Universidade de São Paulo
Faculdade de Medicina de Ribeirão Preto
Programa de Pós-Graduação em Genética

**Ferramentas computacionais para o estudo estrutural e funcional
de genes de dermatófitos potencialmente envolvidos na
patogenicidade**

Pablo Rodrigo Sanches

Ribeirão Preto - SP

2015

PABLO RODRIGO SANCHES

**Ferramentas computacionais para o estudo estrutural e funcional
de genes de dermatófitos potencialmente envolvidos na
patogenicidade**

Tese apresentada ao programa de Pós-graduação em Genética da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, para obtenção do título de Doutor em Ciências.

Área de Concentração: Genética

Orientadora: Prof^a. Dr^a. Nilce Maria Martinez-Rossi

Ribeirão Preto - SP

2015

FICHA CATALOGRÁFICA

Sanches, Pablo Rodrigo

Ferramentas computacionais para o estudo estrutural e funcional de genes de dermatófitos potencialmente envolvidos na patogenicidade.

Tese (Doutorado) - Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2015.

p. 147 : il. ; 30 cm

Tese de Doutorado apresentada à Faculdade de Medicina de Ribeirão Preto-USP. Área de Concentração: Genética.

Orientadora: Profa. Dra. Nilce Maria Martinez-Rossi

1. Dermatófitos. 2. Proteases. 3. Expressão gênica. 4. Genômica comparativa. 5. Bioinformática.

APOIO E SUPORTE FINANCEIRO

Este projeto foi realizado com o apoio financeiro das seguintes entidades e instituições:

Faculdade de Medicina de Ribeirão Preto - FMRP/USP

Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES

Fundação de Apoio ao Ensino, Pesquisa e Assistência - FAEPA (FMRP/USP)

NORMALIZAÇÃO ADOTADA

Este documento foi elaborado de acordo com:

Associação Brasileira de Normas Técnicas (ABNT): NBR 6023, NBR 6024, NBR 6027 e NBR 6028.

Universidade de São Paulo. Sistema Integrado de Bibliotecas. Diretrizes para apresentação de dissertações e teses da USP: documento eletrônico e impresso Parte I (ABNT), São Paulo, 2009.

DEDICATÓRIA

Dedico este trabalho a toda minha família, em especial, aos meus pais Claudinei e Thelma, minha esposa Yanê e meu filho Pedro pelo amor e companheirismo.

AGRADECIMENTOS

À Profa. Dra. Nilce Maria Martinez-Rossi pela orientação dedicada à realização deste trabalho e pela amizade e confiança em mim depositadas. Agradeço por ter me recebido no laboratório como funcionário e aluno, mostrando-me o caminho da ciência.

Ao Prof. Dr. Antonio Rossi Filho, pelas discussões valiosíssimas desde quando cheguei ao laboratório.

A todos os professores do Departamento de Genética, que, de alguma forma, contribuíram para a minha formação.

Aos meus pais, Claudinei e Thelma, pelo apoio, suporte e incentivo que sempre recebi deles, tendo sempre como conselho a calma necessária e a obstinação por aquilo que desejava.

À minha esposa Yanê e ao meu filho Pedro, pelo apoio, incentivo, alegria e amor que recebi mesmo quando tive que dividir meu tempo entre eles e os estudos.

Aos Pós-Graduandos que fizeram parte do laboratório nos últimos 12 anos, com os quais tive oportunidade de conviver desde o momento em que fui contratado pela Universidade: Nalu, Henrique, Jeny, Mônica, Sérgio, Eliane, Segato, Paião, Maranhão, Diana, Juliana, Luciene, Cazzaniga, Gabi, Tiago, Larissa, Rodrigo, Aline, Elza, Niede, Ludmila e Lucia. Agradeço-os pelo apoio, disposição, ensinamentos e incentivo em biologia molecular.

À Pós-Graduanda Maíra pelo suporte na realização dos experimentos de bancada.

Aos amigos e funcionários do laboratório Mendelson, Cuca e Marcos, que me ajudaram sempre que precisei.

Ao amigo e funcionário, Pedro, que “segurou a barra” com relação aos problemas de informática do Departamento, enquanto precisei dividir meu tempo como funcionário e aluno.

À secretária de pós-graduação do Programa, Susie, por sempre me auxiliar e solucionar os problemas burocráticos do Doutorado.

Ao pessoal da secretaria do Departamento: Gustavo, Silvia e Ana Claudia, pelo auxílio e prontidão.

À Faculdade de Medicina de Ribeirão Preto por tornar possível a realização deste Doutorado, já que sou funcionário dessa Faculdade.

A todos os meus amigos, pelos momentos de "espairecimento da mente".

A todos aqueles que colaboraram direta ou indiretamente para a realização deste trabalho.

Aos membros da banca examinadora que tão prontamente se dispuseram a avaliar e discutir o presente trabalho.

A Deus, por tudo...

Your work is going to fill a large part of your life, and the only way to be truly satisfied is to do what you believe is great work. And the only way to do great work is to love what you do. If you haven't found it yet, keep looking. Don't settle. As with all matters of the heart, you'll know when you find it.

Seu trabalho vai preencher uma parte grande da sua vida, e a única maneira de ficar realmente satisfeito é fazer o que você acredita ser um ótimo trabalho. E a única maneira de fazer um excelente trabalho é amar o que você faz. Se você ainda não encontrou o que é, continue procurando. Não sossegue. Assim como todos os assuntos do coração, você saberá quando encontrar.

Steve Jobs

RESUMO

SANCHES, P. R. **Ferramentas computacionais para o estudo estrutural e funcional de genes de dermatófitos potencialmente envolvidos na patogenicidade.** 2015. 147 f. Tese (Doutorado) - Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto-SP, 2015.

Dermatófitos são fungos filamentosos que infectam substratos queratinizados como pele, unha e cabelo em busca de nutrientes para se desenvolverem e permanecerem no hospedeiro. Pertencem aos gêneros *Epidermophyton*, *Microsporum* ou *Trichophyton*, os quais, dependendo de seu habitat natural, são classificados em espécies geofílicas, zoofílicas ou antropofílicas. O uso indiscriminado de antifúngicos levou à seleção de cepas resistentes, e o comportamento invasivo desses patógenos em pacientes imunodeprimidos aumentou nos últimos anos, dificultando o tratamento das dermatofitoses. Há, portanto, a necessidade de estudos para um melhor entendimento da biologia dos dermatófitos devido as suas importâncias médica e/ou veterinária e o escasso conhecimento da interação destes patógenos com os hospedeiros. No presente trabalho, analisamos oito espécies de dermatófitos: *Arthroderma benhamiae*, *Microsporum canis*, *Microsporum gypseum*, *Trichophyton interdigitale*, *Trichophyton equinum*, *Trichophyton rubrum*, *Trichophyton tonsurans* e *Trichophyton verrucosum*. Análises de genômica comparativa e de expressão de genes potencialmente envolvidos na degradação de queratina foram realizadas. Além disso, efetuamos o sequenciamento genômico em larga escala de uma das linhagens. A estrutura dos genes *sub3*, *sub5* e *sub7*, que codificam serina endopeptidases com atividade queratinolítica, *mep3* e *mep4*, que codificam proteínas pertencentes ao grupo das metaloendopeptidases, *dppV*, *lap1* e *lap2*, que codificam exopeptidases, foi analisada por meio de ferramentas computacionais. Essas análises revelaram que os genes que codificam proteases possuem alto grau de conservação em suas estruturas, que é menor quando comparadas apenas suas regiões não codificadoras. As análises permitiram também a identificação em regiões promotoras de consensos específicos a gêneros de dermatófitos. Observamos que o acúmulo de transcritos destes genes, avaliados durante o cultivo em queratina, mimetizando o processo infeccioso, não está correlacionado à similaridade das sequências gênicas entre as espécies. Não encontramos correlação entre o nicho preferencial dos dermatófitos e suas sequências gênicas ou níveis transcricionais. Observamos que, na grande maioria das vezes, genes que codificam endo e exopeptidases, possuem acúmulo de transcritos em períodos iniciais de degradação de queratina. Nossos resultados sugerem que diferenças pontuais na sequência gênica, diferenças em regiões promotoras ou, até mesmo, expressão variável destes genes que codificam um conjunto protéico com funções sinérgicas e provavelmente compensatórias, contribuam para os diferentes graus de reações inflamatórias no hospedeiro, bem como para a especificidade patógeno-hospedeiro.

Palavras-chave: Dermatófitos, Proteases, Expressão gênica, Genômica comparativa, Bioinformática.

ABSTRACT

SANCHES, P. R. **Computational tools for the structural and functional study of dermatophytes genes potentially involved in pathogenicity.** 2015. 147 p. Thesis (Doctorate) - Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto-SP, 2015.

Dermatophytes are filamentous fungi that infect keratinized substrates such as skin, nail and hair, searching for nutrients for their development and permanence in the host. They belong to the genera *Epidermophyton*, *Microsporum* or *Trichophyton*, and, depending on their natural habitat, are classified into geophilics, zoophilics or anthropophilics species. The indiscriminate use of antifungals has led to the selection of resistant strains, and the invasive behaviour of these pathogens in immunocompromised patients increased in the last years, hampering the treatment of the dermatophytoses. Therefore, there is a need of studies for a better understanding of the biology of the dermatophytes due to their medical and/or veterinary importance and the scarce knowledge about the interaction of these pathogens with their hosts. In this work, we analyzed eight species of dermatophytes: *Arthroderma benhamiae*, *Microsporum canis*, *Microsporum gypseum*, *Trichophyton interdigitale*, *Trichophyton equinum*, *Trichophyton rubrum*, *Trichophyton tonsurans*, and *Trichophyton verrucosum*. Comparative genomics and gene expression analyses of genes potentially involved in keratin degradation were performed. Moreover, we performed a large-scale genome sequencing of one of the strains. The structure of the genes *sub3*, *sub5*, and *sub7*, which encode serine endopeptidases with keratinolytic activity, *mep3*, and *mep4*, which encode proteins belonging to the group of the metalloendopeptidases, *dppV*, *lap1*, and *lap2*, encoding exopeptidases, were analyzed by computational tools. These analyses revealed that the genes encoding proteases possesses high degree of conservation in their structures, which are lower when their non-coding regions are compared. The analyses also allowed the identification of consensus in promoter regions, specific of dermatophytes genera. We observed that the transcripts accumulation of these genes, evaluated during the cultivation in keratin, mimicking the infection process, is not correlated to the gene sequence similarities among the species. We have not found any correlation between the preferential niche of dermatophytes and their gene sequences or transcription levels. Most of the times, we observed that genes encoding endo and exopeptidases accumulated transcripts at the beginning of keratin degradation. Our results suggest that specific differences in the genic sequencing, differences in promoter regions, or even variable expression of these genes encoding a set of proteins with synergic and probably compensatory functions, contribute to different levels of inflammatory reactions in the host, as well as to the host-pathogen specificity.

Keywords: Dermatophytes, Proteases, Gene expression, Comparative genomics, Bioinformatics.

LISTA DE FIGURAS

Figura 1: Relação filogenética de dermatófitos com base em sequências da região ITS. Para a construção da árvore, foi utilizado o modelo Kimura-2-parâmetros, que assume taxas diferentes entre transições (A-G, C-T) e transversões (A-C, A-T, C-G, G-T), e o método <i>Neighbour Joining</i> , que presume diferenças nas taxas de evolução de cada táxon. <i>Keratinomyces ceretanicus</i> foi usado como grupo externo. Na cor preta estão as espécies antropofílicas, em vermelho, as espécies zoofílicas e, em verde, as espécies geofílicas (GRASER et al., 2008).	21
Figura 2: Cultura, microconídios e macroconídios de <i>M. canis</i> (ELLIS et al., 2007).	23
Figura 3: Cultura, microconídios e macroconídios de <i>M. gypseum</i> (ELLIS et al., 2007).	24
Figura 4: Cultura, microconídios, macroconídios, clamidósporos e hifas espirais de <i>T. interdigitale</i> (ELLIS et al., 2007).	25
Figura 5: Cultura, microconídios, macroconídios e órgãos nodulares de <i>T. equinum</i> (ELLIS et al., 2007).	26
Figura 6: Cultura, microconídios e macroconídios de <i>T. rubrum</i> (ELLIS et al., 2007).	27
Figura 7: Colônias, hifas, microconídios e macroconídios de <i>T. tonsurans</i> (ELLIS et al., 2007).	28
Figura 8: Cultura, hifa e microconídios de <i>A. benhamiae</i> (SYMONEIS et al., 2013).	28
Figura 9: Cultura, microconídios e macroconídios de <i>T. verrucosum</i> (ELLIS et al., 2007).	29
Figura 10: Representação da via de degradação de tecidos queratinizados pelos dermatófitos em pH neutro (BALDO et al., 2012).	33
Figura 11: Esquema representativo da metodologia utilizada pelo kit Nextera XT para construção de biblioteca genômica (ILLUMINA, 2012).	45
Figura 12: Esquema representativo da metodologia utilizada pelo MiSeq (ILLUMINA, 2010).	47
Figura 13: Exemplo de algoritmo desenvolvido em linguagem <i>Pearl</i>	53
Figura 14: (A) Alinhamento do conteúdo gênico entre os genes que codificam para Sub3. A linha vertical indica o deslocamento dos segmentos alinhados. (B) Representação esquemática dos genes nos oito dermatófitos em estudo. Os organismos estão dispostos na seguinte ordem: <i>M. canis</i> , <i>A. benhamiae</i> , <i>T. interdigitale</i> , <i>M. gypseum</i> , <i>T. equinum</i> , <i>T. rubrum</i> , <i>T. tonsurans</i> e <i>T. verrucosum</i>	62
Figura 15: Representação esquemática dos genes que codificam para Sub5 nos oito dermatófitos em estudo. Os organismos estão dispostos na seguinte ordem: <i>M. canis</i> , <i>A. benhamiae</i> , <i>T. interdigitale</i> , <i>M. gypseum</i> , <i>T. equinum</i> , <i>T. rubrum</i> , <i>T. tonsurans</i> e <i>T. verrucosum</i> . (*) Isoformas descritas em <i>T. rubrum</i> para o gene representado.	63
Figura 16: Representação esquemática dos genes que codificam para Sub7 nos oito dermatófitos em estudo. Os organismos estão dispostos na seguinte ordem: <i>M. canis</i> , <i>A. benhamiae</i> , <i>T. interdigitale</i> , <i>M. gypseum</i> , <i>T. equinum</i> , <i>T. rubrum</i> , <i>T. tonsurans</i> e <i>T. verrucosum</i>	64
Figura 17: Relação filogenética entre os dermatófitos com base nas sequências da proteína subtilisina 3 (Sub3).	69

Figura 18: Relação filogenética entre os dermatófitos com base nas sequências da proteína subtilisina 5 (Sub5).	69
Figura 19: Relação filogenética entre os dermatófitos com base nas sequências da proteína subtilisina 7 (Sub7).	70
Figura 20: Sintenia das regiões próximas ao gene que codifica subtilisina 3 (Sub3). A localização do gene está apontada com uma barra vertical vazada. Ordem dos alinhamentos: <i>T. rubrum</i> , <i>A. benhamiae</i> , <i>M. canis</i> , <i>M. gypseum</i> , <i>T. equinum</i> , <i>T. tonsurans</i> e <i>T. verrucosum</i> .	71
Figura 21: Sintenia das regiões próximas ao gene que codifica subtilisina 5 (Sub5). A localização do gene está apontada com uma barra vertical vazada. Ordem dos alinhamentos: <i>T. rubrum</i> , <i>A. benhamiae</i> , <i>M. canis</i> , <i>M. gypseum</i> , <i>T. equinum</i> , <i>T. tonsurans</i> e <i>T. verrucosum</i> .	72
Figura 22: Número de sítios de ligação para fatores de transcrição para <i>sub3</i> . Análise com base no banco de dados YEASTRACT (Yeast Search for Transcriptional Regulators And Consensus Tracking)	73
Figura 23: Número de sítios de ligação para fatores de transcrição para <i>sub5</i> . Análise com base no banco de dados YEASTRACT (Yeast Search for Transcriptional Regulators And Consensus Tracking)	74
Figura 24: Número de sítios de ligação para fatores de transcrição para <i>sub7</i> . Análise com base no banco de dados YEASTRACT (Yeast Search for Transcriptional Regulators And Consensus Tracking)	74
Figura 25: Software desenvolvido para visualização dos dados de sequenciamento (metodologia SANGER). ...	78
Figura 26: Gráfico dos valores de expressão relativa (<i>fold change</i>) para o gene <i>sub3</i> em cinco dos dermatófitos em estudo. Para cada organismo, o gráfico apresenta os valores em 24 horas e 96 horas.	79
Figura 27: Heatmap construído com os valores de expressão para o gene <i>sub3</i> .	80
Figura 28: Gráfico dos valores de expressão relativa (<i>fold change</i>) para o gene <i>sub5</i> em cinco dos dermatófitos em estudo. Para cada organismo, o gráfico apresenta os valores em 24 horas e 96 horas.	81
Figura 29: Heatmap construído com os valores de expressão para o gene <i>sub5</i> .	81
Figura 30: Gráfico dos valores de expressão relativa (<i>fold change</i>) para o gene <i>sub7</i> em cinco dos dermatófitos em estudo. Para cada organismo, o gráfico apresenta os valores em 24 horas e 96 horas.	82
Figura 31: Heatmap construído com os valores de expressão para o gene <i>sub7</i> .	82
Figura 32: Representação esquemática dos genes <i>mep3</i> nos oito dermatófitos em estudo. Os organismos estão dispostos na seguinte ordem: <i>M. canis</i> , <i>A. benhamiae</i> , <i>T. interdigitale</i> , <i>M. gypseum</i> , <i>T. equinum</i> , <i>T. rubrum</i> , <i>T. tonsurans</i> e <i>T. verrucosum</i> . (*) Isoformas descritas em <i>T. rubrum</i> para o gene representado.	84
Figura 33: Representação esquemática dos genes <i>mep4</i> nos oito dermatófitos em estudo. Os organismos estão dispostos na seguinte ordem: <i>M. canis</i> , <i>A. benhamiae</i> , <i>T. interdigitale</i> , <i>M. gypseum</i> , <i>T. equinum</i> , <i>T. rubrum</i> , <i>T. tonsurans</i> e <i>T. verrucosum</i> . (*) Isoformas descritas em <i>T. rubrum</i> para o gene representado.	84
Figura 34: Relação filogenética entre os dermatófitos com base nas sequências da proteína metaloprotease 3 (Mep3).	88
Figura 35: Relação filogenética entre os dermatófitos com base nas sequências da proteína metaloprotease 4 (Mep4).	89
Figura 36: Número de sítios de ligação para fatores de transcrição para <i>mep3</i> . Análise com base no banco de dados YEASTRACT (Yeast Search for Transcriptional Regulators And Consensus Tracking).	90

Figura 37: Índice O/A (Original/Aleatório). Quanto maior o índice, maior a chance de uma pequena sequência pertencer a posições precisas dentro de um genoma. Foram utilizados genomas de tamanho médio de 5Mb para a confecção dessa análise (o genoma aleatório foi gerado utilizando a frequência de 0,25 para cada nucleotídeo A/C/G/T).	91
Figura 38: Número de sítios de ligação para fatores de transcrição para <i>mep4</i> . Análise com base no banco de dados YEASTRACT (<i>Yeast Search for Transcriptional Regulators And Consensus Tracking</i>)	92
Figura 39: Gráfico dos valores de expressão relativa (<i>fold change</i>) para o gene <i>mep3</i> em cinco dos dermatófitos em estudo. Para cada organismo, o gráfico apresenta os valores em 24 horas e 96 horas.	94
Figura 40: <i>Heatmap</i> construído com os valores de expressão para <i>mep3</i>	95
Figura 41: Gráfico dos valores de expressão relativa (<i>fold change</i>) para o gene <i>mep4</i> em cinco dos dermatófitos em estudo. Para cada organismo, o gráfico apresenta os valores em 24 horas e 96 horas.	96
Figura 42: <i>Heatmap</i> construído com os valores de expressão para <i>mep4</i>	96
Figura 43: (A) Alinhamento do conteúdo gênico entre os genes <i>lap1</i> . A linha vertical indica o deslocamento dos segmentos alinhados. (B) Representação esquemática dos genes nos oito dermatófitos em estudo. Os organismos estão dispostos na seguinte ordem: <i>M. canis</i> , <i>A. benhamiae</i> , <i>T. interdigitale</i> , <i>M. gypseum</i> , <i>T. equinum</i> , <i>T. rubrum</i> , <i>T. tonsurans</i> e <i>T. verrucosum</i> .	97
Figura 44: Representação esquemática dos genes <i>lap2</i> nos oito dermatófitos em estudo. Os organismos estão dispostos na seguinte ordem: <i>M. canis</i> , <i>A. benhamiae</i> , <i>T. interdigitale</i> , <i>M. gypseum</i> , <i>T. equinum</i> , <i>T. rubrum</i> , <i>T. tonsurans</i> e <i>T. verrucosum</i> . (*) Isoformas descritas em <i>T. rubrum</i> para o gene representado.	98
Figura 45: Relação filogenética entre os dermatófitos com base nas sequências da proteína leucina aminopeptidase 1 (Lap1).	101
Figura 46: Relação filogenética entre os dermatófitos com base nas sequências da proteína leucina aminopeptidase 2 (Lap2).	102
Figura 47: Gráfico dos valores de expressão relativa (<i>fold change</i>) para o gene <i>lap1</i> em cinco dos dermatófitos em estudo. Para cada organismo, o gráfico apresenta os valores em 24 horas e 96 horas.	104
Figura 48: <i>Heatmap</i> construído com os valores de expressão para <i>lap1</i> .	105
Figura 49: Gráfico dos valores de expressão relativa (<i>fold change</i>) para o gene <i>lap2</i> em cinco dos dermatófitos em estudo. Para cada organismo, o gráfico apresenta os valores em 24 horas e 96 horas.	105
Figura 50: <i>Heatmap</i> construído com os valores de expressão para <i>lap2</i> .	106
Figura 51: Representação esquemática dos genes que codificam para DppV nos oito dermatófitos em estudo. Os organismos estão dispostos na seguinte ordem: <i>M. canis</i> , <i>A. benhamiae</i> , <i>T. interdigitale</i> , <i>M. gypseum</i> , <i>T. equinum</i> , <i>T. rubrum</i> (Supercontig 11: 290252-292693 +), <i>T. tonsurans</i> e <i>T. verrucosum</i> .	107
Figura 52: Relação filogenética entre os dermatófitos com base nas sequências da proteína dipeptidil peptidase V (DppV).	109
Figura 53: Gráfico dos valores de expressão relativa (<i>fold change</i>) para o gene <i>dppV</i> em cinco dos dermatófitos em estudo. Para cada organismo, o gráfico apresenta os valores em 24 horas e 96 horas.	111
Figura 54: <i>Heatmap</i> construído com os valores de expressão para <i>dppV</i> .	111

LISTA DE TABELAS

Tabela 1: Meio Extrato de Malte-Ágar (MEA)	43
Tabela 2: Lista dos genes/proteínas utilizados no estudo com seus respectivos códigos de identificação.	49
Tabela 3: Tampão de Migração MOPS 5X	53
Tabela 4: Tampão TAE (Tris-Acetato-EDTA) (SAMBROOK et al., 1989).....	54
Tabela 5: Oligonucleotídeos utilizados nas análises do perfil transcricional por PCR em tempo real.	55
Tabela 6: Sequenciamento do genoma de <i>T. interdigitale</i> (H6).	59
Tabela 7: Montagem do genoma de <i>T. interdigitale</i> (H6).	60
Tabela 8: Número de variações entre <i>M. canis</i> e outras espécies de dermatófitos para o gene que codifica subtilisina 3 (Sub3).	64
Tabela 9: Número de variações entre <i>M. canis</i> e outras espécies de dermatófitos para o gene que codifica subtilisina 5 (Sub5).	65
Tabela 10: Número de variações entre <i>M. canis</i> e outras espécies de dermatófitos para o gene que codifica subtilisina 7 (Sub7).	65
Tabela 11: Número de variações entre <i>T. equinum</i> e <i>T. tonsurans</i> e outras espécies do gênero <i>Trichophyton</i> para o gene que codifica subtilisina 3 (Sub3).	66
Tabela 12: Número de variações entre <i>T. equinum</i> e <i>T. tonsurans</i> e outras espécies do gênero <i>Trichophyton</i> para o gene que codifica subtilisina 5 (Sub5).	66
Tabela 13: Número de variações entre <i>T. equinum</i> e <i>T. tonsurans</i> e outras espécies do gênero <i>Trichophyton</i> para o gene que codifica subtilisina 7 (Sub7).	66
Tabela 14: Matriz de distância para a proteína subtilisina 3 (Sub3).	67
Tabela 15: Matriz de distância para a proteína subtilisina 5 (Sub5).	67
Tabela 16: Matriz de distância para a proteína subtilisina 7 (Sub7).	68
Tabela 17: Amostra de sequências consenso obtidas por meio do alinhamento da região 1.000 pares de base <i>upstream</i> do gene <i>sub3</i> dos oito dermatófitos em estudo, incluindo dois organismos como grupo externo.	75
Tabela 18: Amostra de sequências consenso obtidas por meio do alinhamento da região 1.000 pares de base <i>upstream</i> do gene <i>sub5</i> dos oito dermatófitos em estudo, incluindo dois organismos como grupo externo.	76
Tabela 19: Amostra de sequências consenso obtidas por meio do alinhamento da região 1.000 pares de base <i>upstream</i> do gene <i>sub7</i> dos oito dermatófitos em estudo, incluindo dois organismos como grupo externo.	76
Tabela 20: Genes de interesse encontrados no banco de experimentos sobre sequenciamento de genes expressos em <i>T. interdigitale</i> (H6), realizados no laboratório.....	78
Tabela 21: Número de variações entre <i>M. canis</i> e outras espécies de dermatófitos para o gene que codifica metaloprotease 3 (Mep3).	85
Tabela 22: Número de variações entre <i>M. canis</i> e outras espécies de dermatófitos para o gene que codifica metaloprotease 4 (Mep4).	86

Tabela 23: Número de variações entre <i>T. equinum</i> e <i>T. tonsurans</i> e outras espécies do gênero <i>Trichophyton</i> para o gene que codifica metaloprotease 3 (Mep3).	86
Tabela 24: Número de variações entre <i>T. equinum</i> e <i>T. tonsurans</i> e outras espécies do gênero <i>Trichophyton</i> para o gene que codifica metaloprotease 4 (Mep4).	86
Tabela 25: Matriz de distância para a proteína metaloprotease 3 (Mep3).	87
Tabela 26: Matriz de distância para a proteína metaloprotease 4 (Mep4).	87
Tabela 27: Amostra de sequências consenso obtidas por meio do alinhamento da região 1.000 pares de base <i>upstream</i> do gene <i>mep3</i> dos oito dermatófitos em estudo, incluindo dois organismos como grupo externo.	93
Tabela 28: Amostra de sequências consenso obtidas por meio do alinhamento da região 1.000 pares de base <i>upstream</i> do gene <i>mep4</i> dos oito dermatófitos em estudo, incluindo dois organismos como grupo externo.	93
Tabela 29: Número de variações entre <i>M. canis</i> e outras espécies de dermatófitos para o gene que codifica leucina aminopeptidase 1 (Lap1).	99
Tabela 30: Número de variações entre <i>M. canis</i> e outras espécies de dermatófitos para o gene que codifica leucina aminopeptidase 2 (Lap2).	99
Tabela 31: Número de variações entre <i>T. equinum</i> e <i>T. tonsurans</i> e outras espécies do gênero <i>Trichophyton</i> para o gene que codifica leucina aminopeptidase 1 (Lap1).	99
Tabela 32: Número de variações entre <i>T. equinum</i> e <i>T. tonsurans</i> e outras espécies do gênero <i>Trichophyton</i> para o gene que codifica leucina aminopeptidase 2 (Lap2).	100
Tabela 33: Matriz de distância para a proteína leucina aminopeptidase 1 (Lap1).	100
Tabela 34: Matriz de distância para a proteína leucina aminopeptidase 2 (Lap2).	101
Tabela 35: Amostra de sequências consenso obtidas por meio do alinhamento da região 1.000 pares de base <i>upstream</i> do gene <i>lap1</i> dos oito dermatófitos em estudo, incluindo dois organismos como grupo externo.	103
Tabela 36: Amostra de sequências consenso obtidas por meio do alinhamento da região 1.000 pares de base <i>upstream</i> do gene <i>lap2</i> dos oito dermatófitos em estudo, incluindo dois organismos como grupo externo.	103
Tabela 37: Número de variações entre <i>M. canis</i> e outras espécies de dermatófitos para o gene que codifica dipeptidil peptidase V (DppV).	107
Tabela 38: Número de variações entre <i>T. equinum</i> e <i>T. tonsurans</i> e outras espécies do gênero <i>Trichophyton</i> para o gene que codifica dipeptidil peptidase V (DppV).	108
Tabela 39: Matriz de distância para a proteína dipeptidil peptidase V (DppV).	108
Tabela 40: Amostra de sequências consenso obtidas por meio do alinhamento da região 1.000 pares de base <i>upstream</i> do gene <i>dppV</i> dos dermatófitos em estudo, incluindo dois organismos como grupo externo. ...	110

SUMÁRIO

I.	INTRODUÇÃO.....	20
1.1	Dermatófitos.....	22
1.1.1	<i>Microsporum canis</i>	22
1.1.2	<i>Microsporum gypseum</i>	23
1.1.3	<i>Trichophyton interdigitale</i>	24
1.1.4	<i>Trichophyton equinum</i>	25
1.1.5	<i>Trichophyton rubrum</i>	26
1.1.6	<i>Trichophyton tonsurans</i>	27
1.1.7	<i>Arthroderma benhamiae</i>	28
1.1.8	<i>Trichophyton verrucosum</i>	29
1.2	Proteases	30
1.3	Processo de interação patógeno-hospedeiro	32
1.4	Genômica comparativa.....	33
II.	HIPÓTESE DO TRABALHO	38
III.	OBJETIVOS.....	40
IV.	MATERIAIS E MÉTODOS.....	42
4.1	Linhas.....	42
4.2	Meios de cultura.....	43
4.3	Manutenção e estoque das linhagens.....	43
4.4	Preparo das soluções de conídios.....	44
4.5	Sequenciamento e montagem do genoma de <i>T. interdigitale</i> (H6)	44
4.5.1	Extração de DNA genômico.....	44
4.5.2	Construção da biblioteca de sequenciamento	44
4.5.3	Sequenciamento em larga escala	45
4.5.4	Montagem do genoma	48
4.6	Análises de genômica comparativa dos dermatófitos	48
4.6.1	Seleção dos genes de interesse	48
4.6.2	Presença/ausência e conservação entre os genes	50
4.6.3	Análise dos genes (identificação das regiões conservadas)	50
4.6.4	Análise das regiões não codantes	50
4.6.5	Análise das proteínas (comparação do conteúdo protéico)	51
4.6.6	Análise das relações filogenéticas	51
4.6.7	Análise da sintenia genômica	51
4.6.8	Análise de variantes (SNPs)	51
4.6.9	Análise de sítios de ligação nas regiões promotoras.....	52

4.6.10	Algoritmos em Bioinformática	52
4.7	Perfil transcricional comparativo por PCR em tempo real.....	53
4.7.1	Soluções	53
4.7.1.1	Soluções para RNA	53
4.7.1.2	Outras Soluções	54
4.7.2	Condições de cultivo	54
4.7.3	Extração de RNA total	54
4.7.4	Síntese de cDNA	55
4.7.5	Construção de oligonucleotídeos.....	55
4.7.6	Análise dos genes diferencialmente expressos	56
V.	RESULTADOS	59
5.1	Sequenciamento e montagem do genoma de <i>T. interdigitale</i> (H6)	59
5.2	Análise de genes envolvidos no processo de degradação da queratina	61
5.2.1	Serina proteases.....	61
5.2.1.1	Análises por Bioinformática	61
5.2.1.2	Expressão gênica	77
5.2.2	Metaloproteases.....	83
5.2.2.1	Análises por Bioinformática	83
5.2.2.2	Expressão gênica	94
5.2.3	Leucina-aminopeptidases	96
5.2.3.1	Análises por Bioinformática	96
5.2.3.2	Expressão gênica	104
5.2.4	Dipeptidil peptidase V.....	106
5.2.4.1	Análises por Bioinformática	106
5.2.4.2	Expressão gênica	110
VI.	DISCUSSÃO.....	113
VII.	CONCLUSÕES	125
	REFERÊNCIAS BIBLIOGRÁFICAS.....	127
	ANEXO DE PUBLICAÇÕES	138

Introdução

One thing I have learned in a long life: that all our science, measured against reality, is primitive and childlike - and yet it is the most precious thing we have.

Albert Einstein

I. Introdução

Os fungos são organismos eucarióticos que se constituem num reino próprio, possuindo grande importância ambiental, econômica e biotecnológica, além de serem utilizados na decomposição de matéria orgânica, na alimentação, em processos de fermentação e de produção de substâncias de interesse, como por exemplo, os antibióticos. Além disso, algumas espécies patogênicas ou oportunistas causam infecções em humanos, animais e plantas, gerando perdas econômicas na agricultura e na pecuária.

Entre os fungos patogênicos, estão os dermatófitos, que pertencem a um grupo de fungos queratinofílicos, ou seja, capazes de utilizar a queratina como fonte de nutrientes, invadindo tecidos queratinizados em humanos e animais, causando os principais tipos de micoses superficiais denominadas dermatofitoses.

Os dermatófitos são alocados em três gêneros: *Epidermophyton*, *Microsporum* e *Trichophyton*, os quais são classificados em diferentes espécies (geofílicas, zoofílicas e antropofílicas), baseados em seu habitat natural – solo, animais ou homem, respectivamente (WEITZMAN e SUMMERBELL, 1995).

Embora subdivididos em três gêneros, Graser et al. (2008) apontam uma instabilidade na taxonomia e identificação entre as espécies de dermatófitos (Figura 1). Logo, a dificuldade na identificação deles é tema constantemente abordado na investigação biológica molecular moderna.

As dermatofitoses recebem o nome de tineas, reação do hospedeiro às enzimas liberadas pelo fungo durante o processo de digestão dos tecidos queratinizados para a obtenção de nutrientes, seguida do nome da região corpórea onde a infecção está instalada, como por exemplos *tinea unguium* (unhas), *tinea capitis* (cabelo), *tinea pedis* (pele dos pés), *tinea manum* (mãos) e *tinea cruris* (virilha).

Anualmente, as dermatofitoses afetam milhões de indivíduos, estabelecendo um importante problema na saúde pública, já que a severidade e a longevidade da doença refletem na terapia (YANG et al., 2007). Em indivíduos saudáveis, em geral, é caracterizada por uma infecção superficial, mas em indivíduos imunodeprimidos, como HIV positivos e transplantados, podem ocorrer infecções profundas e severas. Tratamento de infecções por

dermatófitos representa uma carga de custos significativa. Estima-se que mais de 500 milhões de dólares por ano em todo o mundo é gasto em medicamentos para tratar dermatofitoses (GRASER et al., 2008; ACHTERMAN e WHITE, 2012).

