenos (como *Candida albicans*) proteínas conhecidas como invasinas interagem com a septina 7 dos hospedeiros e iniciam uma cascata molecular que induz a absorção do patógeno (Phan et al., 2013). Como *O. colligata* não possui invasinas, a presença da *sept7* na superfície celular pode camuflar a entrada do patógeno no hospedeiro (Pombert et al., 2015).

O grupo mais derivado do filo *Microspora*, isto é, aquele em que o grupo atingiu a maior redução genômica, é o gênero *Encephalitozoon*. Esse é um gênero que infecta basicamente vertebrados – *E. cuniculi* tem mamíferos como hospedeiros, principalmente coelhos; *E. hellem* pode infectar anfíbios, répteis, aves e mamíferos; *E. intestinalis* tem mamíferos como hospedeiros principais, infectando inclusive humanos imunocomprometidos. A exceção no gênero é a espécie *E. romaleae* cujo hospedeiro é o gafanhoto (Figura 1).

O ancestral do gênero *Encephalitozoon* provavelmente infectava o trato digestivo de alguma espécie de mamífero. Interessantemente, *O. colligata*, a espécie mais relacionada filogeneticamente a este grupo, infecta o epitélio das alças intestinais anteriores de microcrustáceos do gênero *Daphnia* (Ebert, 2005). Além disso, todas as outras espécies deste clado, que também inclui o gênero *Nosema* (Figura 1), utilizam artrópodes como hospedeiro. Um possível cenário que explicaria o salto de hospedeiros artrópodes para mamíferos seria através de predação do hospedeiro derivado (vertebrado) sobre o hospedeiro ancestral (artrópode).

Este trabalho baseia-se na hipótese de que o salto de hospedeiro invertebrado para vertebrado tenha ocorrido mais de uma vez na história evolutiva dos microsporídeos (Figura 1) e que tenha envolvido uma série de mudanças genômicas adaptativas. A adaptação ao nível molecular dos microsporídeos tem sido primariamente estudada sob a ótica da transferência gênica horizontal. A maior dificuldade em estudar processos adaptativos analisando as sequências dos genes é a antiguidade do grupo – os microsporídeos habitam o planeta há pelo menos 900 milhões de anos (Hedges et al., 2004) – dificultando a busca de rastros de evolução adaptativa. Outro problema são as altas taxas de substituição de nucleotídeos típicas dos microsporídeos, que levam à rápida saturação (Capella-Gutiérrez et al., 2012). Entretanto, comparações entre genomas de espécies fortemente relacionadas podem auxiliar na busca de novas evidências de adaptação no nível das sequências primárias dos genes. O objetivo deste trabalho é investigar quais foram estas mudanças envolvendo passagem do hospedeiro invertebrado para o vertebrado, utilizando como base os genomas das espécies de *Encephalitozoon* e de *O. colligata*.

Dados de Sequências Genômicas

Em nossa análise comparativa foram incluídos cinco genomas de microsporídeos: *Encephalitozoon cuniculi* GB-M1 (Katinka et al., 2001) (RefSeq: NC_003229); *Encephalitozoon hellem* ATCC50504 (Pombert et al., 2012) (RefSeq: NC_018468); *Encephalitozoon intestinalis* ATCC50506 (Corradi et al., 2010) (RefSeq: NC_014415); *Encephalitozoon romaleae* SJ-2008 (Pombert et al., 2012) (RefSeq: NC_024896) e *Ordospora colligata* OC4 (GenBank: GCA_000803265). Todas as sequências foram acessadas entre março e abril de 2014.

Construção dos alinhamentos

Os onze cromossomos de *E. cuniculi* foram alinhados com regiões homólogas dos genomas das outras espécies utilizando o programa LASTZ (versão 1.02.00) (Schwartz, 2003), inserido no pacote Geneius (Versão 7.1.6, Biomatters). O procedimento realizado por LASTZ segue um protocolo em que pequenos fragmentos de identidade de uma sequência de referência em relação às outras sequências a serem alinhadas são utilizados como padrões iniciais (*seeds*). A partir deles, o alinhamento de melhor escore é definido até um comprimento limiar denominado HSP (*High Scoring Segment Pair*), o qual é definido pelo usuário. Por último os segmentos alinhados de HSP são ligados fazendo uso da inserção de *gaps*, se necessário. Este protocolo é adequado para alinhar grandes segmentos de DNA, como cromossomos inteiros. Os parâmetros predefinidos pelo programa foram seguidos: comprimento dos padrões iniciais = 20 bases; número de nucleotídeos idênticos nos *seeds* = 12-19 pb; comprimento do segmento HSP = 3.000 pb. O genoma de *E. cuniculi* foi a referência, por ser o melhor curado, sendo o modelo de estudos com microsporídeos de uma forma geral. Após o alinhamento, foram extraídas somente as sequências dos genes presentes nos cinco genomas contendo mais de 500 nucleotídeos, com intuito de remover todos os *gaps*. Estas sequências foram concatenadas, formando o alinhamento final dos genes de cada cromossomo. Os alinhamentos foram convertidos para o formato FASTA que representa, em formato de texto, as sequências de nucleotídeos e de peptídeos utilizando códigos de letra única. Este formato é necessário para facilitar as análises posteriores.

Análises dos alinhamentos

Para analisarmos os valores de divergência e polimorfismo genético utilizamos o programa DNAsp 5.10.1 (Rozas et al., 2010). Este é um software que estima diversos parâmetros de variação a partir de sequências de DNA alinhadas.

O polimorfismo genético foi calculado com base no número médio de diferenças de nucleotídeos (Nei & Li, 1979) entre duas ou mais sequências, chamado de diversidade nucleotídica. Isso é definido por meio da fórmula:

$$\pi = \sum_{ij}^q x_i x_j d_{ij}$$

onde, q é o número total de alelos (sequência de nucleotídeos diferentes); x_i é a frequência do alelo i na população ($=0,25$) e d_{ij} é o número de diferenças ou substituições de nucleotídeos entre os alelos i e j . Em uma população aleatória π é a heterozigose ao nível de nucleotídeos.

A divergência genética entre *O. colligata* e cada uma das espécies de *Encephalitozoon* é o número médio de substituições de nucleotídeos par-a-par nas sequências dos dois grupos. Esta definição é dada por:

$$d_x = \frac{q}{q-1} \sum_{wz}^q x_w x_z d_{wz}$$

onde, x_w é a frequência do alelo w em *O. colligata* ($=1$), x_z é a frequência do alelo z em uma das espécies de *Encephalitozoon* e d_{wz} é uma estimativa do número de substituições de nucleotídeos entre as sequências dos dois grupos (Nei & Kumar, 2000).

As espécies do gênero *Encephalitozoon* formaram o grupo intraespecífico, já *Ordospora colligata* foi considerada como grupo interespecífico. Uma janela de 500 nucleotídeos com paradas a cada 100 foi deslizada sobre os alinhamentos de cada um dos onze cromossomos a fim de identificar os SNPs (*Single Nucleotide Polymorphism*). Com base nestes dados as taxas de polimorfismo entre as espécies do gênero *Encephalitozoon* ($\text{Pi} = \pi$) e divergência em relação a *O. colligata* ($\text{K} = d_x$) foram obtidas. As regiões com excesso de divergência – calculado a partir da razão $\text{K}/\text{Pi} > 2,0$ – foram interpretadas como sendo potenciais rastros de adaptação no nível molecular.

Os resultados destas análises foram organizados em forma de tabelas, nas quais os genes foram ordenados de forma decrescente, tendo como base o valor de K/Pi , e de gráficos das taxas de divergência e Polimorfismo. Tanto tabelas quanto gráficos foram obtidos a partir do Microsoft Excel 2010, inserido no pacote Office 2010 (Microsoft).

Análises dos genes divergentes

Os genes com alta divergência ($\text{K}/\text{Pi} > 2,0$) foram identificados a partir de seu ponto médio, região do alinhamento que possui maior valor de divergência. As funções dos genes encontrados nestas regiões foram obtidas a partir das anotações nos genomas de *E. cuniculi* e *Saccharomyces cerevisiae* – outro importante fungo modelo – com a utilização das plataformas KEGG (Kyoto Encyclopedia of Genes and Genomes <http://www.keeg.jp>) e OrthoMCL (Ortholog Group of Protein Sequence <http://www.orthomcl.org>).

Na plataforma KEGG, utilizamos o aplicativo BRITE, uma coleção de ferramentas que agrupa um conjunto de proteínas de acordo com uma hierarquia funcional. Cada arquivo de

hierarquia é identificado por códigos de duas a quatro letras e cinco dígitos (Códigos KEGG-Apêndice 1). Em nossa análise utilizamos o prefixo Ko (hierarquia de referência), estes arquivos foram criados manualmente utilizando os códigos KEGG dos genes e convertidos no programa em identificadores de genes para cada família de funções.

O programa OrthoMCL (Li et al., 2003) distingue as redundâncias de genes em um genoma devido a duplicações (paralogia) e divergência de genes com a mesma função em distintos organismos (ortologia). Genes ortólogos são identificados através do algoritmo Markov-Cluster. Um banco de dados derivado de análises de OrthoMCL de uma amostra de 150 genomas representativos de grandes grupos de seres vivos está disponível em <http://www.orthomcl.org> e pode ser utilizado para análise da conservação evolutiva ou função de genes de qualquer organismo. Para determinarmos com maior precisão a função de cada gene divergente utilizamos sua sequência de aminoácidos e executamos o algoritmo BLAST - *Basic Local Alignment Search Tool* - (Altschul et al., 1990) em busca de genes similares no banco OrthoMCL. BLAST nos permite comparar as sequências primárias, tais como sequências de aminoácidos de diferentes proteínas ou nucleotídeos de distintas moléculas de DNA. Isto nos indica o quão semelhantes elas são, e estas semelhanças podem indicar homologia, além possibilitar a inferência de função. Neste estudo utilizamos as funções dos genes da levedura *Saccharomyces cerevisiae* como referência. Foi utilizado um valor limite de e^{-8} (quanto menor o valor de e maior o êxito) para considerar duas sequências como sendo homólogas.

Resultados

O total de nucleotídeos alinhados nos onze cromossomos foi de 375.399 os quais correspondem a 337 genes codificadores de proteínas (Tabela 1). O proteoma total de *O. colligata* contém 1850 genes. A discrepância em relação ao tamanho do proteoma analisado em nosso estudo deve-se a exclusão de genes que continham *gaps*. O comprimento médio dos alinhamentos de cada cromossomo foi de 32.312 nucleotídeos, mas houve grande variação entre os cromossomos. O maior alinhamento obtido foi o do cromossomo XI, com 57.108 nucleotídeos, e o menor foi o IX, com 10.674 (Tabela 1). Esta grande redução no comprimento do alinhamento obtido para o cromossomo IX pode ter ocorrido pois, diferentemente dos demais contigs de *O. colligata* que são estruturalmente semelhantes aos dez cromossomos de *E. cuniculi*, há evidências de numerosos rearranjos neste cromossomo, que é formado por diversos fragmentos dos contigs II, IV, X e XI de *O. colligata* (Pombert et al., 2012). O cromossomo com maior número de genes também foi o XI, com 49 genes, provavelmente devido ao fato de este ser o cromossomo com o mais longo alinhamento de genes. O inverso ocorre com o cromossomo IX que possui um total de dez genes ao longo do alinhamento (Tabela 1). A média do número de genes por alinhamento foi de 27. As porcentagens de sítios idênticos e de conteúdo de GC tiveram pouca variação entre os cromossomos; as médias mantiveram-se em 57,2 e 43,4% (Tabela 1), respectivamente. Os picos nas razões de divergência *versus* polimorfismo (K/Pi) e de polimorfismo *versus* divergência (Pi/K) tiveram distribuição semelhante e homogênea nos alinhamentos (Apêndice 2). As exceções são encontradas no

cromossomo IX que teve uma distribuição mais linear dos valores e apresentou poucos picos (Apêndice 2) e nos cromossomos X e XI que apresentaram um grande número de picos (Apêndice 2), o que pode ter ocorrido devido ao maior comprimento dos alinhamentos destes cromossomos.

Tabela 1 – Estimativas da razão de divergência sobre o polimorfismo (K/Pi), número total de genes (N), número total de genes com excesso de divergência (D), comprimento do alinhamento (L), fração de sítios idênticos (I) e conteúdo de GC nos alinhamentos obtidos para cromossomo, utilizando o genoma de *E. cuniculi* como referência.

Cromossomo	K/Pi	N	D	L (pb)	I (%)	GC (%)
I	2,029	23	11	29.074	60,5	43,9
II	2,014	39	15	39.919	56,1	42,8
III	1,908	27	18	32.312	57,3	43,3
IV	2,068	27	23	33.366	57,5	43,6
V	2,036	27	21	28.825	56,7	43,6
VI	2,43	27	19	28.042	56,3	43,4
VII	2,027	24	18	29.340	56,6	42,2
VIII	2,041	40	31	37.923	57	44
IX	2,235	10	7	10.674	57,9	41,4
X	1,797	44	34	48.816	57,2	43,6
XI	1,777	49	36	57.108	58,4	43,4
Total	na	337	233	375.399	na	na
Médias	2,029	27	19	32.312	57,2	43,4

na = não se aplica

Dos 337 genes incluídos em nossa análise, 233 foram considerados altamente divergentes ($K/Pi > 2,0$), o que corresponde a aproximadamente 2/3 dos genes analisados (Tabela 1 e Apêndice 1). Utilizamos o valor de 2,0 como ponto de corte para determinar quais genes tiveram excesso de divergência, pois este foi o valor médio para a razão K/Pi nos cromossomos (Tabela 1). O cromossomo que teve maior média no valor K/Pi foi o VI com razão de 2,64 (Tabela 1), e o que obteve maior número de genes divergentes foi o XI com 36. Entretanto o cromossomo que teve proporcionalmente maior quantidade de genes altamente divergentes foi o IX onde sete de seus dez genes apresentaram uma razão maior do que dois. Em média o número de genes com excesso de divergência por cromossomo foi de dezenove (Tabela 1).

