

DENIZARD ALLISON SANTOS BUENO

**ANÁLISE GENÔMICA DA RESISTÊNCIA DO TOMATEIRO (*Solanum* spp.) À
REQUEIMA (*Phytophthora infestans* Mont.)**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

Orientador: Derly José Henriques da Silva

Coorientador: Guilherme da Silva Pereira

**VIÇOSA – MINAS GERAIS
2021**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

B928a
2021 Bueno, Denizard Allison Santos, 1989-
Análise genômica da resistência do tomate (*Solanum* spp.) à
requeima (*Phytophthora infestans* Mont.) / Denizard Allison
Santos Bueno. – Viçosa, MG, 2021.

1 dissertação eletrônica (49 f.): il. (algumas color.).

Orientador: Derly José Henriques da Silva.

Dissertação (mestrado) - Universidade Federal de Viçosa,
Departamento de Fitotecnia, 2021.

Inclui bibliografia.

DOI: <https://doi.org/10.47328/ufvbbt.2022.111>

Modo de acesso: World Wide Web.

1. Tomate - Resistência a doenças e pragas - Aspectos
genéticos. 2. Genômica. 3. Tomate - Recursos do germoplasma.
4. *Solanum lycopersicum*. 5. Requeima. 6. *Phytophthora
infestans*. I. Silva, Derly José Henriques da, 1961-.
II. Universidade Federal de Viçosa. Departamento de Fitotecnia.
Programa de Pós-Graduação em Genética e Melhoramento.
III. Título.

CDD 22. ed. 635.6422

DENIZARD ALLISON SANTOS BUENO

**ANÁLISE GENÔMICA DA RESISTÊNCIA DO TOMATEIRO (*Solanum spp.*) À
REQUEIMA (*Phytophthora infestans* Mont.)**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

APROVADA: 02 de agosto de 2021.

Assentimento:



Denizard Allison Santos Bueno
Autor



Derly José Henriques da Silva
Orientador

Derly José Henriques da Silva
Prof. Titular do Depto. Agronomia
Matricula UFV: 6824-1
DAA/CCA/UFV

AGRADECIMENTOS

Agradeço à minha família, meu pai Sérgio Valentim Bueno (*in memoriam*), mãe Cibele de Maria de Oliveira Santos, meus irmãos, Christian e Thamires, minha irmã de outra mãe Thaís Dal’Sasso, meus sogros, Mauro Godoi e Maria Edinauza, e em especial à minha esposa Edna Thaís, por todo apoio.

À Universidade Federal de Viçosa, por me acolher desde a graduação até a pós-graduação.

Aos colegas do laboratório de Manejo de Recursos Genéticos pela boa convivência e pelo conhecimento compartilhado.

Ao meu Orientador pela paciência e conselhos ao longo dessa jornada.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES), Fundação de Apoio e Amparo a Pesquisa do Estado de Minas Gerais (FAPEMIG) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

“Eu questiono porque não basta ser.”
(Pedro Paulo Soares Pereira)

RESUMO

BUENO, Denizard Allison Santos, M.Sc., Universidade Federal de Viçosa, agosto de 2021. **Análise genômica da resistência do tomateiro (*Solanum* spp.) à requeima (*Phytophthora infestans* Mont.)**. Orientador: Derly José Henriques da Silva. Coorientador: Guilherme da Silva Pereira.

O tomateiro, *Solanum lycopersicum*, devido a sua estrutura, é uma planta bastante susceptível a doenças e estresses abióticos. Dentre essas doenças, a requeima (*Phytophthora infestans*) se destaca pelo seu progresso agressivo e sua capacidade destrutiva. Há grande esforço em elucidar os mecanismos genéticos no tomateiro e alguns softwares de análise *in silico* oferecem ferramentas específicas a *S. lycopersicum*. Nesse trabalho visamos identificar polimorfismos associados à resistência à requeima em dois acessos de tomate, BGH-2127 (*S. lycopersicum*) e BGH-6902 (*S. habrochaites*) obtidos no Banco de Germoplasma de Hortaliças (BGH) da UFV. Esses acessos do BGH-UFV foram sequenciados juntamente com o acesso NC-2-CELBR e a cultivar Santa Clara, que serviram de controle resistente e susceptível, respectivamente. Dados obtidos no *Solgenomics* forneceram base para interpretação das leituras obtidas dos sequenciamentos, assim como funções gênicas caracterizadas pelo ITAG 4.0. Sequências de outros acessos foram obtidos na plataforma EBI-EMBL para o teste de associação genômica. BLASTn realizado utilizando a sequência dos genes de resistência *Ph-2* e *Ph-3* indicaram cópias de ambos os genes nos acessos do BGH-UFV, análise filogenética desses genes mostraram cópias mais conservadas do gene *Ph-3* em relação ao *Ph-2*. Além disso, polimorfismos foram encontrados em ambos os genes quando comparados ao controle susceptível. Houve uma maior similaridade do acesso BGH-6902 e NC-2-CELBR em relação ao gene *Ph-2*, e ambos os acessos do BGH apresentaram polimorfismos no gene *Ph-3*, semelhantes aos encontrados no NC-2-CELBR. Esses polimorfismos foram suficientes para geração de marcadores para futura validação. O teste de associação correlacionou 5 SNPs com as variações fenotípicas. O gráfico QQplot evidenciou que os p-valores do teste de associação não seguiram distribuição normal, sendo indispensável futuras elucidaciones.

Palavras-chave: Genômica comparativa. Recursos genéticos. *Solanum lycopersicum*. Requeima. Resistência à *Phytophthora infestans*.

ABSTRACT

BUENO, Denizard Allison Santos, M.Sc., Universidade Federal de Viçosa, August, 2021. **Genomic analysis of tomato (*Solanum* spp.) resistance to late blight (*Phytophthora infestans* Mont.)**. Advisor: Derly José Henriques da Silva. Co-advisor: Guilherme da Silva Pereira.

Tomato, *Solanum lycopersicum*, is the second most produced vegetable in the world, and, due to its traits, it is very susceptible to diseases and abiotic stresses. Among these diseases, late blight (*Phytophthora infestans*) stands out due to its rapid progression and its ability to increase losses. Our goal was to identify polymorphisms associated with late blight resistance in two tomato accessions, BGH-2127 (*S. lycopersicum*) and BGH-6902 (*S. habrochaites*) acquired from the Vegetable Germplasm Bank (BGH) at UFV and compare with NC-2-CELBR accession and the Santa Clara cultivar, which were resistant and susceptible control, respectively. Data obtained from Solgenomics database were basis to manipulate the raw reads, and to characterize gene functions annotated by ITAG 4.0. Another 84 accessions raw reads were downloaded from the EBI-EMBL platform for the genome-wide association study (GWAS). The BLASTn performed, using the Ph-2 and Ph-3 resistance gene sequences, and copies of both genes were found in the accessions. Both genes were identified as CDS or CC-NBS-LRR proteins, which are described in resistance pathways on several plant models. Phylogenetic analysis of these genes showed more conserved copies of the Ph-3 gene in comparison to Ph-2. Polymorphisms were found in both genes when compared to the susceptible control. There was a greater similarity of the BGH-6902 and NC-2-CELBR accessions in relation to the Ph-2 gene, and both BGH accessions presented polymorphisms in the Ph-3 gene, like those found in the NC-2-CELBR. These polymorphisms can be targeted for future marker validation. The genomic association test correlated 5 SNPs with phenotypic variations, however, none of the SNPs matched coding regions in the tomato genome. The QQplot graph showed that the GWAS p-values did not fit a normal distribution, probably due to population sampling, and further elucidation is essential.

Keywords: Comparative genomics. Genetic resources. *Solanum lycopersicum*. Late blight. *Phytophthora infestans* resistance.

SUMÁRIO

1	INTRODUÇÃO	9
2	REVISÃO DE LITERATURA.....	10
2.1	Banco de Germoplasma de Hortaliças da UFV	10
2.2	Requeima (<i>Phytophthora infestans</i>)	11
2.2.1	História da doença	11
2.2.2	Ciclo	12
2.2.3	Controle da doença.....	13
2.3	Resistência do tomateiro à requeima	14
2.4	Estudo de associação genômica ampla (GWAS)	15
2.5	Genômica comparativa	16
3	MATERIAL E MÉTODOS.....	20
3.1	Recursos genéticos	20
3.2	Análise Genômica.....	23
3.2.1	Sequenciamento de DNA.....	24
3.2.2	Montagem do genoma por referência	24
3.2.3	Variantes	24
3.2.4	Análise de componentes principais.....	24
3.2.5	Estudo de associação genômica ampla (GWAS).....	25
3.2.6	Predição gênica	25
3.2.7	Filogenia.....	25
3.2.8	Marcadores Moleculares	26
4	RESULTADOS.....	26
4.1	Montagem por referência.....	26
4.2	Análise de componentes principais (PCA).....	30

4.3 GWAS.....	31
4.4 Filogenia	32
4.4.1 Gene <i>Ph-2</i>	32
4.4.2 Gene <i>Ph-3</i>	33
4.5 Polimorfismos	34
4.5.1 Gene <i>Ph-2</i>	34
4.5.2 Gene <i>Ph-3</i>	38
5 DISCUSSÃO	40
5.1 Gene <i>Ph-2</i>	40
5.2 Gene <i>Ph-3</i>	41
5.3 Estudo de associação genômica ampla	42
6 CONCLUSÕES.....	42
REFERÊNCIAS BIBLIOGRÁFICAS	43

1 INTRODUÇÃO

O tomate (*Solanum lycopersicum*) é a segunda hortaliça mais consumida mundialmente, ficando atrás apenas da batata. Anualmente, são plantados em torno de 5 milhões de ha, que produzem em torno de 180 milhões de toneladas da fruta (FAO Stat, 2016).

Atualmente, existem pouco mais de 4,4 milhões de propriedades rurais no Brasil enquadradas como familiares, representando quase 80% da totalidade. Apesar desse número não representar a distribuição de terras, a produção interna depende em peso desse modelo de propriedade. Esses agricultores são responsáveis por produzir cerca de 70% do feijão nacional, 34% do arroz, 87% da mandioca, e 45% da produção de hortaliças (IBGE Sidra, 2021). A tomaticultura, então, tem grande importância social e faz parte da renda de inúmeras famílias.

O tomateiro é uma cultura de hábito rasteiro e crescimento rápido, cujo ciclo leva de 90 a 150 dias. Temperaturas ótimas para o desenvolvimento da cultura variam entre 18 e 25 °C durante o dia e 10 a 20 °C à noite. Oscilações bruscas de temperaturas diurnas e noturnas notavelmente afetam o desempenho da cultura. O tomateiro é uma planta extremamente sensível a todo tipo de estresse, biótico e abiótico. Por isso, é exigente em tratamentos culturais, principalmente em plantios de tomate mesa, o que acaba implicando em um uso intensivo de defensivos agrícolas.

São poucos os produtos à base de tomate, porém seu uso é bastante diversificado. Do consumo *in natura* em saladas ao uso em preparações na forma de molho, ketchup ou purê de tomate, o tomate é utilizado em incontáveis receitas. As quantidades de antioxidantes, carotenoides, licopeno, vitaminas A e E presentes na fruta tornam-na um alimento funcional (SCHIAVON et al., 2013). Portanto, é importante que exista cautela na aplicação de defensivos agrícolas, para que haja redução da contaminação residual no produto pós-colheita, e, por consequência, menor contaminação ambiental e humana.

Nesse contexto, o controle genético de doenças se destaca. Esse método já é usado amplamente no comércio de sementes, e é a prática menos impactante para o ambiente e manejo do agricultor (QI et al., 2019). Para tal, é imprescindível que haja

compreensão dos processos de resistência e dos elementos responsáveis pela expressão gênica, visando a facilitar a introgressão desses genes de resistência em cultivares elite. Além disso, entender as variadas estratégias de defesa utilizadas pelas plantas pode levar ao desenvolvimento de novos métodos de controle e combate dos fitopatógenos.

A requeima (*Phytophthora infestans*) destaca-se entre os fitopatógenos que afetam o tomate, devido sua rápida progressão e sua grande capacidade de causar perdas. Sete fontes de resistência à requeima já foram descritas, variando quanto ao mecanismo, número de genes e nível de resistência. Dois deles, o gene Ph-2 e Ph-3, foram inseridos com sucesso no acesso NC-2-CELBR (GARDNER; PANTHEE, 2010). O tomate serve como modelo de pesquisa em muitas áreas, e há um grande esforço para elucidar os mecanismos de transcrição e tradução em tomate, razão pela qual a maioria das análises *in silico* oferece ferramentas específicas para *S. lycopersicum* (MUELLER et al., 2009; SCARANO; RAO; CORRADO, 2017). Assim, nosso objetivo foi identificar polimorfismos associados à resistência à requeima em dois acessos de tomateiro, BGH-2127 (*S. lycopersicum*) e BGH-6902 (*S. habrochaites*) adquiridos do Banco de Germoplasma Vegetal (BGH) da UFV para futuramente incorporá-los em cultivares de elite.

2 REVISÃO DE LITERATURA

2.1 Banco de Germoplasma de Hortaliças da UFV

O Banco de Germoplasma de Hortaliças da Universidade Federal de Viçosa (BGH-UFV) foi criado em 1966 com o apoio da Fundação Rockefeller. O BGH-UFV tem como objetivo resgatar o patrimônio genético brasileiro, preservar, documentar e promover o acesso a esse germoplasma no Brasil.

O BGH-UFV tem em sua coleção mais de 100 espécies e 7.000 subamostras de hortaliças, provenientes de coletas e doações dos mais diversos biomas. Cada acesso recebido é avaliado e caracterizado de acordo com as práticas sugeridas para cada espécie, e então é colocado à disposição da comunidade científica brasileira. Variedades desenvolvidas pela UFV, variedades crioulas criadas por produtores da

região em espécies comerciais fazem parte do banco, além de espécies silvestres de gêneros de interesse no melhoramento de hortaliças.

Os recursos genéticos presentes no BGH-UFV representam grande porção da variabilidade genética de hortaliças. Por isso, o BGH-UFV é estratégico para o cenário do melhoramento de hortaliças no Brasil, visando a redução da erosão genética e aumento da capacidade adaptativa das cultivares atualmente utilizadas.

Assim, o objetivo de um banco de germoplasma é possuir um grande acervo de genes de interesse, em vez de genótipos. Genótipos disponíveis aos bancos de germoplasma podem não ser competitivos em relação a variedades comerciais. Genes de interesse, no entanto, podem ser introduzidos em variedades comerciais, conferindo a elas características desejáveis de adaptação, como resistência a pragas, doenças, estresses abióticos, variações no porte, ciclo, entre outras características (BGH - Universidade Federal de Viçosa, 2009).

A Universidade Federal de Viçosa, por meio do curador do BGH-UFV, professor Derly José Henriques Silva, dispõe do direito de intercambiar esse patrimônio genético com instituições de pesquisa, públicas ou privadas.