Figura 1: Relação filogenética de dermatófitos com base em sequências da região ITS. Para a construção da árvore, foi utilizado o modelo Kimura-2-parâmetros, que assume taxas diferentes entre transições (A-G, C-T) e transversões (A-C, A-T, C-G, G-T), e o método *Neighbour Joining*, que presume diferenças nas taxas de evolução de cada táxon. *Keratinomyces ceretanicus* foi usado como grupo externo. Na cor preta estão as espécies antropofílicas, em vermelho, as espécies zoofílicas e, em verde, as espécies geofílicas (GRASER et al., 2008).

Segundo a Organização Mundial da Saúde, os dermatófitos afetam cerca de 25% da população mundial. Estima-se que de 30 a 70% dos adultos sejam portadores assintomáticos desses patógenos e que a incidência dessa doença aumente com a idade. Embora sejam organismos de caráter cosmopolita, fatores geoclimáticos e sociais interferem na distribuição das espécies dermatofíticas nas diferentes regiões do mundo (SEEBACHER et al., 2008).

Alguns estudos epidemiológicos revelaram que entre 35 e 40% dos indivíduos analisados em 16 países da Europa apresentavam infecção nos pés (*tinea pedis*) causada por dermatófitos (BURZYKOWSKI et al., 2003). Nos Estados Unidos, estudos realizados com crianças frequentadoras de creches revelaram que entre 22 e 50% exibiam sintomas de infecção do couro cabeludo (ABDEL-RAHMAN et al., 2006). No Brasil, as regiões Sul e Sudeste têm apresentado alta incidência de infecções causadas pelo dermatófito *Trichophyton rubrum*, seguido por *Microsporum canis* e *Trichophyton mentagrophytes*; na região Nordeste há maior prevalência de *Trichophyton tonsurans*, *T. rubrum* e *M. canis* (SIDRIM e ROCHA, 2004).

Devido à importância médica e veterinária dos dermatófitos, o genoma das espécies *M. canis*, *Microsporum gypseum*, *Trichophyton equinum*, *Trichophyton interdigitale*, *T. rubrum* e *T. tonsurans* foram recentemente sequenciados e disponibilizados pelo *Broad Institute of MIT and Harvard* (MARTINEZ et al., 2012). O Laboratório de Genética e Biologia Molecular de Fungos do Departamento de Genética-FMRP/USP participa do *Steering Committee*, do projeto intitulado *Genomic Determinants of Infection Competence in Dermatophyte Fungi*, financiado pelo NIH e que propõe o sequenciamento do genoma desses fungos. Somando-se a esse projeto, outras duas espécies de dermatófitos (*Arthroderma benhamiae* e *Trichophyton verrucosum*) foram também sequenciadas por uma iniciativa independente (BURMESTER et al., 2011) e disponibilizadas pelo *Broad Institute*.

1.1 Dermatófitos

1.1.1 *Microsporum canis*

Microsporum canis é uma espécie zoofílica transmitida ao homem por animais domésticos e tem os felinos jovens como reservatório principal. É o mais frequente

dermatófito zoofílico em humanos (SILVA et al., 2003), sendo responsável por lesões no couro cabeludo, especialmente, em crianças.

Apesar de ter como hospedeiro natural principal gatos domésticos, existem relatos de ocorrência em outros animais, inclusive cães domésticos. A não ocorrência em seus reservatórios principais, aparentemente, faz com que *M. canis* perca sua virulência após poucas transmissões entre humanos, tornando esse dermatófito incapaz de se disseminar amplamente entre as populações humanas. Seu poder de infecção é restabelecido após novo contato com o hospedeiro animal (NENOFF et al., 2014; GRUNWALD et al., 2015).

Possui crescimento moderadamente rápido e as colônias apresentam textura cotonosa, discreto relevo umbilicado, radiado, de tonalidade branca. O reverso possui tonalidade amarelo-limão, que, com o passar do tempo, tende a tornar-se castanho por difusão do pigmento no meio (SIDRIM e ROCHA, 2004). Os macroconídios são abundantes, grandes, fusiformes e com numerosas septações (de 5 a 15 septos); os microconídios são pequenos e piriformes (ELLIS et al., 2007) (Figura 2).

Figura 2: Cultura, microconídios e macroconídios de *M. canis* (ELLIS et al., 2007).

1.1.2 *Microsporum gypseum*

O fungo geofílico *Microsporum gypseum*, com distribuição mundial, raramente causa infecções superficiais em humanos e animais. Nos seres humanos é responsável por casos de *tinea corporis*, que afetam principalmente crianças. Este fungo geralmente provoca lesões inflamatórias em áreas descobertas do corpo e do couro cabeludo (MACHADO et al., 2005), e raramente invade tecidos subcutâneos (FERNANDEZ-TORRES et al., 2002). As infecções causadas por ele são comumente adquiridas por contato direto com o solo contendo elevado

número de esporos e raramente são transmitidas entre seres humanos, ou de animais infectados aos seres humanos (ROMANO et al., 2009).

As colônias de *M. gypseum* crescem rapidamente, apresentando bordas irregulares e aspecto granular, com a pigmentação da colônia variando de branco a acastanhado e o reverso pode apresentar variações de cores, que vão do amarelo ao marrom. Os macroconídios são abundantes, simétricos, elipsoidal, de paredes finas, com 4 a 6 septos; os microconídios são piriformes (ELLIS et al., 2007) (Figura 3).

Figura 3: Cultura, microconídios e macroconídios de *M. gypseum* (ELLIS et al., 2007).

1.1.3 *Trichophyton interdigitale*

Juntamente a *T. rubrum*, o *Trichophyton interdigitale* é um patógeno predominante de *tinea unguium* dos pés, causando de 80 a 90% dos casos (RICH et al., 2003; KAUR et al., 2008). É um dos fungos mais comuns que causam doenças zoonóticas da pele, isto é, transmitidas dos animais para as pessoas e de pessoas para animais (CAFARCHIA et al., 2012).

Nessa espécie, é feita uma distinção entre linhagens antropofílicas e zoofílicas. As linhagens antropofílicas tendem a causar *tinea unguium* e *tinea pedis*, e, menos frequentemente, *tinea corporis*. Pacientes portadores de linhagens zoofílicas não são candidatos a vetorizar infecções para outras pessoas de seu convívio (SIDRIM e ROCHA, 2004).

Em seu histórico, originalmente, foi feita uma distinção dentro do complexo *T. mentagrophytes* entre diversas variedades antropofílicas e zoofílicas. Com base em análises moleculares de suas sequências de DNA, agora, considera-se que *T. mentagrophytes* seja uma espécie distinta e que corresponde à antiga variedade zoofílica *T. mentagrophytes* var.

quinckeanum. Quase todas as outras variedades antropofílicas e zoofílicas de *T. mentagrophytes* foram identificadas como geneticamente idênticas. Com isso, uma nova classificação taxonômica foi estabelecida e denominada de *T. interdigitale* (GRASER et al., 1999; NENOFF et al., 2014).

As colônias apresentam formas circulares com textura cotonosa e pigmentação de verso variando entre branca a creme. No reverso, apresenta-se com pigmento amarelo amarronzado. Seus microconídios são numerosos, subesféricos a piriformes e poderão surgir hifas em espiral, além de clamidósporos em culturas velhas (Figura 4).

Figura 4: Cultura, microconídios, macroconídios, clamidósporos e hifas espirais de *T. interdigitale* (ELLIS et al., 2007).

1.1.4 *Trichophyton equinum*

A espécie zoofílica associada à infecção em cavalos, *Trichophyton equinum*, pode também causar infecção em cães, gatos e, até mesmo, no homem (HUOVINEN et al., 1998; CHERMETTE et al., 2008). Dermatofitoses causadas em humanos por esse patógeno são raras. E não se sabe o quão difundido *T. equinum* está entre os humanos devido ao pequeno número de casos relatados na literatura (HUOVINEN et al., 1998). Relatos indicam baixa incidência dessa espécie no continente europeu. Durante o período de 1951 a 1987, apenas um em cada 4.571 casos de tinea diagnosticada em Santiago de Compostela, Espanha, foi atribuído à *T. equinum* (*tinea corporis*) (PEREIRO MIGUENS et al., 1991). Na Inglaterra, um

caso de *tinea capitis* foi relatado em 1994, em uma mulher em tratamento sistêmico com corticosteroides, provavelmente, proveniente de cavalos de montaria (BURDEN et al., 1994).

Esta espécie demonstra crescimento rápido em meio Sabouraud ágar. Suas colônias são geralmente planas de cor branca, mas algumas podem desenvolver dobras suaves ou sulcos radiais. O reverso das colônias é amarelo e com o envelhecimento se torna vermelho escuro no centro. Os macroconídios são raros, clavados e os microconídios abundantes e de formas variáveis (esféricos, piriformes ou clavados) (ELLIS et al., 2007) (Figura 5).

Figura 5: Cultura, microconídios, macroconídios e órgãos nodulares de *T. equinum* (ELLIS et al., 2007).

1.1.5 *Trichophyton rubrum*

O fungo *T. rubrum* é uma espécie antropofílica cosmopolita, responsável por cerca de 70% de todas as infecções por dermatófitos em humanos e é o principal agente responsável por infecções superficiais, como onicomicoses e *tinea pedis*. A alta porcentagem está correlacionada à resistência dessa espécie ao tratamento terapêutico e à sua grande adaptação ao hospedeiro (MUKHERJEE et al., 2003).

Embora sejam normalmente restritas às regiões superficiais da epiderme, as infecções causadas por *T. rubrum* podem ocorrer de forma invasiva entre pacientes imunocomprometidos (RODWELL et al., 2008; MATSUZAKI et al., 2012).

É um fungo de crescimento relativamente lento em meio Sabouraud ágar. As colônias caracterizam-se por textura cotonosa, com pregas radiais, formando uma saliência central, características similares às apresentadas pelas colônias de *T. tonsurans*. Estas colônias apresentam tonalidade branca, que pode tornar-se avermelhada e com o passar do tempo, apresentam pigmentação avermelhada no reverso (SIDRIM e ROCHA, 2004). Os

macroconídios, quando presentes, exibem clavas alongadas, quase com um aspecto cilíndrico; os microconídios são mais abundantes, pequenos e piriformes (LARONE, 1996) (Figura 6).

Figura 6: Cultura, microconídios e macroconídios de *T. rubrum* (ELLIS et al., 2007).

1.1.6 *Trichophyton tonsurans*

Trichophyton tonsurans é uma espécie antropofílica responsável por lesões no couro cabeludo (*tinea capitis*). No Brasil, ocorre com mais frequência nos estados acima do Trópico de Capricórnio e está bem adaptada às condições ambientais das regiões Norte e Nordeste, sendo um fungo endêmico na região amazônica (MORAES, 1973; FURTADO et al., 1987; COSTA et al., 1991).

Nem todas as pessoas igualmente expostas a esse agente apresentam sinais clínicos de infecção, pois esta espécie tem a capacidade de permanecer no hospedeiro humano de forma inaparente, ou seja, em estado subclínico, o que torna sua persistência e propagação bastante eficazes na população (ABDEL-RAHMAN et al., 2006).

Apesar de nichos ecológicos distintos, essa espécie possui alto grau de relação filogenética com o dermatófito *T. equinum* (MARTINEZ et al., 2012) e apresenta crescimento lento em meio Sabouraud ágar. As colônias possuem aspectos que podem variar de cotonosas a aveludadas e, com o passar do tempo, pulverulentas. O verso tem coloração variando de branco a diversos tons de bege e o reverso apresenta vários tons de castanho ao avermelhado. Os macroconídios, quando presentes, são de aspecto irregular, e os microconídios abundantes, com forma e dimensões variáveis, podendo apresentar-se com pequenas ou grandes estruturas claviformes ou de aspecto arredondado (LARONE, 1996) (Figura 7).

Figura 7: Colônias, hifas, microconídios e macroconídios de *T. tonsurans* (ELLIS et al., 2007).

1.1.7 *Arthroderma benhamiae*

O fungo *A. benhamiae* é um dermatófito zoofílico pertencente ao complexo de espécies *Trichophyton mentagrophytes*, sendo que seus principais hospedeiros são os animais domésticos. Há diversos relatos de ocorrência de *tinea corporis* em hospedeiros humanos (SHIRAKI et al., 2006; BUDIHARDJA et al., 2010).

A espécie citada possui papel importante na investigação dos fundamentos da biologia e patogenicidade dos dermatófitos (STAIB et al., 2010), e é encontrada, normalmente, em roedores, em particular, nas cobaias, manifestando crescimento relativamente rápido em condições de laboratório. Como uma espécie teleomórfica, esse fungo ainda é capaz de passar por um desenvolvimento sexual, incluindo a formação de frutificações sexuais (denominadas cleistotécios) (FUMEAUX et al., 2004; DROUOT et al., 2009; BURMESTER et al., 2011).

Suas colônias são usualmente planas, da cor branca ao creme, com superfície granular e algumas culturas apresentam dobras/tufos centrais. A pigmentação reversa vai geralmente do amarelo ao marrom avermelhado. Os macroconídios são irregulares em forma e tamanho com septações (de 2 a 6 septos) e os microconídios são alongados e anexados ao lado das hifas (Figura 8).

Figura 8: Cultura, hifa e microconídios de *A. benhamiae* (SYMOENS et al., 2013).

1.1.8 *Trichophyton verrucosum*

Trichophyton verrucosum é uma espécie cosmopolita, parasita de bovinos, que esporadicamente pode causar infecções bastante inflamatórias no couro cabeludo, pele e pelos em humanos. A frequência em humanos aumenta quando há contato com animais doentes ou presença de animais já recuperados da doença, os quais veiculam os conídios do fungo, atuando como disseminadores e perpetuadores da doença (TAKATORI et al., 1993).

Este fungo apresenta diferenças fenotípicas com *A. benhamiae*, porém, com alto grau de relação filogenética entre as espécies, além de ambos induzirem infecções cutâneas altamente inflamatórias em seres humanos, por exemplo, *tinea corporis* (FUMEAUX et al., 2004; GRUNEWALD et al., 2006; BURMESTER et al., 2011).

Seu crescimento é bastante lento, exigindo de 13 a 25 dias para a maturidade. As colônias são pequenas e circulares com relevo rugoso ou cerebriforme, apresentando pigmentação de verso branca ao amarelo ocre; no reverso, tons de amarelo, com pigmento não difusível no meio. Os macroconídios são raramente produzidos e, quando presentes, possuem uma cauda característica com forma alongada. Os microconídios são piriformes, dispostos ao longo das hifas (Figura 9).

Figura 9: Cultura, microconídios e macroconídios de *T. verrucosum* (ELLIS et al., 2007).

Estas espécies de dermatófitos foram selecionadas para serem utilizadas neste trabalho pela disponibilidade dos genomas sequenciados, importância médica e veterinária, diversidade de hospedeiros e relações filogenéticas.

1.2 Proteases

Proteases são enzimas que hidrolisam ligações peptídicas em proteínas e peptídeos. Também denominadas proteinases ou peptidases, formam um complexo grupo de enzimas que possuem características físico-químicas variadas (RAO et al., 1998).

Essas enzimas estão envolvidas em processos metabólicos essenciais, com particular importância em medicina e biotecnologia, representando em torno de 2% do total dos produtos codificados por genes (RAWLINGS e BARRETT, 1999). Possuem importante papel na nutrição por meio da hidrólise de proteínas, facilitando a absorção dos produtos pelas células, além de ter, também, papel fundamental em muitos processos patológicos (RAO et al., 1998).

As proteases são classificadas segundo alguns critérios: (i) tipo de reação catalisada (sendo endo ou exoproteases) dependendo do local de clivagem em relação às regiões terminais do substrato; (ii) pH de atividade ótima, podendo ser classificadas como proteases ácidas, neutras ou alcalinas; (iii) natureza do grupo funcional e mecanismo do sítio ativo, como proteases de serina, aspárticas, cisteínas ou metaloproteases; (iv) aspectos evolutivos relacionados à sua estrutura (RAO et al., 1998; RAWLINGS e BARRETT, 1999).

Proteases secretadas pertencentes às famílias das subtilisinas (SUB) e metaloproteinases (MEP) vem sendo investigadas, fornecendo padrões de expressão e da dinâmica dos genes em *T. rubrum*. Substratos como queratina, colágeno e elastina induziram a expressão de genes de proteases, e os padrões de expressão de proteases em meios contendo pele humana foram diferentes daqueles contendo proteína individual como substrato. De acordo com a dinâmica de expressão desses genes, sugere-se que Sub3, Sub4 e MEP4 podem ser as principais proteases secretadas por *T. rubrum* durante a infecção do hospedeiro, e que estas proteases podem ser importantes alvos para novos agentes antifúngicos e marcadores de diagnósticos moleculares (LENG et al., 2009).

A superexpressão de genes que codificam para Sub3, Sub5, Mep3 e Mep4 foi revelada, após o cultivo de *T. rubrum* em queratina, comparada com o cultivo em glicose, por meio da construção de bibliotecas de hibridização subtrativa por supressão (*Suppression Subtractive Hybridization* - SSH) (MARANHÃO et al., 2007).

Baldo et al. (2010) demonstraram, por meio de experimento realizado com *M. canis*, que uma linhagem com o gene *sub3* silenciado tem dramática perda de capacidade de

aderência aos corneócitos felinos, quando comparado com a linhagem selvagem. Os autores, ainda, promoveram a aderência artificialmente, sendo possível observar então a inexistência de diferenças significativas na capacidade infectante entre a linhagem selvagem e a linhagem *sub3* silenciada, indicando que a secreção de Sub3 é necessária para adesão.

No dermatófito zoofílico *A. benhamiae* também foi observada maior expressão dos genes que codificam para as subtilisinas, Sub3 e Sub4, das fungalisinas, Mep1, Mep3 e Mep4, das exoproteases, leucina-aminopeptidases Lap1 e Lap2, Dipeptidil peptidases DppIV e DppV, metalocarboxipeptidase MCPA e serina carboxipeptidase SCPB, durante crescimento *in vitro* em queratina e soja, quando comparados ao crescimento em meio Sabouraud. (STAIB et al., 2010).

Em *M. canis*, foram isolados três tipos de subtilisinas (*sub1*, *sub2* e *sub3*) e, por meio de análise de sequências, verificou-se que esses genes apresentam alta identidade com as serino proteases de outros fungos filamentosos (DESCAMPS et al., 2002).

Conforme descrito anteriormente, os dermatófitos são capazes de utilizar a queratina como fonte de nutrientes, invadindo tecidos queratinizados em humanos e animais, além disso, algumas espécies geofílicas utilizam debris de queratina (descamação) existentes no solo. As queratinas são proteínas estruturais fibrosas e insolúveis, que compõem pele, pelos, unhas, cascos, chifres, penas e escamas. São proteínas dos filamentos intermediários, os quais são preferencialmente expressos em células epiteliais e apêndices epidérmicos (FUCHS e WEBER, 1994; COULOMBE e OMARY, 2002).

Aparentemente, o conteúdo de ligações dissulfeto é o principal determinante das características de resistência e estabilidade das queratinas. Sendo assim, as queratinas são classificadas como tipo I e tipo II de acordo com o conteúdo de cisteína e ligações dissulfídicas. As primeiras, tipo I (também denominadas como rígidas), possuem alto conteúdo de cisteína, e as do tipo II (conhecidas como macias) possuem menor conteúdo de cisteína (MCLEAN e MOORE, 2011).

Queratinas encontradas em tecidos epiteliais são classificadas como tipo II e apresentam conteúdo de cisteína inferior a 10% em sua estrutura (GRADISAR et al., 2005); como oposição, temos as queratinas classificadas como tipo I, encontradas em componentes de cabelos e unhas, demonstrando taxas maiores (em torno de 17%) no conteúdo de cisteína.

1.3 Processo de interação patógeno-hospedeiro

Os dermatófitos, por meio da germinação de artroconídios, seguidos pela disseminação das hifas no estrato córneo do hospedeiro, aderem-se à superfície do tecido iniciando, assim, o processo infeccioso. Estudos com *T. mentagrophytes* explicitaram que, durante a adesão de artroconídios à superfície do estrato córneo, estruturas fibrilares longas parecem ancorar e conectar o artroconídio à superfície tecidual. Em camadas internas da pele, essas estruturas fibrilares se tornam mais finas e curtas, cobrindo toda a superfície do artroconídio, e geram uma morfologia achatada, fazendo com que ocorra um aumento na superfície de contato com o tecido, além de maior adesão e aquisição de nutrientes (VERMOUT et al., 2008).

Nos estágios iniciais da infecção, em contato com o pH ácido da pele, o patógeno sintetiza proteases e queratinases não específicas, que preferencialmente atuam em pH ácido. Essas enzimas agem tanto em substratos queratinosos, quanto em não queratinosos, gerando peptídeos hidrolisados em aminoácidos e, então, utilizados pelo fungo como fonte de carbono, nitrogênio e enxofre. Alguns aminoácidos metabolizados, como no caso da glicina, fazem com que ocorra a liberação de amônia pelo fungo, elevando o pH do meio. Esse mecanismo permite a ação das queratinases com atividade ótima em pH alcalino, o que possibilita a manutenção da infecção (MARTINEZ-ROSSI et al., 2004; PERES et al., 2010a).

A secreção de enzimas proteolíticas pelos dermatófitos é um fator chave na invasão e utilização do estrato córneo do hospedeiro (MONOD, 2008; MARANHÃO et al., 2009; GRUMBT et al., 2011). O sequenciamento dos genomas de sete dermatófitos revelou em todos eles a presença de um grande número de genes que codificam proteases (BURMESTER et al., 2011; ACHTERMAN e WHITE, 2012; MARTINEZ et al., 2012). No entanto, na queratina dura, presente na unha e cascos, as cadeias peptídicas se encontram firmemente ligadas entre si, formando uma estrutura compacta devido às pontes dissulfeto entre os resíduos de cisteína, altamente abundantes nesta proteína. A digestão proteolítica deste tipo de queratina ocorre após esta estrutura ser enfraquecida pela redução das pontes dissulfeto. Um provável agente redutor é o sulfito, secretado pelos dermatófitos durante o crescimento em queratina, que cliva a molécula em cisteína e S-sulfocisteína (PARK e BAKALINSKY, 2000; LECHENNE et al., 2007; KASPEROVA et al., 2011; GRUMBT et al., 2013). Na Figura 10, é representada a via de degradação de tecidos queratinizados pelos dermatófitos.

Figura 10: Representação da via de degradação de tecidos queratinizados pelos dermatófitos em pH neutro (BALDO et al., 2012).

A capacidade de utilização de queratina, proteína insolúvel, como fonte única de nutrientes, é um dos maiores atributos na virulência dos dermatófitos. Portanto, genes diferencialmente expressos durante a degradação de queratina são alvos atraentes para caracterização funcional, pois ainda pouco se sabe sobre os mecanismos moleculares envolvidos na adaptação ao microambiente durante a interação patógeno-hospedeiro e no processo infeccioso das dermatofitoses. Desta maneira, a degradação de queratina, bem como outros substratos presentes no microambiente hospedeiro, deve ser um processo altamente regulado, que envolve a ativação de diferentes vias de sinalização e fatores de transcrição para garantir que estas enzimas sejam expressas e secretadas corretamente e em resposta ao microambiente hospedeiro.

1.4 Genômica comparativa

Sequências genômicas completas constituem uma fonte de dados relevante porque, em princípio, elas representam tudo o que é necessário para criar um organismo, juntamente com fatores epigenéticos e sua interação com eles (CATANHO et al., 2007). Acredita-se que o

genoma de um organismo, isoladamente, fora de seu contexto evolutivo, não é capaz de nos revelar muito sobre si e que, para obter mais informações, os genomas de diferentes espécies devem ser estudados comparativamente (CLARK, 1999). De fato, análises comparativas entre genomas de diferentes micro-organismos têm contribuído enormemente para a elucidação de aspectos fundamentais da genética, da bioquímica e da evolução de inúmeras espécies (GALPERIN e KOONIN, 1999; KONDRASHOV, 1999; FRASER et al., 2000; GALPERIN e KOONIN, 2000; KOONIN et al., 2000; WEI et al., 2002; HUYNEN et al., 2005; ABBY e DAUBIN, 2007).

Existem diferenças entre os genomas, mesmo em organismos evolutivamente próximos, e tais diferenças são produzidas por processos genéticos, como substituições nucleotídicas, inserções, deleções e rearranjo de grandes fragmentos de DNA. O papel dos genes e de seus produtos pode ser inferido pela comparação de moléculas relacionadas, detectando as diferenças entre elas e computando informações sobre estas diferenças (MUSHEGIAN, 2007).

Dentro desse conjunto de metodologias que visam à identificação de diferenças e semelhanças entre os genomas, existem aquelas que realizam análise comparativa entre genomas de organismos não patógenos com organismos patógenos para tentar identificar os genes ou grupos de genes associados a doenças e infecções (KELLIS et al., 2004). A genômica comparativa está relacionada à predição das funções de proteínas por meio da transferência de informações de proteínas funcionalmente caracterizadas de organismos modelos e à inferência de relações filogenéticas entre os organismos. Algumas aplicações incluem: identificação de genes únicos de uma espécie em particular; identificação de fatores de virulência; reconstrução metabólica (GORDON et al., 2002); caracterização de patógenos; identificação de novos alvos para diagnóstico e para procedimentos terapêuticos (FITZGERALD e MUSSER, 2001); investigação sobre a origem molecular da patogênese; entre outras. A qualidade da inferência desses resultados depende da sensibilidade e robustez de métodos computacionais.

Genômica comparativa entre dermatófitos é um tema relativamente novo. Até 2007 apenas 524 sequências de nucleotídeos relacionados a estes organismos estavam disponíveis em bancos de dados públicos (YANG et al., 2007).

Com a iniciativa do *Broad Institute* de realizar o sequenciamento de cinco linhagens de dermatófitos, que inclui espécies de alta incidência em infecções humanas (MARTINEZ et

al., 2012), somadas a outras duas espécies previamente sequenciadas (BURMESTER et al., 2011) e, mais recentemente, com a colaboração do grupo do Laboratório de Genética e Biologia Molecular de Fungos do Departamento de Genética - FMRP/USP (com contribuições, também, do presente trabalho), no sequenciamento de uma das linhagens de *T. interdigitale*, está sendo possível analisar, do ponto de vista comparativo, eventos de ganhos e perdas de genes entre os fungos, além de sugerir candidatos responsáveis por funções de infecção e escolha de hospedeiros, o que pode vir a ajudar na orientação do desenvolvimento de novas terapias (MARTINEZ et al., 2012). A disponibilidade dos genomas dos dermatófitos aumentará e acelerará o número de estudos no âmbito molecular, elucidando as diferenças evolutivas, fisiológicas e adaptativas das espécies.

Após a comparação do genoma de 7 dermatófitos com o de outros fungos foi revelado que genes pertencentes a quatro grupos gênicos associados à patogenicidade estavam super representados nos dermatófitos. São eles: (i) proteases secretadas, atuando como fator de virulência no processo de degradação de substratos queratinizados; (ii) quinases, que estão envolvidas na sinalização necessária para adaptação à pele e à unha; (iii) genes envolvidos no metabolismo secundário, que atuam também na sinalização na interação entre fungo e hospedeiro; (iv) classe de proteínas (*LysM*), as quais possivelmente se ligam a carboidratos para mascarar componentes da parede celular e hidratos de carbono, evitando a resposta imune do hospedeiro (MARTINEZ et al., 2012).

Proteases secretadas estão entre as famílias de genes mais importantes para os dermatófitos. Estudos sobre genômica comparativa de dermatófitos explicitaram que, para a família das subtilisinas, a maioria dos dermatófitos apresenta 12 subtilisinas. Por outro lado, linhagens do gênero *Aspergillus* possuem apenas de 2 a 4 cópias de genes que codificam subtilisinas. O mesmo ocorre com as fungalisinas (família das metaloendopeptidases) presentes em pelo menos 5 cópias em dermatófitos, porém com 1 cópia, ou até mesmo, ausentes em outros fungos não dermatófitos. É importante destacar que, mesmo entre dermatófitos, diferenças no número de cópias de proteases secretadas foram observadas (MARTINEZ et al., 2012).

Os genomas dos dermatófitos sequenciados possuem tamanhos similares: *T. rubrum* - 22,5Mb (Mega bases); *T. tonsurans* - 23,0Mb; *T. equinum* - 24,1Mb; *M. canis* - 23,1Mb; *M. gypseum* - 23,2Mb; *T. verrucosum* - 22,5Mb; *A. benhamiae* - 22,2Mb; *T. interdigitale* - 21,9Mb (BROAD_INSTITUTE, 2015). Além disto, as espécies de dermatófitos têm

diferenças sutis tanto no conteúdo gênico quanto na organização do genoma. Apesar disto, cada fungo se adaptou a nichos específicos de hospedeiro e de substrato protéico, sugerindo que a regulação de expressão gênica possa ter um papel relevante nestas preferências (MARTINEZ et al., 2012; MARTINS, 2015).

Os genomas dos dermatófitos apresentam uma taxa muito baixa de polimorfismo de nucleotídeo único (SNP), alta similaridade no conteúdo gênico, pequeno número de genes exclusivos em cada espécie (801 no genoma de *T. rubrum*, 340 no genoma de *T. tonsurans*, 467 no genoma de *T. equinum*, 943 no genoma de *M. canis*, 937 no genoma de *M. gypseum*, 532 no genoma de *T. verrucosum* e 480 no genoma de *A. benhamiae*) e um conjunto de 6.168 grupos ortólogos comuns a sete dos oito dermatófitos sequenciados e discutidos neste trabalho (BURMESTER et al., 2011; MARTINEZ et al., 2012). É importante destacar que grande parte das linhagens utilizadas no sequenciamento foi isolada a menos de 10 anos, limitando potenciais perdas ou modificações nas regiões do genoma, associadas à virulência ou a outro determinante genético relacionado ao potencial patogênico (WHITE et al., 2008).

Dessa forma, dada a importância clínica dos dermatófitos e o ainda escasso conhecimento dos fatores que governam o processo patogênico das dermatofitoses, pretende-se, aqui, avaliar alguns genes com potencial para virulência (envolvidos no processo de degradação da queratina) nestes organismos. Como estratégia de análise de genômica comparativa serão utilizadas ferramentas computacionais, além do perfil de expressão gênica. Os resultados poderão contribuir para uma maior compreensão do envolvimento desses genes na especificidade da infecção dos dermatófitos estudados, podendo revelar novos alvos terapêuticos.

Hipótese do Trabalho

There are many hypotheses in science which are wrong. That's perfectly all right: it's the aperture to finding out what's right. Science is a self-correcting process.

Carl Sagan

II. Hipótese do Trabalho

Um número crescente de genomas completamente sequenciados de diversos organismos é depositado anualmente em bancos de dados públicos. Em posse dessas sequências, já foram realizadas várias análises comparativas entre genomas de espécies próximas, ou genomas de linhagens diferentes e pertencentes a uma mesma espécie. Essas análises mostram que a variabilidade genômica encontrada nesses casos pode estar associada a eventos de rearranjos, como deleções e inserções, que ocorrem entre grandes regiões desses genomas e, também, às modificações pontuais, como SNPs (*Single Nucleotide Polymorphisms*) ao longo das sequências. Existe um grande número de exemplos nos quais regiões (que variam entre 10 a 200 kbp (quilobase) presentes em linhagens patogênicas, mas ausentes em membros não patogênicos da mesma espécie ou de espécie relacionada) são determinantes da virulência desses organismos.

Sabe-se que são sutis as diferenças no conteúdo gênico encontradas entre as espécies de dermatófitos disponibilizadas neste trabalho; ao mesmo tempo, cada uma dessas espécies possui adaptação a nichos diferentes, apresentando variedade de hospedeiros e preferência diferenciada em substratos protéicos. Por isso, estruturou-se a hipótese de avaliar se as diferenças encontradas entre esses dermatófitos estão associadas às modificações pontuais no conteúdo gênico ou, até mesmo, às modificações nas regiões promotoras (fundamentais para o processo de transcrição), presentes na estrutura de genes envolvidos no processo de degradação da queratina. Além disso, é relevante avaliar se os padrões de expressão desses genes contribuem para a especificidade de hospedeiros.

Objetivos

O sucesso nasce do querer, da determinação e persistência em se chegar a um objetivo.

Mesmo não atingindo o alvo, quem busca e vence obstáculos, no mínimo fará coisas admiráveis.

José de Alencar

III. Objetivos

Devido à disponibilização dos genomas de sete espécies de dermatófitos pelo *Broad Institute/NIH*, este trabalho visou avaliar, por meio de genômica comparativa, as semelhanças e diferenças no conteúdo de genes que codificam proteínas envolvidas no processo de degradação da queratina e seus perfis transcricionais em dermatófitos.

Estratégias

- Realizar o sequenciamento em larga escala do genoma de uma linhagem de *T. interdigitale*. Após essa etapa, fazer a montagem genômica utilizando técnicas de bioinformática.
- Avaliar por meio de técnicas de genômica comparativa genes que codificam para as proteínas: Sub3, Sub5 e Sub7, serina endopeptidases com atividade queratinolítica; Mep3 e Mep4, do grupo das metaloendopeptidases; Dipeptidil peptidase V (DppV), Leucina aminopeptidase 1 (Lap1) e Leucina aminopeptidase 2 (Lap2), do grupo das exopeptidases. Para isso, comparar os seguintes dermatófitos: *A. benhamiae*, *M. canis*, *M. gypseum*, *T. equinum*, *T. interdigitale*, *T. rubrum*, *T. tonsurans* e *T. verrucosum*.
- Desenvolver algoritmos computacionais que automatizem a análise comparativa de regiões específicas entre os dermatófitos em estudo.
- Avaliar, na presença de queratina, o perfil transcricional por PCR em tempo real, de alguns genes envolvidos no processo de degradação da queratina. Comparar as espécies *M. canis*, *T. equinum*, *T. interdigitale*, *T. rubrum* e *T. tonsurans* com o objetivo de revelar estratégias moleculares específicas ou comuns a esses fungos durante seu desenvolvimento.

Materiais e Métodos

Por vezes, sentimos que aquilo que fazemos não é senão uma gota de água no mar. Mas o mar seria menor se lhe faltasse uma gota.