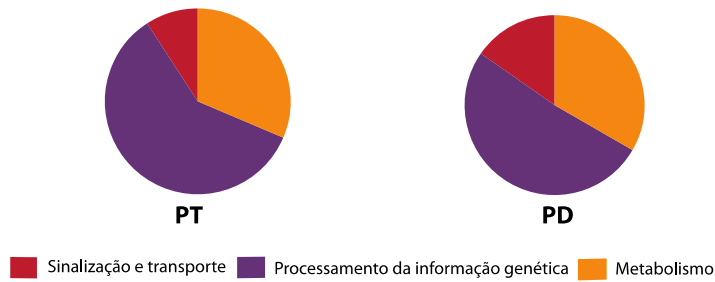
O gene codificador de um transportador ABC (tipo mitocondrial), cuja função está relacionada ao transporte de ATP, foi o que apresentou maior excesso de divergência. Este gene, homólogo a ECU11_1200, está presente no cromossomo XI e teve valor de razão K/Pi igual a 3,08. (Apêndice 1). Outros genes que também obtiveram altos níveis de divergência foram as proteínas quinases (do tipo Serina/Treonina e da família PI3/PI4), envolvidas no controle do ciclo celular, proteínas de reparo de DNA e fatores de transcrição e tradução (Apêndice 1). Proteínas hipotéticas, que são proteínas ainda sem elucidação funcional, também estão representadas em nosso estudo formando um grupo de 38 genes (Apêndice 1).

Quando, com o auxílio da plataforma KEGG, dividimos os genes divergentes de acordo com sua hierarquia funcional, obtivemos três grandes grupos de funções: 1) genes que

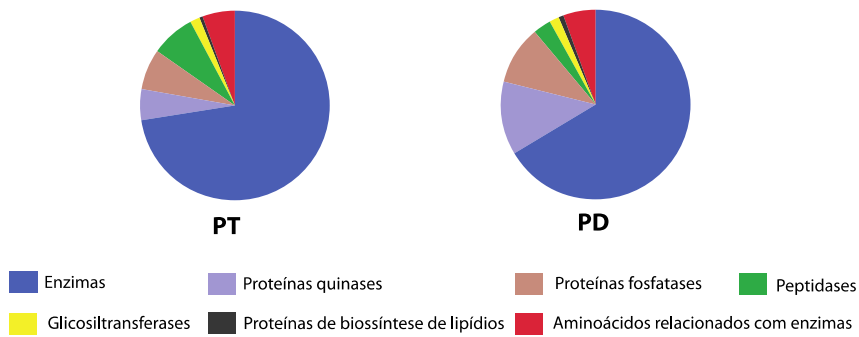
codificam proteínas do metabolismo celular; 2) genes que codificam transportadores ou moléculas sinalizadoras; e 3) genes que codificam proteínas que participam do processamento da informação genética. Entretanto, é importante levar em conta que um mesmo gene pode codificar uma proteína que possua diversas funções e assim, se encaixar em mais de uma categoria hierárquica. Ao compararmos o proteoma altamente divergente de *O. colligata* em relação ao seu proteoma total vemos que há um enriquecimento de funções metabólicas e, principalmente, de funções ligadas ao sistema de sinalização e transporte celular no proteoma divergente (Figura 2).

Ao analisarmos cada um destes grupos de funções individualmente, observamos que nas funções metabólicas, os genes com excesso de divergência mostraram um enriquecimento de funções relacionadas com a fosforilação de proteínas em relação às funções anotadas a partir do proteoma total de *O. colligata* (Figura 2 e Tabela 2). Já no grupo envolvido na função de sinalização e transporte celular, os genes divergentes apresentaram um enriquecimento em proteínas envolvidas com o transporte e proteínas do citoesqueleto (Figura 2 e Tabela 2). No grupo de processamento da informação genética não houve aumento significativo de funções nos genes divergentes em relação ao proteoma total de *O. colligata* (Tabela 2).

Grandes grupos de funções celulares



Metabolismo celular



Sinalização e transporte celulares

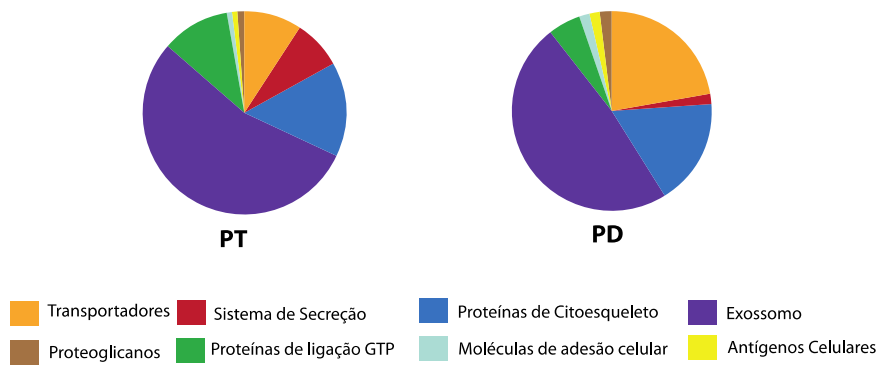


Figura 2 – Distribuição das funções de proteínas codificadas pelos genes divergentes- PD (total = 233) em relação as do proteoma total de *O. colligata* - PT (total =821). a) Grandes módulos de funções; b) Sub-categorias de funções relacionadas ao metabolismo; c) Sub-categorias de funções relacionadas a sinalização e transporte celulares.

Tabela 2: Número de genes relacionados aos distintos grupos de funções no proteoma total de *O.colligata* e do proteoma divergente, com os respectivos códigos de módulos KEGG.

	Genoma total	Divergente	Módulo KEGG
Metabolismo			
Enzima	299	85	Ko01000
Proteínas Quinases	22	16	Ko01001
Proteínas Fosfatases	29	13	Ko01009
Peptidases	30	04	Ko01002
Glicosiltransferases	07	02	Ko01003
Proteínas de Biossíntese de Lipídios	03	01	Ko01004
Aminoácidos relacionados com enzimas	22	07	Ko1007
Total	422	128	na
Sinalização e Processos Celulares			
Transportadores	11	13	Ko02000
Sistema de Secreção	09	01	Ko02044
Proteínas de Citoesqueleto	18	10	Ko04812
Exossomos	64	28	Ko04147
Proteínas de ligação GTP	13	02	Ko04031
Moléculas de adesão e ligantes	01	01	Ko04516
Antígenos celulares	01	01	Ko04090
Proteoglicanos	01	01	Ko00535
Total	118	58	na
Processamento da informação genética			
Fatores de transcrição	10	03	Ko03000
Maquinária de transcrição	78	19	Ko03021
Biogênese de mRNA	69	21	Ko03019
Spliceossoma	45	17	Ko03041
Ribossomo	74	11	Ko03011
Biogênese ribossomal	75	25	Ko03009
Biogênese de tRNA	60	13	Ko03016
Fatores de tradução	26	10	Ko03012
Chaperonas	21	11	Ko03110
SNARES's	04	01	Ko04131
Sistema ubiquitina	37	07	Ko04121
Proteassoma	36	10	Ko03051
Proteínas de replicação do DNA	53	10	Ko03032
Cromossomo	105	30	Ko03036
Reparo e recombinação de DNA	64	14	Ko03400
Biogênese mitocondrial	20	06	Ko03029
Total	777	198	na

na= não se aplica

O filo *Microspora* parece ter dominado o modelo de aquisição de genes por transferência gênica horizontal a partir de seus hospedeiros, bem como das bactérias com as quais divide o microambiente da célula. Esses novos genes adquiridos horizontalmente são utilizados para a sua proteção ou para auxiliar no seu metabolismo (Selman & Corradi, 2011). Em termos de estrutura geral e de densidade genética, os genomas de *O. colligata* e das espécies de *Encephalitozoon* são equivalentes, o que provavelmente denota que o último ancestral comum das espécies era muito semelhante em todos aspectos (Pombert et al., 2015).

Nós verificamos que o genoma de *O. colligata* apresenta 233 genes altamente divergentes ($K/Pi > 2,0$), isso corresponde a aproximadamente 2/3 dos 337 genes perfeitamente alinháveis nas cinco espécies. Esta é uma fração do proteoma de *O. colligata* – que contém 1801 proteínas anotadas (Pombert et al., 2015) – mas outros trabalhos feitos com esse organismo também não utilizaram seu proteoma total. Pombert e colaboradores analisaram 1650 genes, mas ao invés de alinhar as sequências dos cromossomos inteiros, utilizaram as sequências codificadoras de proteínas (CDS) individuais, e estimaram os parâmetros de polimorfismo e divergência a partir da totalidade da sequência dos genes, sem remover os *gaps*. O gene com maior divergência ($K/Pi = 6,914$) em Pombert et al., 2015 foi ECU11_1210 que codifica uma proteína hipotética. Na sequência foram encontrados os genes ECU06_0240 que codifica a proteína II de tubo polar, com razão K/Pi de 5,748 e o gene ECU02_0280 que expressa outra proteína hipotética, com divergência calculada em 4,875. Estes três genes não mostraram excesso de divergência em nosso trabalho.

No presente trabalho regiões contendo *gaps* foram removidas dos alinhamentos, devido ao seu elevado grau de incerteza. Ainda, os parâmetros de divergência (K) e polimorfismo (Pi) foram estimados a partir de uma janela de 100 nucleotídeos deslizada sobre o alinhamento dos cromossomos. Este procedimento evita que pequenos trechos de sequências altamente divergentes sejam diluídos dentro de genes muito longos. De modo inverso, evita também que eles resultem em uma super-estimativa da divergência real em alinhamentos encurtados pela presença de numerosos *gaps*, uma vez que estas regiões não estão incluídas nos cálculos.

No topo da lista, com maior valor de excesso de divergência ($K/Pi = 3,08$), encontramos o gene ECU11_1200, que codifica um transportador ABC de tipo mitocondrial (Apêndice 1). Outro gene codificador de transportador ABC também mostrou excesso de divergência: ECU04_0480 com $K/Pi = 2,168$. Os transportadores ABC formam uma das maiores e mais conhecidas famílias de proteínas presente tanto em procariotos como em eucariotos. Estão localizados nas membranas que acoplam a hidrólise de ATP e ativam o transporte transmembrana (efluxo e influxo) de diversos metabólitos como açúcares, lipídios, esteróis, entre outros (Zhao et al., 2004). Os transportadores também participam da invasão ativa aos hospedeiros manipulando o sistema imune para que não sejam reconhecidos, captando moléculas específicas que interferem nas funções básicas da célula a ser infectada (Zhao et al., 2004). A presença no genoma de um gene que codifica o transportador ABC de tipo mitocondrial confirma que, apesar de atualmente os microsporídeos não possuem

mitocôndrias, seu ancestral ainda as possuía. Esse fato é corroborado por trabalhos realizados com o microsporídeo *Trachipleistophora hominis*, espécie na qual foram descritos pela primeira vez os mitossomos, que são pequenas estruturas de membrana dupla. Sua função ainda não está bem definida, mas acredita-se ter retido algumas propriedades mitocondriais (Heinz et al., 2012).

Três carreadores de ATP/ADP foram localizados por hibridização *in situ* na membrana plasmática de *E. cuniculi* e um na parede dos seus mitossomos, sugerindo que o mitossomo importa, ao invés de exportar, ATP (Tsaousis et al., 2008). Carreadores de ATP/ADP também mostram excesso de divergência no nosso estudo, como, por exemplo ECU08_1300 com K/Pi = 2,567; ECU10_0420 com K/Pi= 2,567 e ECU10_0520 com K/Pi= 2,204. Os transportadores pertencem a um dos subgrupos de funções que mostraram um maior enriquecimento, em relação ao proteoma total de *O.colligata*, no proteoma divergente (Tabela 2 e Figura 2). Foram encontrados genes que codificam transportadores de maltose (ECU11_1870; K/Pi=2,469), de peptídeos (ECU11_1050; K/Pi=2,359) e de zinco (ECU11_1510; K/Pi= 2,100). Nas bactérias os transportadores são fundamentais durante o processo de infecção da célula hospedeira, pois, após o contato entre as superfícies celulares, são necessários para a captação de metabólitos, nucleotídeos, além de transportar as toxinas produzidas durante o metabolismo das bactérias podendo aumentar a patogenicidade do parasita (Schmid-Hempel, 2011). Muitos destes transportadores foram adquiridos pelos microsporídeos através de transferência gênica horizontal (Keeling et al., 2010); logo podemos inferir que eles possuem uma função similar aos das espécies doadoras de genes.

Na segunda e terceira posições dos genes com mais alto nível de divergência em relação ao polimorfismo foram encontrados codificadores de duas proteínas hipotéticas: proteína hipotética YD95 (ECU10_0190 com K/Pi= 3,048) e a proteína hipotética ECU09_1950 (K/Pi= 2,917). A primeira teve sua função inferida a partir de anotações em *S. cerevisiae*: trata-se de proteína reguladora do proteassomo 26S (complexo de capaz de degradar proteínas em oligopeptídeos). O complexo 26S reconhece especificamente proteínas do sistema ubiquitina, que sinaliza o processo de degradação de proteínas defeituosas. A segunda também pode ter a função inferida a partir de *S.cerevisiae*, onde foi mostrado ser uma proteína de membrana que media o tráfego membranar e mantém a morfologia do retículo endoplasmático, fazendo parte do exossomo.