2.2 Requeima (*Phytophthora infestans*)

2.2.1 História da doença

Phytophthora infestans (Mont.) de Bary, causador da requeima em batata e tomate, é um fungo fitopatogênico da classe *Oomycetes*. Esse patógeno tem um papel central na fitopatologia, já que a requeima serve de modelo para estudos epidemiológicos, manejo, e de interação planta-patógeno. Os primeiros relatos da doença ocorreram em 1843 nas cidades de Filadélfia e Nova York. Dois anos depois, a doença já havia se espalhado pela costa oeste dos Estados Unidos mostrando seu grande potencial de dispersão via aérea (READER, 2008). Nesse mesmo ano um carregamento com frutos de batatas-semente infectadas com o fungo cruzou o Atlântico em direção a Europa causando a Grande fome de 1845, matando em torno de 1 milhão de irlandeses (VANHAUTE; PAPING; Ó GRÁDA, 2006). Ainda hoje, *P. infestans* é o principal patógeno das lavouras de batata e tomate ao redor do mundo,

com perdas estimadas em mais de 5 bilhões de dólares anualmente (HAVERKORT et al., 2009).

2.2.2 Ciclo

A maior parte das infecções de *P. infestans* são iniciadas pelo esporângio assexuado. Esses esporângios podem germinar diretamente ou indiretamente, dividindo seu citoplasma multinucleado em zoósporos mononucleados. Os esporângios entram em contato com o hospedeiro e germinam estendendo seu tubo germinativo até um estômato, gema, ferimentos ou outro tipo de abertura natural, podendo também penetrar diretamente através da cutícula utilizando o apressório. Após a entrada do patógeno no apoplasto da planta, o micélio do oomiceto começa a produzir um haustório procurando o espaço intracelular separado pela membrana do hospedeiro. Esse crescimento apoplástico ajuda reduzir a detecção por parte da planta. O dano celular à planta, causado pelo patógeno, pode desencadear uma resposta de defesa, que é ativada pela detecção de moléculas da própria planta oriundas desse dano, também conhecidos como DAMPs (damage-associated molecular patterns). Uma vez que atravessa a membrana celular da planta, o haustório começa a produzir proteínas efetoras (BIRCH et al., 2006) que podem ser reconhecidas por plantas não hospedeiras cessando o processo de infecção ou não reconhecidas pelas plantas susceptíveis onde é perceptível os sintomas característicos da requeima.

Phytophthora infestans é um patógeno hemibiotrófico e pode extrair nutrientes tanto do apoplasto como do meio intracelular, utilizando o haustório. Porém, ele infecta somente tecido vivo, assim a fase terminal de infecção é quando ocorre a esporulação e necrose do hospedeiro.

A esporulação pode ocorrer desde o terceiro dia de infecção, a depender do ambiente e da interação genética entre o patógeno e o hospedeiro. Cada lesão chega a produzir centenas de milhares de esporângios e cada um tem o potencial de gerar novas lesões iniciando novamente o ciclo e potencialmente destruindo campos inteiros em poucas semanas (FRY et al., 2015). A germinação dos esporos de *P. infestans* é inibida com a radiação solar, por isso a esporulação ocorre principalmente à noite (BASHI; BEN-JOSEPH; ROTEM, 1982; XIANG; JUDELSON, 2014). Isso

favorece a sobrevivência do esporo pois a noite proporciona um ambiente mais úmido, ideal para sobrevivência de um esporângio hialino e turgido como os de *P. infestans* (MIZUBUTI; AYLOR; FRY, 2000). Os esporos, então, podem ser dispersos pelo vento ou água (GRANKE et al., 2009).

Zoósporos são propágulos potentes. Eles aumentam o alcance de infecção podendo nadar durante 24 horas alcançando alguns centímetros em solos úmidos via irrigação ou chuva, podendo chegar até as raízes ou tubérculos (PORTER; DASGUPTA; JOHNSON, 2005). Esporos carregados pelo vento conseguem atingir grandes distâncias devido ao seu formato alongado, com estudos mostrando esporos capazes de percorrer até 20 km em menos de 3 horas (AYLOR, 2003). Em poucos minutos os zoósporos são liberados após a maturação do esporângio. Esse processo normalmente ocorre tarde da noite, quando há maior probabilidade de orvalho e baixas temperaturas.

A forma sexuada dos esporos, oósporos, são de grande relevância para o progresso epidemiológico da doença. Eles possuem parede espessa e podem persistir durante anos no solo e sobreviver temperaturas negativas e grande variedade de agentes químicos. (DUNCAN, 1985; TURKENSTEEN et al., 2000).

2.2.3 Controle da doença

O manejo eficiente da requeima é complexo e exige ferramentas conjuntas sobre a doença e conhecimento intrínseco do patógeno. O manejo integrado envolve: a elaboração de boletins meteorológicos, indicando condições propícias ao progresso do patógeno; o conhecimento da raça predominante na região, direcionando a qual método de controle recorrer em face de uma nova epidemia; o uso de marcadores moleculares identificando as diferentes raças para o emprego eficiente do controle genético; uma rede de alerta entre produtores, com indicadores geográficos de novas epidemias, já que devido ao grande alcance dos propágulos de *P. infestans*; conhecer os locais de novas infecções reflete na necessidade de ações preventivas ou não; o controle cultural, removendo plantas doentes, alterando espaçamento da cultura e turnos de irrigação, tornando o microclima desfavorável ao patógeno; e o controle químico, com o uso de fungicidas sistêmico, principalmente o Metalaxyl (MIZUBUTI; JÚNIOR; FORBES, 2007).

Uma plataforma única associando informações meteorológicas, fitopatológicas e alertas da comunidade rural também podem refletir em um controle eficiente com perdas insignificativas de produção (FRY, 2016).

2.3 Resistência do tomateiro à requeima

Os genes de resistência, comercialmente utilizados no controle genético de requeima em tomateiro, são provenientes de espécies selvagens aparentadas do tomate comercial. As principais espécies são *Solanum pimpinellifolium*, *Solanum habrochaites* e *Solanum pennellii*. Resistência proveniente de acessos de *S. pimpinellifolium* são em sua maioria vertical, enquanto as demais espécies normalmente apresentam graus de resistência horizontal. Os quatro principais genes de resistência foram obtidos pelos acessos de *S. pimpinellifolium*: West Virginia 19 e 731 (gene de resistência a requeima *Ph-1*); West Virginia 700 (*Ph-2*); L3708 (ou, LA1269 e PI 365957) (*Ph-3*) e L3707 (ou, PI 365951) (ZHI et al., 2021).

A introdução de fontes de resistência provenientes de acessos silvestres carregam diversos genes indesejáveis, e esse processo, chamado de *linkage drag*, muitas vezes inviabiliza o uso do controle genético em algumas cultivares (BROUWER; CLAIR, 2004). Alguns acessos de *S. lycopersicum* já foram estudados e podem ser considerados fontes de resistência, como o BGH-2127 (LAURINDO et al., 2018) e o NC 2 CELBR (GARDNER; PANTHEE, 2010).

O gene *Ph-1* foi mapeado na década de 70 e confere resistência somente a raça T₀ e, atualmente, é ineficaz contra a principal raça disseminada nos campos de tomate, a raça T₁.

Subsequente a descoberta do *Ph-1*, Gallegly & Marvel em 1955 identificaram um acesso contendo o gene *Ph-2*, esse gene demonstrou ter aproximadamente 20 caracteres quantitativos envolvidos na resistência, e o locus desse gene, Solyc10g08460, indica a produção de uma proteína de sinalização do tipo CC-NBS-LRR (ITAG 4.0) que foi descrito como desencadeante dessa resistência (ZHANG et al., 2013). Esse gene apresenta dominância parcial, e foi localizado no cromossomo 10, entre os marcadores CP105 e TG233 (MOREAU et al., 1998). O *Ph-2* apresenta uma característica distinta dos genes de resistência, pois apresenta um efeito de

redução no progresso da doença (BLACK et al., 1996; GOODWIN; SUJKOWSKI; FRY, 1995) indicando que o gene faz parte de uma rota alternativa de defesa ao invés do reconhecimento e reação de hipersensibilidade, como na maioria das resistências qualitativas.

No início dos anos 2000, o gene *Ph-3* foi caracterizado, e evidenciou-se que se trata de um gene com dominância incompleta. Contudo, essa resistência possui amplo espectro e é considerada uma das principais fontes de resistência de tomateiros à requeima (MERK; ASHRAFI; FOOLAD, 2012). Procurando entender sobre o mecanismo de resistência, Zhang et al. (2013, 2014) mapearam o gene *Ph-3* a uma precisão de 26 kb, e conseguiram demonstrar que essa região codifica uma proteína CC-NBS-LRR (*Coiled-Coil – Nucleotide Binding Site – Leucine Rich Repetition*). Essas proteínas estão envolvidas no reconhecimento do patógeno e ativação das rotas de defesa, impedindo o progresso da doença. O gene *Ph-2* não teve seu mecanismo estudado, porém Moreau et al. (1998) localizaram e caracterizaram fenotipicamente sua resistência, e o gene parece retardar o progresso da doença, indicando que ele atue nos processos de defesa, e não reconhecimento como no caso do gene *Ph-3*.

Existe um grande esforço para associação e pirimidamento dos genes *Ph-2* e *Ph-3* em cultivares elite (NOWICKI et al., 2012). Essa estratégia visa utilizar uma barreira primária de reconhecimento, com o gene de resistência qualitativa (*Ph-3*), e, caso esse gene seja suplantado, a ação do gene *Ph-2* fornece uma barreira secundária ao progresso da doença, reduzindo o propágulo e, assim, a probabilidade de que surjam, na população do patógeno, novos alelos capazes de suplantar essas resistências.

2.4 Estudo de associação genômica ampla (GWAS)

A relação causal entre um polimorfismo e seu fenótipo é o principal alvo no estudo de associação genômica ampla (GWAS). A habilidade de prever características agronômicas por meio de polimorfismos presentes na sequência de DNA possibilita ganhos preciosos no tempo gasto para o desenvolvimento de novas variedades. Essa relação entre genótipo e fenótipo é de interesse dos pesquisadores desde Mendel, quando postulou a existência de um fator interno e transmissível entre

os organismos que determinava suas características. Técnicas de mapeamento surgiram, como os QTL's (locus de caráter quantitativo) e até hoje são de extrema importância para estudos de associação (KORTE; FARLOW, 2013).

Atualmente, foram completamente sequenciados mais de 500 genomas de tomate (BOLGER et al., 2014; CONSORTIUM et al., 2014; LIN et al., 2014; TIEMAN et al., 2017; ZHU et al., 2018). Algumas dessas sequências ficam disponíveis em bancos de dados biológicos e possibilitam seu uso para análises de associação genômica (MARCHINI; HOWIE, 2010). Essas abordagens se mostraram extremamente importantes para os estudos tanto em espécies modelo, como em espécies menos pesquisadas e, assim, aceleraram o progresso de identificação de polimorfismo de único nucleotídeo (SNP's) com o re-sequenciamento de indivíduos utilizando as novas gerações de sequenciadores (TOMATO GENOME CONSORTIUM, 2012).

Atualmente, abordagens utilizando sequenciamento de RNA (RNA-Seq) são amplamente utilizadas, relacionando polimorfismos presentes detectados por GWAS e as diferentes quantidades e transcritos produzidos por fenótipos discrepantes. Porém, essas diferenças em sequências de RNA mensageiro, não necessariamente se traduzem em alterações quantitativas da proteína (VOGEL; MARCOTTE, 2012).

2.5 Genômica comparativa

Assim como no GWAS, a relação entre os polimorfismos genéticos dentro de um acesso e suas diferenças fenotípicas é o foco de interesse na genômica comparativa. A habilidade de prever caracteres fenotípicos, como resistência a doenças, ciclo, porte e outras características de interesse no melhoramento, exige o conhecimento da localização dos loci de uma característica e sua arquitetura genética.

A família *Solanaceae* tem sua estrutura gênica muito bem conservada (BONIERBALE; PLAISTED; TANKSLEY, 1988), e, por isso, culturas como o tomate e a batata são utilizadas frequentemente como modelo de estudos, principalmente sobre genes-R e seus respectivos mecanismos de resistência. O genoma do tomate é estudado extensivamente, e graças a vários trabalhos de confecção de mapas de ligação, o processo de localização de genes de resistência ganhou velocidade

(FOOLAD, 2007). O processo de domesticação reduziu significativamente número de genes em variedades de *S. lycopersicum*, reduzindo sua variabilidade e influenciando também na conservação dos genes (GAO et al., 2019). A genômica comparativa indica presença de regiões conservadas, formando clusters, ou seja, genes de resistência agrupados, essas regiões são denominadas de hot spots (ASHRAFI; KINKADE; FOOLAD, 2009).

Estudo realizado, utilizando proteômica comparativa, mostrou diferença significativa na expressão de proteínas entre acessos de *S. lycopersicum* que diferem em relação a resistência a *P. infestans* (Tabela 1). O estudo comparou o acesso BGH-2127 resistente e a cultivar Santa Clara susceptível. As proteínas endoquitinase ácida de 26 kDa e ribonuclease T2 foram mais abundantes, no momento da infecção do patógeno, no acesso BGH-2127. Atividade enzimática confirmou maior abundância de quitinase no acesso BGH-2127 em relação à Santa Clara (LAURINDO, 2017), evidenciando alta atividade de proteínas do processo oxidativo e da rigidez de parede celular, clara resposta ao reconhecimento do patógeno.

Tabela 1. - Proteínas diferencialmente abundantes em folhas de tomateiro dos genótipos BGH-2127 e Santa Clara, resistente e suscetível à requeima, respectivamente, identificadas por espectrometria de massas.

