

NINA IRIS VERSLYPE

**AVALIAÇÃO E SELEÇÃO DE PORTA-ENXERTOS DE VIDEIRA (*Vitis*
spp.) TOLERANTES AO DÉFICIT HÍDRICO ATRAVÉS DE
APRENDIZAGEM DE MÁQUINA**

RECIFE – PE

2021

NINA IRIS VERSLYPE

**AVALIAÇÃO E SELEÇÃO DE PORTA-ENXERTOS DE VIDEIRA (*Vitis*
spp.) TOLERANTES AO DÉFICIT HÍDRICO ATRAVÉS DE
APRENDIZAGEM DE MÁQUINA**

Dissertação apresentada ao Programa de Pós Graduação em Agronomia – Melhoramento Genético de Plantas, da Universidade Federal Rural de Pernambuco, como parte dos requisitos para obtenção do título de Mestre em Agronomia - Melhoramento Genético de Plantas.

Linha de Pesquisa: Biometria e Melhoramento Vegetal

Orientadora: Profa. Dra. Rosimar dos Santos Musser

Co-orientadores: Prof. Dr. André Câmara Alves do Nascimento

Dra. Patrícia Coelho de Souza Leão

RECIFE – PE

2021

Dados Internacionais de Catalogação na Publicação
Universidade Federal Rural de Pernambuco
Sistema Integrado de Bibliotecas
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

V563a

Verslype, Nina Iris

AVALIAÇÃO E SELEÇÃO DE PORTA-ENXERTOS DE VIDEIRA (*Vitis* spp.) TOLERANTES AO DÉFICIT HÍDRICO ATRAVÉS DE APRENDIZAGEM DE MÁQUINA / Nina Iris Verslype. - 2021.
140 f. : il.

Orientadora: Rosimar dos Santos Musser.

Coorientador: Andre Camara Alves do Nascimento.

Inclui referências e anexo(s).

Dissertação (Mestrado) - Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Agronomia - Melhoramento Genético de Plantas, Recife, 2021.

1. Mudanças climáticas. 2. Algoritmo supervisionado. 3. Algoritmo não supervisionado. 4. Melhoramento vegetal. 5. Inteligência artificial. I. Musser, Rosimar dos Santos, orient. II. Nascimento, Andre Camara Alves do, coorient. III. Título

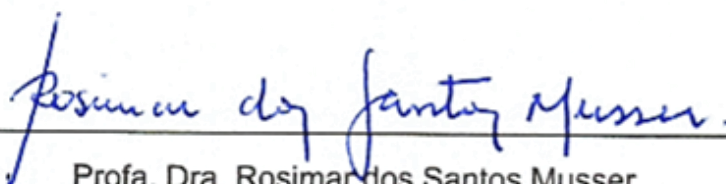
CDD 581.15

NINA IRIS VERSLYPE

**AVALIAÇÃO E SELEÇÃO DE PORTA-ENXERTOS DE VIDEIRA (*Vitis* spp.)
TOLERANTES AO DÉFICIT HÍDRICO ATRAVÉS DE APRENDIZAGEM DE
MÁQUINA**

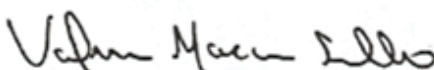
Dissertação defendida e aprovada pela banca examinadora em: 15 de julho de 2021

Banca examinadora:



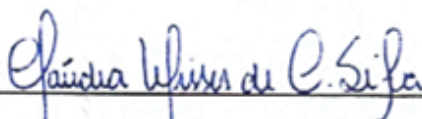
Profa. Dra. Rosimar dos Santos Musser

(Membro Titular Interno – Área de Fitotecnia/DEPA/UFRPE)



Prof. Dr. Valmir Macario Filho

(Membro Titular Externo – Departamento de Computação/UFRPE)



Profa. Dra. Cláudia Ulisses de Carvalho Silva

(Membro Titular Interno – Área de Botânica/DB/UFRPE)

RECIFE – PE

JULHO DE 2021

Dedico este trabalho ao meu pai Christophe Verslype, e aos meus avós, Germain Verslype e Monique Verslype, por todo amor, inspiração, apoio e incentivo.

AGRADECIMENTOS

À minha família, por todo carinho, amor, conselhos, incentivo e apoio nas minhas escolhas, sempre orientando, auxiliando e por sempre acreditarem nos meus objetivos.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão de bolsa de estudos de mestrado.

À Universidade Federal Rural de Pernambuco (UFRPE) e ao Programa de Pós-Graduação em Agronomia - Melhoramento Genético de Plantas (PPGAMGP), por toda a sua estrutura e corpo docente.

A todos os docentes que contribuíram na minha formação. Em especial agradeço a minha orientadora, professora Dra. Rosimar dos Santos Musser pela oportunidade, ensinamentos, conselhos, paciência, aprendizados e por confiar na minha capacidade. Assim como, ao meu co-orientador, professor Dr. André Câmara A. do Nascimento pela paciência, apoio, confiança, disposição, aprendizados e auxílio na condução das análises do trabalho, que contribuíram na minha formação e ampliaram os meus horizontes como engenheira agrônoma.

À Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) Semiárido, por toda a estrutura, aos pesquisadores, aos funcionários e aos colegas, por sua receptividade, aprendizados e experiências. Em especial agradeço a minha co-orientadora Dra. Patrícia C. de S. Leão e ao pesquisador Dr. Agnaldo R. de M. Chaves pela oportunidade.

À empresa Petromudas, por sua recepção e disponibilidade.

Aos meus amigos pela sua companhia, incentivo e amizade durante toda esta fase.

Por fim, a todos aqueles que contribuíram, diretamente ou indiretamente para a concretização deste trabalho.

Muito obrigada!

“Em meio à dificuldade encontra-se a oportunidade.”

Albert Einstein

“Grandes realizações não são feitas por impulso, mas por uma soma de pequenas realizações.”

Vincent Van Gogh

LISTA DE SÍMBOLOS

Ψ_{os} - Potencial osmótico

LISTA DE ABREVIATURAS

A - Taxa de assimilação de dióxido de carbono

ABA - Ácido abscísico

AI – Inteligência artificial

BAG – Banco Ativo de Germoplasma

DT – *Decision Tree*

E - Taxa de transpiração

EIUA - Eficiência instantânea do uso da água

EUA - Eficiência intrínseca do uso da água

gs - Condutância estomática

IPC - *Chlorose Power Index*

Kc - Coeficiente de cultura

KNN - *K-Nearest Neighbors*

LDA - *Linear Discriminant Analysis*

ML – *Machine learning*

Mpa - Mega Pascal

P - Precisão

R - *Recall*

RF – *Random Forest*

SVM – *Support Vector Machines*

XGB - *XGBoost*

LISTA DE TABELAS

CAPÍTULO I - Introdução e referencial teórico

Tabela 1 - Descrição agrônômica das quarenta e cinco cultivares de porta-enxertos de <i>Vitis</i> spp e híbridos empregadas no estudo.....	28
--	----

CAPÍTULO II - Cultivares de porta-enxertos de videira: conjunto de dados descrevendo características fisiológicas, bioquímicas e agrônômicas

Tabela 1 – Origem genética das quarenta e cinco cultivares de porta-enxertos.	81
Tabela 2 – Descrição das variáveis.....	82
Tabela 3 - Estatísticas descritivas para as variáveis categóricas do conjunto de dados.	84
Tabela 4 - Estatísticas descritivas para as variáveis inteiras e numéricas do conjunto de dados.....	84

CAPÍTULO III - Análise da diversidade genética de porta-enxertos de videira através de aprendizado de máquina

Tabela 1 - Quarenta e cinco cultivares de porta-enxertos de videiras (<i>Vitis</i> spp.) avaliadas.	96
Tabela 2 - Distribuição das cultivares de porta-enxertos de videira (<i>Vitis</i> spp) dentro dos cinco agrupamentos.	105
Tabela 3 - Cruzamentos viáveis entre porta-enxertos sugeridos com base na matriz de similaridade genética.....	109

CAPÍTULO IV - Classificação de porta-enxertos de videira tolerantes ao déficit hídrico através de aprendizagem de máquina

Tabela 1 - Quarenta e cinco cultivares de porta-enxertos de videira, origem genética e grau de tolerância ao déficit hídrico.....	118
Tabela 2 – Hiperparâmetros analisados pelo <i>grid search</i> em cada algoritmo.	122
Tabela 3 - Avaliação da acurácia, precisão (P), <i>recall</i> (R) e <i>f1-score</i> dos seis algoritmos obtidos pela validação cruzada.	123
Tabela 4 - Avaliação dos melhores hiperparâmetros para cada modelo obtida pelo <i>Grid search</i>	124
Tabela 5 – Relatório de classificação da validação cruzada com 10 <i>fold</i> do algoritmo RF.	126

ANEXO

Tabela 1 - Parâmetros, descrição e referências bibliográficas.	133
---	-----

LISTA DE ILUSTRAÇÕES

CAPÍTULO I - Introdução e referencial teórico

Figura 1 - Localização do Submédio do Vale do São Francisco. Fonte: Verslype (2021).....	24
Figura 2 - Porta-enxertos de videira avaliados: Harmony (1), IAC 313 (2), IAC 572 (3), IAC 766 (4), SO4 (5), 101-14 MGt (6), Ramsey (7), Paulsen 1103 (8), Freedom (9), Richter (10), 1616 C (11), 3309 C (12), 44-53 M (13), 420 A MGt (14), VR 039-16 (15), Riparia Gloire de Montpellier (16), 140 Ru (17), Schwarzmann (18), 5 C (19), 196-17 Cl (20), 216-3 Cl (21), 157-11 C (22), 161-49 C (23), 1202 C (24), 1613 C (25), 3306 C (26), Dogridge (27), 34 EM (28), Fercal (29), Ganzin 1 (30), 26 G (31), Golia (32), Gravesac (33), 125 AA (34), 41 B MGt (35), 106-8 MGt (36), 99 R (37), Sori (38), 5 BB (39), 8B (40), 93-5 C (41), VR 043-43 (42), Rupestris du lot (43), 62-66C (44) e <i>Vitis champini</i> (45). Fonte: Adaptado de FPS (2020), Embrapa Uva e Vinho (2020) e Maul <i>et al.</i> (2020).	31
Figura 3 - Cultivar 1103P. Fonte: Verslype, 2020.....	32
Figura 4 - Cultivar SO4. Fonte: Verslype, 2020.....	32
Figura 5 - Cultivar IAC 572. Fonte: Verslype (2020).....	33
Figura 6 - Cultivar IAC 313. Fonte: Verslype (2020).....	34
Figura 7 - Cultivar IAC 766. Fonte: Verslype (2020).....	34
Figura 8 - Cultivar Ramsey. Fonte: Verslype (2020).	35
Figura 9 - Cultivar Freedom. Fonte: Verslype (2020).	35
Figura 10 - Cultivar 101-14 MGt. Fonte: Verslype (2020).....	36
Figura 11 - Cultivar Harmony. Fonte: Verslype (2020).	36
Figura 12 - Folhas da cultivar Merlot, com sintomas de estresse hídrico (A) e sem sofrer estresse (B). Fonte: Adaptado de Keller (2015).	41
Figura 13 - Respostas da videira quando submetida a curto e longo prazo sob déficit hídrico. Fonte: Adaptado de Chaves <i>et al.</i> (2003) e Serra <i>et al.</i> (2014).	42
Figura 14 - Distribuição do sistema radicular de quatro cultivares de porta-enxertos de videira, sendo 101-14MGt (Millardet et de Grasset 101–14), Riparia (Riparia Gloire de Montpellier), 110R (Richter 110), Ramsey (Salt creek). Fonte: Fort <i>et al.</i> (2017).....	45
Figura 15 - Diferenças no ângulo geotrópico em função da disposição de emergência de raízes adventícias em diferentes cultivares de porta-enxertos, sendo A - Riparia Gloire de Montpellier, B - 3309C e C - Rupestris du Lot. Fonte: Smart <i>et al.</i> (2006). 46	
Figura 16 - Variação no diâmetro de poros estomáticos e no número de estômatos da cultivar Pinotage. Sendo A - 1103 P sob restrições hídricas e exposto a luz solar (119.1 ± 6.3 poros/mm ² e 13.2 ± 0.31 μ m), B - 1103 P sob restrições hídricas e sombreado (91.0 ± 6.3 poros/mm ² e 20.0 ± 0.3 μ m), C – 140 Ru sob restrições hídricas e exposto a luz solar (113.8 ± 6.3 poros/mm ² e 16.0 ± 0.31 μ m) e D – 140 Ru sob restrições hídricas e sombreado (96.3 ± 6.3 poros/mm ² e 17.2 ± 0.31 μ m). Ampliado 150 vezes. Fonte: Serra <i>et al.</i> (2014).	49
Figura 17 – pH da seiva do xilema sob estresse hídrico. Fonte: Taiz and Zieger (2004).	51

Figura 18 - Fechamento hidroativo dos estômatos sob estresse hídrico. Sendo A ⁻ (ânions), SLAC (canal aniônico lento) e RAC (canal aniônico rápido). Fonte: Adaptado de Schachtman and Goodger, 2008.....	52
Figura 19 - Perda percentual de condutividade hidráulica (%) na <i>Vitis vinifera</i> em suas hastes, pecíolos, folhas e a condutância estomática aparente (Gs) em função do potencial da água nas folhas (MPa). Fonte: Charrier <i>et al.</i> (2016).	55
Figura 20 - Aprendizado de máquina como subcampo da Inteligência Artificial. Fonte: Verslype (2021).	57
Figura 21 - Classificação do Aprendizado de máquina. Fonte: Verslype (2021).	59
Figura 22 - Representação da árvore de decisão. Fonte: Verslype (2021).	60
Figura 23 - Representação do funcionamento do algoritmo <i>Random Forest</i> . Fonte: Verslype (2021).	62
Figura 24 - Exemplo de clustering K-means evidenciando os centróides. Fonte: Verslype (2021).	63

CAPÍTULO II - Cultivares de porta-enxertos de videira: conjunto de dados descrevendo características fisiológicas, bioquímicas e agrônômicas

Figura 1 - Diagrama esquemático explicando a organização do conjunto de dados. Fonte: Verslype <i>et al.</i> (2021).	80
Figura 2- Distribuição dos dados categóricos no conjunto de dados. Fonte: Verslype <i>et al.</i> (2021).	86
Figura 3 - Distribuição das variáveis inteiras dentro de cada classe de tolerância ao déficit hídrico. Fonte: Verslype <i>et al.</i> (2021).	87
Figura 4 - Distribuição das variáveis inteiras dentro de cada classe de tolerância ao déficit hídrico. Fonte: Verslype <i>et al.</i> (2021).	88
Figura 5 - Distribuição das variáveis numéricas dentro de cada classe de tolerância à seca. A prolina sob condições ideais não formou a curva KDE devido à ausência de dados em mais de uma classe de tolerância ao déficit hídrico. Fonte: Verslype <i>et al.</i> (2021).	89
Figura 6 - Distribuição das variáveis numéricas dentro de cada classe de tolerância ao déficit hídrico. Fonte: Verslype <i>et al.</i> (2021).	90

CAPÍTULO III - Análise da diversidade genética de porta-enxertos de videira através de aprendizado de máquina

Figura 1 - Diagrama esquemático da análise de agrupamento do conjunto de dados. Fonte: Verslype <i>et al.</i> (2021).	96
Figura 2 - Distribuição do conjunto de dados original e normalizado. Fonte: Verslype <i>et al.</i> (2021).	99
Figura 3 – Matriz de correlação entre as variáveis, onde as cores escuras representam a correlação negativa e as cores claras a correlação positiva. Fonte: Verslype <i>et al.</i> (2021).	101
Figura 4 - Variáveis com forte correlação positiva e negativa. Fonte: Verslype <i>et al.</i> (2021).	103

Figura 5 - Método do Cotovelo para determinar o número de clusters e o algoritmo Kneedle indicando o número ideal de clusters. Fonte: Verslype <i>et al.</i> (2021).	104
Figura 6 - Agrupamentos formados pelo k-means com base no primeiro e segundo componente principal da PCA. No qual os pontos representam os indivíduos dentro dos agrupamentos representados pelas diferentes cores, e em preto os centróides de cada cluster. Fonte: Verslype <i>et al.</i> (2021).	104
Figura 7 - Distribuição das características agronômicas dentro de cada cluster. Fonte: Verslype <i>et al.</i> (2021).	106
Figura 8 - Matriz de similaridade obtida através da distância euclidiana entre as 45 cultivares de porta-enxertos de videira, onde as cores nos tons de vermelho representam genótipos geneticamente próximos e nos tons de azul genótipos geneticamente distantes. Fonte: Verslype <i>et al.</i> (2021).	108

CAPÍTULO IV - Classificação de porta-enxertos de videira tolerantes ao déficit hídrico através de aprendizagem de máquina

Figura 1 - Conjunto de dados antes e depois da normalização. Fonte: Verslype <i>et al.</i> (2021).	121
Figura 2 - Curva de aprendizagem dos seis algoritmos sobre o conjunto de dados, através da validação cruzada com <i>Kfold</i> igual a 10 e métrica <i>score</i> (R^2). Sendo DT - <i>Decision Tree</i> , RF - <i>Random Forest</i> , XGB - <i>XGBoost</i> , SVM - <i>Support Vector Machines</i> , LDA - <i>Linear Discriminant Analysis</i> e KNN - <i>K-Nearest Neighbors</i> . Fonte: Verslype <i>et al.</i> (2021).	125
Figura 3 - Matriz de confusão do algoritmo RF. Onde 0 - representa a classe de alta tolerância ao déficit hídrico, 1 - baixa e 2 - média. Fonte: Verslype <i>et al.</i> (2021).	126
Figura 4 - Variáveis com maior importância na tomada de decisão do algoritmo RF. Fonte: Verslype <i>et al.</i> (2021).	127

ANEXO

Figura 1 - Distribuição das características agronômicas dentro de cada cluster. Fonte: Verslype <i>et al.</i> (2021).	139
Figura 2 - Distribuição das características agronômicas dentro de cada cluster. Fonte: Verslype <i>et al.</i> (2021).	140

SUMÁRIO

RESUMO.....	xv
ABSTRACT	xvi
CAPÍTULO I	17
1. INTRODUÇÃO.....	18
2. REFERENCIAL TEÓRICO	20
1. REFERÊNCIAS BIBLIOGRÁFICAS	65
CAPÍTULO II	77
1 Descrição dos dados	80
2 Projeto experimental, materiais e métodos.....	81
Referências	91
CAPÍTULO III	92
RESUMO.....	93
ABSTRACT	93
1 Introdução.....	94
2 Material e Métodos	95
3 Resultados e Discussão	100
4 Conclusão.....	110
Referências bibliográficas	110
CAPÍTULO IV.....	114
RESUMO.....	115
ABSTRACT	116
1 INTRODUÇÃO.....	116
2 MATERIAL E MÉTODOS	118
3 RESULTADOS E DISCUSSÃO	122
4 CONCLUSÃO.....	128
REFERÊNCIAS BIBLIOGRÁFICAS	129
ANEXO.....	132
Material suplementar do artigo intitulado: “Cultivares de porta-enxertos de videira: conjunto de dados descrevendo características fisiológicas, bioquímicas e agrônômicas”.....	133
Imagens complementares do artigo intitulado: “Análise da diversidade genética de porta-enxertos de videira através de aprendizado de máquina”.....	139

RESUMO

Os programas de melhoramento genético de plantas buscam selecionar genótipos superiores, para atender as necessidades do homem, no aumento da produtividade, na estabilidade e qualidade das espécies de importância econômica, assim como na redução dos impactos ambientais e nos custos de produção. Dessa forma, a videira (*Vitis* spp.) é considerada uma fruteira perene de grande importância econômica, social e alimentar. Porém face às mudanças climáticas e a limitação de recursos hídricos, tem havido um crescente investimento no desenvolvimento e uso de porta-enxertos tolerantes ao déficit hídrico. No entanto, a obtenção de novas cultivares, tolerantes ao déficit hídrico, trata-se de um processo demorado e difícil por ser uma característica poligênica. Por conta disto, o uso de novas ferramentas, tais como o uso algoritmos de aprendizado de máquina podem facilitar identificação e obtenção de novas cultivares tolerantes ao déficit hídrico, devido a capacidade de gerenciar grandes quantidades de dados e identificar padrões relevantes. Neste sentido, o objetivo deste trabalho foi de avaliar a divergência genética de 45 cultivares de porta-enxertos de videira, assim como, de predizer classes de tolerância ao déficit hídrico nas três cultivares brasileiras IAC313, IAC572 e IAC766, cuja a informação na literatura é desconhecida, através de algoritmos de aprendizagem de máquina. Como método para análise de divergência genética, foram empregados algoritmos de agrupamento *K-means* e a Análise de Componentes Principais. A partir dos resultados obtidos na análise de divergência genética para as características avaliadas, foram formados cinco grupos heteróticos e identificados 37 opções de cruzamentos viáveis, indicando existir divergência entre as cultivares. Para a predição de classes de tolerância ao déficit hídrico, foi comparado o desempenho de seis algoritmos distintos, como o *Decision Tree*, *Random Forest*, *K-Nearest Neighbors*, *XGBoost*, *Support Vector Machines* e *Linear Discriminant Analysis*. E em seguida através do algoritmo com melhor desempenho, predizer o grau de tolerância ao déficit hídrico das três cultivares brasileiras. Os resultados indicaram o Random Forest como melhor algoritmo e pode-se predizer que as cultivares IAC 313 e IAC 766 apresentam alta tolerância ao déficit hídrico e o IAC 572 baixa tolerância. Neste sentido, o uso algoritmos de aprendizado de máquina em nosso trabalho, possibilitaram alcançar resultados de fácil entendimento, mostrando-se como mais uma opção de ferramenta acessível e útil ao melhorista, para a identificação de melhores cruzamentos para uma característica específica e predição de classes.

Palavras-chave: Mudanças climáticas, algoritmo supervisionado, algoritmo não supervisionado, melhoramento vegetal, inteligência artificial.

ABSTRACT

EVALUATION AND SELECTION OF DROUGHT TOLERANT GRAPEVINE ROOTSTOCKS (*Vitis* spp.) THROUGH MACHINE LEARNING

Plant breeding programs seek to select superior genotypes due to human needs, like productivity increase, stability, and quality of economically important species, besides reducing environmental impacts and production costs. Thus, the grapevine (*Vitis* spp.) is considered an economic, social, and food fruit tree important. However, climate change and the scarcity of water resources have provoked a growing investment in the development and use of drought-tolerant rootstocks. However, new drought-tolerant cultivar obtention is a long time-consuming, and complicated process as it's a polygenic characteristic. Because of this, use new tools like machine learning algorithms can leading to the identification and selection of new drought-tolerant cultivars due to the ability to manage large amounts of data and identify relevant patterns. In this sense, this work aimed to assess the genetic divergence of 45 grapevine rootstocks cultivars and predict the drought tolerance classes of the three Brazilian cultivars IAC313, IAC572, and IAC766, whose information is still unknown in the literature, through machine learning algorithms. For genetic divergence analysis, K-means and Principal Components Analysis algorithms were applied. The results obtained indicated five heterotic groups and 37 crossover options viable in the genetic divergence analysis for the evaluated traits, indicating divergence between the cultivars. The performance of six different algorithms was compared, such as Decision Tree, Random Forest, K-Nearest Neighbors, XGBoost, Support Vector Machines, and Linear Discriminant Analysis to predict drought tolerance classes. And the best performing algorithm was used to predict the degree of drought tolerance of the three Brazilian cultivars. The results indicated Random Forest as the best model, which predicted high drought tolerance for IAC 313 and IAC 766 and a low tolerance for IAC 572. In this sense, it was possible to achieve easy-to-understand results with the machine learning algorithms in our study, showing itself as another option helpful and accessible tool to breeders for identifying better crosses for a specific characteristic and predicting classes.

Key words: Climatic changes, supervised algorithm, unsupervised algorithm, plant breeding, artificial intelligence.

CAPÍTULO I

INTRODUÇÃO E REFERENCIAL TEÓRICO

1. INTRODUÇÃO

Entre as espécies agrícolas de maior importância no mundo, a videira (*Vitis* spp.) ocupa lugar de destaque. Por ser uma fruteira perene, mundialmente cultivada sob condições ambientais muito diversas, e por apresentar grande importância alimentar, econômica e social (Tonietto and Carbonneau 2004, Leão and Borges 2009, Zhang *et al.* 2016).

As videiras são geralmente bem adaptadas aos climas áridos e semiáridos principalmente por conta de mecanismos de prevenção à seca em situações de stress hídrico (Chaves *et al.* 2010), porém variam conforme a cultivar, por conta de sua grande diversidade genética (Zarrouk *et al.* 2016). Embora apresentem vários mecanismos de sobrevivência é considerada uma espécie dependente do clima, sendo sensível às variações climáticas bruscas, principalmente nas fases consideradas mais críticas como na brotação, floração e maturação das uvas (Mandelli *et al.* 2009, Jones and Webb 2010, Fraga *et al.* 2013).

Segundo Tonietto and Carbonneau (2004), o caráter climático de uma região, a longo prazo, determina em grande parte o seu potencial vitícola e enológico. Enquanto as variações sazonais, de curto a longo prazo, são considerados fatores-chave de produção, tendo influência na sua quantidade e qualidade de uva e vinho produzido (Jackson and Lombard, 1993), o que representa uma importante limitação para o crescimento e desenvolvimento da espécie (Malheiro *et al.* 2010). Assim, o planejamento da escolha do local de implantação do vinhedo e das cultivares é crucial para o sucesso de um vinhedo e deve seguir as exigências ecofisiológicas da cultura (Mandelli *et al.* 2009). Conforme evidenciam diversos autores, em que o estado hídrico da videira é empregado como um fator de produção, pois a escassez de água, já acomete extensas áreas cultiváveis, e é considerada o fator mais limitante na agricultura, devido aos distúrbios metabólicos ocasionados nas plantas, que afetam tanto no crescimento, distribuição do sistema radicular, produtividade, quanto na qualidade final das uvas (Magalhães Filho *et al.* 2008, Lima Filho *et al.* 2009, Netzer *et al.* 2009, Serra *et al.* 2014).

As projeções das significativas mudanças climáticas em curso indicam uma crescente escassez de água no mundo, notabilizando-se um risco quanto à segurança alimentar. Desta forma, torna-se um desafio inevitável que a sociedade terá de enfrentar nas próximas décadas, exigindo que sejam tomadas medidas apropriadas

de adaptação no setor agrícola, não obstante do setor vitivinícola, em que se preveem potenciais reduções de áreas cultiváveis nas regiões vinícolas tradicionais (Jones *et al.* 2005, Hall and Jones 2009, Malheiro *et al.* 2010, Jones *et al.* 2012, Fraga *et al.* 2013).

A enxertia é uma técnica empregada na maioria das regiões vitícolas do mundo, que consiste na associação entre dois genótipos, a cultivar porta-enxerto e a cultivar copa, caracterizados por apresentarem funções específicas (Camargo 2003, Leão *et al.* 2009). Segundo Leão *et al.* (2016), o uso de cultivares de porta-enxertos adequados as cultivares copa possibilita aos vinhedos atingirem o seu potencial máximo de produção e qualidade das uvas, por agregarem diversas vantagens tais como a sua influência na absorção da água, adaptação a diferentes tipos de solos, entre outras.

A busca por porta-enxertos mais tolerantes à seca deve ser um dos objetivos da vitivinicultura mundial (Serra *et al.* 2014). Como destacam Zhang *et al.* (2016) tem havido um crescente investimento nos últimos anos no desenvolvimento de materiais tolerantes à seca, a exemplo de cultivares porta-enxertos, por conta de determinadas interações entre o porta-enxerto e a copa que podem induzir um melhor desempenho e produção à cultivar copa mesmo quando o suprimento de água é limitado, ou seja, sob estresse hídrico (Souza *et al.* 2001, Serra *et al.* 2014).