Dentre as funções metabólicas, as proteínas quinases aparecem em maior proporção no proteoma divergente (Tabela 2 e Figura 2). As proteínas quinases, são um tipo específico de enzimas que fosforilam outras proteínas, por meio da transferência de um ATP. As proteínas quinases serina/treonina do tipo polo quinases (ECU01_0630 com K/Pi=2,816; ECU05_0630 com K/Pi=2,136; ECU08_1620 com K/Pi=2,00) foram as mais representativas dentre as quinases com excesso de divergência. Elas são responsáveis pelo controle do ciclo celular, regulando o processo da citocinese, por meio da fosforilação das proteínas RhoGEF. Após ativadas, atuam na polaridade celular e na dinâmica dos microtúbulos do citoesqueleto, formando uma espécie de “pescoço” na célula que está em divisão (Gould & Bohnert, 2011).

A família de quinases PI3/PI4 também foi considerada altamente divergente (ECU05_0540 com K/Pi= 2,873; ECU02_1130 com K/Pi= 2,628). Essa família de

quinases faz parte do complexo de histonas acetiltransferases onde atua na ativação da transcrição sendo considerada com um “*checkpoint*” que garante a integridade do genoma. É necessária como transdutor de sinal durante o ciclo celular em resposta a um dano de DNA ou a sua não replicação. Outra família de quinases com excesso de divergência foi CDC2/CDKX (ECU11_1290; K/Pi= 2,027). Trata-se de uma família de quinases dependente de ciclina que é necessária para a progressão do ciclo celular da fase G para a fase S (Silva et al., 2009).

Outra quinase com excesso de divergência foi a de tipo SPK1 (ECU02_0550; K/Pi= 2,733) que se auto-fosforila em resposta ao dano do DNA desempenhando assim um papel importante para iniciar a replicação do DNA. A quinase C (ECU01_1320; K/Pi=2,624) também teve excesso de divergência. Essa é uma quinase que está envolvida em diversas funções celulares como, por exemplo, na regulação da transcrição e do crescimento celular através da fosforilação de outras proteínas envolvidas nestes dois processos (Mellor & Parker, 1998). A família de quinases dependente de calmodulina (ECU03_0630; K/Pi= 2,469) também participa do controle de outras proteínas através da fosforilação. A calmodulina é uma proteína relacionada na regulação do cálcio celular, sua ligação com outras proteínas é dependente de alterações na concentração intracelular de cálcio. Quando se liga a enzima ou proteína alvo modifica atividade de várias enzimas alvo e proteínas transporte da membrana. A última família de quinases com excesso de divergência é a caseína. A cadeia alfa da caseína (ECU05_1510; K/Pi=2,193) tem papel no crescimento e na proliferação celular. Seus substratos contêm todos os fatores de transcrição e RNAs polimerase. Já a caseína 1 (ECU03_0910; K/Pi=2,124) está envolvida em processos de transdução de sinal intracelular .

Ainda no grupo de funções relacionadas ao metabolismo outra sub-categoria com enriquecimento, quando comparamos o proteoma total de *O.colligata* e o proteoma divergente, é o das proteínas fosfatases. Esse grupo de proteínas controla muitas vias de sinalização celular por meio da fosforilação/desfosforilação de proteínas (Aoyama et al., 2003). A proteína nuclear *scaffold* (ECU11_1160; K/Pi = 2,352) foi a proteína com maior divergência neste grupo. Esta é uma proteína que faz parte do complexo promotor da anáfase garantindo a degradação de seus inibidores (incluindo as ciclinas mitóticas). O indutor fosfatase com fase M (ECU04_0710; K/Pi= 2,352) também foi encontrado neste grupo. Esta é uma tirosino-fosfatase envolvida no controle do ciclo celular regulando o estado de fosforilação de Cdc28p (conjunto de proteínas que impede a remontagem de complexos pré-replicativas; Drury et al., 2000).

Outra fosfatase com alto nível de divergência foi a 6-fosfofrutoquinase isoenzima 1 (ECU03_0680; K/Pi= 2,332). Esta é uma subunidade envolvida na glicólise, indispensável para o crescimento anaeróbio sendo ativada pela frutose 2-6 bisfosfato e AMP. Mutações no gene que codifica esta fosfatase inibem a indução da glicose para genes relacionados com ciclo celular. Neste grupo ainda estão incluídas a subunidade beta de calcineurina (ECU08_0160; K/Pi= 2,126), que é uma fosfatase que regula fatores de transcrição de proteínas de resposta ao estresse, e fosfolipídio ATPase (ECU06_0720; K/Pi=2,049) que mantém a membrana lipídica simétrica.

No grupo de funções envolvidas nos processos celulares de transporte e de sinalização temos outra sub-categoria com pequeno enriquecimento na comparação com o proteoma

total de *O. colligata*: as proteínas de citoesqueleto (Figura 2). Essas proteínas formam uma rede complexa no interior da célula, sendo importantes durante o ciclo celular, garantindo a manutenção da forma celular. Participam também do transporte intracelular, e da dinâmica estrutural da célula. A principal família de proteínas do citoesqueleto com excesso de divergência são as cinesinas (ECU05_0060 com K/Pi=2.337; ECU10_0320 com K/Pi=2,249; ECU08_0890 com K/Pi=2,214). Estas proteínas estão envolvidas na formação dos fusos mitóticos e na formação do corpo polar, sendo necessária durante a fusão nuclear. Essa é uma família de proteínas que também está classificada como fosfatase. As proteínas que formam o citoesqueleto estruturalmente também tiveram excesso de divergência: a cadeia de gama tubulina (ECU08_0670; K/Pi=2,177) que está envolvida na montagem de microtúbulos e do corpo polar do fuso; a cadeia pesada de miosina (ECU04_1000; K/Pi= 2,156) proteína que transporta vesículas e organelas. Ainda temos a proteína com domínio calponina (ECU10_1740; K/Pi=2,092) que é essencial para localizar os marcadores axiais e localiza o anel contrátil durante o processo de anáfase.

Os microsporídeos, por serem organismos unicelulares e parasitas obrigatórios, possuem apenas uma única fase de seu ciclo fora da célula hospedeira, que é a de esporos. Quando conseguem invadir algum hospedeiro possuem duas fases distintas no seu desenvolvimento: a merogonia que é uma fase proliferativa, na qual ocorre um aumento maciço no número de parasitos dentro da célula hospedeira; e a esporogonia, onde há a produção de esporoblastos que após fases de amadurecimento tornam-se esporos que irão infectar outros hospedeiros (Bigliardi & Sacchi, 2001). Uma importante diferença entre *O. colligata* e as demais espécies do gênero *Encephalitozoon* ocorre na fase da esporogonia. Os encephalitozoonídeos desenvolvem seus esporos dentro de um vacúolo parasitóforo (membrana lipídica que representa um nicho especializado onde o parasita pode se desenvolver e replicar). Já os esporos de *O. colligata* se desenvolvem em contato direto com o citoplasma da célula hospedeira (Larsson, Ebert & Vávra, 1997). Isso nos sugere que a forma de interação com o hospedeiro é diferente mesmo que os genomas das espécies de *Encephalitozoon* e de *O. colligata* sejam muito parecidos.

Este estudo verificou que os genes mais divergentes entre as espécies de *O. colligata* e o grupo dos encephalitozoonídeos são os que codificam transportadores ou proteínas que, de alguma forma, estão relacionadas com os mecanismos que controlam a divisão celular e a transdução de sinais. Para definirmos se estas mudanças tem caráter adaptativo ou se são oriundas de contingências evolutivas são necessários estudos mais detalhados, mas este trabalho permitiu montar uma lista de genes candidatos para uma investigação mais profunda a respeito da evolução adaptativa dos microsporídeos mais derivados.

Referências Bibliográficas

ALTSCHUL, S.; GISH, W.; Myers, E.; Lipman, D. Basic Local, Alignment Search Tool. **Journal of Molecular Biology**, v.21, n. 5, p.403-410, 1990.

AOYAMA, H. et al. Proteínas tirosina fosfatases: Propriedades e funções biológicas. **Química Nova**, v. 26, n. 6, p. 896–900, 2003.

BIGLIARDI, E.; SACCHI, L. Cell biology and invasion of the microsporidia. **Microbes and Infection**, v. 3, n. 5, p. 373–379, 2001.

CAPELLA-GUTIÉRREZ, S.; MARCET-HOUBEN, M.; GABALDÓN, T. Phylogenomics supports microsporidia as the earliest diverging clade of sequenced fungi. **BMC Biology**, v. 10, n. 1, p. 47, 2012.

CORRADI, N. et al. Draft genome sequence of the *Daphnia* pathogen *Octospora bayeri*: insights into the gene content of a large microsporidian genome and a model for host-parasite interactions. **Genome Biology**, v. 10, n. 10, p. R106, 2009.

CORRADI, N. et al. The complete sequence of the smallest known nuclear genome from the microsporidian *Encephalitozoon intestinalis*. **Nature communications**, v. 1, p. 77, 2010.

CUOMO, C. A et al. Microsporidian genome analysis reveals evolutionary strategies for obligate intracellular growth. **Genome Research**, v. 22 p. 2478–2488, 2012.

DRURY, L. S.; PERKINS, G.; DIFFLEY, J. F. X. The cyclin-dependent kinase Cdc28p regulates distinct modes of Cdc6p proteolysis during the budding yeast cell cycle. **Current Biology**, v. 10, n. 5, p. 231–240, 2000.

EBERT, D. Ecology, epidemiology and evolution of parasitism in *Daphnia*. Disponível em:<<http://www.ncbi.nlm.nih.gov/books/NBK2036>>

HEDGES, S. B. et al. A molecular timescale of eukaryote evolution and the rise of complex multicellular life. **BMC evolutionary biology**, v. 4, p. 2, 2004.

HEINZ, E. et al. The Genome of the Obligate Intracellular Parasite *Trachipleistophora hominis*: New Insights into Microsporidian Genome Dynamics and Reductive Evolution. **PLoS Pathogens**, v. 8, n. 10, 2012.

GOULD, K.L.;BOHNERT, K.A. On the cutting edge: post-translational modifications in cytokinesis. **Trends in cell biology**, v. 21, n. 5, p. 283-292, 2011.

KATINKA, M. D. et al. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. **Nature**, v. 414, n. 6862, p. 450–453, 2001.

KEELING, P. J.; FAST, N. M. Microsporidia: Biology and evolution of highly reduced intracellular parasites. *Annual Review of Microbiology*, v.56, p. 93-116, 2002.

- KEELING, P. J. et al. The reduced genome of the parasitic microsporidian *Enterocytozoon bieneusi* lacks genes for core carbon metabolism. **Genome Biology and Evolution**, v. 2, n. 1, p. 304–309, 2010.
- LARSSON, J. I. R.; EBERT, D.; VÁVRA, J. Ultrastructural study and description of *Ordospora colligata* gen. et sp. nov. (microspora, ordosporidae fam. nov.), a new microsporidian parasite of *Daphnia magna* (Crustacea, Cladocera). **European Journal of Protistology**, v. 33, n. 4, p. 432–443, 1997.
- LI, L. et al. Gene discovery in the Apicomplexa as revealed by EST sequencing and assembly of a comparative gene database. **Genome Research**, v. 13, n. 3, p. 443–454, 2003.
- LIBRADO, P.; ROZAS, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. **Bioinformatics**, v. 25, n. 11, p. 1451–1452, 2009.
- MATHIS, A.; WEBER, R.; DEPLAZES, P. Zoonotic potential of the microsporidia. **Clinical Microbiology Reviews**, v. 18, n. 3, p. 423–445, 2005.
- MELLOR, H.; PARKER, P. J. The extended protein kinase C superfamily. **The Biochemical journal**, v. 332 (Pt 2), p. 281–292, 1998.
- MÈTÈRNIER, G.; VIVARÈS, C. Molecular characteristics and physiology of Microsporidia. **Microbes and Infections**, v.3, p. 407-415, 2001
- NAKJANG, S. et al. Reduction and expansion in microsporidian genome evolution: new insights from comparative genomics. **Genome biology and evolution**, v. 5, n. 12, p. 2285–303, 2013.
- NEI, M.; KUMAR, S. Molecular evolution and phylogenetics. New York: Oxford University Press, 2000. 333p.
- NEI, M.; LI, W. H. Mathematical model for studying genetic variation in terms of restriction endonucleases. **Proceedings of the National Academy of Sciences of the United States of America**, v. 76, n. 10, p. 5269–5273, 1979.
- PEYRETAILLADE, E. et al. Identification of transcriptional signals in *Encephalitozoon cuniculi* widespread among Microsporidia phylum: support for accurate structural genome annotation. **BMC genomics**, v. 10, p. 607, 2009.
- PHAN, Q. T. et al. Role of endothelial cell septin 7 in the endocytosis of *Candida albicans*. **mBio**, v. 4, n. 6, p. 1–9, 2013.
- POMBERT, J. et al. Gain and loss of multiple functionally related, horizontally transferred genes in the reduced genomes of two microsporidian parasites. **Proceedings of the National Academy of Sciences**, v. 109, n. 31, p. 12638–12643, 2012.
- POMBERT, J. et al. The *Ordospora colligata* Genome : Evolution of Extreme Reduction in Microsporidia and Host-To-Parasite Horizontal Gene Transfer. v. 6, n. 1, 2015.

SCHMID-HEMPEL, P. Parasite immune invasion and manipulation of host phenotype. In: SCHMID-HEMPEL, P (ed.). *Evolutionary parasitology: the integrated study of infections, immunology, ecology and genetics*. New York: Oxford University Press, 2011. p. 187-218.

SCHWARTZ, S. et al. Human – Mouse Alignments with BLASTZ. **Genome Research**, v. 13, n. 1, p. 103–107, 2003.