Madre Teresa de Calcutá

IV. Materiais e métodos

4.1 Linhagens

As sequências genômicas e proteicas das seguintes linhagens foram utilizadas neste trabalho:

- *T. rubrum* (CBS118892), *T. tonsurans* (CBS112818), *T. equinum* (CBS127.97), *M. canis* (CBS113480) e *M. gypseum* (CBS118893), cedidas pelo *Centraalbureau voor Schimmelcultures* (CBS) e sequenciadas pelo *Broad Institute*;
- *A. benhamiae* (CBS112371) e *T. verrucosum* (HKI0517), sequenciadas pelo consórcio *Leibniz Institute for Natural Product Research and Infection Biology – Hans Knoell Institute (HKI)*, em cooperação com o *Leibniz Institute for Age Research, Fritz Lipmann Institute (FLI)* e *Friedrich Schiller University Jena (Institute of Microbiology and Department of Dermatology and Allergology)*;
- *T. interdigitale* (H6) (ATCC MYA-3108) isolada de paciente do Hospital das Clínicas de Ribeirão Preto da FMRP-USP e caracterizada segundo (MCGINNIS, 1980) (linhagem cedida pela Profa. Dra. Cláudia M. Leite Maffei, docente do Departamento de Biologia Celular e Molecular e Bioagentes Patogênicos da FMRP-USP).

Os genomas das linhagens aqui utilizadas, incluindo *T. interdigitale* (H6), cujo sequenciamento pertence à lista de objetivos específicos deste trabalho, estão completos ou parcialmente disponíveis no site do *Broad Institute*: http://www.broadinstitute.org/annotation/genome/dermatophyte_comparative/GenomesIndex.html. O sequenciamento genômico da linhagem *T. interdigitale* (H6) será descrito nas próximas seções.

4.2 Meios de cultura

Meio Extrato de Malte Ágar (MEA) (Tabela 1).

Tabela 1: Meio Extrato de Malte-Ágar (MEA)

Descrição	Quantidade
Glicose	2% (m/v)
Extrato de malte	2% (m/v)
Peptona	0,1% (m/v)
Ágar	2% (m/v)
Água destilada (q.s.p.)	1.000 mL

O pH final do meio foi ajustado em 5,7. Autoclavado a 1 atm de pressão e 120°C por 20 minutos (ATLAS, 1993).

Meio com queratina

A queratina utilizada provém de cascos de bois raspados com auxílio de esmerilhadeira até formar um pó fino. Foram autoclavados 0,2 g de queratina (0,4% (m/v) final) a 1 atm de pressão e 120°C por 15 min. Adicionaram-se 50 mL de água destilada estéril por frasco de cultivo e o pH individualmente ajustado em 5 ($\pm 0,1$).

4.3 Manutenção e estoque das linhagens

As linhagens foram estocadas em tubos de ensaios contendo meio Extrato de Malte-Ágar (MEA) e mantidas à temperatura ambiente. Precederam-se repiques periódicos das linhagens em placas de Petri contendo meio MEA (Tabela 1).

Para o sequenciamento do genoma de *T. interdigitale*, o cultivo inicial foi feito em placa de Petri contendo MEA, a 28°C, por 10 a 15 dias.

Para a análise do perfil transcricional comparativo por PCR em tempo real, os cultivos iniciais foram feitos em placas de Petri contendo meio MEA a 28°C, por 15 a 17 dias para as linhagens do gênero *Trichophyton*, e de 13 a 15 dias para o gênero *Microsporum*.

4.4 Preparo das soluções de conídios

Para cada linhagem, após período de crescimento em meio sólido, a massa micelial foi coletada com o auxílio de espátula esterilizada e adicionada em 10 mL de solução salina 0,9% (m/v). Em seguida, a suspensão foi agitada vigorosamente, com auxílio de vórtex para desagregar os conídios, filtrada em funil com lã de vidro, esterilizada para remover as hifas e centrifugada a 4.000 x g por 10 min. Ressuspendeu-se o sedimento em 2 mL de solução salina (0,9%) e a concentração de conídios foi estimada em hematímetro (câmara de Neubauer).

4.5 Sequenciamento e montagem do genoma de *T. interdigitale* (H6)

4.5.1 Extração de DNA genômico

Para a extração do DNA genômico da linhagem *T. interdigitale* (H6), 100 µL de uma suspensão de conídios contendo 1×10^7 conídios/mL foi cultivado em MEA líquido por 72h a 30°C. A massa micelial foi, então, filtrada em funil de Buchner, lavada em solução fisiológica e macerada em nitrogênio líquido. Depois, adicionou-se a cada 100 mg de micélio 1 mL de solução de lise (50 mM EDTA pH 8,0; 1% SDS), e a mistura foi incubada a 68°C por 15 min. Após centrifugação a 10.000 x g por 10 min., acrescentaram-se 40 µL de acetato de potássio 5 M, incubado em gelo por 1h e, a seguir, centrifugado novamente a 10.000 x g por 10 min. A solução contendo o DNA foi transferida para um tubo limpo com 1 mL de isopropanol; coletou-se, então, o DNA com o auxílio de uma pipeta Pasteur em forma de gancho e ele foi lavado com etanol 70%, seco e eluído em tampão TE contendo RNase (2 µg/mL).

4.5.2 Construção da biblioteca de sequenciamento

A amostra de DNA genômico da linhagem *T. interdigitale* (H6) foi quantificada por meio de fluorescência, utilizando o equipamento Qubit (*life technologies*). Para a construção da biblioteca de sequenciamento, utilizou-se 1 ng de DNA em uma concentração de 0,2 ng/µL como entrada para o kit Nextera XT (Illumina) de acordo com as instruções do fabricante. Em

resumo, esse kit usa um sistema de transposomos (complexo transposase/transposon) para simultaneamente fragmentar e “tagmentar” o DNA. Uma reação de PCR é realizada para amplificação desses fragmentos e para a ligação de índices que permitem o sequenciamento multiplexado (Figura 11).

Figura 11: Esquema representativo da metodologia utilizada pelo kit Nextera XT para construção de biblioteca genômica (ILLUMINA, 2012).

4.5.3 Sequenciamento em larga escala

O sistema MiSeq (Illumina) é baseado na tecnologia de sequenciamento por síntese. O DNA é fragmentado aleatoriamente e ligado a adaptadores em ambas as extremidades (Figura

12-A). As moléculas de DNA fita simples são aderidas por afinidade ao suporte sólido onde estão também aderidos em alta densidade oligonucleotídeos complementares aos adaptadores (Figura 12-B). Durante a etapa de anelamento (Figura 12-C), no primeiro ciclo de amplificação da PCR em fase sólida, o adaptador da extremidade livre da molécula aderida ao suporte encontra seu oligonucleotídeo complementar no suporte, formando uma estrutura em ponte. Uma vez fornecidos os reagentes necessários, a PCR é iniciada utilizando a extremidade 3' livre do oligonucleotídeo como primer (Figura 12-D). Na etapa de desnaturação (Figura 12-E), a “ponte” é desfeita mediante elevação de temperatura. Repete-se, então, a etapa de anelamento (Figura 12-F), formando novas estruturas em ponte e iniciando um novo ciclo de amplificação. Após uma série desses ciclos, serão obtidos *clusters* de moléculas idênticas ligadas ao suporte (Figura 12-G). Com a incorporação de nucleotídeos terminadores marcados e excitação a laser (Figura 12-H), é gerado sinal, o qual é captado por dispositivo de leitura e interpretado como um dos quatro possíveis nucleotídeos componentes da cadeia (Figura 12-I). Os processos de incorporação de nucleotídeo marcado, excitação e leitura são repetidos para cada nucleotídeo componente da sequência (Figura 12-J, Figura 12-K). A leitura é feita de forma sequencial, o que permite a montagem da sequência completa de cada *cluster* (Figura 12-L).

Figura 12: Esquema representativo da metodologia utilizada pelo MiSeq (ILLUMINA, 2010).

4.5.4 Montagem do genoma

Foram usadas duas estratégias de montagem, a montagem “*de novo*” e a montagem por referência. Na montagem por referência, o *software MiSeq Reporter* (Illumina - company) foi utilizado para a extração das *reads* de leitura e, posteriormente, o alinhamento com o genoma de referência, e, nesse caso, optou-se por utilizar *T. rubrum* (CBS118892) disponível no site do *Broad Institute*, já que até o momento não havia disponível o genoma ou parte dele para a espécie *T. interdigitale*.

Em seguida, para obter-se uma comparação com outra estratégia de montagem, realizou-se a montagem “*de novo*”, na qual, com as *reads* já extraídas pelo *software MiSeq Reporter*, fez-se o uso do *software Velvet* (ZERBINO e BIRNEY, 2008), que realiza a montagem de *reads* curtos utilizando-se de grafos de Bruijn, útil na eliminação de erros e repetições.

4.6 Análises de genômica comparativa dos dermatófitos

4.6.1 Seleção dos genes de interesse

Foram selecionadas 5 proteínas identificadas em dermatófitos, provavelmente envolvidas na virulência fúngica (PERES et al., 2010b), e outras 3 proteínas identificadas em um secretoma de *A. benhamiae* obtido a partir do sobrenadante de cultura após 48 horas de cultivo com glicose e queratina (BURMESTER et al., 2011). Como fase inicial deste trabalho, as sequências dos genes que codificam cada uma das proteínas listadas na Tabela 2 foram capturadas do banco de dados do *Broad Institute*. Para cada um desses genes, buscou-se a identificação dos genes ortólogos entre os dermatófitos em estudo com o uso da ferramenta *BLASTn* (ALTSCHUL et al., 1990).

Tabela 2: Lista dos genes/proteínas utilizados no estudo com seus respectivos códigos de identificação.

Gene/Proteína	<i>A. benhamiae</i>	<i>M. canis</i>	<i>M. gypseum</i>	<i>T. equinum</i>	<i>T. interdigitale</i> (H6)	<i>T. rubrum</i>	<i>T. tonsurans</i>	<i>T. verrucosum</i>
Dipeptidil peptidase V (DppV)	ARB_06651	MCYG_01626	MGYG_04446	TEQG_02326	H101_03870	(*)	TESG_00772	TRV_02418
Leucina aminopeptidase 1 (Lap1)	ARB_03568	MCYG_02286	MGYG_08250	TEQG_06332	H101_03813	TERG_05652	TESG_07124	TRV_06599
Leucina aminopeptidase 2 (Lap2)	ARB_00494	MCYG_06199	MGYG_06635	TEQG_05395	H101_04588	TERG_08405	TESG_06750	TRV_01590
Metaloprotease 3 (Mep3)	ARB_05085	MCYG_04620	MGYG_08200	TEQG_02420	H101_04475	TERG_03248	TESG_02775	TRV_06691
Metaloprotease 4 (Mep4)	ARB_00762	MCYG_02088	MGYG_03984	TEQG_04323	H101_01910	TERG_04324	TESG_00401	TRV_00081
Protease subtilisina 3 (Sub3)	ARB_00701	MCYG_04040	MGYG_02570	TEQG_06704	H101_01018	TERG_03815	TESG_03300	TRV_07976
Protease subtilisina 5 (Sub5)	ARB_02223	MCYG_08753	MGYG_07290	TEQG_05834	H101_05652	TERG_08201	TESG_07246	TRV_00550
Protease subtilisina 7 (Sub7)	ARB_06076	MCYG_02345	MGYG_07342	TEQG_06283	H101_00208	TERG_12327	TESG_06365	TRV_00296

(*)*T. rubrum* CBS 118892: Supercontig 11: 290252-292693 +

4.6.2 Presença/ausência e conservação entre os genes

Por meio da ferramenta *BLASTn* (ALTSCHUL et al., 1990) e do banco de dados disponível no *Broad Institute*, foram verificados quais dos 8 genes em estudo estavam presentes entre os 8 dermatófitos analisados. Nessa fase de análise, também foi possível avaliar o grau de conservação entre cada uma das espécies, usando como base as sequências gênicas estudadas.

4.6.3 Análise dos genes (identificação das regiões conservadas)

Para determinar as regiões comuns, mapear e analisar as alterações entre cada um dos genes das espécies de dermatófitos estudadas, os alinhamentos múltiplos entre sequências foram construídos com a utilização da ferramenta *MALVE* (DARLING et al., 2004).

4.6.4 Análise das regiões não codantes

Sabe-se que regiões não codantes podem ser importantes para a regulação da expressão gênica. Sendo assim, fez-se uma análise das regiões de íntrons de cada um dos genes.

Os arquivos em formato GFF (*General Feature Format*) disponibilizados pelo *Broad Institute* foram utilizados e é importante destacar que GFF é um formato padrão de arquivo texto para armazenamento de características genômicas. Além disso, desenvolveu-se um algoritmo na linguagem *Perl* (*Practical Extraction and Reporting Language*) com o objetivo de extrair dados importantes desses arquivos. Tal algoritmo foi utilizado para, por meio dos dados extraídos dos arquivos em formato GFF, gerar um arquivo que sirva de entrada para o *software Fancygene* (RAMBALDI e CICCARELLI, 2009). Esse software foi então utilizado para criar a representação gráfica das estruturas de cada um dos genes.

4.6.5 Análise das proteínas (comparação do conteúdo protéico)

Os alinhamentos das proteínas foram feitos com o auxílio do *software ClustalW* (THOMPSON et al., 1994), as sequências foram editadas no *software Bioedit* (<http://jwbrown.mbio.ncsu.edu/BioEdit/bioedit.html>) e, para cada um dos alinhamentos, construiu-se uma matriz identidade.

4.6.6 Análise das relações filogenéticas

As sequências das proteínas de todas as subfamílias de interesse foram alinhadas a fim de verificar a similaridade delas entre as oito espécies estudadas. Os alinhamentos foram efetuados com o auxílio do programa *ClustalW* (THOMPSON et al., 1994) e, para a inferência de informação filogenética, utilizou-se o programa *Mega 5.0* (TAMURA et al., 2011). Também usou-se o método de Máxima Verossimilhança com modelo de substituição de aminoácidos LG, com *Bootstrap* de 1.000 repetições para a construção dos cladogramas (FELSENSTEIN, 1985).

4.6.7 Análise da sintenia genômica

Com o objetivo de identificar agrupamentos, ordem dos genes e rearranjos gênicos, desenvolveu-se um algoritmo em linguagem *Perl*, que permite identificar o *supercontig* ao qual um determinado gene pertence. Esses *supercontigs* são, então, separados por subfamílias para serem submetidos a um alinhamento múltiplo por meio do software *MALVE* (DARLING et al., 2004).

4.6.8 Análise de variantes (SNPs)

Após a seleção das regiões que se pretendia analisar (éxons, íntrons e 1.000 pares de base *upstream* de cada gene), um algoritmo foi desenvolvido almejando-se gerar arquivos VCF (*Variant Call Format*) contendo as variações encontradas nas comparações entre as diversas possibilidades de pares entre os organismos em estudo. Um filtro desenvolvido por

meio de um algoritmo foi aplicado, visando identificar as variações de melhor qualidade (eliminação de bases representadas como N). Depois, os arquivos VCF e GFF foram submetidos a um segundo processo com o objetivo de gerar tabelas com o número de variações encontradas para cada uma das combinações selecionadas.

4.6.9 Análise de sítios de ligação nas regiões promotoras

As buscas por motivos de ligação foram feitas com o auxílio das ferramentas *Yeasttract*, disponível em <http://www.yeasttract.com/formtfsbindingsites.php> e *MEME Suite*, disponível em <http://meme.nbcr.net/meme/>.

Nas buscas com o pacote *MEME Suite*, sequências 1.000 pares de base *upstream* de cada um dos genes e de cada um dos organismos, incluindo dois outros organismos externos (*Coccidioides immitis* e *Coccidioides posadasii*), foram alinhadas visando encontrar sequências consenso de tamanho 6-15 pares de base. Esses consensos foram, então, analisados com o objetivo de verificar especificidades entre gêneros, espécies e nichos hospedeiros.

4.6.10 Algoritmos em Bioinformática

Vários foram os algoritmos construídos no decorrer deste trabalho e alguns focaram realizar o *parsing* (análise sintática ou processo que analisa uma sequência de entrada para determinar sua estrutura gramatical segundo uma determinada gramática formal). E essa análise sintática transforma um texto de entrada em uma estrutura de dados. Outros, visando identificar padrões por expressões regulares (provê uma forma concisa e flexível de identificar cadeias de caracteres de interesse em sequências de DNA, por exemplo). Na Figura 13, temos um exemplo de algoritmo desenvolvido, cujo foco é selecionar sequências de interesse em um arquivo de formato FASTA contendo a lista completa de genes/proteínas de um organismo.

```
#!/usr/bin/perl
use Bio::SeqIO;
my $seqio = Bio::SeqIO->new(-file => $ARGV[1], '-format' => 'Fasta');
open(INFO, ">$ARGV[2]");

while(my $seq = $seqio->next_seq) {
    @idseqs = `cat $ARGV[0]`;
    foreach $idseq (@idseqs)
    {
        chomp $idseq;
        if($seq->desc =~ /$idseq/)
        {
            if($seq->desc =~ m/$ARGV[3]/) {
                $pos = $-[0];
            }
            $desc = substr($seq->desc,$pos);
            print INFO ">" . $idseq . " " . $desc . "\n";
            print INFO $seq->seq . "\n";
        }
    }
}
close(INFO);
```

Figura 13: Exemplo de algoritmo desenvolvido em linguagem *Pearl*.

4.7 Perfil transcricional comparativo por PCR em tempo real

4.7.1 Soluções

4.7.1.1 Soluções para RNA

Água livre de RNase (SAMBROOK et al., 1989)

A água utilizada durante extração e preparo de soluções para RNA foi tratada com dietilpirocarbonato (DEPC) 0,1% (v/v) a 37°C, por 18 a 20h, tendo sido, posteriormente, autoclavada por 30 min. a 1 atm de pressão e 120°C.

Tabela 3: Tampão de Migração MOPS 5X

Descrição	Quantidade
MOPS pH 7,0	100 mM
Acetato de Sódio	40 mM
EDTA pH 8,0	5 mM

O tampão foi preparado com água livre de RNases e mantido em frasco âmbar à temperatura ambiente.

4.7.1.2 Outras Soluções

Tabela 4: Tampão TAE (Tris-Acetato-EDTA) (SAMBROOK et al., 1989)

Descrição	Quantidade
Tris-base	242 g
Ácido acético glacial	57,1 mL
EDTA 0,5M	100 mL
Água destilada (q.s.p)	1.000 mL

pH final do meio ajustado em 8,0. Solução autoclavada a 1 atm de pressão e 120°C por 20 min.

4.7.2 Condições de cultivo

Para cada linhagem, aproximadamente, 1×10^6 conídios foram inoculados em 50 mL de meio contendo H₂O acrescido de queratina bovina 0,4% (m/v) e mantidos sob agitação (100 rpm) a 28°C. Incubaram-se os conídios por 24 e 96 horas, e, depois disso, em cada um desses períodos, o micélio obtido foi filtrado, seco em papel de filtro, congelado em nitrogênio líquido e armazenado a -80°C até a extração de RNA.

Depois de cada tempo de cultivo, fez-se a verificação do pH do meio de cultura em pHmetro e os experimentos foram efetuados em triplicata.

4.7.3 Extração de RNA total

O RNA total foi extraído com 1 mL de TRI Reagent® (Sigma-Aldrich, cat.#T9424) e adicionado a, aproximadamente, 20 mg da massa micelial seca obtida; as amostras, maceradas e mantidas à temperatura ambiente por 5 min. Posteriormente, acrescentou-se 0,3 mL de clorofórmio a cada amostra, submetida a um Vórtex por 10 s., seguindo de incubação por 10 min., à temperatura ambiente. Centrifugaram-se, então, as amostras a 12.000 x g por 15 min., a 4°C, a fase aquosa foi transferida para um novo tubo e submeteram-nas novamente à etapa de clorofórmio. O RNA foi precipitado com 0,25 mL de isopropanol por 10 min., à temperatura ambiente. As amostras foram centrifugadas a 12.000 x g por 15 min., a 4°C. Descartou-se o sobrenadante e o RNA sedimentado foi lavado com 1 mL de etanol 75% (v/v), passando por uma nova etapa de centrifugação a 8.000 x g por 5 min., a 4°C. Descartado o

sobrenadante, o RNA foi ressuspensionado em 40 µL de água deionizada, livre de RNase e incubado por 5 min., a 65°C. A integridade do RNA foi verificada em gel de agarose 1,5% (m/v) em condição desnaturante e a concentração determinada por meio da leitura de absorbância a 260 nm. As amostras foram congeladas a -80°C até sua utilização.

4.7.4 Síntese de cDNA

Para a síntese de cDNA, 1 µg do RNA total de cada uma das condições em análise foi tratado com 1 U de enzima Deoxyribonuclease I (DNase I), RNase-free (*Thermo Scientific, CA, USA, cat.#EN0521*) e convertido em cDNA fita simples, utilizando o *Kit High Capacity cDNA Reverse Transcription Kit* (*Applied Biosystems, Foster City, CA, USA, cat.# 4368814*), de acordo com protocolo do fabricante. As amostras foram verificadas por PCR quanto à presença de contaminação com DNA genômico e estocadas a -20°C.

4.7.5 Construção de oligonucleotídeos

Os oligonucleotídeos usados nas análises do perfil transcricional por PCR em tempo real foram desenhados a partir das sequências dos genomas de *T. rubrum*, *T. tonsurans* e *M. canis* (Tabela 5). A confirmação de genes ortólogos deu-se por meio da ferramenta *Blast* nos genomas de interesse, disponibilizados pelo *Broad Institute*. Foi utilizado *software Primer3* (http://biotools.umassmed.edu/bioapps/primer3_www.cgi) para o desenho dos oligonucleotídeos.

Tabela 5: Oligonucleotídeos utilizados nas análises do perfil transcricional por PCR em tempo real.

Gene	Tipo	Organismo base	Sequência 5' - 3'
Dipeptidil peptidase V (<i>dppV</i>)	F	<i>M. canis</i>	AAC TAC GAC CGAT GGG ACT G
Dipeptidil peptidase V (<i>dppV</i>)	R	<i>M. canis</i>	ACG GAAG TCG TAG TCG CTGT
Dipeptidil peptidase V (<i>dppV</i>)	F	<i>T. tonsurans</i>	AAC TAT GAC CGC TGG GATT G
Dipeptidil peptidase V (<i>dppV</i>)	R	<i>T. tonsurans</i>	ACG TTG AAG AGAC CGAC ACC
Leucina aminopeptidase 1 (<i>lap1</i>)	F	<i>M. canis</i>	TGG CGT TGC TTAT GACA AGA
Leucina aminopeptidase 1 (<i>lap1</i>)	R	<i>M. canis</i>	CGC AGC GAA CTA GCATA CTG
Leucina aminopeptidase 1 (<i>lap1</i>)	F	<i>T. rubrum</i>	TGAG GAAT TCG GTCT CCTTG
Leucina aminopeptidase 1 (<i>lap1</i>)	R	<i>T. rubrum</i>	GAAC CGT CAC CGT CGT AGAT

Leucina aminopeptidase 2 (<i>lap2</i>)	F	<i>M. canis</i>	ATCATCACCGACAACGTCAA
Leucina aminopeptidase 2 (<i>lap2</i>)	R	<i>M. canis</i>	ATGGTCAGAGCAGGCATAGC
Leucina aminopeptidase 2 (<i>lap2</i>)	F	<i>T. rubrum</i>	TCGGAGCTAGCATTCCTGTT
Leucina aminopeptidase 2 (<i>lap2</i>)	R	<i>T. rubrum</i>	ATCCAACCTCCTCGTCAGTG
Metaloprotease 3 (<i>mep3</i>)	F	<i>M. canis</i>	GAGCTCATGTGGGGTCTGAT
Metaloprotease 3 (<i>mep3</i>)	R	<i>M. canis</i>	GACGATCTTCATGGCAAGGT
Metaloprotease 3 (<i>mep3</i>)	F	<i>T. rubrum</i>	GCCATCTTTCCATGAAGCTC
Metaloprotease 3 (<i>mep3</i>)	R	<i>T. rubrum</i>	GTGCACTTGTTGGCACTGTT
Metaloprotease 3 (<i>mep3</i>)	F	<i>T. tonsurans</i>	CTACTCCTTCGGCTTCAACG
Metaloprotease 3 (<i>mep3</i>)	R	<i>T. tonsurans</i>	TGGCATTGTTAGTGCCAGAG
Metaloprotease 4 (<i>mep4</i>)	F	<i>M. canis</i>	CGGCACTATCTGGTGTACGA
Metaloprotease 4 (<i>mep4</i>)	R	<i>M. canis</i>	CCGTCAAGGACGAGTTTCAT
Metaloprotease 4 (<i>mep4</i>)	F	<i>T. tonsurans</i>	AGAATGGTGTCCCGACTGAC
Metaloprotease 4 (<i>mep4</i>)	R	<i>T. tonsurans</i>	AAGTTGGGGAGACATGGTTG
Protease subtilisina 3 (<i>sub3</i>)	F	<i>M. canis</i>	TGCTGGTATCTTCGTTGCTG
Protease subtilisina 3 (<i>sub3</i>)	R	<i>M. canis</i>	GATGGTACCCAGTTGGAGA
Protease subtilisina 3 (<i>sub3</i>)	F	<i>T. rubrum</i>	CCAGGAAGCTGGTATCTTCG
Protease subtilisina 3 (<i>sub3</i>)	R	<i>T. rubrum</i>	TTCTGGTCATCAATGGTGGA
Protease subtilisina 5 (<i>sub5</i>)	F	<i>M. canis</i>	CAGCCTACCTAATCGCGAAG
Protease subtilisina 5 (<i>sub5</i>)	R	<i>M. canis</i>	AACGACCGCTGCCATTATAG
Protease subtilisina 5 (<i>sub5</i>)	F	<i>T. rubrum</i>	AAGGCTGGCATGTTTGTTC
Protease subtilisina 5 (<i>sub5</i>)	R	<i>T. rubrum</i>	AAGCCTTTCCGTCATTCTCA
Protease subtilisina 7 (<i>sub7</i>)	F	<i>M. canis</i>	GACAACGTCGATGCTGCTAA
Protease subtilisina 7 (<i>sub7</i>)	R	<i>M. canis</i>	GTTGGAGAAGTCTGCCTTGC
Protease subtilisina 7 (<i>sub7</i>)	F	<i>T. tonsurans</i>	CCCTCAACTCTGTGACACCA
Protease subtilisina 7 (<i>sub7</i>)	R	<i>T. tonsurans</i>	ATGCCAGATCCGTTGTTGAT

4.7.6 Análise dos genes diferencialmente expressos

A análise da expressão dos genes estudados neste trabalho foi feita por meio da técnica de RT-qPCR com metodologia de detecção por *SYBR*[®] *Green* (Applied Biosystems, Cambridge, UK, cat.#4367659), de acordo com protocolo do fabricante, utilizando o equipamento *StepOnePlus Real Time PCR* (Applied Biosystems). Três amostras de RNA provenientes de experimentos independentes foram usadas para cada um dos pontos analisados e as reações feitas em triplicata técnica. Cada reação, realizada com volume final de 12,5 µL, conteve 6,25 µL de *SYBR*[®] *Green* (PCR mix) e as quantidades de cDNA e primers necessárias.

As condições de ciclagem foram: 95°C por 10 min., seguido de 40 ciclos de 95°C por 15 s., e 60°C por 1 min. A curva de dissociação foi gerada ao final de cada corrida para

verificação da especificidade dos oligonucleotídeos utilizados. Todos os procedimentos e metodologias adotados para a realização de RT-qPCR tiveram como base estudos prévios (BUSTIN e NOLAN, 2004; DHEDA et al., 2005; HUGGETT et al., 2005; NOLAN et al., 2006; BUSTIN et al., 2009).

Focou-se a quantificação relativa dos genes de interesse, normalizando os dados com dois controles endógenos, *gapdh* e *actina* (JACOB et al., 2012). E as análises de expressão gênica foram realizadas pelo método de quantificação relativa $2^{-\Delta\Delta C_q}$ (LIVAK e SCHMITTGEN, 2001), além de os gráficos serem gerados com o auxílio do programa *GraphPad Prims v5* (www.graphpad.com/prism/Prism.htm). Como amostra referência, para os cultivos com queratina, utilizou-se *M. canis* no tempo 24 horas. Os resultados foram submetidos à análise de Variância (one-way ANOVA, com $p < 0,05$), seguida do teste de Tukey com nível de significância $p < 0,05$.

Para auxiliar e complementar a análise, gráficos do tipo *Heatmap* foram construídos com o uso do *software* estatístico *R* (<http://www.r-project.org/>), incluindo a utilização do pacote gráfico *gplots* (<http://cran.r-project.org/web/packages/gplots/gplots.pdf>).

Resultados

Discovery consists of seeing what everybody has seen and thinking what nobody has thought.

Albert Szent-Györgyi

V. Resultados

A obtenção das sequências genômicas completas dos organismos em estudo, de dados de expressão gênica e proteica, aliada ao desenvolvimento de tecnologias computacionais de alto desempenho e de algoritmos eficientes, permite o uso de abordagens holísticas no estudo da estrutura, organização e evolução de genomas.

Os resultados de cada uma das etapas de análise estão apresentados nesta seção e dispostos da seguinte forma:

5.1 Sequenciamento e montagem do genoma de *T. interdigitale* (H6);

5.2 Análise de genes envolvidos no processo de degradação da queratina.

Vale ressaltar que será utilizada nesta seção a identificação dada pelo *Broad Institute* para cada uma das espécies, como: *M. canis* (MCYG), *A. benhamiae* (ARB), *T. interdigitale* (H101), *M. gypseum* (MGYG), *T. equinum* (TEQG), *T. rubrum* (TERG), *T. tonsurans* (TESG) e *T. verrucosum* (TRV). Dentre os fungos do grupo externo, teremos: *C. immitis* (CIMG) e *C. posadasii* (CPAG).

5.1 Sequenciamento e montagem do genoma de *T. interdigitale* (H6)

O sequenciamento do genoma de *T. interdigitale* (H6) ocorreu em duas tentativas e os resultados obtidos estão apresentados na Tabela 6, a seguir.

Tabela 6: Sequenciamento do genoma de *T. interdigitale* (H6).

#	Quantidade de reads (<i>paired-end</i>)	Cobertura (*)
Sequenciamento 1	791.044	1,8
Sequenciamento 2	5.872.423	16,5

(*) Cobertura do alinhamento em relação ao genoma de referência utilizado (*T. rubrum* CBS 118892).

Na primeira tentativa de sequenciamento, ocorreu um número reduzido de *reads* e, consequentemente, uma cobertura bastante limitada no alinhamento com o genoma de referência. Credita-se o pequeno número de *reads* obtidos nesse sequenciamento ao fato de ter ocorrido como projeto piloto, logo após a aquisição do equipamento *MiSeq* pela Faculdade

de Medicina de Ribeirão Preto (USP). Na segunda tentativa, com um número bem mais expressivo de *reads*, a cobertura alcançou 16,5 vezes o tamanho do genoma.

A escolha do genoma de referência *T. rubrum* CBS 118892 deve-se ao fato da linguagem H6 ter sido identificada previamente como *T. rubrum*; sendo posteriormente reclassificada como *T. interdigitale* por consequência das análises de similaridade entre os diferentes genomas sequenciados pelo *Broad Institute*.

A montagem do genoma foi realizada da seguinte forma: após a etapa do processo de *base calling* (processo de nomeação das bases), fez-se a montagem (*assembly*) das sequências visando gerar *contigs* (conjunto de *reads* agrupados). O objetivo final em projetos genoma é unir os *reads* gerados, criando tantos *contigs* quanto for o número de cromossomos.

Como tentativa para a obtenção de melhores resultados, foram utilizados dois *softwares* no processo de montagem; o primeiro (*MiSeq Reporter*) obteve menor eficiência em relação ao segundo (*Velvet*) como apresentado na Tabela 7. Vale ressaltar que na utilização do *software MiSeq Reporter*, a montagem foi baseada em um genoma de referência. No caso do *software Velvet*, ele foi utilizado com estratégia de montagem “*de novo*” (do latim “desde o princípio”, e determina a sequência de DNA quando não há sequência genômica conhecida anteriormente e disponível para uso como referência).

Tabela 7: Montagem do genoma de *T. interdigitale* (H6).

Software	# sequenciamento	# <i>contigs</i>
<i>MiSeq Reporter</i>	Sequenciamento 1	10.690
<i>MiSeq Reporter</i>	Sequenciamento 2	4.746
<i>Velvet</i>	Sequenciamento 1 e 2 ^(a)	1.068 ^(b)

(a) Montagem utilizando as *reads* obtidas nos dois sequenciamentos; (b) Número de *contigs* com tamanho mínimo de 500 pares de base.

Pode-se constatar que o método de montagem com o *software Velvet* gerou 1.068 *contigs* (todos com tamanho acima de 500 pares de base). Apesar disso, o número de *gaps* ("buracos") entre os vários *contigs* obtidos é grande, o que poderia dificultar as próximas etapas de análises comparativas entre as sequências genômicas. Dessa forma, os 1.068 *contigs* gerados foram enviados ao *Broad Institute* para que pudessem se unir ao sequenciamento que o grupo havia realizado e, conseqüentemente, melhorar a cobertura no genoma. O *Broad*

Institute também utilizou técnicas de sequenciamento de nova geração, porém, com o equipamento *Hiseq 2000* (Illumina).

Atualmente, o genoma de *T. interdigitale* (H6) ainda se encontra em um estado *draft* (rascunho) e está disponível no site do *Broad Institute* no seguinte endereço: (http://www.broadinstitute.org/annotation/genome/dermatophyte_comparative/GenomesIndex.html).

5.2 Análise de genes envolvidos no processo de degradação da queratina

5.2.1 Serina proteases

5.2.1.1 Análises por Bioinformática

Com o intuito de verificar a similaridade entre os genes que codificam para subtilisina 3 (Sub3), subtilisina 5 (Sub5) e subtilisina 7 (Sub7) entre os organismos em estudo, cujas proteínas apresentam características de endoproteases, efetuaram-se alinhamentos entre os ortólogos, além da comparação de suas estruturas éxon/intron.

Para *sub3* (Figura 14), os resultados mostraram um grau de paridade quando observada a quantidade de regiões intrônicas entre os organismos. Foi possível verificar um alto grau de similaridade entre as sequências dos genes, com evidência de três regiões menos similares (menor taxa de conservação encontrada; faixas entre 300-400, 900-1.000 e 1.100-1.200) (Figura 14-A). A representação esquemática dos genes (Figura 14-B) mostra que essas quatro regiões são identificadas como íntrons para cada um dos genes alinhados em cada organismo do estudo.

Figura 14: (A) Alinhamento do conteúdo gênico entre os genes que codificam para Sub3. A linha vertical indica o deslocamento dos segmentos alinhados. (B) Representação esquemática dos genes nos oito dermatófitos em estudo. Os organismos estão dispostos na seguinte ordem: *M. canis*, *A. benhamiae*, *T. interdigitale*, *M. gypseum*, *T. equinum*, *T. rubrum*, *T. tonsurans* e *T. verrucosum*.

Visando verificar uma possível colinearidade entre os genes que codificam o grupo das serinas proteases, genes que codificam para Sub5 e Sub7 foram submetidos ao mesmo processo de análise.

Para *sub5* (Figura 15), verificou-se o mesmo grau de paridade quando observada a quantidade de regiões intrônicas entre os organismos. Nas análises de similaridade entre os genes, constatou-se que as regiões menos similares se encontram nas faixas entre 300-400, 900-1.000 e 1.100-1.200 (regiões de íntrons). Recentemente a anotação desse gene, em *T. rubrum*, foi modificada pelo *Broad Institute*, na qual foram propostas três isoformas.