Numerosos trabalhos desenvolvidos em diversas localidades no mundo, demonstram o potencial do uso de porta-enxertos frente a problemática da escassez da água, buscando-se melhorar e desenvolver porta-enxertos tolerantes à seca, como os da autoria de Alsina *et al.* (2011), Marguerit *et al.* (2011), Serra *et al.* (2014), Tsegay *et al.* (2014), Zhang *et al.* (2016), Prinsi *et al.* (2018), entre outros. Entranto, poucos são os estudos sobre esse tema no Brasil. Além disso, a tolerância ao estresse hídrico nas videiras é considerada uma característica poligênica (Chaves *et al.* 2003, Ashraf 2010, Marguerit *et al.* 2012, Serra *et al.* 2014, Dayer *et al.* 2019), o que dificulta a seleção de genótipos para esse fim em programas de melhoramento genético de porta-enxertos de videira, tendo em vista a interação genótipo-ambiente.

Para contornar esta problemática, o *machine learning* (aprendizado de máquina) é uma nova metodologia que vem sendo cada vez mais empregada em diversas áreas de estudo, como no melhoramento genético de plantas, por possibilitar identificar, predizer e classificar genótipos conforme a necessidade do programa de melhoramento (Beiki 2019, Etminan *et al.* 2019). Conforme pode ser observado no

trabalho de Beiki (2019) o *machine learning* possibilita acelerar a liberação de novos genótipos tolerantes à seca, por alcançar uma alta precisão na identificação de linhagens puras de milho tolerantes ao déficit hídrico. Assim como, no trabalho de Etminan *et al.* (2019) que conseguiram identificar os melhores índices de tolerância à seca para o trigo duro e também indicar os melhores genótipos para serem cultivados em ambientes propensos à seca através do uso de rede neural artificial (RNA).

Diante do exposto, o objetivo deste trabalho foi utilizar seis algoritmos de aprendizado de máquina supervisionado, o *Decision Tree*, *Random Forest*, *K-Nearest Neighbors*, *XGBoost*, *Support Vector Machines* e *Linear Discriminant Analysis*, como ferramenta na predição de classes de tolerância ao déficit hídrico em três cultivares de porta-enxertos de videira brasileiras IAC313, IAC572 e IAC766, cuja a informação na literatura é desconhecida, e determinar a divergência genética entre as quarenta e cinco cultivares estudadas com os algoritmos de agrupamento *K-means* e a Análise de Componentes Principais de aprendizado não-supervisionado, para auxiliar na seleção de porta-enxertos tolerantes ao déficit hídrico para cultivo na região do Submédio do Vale do São Francisco.

2. REFERENCIAL TEÓRICO

A videira (*Vitis* spp.), é considerada uma das culturas mais cultivadas no mundo e de maior valor agregado (Walker *et al.* 2019). Por conta de sua grande importância e os desafios frente às mudanças climáticas, no referencial teórico serão abordados a origem, domesticação e aspectos botânicos, assim como, a sua importância socioeconômica a nível mundial e nacional, com enfoque na região do Vale do Submédio São Francisco que é considerada a segunda maior região vinícola do Brasil (Kist *et al.* 2019), também será abordado as características climáticas desta região vinícola. Além destes temas, o destaque para a importância do uso de porta-enxertos na Vitivinicultura, descrição agrônoma das quarenta e cinco cultivares de porta-enxertos estudadas, os desafios do setor Vitivinícola frente às mudanças climáticas, a questão do estresse hídrico que afeta as videiras e quais são os seus principais mecanismos de tolerância a esse estresse, a importância do melhoramento genético de porta-enxertos e finalizando o *machine learning* aplicados ao melhoramento genético de plantas.

2.1 Origem e domesticação da videira

O centro de origem paleontológico das videiras segundo Giovannini (2014), é na Groenlândia, onde pode-se encontrar os fósseis mais antigos. A videira teria se dispersado nas direções Américo-Asiática e Euro-Asiática, formando três centros de origem distintos, localizados no Sul da Europa e Ásia Menor, no Leste da Ásia e nas Américas do Norte e Central (Leão and Borges 2009, Giovannini 2014).

A videira possui uma longa e rica história cultural e geográfica, por ter sido dispersa, cultivada e moldada pelo homem desde a antiguidade (Dougherty 2012). Albuquerque (2003) destaca que, na civilização ocidental o seu cultivo e produção de vinho desempenhou um importante papel. Impactou significativamente nas atividades humanas, econômicas e culturais (Dougherty 2012). No entanto, de acordo com This *et al.* (2006), permanece incerto o seu local e período de domesticação original, assim como, se houveram domesticações secundárias independentes, porém evidências sugerem que a sua domesticação teria ocorrido no Oriente Próximo. Para Leão *et al.* (2009), a sua domesticação teria ocorrido há cerca de 6000 mil anos e cultivada há milênios, sendo assim, considerada a fruta de domesticação mais antiga de que se tem conhecimento (Radmann and Bianchi 2008).

2.2 Aspectos botânicos da videira

Dentro do Reino Vegetal, a videira (*Vitis* spp.) encontra-se classificada na Divisão *Magnoliophyta* Cronq., Takht. & W. Zimm. ex Reveal, na Classe *Magnoliopsida* Brongn., Subclasse *Rosidae* Takht., na Ordem *Rhamnales* Lindl. e na família *Vitaceae* Juss. (Albuquerque 2003, Mandelli *et al.* 2009, Villa 2018, Roskov *et al.* 2019, Tropicos 2020). A família *Vitaceae* Juss. engloba principalmente espécies de arbustos que são lianas do tipo trepadeiras ou cipós lenhosos perenes que possuem folhas originadas de nós, dispostas alternadamente e com gavinhas opostas às folhas (Leão and Borges 2009, Giovannini 2014, Keller 2015).

De acordo com Giovannini (2014) e Roskov *et al.* (2019), a família *Vitaceae* Juss. é constituída por quinze gêneros vivos e dois gêneros fósseis, abrangendo ao total cerca de 1000 espécies. Dentre estes gêneros, o gênero *Vitis* L., é considerado o mais antigo e o único com importância econômica por produzir frutos comestíveis, compreendendo mais de 70 espécies, entre as espécies de maior interesse econômico destacam-se as videiras européias também conhecidas como viníferas

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

(*Vitis vinifera* L.), as videiras americanas (*Vitis labrusca* L.) e as videiras híbridas (*Vitis* spp.) (Albuquerque 2003, This *et al.* 2006, Camargo *et al.* 2009, Leão and Borges 2009, Mandelli *et al.* 2009, Giovannini 2014).

A respeito disso Camargo *et al.* (2009) afirmam que não existe consenso entre os autores quanto à classificação das espécies de videira no gênero *Vitis* L., no entanto, para Keller (2015) a classificação da *Muscadinia* (Planch.) Small em subgêneros ou seções dentro do gênero *Vitis* L. está caindo em desuso. Dessa forma a *Muscadinia* (Planch.) Small (2n = 40 cromossomos), que compreende apenas três espécies, e *Vitis* L. (2n = 38 cromossomos) seriam dois gêneros distintos na família *Vitaceae*. Esta classificação é cada vez mais apoiada por alguns melhoristas e pesquisadores, por conta dos cruzamentos entre os dois táxons serem difíceis e os híbridos obtidos são quase todos estéreis (2n = 39 cromossomos), além de certas diferenças morfológicas como as gavinhas ramificadas nas espécies pertencentes ao gênero *Vitis* L. e as gavinhas simples na *Muscadinia* (Planch.) Small (Walker, 2019).

As folhas da videira variam de forma, cor, pilosidade, brilho e tamanho (Giovannini 2014), em função da espécie e cultivar (Albuquerque 2003), enquanto as suas inflorescências são opostas do tipo racimo, com flores que podem ser unissexuais ou hermafroditas, e cachos e bagas que podem apresentar diversos formatos, sendo as suas bagas macias, bicarpelares, compostas pela película, e pruína (cera) na parte externa que retém as leveduras e outros microrganismos (Leão and Borges 2009, Giovannini 2014, Keller 2015).

2.3 Importância socioeconômica da videira

Segundo Albuquerque (2003) e Moura *et al.* (2009), as principais áreas de cultivo da videira concentram-se entre as latitudes 30°N e 50°N e entre as latitudes 30°S e 40°S, e inclusive em regiões próximas ao equador. Apesar de ser originária de regiões de clima temperado (Moura *et al.* 2009). Para Walker *et al.* (2019), ela pode ser cultivada em diversas regiões do mundo, que variam desde regiões com ambientes quente e seco desérticos até muito frios e climas tropicais, além de ser explorada em todas as regiões de clima temperado. Conforme pode ser observado com os dados da FAOSAT (Food and Agriculture Organization Corporate Statistical Database, 2020), na safra de 2019, foram colhidas mundialmente 77.137.016 milhões de toneladas de uvas frescas em 6.925.972 hectares, devido a cultura ter sido explorada em diversas localidades, desde o continente Asiático responsável por 37,8% da

produção mundial, assim como, no continente Europeu responsável por 34,6% da produção, na América com 18,7%, África com 6,3% e até a Oceania que deteve 2,6% da produção mundial de uva neste ano.

Apesar de sua exploração a nível mundial, em 2019, 45% da superfície mundial ocupada pela cultura encontrava-se concentrada em apenas quatro países: Espanha (936.890 ha), China (745.907 ha), França (755.470 ha) e Itália (697.910 ha). Sendo os cinco maiores produtores mundiais, a China com 14.283.532 t, a Itália com 7.900.120 t, os Estados Unidos com 6.233.270 t, a Espanha com 5.745.450 t e a França com 5.489.650 t.

De acordo com Zhang *et al.* (2016), a videira é uma das culturas perenes mais valiosas do mundo, por ser considerada uma cultura versátil com alto valor agregado aos seus produtos (Walker *et al.* 2019), podendo ser utilizada de diversas formas, tais como para a produção de sucos, vinhos, espumantes, passas, vinagres, geléias, fármacos, pigmentos, óleo a partir de suas sementes, produção de etanol, destilados (brandy e graspa), entre outros, assim como o seu potencial para o Enoturismo (Mullins *et al.* 1992, Nachtigal and Mazzarolo 2008, Silva *et al.* 2009). A respeito disso, na safra de 2018, dos 77,8 milhões de toneladas de uvas colhidas no mundo, 57% foram destinados para a produção de vinhos, 36% para o consumo *in natura* (uva de mesa) e apenas 7% para a produção de passa (OIV - Organización Internacional de la Viña y el Vino, 2020).

No Brasil a viticultura é um dos setores mais representativos na fruticultura nacional, quanto a geração de emprego, renda e valor bruto, pelos múltiplos usos da matéria-prima, do consumo *in natura* até a produção de sucos, vinhos, espumantes e doces (Kist *et al.* 2019). Na safra de 2019, o Brasil foi o décimo quinto maior produtor mundial de uva, com uma produção de 1.485.292 t em uma superfície cultivada de 74.625 hectares, contribuindo com cerca de 1,9% da produção mundial de uva, abrangendo 1,1% da superfície mundial cultivada por videira e cerca de 7,8% da área cultivada no continente Americano (IBGE - Instituto Brasileiro de Geografia e Estatística 2020, FAOSAT 2020).

A produção de uvas vêm ganhando, cada vez mais destaque na economia do país, sendo as regiões Sul e Nordeste as que concentram o cultivo da videira (Kist *et al.* 2019), em que a região Sul deteve 73,3% da viticultura nacional, participando com 53,5% do total de uvas colhidas no país, e a região Nordeste deteve 13,9% da área

nacional cultivada com uva, sendo considerada a segunda região mais representativa do país e contribuiu com 34,5% da safra de 2019 (IBGE, 2020).

Kist *et al.* (2019) destacam que na Região Nordeste a viticultura encontra-se concentrada na região do Vale do São Francisco, mais precisamente no pólo Petrolina – PE e Juazeiro - BA. Segundo Silva *et al.* (2009), a viticultura irrigada nesta região, destaca-se por sua importância econômica e social, que engloba um grande volume anual de negócios e se destaca entre as culturas irrigadas, com a maior geração de empregos diretos e indiretos. De acordo com Silva and Correia (2016), no Submédio do Vale do São Francisco, são gerados em média dois empregos diretos por hectare no campo e quatro empregos indiretos decorrentes da dinâmica dos serviços dessa atividade, representando cerca de 72 mil ocupações diretas ou indiretas por ano. Dessa forma, a vitivinicultura nesta região vem ganhando cada vez mais destaque no país, e a sua intensificação na fruticultura de exportação beneficiou o desenvolvimento regional (Kist *et al.* 2018).

2.4 Características da Vitivinicultura no Submédio do Vale do São Francisco

O semiárido do Nordeste brasileiro, localiza-se entre os paralelos 8°S e 10°S (Figura 01), nesta região encontra-se o Submédio do Vale do São Francisco, que apresenta condições favoráveis para a produção de frutas de alta qualidade e produtividade (Netzer *et al.* 2009, Kist *et al.* 2018). As uvas e vinhos produzidos adquirem características muito peculiares e exclusivas desta região vinícola, única no mundo (Moura *et al.* 2009, Silva *et al.* 2009, Kist *et al.* 2018).

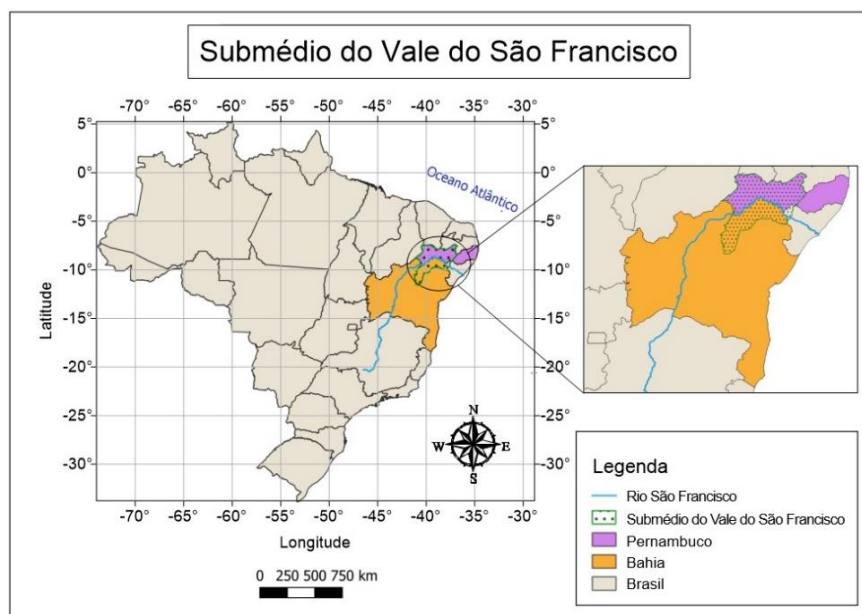


Figura 1 - Localização do Submédio do Vale do São Francisco. Fonte: Verslype (2021).

O clima da região segundo Köeppen, é do tipo BswH, que corresponde a um clima tropical semiárido muito quente (Leão and Silva 2014). Se caracterizando pela insuficiência de chuvas que são inferiores a 800 mm por ano com irregularidade temporal e espacial, assim como, pelas temperaturas elevadas cujas médias térmicas são sempre superiores a 18°C, e por fortes taxas de evaporação (CODEVASF - Companhia de Desenvolvimento do Vale do São Francisco, 1996).

As condições climáticas áridas e semiáridas da região, representam grandes impactos no desenvolvimento das culturas, no que se refere à produtividade e à qualidade da uva obtida (Soares and Costa 2009). A videira consome entre 600 e 900 mm de água por ciclo, podendo chegar até 1200 mm, dependendo principalmente do clima e da duração do período de crescimento (Ojeda *et al.* 2004, FAO - Food and Agriculture Organization of the United Nations, 2020). Por conta das deficiências hídricas ocasionadas pela elevada evapotranspiração e disponibilidade de água no solo para as videiras, torna-se indispensável o uso de irrigação nesta região vinícola (CODEVASF 1996, Netzer *et al.* 2009).

2.5 Uso de porta-enxertos na Vitivinicultura

A enxertia é uma técnica empregada na maioria das regiões vitícolas do mundo, que teve início a partir do século XIX, quando a filoxera acometeu os vinhedos europeus obrigando a adoção de porta-enxertos (Camargo 2003, Leão *et al.* 2009). Apesar de quase todos os porta-enxertos de videira utilizados atualmente terem sido desenvolvidos em sua maioria há mais de 100 anos para resistir à infestação por filoxera (Riaz *et al.* 2019). Para Peterson *et al.* (2019), o seu uso na viticultura é relativamente recente, quando se leva em conta a extensa história de cultivo de uvas pelo homem.

Os porta-enxertos são utilizados nas mais diversas regiões vitivinícolas no mundo, sendo poucas exceções de regiões em que é possível cultivar sem o seu uso, e mesmo assim vem sendo adotado por conta de se tratar de uma tecnologia simples, que associa duas cultivares e agrega diversas vantagens (Hernandes *et al.* 2011, Ollat *et al.* 2016).

A base genética da maioria dos porta-enxertos, são provenientes de espécies de videiras americanas selvagens resistentes à filoxera, como a *Vitis rupestris* Scheele, *V. riparia* Michx., *V. berlandieri* Planch. e *V. champinii* e seus híbridos (Goldammer 2018, Riaz *et al.* 2019, Maul *et al.* 2020).

Normalmente, os porta-enxertos são propagados vegetativamente por estaquia (Bruckner e Albuquerque 2005). E geralmente não chegam a produzir frutos e quando produzem não apresentam valor comercial devido ao tamanho diminuto de cachos e bagas (Hernandes *et al.* 2011). Porém, a sua importância na viticultura mundial é amplamente conhecida (Leão *et al.* 2016).

Os porta-enxertos possibilitam que haja o elo entre o ambiente edáfico e a cultivar copa, por conta do seu sistema radicular resistente à pragas e às situações adversas do solo, exercendo influência direta no porte da planta, e na capacidade de transmitir o seu vigor, na qualidade final dos seus frutos, além de outras características como na resistência às cultivares copa enxertadas, sendo responsáveis pela captação de água e minerais (Bruckner e Albuquerque 2005, Hernandez *et al.* 2011, Leão *et al.* 2016, Ollat *et al.* 2016). Dessa forma, essa prática, é capaz de agregar diversas vantagens aos vinhedos, tais como o controle de pragas, a melhoria da qualidade da uva, indução de maior vigor à copa, maior desenvolvimento inicial das plantas, precocidade de produção, aumento da produtividade, maior resistência à doenças de solo, influência na absorção de água e nutrientes, adaptação a diferentes tipos de solos, conferindo assim resistência às condições abióticas adversas (Nachtigal and Mazzarolo 2008, Leão *et al.* 2009, Leão *et al.* 2016). Por conta de suas vantagens, pesquisadores comentam que os porta-enxertos vem sendo empregados inclusive em regiões privilegiadas como no Chile, que apresentam algum isolamento geográfico ou climático ou condições edáficas que desfavorecem o desenvolvimento de pragas e dessa forma não seria necessário o seu uso (Hernandes *et al.* 2011).

A escolha do porta-enxerto deve ser realizada de forma cautelosa, uma vez que existem centenas de cultivares e cada porta-enxerto se comporta diferentemente com a cultivar copa enxertada, assim como se adapta a determinadas condições de solo e clima (Camargo 2003, Leão 2010). Além disso, a sua escolha é uma decisão importante, não apenas pelo benefício potencial que os viticultores esperam, mas também por conta do estabelecimento de um vinhedo ser um investimento a longo prazo (Serra *et al.* 2014). Em razão disto, devem ser levados em consideração alguns fatores tais como a susceptibilidade do porta-enxerto a doenças e pragas ocorrentes na região em que se encontra o vinhedo, assim como a sua adaptação aos solos da região, a tolerância à deficiência nutricional, a cultivar copa que deseja-se usar, que deve apresentar boa compatibilidade na enxertia; além disso, o porta-enxerto deve

apresentar facilidade de enraizamento e de pegamento na enxertia (Camargo 2003, Leão 2010).

No Submédio do Vale do São Francisco, o porta-enxerto ideal deve reunir características como vigor, fácil enraizamento e cicatrização na enxertia, devendo minimizar às condições adversas do solo, como baixa fertilidade, salinidade, alcalinidade e déficit hídrico, assim como, devem apresentar resistência a pragas e doenças, principalmente nematóides presentes nos solos arenosos da região e ter boa afinidade com as cultivares copa que se deseja enxertar (Leão *et al.* 2009, Leão 2010, Leão and Silva 2014).

Neste estudo, optou-se pela escolha de quarenta e cinco cultivares de porta-enxertos de videira, conforme pode ser visualizado na descrição agronômica resumida da Tabela 1 e Figura 2. Para a escolha destas quarenta e cinco cultivares de porta-enxertos, foram considerados três fatores:

- Cultivares que se destacam no Submédio do Vale do São Francisco, como as cultivares Paulsen 1103, SO4, IAC 572 Jales, IAC 313 Tropical, IAC 766 Campinas, Salt Creek, Freedom e 101-14 MGt por apresentarem resultados promissores na região com cultivares de uva de mesa (Camargo *et al.* 2011, Leão *et al.* 2020a, Leão 2020b); assim como, optamos por incluir a cultivar Harmony devido ao seu bom desempenho com cultivares de uva de mesa sem semente (Leão e Silva 2014);
- Disponibilidade de acessos do Banco Ativo de Germoplasma de videira da Embrapa Uva e Vinho (Camargo *et al.* 2017);
- Disponibilidade de informações na literatura sobre o grau de tolerância à seca de cada porta-enxerto.

Tabela 1 - Descrição agrônômica das quarenta e cinco cultivares de porta-enxertos de *Vitis* spp e híbridos empregadas no estudo.

Porta-enxerto	Origem genética	Flor	Tolerância à seca	Vigor
OLMO VR 039-16	<i>Vitis vinifera</i> L. x <i>Muscadinia rotundifolia</i> Michaux ^{1,5}	Feminina ¹	Baixa ^{5,7,18,24}	Alto ^{5,6,24}
OLMO VR 043-43	<i>Vitis vinifera</i> L. x <i>Muscadinia rotundifolia</i> Michaux ¹	Masculina ³	Baixa ³	Alto ^{3,6}
101-14 MGt (Millardet et Grasset 101- 14)	<i>Vitis riparia</i> Michaux x <i>Vitis rupestris</i> Scheele ^{1,3,5,6}	Feminina ^{1,2,3,7}	Baixa ^{3,6,10,14,17,19,21,22,23,24}	Médio ^{5,8,10,18}
106-8 MGt (Millardet et Grasset 106- 8)	<i>Vitis riparia</i> Michaux x <i>Cordifolia rupestris</i> de grasset n ^o 1 ¹	Feminina ^{1,2,15}	Alta ¹⁵	Alto-médio ^{14,15}
110 R (Richter 110)	<i>Vitis berlandieri</i> Planchon x <i>Vitis rupestris</i> Scheele ^{1,5,6,8}	Masculina ^{1,2,7}	Alta ^{5,6,7,11,13,14,15,16,17,19,21,22,23,24}	Alto-médio ^{6,8,21,24}
1103 P (Paulsen 1103)	<i>Vitis berlandieri</i> Planchon x <i>Vitis rupestris</i> Scheele ^{1,5,6,8}	Masculina ^{1,2,3,7}	Alta ^{3,6,7,11,16,17,19,21,22,23}	Alto-médio ^{5,21,24}
1202 C (Couderc 1202)	<i>Vitis vinifera</i> L. x <i>Vitis rupestris</i> Scheele ¹	Monóica ^{1,2}	Alta ^{7,18}	Alto ^{11,15}
125 AA (Kober 125 AA)	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux ¹	Feminina ^{1,2,7,15}	Baixa ^{21,23}	Alto-médio ^{7,11,12}
140 Ru (Ruggeri 140)	<i>Vitis berlandieri</i> Planchon x <i>Vitis rupestris</i> Scheele ^{1,5,6}	Masculina ^{1,2,7}	Alta ^{5,6,7,11,12,13,14,16,17,19,21,22,23,24}	Alto ^{5,21}
157-11 C (Couderc 157-11)	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux ¹	Feminina ^{1,2}	Media ^{21,23}	Alto-médio ¹⁵
1613 C (Couderc 1613)	SOLONIS (<i>Vitis riparia</i> Michaux x <i>Vitis longii</i> Prince) x OTHELLO ^{1,5}	Feminina ^{1,2}	Media ⁹	Médio ⁵
161-49 C (Couderc 161-49)	<i>Vitis riparia</i> Michaux x <i>Vitis berlandieri</i> Planchon ^{1,6}	Feminina ^{1,2,7}	Baixa ^{6,8,17,22,24}	Médio ^{6,7,8}
1616 C (Couderc 1616)	SOLONIS x <i>Vitis riparia</i> Michaux ¹	Masculina ^{1,2}	Baixa ^{5,22,24}	Alto-médio ⁸
196-17 CI (Castel 196-17)	1203C x <i>Vitis riparia</i> Michaux	Masculina ^{1,2,7}	Alta ^{7,15,22}	Médio ⁷
216-3 CI (Castel 216-3)	1616C x <i>Vitis rupestris</i> Scheele ¹	Masculina ^{1,2,7,15}	Média ^{7,15,17}	Alto ⁷

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

26 G (Geisenheim 26)	<i>Vitis vinifera</i> L. X <i>Vitis riparia</i> Michaux ¹	Feminina ^{1,2}	Média ¹²	Alto ¹²
3306 C (Couderc 3306)	<i>Vitis riparia</i> Michaux x <i>Vitis rupestris</i> Scheele ¹	Masculina ^{1,2}	Média ¹⁹	Médio-baixo ¹¹
3309 C (Couderc 3309)	<i>Vitis riparia</i> Michaux x <i>Vitis rupestris</i> Scheele ^{1,5,6}	Masculina ^{1,2,7,15}	Baixa ^{7,13,14,17,19,20,21,22,23}	Médio-baixo ^{5,8,18,24}
34 EM (Ecole Montpellier 34)	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux ¹	Masculina ^{1,2,7}	Baixa ^{11,17,23}	Médio ⁸
41 B MGt (Millardet et Grasset 41 B)	<i>Vitis vinifera</i> L. x <i>Vitis berlandieri</i> Planchon ¹	Monóica ^{1,2}	Alta ¹⁰	Alto-médio ^{7,12,21}
420 A MGt (Millardet et Grasset 420 A)	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux ^{1,3,5,6,9}	Masculina ^{1,2,3,7}	Média ^{5,6,10,12,16,17,21,22,23}	Médio-baixo ^{5,6,8,10,21}
44-53 M (44-53 Malegue)	<i>Vitis riparia</i> Michaux x MALEGUE 144 ¹	Masculina ^{1,2,7,15}	Alta ^{5,7,15,16,17,24}	Médio ^{5,8}
5 BB (Kober 5 BB)	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux ^{1,5,6}	Feminina ^{1,2,7,15}	Média ^{5,6,23,24}	Alto-médio ^{8,21}
5 C (Teleki 5C)	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux ^{1,5}	Masculina ^{1,2,7}	Baixa ^{5,8,11,19,24}	Alto-médio ^{8,10,12}
62-66 C (Couderc 62-66)	<i>Vitis vinifera</i> L. X <i>Vitis cordifolia</i> Michaux ¹	Masculina ^{1,2}	Alta ²⁰	-
8 B (Teleki 8 B)	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux ¹	Masculina ^{1,2,15}	Média ^{11,17}	Alto-médio ^{4,6,11}
93-5 C (Couderc 93-5)	<i>Vitis vinifera</i> L. x <i>Vitis rupestris</i> Scheele ¹	Feminina ^{1,2}	Baixa ¹⁵	Alto ²⁰
99 R (Richter 99)	<i>Vitis berlandieri</i> Planchon x <i>Vitis rupestris</i> Scheele ^{1,5,6,8}	Masculina ^{1,2,7}	Alta ^{11,16,19}	Alto-médio ^{5,6,8}
ARG 1 (Ganzin 1)	ARAMON NOIR x <i>Vitis rupestris</i> Scheele ¹	Masculina ^{1,2}	Alta ⁷	Alto-médio ^{5,11}
Dogridge (Dog Ridge)	<i>Vitis rupestris</i> Scheele x <i>Vitis candicans</i> Engelm ^{1,5}	Feminina ^{1,2}	Alta ⁹	Alto ⁵
Fercal	<i>Vitis berlandieri</i> x Planchon 31R ¹	Feminina ^{1,2,7}	Média ^{7,16,17,19,21,24}	Alto-médio ^{7,18,21}
Freedom	FRESNO 1613-59 x DOGRIDGE ¹	Monóica ²	Média ^{5,19,24}	Alto-médio ^{5,13,15}

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

Golia	CASTEL 156-12 x <i>Vitis rupestris</i> Scheele ¹	Monóica ^{1,2}	Baixa ²³	Alto ¹⁴
Gravesac	161-49C x 3309C ¹	Masculina ^{1,2,7}	Media ^{7,8,15,18,19}	Alto-médio ^{7,10,12,21}
Harmony	1613C x DOGRIDGE ¹	Feminina ^{1,2}	Média ²⁴	Alto-médio ^{5,24}
Riparia Gloire de Montpellier	<i>Vitis riparia</i> Michaux ^{1,5,6}	Masculina ^{1,2,7}	Baixa ^{5,6,7,10,12,13,16,17,22,24}	Baixo ^{4,6,7,8,18,24}
Rupestris du lot	<i>Vitis rupestris</i> Scheele ^{1,5,6,8}	Masculina ^{1,2,7,15}	Média ^{7,8,9,17}	Alto ^{4,5}
Ramsey (Salt Creek)	<i>Vitis champinii</i> Planchon ^{1,5}	Feminina ^{1,2}	Alta ^{9,19}	Alto-médio ^{5,6,8,13}
Schwarzmann	<i>Vitis riparia</i> Michaux x <i>Vitis rupestris</i> Scheele ^{1,5}	Masculina ^{1,2}	Baixa ^{17,19}	Médio ^{5,8,11,15,24}
SO4	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux ^{1,5,6}	Masculina ^{1,2,3,7,15}	Baixa ^{3,6,8,10,17,18,19,22}	Alto-médio ^{4,10,21}
Sori	SOLOINIS x <i>Vitis riparia</i> Michaux ¹	Feminina ^{1,15}	Média ^{7,12}	Médio ¹²
<i>Vitis champini</i>	<i>Vitis champinii</i> Planchon ¹	Feminina ²	Alta ¹³	Alto ¹³
IAC 313 (Tropical)	GOLIA x <i>Vitis cinerea</i> Engelmann ^{1,3}	Masculina ³	-	Alto ^{3,4,6}
IAC 572 (Jales)	<i>Vitis caribaea</i> De Candolle x 101-14MGt ^{1,3}	Masculina ^{2,3}	-	Alto ^{4,6}
IAC 766 (Campinas)	106-8MGt x <i>Vitis caribaea</i> De Candolle ^{1,3}	Masculina ^{2,3}	-	Alto ^{4,6}

Informações obtidas em: ¹Maul *et al.* (2020); ²Camargo *et al.* (2017); ³Embrapa Uva e Vinho (2016); ⁴Nachtigal *et al.* (2008); ⁵Bettiga *et al.* (2003); ⁶Giovannini (2014); ⁷Audeguin *et al.* (2020); ⁸Keller (2015); ⁹Satisha *et al.* (2006); ¹⁰ATVB (2013); ¹¹Clarke (2004); ¹²Rebschule Mueller (2020); ¹³Carroll (2016); ¹⁴Pecile *et al.* (2020); ¹⁵Storm and Krasokhina (2020); ¹⁶Carbonneau (1985); ¹⁷Villa (2018); ¹⁸IFV (2007); ¹⁹Wine Australia (2016); ²⁰Chevalier (1925); ²¹Brancadoro *et al.* (2014); ²²Morton (1979); ²³VICOPAD (2020); ²⁴Sunridge (2020).