SELMAN, M.; CORRADI, N. Microsporidia: Horizontal gene transfers in vicious parasites. **Mobile Genetic Elements**, v. 1, n. 4, p. 251–255, 2011.

SILVA, B. V. et al. Proteínas quinases: Características estruturais e inibidores químicos. **Química Nova**, v. 32, n. 2, p. 453–462, 2009.

TSAOUSIS, A. D. et al. A novel route for ATP acquisition by the remnant mitochondria of *Encephalitozoon cuniculi*. **Nature**, v. 453, n. 7194, p. 553–556, 2008.

WITTNER, M. Historic perspective on the Microsporidia: Expanding horizons. In: WITTNER, M (ed.). *Microsporidia and microsporidiosis*. Washington D.C: American Society for Microbiology, 1999. p. 1-6.

XU, YANJI & WEISS, L. The microsporidian polar tube: A highly specialised invasion organelle. **National Institutes of Health**, v. 35, n. 9, p. 941–953, 2005.

ZHAO, H. et al. Host-to-Pathogen Gene Transfer Facilitated Infection of Insects by a Pathogenic Fungus. **PLoS Pathogens**, v. 10, n. 4, 2014.

Apêndice 1: Tabela dos genes divergentes com suas respectivas funções e códigos KEGG.

Cromossomo	Ponto Médio	Razão	Gene	Proteína(s)	Função	Código KEGG	Categorias de funções
XI	31350	3.082	ECU11_1200	Transportador ABC (tipo mitocondrial)	Transportador mitocondrial de ATP para o citosol (ABC). Precusores Fe/S	K05663	Transportador
X	1450	3.048	ECU10_0190	Proteína Hipotética YD95	Proteína reguladora do proteassomo 26S.	K03037	Proteassomo
IX	10437	2.917	ECU09_1950	Proteína Hipotética	Proteína de membrana que media o tráfego membranar e mantém a morfologia do RE. Superexpressão leva a morte celular.	K17279	Exossomo Exossomo Biogênese de mRNA
IV	33133	2.885	ECU04_1560	Proteína Nuclear de Ligação GTP	Proteína envolvida na manutenção da organização nuclear	K07936	Proteína cromossomal Biogênese ribossomal Biogênese de tRNA Proteína Quinase Proteína
V	11550	2.873	ECU05_0540	Proteína Quinase da Família PI3/PI4	Subunidade NuA4 do complexo de Histonas Acetiltransferase, conduz a ativação da Transcrição.	K08874	Cromossomal Maquinária de Transcrição
IV	11250	2.854	ECU04_0490	Proteína Ubiquitina Ligase	Subunidade E3 enzimática, responsável pelo reconhecimento de proteínas do substrato, bem como de transferi-las para as próximas fases do processo.	K10587	Enzima Sistema Ubiquitina
XI	36950	2.845	ECU11_1370	Regulador Negativo Geral da Transcrição (Subunidade 1)	Possui múltiplos papéis na regulação dos níveis de mRNA, incluindo regulação na transcrição e na desestabilização de mRNA.	K12604	Biogênese de mRNA
I	8150	2,816	ECU01_0630	Proteína Quinase Serina/Treonina	Proteína Polo-Quinase com múltiplas funções durante a mitose e a citocinese, atuando por meio da fosforilação do substrato.	K06631	Enzima Proteína Quinase Proteína cromossomal
II	6950	2,796	ECU02_0430	Proteína de Ligação tipo SEC1	Transportador de proteínas de membrana.	K15292	Exossomo
VII	9550	2.771	ECU07_0590	Proteína Hipotética YMJ5 CAEEL	Subunidade Essencial na regulação do proteassoma 26S, sendo necessário na sua montagem e na manutenção de sua atividade.	K03036	Proteassomo
VII	12150	2.770	ECU07_0790	Proteína Hipotética	Proteína sem função definida.	K03546	Reparo e Recombinação

I	12750	2,733	ECU01_1060	Reparo de DNA Helicase	Componente da holoenzima RNA polimerase e do Fator de Transcrição TFIIH, tem atividade ATPase dependente. É necessária para desenrolar o DNA promotor, bem como está envolvida no reparo do DNA.	K10843	Enzima Maquinária de Transcrição Reparo e recombinação
II	10950	2.733	ECU02_0550	Proteína Quinase tipo SPK1	Proteína quinase necessária durante o ciclo celular sendo autofosforilada em resposta ao dano do DNA. Desempenha papel importante para iniciar a replicação do DNA.	K08863	Enzima Proteína quinase Proteína cromossomal
VIII	13850	2,706	ECU08_0830	Proteína Ribossomal 60S L4	Componente proteico da subunidade 60S. Participa do processo de tradução.	K02930	Proteína ribossomal
XI	9550	2.686	ECU11_0400	Proteína Hipotética YHY1	Proteína sem função definida.	K05016	Proteína fosfatase Canal de íons
VIII	20850	2.684	ECU08_1100	Complexo Coatomer Subunidade Beta	Beta-proteína essencial no complexo COPI. Está envolvida na manutenção da morfologia normal do RE e das proteínas do Complexo de Golgi.	K17301	Exossomo
VIII	7150	2.666	ECU08_0360	Ornitina Descarboxilase DCOR_yeast	Enzima necessária para a síntese de poliamina, importante para estabilizar a estrutura do DNA.	K07178	Enzima Proteína quinase Biogênese de ribossomos
III	15250	2.664	ECU03_0590	Proteína Hipotética integral de membrana Yipa_yeast	Proteína de membrana necessária para biogênese de RE e para as vesículas transportadoras COPII.	K00102	Enzima
II	23050	2.628	ECU02_1130	Proteína Quinase da família PI3/PI4	Proteína Quinase de checkpoint que garante a integridade do genoma. É necessária como transdutor de sinal em uma parada no ciclo celular em resposta a um dano de DNA ou a sua não replicação. Monitora e participa da recombinação meiótica.	K06640	Enzima Proteína quinase Reparo e recombinação
I	10150	2.628	ECU01_0700	Fator 2 de tradução em Eucariotos	Subunidade Gama do fator de tradução eIFII, está envolvido na identificação do códon de iniciação. Se liga ao GTP, formando um complexo ternário com o ATP e o tRNA-Met.	K03242	Fator de tradução
I	26550	2,624	ECU01_1320	Proteína Quinase C	Proteína envolvida no controle das funções de outras proteínas, através da fosforilação.	K04515	Enzima Proteína quinase
VIII	28750	2,622	ECU08_1440	Proteína Hipotética	Proteína de membrana com função desconhecida. Superexpressão desta proteína suprime a letalidade da deficiência da proteína G-alfa em leveduras.	K14430	Proteína transportadora
XI	50050	2.622	ECU11_1760	Carioferina (Subunidade alfa)	Proteína que forma um dímero com a Carioferina Beta para mediar a importação de proteínas nucleares. Também pode desempenhar papel na regulação e degradação de proteínas.	K15043	Proteína cromossomal

VII	9450	2.615	ECU07_0580	Proteína Hipotética	Proteína sem função definida.	K15732	Enzima Maquinária de transcrição Proteína fosfatase
IV	350	2.597	ECU04_0260	Proteína Hipotética YB3F	Proteína sem função definida.	K03263	Fator de tradução
X	41350	2.589	ECU10_1450	Proteassoma Tipo B (Subunidade cadeia Delta)	Subunidade do proteassoma responsável pela clivagem de resíduos ácidos em peptídeos.	k02738	Proteassomo
XI	9850	2.573	ECU11_0430	Proteína SKT5	Ativador de quitina sintetase III durante o crescimento vegetativo.	K07126	–
V	24050	2.569	ECU05_0990	DNA Polimerase Alfa	Região da DNA Polimerase que possui atividade de alongamento, ligação de DNA e ligação de dNTPs.	K02320	Enzima Replicação do DNA
X	9450	2.567	ECU10_0420	Proteína Carreadora de ADP/ATP	Proteína de membrana que realiza o transporte de ATP para célula.	K03301	Proteína transportadora
VIII	27150	2.567	ECU08_1300	Proteína Carreadora de ATP/ADP	Proteína de membrana que realiza o transporte de ATP para célula.	K03301	Proteína transportadora
XI	14450	2.564	ECU11_0610	Fator Específico de Clivagem e Poliadenilação	Subunidade de clivagem do mRNA e fator de poliadenilação do RNA. Envolvida no reconhecimento do sítio para clivagem e poliadenilação do pré-mRNA.	K14401	Biogênese de mRNA Maquinaria de transcrição
V	2550	2.555	ECU05_0190	Proteína Triptofan 2 (Repetição WD)	Componente pré-ribossomal essencial para a clivagem endonucleotídica. É também um precursor do rRNA, onde elimina possíveis defeitos no ciclo celular e na morfogênese.	K14558	Biogênese de ribossomos
III	31850	2.550	ECU03_1560	Proteína RAS-L42 relacionada com a divisão celular	Proteína essencial para a manutenção da polaridade celular. Em caso de mutação, ocorre defeitos na organização dos filamentos de actina.	K04393	Exossomo Proteína cromossomal Proteína de ligação GTP
IV	1650	2.542	ECU04_0310	Proteassoma 26S Subunidade Reguladora 4	Subunidade que participa do reconhecimento dos ligantes no proteassoma. Contém um domínio de repetição rica em leucina (LRR), onde ocorre a interação proteína-proteína.	K03028	Proteassomo
VIII	23750	2.536	ECU08_1250	Complexo Coatomer Subunidade Beta Principal	Proteína Beta essencial para o complexo COPI, envolvida no transporte do RE para o Complexo de Golgi e de Golgi para o RE. Possui domínios que medeiam as interações seletivas entre cargas.	K17302	Exossomo
VIII	9650	2.515	ECU08_0520	Proteína de Ligação Poliadenilato	Proteína conservada essencial que faz parte da subunidade 90S pré-ribossomal, necessária na produção de rRNA. Possui cinco domínios de ligação com o RNA consenso.	K14787	Biogênese de ribossomos

XI	31650	2.504	ECU11_1230	Proteína Hipotética YDD5	Proteína sem função definida.	K11558	Proteína cromossomal
V	23950	2.492	ECU05_0940	RNA-Helicase subfamília SKI2	Componente do complexo RNA helicase que medeia a degradação do RNA 3-5. Mutantes possuem efeito nulo, levando a um aumento de dsRNA.	K01529	Enzima
VII	25150	2.483	ECU07_1600	Fragmentos de Treonina tRNA Sintetase	tRNA ligado ao aminoácido Treonina. Corrige erros para evitar a incorporação de um aminoácido incorreto ao tRNA.	K00064	Enzima Biogênese de tRNA
III	12050	2.474	ECU03_0520	Proteína de Choque Térmico 70kDa (família HSP 70)	Proteína que se liga a cadeias hidrofóbicas expostas, mantendo a cadeia desenovelada até que ela possa assumir a conformação correta evitando assim, proteínas malformadas.	K03283	Proteossomo Exossomo Biogênese de ribossomos
XI	12550	2.470	ECU11_0450	Fosfatidilinositol 4-Quinase (Região C-Terminal)	Enzima necessária para a manter a morfologia normal do vacúolo, a integridade da parede celular e a organização do citoesqueleto de actina.	K00888	Enzima
XI	51750	2.469	ECU11_1870	Maltose Permease	Proteína de membrana plasmática que participa do processo de difusão da maltose.	K05820	Proteína transportadora
III	15450	2.469	ECU03_0630	Proteína Quinase dependente de Calmodulina	Proteína envolvida no controle das funções de outras proteínas, através da fosforilação. Depende de Calmodulina para iniciar sua ativação.	K04409	Enzima Proteína quinase Proteína cromossomal
IX	9750	2.465	ECU09_1910	Proteína Segregadora cromossomal	Subunidade do complexo multiproteico necessário para coesão das cromátides-irmãs em células mitóticas. É uma proteína necessária durante a recombinação meiótica.	K06669	Proteína cromossomal Proteoglicano
VIII	22150	2.464	ECU08_1180	Proteína Hipotética	Proteína sem função definida.	K11798	Sistema ubiquitina
IV	14850	2.461	ECU04_0760	Fator PRP8 de Splicing do Pré-mRNA	Componente do complexo U4/U5-U6 snRNP, envolvido no segundo passo catalítico de splicing.	K12856	Spliceossomo
I	17750	2.453	ECU01_1230	ATPase	ATPase presente no RE que participa da translocação de proteínas do RE para o citosol para degradação pelo proteossoma.	K13525	Exossomo Biogênese de mRNA
X	34550	2.451	ECU10_1240	Importador Beta 1 (Carioferina)	Complexo que interage com as nucleoporinas, mediando a importação nuclear de proteínas.	K14293	Proteína transportadora
X	41250	2.449	ECU10_1400	Proteína Hipotética	Proteína sem função definida.	K17257	Exossomo
II	15050	2.444	ECU02_0730	Polimerase Poli-A Tipo I	Poli (A) polimerase, um dos fatores necessários para a poliadenilação do mRNA. É necessária também para a exportação do mRNA nuclear.	K14376	Enzima Biogênese de mRNA
XI	17350	2.443	ECU11_0660	Proteína Fosfatase SER/THR (Subunidade Catalítica PPI-1)	Subunidade catalítica envolvida em diversos processos, como metabolismo de glicogênio, esporulação e mitose. Proteína que interage com muitas subunidades reguladoras.	K06269	Enzima Biogênese de mRNA Maquinaria de transcrição