O grupo de pesquisa do Laboratório de Genética e Biologia Molecular de Fungos do Departamento de Genética (FMRP/USP) vem contribuindo para revelar isoformas de várias

proteínas em dermatófitos através do uso da técnica de RNA-Seq (PERSINOTI et al., 2014) (MENDES, 2015) (informação pessoal)¹.

Figura 15: Representação esquemática dos genes que codificam para Sub5 nos oito dermatófitos em estudo. Os organismos estão dispostos na seguinte ordem: *M. canis*, *A. benhamiae*, *T. interdigitale*, *M. gypseum*, *T. equinum*, *T. rubrum*, *T. tonsurans* e *T. verrucosum*. (*) Isoformas descritas em *T. rubrum* para o gene representado.

Em *sub7* (Figura 16), constatou-se diversidade do número de regiões intrônicas entre algumas das espécies em estudo. Foi possível verificar três padrões de ocorrência de íntrons. O primeiro, com três íntrons, é formado por um grupo maior, incluindo as espécies *M. canis*, *T. interdigitale*, *M. gypseum*, *T. equinum* e *T. tonsurans*. O segundo, com cinco íntrons, é formado pelas espécies *A. benhamiae* e *T. verrucosum*. E, por último, *T. rubrum* não apresentando íntrons identificados em sua anotação. Outra característica importante é que, apesar das diversidades encontradas, essas não estão associadas ao gênero de cada espécie, tendo em vista que organismos do gênero *Microsporum* e *Trichophyton* apresentaram similaridade em alguns casos.

¹ Informação fornecida por Nieve Silva Mendes em Ribeirão Preto, em 2015.

Figura 16: Representação esquemática dos genes que codificam para Sub7 nos oito dermatófitos em estudo. Os organismos estão dispostos na seguinte ordem: *M. canis*, *A. benhamiae*, *T. interdigitale*, *M. gypseum*, *T. equinum*, *T. rubrum*, *T. tonsurans* e *T. verrucosum*.

Para verificar possíveis polimorfismos de um único nucleotídeo, foram feitas análises de bioinformática com um algoritmo construído no decorrer deste trabalho, sendo possível por meio dele identificar e disponibilizar o número de variações e a relação destas com o tamanho dos segmentos dos genomas analisados. Inicialmente, usou-se como referência *M. canis*, a qual foi comparada às outras sete espécies do estudo. Tais análises foram realizadas para *sub3*, *sub5* e *sub7* e estão dispostas nas Tabelas 8, 9 e 10, respectivamente.

Tabela 8: Número de variações entre *M. canis* e outras espécies de dermatófitos para o gene que codifica subtilisina 3 (Sub3).

	Referência MCYG_04040						
	ARB_00701	H101_01018	MGYG_02570	TEQG_06704	TERG_03815	TESG_03300	TRV_07976
Éxon	127 (10,64%)	150 (12,56%)	131 (10,97%)	153 (12,81%)	138 (11,56%)	155 (12,98%)	133 (11,14%)
Íntron	76 (36,54%)	66 (31,73%)	69 (33,17%)	70 (33,65%)	68 (32,69%)	70 (33,65%)	80 (38,46%)
1.000upstream	206 (20,60%)	175 (17,50%)	211 (21,10%)	211 (21,10%)	188 (18,80%)	201 (20,10%)	204 (20,40%)

Os valores apresentados entre parênteses como porcentagem representam o número de variações em relação ao tamanho do segmento analisado.

Tabela 9: Número de variações entre *M. canis* e outras espécies de dermatófitos para o gene que codifica subtilisina 5 (Sub5).

Ref. MCYG_08753							
	ARB_02223	H101_05652	MGYG_07290	TEQG_05834	TERG_08201	TESG_07246	TRV_00550
Éxon	236 (19,87%)	228 (19,19%)	246 (20,71%)	224 (18,86%)	237 (19,95%)	227 (19,11%)	236 (19,87%)
Íntron	65 (36,93%)	74 (42,05%)	64 (36,36%)	76 (43,18%)	69 (39,20%)	76 (43,18%)	76 (43,18%)
1.000upstream	179 (17,90%)	71 (7,10%)	197 (19,70%)	141 (14,10%)	109 (10,90%)	70 (7,00%)	188 (18,80%)

Os valores apresentados entre parênteses como porcentagem representam o número de variações em relação ao tamanho do segmento analisado.

Tabela 10: Número de variações entre *M. canis* e outras espécies de dermatófitos para o gene que codifica subtilisina 7 (Sub7).

MCYG_02345							
	ARB_06076	H101_00208	MGYG_07342	TEQG_06283	TERG_12327	TESG_06365	TRV_00296
Éxon	163 (13,55%)	179 (14,88%)	185 (15,38%)	179 (14,88%)	124 (10,31%)	180 (14,96%)	160 (13,30%)
Íntron	73 (36,68%)	75 (37,69%)	55 (27,64%)	76 (38,19%)	76 (38,19%)	76 (38,19%)	71 (35,68%)
1.000upstream	128 (12,80%)	210 (21,00%)	211 (21,10%)	207 (20,70%)	ERRO*	208 (20,80%)	100 (10,00%)

Os valores apresentados entre parênteses como porcentagem representam o número de variações em relação ao tamanho do segmento analisado. (*) Erro no alinhamento entre as sequências – Não ocorrência de similaridade.

O número de variações, aparentemente alto, encontrado na análise, está relacionado à escolha do genoma de referência (*M. canis* - MCYG). Por outro lado, quando utilizado *T. equinum* (TEQG) como referência, os resultados obtidos, quando comparados a *T. tonsurans* ou a outras espécies do gênero *Trichophyton*, apresentam redução importante no número de variações (Tabela 11 a Tabela 13).

Além de analisar as variações em regiões de íntrons e de éxons, preocupou-se com a verificação do número de variações encontradas a 1.000 pares de base *upstream* de cada um dos genes. Sabe-se que pequenas modificações nessas regiões (regiões promotoras) podem influenciar a regulação dos genes e, por conseguinte, a especificidade de cada organismo. Essas análises foram realizadas para *sub3*, *sub5* e *sub7*, e estão dispostas nas Tabelas 11, 12 e 13, respectivamente.

Tabela 11: Número de variações entre *T. equinum* e *T. tonsurans* e outras espécies do gênero *Trichophyton* para o gene que codifica subtilisina 3 (Sub3).

Referência TEQG_06704		
	TESG_03300	<i>Trichophyton spp.</i>
Éxon	2 (0,17%)	76 (6,37%)
Íntron	0 (0,00%)	45 (20,09%)
1.000upstream	5 (0,50%)	83 (8,30%)

Os valores apresentados entre parênteses como porcentagem representam o número de variações em relação ao tamanho do segmento analisado.

Tabela 12: Número de variações entre *T. equinum* e *T. tonsurans* e outras espécies do gênero *Trichophyton* para o gene que codifica subtilisina 5 (Sub5).

Referência TEQG_05834		
	TESG_07246	<i>Trichophyton spp.</i>
Éxon	3 (0,25%)	79 (6,63%)
Íntron	0 (0,00%)	34 (17,26%)
1.000uptream	5 (0,50%)	120 (12,00%)

Os valores apresentados entre parênteses como porcentagem representam o número de variações em relação ao tamanho do segmento analisado.

Tabela 13: Número de variações entre *T. equinum* e *T. tonsurans* e outras espécies do gênero *Trichophyton* para o gene que codifica subtilisina 7 (Sub7).

Referência TEQG_06283		
	TESG_06365	<i>Trichophyton spp.</i>
Éxon	1 (0,08%)	93 (7,71%)
Íntron	0 (0,00%)	31 (16,58%)
1.000uptream	2 (0,20%)	40 (4,00%)

Os valores apresentados entre parênteses como porcentagem representam o número de variações em relação ao tamanho do segmento analisado.

Percebe-se uma redução do número de variações quando comparados *T. equinum* com outras espécies do gênero *Trichophyton* (exemplo: 6,37% de variações em regiões de éxons para *sub3*) e um número muito baixo de variações quando comparados apenas *T. equinum* com *T. tonsurans* – TESS (exemplo: 2 variações em regiões de éxons; 0 variação em regiões de íntrons; 5 variações encontradas a 1.000 pares de base *upstream* de *sub3*).

Outro aspecto importante a se observar, é que o valor absoluto das variações é maior em regiões de éxons (exemplo: 93 variações na região de éxon e 31 na região de íntron, quando comparados *T. equinum* e outros organismos do gênero *Trichophyton* (Tabela 13)), contudo, quando se leva em conta o tamanho de cada região (éxon/íntron), o percentual de variações é mais representativo em regiões de íntrons (exemplo: 7,71% de variações na região de éxon e 16,58% na região de íntron (Tabela 13)).

As proteínas preditas, codificadas por cada um dos genes, foram submetidas a uma análise de identidade, observando-se o grau de divergência entre as espécies. Matrizes de distância foram construídas a partir do alinhamento de sequências (Tabela 14 a Tabela 16).

Tabela 14: Matriz de distância para a proteína subtilisina 3 (Sub3).

	MCYG_04040	ARB_00701	H101_01018	MGYG_02570	TEQG_06704	TERG_03815	TESG_03300	TRV_07976
MCYG_04040	ID							
ARB_00701	20	ID						
H101_01018	31	17	ID					
MGYG_02570	26	26	25	ID				
TEQG_06704	33	19	2	27	ID			
TERG_03815	22	5	20	26	22	ID		
TESG_03300	34	20	3	28	1	23	ID	
TRV_07976	20	0	17	26	19	5	20	ID

Nas linhas e colunas, estão representadas as proteínas para cada uma das espécies analisadas. O cruzamento de uma linha com uma coluna representa a distância estimada entre as sequências de aminoácidos correspondentes. ID=Identidade.

Tabela 15: Matriz de distância para a proteína subtilisina 5 (Sub5).

	MCYG_08753	ARB_02223	H101_05652	MGYG_07290	TEQG_05834	TERG_08201	TESG_07246	TRV_00550
MCYG_08753	ID							
ARB_02223	42	ID						
H101_05652	40	6	ID					
MGYG_07290	44	35	36	ID				
TEQG_05834	40	6	0	36	ID			
TERG_08201	40	2	4	33	4	ID		
TESG_07246	41	7	1	37	1	5	ID	
TRV_00550	40	2	6	35	6	2	7	ID

Nas linhas e colunas, estão representadas as proteínas para a cada uma das espécies analisadas. O cruzamento de uma linha com uma coluna representa a distância estimada entre as sequências de aminoácidos correspondentes. ID=Identidade.

Tabela 16: Matriz de distância para a proteína subtilisina 7 (Sub7).

	MCYG_02345	ARB_06076	H101_00208	MGYG_07342	TEQG_06283	TERG_12327	TESG_06365	TRV_00296
MCYG_02345	ID							
ARB_06076	18	ID						
H101_00208	22	11	ID					
MGYG_07342	15	14	17	ID				
TEQG_06283	22	11	0	17	ID			
TERG_12327	123	122	124	121	124	ID		
TESG_06365	22	11	0	17	0	124	ID	
TRV_00296	17	2	10	12	10	122	10	ID

Nas linhas e colunas, estão representadas as proteínas para cada uma das espécies analisadas. O cruzamento de uma linha com uma coluna representa a distância estimada entre as sequências de aminoácidos correspondentes. ID=Identidade.

Nessa análise, foi possível observar que o alto grau de conservação das sequências de aminoácidos está relacionado à conservação das estruturas dos genes entre cada uma das espécies.

Para as três subtilisinas analisadas, as menores distâncias observadas ocorreram entre as espécies *T. interdigitale* (H101), *T. equinum* e *T. tonsurans*. Outras duas espécies muito próximas observadas foram *A. benhamiae* (ARB) e *T. verrucosum* (TRV). Por outro lado, com exceção da Sub7, que apresenta comportamento diferenciado das outras duas subtilisinas, *M. canis* e *M. gypseum* (MGYG) possuem maiores distâncias quando comparadas às outras espécies.

A existência de domínios (módulos que constituem unidades distintas do ponto de vista evolutivo, funcional e estrutural) em proteínas é um fator complicador em tais análises, que deve ser tratado com atenção posteriormente.

Visando representar as relações filogenéticas entre as subfamílias das proteínas de interesse, foram construídos cladogramas para cada um dos grupos de sequências de aminoácidos.

As representações foram realizadas para Sub3, Sub5 e Sub7 e estão dispostas nas Figuras 17, 18 e 19, respectivamente.

Figura 17: Relação filogenética entre os dermatófitos com base nas sequências da proteína subtilisina 3 (Sub3).

Figura 18: Relação filogenética entre os dermatófitos com base nas sequências da proteína subtilisina 5 (Sub5).

Figura 19: Relação filogenética entre os dermatófitos com base nas sequências da proteína subtilisina 7 (Sub7).

Pode-se observar nos cladogramas que as subfamílias analisadas formam agrupamentos distintos, mas não característicos para os grupos antropofílicos, zoofílicos ou geofílicos. É visto que os grupos formados estão mais associados ao gênero de cada espécie, comprovando a existência de sinal filogenético e derivação de um mesmo ancestral. Ocorrências de duplicações mais recentes poderiam explicar o surgimento de alguns subgrupos formados, todavia, dados de outras espécies relacionadas às espécies aqui estudadas teriam que ser incluídos nesta análise para confirmação dessa hipótese.

Para Sub3 e Sub5 observaram-se quatro subgrupos principais: um subgrupo entre as espécies *T. verrucosum*, *A. benhamiae* e *T. rubrum*, um segundo subgrupo entre *T. equinum*, *T. tonsurans* e *T. interdigitale*, um terceiro subgrupo entre *M. canis* e *M. gypseum* e, por último, um subgrupo que representa o grupo externo formado pelas espécies *C. immitis* e *C. posadasii*. A existência desses subgrupos também foi observada quando da análise de Sub7, sendo que a única diferença encontrada foi o deslocamento de *T. rubrum* para um ramo de maior distância. Essa diferença se deve a não paridade entre as estruturas dos genes na espécie (não ocorrência de regiões de íntrons em *T. rubrum*) e, conseqüentemente, a formação de uma proteína com maior índice de distância das demais (Tabela 16).

Entre as espécies deste estudo, procurou-se também verificar aspectos relacionados à vizinhança entre os genes analisados e posicionamento cromossômico. Observou-se uma maior divergência relacionada ao posicionamento de cada um dos genes analisados entre cada

uma das espécies em estudo. Mesmo se tratando de genomas relativamente parecidos em seus tamanhos, aspectos sobre mudanças de vizinhança entre os genes foram detectados.

Nas Figuras 20 e 21, partindo da localização dos genes que codificam para Sub3 e Sub5, respectivamente (representados por uma barra vertical vazada), observa-se que *T. rubrum* e *A. benhamiae* possuem a mesma vizinhança (representadas pela diferenciação entre cores). As espécies *M. canis*, *M. gypseum*, *T. equinum*, *T. tonsurans* e *T. verrucosum* possuem alguns vizinhos em comum, mas alguns de seus *supercontigs* estão em sentidos opostos.

Figura 20: Sintenia das regiões próximas ao gene que codifica subtilisina 3 (Sub3). A localização do gene está apontada com uma barra vertical vazada. Ordem dos alinhamentos: *T. rubrum*, *A. benhamiae*, *M. canis*, *M. gypseum*, *T. equinum*, *T. tonsurans* e *T. verrucosum*.

Para as espécies *T. tonsurans* e *T. equinum*, quando comparadas às regiões adjacentes aos genes que codificam Sub3 e Sub5 (Figura 20 e Figura 21), verificou-se alguns vizinhos preservados, porém observa-se também que algumas destas regiões apresentaram tamanhos diferentes entre estas espécies.

Figura 21: Sintenia das regiões próximas ao gene que codifica subtilisina 5 (Sub5). A localização do gene está apontada com uma barra vertical vazada. Ordem dos alinhamentos: *T. rubrum*, *A. benhamiae*, *M. canis*, *M. gypseum*, *T. equinum*, *T. tonsurans* e *T. verrucosum*.

Fatores de transcrição se ligam às sequências gênicas reguladoras, geralmente na região promotora do gene alvo, para aumentar ou diminuir a taxa de transcrição do gene. Buscas por motivos de ligação em regiões promotoras foram realizadas de duas formas: na primeira, usou-se um banco de dados de sítios de ligação para fator de transcrição de *Saccharomyces cerevisiae*; na segunda, utilizou-se uma metodologia por nós desenvolvida para este projeto (vide item 4.3.9 - Análise de sítios de ligação nas regiões promotoras), que visa identificar pequenas sequências consenso por meio do alinhamento de regiões 1.000 pares de base *upstream* dos genes de interesse.

A Figura 22 apresenta as quantificações de sítios de ligação para fatores de transcrição em cada uma das espécies do estudo. Sequências de 1.000 pares de base *upstream* de *sub3* para as oito espécies foram submetidas à busca por meio da ferramenta *Yeasttract*, disponível em <http://www.yeasttract.com/formtfsbindingsites.php>.

Na região promotora do gene *sub3* (Figura 22), observou-se grande quantidade de possíveis sítios de ligação para os fatores transcricionais, entre eles se destacam: Ash1p, Stb5p, Fkh1p, Fkh2p e Gcr1p. Para *sub5* (Figura 23), as maiores quantidades de sítios observadas foram: Fkh1p, Fkh2p, Gcr1p, Mot3p e Stb5p. Por último, para *sub7* (Figura 24), destacam-se os fatores transcricionais: Ash1p, Gcr1p, Mot3p, Stb5p e Tec1p. De acordo com

esse resultado, é possível sugerir que *sub3*, *sub5* e *sub7* tem potencial para serem regulados por cada um desses fatores transcricionais.

Figura 22: Número de sítios de ligação para fatores de transcrição para *sub3*. Análise com base no banco de dados YEASTRACT (Yeast Search for Transcriptional Regulators And Consensus Tracking)

A possível redundância de bases em algumas posições da sequência consenso, o pequeno tamanho destas sequências e a metodologia utilizada no mecanismo de busca (identificação do sítio de ligação descrito em *S. cerevisiae* no sentido *forward* e *reverse* da região 1.000 pares de base *upstream* de cada gene) podem ter contribuído para que o número de sítios de ligação encontrados atingisse altos valores.

Outra característica importante observada foi a similaridade do número de sítios de ligação encontrados, entre os organismos analisados, para os genes homologos. Isso também se deve ao fato das regiões 1.000 pares de base *upstream* desses genes apresentarem um alto grau de conservação.

Nas Figuras 23 e 24, são apresentados os resultados das quantificações de sítios de ligação para fatores de transcrição de *sub5* e *sub7*, respectivamente.

Figura 23: Número de sítios de ligação para fatores de transcrição para *sub5*. Análise com base no banco de dados YEASTRACT (Yeast Search for Transcriptional Regulators And Consensus Tracking)

Figura 24: Número de sítios de ligação para fatores de transcrição para *sub7*. Análise com base no banco de dados YEASTRACT (Yeast Search for Transcriptional Regulators And Consensus Tracking)

Uma segunda proposta de análise apresentou sequências consenso que podem estar relacionadas à classe, ao gênero ou até mesmo ao nicho ecológico. Nessa análise, sequências de 1.000 pares de base upstream de cada um dos genes de interesse em cada uma das espécies foram alinhadas. Após o alinhamento, pequenas regiões de consenso com tamanhos entre 6 e 15 nucleotídeos foram selecionadas como possíveis alvos. A Tabela 17 apresenta parte da análise realizada para subtilisina 3 (*sub3*).

Tabela 17: Amostra de sequências consenso obtidas por meio do alinhamento da região 1.000 pares de base *upstream* do gene *sub3* dos oito dermatófitos em estudo, incluindo dois organismos como grupo externo.

1.000upstream	Motivos identificados				
	M1	M2	M3	M4	M5
MCYG_04040	3	2	1	-	1
ARB_00701	5	7	1	1	1
H101_01018	5	6	1	1	-
MGYG_02570	2	4	1	-	2
TEQG_06704	5	7	1	1	-
TERG_03815	4	7	1	1	-
TESG_03300	5	7	1	1	-
TRV_07976	5	7	1	1	-
CPAG_00038	4	-	-	1	2
CIMG_02881	1	-	-	1	2

M1=T[GC][TA]CC[CA][GT][CG]TC[AT]C[CT]CT; M2=[ATC][TAC][CT][TC]CGCC[GTA][GAT][TCA]G[AG]TA;

M3=GTCTAGTC[GA]TGAAGA; M4=GCTCG[AG]TGT[AG]T[GC][TC][CGA]T;

M5=G[GA][TA][CG]AG[AG]GC[TAG][TA]TTTG.

(*) O colchete é usado para denominar um conjunto e representa a ocorrência de uma de suas bases na posição.

Na Tabela 17, há duas sequências consenso (M2 e M3) encontradas em todos os dermatófitos selecionados, porém, ausentes nos organismos pertencentes ao grupo externo. Esse fato pode sugerir que elas estejam relacionadas às espécies de dermatófitos. Pode-se observar, também, a ausência de M4 nos organismos do gênero *Microsporum* e M5 no gênero *Trichophyton*, ou seja, esses consensos estariam relacionados aos gêneros específicos.

Estas análises também foram realizadas para subtilisina 5 (*sub5*) e subtilisina 7 (*sub7*) e estão representadas na Tabela 18 e Tabela 19, respectivamente.

Tabela 18: Amostra de sequências consenso obtidas por meio do alinhamento da região 1.000 pares de base *upstream* do gene *sub5* dos oito dermatófitos em estudo, incluindo dois organismos como grupo externo.

Motivos identificados					
1.000upstream	M1	M2	M3	M4	M5
MCYG_08753	1	5	2	-	1
ARB_02223	5	1	7	4	1
H101_05652	5	2	8	4	-
MGYG_07290	1	1	4	-	2
TEQG_05834	5	2	8	4	-
TERG_08201	5	1	6	4	-
TESG_07246	5	2	8	4	-
TRV_00550	5	1	6	4	1
CPAG_08024	1	-	-	1	2
CIMG_09106	1	-	-	1	2

M1=G[CA][AT]G[TA][TC]CCT[GC][TA][CG]A[TAC][GT]; M2=T[TA]TT[GT]CT[CT][TC]C[CG];

M3=[CTG]C[TC]T[TAC]T[ATC]CA[CT][TG]C[TCGA]TT; M4=[CG][AT][AC]GA[AT][GC][GAT]A[TGC][TG]T[TA][GT]T;

M5=G[AC][AG]G[CA][AG][ACT][GT]G[GT]A[TA]TGC.

(*) O colchete é usado para denominar um conjunto e representa a ocorrência de uma de suas bases na posição.

Tabela 19: Amostra de sequências consenso obtidas por meio do alinhamento da região 1.000 pares de base *upstream* do gene *sub7* dos oito dermatófitos em estudo, incluindo dois organismos como grupo externo.

Motivos identificados					
1.000upstream	M1	M2	M3	M4	M5
MCYG_02345	3	2	-	-	2
ARB_06076	6	5	2	1	-
H101_00208	5	6	1	1	2
MGYG_07342	6	6	1	-	1
TEQG_06283	5	8	1	1	2
TERG_12327	3	-	-	-	-
TESG_06365	5	8	1	1	2
TRV_00296	5	5	2	1	-
CPAG_01454	3	-	-	1	1
CIMG_01394	3	-	-	1	1

M1=T[CT][CG][TG][CT]TC[ATC]T[CT][TCG]CT[TC][CG]; M2=[GC]GG[CA][CT][TC][CT]CG[AC][AT]A[TC]C[AC];

M3=C[GC]TG[TG][TC][TG]T[CT]ATGGTT; M4=[GT]G[TC]C[CG]GC[CG]AA[TC]TCTC;

M5=G[TC][GA]AA[CT]TGGT[AG][CT]C[CT][AT].

(*) O colchete é usado para denominar um conjunto e representa a ocorrência de uma de suas bases na posição.

Alguns consensos foram identificados em todas as sequências analisadas (incluindo o grupo externo), o que aumenta a possibilidade de estarem relacionados à função básica do gene (exemplo: M1 para as Tabelas 17, 18 e 19). Como também é possível constatar, o número de sequências consenso pode variar de acordo com o organismo.

5.2.1.2 Expressão gênica

Estudos realizados no Laboratório de Genética e Biologia Molecular de Fungos detectaram a modulação da expressão de genes da linhagem H6 (identificada pelo *Broad Institute* como H101) em diversas condições de cultivo (PERES et al., 2010b). Um banco de dados relacional² com interface *web*, que oferece funções analíticas e acesso integrado aos dados de expressão gênica, obtidos pelo laboratório, foi criado para o armazenamento de sequências. Na Figura 25, é apresentada uma das interfaces do *software* desenvolvido para visualização das sequências obtidas por cada condição, assim como um detalhamento sobre a qualidade do sequenciamento e comparação com outros bancos de dados públicos em biologia molecular.

O *software* possui vários módulos de busca, agilizando as pesquisas por dados obtidos em experimentos do laboratório. Foi desenvolvido utilizando a linguagem de programação Perl e o Sistema Gerenciador de Banco de Dados Mysql (<https://www.mysql.com>).

A criação e a manutenção de bancos de dados biológicos são por si só um desafio, devido não só à imensa quantidade de dados, mas, sobretudo, à dificuldade de desenvolver esquemas e estruturas que representem de forma exata ou bastante aproximada a complexa relação existente entre os diversos componentes dos sistemas biológicos (MACÊDO, 2003).

Os genes de interesse, neste trabalho, foram submetidos a uma análise de similaridade com os dados armazenados nesse banco de dados, a fim de prever possíveis funções compatíveis entre as espécies. Na Tabela 20, é apresentada uma lista contendo as bibliotecas de expressão em *T. interdigitale* (H6), na qual cada um dos genes de interesse foi encontrado.

Grande parte dessas bibliotecas de expressão foi obtida utilizando a técnica de Hibridização Subtrativa Supressiva (SSH), que permite uma análise de expressão diferencial quando comparadas duas condições de cultivo (controle/teste).

² Conceito abstrato que define maneiras de organizar, armazenar, manipular e recuperar dados estruturados na forma de tabelas ou, mais formalmente, relações.

Gene	Virulência	Biblioteca *
TERG_03815	Protease subtilisina (<i>sub3</i>)	7
TERG_08201	Protease subtilisina (<i>sub5</i>)	6, 7, 10
TERG_03248	Metaloprotease (<i>mep3</i>)	7
TERG_04324	Metaloprotease (<i>mep4</i>)	7
TERG_07861	Dipeptidil peptidase V (<i>dppV</i>)	1, 9

(*) 1. Transcritos enriquecidos em fase de desenvolvimento; 6. Transcritos enriquecidos em cultivos com glicose; 7. Transcritos enriquecidos em cultivos com queratina; 9. Transcritos enriquecidos em pH 5,0 (exposição de 60 min.); 10. Transcritos enriquecidos em pH 8,0 (exposição de 60 min.)

Essas análises sugerem que os genes *sub3*, *sub5*, *mep3*, *mep4* e dipeptidil peptidase V (*dppV*) estariam envolvidos na resposta à presença de queratina, o que nos levou a verificar a expressão de cada um dos genes de interesse em resposta à queratina.

Definiram-se dois tempos de cultivo (24 horas e 96 horas) em substrato de queratina para a extração do RNA, almejando identificar o momento de expressão, os níveis transcricionais do gene e o envolvimento dessas proteínas na regulação do processo de degradação da queratina. Os gráficos de expressão relativa foram construídos de acordo com os valores de quantificação relativa (*RQ - Relative Quantification*) e calculados utilizando como calibrador o menor dos valores de ΔC_t obtido para uma das linhagens (*M. canis* – 24 horas). As reações de PCR foram padronizadas pela construção de uma curva padrão para o cálculo de eficiência dos oligonucleotídeos e pela análise da curva de dissociação, observando a ocorrência de amplificação de apenas um fragmento. Na Figura 26, é apresentada a análise quantitativa da expressão gênica de subtilisina 3 (*sub3*) pela técnica de PCR em tempo real.

Figura 26: Gráfico dos valores de expressão relativa (*fold change*) para o gene *sub3* em cinco dos dermatófitos em estudo. Para cada organismo, o gráfico apresenta os valores em 24 horas e 96 horas.

Figura 27: Heatmap construído com os valores de expressão para o gene *sub3*.

Como observado na Figura 26, o gene *sub3* apresenta padrões muito diferentes de níveis de transcrição em relação aos tempos de cultivo. *T. rubrum* e *T. interdigitale* apresentam níveis de transcritos aumentados significativamente em tempos iniciais de exposição à queratina (24 horas) e, aparentemente, com o passar do tempo (96 horas), a atuação desse gene é reduzida em resposta ao meio de cultivo. O inverso ocorre em *T. equinum*. Quando comparados apenas os tempos iniciais de exposição (24 horas), *T. rubrum*, *T. equinum* e *T. interdigitale* apresentam aumento significativo no nível de transcritos de *sub3* em relação a *M. canis*, sugerindo efetiva atividade desse gene no processo de degradação da queratina nos organismos do gênero *Trichophyton* (com exceção de *T. tonsurans*).

Para verificar se os padrões de expressão gênica poderiam estar correlacionados às relações filogenéticas descritas para *Sub3*, uma matriz de expressão gênica com análise classificatória hierárquica (gráfico de *Heatmap*) foi construída a partir dos valores de expressão obtidos (Figura 27). Percebe-se que *M. canis* possui maior distância em relação aos organismos do gênero *Trichophyton*. E um aspecto importante observado é que as espécies antropofílicas do gênero *Trichophyton* estão mais próximas, se comparadas à única espécie zoofílica do gênero (*T. equinum*).

Na Figura 28, é apresentada a análise quantitativa da expressão gênica de subtilisina 5 (*sub5*) pela técnica de PCR em tempo real.

Figura 28: Gráfico dos valores de expressão relativa (*fold change*) para o gene *sub5* em cinco dos dermatófitos em estudo. Para cada organismo, o gráfico apresenta os valores em 24 horas e 96 horas.

O gene *sub5* apresenta níveis de transcritos significativamente aumentados nas linhagens *T. rubrum*, *T. equinum* e *T. interdigitale* após o cultivo em queratina por 24 horas, sendo que estes níveis reduzem significativamente nos tempos de 96 horas. Novamente, constata-se uma exceção, em *T. tonsurans*, que apresenta um perfil transcricional diferente das demais (Figura 28 e Figura 29).

Quando observados os padrões de níveis de transcrição identificados para *sub5* (Figura 29), constata-se *T. interdigitale* com maior distância em relação às demais espécies e isso se deve ao fato de sua quantidade de transcritos apresentar-se, aproximadamente, duas vezes maior, quando comparados os tempos de cultivo em 24 horas, e duas vezes menor, quando comparados no tempo mais tardio, em relação às outras espécies.

Figura 29: Heatmap construído com os valores de expressão para o gene *sub5*.

Na Figura 30 é exibida a análise quantitativa da expressão gênica de subtilisina 7 (*sub7*) pela técnica de PCR em tempo real.

Figura 30: Gráfico dos valores de expressão relativa (*fold change*) para o gene *sub7* em cinco dos dermatófitos em estudo. Para cada organismo, o gráfico apresenta os valores em 24 horas e 96 horas.

O gene *sub7* responde de forma expressiva em cultivo com queratina em 24 horas para as linhagens *T. rubrum* e *T. equinum*, sendo que essa resposta decai significativamente em 96 horas de cultivo, mas ainda mantem uma alta expressão (Figura 30 e Figura 31).

Figura 31: Heatmap construído com os valores de expressão para o gene *sub7*.

Apesar da estrutura gênica de *sub7* de *T. equinum* e *T. tonsurans* apresentar apenas uma SNP em seus éxons, o perfil de expressão deste gene nestas espécies se mostrou significativamente diferente. Isto sugere que a estrutura gênica de *sub7* não tenha influencia no perfil de expressão do próprio gene, como ocorre para vários outros genes.

5.2.2 Metaloproteases

5.2.2.1 Análises por Bioinformática

Visando verificar a similaridade entre os genes que codificam para metaloprotease 3 (Mep3) e metaloprotease 4 (Mep4) entre os organismos em estudo, cujas proteínas possuem papel importante durante a interação patógeno-hospedeiro, foram realizados alinhamentos entre os ortólogos, além da comparação de suas estruturas éxon/íntron.

Para *mep3* (Figura 32), verificou-se que algumas espécies apresentam o mesmo grau de paridade quando observada a quantidade de regiões intrônicas. Por outro lado, *T. interdigitale* e *T. equinum* evidenciam números diferenciados em relação à quantidade de íntrons. Os resultados mostraram que *T. interdigitale* possui três íntrons (um a menos que a maioria das espécies do estudo). Sugere-se que o tamanho do gene nessa espécie possa estar relacionado a esse fator, tendo em vista que *mep3* possui em torno de 2.300 pares de base em todas as espécies, com exceção de *T. interdigitale*, que possui em torno de 1.700 pares de base. Os resultados para *T. equinum* mostraram a ocorrência de cinco regiões intrônicas, porém, a sequência disponibilizada para um de seus íntrons possui grande região de baixa qualidade, o que sugere uma possível falha no sequenciamento.

Figura 32: Representação esquemática dos genes *mep3* nos oito dermatófitos em estudo. Os organismos estão dispostos na seguinte ordem: *M. canis*, *A. benhamiae*, *T. interdigitale*, *M. gypseum*, *T. equinum*, *T. rubrum*, *T. tonsurans* e *T. verrucosum*. (*) Isoformas descritas em *T. rubrum* para o gene representado.

Nas análises de similaridade entre os genes, constatou-se que as regiões menos similares se encontram nas faixas próximas a 400, 900, 1.500 e 2.000 pares de base (regiões de íntrons na maioria das espécies). Para *mep3*, também se notou a presença de duas isoformas anotadas em *T. rubrum*.

Figura 33: Representação esquemática dos genes *mep4* nos oito dermatófitos em estudo. Os organismos estão dispostos na seguinte ordem: *M. canis*, *A. benhamiae*, *T. interdigitale*, *M. gypseum*, *T. equinum*, *T. rubrum*, *T. tonsurans* e *T. verrucosum*. (*) Isoformas descritas em *T. rubrum* para o gene representado.

Os resultados para *mep4* (Figura 33) mostraram o mesmo padrão de quantidade e posicionamento dos íntrons entre os organismos, exceto para *T. rubrum*, na qual foi verificada a presença de três isoformas descritas. Nas análises de similaridade entre os genes, verificou-se que as regiões menos similares se encontram nas faixas entre 400-500, 800-900, 1.400-1.500 e 1.900-2.000 (regiões de íntrons).

Com o objetivo de avaliar modificações pontuais nas regiões de éxon, íntron e, até mesmo, a 1.000 pares de base upstream dos genes que codificam para Mep3 e Mep4, utilizou-se como referência uma das espécies do estudo, comparando às outras espécies.

