

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

DANIELE RIBEIRO DE SOUZA

**ANÁLISE GENÔMICA COMPARATIVA DA *Xanthomonas arboricola* pv. *juglandis*
417 E ESTUDO *IN SILICO* DO GENE *copb* DE RESISTÊNCIA AO COBRE**

UBERLÂNDIA

2021

Daniele Ribeiro de Souza

**ANÁLISE GENÔMICA COMPARATIVA DA *Xanthomonas arboricola* pv. *juglandis*
417 E ESTUDO *IN SILICO* DO GENE *copb* DE RESISTÊNCIA AO COBRE**

Trabalho de Conclusão de Curso II apresentado ao Curso de Ciências Biológicas como pré-requisito para a obtenção do Título de Bacharel em Ciências Biológicas pela Universidade Federal de Uberlândia.

Orientador: Prof. Dr. Nilson Nicolau Junior (Instituto de Biotecnologia).

UBERLÂNDIA

2021

**ANÁLISE GENÔMICA COMPARATIVA DA *Xanthomonas arboricola* pv. *juglandis*
417 E ESTUDO *IN SILICO* DO GENE *copb* DE RESISTÊNCIA AO COBRE**

Trabalho de Conclusão de Curso aprovado para a
obtenção do Título de Bacharel em Ciências
Biológicas pela Universidade Federal de
Uberlândia pela banca examinadora formada por:

Uberlândia, 20 de outubro de 2021.

Prof. Dr Nilson Nicolau Junior, UFU

Prof. Dra. Viviane Rodrigues Alves de Moraes, UFU

Profa. Dra. Lizandra Ferreira de Almeida e Borges, UFU

AGRADECIMENTOS

Agradeço ao meu orientador, Prof. Dr. Nilson Nicolau Junior, pela oportunidade, conhecimento, paciência e dedicação.

Aos professores, pelo incentivo e aprendizado ao longo do curso.

Agradeço Evelin Silva e Bruna Gomes, pelo companheirismo e apoio.

À Universidade Federal de Uberlândia e o Instituto de Biologia pelas oportunidades e recursos.

À minha mãe Adilma (*in memoriam*) e meu pai Jair que são as minhas maiores motivações para conclusão desta etapa em minha vida.

RESUMO

A *Xanthomonas arboricola* pv. *juglandis* é um organismo com presente resistência ao metal Cobre e o causador da ferrugem bacteriana, fitopatogenia que acomete Nogueiras, em todo o mundo. Através de estudos *in silico* da espécie, foram realizadas análises, em nível molecular, de hipóteses acerca do operon *cop* e do gene de resistência *copB*, além de comparações genômicas entre a espécie bacteriana e outras espécies filogeneticamente próximas. Por meio da utilização de softwares de comparação genômica como: EDGAR 3.0 e MAUVE, foi possível identificar os organismos mais próximos filogeneticamente do organismo referencial deste estudo, sendo eles: *Xanthomonas campestris* pv. *campestris*, com 253 sequências compatíveis e *Xanthomonas hortorum* pv. *carotae* com 148 sequências compatíveis. Baseado no resultado de identificação de similaridade em alinhamentos, gerado pelo software BLAST2GO, foram filtrados 5 genes relacionados à resistência ao cobre: *copA*, *copB*, *copC* e *copD*, sendo o gene *copB* expressado duas vezes. Juntamente com os genes citados, outros genes relacionados a funcionalidade da resistência também foram identificados: expressão de amina oxidase, expressão de transportador de cobre ABC ATPase, expressão de homeostase do cobre *cutC*, expressão de multicobre oxidase (MCO), expressão de regulador transcricional da família *copG* e expressão de chaperona de cobre PCu(A)C. A predição de processo biológico associado à proteína *copb* 119 indicando paridade com a *copb* 3389 no que se refere ao processo biológico. Alinhamento entre os genes *copB* apresentou taxa de similaridade residual igual à 58,9% e Identidade residual igual à 47,0%. RMSD de alinhamento estrutural entre os genes *copB* foi igual à 3.22.

Palavras-chave: Bioinformática. Análise molecular. Predição estrutural. *Xaj* 417. Alinhamento estrutural.

ABSTRACT

Xanthomonas arboricola pv. *juglandis* is an organism with present resistance to the metal Copper and the cause of bacterial blight, phytopathogen that affects Walnut trees worldwide. Via *in silico* studies of the species, molecular level analyses of hypotheses about the *cop* operon and the *copB* resistance gene and in addition, genomic comparisons between the bacterial species and two other phylogenetically close species were performed. Using genomic comparison softwares such as EDGAR 3.0 and MAUVE, it was possible to identify the organisms phylogenetically closest to the reference organism of this study: *Xanthomonas campestris* pv. *campestris*, with 253 compatible sequences and *Xanthomonas hortorum* pv. *carotae* with 148 compatible sequences. Based on the result of similarity identification in alignments, generated by BLAST2GO software, 5 genes related to copper resistance were expressed: *copA*, *copB*, *copC* and *copD*, where the *copB* gene was expressed twice. Along with the mentioned genes, other genes related to resistance functionality were also identified: amine oxidase, copper ABC transporter ATPase, copper homeostasis protein *cutC*, multicopper oxidase (MCO), *copG* family transcriptional regulator and copper chaperone PCu(A)C. The *copB* 119 prediction of biological process indicates response to copper ion, a parity with *copB* 3389 regarding the biological process. Similarity between *copB* genes: 58,9%, Identity between *copB* genes: 47,0%, and RMSD value of *copB* genes structural alignments is 3.22.

Keywords: Bioinformatics. Molecular analysis. Structural Prediction. *Xaj* 417. Structural alignment.

LISTA DE ILUSTRAÇÕES

Figura 1 - Pontos necróticos nas folha de <i>J. regia</i> causadas pela bactéria <i>Xanthomonas arboricola</i> pv. <i>juglandis</i>	12
Figura 2 - Manchas no fruto da <i>J. regia</i> causadas pela bactéria <i>Xanthomonas arboricola</i> pv. <i>juglandis</i>	13
Figura 3 - <i>Juglans regia</i> L. localizada no Jardim Botânico da Universidade de Trás-os-Montes e Alto Douro (UTAD).....	14
Figura 4 - Noz da <i>Juglans regia</i>	15
Figura 5 - Sequência da <i>copB</i> 119 submetida ao servidor C-I-TASSER.....	26
Figura 6 - Sequência da <i>copB</i> 3389 submetida ao servidor C-I-TASSER.....	26
Figura 7 - Pontuação de confiabilidade - Cscore ^{GO}	27
Figura 8 - Distribuição de dados genômicos da <i>Xanthomonas arboricola</i> pv. <i>juglandis</i> 417..	28
Figura 9 - Representação circular do genoma bacteriano da <i>Xanthomonas arboricola</i> pv. <i>juglandis</i> 417.....	29
Figura 10 - Árvore Filogenética projetada ao selecionar o gênero <i>Xanthomonas</i> e construída pelo método neighbor-joining.....	30
Figura 11 - <i>Xanthomonas arboricola</i> pv. <i>juglandis</i> 417 e os dois organismos mais próximos filogeneticamente do referencial de estudo: <i>Xanthomonas campestris</i> pv. <i>campestris</i> e <i>Xanthomonas hortorum</i> pv. <i>carotae</i>	32
Figura 12 - Diagrama de Venn: 1- <i>Xanthomonas arboricola</i> pv. <i>juglandis</i> 417. 2 - <i>Xanthomonas campestris</i> pv. <i>campestris</i> . 3 - <i>Xanthomonas hortorum</i> pv. <i>carotae</i>	33

Figura 13 - Alinhamento de sequências em formato de blocos colineares 1 - <i>Xanthomonas hortorum</i> pv. <i>carotae</i> 2 - <i>Xanthomonas arboricola</i> pv. <i>juglandis</i> 417 3 - <i>Xanthomonas campestris</i> pv. <i>campestris</i>	34
Figura 14 - Alinhamento global das proteínas <i>copB</i> 119 e <i>copB</i> 3389 de resistência ao cobre.	38
Figura 15 - Predição estrutural da proteína <i>copB</i> 119 de resistência ao cobre.....	40
Figura 16 - Predição estrutural da proteína <i>copB</i> 3389 de resistência ao cobre.....	41
Figura 17 - Diagrama da predição de processo biológico da proteína <i>copB</i> 119 de resistência ao cobre.....	43
Figura 18 - Diagrama da predição de processo biológico da proteína <i>copB</i> 3389 de resistência ao cobre.....	44
Figura 19 - Alinhamento estrutural das proteínas <i>copB</i> 119 e <i>copB</i> 3389 de resistência ao cobre.....	45

SUMÁRIO

1.	INTRODUÇÃO.....	12
1.1	Gênero <i>Xanthomonas</i>	12
1.1.1	<i>Xanthomonas arboricola</i> pv. <i>juglandis</i> e patogenia relacionada.....	12
1.1.2	Cepa 417 da <i>Xanthomonas arboricola</i> pv. <i>juglandis</i>	13
1.2	<i>Juglans regia</i> L e sua importância.....	14
1.3	Cobre.....	16
1.3.1	Atributos antimicrobianos.....	16
1.3.2	O uso do cobre como biocida em plantas.....	17
1.4	Análise Genômica e o uso de banco de dados biológicos.....	17
1.5	<i>Deep learning</i> e bioinformática estrutural.....	18
1.5.1	Níveis de estruturação proteica.....	19
1.5.2	Modelagem Molecular: método <i>threading</i>	20
1.6	Resistência bacteriana.....	20
1.6.1	Operon <i>cop</i>	21
2.	OBJETIVOS.....	22
2.1	Objetivos gerais.....	22

2.2	Objetivos específicos.....	22
3.	MATERIAL E MÉTODOS.....	23
3.1	Organismo e levantamento de dados.....	23
3.1.1	Organismo.....	23
3.1.2	Distribuição de dados.....	23
3.2	Genômica Comparativa.....	23
3.2.1	Visualização do genoma bacteriano.....	23
3.2.2	Árvore Filogenética e Diagrama de Venn.....	24
3.2.3	Alinhamento de sequências.....	24
3.3	Genes referentes à resistência ao cobre.....	24
3.4	Análise comparativa estrutural e funcional das <i>copBs</i>	25
3.4.1	Análise estrutural.....	25
3.4.1.1	Alinhamento global de sequências.....	25
3.4.1.2	Predição estrutural e funcional.....	25
3.4.3	Alinhamento estrutural das <i>copBs</i>	27
4.	RESULTADOS E DISCUSSÃO.....	27
4.1	Organismo.....	27

4.2	Distribuição de dados.....	27
4.3	Genômica Comparativa.....	28
4.3.1	Visualização do genoma bacteriano.....	28
4.3.2	Árvore Filogenética e Diagrama de Venn.....	29
4.3.3	Alinhamento de sequências.....	34
4.4	Genes de Resistência.....	35
4.4.1	Genes referentes à resistência ao cobre.....	35
4.5	Análise comparativa estrutural e funcional da <i>copB</i> 119 e a <i>copB</i> 3389.....	38
4.5.1	Alinhamento global de sequências.....	38
4.5.2	Predição estrutural.....	40
4.5.3	Anotação funcional comparativa: Processo biológico.....	42
4.5.4	Alinhamento estrutural das <i>copBs</i>	45
5.	CONCLUSÃO.....	46
	REFERÊNCIAS BIBLIOGRÁFICAS.....	47

1. INTRODUÇÃO

1.1 Gênero *Xanthomonas*

O gênero *Xanthomonas*, pertencente à família *Xanthomonadaceae*, é um gênero de bactérias gram-negativas, baciliformes e aeróbicas, que compõe 27 espécies, sendo a maioria dessas espécies designadas como fitopatógenos (SHARMA; GAUTAM & WADHAWAN, 2014) (ANDREANI; FASOLATO, 2017).

Mesmo com a capacidade de colonizar até 400 espécies de plantas, o gênero *Xanthomonas*, apresenta espécies e patovares de grande especificação, que conseguem colonizar tecidos da espécie hospedeira (RYAN et al., 2011).

1.1.1 *Xanthomonas arboricola* pv. *juglandis* e patogenia relacionada

Xanthomonas arboricola pv. *juglandis* (*Xaj*) (Pierce 1901) é designada como uma proteobacteria patogênica associada à plantas (SWINGS; CIVEROLO, 1993). Classificada como gammaproteobacteria, a *Xaj* acomete principalmente a espécie de planta *Juglans regia*, causando a ferrugem bacteriana, como é popularmente conhecida, no limbo e no fruto da espécie (FERNANDES et al., 2017).