ID ^a	Acesso (UNIPROT)	Proteína (Organismo)	pI/MM ^b		Nível de abundância ^c	Anova (p<0.05)	Cobertura (%) ^d	Score ^e
			Experimental	Teórico				
Energia e metabolismo								
8	K4AYT8	Nucleosídeo difosfatoquinase (<i>Solanum lycopersicum</i>)	6,66/12,10	6,32/16,24	-2,07	0,008	10,8	100
24	Q2MI94	ATP sintase subunidade beta (<i>Solanum lycopersicum</i>)	5,74/16,51	5,43/14,57	-4,18	0,001	19,5	169
122	Q2MI93	ATP sintase subunidade beta (<i>Solanum lycopersicum</i>)	5,50/54,28	5,28/53,49	-2,19	0,025	4,6	183
256	Q645M9	Malato desidrogenase glioxissomal (<i>Solanum lycopersicum</i>)	9,27/35,37	8,40/37,94	-2,43	0,006	5,37	60
Fotossíntese								
5	A9S3R8	Ribulose bifosfato carboxilase subunidade menor (<i>Physcomitrella patens</i>)	4,81/11,80	8,20/20,16	-2,67	0,045	6,63	178
106	K4B4L4	Estabilidade do FSII/Fator de montagem HCF136 (<i>Solanum lycopersicum</i>)	5,12/40,68	6,25/41,16	-1,94	0,03	15,9	236
121	P27065	Ribulose bifosfato carboxilasesubunidade maior (<i>Solanum lycopersicum</i>)	7,09/54,22	6,55/53,43	-2,95	0,035	6,29	243
140	P12372	Subunidade II do centro de reação FSI (<i>Solanum lycopersicum</i>)	9,73/17,98	9,71/22,96	-10,09	0,038	3,37	50
185	P23322	Proteína envolvida no enriquecimento de oxigênio 1 (<i>Solanum lycopersicum</i>)	5,23/32,50	5,91/35,15	-1,64	0,024	2,43	81
191	K4B876	Clorofila de ligação a/b (<i>Solanum lycopersicum</i>)	4,87/28,87	5,15/28,20	-2,31	0,018	7,92	76
268	P12372	Subunidade II do centro de reação FSI (<i>Solanum lycopersicum</i>)	9,76/21,56	9,71/22,96	-2,58	0,004	48	3,37
Estresse e defesa								
44	K4D389	2-cis peroxirredoxina BAS1 (<i>Solanum lycopersicum</i>)	4,53/23,80	6,00/29,73	-3,03	0,006	14,2	155
48	K0I7G7	Tiorredoxina peroxidase (<i>Nicotiana tabacum</i>)	4,70/24,67	8,25/39,97	-2,16	0,006	5,19	90
56	Q9M5A8	Chaperonina 21 (<i>Solanum lycopersicum</i>)	4,66/36,59	5,84/56,02	-1,53	0,039	7,03	394
65	Q05539	Endoquitinase ácida de 26 kDa (<i>Solanum lycopersicum</i>)	5,52/28,15	5,93/28,01	1,77	0,036	3,95	65
131	K4B172	Calreticulina (<i>Solanum lycopersicum</i>)	4,09/57,73	4,50/40,80	-1,73	0,011	3,36	53
261	K4CBC7	Ribonuclease T2 (<i>Solanum lycopersicum</i>)	9,95/30,89	8,52/17,34	2,43	0,006	5,37	66
Transcrição								
67	K4BK45	Fator de splicing 3B, subunidade IV (<i>Solanum lycopersicum</i>)	4,36/28,97	4,70/32,68	-1,65	0,035	4,73	68
Outras								
11	K4CA08	Repetições de pentapéptido (<i>Solanum lycopersicum</i>)	5,23/13,53	8,24/24,76	-1,58	0,026	6,58	70

^a ID representa o número de identificação das proteínas nos géis bidimensionais.

^b Pontos isoeletricos (pI) e massas moleculares (MM em KDa) experimentais e teóricos.

^c Nível de abundância representados pela % de volume dos spots: os sinais (+) e (-) indicam maior e menor abundância da proteína, respectivamente, para o genótipo BGH-2127(resistente) em relação ao cultivar Santa Clara (suscetível) à requeima.

^d Porcentagem de cobertura.

^e Score Mascot.

^f Peptídeos identificados e confirmados com pelo menos de 90% de certeza pelo software Scaffold.

Fonte: Laurindo, 2017, p. 33.

A análise estrutural do locus de resistência resultado da construção de mapas de ligação, acelerou significativamente o processo de localização de genes R. Mais de 100 loci de resistência em tomate já foram caracterizados (FOOLAD, 2007). Utilizando estratégias como marcadores moleculares, caminhamento em cromossomo e mapas de ligação, vários desses genes já foram isolados, como por exemplo o ASC, BS4, CF5, CF9, HERO, I2, MI1-2, PTO, SW5 e VE (BRANDWAGT et al., 2000; DIXON et al., 1998; ERNST et al., 2002; KAWCHUK et al., 2001; MILLIGAN et al., 1998; ORI et al., 1997; PARNISKE et al., 1997; SCHORNACK et al., 2004). Graças a esses estudos de associação realizados, há um grande entendimento sobre os hotspots de genes de resistência, permitindo a identificação de genes homólogos até em espécies diferentes. Isso permite que as mais diversas fontes de dado a respeito de uma doença, como um proteoma por exemplo, sejam também inclusos para análise comparativa e tenham resultados positivos.

As novas tecnologias computacionais e de sequenciamento (NGSs), como HiSeq2500 (Illumina), a GS FLX+ system (Roche), 5500xl SOLiD (Life Technologies), Ion Proton (Life Technologies), e Nanopore (Oxford) criaram um arcabouço tecnológico que favoreceu em peso o avanço desses estudos, assim o re-sequenciamento de genomas já estudados pode chegar ao nível de diferenciação em um genoma de apenas um par de base.

Os recursos genéticos proveniente de acessos de tomate, devido a essa tecnologia hoje são guardados em bancos de dados biológicos digitais, como o European Bioinformatics Institute (EMBL-EBI), Sol Genomics Network (SNG), Tomato Genetic Resource Center (TGRC), USA (<http://tgrc.ucdavis.edu>); o National Institute of Agrobiological Sciences (NIAS) Genbank, Japan (<http://www.gene.affrc.go.jp>); e o NARO Institute of Vegetable and Tea Science (NIVTS), Japan (<http://www.naro.affrc.go.jp/vegetea>) e seus caracteres morfoagronômicos associados podem ser encontrados nesses bancos, independentemente de ter sido ou não avaliado sua origem genômica (SHIRASAWA et al., 2013).

3 MATERIAL E MÉTODOS

3.1 Recursos genéticos

A Universidade Federal de Viçosa, por meio do curador do BGH-UFV, professor Derly José Henriques Silva, dispõe do direito de intercambiar esse patrimônio genético com instituições de pesquisa, públicas ou privadas e disponibilizou os acessos de interesse dessa pesquisa.

Foram utilizados nessa pesquisa dados do sequenciamento WGS dos acessos de *Solanum lycopersicum*, BHG-2127, Santa Clara e o NC-2-CELBR, e o acesso de *Solanum habrochaites* código BGH-6902 (LAURINDO, 2017). A cultivar Santa Clara foi utilizada como controle susceptível. O acesso NC-2-CELBR foi utilizado como controle resistente por possuir ambos os genes *Ph-2* e *Ph-3*. Os acessos BGH-2127 e BHG-6902 resistentes, foram focos dessa análise.

Além destes, 84 sequências (Tabela 2) de genoma obtidas no European Bioinformatics Institute (EBI), de acessos de *S. lycopersicum*, *S. arcanum*, *S. cheesmaniae*, *S. chilense*, *S. chmielewskii*, *S. corneliomuelleri*, *S. galapagense*, *S. habrochaites*, *S. huaylasense*, *S. neorickii*, *S. pennellii*, *S. peruvianum* e *S. pimpinellifolium* já caracterizados em relação a resistência à *P. infestans* foram adicionadas para o estudo de associação genômica ampla. Dos acessos obtidos no banco de dados biológicos, apenas dois acessos foram fenotipados como resistentes à *Phytophthora infestans*, os acessos LA1777 e LA2093 (Tomato Genetics Resource Center), enquanto todos os outros acessos foram caracterizados como susceptíveis.

Tabela 2. – Lista de acessos obtidos no banco de dados biológicos EBI-EMBL utilizados nesse estudo.

Ref	Nome Acesso	Nome Científico	Cód. EBI	Amostra	Cultivar	Fenótipo
RF_001	LA2706	<i>Solanum lycopersicum</i>	ERR418039	SAMEA2340764	Solanum lycopersicum cv Moneymaker LA2706/PI262996	Susceptível
RF_002	LA2838A	<i>Solanum lycopersicum</i>	ERR418040	SAMEA2340765	Solanum lycopersicum cv Alisa Craig LA2838A/EA01101	Susceptível
RF_003	PI406760	<i>Solanum lycopersicum</i>	ERR418041	SAMEA2340766	Solanum lycopersicum cv Gardeners delight EA06086/PI406760	Susceptível
RF_004	LA1090	<i>Solanum lycopersicum</i>	ERR418042	SAMEA2340767	Solanum lycopersicum cv Rutgers LA1090/EA00465	Susceptível
RF_005	EA00325	<i>Solanum lycopersicum</i>	ERR418043	SAMEA2340768	Solanum lycopersicum cv Galina EA00325	Susceptível
RF_006	EA00488	<i>Solanum lycopersicum</i>	ERR418044	SAMEA2340769	Solanum lycopersicum cv Taxi EA00488	Susceptível
RF_007	EA00375	<i>Solanum lycopersicum</i>	ERR418045	SAMEA2340770	Solanum lycopersicum cv Katinka Cherry EA00375	Susceptível
RF_008	EA00371	<i>Solanum lycopersicum</i>	ERR418046	SAMEA2340771	Solanum lycopersicum cv John's big orange EA00371	Susceptível
RF_011	LA2463	<i>Solanum lycopersicum</i>	ERR418047	SAMEA2340772	Solanum lycopersicum cv Allround LA2463/LYC1365/EA02617	Susceptível
RF_012	LYC1969	<i>Solanum lycopersicum</i>	ERR418048	SAMEA2340773	Solanum lycopersicum cv Sonata LYC1969/EA02724	Susceptível
RF_013	LYC3897	<i>Solanum lycopersicum</i>	ERR418049	SAMEA2340774	Solanum lycopersicum cv Cross Country LYC1738/EA03701	Susceptível
RF_014	LYC3476	<i>Solanum lycopersicum</i>	ERR418050	SAMEA2340775	Solanum lycopersicum cv lidi LYC3476/EA03362	Susceptível
RF_015	TR00003	<i>Solanum lycopersicum</i>	ERR418051	SAMEA2340776	Solanum lycopersicum cv Momatero TR00003	Susceptível
RF_016	LYC11	<i>Solanum lycopersicum</i>	ERR418052	SAMEA2340777	Solanum lycopersicum cv Rote Beere LYC1343/CGN15465	Susceptível
RF_017	LYC3340	<i>Solanum lycopersicum</i>	ERR418053	SAMEA2340778	Solanum lycopersicum LYC3306/T1039	Susceptível
RF_018	EA01155	<i>Solanum lycopersicum</i>	ERR418054	SAMEA2340779	Solanum lycopersicum cv Dana EA01155	Susceptível
RF_019	EA01049	<i>Solanum lycopersicum</i>	ERR418055	SAMEA2340780	Solanum lycopersicum cv Large Pink EA01049	Susceptível
RF_020	LYC3153	<i>Solanum lycopersicum</i>	ERR418056	SAMEA2340781	Solanum lycopersicum LYC3153/EA03221	Susceptível
RF_021	LYC3155	<i>Solanum lycopersicum</i>	ERR418057	SAMEA2340782	Solanum lycopersicum EA03222/LYC3155/LYC2513/T828	Susceptível
RF_022	PI129097	<i>Solanum lycopersicum</i>	ERR418058	SAMEA2340783	Solanum lycopersicum PI129097/EA04710	Susceptível
RF_023	PI272654	<i>Solanum lycopersicum</i>	ERR418059	SAMEA2340784	Solanum lycopersicum PI272654/EA05170	Susceptível
RF_024	EA00990	<i>Solanum lycopersicum</i>	ERR418060	SAMEA2340785	Solanum lycopersicum cv Jersey devil EA00990	Susceptível
RF_025	LA0118	<i>Solanum corneliomulleri</i>	ERR418061	SAMEA2340786	Solanum corneliomulleri LA0118/EA03384/T1248	Susceptível
RF_026	EA00157	<i>Solanum lycopersicum</i>	ERR418062	SAMEA2340787	Solanum lycopersicum cv Polish Joe EA00157	Susceptível
RF_027	CGN20815	<i>Solanum lycopersicum</i>	ERR418063	SAMEA2340788	Solanum lycopersicum cv Cal J Tm VF EA02054/CGN20815	Susceptível
RF_028	PI303721	<i>Solanum lycopersicum</i>	ERR418064	SAMEA2340789	Solanum lycopersicum cv The Dutchman PI303721	Susceptível
RF_029	LA4451	<i>Solanum lycopersicum</i>	ERR418065	SAMEA2340790	Solanum lycopersicum cv Black Cherry LA4451(?)	Susceptível
RF_030	V710092	<i>Solanum lycopersicum</i>	ERR418066	SAMEA2340791	Solanum lycopersicum cv Anto EA01835/V710029	Susceptível
RF_031	PC11029	<i>Solanum lycopersicum</i>	ERR418067	SAMEA2340792	Solanum lycopersicum cv Winter Tipe PC11029	Susceptível
RF_032	PI093302	<i>Solanum lycopersicum</i>	ERR418068	SAMEA2340793	Solanum lycopersicum cv Chang Li PI93302/EA04243	Susceptível
RF_033	EA00892	<i>Solanum lycopersicum</i>	ERR418069	SAMEA2340794	Solanum lycopersicum cv Belmonte EA00892/SG16	Susceptível
RF_034	EA01088	<i>Solanum lycopersicum</i>	ERR418070	SAMEA2340795	Solanum lycopersicum cv Tiffen Mennonite EA1088	Susceptível
RF_035	PI203232	<i>Solanum lycopersicum</i>	ERR418071	SAMEA2340796	Solanum lycopersicum cv Wheatleys Frost Resistant PI203232/EA04939	Susceptível
RF_036	PI311117	<i>Solanum lycopersicum</i>	ERR418072	SAMEA2340797	Solanum lycopersicum PI311117/EA05701	Susceptível
RF_037	LA1324	<i>Solanum lycopersicum</i>	ERR418073	SAMEA2340798	Solanum lycopersicum PI365925/LA1324/EA05891	Susceptível
RF_038	PI158760	<i>Solanum lycopersicum</i>	ERR418074	SAMEA2340799	Solanum lycopersicum cv Chih Mu Tao Se PI158760/EA04828	Susceptível
RF_039	LA0113	<i>Solanum lycopersicum</i>	ERR418075	SAMEA2340800	Solanum lycopersicum LA0113/EA00526	Susceptível
RF_040	LYC1410	<i>Solanum lycopersicum</i>	ERR418076	SAMEA2340801	Solanum lycopersicum cv ES 58 Heinz LYC1410/EA02655	Susceptível
RF_041	PI169588	<i>Solanum lycopersicum</i>	ERR418077	SAMEA2340802	Solanum lycopersicum cv Dolmaik PI169588/EA04861	Susceptível
RF_042	LYC2962	<i>Solanum lycopersicum</i>	ERR418078	SAMEA2340803	Solanum lycopersicum LYC2962/EA03107/T556	Susceptível
RF_043	LYC2910	<i>Solanum lycopersicum</i>	ERR418079	SAMEA2340804	Solanum lycopersicum LYC2910/EA03058/T115	Susceptível
RF_044	LYC2798	<i>Solanum pimpinellifolium</i>	ERR418080	SAMEA2340805	Solanum pimpinellifolium LYC2798/EA02994	Susceptível
RF_045	LYC2740	<i>Solanum pimpinellifolium</i>	ERR418081	SAMEA2340806	Solanum pimpinellifolium LYC2740/EA02960	Susceptível