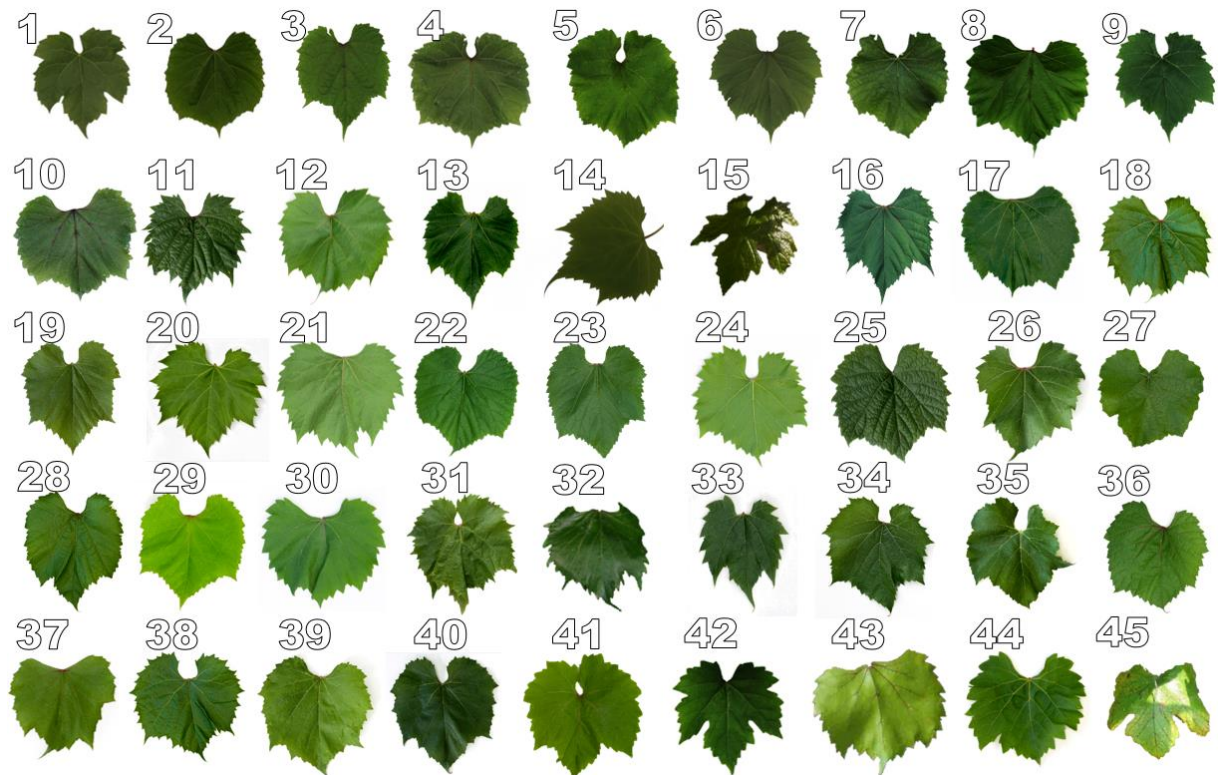


Figura 2 - Porta-enxertos de videira avaliados: Harmony (1), IAC 313 (2), IAC 572 (3), IAC 766 (4), SO4 (5), 101-14 MGt (6), Ramsey (7), Paulsen 1103 (8), Freedom (9), Richter (10), 1616 C (11), 3309 C (12), 44-53 M (13), 420 A MGt (14), VR 039-16 (15), Riparia Gloire de Montpellier (16), 140 Ru (17), Schwarzmann (18), 5 C (19), 196-17 CI (20), 216-3 CI (21), 157-11 C (22), 161-49 C (23), 1202 C (24), 1613 C (25), 3306 C (26), Dogridge (27), 34 EM (28), Fercal (29), Ganzin 1 (30), 26 G (31), Golia (32), Gravesac (33), 125 AA (34), 41 B MGt (35), 106-8 MGt (36), 99 R (37), Sori (38), 5 BB (39), 8B (40), 93-5 C (41), VR 043-43 (42), Rupestris du lot (43), 62-66C (44) e *Vitis champini* (45). Fonte: Adaptado de FPS (2020), Embrapa Uva e Vinho (2020) e Maul *et al.* (2020).

A seguir a descrição agrônômica das principais cultivares de porta-enxerto empregadas no Submédio do Vale do São Francisco.

2.5.1 Paulsen 1103 (1103P)

A cultivar Paulsen 1103 também conhecida como 1103P ou simplesmente 1103 (Figura 03) foi obtida na Sicília, Itália, em 1896 pelo melhorista Federico Paulsen, sendo registrada como cultivar no ano de 1971 (Giovannini 2014, Audeguin *et al.* 2020). Caracteriza-se pela alta resistência à fusariose, ao míldio e a filoxera nas raízes, e tolerância moderada a solos salinos, baixa resistência à antracnose e a nematóides (Embrapa Uva e Vinho 2016). Além disso, apresenta compatibilidade geral com as cultivares copa, e fácil enraizamento e pegamento na enxertia (Camargo 2003, Giovannini 2014). Destaca-se como um dos mais importantes porta enxertos

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

utilizados no Submédio do Vale do São Francisco com diferentes cultivares de videira, tais como a Crimson Seedless e Sugraone (Leão *et al.* 2009).



Figura 3 - Cultivar 1103P. Fonte: Verslype, 2020.

2.5.2 SO4 (Selektion Oppenheim 4)

A cultivar SO4 também conhecida como Selektion Oppenheim 4 (Figura 04), foi selecionado na Alemanha em 1941 (Leão *et al.* 2009, Embrapa Uva e Vinho 2016, Maul *et al.* 2020). Apresenta tolerância moderada a solos salinos e ácidos, bom desenvolvimento em solos de textura franca a argilosa, alta resistência ao míldio, moderada a alta resistência à filoxera nas raízes e moderada aos nematóides, e baixa resistência à antracnose e fusariose (Giovannini 2014, Embrapa Uva e Vinho 2016). Trata-se de uma cultivar facilmente propagada e em geral apresenta fácil enraizamento e boa compatibilidade de enxertia, induz altos rendimentos de produtividade com bons teores de açúcares e necessita as vezes de desbaste (Camargo 2003, Giovannini 2014, Audeguin *et al.* 2020). Porém, os vinhos obtidos apresentam pH alto, e muitas vezes não têm corpo e, podem às vezes apresentarem notas herbáceas devido ao seu alto rendimento (Audeguin *et al.* 2020). No Submédio do Vale do São Francisco vem apresentando um comportamento satisfatório com uvas de mesa sem semente, tais como a cultivar BRS Vitória e BRS Ísis (Leão *et al.* 2020a, Leão *et al.* 2020b).



Figura 4 - Cultivar SO4. Fonte: Verslype, 2020.

2.5.3 IAC 572 (Jales)

A cultivar Jales, também conhecida como IAC 572 (Figura 05) é a mais empregada nas principais regiões tropicais produtoras de uvas de mesa (Nachtigal 2003). Foi desenvolvida no Brasil em Campinas pelo melhorista Santos Neto, no ano de 1958 sendo lançada como cultivar em 1970 (Pommer 2000, Embrapa Uva e Vinho 2016). Apresenta alta sobrevivência ao ser transplantada ao campo (Nachtigal 2003), ótimo enraizamento e pegamento, se desenvolvendo bem em solos ácidos, argilosos e arenosos assim como, possui alta resistência ao míldio, fusariose, filoxera e nematóides, e baixa à antracnose (Pommer 2000, Giovannini 2014, Embrapa Uva e Vinho 2016). Por conta destas características, esta cultivar vem sendo empregada nas regiões vinícolas do Estado de São Paulo e no Vale do São Francisco (Pommer 2000), enxertadas em uvas de mesa fina com semente como a Itália e suas mutações, Red Globe, Benitaka, e com cultivares de uva de vinho devido a sua excelente afinidade, porém não é recomendado seu uso com cultivares copa de uvas sem sementes (Leão *et al.* 2009, Hernandez *et al.* 2011, Leão e Silva 2014).



Figura 5 - Cultivar IAC 572. Fonte: Verslype (2020).

2.5.4 IAC 313 (Tropical)

A cultivar Tropical também conhecida como IAC 313 (Figura 06), foi obtida no Brasil em Campinas pelo melhorista Santos Neto, sendo lançada em 1950 (Embrapa Uva e Vinho 2016). Esta cultivar apresenta alto potencial de enraizamento, bom índice de pegamento de estacas inferiores a 1 cm de diâmetro, e caracteriza-se por apresentar alta resistência ao míldio, filoxera nas raízes e nematóides, assim como, média resistência a antracnose, se adaptado a solos argilosos, arenosos e ácidos (Leão 2010, Giovannini 2014, Embrapa Uva e Vinho 2016). Segundo Hernandez *et al.* (2011) os seus ramos dificilmente perdem as folhas e lignificam tardiamente. Denota boa afinidade com as cultivares copa de uva de mesa com e sem sementes

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

exploradas no Submédio do Vale do São Francisco, como a Sograone, Thompson Seedless, Red Globe e Itália e suas mutações (Leão *et al.* 2009, Leão 2010, Hernandez *et al.* 2011, Leão and Silva 2014).



Figura 6 - Cultivar IAC 313. Fonte: Verslype (2020).

2.5.5 IAC 766 (Campinas)

A cultivar Campinas também é conhecida como IAC 766 (Figura 07) (Leão *et al.* 2009, Maul *et al.* 2020). De acordo com Leão *et al.* (2009) esta cultivar se destaca pelos seus bons índices de pegamento na enxertia e enraizamento de estacas. Apresentando também alta resistência ao míldio, à fusariose, à filoxera nas raízes e aos nematóides, e moderada à antracnose, se adaptando bem aos solos argilosos, arenosos, ácidos e salinos (Leão 2010, Embrapa Uva e Vinho 2016). Apresenta boa afinidade com diferentes cultivares copa, tais como as Niagaras Branca e Rosada, Itália e suas mutações, Red Globe, Centennial Seedless, Patrícia, Maria, Paulistinha e Máximo (Hernandes *et al.* 2011).



Figura 7 - Cultivar IAC 766. Fonte: Verslype (2020).

2.5.6 Salt Creek (Ramsey)

A cultivar Salt Creek também conhecida como Ramsey (Figura 08), foi obtida pelo melhorista Thomas Volney Munson, nos Estados Unidos da América. Trata-se de um híbrido natural originado do cruzamento entre as espécies *Vitis candicans* Engelmann e *Vitis rupestris* Scheele (Riaz *et al.* 2019, Maul *et al.* 2020). Para Bettiga

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

et al. (2003) esta cultivar se destaca pela sua alta resistência à filoxera e aos nematóides, devido ao seu sistema radicular profundo, e possui alta tolerância a salinidade e alta eficiência no uso da água, assim como apresenta um bom desempenho em solos levemente ácidos (Bettiga *et al.* 2003, Goldammer 2018). Outra característica é a sua baixa aptidão de enraizamento e boa enxertia (Giovanni 2014).



Figura 8 - Cultivar Ramsey. Fonte: Verslype (2020).

2.5.7 Freedom

A cultivar Freedom (Figura 09) foi obtida em 1967 pelos melhoristas Weinberger e Harmon em Fresno na Califórnia, EUA (Bettiga *et al.* 2003, Garris *et al.* 2009, Peterson *et al.* 2019). De acordo com Bettiga *et al.* (2003), esta cultivar apresenta alta resistência à filoxera e aos nematóides, média a baixa tolerância à salinidade, sendo considerada uma cultivar sensível à vírus. Apresenta bom desenvolvimento em solos de textura franca, e baixa em solos argilosos, assim como, moderada aptidão de enxertia e enraizamento (Keller 2015).



Figura 9 - Cultivar Freedom. Fonte: Verslype (2020).

2.5.8 Millardet et Grasset 101-14 (101-14 MGt)

A cultivar Millardet et Grasset 101-14, também conhecida como 101-14 MGt (Figura 10), foi obtida em 1882, na França pelos melhoristas Alexis Millardet e Marques du Grasset (Embrapa Uva e Vinho 2016). Esta cultivar apresenta alta resistência ao míldio e à filoxera nas raízes, moderada aos nematóides, média a baixa

à fusariose, e baixa à antracnose e fusariose (Embrapa Uva e Vinho 2016). Também apresentam boa capacidade de enraizamento e enxertia, exibindo uma afinidade muito boa com os enxertos, se adaptando bem a solos de textura franca a arenosa e moderada tolerância a solos salinos (Giovannini 2014, Embrapa Uva e Vinho 2016, Audeguin *et al.* 2020). Para a Audeguin *et al.* (2020), também apresenta um crescimento inicial rápido e produz uvas balanceadas com boa maturidade o que possibilita blends bastante apreciados nas uvas de vinho. Sendo aconselhado seu uso tanto com cultivares copa de uva de vinho finos quanto uvas de mesa precoces (Giovannini 2014).



Figura 10 - Cultivar 101-14 MGt. Fonte: Verslype (2020).

2.5.9 Harmony

A cultivar Harmony (Figura 11) foi selecionada em 1956 em Fresno, na Califórnia, EUA, pelos melhoristas Weinberger e Harmon (Bettiga *et al.* 2003, Garris *et al.* 2009, Peterson *et al.* 2019). Conforme relata Leão e Silva (2014) esta cultivar tem possibilitado bom desempenho quando usada com cultivares copa de uvas de mesa sem semente, muito vigorosas como 'Thompson Seedless', por conta do equilíbrio entre o desenvolvimento vegetativo e qualidade de frutos. Além disso, apresenta moderada aptidão na enxertia e bom enraizamento (Keller 2015). Bettiga *et al.* (2003) destacam que esta cultivar também apresenta resistência aos nematóides que varia de média à alta, assim como, resistência média a baixa à filoxera, sendo sensível a solos salinos.



Figura 11 - Cultivar Harmony. Fonte: Verslype (2020).

2.6 A Vitivinicultura e as mudanças climáticas

A influência do clima no agronegócio nunca foi tão evidente quanto no setor vitivinícola, pois encontra-se intimamente ligado a geografia do vinho, interferindo diretamente na escolha do local de plantio e cultivares, potencial vegetativo, desenvolvimento, comportamento fitossanitário, rendimento e inclusive no amadurecimento e qualidade final das uvas (Moura *et al.* 2009, Fraga *et al.* 2013, Jones *et al.* 2012, Jones 2016). Dessa forma, o sucesso de todos os sistemas agrícolas foi ajustado a determinado nicho climático, permitindo assim obter uma produção economicamente sustentável (Jones *et al.* 2012, Jones 2016).

Nesse sentido, qualquer variação que ocorra no clima para a produção de vinhos, deve levar em conta vários fatores que operam em escalas temporais e espaciais (Jones *et al.* 2012). Por conta das plantas, serem imóveis e autotróficas, necessitando do suprimento de água e nutrientes do solo, assim como, do CO₂ da atmosfera e da luz para se desenvolverem (Kathpalia and Bhatla 2018). Dessa forma, às mudanças climáticas tornam a produção agrícola extremamente vulnerável, principalmente no quesito de disponibilidade hídrica (Jones 2016), devido a água ser essencial para a vida, que dependendo da espécie representa até 99% da sua composição, permitindo assim o seu crescimento e reprodução (Taiz and Zieger 2004, Soares and Costa 2009, Kathpalia and Bhatla 2018).

Para Ashraf (2010), a seca é sem dúvida um dos principais estresses abióticos do mundo, ocasionando perdas de produção consideráveis, devido as projeções indicarem alterações significativas, tanto nas temperaturas quanto na intensidade e regime das precipitações nas próximas décadas (Marguerit *et al.* 2011, Fraga *et al.* 2013). O aquecimento global impactará em um aumento de 20% na escassez de água no mundo, acometendo desde áreas atualmente propensas à seca, assim como, diversos ecossistemas (Silva *et al.* 2011). Apesar destas projeções de mudanças espaciais exatas do clima no futuro sejam especulativas neste momento, para Jones e Webb (2010), tem-se a certeza que os climas do futuro, tanto a curto como a longo prazo, serão diferentes dos de hoje. E estas mudanças, podem gerar impactos econômicos significativos sobre o setor, principalmente nas regiões vinícolas mais famosas e tradicionais, que se adequam a determinadas cultivares e produzem diferentes estilos de vinhos, apresentando terroir próprio, por geralmente possuírem

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

características ambientais específicas diretamente interferidas pelo clima (Jones *et al.* 2012, Fraga *et al.* 2013).

Entre estes impactos Jones e Webb (2010) relatam, a pressão adicional sobre o abastecimento de água cada vez mais escasso, mudanças no ciclo fenológico da videira, interrupção ou alteração da composição equilibrada das uvas e de vinhos regionais, necessidade de alterar as variedades cultivadas, mudanças nos estilos regionais de vinho e mudanças espaciais nas regiões viáveis de cultivo de uvas no mundo. Dessa forma, o maior desafio de adaptação será provavelmente a forma como os produtores e consumidores responderão culturalmente às mudanças nas identidades regionais devido a mudanças nas cultivares ou no estilo do vinho (Jones *et al.* 2012).

Por conta de grande parte dos vinhedos já se encontrarem localizados em regiões com seca sazonal, que ocasionam grandes restrições no rendimento e na qualidade final das uvas (Zarrouk *et al.* 2016), e para obter-se uma produção economicamente sustentável nestas regiões vinícolas, torna-se essencial o uso da irrigação (Basso *et al.* 1999). No entanto, nestas regiões existe uma crescente demanda por irrigação, para possibilitar que os vinhedos resistam a um leve estresse hídrico, com pouca ou nenhuma redução no rendimento e um potencial impacto positivo na qualidade da baga (Zarrouk *et al.* 2016). Conforme Fort *et al.* (2017), relatam que tanto nos Estados Unidos quanto na Austrália, recentemente as principais regiões produtoras de uvas foram acometidas por secas severas. Dessa forma, autores como Chaves *et al.* (2010), Zarrouk *et al.* (2016) e Fort *et al.* (2017), enfatizam a necessidade de melhoria na eficiência do uso da água, visto a crescente demanda por irrigação nos vinhedos, para aumentar a competitividade destas regiões, face à escassez dos recursos hídricos e o alto custo dos insumos (Netzer *et al.* 2009). A água é um elemento natural precioso, que deve ser utilizado de maneira eficiente pela agricultura intensiva, para que não venha a causar impactos ao ambiente, assim como, ao próprio homem (Soares and Costa 2009). Desta forma, torna-se crucial o uso de medidas de adaptação à seca (Malheiro *et al.* 2010).

Embora várias abordagens possam ser empregadas para aliviar o problema da seca na agricultura, o melhoramento genético de plantas, apresenta-se como um meio eficiente e economicamente sustentável que possibilita adaptar as culturas para produzirem com sucesso em ambientes propensos à seca (Ashraf 2010), como a mudança de cultivares (Malheiro *et al.* 2010), visto que as videiras segundo Zarrouk

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

et al. (2016), possuem uma imensa diversidade genética. Dessa forma, a obtenção de novos materiais vegetais podem reduzir a necessidade de irrigação e aumentar a conservação da água, mantendo a produtividade, e protegendo as videiras de danos ocasionados pela seca a longo prazo (Zhang *et al.* 2016).

No entanto, os mercados locais e globais frequentemente restringem as cultivares que podem ser cultivadas e, desta forma, tem-se procurado cada vez mais conferir características particulares de interesse como a tolerância à seca, por meio de enxertia para não alterar as cultivares copas exploradas tradicionalmente (Zhang *et al.* 2016). Nesse sentido, os porta-enxertos podem ser o melhor meio de adaptação dos vinhedos face às mudanças climáticas (Marguerit *et al.* 2011), sendo inclusive considerado por Ollat *et al.* (2016), como um elemento-chave de adaptação. Inúmeros trabalhos demonstraram o papel positivo dos porta-enxertos tolerantes à seca no controle da condutância estomática foliar e, portanto, na transpiração da cultivar copa (Serra *et al.* 2014). Desta forma, para Marguerit *et al.* (2011), os programas de melhoramento genético de porta-enxertos de videira devem priorizar entre os seus objetivos à sua adaptação ao déficit hídrico, o que para Fort *et al.* (2017), pode contribuir significativamente na continuidade da produção sustentável dos vinhedos, à medida que diminui o suprimento de água disponível para a irrigação.

2.7 Estresse hídrico nas videiras

As chuvas, a umidade do solo e o balanço hídrico são características que devem ser levadas em conta, dada a sua influencia no equilíbrio da videira, na qualidade dos seus frutos e na pressão de doenças, sendo muito importante compreender as relações d'água em qualquer região vinícola (Jones *et al.* 2012). Principalmente em regiões que apresentam seca sazonal com baixa disponibilidade d'água no solo associado a altas temperaturas e baixa umidade do ar, que ocasionam déficit hídrico nos vinhedos (Zarrouk *et al.* 2016), como pode ser observado no Submédio do Vale do São Francisco.

De acordo com Lima Filho *et al.* (2009) quando o conteúdo de água disponível no solo não é suficiente para repor as perdas de água da videira por evapotranspiração, ocorre o estresse hídrico. Segundo Dane and Hopmans (2002), a água disponível no solo para as plantas depende do tipo de poros e da energia com que a água se encontra retida nestes, estando disponível para as plantas entre a capacidade de campo e o ponto de murcha permanente. Dessa forma, solos com texturas distintas

podem apresentar valores de potencial mátrico diferentes, sendo de -0,01 MPa para solos de textura mais grosseira (Cassel and Nielsen 1986), e de -0,033 MPa para solos argilosos (Colman 1947). Segundo Taiz and Zieger (2004), o ponto de murcha permanente, em geral ocorre quando o conteúdo d'água disponível no solo encontra-se retido em um potencial mátrico de -1,5 MPa, ocasionando assim uma murcha irreversível nas plantas, no qual não conseguem readquirir a pressão de turgor das folhas.

As videiras, de acordo com Ojeda *et al.* (2004), apresentam uma demanda média de 1 litro de água para produção de 2 gramas de massa seca, destes 82% são transpirados, 17% armazenado e apenas 1% utilizado na fotossíntese para a produção de açúcares. Para repor as grandes quantidades de água perdidas pela transpiração para a atmosfera, a videira necessita retirar do solo um volume equivalente (Giovannini 2014). Segundo Lima Filho *et al.* (2009), o estado hídrico da videira é considerado uma variável de grande importância por afetar a qualidade da uva, o seu crescimento vegetativo, e inclusive seu metabolismo, devido a fotossíntese ser um dos processos fisiológicos sensíveis a esta condição. Ojeda *et al.* (2004) destacam que, um potencial hídrico do solo entre -0,2 e -0,6 MPa é suficiente para ocasionar um estresse hídrico moderado na videira.

Basso *et al.* (1999) evidenciam, que a deficiência hídrica pode, simultaneamente afetar os tecidos e os seus processos reprodutivos por se tratar de uma espécie perene. Dessa forma, o momento e a intensidade do estresse hídrico, podem afetar negativamente ou induzir um impacto positivo na qualidade final da baga, no entanto, isso pode variar de acordo com a cultivar empregada (Zarrouk *et al.* 2016). Um estresse hídrico moderado após o início da maturação, para Ojeda *et al.* (2004), pode favorecer a qualidade enológica da uva por influenciar no metabolismo das bagas, e, como consequência, nas características sensoriais do vinho (Zarrouk *et al.* 2016). Devido ao controle do vigor vegetativo, pela manutenção plena da função foliar em conjunto com a restrição no crescimento vegetativo, promovem maior direcionamento de fotoassimilados das folhas para síntese de pigmentos e compostos aromáticos na uva (Ojeda *et al.* 2004). No entanto, quando as videiras sofrem alguma deficiência hídrica antes do início da maturação, o crescimento do fruto é afetado negativamente, tanto no seu volume quanto em tamanho (Basso *et al.* 1999). Principalmente quando as videiras são submetidas a uma condição com excesso de restrição hídrica, em conjunto a paralisação do crescimento e maturação da uva, ocorre uma redução

drástica da qualidade enológica por influenciar na concentração de metabólitos das bagas. (Ojeda *et al.* 2004, Zarrouk *et al.* 2016). Além disso, compromete também o acúmulo de reservas na planta, e é possível observar a paralisação da maturação dos ramos (Ojeda *et al.* 2004).