A infecção pela bactéria *Xanthomonas arboricola* pv. *juglandis* depende das condições ambientais, da quantidade de patógeno em gemas individuais (inóculo) e da quantidade de cancrios presentes (UC ANR STATEWIDE INTEGRATED PEST MANAGEMENT, 2017).

A presença da bactéria pode ser identificada pela aparição de pontos necróticos nas folhas (Fig. 1) e manchas (Fig. 2) nos galhos e frutos. Considerada um fator limitante na produção de *Juglans regia*, a *Xaj* pode prejudicar 50% dos frutos, ocasionando perda severa na produtividade da espécie (GIOVANARDI et al., 2015).

Figura 1 - Pontos necróticos nas folha de *J. regia* causadas pela bactéria *Xanthomonas arboricola* pv. *juglandis*.



Fonte: Monika Katuzna et al.

Figura 2 - Manchas no fruto da *J. regia* causadas pela bactéria *Xanthomonas arboricola* pv. *juglandis*.



Fonte: Monika Katuzna et al.

1.1.2 Cepa 417 da *Xanthomonas arboricola* pv. *juglandis*

Descrita em 2015, a cepa 417 da espécie *X. arboricola* pv. *juglandis*, é uma cepa resistente ao cobre isolada de um cultivar de noz Chandler afetado pela ferrugem bacteriana, na cidade de Chico, estado da Califórnia, em 2012 (PEREIRA et al., 2015).

1.2 *Juglans regia* L e sua importância

A Nogueira (*Juglans regia* L.), família *Juglandaceae*, também designada como noqueira-comum, é uma espécie de angiosperma descrita em 1753 por Linnaeus (IPNI, 20--) e de origem na Ásia Central.

A *J. regia* (Fig. 3) possui folhas alternas, imparipinadas com 3-9 folíolos, sua copa é densa e ampla e consegue chegar até 30 metros de altura (FLORESTAR, 20--). Seu fruto, a noz, é de grande valor econômico assim como sua madeira (EBRAHIMI, 2016), além disso a espécie é utilizada na fabricação de licores, cosméticos, tintas e na medicina tradicional (CABRAL, 2008).

Figura 3 - *Juglans regia* L. localizada no Jardim Botânico da Universidade de Trás-os-Montes e Alto Douro (UTAD).



Fonte: Tim Sheerman-Chase, 2017.

No que se refere as nozes (Fig. 4), consumidas mundialmente, estão relacionadas benéficamente a saúde, onde a ingestão do fruto traz como vantagens o controle de deslipidemia e hipertensão (CABRAL, 2008) e a espécie também se encontra presente em diversos estudos voltados para o tratamento de diabetes e condições inflamatórias juntamente com outras plantas medicinais (PANTH, 2016). O cultivo da *Juglans regia* atualmente é feita em vários países sendo os três principais: Estados Unidos, China e Turquia (RIGO et al., 2016).

Figura 4 - Noz da *Juglans regia*.



Fonte: Jonson22, commons.wikimedia.org.

1.3 Cobre

O cobre é um elemento metálico de número atômico 29 e peso atômico 63,54, com coloração marrom avermelhado, quando encontrado em sua forma pura e com densidade de 8,94 g/cm³. São presentes no cobre dois isótopos estáveis, ⁶³Cu e ⁶⁵Cu e quatro estados de oxidação assumíveis: +0,+1,+2 e +3 (ELLINGSEN et al, 2015).

No século XIX sua utilização era favorável devido às propriedades mecânicas desse metal ou seja maleabilidade, trabalhabilidade e sua adequação para fundição (CDA, 2010). Neste século a utilização do cobre continua sendo muito requisitada nas mais variadas áreas do mercado como: produtos antimicrobianos, componentes eletrônicos, energia renovável, permutadores de calor e tecnologias de informação e telecomunicação (CDA, 2010)

1.3.1 Atributos antimicrobianos

Desde 2008, quando a Agência de Proteção Ambiental dos Estados Unidos (EPA) listou aproximadamente 300 ligas metálicas contendo cobre, considerando-as como eficientes agentes antimicrobianos, o número de estudos voltado às propriedades antimicrobianas do cobre cresceram significativamente (VINCENT et al., 2016).

Mesmo com seu grande potencial como agente antimicrobiano, condições para essa potencialização devem ser consideradas, entre elas a aplicação por via seca ou úmida do cobre. Estudos relacionados a via de aplicação apresentam resultados semelhantes, indicando que a eficácia na aplicação por via seca é mais rápida do que na aplicação por via úmida (VINCENT et al., 2016).

1.3.2 O uso do cobre como biocida em plantas

O termo Biocida pode ser definido como uma preparação formada por uma ou mais substâncias ativas e que tem como objetivo exercer um efeito de controle sob um organismo, destruindo-o ou tornando-o menos prejudicial por meios químicos ou biológicos (MICHALAK & CHOJNACKA, 2014).

A preparação pode exercer funcionalidade como um agente eletrofílico, causando dano em enzimas para impedir crescimento e metabolismo e conseqüentemente ocasionando morte celular ou ainda, funcionar como um agente que atue diretamente na membrana celular (MICHALAK & CHOJNACKA, 2014).

Utilizado em grandes concentrações como biocida, o cobre interage com ácidos nucleicos, rompimento de sítios ativos em enzimas, interfere no sistema de transporte de energia e rompe a integridade de membranas celulares (LAMICHHANE et al., 2018). Existem estudos que indicam a frequente aplicação de cobre como causa de resistência em bactérias. (LAMICHHANE et al., 2018).

1.4 Análise Genômica e o uso de banco de dados biológicos

Em *The Delphic Boat: What Genome Tell us* (DANCHIN, 2002), a análise de genomas é acentuada como uma ramificação da bioinformática, que acompanhada do termo *in*

silico define uma sequência de técnicas computacionais voltadas a hipóteses biológicas por meio de estudos moleculares (DANCHIN, 2002). Para a análise *in silico* são utilizadas ferramentas que identificam e caracterizam genes, identificam elementos transponíveis e também a correlação entre diferentes genomas na busca de pontos evolutivos comuns (VERLI, 2014).

Comumente, a informática está inserida do início ao fim no que se refere à dados genômicos, sendo utilizada desde a coleta de dados até a condução dos mesmos para utilização em próximas etapas (DANCHIN, 2002). O investimento na busca de novos algoritmos para a análise genômica é constante, tornando-os cada vez mais amplos devido à complexidade e aumento do número de sequências analisadas (VERLI, 2014).

Baseado na análise genômica, é possível explorar uma variedade de tópicos, sendo eles: montagem de genomas, montagem de transcriptomas, anotação gênica e anotação de RNAs não codificantes (VERLI, 2014).

Um banco de dados biológicos, para a bioinformática, é uma fonte de dados de extrema importância devido a sua capacidade de armazenamento de informações, organização e compartilhamento (ZOU et al., 2015). A partir da facilidade presente nas tecnologias sequenciais, existe um crescimento exponencial quanto a criação de bancos de dados para gerir a grande quantidade de novas informações (ZOU et al., 2015).

Presente neste estudo, o banco de dados nucleotídicos, NCBI (National Center for Biotechnology Information), recolhe de repositórios online: genes, genomas e sequências de transcrição, unindo os dados em apenas um banco (MUKHOPADHYAY et al., 2018) para auxiliar a coleta de informações.

1.5 *Deep learning* e bioinformática estrutural

Em consequência da grande quantidade de dados gerados atualmente, o *Deep Learning*, também denominado como Aprendizagem Profunda, encontra-se incluído no campo de Inteligência Artificial, mais especificamente na metodologia de Aprendizado de Máquina que está presente na construção de ferramentas computacionais usadas na

Bioinformática (LECUN, 2019; KOUMAKIS, 2020).

Agora, o mesmo algoritmo, *Deep Learning*, utilizado em processamento de imagens e reconhecimento oral está sendo implementado na modelagem de proteínas de forma que possa ser uma vantagem na análise de banco de dados sequenciais e estruturais (KUHLMAN & BRADLEY, 2019).

A bioinformática estrutural, quando voltada para estudo de proteínas, busca descobrir o vínculo entre aminoácidos e suas funções e conseqüentemente auxiliar na identificação ou delineamento de proteínas capazes de se ligarem em sítios específicos (TORRISI; POLLASTRI & LE, 2020).

Sendo assim, o rápido avanço de softwares e automação causaram aumento no ritmo de determinação de estruturas experimentais e conseqüentemente acelerou o crescimento do banco de dados voltado à estrutura de proteínas, o Protein Data Bank (PDB) (BERMAN et al., 2000), que hoje contém 181.969 macromoléculas estruturadas.

1.5.1 Níveis de estruturação proteica

Aminoácidos são unidades estruturais formadoras de uma proteína e são precisamente denominadas como α -aminoácidos (CHOUDHURI, 2014). A localização de aminoácidos na conformação proteica conseqüentemente dita a função da proteína e a interação com o meio: proteínas em meio hidrofóbico dispõem de cadeias laterais apolares que interagem com membranas lipídicas, enquanto proteínas em meio aquoso, como citosol, possuem cadeias laterais polares para interação com o meio (CHOUDHURI, 2014).

Uma proteína compreende quatro níveis de estruturação: estrutura primária, estrutura secundária, estrutura terciária e estrutura quaternária (CHOUDHURI, 2014). A estrutura primária abrange a sequência de aminoácidos enquanto a secundária refere-se à conformação polipeptídica, sendo α -hélice e a folha- β as principais estruturas (CHOUDHURI, 2014).

A estrutura terciária consiste na estrutura secundária em espaço tridimensional causado pelo enovelamento da cadeia e a estrutura quaternária se dá por uma estruturação

alcançada por proteínas com mais de uma cadeia polipeptídica onde cada cadeia é uma subunidade com estruturas específicas (CHOUDHURI, 2014).

A diversidade de funções moleculares realizada por proteínas sucede da sua estrutura tridimensional, determinada por sequências de aminoácidos codificadas geneticamente (KUHLMAN & BRADLEY, 2019).

1.5.2 Modelagem Molecular: método *threading*

Entende-se por modelagem molecular como um conjunto técnico e teórico que gera representações de estruturas moleculares e simula seu funcionamento (GU & LI., 2011). Atualmente a modelagem computacional é estabelecida pela união da bioinformática estrutural e a física molecular computacional (BRINI; SIMMERLING & DILL, 2020).

No que se refere à predição estrutural 3D de uma proteína, seus métodos são categorizados em três principais vias: métodos *ab initio*, *Threading* e Modelagem por Homologia/ Comparativa (KUMAR et al., 2019).

A predição de enovelamento de proteínas (*Threading*), consiste no reconhecimento de similaridade do modelo de estrutura enovelada (JONES; TAYLOR; & THORNTON, 1992) onde o alinhamento comparativo considera fatores como por exemplo: a preferência à acessibilidade ao solvente, preferência por uma estrutura secundária específica e interação entre segmentos vizinhos e conseqüentemente pontua esses alinhamentos (ALMEIDA, 2016).

1.6 Resistência bacteriana

O cobre pode acontecer em duas formas iônicas em um sistema biológico, sendo elas: Cu(I), dominante no ambiente intracelular de natureza redutora e Cu(II), dominante no meio extracelular de natureza oxidativa (PALUMAA, 2013).

De modo geral, no que se refere ao termo resistência bacteriana, diversos mecanismos estão relacionados, sendo esses mecanismos: alterações de permeabilidade na parede/membrana das células, restringindo o acesso antimicrobiano à sítios de ação; efluxo

ativo do antimicrobiano para o meio externo; mutações no sítio de ação; modificação ou degradação do antimicrobiano por enzimas e obtenção de caminhos metabólicos alternativos para os já inibidos (MCDERMOTT et al, 2003).

Os primeiros padrões de resistência e homeostase relacionada ao cobre em bactérias Gram-negativas foram oriundos do sistema de efluxo de cobre, que ocorre no cromossomo da *E. coli* e sistema *pco* da *E. coli* que ocorre no plasmídeo assim como o sistema *cop* de *Pseudomonas syringae* (COOKSEY, 1994) (RENSING & GRASS, 2003) (BONDARCZUK & PIOTROWSKA-SEGET, 2013).