Ref	Nome Acesso	Nome Científico	Cód. EBI	Amostra	Cultivar	Fenótipo
RF_046	LA1584	<i>Solanum pimpinellifolium</i>	ERR418082	SAMEA2340807	Solanum pimpinellifolium LA1584/PI407541/EA00676	Susceptível
RF_047	LA1578	<i>Solanum pimpinellifolium</i>	ERR418083	SAMEA2340808	Solanum pimpinellifolium LA1578/EA00674	Susceptível
RF_049	LA1278	<i>Solanum peruvianum</i>	ERR418084	SAMEA2340809	Solanum peruvianum LA1278/PI365941/TR00005	Susceptível
RF_051	LA2663	<i>Solanum chmielewskii</i>	ERR418085	SAMEA2340810	Solanum chmielewskii LA2663/TR00007	Susceptível
RF_052	LA2695	<i>Solanum chmielewskii</i>	ERR418086	SAMEA2340811	Solanum chmielewskii LA2695/EA00759	Susceptível
RF_053	LA0483	<i>Solanum cheesmaniae</i>	ERR418087	SAMEA2340812	Solanum cheesmaniae-f-minor/S. Galapagense LA0483/EA00581	Susceptível
RF_054	CGN15820	<i>Solanum lycopersicum</i>	ERR418088	SAMEA2340813	Solanum lycopersicum CGN15820/TR00024	Susceptível
RF_055	LA1401	<i>Solanum cheesmaniae</i>	ERR418089	SAMEA2340814	Solanum cheesmaniae-f-minor/S. Galapagense LA1401/PI 365897/EA00652	Susceptível
RF_056	LA2133	<i>Solanum neorickii</i>	ERR418090	SAMEA2340815	Solanum neorickii LA2133/EA00729	Susceptível
RF_057	LA0735	<i>Solanum neorickii</i>	ERR418091	SAMEA2340816	Solanum neorickii LA0735/CGN24193/TR00025	Susceptível
RF_058	LA2157	<i>Solanum arcanum</i>	ERR418092	SAMEA2340817	Solanum arcanum LA2157/TR00008	Susceptível
RF_059	LA2172	<i>Solanum arcanum</i>	ERR418093	SAMEA2340818	Solanum arcanum LA2172/TR00009	Susceptível
RF_060	LA1954	<i>Solanum peruvianum</i>	ERR418094	SAMEA2340819	Solanum peruvianum LA1954/EA00713	Susceptível
RF_062	LA1983	<i>Solanum huaylasense</i>	ERR418095	SAMEA2340820	Solanum huaylasense LA1983/TR00010	Susceptível
RF_063	LA1365	<i>Solanum huaylasense</i>	ERR418096	SAMEA2340821	Solanum huaylasense LA1365/PI 365953/TR00011	Susceptível
RF_064	CGN15532	<i>Solanum chilense</i>	ERR418097	SAMEA2340822	Solanum chilense CGN15532/TR00012	Susceptível
RF_065	CGN15530	<i>Solanum chilense</i>	ERR418098	SAMEA2340823	Solanum chilense CGN15530/TR00013	Susceptível
RF_066	CGN15791	<i>Solanum habrochaites</i>	ERR418099	SAMEA2340824	Solanum habrochaites F glabratum PI 127827 (?)/CGN157591/TR00014	Susceptível
RF_067	PI134418	<i>Solanum habrochaites</i>	ERR418100	SAMEA2340825	Solanum habrochaites F glabratum PI134418/TR00015	Susceptível
RF_068	CGN15792	<i>Solanum habrochaites</i>	ERR418101	SAMEA2340826	Solanum habrochaites F glabratum CGN157592/TR00015	Susceptível
RF_069	LA1718	<i>Solanum habrochaites</i>	ERR418102	SAMEA2340827	Solanum habrochaites F glabratum LA1718/PI 390663/EA00699	Resistente
RF_070	LA1777	<i>Solanum habrochaites</i>	ERR418103	SAMEA2340828	Solanum habrochaites LA1777/EA00703	Susceptível
RF_071	LA0407	<i>Solanum habrochaites</i>	ERR418104	SAMEA2340829	Solanum habrochaites LA0407/PI 251304/EA0558	Susceptível
RF_072	LYC4	<i>Solanum habrochaites</i>	ERR418105	SAMEA2340830	Solanum habrochaites LYC4/TR00017	Susceptível
RF_073	LA1272	<i>Solanum pennellii</i>	ERR418106	SAMEA2340831	Solanum pennellii LA1272/PI 365970/EA02701/LYC1831	Susceptível
RF_074	LA0716	<i>Solanum pennellii</i>	ERR418107	SAMEA2340832	Solanum pennellii LA0716/PI 246502/EA00585	Susceptível
RF_075	LA1364	<i>Solanum huaylasense</i>	ERR418108	SAMEA2340833	Solanum huaylasense LA1364	Susceptível
RF_077	TR00018	<i>Solanum lycopersicum</i>	ERR418109	SAMEA2340834	Solanum lycopersicum cv Large Red Cherry TR00018	Susceptível
RF_078	EA00940	<i>Solanum lycopersicum</i>	ERR418110	SAMEA2340835	Solanum lycopersicum cv Porter EA00940	Susceptível
RF_088	TR00019	<i>Solanum lycopersicum</i>	ERR418111	SAMEA2340836	Solanum lycopersicum cv Bloody Butcher TR00019	Susceptível
RF_089	EA01019	<i>Solanum lycopersicum</i>	ERR418112	SAMEA2340837	Solanum lycopersicum cv Brandywine EA01019	Susceptível
RF_090	TR00020	<i>Solanum lycopersicum</i>	ERR418113	SAMEA2340838	Solanum lycopersicum cv Dixie Golden Giant TR00020	Susceptível
RF_091	EA01037	<i>Solanum lycopersicum</i>	ERR418114	SAMEA2340839	Solanum lycopersicum cv Giant Belgium EA01037	Susceptível
RF_093	TR00021	<i>Solanum lycopersicum</i>	ERR418115	SAMEA2340840	Solanum lycopersicum cv Kentucky Beefsteak TR00021	Susceptível
RF_094	TR00022	<i>Solanum lycopersicum</i>	ERR418116	SAMEA2340841	Solanum lycopersicum cv Marmande TR00022/PI647486	Susceptível
RF_096	TR00023	<i>Solanum lycopersicum</i>	ERR418117	SAMEA2340842	Solanum lycopersicum cv Thessaloniki TR00023	Susceptível
RF_097	EA01640	<i>Solanum lycopersicum</i>	ERR418118	SAMEA2340843	Solanum lycopersicum cv Watermelon beefsteak EA01640	Susceptível
RF_102	LA4133	<i>Solanum lycopersicum</i>	ERR418119	SAMEA2340844	Solanum lycopersicum LA4133/TR00026	Susceptível
RF_103	LA1421	<i>Solanum lycopersicum</i>	ERR418120	SAMEA2340845	Solanum lycopersicum LA1421/TR00027	Susceptível
RF_104	LA1044	<i>Solanum galapagense</i>	ERR418121	SAMEA2340846	Solanum galapagense LA1044	Susceptível
RF_105	LA1479	<i>Solanum lycopersicum</i>	ERR418122	SAMEA2340847	Solanum lycopersicum LA1479/TR00028	Susceptível
RF_201	LA2093	<i>Solanum pimpinellifolium</i>	SRR12039812	SAMN15293845	Solanum pimpinellifolium LA2093	Resistente

3.2.1 Sequenciamento de DNA

Dados do sequenciamento do genoma dos acessos BGH-2127, NC 2 CELBR (*S. lycopersicum*) e BGH-6902 (*S. habrochaites*), obtidos de experimentos anteriores, sequenciados pela plataforma DNBseq (BGI Americas Corporation, Cambridge, EUA) (LAURINDO, 2017), foram montados e utilizados nesse experimento.

3.2.2 Montagem do genoma por referência

Para montagem, primeiramente avaliou-se a qualidade dos fragmentos de leitura obtidos pelo sequenciamento, utilizando o software FastQC (ANDREWS, 2010). As amostras que tiveram trechos com leitura de PHRED menor 20, e menores que 50pb foram removidos utilizando o software Trimmomatic (BOLGER; LOHSE; USADEL, 2014).

O software utilizado na montagem por referência dos dados sequenciados foi o Bowtie2 (LANGMEAD; SALZBERG, 2012), O acesso utilizado como referência foi o cultivar 'Heinz 1706' versão SL4.0, variedade que teve seu genoma completamente montado em consórcio entre inúmeros pesquisadores (TOMATO GENOME CONSORTIUM, 2012), e que é amplamente utilizado em trabalhos de genômica comparativa em plantas de tomate (ARAFÁ et al., 2017a; PHAN et al., 2019).

3.2.3 Variantes

Para chamada de variantes, o programa utilizado foi o Freebayes 1.3.5 (GARRISON; MARTH, 2012) com parâmetros default, primeiramente as variantes foram chamadas individualmente contrastando com a referência (var. Heinz 1706 SL4.0). Depois, uma chamada de variantes, em conjunto, foi realizada com as 85 amostras do banco de dados biológicos e os 4 acessos do BGH-UFV para os estudos de associação. Nessa última análise, foi inserido parâmetro de exclusão de regiões com cobertura maior que 1.000 vezes.

3.2.4 Análise de componentes principais

Para análise de estratificação populacional foi utilizado o software PLINK 1.9 (PURCELL et al., 2007), foram filtradas variantes ligadas a menos de 50 kb, janela de

leitura de 10 pb e correlação maior que 0.1. Os gráficos foram gerados com auxílio do software R e pacotes ggplot2 (TEAM, 2013; WICKHAM, 2017).

3.2.5 Estudo de associação genômica ampla (GWAS)

O resultado da chamada de variantes conjunta foi, então, utilizado no software PLINK 1.9 (PURCELL et al., 2007), para o cálculo de associação das variantes com os fenótipos resistentes, BGH-2127, BGH-6902, LA1777, LA2093, e NC-2-CELBR, por grupo de espécies. Subsequentemente, para obtenção dos p-valores associados a cada SNP, foi utilizado o teste de Fisher exato e permutação de Monte Carlo como parâmetros do modelo. Foram filtrados SNPs com baixa qualidade e monomórficos, devido a amostra conter alelos raros nenhum parâmetro de frequência do alelo de menor frequência (MAF) foi estabelecido. Os p-valores obtidos foram submetidos a scripts utilizando o software R para definir o limiar significativo e de ruído com base no controle adaptativo de falso negativo (AFNC), descrito por Jeng et al., 2016, para extrair significância em testes de associação com alelos raros. O pacote qqman (TURNER, 2014) foi utilizado na elaboração do QQplot e Manhattan plot.

3.2.6 Predição gênica

Com os dados de variantes contrastados com a referência, utilizou-se o software BCFTools 1.10.2 (LI et al., 2009) para normalizar os indels e gerar a sequência consenso excluindo bases com qualidade inferior a 20.

A predição gênica foi feita pelo software AUGUSTUS 3.3.3 (STANKE et al., 2008), e extraiu genes codificantes (CDS), e os transcritos. O software foi utilizado em default, os genomas foram subdivididos por cromossomos por meio de scripts de Python 3.7 para melhorar a performance da análise.

3.2.7 Filogenia

Foi realizado um BLASTn (ALTSCHUL et al., 1990) com as sequências codificadoras (CDS) preditas para cada acesso, utilizando como query o loci dos genes R *Ph-2* e *Ph-3* presentes no acesso NC-2-CELBR, obtidos na anotação do genoma de referência. Os BLAST's com mais de 90% de identidade foram então alinhados com auxílio do software MUSCLE (EDGAR, 2004), então a relação

filogenética obtida pelo método da máxima verossimilhança calculada pelo software IQ-TREE (MINH et al., 2020), o cladograma foi desenhado com auxílio do software FigTree (RAMBAUT, 2012). Os genes em clados distintos foram analisados por sua proximidade ao gene de referência.

3.2.8 Marcadores Moleculares

Com os polimorfismos obtidos do alinhamento o software Primer3 (ROZEN; SKALETSKY, 2000) foi utilizado, recebendo o alinhamento local entre os acessos.

Para avaliação da viabilidade dos primers uma PCR *in silico*, foi executada com a ferramenta do Solgenomics Network com base no genoma de referência. As sequências forward e reverse dos primers, foram desenhados, e descartados primers em que houve múltiplos anelamentos (RAMIREZ-GONZALEZ; UAUY; CACCAMO, 2015).

4 RESULTADOS

4.1 Montagem por referência

O conteúdo de GC das amostras variaram entre 33% e 37%. Todas as amostras utilizadas passaram o teste de qualidade das reads pelo programa FASTQC (ANDREWS, 2010). O menor número de reads por acesso foi de 80.488.422 reads e o maior de 267.332.134 reads. A taxa geral de alinhamento variou entre 79,01% e 99,69% (Tabela 3).

Tabela 3. – Resultados de qualidade das reads e montagem por alinhamento dos acessos do banco de dados biológico EBI-EMBL.