Quando a deficiência hídrica ocorre em estádios iniciais do ciclo vegetativo, o crescimento dos ramos laterais podem ser menores, e praticamente nulos quando ocorre entre a fase de pegamento dos frutos e intermediária de crescimento da baga (Basso *et al.* 1999). Outra sintomatologia nas videiras bastante observada é a murcha de suas folhas e de gavinhas mais novas, por conta da redução do conteúdo de água após alguns dias de sua falta provocando alterações fisiológicas em suas folhas. E caso a deficiência se estenda por um período maior, evoluem posteriormente, para alterações físicas nas folhas como o aparecimento de áreas necrosadas e a sua senescência (Basso *et al.* 1999, Lima Filho *et al.* 2009), conforme pode ser visualizado na (Figura 12).

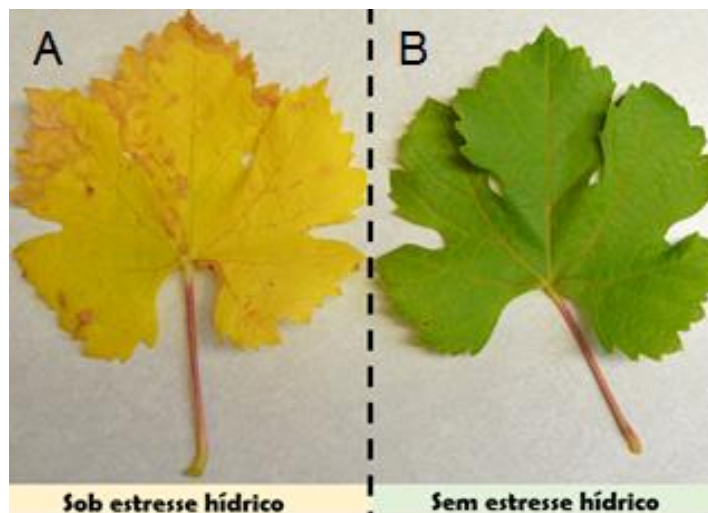


Figura 12 - Folhas da cultivar Merlot, com sintomas de estresse hídrico (A) e sem sofrer estresse (B). Fonte: Adaptado de Keller (2015).

Para Alsina *et al.* (2011), as plantas perenes respondem ao estresse hídrico de várias maneiras importantes que variam em escala temporal. Nas videiras, podemos observar que as cultivares de porta-enxertos diferem em sua capacidade de fornecer água a copa em situações de estresse hídrico (Serra *et al.* 2014). Esta variação pode ser explicada, devido aos avanços recentes em pesquisa nesta área demonstrarem que, a tolerância à seca na videira é uma característica controlada por um grande número de genes, e muitos destes genes são quantitativos, dificultando assim, a compreensão de suas respostas aos déficits hídricos (Ashraf 2010, Marguerit *et al.* 2012, Serra *et al.* 2014, Dayer *et al.* 2019).

Foram detectadas que 17 proteínas na cultivar 101-14 MGt e 23 proteínas na cultivar M4, foram afetadas, estando envolvidas nas respostas ao estresse hídrico, demonstrando à nível proteômico diferenças entre genótipos (Prinsi *et al.* 2018). Dessa forma, as videiras reagem tanto em níveis fisiológicos e bioquímicos, quanto genéticos para tolerar as restrições hídricas, apresentando algumas características adaptativas de prevenção na perda de água (Serra *et al.* 2014) (Figura 13).

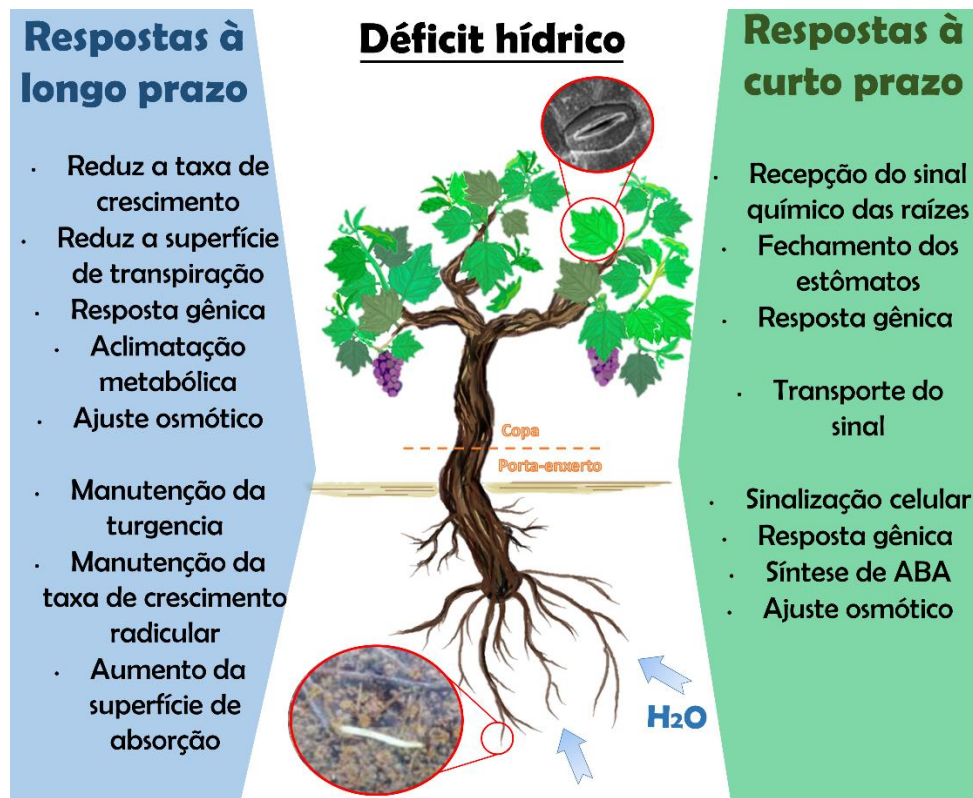


Figura 13 - Respostas da videira quando submetida a curto e longo prazo sob déficit hídrico. Fonte: Adaptado de Chaves *et al.* (2003) e Serra *et al.* (2014).

A tolerância à seca deve ser entendida como um compromisso entre a sobrevivência da planta e a sua produção (Serra *et al.* 2014). As raízes desempenham um papel central na resposta das videiras ao estresse hídrico, atuando desde a sua percepção e sinalização para as folhas, até possibilitar que a planta se adapte para manter um balanço hídrico adequado (Prinsi *et al.* 2018). Dessa forma, a máquina genética permite que os porta-enxertos controlem a capacidade de extração de água e a transpiração da cultivar copa (Serra *et al.* 2014).

Os porta-enxertos de videira tolerantes à seca, devem possibilitar que a cultivar copa se desenvolva e produza mesmo quando o suprimento de água é limitado (Serra *et al.* 2014). A compreensão e conhecimento de todos os mecanismos envolvidos na

influência dos porta-enxertos que confere tolerância à seca e as suas interações entre porta-enxerto e copa estão longe de estarem completos (Zhang *et al.* 2016). Dessa forma, estudos que possibilitem compreender os mecanismos de tolerância à seca induzidos e regulados por porta-enxertos de videira, assim como, as bases fisiológicas das respostas em função da intensidade do estresse que variam de leves a moderados, é fundamental para otimizar o manejo da irrigação por déficit, assim como, para facilitar na identificação de genótipos mais adequados a essas condições em programas de melhoramento, a fim de desenvolver cultivares de porta-enxertos tolerantes à seca para cada região vitivinícola, principalmente aquelas que apresentam climas áridos e semiáridos (Chaves *et al.* 2010, Marguerit *et al.* 2012, Serra *et al.* 2014, Zarrouk *et al.* 2016, Zhang *et al.* 2016).

Entre as características conhecidas, que contribuem para a resposta à seca dos porta-enxertos, destacam-se o ajuste osmótico nas raízes por meio de padrões de crescimento e anatomia da raiz, e o controle de perda de água por meio de sinais químicos e físicos relacionados à condutância estomática (Serra *et al.* 2014, Tsegay *et al.* 2014), conforme descritas a seguir:

2.7.1 Arquitetura do sistema radicular

O sistema radicular das videiras desempenha como principais funções a sustentação da planta, assim como, atua na absorção e transporte de água e nutrientes, no acúmulo de reservas e como fonte de fitormônios (Smart *et al.* 2006, Keller 2015). Para Magalhães Filho *et al.* (2008), a forma como se desenvolve o sistema radicular, tanto na sua distribuição como no seu crescimento, afeta diretamente nas respostas das plantas à ocorrência de deficiência hídrica. Quando bem desenvolvido, pode representar um terço do peso seco total da planta (Giovannini 2014).

Nas videiras, de acordo com Ollat *et al.* (2016), os efeitos causados pelo porta-enxerto no desempenho das cultivares copas estão associados às suas propriedades fisiológicas e estruturais como a morfologia de suas raízes. Conforme Fernández-Cano and Togores (2011) comentam, as raízes das cultivares provenientes da espécie *Vitis berlandieri*, assim como, seus híbridos e inclusive nas espécies de *Vitis vinifera* é possível observar raízes grossas e carnosas, enquanto nas cultivares originadas da espécie *Vitis rupestris* tendem a apresentar raízes finas, fibrosas e duras. Para Fort *et*

al. (2017), quando as cultivares de porta-enxertos tolerantes à seca, são submetidas a condições sob estresse hídrico o seu sistema radicular apresenta como características a manutenção da taxa de crescimento de raízes e raízes profundas (Villa, 2018).

No entanto o papel desempenhado pelo sistema radicular na tolerância à seca tem informações muito limitadas, em comparação com as respostas observadas acima do solo, e ainda não está totalmente esclarecido como os porta-enxertos podem melhorar a tolerância à seca da cultivar copa (Alsina *et al.* 2011, Serra *et al.* 2014). Segundo Tsegay *et al.* (2014), a anatomia da raiz e os seus padrões de crescimento podem contribuir significativamente para a resposta à seca dos porta-enxertos. Em razão de apresentarem diferenças, tanto na constituição anatômica das raízes quanto nos padrões de crescimento, que variam desde o calibre dos vasos até a proporção de fibras liberianas em relação às xilemáticas, assim como, no ângulo geotrópico, as espécies de videira possuem diferentes graus de adaptação aos diversos tipos de solo (Albuquerque 2003, Giovannini 2014). Dessa forma, pode-se citar entre as variáveis que afetam na indução de tolerância à seca o ângulo geotrópico e a profundidade de suas raízes.

Os genótipos de videira e o sistema de cultivo determinam a distribuição da profundidade de enraizamento (Smart *et al.* 2006). Dessa forma, genótipos distintos possuem diferentes padrões de crescimento do sistema radicular que são controlados a nível genético, conforme pode ser visualizado na (Figura 14), no estudo de Fort *et al.* (2017), pode-se observar que todas as cultivares submetidas a um período de seis dias sob estresse hídrico, apresentaram diferentes padrões de crescimento do sistema radicular e tolerância à seca, como no caso da cultivar 101-14MGt que apresentou uma distribuição radicular muito superficial e susceptível ao estresse hídrico junto com a Riparia, enquanto a cultivar Ramsey apresentou a maior distribuição de raízes e mostrou-se com alta tolerância à seca junto com o 110R, mantendo a sua taxa produção de novas raízes e induzindo uma menor taxa de crescimento à copa Merlot. Porém, os genótipos são afetados pelas condições edafoclimáticas do local onde encontra-se o vinhedo, devido a interação genótipo x ambiente, interferindo assim, no padrão de crescimento de novas raízes laterais (Serra *et al.* 2014), principalmente quando se tem a presença de uma camada de impedimento no solo.

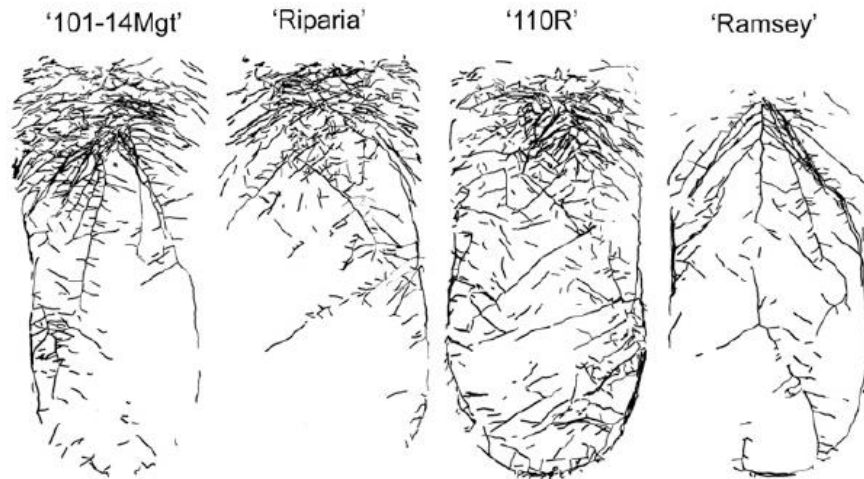


Figura 14 - Distribuição do sistema radicular de quatro cultivares de porta-enxertos de videira, sendo 101-14MGt (Millardet et de Grasset 101–14), Riparia (Riparia Gloire de Montpellier), 110R (Richter 110), Ramsey (Salt creek). Fonte: Fort *et al.* (2017).

Geralmente a área coberta pelas raízes das plantas é muito grande (Kathpalia and Bhatla 2018). Não obstante das videiras, em que o seu sistema radicular se desenvolve tanto horizontalmente como verticalmente, as suas raízes podem atingir profundidades entre dois e três metros ou até superiores a cinco metros quando em solos arenosos sem camadas de impedimento, tais como, a presença de horizontes compactados no solo, pedregosidade, rochas, barreiras químicas e ou lençol freático superficial (Albuquerque 2003, Smart *et al.* 2006, Mandelli *et al.* 2009, Soares and Costa 2009, FAO, 2020). Concentrando a sua região de captação máxima de água, onde ocorre o desenvolvimento ativo das raízes laterais (Kathpalia and Bhatla 2018), entre 1 e 2 m de profundidade, onde são absorvidos normalmente 100% da água extraída (FAO, 2020).

De acordo com Magalhães Filho *et al.* (2008), a maior razão entre a raiz e parte aérea é considerado uma característica importante de escape dos porta-enxertos ao déficit hídrico, pois o balanço funcional entre a copa e as raízes pode ser alterado de acordo com a disponibilidade hídrica (Taiz and Zieger 2004). As cultivares 110R e Ramsey, consideradas com alta tolerância à seca, mantiveram a taxa de produção de novas raízes sob estresse hídrico, enquanto as cultivares 101-14MGt e Riparia Gloire de Montpellier ambas classificadas com baixa tolerância à seca, apresentaram um declínio na taxa de produção de novas raízes (Fort *et al.* 2017).

Estes ajustes na arquitetura do sistema radicular, para Alsina *et al.* (2011), é um dos principais meios das plantas lenhosas perenes tolerarem à seca, e geralmente nas videiras, ocorre em dias ou até semanas (Serra *et al.* 2014). Através da inibição da expansão foliar, reduzem a superfície de transpiração e o consumo de carbono e

energia, possibilitando direcionar uma maior porção de fotoassimilados para o sistema radicular para sustentar o seu crescimento em profundidade e ramificações, aumentando a densidade de raízes e a capacidade em fornecer água para a parte aérea (Chaves *et al.* 2003, Taiz and Zieger 2004, Alsina *et al.* 2011, Serra *et al.* 2014, Ollat *et al.* 2016). Desta forma, a densidade e extensão das raízes são consideradas características importantes para a extração de água no solo (Taiz and Zieger 2004, Kathpalia and Bhatla 2018). Pois, Mandelli *et al.* (2009) mencionam, que vinhedos com bom desenvolvimento e aprofundamento de suas raízes, podem suportar, por curtos períodos, potenciais de água no solo próximos ao do ponto de murcha de -1,5 MPa. Devido as suas raízes em camadas mais profundas do solo apresentarem uma vida útil mais longa do que em solo raso (Ollat *et al.* 2016), e a possibilidade de alcançar regiões mais profundas do solo que permanecem úmidas (Serra *et al.* 2014).

Nem todas as plantas possuem a mesma tendência em aprofundar o seu sistema radicular no solo, nas videiras, esta capacidade varia em função do seu ângulo geotrópico (Fernández-Cano and Togores 2011). O ângulo geotrópico, de acordo com Giovannini (2014), se caracteriza por ser o ângulo formado entre a direção da raiz e o tronco da videira com a vertical, conforme pode ser visualizado na (Figura 15).

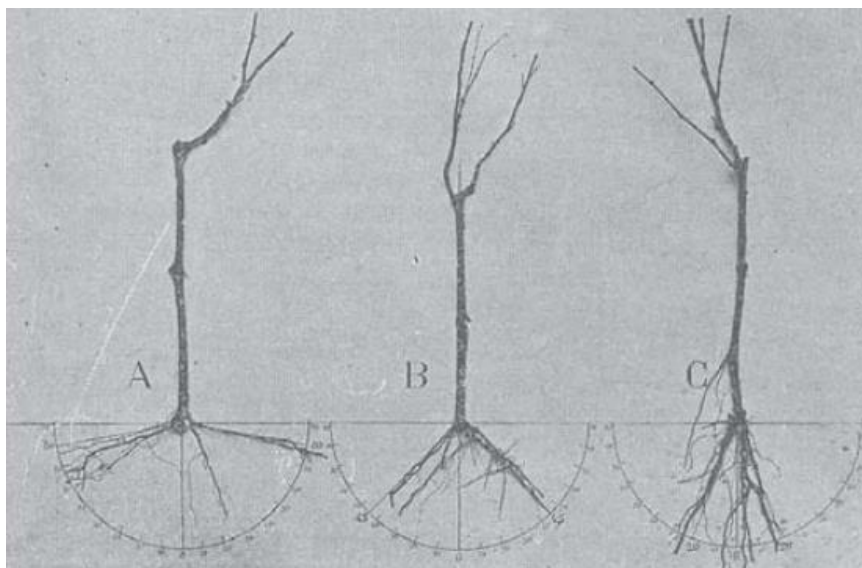


Figura 15 - Diferenças no ângulo geotrópico em função da disposição de emergência de raízes adventícias em diferentes cultivares de porta-enxertos, sendo A - Riparia Gloire de Montpellier, B - 3309C e C - Rupestris du Lot. Fonte: Smart *et al.* (2006).

De acordo com Giovannini (2014), o ângulo geotrópico é muito influenciado pelo regime de chuvas, pela textura e manejo do solo, e apesar de ser considerado uma característica genética da cultivar, percebeu-se que cultivares de porta-enxertos com ângulo geotrópico aberto seriam mais sensíveis à seca do que as de ângulo mais

fechado. Devido às plantas com ângulo geotrópico aberto favorecerem o crescimento superficial das raízes, conforme pode ser visualizado no caso da cultivar Riparia Gloire de Montpellier apresentada na (Figura 14A) (Smart *et al.* 2006, Magalhães Filho *et al.* 2008, Fernández-Cano and Togores 2011, Serra *et al.* 2014, Keller 2015), que é classificada por diversos autores como uma cultivar muito sensível ao estresse hídrico (Morton 1979, Carbonneau 1985, Bettiga *et al.* 2003, ATVB 2013, Giovannini 2014, Villa 2018, Audeguin *et al.* 2020, Carroll 2020, Rebschule Mueller 2020).

O ângulo geotrópico mais fechado favorece o aprofundamento das raízes conferindo uma maior tolerância à seca por possibilitar que as raízes alcancem camadas mais profundas e úmidas do solo absorvendo nutrientes móveis solubilizados em água, como no caso da cultivar Rupestris du Lot (Smart *et al.* 2006, Magalhães Filho *et al.* 2008, Fernández-Cano and Togores 2011, Serra *et al.* 2014, Keller 2015), conforme pode ser observado na (Figura 14C), que é classificada na literatura como uma cultivar de média tolerância à seca (Satisha *et al.* 2006, Keller 2015, Villa 2018, Audeguin *et al.* 2020). Enquanto a cultivar 3309C, com um ângulo intermediário (Figura 14B), é classificada como uma cultivar sensível ao estresse hídrico (Chevalier 1925, Morton 1979, Brancadoro *et al.* 2014, Wine Australia 2016, Villa 2018, Audeguin *et al.* 2020, Carroll 2020, Pecile *et al.* 2020, VICOPAD 2020).

2.7.2 Controle estomático

Os estômatos são estruturas microscópicas presente na epiderme das folhas das plantas, formada pelo poro estomático e por duas células guarda especializadas, que controlam a troca de vapor d'água e CO₂ com a atmosfera (Franks and Beerling 2009). De acordo com Kathpalia and Bhatla (2018), os estômatos são responsáveis por mais de 90% da perda de água devido à transpiração nas plantas. Dessa forma, segundo Marguerit *et al.* (2012), uma das principais estratégias pelas quais as plantas lidam com o estresse hídrico é através do controle estomático.

Para Chaves *et al.* (2003) o fechamento dos estômatos está entre as respostas iniciais das plantas sob déficits hídricos leves a moderados, a sua abertura e fechamento é ocasionado pelo movimento das células guarda (Kathpalia and Bhatla 2018). Nas videiras, segundo Giovannini (2014), o déficit hídrico ocasiona o fechamento progressivo dos estômatos, diminuindo assim a transpiração, a absorção e fixação de carbono (Chaves *et al.* 2003), e conseqüentemente a taxa fotossintética (Albuquerque 2003).

Em situações de estresses hídrico muito intenso, ou seja, inferior a $-1,2$ Mpa, ocorre o fechamento estomático (Giovannini 2014), ocasionando um alto custo para a videira (Serra *et al.* 2014), devido a transpiração ser um dos processos metabólicos de maior importância e de efeito benéfico para a planta, por desencadear o mecanismo de absorção de água e nutrientes minerais, produção de energia, assim como, a sua atuação no resfriamento e manutenção da sua temperatura de suas folhas possibilitando tolerar pressões ambientais severas (Giovannini 2014, Kathpalia and Bhatla 2018).

Segundo Serra *et al.* (2014), um ajuste na condutância estomática pode ocorrer dentro de minutos ou segundos na videira. Para Chaves *et al.* (2010), já se reconhece que o controle estomático depende muito do genótipo, principalmente a respeito da anatomia, estado hídrico da planta, momento e intensidade da resposta aos déficits hídricos no solo e na atmosfera. O sistema radicular das videiras não só pode melhorar a captação e o transporte de água, mas também pode detectar déficits de água no solo e enviar sinais que regulam o funcionamento estomático (Serra *et al.* 2014), ocasionando alterações fisiológicas nas videiras, que possibilitam reduzir a transpiração (Basso *et al.* 1999).

Evidenciam-se assim, o papel fundamental do porta-enxerto no controle da transpiração da cultivar copa (Marguerit *et al.* 2012). Nesse sentido, os estômatos podem sofrer o fechamento hidropassivo ou hidroativo, sendo o primeiro ocasionado pela perda da turgidez da célula-guarda, enquanto o fechamento hidroativo é mediado pelo ácido abscísico (ABA) (Kathpalia and Bhatla 2018).

As plantas enfrentam uma dissecação potencialmente letal através do poro estomático, por conta do conflito entre a necessidade de assimilação de CO_2 e a conservação de água, ocasionando assim perda de água para a atmosfera (Taiz and Zieger 2004). Devido a esta problemática, segundo Kathpalia and Bhatla (2018), as plantas desenvolveram vários mecanismos adaptativos para evitar esta perda de água para o ambiente através da transpiração, como por exemplo a presença de estômatos apenas no lado inferior das folhas, conforme pode ser observado nas videiras (Albuquerque 2003).

A densidade de estômatos e o tamanho do poro estomático são características variáveis entre as espécies, de acordo com Franks and Beerling (2009), ambas as características determinam a condutância estomática máxima que pode ser alcançada pela folha para os locais de assimilação de CO_2 . Para Giovannini (2014), nas videiras

em geral este número varia entre 50 e 400 estômatos/mm² nas folhas das viníferas, enquanto na grande maioria dos porta-enxertos o número médio de estômatos varia entre 125 e 250 estômatos/mm², enquanto as videiras americanas em geral apresentam mais de 300 estômatos/mm². Um grande número de estômatos com pequeno tamanho do poro estomático possibilitam alcançar taxas mais altas de fotossíntese, porém também podem aumentar as taxas de transpiração (Franks and Beerling 2009).

A densidade de estômatos por unidade de área e o seu tamanho podem ser afetados significativamente por restrições hídricas e de intensidade luminosa, assim como, pelo uso de diferentes cultivares de porta-enxertos em uma mesma cultivar copa (Serra *et al.* 2014). Conforme pode ser visualizado na (Figura 16), em que a cultivar copa Pinotage, quando enxertada na cultivar 140 Ru apresentou menor densidade estomática e maior diâmetro do poro estomático do que quando enxertada nas cultivares 110 R e 1103 P. Enquanto, o déficit hídrico induziu uma redução do diâmetro de poros, e a menor intensidade de luz aumentou o diâmetro de poros e diminuiu a densidade estomática. Para Franks and Beerling (2009), estas reduções de tamanho do poro estomático ocasionam uma redução nas perdas d'água das folhas para a atmosfera.

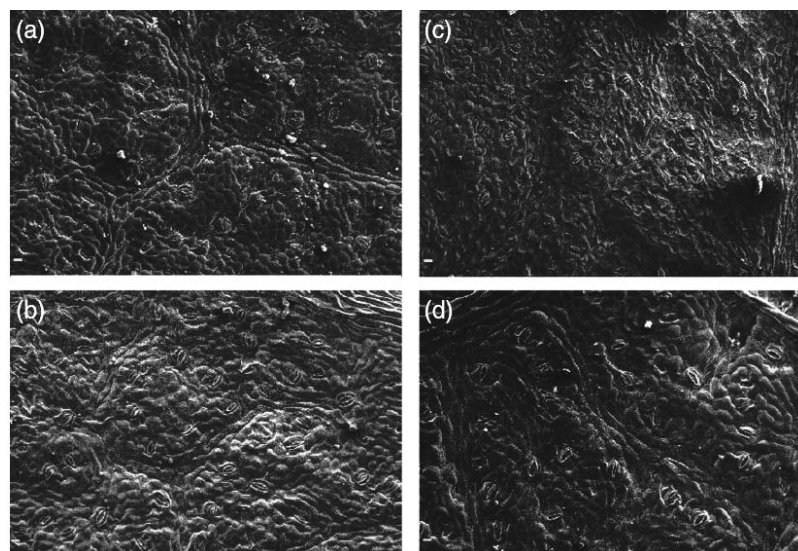


Figura 16 - Variação no diâmetro de poros estomáticos e no número de estômatos da cultivar Pinotage. Sendo A - 1103 P sob restrições hídricas e exposto a luz solar (119.1 ± 6.3 poros/mm² e $13.2 \pm 0.31 \mu\text{m}$), B - 1103 P sob restrições hídricas e sombreado (91.0 ± 6.3 poros/mm² e $20.0 \pm 0.3 \mu\text{m}$), C – 140 Ru sob restrições hídricas e exposto a luz solar (113.8 ± 6.3 poros/mm² e $16.0 \pm 0.31 \mu\text{m}$) e D – 140 Ru sob restrições hídricas e sombreado (96.3 ± 6.3 poros/mm² e $17.2 \pm 0.31 \mu\text{m}$). Ampliado 150 vezes. Fonte: Serra *et al.* (2014).

Apesar de ainda não serem explicados como os porta-enxertos afetam no desenvolvimento estomático das copas, Serra *et al.* (2014) mencionam, que devem

estar relacionadas com a interação do vigor induzido pelo porta-enxerto na copa e na sua demanda de água. Dessa forma, segundo Fregoni *et al.* (1978), a densidade de estômatos é um parâmetro que pode ajudar na seleção precoce de genótipos de porta-enxertos tolerantes à seca, por ser uma característica de alta herdabilidade e geneticamente relacionada com a seca.