A identificação de resistência ao cobre por uma bactéria em plantas teve seu primeiro caso apenas em 1980. A partir desta identificação, começaram a surgir inúmeros estudos voltados para a caracterização de cepas sensíveis ao cobre e os possíveis mecanismos por trás dessa resistência em bactérias causadoras de patógenos em plantas (MARQUES et al., 2009)

1.6.1 Operon *cop*

Em 1961, proposto por Jacob e Monod, o termo operon refere-se à genes controlados por um único promotor (JACOB; MONOD, 1961) sendo o operon *cop* e seus quatro genes (*copABCD*) de resistência ao cobre pela primeira vez identificados em 1986, por meio de estudos voltados à espécie *Pseudomonas syringae* (COOKSEY, 1994). No mesmo ano, o operon *cop* também foi identificado pela primeira vez no gênero *Xanthomonas*, espécie *X. campestris* (COOKSEY, 1994).

Designadas como proteínas do espaço periplasmático, a *copA* e a *copC* são codificadas pelo operon *cop* assim como a *copD*, que é considerada uma proteína da membrana interna e a *copB*, proteína localizada na membrana externa bacteriana (COOKSEY, 1994). Em estudos realizados em 1991 e 1992, foram atribuídas a essas proteínas as funções de sequestro e regulação de comportamento do cobre no espaço periplasmático e membrana externa (CHA; COOKSEY, 1991) (COOKSEY; AZAD, 1992).

2. OBJETIVOS

2.1 Objetivos gerais

Realizar uma comparação genômica da *Xanthomonas arboricola* pv. *juglandis* 417 com outras duas espécies filogeneticamente mais próximas, identificar os genes resistentes ao cobre na *Xaj* 417 por meio de análise genômica e expor possíveis fundamentos para a resistência do organismo ao cobre com o estudo *in silico*.

2.2 Objetivos específicos

- Coleta de dados do organismo *Xanthomonas arboricola* pv. *juglandis* 417.
- Visualização genômica da *Xanthomonas arboricola* pv. *juglandis* 417 utilizando o software CGView.
- Construção da árvore filogenética e diagrama de Venn no software EDGAR 3.0 para comparação genômica do organismo com outras duas espécies mais próximas.
- Alinhamento comparativo da *Xanthomonas arboricola* pv. *juglandis* 417 com as sequências das duas outras espécies mais próximas utilizando o software Mauve para encontro de similaridade genômica.
- Alinhamento comparativo entre *copBs* utilizando o software EMBOSS.
- Predição estrutural e funcional das *copBs* por meio do servidor C-I-TASSER.
- Alinhamento estrutural pelo software UCSF Chimera.

3. MATERIAL E MÉTODOS

3.1 Organismo e levantamento de dados

3.1.1 Organismo

O genoma da proteobacteria *Xanthomonas arboricola* pv. *juglandis* 417, foi o referencial escolhido para o desenvolvimento deste estudo.

3.1.2 Distribuição de dados

Inicialmente, foi utilizado o banco de dados do NCBI (National Center for Biotechnology Information) que fornece a qualquer um que o acesse, informações genômicas de qualquer organismo já estudado.

A sequência de nucleotídeos referente a *Xaj* 417, foi baixada em formato FASTA (LIPMAN & PEARSON, 1988) que é um algoritmo para comparação de sequências e depois submetida ao BLAST (Basic Local Alignment Search Tool) (ALTSCHUL et al, 1990) por meio do software Blast2GO com o objetivo de identificar locais de grande similaridade em alinhamentos com outros organismos através de sequências de nucleotídeos ou proteínas (FRANKEL;RICHTER, 1998).

3.2 Genômica Comparativa

Nesta etapa foi trabalhado o grau de relação entre o genoma da *X. juglandis* 417 com genomas de outras espécies pertencentes ao gênero *Xanthomonas*.

3.2.1 Visualização do genoma bacteriano

Circular Genome Visualization (CGView) (<http://cgview.ca/>) é uma ferramenta de comparação genômica que gera mapas de alta qualidade de genomas circulares como o de bactérias.

A sequência em formato FASTA foi submetida ao programa que por meio de BLAST, comparou a sequência genômica da *Xaj* 417 com outras sequências, convertendo os resultados em um mapa da sequência genômica completa.

3.2.2 Árvore Filogenética e Diagrama de Venn

O software EDGAR 3.0 (<https://edgar3.computational.bio.uni-giessen.de>), auxilia na análise comparativa entre genes combinada à representação visual desta comparação (BLOM et al, 2016).

Nesta etapa, foi originada pela ferramenta, uma árvore filogenética contendo o gênero *Xanthomonas* e conseqüentemente os dois organismos mais próximos à *Xanthomonas arboricola* pv. *juglandis* 417, por meio do método *neighbor-joining*, proposto por Saitou e Nei em 1987, onde o par de sequências que apresentar a menor distância é identificado, unido por um nó que representa o ancestral comum deste par de sequências e incorporado na árvore (VERLI, 2014).

Já na construção do diagrama de Venn, os genomas das duas espécies mais próximas filogeneticamente da *Xaj* 417, foram utilizados no cálculo da equiparação entre os genomas para representação visual pelo diagrama.

3.2.3 Alinhamento de sequências

Nesta etapa, com o uso do software Mauve (Multiple Genome Alignment) (<http://darlinglab.org/mauve/mauve.html>), o genoma da *Xaj* 417 juntamente com os genomas de outras duas *Xanthomonas* que se encontram próximas na árvore filogenética, foram submetidos a um alinhamento com o objetivo de identificar o grau de similaridade entre esses organismos além da visualização das sequências genômicas e regiões específicas das bactérias.

3.3 Genes referentes à resistência ao cobre

Mediante a distribuição de dados gerada pelo software BLAST2GO, os genes relacionados à resistência ao cobre foram manualmente filtrados para estudo conceitual dos

mesmos.

3.4 Análise comparativa estrutural e funcional das *copBs*

3.4.1 Análise estrutural

3.4.1.1 Alinhamento global de sequências

EMBOSS (The European Molecular Biology Open Software Suite) (https://www.ebi.ac.uk/Tools/psa/emboss_needle/), é um pacote de software que pode ser operado por meio de linha de comando no navegador web e que contém mais de 100 aplicações, sendo uma das principais o alinhamento global de sequências que utiliza o algoritmo Needleman-Wunsch (RICE; LONGDEN & BLEASBY, 2000).

As sequências de entrada das proteínas *copB* 119 e *copB* 3389, foram enviadas de modo respectivo e as opções de alinhamento global do software mantidas como padrão.

3.4.1.2 Predição estrutural e funcional

O C-I-TASSER (<https://zhanggroup.org/C-I-TASSER/>), servidor online utilizado nesta etapa do estudo, consiste em um método com alta precisão de predição de estrutura proteica e função, de modo que à partir de uma sequência submetida é gerada a estrutura e consequentemente funções biológicas relacionadas (ZHENG et al., 2021).

Anteriormente geradas pelo software BLAST2GO, as duas sequências submetidas (Fig. 5) (Fig. 6) ao servidor correspondem à proteína de resistência ao cobre: *copB*, que ocorre duas vezes durante a anotação genômica da *Xaj* 417.

***copB* 119** - 253 resíduos

Figura 5 - Sequência da *copB* 119 submetida ao servidor C-I-TASSER.

C-I-TASSER On-line Server (View example output):

Copy and paste your sequence below ([10, 750] residues in FASTA format). [Click here for a sample input:](#)

```
megmdhaamghsattseprtpipkvtdidraaavpppsdhpvhndntiqhyvlfnrleawnadpgtgqawegqs
wigtdlnrlwlrsegerndgrteaadlevlygrsvapwwdlvagirqdfkpgssqtyaafgvqglapmkfeva
atayigergqtaarfeaeeyelllnrwilqplaeielfgkddpargigsglstaelglrlryeftrqfapyig
lvyerafgntadlrrdeservddtrivvgirtwf
```

Fonte: C-I-TASSER.

***copB* 3389 - 212 resíduos**

Figura 6 - Sequência da *copB* 3389 submetida ao servidor C-I-TASSER.

C-I-TASSER On-line Server (View example output):

Copy and paste your sequence below ([10, 750] residues in FASTA format). [Click here for a sample input:](#)

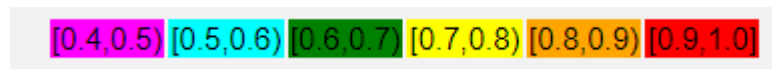
```
mhappginhyvllldrleafdtrrgsgqdwearawiggdtdrlwlrsegerqdgrtqaasvealyghaispwwdl
lvgarqdigadehrsaaafgvqglapykfeteatlyvgsgsraalrlegeyevlltnrlilqprveadialtd
ddrrngvsggleqvqaglrlyeitrrfapyigwvhsrsfgdtarraaiddepardsrfvagvriwf
```

Fonte: C-I-TASSER.

3.4.2 Anotação funcional: processo biológico

A predição foi baseada no sistema de pontuação Cscore^{GO} de predição de termos da Ontologia Genética. A pontuação em questão tem uma faixa de valores (Fig. 7) entre 0 e 1, que, quanto maior o valor, maior a confiabilidade da predição funcional, sendo o intervalo de 0.4 a 1.0 de confiabilidade codificado por cores (ZHENG et al., 2021).

Figura 7 - Pontuação de confiabilidade - Cscore^{GO}.



Fonte: C-I-TASSER.

3.4.3 Alinhamento estrutural das *copBs*

O sistema de modelagem molecular, UCSF Chimera, é um programa de visualização interativa e que analisa estrutura moleculares e aspectos relacionados (HUANG et al., 2014). Ao submeter as predições estruturais de ambas *copBs*, à ferramenta MatchMaker, que é uma extensão de alinhamento estrutural encontrada no programa, sobrepôs as estruturas submetidas (MENG et al., 2006).

A sobreposição pode ser baseada nos algoritmos de Needleman-Wunsch e Smith-Waterman, de alinhamento global e local ou ainda, baseada no score do alinhamento para assim identificar a melhor combinação entre cadeias (MENG et al., 2006).

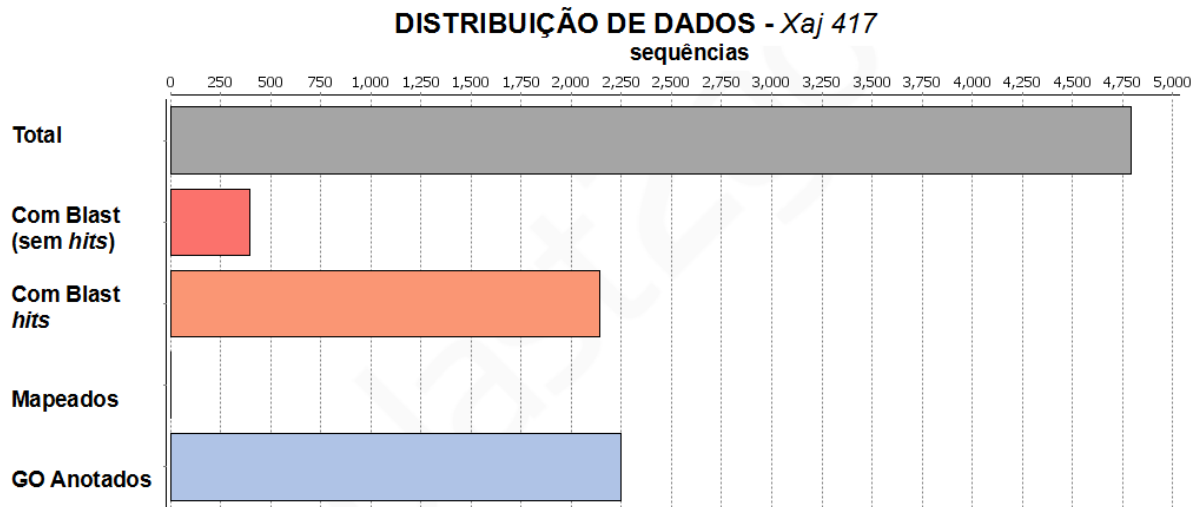
4. RESULTADOS E DISCUSSÃO

4.1 Organismo

Segundo o banco de dados usado como fonte, NCBI, o organismo *Xanthomonas arboricola* pv. *juglandis* 417, de nome infraespecífico: *Xaj* 417, foi submetido pela Universidade Federal de Londrina no ano de 2015, sendo o comprimento total da sequência nucleotídica de 5.218.943 pares de base. Seu código de acesso CP012251 atualmente encontra-se na versão CP012251.1.