Inicialmente, utilizou-se como referência *M. canis* para as comparações dos genes que codificam para Mep3 (Tabela 21) e Mep4 (Tabela 22). E, mais uma vez, foi possível observar que o número de variações encontradas é relativamente alto e o motivo é a escolha do genoma de referência (*M. canis*). Por outro lado, quando utilizado *T. equinum* como referência, os resultados obtidos, quando comparados à *T. tonsurans* ou às outras espécies do gênero *Trichophyton*, apresentam redução importante no número de variações (Tabela 23 e Tabela 24).

Tabela 21: Número de variações entre *M. canis* e outras espécies de dermatófitos para o gene que codifica metaloprotease 3 (Mep3).

	Referência MCYG_04620						
	ARB_05085	H101_04475	MGYG_08200	TEQG_02420	TERG_03248	TESG_02775	TRV_06691
Éxon	193 (10,15%)	159 (8,36%)	218 (11,46%)	186 (9,78%)	209 (10,99%)	205 (10,78%)	194 (10,20%)
Íntron	73 (25,70%)	77 (27,11%)	85 (29,93%)	65 (22,89%)	88 (30,99%)	91 (32,04%)	81 (28,52%)
1.000upstream	153 (15,30%)	ERRO*	113 (11,30%)	95 (9,50%)	133 (13,30%)	129 (12,90%)	153 (15,30%)

Os valores apresentados entre parênteses como porcentagem representam o número de variações em relação ao tamanho do segmento analisado. (*) Erro no alinhamento entre as sequências – *T. interdigitale* possui descrita uma sequência de apenas 58 pares de base na região *upstream* do gene.

Tabela 22: Número de variações entre *M. canis* e outras espécies de dermatófitos para o gene que codifica metaloprotease 4 (Mep4).

Ref. MCG_02088							
	ARB_00762	H101_01910	MGYG_03984	TEQG_04323	TERG_04324	TESG_00401	TRV_00081
Éxon	476 (24,41%)	480 (24,62%)	493 (25,28%)	483 (24,77%)	487 (24,97%)	483 (24,77%)	475 (24,36%)
Íntron	82 (35,04%)	59 (25,21%)	62 (26,50%)	58 (24,79%)	67 (28,63%)	58 (24,79%)	73 (31,20%)
1.000upstream	205 (20,50%)	185 (18,50%)	195 (19,50%)	268 (26,80%)	ERRO*	238 (23,80%)	210 (21,00%)

Os valores apresentados entre parênteses como porcentagem representam o número de variações em relação ao tamanho do segmento analisado. (*) Erro no alinhamento entre as sequências – Não ocorrência de similaridade entre os pares.

Tabela 23: Número de variações entre *T. equinum* e *T. tonsurans* e outras espécies do gênero *Trichophyton* para o gene que codifica metaloprotease 3 (Mep3).

Referência TEQG_02420		
	TESG_02775	<i>Trichophyton spp.</i>
Éxon	5 (0,30%)	90 (5,41%)
Íntron	1 (0,16%)	56 (8,68%)
1.000upstream	18 (1,80%)	118 (11,80%)

Os valores apresentados entre parênteses como porcentagem representam o número de variações em relação ao tamanho do segmento analisado.

Tabela 24: Número de variações entre *T. equinum* e *T. tonsurans* e outras espécies do gênero *Trichophyton* para o gene que codifica metaloprotease 4 (Mep4).

Referência TEQG_04323		
	TESG_00401	<i>Trichophyton spp.</i>
Éxon	0 (0,00%)	149 (7,71%)
Íntron	0 (0,00%)	50 (19,76%)
1.000upstream	8 (0,80%)	102 (10,20%)

Os valores apresentados entre parênteses como porcentagem representam o número de variações em relação ao tamanho do segmento analisado.

As proteínas preditas codificadas por cada um dos genes foram submetidas a uma análise de identidade, observando-se o grau de divergência entre as espécies. Nas Tabelas 25 e 26, estão expostos os resultados obtidos a partir do alinhamento das sequências.

Tabela 25: Matriz de distância para a proteína metaloprotease 3 (Mep3).

	MCYG_04620	ARB_05085	H101_04475	MGYG_08200	TEQG_02420	TERG_03248	TESG_02775	TRV_06691
MCYG_04620	ID							
ARB_05085	38	ID						
H101_04475	37	13	ID					
MGYG_08200	43	27	26	ID				
TEQG_02420	56	34	21	47	ID			
TERG_03248	41	6	13	27	34	ID		
TESG_02775	37	13	0	26	21	13	ID	
TRV_06691	40	3	13	26	34	7	13	ID

Nas linhas e colunas, estão representadas as proteínas para cada uma das espécies analisadas. O cruzamento de uma linha com uma coluna representa a distância estimada entre as sequências de aminoácidos correspondentes. ID=Identidade.

Os resultados dessa análise mostraram que, para Mep3, a distância entre as espécies *T. tonsurans* e *T. equinum* foi maior que ao serem comparados *T. tonsurans* e *T. interdigitale*. Esse padrão é diferenciado dos demais genes do estudo, pois o motivo é a grande região de baixa qualidade encontrada na sequência gênica de *T. equinum*, que posteriormente é traduzida em proteína. Em sua sequência gênica, foi observada a ocorrência de 351 bases identificadas como “N”, o que sugere uma possível falha no sequenciamento. Mais uma vez, verificou-se que a espécie de maior distância foi *M. canis*.

Tabela 26: Matriz de distância para a proteína metaloprotease 4 (Mep4).

	MCYG_02088	ARB_00762	H101_01910	MGYG_03984	TEQG_04323	TERG_04324	TESG_00401	TRV_00081
MCYG_02088	ID							
ARB_00762	158	ID						
H101_01910	159	22	ID					
MGYG_03984	167	52	52	ID				
TEQG_04323	162	23	3	53	ID			
TERG_04324	164	20	28	56	29	ID		
TESG_00401	162	23	3	53	0	29	ID	
TRV_00081	159	4	23	52	24	21	24	ID

Nas linhas e colunas, estão representadas as proteínas para cada uma das espécies analisadas. O cruzamento de uma linha com uma coluna representa a distância estimada entre as sequências de aminoácidos correspondentes. ID=Identidade.

Para Mep4, foi possível verificar que, em *M. canis*, apesar de possuir estruturas gênicas similares (Figura 33), o número de variações de nucleotídeos encontrados, quando

comparado aos das outras espécies, é relativamente alto (Tabela 22), e isso reflete sua maior distância quando comparadas suas proteínas preditas (Tabela 26).

Nessa análise, mais uma vez, constatou-se que o grau de conservação das sequências de aminoácidos está relacionado à conservação das estruturas dos genes e à composição de seus nucleotídeos.

Visando verificar as relações filogenéticas entre as subfamílias das proteínas Mep3 e Mep4, foram construídos cladogramas para cada um dos grupos de sequências de aminoácidos (Figura 34 e Figura 35).

Percebeu-se nos cladogramas que as subfamílias analisadas formam agrupamentos distintos, mas possivelmente característicos do grupo das metaloproteases, tendo em vista que os clados formados possuem padrões similares para Mep3 e Mep4.

Novamente, não foi possível identificar agrupamentos característicos para os grupos antropofílicos, zoofílicos e geofílicos. Observa-se que os grupos formados estão mais associados ao gênero de cada espécie, comprovando a existência de sinal filogenético e derivação de um mesmo ancestral. Fica evidente, assim, a similaridade encontrada para Mep3 entre as espécies *T. tonsurans* e *T. interdigitale*. Os resultados mostraram que elas estão no mesmo ramo e não apresentam diferenças em relação ao tempo de surgimento da espécie.

Figura 34: Relação filogenética entre os dermatófitos com base nas sequências da proteína metaloprotease 3 (Mep3).

Figura 35: Relação filogenética entre os dermatófitos com base nas sequências da proteína metaloprotease 4 (Mep4).

Para Mep3 e Mep4 constataram-se cinco subgrupos principais: um subgrupo entre as espécies *T. verrucosum*, *A. benhamiae* e *T. rubrum*; um segundo subgrupo entre *T. equinum*, *T. tonsurans* e *T. interdigitale*; um terceiro subgrupo contendo apenas *M. gypseum*; um quarto com *M. canis*; por último, um subgrupo que representa o grupo externo formado pelas espécies *C. immitis* e *C. posadasii*. A distância entre *M. canis* e as outras espécies (mesmo pertencentes ao mesmo gênero) foi mais evidente nas proteínas do grupo das metaloproteases que nas subtilisinas.

Nas buscas por motivos de ligação em regiões promotoras, realizadas por meio do banco de dados de sítios de ligação para fator de transcrição de *S. cerevisiae* (YEASTRACT), foi possível verificar que a frequência encontrada para alguns tipos de sítios de ligações é bastante aumentada, independentemente do gene estudado. As Figuras 36 e 38 apresentam as análises feitas para *mep3* e *mep4*, respectivamente.

Figura 36: Número de sítios de ligação para fatores de transcrição para *mep3*. Análise com base no banco de dados YEASTRACT (*Yeast Search for Transcriptional Regulators And Consensus Tracking*).

Novamente, verificou-se grande quantidade de possíveis sítios de ligação para os fatores transcripcionais, entre eles destacam-se: Gcr1p, Hac1p, Mot3p, Skn7p, Stb5p e Xbp1p. De acordo com esse resultado, sugere-se que tais genes podem ser regulados por cada um desses fatores transcripcionais (Figura 36 e Figura 38).

O tamanho das sequências consenso, como no caso de Stb5p (5 pares de base), é um fator determinante para que os números de sítios de ligação encontrados atingissem altos valores. Em estudo realizado durante desenvolvimento deste trabalho, definiu-se um índice relacionado à chance de uma pequena sequência consenso pertencer a posições precisas dentro de um genoma. Para isso, utilizaram-se sequências completas de genomas disponíveis em bancos de dados públicos em conjunto com genomas gerados aleatoriamente, mas de tamanhos similares aos anteriores. Os genomas aleatórios foram gerados utilizando a frequência de 0,25 para cada tipo de nucleotídeo (A/C/G/T). Após a construção de duas bases de dados denominadas Genoma original e Genoma aleatório, buscas utilizando sequências consenso de sítios de ligações conhecidos na literatura (de tamanhos variados) foram realizadas, almejando identificar o número de ocorrências em cada um dos tipos de genomas.

Por último, um índice foi determinado a partir da relação Genoma original / Genoma aleatório (Figura 37).

Figura 37: Índice O/A (Original/Aleatório). Quanto maior o índice, maior a chance de uma pequena sequência pertencer a posições precisas dentro de um genoma. Foram utilizados genomas de tamanho médio de 5Mb para a confecção dessa análise (o genoma aleatório foi gerado utilizando a frequência de 0,25 para cada nucleotídeo A/C/G/T).

O gráfico sugere que a partir de 14 pares de base, aumenta-se consideravelmente a chance de uma pequena sequência pertencer a posições precisas dentro de um genoma.

Figura 38: Número de sítios de ligação para fatores de transcrição para *mep4*. Análise com base no banco de dados YEASTRACT (Yeast Search for Transcriptional Regulators And Consensus Tracking)

Visando encontrar melhores consensos nas regiões promotoras dos genes que codificam para Mep3 e Mep4, sequências de 1.000 pares de base *upstream* de cada uma das espécies foram alinhadas. A inclusão de organismos pertencentes a um grupo externo foi necessária, tendo em vista que a similaridade entre os dermatófitos, mesmo nessa região, é relativamente alta. A Tabela 27 apresenta parte da análise realizada para metaloprotease 3 (*mep3*) e há nessa tabela uma sequência consenso (M2) encontrada em todos os dermatófitos selecionados (com exceção de *T.interdigitale*, que não possui sequência de 1.000 pares de base *upstream* completa para esse gene), porém, ausentes nos organismos pertencentes ao grupo externo. Constatou-se, também, a ausência de M3 nos organismos do gênero *Microsporum*. E não foi possível identificar consensos pertencentes apenas ao gênero *Microsporum*.

Tais análises também foram realizadas para metaloprotease 4 (*mep4*) e estão representadas na Tabela 28.

Tabela 27: Amostra de sequências consenso obtidas por meio do alinhamento da região 1.000 pares de base *upstream* do gene *mep3* dos oito dermatófitos em estudo, incluindo dois organismos como grupo externo.

1.000upstream	Motivos identificados		
	M1	M2	M3
MCYG_04620	2	1	-
ARB_05085	4	1	2
H101_04475 ^(a)	1	-	-
MGYG_08200	3	1	-
TEQG_02420	2	1	1
TERG_03248	4	1	2
TESG_02775	2	1	1
TRV_06691	5	1	2
CPAG_07716	2	-	2
CIMG_10191	3	-	2

M1=GCT[TC]C[TA][CG]CAG[AC]; M2=TCCAGCCGTGG; M3=A[CG]CACA[GC]AG[TG]GAC[AT]G.

^(a) Possui região com apenas 58 pares de base.

(*) O colchete é usado para denominar um conjunto e representa a ocorrência de uma de suas bases na posição.

Tabela 28: Amostra de sequências consenso obtidas por meio do alinhamento da região 1.000 pares de base *upstream* do gene *mep4* dos oito dermatófitos em estudo, incluindo dois organismos como grupo externo.

1.000upstream	Motivos identificados		
	M1	M2	M3
MCYG_02088		2	
ARB_00762	5	3	2
H101_01910	4	3	2
MGYG_03984	2	2	
TEQG_04323	3	3	2
TERG_04324	1	1	2
TESG_00401	4	3	2
TRV_00081	4	2	1
CPAG_07716	1		2
CIMG_06073	1		

M1=[CTA][AC][TA]TC[TA]G[CG][AT][CGT][AG][AGT][TC][AT]G; M2=[TC]G[AG][GT][ATC]T[CG][CT]CGT[GC]TG[CG];

M3=[TC]CC[AC]CGGCCA[CT][AT]T[GTC]C.

(*) O colchete é usado para denominar um conjunto e representa a ocorrência de uma de suas bases na posição.

Alguns consensos foram identificados em todas as sequências analisadas (incluindo o grupo externo), o que aumenta a possibilidade de estarem relacionados à função básica do gene. Também foi possível constatar que o número de sequências consenso pode variar de acordo com o organismo.

5.2.2.2 Expressão gênica

A expressão dos genes *mep3* e *mep4*, que codificam as proteínas metaloproteases Mep3 e Mep4, respectivamente, foram analisadas após o cultivo das linhagens com queratina nos tempos de 24 e 96 horas. Após a quantificação dos transcritos, todas as linhagens e condições foram normalizadas pelo menor valor de expressão observado, que foi o da espécie *M. canis* no tempo de 24 horas (Figura 39 a Figura 42).

Notou-se um maior acúmulo de mRNA de *mep3* em *T. rubrum* (próximo a 500 vezes), *T. equinum* (cerca de 45 vezes) e *T. tonsurans* (cerca de 75 vezes), todos em relação ao controle no tempo de 24 horas. Quando comparados os tempos entre as espécies, os resultados mostraram que para *T. rubrum* o acúmulo de transcritos é reduzido significativamente nos tempos mais prolongados do cultivo (96 horas). E *T. tonsurans* apresentou um aumento dos transcritos nessas mesmas condições (96 horas) (Figura 39).

Figura 39: Gráfico dos valores de expressão relativa (*fold change*) para o gene *mep3* em cinco dos dermatófitos em estudo. Para cada organismo, o gráfico apresenta os valores em 24 horas e 96 horas.

Quando observados os padrões de níveis de transcrição identificados para *mep3* (Figura 40), constata-se *T. rubrum* com maior distância em relação às demais espécies e isso ocorre devido ao acúmulo significativo da quantidade de transcritos nos tempos de 24 horas. Por outro lado, um agrupamento é observado para *M. canis* e *T. interdigitale* devido aos baixos níveis de transcrição identificados tanto nos tempos de 24 horas como em 96 horas.

Figura 40: Heatmap construído com os valores de expressão para *mep3*.

Na análise de expressão do gene *mep4* entre os dermatófitos cultivados em queratina (Figura 41), apenas na espécie *T. rubrum* no tempo de 24 horas (quando comparado à condição controle) foi observado um acúmulo significativo da quantidade de transcritos. Na matriz de expressão gênica com análise classificatória hierárquica (Figura 42), foi possível observar maior distanciamento da espécie *T. rubrum* em relação às demais.

Figura 41: Gráfico dos valores de expressão relativa (*fold change*) para o gene *mep4* em cinco dos dermatófitos em estudo. Para cada organismo, o gráfico apresenta os valores em 24 horas e 96 horas.

Figura 42: Heatmap construído com os valores de expressão para *mep4*.

5.2.3 Leucina-aminopeptidases

5.2.3.1 Análises por Bioinformática

Com o intuito de verificar a similaridade entre os genes que codificam para leucina aminopeptidase 1 (Lap1) e leucina aminopeptidase 2 (Lap2) entre os organismos em estudo, cujas proteínas atuam na extremidade da cadeia peptídica, foram realizados alinhamentos entre os ortólogos, além da comparação de suas estruturas éxon/intron.

Figura 43: (A) Alinhamento do conteúdo gênico entre os genes *lap1*. A linha vertical indica o deslocamento dos segmentos alinhados. (B) Representação esquemática dos genes nos oito dermatófitos em estudo. Os organismos estão dispostos na seguinte ordem: *M. canis*, *A. benhamiae*, *T. interdigitale*, *M. gypseum*, *T. equinum*, *T. rubrum*, *T. tonsurans* e *T. verrucosum*.

Para *lap1* (Figura 43), verificou-se que a maioria das espécies, com exceção de *T. tonsurans*, apresenta o mesmo número de íntrons. Outro aspecto importante foi que para *T. rubrum* o gene tem maior tamanho e apresenta diferenças nos tamanhos de éxons em relação às outras espécies. Nas análises de similaridade entre os genes, constatou-se que as regiões menos similares se encontram nas faixas próximas a 200 e 1.000 pares de base (regiões de íntrons na maioria das espécies). Observa-se, também, um deslocamento da região de

similaridade na espécie *T. rubrum* e que os primeiros 750 pares de base, aproximadamente, não fazem parte do alinhamento com as outras espécies.

Figura 44: Representação esquemática dos genes *lap2* nos oito dermatófitos em estudo. Os organismos estão dispostos na seguinte ordem: *M. canis*, *A. benhamiae*, *T. interdigitale*, *M. gypseum*, *T. equinum*, *T. rubrum*, *T. tonsurans* e *T. verrucosum*. (*) Isoformas descritas em *T. rubrum* para o gene representado.

Os resultados para *lap2* (Figura 44) mostraram posicionamento e tamanho similar dos íntrons na maioria das espécies. A espécie *M. gypseum* possui um íntron a mais próximo à região 3' UTR (*Untranslated Region*) do gene, todavia, a sequência disponibilizada nessa região possui baixa qualidade, o que sugere uma possível falha no sequenciamento. Nota-se também a presença de três isoformas descritas para *T. rubrum*. Uma pequena diferença no tamanho do gene, assim como do primeiro éxon, foi observada quando comparadas as espécies *T. equinum* e *T. tonsurans*. As regiões próximas a 100-200, 500-600 e 900-1.000 obtiveram menor taxa de similaridade entre os genes (regiões de íntrons). Assim como em *lap1*, o gene *lap2* tem maior tamanho na espécie *T. rubrum*.

Para avaliar as modificações pontuais nas regiões de éxon, íntron e 1.000 pares de base *upstream* dos genes que codificam para Lap1 e Lap2, realizou-se uma comparação entre as espécies, utilizando como referência *M. canis* (Tabela 29 e Tabela 30).

Tabela 29: Número de variações entre *M. canis* e outras espécies de dermatófitos para o gene que codifica leucina aminopeptidase 1 (Lap1).

Referência MCYG_02286							
	ARB_03568	H101_03813	MGYG_08250	TEQG_06332	TERG_05652	TESG_07124	TRV_06599
Éxon	132 (11,76%)	145 (12,92%)	108 (9,63%)	145 (12,92%)	114 (10,16%)	146 (13,01%)	138 (12,30%)
Íntron	40 (32,79%)	43 (32,25%)	40 (32,79%)	49 (40,16%)	24 (19,67%)	51 (41,80%)	37 (30,33%)
1.000upstream	225 (22,50%)	220 (22,00%)	208 (20,80%)	185 (18,50%)	56 (5,60%)	221 (22,10%)	228 (22,80%)

Os valores apresentados entre parênteses como porcentagem representam o número de variações em relação ao tamanho do segmento analisado.

Tabela 30: Número de variações entre *M. canis* e outras espécies de dermatófitos para o gene que codifica leucina aminopeptidase 2 (Lap2).

Referência MCYG_06199							
	ARB_00494	H101_04588	MGYG_06635	TEQG_05395	TERG_08405	TESG_06750	TRV_01590
Éxon	218 (14,65%)	226 (15,19%)	246 (16,53%)	231 (15,52%)	228 (15,32%)	201 (13,51%)	229 (15,39%)
Íntron	75 (30,99%)	71 (29,34%)	59 (24,38%)	64 (26,45%)	70 (28,93%)	66 (27,27%)	77 (31,82%)
1.000upstream	141 (14,10%)	197 (19,70%)	135 (13,50%)	195 (19,50%)	86 (8,60%)	192 (19,20%)	194 (19,40%)

Os valores apresentados entre parênteses como porcentagem representam o número de variações em relação ao tamanho do segmento analisado.

Foi possível, de novo, observar que o número de variações encontradas é relativamente alto e o motivo é a escolha do genoma de referência (*M. canis*). Por outro lado, quando utilizado *T. equinum* como referência, os resultados obtidos quando comparados à *T. tonsurans*, ou às outras espécies do gênero *Trichophyton*, apresentam redução importante no número de variações (Tabela 31 e Tabela 32).

Tabela 31: Número de variações entre *T. equinum* e *T. tonsurans* e outras espécies do gênero *Trichophyton* para o gene que codifica leucina aminopeptidase 1 (Lap1).

Referência TEQG_06332		
	TESG_07124	<i>Trichophyton spp.</i>
Éxon	1 (0,09%)	85 (7,58%)
Íntron	1 (0,74%)	29 (21,32%)
1.000upstream	2 (0,20%)	249 (24,90%)

Os valores apresentados entre parênteses como porcentagem representam o número de variações em relação ao tamanho do segmento analisado.

Tabela 32: Número de variações entre *T. equinum* e *T. tonsurans* e outras espécies do gênero *Trichophyton* para o gene que codifica leucina aminopeptidase 2 (Lap2).

Referência TEQG_05395		
	TESG_06750	<i>Trichophyton spp.</i>
Éxon	5 (0,34%)	104 (6,99%)
Íntron	1 (0,35%)	54 (19,08%)
1.000upstream	2 (0,20%)	105 (10,50%)

Os valores apresentados entre parênteses como porcentagem representam o número de variações em relação ao tamanho do segmento analisado.

Os resultados apresentados nessa análise mostraram que os percentuais de variações, mesmo entre espécies do mesmo gênero, foram maiores em regiões não codificantes (regiões de íntrons) e regiões promotoras dos genes.

As proteínas preditas codificadas por cada um dos genes *lap1* e *lap2* foram, então, submetidas a uma análise de identidade, almejando-se verificar a conservação de suas estruturas de aminoácidos. Para isso, uma matriz de distância foi construída indicando as divergências encontradas quando comparadas as espécies por pares (Tabela 33 e Tabela 34).

Os resultados dessa análise mostraram que, para Lap1, *T. interdigitale*, *T. tonsurans* e *T. equinum* não possuem diferenças na composição de seus aminoácidos. Por outro lado, a maior diferença persiste na espécie *M. canis* em relação às demais (Tabela 33).

Para Lap2, foi possível verificar que não há diferença entre as proteínas de *T. interdigitale* e *T. equinum*. Já em *T. tonsurans*, uma única variação de aminoácidos foi encontrada quando comparada a essas duas outras espécies (Tabela 34).

Tabela 33: Matriz de distância para a proteína leucina aminopeptidase 1 (Lap1).

	MCYG_02286	ARB_03568	H101_03813	MGYG_08250	TEQG_06332	TERG_05652	TESG_07124	TRV_06599
MCYG_02286	ID							
ARB_03568	28	ID						
H101_03813	27	15	ID					
MGYG_08250	18	19	21	ID				
TEQG_06332	27	15	0	21	ID			
TERG_05652	32	13	24	22	24	ID		
TESG_07124	27	15	0	21	0	24	ID	
TRV_06599	30	2	17	21	17	15	17	ID

Nas linhas e colunas, estão representadas as proteínas para cada uma das espécies analisadas. O cruzamento de uma linha com uma coluna representa a distância estimada entre as sequências de aminoácidos correspondentes. ID=Identidade.

Tabela 34: Matriz de distância para a proteína leucina aminopeptidase 2 (Lap2).

	MCYG_06199	ARB_00494	H101_04588	MGYG_06635	TEQG_05395	TERG_08405	TESG_06750	TRV_01590
MCYG_06199	ID							
ARB_00494	46	ID						
H101_04588	46	16	ID					
MGYG_06635	70	58	56	ID				
TEQG_05395	46	16	0	56	ID			
TERG_08405	44	8	18	57	18	ID		
TESG_06750	47	17	1	56	1	19	ID	
TRV_01590	46	4	18	58	18	8	19	ID

Nas linhas e colunas, estão representadas as proteínas para cada uma das espécies analisadas. O cruzamento de uma linha com uma coluna representa a distância estimada entre as sequências de aminoácidos correspondentes. ID=Identidade.

Visando verificar as relações filogenéticas entre as subfamílias das proteínas Lap1 e Lap2, cladogramas foram construídos para cada um dos grupos de sequências de aminoácidos (Figura 45 e Figura 46).

Os resultados mostraram padrões de agrupamentos similares entre Lap1 e Lap2. E, mais uma vez, não foi possível identificar agrupamentos característicos para os grupos antropílicos, zoofílicos e geofílicos. Observa-se que os grupos formados estão mais associados ao gênero de cada espécie.

Figura 45: Relação filogenética entre os dermatófitos com base nas sequências da proteína leucina aminopeptidase 1 (Lap1).

Figura 46: Relação filogenética entre os dermatófitos com base nas sequências da proteína leucina aminopeptidase 2 (Lap2).

Para Lap1 e Lap2, observaram-se cinco subgrupos principais: um subgrupo entre as espécies *T. verrucosum*, *A. benhamiae* e *T. rubrum*; um segundo subgrupo entre *T. equinum*, *T. tonsurans* e *T. interdigitale*; um terceiro subgrupo contendo apenas *M. gypseum*; um quarto subgrupo com *M. canis*; por último, um subgrupo que representa o grupo externo formado pelas espécies *C. immitis* e *C. posadasii*. Outro ponto importante a se destacar é que, ao ser apresentadas as estruturas dos genes *lap1* e *lap2* para *T. rubrum*, foi identificada maior divergência quando comparadas às outras espécies. Essa divergência, inclusive de tamanho, não se refletiu em sua estrutura proteica.

Para encontrar consensos específicos nas regiões promotoras dos genes que codificam para Lap1 e Lap2, sequências de 1.000 pares de base *upstream* de cada uma das espécies, para cada um dos genes, foram alinhadas. As Tabelas 35 e 36 apresentam parte da análise realizada para leucina aminopeptidase 1 (*lap1*) e leucina aminopeptidase 2 (*lap2*), respectivamente.

Tabela 35: Amostra de sequências consenso obtidas por meio do alinhamento da região 1.000 pares de base *upstream* do gene *lap1* dos oito dermatófitos em estudo, incluindo dois organismos como grupo externo.

1.000upstream	Motivos identificados		
	M1	M2	M3
MCYG_02286	1	2	-
ARB_03568	1	2	1
H101_03813	1	2	1
MGYG_08250	1	2	-
TEQG_06332	1	2	1
TERG_05652	1	1	1
TESG_07124	1	2	1
TRV_06599	1	2	1
CPAG_06313	2	-	-
CIMG_08650	1	-	-

M1=TTCAATTG[CT]TTTCT; M2=CCAG[GA][CT]CC[AG][GC]ATCAC; M3=GAAG[AG]CCTATC[GT]TGC.

(*) O colchete é usado para denominar um conjunto e representa a ocorrência de uma de suas bases na posição.

Tabela 36: Amostra de sequências consenso obtidas por meio do alinhamento da região 1.000 pares de base *upstream* do gene *lap2* dos oito dermatófitos em estudo, incluindo dois organismos como grupo externo.

1.000upstream	Identificador do Motivo		
	M1	M2	M3
MCYG_06199	3	2	
ARB_00494	5	3	1
H101_04588	6	2	1
MGYG_06635	4	1	
TEQG_05395	6	2	1
TERG_08405	4	3	1
TESG_06750	6	2	1
TRV_01590	6	3	1
CPAG_05894	2		2
CIMG_07629	2		1

M1=[GC][GT]G[ACT][CA]G[AC]TCCCTG[CG][CA]; M2=[GC][CA]AAG[TG][TC]A[TA][TA][TC]GCT[CG];

M3=AGT[TC][GT]GGAAG[GA][ACT]AT[GT].

(*) O colchete é usado para denominar um conjunto e representa a ocorrência de uma de suas bases na posição.

Os resultados mostraram que os consensos denominados como M1 se encontram em todas as espécies analisadas, incluindo o grupo externo. No caso dos consensos denominados M2, eles são encontrados apenas em dermatófitos. Por último, observam-se os consensos M3 ausentes nos organismos do gênero *Microsporum*. Para *lap1* e *lap2* não foi possível identificar consensos pertencentes apenas ao gênero *Microsporum* (Tabelas 35 e 36).

5.2.3.2 Expressão gênica

As expressões dos genes *lap1* e *lap2* também foram analisadas após o cultivo das linhagens com queratina nos tempos de 24 e 96 horas. Após a quantificação dos transcritos, todas as linhagens e condições foram normalizadas pelo valor de expressão observado na espécie *M. canis* no tempo de 24 horas, seguindo o padrão de análise utilizado nos outros genes do trabalho (Figura 47 a Figura 50).

Pode-se observar que, ao ser comparado o tempo de 24 horas de *M. canis*, o acúmulo de transcritos nas espécies do gênero *Trichophyton* apresenta níveis significativamente menores, tanto no tempo de 24 horas como no de 96 horas (Figura 47).

Figura 47: Gráfico dos valores de expressão relativa (*fold change*) para o gene *lap1* em cinco dos dermatófitos em estudo. Para cada organismo, o gráfico apresenta os valores em 24 horas e 96 horas.

Figura 48: Heatmap construído com os valores de expressão para *lap1*.

Quando observados os padrões de níveis de transcrição identificados para *lap1*, constata-se *M. canis* com maior distância em relação às demais espécies, conseqüentemente, os organismos do gênero *Trichophyton* formam melhores agrupamentos em relação à expressão desse gene (Figura 48).

A regulação da expressão de *lap2* no cultivo em queratina foi diferenciada significativamente apenas nas espécies *T. rubrum* e *T. tonsurans*, ambas no tempo de 24 horas (quando comparadas à condição controle). Observa-se nestas duas espécies um acúmulo significativo da quantidade de transcritos nos tempos de 24 horas e uma redução da quantidade de transcritos no tempo mais tardio (Figura 49).

Figura 49: Gráfico dos valores de expressão relativa (*fold change*) para o gene *lap2* em cinco dos dermatófitos em estudo. Para cada organismo, o gráfico apresenta os valores em 24 horas e 96 horas.

Na matriz de expressão gênica com análise classificatória hierárquica (Figura 50), foi possível observar maior distanciamento da espécie *T. tonsurans* em relação às demais, devido à quantidade expressiva do número de transcritos observadas no tempo de 24 horas em relação às outras espécies.

Figura 50: Heatmap construído com os valores de expressão para *lap2*.

5.2.4 Dipeptidil peptidase V

5.2.4.1 Análises por Bioinformática

Com o intuito de verificar a similaridade entre os genes que codificam para dipeptidil peptidase 5 (DppV) entre os organismos em estudo, foram realizados alinhamentos entre os ortólogos, além da comparação de suas estruturas éxon/intron.

Para *dppV* (Figura 51), verificou-se um alto grau de paridade quando observada a quantidade e a posição de íntrons. Além disso, é importante ressaltar que embora o gene *dppV* de *T. rubrum* não esteja identificado no site do *Broad Institute*, nós o identificamos a partir da TERG_12606, após sua realocação e reanotação (*Supercontig* 11: 290252-292693 +). Para isto, procedeu-se com o alinhamento da região do *supercontig* de *T. rubrum*, posteriormente, com a predição gênica por meio do *software* Augustus [*gene prediction*] (<http://bioinf.uni-greifswald.de/augustus/>). Nas análises de similaridade entre os genes, também se constatou que as regiões menos similares se encontram nas regiões de íntrons na maioria das espécies.

Figura 51: Representação esquemática dos genes que codificam para DppV nos oito dermatófitos em estudo. Os organismos estão dispostos na seguinte ordem: *M. canis*, *A. benhamiae*, *T. interdigitale*, *M. gypseum*, *T. equinum*, *T. rubrum* (Supercontig 11: 290252-292693 +), *T. tonsurans* e *T. verrucosum*.

Com vistas a avaliar as modificações pontuais nas regiões de éxon, íntron e 1.000 pares de base *upstream* dos genes que codificam para DppV, realizou-se uma comparação entre as espécies utilizando como referência *M. canis* (Tabela 37).

Tabela 37: Número de variações entre *M. canis* e outras espécies de dermatófitos para o gene que codifica dipeptidil peptidase V (DppV).

Referência MCG_01626							
	ARB_06651	H101_03870	MGYG_04446	TEQG_02326	(*)	TESG_00772	TRV_02418
Éxon	265 (12,15%)	274 (12,56%)	317 (14,53%)	275 (12,61%)	275 (12,61%)	274 (12,56%)	236 (10,82%)
Íntron	61 (24,02%)	74 (29,13%)	55 (21,65%)	68 (26,77%)	80 (31,50%)	68 (26,77%)	70 (27,56%)
1.000upstream	118 (11,80%)	135 (13,50%)	151 (15,10%)	123 (12,30%)	130 (13,00%)	121 (12,10%)	116 (11,60%)

(*) *T. rubrum* CBS 118892: Supercontig 11: 290252-292693 +

Os valores apresentados entre parênteses como porcentagem representam o número de variações em relação ao tamanho do segmento analisado.

O número de variações encontradas é relativamente alto e o motivo é a escolha do genoma de referência (*M. canis*). Quando utilizado *T. equinum* como referência, os resultados obtidos, quando comparados a *T. tonsurans* ou às outras espécies do gênero *Trichophyton*, apresentam redução no número de variações (Tabela 38).

Tabela 38: Número de variações entre *T. equinum* e *T. tonsurans* e outras espécies do gênero *Trichophyton* para o gene que codifica dipeptidil peptidase V (DppV).

Referência TEQG_02326		
	TESG_00772	<i>Trichophyton spp.</i>
Éxon	1 (0,05%)	103 (4,72%)
Íntron	0 (0,00%)	38 (14,02%)
1.000upstream	2 (0,20%)	67 (6,70%)

Os valores apresentados entre parênteses como porcentagem representam o número de variações em relação ao tamanho do segmento analisado.