II	23750	2.443	ECU02_1150	Fator de Replicação de DNA da Família MCM (MCM4)	Componente essencial da helicase, se liga ao complexo de pré replicação do DNA e separa a dupla hélice antes da replicação. Acumula-se no núcleo durante a fase G1.	K02212	Enzima Replicação de DNA
VII	9350	2.440	ECU07_0450	Proteína Hipotética de Membrana YAB9	Proteína sem função definida, podendo estar envolvida na desintoxicação celular,	K01196	Enzima
XI	17450	2.433	ECU11_0670	Proteína Hipotética YG22	Fator de Iniciação Eucariótica 2A (eIF), que está associado especificamente com as subunidades 40S e 80S.	K15026	Fator de tradução
XI	51650	2.433	ECU11_1830	Proteína Tipo DNAK (Proteína de Choque da Família HSP70)	Proteína envolvida na translocação e no dobramento de outras proteínas. Responde ao estresse por calor, etanol, entre outros.	K09485	Chaperona
II	17050	2.428	ECU02_0800	Proteína Ribossomal subunidade maior L9e	Componente proteico da subunidade 60S ribossômica.	K02940	Proteína ribossomal
III	9250	2.428	ECU03_0420	Similar a Proteína Yeast CDC68	Subunidade do complexo FACT que auxilia a RNA polimerase II no alongamento e na transcrição, desestabilizando o nucleossoma e remontando sua estrutura.	K01262	Enzima Peptidase
I	27150	2.426	ECU01_1390	Quitina Sintetase I	Catalizador na transferência de N-acetilglucosamina, necessária para síntese da maioria das paredes celulares de quitina. Necessária também para a formação do anel de quitina durante o brotamento das gemas e para a parede dos esporos.	K00698	Enzima Glicosiltransferase
IX	2750	2.425	ECU09_1680	Asparagina-tRNA Ligase	Necessária nas sínteses de proteínas, catalisa a ligação da asparagina ao tRNA.	K01893	Enzima Biogênese de tRNA
VIII	15250	2.417	ECU08_0850	Fator 1 Associado a CCR4	Rnase da superfamília DEDD, subunidade do complexo CCR4 que medeia a desadenilação 3' a 5' do mRNA	K12581	Biogênese de mRNA
VIII	27450	2.415	ECU08_1420	Fragmento similar a Fosfatidilinositol Fosfato Fosfatase	Desfosforila uma série de fosfatidilinositóis. Está envolvida na endocitose. Durante o estresse hiperosmótico provoca translocações para correção dos filamentos de actina.	K01106	Enzima
IV	31450	2.409	ECU04_1400	RNA Polimerase I	Subunidade maior da RNA Polimerase I. Tem por função transcrever as regiões do DNA que contém os genes para o rRNA.	K02999	Enzima Maquinaria de transcrição
X	29050	2.408	ECU10_1060	Fosfoglicerato Mutase	Enzima responsável pela conversão do 3-Fosfoglicerato em 2-Fosfoglicerato. O termo mutase é designado enzimas que catalizam os deslocamentos intramoleculares de grupos funcionais.	K15633	Enzima

X	35350	2.407	ECU10_1320	Fator de Remodelação da Cromatina componente ATPase da segunda subunidade	DNA translocase dependente de ATP, envolvida na remodelação da cromatina. Complexo necessário para repressão de genes específicos durante o início da meiose.	K11654	Enzima Proteína comossomal Maquinaria de transcrição
VIII	650	2.394	ECU08_0120	Domínio de GTP	Domínio responsável pela hidrólise de GTP para GDP sem a necessidade de proteínas acessórias.	K12013	Sistema ubiquitina
VIII	13550	2.391	ECU08_0730	Proteína RAS- relacionada com Rab5	GTPase necessária para o transporte durante a endocitose e para a correta separação das hidrolases vacuolares.	K07889	Exossomo Proteína de ligação GTP
V	7750	2,385	ECU05_0370	Proteína Histona de Ligação N1/N2	Atua na compactação do DNA.	K11291	Proteína cromossomal Maquinaria de transcrição
IV	1550	2,382	ECU04_0280	Supressor da Proteína "LOOP" I	Componente da RNA polimerase, é essencial para a reparação e excisão de nucleotídeos	K03142	Reparo e recombinação
VIII	550	2,382	ECU08_0100	Proteína Hipotética	Proteína sem função definida.	K06071	Enzima Proteína quinase
VIII	29350	2.379	ECU08_1480	Proteína Quinase Relacionada com SNF1	Proteína necessária para a fosforilação do repressor transcricional Mig1.	K08798	Enzima Proteína quinase Proteína de citoesqueleto
V	27650	2.375	ECU05_1190	Superfamília de transportadores ABC	Regulador positivo da atividade da quinase Gcn2p, estimulando sua ativação.	K06158	Fator de tradução
VII	12350	2.372	ECU07_0800	Gliceraldeído 3-Fosfato Desidrogenase	Enzima envolvida na glicólise e na gliconeogênese. Trata-se de um tetrâmero, detectado no citoplasma e na parede celular, que catalisa a reação de Gliceraldeído 3-Fosfato a 1,3 Bisfosfoglicerato.	K00134	Enzima Exossomo
III	29850	2.365	ECU03_1460	Histona H3	Atua na compactação do DNA, auxiliando na formação do octâmero central do nucleossomo.	K11253	Exossomo Proteína cromossomal Maquinaria de transcrição
IV	31650	2.360	ECU04_1440	Fator de Início da Transcrição (TFII-I)	Proteína de ligação TATA. Fator geral de transcrição, que interage com os demais fatores para formar o complexo de pré-iniciação nos promotores.	K03120	Fator de transcrição
XI	26350	2.359	ECU11_1050	Transportador Oligopeptídico	Peptideo transportador integral de membrana. Medeia o transporte de di e tripeptídeos.	K03305	Proteína transportadora
VII	14850	2.358	ECU07_1010	Proteína de Reconhecimento Estrutural Específica	Subunidade do complexo heterodimérico FACT que facilita o alongamento transcricional pela RNA Pol II. Desestabiliza e remonta a estrutura do nucleossomo, interagindo com a DNA Polimerase Alfa.	K09272	Fator de transcrição

XI	29150	2.352	ECU11_1160	Possível Proteína Nuclear Scaffold	Subunidade do complexo promotor da anáfase, sendo uma ubiquitina ligase necessária para a degradação dos inibidores de anáfase (incluindo as ciclinas mitóticas).	K03350	Proteína cromossomal Proteína fosfatase Sistema ubiquitina
IV	14550	2.352	ECU04_0710	Indutor de Fosfatase com fase M	Proteína tirosino-fosfatase envolvida no controle do ciclo celular. Regula o estado de fosforilação de Cdc28p.	K02555	Enzima Proteína fosfatase
II	11150	2.348	ECU02_0570	Proteína Hipotética	Proteína sem função definida.	K14403	Enzima Maquinaria de transcrição
V	1250	2.337	ECU05_0060	Proteína tipo Quinesina	Proteína Carreadora, com importante função na formação dos fusos meióticos e mitóticos.	K10395	Biogênese de mRNA Proteína de citoesqueleto
VIII	17150	2.337	ECU08_0910	Proteína Hipotética	Proteína sem função definida.	K02051	Proteína transportadora
III	16550	2.332	ECU03_0680	6-Fosfofrutoquinase isoenzima 1	Subunidade envolvida na glicólise, indispensável para o crescimento anaeróbico, ativada pela frutose 2-6 bisfosfato e AMP. Mutação inibe a indução da glucose para genes relacionados com ciclo celular.	K00850	Enzima Proteína fosfatase Biogênese de mRNA
VII	6350	2.332	ECU07_0380	Exonuclease Putativa	Exonuclease necessária para a maturação da extremidade 3' do rRNA.	K14570	Enzima Biogênese ribossomal
VIII	20650	2.325	ECU08_1090	Similar a Proteína Hipotética YB36	Subunidade necessária para a modificação de nucleotídeos. Apresenta atividade de histona acetiltransferase.	K07739	Enzima Proteína cromossomal Biogênese de tRNA
III	22350	2.318	ECU03_0950	Proteína Ribossomal 60S L7e	Componente proteico da subunidade 60S. Apresenta um domínio de ligação altamente conservado	K02937	Proteína ribossomal
VII	21450	2.315	ECU07_1440	Similar a Proteína Hipotética Humana C215	Proteína sem função definida, mas essencial para a viabilidade celular. Pode estar envolvida no estabelecimento da polaridade celular e na morfogênese.	K17710	Biogênese de tRNA Biogênese mitocondrial
VII	11450	2.314	ECU07_0760	Proteína DNAj	Chaperona envolvida na regulação de HSP90 e HSP70. Também é responsável pela translocação de proteínas através da membrana.	K09503	Chaperona Biogênese mitocondrial
X	14050	2.314	ECU10_0620	Mediador "nonsense" de mRNA (proteína 5)	Proteína transportadora, envolvida na importação nuclear de proteínas.	K18755	Biogênese de mRNA
VIII	8550	2.313	ECU08_0490	Cisteína tRNA Sintetase	Proteína que pode interagir com ribossomos, com base em experiências de co-purificação.	K01883	Enzima Biogênese de tRNA
XI	39050	2.312	ECU11_1430	Diptina Sintase	Resíduo de histidina modificado pelo fator de alongação 2.	K00586	Enzima Fator de tradução

V	20750	2.310	ECU05_0860	6-Fosfogluconato desidrogenase	Catalisa a reação de regeneração do NAPH na via das pentoses.	K00033	Enzima
V	19950	2.307	ECU05_0820	Sintaxina (Homologo)	Proteína envolvida na especificidade de tráfego celular e na catálise dos processos de fusão.	K12559	Proteína de citoesqueleto
V	15650	2.303	ECU05_0730	Serina Palmitoil Transferase Subunidade 2	Componente da serina palmitoil responsável pela primeira etapa na síntese de esfingolipídios que é a condensação da serina palmitoil com a Co-A	K00654	Enzima
IV	16450	2.298	ECU04_0850	Fator MCM2 de Replicação do DNA	Componente do complexo hexamérico MCM-2 que se liga a cromatina como parte do complexo pré-replicativo.	K02540	Enzima Proteína cromossomal Replicação de DNA
V	14150	2.297	ECU05_0670	Proteína Ribossomal 40S S6	Componente proteico da subunidade 40S ribossomal, idêntica a Rps6Ap.	K02991	Proteína ribossomal
VII	25550	2.293	ECU07_1610	Proteína Hipotética YM10	Proteína que modifica o pré-tRNA no U17.	K05542	Enzima Biogênese de tRNA
V	15850	2.281	ECU05_0780	Fator de Replicação de DNA da Família MCM	Proteína envolvida na replicação do DNA sendo componente do complexo hexamérico Mcm2-7 que se liga a cromatina, como parte do complexo pré-replicativo.	K02542	Enzima Replicação de DNA
I	20450	2.281	ECU01_1250	Proteína Hipotética	Proteína sem função definida.	K17764	Biogênese mitocondrial
III	25750	2.279	ECU03_1330	Proteassoma 26S Subunidade regulatória T1	Uma das seis ATPases do proteassoma 26S que degrada substratos de ubiquitina. Elemento necessário para transcrição.	K03061	Proteassoma
II	17250	2.278	ECU02_0840	Proteína Hipotética	Proteína sem função definida.	K10362	Proteína de citoesqueleto
XI	22050	2.277	ECU11_0800	Fator de Replicação do DNA MCM4	Complexo heterohexamérico que se liga aos complexos de pré-replicação do DNA. Acumula-se no núcleo durante a fase G1.	K10373	Exossomo Proteína de citoesqueleto
VIII	32050	2.274	ECU08_1550	Proteína Hipotética	Proteína sem função definida.	K06564	Exossomo Lectina Antígeno celular
VIII	33750	2.268	ECU08_1640	Fator Geral de Transcrição	Proteínas envolvidas no início da transcrição (TATA), contendo os promotores. Associa-se ao fator de transcrição TFIID.	K11684	Proteína cromossomal
IV	8350	2.267	ECU04_0450	Proteína Ribossomal 40S AS ou P40	Subunidade menor ribossomal, necessária para a maturação do rRNA- subunidade 18S. Mutação no gene reduz a taxa de crescimento e sua deleção é fatal.	K02928	Proteína ribossomal
III	14250	2.267	ECU03_0560	Proteína Hipotética	Proteína Ativadora de GTPase.	K18469	-
X	21750	2.264	ECU10_0870	Glicerol-3-Fosfato Desidrogenase Mitocondrial	Sua expressão é reprimida por glicose e cAMP e é ativada por uma fonte de carbono não fermentável.	K00111	Enzima