4.2 Distribuição de dados

Figura 8 - - Distribuição de dados genômicos da *Xanthomonas arboricola* pv. *juglandis* 417.



Fonte: Software BLAST2GO.

Graficamente (Fig. 8), a distribuição de dados pelo software BLAST2GO apresentou um total de 4.791 sequências avaliadas, denominadas como “Total” e representadas pela barra horizontal de cor cinza. Partindo do valor total, 395 destas sequências não obtiveram alinhamentos identificados e foram classificadas como “Blast (sem hits)”, valor representado pela barra em cor vermelha. Já a barra de cor laranja e intitulada como “Blast hits” encontra-se 2.142 sequências alinhadas a outras espécies.

No processo de notação, cinco não conseguiram ser anotadas e portanto apenas mapeadas. Por fim, 2.249 sequências do valor total foram identificadas com genes funcionais através da análise de Gene Ontology (GO) feita pelo software. Análise essa que é subdividida em Função Molecular, Processo Biológico e Componente Celular (ASHBURNER, 2000).

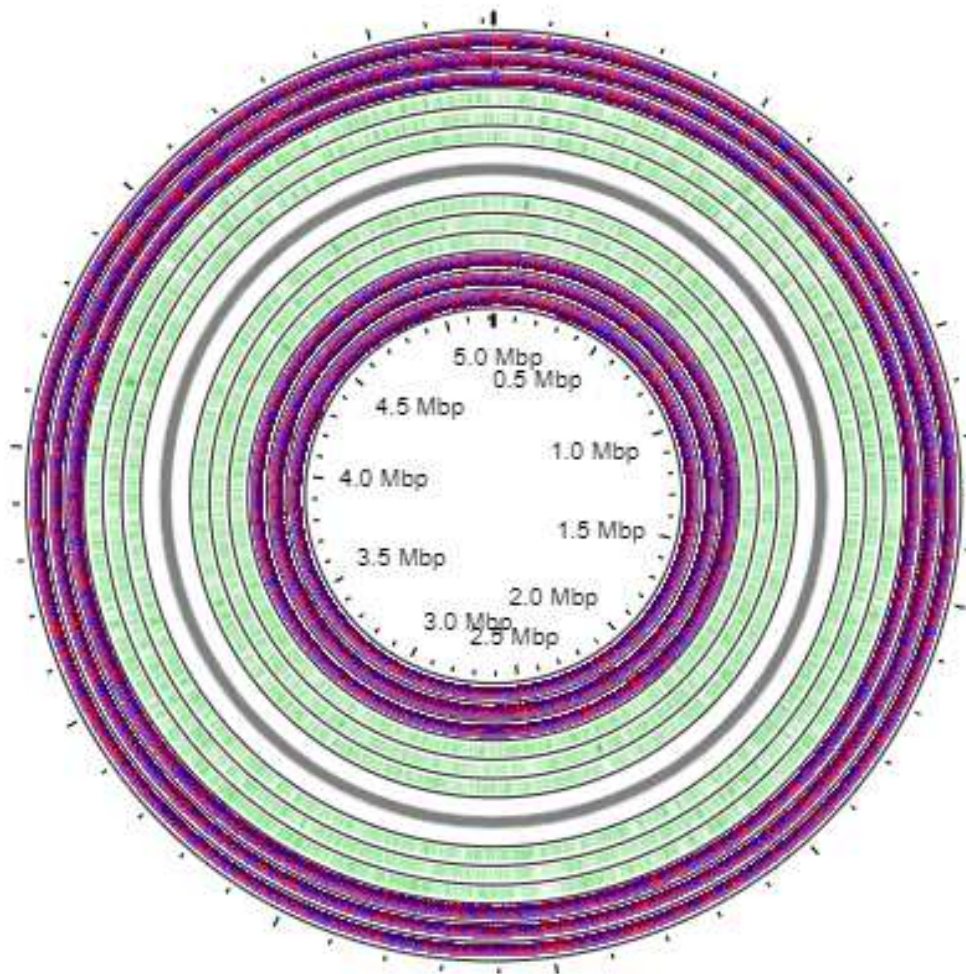
4.3 Genômica Comparativa

4.3.1 Visualização do genoma bacteriano

A informação requerida para geração do genoma bacteriano foi a sequência genômica do organismo, neste caso utilizada em formato FASTA. A interface do programa, então,

ofereceu opções que podem ser calculadas e onde foram escolhidas: códon de partida, códon de parada e open reading frames (ORFs), além de configurações voltadas à organização e aparência do mapa gerado.

Figura 9 - Representação circular do genoma bacteriano da *Xanthomonas arboricola* pv. *juglandis* 417.



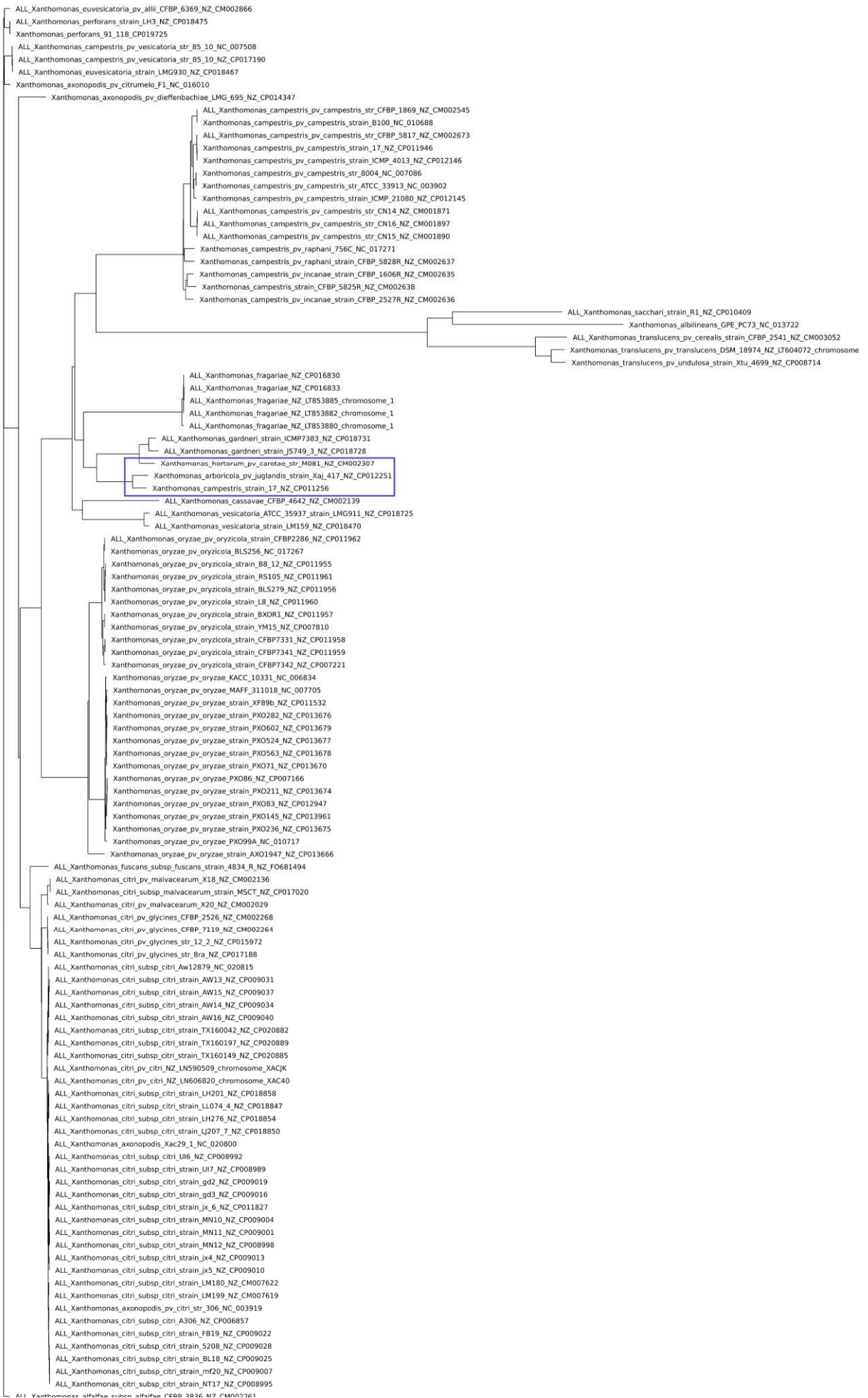
Fonte: Software CGView.

A representação gerada (Fig. 9) apresenta em azul e vermelho, respectivamente, os códons de partida e parada e na cor verde os ORFs.

4.3.2 Árvore Filogenética e Diagrama de Venn

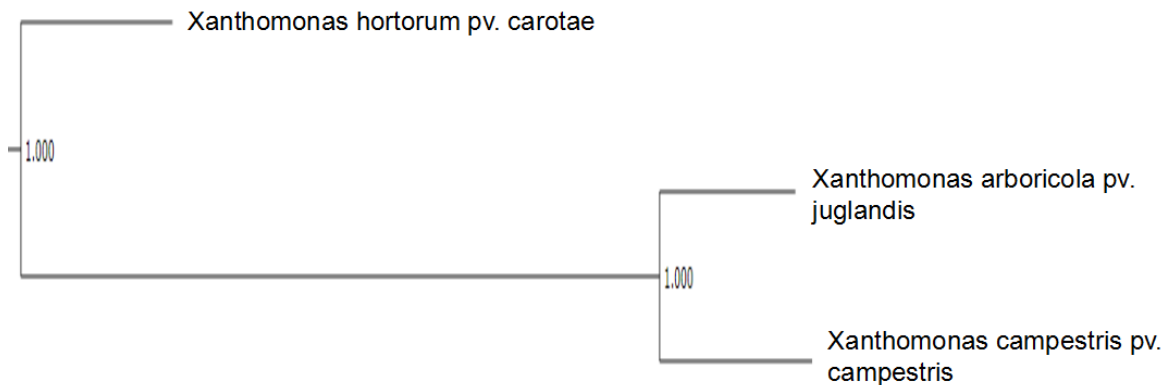
Selecionado o organismo, gênero *Xanthomonas*, o software EDGAR 3.0 utilizou sua própria ferramenta de alinhamento múltiplo onde foi calculado uma matriz de distância juntamente com o auxílio do algoritmo neighbor-joining, sendo essa a principal base para construção da árvore filogenética (Fig. 10). Segundo a descrição contida no software, a utilização do algoritmo neighbor-joining, minimiza o comprimento de ramos na árvore filogenética a partir da junção de nós adjacentes que contenham maior similaridade entre suas sequências. (Fig. 11).

Figura 10 - Árvore Filogenética projetada ao selecionar o gênero *Xanthomonas* e construída pelo método neighbor-joining.



Fonte: Software EDGAR 3.0.

Figura 11 - *Xanthomonas arboricola* pv. *juglandis* 417 e os dois organismos mais próximos filogeneticamente do referencial de estudo: *Xanthomonas campestris* pv. *campestris* e *Xanthomonas hortorum* pv. *carotae*.



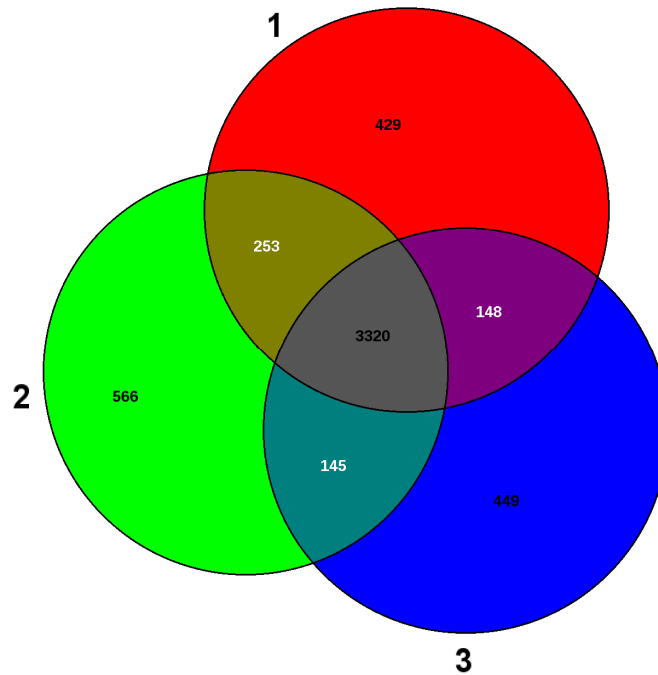
Fonte: Software EDGAR 3.0.