Os resultados apresentados nessa análise mostraram que os percentuais de variações, ainda que entre espécies do mesmo gênero, foram maiores em regiões não codificantes (regiões de íntrons) e regiões promotoras dos genes.

As proteínas preditas codificadas por DppV foram submetidas a uma análise de identidade. Uma matriz de distância foi construída indicando as divergências encontradas (Tabela 39).

Tabela 39: Matriz de distância para a proteína dipeptidil peptidase V (DppV).

	MCYG_01626	ARB_06651	H101_03870	MGYG_04446	TEQG_02326	(*)	TESG_00772	TRV_02418
MCYG_01626	ID							
ARB_06651	59	ID						
H101_03870	65	18	ID					
MGYG_04446	85	58	56	ID				
TEQG_02326	67	20	2	58	ID			
(*)	66	12	20	57	22	ID		
TESG_00772	66	19	1	57	1	21	ID	
TRV_02418	60	3	15	58	17	13	16	ID

(*) Proteína predita para *T. rubrum* CBS 118892: *Supercontig* 11: 290252-292693 +

Nas linhas e colunas, estão representadas as proteínas para cada uma das espécies analisadas. O cruzamento de uma linha com uma coluna representa a distância estimada entre as sequências de aminoácidos correspondentes. ID=Identidade.

A sequência da proteína DppV de *T. rubrum* foi gerada por meio da retirada dos íntrons preditos pelo *software* Augustus e posteriormente traduzida pelo *software* ExPASy Translate (<http://web.expasy.org/translate/>).

Os resultados da análise de identidade mostraram que, para DppV, comparações entre as espécies do gênero *Trichophyton* apresentam menor diferença na composição de seus

aminoácidos. Por outro lado, *M. canis* e *M. gypseum* possuem valores expressivos no número de aminoácidos divergentes (Tabela 39).

As relações filogenéticas entre as proteínas DppV das espécies analisadas foram preditas e representadas na Figura 52.

Figura 52: Relação filogenética entre os dermatófitos com base nas sequências da proteína dipeptidil peptidase V (DppV).

Os resultados mostraram padrões de agrupamentos similares às outras proteínas do estudo. De novo, não foi possível identificar agrupamentos característicos para os grupos antropofílicos, zoofílicos e geofílicos.

Visando encontrar consensos específicos nas regiões promotoras dos genes que codificam para DppV, sequências de 1.000 pares de base *upstream* de cada uma das espécies foram alinhadas. A Tabela 40 apresenta parte da análise realizada para dipeptidil peptidase V (*dppV*).

Os resultados revelaram que os consensos denominados como M1 se encontram em todas as espécies analisadas, incluindo o grupo externo. No caso dos consensos denominados M2, eles são encontrados apenas em dermatófitos. Por último, observavam-se os consensos M3 ausentes nos organismos do gênero *Microsporium*, além do grupo externo. Não foi possível identificar consensos pertencentes apenas ao gênero *Microsporium* (Tabela 40).

Tabela 40: Amostra de sequências consenso obtidas por meio do alinhamento da região 1.000 pares de base *upstream* do gene *dppV* dos dermatófitos em estudo, incluindo dois organismos como grupo externo.

1.000upstream	Motivos identificados		
	M1	M2	M3
MCYG_01626	4	1	-
ARB_06651	5	1	1
H101_03870	5	1	1
MGYG_04446	7	1	-
TEQG_02326	6	1	1
<i>T. rubrum</i> CBS 118892: Supercontig 11: 289252-290251 +	5	1	1
TESG_00772	6	1	1
TRV_02418	6	1	1
CPAG_04152	1	-	-
CIMG_10883	1	-	-

M1=C[CA][GC]AA[GT][CG]T[CT][GA][CT]; M2=ACCAGGTTGAAATCC; M3= GAACATGAGCT[AG]GAG.

(*) O colchete é usado para denominar um conjunto e representa a ocorrência de uma de suas bases na posição.

5.2.4.2 Expressão gênica

A expressão do gene *dppV* foi analisada após o cultivo das linhagens com queratina nos tempos de 24 e 96 horas. Assim como nos outros experimentos do trabalho, após a quantificação dos transcritos, todas as linhagens e condições foram normalizadas pelo valor de expressão observado na espécie *M. canis*, no tempo de 24 horas (Figura 53).

Ao ser comparado o tempo de 24 horas de *M. canis*, *T. rubrum* apresenta níveis significativamente aumentados no acúmulo de transcritos para os tempos de 24 e 96 horas, o mesmo ocorre com *T. equinum* (24 horas). Os resultados para *T. interdigitale* (24 horas) também demonstram quantidade de transcritos maiores quando comparados à condição controle. Os resultados mostraram que *T. equinum* e *T. interdigitale* apresentam maiores valores em tempos iniciais (24 horas), que é reduzido significativamente no tempo mais prolongado do cultivo. Por outro lado, para *T. rubrum* foi observada quantidade expressiva do número de transcritos no tempo de 96 horas (Figura 53 e Figura 54).

Figura 53: Gráfico dos valores de expressão relativa (*fold change*) para o gene *dppV* em cinco dos dermatófitos em estudo. Para cada organismo, o gráfico apresenta os valores em 24 horas e 96 horas.

Figura 54: *Heatmap* construído com os valores de expressão para *dppV*.

Na matriz de expressão gênica com análise classificatória hierárquica (Figura 54), foi possível observar maior distanciamento da espécie *T. rubrum* em relação às demais, devido à quantidade expressiva do número de transcritos observados (96 horas), em relação às outras espécies.

Discussão

Thus, the task is not so much to see what no one yet has seen, but to think what nobody yet has thought about that which everybody sees.

Arthur Schopenhauer

VI. Discussão

As técnicas utilizadas aqui possibilitaram a análise comparativa dos dermatófitos, proporcionando uma comparação molecular e evolucionária entre esses fungos no que diz respeito aos genes envolvidos no processo de degradação de queratina e esta visão holística é necessária para que haja maior entendimento dos mecanismos envolvidos na patogenicidade desses dermatófitos.

Os avanços nas tecnologias de sequenciamento estão mudando o ritmo de projetos de genoma e têm contribuído para o crescente volume de dados genômicos. Isso é acompanhado pelo aumento da capacidade dos recursos computacionais para processar e traduzir os dados brutos em informações significativas.

Neste estudo, dados de sequenciamento NGS (*Next-generation sequencing*) foram utilizados em procedimentos de montagem "*de novo*" para fins de comparação com resultados obtidos no processo de montagem, usando um genoma como referência. Verificou-se na montagem "*de novo*" que *T. interdigitale* (H6) apresentou melhores resultados quando comparado com a montagem realizada por referência. O número de *contigs* resultantes do sequenciamento "*de novo*" foi cerca de 4 vezes menor do que o obtido para a tentativa anterior, sendo que, quanto maior a cobertura, menor o número de *contigs* (significando que os fragmentos de DNA foram montados por completo). A montagem de pequenos segmentos de DNA é uma estratégia desafiadora, mas que se apresentou rápida e eficiente na montagem deste genoma. Para que esta montagem atingisse melhores resultados, optou-se pelo envio dos *contigs*, obtidos na confecção deste trabalho, ao *Broad Institute*, assim, foi possível unir os sequenciamentos obtidos pelos dois grupos, resultando em uma melhor disponibilização do genoma de *T. interdigitale* (H6).

Os demais resultados aqui apresentados podem ser enquadrados em duas frentes principais: a análise comparativa da relação entre os dermatófitos, no que diz respeito à estrutura genômica de genes que codificam proteínas envolvidas no processo de degradação de queratina, e a análise comparativa da expressão desses genes durante a mimetização do processo infeccioso. A modulação da expressão dos genes é um importante mecanismo, pois tem um papel fundamental na resposta aos estímulos ambientais, ao controle do metabolismo

celular frente às reações adversas, às mudanças abióticas e/ou nutricionais e, finalmente, à adaptação em novos nichos parasitados.

A análise de endo e exoproteases (que constituem uma importante maquinaria capaz de digerir uma cadeia polipeptídica intacta em peptídios curtos e aminoácidos) foi o objetivo principal deste trabalho, que culminou na identificação de padrões diferenciados entre as espécies no que diz respeito à estrutura e à expressão dos genes analisados.

Os resultados obtidos no presente trabalho demonstram que tanto para o grupo das endoproteases, como para o das exoproteases, ocorre variabilidade minuciosa em relação à estrutura genômica entre as espécies, todavia, diferenças significativas acontecem quando comparados seus níveis transcricionais em resposta à queratina.

Inicialmente, foram realizadas análises comparativas do conteúdo gênico. É possível analisar os genes alinhando-os e utilizando um dos algoritmos de análise de sequências descritos na literatura. No entanto, isso só deve ser feito com genomas de espécies próximas, uma vez que mudanças na estrutura do DNA (inserções, deleções, inversões, rearranjos, trocas e duplicações) ocorrem com uma taxa elevada e dificultam o processo de análise.

A etapa crucial desse tipo de análise é determinar se as sequências comparadas são ou não homólogas, isto é, descendem ou não de uma sequência ancestral comum, estabelecendo-se equivalência entre as partes comparadas.

Pôde-se verificar por meio do algoritmo escolhido um alto grau de similaridade entre todas as espécies selecionadas. As endo e exoproteases secretadas por dermatófitos também são similares às de outras espécies, como por exemplo, *Aspergillus spp.* Porém, as endoproteases secretadas pelos dermatófitos são múltiplas (JOUSSON et al., 2004a; MARTINEZ et al., 2012). Jousson et al. (2004a) afirmaram que eventos de duplicações de genes ancestrais geraram famílias codificadoras de proteases secretadas e necessárias para a instalação da dermatofitose.

A comparação de ORFs (*Open Reading Frame*) preditas para as subtilisinas *sub3*, *sub5* e *sub7*, para as metaloproteases *mep3* e *mep4*, para as leucina-aminopeptidases *lap1* e *lap2* e para a dipeptidil peptidase V (*dppV*), entre os oito dermatófitos em estudo, revelou uma colinearidade das estruturas íntron-éxon na grande maioria das vezes.

As subtilisinas são serino proteases características como fator de virulência de dermatófitos. Brouta et al. (2001) identificaram sua presença em *M. canis* por meio de *screening* em biblioteca de DNA utilizando sonda heteróloga, baseada na sequência de outras

serino proteases fúngicas. No presente trabalho, para os genes desse grupo, foi observada a presença de três íntrons em *sub3*, *sub5* e *sub7* (com exceção de *sub7* com padrões diferenciados para *T. rubrum*, *A. benhamiae* e *T. verrucosum*). Jousson et al. (2004a) clonaram 7 genes de subtilisinas de *T. rubrum* (incluindo *sub3*, *sub5* e *sub7*), avaliando a atividade queratinolítica dessas SUBs e destacando a ótima atividade em pH alcalino.

Outra classe de proteases secretadas por dermatófitos é a de metaloproteases (MEPs) – enzimas que possuem um íon acoplado em seu sítio catalítico. Jousson et al. (2004b) caracterizaram estes genes para as espécies *T. rubrum*, *T. mentagrophytes* e *M. canis* e apontaram que, com exceção de *mep2*, todos os genes MEPs possuem quatro íntrons localizados em posições conservadas, quando comparados 5 tipos de metaloproteases (*mep1-5*). No presente trabalho, a presença de quatro íntros em *mep3* e *mep4* para a grande maioria das espécies do estudo corrobora com os resultados dos autores citados. Vale ressaltar que o gene *mep3* de *T. interdigitale* apresentou tamanho inferior em relação à sua ORF predita e que a espécie *T. equinum* apresentou grande região de baixa qualidade na sequência disponibilizada para *mep3*.

Em uma investigação detalhada dos alinhamentos das sequências nucleotídicas, realizados para os genes MEPs e para os genes SUBs, observou-se que ambos possuem conteúdo gênico altamente conservado, como mencionado anteriormente, principalmente em regiões codificadoras. Ao serem comparadas as oito espécies do estudo e com a utilização de *M. canis* como referência, verifica-se uma divergência relativa³ em torno de 10% a 20%, levando-se em conta apenas as regiões de éxons. Por outro lado, a divergência relativa é em média de 30% a 40% quando analisadas apenas as regiões de íntrons. Estudos sugerem que uma grande parcela dos fragmentos pertencentes às regiões não codificadoras intragênicas produzam RNAs não codificadores (RNAs intrônicos), os quais podem contribuir para o controle das proteínas que regulam a transcrição do DNA em RNA. RNAs não codificadores podem constituir mais um fator que determina as diferenças entre os diversos organismos (OLIVEIRA et al., 2011).

As subtilisinas e as metaloproteases das espécies *T. tonsurans* e *T. equinum* compartilham alto grau de similaridade (em regiões de íntrons e de éxons). Os resultados demonstraram as maiores divergências relativas em *sub5* (0,25% em regiões de exon) e *mep3* (0,16% em regiões de íntrons), o que deve ser considerado muito pequeno. Já foi citado que

³ Representa o número de variações em relação ao tamanho do segmento analisado.

essas duas espécies têm pouca divergência (0,2%) apesar de pertencerem a distintos nichos ecológicos (uma é antropofílica e a outra zoofílica) (WHITE et al., 2008). Análises comparativas de genomas e busca por polimorfismos de um único nucleotídeo têm revelado cada vez mais a natureza complexa da estrutura e organização dos genomas, além da enorme diversidade genética entre organismos (muito acima daquela esperada, mesmo entre isolados de uma mesma espécie), levando a questionamentos importantes sobre os mecanismos pelos quais esses micro-organismos evoluem e como devem ser classificados taxonomicamente (COENYE et al., 2005; BINNEWIES et al., 2006).

Análises de sintenia foram realizadas com o objetivo de estudar vizinhanças entre os genes pertencentes ao estudo. Sintenia foi um termo originalmente cunhado para designar a presença de dois ou mais *loci* gênicos (próximos ou não) no mesmo cromossomo. Refere-se também a duas regiões de genomas distintos que mostram considerável grau de similaridade de sequência entre si e algum grau de conservação da ordem dos genes nessas regiões e que, portanto, têm probabilidade de descender de um ancestral comum (MORENO-HAGELSIEB et al., 2001).

Mais uma vez, constatou-se colinearidade entre os genes vizinhos quando comparadas algumas das espécies, principalmente, quando se tratam do mesmo gênero. Entre os maiores graus de identidade observados, quando analisamos a sintenia gênica estão as espécies *T. rubrum* e *A. benhamiae*, a primeira antropofílica e a segunda zoofílica. Novamente, foi possível observar grande similaridade entre as espécies *T. equinum* e *T. tonsurans* (ambas também distintas em relação ao nicho do hospedeiro).

As relações filogenéticas, no passado, foram construídas com uma visão fenotípica e das características anatômicas dos organismos. Hoje em dia, dados moleculares, que incluem sequências de proteínas e de DNA, são utilizados para a construção de árvores filogenéticas.

Segundo (KONDRASHOV, 1999), os produtos dessas análises têm fornecido as melhores evidências disponíveis para alguns fenômenos evolutivos e, em alguns casos, levado ao refinamento de antigos conceitos.

A duplicação de genes em genomas de eucariotos leva à formação de famílias gênicas variáveis em relação ao número e tamanho de membros (diferentes tecidos, células ou organelas), e em relação ao tempo (diferentes estágios de desenvolvimento). A homologia dessas sequências permanece maior ou menor de acordo com a taxa de mutação e “erros” de *crossing-over* (WAGNER, 2002).

Entre as proteínas ensaiadas neste trabalho foi possível detectar variação de sequência entre as espécies de dermatófitos. Estes dados possibilitaram a construção de árvores filogenéticas, nas quais observou-se que *T. equinum*, *T. tonsurans* e *T. interdigitale* pertencem ao mesmo grupo monofilético, *A. benhamiae* e *T. verrucosum* formam um segundo grupo e que *T. rubrum*, muitas vezes, possui um pequeno número de substituições, fazendo com que essa espécie se aproxime desse último grupo. Foi visto, ainda, que *M. canis* é facilmente separado desses grupos com base nas mutações observadas. A diferenciação entre *T. tonsurans* e *T. equinum* é muito limitada nessas regiões, referentes a um pequeno número de substituições de aminoácidos (entre 0 e 1 aminoácidos, com exceção da Mep3, com 21 substituições). A extensão das linhas entre as relações filogenéticas representa a duração no tempo em relação à origem das espécies analisadas. Foi possível observar na grande maioria dos cladogramas apresentados que *T. tonsurans* e *T. equinum* possuem as mesmas origens, estão no mesmo ramo e não apresentam diferenças em relação ao tempo de surgimento das espécies. Várias regiões podem não ter resolução suficiente para a discriminação de espécies dentro do mesmo gênero, especialmente, quando a taxa de substituição de cada uma delas é baixa e/ou o período de especiação dos acessos testados é recente (SILVA et al., 2013).

A identificação computacional de sítios de ligação de fatores de transcrição é a primeira etapa no estudo de uma região promotora. Diversos *softwares* e bancos de dados públicos estão disponíveis para a predição de sítios de ligação putativos em sequências de DNA, por meio de análise comparativa entre sequências. Em buscas realizadas em um banco de dados de sítios de ligação identificados para *S. cerevisiae*, foi possível detectar variação do número de sítios observados entre as espécies de dermatófitos testadas. Em *sub7*, pôde-se destacar uma diferença significativa no número de sítios de ligação para o fator transcricional Tec1p (CATTCY) em *M. gypseum*, comparado aos outros dermatófitos. Segundo Borneman et al. (2006), entre outros fatores, Tec1p (pertencente à via das MAP quinases) é um regulador chave no crescimento de pseudo-hifas. Esse experimento foi realizado utilizando linhagens diploides de *S. cerevisiae*, quando expostas a ambientes pobres em nitrogênio.

Outro sítio de ligação com grande ocorrência na região promotora de vários genes analisados foi para Stb5p, que é um ativador transcricional de resistência multidrogas. Porém, acredita-se que sua sequência consenso (CGGNS) tenha contribuído para isso, haja vista que a possível redundância de bases em algumas posições da sequência consenso e o pequeno tamanho da sequência facilitem o alinhamento ao acaso com as sequências das regiões

promotoras dos organismos. Para verificar esta possibilidade, buscas por sítios de ligação de Stb5p também foram realizadas nas regiões promotoras dos genes actin (*act*), glyceraldehyde 3-phosphate dehydrogenase (*gapdh*) DNA-dependent RNA polymerase II RPB140 (*rpb2*) e Tubulin beta chain (β -*tub*). Realmente, foi possível observar que o consenso para Stb5p ocorre com grande frequência na região promotora de todos estes genes definidos como controle (resultados não mostrados).

O consenso para o fator de transcrição Ash1p é um dos sítios de ligação mais presentes em *sub3*. Ash1p também é necessária para o crescimento de pseudo-hifas em *S. cerevisiae*; associado a uma cascata reguladora da transcrição que, em resposta à privação de nitrogênio, ativa a expressão do gene *flo11* (codifica uma proteína da superfície celular que é necessária para o crescimento de pseudo-hifas) (CHANDARLAPATY e ERREDE, 1998; PAN e HEITMAN, 2000). Como observado, quando analisado o sítio de ligação Stb5p (CGGNS), a possível redundância de bases nas posições 4 e 5 da sequência consenso, além, de seu tamanho (apenas cinco pares de base), dificulta esse tipo de análise e propicia a ocorrência de falsos positivos.

Sendo assim, com o objetivo de verificar o tamanho ideal para se detectar uma sequência consenso não aleatória, realizou-se um teste que demonstrou que sequências com tamanho de pelo menos 14 pb (pares de base) aumentam consideravelmente sua eficiência quando são analisadas regiões de 5 Mb. Vale ressaltar que esse tamanho de sequência consenso, possivelmente, será inferior quando o estudo for realizado em sequências de 1.000 pb (tamanho médio de regiões promotoras). Estes estudos não foram realizados até o momento, devido aos cuidados que se deve ter quando são focadas as probabilidades de ocorrência de cada uma das bases de nucleotídeo para essas regiões específicas. Sabe-se que as regiões promotoras são ricas em G+C e, dessa maneira, probabilidades diferenciadas deverão ser testadas para cada base nucleotídica.

Ainda em relação às regiões promotoras, foi possível identificar sequências consenso após alinhamento das regiões promotoras dos genes homólogos de todos os dermatófitos. Alguns destes consensos foram encontrados em todos os dermatófitos selecionados, porém, ausentes nos organismos pertencentes ao grupo externo, sugerindo que os mesmos estejam relacionadas às espécies de dermatófitos. Algumas sequências consenso, previstas no estudo, foram encontradas nos organismos do gênero *Microsporum*, por outro lado, ausentes no gênero *Trichophyton*. O inverso também foi observado. O que sugere que esses consensos

estariam relacionados aos gêneros específicos. O número de ocorrência de cada sequência consenso variou de acordo com o organismo, sugerindo que possa ser um dos fatores que contribuam na especificidade ao hospedeiro. Alguns consensos foram identificados em todas as sequências analisadas (incluindo o grupo externo), o que aumenta a possibilidade de estarem relacionados à função básica do gene. É importante ressaltar que a união de fatores de transcrição aos seus motivos específicos de ligação na região promotora pode alterar a transcrição e determinar características fenotípicas.

Diante dos dados aqui analisados por meio da bioinformática e segundo estudos realizados por outros grupos, as proteases possuem sequências altamente conservadas entre os dermatófitos, mas com seu nível de secreção diferenciado entre as espécies (GIDDEY et al., 2007). Nesse caso, sugere-se que as diferentes escolhas de habitat poderiam estar relacionadas com uma expressão diferencial de genes que codificam proteínas proteolíticas homólogas (GIDDEY et al., 2007; MARTINEZ et al., 2012).

Enzimas secretadas, tais como fosfolipases, peptidases e proteases, são cruciais para a virulência dos dermatófitos, uma vez que esses agentes patogênicos infectam o estrato córneo, unhas ou cabelo (GHANNOUM, 2000; JOUSSON et al., 2004b). Durante a infecção, as carboxipeptidases podem contribuir para a virulência de *T. rubrum*, cooperando com endoproteases e aminopeptidases para degradar tecidos queratinizados compactos em péptidos curtos e aminoácidos, que podem ser assimilados (ZAUGG et al., 2008).

Em relação ao desenvolvimento radial dos dermatófitos, Martins (2015) observou padrões diferenciados de crescimento quando comparados meio MEA e meio mínimo suplementado com queratina de origem bovina. A autora verificou que em meio MEA a espécie *T. equinum* apresenta rápida progressão radial, seguida por *T. interdigitale*, *T. rubrum*, *M. canis* e, por último, *T. tonsurans*. Entretanto, em meio mínimo suplementado com a mesma queratina utilizada neste trabalho, *T. interdigitale* apresenta maior taxa de crescimento, seguida de *T. rubrum*, *T. equinum*, *M. canis* e *T. tonsurans*. Por último, ela observou que as duas maiores taxas de crescimento foram de espécies antropofílicas (*T. interdigitale* e *T. rubrum*), demonstrando não haver restrição com relação à fonte de carbono ser de origem animal.

Análises de um secretoma de *A. benhamiae*, durante o crescimento em queratina, confirmaram que as proteases pertencem ao maior grupo de proteínas secretadas. A expressão de alguns genes que codificam tais proteínas foi confirmada, induzida ou reprimida durante a

infecção de *A. benhamiae*, utilizando queratinócitos (BURMESTER et al., 2011). Nos resultados aqui levantados, constatou-se níveis de acúmulo de transcritos diferenciais entre os organismos testados e entre tempos de cultivo (24 horas e 96 horas). Vale ressaltar que não foi possível realizar as análises utilizando a técnica de PCR quantitativa em tempo real, a partir de transcrição reversa (RT-qPCR), para alguns dos dermatófitos selecionados. Entre as espécies não testadas, temos: *M. gypseum*, que apresentou quantidade de microconídios insuficiente para a realização dos experimentos, além das espécies *A. benhamiae* e *T. verrucosum*, que não foram adquiridas.

Vários genes de *A. benhamiae* que codificam endo e exoproteases foram induzidos durante crescimento *in vitro* em queratina e soja, quando comparados ao crescimento em meio Sabouraud. Membros das famílias das subtilisinas, Sub3 e Sub4 e das fungalisinas, Mep1, Mep3 e Mep4 foram fortemente expressos. Uma expressão significativa de genes que codificam exoproteases também foi detectada, e entre elas estão: leucina-aminopeptidases Lap1 e Lap2, Dipeptidil peptidases DppIV e DppV, metalocarboxipeptidase MCPA e serina carboxipeptidase SCPB. Surpreendentemente, nenhum dos genes que codificam essas proteínas (com exceção das subtilisinas) foi induzido durante testes realizados por meio da infecção de porquinhos da índia (análise *in vivo*). Nas comparações entre os experimentos realizados *in vitro* e *in vivo*, Sub3 foi ativada durante a infecção, mesmo que a um nível comparativamente baixo. Em contraste, verificou-se que Sub6 não foi detectada durante o crescimento *in vitro* sobre a queratina, mas apresentou regulação positiva *in vivo*. Além desses, genes que codificam Sub1, Sub2 e Sub7 também foram induzidos nos experimentos *in vivo* (STAIB et al., 2010).

Nos resultados obtidos no presente trabalho verificou-se elevados acúmulos de transcritos de *sub3*, *sub7* e *mep3*, principalmente em organismos do gênero *Trichophyton*. Em *T. equinum* (meio de cultivo contendo queratina bovina como única fonte de nutrientes no tempo de 96 horas) *sub3* apresentou um nível de acúmulo de transcritos 800 vezes maior que em *M. canis* (mesmo meio em 24 horas). *T. rubrum* exibiu 8.000 vezes mais transcritos de *sub7* acumulados em relação a *M. canis* (mesmo meio e tempo de cultivo - 24 horas) e que a quantidade de transcritos de *mep3* acumulado foi superior a 500 vezes também em relação a *M. canis* (mesmo meio e tempo de cultivo - 24 horas).

Em biblioteca de cDNA de *T. rubrum* cultivado em meio com soja foi encontrado um número abundante de ESTs que codificam as subtilisinas Sub3 e Sub5, as fungalisinas Mep1

e Mep3, além de leucina aminopeptidase 2 (Lap2) (ZAUGG et al., 2009). Os autores também analisaram *T. rubrum* crescido em meio com queratina mais soja comparado a *T. rubrum* crescido em Sabouraud. Foi possível identificar que *sub3* e *sub4*, *mep3* e *mep4* foram fortemente induzidos, por outro lado, *sub5* apresentou indução moderada no meio contendo a proteína (ZAUGG et al., 2009). No presente trabalho, para os genes que codificam as subtilisinas Sub3 e Sub5, foi observado padrão similar de expressão, com *sub3* apresentando altos níveis de expressão relativa em organismos do gênero *Trichophyton*, quando comparados à *Microsporum*, e *sub5* com indução moderada quando comparados os dois gêneros. Dados que corroboram com os resultados dos autores citados.

Nos tempos de 24 horas, verificou-se uma colinearidade quando comparados os níveis de acúmulo de transcritos *sub5* com o desenvolvimento radial observado por Martins (2015). Sugere-se que, quanto maior a taxa de germinação nas condições avaliadas, maior será a necessidade de mRNA de *sub5* em tempos iniciais (24 horas).

Neste estudo, a espécie *T. tonsurans* apresentou níveis reduzidos do acúmulo de mRNA no tempos iniciais de cultivo (24 horas) para a maioria das proteases analisadas. Achterman et al. (2011) indicaram menor conidiação para *T. tonsurans* em relação às espécies *M. canis*, *T. rubrum* e *T. equinum*.

T. rubrum apresentou acúmulo significativo da quantidade de transcritos nos tempos de 24 horas para as endopeptidases (SUBs e MEPs) do presente estudo, quando comparados aos tempos de 96 horas. Sugere-se que a maior produção destes transcritos possa se relacionar com a atividade das enzimas e, conseqüentemente, aumentar seu poder de infecção.

O início do processo de infecção por dermatófitos depende do pH ácido da pele, e a instalação e manutenção da infecção dependem da alcalinização do ambiente (FERREIRA-NOZAWA et al., 2003; MARTINEZ-ROSSI et al., 2004). Martins (2015) indica que durante o crescimento dos dermatófitos o pH dos meios de cultivo atinge, aproximadamente 8,0 para as espécies *M. canis*, *T. equinum*, *T. rubrum* e *T. tonsurans*, ultrapassando 8,0 apenas para *T. interdigitale* ao final do mesmo tempo analisado neste estudo (96 horas). Vale ressaltar que a alcalinização do meio proporciona um ambiente adequado para a atuação da maioria das proteases queratinolíticas (TSUBOI et al., 1989; JOUSSON et al., 2004a), que são enzimas essenciais para os processos infectivos. De modo geral, verificou-se que os genes que codificam proteases secretadas foram induzidos nos dermatófitos em estudo. Sugere-se que os aumentados níveis de transcritos, mesmo em tempos inferiores a 96 horas, estejam

relacionados à tentativa de produzir material suficiente para garantir a quebra da queratina por essas enzimas, com o objetivo de adesão dos dermatófitos ao substrato.

O meio Sabouraud é conhecido por induzir uma menor atividade proteolítica por dermatófitos do que os meios contendo proteínas como única fonte de carbono e nitrogênio (JOUSSON et al., 2004a; JOUSSON et al., 2004b; MONOD et al., 2005). Embora a utilização de queratina supostamente tenha papel importante na patogenicidade dos dermatófitos, todo processo de adaptação do hospedeiro durante a infecção parece ser mais complexo. Análises do transcriptoma de *A. benhamiae* mostraram que apenas algumas das proteases normalmente induzidas por queratina foram encontradas fortemente induzidas durante a interação fungo-queratinócito (BURMESTER et al., 2011).

A sinergia entre endo e exoproteases é provável que seja essencial para a virulência dos dermatófitos, uma vez que estes fungos crescem preferencialmente em tecidos queratinizados (MONOD et al., 2005). Estudos demonstraram que a atividade de aminopeptidases (da classe das exoproteases) foi detectada a partir do sobrenadante de cultura e micélio de diferentes espécies de fungos, e que essa atividade exoproteolítica poderia desempenhar um papel importante no desenvolvimento dos fungos durante a infecção. É altamente provável que apenas os aminoácidos ou peptídeos curtos digeridos de queratina possam ser usados pelos dermatófitos como fonte de nutriente para o crescimento *in vivo* (DE BERSAQUES e DOCKX, 1973; MONOD et al., 2005). A natureza das aminopeptidases secretadas em *T. rubrum* foi investigada em experimentos de Monod et al. (2005) e os autores sugerem atividade ótima com pH 8,5 para Lap1 e pH 7,0 para outras LAPs (entre elas, Lap2).

Por meio do presente trabalho, constatou-se que *lap1* apresentou níveis bastante reduzidos do acúmulo de transcritos para as espécies do gênero *Trichophyton* (tempos de 24 horas e 96 horas), quando comparados à *M. canis* (24 horas). Em relação à *lap2*, observou-se redução do acúmulo de mRNA nos tempos de 96 horas para todas as espécies, quando comparadas aos tempos de 24 horas. Sugere-se que o acúmulo de transcritos de *lap2* não é tão necessário para o crescimento de todas as espécies de dermatófitos em queratina, após 96 horas de cultivo, quando o pH do meio já está alcalino.

Beauvais et al. (1997) relataram em estudos com *Aspergillus fumigatus* que DppV foi ativada ao longo de um grande intervalo de valores de pH (entre 6 e 8), com atividade ótima em pH 6,5, independentemente do substrato utilizado. Os autores ainda apontaram, em relação à influência do pH na atividade de DppV, maiores valores de atividades em pH 6,0,

atingindo o pico de atividade (para a enzima) em pH 6,5 e redução significativa em pH 8,0. Na análise do gene que codifica a proteína DppV foi observado, no presente trabalho, que nos tempos de cultivo de 24 horas, as espécies *T. equinum* e *T. interdigitale*, apresentam um expressivo acúmulo de mRNA, diminuindo nos tempos tardios (96 horas). Por outro lado, uma inversão deste padrão de modulação ocorre para *T. tonsurans* e *T. rubrum*. Este último alcançando significativo aumento no acúmulo de mRNA (cerca de 35 mil vezes) em relação ao controle. Foi possível observar também a baixa expressão deste gene em *M. canis*, independente do tempo de cultivo.

Portanto, embora a sequência das proteases seja altamente conservada entre as espécies de dermatófitos, a expressão destes genes é variável, dependendo da espécie. A variação do nível de secreção de algumas destas proteases entre os dermatófitos já havia sido observada, conforme mencionado acima, o que corrobora nossos resultados. Consequentemente, este conjunto protéico com funções sinérgicas e provavelmente compensatórias, que não é idêntico para todas as espécies, contribui para os diferentes graus de reações inflamatórias no hospedeiro, bem como para a especificidade patógeno-hospedeiro.

Os resultados alcançados neste trabalho podem se tornar atraentes para estudos futuros sobre a relação patógeno-hospedeiro e auxiliar no entendimento da biologia desses fungos, representando um avanço no conhecimento desses dermatófitos, no que diz respeito a genes envolvidos no processo de degradação de queratina e, consequentemente, na virulência.

Conclusões

The scientist is not a person who gives the right answers, he's the one who asks the right questions.

Claude Lévi-Strauss

VII. Conclusões

É notável o alto grau de conservação entre as sequências dos genes selecionados neste estudo, comprovando a existência de sinal filogenético e derivação de um mesmo ancestral. Apesar desta conservação, observaram-se maiores índices de polimorfismo de um único nucleotídeo em regiões intrônicas, o que pode se constituir em um fator que determina as diferenças entre os organismos.

Através do alinhamento da região promotora de genes que codificam cada protease foram revelados consensos que podem ser específicos para dermatófitos e consensos gênero-específico.

As sequências das proteínas subtilisinas, metaloproteases, leucina-aminopeptidases e dipeptidil peptidase estão associadas aos gêneros e não ao habitat de cada dermatófito.

O uso do PCR em tempo real possibilitou a análise da expressão diferencial comparativa entre os genes que codificam proteases em dermatófitos, durante a mimetização do processo infeccioso, estabelecendo um padrão de resposta diferenciada mesmo entre espécies pertencentes ao mesmo gênero e que possuem o mesmo nicho preferencial.

Para *lap1*, por exemplo, o acúmulo de transcritos está preferencialmente relacionado à espécie *M. canis*, quando comparado aos organismos do gênero *Trichophyton*.

Linhagens das diferentes espécies de dermatófitos analisadas modulam o acúmulo de transcritos dos genes *sub3*, *sub5*, *sub7*, *mep3*, *mep4*, *dppV*, *lap1* e *lap2*, em função do tempo de cultivo em queratina e, conseqüentemente, da alcalinização do meio de cultivo.

O presente trabalho contribuiu para o desenvolvimento de algoritmos computacionais, agregando mecanismos acerca de genômica comparativa que podem auxiliar em estudos futuros.