IV	8650	2.256	ECU04_0460	Dimetil-adenosina Transferase	Envolvida no processamento do RNA pré-ribossomal, por meio da dimetilação da extremidade 3' terminal da subunidade 18S do rRNA.	K14191	Enzima Biogênese ribossomal
XI	38950	2.256	ECU11_1420	Proteína de Choque HSP101	Proteína de resposta ao estresse por calor que redobra e reativa proteínas desnaturadas. Trabalha juntamente com HSP40 e HSP70.	K03695	Proteassomo Exossomo Biogênese ribossomal
V	4950	2.254	ECU05_0310	Acetil coenzima A Sintetase	Acetil-CoA sintetase que, juntamente com Acs2p, é a fonte nuclear de acetil-CoA para a acetilação de histonas. É expressa durante o crescimento em fontes de carbono não fermentável e em condições aeróbias.	K01895	Enzima Biogênese de proteínas lipídicas
VII	13750	2.253	ECU07_0900	Proteína Hipotética YM60	Aciltransferase envolvida no metabolismo de lipídios e na detoxificação.	K07019	Enzima
XI	12650	2.253	ECU11_0540	Proteína de Choque HSP70	Proteína envolvida na translocação e no dobramento de outras proteínas. Responde ao estresse por calor, etanol, entre outros.	K03283	Exossomo Proteína ribossomal Spliceossoma
X	6850	2.249	ECU10_0320	Proteína Quinesina Tipo A	Localiza o corpo polar do fuso durante a mitose e a meiose. É necessária durante a fusão nuclear.	K10405	Proteína cromossomal Proteína de citoesqueleto
IX	5350	2.246	ECU09_1790	ATP Sintase	Subunidade ATPase vacuolar, localizada no complexo de Golgi e nos endossomas.	K02154	Enzima
IV	16350	2.237	ECU04_0780	Fator 6 de Início da Tradução Eucariótica	Envolvida na biogênese e na estabilidade de subunidades ribossomais 60S.	K03264	Biogênese ribossomal Fator de tradução
VIII	3850	2.234	ECU08_0240	Proteína de ligação GTP	GTPase envolvida na biogênese da subunidade 60S ribossomica. Necessária para a liberação de Nmd3p da subunidade 60S para o citoplasma	K14539	Enzima Proteína ribossomal
I	17650	2.229	ECU01_1200	Proteína Hipotética	Proteína sem função definida.	K10866	Enzima Reparo e recombinação
VIII	30150	2.225	ECU08_1510	Proteína Hipotética	Proteína sem função definida.	K14852	Biogênese ribossomal
II	20150	2.223	ECU02_0930	Supressor de Proteína Bifurcada (forked)	Componente envolvido na clivagem e na poliadenilação do mRNA (extremidade 3'). Matém as pontes de interação entre o Rna15p e Hrp1p.	K14408	Biogênese de mRNA
X	45850	2.223	ECU10_1700i	Proteína Vacuolar Tipo Dinamina	GTPase necessária para triagem vacuolar. Está envolvida na organização do citoesqueleto de actina e na retenção de algumas proteínas.	K17065	Enzima Exossomo Biogênese mitocondrial
II	16650	2.222	ECU02_0780	Fator de Início de Transcrição TFIIIF Subunidade Beta	Fator basal de Transcrição	K03139	Enzima Maquinaria de transcrição

II	8750	2.221	ECU02_0480	Proteassoma 26S subunidade Regulatória N2	Subunidade do proteassoma 26S que é substrato para N-acetiltransferase Nat1p	K03032	Proteassomo
XI	47150	2.215	ECU11_1700	Proteína Hipotética Transmembrana	Proteína sem função definida.	K08869	Proteína quinase
VIII	16850	2.214	ECU08_0890	Proteína tipo Quinesina	Proteína envolvida no transporte de organelas.	K10400	Proteína de citoesqueleto
I	16550	2.209	ECU01_1170	Proteína Hipotética	Proteína de pequena condução de canal mecano-sensível	K16052	Proteína transportadora
XI	9150	2.208	ECU11_0390	Proteína Hipotética	Proteína que auxilia nos processos de comunicação celular.	K06110	
VIII	32450	2.207	ECU08_1570	Proteassoma 20S Componente C3	Componente responsável por unir as subunidades ribossomais 40S e 60S. Regula também o início da tradução	K02866	Proteassomo Biogênese ribossomal
VII	19450	2.206	ECU07_1220	Similar a Formina	Regulador de diversos processos celulares incluindo a apoptose.	K04512	Exossomo
III	24250	2.205	ECU03_1230	Proteína Periódica Tritofan 1	Proteína associada ao complexo CCR4-NOT, desempenhando papel na fusão mitocondrial.	K14791	Biogênese ribossomal
IV	29050	2.205	ECU04_1210	Proteína de Degradação da Fusão de Ubiquitina	Fator de montagem da cadeia E4 de Ubiquitina, que coopera com a enzima ativadora de Ubiquitina E1, com a enzima de conjugação E2 e com a com Ubiquitina Ligase E3 afim de conjugar a Ubiquitina nos substratos.	K10587	Sistema ubiquitina
X	11550	2.204	ECU10_0520	Proteína Carreadora de ATP/ADP	Proteína de membrana que realiza o transporte de ATP para célula.	K03301	Proteína transportadora
V	19150	2.201	ECU05_0800	Proteína de Ligação GTP da Família GTP/OBG	GTPase que associa-se a subunidade 60S, sendo necessária para a biogênese desta subunidade.	K06943	Proteína de ligação GTP Biogênese ribossomal
XI	2850	2.197	ECU11_0240	Exportina 1	Carioferina envolvida na exportação de proteínas, RNAs ribossomais.	K14290	Biogênese de mRNA Biogênese ribossomal
IV	6750	2.194	ECU04_0350	DNA topoisomerase II	Alivia a tensão na torção do DNA durante a clivagem. Cliva as cadeias complementares e localiza os núcleos axiais na meiose.	K03164	Enzima Proteína cromossomal Replicação do DNA
V	28050	2.193	ECU05_1510	Caseína Quinase II (Cadeia alfa)	Tem papel no crescimento e na proliferação celular.Seus substratos contém todos os fatores de transcrição e RNAs polimerase.	K03097	Enzima Proteína quinase Biogênese ribossomal
VII	17450	2.192	ECU07_1190	Cadeia Alfa Tubulina	Principal responsável pela formação do microtubulo.	K07374	Exossomo Biogênese de mRNA Proteína cromossomal

XI	53750	2.191	ECU11_1920	Proteína de Membrana Endoplasmática Degradadora de Proteínas Deformadas do Lúmem	Proteína cuja função é degradar outras proteínas mal formadas no lúmem.	K13989	-
X	26850	2.189	ECU10_1000	Proteína Quinesina	Proteína envolvida no transporte de organelas.	K11498	Proteína cromossomal Proteína fosfatase Enzima
III	19850	2.188	ECU03_0900	Proteína Vacuolar (Família SEC1)	Proteína envolvida no tráfego de vesículas entre RE e o Complexo de Golgi. Estimula o complexo membrana-proteína trans-SNARE.	K03763	Reparo e recombinação Replicação de DNA
X	6950	2.183	ECU10_0340	Pirofosfatase Inorganica	Promove as trocas rápidas de Pi com a água. Altamente expressa, é essencial para a viabilidade celular.	K01507	Enzima
VIII	37550	2.182	ECU08_1970	Proteassoma 26S Subunidade Reguladora 6	Uma das seis ATPases do proteassoma 26S que degrada substratos de ubiquitina. Substrato para N-acetiltransferase B.	K03063	Proteassomo
II	14250	2.180	ECU02_0700	Proteína Hipotética	Proteína sem função definida.	K14832	Biogênese ribossomal
VIII	30550	2.177	ECU08_1520	ATP Vacuolar Sintase Subunidade AC39	Domínio responsável pela bomba de prótons. É encontrado no sistema endomembrana, onde é necessário para a montagem do domínio VI.	K02146	Enzima
VIII	12750	2.177	ECU08_0670	Cadeia Gama Tubulina	Envolvida da montagem dos microtúbulos e do corpo polar do fuso.	K10389	Proteína cromossomal Proteína de citoesqueleto
X	950	2.175	ECU10_0160	Proteína Ribossomal 60S L23A	Componente proteico da subunidade 60S. Participa do processo de tradução.	K02893	Proteína ribossomal
X	20350	2.175	ECU10_0760	Fator de Início da Transcrição TFIID Subunidade 11kDa	Possui atividade de histona acetiltransferase, estando envolvida na ligação do promotor G1/S.	K03125	Proteína cromossomal Maquinaria de transcrição
XI	44450	2.173	ECU11_1580	Transportador ABC Hipotético Y719	Proteína essencial para a biogênese do ribossomo e início da tradução. Facilita a ligação do complexo multifatorial (MEC) e de fatores de início de tradução.	K06174	Biogênese ribossomal Proteína transportadora
II	28836	2.172	ECU02_1230	Similar a Proteína SEC24	Componente do complexo heterodimérico de revestimento da vesícula COPII. Está envolvida no transporte de vesículas de RE para Golgi	K14007	Proteína transportadora
XI	29050	2.172	ECU11_1100	Isoleucina tRNA Sintetase	Enzima citoplasmática inibidora de alvos específicos de G1.	K01870	Enzima Biogênese de tRNA
XI	41750	2.169	ECU11_1460	Fator 2 de Alongamento da Tradução	Catalisa a translocação ribossomal durante a síntese de proteínas.	K03234	Exossomo Fator de tradução

IV	10850	2.168	ECU04_0480	Transportador ABC	Transporte de substratos através da membrana plasmática.	K06147	Proteína transportadora
IV	19250	2.165	ECU04_0910	RNA Helicase dependente de ATP	Envolvida no processamento e na degradação do RNA nuclear, além de ser um dos componentes do complexo TRAMP e membro da família DEAD-box helicases.	K12598	Enzima Biogênese de mRNA Biogênese ribossomal
IV	7950	2.164	ECU04_0420	Proteassoma 26S Subunidade Reguladora 12	Subunidade essencial na regulação do proteassoma 26S.	K03038	Proteassomo
III	31650	2.162	ECU03_1530	Helicase MOT1	Proteína essencial envolvida na regulação da transcrição. Remove Spt15p do DNA a partir de sua atividade C-terminal, formando um complexo que se liga a TATA com elevada afinidade.	K15192	Enzima Maquinaria de transcrição
I	8650	2.162	ECU01_0670	Proteína Hipotética	Subunidade do fator de reparação TFIIH e excisão de nucleotídeos, envolvidos no início da tatranscrição. Necessário para o reparo por excisão de nucleotídeos.	K03144	Maquinaria de transcrição Reparo e recombinação
X	15050	2.161	ECU10_0630	Proteína 1 do Complexo T (subunidade ETA)	Subunidade citosólica do complexo chaperonina Cct, necessária para a montagem de actinas e tubulinas.	K09499	Exossomo Chaperona
X	43950	2.160	ECU10_1680	RNA Helicase dependente de ATP (Família DEAD box)	Possui papel na síntese de rRNA 18S.	K14776	Enzima Biogênese ribossomal
II	7350	2.157	ECU02_0450	Fator de Início de Transcrição TFIIA Subunidade menor	Fator de Transcrição específico da RNA Polimerase II que interage com a Proteína de Ligação a TATA-box.	K03123	Maquinaria de transcrição Fator de transcrição
IV	22050	2.156	ECU04_1000	Cadeia Pesada de Miosina	Proteína envolvida no transporte de vesículas e organelas.	K10362	Proteína de citoesqueleto
VIII	8250	2.155	ECU08_0380	Fator de Estímulo de Clivagem	Proteína citoplasmática de ligação de RNA, tem papel importante na tradução do mRNA, promovendo interações alélicas com os genes que codificam as proteínas envolvidas no início do processo de tradução.	K14407	Biogênese de mRNA
X	29550	2.153	ECU10_1110	Proteína 2 de Ligação Poliadenilato	Parte do complexo de processamento de RNA 3'. Está envolvida no controle do comprimento da cauda de Poli(A) e interage com o fator eIF-4G.	K13126	Biogênese de mRNA Spliceossomo
X	1850	2.149	ECU10_0240	Proteína do Complexo-T1 (subunidade Gama)	Subunidade do complexo citosólico chaperonina Cct necessário para montagem de actinas e tubulinas.	K09495	Exossomo Chaperona
VIII	6850	2.146	ECU08_0290	Fator de Replicação de DNA da família MCM3	Proteína envolvida no complexo de replicação do DNA, componente do complexo hexamérico McM que se liga a cromatina como parte do complexo pré-replicação.	K02541	Enzima Proteína cromossomal Replicação de DNA

V	5950	2.144	ECU05_0340	Proteína Hipotética	Proteína sem função definida. Homólogo da proteína humana BAP31.	K14009	–
X	39050	2.140	ECU10_1380	Ubiquitina Ligase	Possui papel na exportação do mRNA do núcleo, regulando a transcrição de coativadores. Está envolvida na degradação de excesso de histonas.	K10592	Enzima Sistema ubiquitina
V	12750	2.140	ECU05_0580	Transportador de Aminoácidos	Transportador putativo relacionado com os transportadores vesiculares GABA-glicina.	K14993	Proteína transportadora
X	20050	2.140	ECU10_0750	Metionina Aminopeptidase II	Catalisa a remoção da metionina N-terminal a partir de polipeptídeos nascentes.	K01265	Enzima Peptidase
V	25550	2.138	ECU05_1010	Proteína de Repetição Beta-Transducina	Proteína nuclear essencial envolvidos nas etapas iniciais da biogênese do ribossomo. Interage fisicamente com a proteína ribossômica Rpl3p.	K14848	Biogênese de ribossomo
V	12850	2.136	ECU05_0630	Proteína Quinase SER/THR	Proteína Polo-Kinase com múltiplas funções durante a mitose e a citocinese, atuando por meio da fosforilação do substrato.	K08867	Enzima Proteína quinase Proteína fosfatase
VII	250	2.133	ECU07_0090	RNA polimerase III 160k (Subunidade Maior)	Componente do núcleo da RNA polimerase III, semelhante a subunidade bacteriana beta-prime.	K03018	Enzima Maquinaria de transcrição
IX	2950	2.134	ECU09_1690	Proteína Hipotética	Proteína sem função definida.	K11548	Proteína cromossomal
VIII	10050	2.133	ECU08_0550	Arginina tRNA Sintetase	Proteína citoplasmática também encontrada em estudos com mitocôndrias altamente purificadas.	K01887	Enzima Biogênese de tRNA Biogênese mitocondrial
XI	55050	2.132	ECU11_2000	Proteína Nuclear da Família SMC	Proteína de manutenção estrutural dos cromossomos (SMC). Subunidade essencial do complexo Mms21-SMC5-SMC6, necessária para o crescimento e a manutenção do DNA.	K03529	Proteína cromossomal
II	550	2.130	ECU02_0100	Proteína HSP70	Chaperona associada ao ribossomo envolvida no dobramento de cadeias polipeptídicas recém-formadas.	K09490	Exossomo Chaperona Proteína transportadora
XI	54950	2.130	ECU11_1950	Homólogo a Septina CDC10	Proteína de controle da divisão celular.	K16994	Exossomo
VIII	2450	2.126	ECU08_0160	Subunidade Beta de Calcineurina	Proteína fosfatase que regula Crz1p, fator de transcrição em resposta ao estresse.	K17610	Proteína fosfatase
V	3450	2.125	ECU05_0270	Glicerol-3-Fosfato Desidrogenase	Enzima chave para a síntese de glicerol, essencial para o crescimento sob estresse osmótico.	K00006	Enzima