A *Xanthomonas hortorum* pv. *carotae* é responsável pela ferrugem bacteriana desenvolvida em culturas de cenouras, causando lesões amareladas em folhagens, caules e pecíolos (GILBERTSON, 2002). A maior problemática relacionada à patogenicidade é sua transmissão e associação à sementes, afetando a produtividade (UMESH; DAVIS; GILBERTSON, 1998).

Compartilhando o mesmo nó, o patógeno *Xanthomonas campestris* pv. *Campestris* é o principal causador da bacteriose comumente chamada de podridão negra que acomete o gênero *Brassica* que engloba hortaliças de importância econômica (WILLIAMS, 1980).

Posteriormente à geração da árvore filogenética (Fig. 10) e identificação dos dois organismos mais próximos da *Xaj* 417 (Fig. 11), foi projetado o diagrama de Venn das três espécies envolvidas na etapa de genômica comparativa utilizando suas sequências nucleotídicas.

Figura 12 - Diagrama de Venn: 1- *Xanthomonas arboricola* pv. *juglandis* 417. 2 - *Xanthomonas campestris* pv. *campestris*. 3 - *Xanthomonas hortorum* pv. *carotae*.



Fonte: Software EDGAR 3.0

O diagrama de Venn (Fig. 12) gerado pelo software apresentou uma equiparação genômica entre a *Xanthomonas arboricola* pv. *juglandis* 417 e as duas espécies mais próximas filogeneticamente: *Xanthomonas campestris* pv. *campestris* e *Xanthomonas hortorum* pv. *carotae*.

Por meio da representação foi possível notar maior registro de compatibilidade genômica entre *X. arboricola* e *X. campestris*, com um total de 253 sequências compatíveis, enquanto a compatibilidade entre a *X. arboricola* e *X. hortorum*, corresponde à apenas 148 sequências.

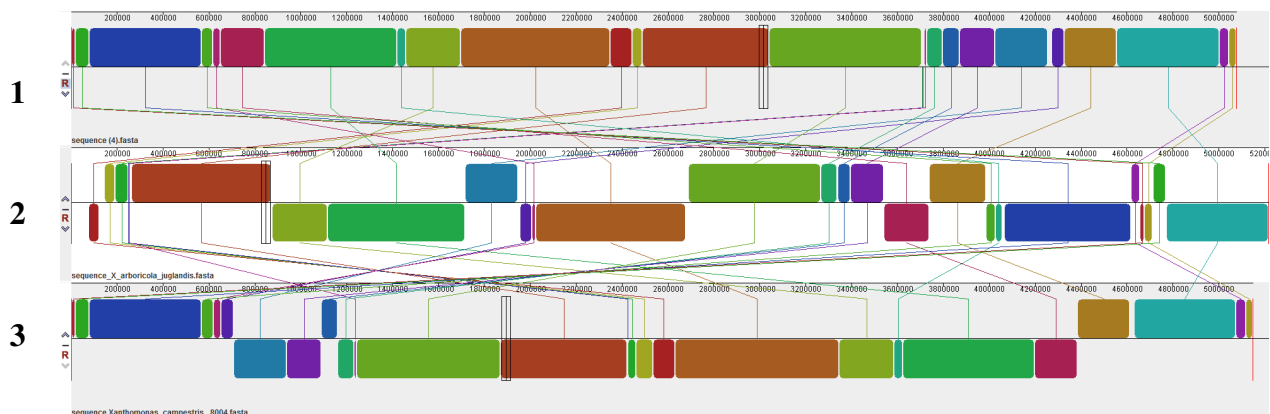
Um total de 3.320 sequências apresentam-se como comuns (*Core Genoma*) às três espécies e analisando apenas o referencial deste estudo, a *Xaj* 417, existem 429 sequências singulares à esse organismo. Em geral, genes específicos de uma cepa ocorrem por meio de

transferência horizontal entre espécies e sofrem mutações frequentes (LEFÉBURE & STANHOPE, 2007) (LAWRENCE & OCHMAN, 1997), possibilitando a adaptação à meios e condições específicas (LAPIERRE & GOGARTEN, 2009).

4.3.3 Alinhamento de sequências

Inicialmente foram fornecidos ao software Mauve as sequências em formato FASTA dos três organismos a serem comparados: *Xanthomonas arboricola* pv. *juglandis* 417, *Xanthomonas hortorum* pv. *carotae* e *Xanthomonas campestris* pv. *campestris*. O parâmetro de alinhamento utilizado pelo software é por padrão apropriado para o alinhamento de genomas e portanto não foram alterados.

Figura 13 - Alinhamento de sequências em formato de blocos colineares 1 - *Xanthomonas hortorum* pv. *carotae* 2 - *Xanthomonas arboricola* pv. *juglandis* 417 3 - *Xanthomonas campestris* pv. *campestris*.



Fonte: Software MAUVE.

Gerado o alinhamento (Fig. 13), foi possível observar um painel horizontal contendo as sequências genômicas das três espécies: *X.hortorum*, *X arboricola*, *X. campestris*, respectivamente, em uma escala mostrando as coordenadas da sequência.

Os blocos em diferentes colorações envolvem regiões das sequências que alinham-se às regiões das respectivas espécies em comparação, mostrando presumidamente a presença de homologia. Sendo assim, a presença de regiões homólogas entre as três espécies é apresentada

por meio das linhas interligando as respectivas regiões de coloração iguais. De acordo com o esquema padrão do software, o não preenchimento de blocos pela coloração indicaria baixa colinearidade entre as espécies analisadas.

4.4 Genes de Resistência

4.4.1 Genes referentes à resistência ao cobre

A anotação genômica da *Xaj* 417, realizada por meio do software BLAST2GO e baixada em arquivo FASTA, apresentou os seguintes genes relacionados à resistência ao cobre: *copA* (575 aa), *copB* 119 (253 aa) *copB* 3389 (212 aa), *copC* (127 aa) e *copD* (242 aa) e possíveis genes relacionados a funcionalidade da resistência: expressão de amina oxidase, expressão de transportador de cobre ABC ATPase, expressão de homeostase do cobre *cutC*, expressão de multicobre oxidase (MCO), expressão de regulador transcricional da família *copG* e expressão de chaperona de cobre PCu(A)C.

Segundo estudos de 1994, em cepas de *Pseudomonas syringae*, a *copA* é uma proteína periplasmática que liga múltiplos átomos de cobre, diferentemente da *copC* que mesmo sendo também uma proteína periplasmática liga apenas um átomo de cobre por polipeptídeo. A abundância dessas proteínas no periplasma de células induzidas por cobre sugere uma significativa contribuição na acumulação do mesmo (COOKSEY, 1994).

Sugere-se que as proteínas *copC* e *copD* não sejam exatamente necessárias no funcionamento da resistência, porém necessárias em uma máxima resistência ao cobre onde seu propósito está relacionado ao influxo de pequenas quantidades de cobre para auxílio do crescimento celular em casos onde as proteínas *copA* e *copB* já tenham alcançado o máximo de indução (BROWN et al., 1995). Apenas a presença da expressão de *copA* e *copB* em fitopatógenos do gênero *Xanthomonas* e *Pseudomonas* já viabiliza o desenvolvimento de resistência ao cobre nesses organismos (TEIXEIRA et al., 2008).

O gene *copB* é o codificador da proteína *copB* que está envolvida no funcionamento da resistência ao cobre (TEIXEIRA et al., 2008) e segundo estudos com a cepa PT23.2 de *P. syringae* pv. *tomato*, está fortemente associada à membrana externa (CHA; COOKSEY,

1991). Segundo Cooksey (1990), a *copB* possui região hidrofóbica na porção inicial que se assemelha à uma sequência de peptídeo sinal, enquanto a estrutura remanescente é predominantemente hidrofílica.

A anotação genômica da *Xaj* 417 apontou a duplicação dos seguintes genes parálogos: *copB* 119 e *copB* 3389. Na predição gênica da *Xaj* 417, por meio do programa FGENESB, a *copB* 119 apresentou ponto de localização entre as bases 13551 - 136312 em operon com a *copC* (gene 118) na fita negativa. Já a *copB* 3389 aparece localizada entre as bases 3704393 - 3705031 como unidade de transcrição e na fita positiva.

A duplicação de um gene consegue acelerar de forma significativa a evolução de material genético, tornando ausentes as restrições que possam impedir o desenvolvimento de novas funções (OHNO, 1970). A duplicação do gene *copB* pode estar associada à: Neofuncionalização (um dos genes duplicados evolui uma nova função), Subfuncionalização (funções ancestrais são divididas entre os genes duplicados) ou à Conservação de Genes (todas as funções são conservadas pelos duplicados) (HAHN, 2009).

Um gene duplicado só é mantido em um genoma se vantajoso para o organismo, ou seja, quando a ausência do gene causa decréscimo na aptidão do organismo (ZHANG, 2014) que neste caso refere-se à resistência ao cobre. Em circunstâncias de ausência de vantagem em perpetuar o gene, existiria a possibilidade de inativação do gene, por meio de mutações, reduzindo-o à um pseudogene e posteriormente a ocorrência de sua deleção (LYNCH & CONERY, 2000).

Um estudo voltado à investigação de duplicação de genes e retenção de parálogos em 106 genomas bacterianos afirma que genes envolvidos em constantes adaptações ao meio são preservados, salientando a importância da duplicação de genes na biologia evolutiva (GEVERS et al, 2004).

Aminas oxidases contendo cobre são consideradas ubíquas em organismos aeróbicos e são responsáveis pela catalisação da desaminação oxidativa de aminas primárias (CHANG et al., 2010). Os genes da homeostase de cobre da família *Cut*: *CutA*, *CutB*, *CutC*, *CutD*, *CutE* e

CutF, indicam envolvimento na absorção, armazenamento intracelular, transporte e efluxo de cobre (LI et al., 2010).

Em estudo voltado à *Xylella fastidiosa*, também patógeno de plantas economicamente importantes, o gene relacionado a homeostase de cobre, *CutC*, indica envolvimento na difusão de cátions (RODRIGUES et al., 2008) e em estudo referente a *Escherichia coli* e a identificação de tolerância ao cobre, a mutação ou deleção do gene *CutC* consequentemente aumentou a sensibilidade ao cobre quando em concentrações altas (GUPTA et al., 1995).

O transportador *ABC ATPase* consiste em quatro domínios associados à membrana onde dois destes domínios são altamente hidrofóbicos e cada um deles integra normalmente seis segmentos transmembranares. Consequentemente esses domínios configuram caminhos através da membrana e acredita-se que também determina a especificidade do substrato do transportador. Os outros dois domínios localizam-se de forma periférica na face citoplasmática da membrana, ligando ATP e acoplando hidrólise de ATP ao processo de transporte (HIGGINS, 1992).

As chaperonas periplasmáticas de cobre da família *PCu(A)C* estão presentes em várias bactérias, sendo sua função possivelmente relacionada à captura e transferência do cobre para outras chaperonas (THOMPSON et al., 2012). Juntamente com chaperonas, a expressão de multicobre oxidase configura uma das estratégias das bactérias em lidar com a toxicidade por metais, estratégia essa chamada de: armazenamento e desintoxicação (GILLET; LAWARÉE; MATROULE, 2019) a partir de ancoragens realizadas em três sítios de ligação de cobre: T1, T2 e T3 (GALLI; MUSCI; PATTI, 2004).

Visando estudos aprofundados sobre o funcionamento da resistência ao cobre em organismos causadores de patógenos em plantas economicamente importantes, testes laboratoriais, realizado por Cooksey e Lim, utilizaram por meio de clonagem, uma cepa sensível ao cobre de uma *Pseudomonas syringae* e conjugaram à uma patovar de *P. syringae*, também sensível ao cobre mas com níveis baixos quanto a sua frequência de mutações relacionados à resistência. Como consequência, o transconjugante obteve uma maior frequência de mutações espontâneas, indicando assim que na espécie em questão, a resistência

ao cobre pode evoluir por meio da modificação do *cop* cromossomal (COOKSEY, 1993).

O mesmo estudo laboratorial, realizado com *Xanthomonas campestris* não detectou forte homologia entre os genes de plasmídeos carreadores de resistência ao cobre e seu DNA cromossômico, sugerindo que os genes podem ter sido recebidos por *Pseudomonas* ou outra bactéria (COOKSEY, 1993). O estudo de relações evolutivas entre genes cromossômicos e o sistema de resistência de plasmídeos indicam um possível caminho no esclarecimento de transporte e gerenciamento de cobre (COOKSEY, 1993).