Nome acesso	Cód. EBI	Conteúdo GC	Qualidade por base FASTQC	Número de reads	Alinhamentos pareados			Alinhamentos não pareados		Taxa geral de alinhamento
					Concordantes mais de uma vez	Concordantes apenas uma vez	Discordantes	Mais de uma vez	Apenas uma vez	
LA2706	ERR418039	35%	APROV	143.048.409	46.196.662	95.103.775	246.886	1.157.028	542.660	99,54%
LA2838A	ERR418040	34%	APROV	81.844.411	24.547.821	48.112.736	3.720.716	7.149.159	2.714.456	99,35%
PI406760	ERR418041	35%	APROV	118.177.838	33.718.193	65.887.924	8.152.568	14.551.354	4.188.380	99,11%
LA1090	ERR418042	35%	APROV	129.353.888	40.982.656	80.080.584	3.347.357	6.554.994	2.351.154	99,62%
EA00325	ERR418043	35%	APROV	117.144.896	23.154.372	45.355.701	23.296.571	38.786.567	10.296.652	99,32%
EA00488	ERR418044	35%	APROV	129.719.109	43.478.955	83.040.220	1.053.045	2.494.679	813.513	99,62%
EA00375	ERR418045	35%	APROV	118.648.063	35.261.404	64.643.754	5.936.826	13.067.933	4.472.544	96,60%
EA00371	ERR418046	35%	APROV	127.679.675	41.876.707	78.667.429	2.726.503	5.773.445	1.774.527	99,50%
LA2463	ERR418047	35%	APROV	149.476.422	48.809.167	95.049.262	1.960.991	4.343.694	1.525.595	99,52%
LYC1969	ERR418048	35%	APROV	143.309.610	38.969.177	68.482.636	13.291.178	28.170.987	9.423.047	97,37%
LYC3897	ERR418049	35%	APROV	144.473.838	48.661.748	92.568.966	1.002.552	2.590.330	1.004.149	99,69%
LYC3476	ERR418050	35%	APROV	135.095.010	42.187.266	80.688.100	5.212.282	9.728.364	2.774.792	99,44%
TR00003	ERR418051	35%	APROV	145.193.222	49.831.624	88.216.050	678.448	3.965.342	2.073.864	97,63%
LYC11	ERR418052	35%	APROV	150.032.336	51.869.603	92.314.497	835.848	3.310.386	1.660.633	98,32%
LYC3340	ERR418053	35%	APROV	155.640.727	52.521.756	93.680.837	1.352.677	5.315.929	2.605.668	97,35%
EA01155	ERR418054	36%	APROV	136.938.731	49.199.209	86.056.060	224.549	1.276.773	559.860	99,61%
EA01049	ERR418055	35%	APROV	147.796.558	50.187.675	92.168.916	1.434.146	4.152.703	2.184.604	99,43%
LYC3153	ERR418056	34%	APROV	148.394.267	48.664.379	96.090.406	1.127.098	2.756.458	972.577	99,56%
LYC3155	ERR418057	34%	APROV	133.665.842	42.543.940	87.793.133	808.755	2.132.075	891.226	99,25%
PI129097	ERR418058	35%	APROV	146.048.582	45.654.256	86.086.589	5.270.443	10.432.762	3.406.768	98,55%
PI272654	ERR418059	35%	APROV	152.866.000	50.354.122	92.837.679	2.463.370	6.246.008	2.914.370	98,28%
EA00990	ERR418060	36%	APROV	142.565.456	42.857.457	77.956.587	9.000.351	17.433.927	5.560.788	99,12%
LA0118	ERR418061	36%	APROV	147.432.175	38.100.217	62.474.808	4.026.660	26.552.598	14.352.771	84,82%
EA00157	ERR418062	36%	APROV	152.746.309	52.101.873	92.841.473	2.295.346	5.654.679	2.735.578	99,14%
CGN20815	ERR418063	36%	APROV	157.538.153	54.971.329	95.484.624	688.244	2.955.065	1.438.162	97,34%
PI303721	ERR418064	36%	APROV	147.920.197	55.469.803	87.973.831	984.192	3.266.838	1.292.014	99,18%
LA4451	ERR418065	37%	APROV	126.027.161	42.421.111	79.989.703	812.180	2.254.290	921.764	99,04%
V710092	ERR418066	35%	APROV	149.029.723	47.786.520	87.726.603	3.385.723	9.816.446	6.224.348	98,58%
PC11029	ERR418067	35%	APROV	148.971.559	48.596.207	93.193.365	2.922.884	5.585.788	1.605.121	99,55%
PI093302	ERR418068	35%	APROV	155.233.701	49.297.635	95.925.065	3.499.402	7.309.045	2.975.346	99,12%
EA00892	ERR418069	35%	APROV	161.983.887	53.842.428	101.335.815	1.724.599	4.185.241	1.448.901	98,60%
EA01088	ERR418070	35%	APROV	157.656.335	53.342.663	97.684.310	2.343.960	5.354.915	1.758.240	99,54%
PI203232	ERR418071	36%	APROV	162.436.627	57.244.083	102.428.504	534.733	2.134.329	1.067.732	99,61%
PI311117	ERR418072	35%	APROV	159.284.916	48.815.878	91.908.136	7.980.396	15.018.657	4.185.471	99,39%
LA1324	ERR418073	35%	APROV	138.705.470	45.994.818	87.431.911	445.307	2.698.489	1.585.130	98,06%
PI158760	ERR418074	35%	APROV	159.202.970	52.124.397	93.696.593	4.122.821	10.429.871	5.704.881	99,25%
LA0113	ERR418075	34%	APROV	147.323.253	47.851.640	91.336.661	995.626	3.998.451	2.163.967	97,25%
LYC1410	ERR418076	35%	APROV	147.014.168	48.237.379	87.701.661	3.845.421	8.868.007	3.718.403	99,36%
PI169588	ERR418077	36%	APROV	154.894.250	51.144.495	87.932.753	6.311.681	13.249.973	4.287.410	99,52%
LYC2962	ERR418078	35%	APROV	154.247.813	50.582.529	87.901.155	2.551.102	8.734.223	4.957.742	95,87%

Nome acesso	Cód. EBI	Conteúdo GC	Qualidade por base FASTQC	Número de reads	Alinhamentos pareados			Alinhamentos não pareados		Taxa geral de alinhamento
					Concordantes mais de uma vez	Concordantes apenas uma vez	Discordantes	Mais de uma vez	Apenas uma vez	
LYC2910	ERR418079	34%	APROV	153.157.476	49.007.810	99.089.719	1.475.479	3.456.705	1.382.621	99,24%
LYC2798	ERR418080	36%	APROV	145.244.803	25.761.814	40.349.019	32.871.178	70.310.728	17.825.159	98,49%
LYC2740	ERR418081	35%	APROV	153.603.405	48.261.484	83.146.337	5.221.272	16.042.462	7.372.370	96,57%
LA1584	ERR418082	34%	APROV	132.404.845	46.259.983	80.872.640	400.184	3.444.088	1.655.136	98,25%
LA1578	ERR418083	36%	APROV	151.237.889	50.651.991	82.530.205	5.379.705	14.351.648	4.837.719	97,96%
LA1278	ERR418084	35%	APROV	143.419.405	38.665.541	61.310.461	2.910.341	22.902.204	13.318.287	84,37%
LA2663	ERR418085	35%	APROV	152.281.578	35.295.140	77.688.026	3.461.476	21.449.744	12.676.190	87,67%
LA2695	ERR418086	36%	APROV	151.086.336	34.816.891	75.720.155	3.658.486	22.651.081	13.793.291	87,64%
LA0483	ERR418087	36%	APROV	152.081.255	51.925.923	88.884.729	1.994.615	6.603.841	3.049.858	97,07%
CGN15820	ERR418088	35%	APROV	154.967.111	54.849.113	91.748.412	1.251.084	4.858.312	2.290.151	97,71%
LA1401	ERR418089	35%	APROV	147.052.584	48.566.744	86.786.860	2.696.574	7.205.518	3.101.939	97,38%
LA2133	ERR418090	36%	APROV	146.343.150	34.580.471	71.200.826	3.972.365	22.766.308	13.066.170	87,24%
LA0735	ERR418091	37%	APROV	157.987.683	39.934.345	81.304.587	1.983.898	19.382.356	11.012.051	87,61%
LA2157	ERR418092	36%	APROV	139.769.028	34.527.581	79.350.303	4.101.970	22.632.622	8.583.249	95,58%
LA2172	ERR418093	35%	APROV	141.835.429	34.212.493	71.370.354	2.493.628	18.652.663	10.789.275	86,58%
LA1954	ERR418094	36%	APROV	155.141.729	43.246.932	65.054.694	3.530.025	26.532.229	14.275.311	85,24%
LA1983	ERR418095	36%	APROV	156.508.615	48.784.677	59.250.429	2.022.952	20.671.443	13.840.936	81,35%
LA1365	ERR418096	36%	APROV	134.921.516	46.025.442	46.935.979	1.569.652	17.365.263	12.262.610	81,04%
CGN15532	ERR418097	36%	APROV	134.139.679	38.266.833	49.154.666	3.621.215	28.274.393	15.553.620	84,21%
CGN15530	ERR418098	35%	APROV	140.897.300	39.715.105	60.237.502	2.388.059	22.266.675	13.132.936	85,20%
CGN15791	ERR418099	34%	APROV	145.709.843	39.860.362	94.207.548	849.167	10.718.414	4.073.455	97,67%
PI134418	ERR418100	35%	APROV	146.880.446	38.120.325	78.181.734	3.517.895	20.002.422	9.888.828	91,75%
CGN15792	ERR418101	35%	APROV	151.530.514	44.476.045	93.070.202	940.621	11.367.365	4.744.714	96,71%
LA1718	ERR418102	35%	APROV	146.641.566	35.726.432	71.643.646	6.948.935	27.367.987	11.629.149	91,25%
LA1777	ERR418103	34%	APROV	121.613.471	33.003.160	64.001.766	1.643.168	15.293.651	8.174.463	90,76%
LA0407	ERR418104	36%	APROV	156.291.729	44.505.786	93.464.623	2.261.316	14.833.296	5.979.719	96,38%
LYC4	ERR418105	35%	APROV	148.618.639	41.021.201	96.010.577	758.839	10.585.881	3.883.132	97,58%
LA1272	ERR418106	36%	APROV	150.806.550	25.770.511	31.550.762	19.091.484	74.389.133	22.944.554	82,94%
LA0716	ERR418107	35%	APROV	113.901.246	45.051.447	64.052.146	289.019	3.834.642	1.479.128	98,37%
LA1364	ERR418108	35%	APROV	151.498.167	35.810.698	41.127.209	10.552.972	44.976.596	19.451.203	79,01%
TR00018	ERR418109	36%	APROV	153.424.250	55.014.564	93.863.464	978.471	3.280.746	1.864.060	99,35%
EA00940	ERR418110	36%	APROV	153.958.385	25.393.290	45.936.965	39.106.510	68.693.329	16.776.722	99,49%
TR00019	ERR418111	35%	APROV	141.881.039	40.614.910	78.951.524	10.144.505	17.958.150	4.693.528	99,40%
EA01019	ERR418112	36%	APROV	139.470.839	49.779.140	86.904.394	680.118	2.084.593	689.870	99,48%
TR00020	ERR418113	35%	APROV	157.562.014	52.402.150	99.401.734	1.540.900	4.314.707	1.630.899	99,21%
EA01037	ERR418114	35%	APROV	127.601.507	46.178.163	80.152.318	102.464	803.950	354.051	99,54%
TR00021	ERR418115	36%	APROV	150.987.601	43.545.556	75.074.707	14.231.035	27.766.306	6.801.651	99,44%
TR00022	ERR418116	33%	APROV	132.634.642	43.863.319	87.657.319	117.282	778.417	367.116	99,68%
TR00023	ERR418117	36%	APROV	154.143.955	29.062.345	51.607.385	33.454.727	61.898.273	16.452.604	99,45%
EA01640	ERR418118	35%	APROV	150.115.060	43.204.073	85.905.751	9.428.800	16.559.742	4.411.259	99,27%
LA4133	ERR418119	36%	APROV	139.011.074	40.268.676	66.599.189	11.643.164	25.877.328	10.960.067	98,50%
LA1421	ERR418120	36%	APROV	154.620.421	49.887.908	88.031.227	5.371.898	12.398.115	3.882.924	97,94%

Nome acesso	Cód. EBI	Conteúdo GC	Qualidade por base FASTQC	Número de reads	Alinhamentos pareados			Alinhamentos não pareados		Taxa geral de alinhamento
					Concordantes mais de uma vez	Concordantes apenas uma vez	Discordantes	Mais de uma vez	Apenas uma vez	
LA1044	ERR418121	33%	APROV	136.693.393	44.332.087	85.005.077	532.714	3.161.384	2.170.278	96,96%
LA1479	ERR418122	35%	APROV	152.452.025	46.295.473	75.817.275	8.802.663	21.714.616	9.227.013	96,02%
LA2093	SRR12039812	36%	APROV	267.332.134	93.821.815	92.179.300	19.492.126	67.621.926	10.561.203	91,49%
BGH2127	n/a	36%	APROV	106.658.609	52.613.756	51.503.571	276.308	1.976.016	909.319	99,23%
BGH6902	n/a	36%	APROV	80.488.422	29.643.102	43.627.111	600.231	5.047.876	2.547.145	96,50%
NC2CELBR	n/a	36%	APROV	96.206.349	41.368.170	50.815.087	431.447	2.145.448	1.028.563	97,92%
STACLARA	n/a	36%	APROV	94.216.420	41.782.273	49.521.791	339.850	1.897.525	944.231	98,78%

4.2 Análise de componentes principais (PCA)

A análise de PCA indicou que 39,8% da variância dos SNPs foi explicada pelos componentes principais um e dois (Figura 2).

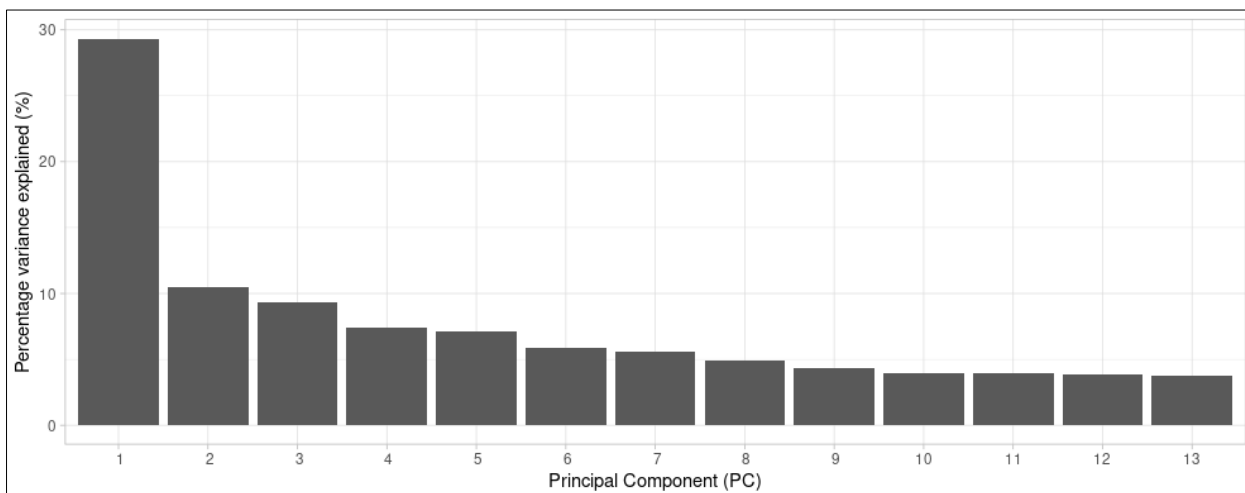


Figura 2. - Percentual da variância explicada por componente principal. Eixo X: número do componente principal. Eixo Y: porcentagem da variância explicada pelo componente principal.

A análise de componentes principais (Figura 3) indicou dois grupos predominantemente distintos, contendo principalmente os acessos da espécie *S. lycopersicum* e *S. habrochaites*.

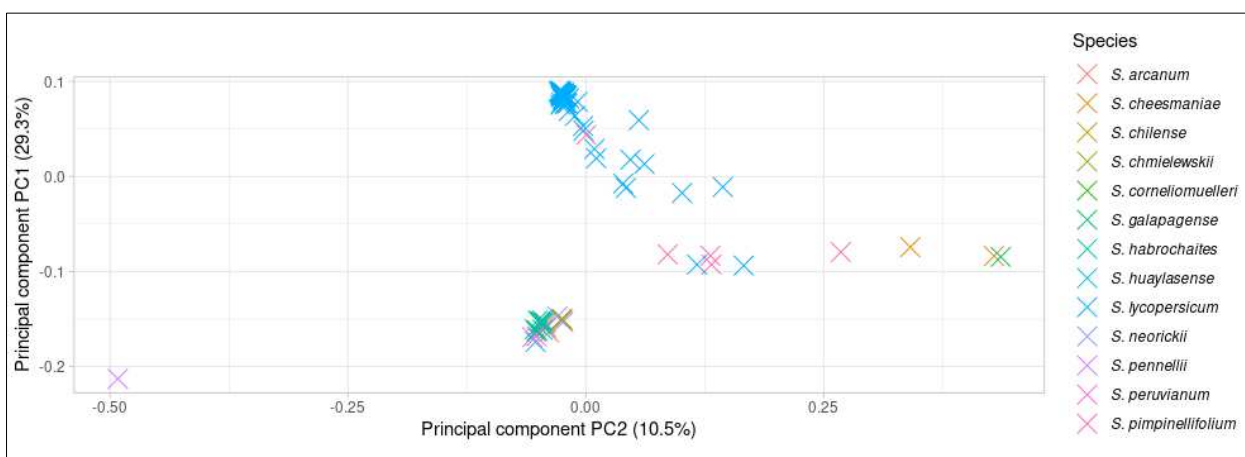


Figura 3. Dispersão dos acessos alvo em relação aos componentes principais um e dois. As marcas indicam os p-valores associados aos

componentes principais um e dois de cada um dos acessos. As cores agrupam os acessos por espécie.