III	20550	2.124	ECU03_0910	Homólogo da Proteína Quinase Caseína 1	Proteína envolvida na transdução de sinal intracelular.	K14758	Enzima Proteína quinase Biogênese ribossomal Biogênese de tRNA
XI	43050	2.123	ECU11_1560	Proteína da Família Epimorfina	Família de proteínas que promove o polaridade e o correto arranjo celular.	K08486	Exossomo
X	9850	2.121	ECU10_0440	RNA -Helicase dependente de ATP	Proteína envolvida no processamento de rRNA, necessária para a maturação e clivagem do pré-rRNA.	K14777	Enzima Biogênese ribossomal
VII	17150	2.119	ECU04_1130	Manosiltransferase da família KTR	Enzima envolvida na glicosilação de proteínas N-ligadas. Proteína de Membrana tipo II	K10967	Enzima Glicosiltransferase
X	11750	2.117	ECU10_0550	Proteassoma Subunidade Alfa C6	Pode substituir a subunidade alfa 3, sob condições de estresse, para formar uma isoforma mais ativa.	K02730	Enzima Peptidase Proteassomo
III	2550	2.117	ECU03_0220	Proteína do Complexo-T1 Subunidade Alfa	Medeia o dobramento das proteínas no citosol, está envolvida também na manutenção dos filamentos de actina no citoesqueleto.	K09493	Exossomo Chaperona
VIII	3350	2.116	ECU08_0210	Proteína tipo SEC31	Fosfoproteína essencial para formação de vesículas de transporte derivado do RE.	K14005	Proteína fosfatase
VII	22750	2.115	ECU07_1500	ATPase Transportadora de Cátions	Proteína vacuolar com possível papel no sequestro de materiais pesados.	K14951	Enzima
X	18850	2.111	ECU10_0690	RNA Helicase dependente de ATP (Família DEAD box)	Possui papel na síntese de rRNA 18S.	K14809	Enzima Biogênese ribossomal
XI	20850	2.103	ECU11_0790	Proteína Homóloga ao Transportador SEC23 (COPII COAT)	Proteína GTPase de ativação. Proteína envolvida no transporte de RE para Golgi e na autofagia.	K14006	Proteína transportadora
IV	21850	2.102	ECU04_0980	Partícula de Reconhecimento de Sinal (Subunidade 54kDa)	Subunidade que contém um domínio GTPase que medeia a ligação do receptor de sinal SRP.	K03106	
XI	42950	2.100	ECU11_1510	Transportador de Zinco	Envolvido no transporte de íons de zinco para a célula.	K14689	Proteína transportadora
I	4950	2.098	ECU01_0410	Proteína Hipotética	Proteína sem função definida.	K15152	Maquinaria de transcrição
X	38250	2.098	ECU10_1350	Proteína Hipotética Y047	Endorribonuclease, subunidade do complexo de clivagem de mRNA. Necessário para o processamento, o splicing e terminação de transcrição de mRNA.	K13148	Enzima Spliceossomo
III	12750	2.094	ECU03_0540	Proteína de reparo de DNA	Heterodímero que se liga ao DNA para iniciar o processo de reparo.	K08735	Reparo e recombinação
X	46150	2.092	ECU10_1740	Proteína com domínio calponina	Proteína essencial necessária para definir o padrão de brotamento, promove a localização de marcadores axiais e localiza o anel contrátil durante a anáfase.	K05767	Exossomo Proteína de citoesqueleto
X	22550	2.092	ECU10_0900	Proteína Hipotética Y162	Proteína sem função definida.	K14403	Enzima

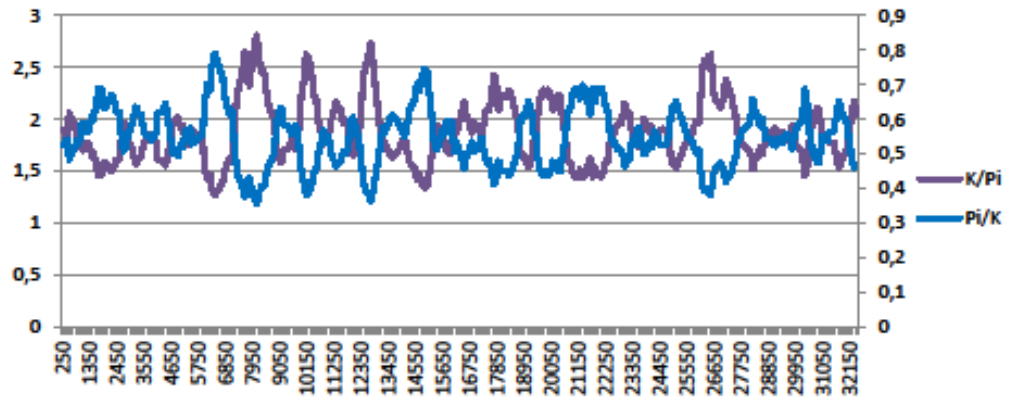
VIII	33350	2.090	ECU08_1620	Proteína Quinase SER/TRE	Proteína Polo-Quinase com multiplas funções durante a mitose e a citocinese, atuando por meio da fosforilação do substrato.	K03114	Biogênese de mRNA Maquinaria de transcrição Enzima Proteína quinase Proteína cromossomal
III	29450	2.088	ECU03_1410	Proteína de Fusão Vesicular Tipo SEC-18	ATPase necessária para a liberação de SEC17p durante a etapa de fusão vacuolar e para o transporte de RE para o Complexo de Golgi	K06027	Enzima
V	21050	2.086	ECU05_0920	Proteína Ribossomal 40S S9	Componente proteico da subunidade 40S ribossomal, idêntica a Rps9Ap.	K02997	Proteína ribossomal
XI	35150	2.086	ECU11_1350	Fator 4E de Início da Tradução Eucariótica	Fator de reconhecimento do complexo de tradução (eIF-1; eIF-1A e eIF-3 associados a subunidade 40S ribossomal) .	K03259	Biogênese de mRNA Fator de tradução
III	23550	2.080	ECU03_1220	Proteína Ribossomal 60s L3e	Envolvido na replicação e manutenção do rRNA.	K02925	Proteína ribossomal
XI	27050	2.079	ECU11_1080	Manganês Superóxido Dismutase	Enzima antioxidante celular, participa da defesa contra reações oxidativas.	K04564	Enzima
VI	850	2.079	ECU06_0150	Componente do complexo 20S promotor de Ciclossomo/Anáfase	Controle do ciclo celular. Destroi a coesão entre as cromátides irmãs e inicia a fase da anáfase.	K03353	Sistema ubiquitina Proteína cromossomal
VI	6850	2.078	ECU06_0450	Proteína Hipotética	Proteína sem função definida.	K03253	Proteína fosfatase Fator de tradução Biogênese de mRNA
XI	22250	2.077	ECU11_0810	Proteína NSA2 de Biogênese Ribossomal	Proteína constituinte das particulas pré-ribossomais 66S que contribui para a formação do pré-RNA 27S.	K14842	Biogênese ribossomal
VI	7250	2.076	ECU06_0510	Proteína do Complexo T Subunidade Zeta	Proteína essencial para montagem de filamentos de actina e tubulinas.	K09498	Exossomo Chaperona
VI	8250	2.075	ECU06_0600	RNA Polimerase II Terceira Cadeia	Terceira maior subunidade da RNA polimerase, sendo parte do núcleo central. Semelhante a subunidade alfa procariótica	K03011	Maquinaria de transcrição Reparo e recombinação
IV	29850	2.074	ECU04_1290	RNA Helicase dependente de ATP (Família SK12)	Envolvida no processamento e na degradação do RNA nuclear, além de ser um dos componentes do complexo TRAMP e membro da família DEAD-box helicases.	K12599	Biogênese de mRNA
VI	9350	2.074	ECU06_0630	Proteína Hipotética	Proteína sem função definida.	K02073	Proteína transportadora
IV	24150	2.070	ECU04_1020	Proteína 1 do Complexo T (Subunidade Theta)	Subunidade do complexo chaperonina Cct é necessária na montagem de actinas e tubulinas.	K09500	Chaperona

X	41950	2.068	ECU10_1460	Glutamina tRNA Sintetase	Cataliza a glutaminação específica de tRNA (Glu)	K01886	Enzima Biogênese de tRNA
VI	9550	2.065	ECU06_0720	Proteína Hipotética YIP3	Ativador da fosfatase PP2A regulando a progressão da fase G1.	K17605	Proteína fosfatase
VI	10350	2.064	ECU06_0730	Ribonucleosídeo Difosfato Redutase - Cadeia Pequena	Catalisador da atividade de síntese de dNTPs, é regulado pela replicação de DNA e pelas vias de reparo.	K10808	Enzima Reparo e recombinação
VI	11250	2.062	ECU06_0790	Asparatil tRNA Sintetase	Enzima catalisadora que se autorregula, por meio de seu próprio mRNA.	K01879	Enzima Biogênese de tRNA
VI	11650	2.061	ECU06_0880	Homólogo a culina	Proteína que participa de modificações pós-traducionais.	K03347	Sistema ubiquitina
X	31950	2.057	ECU10_1210	DNA Polimerase Epsilon	Subunidade catalítica da DNA polimerase, exibindo atividade de exonuclease. Também está envolvida na síntese de DNA.	K02324	Enzima Reparo e recombinação Replicação de DNA
VI	12350	2.056	_	Intron no gene L5 (ECU06_0990i)	Intron do gene ECU06_0990i que expressa a proteína 1 do complexo T (Subunidade epsilon).		-
VI	12450	2.054	ECU06_0900i	Proteína Ribossomal 60S Subunidade L5	Componente ribossomal que se liga ao rRNA 5S, sendo necessário na montagem da subunidade 60S.	K02932	Proteína ribossomal Proteína fosfatase
VI	13650	2.049	ECU06_0930	Fosfolipídio transporador ATPase	Fosfolipídio que mantém a membrana lipídica ssimétrica.	K14802	Proteína ribossomal Proteína fosfatase
XI	750	2.046	ECU11_0180	Proteína Quinase SER/THR (Subfamília MNB/DYRK)	Sistema de detecção de glicose envolvido com o controle do ciclo celular em resposta a disponibilidade de glicose no meio.	K18670	Enzima Proteína quinase Proteína cromossomal
X	26350	2.036	ECU10_0920	Ribonucleosídeo Difosfato Redutase	Uma das duas subunidades reguladoras maiores do complexo RNR. Catalisa a síntese de dNTP, regulada pela replicação de DNA. Vias de checkpoint são reguladas pelas subunidades menores.	K10807	Enzima Reparo e recombinação
VI	15750	2.033	ECU06_0980	Proteína Hipotética	Proteína sem função definida.	K03547	Reparo e recombinação
XI	34950	2.027	ECU11_1290	Proteína Quinase de Ciclo Celular (Família CDC2/CDKX)	Quinase dependente de ciclina que auxilia na regulação do ciclo celular.	K08818	Enzima Proteína quinase Spliceossomo
VI	16650	2.024	ECU06_0990	Proteína 1 do Complexo T - Subunidade Epsilon	Subunidade do complexo citosólico chaperonina Cct necessário para montagem de actinas e tubulinas.	K09497	Exossomo Chaperona