4.5 Análise comparativa estrutural e funcional da *copB* 119 e a *copB* 3389

4.5.1 Alinhamento global de sequências

Figura 14 - Alinhamento global das proteínas *copB* 119 e *copB* 3389 de resistência ao cobre.

```

# Length: 253
# Identity: 119/253 (47.0%)
# Similarity: 149/253 (58.9%)
# Gaps: 41/253 (16.2%)
# Score: 640.0
#
#
#=====
copB119 1 megmdhaamghsattseprtpipkvtdidraaavpppsdhvhdntiqhy 50
                                         :|...|.||
copB3389 1 -----mhapginhy 9

copB119 51 vlnrleawnadpgtgawegqswigtdlnrlwlrsegerndgrteaadl 100
      ||.:|||||:|...|:|.|.:.:|.|.|.:|.|.|.|.|.|.|.|.|.|.|.:
copB3389 10 vllrleafdtrrgsgdwearawiggdtdrlwlrsegerqdgrtqaasv 59

copB119 101 evlygrsvapwwdlvagirqdfkpgssqtyaafgvqglapmkfevaatay 150
      |.||||.:|:|.|.|.:|.|||.:.:|.|.|.|.|.|.|.|.|.|.|.|.|.|
copB3389 60 ealyghaispwwdlvgarqdigadehrsaaafgvqglapykfeteatly 109

copB119 151 igergqtaarfeaeeyellltnrwilqplaeielfgkddpargigsglsta 200
      :|...:.|.|.|.|||:|.|.|.|.|.|.|.|.|.|.:|...|.||:|.|.|...
copB3389 110 vsgsraalrlegeyevlltnrlilqprveadialtddrrrgvgsgleqv 159

copB119 201 elglrlryeftrqfapyiglvyerafgntadlrrdeservddtrivvgir 250
      |.|||||.|.|:|.|.|.|.:|.:|.|.|.|.|.|.|.|.|.|.|.|.|.|
copB3389 160 eaglrlryeitrrfapyigwvhsrsfgdtarraaiddepardsrfvagvr 209

copB119 251 twf 253
      .||
copB3389 210 iwf 212

#-----
#-----

```

Fonte: Software EMBOSS.

O alinhamento global (Fig. 14) entre as duas *copBs* obteve Identidade de resíduos igual a 47,0%, ou seja, 119 resíduos compartilhados de um total de 253 e que estão representados por barras verticais. A presença de homologia entre sequências é definida pela porcentagem de identidade maior ou igual a 30% (PEARSON, 2013).

A Similaridade de resíduos, que engloba resíduos idênticos e resíduos com presença de propriedades químicas semelhantes, apresentou maior valor: 58,9% ou 149 resíduos de um total de 253, representado no alinhamento por “:” e “.”.

Por fim, o número de *gaps* presentes no alinhamento corresponde à 41 de um total de

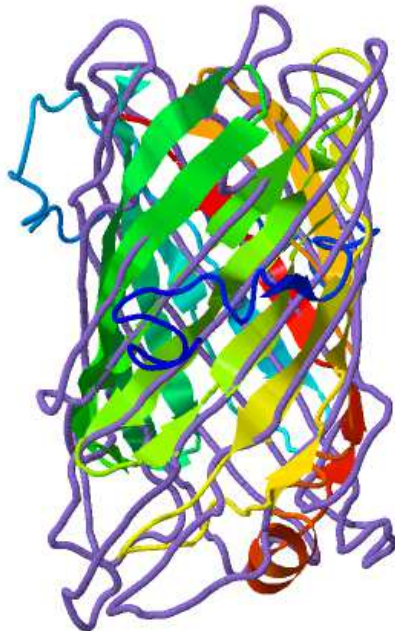
16,2%, representado por linhas horizontais. A presença de *gaps* no n-terminal indica mutações ocorridas na sequência da *copB* 3389, sendo essa mutação a deleção de aminoácidos.

4.5.2 Predição estrutural

As predições estruturais da *copB* 119 (Fig. 15) e *copB* 3389 (Fig. 16) foram realizadas pelo método de análise de agrupamentos (*clustering*), método esse usado pelo algoritmo SPIKER de identificação de estruturas conformacionais de proteínas com base em equiparações estruturais por similaridade (ZHANG & SKOLNICK, 2004a).

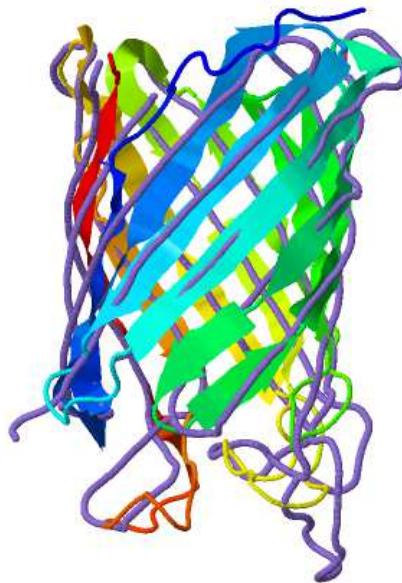
Os modelos de confiança são quantificados pelo C-score, responsável por estimar a qualidade dos modelos preditos em um intervalo de -5 a 2 onde quanto maior o valor, maior a confiança (ROY; KUCUKURAL & ZHANG, 2010). O valor de C-score correspondente à predição estrutural da *copB* 119 (Fig. 15) foi de -2.44 e o da *copB* 3389 (Fig. 16) foi de -0.47.

Figura 15 - Predição estrutural da proteína *copB* 119 de resistência ao cobre.



Fonte: C-I-TASSER.

Figura 16 - Predição estrutural da proteína *copB* 3389 de resistência ao cobre.



Fonte: C-I-TASSER.

Cada estrutura predita, quando comparada às estruturas análogas presentes no Protein Data Bank, gera um ranking com base em uma métrica de avaliação da similaridade de estruturas proteicas denominada TM-score. O alinhamento avaliado se dá entre a estrutura requerida e estruturas encontradas no Protein Data Bank e o valor encontrado é analisado com base no intervalo entre 0 e 1, sendo 1 o valor indicativo de combinação perfeita entre as duas estruturas. (ZHANG & SKOLNICK, 2004b).

Além desta métrica outro valor também atribuído é denominado como RMSD (Desvio Médio Quadrático), que avalia as posições atômicas e suas distâncias. Quanto menor o valor de RMSD maior a similaridade entre as estruturas (REVA; FINKELSTEIN & SKOLNICK, 1998).

Com base no ranking de estruturas similares à *copB* 119, a estrutura com maior valor de TM-score encontrada no PDB foi uma estrutura proteica cristalizada “COG4313” que está localizada na membrana, presente bactéria *Pseudomonas putida* F1 e que tem como sistema de expressão a *Escherichia coli* e é responsável pelo influxo de moléculas hidrofóbicas. Seu

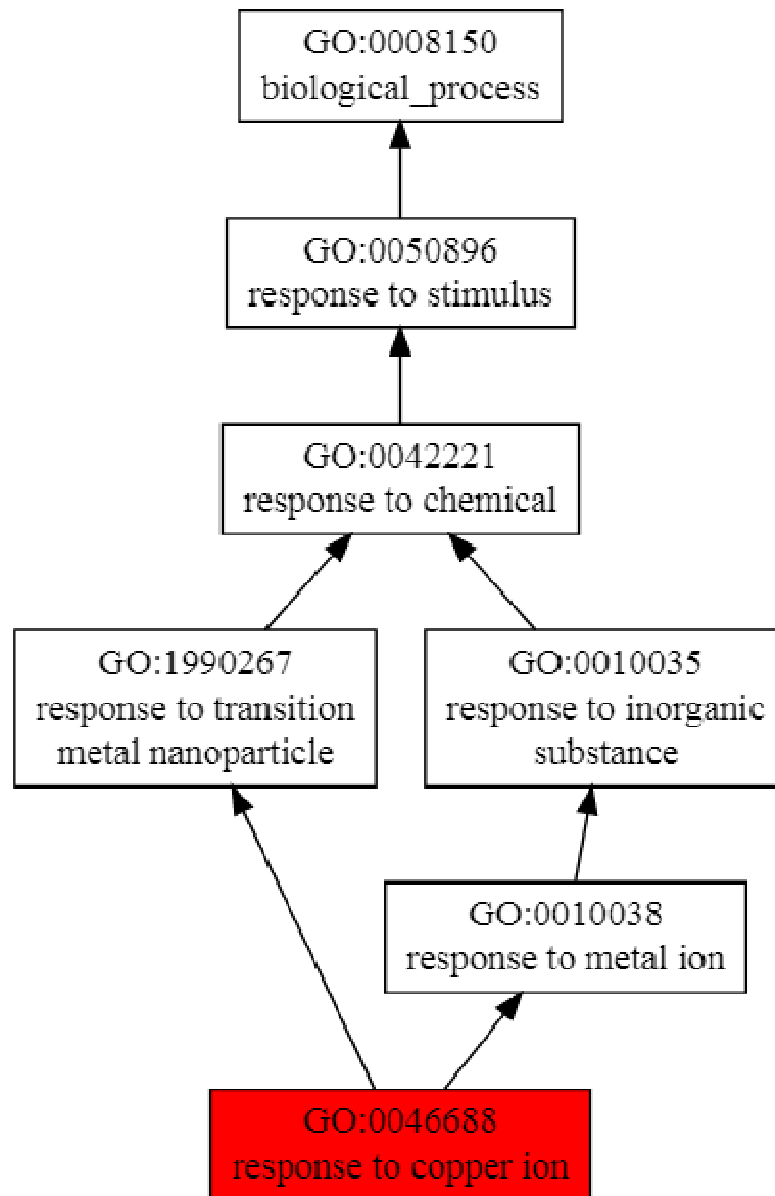
código de entrada no PDB é: 4RL8 (VAN DEN BERG et al., 2015) e um TM-score de: 0.66 e RMSD de: 3.34.

A estrutura análoga à *copB* 3389 com maior valor de TM-score foi a porina NanC, proteína essa de transporte encontrada na *Escherichia coli* K-12 e que tem como sistema de expressão a *Escherichia coli* BL21 (DE3). Seu código de entrada no PDB é: 2WJQ (WIRTH et al, 2009) e um TM-score de: 0.80 e RMSD de: 2:36.

4.5.3 Anotação funcional comparativa: Processo biológico

Baseado na ontologia genética e utilizando conceitos denominados *Go terms*, o C-I-TASSER gerou um diagrama hierárquico (Fig. 17 e 18) contendo termos basais e termos mais específicos relacionados ao processo biológico de cada *copB*.

Figura 17 - Diagrama da predição de processo biológico da proteína *copB* 119 de resistência ao cobre.

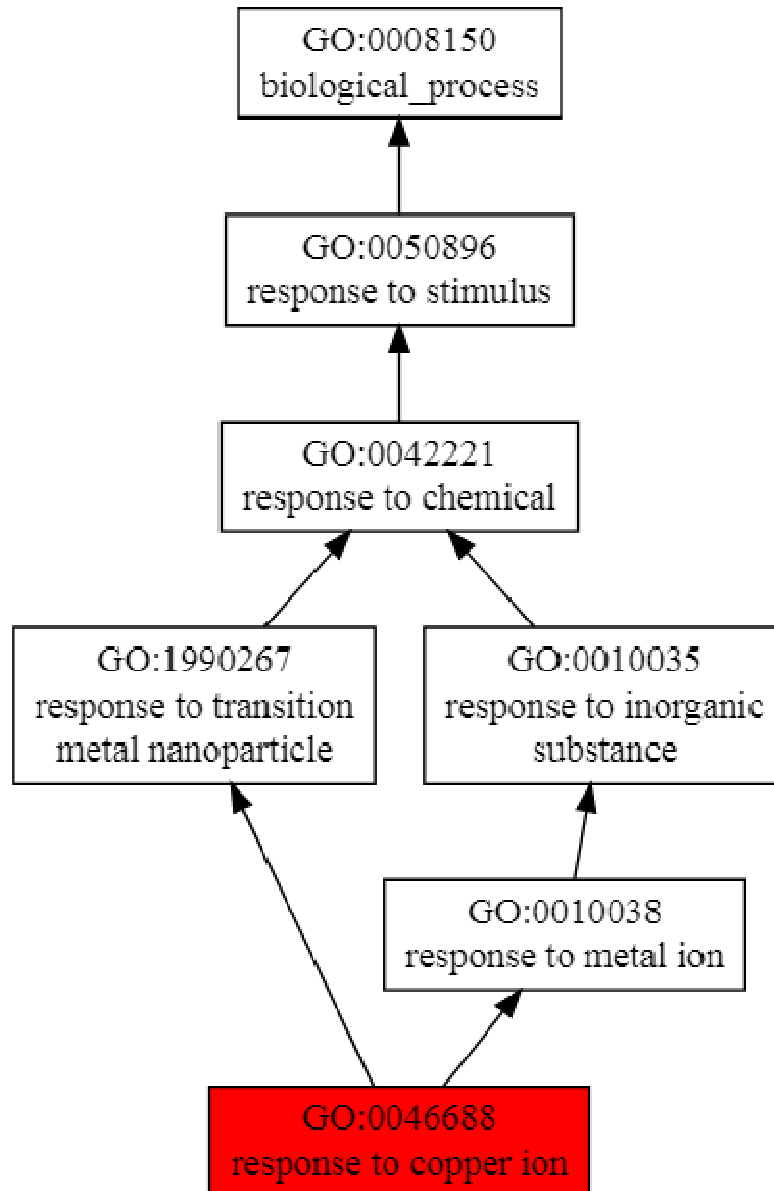


Fonte: C-I-TASSER.

