

III Workshop de Dissertações de Mestrado do PPgSI (2017)

FICHA DA PESQUISA¹

DADOS GERAIS					
Título do projeto de pesquisa	Uma nova abordagem para identificar a provável origem de genes específicos				
Orientando	Priscilla Koch Wagner				
Orientador(es)	Luciano Antonio Digiampietri				
Semestre no curso, na data do workshop	<input type="checkbox"/> 2º semestre	<input type="checkbox"/> 3º semestre	<input type="checkbox"/> 4º semestre	<input checked="" type="checkbox"/> 5º semestre	<input type="checkbox"/> 6º semestre
Qualificação	<input checked="" type="checkbox"/> Qualificação já realizada em: 27/06/2016		<input type="checkbox"/> Qualificação planejada para: ___/___/___		
Defesa	Prazo máximo para depósito: 05/02/2018		Depósito planejado para: 31/10/2017		
Linha e Área de pesquisa	Gestão e desenvolvimento de Sistemas: <input type="checkbox"/> BD <input type="checkbox"/> Eng. de Software		Inteligência de Sistemas: <input checked="" type="checkbox"/> IA <input type="checkbox"/> Proc. Gráfico		<input type="checkbox"/> Gestão de SI <input type="checkbox"/> IHC <input type="checkbox"/> Rec. de Padrões
Área de aplicação	<input type="checkbox"/> Ambientes Corporativos <input checked="" type="checkbox"/> Bioinformática <input type="checkbox"/> Biometria <input type="checkbox"/> Economia	<input type="checkbox"/> Educação <input type="checkbox"/> Educação a Distância <input type="checkbox"/> Internet <input type="checkbox"/> Jogos	<input type="checkbox"/> Linguagem Natural <input type="checkbox"/> Linguística <input type="checkbox"/> Processos de Negócio <input type="checkbox"/> Química	<input type="checkbox"/> Redes Sociais <input type="checkbox"/> Robótica <input type="checkbox"/> Saúde <input type="checkbox"/> [outro – escrever]	
Publicações associadas ao projeto de mestrado	<p>Até o momento, a aluna já participou da publicação de dois resumos no formato de pôster e um artigo:</p> <p>- Pereira, V. M. Y.; WAGNER, P. K.; Digiampietri, L.A. Bacterial whole genome comparison: a systematic literature review. In: X-Meeting 2015 - 11th International Conference of the AB3C + Brazilian Symposium of Bioinformatics. 2015.</p> <p>- WAGNER, P. K.; Pereira, V. M. Y.; Digiampietri, L. A. A systematic review of phylogenetic analysis of a special gene. In: X-Meeting 2015 - 11th International Conference of the AB3C + Brazilian Symposium of Bioinformatics. 2015.</p> <p>- WAGNER, P. K.; Mattos, G. D.; Digiampietri, L. A. Análise de erros de anotação de genes exclusivos. In: XI Brazilian e-Science Workshop (BreSci 2017), 2017, São Paulo, SP. Anais do XXXVII Congresso da Sociedade Brasileira de Computação (CSBC 2017), 2017</p>				

DESCRIÇÃO DO PROJETO DE PESQUISA	
Contextualização / motivação	A identificação da origem de um gene, especialmente de genes exclusivos, é fundamental para entender as características específicas dessa espécie e ajuda na identificação de doenças ou no desenvolvimento de vacinas.
Problema de pesquisa	De forma geral, o problema abordado neste trabalho é a identificação da filogenia de genes exclusivos de um genoma de um grupo de espécies próximas, levando em consideração análises da anotação desses genes.
Objetivo geral da pesquisa	O objetivo deste projeto é a aplicação de uma nova abordagem para análise da filogenia de genes exclusivos, considerando os genomas do banco mundial, melhorando e integrando ferramentas já existentes; permitindo o entendimento das possíveis relações evolutivas de um dado gene exclusivo em relação a diferentes espécies.
Trabalhos relacionados	Tan e Wu (2011) propõem o Phytree, um workbench que integra as ferramentas Blast, ClustalW e Tree-Puzzle, para pesquisa de sequências homólogas, alinhamento múltiplo e reconstrução da árvore filogenética. Moreira et al. (2010) estuda genomas de <i>Xanthomonas</i> , com ênfase nos genes exclusivos presentes em determinadas espécies. Apesar de não ser um trabalho da área de computação, poderá ser utilizado para comparação e validação dos resultados obtidos neste projeto. Por fim, Pereira (2014) desenvolveu as ferramentas utilizadas nesse trabalho para seleção de genes exclusivos.
Justificativa e relevância	Apesar da importância da identificação da filogenia de genes exclusivos, não foram encontradas abordagens robustas para a realização dessa tarefa em específico. As ferramentas existentes atualmente são genéricas, não apresentando bons resultados para análise de genes exclusivos. Além disso, são capazes de apresentar resultados