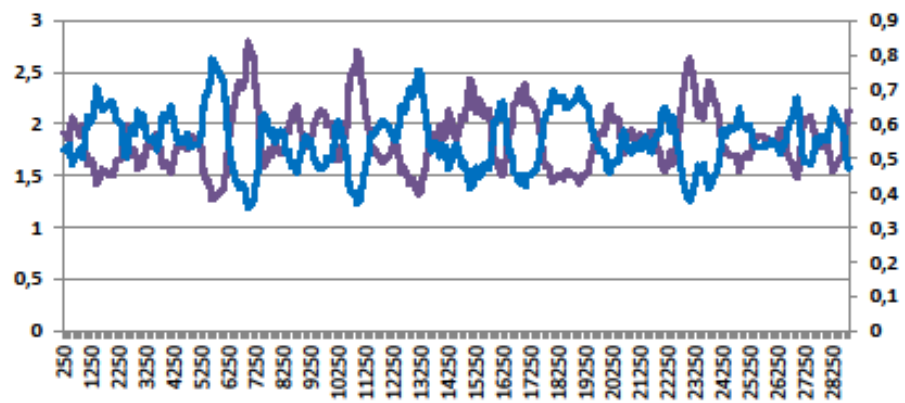
X	31750	2.023	ECU10_1180	Proteína 1 de Degradação da Fusão de Ubiquitina	Proteína que interage com Cdc48p e Npl4p. Envolvida no reconhecimento de proteínas para degradação pelo proteossoma 26S e no transporte de RE para o citosol.	K14016	Sistema ubiquitina
VI	18650	2.020	ECU06_1020	Fator de Transcrição tipo MYB	Família de genes envolvida na proliferação, diferenciação e apoptose celular.	K09422	Fator de transcrição
VII	15050	2.017	ECU07_1040	Proteossoma 26S Subunidade Alfa 4	Subunidade alfa 4 do proteossoma 26S que pode substituir a subunidade 3 (sob condições de estresse), para criar uma isoforma mais ativa. Realoca a superfície do citosol durante o estresse oxidativo.	K02731	Enzima Peptidase Proteassomo
IV	25350	2.016	ECU04_1070	DNA Topoisomerase	Proteína altamente conservada que relaxa a cadeia de DNA superenrolada, estabiliza os telômeros e regula a recombinação mitótica.	K03165	Enzima Reparo e recombinação Replicação de DNA
VI	21750	2.015	ECU06_1130	Proteína Hipotética	Proteína sem função definida.	K03699	Toxina bacteriana
VI	22850	2.013	ECU06_1250	RNA Helicase Dependente de ATP	Envolvida no processamento e na degradação do RNA nuclear, além de ser um dos componentes do complexo TRAMP e membro da família DEAD-box helicases.	K14808	Enzima Biogênese ribossomal
XI	37250	2.009	ECU11_1380i	Proteína Ribossomal 60S L15	Componente proteico da subunidade ribossomal 60S.	K02877	Proteína ribossomal
X	30650	2.009	ECU10_1120	Fator de Início da Transcrição TFIID Subunidade 70kDa	Envolvida no início da transcrição da RNA polimerase II e na modificação da cromatina. Semelhante a Histona H4	K03131	Proteína cromossomal Maquinaria de transcrição
VI	24850	2.007	ECU06_1290	Proteína Hipotética Integral de Membrana YMB8	Nucleotídeo transportador de açúcar, semelhante a VRG4p.	K15280	Proteína transportadora
VI	27250	2.006	ECU06_1440	Fator de Elongação de Tradução Alfa-1	Proteína que altera a fidelidade na tradução criando um fenótipo supressor.	K03267	Fator de tradução Biogênese de mRNA
IX	9450	2.006	ECU09_1890	Ativador Transcricional	Atividade de remodelação da cromatina, além de modular a transcrição.	K11665	Enzima Proteína cromossomal
IX	650	2.002	ECU09_1650	Proteína de Ligação GTP	GTPase que associa-se a subunidade 60S, sendo necessária para sua maturação e exportação.	K14537	Biogênese ribossomal
VI	27821	2.001	ECU06_1480	Metionina Permease	Um dos cinco genes de metabolismo de enxofre durante a fase de crescimento.	K13868	Proteína transportadora

Apêndice 2: Gráficos Divergencia *versus* Polimorfismo (K/Pi) e Polimorfismo *versus* Divergencia (Pi/K) dos alinhamentos dos onze cromossomos analisados.

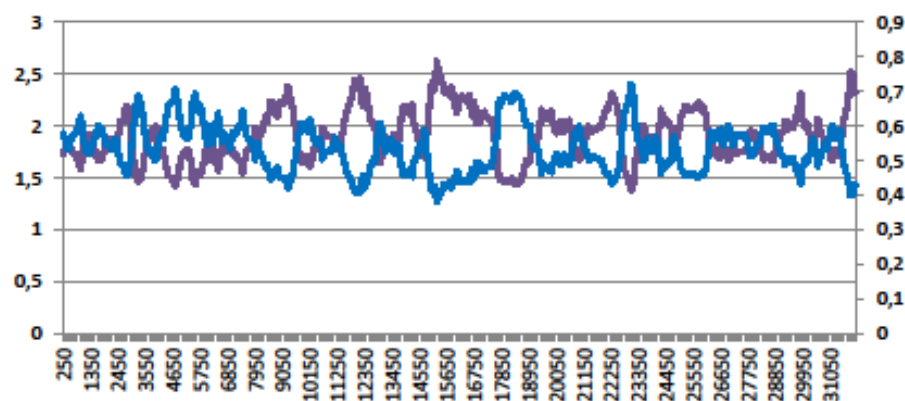
Cromossomo I



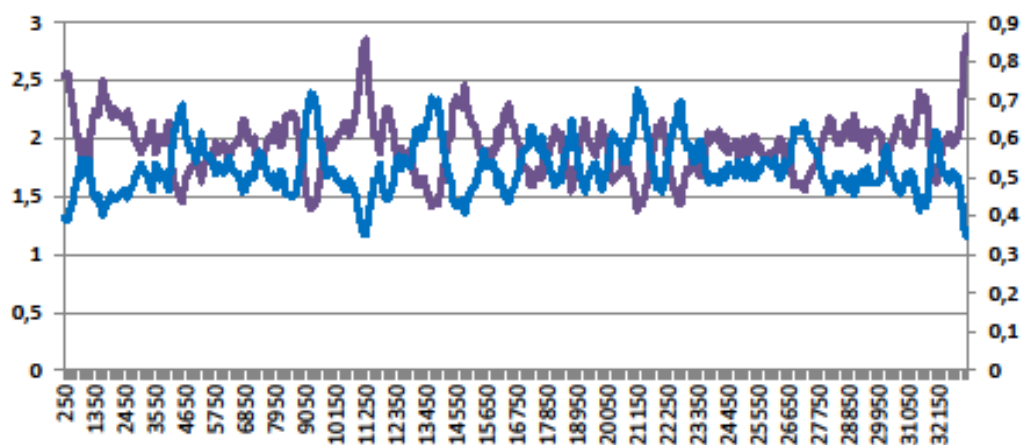
Cromossomo II



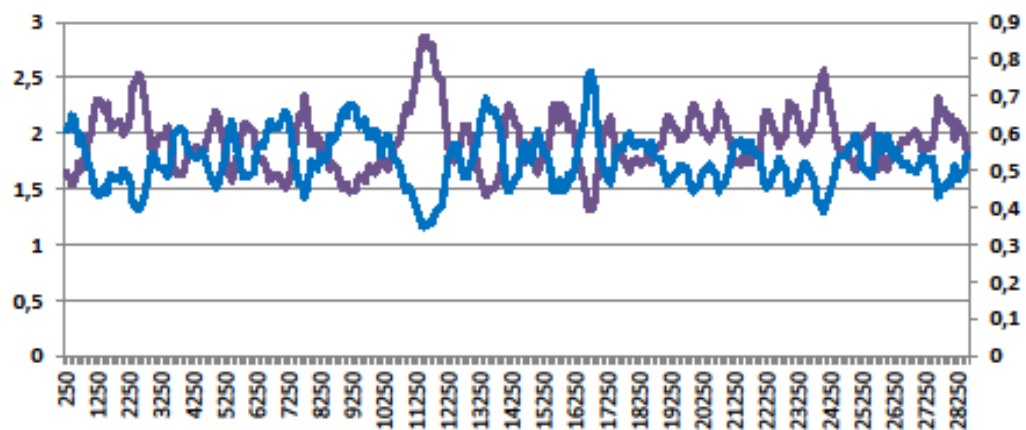
Cromossomo III



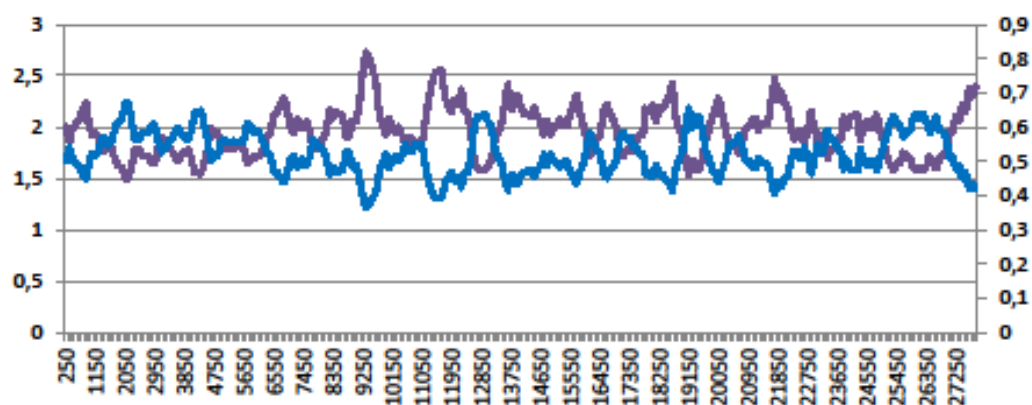
Cromossomo IV



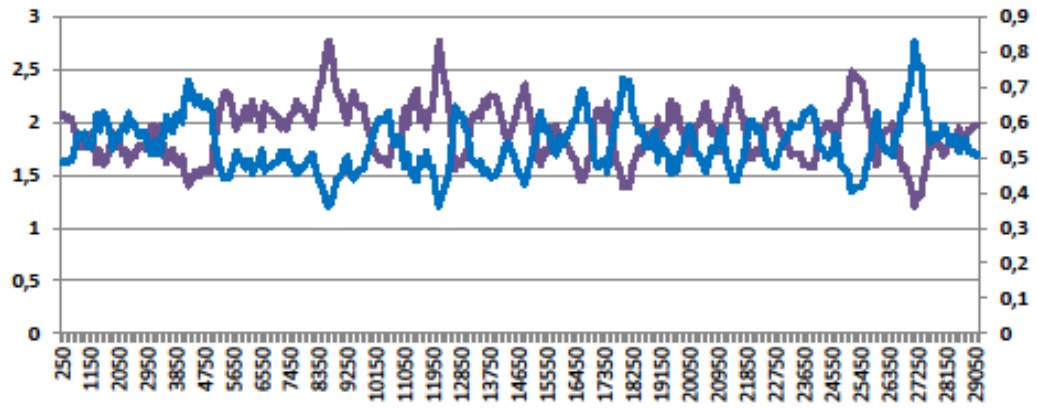
Cromossomo V



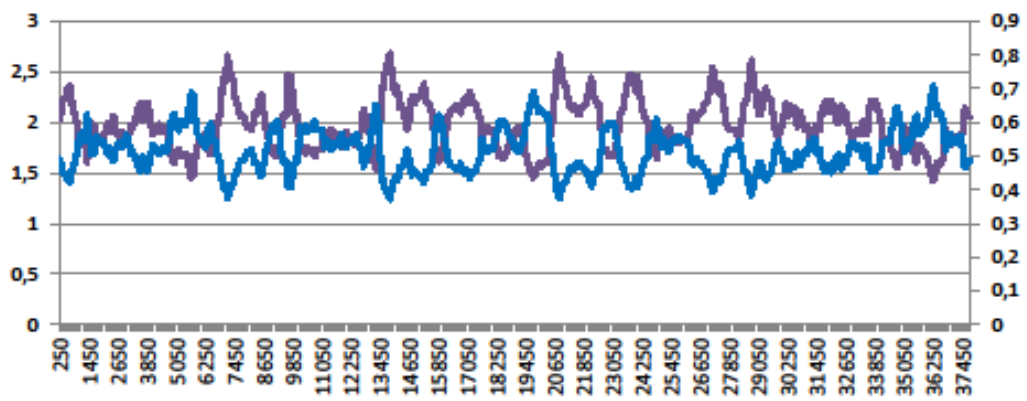
Cromossomo VI



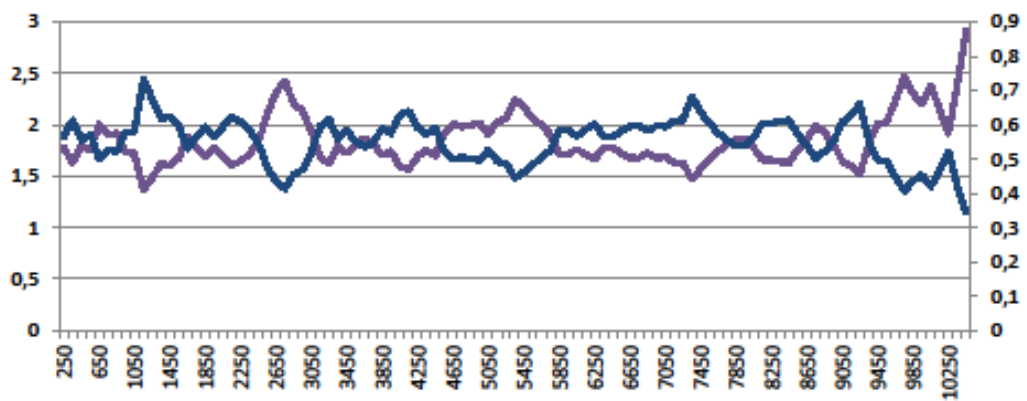
Cromossomo VII



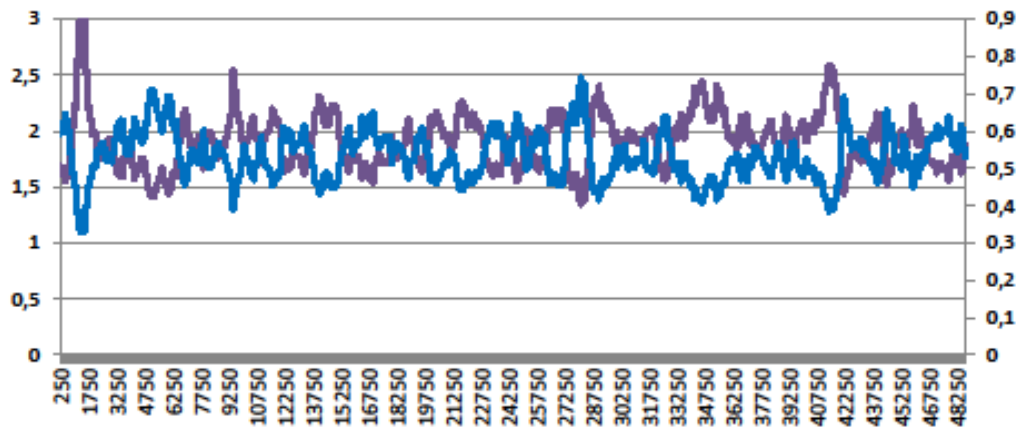
Cromossomo VIII



Cromossomo IX



Cromossomo X



Cromossomo XI