Serra *et al.* (2014) comentam que essas diferenças na condutância estomática entre os porta-enxertos interferem no estado hídrico da planta, afetando o crescimento das folhas e, conseqüentemente, causando variabilidade na densidade e tamanho estomático da cultivar copa, estando intimamente relacionado à troca de gases nas folhas e à eficiência no uso da água.

2.7.3 Fitormônios x Estresse hídrico

A medida que o solo seca, sinais químicos são produzidos pelas raízes das videiras, sendo importantes principalmente em estágios iniciais do estresse hídrico, por possibilitar que se adaptem a este estresse abiótico (Schachtman and Goodger 2008). Entre os fitormônios que podem estar associados ao estresse hídrico nas plantas pode-se citar o etileno, as citocininas e o ácido abscísico (Tsegay *et al.* 2014, Keller 2015, Kathpalia and Bhatla 2018). No geral, o ácido abscísico (ABA) é um sinal químico dominante no controle do crescimento e da transpiração das espécies vegetais (Schachtman and Goodger 2008, Tsegay *et al.* 2014), atuando como sinalizador de estresses nas videiras (Giovannini 2014), pois em perdas d'água da ordem de 10%, que ocasionam o murchamento de suas folhas, desencadeiam um aumento rápido na concentração de ABA, em média 40 vezes maior, promovendo assim o fechamento rápido dos estômatos (Lima Filho *et al.* 2009), e geralmente essa resposta da videira ao ABA ocorre em horas (Serra *et al.* 2014).

Por conta disso, o ABA vem sendo referenciado na literatura como um antitranspirante, devido ao seu potencial em reduzir perdas d'água pelas plantas em situações de estresse hídrico (Kathpalia and Bhatla 2018). No entanto, podem ocorrer diferenças significativas nas respostas estomáticas em função do porta-enxerto empregado, devido a capacidade de produção do ABA ser parcialmente determinadas por sua genética (Dayer *et al.* 2019).

Apesar de seu papel reconhecido de indução do fechamento dos estômatos em situações de estresse hídrico, para Giovannini (2014), o ABA também atua na videira em condições normais, através da iniciação do processo de maturação da uva, a

medida que aumenta a sua concentração nos frutos em sintonia com a redução das auxinas. Dessa forma, o ABA normalmente é biossintetizado a partir de B-caroteno presente em folhas adultas, no entanto, em situações de déficit hídrico ou sob outros estresses abióticos, é sintetizado nas raízes sendo transportado através do xilema rapidamente para as folhas, e induzindo assim o fechamento estomático, que permite a videira regular sua perda de água e reduzir o crescimento de suas folhas (Schachtman and Goodger 2008, Giovannini 2014, Keller 2015, Kathpalia and Bhatla 2018, Dayer *et al.* 2019).

O transporte do ABA nas videiras em situação de estresse hídrico ocorre via xilema, que simultaneamente sofre normalmente uma alteração no pH da seiva do xilema (Schachtman and Goodger 2008, Tsegay *et al.* 2014). Esta alcalinização de seu pH atua sinergicamente com o ABA, interferindo na sua distribuição nas folhas da planta, favorecendo a produção da forma dissociada do ABA, que dificulta a travessia na membrana e reduz a sua absorção pelas células do mesofilo, propiciando assim, um aumento na quantidade de ABA que alcança as células guarda e conseqüentemente aprimorando o fechamento estomático induzido pelo ácido (Taiz and Zieger 2004, Schachtman and Goodger 2008), conforme pode ser visualizado na (Figura 17).

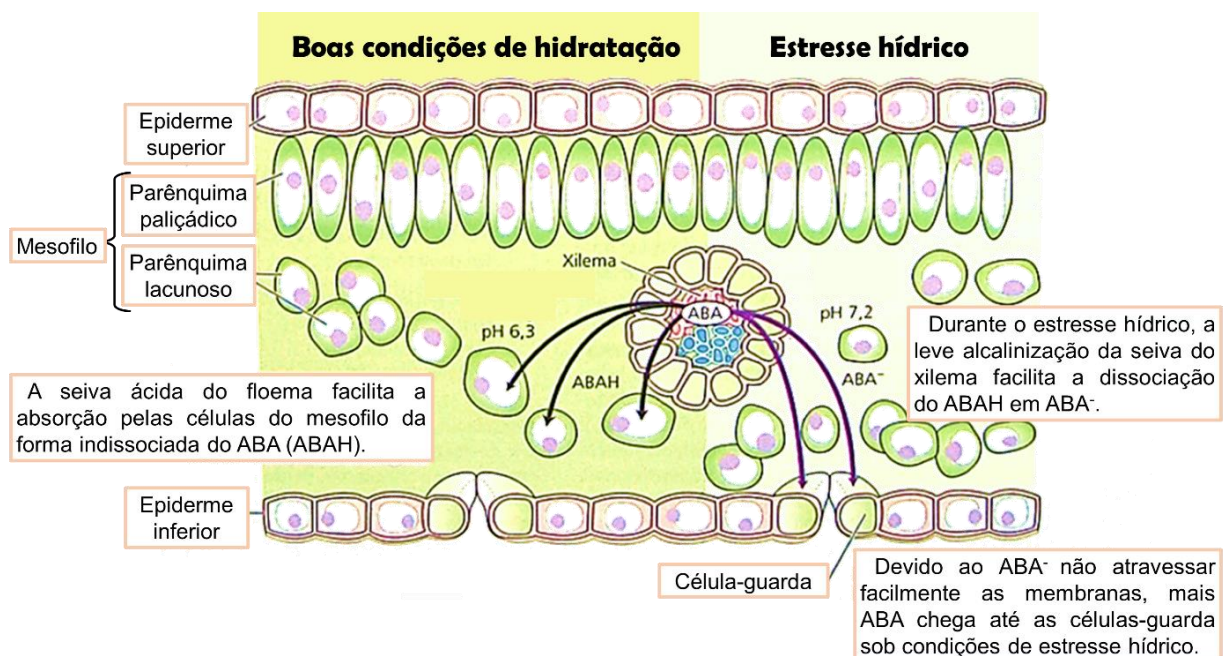


Figura 17 – pH da seiva do xilema sob estresse hídrico. Fonte: Taiz and Zieger (2004).

Os fitohormônios atuam inicialmente ao nível das membranas plasmáticas, onde existem proteínas receptoras de ABA nas células-guardas tanto intracelulares quanto extracelulares (Taiz and Zieger, 2004). Schachtman and Goodger (2008), destacam a

importância de receptores ligados à membrana plasmática como o GRC2 (receptor acoplado à proteína G2) dentre outros ainda desconhecidos, e de receptores intracelulares como o CHLH localizado no cloroplasto (subunidade H da protoporfirina-IX quelatase de magnésio), que são importantes na transdução das concentrações aumentadas de ABA nos apoplastos, ocasionando o fechamento estomático quando as células-guardas identificam o ABA.

O ABA causa uma despolarização da membrana plasmática da célula-guarda, que seria causada pelo aumento de Ca^{2+} (Taiz and Zieger, 2004). Dessa forma quando o fechamento estomático é induzido pelo ABA, um grande número de canais de íons é regulado por sinais de Ca^{2+} . O ABA então induz o fechamento estomático através da abertura dos canais aniônicos, lento e rápido presentes na membrana plasmática das células-guardas, permitindo o efluxo de Cl^- das células-guardas, além disso, também inibe a bomba H^+ e abre ainda mais os canais de K^+ , possibilitando que saiam das células-guardas, ocasionando a perda de íons (ânions e cátions) (Schachtman and Goodger 2008, Kathpalia and Bhatla 2018). E esta redução na concentração de solutos nas células-guardas é seguida pela perda de água, para as células subsidiárias, fechando assim os estômatos (Kathpalia and Bhatla 2018), conforme pode ser visualizado de forma esquematizada na (Figura 18).

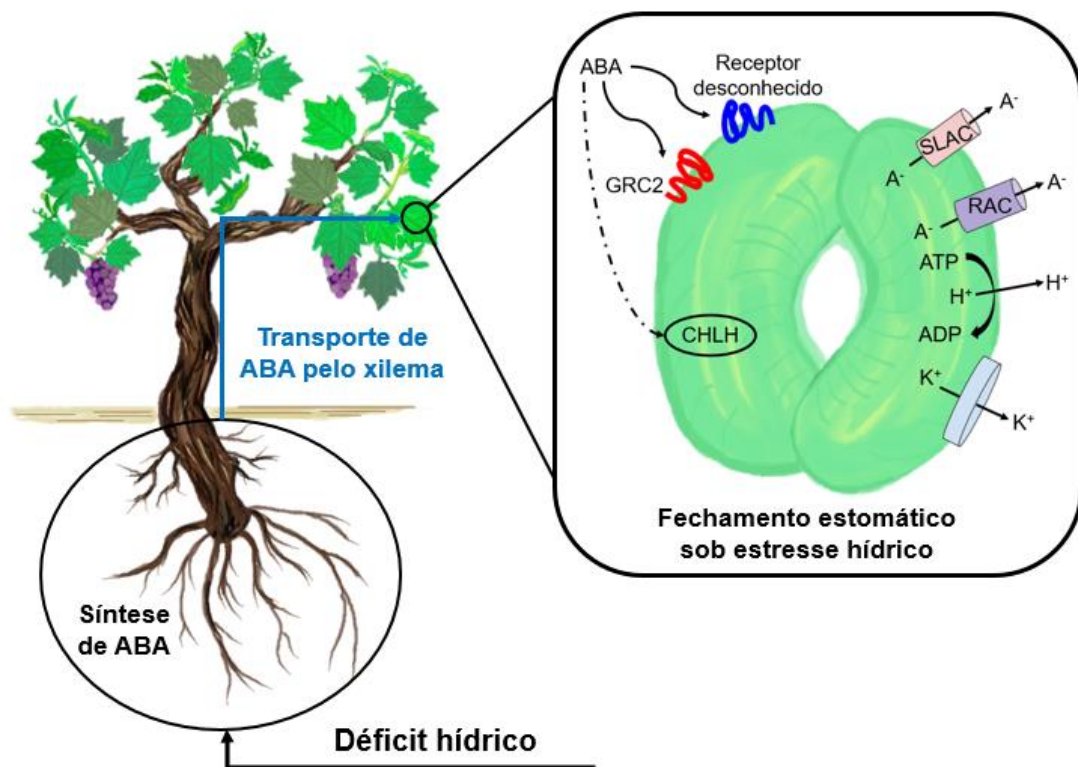


Figura 18 - Fechamento hidroativo dos estômatos sob estresse hídrico. Sendo A^- (ânions), SLAC (canal aniônico lento) e RAC (canal aniônico rápido). Fonte: Adaptado de Schachtman and Goodger, 2008.

2.7.4 Potencial osmótico

As plantas como as videiras, são compostas principalmente por água que atua como solvente para solutos orgânicos e inorgânicos (Keller 2015). A osmose é considerada como um processo biológico no qual as moléculas de solvente se deslocam através de membrana semipermeável, controlada tanto pela concentração quanto pelo gradiente de pressão, no qual o solvente se desloca da célula vegetal com menor concentração de soluto para a célula com maior concentração de soluto aumentando assim a sua pressão de turgor (Keller 2015, Kathpalia and Bhatla 2018). Essa tensão provocada pela presença de solutos em uma célula nas moléculas de água ao redor da célula, é denominada potencial osmótico (Keller 2015).

Dessa forma, esse tipo especial de difusão mantém a taxa de crescimento, desenvolvimento e turgidez das células vegetais através da osmorregulação, pois a medida que as células perdem água para a atmosfera, principalmente através dos poros estomáticos, vão recebendo, por osmose, mais água das células adjacentes, até atingirem os vasos que contêm água livre induzindo um déficit nas raízes. Isso ocasiona um rápido ajuste osmótico, e auxilia o restabelecimento da pressão de turgor que leva a captação de mais água do solo pelas raízes, estabelecendo um gradiente hídrico entre a atmosfera e o sistema de condução de seiva da videira, e permite a manutenção do alongamento celular (Magalhães Filho *et al.* 2008, Giovannini 2014, Keller 2015, Kathpalia and Bhatla 2018).

Normalmente o ajuste osmótico nas videiras, quando sob estresse hídrico, pode ocorrer em horas, induzindo o acúmulo de solutos (Chaves *et al.* 2010), e reduzindo o potencial osmótico nas raízes da videira, atraindo assim água para as células, e possibilitando a absorção de água e a pressão do turgor celular (Serra *et al.* 2014, Keller 2015). Dessa forma, a tolerância à seca com baixo potencial hídrico da planta envolve a tolerância à dessecação e a manutenção do turgor, principalmente por ajuste osmótico, restringindo a produção de ABA pelas raízes (Serra *et al.* 2014).

2.7.5 Osmorreguladores

De acordo com Keller (2015), o crescimento da videira é extremamente sensível a reduções no potencial hídrico. No entanto, a sua taxa de osmose pode ser aumentada pela adição de substâncias osmoticamente ativas, conhecidas como osmorreguladores, que envolvem solutos inorgânicos retirados do substrato e ou solutos orgânicos sintetizados pela própria planta, tais como a sacarose, o potássio,

o cálcio e a prolina (Serra *et al.* 2014, Keller 2015, Kathalia and Bhatla 2018). Dentre estes osmorreguladores, a prolina é um exemplo, pois trata-se de um aminoácido bastante estudado nas videiras por conta de sua relação com o estresse hídrico nas plantas (Giovannini 2014). De acordo com Keller (2015), durante o estresse hídrico, grande parte da síntese de prolina resulta da glutamina produzida pela forma GS1 da glutamina sintetase nas células companheiras do floema. Dessa forma, quando a planta encontra-se sob situações de estresse hídrico a prolina por ser altamente solúvel, é acumulada no citoplasma em níveis elevados, sem interromper o metabolismo da planta, possibilitando assim, manter em equilíbrio o potencial hídrico dentro da célula vegetal e a pressão de turgor durante os períodos de seca (Taiz and Zieger 2004, Keller 2015).

2.7.6 Condutividade hidráulica

Os severos déficits hídricos nas videiras, provocam respostas mais drásticas, tais como a abscisão de folhas e, em casos graves a embolia (Dayer *et al.* 2019). Ocasionalmente por respostas físicas tais como a condutividade hidráulica, que para Schachtman and Goodger (2008), demonstraram ter mais significado durante condições severas de déficits hídricos, enquanto os sinais químicos desempenham uma maior importância em estágios iniciais do estresse.

Nas plantas o sistema hidráulico encontra-se localizado na interface entre a água disponível no solo e a atmosfera (Charrier *et al.* 2016). Apesar da sinalização hidráulica ser um mecanismo bastante discutido, segundo Zhang *et al.* (2016), ainda é pouco estudado nas videiras. No entanto, sabe-se que reduções na condutividade hidráulica das raízes levam a reduções no turgor e potencial d'água das folhas. Conforme foi observado por Serra *et al.* (2014), que os porta-enxertos mais vigorosos, conferem um maior crescimento da copa e conseqüente demandam mais água, parecem estar balanceados por ajustes na condutividade hidráulica de suas raízes, por meio de suas raízes finas e maior área de alcance.

O aumento da resistência ao fluxo d'água provoca alterações na resistência em vários órgãos das videiras tais como nas raízes, caule, ramos, pecíolos e folhas, controlados pelo fechamento estomático e perda de condutividade hidráulica do xilema (Charrier *et al.* 2016). Devido ao grande calibre dos vasos do xilema, quando comparado com o seu caule e outras espécies, o que acaba favorecendo a cavitação sob potenciais inferiores a -1MPa, gerando assim o colapso da tensão da coluna de

água e o preenchimento do vaso ou traqueídeos com ar (Taiz and Zieger 2004, Serra *et al.* 2014, Kathpalia and Bhatla 2018).

Os efeitos da cavitação são mais evidentes em órgãos distais produzidos anualmente como as folhas, pecíolos e raízes finas, que são mais vulneráveis à embolia, por meio deste mecanismo, permite as videiras se protegerem o máximo possível de embolias na proximidade de seus órgãos vitais, como hastes e o caule considerados perenes, impedindo assim que sofram perdas mais severas (Serra *et al.* 2014, Charrier *et al.* 2016, Tramontini and Lovisolo 2016, Dayer *et al.* 2019). De acordo com Charrier *et al.* (2016), a progressão da mortalidade foliar reflete o aumento da embolia foliar e peciolar, conforme pode ser observado na (Figura 19), a perda percentual de condutividade hidráulica nos órgãos das videiras, a medida que aumenta a tensão ocasionada pelo estresse hídrico.

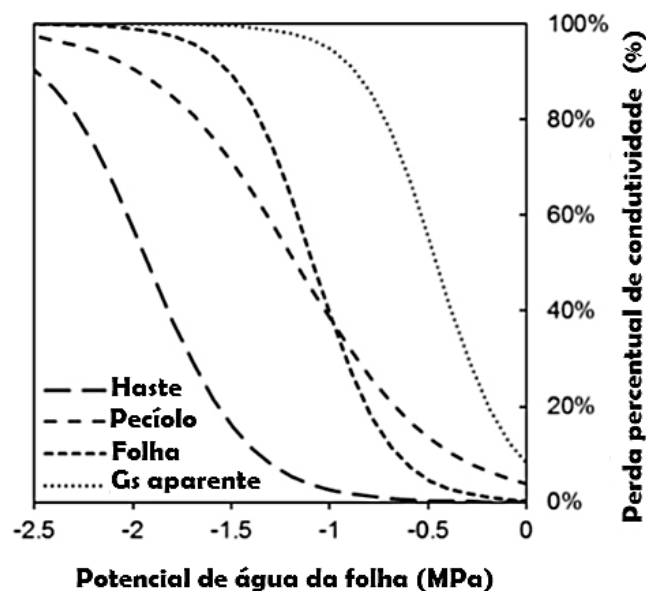


Figura 19 - Perda percentual de condutividade hidráulica (%) na *Vitis vinifera* em suas hastes, pecíolos, folhas e a condutância estomática aparente (Gs) em função do potencial da água nas folhas (MPa). Fonte: Charrier *et al.* (2016).

Para Serra *et al.* (2014), a maior condutividade hidráulica está associada aos porta-enxertos mais tolerantes à seca, devido a menor ocorrência de embolização dos seus vasos. Desta forma, algumas cultivares de porta-enxertos são menos propensas a cavitarem do que outras, devido à crescente elasticidade de suas paredes celulares sob condições de estresse hídrico, assim como, por sua sensibilidade diferencial ao ABA. Também observa-se um aumento da suberização radicular sob estresse hídrico, reduzindo assim a condutância hidráulica total do sistema radicular (Tramontini and Lovisolo 2016, Dayer *et al.* 2019).

2.8 A importância do melhoramento genético de porta-enxertos

O melhoramento genético da videira de forma geral busca atender ao aumento da produtividade e da qualidade da uva, assim como, também persegue objetivos específicos, tais como resistência aos principais patógenos que afetam a cultura e a adaptação as condições ambientais diversas (Camargo *et al.* 2009).

No contexto das mudanças climáticas, se busca cada vez mais obter porta-enxertos tolerantes à seca com a capacidade de proteger a videira contra danos ocasionados pelo estresse hídrico à longo prazo mantendo a produtividade, e simultaneamente as características agronômicas desejáveis para atender às demandas dos produtores e enfrentar de forma sustentável as mudanças climáticas na viticultura (Tsegay *et al.* 2014, Zhang *et al.* 2016).

Embora seja de grande importância para a vitivinicultura, os programas de melhoramento de porta-enxertos são relativamente escassos, porém com as mudanças climáticas tem-se renovado o interesse no melhoramento genético de porta-enxertos para se adaptar aos problemas atuais e futuros (Bruckner and Albuquerque 2005, Riaz *et al.* 2019). Pois a grande maioria das cultivares de porta-enxertos disponíveis atualmente, segundo Riaz *et al.* (2019), foram obtidas há mais de 100 anos, e apresentam uma estreita base genética, onde apenas três acessos de três espécies de videira contribuíram com 39% do conteúdo genético presente em todos os porta-enxertos avaliados em seu estudo. Além disso, pode levar de 7 a 25 anos para se obter novas cultivares de porta-enxertos melhorados (Bruckner and Albuquerque 2005, Zhang *et al.* 2016), e isso pode se tornar uma limitação para prevenir estresses bióticos e abióticos em ambientes que estão sujeitos a mudanças contínuas (Ollat *et al.* 2016, Riaz *et al.* 2019).

Além disso, o desenvolvimento de porta-enxertos tolerantes ao déficit hídrico encontra outras dificuldades pois esta característica tem caráter poligênico, apresentando forte interação entre a expressão deles e os níveis de disponibilidade hídrica, o que dificulta a identificação e seleção de materiais superiores (Ashraf 2010, Marguerit *et al.* 2012, Serra *et al.* 2014, Dayer *et al.* 2019). Segundo Silva *et al.* (2011) e Zhang *et al.* (2016), o conhecimento dos mecanismos genéticos envolvidos na resposta à deficiência hídrica pelos quais os porta-enxertos influenciam no comportamento de suas copas, é importante, pois através da manipulação dessas informações obtidas é possível aprimorar a seleção de materiais tolerantes.

O uso adequado das informações em programas de melhoramento, podem resultar em ferramentas importantes na identificação de materiais genéticos promissores (Cavatte *et al.* 2011). Entre estas ferramentas pode-se citar os algoritmos de aprendizado de máquina, os quais segundo Shalev-Shwartz and Ben-David (2014), podem auxiliar na descoberta de padrões significativos, e que podem passar despercebidos pelo melhorista, por serem capazes de gerenciar grandes quantidades de dados, que podem levar a resultados mais precisos e melhor interpretação do que os métodos estatísticos clássicos (Niazian and Niedbala 2020).

2.9 Aprendizado de máquina

Aprendizagem de máquina, ou no inglês, *Machine learning* (ML), refere-se a um subcampo da Ciência da Computação, que para Shalev-Shwartz and Ben-David (2014) pode ser visto como um ramo da Inteligência Artificial (IA). Dentro do aprendizado de máquina, concentra-se o subcampo do *deep learning* também conhecido como aprendizado profundo, que engloba o campo das redes neurais (Ertel 2017) conforme pode ser visualizado na Figura 20.

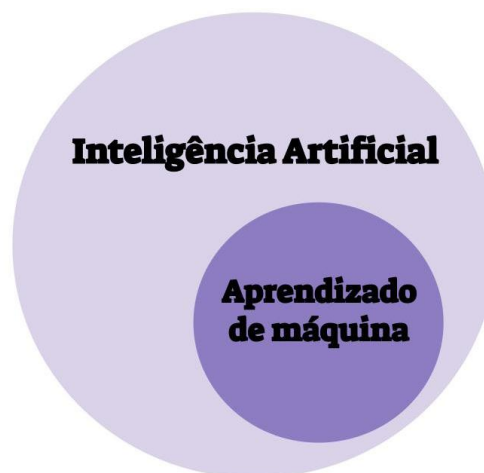


Figura 20 - Aprendizado de máquina como subcampo da Inteligência Artificial. Fonte: Verslype (2021).

O ML pode ser considerado como um ramo da IA, pois tem o objetivo de programar máquinas para que aprendam e se adaptem a mudanças, através da transformação da experiência em especialização ou detecção de padrões significativos em dados sensoriais complexos, sem precisar fornecer soluções para todas as situações possíveis, por ter esta capacidade de transformar o fundamento da inteligência humana (Alpaydın 2010, Shalev-Shwartz and Ben-David 2014).

Nesse sentido, o ML pode ser visto como um campo interdisciplinar, que compartilha linhas comuns com os campos matemáticos da estatística, teoria da informação, teoria dos jogos e otimização (Shalev-Shwartz and Ben-David 2014). O aprendizado de máquina usa a teoria da estatística na construção de modelos matemáticos, devido a sua tarefa principal seja fazer inferências a partir de uma amostra de dados (Alpaydin 2010).

Apesar de sua semelhança com a estatística tradicional, eles diferem em alguns quesitos, entre estes pode-se dar ênfase a questão do ML trabalhar sob uma configuração "livre de distribuição normal", assumindo assim o mínimo possível sobre a natureza da distribuição dos dados, e permitindo que o algoritmo de aprendizagem descubra quais modelos que melhor se aproximam do processo de geração de dados. Outra diferença do ML é concentrar-se em limites de amostra finitos, ou seja, dado o tamanho das amostras disponíveis, a teoria do ML visa descobrir o grau de precisão que se pode esperar com base em tais amostras, diferentemente da estatística que visa expandir para situações infinitas (Shalev-Shwartz e Ben-David 2014).

De acordo com Ertel (2017) o ML é considerado por muitos como um dos ramos mais importantes da IA. No entanto, deve-se notar que, em contraste com a IA tradicional, o ML não tenta construir uma imitação automatizada de comportamento inteligente, mas sim usar os pontos fortes e habilidades especiais dos computadores para complementar a inteligência humana, muitas vezes realizando tarefas que vão muito além das capacidades do homem (Shalev-Shwartz and Ben-David 2014). Dessa forma, o ML torna-se interessante para ser aplicado em grandes bancos de dados, em um processo conhecido como mineração de dados (Alpaydin 2010), permitindo detectar padrões que estão fora do escopo da percepção humana ao escanear e processar esses bancos de dados (Shalev-Shwartz and Ben-David 2014).

Dessa forma, o ML pode ser definido como um método computacional que se baseia na experiência para melhorar o seu desempenho e/ou fazer previsões acuradas (Mohri *et al.* 2018), ou seja, usa a experiência para se tornar especialista (Shalev-Shwartz and Ben-David 2014). Com isso o modelo pode ser preditivo para fazer previsões no futuro ou descritivo para obter conhecimento através dos dados, ou ambos (Alpaydin 2010). Por conta do ML ser uma área de domínio muito amplo, segundo Shalev-Shwartz and Ben-David (2014), foi ramificada em vários subcampos que lidam com diferentes tipos de tarefas de aprendizado, podendo ser classificada de forma simplificada em aprendizado supervisionado, aprendizado não-

supervisionado e aprendizado por reforço, conforme pode ser visualizado na Figura 21.



Figura 21 - Classificação do Aprendizado de máquina. Fonte: Verslype (2021).

Apesar da existência dos três subcampos do aprendizado de máquina, neste trabalho teve-se como objetivo avaliar a divergência genética em 45 cultivares de porta-enxertos de videira através do aprendizado não-supervisionado, e de encontrar as classes de tolerância ao déficit hídrico dos porta-enxertos IAC 313, IAC 572 e IAC 766, cuja a informação na literatura é desconhecida, através da capacidade de predição dos algoritmos de aprendizado supervisionado. Nesse sentido, neste trabalho será dado enfoque apenas ao aprendizado supervisionado e aprendizado não-supervisionado, por englobarem todos os algoritmos que serão usados para desenvolver esta pesquisa, conforme estão descritas a seguir:

2.9.1 Aprendizado supervisionado

No aprendizado supervisionado, segundo Alpaydın (2010), o objetivo é que haja o aprendizado através do mapeamento da entrada para uma saída, cujos valores corretos são fornecidos por um supervisor. Dessa forma, a experiência adquirida é através de um treinamento que contém informações significativas rotuladas e o conjunto teste que falta a estas informações rotuladas, e com base na experiência adquirida no conjunto de treinamento pode-se prever a falta de informações para o conjunto de dados testado (Shalev-Shwartz and Ben-David 2014). Dentre os algoritmos supervisionados citam-se o *decision tree* e o *random forest*, ambos

empregados neste estudo para predizer o grau de tolerância ao déficit hídrico dos porta-enxertos de videira.