4.3 GWAS

O modelo estratificado por espécie apresentou o melhor ajuste de p-valores seguindo distribuição normal (Figura 4) comparado aos outros modelos.

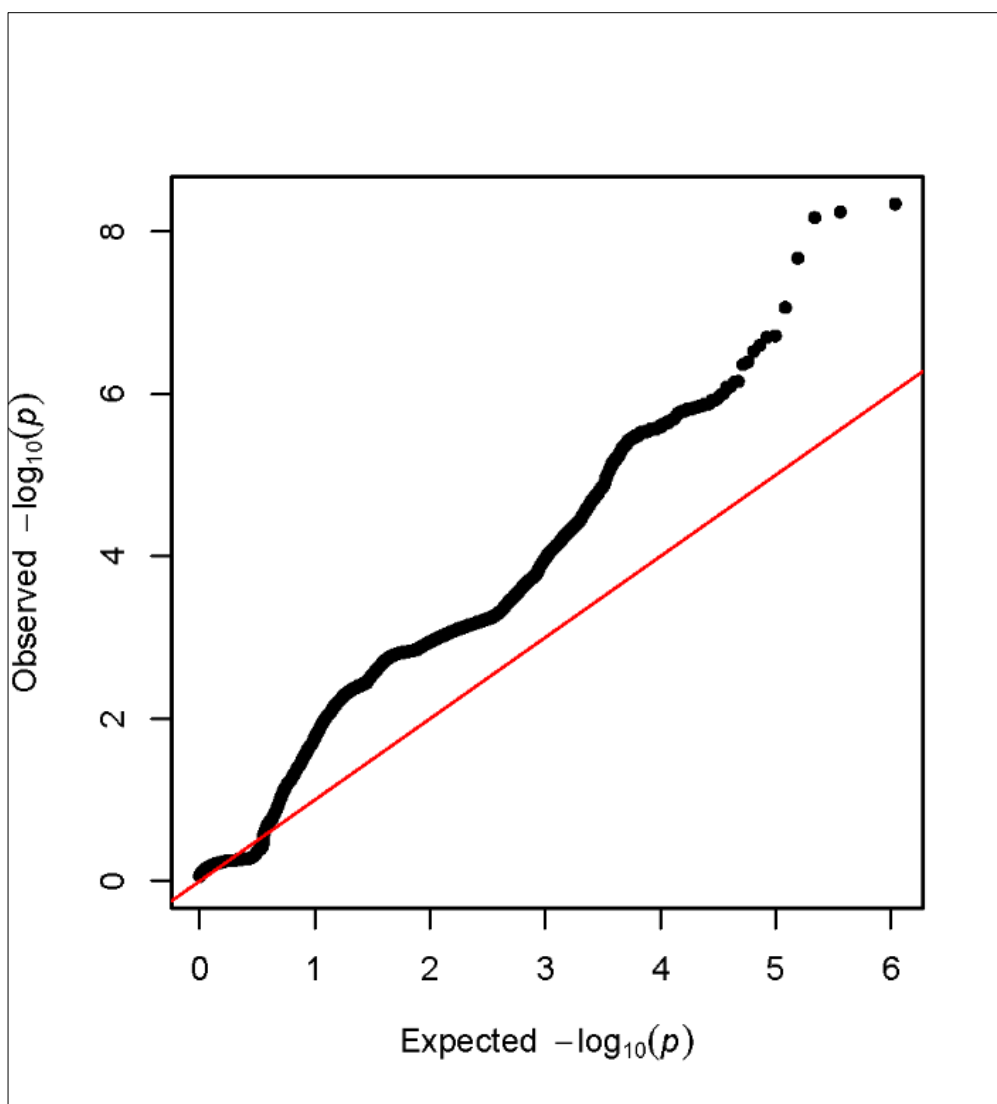


Figura 4. - Gráfico Q-Q plot. Dispersão de p-valor de cada SNP em relação ao esperado. Relação esperada entre o logaritmo negativo do p-valor observado e esperado entre SNPs sem associação, estão indicados pela linha vermelha. Os pontos indicam os valores observados e esperados do logaritmo negativo do p-valor nos SNPs do painel de acessos.

Após a correção de AFNC os p-valores associados a cada SNP foram significativos nos cromossomos 2, 3, 5, 6 e 7 (Figura 5) nas bases de posição: 1.730.874; 64.034.157; 56.537.071; 40.681.997; e 62.210.127 em cada cromossomo respectivamente.

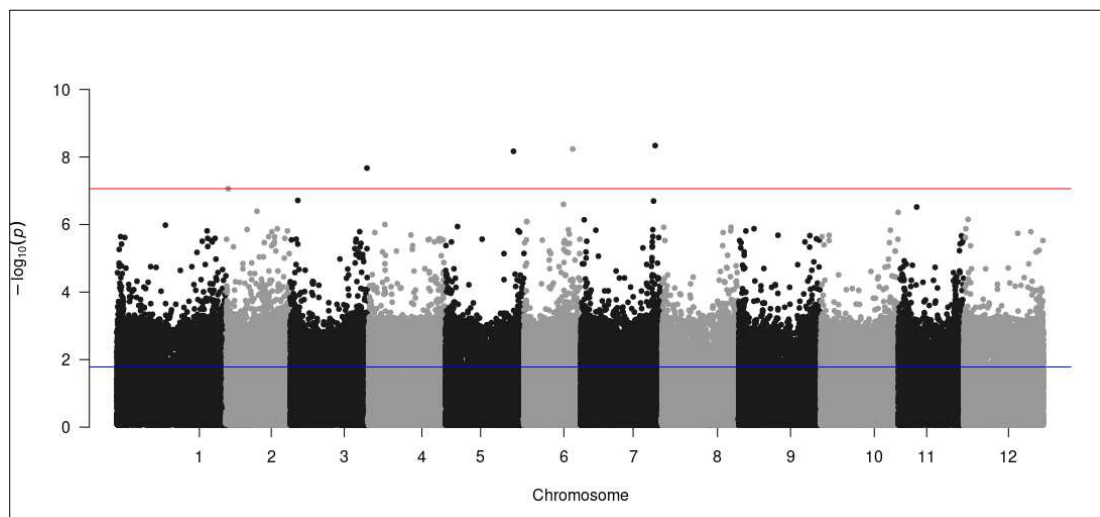


Figura 5. – Gráfico Manhattan plot. Dispersão dos logaritmos negativo dos p-valores de cada SNP em relação a posição no cromossomo. Linha vermelha indica o valor de significância pelo teste de AFNC. Linha azul delimita o nível de ruído.

4.4 Filogenia

4.4.1 Gene *Ph-2*

O BLASTn para região do gene *Ph-2* (Solyc10g08460) encontrou nas CDS dos acessos BGH-2129, BGH-6902, NC-2-CELBR, e Santa Clara, 29 cópias sendo que cada acesso esteve duplicado oito, cinco, oito e oito vezes no cromossomo 10, respectivamente.

A filogenia (Figura 6) dividiu essas cópias em oito clados, sendo os genes do clado verde e laranja os genes com maiores similaridades em relação ao locus de referência no acesso Heinz 1706, respectivamente.

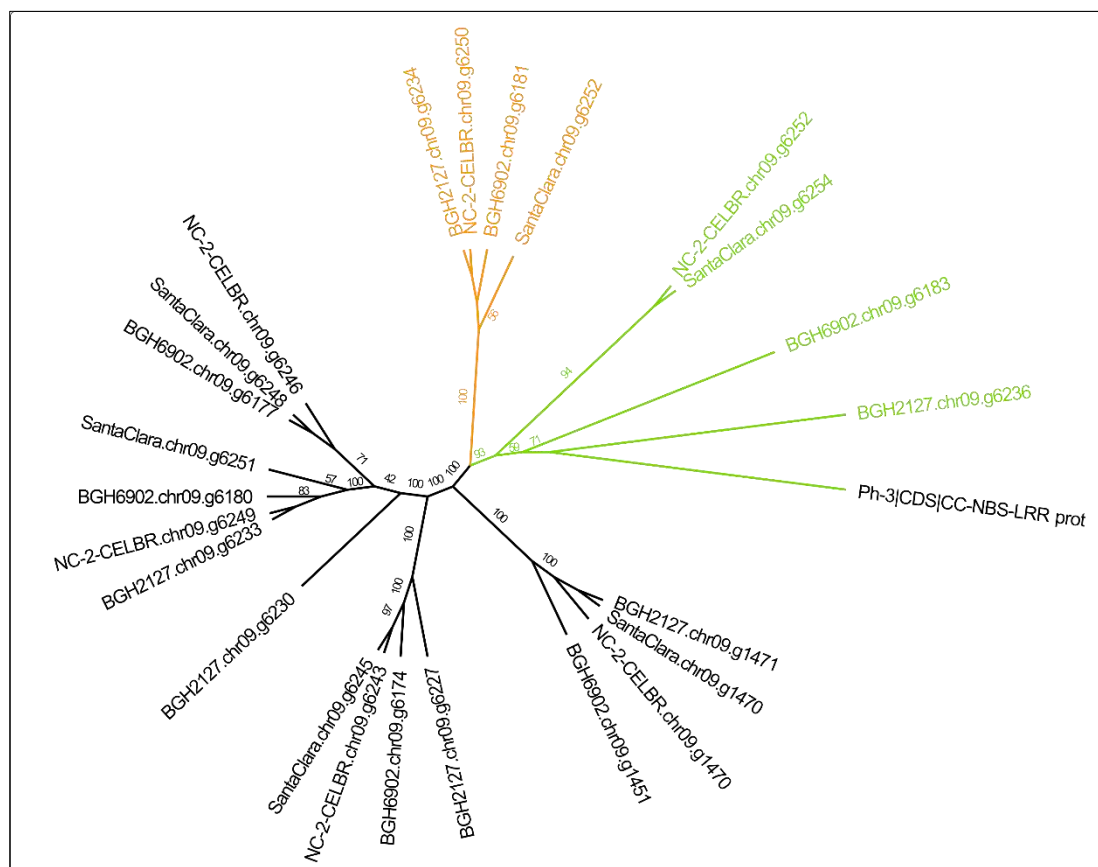


Figura 7. – Filogenia gene *Ph-3*. Cladograma indicando cópias semelhantes (coloridas) do gene *Ph-3*. Gene de referência *Ph-3* agrupado com controle suscetível (cv. Santa Clara) e resistente (NC-2-CELBR) juntamente com acessos BGH no clado verde.

4.5 Polimorfismos

4.5.1 Gene *Ph-2*

4.5.1.1 Clado destacado em verde

O gene g6035 do acesso NC-2-CELBR foi alinhado, devido sua alta similaridade, aos demais genes do clado destacado em verde no cladograma (Figura 6). Nos acessos de *S. lycopersicum* nenhum polimorfismo foi encontrado simultaneamente nos acessos resistentes, NC-2-CELBR e BGH-2127, em contraste ao controle susceptível Santa Clara.

Foram encontrados 6 polimorfismos de base única (Figura 8) simultaneamente nos acessos resistentes NC-2-CELBR (*S. lycopersicum*) e BGH-6902 (*S. habrochaites*) em contraste ao acesso susceptível Santa Clara

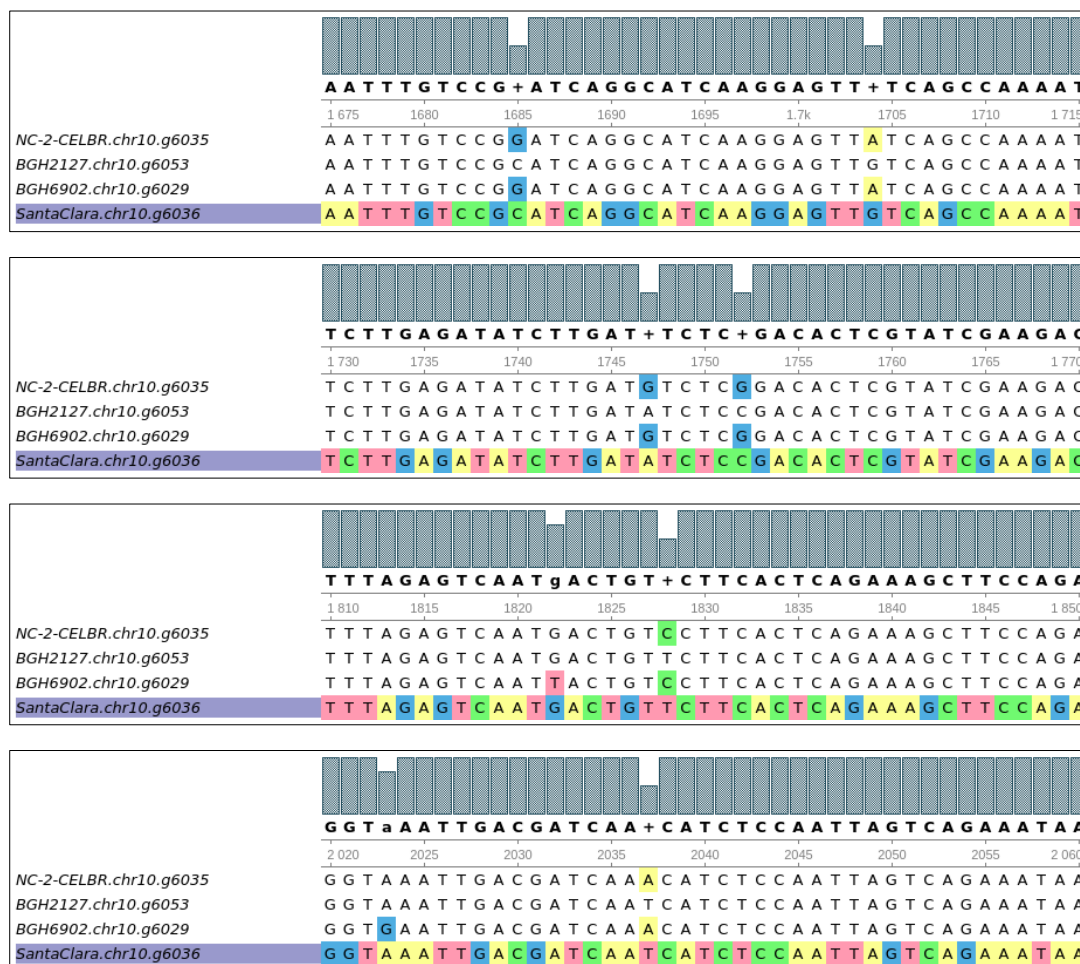


Figura 8. - Alinhamento e polimorfismos em relação ao controle susceptível Santa Clara.

Dos seis polimorfismos quatro foram não sinônimos (Figura 9).