A predição do processo biológico da *copB* 119 (Fig. 17) apresentou um diagrama onde a sensibilidade à íon de cobre apareceu como o único resultado com nível de confiabilidade, sendo este resultado com $Cscore^{GO}$ de 1.0, maior taxa de confiabilidade possível de ser encontrada. O termo GO em questão é definido como: processos que possam de alguma forma alterar o estado ou atividade da célula ou organismo (movimento, secreção, produção de

enzimas, expressão gênica, etc.) quando estimulado por íon de cobre (QUICKGO, 2021).

Figura 18 - Diagrama da predição de processo biológico da proteína *copB* 3389 de resistência ao cobre.



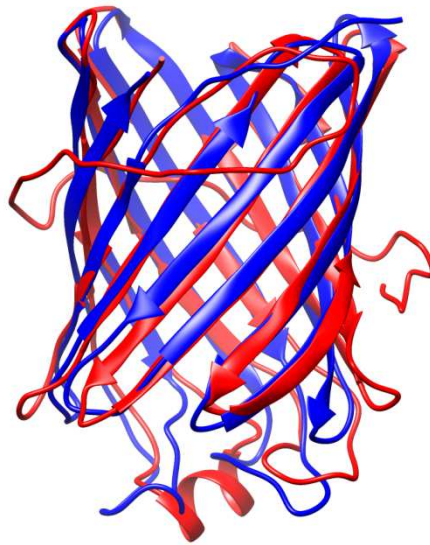
Fonte: C-I-TASSER.

Assim como a predição do processo biológico relacionado à *copB* 119, a predição da *copB* 3389 (Fig. 18) indica resposta à íon de cobre, sendo a taxa de confiabilidade também a mesma: 1.0. Sendo assim, o processo biológico das duas *copBs* submetidas para análise,

indicam alta taxa de paridade quando comparadas.

4.5.4 Alinhamento estrutural das *copB*s

Figura 19 - Alinhamento estrutural das proteínas *copB* 119 e *copB* 3389 de resistência ao cobre.



Fonte: Software Chimera.

O alinhamento estrutural (Fig. 19) indica em coloração vermelha a *copB* 119 e em azul a *copB* 3389. A partir do alinhamento foi possível identificar o n-terminal como não transmembranar. O desvio médio quadrático (RMSD) entre átomos foi de 3.22.

Dado o exposto, a *Xanthomonas arboricola* pv. *juglandis* 417 encontra na duplicação de genes parálogos como a *copB*, o que provavelmente pode estar associado à vantagens para evolução da resistência bacteriana, tendo em vista a grande diversidade adaptativa gerada como consequência da variedade funcional possível de ser encontrada em genes duplicados. (FARES et al., 2013).

5. CONCLUSÃO

A *Xanthomonas arboricola* pv. *juglandis* 417 possui uma sequência nucleotídica de 5.218.943 pares de base. Os organismos mais próximos filogeneticamente da *Xaj* 147 são as espécies *X. hortorum* pv. *carotae* e *X. campestris* pv. *campestris*. As espécies possuem alta taxa de colinearidade entre si.

O fitopatógeno *Xanthomonas arboricola* pv. *juglandis* 417, possui um sistema de resistência ao cobre que compreende a duplicação do gene de resistência: *copB*, que possui um processo biológico relacionado à sensibilidade à íon de cobre.

Houve constatação de deleções de aminoácidos no n-terminal da *copB* 3389, que obtém como estrutura análoga, a porina NanC, proteína essa de transporte. A estrutura identificada como análoga à *copB* 119 foi a proteína de membrana: “COG4313”. O alinhamento estrutural das *copBs* contém RMSD de 3.22, sendo o n-terminal não transmembranar.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALMEIDA, A. B. **Predição de Estrutura Terciária de Proteínas com Técnicas Multiobjetivo no Algoritmo de Monte Carlo**. Instituto de Informática, Universidade Federal de Goiás, 2016. Disponível em: <https://repositorio.bc.ufg.br/tede/bitstream/tede/5872/5/Disserta%C3%A7%C3%A3o%20-%20Alexandre%20Barbosa%20de%20Almeida%20-%202016.pdf>.
- ALTSCHUL, S.F et al. **Basic local alignment search tool**. Journal of Molecular Biology, 1990. DOI 10.1016/S0022-2836(05)80360-2.
- ANDREANI, N.A, FASOLATO, L. **Chapter 2 - Pseudomonas and Related Genera**. The Microbiological Quality of Food, 2017. p. 25-59. DOI 10.1016/B978-0-08-100502-6.00005-4.
- ASHBURNER, M.A et al. **Gene ontology: tool for the unification of biology**. The Gene Ontology Consortium. 2000. Nature Genetics. DOI 10.1038/75556.
- BERMAN, H et al. **The Protein Data Bank**. Nucleic Acids Research, 2000. v. 28, p. 235–242, DOI 10.1093/nar/28.1.235.
- BLOM, J et al. **EDGAR 2.0: an enhanced software platform for comparative gene content analyses**. Nucleic Acids Research, 20 abr. 2016. DOI 10.1093/nar/gkw255.
- BONDARCZUK, K. PIOTROWSKA-SEGET, Z. **Molecular basis of active copper resistance mechanisms in Gram-negative bacteria**. Cell Biology and Toxicology, 2013. v. 29 p. 397-405 DOI 10.1007/s10565-013-9262-1.
- BRINI, E; SIMMERLING, C; & DILL, K. **Protein storytelling through physics**. Science, 2020. v. 370, i 6520 DOI:10.1126/science.aaz3041.
- BROWN, N. L. et al. **Molecular genetics and transport analysis of the copper-resistance determinant (pco) from Escherichia coli plasmid pRJ1004**. Molecular Microbiology, 1995. v. 17, p. 1153-1166. DOI 10.1111/j.1365-2958.1995.mmi_17061153.x
- CABRAL, E. M. F. M. **Caracterização Molecular de Variedades de Nogueiras (*Juglans regia L.*) Portuguesas**: Estudo comparativo com cultivares internacionais Autenticidade de nozes comerciais e produtos derivados. 2008. 134 p. Dissertação de Mestrado (Mestre em Controle de Qualidade na área da Especialidade: Água e Alimentos) - Universidade do Porto, Porto, 2008. Disponível em: <https://repositorio->

aberto.up.pt/bitstream/10216/20783/2/DISSERTA%C3%83O.pdf.

CDA. **THE COPPER ADVANTAGE: A Guide to Working With Copper and Copper Alloys.** 2010. 27 p.

CHA, J.S; COOKSEY, D.A. **Copper resistance in *Pseudomonas syringae* mediated by periplasmic and outer membrane proteins.** Proceedings of the National Academy of Sciences of The United States of America. 1991. p. 8915- 8919 DOI 10.1073/pnas.88.20.8915.

CHANG, C.M et al. **Kinetic and Structural Analysis of Substrate Specificity in Two Copper Amine Oxidases from *Hansenula polymorpha*.** American Chemical Society, 2010. p. 2540 -2550. DOI 10.1021/bi901933d.

CHOUDHURI, S. **Bioinformatics for Beginners: Genes, Genomes, Molecular Evolution, Databases and Analytical Tools.** Elsevier, 2014. Primeira edição, p. 222, ISBN: 978-0-12-410471-6.

COOKSEY, D.A; AZAD, H.R. **Accumulation of copper and other metals by copper-resistant plant-pathogenic and saprophytic pseudomonads.** Applied and Environmental Microbiology. 1992. v 58, p. 274-278. DOI 10.1128/aem.58.1.274-278.1992.

COOKSEY, D. A. **Copper uptake and resistance in bacteria.** Molecular Microbiology, 1993. v. 7. p 1-5 DOI 10.1111/j.1365-2958.1993.tb01091.x.

COOKSEY, D. A. **Genetics of Bactericide Resistance in Plant Pathogenic Bacteria.** Annual Review of Phytopathology, 1990. v. 28, p. 201-219. DOI 10.1146/annurev.py.28.090190.001221.

COOKSEY, D. A. **Molecular mechanisms of copper resistance and accumulation in bacteria.** Microbiology Reviews, FEMS, ago. 1994, v. 14, p. 381-386. DOI 10.1111/j.1574-6976.1994.tb00112.x.

DANCHIN, A. **The Delphic Boat: What Genomes Tell Us.** Harvard University Press, 2002. 380 p. ISBN 978-0674009301.

EBRAHIMI, A. **Genetic diversity and genetic structure of Persian walnut (*Juglans regia*) accessions from 14 European, African, and Asian countries using SSR markers.** Tree Genetics & Genomes, 12 dez. 2016. DOI 10.1007/s11295-016-1075-y.

ELLINGSEN, D. G. et al. **Handbook of the Toxicology of Metals: Chapter 35 - Copper.** 2015, p. 765-786, v. 3. DOI 10.1016/B978-0-444-59453-2.00035-4.

EUFORGEN. *Juglans regia*: common walnut. Disponível em: <http://www.euforgen.org/species/juglans-regia/>

FARES, M. A. **The Roles of Whole-Genome and Small-Scale Duplications in the Functional Specialization of *Saccharomyces cerevisiae* Genes.** Plos Genetics, 2013. v. 13 p. DOI 10.1371/journal.pgen.1003176.

FERNANDES, C. et al. **Multiple DNA Markers for Identification of *Xanthomonas arboricola* pv. *juglandis* Isolates and its Direct Detection in Plant Samples.** Plant Disease, 20 mar. 2017. DOI 10.1094/PDIS-10-16-1481-RE.

FLORESTAR. **Nogueira (nogueira-comum)**, jan/dez 20---. Disponível em: <http://www.florestar.net/nogueira/nogueira.html>. Acesso em: 1 dez. 2019.

FRANKEL, M. RICHTER, Y. **Algorithms for Molecular Biology.** 1998. Disponível em: <http://www.cs.tau.ac.il/~rshamir/algmb/98/scribe/pdf/lec03.pdf>

GALLI, I. MUSCI, G. PATTI, M.C.B. **Sequential reconstitution of copper sites in the multicopper oxidase CueO.** Journal of Biological Inorganic Chemistry, 2004. v. 9 p 90-95. DOI 10.1007/s00775-003-0501-4.

GEVERS, D. et al., **Gene duplication and biased functional retention of paralogs in bacterial genomes.** Trends in Microbiology, 2004. v.12, p. 148-154. DOI 10.1016/j.tim.2004.02.007.

GILBERTSON, R. L. **Compendium Of Umbelliferous Crop Diseases: Bacterial leaf blight of Carrot.** American Phytopathological Society, 2002. p. 11-12.

GILLET, S; LAWARÉE, E; MATROULE J.-Y. **Functional Diversity of Bacterial Strategies to Cope With Metal Toxicity.** Microbial Diversity in the Genomic Era, 2019. p. 409–426. DOI 10.1016/b978-0-12-814849-5.00023-x.

GIOVANARDI, D. **Morphological and genotypic features of *Xanthomonas arboricola* pv. *juglandis* populations from walnut groves in Romagna region, Italy.** Eur J Plant Patho, [S. l.], p. 1-1, 27 out. 2015. DOI 10.1007/s10658-015-0809-2.