2.9.1.1 Árvore de decisão

A árvore de decisão ou *decision tree* é um algoritmo extraordinariamente importante para inteligência artificial por ser simples, poderoso e eficiente na extração de informações do conjunto de dados (Ertel 2017). Podendo ser empregado para determinar classes, mas também em outros problemas de aprendizado, como regressão e agrupamento (Mohri *et al.* 2018), sendo as informações extraídas a partir deste algoritmo facilmente interpretadas, compreendidas e controladas pelo homem (Ertel 2017). Além disso, também são normalmente rápidas de treinar e avaliar (Mohri *et al.* 2018), principalmente quando comparado com outros algoritmos e, por conta destas características, torna-se uma ferramenta importante para mineração de dados (Ertel 2017) conforme podemos visualizar a sua representação (Figura 22).

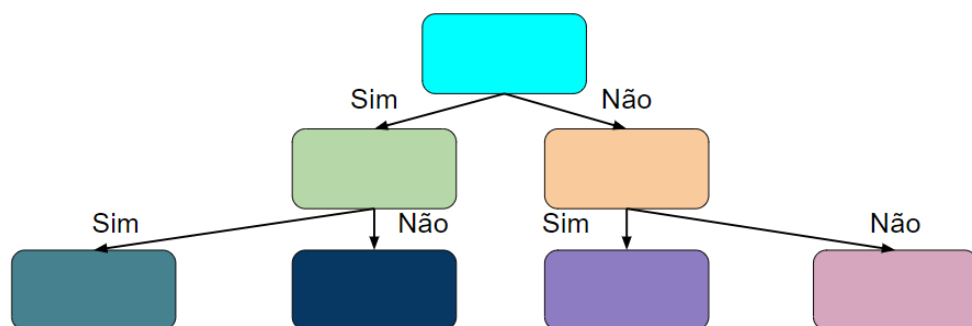


Figura 22 - Representação da árvore de decisão. Fonte: Verslype (2021).

A *decision tree* pode ser usada na classificação de várias classes, mas também em outros problemas de aprendizado, como regressão e agrupamento (Mohri *et al.* 2018). Conforme pode ser observado na (Figura 22), cada folha da árvore é representada pelos quadrados, que contém um rótulo específico associado a uma instância x , ao viajar de um nó da raiz para uma folha da árvore de decisão, sendo o sucessor escolhido com base em uma divisão do espaço de entrada (Shalev-Shwartz and Ben-David 2014).

Algoritmos de aprendizado poderosos, como por exemplo o aprendizado da árvore de decisão, podem adaptar a complexidade do modelo aprendido à complexidade dos dados de treinamento. Isso leva a um ajuste excessivo se houver ruído nos dados (Ertel 2017). Apesar de não serem considerados de última geração por Mohri *et al.* (2018), também é possível associá-los com outros algoritmos de aprendizagem. Para

Ertel (2017), muitos algoritmos de aprendizagem têm o problema de ajuste excessivo, no entanto podem ser corrigidos com o uso da validação cruzada que é o método de otimização automática mais importante e mais usado para a complexidade do modelo. Dessa forma, o algoritmo de *decision tree* pode ser empregado como ferramenta de estudo na área de melhoramento de plantas Zhao *et al.* (2016), demonstraram a eficiência do algoritmo, pois ao avaliarem as variedades de soja, os autores encontraram resultados promissores, e inclusive indicam a sua relevância e eficácia para avaliações no melhoramento genético de plantas.

2.9.1.2 Floresta Aleatória

Random forest traduzida em português como floresta aleatória é uma abordagem não paramétrica, criada por Breiman (2001), com o intuito de resolver problemas de regressão e classificação. Trata-se de um classificador estável e forte por construir um grande número de árvores, sendo desta forma, mais eficiente do que uma árvore avaliada de forma individual (Dean *et al.* 2014). Por conta deste embasamento, em uma coleção de árvores de decisão, assim como, da aplicação de um algoritmo no conjunto de treinamento para construir cada árvore (Shalev-Shwartz and Ben-David 2014). Breiman (2001) destaca que o *random forest* é considerado uma ferramenta eficaz de previsão, que apresenta melhorias significativas na precisão da sua classificação em relação ao *decision tree*.

Esse processo de construção de uma coleção de árvores de decisão, segundo Parmley *et al.* (2019), ocorre através do processo de ensacamento, onde um subconjunto dos dados é usado para o treinamento de cada árvore de decisão e um nível adicional de aleatoriedade é incluído selecionando aleatoriamente, em cada nó, um subconjunto de preditores. Dessa forma, uma das ideias fundamentais na floresta aleatória, para Dean *et al.* (2014), é que um subconjunto das observações e um subconjunto das variáveis sejam levados em conta, desta forma, essa amostragem em ambas as dimensões garante que todas as variáveis sejam consideradas durante a análise, e não apenas as variáveis dominantes, como geralmente é observado nas árvores de decisão, possibilitando assim, que haja uma votação na classe mais popular (Breiman 2001). A representação esquemática do algoritmo *random forest* de aprendizado não-supervisionado pode ser observada na (Figura 23).

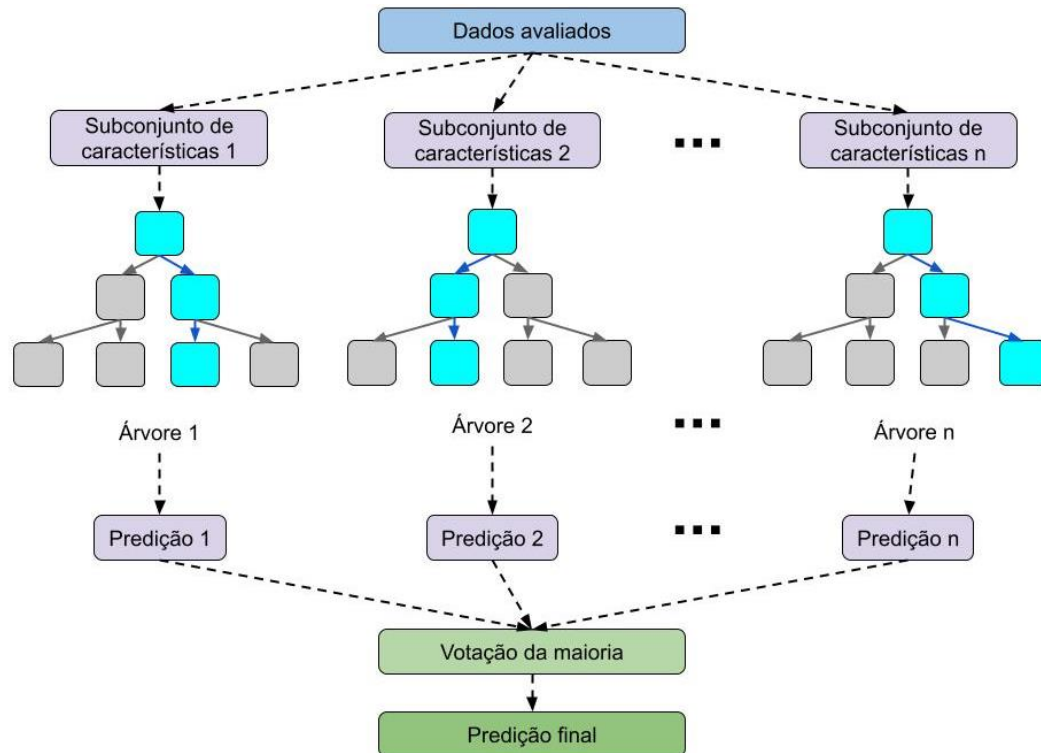


Figura 23 - Representação do funcionamento do algoritmo *Random Forest*. Fonte: Verslype (2021).

Na literatura, é possível observar a utilização do *random forest* em diversas áreas de estudo, tais como no melhoramento genético de plantas. Rad and Rafezi (2020), utilizaram o algoritmo *random forest* para, realizar a descrição morfológica e a análise dos componentes genéticos de características e produtividade dos frutos de 10 variedades de meloeiro, sendo possível indicar a importância do diâmetro da cavidade do melão como variável adequada para separar os genótipos. Por sua vez, Parmley *et al.* (2019), usaram o *random forest* para treinar modelos de previsão de crescimento e rendimento de grãos de soja, possibilitando orientar programas de melhoramento de plantas para o desenvolvimento de cultivares.

2.9.2 Aprendizado não supervisionado

No aprendizado não-supervisionado, segundo Alpaydın (2010), não existe supervisor e temos apenas dados de entrada, não rotulados sobre o qual o algoritmo realiza previsões (Mohri *et al.* 2018). Desta forma, não se tem distinção entre o conjunto de dados de treinamento e de teste (Shalev-Shwartz and Ben-David 2014). Logo, os dados de entrada são processados com o intuito de obter uma versão resumida ou compacta, o que possibilita compreender e visualizar padrões e tendências mais frequentes do que outras dentro de um conjunto de dados, como um exemplo típico pode-se citar a análise de agrupamentos (Alpaydın 2010, Shalev-

Shwartz and Ben-David 2014). Entre estes algoritmos de aprendizado não-supervisionado pode-se citar o *K-means*, que foi empregado neste estudo para determinar a divergência genética entre as 45 cultivares de porta-enxertos de videira.

2.9.2.1 K-means

O algoritmo de agrupamento *K-means* (MacQueen 1967), é um dos mais populares em aplicações práticas de agrupamento (Shalev-Shwartz and Ben-David 2014). Na área do melhoramento genético de plantas, este algoritmo vem sendo cada vez mais empregado, principalmente em trabalhos de estudo de diversidade genética em acessos de Bancos Ativos de Germoplasma (BAG), tais como os da autoria de Nandini *et al.* (2020), que classificaram de forma efetiva 494 acessos de milho ou painço japonês (*Echinochloa frumentaceae* (Roxb.) Link) em 23 grupos; assim como na pesquisa da autoria de Kanavi *et al.* (2020), estudaram 205 acessos de feijão moyashi (*Vigna radiata* L.) sob condição de estresse hídrico e conseguiram agrupar todos os acessos em sete clusters diferentes, com o intuito de identificar genótipos tolerantes à seca, geneticamente divergentes, para serem usados posteriormente como linhagens parentais em programas de melhoramento genético.

Segundo Kanavi *et al.* (2020), o *K-means* é uma técnica muito poderosa para identificar linhagens parentais promissoras, por possibilitar formar grupos heteróticos geneticamente divergentes (clusters) com base nas distâncias genéticas entre os acessos do germoplasma. De acordo com Ertel (2017), na nomenclatura do algoritmo de agrupamento *K-means*, o 'K' refere-se aos clusters, enquanto o 'means' ao ponto médio, representado por um centróide (C_i) que é o ponto em vermelho destacado na (Figura 24), sendo calculado através de seu valor médio do *cluster*, ou seja, a função tem como objetivo minimizar a soma da distância quadrática do *cluster* em relação ao centro entre cada ponto em X ao centróide de seu cluster (Shalev-Shwartz and Ben-David 2014).

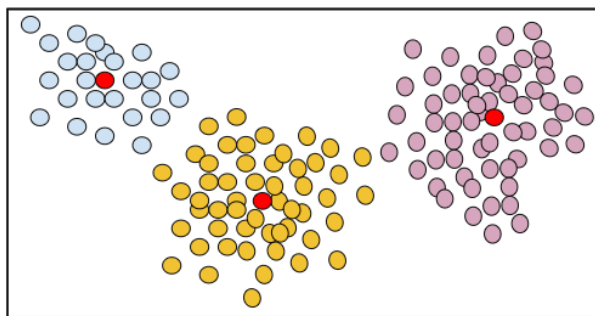


Figura 24 - Exemplo de clustering K-means evidenciando os centróides. Fonte: Verslype (2021).

2.10 Aplicações do Aprendizado de máquina no melhoramento genético de plantas

Diante disso, o aprendizado de máquina mostra-se como uma ferramenta útil em diversas áreas de estudo, devido a sua capacidade de fazer previsões acuradas, gerenciar grandes bancos de dados, identificar padrões e tendências que muitas vezes passariam imperceptíveis pelo homem, permitindo assim visualizar e compreender facilmente estas tendências dentro de um conjunto de dados (Alpaydin 2010, Shalev-Shwartz and Ben-David 2014, Mohri *et al.* 2018).

Esta capacidade do aprendizado de máquina pode trazer diversos benefícios na área do melhoramento genético de plantas, se for empregada corretamente, entre estes benefícios, pode-se destacar a possibilidade de gerenciar grandes números de acessos em BAG, tais como, realizar análises de agrupamento para estudos de diversidade genética, identificar linhagens parentais promissoras e os melhores cruzamentos, conforme a necessidade do programa de melhoramento genético de plantas, através do aprendizado não-supervisionado (Beiki 2019, Etminan *et al.* 2019, Kanavi *et al.* 2020, Nandini *et al.* 2020).

Além disso, pode-se dar destaque a capacidade dos algoritmos de aprendizado supervisionado em realizar previsões acuradas de características qualitativas ou quantitativas nas plantas, como por exemplo, em previsões de crescimento, produtividade, tolerância a estresses abióticos ou bióticos, no desempenho de futuras cultivares e inclusive em estudos de fenotipagem, com bancos de imagens (Zhao *et al.* 2016, Etminan *et al.* 2019, Parmley *et al.* 2019, Niazian and Niedbala 2020, Rad and Rafezi 2020). Neste contexto, algoritmos de aprendizado de máquina podem trazer avanços e auxiliar os melhoristas em suas pesquisas, uma vez que se trata de uma técnica muito poderosa e útil em problemas de classificação, previsão e identificação, tornando assim, o processo de seleção mais preciso por trazer informações precisas de rápida obtenção e fácil interpretação (Etminan *et al.* 2019, Kanavi *et al.* 2020, Nandini *et al.* 2020, Niazian and Niedbala 2020).

1. REFERÊNCIAS BIBLIOGRÁFICAS

Albuquerque TCS (2003) Videira (*Vitis* sp). In Castro PRC and Kluge RA (Coord.). **Ecofisiologia de fruteiras: abacateiro, aceroleira, macieira, pereira e videira.** Agronômica Ceres, Piracicaba, p. 93-119.

Alpaydin E (2020) **Introduction to machine learning.** MIT press, Cambridge, Massachusetts, 682p.

Alsina MM, Smart DR, Bauerle T, Herralde F, Biel C, Stockert C, Negrón C and Save R (2011) Seasonal changes of whole root system conductance by a drought-tolerant grape root system. **Journal of Experimental Botany** **62**: 99-109.

Ashraf M (2010) Inducing drought tolerance in plants: recent advances. **Biotechnology advance** **28**: 169-183.

ATVB – Association Technique Viticole de Bourgogne (2013) **Caractéristiques des porte-greffes et greffons utilisés en côte-d’or.** Disponível em <[http://www.atvb-bourgogne.com/sites/default/files/catalogue%20ATV Bsept2015%20\(2\).pdf](http://www.atvb-bourgogne.com/sites/default/files/catalogue%20ATV%20Bsept2015%20(2).pdf)>. Acesso em agosto 2020.

Audeguin L, Bécart V, Besnard E, Boursiquot JM, Ciccolini G, Dedet S, Delorme G, Follet J, Froelhy A, Garcin L, Gonin N, Lacombe T, Laffargue F, Marchal C, Pesato M, Sereno C, Uriel G, Wentzel M and Yobregat O (2020) **PI@ntGrape: le catalogue des vignes cultivées en France.** Montpellier, France: IFV – INRA – Montpellier SupAgro, 2009-2020. Disponível em <<http://plantgrape.plantnet-project.org/en/porte-greffe/>>. Acesso em setembro 2020.

Basso LH, Assis JD, Lima Filho JMP, Ribeiro HA, Silva MR and Miranda AA (1999) **Interrupções da irrigação no período de maturação da uva cv Itália.** Embrapa-CPATSA, Petrolina, 5p.

Beiki AH (2019) Molecular Predicting Drought Tolerance in Maize Inbred Lines by Machine Learning Approaches. **BioRxiv** **10**: 1-33.

Bettiga LJ, Christensen LP, Dokoozilian NK, Golino DA, McGourty G, Smith RJ, Verdegaal PS, Walker MA, Wolpert JA and Weber E (2003) **Wine grape varieties in California.** UCANR Publications, Califórnia, 188p.

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

Brancadoro L, Failla O, Gardiman M, Rossoni M, Scienza A, and Zamboni M (2014) I nuovi portainnesti della serie M a confronto con altri in commercio. In Scienza A, Failla O and Espen L. (eds) **Innovative strategies for the selection of new rootstocks of vines**, L'Informatore Agrario, Verona, p. 25-58.

Breiman L (2001) Random forests. **Machine learning** 45: 5-32.

Bruckner CH and Albuquerque AS (2005) Melhoramento de fruteiras. In: Borém A (ed.). **Melhoramento de espécies cultivadas**. Viçosa: Ed. UFV, p. 813-863.

Camargo UA, Bernd RB and Revers LF (2009) Melhoramento genético. In Soares JM and Leão PCS (eds) **A vitivinicultura no Semiárido Brasileiro**. Embrapa Informação Tecnológica, Brasília, Embrapa Semi-Árido, Petrolina, p.151-214.

Camargo UA (2003) Porta-enxerto e cultivares. Embrapa Uva e Vinho, Bento Gonçalves, **Sistema de produção 2**. Disponível em <<https://sistemasdeproducao.cnptia.embrapa.br/FontesHTML/Uva/UvaAmericanaHibridaClimaTemperado/cultivar.htm>>. Acesso em maio 2020.

Camargo UA, Tonietto J and Hoffmann (2011) Progressos na viticultura brasileira. **Rev. Bras. Frutic.** 33: 144-149.

Camargo UA, Ritschel PS, Maia JDG, Quecini V, Machado CAE, Bosco DD, Sinski I, Comachio V and Zilio R (2017) **Banco ativo de germoplasma de uva**. Embrapa Uva e Vinho, Bento Gonçalves. Disponível em <<https://www.embrapa.br/en/uva-e-vinho/banco-ativo-de-germoplasma-de-uva/pesquisa>>. Acesso em junho 2020.

Carbonneau A (1985) The early selection of grapevine rootstocks for resistance to drought conditions. **Am. J. Enol. Vitic.** 36: 195-198.

Carroll B (2020) **Characteristics of Vitis grape rootstocks and adaptability**. Oklahoma State University, Oklahoma, 4p.

Cassel DK and Nielsen DR (1986) Field capacity and available water capacity. In Klute A (eds) **Methods of soil analysis**. Physical and mineralogical methods. 2nd ed., SSSA, Madison, p. 901-926.

Cavatte PC, Martins SCV, Morais LE, Silva PEM, Souza LT and Damatta FM (2011) A fisiologia dos estresses abióticos. In Fritsche-Neto R and Borém A (eds)

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

Melhoramento de plantas para condições de estresses abióticos. Suprema, Visconde do Rio Branco, p. 39-47.

Chaves MM *et al.* (2010) Grapevine under deficit irrigation: hints from physiological and molecular data. **Annals Appl. Biol.** 1: 661-676.

Chaves MM, Maroco JP and Pereira JS (2003) Understanding plant responses to drought from genes to the whole plant. **Functional Plant Biology** 30: 239-264.

Charrier G, Torres-Ruiz JM, Badel E, Burlett R, Choat B, Cochard H, Delmas CEL, Domec JC, Jansen S, King A, Lenoir N, Martin-StPaul N, Gambetta GA, and Delzon S (2016) Evidence for hydraulic vulnerability segmentation and lack of xylem refilling under tension. **Plant Physiol** 172: 1657-1668.

Chevalier A (1925) L'amélioration de la vigne en France et les Travaux de G. Couderc sur l'Hybridation et le Greffage. **Revue de botanique appliquée et d'agriculture coloniale** 52: 926-945.

Clarke AD (2004) **Suggested rootstocks for New Zealand vineyards.** Blenheim, Grovetown, 9p.

CODEVASF – Companhia de Desenvolvimento do Vale do São Francisco (1996) **Projeto semi-árido:** proposta de desenvolvimento sustentável da bacia do rio São Francisco e do semi-árido nordestino. CODEVASF, Brasília, 60p.

Colman EA (1947) A laboratory procedure for determining the field capacity of soils. **Soil Science** 63: 277-283.

Dane JH and Hopmans JW (2002) Water retention and storage. In Dane JH and Topp GC (eds) **Methods of soil analysis:** physical methods. American Society of America, Madison, p. 671-720.

Dayer S, Reingwartz I, McElrone AJ and Gambetta GA (2019) Response and Recovery of Grapevine to Water Deficit: From Genes to Physiology. In Cantu D and Walker M (eds) **The Grape Genome.** Compendium of Plant Genomes. Springer, Switzerland, p. 223-245.

Dean J. (2014) **Big data, data mining, and machine learning:** value creation for business leaders and practitioners. John Wiley & Sons, Hoboken, 288p.

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

Dougherty PH (2012) Introduction to the geographical study of viticulture and wine production. In Dougherty, Percy H (eds) **The Geography of Wine: Regions, Terroir and Technique**. Springer, Netherlands, p. 3-36

Embrapa Uva e Vinho (2016) **Cultivares de porta-enxertos**. Disponível em <<https://www.embrapa.br/en/uva-e-vinho/cultivares-e-porta-enxertos/porta-enxertos> > . Acesso em abril 2020.

Ertel W (2017) **Introduction to artificial intelligence**. Springer International Publishing AG, Switzerland, 365p.

Etminan A, Pour-Aboughadareh A, Mohammadi R, Shooshtari L, Yousefiazarkhanian M, and Moradkhani H (2019) Determining the best drought tolerance indices using artificial neural network (ANN): Insight into application of intelligent agriculture in agronomy and plant breeding. **Cereal Research Communications**, **47**: 170-181.

FAO - Food and Agriculture Organization of the United Nations (2020) Crop information: grape. Disponível em <<http://www.fao.org/land-water/databases-and-software/crop-information/grape/en/>> . Acesso em março 2020.

FAOSAT - Food and Agriculture Organization Corporate Statistical Database (2020) **Crops statistics**: data about production quantities of grapes in 2019. Disponível em <<http://www.Fao.org/faostat/en/#data/QC/visualize>>. Acesso em dezembro 2020.

Fernández-Cano LH and Togores JH (2011) **Tratado de viticultura**. Ediciones Mundi prensa, Madrid, 1031p.

Fort K, Fraga H, Grossi D, Walker MA (2017) Early measures of drought tolerance in four grape rootstocks. **J. Amer. Soc. Hort. Sci.** **1**: 36-46.

Fraga H, Malheiro AC, Moutinho-Pereira J and Santos JA (2013) An overview of climate change impacts on European viticulture. **Food Energy Security** **1**: 94-110.

Franks PJ and Beerling DJ (2009) Maximum leaf conductance driven by CO₂ effects on stomatal size and density over geologic time. **Proceedings of the National Academy of Sciences** **106**: 10343-10347.

Fregoni M, Scienza A and Miravalle R (1978) Evaluation précoce de la résistance des porte-greffes à la secheresse. In: POUGET R and DOAZAN JP (eds.) **Proceedings**

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

of the II^e Symposium International sur l'Amélioration de la Vigne Bordeaux, 14-18 Juin 1977. INRA, Paris, p. 287-96.

Garris A, Cousins P, Ramming D and Bald A (2009) Parentage Analysis of Freedom Rootstock. **Am. J. Enol. Vitic. 60:** 357-361.

Giovannini E (2014) **Manual de viticultura.** Bookman, Porto Alegre, 264p.

Goldammer T (2018) Wine Grape Rootstocks In: Goldammer T (ed) **Grape Grower's Handbook:** a guide to viticulture for wine production. Apex Publishers, Virginia, p. 25-34.

Hall A and Jones GV (2009) Effect of potential atmospheric warming on temperature based indices describing Australian wine grape growing conditions. **Aust.J. Grape Wine Res. 15:** 97–119.

Hernandes JL, Martins FP and Pedro Júnior MJ (2011) **Uso de porta-enxertos: Tecnologia simples e fundamental na cultura da videira.** Jundiaí, SP: Instituto Agrônômico de Campinas, 19p.

IBGE - Instituto Brasileiro de Geografia e Estatística (2020) **Levantamento Sistemático da Produção Agrícola.** Disponível em <<https://sidra.ibge.gov.br/tabela/1618#resultado>>. Acesso em janeiro 2020.

IFV - Institut Français de la Vigne et du Vin (2007) **Catalogue des variétés et clones de vigne cultivés en France.** Disponível em <<https://www.vignevin-occitanie.com/fiches-pratiques/cepages-et-porte-greffes-utilises-dans-le-sud-ouest/>>. Acesso em junho 2020.

Jackson DI and Lombard PB (1993) Environmental and management practices affecting grape composition and wine quality-a review. **American Journal of Enology and Viticulture 44:** 409-430.

Jones GV, Reild R and Vilks A (2012) Climate, grapes, and wine: structure and suitability in a variable and changing climate In Dougherty, Percy H (eds) **The Geography of Wine:** regions, terroir and technique. Springer, Netherlands, p. 109-36

Jones GV (2016) Grapevines in a changing environment: a global perspective. In Gerós H, Chaves MM, Gil HM and Delrot S (eds) **Grapevine in a changing**

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

environment: a molecular and ecophysiological perspective. John Wiley & Sons, UK, p. 23-34.

Jones GV and Webb LB (2010) Climate change, viticulture, and wine: challenges and opportunities. **Journal of Wine Research** **21**: 103-106.

Jones GV, White MA, Cooper OR and Storchmann K (2005) Climate change and global wine quality. **Clim. Change** **73**: 319-343.

Kanavi MSP, Koler P, Somu G and Marappa N (2020) Genetic diversity study through k-means clustering in germplasm accessions of green gram [*Vigna radiata* (L.)] under drought condition. **International Journal of Bio-resource and Stress Management** **18**: 138-147.

Kathpalia R and Bhatla SC (2018) Plant Water Relations. In Bhatla SC and Lal MA (eds) **Plant physiology, development and metabolism**. Springer, Singapore, p. 3-36p.

Keller M (2015) **The science of grapevines: anatomy and physiology**. 2nd ed., Academic press, San Diego, 509p.

Kist BB, Carvalho C, Treichel M and Santos CE (2018) **Anuário brasileiro de fruticultura 2018**. Editora Gazeta Santa Cruz, Santa Cruz do Sul, 88p.

Kist BB, Santos CE, Carvalho C and Beling RR (2019) **Anuário Brasileiro de Horti e Fruti 2019**. Editora Gazeta Santa Cruz, Santa Cruz do Sul, 96p.

Leão PCDS, Nascimento JHBD, Moraes DSD and Souza ERD (2020a). Rootstocks for the new seedless table grape 'BRS Vitória' under tropical semi-arid conditions of São Francisco Valley. **Ciência e Agrotecnologia**, **44**: 2-8.

Leão PCDS, Nascimento JHBD, Moraes DSD and Souza ERD (2020b). Yield components of the new seedless table grape 'BRS Ísis' as affected by the rootstock under semi-arid tropical conditions. **Scientia Horticulturae**, **263**: 109-114.

Leão PCS and Silva DJ (2014) Cultivo da videira no Semiárido brasileiro. In Pio R (ed) **Cultivo de fruteiras de clima temperado em regiões subtropicais e tropicais**. UFLA, Lavras, p. 578-618.

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

Leão PCS (2010) **Cultivo da videira**. 2nd ed. Embrapa Semi-Árido, Petrolina, Sistemas de Produção, Disponível em <http://www.cpatsa.embrapa.br:8080/sistema_producao/spuva/cultivares.html> . Acesso em maio 2020.

Leão PCS and Borges RME (2009) **Melhoramento genético da videira**. Embrapa Semiárido, Petrolina, 61p.

Leão PCS, Soares JM and Rodrigues BL (2009) Principais Cultivares. In Soares JM and Leão PCS (eds) **A vitivinicultura no Semiárido Brasileiro**. Embrapa Informação Tecnológica, Brasília, Embrapa Semi-Árido, Petrolina, p.151-214.

Leão PCS, Rego JIS, Nascimento JHB and Souza EMC (2016) Produção e características físico-químicas dos frutos de novas cultivares de uvas de mesa sobre diferentes porta-enxertos no Vale do Submédio São Francisco. In: **Proceedings of the XXIV Congresso brasileiro de fruticultura**, 24. SBF, São Luis, 1-7p.