¹ Esta ficha é uma adaptação da usada no “VIII Workshop de Teses e Dissertações em Sistemas de Informação (WTDSI 2015)” realizado como parte do “XI Simpósio Brasileiro de Sistemas de Informação (SBSI 2015)”

	detalhados para pequenos volumes de dados. Porém, encontram dificuldades ao lidar com grandes volumes de dados, onde existe cruzamento de informações de espécies muito distantes evolutivamente.
Proposta para Solução	A proposta visa a especificar uma abordagem suficientemente robusta e adequada, que integre as tarefas do processo de análise filogenética, melhorando as ferramentas existentes e lidando com grandes volumes de dados. Além disso, o projeto analisa algumas hipóteses em relação à origem de genes exclusivos de bactérias, bem como divergências no processo de anotação de genes.
Dados	O presente projeto utiliza em seus estudos de caso dados do banco de genomas do NCBI (National Center for Biotechnology Information), especialmente genomas de bactérias do gênero <i>Xanthomonas</i> . Essa base de genomas contém as sequências de aminoácidos e nucleotídeos de diversas espécies.
Validação	Os resultados e as ferramentas desenvolvidas serão validados com especialistas do domínio. Outra forma de avaliar será através da busca por possíveis literaturas sobre o gene estudado. O objetivo é descobrir se a filogenia obtida reflete a real origem dos genes exclusivos dos indivíduos e a qual família pertencem.
Limitações	Apesar da abordagem ser genérica, o projeto estudará apenas genes exclusivos de bactérias, por conta da variedade de genes existentes. Por possuírem poucas pesquisas a respeito de genes exclusivos de bactérias, uma possível limitação será a validação dos resultados obtidos a partir da comparação com filogenias existentes.
Resultados esperados	Contribuições científicas: facilitar pesquisas de biologia comparativa identificando possíveis filogenias, além de apresentar resultados importantes a respeito da origem de genes exclusivos.
	Contribuições tecnológicas: implementação de uma ferramenta robusta e integrada das etapas do processo de análise filogenética, para obtenção de filogenias a partir de um dado gene exclusivo.

MÉTODO DE PESQUISA [para cada item, marque todas as opções adequadas para seu projeto]				
Gênero	<input type="checkbox"/> Pesquisa teórica	<input type="checkbox"/> Pesquisa prática	<input checked="" type="checkbox"/> Pesquisa empírica	<input type="checkbox"/> Pesquisa metodológica
Natureza	<input type="checkbox"/> Pesquisa básica/pura	<input checked="" type="checkbox"/> Pesquisa aplicada		
Objetivo	<input checked="" type="checkbox"/> Pesquisa exploratória	<input type="checkbox"/> Pesquisa descritiva	<input type="checkbox"/> Pesquisa explicativa	<input type="checkbox"/> [outro – escrever]
Abordagem	<input type="checkbox"/> Pesquisa quantitativa	<input type="checkbox"/> Pesquisa qualitativa	<input checked="" type="checkbox"/> Pesquisa mista (quali-quant)	
Procedimento(s) técnico(s)	<input checked="" type="checkbox"/> Pesquisa experimental <input checked="" type="checkbox"/> Pesquisa bibliográfica <input type="checkbox"/> Pesquisa documental <input type="checkbox"/> Pesquisa <i>ex-post-facto</i> <input type="checkbox"/> Pesquisa de levantamento	<input type="checkbox"/> Pesquisa com <i>survey</i> <input checked="" type="checkbox"/> Estudo de caso <input type="checkbox"/> Pesquisa participante <input type="checkbox"/> Pesquisa-ação <input type="checkbox"/> Pesquisa etnográfica	<input type="checkbox"/> Pesquisa netnográfica <input type="checkbox"/> Teoria fundamentada em dados (<i>grounded theory</i>) <input type="checkbox"/> Ciência do projeto (<i>Design science research</i>) <input type="checkbox"/> [outro – escrever]	
Fonte(s) de dados	<input type="checkbox"/> pesquisa de laboratório	<input type="checkbox"/> pesquisa de campo	<input type="checkbox"/> pesquisa bibliográfica	<input checked="" type="checkbox"/> banco de dados de genomas do NCBI
Técnica(s) / Instrumento(s) de coleta de dados	<input type="checkbox"/> medição <input type="checkbox"/> questionário <input type="checkbox"/> entrevista <input type="checkbox"/> grupos focais	<input type="checkbox"/> formulário <input type="checkbox"/> <i>benchmark</i>	<input type="checkbox"/> observação (direta / participante) <input type="checkbox"/> diário de campo / notas de campo <input type="checkbox"/> análise documental (ou de artefatos) <input checked="" type="checkbox"/> pesquisa em banco de dados de genomas do NCBI	
Técnica(s) de análise de dados	Análise quantitativa: <input type="checkbox"/> Estatística descritiva <input type="checkbox"/> Estatística inferencial <input type="checkbox"/> [outro – escrever]		Análise qualitativa: <input checked="" type="checkbox"/> Análise de conteúdo <input type="checkbox"/> Análise do discurso <input type="checkbox"/> [outro – escrever]	

CRONOGRAMA [altere conforme necessário]																								
	2015				2016								2017											
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
Estudo da literatura					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X				
Revisão sistemática								X	X	X	X	X												
Definição da proposta					X	X	X	X	X	X														
Execução da proposta													X	X	X	X	X	X	X	X	X	X	X	X
Coleta/obtenção de dados													X	X	X	X	X	X	X	X				