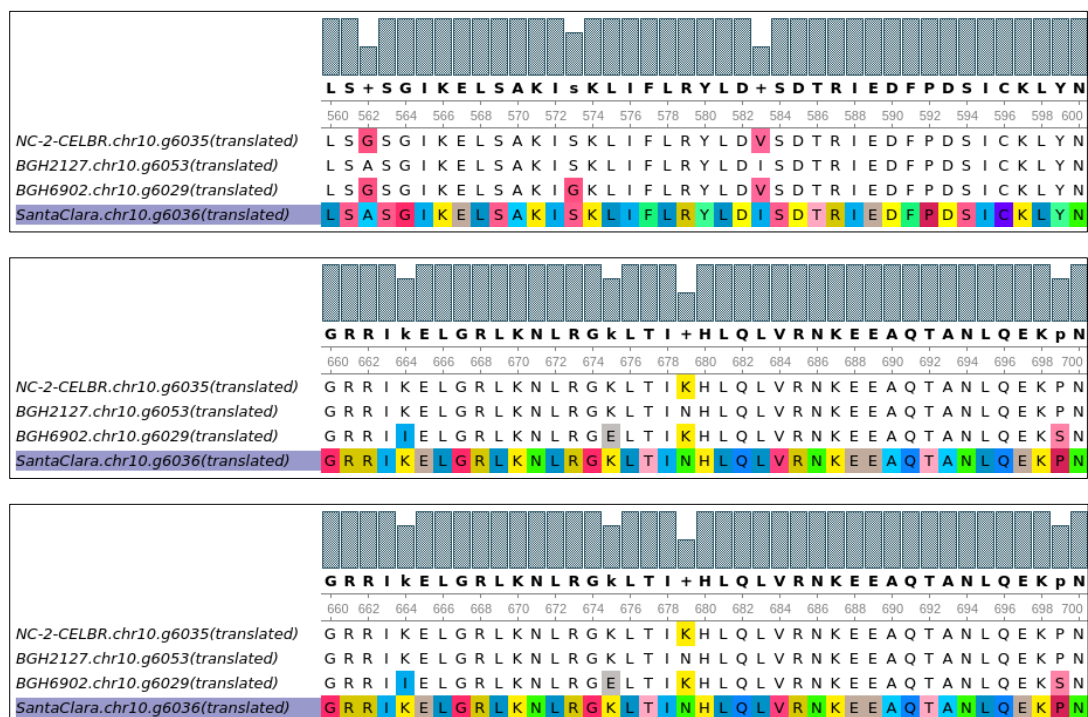


Figura 9. - Alinhamento traduzido e polimorfismos não sinônimos em relação ao controle susceptível Santa Clara.

4.5.1.2 Clado destacado em laranja

No primeiro cladograma (Figura 6) foram agrupadas, em laranja, duas cópias do gene *Ph-2* do acesso NC-2-CELBR. Entretanto a cópia g4751 é 2,6kb menor que o gene g6035 (Figura 10). Os genes g4734, g4751 e g4735 dos acessos NC-2-CELBR, BGH-2127 e Santa Clara, respectivamente, não apresentaram polimorfismos. Já o gene g6035 do acesso NC-2CELBR (*S. lycopersicum*) e o g6028 do acesso BGH-6902 (*S. habrochaites*) compartilharam mais de 200 polimorfismos em relação ao controle Santa Clara, e grande parte desses polimorfismos foram não sinônimos (Figura 11).



Figura 11. - Alinhamento traduzido e polimorfismos não sinônimos em relação ao controle susceptível Santa Clara.

Os demais clados apresentaram genes conservados sem polimorfismos, porém significativamente menores que o gene de referência.

4.5.1.3 Marcadores moleculares para o gene *Ph-2*

Com base nos polimorfismos de interesse encontrados no acesso BGH-6902, foram desenhados primers para futura validação (Tabela 4).

Tabela 4. - Primers acesso BGH-6902.

Oligo	start	len	tm	GC%	Any	3'	seq
ESQ	1132	20	60.08	45.0	8.00	2.0	GAATGGCAGGCAATTCTTGT
DIR	1349	20	59.96	55.0	4.00	2.0	ACAGGGGTCTCTTGACATGG

4.5.2 Gene *Ph-3*

4.5.2.1 Clado destacado em verde

Cinco polimorfismos foram simultaneamente detectados ao comparar os acessos resistentes, NC-2-CELBR, BGH-2127 e BGH-6902, e o controle

susceptível Santa Clara (Figura 12). Destes polimorfismos quatro foram sinônimos (Figura 13).

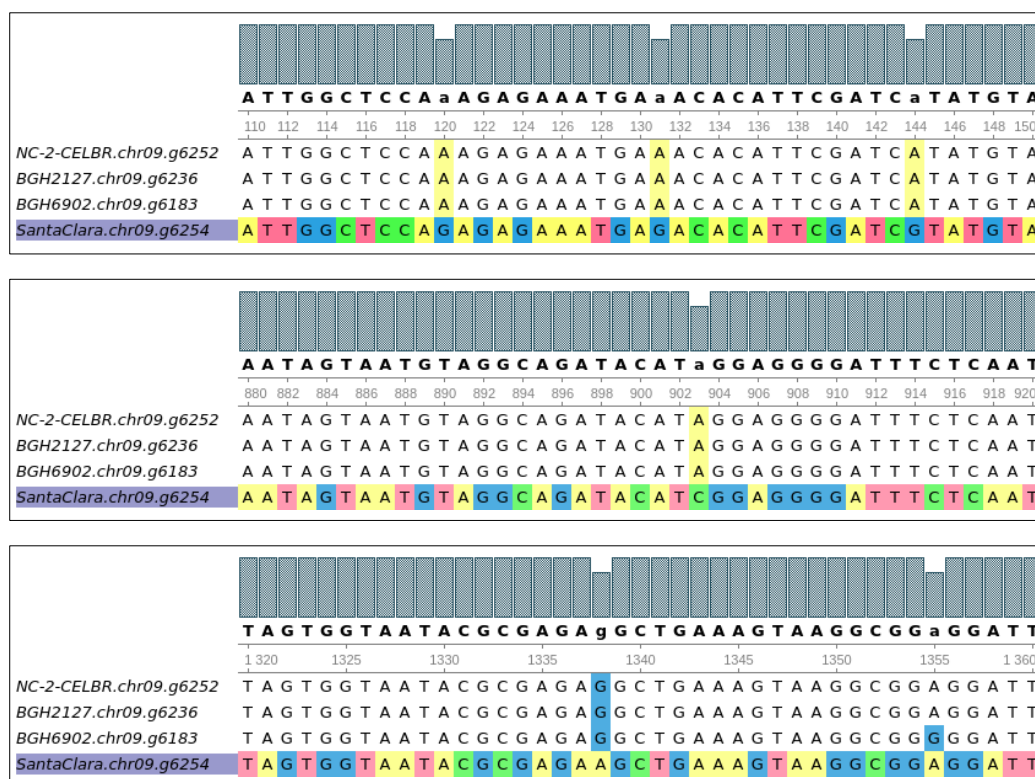


Figura 12. - Alinhamento e polimorfismos em relação ao controle susceptível Santa Clara.

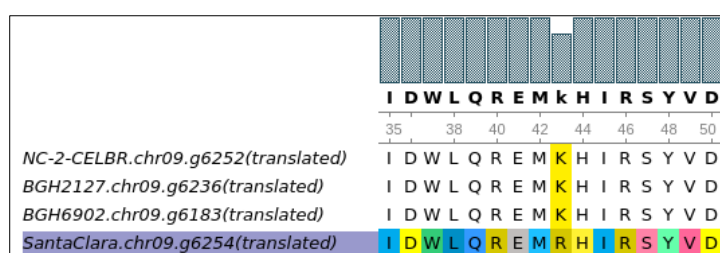


Figura 13. - Alinhamento traduzido e polimorfismos não sinônimos em relação ao controle susceptível Santa Clara.

4.5.2.2 Clado destacado em laranja

Nas cópias do gene *Ph-3* do clado laranja apresentados no segundo cladograma (Figura 7) foi detectado apenas um polimorfismo, compartilhado entre os acessos resistentes, NC-2-CELBR, BGH-2127 e BGH-6902, em comparação ao controle susceptível Santa Clara (Figura 14), esse polimorfismo foi sinônimo.

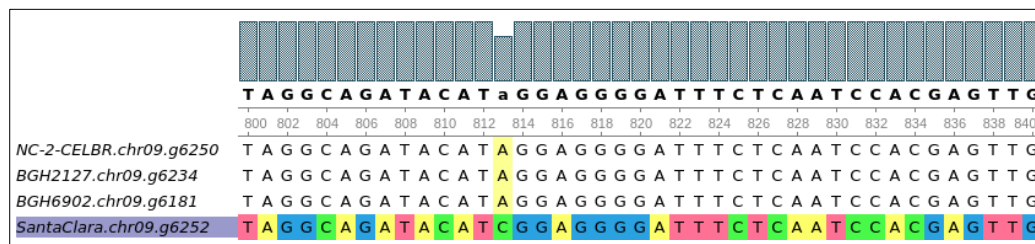


Figura 14. - Alinhamento e polimorfismos em relação ao controle susceptível Santa Clara.

4.5.2.3 Marcadores moleculares para o gene *Ph-3*

Assim como no gene *Ph-2*, foram desenhados marcadores com base nos polimorfismos de interesse encontrados nos acessos BGH-6902 e BHG-2127, foram desenhados primers para futura validação (Tabela 5)

Tabela 5. - Primers acessos BGH-6902 e BGH-2127.

Oligo	start	len	tm	GC%	Any	3'	seq
ESQ	1764	20	60.02	50.0	2.00	0.0	CTTCTTTCACACCGAGCACA
DIR	1951	20	59.98	50.0	4.00	1.0	CATGACGAGCCAGTTGAGAA

5 DISCUSSÃO

Depois do surto de requeima na década de 1940, afetando severamente plantações de batata e tomate, fontes de resistência a essa doença começaram, então, a ser procuradas. O primeiro gene descoberto foi o *Ph-1*, resistente as primeiras estirpes de *P. infestans*, T-0 (BONDE; MURPHY, 1952; PEIRCE, 1971), esse gene estava presente em acessos de *S. pimpinellifolium* e até hoje é uma espécie importante na obtenção de fontes de resistência a esse patógeno. O *Ph-1* foi localizado, utilizando marcadores morfológicos, no braço longo do cromossomo 7 e apresentava dominância completa (PEIRCE, 1971).

5.1 Gene *Ph-2*

O *Ph-2* foi incorporado em uma variedade comercial de tomate, NC-2-CELBR (GARDNER; PANTHEE, 2010) e ao ser comparado com os acessos do BGH-UFV demonstrou não se agrupar ao lócus de referência do gene descrito (Figura 6 – clado

destacado em verde). Entretanto ao compará-lo aos genes desse clado, o alinhamento se deu com alta similaridade, e um dos acessos, BGH-6902, compartilhou inúmeros polimorfismos não sinônimos com NC-2-CELBR em relação ao controle susceptível Santa Clara (Figura 11), indicando a presença do *Ph-2* nesse indivíduo. Apesar do acesso BGH-6902 ser da espécie *S. habrochaites*, essa espécie também é fonte de resistência a requeima, estudos conduzidos com o acesso LA1777, também dessa espécie, indicaram a presença de múltiplos genes de resistência (LI et al., 2011).

O acesso BGH-2127 não apresentou nenhum polimorfismo comparado ao controle susceptível, e por tanto reforça que não há a presença desse gene no indivíduo.

5.2 Gene *Ph-3*

Relatos de isolados de *P. infestans* quebrando a resistência dos *Ph-1* e *Ph-2* na Ásia levaram ao descobrimento de mais um gene o *Ph-3*. Esse gene foi encontrado em outra espécie de *S. pimpinellifolium*, LA1269 (ARAFÁ et al., 2017b). Localizado no braço longo do cromossomo 9 (ZHANG et al., 2013), esse locus também é responsável pela produção de uma proteína de reconhecimento fitopatológico (CC-NBS-LRR). O gene apresenta dominância parcial, conferindo resistência a um amplo número de isolados de *P. infestans*. O *Ph-3* também foi incorporado no acesso NC-2-CELBR (GARDNER; PANTHEE, 2010), utilizando de uma estratégia de piramidização de resistência, conferindo maior durabilidade a esses alelos.

O locus de resistência foi encontrado em quatro homólogos (Figura 7) no DNA dos acessos, esses homólogos se mostraram altamente conservados, e não tiveram polimorfismos não sinônimos em suas sequências. Porém, esse fato leva a crer que há a possibilidade que esses genes também estejam ativos durante o processo de infecção pelo patógeno e que sejam, portanto, susceptíveis a alterações que poderiam levar ao surgimento de novos alelos de resistência.

Entre os genes do clado verde do segundo cladograma (Figura 7) os acessos do BGH-UFV compartilharam alguns dos polimorfismos do acesso NC-2-CELBR em comparação ao controle susceptível (Figura 12). No entanto, apenas um

desses polimorfismos foi não sinônimo (Figura 13). Esse polimorfismo pode desencadear mudanças estruturais na proteína levando ao fenótipo conferido por *Ph-3*. O emprego dos marcadores moleculares específicos desenvolvidos podem evidenciar essa associação e presença do alelo, validando-o.

5.3 Estudo de associação genômica ampla

A análise de componentes principais indicou agrupamentos claros em relação as espécies de tomate analisadas. O uso do modelo agrupado por espécies, no cálculo de associação entre os fenótipos resistentes e os SNPs, resultou no melhor ajuste do Q-Q plot (Figura 4). Esse fato pode ser causado devido ao pequeno espaço amostral (89 amostras e 5 resistentes), nenhuma informação de parentesco entre os acessos, e um grande número de SNPs (aproximadamente 500 mil) para essa análise. Jeng et al. (2016) ressalta que o tamanho amostral influencia negativamente na caracterização de alelos raros, onde pequenas amostras exigem maiores proporções de indivíduos afetados e sugere o método de controle adaptativo de falso negativo (AFNC). Paralelamente, nesses casos não é recomendado filtrar as variantes devido à alta probabilidade de se excluir as variantes causais das alterações fenotípicas.

O Manhattan plot (Figura 5) mostrou cinco variantes estatisticamente significativas, utilizando o AFNC. Essas variantes não estiveram localizadas dentro das CDS do genoma descrito, apesar de estarem localizadas em cromossomos onde já foram caracterizados alelos de resistência. O estudo de associação necessita, então, ser refinado, trabalhando com um espaço amostral maior, e possivelmente com plantas recombinantes provenientes do cruzamento entre resistente e susceptível.