Lima Filho JMP, Dantas BF, Assis JS, Souza CR and Albuquerque TCS (2009) Aspectos fisiológicos. In Soares JM and Leão PCS (eds) **A vitivinicultura no Semiárido Brasileiro**. Embrapa Informação Tecnológica, Brasília, Embrapa Semi-Árido, Petrolina, p. 73-108.

MacQueen J (1967) Some methods for classification and analysis of multivariate observations. In: **Proceedings of the fifth Berkeley symposium on mathematical statistics and probability**, 14.281-297p.

Magalhães Filho JR, Amaral LR, Machado DFSP, Medina CL and Machado EC (2008) Deficiência hídrica, trocas gasosas e crescimento de raízes em laranjeira 'Valência' sobre dois tipos de porta-enxerto. **Bragantia** **67**: 75-82.

Malheiro AC, Santos JA, Fraga H and Pinto JG (2010) European viticulture geography in a changing climate. **Climate Research** **43**: 163–177.

Mandelli F, Miele A and Tonietto J (2009) Uva em clima temperado. In Monteiro JEBA (ed) **Agrometeorologia dos cultivos**: o fator meteorológico na produção agrícola. INMET, Brasília, p. 503-515.

Marguerit E, Tandonnet JP, Golard B, Peccoux A, Leeuwen KV and Ollat N (2011) Le porte-greffe, outil original et pertinent d'adaptation au changement climatique. In: **Proceedings of the 10 Journée Technique du Conseil Interprofessionnel du Vin de Bordeaux (CIVB)**. INRAE, Bordeaux, p. 1-4.

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

Marguerit E, Brendel O, Lebon E, Van Leeuwen C and Ollat N (2012) Rootstock control of scion transpiration and its acclimation to water deficit are controlled by different genes. **New Phytologist** **194**: 416-429.

Maul E, Töpfer R, Röckel F, Brühl U, Hundemer M, Mahler-Ries A, Walk M, Kecke S, Wolck A and Ganesch A (2020) **Vitis International Variety Catalogue (VIVC)**. Julius Kühn-Institut - Federal Research Centre for Cultivated Plants (JKI), Institute for Grapevine Breeding - Geilweilerhof (ZR). Disponível em <<http://www.vivc.de/>>. Acesso em abril 2020.

Mohri M., Rostamizadeh A. and Talwalkar A. (2018) **Foundations of machine learning**. 2nd ed., The MIT Press, Cambridge, 505p.

Morton LT (1979) The myth of the universal rootstock. **Wines & vines** **60**: 24-26.

Moura MSB, Teixeira AHC and Soares JM (2009) Exigências climáticas. In Soares JM and Leão PCS (eds) **A vitivinicultura no Semiárido Brasileiro**. Embrapa Informação Tecnológica, Brasília, Embrapa Semi-Árido, Petrolina, p.36-69.

Mullins MG, Bouquet A and Williams LE (1992) **Biology of the Grapevine**. Cambridge University Press, Cambridge, 239p.

Nachtigal JC (2003) **Produção de mudas de videira em regiões tropicais e subtropicais do Brasil**. Bento Gonçalves, Embrapa Uva e Vinho, 22 p.

Nachtigal JC and Mazzarolo A (2008) **Uva: o produtor pergunta, a Embrapa responde**. Embrapa Informação Tecnológica, Brasília, 202p.

Nandini C, Bhat S, Saritha HS, Pandey CD, Sushil Pandey P, Bai L and Gowda J (2020) Characterization of barnyard millet (*Echinochloa frumentaceae* (Roxb.) Link) germplasm for quantitative traits to enhance its utilization. **Electronic Journal of Plant Breeding** **11**: 1066-1072.

Netzer Y, Yao C and Shenker M (2009) Water use and the development of seasonal crop coefficients for Superior Seedless grapevines trained to an open-gable trellis system. **Irrigation Science** **27**: 109-120.

Niazian M and Niedbała G (2020) Machine learning for plant breeding and biotechnology. **Agriculture** **10**: 436-459.

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

OIV - Organización Internacional de la Viña y el Vino (2020). **Estadísticas: informe sobre la situación de la viticultura em 2019**. Disponível em < <http://www.oiv.int/es/statistiques/recherche>>. Acesso em dezembro 2020.

Ojeda H, Deloire A, Wang Z and Carbonneau A (2004) Determinación y control del estado hídrico de la vid: efectos morfológicos y fisiológicos de la restricción hídrica en vides. **Viticultura/Enología Profesional 90**: p. 27-43.

Ollat N, Peccoux A, Papura D, Esmenjaud D, Marguerit E, Tandonnet JP, Bordenave L, Cookson SJ, Barrieu F, Rossdeutsch L, Lecourt J, Lauvergeat V, Vivin P, Bert PF and Delrot S (2016) Rootstocks as a component of adaptation to environment. In Gerós H, Chaves MM, Gil HM and Delrot S (eds) **Grapevine in a changing environment: a molecular and ecophysiological perspective**. John Wiley & Sons, UK, p. 102-152.

Parmley KA, Higgins RH, Ganapathysubramanian B, Sarkar S and Singh AK (2019) Machine learning approach for prescriptive plant breeding. **Scientific reports 9**: 1-12.

Pecile M, Zavaglia C and Ciardi A (2020) **Registro nazionale delle varietà di vite**. MiPAAF, Itália. Disponível em <<http://catalogoviti.politicheagricole.it/catalogo.php>>. Acesso em agosto 2020.

Peterson JCD, Duncan R, Hirschfeld D, Ingels C, McGourty G, Smith R, Weber E, Wolpert J, Anderson M, Benz J and Walker MA (2019) Grape rootstock breeding and their performance based on the wolpert trials in California. In Cantu D and Walker M (eds) **The Grape Genome**. Compendium of Plant Genomes. Springer, Switzerland, p. 301-318.

Pommer CV (2000) Cultivares de uva produzidos ou introduzidos pelo IAC. In: Boliani AC and Corrêa L de S (eds.) **Proceedings of the Simpósio Brasileiro sobre Uvas de Mesa 1**. Ilha Solteira, p. 51-67.

Prinsi B, Negri AS, Failla O, Scienza A and Espen L (2018) Root proteomic and metabolic analyses reveal specific responses to drought stress in differently tolerant grapevine rootstocks. **BMC plant biology 18**: 1-26.

Rad MRN and Rafezi R (2020) Integrated genetic components and machine learning approaches for better selection of traits in breeding of melon under high tunnel cultivation condition. **Plant Cell Biotechnology and Molecular Biology 21**: 36-46.

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

Radmann EB and Bianchi VJ (2008) Uva: da antiguidade a mesa de nossos dias. In Barbieri RL and Steempt ER (eds) **Origem e evolução de plantas cultivadas**. Embrapa Informação Tecnológica, Brasília, p. 891 -909.

Rebschule Mueller (2020) **Unterlagen für den Weinbau**. Germany. Disponível em <<http://www.rebschule-mueller.de/Unterlagen-Info.html>>. Acesso em setembro 2020.

Riaz S, Pap D, Uretsky J, Laucou V, Boursiquot JM, Kocsis L and Walker MA (2019) Genetic diversity and parentage analysis of grape rootstocks. **Theoretical and Applied Genetics** **132**: 1847–1860.

Roskov Y, Ower G, Orrell T, Nicolson D, Bailly N, Kirk PM, Bourgoin T, DeWalt RE, Decock W, Nieukerken EV, Zarucchi J and Penev L (2019) **Species 2000 & ITIS Catalogue of Life, 2019 Annual Checklist**. Netherlands, Species 2000. Disponível em <www.catalogueoflife.org/annual-checklist/2019>. Acesso em maio 2020.

Satisha J, Prakash GS, Murti GSR and Upreti KK (2006) Response of grape rootstocks to soil moisture stress. **Journal of Horticultural Sciences** **1**: 19-23p.

Schachtman DP and Goodger JQ (2008) Chemical root to shoot signalling under drought. **Trends Plant Sci.** **13**: 281-287.

Serra I, Strever A, Myburgh PA and Deloire A (2014) The interaction between rootstocks and cultivars (*Vitis vinifera* L.) to enhance drought tolerance in grapevine. **Australian Journal of Grape and Wine Research** **20**: 1-14.

Shalev-Shwartz S. and Ben-David S (2014) **Understanding machine learning: from theory to algorithms**. Cambridge university press, New York, 409 p.

Silva PCG and Correia RC (2016) Caracterização social e econômica da cultura da videira. In Leão PCS (eds) **Sistemas de Produção Embrapa: cultivo de videira**. 2nd ed., Embrapa Informação Tecnológica, Brasília, Embrapa Semi Árido, Petrolina.

Silva PCG, Correia RC and Soares JM (2009) Histórico e Importância Socioeconômica. In Soares JM and Leão PCS (eds) **A vitivinicultura no Semiárido Brasileiro**. Embrapa Informação Tecnológica, Brasília, Embrapa Semi-Árido, Petrolina, p.20-34.

Silva MA, Santos CM, Labate CA, Guidetti-Gonzalez S, Borges JS, Ferreira LC, Lima RO and Fritsche-Neto R (2011) Melhoramento para eficiência no uso da água. In:

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

Fritsche-Neto R and Borém A (eds) **Melhoramento de plantas para condições de estresses abióticos**. Suprema, Visconde do Rio Branco, p.126-149.

Smart DR, Schwass E, Lakso A and Morano L (2006) Grapevine rooting patterns: a comprehensive analysis and a review. **Am. J. Enol. Vitic.** **57**: 89 -104.

Soares JM and Costa FF (2009) Irrigação. In Soares JM and Leão PCS (eds) **A vitivinicultura no Semiárido Brasileiro**. Embrapa Informação Tecnológica, Brasília, Embrapa Semi-Árido, Petrolina, p.351-427.

Souza CR, Soares AM and Regina MA (2001) Trocas gasosas de mudas de videira, obtidas por dois porta-enxertos, submetidas à deficiência hídrica. **Pesquisa Agropecuária Brasileira** **36**: 1221-1230.

Storm B., and Krasokhina C. (2020) **Остальное**: Подвойные сорта винограда, сорта-озеленители, дикие виды винограда, сорта для селекции. Rostov, Rússia. Disponível em <<https://vinograd.info/sorta/ostalnoe/>>. Acesso em: 05 Ago. 2020.

Sunridge Nurseries (2020) **Rootstock chart**. Bakersfield, California. Disponível em <<https://www.sunridgenurseries.com/index.php/clonalselections/rootstock-chart>>. Acesso em: Agosto 2020.

Taiz L and Zieger E (2004) **Fisiologia vegetal**. Artmed, Porto Alegre, 723 p.

This P, Lacombe T and Thomas MR (2006) Historical origins and genetic diversity of wine grapes. **Trends in Genetics** **22**: 511-519.

Tonietto J and Carbonneau AA (2004) multicriteria climatic classification system for grape-growing regions worldwide. **Agricultural and forest meteorology** **124**: 81-97.

Tramontini S and Lovisolo C (2016) Embolism formation and removal in grapevines: a phenomenon affecting hydraulics and transpiration upon water stress. In Gerós H, Chaves MM, Gil HM and Delrot S (eds) **Grapevine in a changing environment: a molecular and ecophysiological perspective**. John Wiley & Sons, UK, p. 185 – 200.

TROPICOS - Missouri Botanical Garden (2020) **Vitis spp**. Missouri, Shaw Boulevard, Disponível em <<http://www.tropicos.org/Name/2710604>>. Acesso em março 2020.

Tsegay D, Amsalem D, Almeida M and Crandles M (2014) Responses of grapevine rootstocks to drought stress. **International Journal of Plant Physiology and Biochemistry** **6**: 1-6.

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

Villa P (2018) **Cultivar la vid**. Editorial de Vecchi, USA, 160p.

VICOPAD - Vivai Cooperativi di Padergnone (2020) **Scelta del portinnesto**. Padergnone, Italy. Disponível em <<http://www.vicopad.it/consigli-per-limpianto/>>. Acesso em: agosto 2020.

Walker MA, Heinitz C, Riaz S and Uretsky J (2019) Grape Taxonomy and Germplasm. In Cantu D and Walker MA (eds) **The grape genome**. Springer, Switzerland, p. 25-38.

Wine Australia (2016) **Grapevine rootstock selector tool**. Kent Town, Australia: Industry House, Cnr Botanic and Hackney Roads. Disponível em <<http://www.grapevinerootstock.com/about/>>. Acesso em setembro 2020.

Zarrouk O, Costa JM, Francisco R, Lopes C and Chaves MM (2016) Drought and water management in Mediterranean vineyards In Gerós H, Chaves MM, Gil HM and Delrot S (eds) **Grapevine in a changing environment**: a molecular and ecophysiological perspective. John Wiley & Sons, UK, p. 67 – 88.

Zhang L, Marguerit E, Rossdeutsch L, Ollat N and Gambetta GA (2016) The influence of grapevine rootstocks on scion growth and drought resistance. **Theoretical and Experimental Plant Physiology** 28: 143-157.

Zhao X, Liu Z, Dan F and Wang K (2016) Plant breeding evaluation with rank entropy-based decision tree. **IFAC-PapersOnLine** 49: 336-340.

CAPÍTULO II

CULTIVARES DE PORTA-ENXERTOS DE VIDEIRA: CONJUNTO DE DADOS DESCREVENDO CARACTERÍSTICAS FISIOLÓGICAS, BIOQUÍMICAS E AGRONÔMICAS

Cultivares de porta-enxertos de videira: conjunto de dados descrevendo características fisiológicas, bioquímicas e agronômicas

Grapevine rootstock cultivars: Dataset describing physiological, biochemical, and agronomic characteristics

Nina Iris Verslype^{a*}; Rosimar dos Santos Musser^a; André Câmara Alves do Nascimento^b; Raphael Miller de Souza Caldas^a; Luiza Suely Semen Martins^d; Patrícia Coelho de Souza Leão^c

^a Departamento de Agronomia, Universidade Federal Rural de Pernambuco, rua Dom Manoel de Medeiros, s/n - Dois Irmãos, Recife, 52171-900, Brasil

^b Departamento de Computação, Universidade Federal Rural de Pernambuco, rua Dom Manoel de Medeiros, s/n - Dois Irmãos, Recife, 52171-900, Brasil

^c Embrapa Semiárido, Zona Rural, s/n, Petrolina, 56302-970, Brasil.

^d Departamento de Biologia, Universidade Federal Rural de Pernambuco, rua Dom Manoel de Medeiros, s/n - Dois Irmãos, Recife, 52171-900, Brasil

* Autor correspondente.

E-mail: nina.iris.v@gmail.com (N. I. Verslype).

Artigo a ser enviado para publicação na “Data in Brief”.

INFORMAÇÃO DO ARTIGO	RESUMO
<p><i>Palavras chaves:</i></p> <p><i>Vitis</i> spp. Videira Aprendizado de máquina Mudanças climáticas Eficiência no uso da água Escassez hídrica</p>	<p>Este artigo descreve informações sobre 45 cultivares de porta-enxertos de videira, com enfoque em variáveis relacionados à tolerância à seca. Neste banco de dados, é possível encontrar cinco repetições para cada uma das 45 cultivares, o que totaliza 225 observações, descritas por 61 variáveis morfoagronômicas, bioquímicas e fisiológicas. Estas informações foram coletadas manualmente, entre os meses de junho e novembro de 2020, em artigos, livros, sites de produtores agrícolas e centros de pesquisa, disponíveis online na internet. Os dados coletados são de materiais datados entre os anos de 1925 e 2020. O objetivo de reunir estas informações em um único banco de dados, foi tornar mais acessível ao público interessado e contribuir. Portanto, este conjunto de dados pode ter um papel importante para a pesquisa e ensino sobre porta-enxertos de videira, aprendizado de máquina, melhoramento genético de videira para tolerância à seca, até mesmo como em outras áreas.</p>
ARTICLE INFO	ABSTRACT

Keywords:

Vitis spp.
Drought tolerance
Grapevine
Machine learning
Climate change
Water use efficiency
Water scarcity

This data article describes 45 grapevine rootstocks cultivars, with a focus on drought tolerance parameters. This dataset contains information about 61 parameters describing agronomic, biochemical, and physiological characteristics of the 45 cultivars with five replicates, a total of 225 observations. All the information obtained was collected manually, between June and November 2020, in scientific articles, books, websites of grapevine nurseries, and research centers, all available online on the internet and dated between 1925 and 2020. The aim of gather all this information in a single database to be more accessible for the interested public. So this dataset can have an important role for research and education in grapevine rootstocks, machine learning, plant breeding, even as in other fields.

Tabela de especificações

Área de estudo	<i>Agronomia e ciência da produção</i>
Área de assunto mais específica	<i>Informações sobre caracterização agrônômica e tolerância à seca de 45 cultivares de porta-enxertos</i>
Tipo de dados	<i>Texto e numérica</i>
Como os dados foram adquiridos	<i>Coletada manualmente em artigos, livros, sites de produtores agrícolas e centros de pesquisa</i>
Formato de dados	<i>Misto (bruta e pré-processado)</i>
Parâmetros para coleta de dados	<i>Informações sobre 45 cultivares de porta-enxertos a respeito de suas características morfoagronômicas, bioquímicas e fisiológicas</i>
Descrição da coleção de dados	<i>Coletada manualmente em artigos, livros, sites de produtores agrícolas e centros de pesquisa, todos disponíveis online na internet e foi empregado o Python para realizar análise de dados</i>
Localização da fonte de dados	<i>A lista das fontes de retirada dos dados originários é fornecida com o artigo.</i>
Acessibilidade aos dados	<i>Os dados são fornecidos com o artigo</i>

Valor dos dados

- O conjunto de dados engloba informações sobre 45 cultivares de porta-enxertos a respeito de suas características morfoagronômicas, bioquímicas e fisiológicas. Neste sentido, este conjunto de dados sobre os porta-enxertos possibilita fácil acesso e o entendimento das diferenças entre suas características.
- Pode ser empregado para realizar pesquisas para responder diferentes problemas tais como: prever classes de tolerância à seca em porta-enxertos de videira, análise de divergência genética, seleção e escolha das cultivares mais adequadas para cada região e situação que potencialmente poderia levar à resolução de problemas locais, entre outros.

- Este conjunto de dados pode ser usado para treinamento de mineração de dados, teste e validação de algoritmos de classificação de aprendizado de máquina. Também pode ser usado para seleção indireta em programas de melhoramento genético de porta-enxertos de videira, entre outros.
- O conjunto de dados também pode ser empregado por estudantes, para aprenderem a lidar com dados faltantes.

1 Descrição dos dados

Face às mudanças climáticas, os porta-enxertos de videira tolerantes à seca podem ser considerados como o melhor meio de adaptação dos vinhedos, devido ao seu papel positivo na indução da tolerância à seca na cultivar copa [1, 2]. No entanto, sua identificação e seleção é considerada difícil, uma vez que a tolerância à seca é considerada uma característica poligênica [2, 3]. Neste sentido, o uso adequado de informações associado aos algoritmos de aprendizado de máquina, podem vir a auxiliar de forma significativa os programas de melhoramento, com o intuito de selecionar materiais genéticos promissores, pois os algoritmos são capazes de gerenciar grandes quantidades de dados, assim como de identificar padrões [4, 5].

Portanto, todos os dados presentes no artigo foram coletados manualmente para o desenvolvimento de modelos de previsão de classes de tolerância à seca em porta-enxertos de videira ainda não estudados para esta característica. A Figura 1 mostra um diagrama esquemático descrevendo a organização do conjunto de dados.

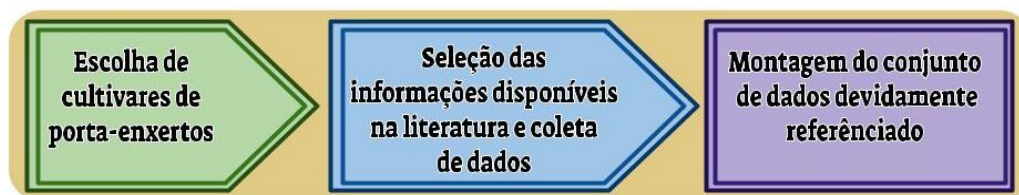


Figura 1 - Diagrama esquemático explicando a organização do conjunto de dados. Fonte: Verslype *et al.* (2021).

Embora exista uma ampla gama de informações disponíveis na literatura sobre porta-enxertos de videira, estas informações nem sempre eram facilmente acessíveis, pois às vezes só estavam disponíveis em diferentes idiomas e, muitas vezes, os materiais encontravam-se dispersos. Neste sentido, o artigo busca disponibilizar um conjunto de dados facilmente acessível ao público interessado, sobre as características agrônômicas, bioquímicas e fisiológicas de 45 cultivares de porta-

enxertos de videira (*Vitis* spp.), sendo cinco repetições, totalizando assim 225 observações.

Foram incluídos 61 variáveis, dentro deste conjunto de dados descritos de forma detalhada na próxima sessão, o que possibilita expandir seu uso em outros campos de estudo, tais como divergência genética ou, na seleção da cultivar de porta-enxerto melhor adaptada para uma região específica.

2 Projeto experimental, materiais e métodos

Entre os meses de junho e outubro de 2020, foram coletados dados manualmente em diversas fontes disponíveis na internet datadas entre os anos de 1925 à 2020, tais como artigos científicos, livros, teses, anais de congressos, sites de produtores de mudas de videira e centros de pesquisa do Brasil, Espanha, Itália, Rússia, Europa, Austrália, Nova Zelândia e Estados Unidos. As informações reunidas descrevem 45 cultivares de porta-enxertos de videira (Tabela 1), com a lista completa das cultivares disponíveis no dataset e a sua respectiva origem genética de acordo com a classificação do Catalogo Internacional de Variedades de *Vitis* – VIVC [6].

Tabela 1 – Origem genética das quarenta e cinco cultivares de porta-enxertos.

Cultivar	Origem genética
VR 039-16	<i>Vitis vinifera</i> L. x <i>Muscadinia rotundifolia</i> Michaux
VR 043-43	<i>Vitis vinifera</i> L. x <i>Muscadinia rotundifolia</i> Michaux
101-14 MGt	<i>Vitis riparia</i> Michaux x <i>Vitis rupestris</i> Scheele
106-8 MGt	<i>Vitis riparia</i> Michaux x <i>Cordifolia rupestris</i> de grasset nº1
110 R	<i>Vitis berlandieri</i> Planchon x <i>Vitis rupestris</i> Scheele
1103 P	<i>Vitis berlandieri</i> Planchon x <i>Vitis rupestris</i> Scheele
1202 C	<i>Vitis vinifera</i> L. x <i>Vitis rupestris</i> Scheele
125 AA	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux
140 Ru	<i>Vitis berlandieri</i> Planchon x <i>Vitis rupestris</i> Scheele
157-11 C	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux
1613 C	Solonis (<i>Vitis riparia</i> Michaux x <i>Vitis longii</i> Prince) x Othello
161-49 C	<i>Vitis riparia</i> Michaux x <i>Vitis berlandieri</i> Planchon
1616 C	Solonis x <i>Vitis riparia</i> Michaux
196-17 CI	1203C x <i>Vitis riparia</i> Michaux
216-3 CI	1616C x <i>Vitis rupestris</i> Scheele
26 G	<i>Vitis vinifera</i> L. x <i>Vitis riparia</i> Michaux
3306 C	<i>Vitis riparia</i> Michaux x <i>Vitis rupestris</i> Scheele
3309 C	<i>Vitis riparia</i> Michaux x <i>Vitis rupestris</i> Scheele
34 EM	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux
41 B MGt	<i>Vitis vinifera</i> L. x <i>Vitis berlandieri</i> Planchon
420 A MGt	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux
44-53 M	<i>Vitis riparia</i> Michaux x Malegue 144
5 BB	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux
5 C	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

62-66 C	<i>Vitis vinifera</i> L. X <i>Vitis cordifolia</i> Michaux
8 B	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux
93-5 C	<i>Vitis vinifera</i> L. x <i>Vitis rupestris</i> Scheele
99 R	<i>Vitis berlandieri</i> Planchon x <i>Vitis rupestris</i> Scheele
Ganzin 1	<i>Vitis vinifera</i> L. x <i>Vitis rupestris</i> Scheele
Dogridge	<i>Vitis rupestris</i> Scheele x <i>Vitis candicans</i> Engelmann
Fercal	<i>Vitis berlandieri</i> x Planchon 31R
Freedom	Fresno 1613-59 x Dogridge
Golia	Castel 156-12 x <i>Vitis rupestris</i> Scheele
Gravesac	161-49C x 3309C
Harmony	1613C x Dogridge
Riparia Gloire	<i>Vitis riparia</i> Michaux
Rupestris du lot	<i>Vitis rupestris</i> Scheele
Salt Creek	<i>Vitis champinii</i> Planchon
Schwarzmann	<i>Vitis riparia</i> Michaux x <i>Vitis rupestris</i> Scheele
SO4	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux
Sori	Solonis x <i>Vitis riparia</i> Michaux
<i>Vitis champini</i>	<i>Vitis champinii</i> Planchon
IAC 313	Golia x <i>Vitis cinerea</i> Engelmann
IAC 572	<i>Vitis caribaea</i> De Candolle x 101-14MGt
IAC 766	106-8MGt x <i>Vitis caribaea</i> De Candolle

Inicialmente foram coletados informações sobre a disponibilidade de acessos no Banco Ativo de Germoplasma da Embrapa Uva e Vinho, Brasil [7], incluindo-se prioritariamente as cultivares de porta-enxerto de videira mais usadas no Vale do Submédio São Francisco. Posteriormente, foram pesquisadas informações disponíveis na literatura consultada sobre a sua classe de tolerância à seca, e a partir destas informações optou-se pela inclusão de 14 cultivares em cada um dos três grupos de tolerância à seca (alta, média e baixa) para se ter a mesma quantidade de cultivares em cada grupo, e também foram incluídas três cultivares amplamente usadas no Vale do Submédio São Francisco cuja classe de tolerância à seca ainda é desconhecida.

Foram consideradas as informações sobre características agrônômicas, bioquímicas e fisiológicas relacionadas à tolerância à seca (Tabela 2) além de uma descrição das 61 variáveis consideradas no estudo, e suas respectivas porcentagens de dados ausentes no conjunto de dados. Essas porcentagens foram obtidas usando a função '*isnull*' disponível na biblioteca do Pandas [8].

Tabela 2 – Descrição das variáveis.