Referências bibliográficas

El conocimiento nos hace responsables.

Che Guevara

Referências bibliográficas

- ABBY, S. e DAUBIN, V. Comparative genomics and the evolution of prokaryotes. **Trends Microbiol**, v.15, n.3, Mar, p.135-41. 2007.
- ABDEL-RAHMAN, S. M., SIMON, S., WRIGHT, K. J., NDJOUNTCHE, L. e GAEDIGK, A. Tracking *Trichophyton tonsurans* through a large urban child care center: defining infection prevalence and transmission patterns by molecular strain typing. **Pediatrics**, v.118, n.6, Dec, p.2365-73. 2006.
- ACHTERMAN, R. R., SMITH, A. R., OLIVER, B. G. e WHITE, T. C. Sequenced dermatophyte strains: growth rate, conidiation, drug susceptibilities, and virulence in an invertebrate model. **Fungal Genet Biol**, v.48, n.3, Mar, p.335-41. 2011.
- ACHTERMAN, R. R. e WHITE, T. C. Dermatophyte virulence factors: identifying and analyzing genes that may contribute to chronic or acute skin infections. **Int J Microbiol**, v.2012, p.358305. 2012.
- ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. e LIPMAN, D. J. Basic local alignment search tool. **J Mol Biol**, v.215, n.3, Oct 5, p.403-10. 1990.
- ATLAS, R. M. **Handbook of microbiological media**: CRC Press. 1993. 786 p. (Perks, L. C.)
- BALDO, A., MATHY, A., TABART, J., CAMPONOVA, P., VERMOUT, S., MASSART, L., MARECHAL, F., GALLENI, M. e MIGNON, B. Secreted subtilisin Sub3 from *Microsporum canis* is required for adherence to but not for invasion of the epidermis. **Br J Dermatol**, v.162, n.5, May, p.990-7. 2010.
- BALDO, A., MONOD, M., MATHY, A., CAMBIER, L., BAGUT, E. T., DEFAWEUX, V., SYMOENS, F., ANTOINE, N. e MIGNON, B. Mechanisms of skin adherence and invasion by dermatophytes. **Mycoses**, v.55, n.3, May, p.218-23. 2012.
- BEAUVAIS, A., MONOD, M., DEBEAUPUIS, J. P., DIAQUIN, M., KOBAYASHI, H. e LATGE, J. P. Biochemical and antigenic characterization of a new dipeptidyl-peptidase isolated from *Aspergillus fumigatus*. **J Biol Chem**, v.272, n.10, Mar 7, p.6238-44. 1997.
- BINNEWIES, T. T., MOTRO, Y., HALLIN, P. F., LUND, O., DUNN, D., LA, T., HAMPSON, D. J., BELLGARD, M., WASSENAAR, T. M. e USSERY, D. W. Ten years of bacterial genome sequencing: comparative-genomics-based discoveries. **Funct Integr Genomics**, v.6, n.3, Jul, p.165-85. 2006.
- BORNEMAN, A. R., LEIGH-BELL, J. A., YU, H. Y., BERTONE, P., GERSTEIN, M. e SNYDER, M. Target hub proteins serve as master regulators of development in yeast. **Genes & Development**, v.20, n.4, Feb 15, p.435-448. 2006.

BROAD_INSTITUTE. Dermatophyte Comparative Sequencing Project, Broad Institute of Harvard and MIT. Disponível em: <http://www.broadinstitute.org>. Acesso em: 10 jan. 2015.

BROUTA, F., DESCAMPS, F., FETT, T., LOSSON, B., GERDAY, C. e MIGNON, B. Purification and characterization of a 43.5 kDa keratinolytic metalloprotease from *Microsporum canis*. **Med Mycol**, v.39, n.3, JUN, p.269-275. 2001.

BUDIHARDJA, D., FREUND, V. e MAYSER, P. Widespread erosive tinea corporis by *Arthroderma benhamiae* in a renal transplant recipient: case report. **Mycoses**, v.53, n.6, Nov, p.530-2. 2010.

BURDEN, A. D., TILLMAN, D. M. e RICHARDSON, M. D. Human *Trichophyton equinum* infection treated with terbinafine. **Clin Exp Dermatol**, v.19, n.4, Jul, p.359-60. 1994.

BURMESTER, A., SHELEST, E., GLOCKNER, G., HEDDERGOTT, C., SCHINDLER, S., STAIB, P., HEIDEL, A., FELDER, M., PETZOLD, A., SZAFRANSKI, K., FEUERMANN, M., PEDRUZZI, I., PRIEBE, S., GROTH, M., WINKLER, R., LI, W., KNIEMEYER, O., SCHROECKH, V., HERTWECK, C., HUBE, B., WHITE, T. C., PLATZER, M., GUTHKE, R., HEITMAN, J., WOSTEMEYER, J., ZIPFEL, P. F., MONOD, M. e BRAKHAGE, A. A. Comparative and functional genomics provide insights into the pathogenicity of dermatophytic fungi. **Genome Biol**, v.12, n.1, Jan 19, p.R7. 2011.

BURZYKOWSKI, T., MOLENBERGHS, G., ABECK, D., HANEKE, E., HAY, R., KATSAMBAS, A., ROSEEuw, D., VAN DE KERKHOF, P., VAN AELST, R. e MARYNISSEN, G. High prevalence of foot diseases in Europe: results of the Achilles Project. **Mycoses**, v.46, n.11-12, Dec, p.496-505. 2003.

BUSTIN, S. A., BENES, V., GARSON, J. A., HELLEMANS, J., HUGGETT, J., KUBISTA, M., MUELLER, R., NOLAN, T., PFAFFL, M. W., SHIPLEY, G. L., VANDESOMPELE, J. e WITTEWER, C. T. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. **Clin Chem**, v.55, n.4, Apr, p.611-22. 2009.

BUSTIN, S. A. e NOLAN, T. Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. **J Biomol Tech**, v.15, n.3, Sep, p.155-66. 2004.

CAFARCHIA, C., WEIGL, S., FIGUEREDO, L. A. e OTRANTO, D. Molecular identification and phylogenesis of dermatophytes isolated from rabbit farms and rabbit farm workers. **Vet Microbiol**, v.154, n.3-4, Jan 27, p.395-402. 2012.

CATANHO, M., DEGRAVE, W. e MIRANDA, A. B. Análise Comparativa de Genomas Procarióticos. **Biotecnologia Ciência & Desenvolvimento**, v.37, p.20-9. 2007.

CHANDARLAPATY, S. e ERREDE, B. Ash1, a daughter cell-specific protein, is required for pseudohyphal growth of *Saccharomyces cerevisiae*. **Mol Cell Biol**, v.18, n.5, May, p.2884-91. 1998.

CHERMETTE, R., FERREIRO, L. e GUILLOT, J. Dermatophytoses in animals. **Mycopathologia**, v.166, n.5-6, Nov-Dec, p.385-405. 2008.

- CLARK, M. S. Comparative genomics: the key to understanding the Human Genome Project. **Bioessays**, v.21, n.2, Feb, p.121-30. 1999.
- COENYE, T., GEVERS, D., VAN DE PEER, Y., VANDAMME, P. e SWINGS, J. Towards a prokaryotic genomic taxonomy. **FEMS Microbiol Rev**, v.29, n.2, Apr, p.147-67. 2005.
- COSTA, E. F., WANKE, B. e MARTINS, E. C. Micoses superficiais e cutâneas. Estudo comparativo entre duas populações: Rio de Janeiro (RJ) e Aracaju (SE). **An bras Dermatol**, v.66, n.3, p.119-22. 1991.
- COULOMBE, P. A. e OMARY, M. B. 'Hard' and 'soft' principles defining the structure, function and regulation of keratin intermediate filaments. **Curr Opin Cell Biol**, v.14, n.1, Feb, p.110-22. 2002.
- DARLING, A. C., MAU, B., BLATTNER, F. R. e PERNA, N. T. Mauve: multiple alignment of conserved genomic sequence with rearrangements. **Genome Res**, v.14, n.7, Jul, p.1394-403. 2004.
- DE BERSAQUES, J. e DOCKX, P. Proteolytic and leucyl-naphthylamidase activity in some dermatophytes. Preliminary results. **Arch Belg Dermatol Syphiligr**, v.29, n.2, Apr-Jun, p.135-40. 1973.
- DESCAMPS, F., BROUTA, F., MONOD, M., ZAUGG, C., BAAR, D., LOSSON, B. e MIGNON, B. Isolation of a *Microsporum canis* gene family encoding three subtilisin-like proteases expressed in vivo. **J Invest Dermatol**, v.119, n.4, Oct, p.830-5. 2002.
- DHEDA, K., HUGGETT, J. F., CHANG, J. S., KIM, L. U., BUSTIN, S. A., JOHNSON, M. A., ROOK, G. A. e ZUMLA, A. The implications of using an inappropriate reference gene for real-time reverse transcription PCR data normalization. **Anal Biochem**, v.344, n.1, Sep 1, p.141-3. 2005.
- DROUOT, S., MIGNON, B., FRATTI, M., ROOSJE, P. e MONOD, M. Pets as the main source of two zoonotic species of the *Trichophyton mentagrophytes* complex in Switzerland, *Arthroderma vanbreuseghemii* and *Arthroderma benhamiae*. **Vet Dermatol**, v.20, n.1, Feb, p.13-8. 2009.
- ELLIS, D., DAVIS, S., ALEXIOU, H., HANDKE, R. e BARTLEY, R. **Descriptions of medical fungi**. Australia: Adelaide. 2007. 198 p.
- FELSENSTEIN, J. Confidence-Limits on Phylogenies - an Approach Using the Bootstrap. **Evolution**, v.39, n.4, p.783-791. 1985.
- FERNANDEZ-TORRES, B., MAYAYO, E., BORONAT, J. e GUARRO, J. Subcutaneous infection by *Microsporum gypseum*. **Br J Dermatol**, v.146, n.2, Feb, p.311-3. 2002.

- FERREIRA-NOZAWA, M. S., NOZAWA, S. R., MARTINEZ-ROSSI, N. M. e ROSSI, A. The dermatophyte *Trichophyton rubrum* secretes an EDTA-sensitive alkaline phosphatase on high-phosphate medium. **Braz J Microbiol**, v.34, n.2, APR-JUN, p.161-164. 2003.
- FITZGERALD, J. R. e MUSSER, J. M. Evolutionary genomics of pathogenic bacteria. **Trends Microbiol**, v.9, n.11, Nov, p.547-53. 2001.
- FRASER, C. M., EISEN, J., FLEISCHMANN, R. D., KETCHUM, K. A. e PETERSON, S. Comparative genomics and understanding of microbial biology. **Emerg Infect Dis**, v.6, n.5, Sep-Oct, p.505-12. 2000.
- FUCHS, E. e WEBER, K. Intermediate filaments: structure, dynamics, function, and disease. **Annu Rev Biochem**, v.63, p.345-82. 1994.
- FUMEAUX, J., MOCK, M., NINET, B., JAN, I., BONTEMPS, O., LECHENNE, B., LEW, D., PANIZZON, R. G., JOUSSON, O. e MONOD, M. First report of *Arthroderma benhamiae* in Switzerland. **Dermatology**, v.208, n.3, p.244-50. 2004.
- FURTADO, M. S. S., THÁRA, L. T., MAROJA, M. F., JOSÉ, J. I. N. S. e CASTRILLÓN, A. L. Dermatofitoses na cidade de Manaus (AM). **An bras Dermatol**, v.62, n.4, p.195-6. 1987.
- GALPERIN, M. Y. e KOONIN, E. V. Functional genomics and enzyme evolution. Homologous and analogous enzymes encoded in microbial genomes. **Genetica**, v.106, n.1-2, p.159-70. 1999.
- GALPERIN, M. Y. e KOONIN, E. V. Who's your neighbor? New computational approaches for functional genomics. **Nat Biotechnol**, v.18, n.6, Jun, p.609-13. 2000.
- GHANNOUM, M. A. Potential role of phospholipases in virulence and fungal pathogenesis. **Clin Microbiol Rev**, v.13, n.1, Jan, p.122-43. 2000.
- GIDDEY, K., FAVRE, B., QUADRONI, M. e MONOD, M. Closely related dermatophyte species produce different patterns of secreted proteins. **FEMS Microbiol Lett**, v.267, n.1, Feb, p.95-101. 2007.
- GORDON, S. V., BROSCHE, R., EIGLMEIER, K., GARNIER, T., HEWINSON, R. G. e COLE, S. T. Royal Society of Tropical Medicine and Hygiene Meeting at Manson House, London, 18th January 2001. Pathogen genomes and human health. Mycobacterial genomics. **Trans R Soc Trop Med Hyg**, v.96, n.1, Jan-Feb, p.1-6. 2002.
- GRADISAR, H., FRIEDRICH, J., KRIZAJ, I. e JERALA, R. Similarities and specificities of fungal keratinolytic proteases: comparison of keratinases of *Paecilomyces marquandii* and *Doratomyces microsporus* to some known proteases. **Appl Environ Microbiol**, v.71, n.7, Jul, p.3420-6. 2005.
- GRASER, Y., KUIJPERS, A. F., PRESBER, W. e DE HOOG, G. S. Molecular taxonomy of *Trichophyton mentagrophytes* and *T. tonsurans*. **Med Mycol**, v.37, n.5, Oct, p.315-30. 1999.

- GRASER, Y., SCOTT, J. e SUMMERBELL, R. The new species concept in dermatophytes-a polyphasic approach. **Mycopathologia**, v.166, n.5-6, Nov-Dec, p.239-56. 2008.
- GRUMBT, M., MONOD, M. e STAIB, P. Genetic advances in dermatophytes. **FEMS Microbiol Lett**, v.320, n.2, Jul, p.79-86. 2011.
- GRUMBT, M., MONOD, M., YAMADA, T., HERTWECK, C., KUNERT, J. e STAIB, P. Keratin degradation by dermatophytes relies on cysteine dioxygenase and a sulfite efflux pump. **J Invest Dermatol**, v.133, n.6, Jun, p.1550-5. 2013.
- GRUNEWALD, S., PAASCH, U., GRASER, Y., GLANDER, H. J., SIMON, J. C. e NENOFF, P. [Scarring tinea profunda in the pubic area due to *Trichophyton verrucosum*]. **Hautarzt**, v.57, n.9, Sep, p.811-3. 2006.
- GRUNWALD, M. H., AMICHAÏ, B. e SHEMER, A. Fingertip contamination after a brief touch of tinea capitis lesions caused by *Microsporum canis*. **Br J Dermatol**, v.172, n.1, Jan, p.291-2. 2015.
- HUGGETT, J., DHEDA, K., BUSTIN, S. e ZUMLA, A. Real-time RT-PCR normalisation; strategies and considerations. **Genes Immun**, v.6, n.4, Jun, p.279-84. 2005.
- HUOVINEN, S., TUNNELA, E., HUOVINEN, P., KUIJPERS, A. F. e SUHONEN, R. Human onychomycosis caused by *Trichophyton equinum* transmitted from a racehorse. **Br J Dermatol**, v.138, n.6, Jun, p.1082-4. 1998.
- HUYNEN, M. A., GABALDON, T. e SNEL, B. Variation and evolution of biomolecular systems: searching for functional relevance. **FEBS Lett**, v.579, n.8, Mar 21, p.1839-45. 2005.
- ILLUMINA. Illumina Sequencing Technology: Illumina Inc. 2010.
- ILLUMINA. Nextera XT DNA - Sample Preparation Guide: Illumina Inc. 2012.
- JACOB, T. R., PERES, N. T., PERSINOTI, G. F., SILVA, L. G., MAZUCATO, M., ROSSI, A. e MARTINEZ-ROSSI, N. M. *rpb2* is a reliable reference gene for quantitative gene expression analysis in the dermatophyte *Trichophyton rubrum*. **Med Mycol**, v.50, n.4, May, p.368-77. 2012.
- JOUSSEON, O., LECHENNE, B., BONTEMS, O., CAPOCCIA, S., MIGNON, B., BARBLAN, J., QUADRONI, M. e MONOD, M. Multiplication of an ancestral gene encoding secreted fungalysin preceded species differentiation in the dermatophytes *Trichophyton* and *Microsporum*. **Microbiology**, v.150, n.Pt 2, Feb, p.301-10. 2004a.
- JOUSSEON, O., LECHENNE, B., BONTEMS, O., MIGNON, B., REICHARD, U., BARBLAN, J., QUADRONI, M. e MONOD, M. Secreted subtilisin gene family in *Trichophyton rubrum*. **Gene**, v.339, Sep 15, p.79-88. 2004b.

- KASPEROVA, A., KUNERT, J., HORYNOVA, M., WEIGL, E., SEBELA, M., LENOBEL, R. e RASKA, M. Isolation of recombinant cysteine dioxygenase protein from *Trichophyton mentagrophytes*. **Mycoses**, v.54, n.5, Sep, p.e456-62. 2011.
- KAUR, R., KASHYAP, B. e BHALLA, P. Onychomycosis--epidemiology, diagnosis and management. **Indian J Med Microbiol**, v.26, n.2, Apr-Jun, p.108-16. 2008.
- KELLIS, M., PATTERSON, N., BIRREN, B., BERGER, B. e LANDER, E. S. Methods in comparative genomics: genome correspondence, gene identification and regulatory motif discovery. **J Comput Biol**, v.11, n.2-3, p.319-55. 2004.
- KONDRASHOV, A. S. Comparative genomics and evolutionary biology. **Curr Opin Genet Dev**, v.9, n.6, Dec, p.624-9. 1999.
- KOONIN, E. V., ARAVIND, L. e KONDRASHOV, A. S. The impact of comparative genomics on our understanding of evolution. **Cell**, v.101, n.6, Jun 9, p.573-6. 2000.
- LARONE, D. H. Culture and Identification of Dermatophytes. **Clinical Microbiology Newsletter**, v.18, n.5, p.33-8. 1996.
- LECHENNE, B., REICHARD, U., ZAUGG, C., FRATTI, M., KUNERT, J., BOULAT, O. e MONOD, M. Sulphite efflux pumps in *Aspergillus fumigatus* and dermatophytes. **Microbiology**, v.153, n.Pt 3, Mar, p.905-13. 2007.
- LENG, W., LIU, T., WANG, J., LI, R. e JIN, Q. Expression dynamics of secreted protease genes in *Trichophyton rubrum* induced by key host's proteinaceous components. **Med Mycol**, v.47, n.7, Nov, p.759-65. 2009.
- LIVAK, K. J. e SCHMITTGEN, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. **Methods**, v.25, n.4, Dec, p.402-8. 2001.
- MACÊDO, J. A. A Molecular Biology Conceptual Model for Information Integration. **Revista Tecnologia da Informação**, v.3, n.2, p.41-8. 2003.
- MACHADO, A. P., HIRATA, S. H., OGAWA, M. M., TOMIMORI-YAMASHITA, J. e FISCHMAN, O. Dermatophytosis on the eyelid caused by *Microsporum gypseum*. **Mycoses**, v.48, n.1, Jan, p.73-5. 2005.
- MARANHÃO, F. C. A., PAIÃO, F. G., FACHIN, A. L. e MARTINEZ-ROSSI, N. M. Membrane transporter proteins are involved in *Trichophyton rubrum* pathogenesis. **J Med Microbiol**, v.58, n.Pt 2, Feb, p.163-8. 2009.
- MARANHÃO, F. C. A., PAIÃO, F. G. e MARTINEZ-ROSSI, N. M. Isolation of transcripts over-expressed in human pathogen *Trichophyton rubrum* during growth in keratin. **Microb Pathog**, v.43, May 21, p.166-72. 2007.

- MARTINEZ-ROSSI, N. M., FERREIRA-NOZAWA, M. S., GRAMINHA, M. A. S., NOZAWA, S. R., FACHIN, A. L., CERVELATTI, E. P., PRADE, R. A. e ROSSI, A. Molecular aspects of dermatophyte-host interactions. In: R. K. S. Kushwaha (Ed.). **Fungi in human and animal health**. Jodhpur, India: Scientific Publishers, v.9, 2004. Molecular aspects of dermatophyte-host interactions, p.143-165
- MARTINEZ, D. A., OLIVER, B. G., GRASER, Y., GOLDBERG, J. M., LI, W., MARTINEZ-ROSSI, N. M., MONOD, M., SHELEST, E., BARTON, R. C., BIRCH, E., BRAKHAGE, A. A., CHEN, Z., GURR, S. J., HEIMAN, D., HEITMAN, J., KOSTI, I., ROSSI, A., SAIF, S., SAMALOVA, M., SAUNDERS, C. W., SHEA, T., SUMMERBELL, R. C., XU, J., YOUNG, S., ZENG, Q., BIRREN, B. W., CUOMO, C. A. e WHITE, T. C. Comparative genome analysis of *Trichophyton rubrum* and related dermatophytes reveals candidate genes involved in infection. **MBio**, v.3, n.5, p.e00259-12. 2012.
- MARTINS, M. P. **Expressão comparativa de genes em dermatófitos durante o processo de interação com moléculas do hospedeiro e em resposta a agentes antifúngicos**. (Tese - Doutorado). Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto - SP, 2015. 159 p.
- MATSUZAKI, Y., OTA, K., SATO, K., NARA, S., YAGUSHI, T., NAKANO, H. e SAWAMURA, D. Deep pseudocystic dermatophytosis caused by *Trichophyton rubrum* in a patient with myasthenia gravis. **Acta Derm Venereol**, v.93, n.3, May, p.358-9. 2012.
- MCGINNIS, M. R. **Laboratory handbook of Medical Mycology**: Academic Press. 1980. 678 p.
- MCLEAN, W. H. e MOORE, C. B. Keratin disorders: from gene to therapy. **Hum Mol Genet**, v.20, n.R2, Oct 15, p.R189-97. 2011.
- MONOD, M. Secreted proteases from dermatophytes. **Mycopathologia**, v.166, n.5-6, Nov-Dec, p.285-94. 2008.
- MONOD, M., LECHENNE, B., JOUSSON, O., GRAND, D., ZAUGG, C., STOCKLIN, R. e GROUZMANN, E. Aminopeptidases and dipeptidyl-peptidases secreted by the dermatophyte *Trichophyton rubrum*. **Microbiology**, v.151, n.Pt 1, Jan, p.145-55. 2005.
- MORAES, M. A. P. Dermatófitos no Estado do Amazonas - Brasil. **Acta Amazônica**, v.3, n.1, p.65-9. 1973.
- MORENO-HAGELSIEB, G., TREVINO, V., PEREZ-RUEDA, E., SMITH, T. F. e COLLADO-VIDES, J. Transcription unit conservation in the three domains of life: a perspective from *Escherichia coli*. **Trends Genet**, v.17, n.4, Apr, p.175-7. 2001.
- MUKHERJEE, P. K., LEIDICH, S. D., ISHAM, N., LEITNER, I., RYDER, N. S. e GHANNOUM, M. A. Clinical *Trichophyton rubrum* strain exhibiting primary resistance to terbinafine. **Antimicrob Agents Chemother**, v.47, n.1, Jan, p.82-6. 2003.

- MUSHEGIAN, A. R. **Foundations of Comparative Genomics**: Academic Press. 2007. 280 p.
- NENOFF, P., KRUGER, C., GINTER-HANSELMAYER, G. e TIETZ, H. J. Mycology - an update. Part 1: Dermatomycoses: causative agents, epidemiology and pathogenesis. **J Dtsch Dermatol Ges**, v.12, n.3, Mar, p.188-209; quiz 210, 188-211; quiz 212. 2014.
- NOLAN, T., HANDS, R. E. e BUSTIN, S. A. Quantification of mRNA using real-time RT-PCR. **Nat Protoc**, v.1, n.3, p.1559-82. 2006.
- OLIVEIRA, K. C., CARVALHO, M. L., MARACAJA-COUTINHO, V., KITAJIMA, J. P. e VERJOVSKI-ALMEIDA, S. Non-coding RNAs in schistosomes: an unexplored world. **An Acad Bras Cienc**, v.83, n.2, Jun, p.673-94. 2011.
- PAN, X. e HEITMAN, J. Sok2 regulates yeast pseudohyphal differentiation via a transcription factor cascade that regulates cell-cell adhesion. **Mol Cell Biol**, v.20, n.22, Nov, p.8364-72. 2000.
- PARK, H. e BAKALINSKY, A. T. SSU1 mediates sulphite efflux in *Saccharomyces cerevisiae*. **Yeast**, v.16, n.10, Jul, p.881-8. 2000.
- PEREIRO MIGUENS, M., PEREIRO, M. e PEREIRO, M., JR. Review of dermatophytoses in Galicia from 1951 to 1987, and comparison with other areas of Spain. **Mycopathologia**, v.113, n.2, Feb, p.65-78. 1991.
- PERES, N. T. A., MARANHÃO, F. C. A., ROSSI, A. e MARTINEZ-ROSSI, N. M. Dermatofitos: interação patógeno-hospedeiro e resistência a antifúngicos. **An Bras Dermatol**, v.85, n.5, p.657-67. 2010a.
- PERES, N. T. A., SANCHES, P. R., FALCÃO, J. P., SILVEIRA, H. C. S., PAIÃO, F. G., MARANHÃO, F. C. A., GRAS, D. E., SEGATO, F., CAZZANIGA, R. A., MAZUCATO, M., CURSINO-SANTOS, J. R., AQUINO-FERREIRA, R., ROSSI, A. e MARTINEZ-ROSSI, N. M. Transcriptional profiling reveals the expression of novel genes in response to various stimuli in the human dermatophyte *Trichophyton rubrum*. **BMC Microbiol** v.10, p.39. 2010b.
- PERSINOTI, G. F., DE AGUIAR PERES, N. T., JACOB, T. R., ROSSI, A., VENCIO, R. Z. e MARTINEZ-ROSSI, N. M. RNA-sequencing analysis of *Trichophyton rubrum* transcriptome in response to sublethal doses of acriflavine. **BMC Genomics**, v.15 Suppl 7, p.S1. 2014.
- RAMBALDI, D. e CICCARELLI, F. D. FancyGene: dynamic visualization of gene structures and protein domain architectures on genomic loci. **Bioinformatics**, v.25, n.17, Sep 1, p.2281-2. 2009.
- RAO, M. B., TANKSALE, A. M., GHATGE, M. S. e DESHPANDE, V. V. Molecular and biotechnological aspects of microbial proteases. **Microbiol Mol Biol Rev**, v.62, n.3, Sep, p.597-635. 1998.

- RAWLINGS, N. D. e BARRETT, A. J. MEROPS: the peptidase database. **Nucleic Acids Res**, v.27, n.1, Jan 1, p.325-31. 1999.
- RICH, P., HARKLESS, L. B. e ATILLASOY, E. S. Dermatophyte test medium culture for evaluating toenail infections in patients with diabetes. **Diabetes Care**, v.26, n.5, May, p.1480-4. 2003.
- RODWELL, G. E., BAYLES, C. L., TOWERSEY, L. e ALY, R. The prevalence of dermatophyte infection in patients infected with human immunodeficiency virus. **Int J Dermatol**, v.47, n.4, Apr, p.339-43. 2008.
- ROMANO, C., MASSAI, L., GALLO, A. e FIMIANI, M. *Microsporum gypseum* infection in the Siena area in 2005-2006. **Mycoses**, v.52, n.1, Jan, p.67-71. 2009.
- SAMBROOK, J., FRITSCH, E. F. e MANIATIS, T. **Molecular cloning: a laboratory manual**. New York: Cold Spring Harbor Laboratory Press. 1989
- SEEBACHER, C., BOUCHARA, J. P. e MIGNON, B. Updates on the epidemiology of dermatophyte infections. **Mycopathologia**, v.166, n.5-6, Nov-Dec, p.335-52. 2008.
- SHIRAKI, Y., HIRUMA, M., MATSUBA, Y., KANO, R., MAKIMURA, K., IKEDA, S. e OGAWA, H. A case of tinea corporis caused by *Arthroderma benhamiae* (teleomorph of *Tinea mentagrophytes*) in a pet shop employee. **J Am Acad Dermatol**, v.55, n.1, Jul, p.153-4. 2006.
- SIDRIM, J. J. C. e ROCHA, M. F. G. **Micologia Médica à Luz de autores contemporâneos**: Guanabara Koogan. 2004. 396 p.
- SILVA, P. I., MARTINS, A. M., GOUVEA, E. G., PESSOA-FILHO, M. e FERREIRA, M. E. Development and validation of microsatellite markers for *Brachiararia ruziziensis* obtained by partial genome assembly of Illumina single-end reads. **BMC Genomics**, v.14, p.17. 2013.
- SILVA, V., THOMSON, P., MAIER, L. e ANTICEVIC, S. [Dermatophyte infection and colonization in dogs from South Santiago, Chile]. **Rev Iberoam Micol**, v.20, n.4, Dec, p.145-8. 2003.
- STAIB, P., ZAUGG, C., MIGNON, B., WEBER, J., GRUMBT, M., PRADERVAND, S., HARSHMAN, K. e MONOD, M. Differential gene expression in the pathogenic dermatophyte *Arthroderma benhamiae* in vitro versus during infection. **Microbiology**, v.156, n.Pt 3, Mar, p.884-95. 2010.
- SYMOENS, F., JOUSSON, O., PACKEU, A., FRATTI, M., STAIB, P., MIGNON, B. e MONOD, M. The dermatophyte species *Arthroderma benhamiae*: intraspecies variability and mating behaviour. **J Med Microbiol**, v.62, n.Pt 3, Mar, p.377-85. 2013.

- TAKATORI, K., TAKAHASHI, A., KAWAI, S., ICHIJO, S. e HASEGAWA, A. Isolation of *Trichophyton verrucosum* from lesional and non-lesional skin in calves. **J Vet Med Sci**, v.55, n.2, Apr, p.343-4. 1993.
- TAMURA, K., PETERSON, D., PETERSON, N., STECHER, G., NEI, M. e KUMAR, S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. **Mol Biol Evol**, v.28, n.10, Oct, p.2731-9. 2011.
- THOMPSON, J. D., HIGGINS, D. G. e GIBSON, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Res**, v.22, n.22, Nov 11, p.4673-80. 1994.
- TSUBOI, R., KO, I. J., TAKAMORI, K. e OGAWA, H. Isolation of a keratinolytic proteinase from *Trichophyton mentagrophytes* with enzymatic activity at acidic pH. **Infect Immun**, v.57, n.11, NOV, p.3479-3483. 1989.
- VERMOUT, S., BALDO, A., TABART, J., LOSSON, B. e MIGNON, B. Secreted dipeptidyl peptidases as potential virulence factors for *Microsporum canis*. **FEMS Immunol Med Microbiol**, v.54, n.3, Dec, p.299-308. 2008.
- WAGNER, A. Selection and gene duplication: a view from the genome. **Genome Biol**, v.3, n.5, p.reviews1012. 2002.
- WEI, L., LIU, Y., DUBCHAK, I., SHON, J. e PARK, J. Comparative genomics approaches to study organism similarities and differences. **J Biomed Inform**, v.35, n.2, Apr, p.142-50. 2002.
- WEITZMAN, I. e SUMMERBELL, R. C. The Dermatophytes. **Clin Microbiol Rev**, v.8, n.2, APR, p.240-259. 1995.
- WHITE, T. C., OLIVER, B. G., GRASER, Y. e HENN, M. R. Generating and testing molecular hypotheses in the dermatophytes. **Eukaryot Cell**, v.7, n.8, Aug, p.1238-45. 2008.
- YANG, J., CHEN, L., WANG, L., ZHANG, W., LIU, T. e JIN, Q. TrED: the *Trichophyton rubrum* Expression Database. **BMC Genomics**, v.8, p.250. 2007.
- ZAUGG, C., JOUSSON, O., LECHENNE, B., STAIB, P. e MONOD, M. *Trichophyton rubrum* secreted and membrane-associated carboxypeptidases. **Int J Med Microbiol**, v.298, n.7-8, Oct, p.669-82. 2008.
- ZAUGG, C., MONOD, M., WEBER, J., HARSHMAN, K., PRADERVAND, S., THOMAS, J., BUENO, M., GIDDEY, K. e STAIB, P. Gene expression profiling in the human pathogenic dermatophyte *Trichophyton rubrum* during growth on proteins. **Eukaryot Cell**, v.8, n.2, Feb, p.241-50. 2009.
- ZERBINO, D. R. e BIRNEY, E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. **Genome Res**, v.18, n.5, May, p.821-9. 2008.

Anexo de Publicações

Far better it is to dare mighty things, to win glorious triumphs, even though checkered by failure, than to take rank with those poor spirits who neither enjoy much nor suffer much, because they live in the gray twilight that knows neither victory nor defeat.

Theodore Roosevelt

Anexo de Publicações

Biocatalysis and Biotransformation, 2014; 32(1): 74–83informa
healthcare

ORIGINAL ARTICLE

Screening of filamentous fungi for lipase production: *Hypocrea pseudokoningii* a new producer with a high biotechnological potentialMARITA GIMENEZ PEREIRA¹, ANA CLAUDIA VICI²,
FERNANDA DELL ANTONIO FACCHINI², ALAN PADUA TRISTÃO¹,
JENY RACHID CURSINO-SANTOS³, PABLO RODRIGO SANCHES³,
JOÃO ATÍLIO JORGE¹ & MARIA DE LOURDES TEIXEIRA DE MORAES POLIZELI¹¹Departamento de Biologia, Faculdade de Filosofia Ciências e Letras, Universidade de São Paulo, SP, Brazil,²Departamento de Bioquímica e Imunologia, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, SP, Brazil, and ³Departamento de Genética, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, SP, Brazil**Abstract**

Filamentous fungi isolated from soil samples were screened for extracellular lipase production. The best producer was *Hypocrea pseudokoningii* identified by taxonomical criteria, and by rDNA sequencing of the variable internal transcribed spacers (ITS I and II) and the intervening 5.8S gene. The fungus was grown in a complex medium supplemented with 1% Tween 80 and 0.2% yeast extract, for 4 days. The optimum pH for extracellular and intracellular lipases was 7.0 and 8.0, respectively. Both enzymes exhibited maximum activity at 40°C. Extracellular and intracellular lipase activities were highly stable in the pH range 3.0–8.0 at room temperature. The intracellular lipase was thermostable up to 60°C, for 15 min and the extracellular, for 107 min, at the same temperature. The intracellular lipase was stimulated by silver ions. Extracellular lipase was stable in organic solvents, such as DMSO, alcohols, acetone, and acetonitrile, for 24 hours. Lipase activity increased around 80% when detergents were added to the enzymatic assay, such as Tween 80, Triton X-100, and SDS.

Keywords: Screening, *Hypocrea pseudokoningii*, lipase, molecular characterization, ITS**Abbreviations:** DMSO, Dimethyl sulfoxide; SDS, sodium dodecyl sulfate**Introduction**

Interest in lipases is increasing, mainly due to the wide variety of potential applications of these enzymes (Hasan et al. 2009). For instance, lipases play an important role in the processing of polyunsaturated fatty acids (PUFA), in the synthesis of the food colorant astaxanthin and flavor molecules, as well as in the interesterification of cheaper glycerides into more valuable forms (Undurraga et al. 2001). Lipid esters, used in cosmetic formulations, and monoglycerides used in pharmaceutical applications and as emulsifiers in food products, can be obtained by enzymatic synthesis using lipases (Sharma et al.