6 CONCLUSÕES

Os polimorfismos encontrados nos acessos do BGH-UFV indicam a presença dos genes *Ph-2* e *Ph-3* no acesso BGH-6902, e do gene *Ph-3* no acesso BGH-2127. A validação dos marcadores (Tabela 4, e Tabela 5) desenvolvidos se faz necessária para confirmação da presença desses genes.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ALTSCHUL, S. F. et al. Basic local alignment search tool. **Journal of molecular biology**, v. 215, n. 3, p. 403–410, 1990.
2. ANDREWS, S. Babraham bioinformatics-FastQC a quality control tool for high throughput sequence data. URL: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>, 2010.
3. ARAFA, R. A. et al. Rapid identification of candidate genes for resistance to tomato late blight disease using next-generation sequencing technologies. **PLoS One**, v. 12, n. 12, p. e0189951, 2017a.
4. ARAFA, R. A. et al. Resistance to *Phytophthora infestans* in tomato wild relatives. **African Journal of Agricultural Research**, v. 12, n. 26, p. 2188–2196, 2017b.
5. ASHRAFI, H.; KINKADE, M.; FOOLAD, M. R. A new genetic linkage map of tomato based on a *Solanum lycopersicum* × *S. pimpinellifolium* RIL population displaying locations of candidate pathogen response genes. **Genome**, v. 52, n. 11, p. 935–956, 2009.
6. AYLOR, D. E. Spread of plant disease on a continental scale: role of aerial dispersal of pathogens. **Ecology**, v. 84, n. 8, p. 1989–1997, 2003.
7. BASHI, E.; BEN-JOSEPH, Y.; ROTEM, J. Inoculum potential of *Phytophthora infestans* and the development of potato late blight epidemics. **Phytopathology**, v. 72, n. 8, p. 1043–1047, 1982.
8. BIRCH, P. R. J. et al. Trafficking arms: oomycete effectors enter host plant cells. **Trends in microbiology**, v. 14, n. 1, p. 8–11, 2006.
9. BLACK, L. L. et al. Late blight resistance in four wild tomato accessions: effectiveness in diverse locations and inheritance of resistance. **Phytopathology**, v. 86, n. 11, p. S24, 1996.
10. BOLGER, A. et al. The genome of the stress-tolerant wild tomato species *Solanum pennellii*. **Nature genetics**, v. 46, n. 9, p. 1034–1038, 2014.
11. BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114–2120, 2014.
12. BONDE, R.; MURPHY, E. F. Resistance of certain tomato varieties and crosses to late blight. 1952.

13. BONIERBALE, M. W.; PLAISTED, R. L.; TANKSLEY, S. D. RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. **Genetics**, v. 120, n. 4, p. 1095–1103, 1988.
14. BRANDWAGT, B. F. et al. A longevity assurance gene homolog of tomato mediates resistance to *Alternaria alternata* f. sp. *lycopersici* toxins and fumonisin B1. **Proceedings of the National Academy of Sciences**, v. 97, n. 9, p. 4961–4966, 2000.
15. BROUWER, D. J.; CLAIR, D. A. S. Fine mapping of three quantitative trait loci for late blight resistance in tomato using near isogenic lines (NILs) and sub-NILs. **Theoretical and Applied Genetics**, v. 108, n. 4, p. 628–638, 2004.
16. CONSORTIUM, 100 TOMATO GENOME SEQUENCING et al. Exploring genetic variation in the tomato (*Solanum* section *Lycopersicon*) clade by whole-genome sequencing. **The Plant Journal**, v. 80, n. 1, p. 136–148, 2014.
17. DIXON, M. S. et al. The tomato Cf-5 disease resistance gene and six homologs show pronounced allelic variation in leucine-rich repeat copy number. **The Plant Cell**, v. 10, n. 11, p. 1915–1925, 1998.
18. DUNCAN, J. M. Effect of fungicides on survival, infectivity and germination of *Phytophthora fragariae* oospores. **Transactions of the British Mycological Society**, v. 85, n. 4, p. 585–593, 1985.
19. EDGAR, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic acids research**, v. 32, n. 5, p. 1792–1797, 2004.
20. ERNST, K. et al. The broad-spectrum potato cyst nematode resistance gene (Hero) from tomato is the only member of a large gene family of NBS-LRR genes with an unusual amino acid repeat in the LRR region. **The Plant Journal**, v. 31, n. 2, p. 127–136, 2002.
21. FOOLAD, M. R. Genome mapping and molecular breeding of tomato. **International journal of plant genomics**, v. 2007, 2007.
22. FRY, W. E. et al. Five reasons to consider *Phytophthora infestans* a reemerging pathogen. **Phytopathology**, v. 105, n. 7, p. 966–981, 2015.
23. FRY, W. E. *Phytophthora infestans*: New tools (and old ones) lead to new understanding and precision management. **Annual review of phytopathology**, v. 54, p. 529–547, 2016.

24. GALLEGLY, M. E.; MARVEL, M. E. Inheritance of resistance to tomato race 0 of *Phytophthora infestans*. **Phytopathology**, v. 45, n. 2, p. 103–109, 1955.
25. GAO, L. et al. The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. **Nature genetics**, v. 51, n. 6, p. 1044–1051, 2019.
26. GARDNER, R. G.; PANTHEE, D. R. NC 1 CELBR and NC 2 CELBR: Early blight and late blight-resistant fresh market tomato breeding lines. **HortScience**, v. 45, n. 6, p. 975–976, 2010.
27. GARRISON, E.; MARTH, G. Haplotype-based variant detection from short-read sequencing. **arXiv preprint arXiv:1207.3907**, 2012.
28. GOODWIN, S. B.; SUJKOWSKI, L. S.; FRY, W. E. Rapid evolution of pathogenicity within clonal lineages of the potato late blight disease fungus. **Phytopathology**, v. 85, n. 6, p. 669–676, 1995.
29. GRANKE, L. L. et al. Dispersal and movement mechanisms of *Phytophthora capsici* sporangia. **Phytopathology**, v. 99, n. 11, p. 1258–1264, 2009.
30. HAVERKORT, A. J. et al. Applied biotechnology to combat late blight in potato caused by *Phytophthora infestans*. **Potato Research**, v. 52, n. 3, p. 249–264, 2009.
31. JENG, X. J. et al. Rare variants association analysis in large-scale sequencing studies at the single locus level. **PLoS computational biology**, v. 12, n. 6, p. e1004993, 2016.
32. KAWCHUK, L. M. et al. Tomato Ve disease resistance genes encode cell surface-like receptors. **Proceedings of the National Academy of Sciences**, v. 98, n. 11, p. 6511–6515, 2001.
33. KORTE, A.; FARLOW, A. The advantages and limitations of trait analysis with GWAS: a review. **Plant methods**, v. 9, n. 1, p. 1–9, 2013.
34. LANGMEAD, B.; SALZBERG, S. L. Fast gapped-read alignment with Bowtie 2. **Nature methods**, v. 9, n. 4, p. 357–359, 2012.
35. LAURINDO, B. S. **Análise proteômica da resistência à requeima em tomateiro**. Viçosa - MG: [s.n.].
36. LAURINDO, B. S. et al. Comparative analysis of constitutive proteome between resistant and susceptible tomato genotypes regarding to late blight. **Functional & integrative genomics**, v. 18, n. 1, p. 11–21, 2018.

37. LI, H. et al. The sequence alignment/map format and SAMtools. **Bioinformatics**, v. 25, n. 16, p. 2078–2079, 2009.
38. LI, J. et al. Identification and mapping of quantitative resistance to late blight (*Phytophthora infestans*) in *Solanum habrochaites* LA1777. **Euphytica**, v. 179, n. 3, p. 427–438, 2011.
39. LIN, T. et al. Genomic analyses provide insights into the history of tomato breeding. **Nature genetics**, v. 46, n. 11, p. 1220–1226, 2014.
40. MARCHINI, J.; HOWIE, B. Genotype imputation for genome-wide association studies. **Nature Reviews Genetics**, v. 11, n. 7, p. 499–511, 2010.
41. MERK, H. L.; ASHRAFI, H.; FOOLAD, M. R. Selective genotyping to identify late blight resistance genes in an accession of the tomato wild species *Solanum pimpinellifolium*. **Euphytica**, v. 187, n. 1, p. 63–75, 2012.
42. MILLIGAN, S. B. et al. The root knot nematode resistance gene *Mi* from tomato is a member of the leucine zipper, nucleotide binding, leucine-rich repeat family of plant genes. **The Plant Cell**, v. 10, n. 8, p. 1307–1319, 1998.
43. MINH, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. **Molecular biology and evolution**, v. 37, n. 5, p. 1530–1534, 2020.
44. MIZUBUTI, E. S. G.; AYLOR, D. E.; FRY, W. E. Survival of *Phytophthora infestans* sporangia exposed to solar radiation. **Phytopathology**, v. 90, n. 1, p. 78–84, 2000.
45. MIZUBUTI, E. S. G.; JÚNIOR, V. L.; FORBES, G. A. Management of late blight with alternative products. **Pest technology**, v. 1, n. 2, p. 106–116, 2007.
46. MOREAU, P. et al. Genetic mapping of *Ph-2*, a single locus controlling partial resistance to *Phytophthora infestans* in tomato. **Molecular Plant-Microbe Interactions**, v. 11, n. 4, p. 259–269, 1998.
47. MUELLER, L. et al. A snapshot of the emerging tomato genome sequence. **Plant Genome**, v. 2, n. 1, p. 78–92, 2009.
48. NOWICKI, M. et al. Potato and tomato late blight caused by *Phytophthora infestans*: an overview of pathology and resistance breeding. **Plant disease**, v. 96, n. 1, p. 4–17, 2012.

49. ORI, N. et al. The I2C family from the wilt disease resistance locus I2 belongs to the nucleotide binding, leucine-rich repeat superfamily of plant resistance genes. **The Plant Cell**, v. 9, n. 4, p. 521–532, 1997.
50. PARNISKE, M. et al. Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the Cf-4/9 locus of tomato. **Cell**, v. 91, n. 6, p. 821–832, 1997.
51. PEIRCE, L. C. Linkage tests with Ph conditioning resistance to race 0, *Phytophthora infestans*. **Rep. Tomato Genet. Coop**, v. 21, p. 30, 1971.
52. PHAN, N. T. et al. Identification of loci associated with fruit traits using genome-wide single nucleotide polymorphisms in a core collection of tomato (*Solanum lycopersicum* L.). **Scientia Horticulturae**, v. 243, p. 567–574, 2019.
53. PORTER, L. D.; DASGUPTA, N.; JOHNSON, D. A. Effects of tuber depth and soil moisture on infection of potato tubers in soil by *Phytophthora infestans*. **Plant Disease**, v. 89, n. 2, p. 146–152, 2005.
54. PURCELL, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. **The American journal of human genetics**, v. 81, n. 3, p. 559–575, 2007.
55. QI, T. et al. Host-induced gene silencing: a powerful strategy to control diseases of wheat and barley. **International journal of molecular sciences**, v. 20, n. 1, p. 206, 2019.
56. RAMBAUT, A. FigTree v1. 4. Molecular evolution, phylogenetics and epidemiology. **Edinburgh: University of Edinburgh, Institute of Evolutionary Biology**, 2012.
57. RAMIREZ-GONZALEZ, R. H.; UAUY, C.; CACCAMO, M. PolyMarker: a fast polyploid primer design pipeline. **Bioinformatics**, v. 31, n. 12, p. 2038–2039, 2015.
58. READER, J. The fungus that conquered Europe. **The New York Times, London**, 2008.
59. ROZEN, S.; SKALETSKY, H. Primer3 on the WWW for general users and for biologist programmers. In: **Bioinformatics methods and protocols**. [s.l.] Springer, 2000. p. 365–386.

60. SCARANO, D.; RAO, R.; CORRADO, G. In Silico identification and annotation of non-coding RNAs by RNA-seq and De Novo assembly of the transcriptome of Tomato Fruits. **PloS one**, v. 12, n. 2, p. e0171504, 2017.
61. SCHIAVON, M. et al. Selenium fertilization alters the chemical composition and antioxidant constituents of tomato (*Solanum lycopersicon* L.). **Journal of Agricultural and Food Chemistry**, v. 61, n. 44, p. 10542–10554, 2013.
62. SCHORNACK, S. et al. The tomato resistance protein Bs4 is a predicted non-nuclear TIR-NB-LRR protein that mediates defense responses to severely truncated derivatives of AvrBs4 and overexpressed AvrBs3. **The Plant Journal**, v. 37, n. 1, p. 46–60, 2004.
63. SHIRASAWA, K. et al. Genome-wide association studies using single nucleotide polymorphism markers developed by re-sequencing of the genomes of cultivated tomato. **DNA research**, v. 20, n. 6, p. 593–603, 2013.
64. SIDRA, I. Censo Agropecuário, 2017. **Acesso em junho de 2021**, v. 4, n. 10, 2021.
65. STANKE, M. et al. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. **Bioinformatics**, v. 24, n. 5, p. 637–644, 2008.
66. STAT, F. A. O. Food and Agriculture Organization of the United Nations: Statistics Division. **CropAvailable online: <http://www.fao.org/faostat/en/#data/QC> (acesso em Janeiro 2021)**, 2016.
67. TEAM, R. C. R: A language and environment for statistical computing. 2013.
68. TIEMAN, D. et al. A chemical genetic roadmap to improved tomato flavor. **Science**, v. 355, n. 6323, p. 391–394, 2017.
69. TOMATO GENOME CONSORTIUM, XX. The tomato genome sequence provides insights into fleshy fruit evolution. **Nature**, v. 485, n. 7400, p. 635, 2012.
70. TURKENSTEEN, L. J. et al. Production, survival and infectivity of oospores of *Phytophthora infestans*. **Plant Pathology**, v. 49, n. 6, p. 688–696, 2000.
71. TURNER, S. D. qqman: an R package for visualizing GWAS results using QQ and manhattan plots. **Biorxiv**, p. 005165, 2014.
72. UNIVERSIDADE FEDERAL DE VIÇOSA. **Banco de Germoplasma de Hortaliças da Universidade Federal de Viçosa**.

73. VANHAUTE, E.; PAPIING, R.; Ó GRÁDA, C. **The European subsistence crisis of 1845-1850: a comparative perspective**. [s.l.] UCD Centre for Economic Research Working Paper Series, 2006.
74. VOGEL, C.; MARCOTTE, E. M. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. **Nature reviews genetics**, v. 13, n. 4, p. 227–232, 2012.
75. WICKHAM, H. The tidyverse. **R package ver**, v. 1, n. 1, p. 836, 2017.
76. XIANG, Q.; JUDELSON, H. S. Myb transcription factors and light regulate sporulation in the oomycete *Phytophthora infestans*. **PloS one**, v. 9, n. 4, p. e92086, 2014.
77. ZHANG, C. et al. Fine mapping of the Ph-3 gene conferring resistance to late blight (*Phytophthora infestans*) in tomato. **Theoretical and applied genetics**, v. 126, n. 10, p. 2643–2653, 2013.
78. ZHANG, C. et al. The Ph-3 gene from *Solanum pimpinellifolium* encodes CC-NBS-LRR protein conferring resistance to *Phytophthora infestans*. **Theoretical and applied genetics**, v. 127, n. 6, p. 1353–1364, 2014.
79. ZHI, X. et al. Fine Mapping of the Ph-2 Gene Conferring Resistance to Late Blight (*Phytophthora infestans*) in Tomato. **Plant Disease**, v. 105, n. 4, p. 851–858, 2021.
80. ZHU, G. et al. Rewiring of the fruit metabolome in tomato breeding. **Cell**, v. 172, n. 1–2, p. 249–261, 2018.