GUPTA, S. D et al. **Identification of cutC and cutF (nlpE) genes involved in copper tolerance in *Escherichia coli*.** Journal of Bacteriology, 1995. v. 177, p. 4207–4215 DOI 10.1128/jb.177.15.4207-4215.1995.

GU, Y; LI, M. **Chapter 3 - Molecular Modeling.** Handbook of Benzoxazine Resins, 2011. p. 103-110. DOI 10.1016/B978-0-444-53790-4.00048-5.

HAHN, M. W. **Distinguishing Among Evolutionary Models for the Maintenance of Gene Duplicates.** Journal of Heredity, 2009. v. 100, p. 605-617. DOI 10.1093/jhered/esp047.

HIGGINS, C.F. **ABC TRANSPORTERS: From Microorganisms to Man.** Annual Review of Cell Biology, 1992. v.8. p. 67-113. DOI 10.1146/annurev.cb.08.110192.000435.

HUANG, C.C. et al. **Enhancing UCSF Chimera through web services.** Nucleic Acids Research, 2014. DOI 10.1093/nar/gku377.

IPNI. **INTERNATIONAL PLANT NAMES INDEX (IPNI): JUGLANS REGIA L.** Disponível em: <https://www.ipni.org/n/442427-1>.

JACOB, F; MONOD, J. **Genetic regulatory mechanisms in the synthesis of proteins.** Journal of Molecular Biology, jun. 1961, v. 3, p. 318 - 356. DOI 10.1016/S0022-2836(61)80072-7

JONES, D; TAYLOR, W. & THORNTON, J. **A new approach to protein fold recognition.** Nature, 1992. v. 358, p. 86–89 DOI 10.1038/358086a0.

KATUZNA, M. **Xanthomonas arboricola pv. juglandis and pv. corylina: Brothers or distant relatives? Genetic clues, epidemiology, and insights for disease management.** Molecular Plant Pathology, 2021. p. 1-19. DOI 10.1111/mpp.13073.

KONDRASHOV, F. A. **Gene duplication as a mechanism of genomic adaptation to a changing environment.** Proceedings of The Royal Society B. 2012. v. 279, p. 5048-5057. DOI 10.1098/rspb.2012.1108.

KOUMAKIS, L. **Deep learning models in genomics; are we there yet?** Computational and Structural Biotechnology Journal, 2020, p. 1466–1473 DOI 10.1016/j.csbj.2020.06.017.

KUHLMAN, B. BRADLEY, P. **Advances in protein structure prediction and design.** Nature Reviews: Molecular Cell Biology, 2019. v. 20, p. 681-697. DOI 10.1038/s41580-019-0163-x.

KUMAR, P; HALDER, S. & BANSAL, M. **Biomolecular Structures: Prediction, Identification and Analyses.** Reference Module in Life Sciences, 2018. v. 3, p. 504-534. DOI 10.1016/b978-0-12-809633-8.20141-6

LAMICHHANE, J. R. et al. **Thirteen decades of antimicrobial copper compounds applied in agriculture. A review.** Agronomy for Sustainable Development, [S. l.], 24 maio 2018. DOI 10.1007/s13593-018-0503-9.

- LAPIERRE, P. & GOGARTEN, J. P. **Estimating the size of the bacterial pan-genome.** Trends in Genetics, 2009. DOI 10.1016/j.tig.2008.12.004.
- LAWRENCE, J. G. & OCHMAN, H. **Amelioration of bacterial genomes: rates of change and exchange.** Journal of Molecular Evolution, 1997. v. 44. DOI 10.1007/pl00006158.
- LECUN, Y. 1.1 **Deep Learning Hardware: Past, Present, and Future.** IEEE International Solid-State Circuits Conference, 2019. p. 12-19. DOI 10.1109/isscc.2019.8662396.
- LEFÉBURE, T & STANHOPE, M. J. **Evolution of the core and pan-genome of *Streptococcus*: positive selection, recombination, and genome composition.** Genome Biology, 2007. v. 8 DOI 10.1186/gb-2007-8-5-r71.
- LI, Y et al. **Crystal structure of human copper homeostasis protein CutC reveals a potential copper-binding site.** Journal of Structural Biology. 2010, p 399-405. DOI: 10.1016/j.jsb.2009.10.012.
- LYNCH, M. CONERY, J.S. **The Evolutionary Fate and Consequences of Duplicate Genes.** Science, 2000. v. 290, p. 1151-1155. DOI 10.1126/science.290.5494.1151.
- MARQUES, E. et al. **Sensitivity to copper in *Xanthomonas campestris* pv. *viticola*.** Tropical Plant Pathology, Brasília, v. 34, n. 6, p. 1-18, 28 nov. 2009. DOI 10.1590/S1982-56762009000600007.
- MCDERMOTT, PATRICK F. et al. **Antimicrobials: Modes of Action and Mechanisms of Resistance.** International Journal of Toxicology. Mar. 2003. DOI 10.1080/2F10915810305089.
- MENG, E.C. PETERSEN, E.F. COUCH, G.S. et al. **Tools for integrated sequence-structure analysis with UCSF Chimera.** BMC Bioinformatics, 2006 DOI 10.1186/1471-2105-7-339.
- MICHALAK, I; CHOJNACKA, K. **Encyclopedia of Toxicology: Biocides.** Elsevier, 2014. p. 461-463. DOI 10.1016/B978-0-12-386454-3.00472-3.
- MUKHOPADHYAY, C. S et al. **Basic Applied Bioinformatics.** Hoboken, NJ: Wiley, 2018.
- OHNO, S. **Evolution by Gene Duplication.** Springer, 1970. p.160. ISBN: 978-3-642-86659-3.
- PALUMAA, P. **Copper chaperones. The concept of conformational control in the metabolism of copper.** Febs Letters, 2013. v. 587. p. 1902 - 1910. DOI 10.1016/j.febslet.2013.05.019.

PANTH, Nisha et al. **Phytochemical profile and biological activity of Juglans regia**. Journal of Integrative Medicine, set. 2016, v. 14, p. 359-373. DOI 10.1016/S2095-4964(16)60274-1.

PEARSON, W. R. **An Introduction to Sequence Similarity (“Homology”) Searching**. Current Protocols in Bioinformatics, 2013. DOI 10.1002/0471250953.bi0301s42.

PEARSON, W. R; LIPMAN, David. L. **Improved tools for biological sequence comparison**. PNAS. Abril, 1988. DOI 10.1073/pnas.85.8.2444.

PEREIRA, U.P. et al. **Complete Genome Sequence of *Xanthomonas arboricola* pv. *juglandis* 417, a Copper-Resistant Strain Isolated from *Juglans regia* L.** Genome Announcements, 2015. v. 3. DOI 10.1128/genomeA.01126-15.

QUICKGO. **Gene Ontology and GO Annotations**. Disponível em: <https://www.ebi.ac.uk/QuickGO/term/GO:0046688>.

RENSING, C. GRASS, G. **Escherichia coli mechanisms of copper homeostasis in a changing environment**. FEMS Microbiology Reviews, 2003. v.27, p. 197-213. DOI 10.1016/S0168-6445(03)00049-4.

REVA, B.A. FINKELSTEIN, A.V. SKOLNICK, J. **What is the probability of a chance prediction of a protein structure with an rmsd of 6 Å?** Folding and Design, 1998. v. 3, p. 141-147. DOI 10.1016/S1359-0278(98)00019-4.

RICE, P., LONGDEN, I., & BLEASBY, A. **EMBOSS: The European Molecular Biology Open Software Suite**. Trends in Genetics, 2000, v. 16(6), p 276–277. DOI:10.1016/s0168-9525(00)02024-2.

RIGO, D. et al. **European Atlas of Forest Tree Species: Juglans regia in Europe: distribution, habitat, usage and threats**. [S. l.: s. n.], 2016. Disponível em: https://www.researchgate.net/publication/299470113_Juglans_regia_in_Europe_distribution_habitat_usage_and_threats.

RODRIGUES, C.M. et al. **Copper resistance of biofilm cells of the plant pathogen Xylella fastidiosa**. Applied Microbiology and Biotechnology, 2008. p. 1145-1157. DOI 10.1007/s00253-007-1232-1.

ROY, A. KUCUKURAL, A. ZHANG, Y. **I-TASSER: a unified platform for automated protein structure and function prediction**. Nature Protocols, 2010. p 725-738. DOI 10.1038/NPROT.2010.5.

RYAN, R.P et al., **Pathogenomics of *Xanthomonas*: understanding bacterium–plant interactions**. Nature Reviews Microbiology, 2011. v. 9, p. 344-355. DOI 10.1038/nrmicro2558.

SHARMA, A. GAUTAM, S. & WADHAWAN, S. *Xanthomonas*. Encyclopedia of Food Biology, 2014. p. 811-817. DOI 10.1016/B978-0-12-384730-0.00359-1.

SHEERMAN-CHASE, T. *Juglans regia L.* Disponível em: https://jb.utad.pt/especie/Juglans_regia.

SWINGS, J.G.; CIVEROLO, E.L. *XANTHOMONAS*. 1993.

TEIXEIRA, E.C et al. **The copper resistance operon *copAB* from *Xanthomonas axonopodis* pathovar citri: gene inactivation results in copper sensitivity**. Microbiology, 2008. v. 154 p. 402-412. DOI 10.1099/mic.0.2007/013821-0.

TORRISI, M. POLLASTRIG, G. & LEE, Q. **Deep learning methods in protein structure prediction**. Computational and Structural Biotechnology Journal, 2020. v. 18 p. 1301-1310. DOI 10.1016/j.csbj.2019.12.011.

THOMPSON, A.K. et al. **The roles of *Rhodobacter sphaeroides* copper chaperones PCuAC and Sco (PrrC) in the assembly of the copper centers of the aa3-type and the cbb3-type cytochrome c oxidases**. Biochimica et Biophysica Acta (BBA) - Bioenergetics, 2012. v. 1817, p 955-964. DOI 10.1016/j.bbabi.2012.01.003.

UC ANR STATEWIDE INTEGRATED PEST MANAGEMENT. **Walnut Blight**. [S. l.], 2017. Disponível em: <https://www2.ipm.ucanr.edu/agriculture/walnut/Walnut-Blight/#MANAGEMENT>

UMESH, K.C. DAVIS, R. M. & GILBERTSON, R. L. **Seed Contamination Thresholds for Development of Carrot Bacterial Blight Caused by *Xanthomonas campestris* pv. *carotae***. Plant Disease, 1998. p. 1271-1275. DOI 10.1094/PDIS.1998.82.11.1271.

VAN DEN BERG, B. et al. **Crystal structure of a COG4313 outer membrane channel**. Scientific Reports, 2015. v. 5. 10.1038/srep11927.

VERLI, H. **Bioinformática: da Biologia à Flexibilidade Molecular**. Primeira. ed. São Paulo: [s. n.], 2014. 282 p. ISBN 978-85-69288-00-8.

VINCENT, M et al. **Antimicrobial applications of copper**. International Journal of Hygiene and Environmental Health, Out, 2016, p. 585 - 591, v. 219, Elsevier. DOI 10.1016/j.ijheh.2016.06.003.

WILLIAMS, P. H. **Black Rot: A Continuing Threat to World Crucifers.** Plant Disease, 1980. American Phytopathological Society. DOI: 10.1094/PD-64-736.

WIRTH, C. et al. **NanC Crystal Structure, a Model for Outer-Membrane Channels of the Acidic Sugar-Specific KdgM Porin Family.** Journal of Molecular Biology, 2009. v. 394 p. 718-731. DOI 10.1016/j.jmb.2009.09.054.

ZHANG, J. **The Princeton Guide to Evolution.** Princeton University Press, 2014. p. 848.

ZHANG, Y. SKOLNICK, J. **Scoring function for automated assessment of protein structure template quality.** Proteins, 2004b. p 702-710. DOI 10.1002/prot.20264.

ZHANG, Y. SKOLNICK, J. **SPICKER: A Clustering Approach to Identify Near-Native Protein Folds.** Journal of Computational Chemistry, 2004a. p.865-871. DOI 10.1002/jcc.20011.

ZHENG, W. **Folding non-homologous proteins by coupling deep-learning contact maps with I-TASSER assembly simulations.** Cell Reports Methods, 2021 1: 100014.

ZOU, D. et al. **Biological Databases for Human Research.** Genomics, Proteomics and Bioinformatics, Feb. 2015, v. 13, p. 55-63. Elsevier. DOI 10.1016/j.gpb.2015.01.006.