2001). Other lipase applications include the production of dicarboxylic acids, used as prepolymers (Gross et al. 2001); in the detergent industries in order to remove fats and oils; in the biodiesel synthesis (Parawira 2009) textile, food, pulp and paper, and pharmaceutical industries; in cosmetics; tea processing; medical applications, as biosensors and in waste treatment (Singh and Mukhopadhyay 2012).

Lipases are also known as triacylglycerol acyl hydrolases (EC 3.1.1.3). They constitute a group of enzymes showing interfacial activation, due to their ability to hydrolyze hydrophobic substrates in hydrophilic media, at the lipid–water interface (Reis et al.

Correspondence: Dr Maria de Lourdes Teixeira de Moraes Polizeli, Departamento de Biologia, Faculdade de Filosofia Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Av. Bandeirantes, 3900, 14040-901 Ribeirão Preto, SP, Brazil. Tel. +55-16-3602-4680. Fax: +55-16-3602-4886. E-mail: polizeli@fclrp.usp.br

(Received 4 September 2013; revised 11 October 2013; accepted 5 December 2013)

ISSN 1024-2422 print/ISSN 1029-2446 online © 2014 Informa UK, Ltd.
DOI: 10.3109/10242422.2013.873417RIGHTS LINK
Copyright Clearance Center

2009). Researchers also believe that the presence of a loop covering the active site (lid) is another characteristic that distinguishes lipases from esterases (Fuciños et al. 2005). Due to their ability to utilize a broad spectrum of substrates with different regio-, chemo-, and enantioselectivities, lipases can also perform esterification reactions in anhydrous organic media (Fernandez-Lafuente 2010), an important feature in the development of efficient methods for the industrial synthesis of pure enantiomers including chiral drugs (Wu et al. 2007).

Although they are found in various organisms, only microbial lipases are commercially attractive (Sharma et al. 2001). As expected, enzymes from diverse microbial sources show different catalytic properties, stimulating the constant search for new lipase-producing microorganisms. One of the required properties for industrial use is reasonable thermostability. Thermostability is essential to enable lipases to be used with high melting point substrates (lipids) (Ahmed et al. 2010). The aim of the present study was to identify filamentous fungi which are lipase producers and select the best one. *Hypocrea pseudokoningii* was chosen for further study. Intracellular and extracellular lipase biochemical characterization was performed in order to evaluate their application in industrial processes.

Materials and methods

Microorganisms and maintenance medium

Microorganisms were previously collected and isolated from soil samples from different regions of São Paulo State – Brazil. All fungal strains were taxonomically identified and deposited at the Federal University of Pernambuco, PE, Brazil. Fungal identification was primarily achieved from morphological characteristics, followed by molecular analysis as described. Microorganisms were maintained in solid oatmeal baby food medium (4.0% oatmeal baby food and 2.0% agar) at 4°C.

Molecular characterization of the isolated fungi

Analysis based on rDNA sequences coding for internal transcribed spacers (ITS I and II) and the intervening 5.8S gene was used for the molecular identification of the isolated fungal species. Mycelia, developed in SR (Segato-Rizzatti) medium cultures (Rizzatti et al. 2001), were incubated at 30°C, for 4 days and used for genomic DNA extraction. Forty milligram of fine powdered mycelium obtained using acetone-drying protocol were suspended in 1 mL of the extraction buffer (1% SDS; 50 mM EDTA, 200 µg/mL RNase) and subjected to a conventional phenol extraction method (Punekar et al. 2003).

Screening of filamentous fungi for lipase production 75

The internal transcribed spacer sequence (ITSI-5.8S-ITSII) was amplified using universal primers, ITS1 (5'-TCCGTAGGTGAACCTGCGG-3') and ITS4 (5'-TCCTCCGCTTATTGATATGC-3') according to White et al. (1990). The PCR reaction mixture (50 µL) was composed by 2 mM MgCl₂, 2 µM of each primer, 100 µM dNTP mix, 1U of Taq DNA polymerase and 100 ng of DNA template. Thermal cycling conditions were initial denaturation (10 min at 96°C), followed by 30 cycles of denaturation (95°C for 1 min), annealing (60°C for 1 min), and primer extension (72°C for 1 min), followed by a final extension step for 10 min at 72°C. Amplification products were electrophoretically resolved on a 1.8% (w/v) agarose gel.

Both strands of the PCR products were directly sequenced using ITS 1 and ITS 4 PCR primers as described above and two additional internal primers ITS2 (5' GCTGCGTTCATCGATGC 3') and ITS3 (5' GCATCGATGAAGAACGCAGC 3') (White et al. 1990). DNA sequencing was performed using an ABI PRISM 377 sequencer (Applied Biosystem, USA) following the manufacturer's protocol from the DYEnamic ET terminator cycle sequencing kits (GE Healthcare, UK). The quality of the sequences was examined using the Phred/Phrap/Consed package (University of Washington, Seattle, Wash.).

Species identification was made by searching databases using the BLASTn (Altschul et al. 1990) and ClustalW2 (Thompson et al. 1990) algorithms, with default settings, for sequence comparison with reference sequences from NCBI GenBank. Species identification was determined from the best-scoring reference sequence of the similarity output that had 97% or more identity with the query sequence. The sequences determined in this study were deposited in GenBank. The accession numbers assigned are listed in Table I.

Lipase production

For examining the influence of medium-composition on lipase production, *H. pseudokoningii* was cultivated in seven different media supplemented with 1% olive oil: Czapek (Wiseman 1975), Emerson (1941), Adams (1990), Khanna et al. (1995), M-5 (Peralta et al. 1990), SR (Rizzatti et al. 2001), and Vogel et al. (1964). The best liquid medium was used in further experiments studying different incubation periods.

Enzyme production

In order to obtain a final concentration of 10⁵ spores/mL, a solution of 2.0 mL of a conidial suspension

Table I. Screening of microorganisms for lipase producing.

Fungus	GenBank accession number	Activity (U/mL)	
		Extracellular	Intracellular
<i>Aspergillus niger</i>	FJ810501	0.28 ± 0.05	0.97 ± 0.04
<i>Aspergillus fumigatus</i>	FJ810502	0 ± 0	0.51 ± 0.05
<i>Aspergillus ochraceus</i>	FJ810503	0 ± 0	1.42 ± 0.09
<i>Aspergillus phoenicis</i>	FJ810504	1.28 ± 0.09	1.39 ± 0.07
<i>Beauveria bassiana</i>	HQ665468	3.37 ± 0.08	0.34 ± 0.03
<i>Paeclomyces variotii</i>	FJ895878	0 ± 0	0.40 ± 0.05
<i>Rhizopus microsporus</i>	FJ810505	0.72 ± 0.09	0.77 ± 0.06
<i>Hypocrea pseudokoningii</i>	FJ810506	5.28 ± 1.03	0.65 ± 0.05

Enzymatic assay was performed at 40°C and pH 5.0.

was inoculated into 125 mL Erlenmeyer flasks containing 25 mL of Adams medium (0.2% yeast extract, 0.1% KH_2PO_4 , 0.05% $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 1% sunflower oil and distilled water to reach a final volume of 100 mL). The medium was also supplemented with 1% different oils and surfactants (Tween 80, olive, macauba pulp (*Acrocomia aculeata*), fried oil, neem, pequi, sesame, *Jatropha curcas* seed cake, canola, palm, macauba almond, macauba seed, sunflower, castor, corn, and soybean oils). In order to find the best nitrogen source, peptone was replaced with yeast extract, soybean meal, wheat bran, ammonium nitrate, ammonium sulfate, ammonium chloride urea, ammonium phosphate monobasic, ammonium phosphate dibasic, tryptone, or ammonium acetate, at the same concentration (0.2%). The cultures were incubated under orbital stirring (100 rpm), for 96 h, at 35°C.

After this period the liquid culture medium was harvested by vacuum filtration on n°1 Whatman filter paper, and the crude filtrate used as a source of extracellular lipase. The mycelial pads were ground in a porcelain mortar, at 4°C (with twice their weight of glass beads 425–600 microns, Sigma™), extracted with 10 mL of McIlvaine buffer [0.1M phosphate-citrate buffer, pH 5.5 (McIlvaine 1921)] and centrifuged at 21,000 g for 10 min, 4°C. The supernatant was used as a source of crude intracellular lipase activity.

Measurement of lipase activity and protein

Extracellular and intracellular lipase activities were determined discontinuously using *p*-nitrophenyl palmitate as substrate (Pencrac'h and Baratti 1996). Standard assay conditions were 25 µL of enzymatic sample, 25 µL of distilled water, 450 µL of McIlvaine buffer 0.1M, pH 5.0 containing 0.77 mM *p*-nitrophenyl palmitate, 10% isopropanol, 0.5 mg/mL gum Arabic, and 0.25% Triton X-100. The mixture was incubated at 40°C for different periods, the

assay was stopped with 0.5 mL of saturated sodium tetraborate solution and the liberated *p*-nitrophenolate ion was estimated at 410 nm. One enzyme unit was defined as the amount of enzyme that produced 1 µmol of *p*-nitrophenol per minute under the assay conditions. Proteins were measured using the Bradford method (Bradford 1976), with bovine serum albumin as standard.

In order to determine constitutive lipase production, the activity was also determined in the spores of *H. pseudokoningii*. Hence, the filamentous fungus was cultivated on slants of solid oatmeal baby food medium, for 7 days. After that, the spores were scraped of the culture medium with a platinum handle and resuspended in 10 mL of sterile distilled water. After vigorous stirring, the solution was filtered in gauze to obtain a final concentration of 10^7 spores/mL.

Reproducibility of experiments

All experiments were performed in triplicate and the Standard Deviation was calculated.

Results

Screening

The isolated fungi were screened to identify the best lipase producers (Table I). Molecular analyses of PCR amplicons for the region sequences of ITS-1-5.8S-ITS2 was started in parallel with the morphological analyses in order to provide species identification of eight isolated fungal strains. Their ITS-1-5.8S-ITS2 sequences and the respective reference sequences from NCBI GenBank obtained by BLASTn searching (Altschul et al. 1990) were subjected to a multiple alignment using CLUSTAL W2 (Thompson et al. 1994) (data not shown). The topology of the resulting dendrogram is shown in Figure 1.

Molecular analyzes of the ITS sequences from *A. ochraceus* (FJ810503) and *A. phoenicis* (FJ810504) were restricted due to the lack of information available in databases. There was no reliable reference sequence that could be used to compare FJ810504, isolating it from the other species in the dendrogram. The higher similarity of the FJ810503 sequence to ITS1-5.8S-ITS2 of *A. fumigatus* than to other species considered in this multiple alignment showed that *A. ochraceus* strains isolated here belonged to the same cluster of *A. fumigatus*. However, it does not mean that FJ810503 is the same species of *A. fumigatus* since they are morphologically distinct. In addition, the BLASTn search showed that the strain of *Beauveria bassiana* isolated

Screening of filamentous fungi for lipase production 77

Figure 1. Dendrogram after CLUSTAL W2 multiple alignment of the ITS-1-5.8S-ITS2 sequence regions of the eight fungal strains studied in this work in comparison to their respective reference sequences (*) obtained from NCBI GenBank by BLASTn searching. Every sequence access numbers are indicated according to Table I.

here (HQ665468) has high similarity not only to other strains of *Beauveria bassiana* from the database, but also with strains of *B. brongniartii*. This is confirmed by the small variation obtained between the ITS1-5.8S-ITS2 regions of these two species.

Among these isolated fungi, *H. pseudokoningii* was the best enzyme producer (Table I). Extracellular activity was 8-fold higher than the intracellular activity and almost 2-fold higher than the lipase activity from *Beauveria bassiana*, the second best producer.

Improvement in lipase production

Effect of media composition and time-course on lipase production. Seven media with different compositions, all supplemented with 1% olive oil, initial pH 6.0, were used to evaluate the growth of *H. pseudokoningii* and lipase production. The incubation was carried out on liquid stirred cultures for 4 days, at 100 rpm and 35°C. Although *H. pseudokoningii* grew in all culture media, no enzymatic activity was observed in M-5 medium. Maximum extracellular lipase activity was observed in Adams medium, but high levels of intracellular activity were detected in SR medium (Figure 2). As extracellular activity was preferred as it minimizes the extraction time compared to intracellular activities (from mycelia extracts), Adams medium was chosen to continue the experimental optimization.

A time-course of lipase production was carried out using Adams medium supplemented with 1% oil olive and the maximum extracellular lipase activity was obtained on the fourth day of the fermentation process (Figure 3). The intracellular levels of lipase did not show significant variation from the 4th to the 5th days of incubation and represented only 12% of the maximum level obtained for the extracellular form. Interestingly, high lipase levels

(6 U/mL) were also detected with a solution of 10^7 spores/mL of *H. pseudokoningii*.

Influence of carbon and nitrogen sources on lipase production. Different vegetable oils and Tween 80 were tested for lipase production by *H. pseudokoningii*.

Figure 2. Effect of medium composition on (A) lipase activity and (B) protein production. The fungus was grown in different media for 4 days. Black bars correspond to the extracellular lipase activity and gray bars to the intracellular lipase activity.

78 M. G. Pereira et al.

Figure 3. Time course of lipolytic activity of *H. pseudokoningii*. Gray bars correspond to the intracellular lipase activity and black bars to the extracellular lipase activity.

All the sources tested induced lipase production, but the best source for extracellular lipase production was Tween 80 (Figure 4A) followed by olive and sunflower oils. On the other hand, for intracellular lipase production the best carbon source was palm oil (Figure 4A) followed by soybean, castor, and fried oil. Extracellular lipase production in the medium supplemented with oil was much lower than in the medium containing Tween 80. The best growth of *H. pseudokoningii* obtained was when pie jatropha and seed of macauba oil were used as carbon sources (Figure 4B).

Different nitrogen sources (0.2%) such as soybean meal, wheat bran, yeast extract, peptone, tryptone, NH_4NO_3 , $(\text{NH}_4)_2\text{SO}_4$, NH_4Cl , urea, KH_2PO_4 , $(\text{NH}_4)_2\text{HPO}_4$, and $\text{CH}_3\text{COONH}_4$, were tested as lipase inducers. The inorganic nitrogen sources did not improve lipase production, but a slight increase in lipase activity was observed when yeast extract was used (Figure 5).

Enzymatic characterization

Effect of pH and temperature on lipase activity. All biochemical characterizations were carried out with lipase produced in Adams medium supplemented with 1% olive oil. The optimum pH and pH stability were determined by assaying the lipase produced in McIlvaine's buffer over the pH range of 4.0–8.0. Other buffers as Tris–HCl buffer at pH 8.5, glycine buffer at pH 9.0, and CAPS buffer at pH 10.0, were tested in order to achieve a larger pH range. The optimum pH for the extracellular lipase was observed in the range pH 4.0–5.0, but a peak was also detected at pH 7.0. The maximum intracellular activity also showed two optima, at pH 5.5 and 8.0 (Figure 6A).

Figure 4. Effect of different oils and surfactants (1%) (A) on lipase production (B) Growth (mg/mL). Symbols: gray bars indicate the intracellular activity and black bars indicate the extracellular activity. 1. Canola oil, 2. Sunflower oil, 3. Soybean oil, 4. Tween 80, 5. Olive oil, 6. Macauba pulp oil, 7. Castor oil, 8. Corn oil, 9. Fried oil, 10. Neem oil, 11. Pequi oil, 12. Sesame oil, 13. *Jatropha curcas* seed cake, 14. Palm oil, 15. macauba almond oil, 16. Macauba seed oil.

The pH stability was investigated by incubating the enzyme at the same pH range for 1h, at room temperature. The remaining activities were measured under standard conditions. Extracellular and intracellular lipases showed identical profiles of pH stability in the range pH 3.0–10.0 (Figure 6B). To determine the temperature effect on lipase activity, the enzyme assays were carried out over a range of temperature varying from 30°C to 65°C. The extracellular and intracellular lipases had maximal activity at 40°C (Figure 7A).

Thermostability was investigated by incubating the enzyme at temperatures from 30°C to 60°C for 2h. Immediately afterwards the enzyme was immersed in an ice bath and then the activity was tested under standard conditions. The half-life of the intracellular lipase was 100 min, at 50°C, while the extracellular lipase did not reach its half-life period

Figure 5. Effect of different nitrogen sources on the extracellular lipase activity. 1. yeast extract + soybean meal, 2. yeast extract + wheat bran, 3. yeast extract, 4. NH_4NO_3 , 5. NH_4Cl , 6. $(\text{NH}_4)_2\text{CO}_3$, 7. wheat bran, 8. soybean meal, 9. KH_2PO_4 , 10. $(\text{NH}_4)_2\text{HPO}_4$, 11. $(\text{NH}_4)_2\text{SO}_4$, 12. $\text{CH}_3\text{COONH}_4$, 13. peptone, 14. tryptone.

Figure 6. Effect of pH on (A) lipase activity and (B) lipase stability. Symbols (○) indicate the extracellular activity and (●) indicate the intracellular activity.

Screening of filamentous fungi for lipase production 79

after 2 hours at this temperature. The half-life of the extracellular and intracellular lipases was 107 min and 15 min, respectively, at 60°C (Figure 7B, C).

Influence of metal ions and EDTA. In order to examine the influence of metal ions and EDTA on lipase activity the enzyme samples were incubated with

Figure 7. Effect of temperature on lipase activity (A); extracellular lipase stability after 2 hours (B) and intracellular lipase stability after 2 hours (C). Symbols: Figure A (○) 30°C; (●) 40°C; (◐) 50°C and (◑) 60°C.

80 M. G. Pereira et al.

different metal ions (Fe^{+2} , Na^+ , K^+ , NH_4^+ , Mg^{+2} , Ca^{+2} , Ba^{+2} , Mn^{+2} , Ag^+ , Co^{+2} , Cu^{+2} , Zn^{+2} , Hg^{+2} , and Al^{+3} used as chloride salts) or EDTA, at 1mM final concentration for 30 minutes. The reactions were measured under standard assay conditions. Extracellular lipase was inhibited by the following ions: Na^+ (48.1%), Ba^{+2} (56.4%), Mn^{+2} (54.5%), Co^{+2} (52.3%), Cu^{+2} (53.8%), Zn^{+2} (52.5%), Hg^{+2} (56.2%), and Al^{+3} (56.4%). The enzyme was not activated by any ions. On the other hand, the intracellular lipase was inhibited by Na^+ (30.8%) and Ba^{+2} (33.9%), but was activated (22.2%) by silver ions (Table II).

Influence of organic solvents and detergent on lipase activity. To analyze the stability in the presence of organic solvents, the lipase was incubated in the presence of dimethyl sulfoxide, methanol, acetonitrile, ethanol, acetone, isopropanol, glycerol, or butanol at a final concentration of 50%. After 24-h incubation at room temperature, the residual activity was determined as previously described (Figure 8). The results showed that extracellular lipase was activated by dimethyl sulfoxide, acetone and ethanol by 90%, methanol activated the enzyme by about 150%, and the extracellular lipase was only inhibited by butanol.

In order to investigate the effect of different detergents on lipase activity, polyoxyethylene sorbitan monooleate (Tween 80®), polyoxyethylene glycol octylphenol ether (Triton X-100®) and sodium dodecyl sulfate (SDS™) were added to the reaction mixture at a final concentration of 5 and 10 mM. Lipase activity was measured as previously described. All detergents enhanced the extracellular lipase

Figure 8. Effect of organic solvents on lipase activity. 1. DMSO (Dimethyl sulfoxide), 2. Acetone, 3. Glycerin, 4. Methanol, 5. Butanol, 6. Acetonitrile, 7. Ethanol, 8. isopropanol, 9. Control. Gray bars correspond to the extracellular activity and black bars to the intracellular activity.

activity, but only sodium dodecyl sulfate enhanced (60%) the intracellular lipase (Figure 9).

Discussion

Eight different fungi with lipolytic properties were selected, and later identified by ITS. Although ITS1-5.8S-ITS2 sequences have been widely used in phylogenetic studies and species identification from environmental and clinical samples (White et al. 1990; Ciardo et al. 2010), some studies have combined ITS sequences with other nuclear gene sequences or with morphological analyses to solve taxonomic and phylogenetic issues (Ghikas

Table II. Influence of metal ions and EDTA on lipase activity (%).

Supplement [1mM]	Extracellular	Intracellular
Control*	100.0 ± 0.4	100.0 ± 1.0
EDTA	103.5 ± 1.8	91.9 ± 0.6
Na^+	51.9 ± 2.7	69.2 ± 3.0
K^+	103.5 ± 5.1	86.6 ± 1.6
NH_4^+	90.1 ± 6.6	89.2 ± 2.2
Mg^{+2}	95.9 ± 6.2	89.7 ± 1.2
Ca^{+2}	95.6 ± 3.6	81.7 ± 1.6
Ba^{+2}	43.6 ± 3.6	66.1 ± 0.2
Mn^{+2}	45.5 ± 1.1	76.7 ± 1.6
Co^{+2}	47.7 ± 1.9	75.5 ± 1.2
Cu^{+2}	46.2 ± 3.5	74.0 ± 2.3
Zn^{+2}	47.5 ± 2.7	73.8 ± 3.1
Hg^{+2}	43.8 ± 1.5	91.6 ± 2.8
Al^{+3}	43.6 ± 1.1	88.7 ± 3.8
Ag^+	83.1 ± 3.2	122.2 ± 0.2
Fe^{+2}	91.1 ± 3.2	96.5 ± 1.3

*Control = without addition.

Figure 9. Effect of detergents on lipase activity. Black bars correspond to the extracellular activity with 5 mM detergents, gray bars to the extracellular activity with 10 mM detergents, narrow stripes bars correspond to the intracellular activity with 5 mM detergents and wide stripes bars to the intracellular activity with 10 mM detergents.

et al. 2010; Mansfield and Kuldau 2007). In the present work, the approach using ITS-1-5.8S-ITS2 region sequences in combination with morphological analyses proved to be efficient, giving us a definite identification of these eight fungi. To our knowledge there are no publications regarding lipase production by *H. pseudokoningii*. However, there are reports in the literature in which a new strain of lipolytic *Trichoderma viride* was isolated from the soil on a selective medium that contained olive oil as the only carbon and energy sources (Kashmiri et al. 2006).

Interestingly, high lipase levels were also detected with a solution of 10^7 spores/mL of *H. pseudokoningii*. Evidence for lipolytic activity in fungal spores has been documented by Berto et al. (1999) using *Alternaria brassicicola* which is responsible for an infection in cauliflower leaves. Lipases, as well as other enzymes that act on metabolic reserve molecules are reported in fungi (Aquino et al. 2005) since they facilitate the germination process of these dormant cells.

High yields of lipase were obtained when Tweens were used as carbon source (Nahas 1988). In this work, Tween 80 proved to be a better inducer of lipase production when compared to oils. Chander et al. (1980) reported that butter oil and olive oil inhibited *Aspergillus wentii* lipase production by 53 and 63%, respectively.

The biochemical characterization of the new lipases showed pH profiles with several activity peaks, suggesting the existence of more than one lipase form, as reported by Pernas et al. (2000). It was evident that the lipase produced by *H. pseudokoningii* was stable over a wide pH range, including acidic pHs, in contrast to *Talaromyces thermophilus* which was more stable at alkaline pHs (Romdhane et al. 2010) and unlike the one produced by *Pseudomonas aeruginosa* which is stable in an extremely alkaline environment (Karadzic et al. 2006). The lipases presented in this work showed a thermal stability comparable to the *Penicillium cyclopium* extracellular lipase that had a half-life of 55 minutes at 50°C (Vanot et al. 2002).

The effects of various environmental conditions and chemical agents on the lipase activity and stability were investigated in order to evaluate their potential for industrial application. The effects of different metal ions on activity revealed that the extracellular lipase was inhibited by many ions and was not subject to ion activation, while the intracellular lipase was only inhibited by Na^+ and Ba^{+2} , and was significantly activated by silver. Probably, these ions impose some conformational change toward a less stable structure (Sharon et al. 1998; Akbari et al. 2011). Some reports have shown strong lipase

Screening of filamentous fungi for lipase production 81

activation in the presence of metal ions such as Ca^{2+} , Zn^{2+} , Na^+ , Mg^{2+} , Mn^{2+} , Fe^{2+} , and Hg^{2+} (Kanwar et al. 2006; Hasan et al. 2009). Ca^{2+} stabilized the structure by bridging the active site region to the second subdomain of the protein, thus maintaining the integrity of the complete enzyme and acting as a cofactor for catalytic activity (Sabri et al. 2009).

Enzyme activities in organic solvents depend on the properties and concentration of the solvents and the nature of the enzymes (Akbari et al. 2011). In this work it was observed that, the extracellular lipase was activated by dimethyl sulfoxide, acetone, ethanol, methanol, and only inhibited by butanol (of the solvents tested). So, this lipase was not only stable, but also activated in the presence of organic solvents. The ability of solvents to increase the solubility of substrates, thus facilitating the reaction or to maintain the active structural conformational of the enzyme, might be the basis for higher lipase activity in exposure to organic solvents (Akbari et al. 2011). However, the intracellular lipase activity decreased in the presence of all the organic solvents tested. Banczerz et al. (2005) demonstrated that lipase activity was dependent on the hydrophobicity of the medium.

All detergents tested enhanced the extracellular lipase activity, but only sodium dodecyl sulfate enhanced the intracellular lipase. The activity of the lipase purified from *Aspergillus terreus* was inhibited by ionic detergents, whereas, non-ionic detergents stimulated the enzyme activity (Hasan et al. 2009). Common surfactants except Triton X-100 and cetyltrimethylammonium bromide have no or very little inhibitory effects on the activity of lipase produced by *Bacillus cereus* (Hasan et al. 2009).

Therefore, our study showed that *H. pseudokoningii* lipase has significant potential to be used in industrial processes due to its great stability in different pHs, temperature, and organic solvents.

Acknowledgments

IAJ and MLTMP are Research Fellows of CNPq. APT, FDAF were recipient FAPESP Fellowship. MGP and ACV are supported by CNPq. We thank Ricardo F. Alarcon, Mariana Cereia and Mauricio de Oliveira for technical assistance.

Declaration of interest: The authors report no declarations of interest. The authors alone are responsible for the content and writing of the paper.

This work was supported by grants from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Conselho de Desenvolvimento Científico e Tecnológico (CNPQ) and National System for

82 M. G. Pereira et al.

Research on Biodiversity (Sisbiota-Brazil, CNPq 563260/2010-6/FAPESP n° 2010/52322-3).

References

- Adams PR. Mycelial amylase activities of thermophilic species of *Rhizomucor*, *Humicola* and *Papulaspora*. *Mycopathologia* 1990;112:35–37.
- Ahmed EH, Raghavendra T, Madamwar D. A thermostable alkaline lipase from a local isolate *Bacillus subtilis* EH 37: characterization, partial purification, and application in organic synthesis. *Appl Biochem Biotechnol* 2010;160:2102–2113.
- Akbari N, Daneshjoo S, Akbari J, Khajeh K. Isolation, characterization, and catalytic properties of a novel lipase which is activated in ionic liquids and organic solvents. *Appl Biochem Biotechnol* 2011;165:785–794.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990;215:403–410.
- Aquino ACMM, Peixoto-Nogueira SC, Jorge JA, Terenzi HF, Polizeli MLTM. Characterization of an acid trehalase produced by the thermotolerant fungus *Rhizopus microsporus* var. *rhizopodiformis*: Biochemical properties and immunochemical localization. *FEMS Microbiol Lett* 2005;251:169–175.
- Banczer R, Ginalska G, Fiedurek J, Gromada A. Cultivation conditions and properties of extracellular crude lipase from the psychrotrophic fungus *Penicillium chrysogenum* 9'. *J Ind Microbiol Biotechnol* 2005;32:253–260.
- Berto P, Comménil P, Belingheri L, Dehorter B. Occurrence of a lipase in spores of *Alternaria brassicicola* with a crucial role in the infection of cauliflower leaves. *FEMS Microbiol Lett* 1999;180:183–189.
- Bradford MM. Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-Dye binding. *Anal Biochem* 1976;72:248–254.
- Chander H, Batish VK, Sannabhadri SS, Srinivasan RA. Factors affecting lipase production in *Aspergillus wentii*. *J Food Sci* 1980;45:598–600.
- Ciarro DE, Lucke K, Imhof A, Bloemberg GV, Böttger EC. Systematic internal transcribed spacer sequence analysis for identification of clinical mold isolates in diagnostic mycology: a 5-year study. *J Clin Microbiol* 2010;48:2809–2813.
- Emerson R. An Experimental study of the life cycles and taxonomy of Allomyces. *Lloydia* 1941;4:77–144.
- Fernandez-Lafuente R. Lipase from *Thermomyces lanuginosus*: uses and prospects as an industrial biocatalyst. *J Mol Catal B Enzym* 2010;62:197–212.
- Fuciños P, Abadín CM, Sanromán A, Longob MA, Pastrana L, Rúa ML. Identification of extracellular lipases/esterases produced by *Thermus thermophilus* HB27: Partial purification and preliminary biochemical characterization. *J Biotechnol* 2005;117:233–241.
- Ghikas DV, Kouvelis VN, Typas MA. Phylogenetic and biogeographic implications inferred by mitochondrial intergenic region analyses and ITS1–5.8S-ITS2 of the entomopathogenic fungi *Beauveria bassiana* and *B. brongniartii*. *BMC Microbiol* 2010;10:174–189.
- Gross RA, Kalra B, Kumar A. Polyester and polycarbonate synthesis by in vitro enzyme catalysis. *Appl Microbiol Biotechnol* 2001;55:655–660.
- Hasan F, Shah AA, Hamed A. Methods for detection and characterization of lipases: a comprehensive review. *Biotechnol Adv* 2009;27:782–798.
- Kanwar SS, Ghazi IA, Chimni SS, Joshi GK, Rao GV, Kaushal RK, Gupta R, Punj V. Purification and properties of a novel extracellular thermotolerant metalloproteinase of *Bacillus coagulans* MTCC-6375 isolate. *Protein Expr Purif* 2006;46:421–428.
- Karadzic I, Masui A, Zivkovic LJ, Fujiwara N. Purification and characterization of an alkaline lipase from *Pseudomonas aeruginosa* isolated from putrid mineral cutting oil as component of metalworking fluid. *J Biosci Bioeng* 2006;102:82–89.
- Kashmiri MA, Adnan A, Butt BW. Production, purification and partial characterization of lipase from *Trichoderma viride*. *Afr J Biotechnol* 2006;5:878–882.
- Khanna P, Sundari SS, Kumar NJ. Production, isolation and partial purification of xylanase from *Aspergillus* sp. *World J Microbiol Biotechnol* 1995;11:242–243.
- Mansfield MA, Kuldau GA. Microbial and molecular determination of mycobiota in fresh and ensiled maize silage. *Mycologia* 2007;99:269–278.
- Mellvaine TC. A buffer solution for colorimetric comparison. *J Biol Chem* 1921;49:183–186.
- Nahas E. Control of lipase production by *Rhizopus oligosporus* under various growth conditions. *J Gen Microbiol* 1988;134:227–233.
- Parawira W. Biotechnological production of biodiesel fuel using biocatalysed transesterification: a review. *Crit Rev Biotechnol* 2009;29:82–93.
- Pencreac'h G, Baratti JC. Hydrolysis of p-nitrophenyl palmitate in n-heptane by *Pseudomonas cepacia* lipase: a simple test for the determination of lipase activity in organic media. *Enzyme Microb Technol* 1996;18:417–422.
- Peralta RM, Terenzi HF, Jorge JA. Beta-D-glycosidase activities of *Humicola grisea*: biochemical and kinetic characterization of a multifunctional enzyme. *Biochim Biophys Acta* 1990;1033:243–249.
- Pernas MA, López C, Pastrana L, Rúa ML. Purification and characterization of Lip2 and Lip3 isoenzymes from a *Candida rugosa* pilot-plant scale fed-batch fermentation. *J Biotechnol* 2000;84:163–174.
- Punekar NS, Suresh Kumar SV, Jayashri TN, Anuradha R. Isolation of genomic DNA from acetone-dried *Aspergillus nycelia*. *Fungal Genet News* 2003;50:15–16.
- Reis P, Holmberg K, Watzke H, Leser ME, Miller R. Lipases at interfaces: a review. *Adv Colloid Interface Sci* 2009;147:237–250.
- Rizzatti ACS, Jorge JA, Terenzi HF, Rechia CGV, Polizeli MLTM. Purification and properties of a thermostable extracellular-D-xylosidase produced by thermotolerant *Aspergillus phoenicis*. *J Ind Microbiol Biotechnol* 2001;26:156–160.
- Romdhane IBB, Fendri A, Gargouri A, Belghith H. A novel thermoactive and alkaline lipase from *Talaromyces thermophilus* fungus for use in laundry detergents. *Biochem Eng J* 2010;53:112–120.
- Sabri S, Rahman RNZRA, Leow TC, Basri M, Salleh AB. Secretory expression and characterization of a highly Ca²⁺-activated thermostable L2 lipase. *Protein Expr Purif* 2009;62:161–166.
- Sharma R, Chisti Y, Banerjee UC. Production, purification, characterization, and applications of lipases. *Biotechnol Adv* 2001;19:627–662.
- Sharon C, Furugoh S, Yamakido T, Ogawa H, Kato Y. Purification and characterization of a lipase from *Pseudomonas aeruginosa* KKA-5 and its role in castor oil hydrolysis. *J Ind Microbiol Biotechnol* 1998;20:304–307.
- Singh AK, Mukhopadhyay M. Overview of fungal lipase: a review. *Appl Biotechnol* 2012;166:486–520.
- Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994;22:4673–4680.

- Undurraga D, Markovits A, Erazo S. Cocoa butter equivalent through enzymatic interesterification of palm oil midfraction. *Process Biochem* 2001;36:933-939.
- Vanot G, Valérie D, Guilhem MC, Luu RPT, Comeau LC. Maximizing production of *Penicillium cyclopium* partial acylglycerol lipase. *Appl Microbiol Biotechnol* 2002;60: 417-419.
- Vogel HF. Distribution of lysine pathways among fungi: evolutionary implications. *Amer Naturalist* 1964;98:435-446.
- White TJ, Bruns T, Lee S, Taylor JW. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, eds. *PCR protocols: a guide to methods and applications*. New York: Academic Press. pp. 315-322.
- Wiseman A. 1975. *Handbook of enzyme biotechnology*. Chichester: Ellis Horwood Ltd, John Wiley & Sons. pp. 148.
- Wu XM, Xin JY, Sun W, Xi CG. Environmentally friendly, efficient resolution of racemic secondary alcohols by lipase-catalyzed enantioselective transesterification in ionic liquids in the presence of organic bases. *Chem Biodivers* 2007;4: 183-188.

Screening of filamentous fungi for lipase production 83