Variável	Tipo	Valores nulos(%)	Variável	Tipo	Valores nulos(%)
Genótipo	Catagórico	0	ABA	Numérico	81,78
Espécie	Catagórico	0	ABA sob estresse hídrico	Numérico	84
Genitor 1	Catagórico	0	ABA sob reidratação	Numérico	81,33

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

Genitor 2	Categórico	0	Prolina	Numérico	96
Flor	Categórico	0	Prolina sob estresse hídrico	Numérico	96,44
Ciclo vegetativo	Categórico	0	Densidade de estômatos	Numérico	76,89
Brotação	Inteiro	0	Potencial osmótico (Ψ_{os})	Numérico	96,44
Maturação	Inteiro	0	Ψ_{os} sob estresse hídrico	Numérico	96,44
Queda folhas	Inteiro	0	Condutividade hidráulica (K)	Numérico	62,22
Resistência à antracnose	Inteiro	0	Tolerância ao déficit hídrico	Inteiro	7,04
Resistência ao míldio	Inteiro	0	Tolerância ao calcário	Inteiro	17,78
Produção de estacas	Inteiro	0	Tolerância ao Calcário ativo	Numérico	24,44
Resistência à fusariose	Inteiro	49,29	Tolerância ao calcário total	Inteiro	60
Resistência a nematóides	Inteiro	11,74	Taxa fotossintética (A)	Numérico	53,05
Resistência à filoxera	Inteiro	1,41	A sob estresse hídrico	Numérico	87,56
Pega na enxertia	Inteiro	2,35	Taxa de transpiração (E)	Numérico	64,32
Facilidade de enraizar	Inteiro	1,88	E sob estresse hídrico	Numérico	94,67
Tolerância a salinidade	Inteiro	15,96	Condutância estomática (gs)	Numérico	64,32
Tolerância ao encharcamento	Inteiro	14,08	gs sob estresse hídrico	Numérico	85,78
Vigor	Inteiro	1,41	Eficiência no uso da água (EUA)	Numérico	64,78
Tolerância ao solo argiloso	Inteiro	11,73	EUA sob estresse hídrico	Numérico	89,33
Tolerância ao solo arenoso	Inteiro	15,56	Eficiência instantânea do uso da água (EIUA)	Numérico	66,2
Tolerância ao solo ácido	Inteiro	28,63	EIUA sob estresse hídrico	Numérico	95,56
Tolerância ao solo alcalino	Inteiro	30,52	Teor de N	Numérico	57,33
Tolerância ao solo compactado	Inteiro	37,56	Teor de P	Numérico	63,56
Profundidade de raízes	Inteiro	26,29	Teor de K	Numérico	87,56
Ângulo geotrópico	Numérico	43,19	Tolerância à clorose férrica	Inteiro	26,67
Distribuição de raízes até 60cm	Numérico	37,56	Índice de poder clorosante (IPC)	Numérico	44,44

Distribuição de raízes até 100cm	Numérico	40	Absorção de K	Inteiro	44,44
Absorção de N	Inteiro	37,78	Absorção de Mg	Inteiro	57,78
Absorção de P	Inteiro	46,67			

Para todas as variáveis inteiras foram determinadas uma escala de notas variando de 1 a 5 (5 = alto, 4 = alto-médio, 3 = médio, 2 = médio-baixo e 1 = baixo), com exceção da característica de profundidade do sistema radicular, cuja escala de notas variou de 1 a 6 (muito profundo = 6, profundo = 5, profundo a moderado = 4, moderado = 3, superficial = 2 e muito superficial = 1). Para as variáveis categóricas foi adotada a classificação do Catálogo Internacional de Variedades de *Vitis* - VIVC [6]. Para o sexo das flores foram consideradas três categorias; flor masculina, hermafrodita e feminina, enquanto o ciclo vegetativo pode ser precoce, intermediário e tardio. No material complementar, encontra-se disponível uma descrição detalhada das variáveis, sua origem e os respectivos procedimentos empregados para a sua obtenção. As estatísticas descritivas das variáveis categóricas foram apresentadas na (Tabela 3), e das variáveis inteiras e numéricas na (Tabela 4). As estatísticas descritivas foram obtidas no python utilizando-se a biblioteca do Pandas, no qual foi empregada a função 'describe' [1].

Na (Tabela 3), pode ser visualizado a estatística descritiva das variáveis categóricas e na (Tabela 4) das variáveis inteiras e numéricas do conjunto de dados. A estatística descritiva de ambas as tabelas foi obtida no python através da biblioteca do Pandas, no qual foi empregada a função 'describe' [1].

Tabela 3 - Estatísticas descritivas para as variáveis categóricas do conjunto de dados.

Variável	Qnt.	Único	Topo	Freq.
Genótipo	225	45	110R	5
Espécie	225	5	Cruzamento Interspecífico de <i>Vitis</i>	195
Genitor 1	225	15	<i>Vitis berlandieri</i> Planchon	65
Genitor 2	225	16	<i>Vitis rupestris</i> Scheele	70
Flor	225	3	Masculina	125
Ciclo vegetativo	225	3	Intermediário	195

Tabela 4 - Estatísticas descritivas para as variáveis inteiras e numéricas do conjunto de dados.

Variável	Qnt.	\bar{X}	s	Min.	25%	50%	75%	Max.
Brotação	225	117,7	13,44	90	108	116	126	160
Maturação	225	118,1	12,18	88	111	117	125	149
Queda folhas	225	236,8	19,36	196	224	232	248	293

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

Resistência à antracnose	225	2,51	1,16	1	2	2	3	5
Resistência ao míldio	225	3,97	0,78	1	4	4	4	5
Tolerância ao calcário	185	3,08	1,36	1	2	3	4	5
Tolerância ao Calcário ativo	170	18,12	8,79	6	130	17	20	50
Produção de estacas	225	35,27	17,24	7	22	32	48	101
Resistência à fusariose	110	2,48	1,04	1	2	2	3	5
Resistência a nematóides	190	3,12	1,6	1	1	3	5	5
Resistência à filoxera	220	3,96	1,2	1	3	4	5	5
Pega na enxertia	219	3,49	1,33	1	3	4	4	5
Facilidade de enraizar	220	3,53	1,33	1	3	4	5	5
Tolerância a salinidade	180	2,77	1,38	1	1	3	4	5
Tolerância ao encharcamento	188	2,64	1,39	1	1	3	4	5
Vigor	220	3,83	1,1	1	3	4	5	5
Tolerância ao solo argiloso	196	3,05	1,3	1	2	3	4	5
Tolerância ao solo arenoso	190	3,13	1,01	1	3	3	4	5
Tolerância ao solo ácido	155	2,99	1,22	1	2	3	4	5
Tolerância ao solo alcalino	153	3,12	1,07	1	3	3	4	5
Tolerância ao solo compactado	135	2,64	1,36	1	1	3	4	5
Profundidade de raízes	160	3,17	1,58	1	2	3	5	6
Ângulo geotrópico	122	59,62	15,88	20	48,50	60	71	90
Tolerância à clorose férrica	165	3,27	1,47	1	2	3	5	5
IPC	125	32,88	26,63	5	10	30	40	120
Distribuição de raízes até 60cm	135	70,75	15,93	32,7	59,55	71,6	79,8	98,9
Distribuição de raízes até 100cm	135	85,83	13,01	46	76,9	88,2	97,48	100
Absorção de N	140	3,33	1,43	1	2	4	5	5
Absorção de P	120	3,35	1,55	1	2	3	5	5
Absorção de k	125	3,39	1,47	1	2	4	5	5
Absorção de Mg	95	2,42	1,33	1	1	3	3,50	5
Tolerância ao calcário total	90	31,69	15,72	15	20	30	36,5	80
Teor de N	96	1,34	0,94	0,25	0,84	0,95	1,39	3,42
Teor de P	82	0,42	0,48	0,09	0,21	0,28	0,41	2,21
Teor de K	120	1,28	0,68	0,38	0,79	1,17	1,53	3,37
A	100	12,84	5,04	5,27	8,51	12,55	15,1	40,6
A sob estresse hídrico	28	6,94	4,02	0,3	4,38	7,75	8,5	15,1
E	76	6,22	1,43	1,66	5,4	6,14	6,93	9,56
E sob estresse hídrico	12	5,16	3,28	0,09	2,66	6,5	7,71	9,4
gs)	76	0,79	0,91	0,07	0,26	0,36	0,66	3
gs sob estresse hídrico	32	0,19	0,15	0,03	0,09	0,1	0,28	0,58
EUA	75	40,41	33,85	3,14	13,14	37,54	62,95	183,6
EUA sob estresse hídrico	24	39,36	29,16	5,56	16,69	25,21	72,33	91,11

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

EIUA	76	2,38	2,7	0,87	1,37	2,18	2,63	24,46
EIUA sob estresse hídrico	10	1,66	0,83	0,58	1,1	1,635	2,12	3,33
ABA	41	23,7	13,8	6	16,4	21,4	27,6	73,8
ABA sob estresse hídrico	36	70,67	43,63	17,9	37,15	56,90	93,1	192,2
ABA sob reidratação	42	22,02	9,05	13,7	15,67	18,12	25,84	40,85
Prolina	9	7,16	21,11	0,08	0,09	0,13	0,14	63,44
Prolina sob estresse hídrico	8	12,41	2,82	7,96	10,55	12,68	14,21	15,88
Densidade estômatos	52	138,4	25,43	61	127,5	134,5	153	188
Ψ_{os}	8	1,68	0,24	1,49	1,56	1,59	1,69	2,24
Ψ_{os} sob estresse hídrico	8	2,14	0,96	1,25	1,63	1,745	2,45	4,26
K	85	1,79	0,63	0,71	1,39	1,65	2,27	4,08
Tolerância ao déficit hídrico	210	2	0,82	1	1	2	3	3

Para visualizar a distribuição de todas as variáveis categóricas no conjunto de dados, foi empregado um gráfico de barras disponível na biblioteca Matplotlib em Python [9] (Figura 2). Já para a visualização da distribuição de todas as variáveis numéricas e inteiras em cada classe de tolerância à seca foi utilizado o gráfico Kernel densidade estimation (KDE), disponível na biblioteca Seaborn em Python [10] (Figuras 3, 4, 5 e 6). O KDE é um método que possibilita visualizar a distribuição das observações em um conjunto de dados, através da representação dos dados por meio de uma curva de densidade de probabilidade contínua em uma ou mais dimensões [10].

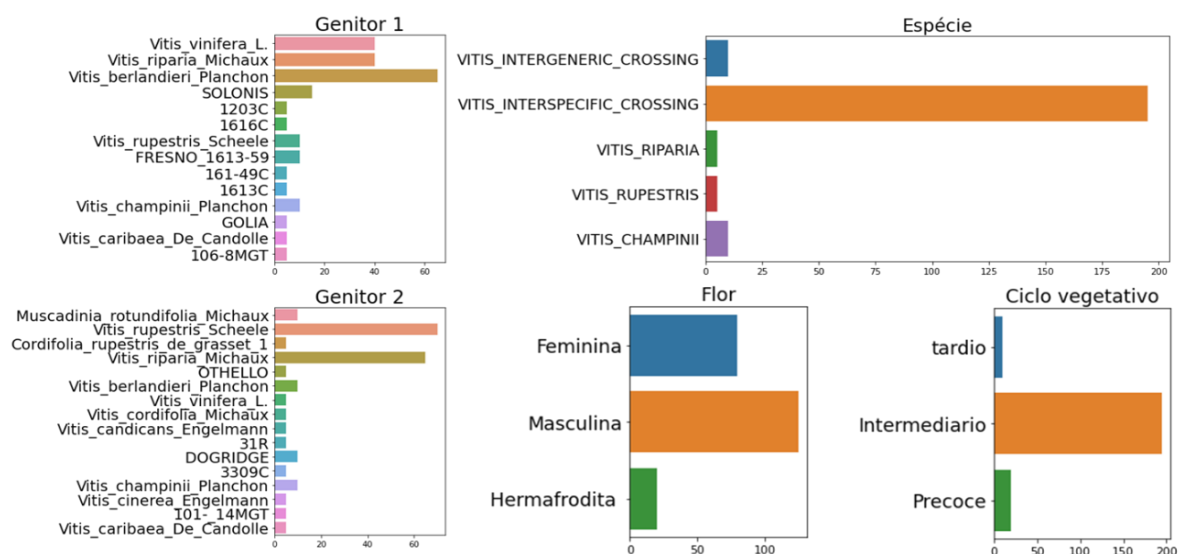


Figura 2- Distribuição dos dados categóricos no conjunto de dados. Fonte: Verslype *et al.* (2021).

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

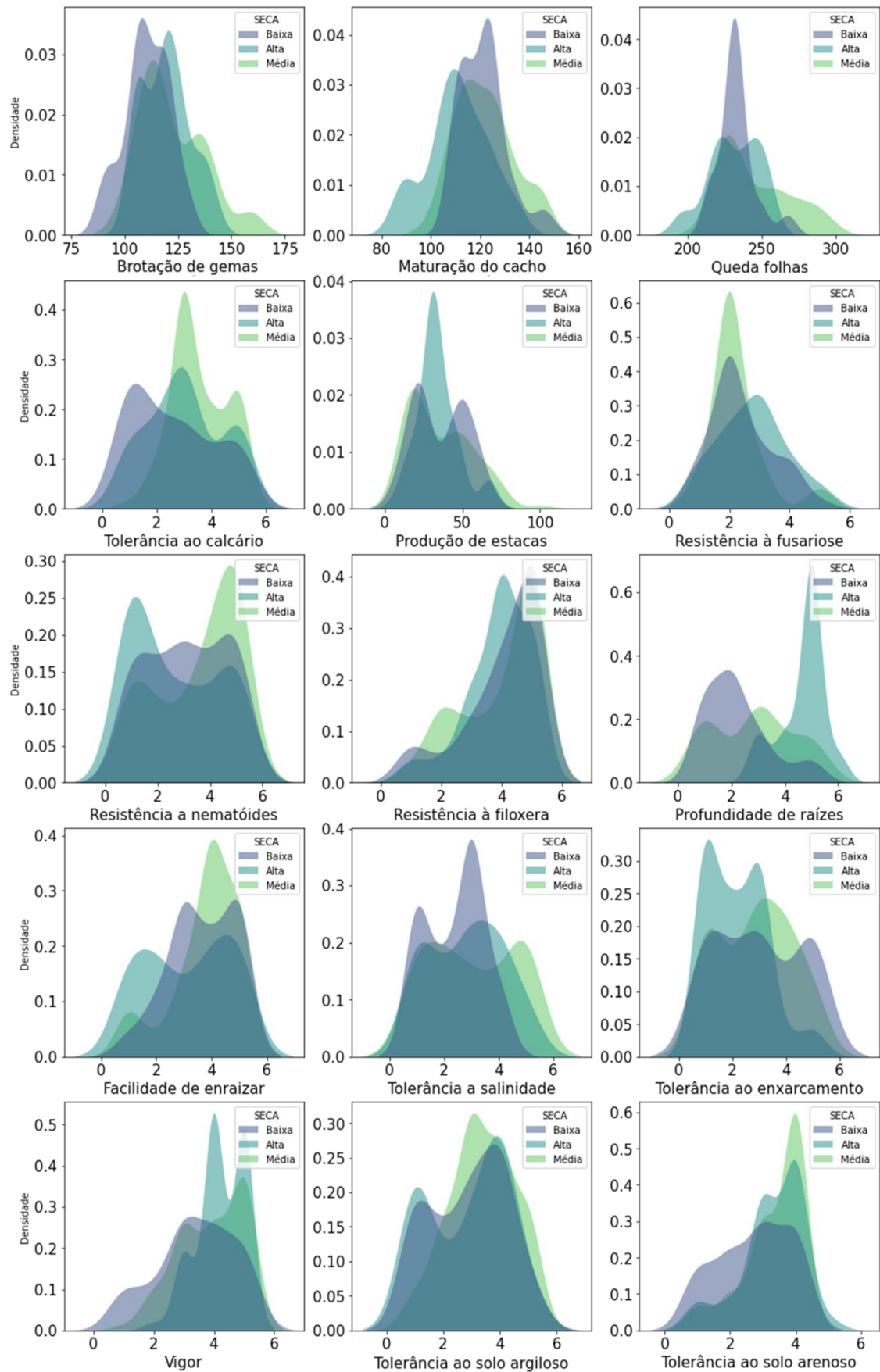


Figura 3 - Distribuição das variáveis inteiras dentro de cada classe de tolerância ao déficit hídrico. Fonte: Verslype *et al.* (2021).

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

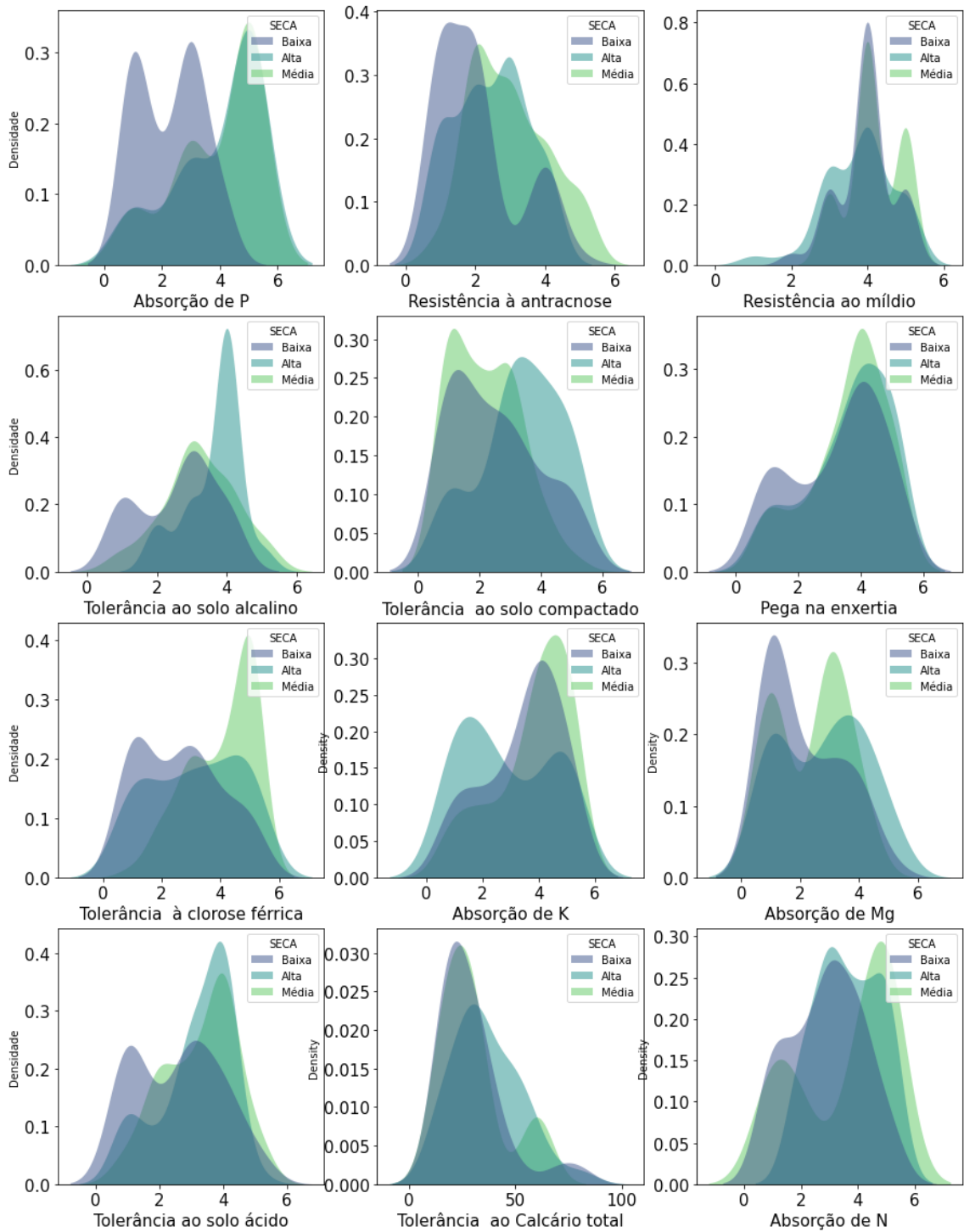


Figura 4 - Distribuição das variáveis inteiras dentro de cada classe de tolerância ao déficit hídrico. Fonte: Verslype *et al.* (2021).

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

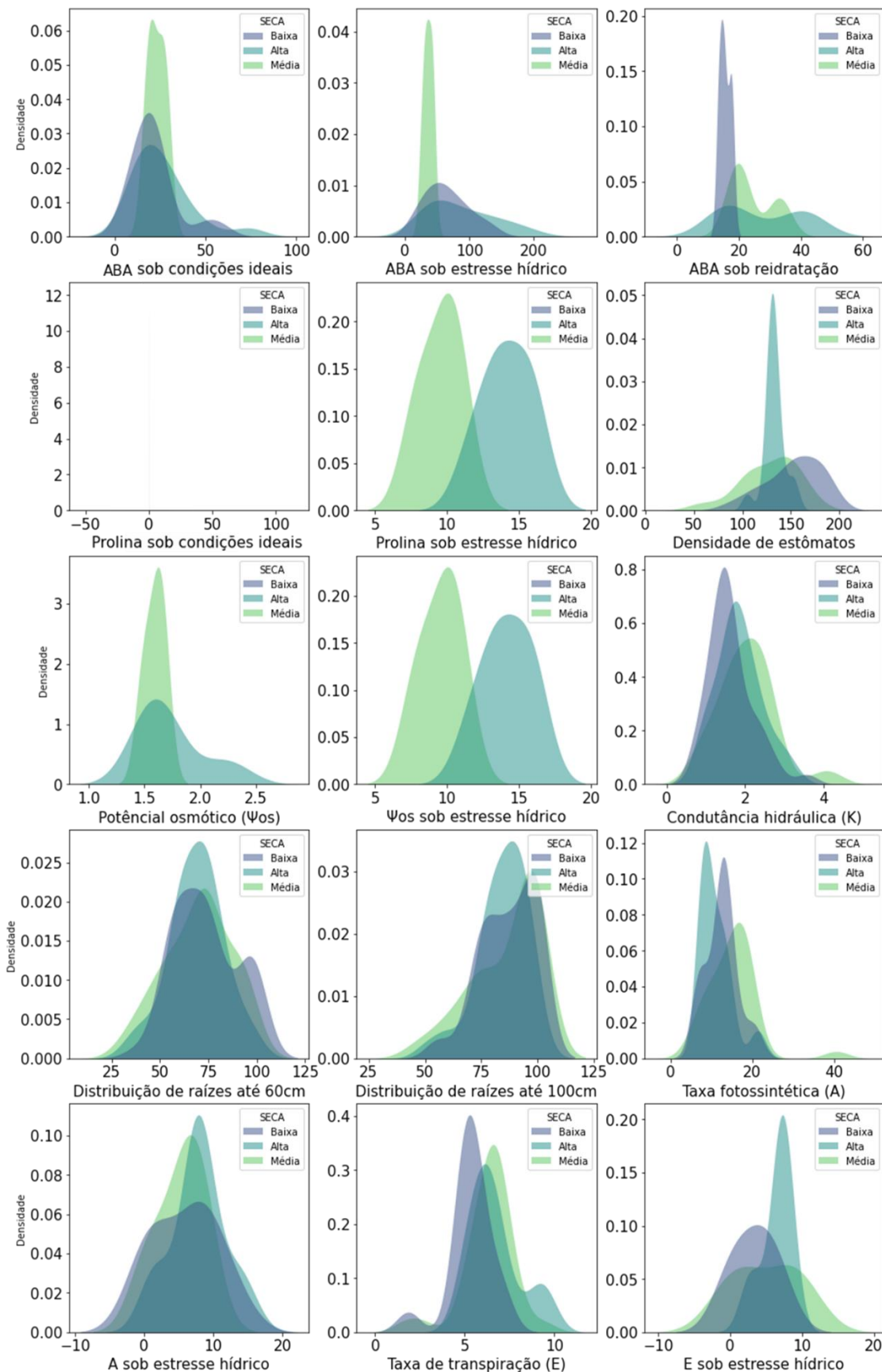


Figura 5 - Distribuição das variáveis numéricas dentro de cada classe de tolerância à seca. A prolina sob condições ideais não formou a curva KDE devido à ausência de dados em mais de uma classe de tolerância ao déficit hídrico. Fonte: Verslype *et al.* (2021).

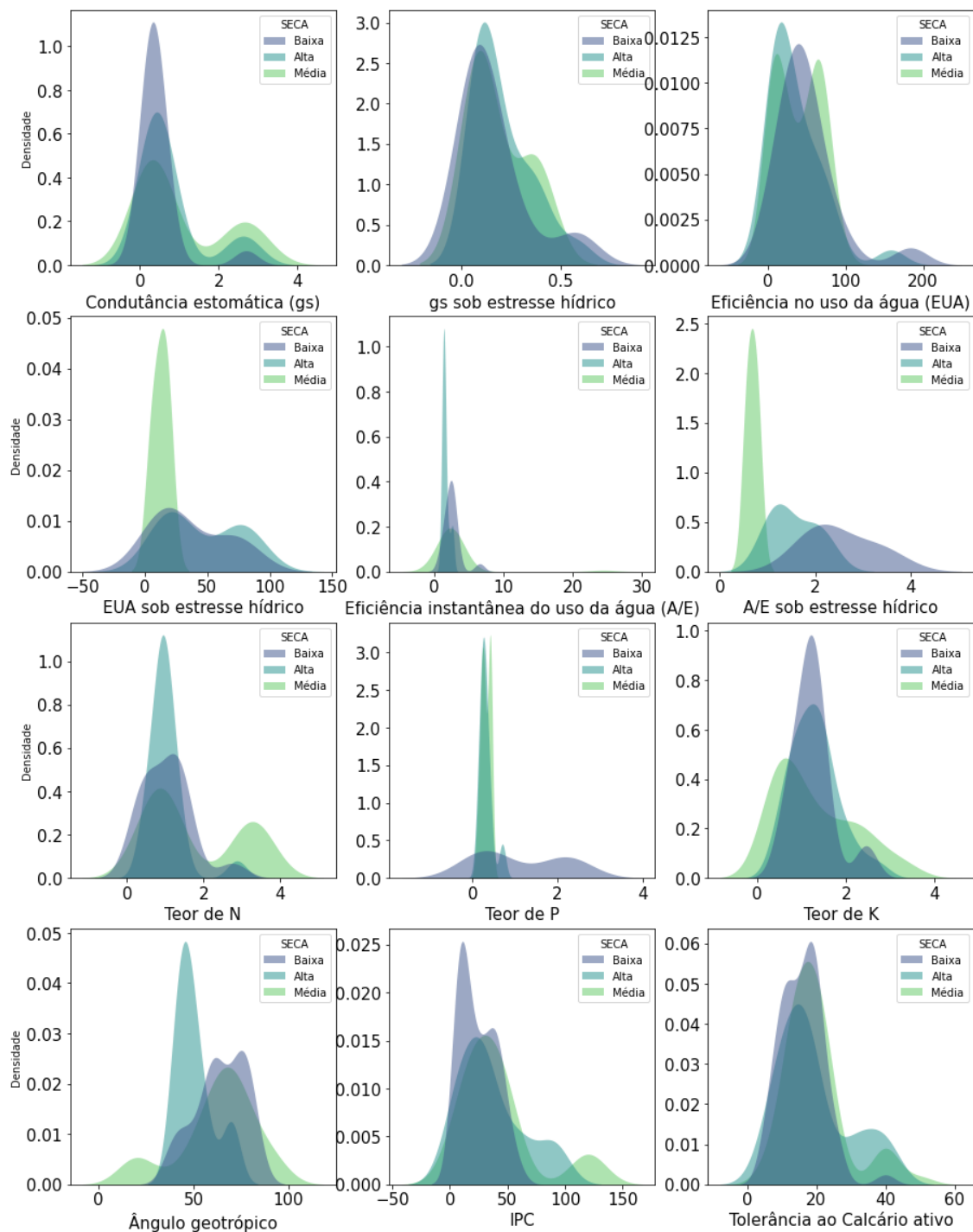


Figura 6 - Distribuição das variáveis numéricas dentro de cada classe de tolerância ao déficit hídrico. Fonte: Verslype *et al.* (2021).

Agradecimentos

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão de bolsa de mestrado à primeira autora.

Referências

- [1] Elisa Marguerit, Jean-Pascal Tandonnet, Bastien Golard, Anthony Peccoux, Kees Van Leeuwen, and Nathalie Ollat. Le porte-greffe, outil original et pertinent d'adaptation au changement climatique. In 10. Journée Technique du Conseil Interprofessionnel du Vin de Bordeaux (CIVB), pages 4–p, 2011.
- [2] I Serra, A Strever, PA Myburgh, and Alain Deloire. the interaction between rootstocks and cultivars (*Vitis vinifera*.) to enhance drought tolerance in grapevine. *Australian Journal of Grape and Wine Research*, 20(1):1–14, 2014.
- [3] Silvina Dayer, Idan Reingwartz, Andrew J McElrone, and Gregory A Gambetta. Response and recovery of grapevine to water deficit: from genes to physiology. In *The grape genome*, pages 223–245. Springer, 2019.
- [4] PC Cavatte, SCV Martins, LE Morais, PEM Silva, LT Souza, FM da Matta, R Fritsche Neto, and A Borém. Afisiologia dos estresses abióticos. *Melhoramento de plantas para condições de estresses abióticos*. Visconde do Rio Branco: Suprema, pages 39–79, 2011.
- [5] Shai Shalev-Shwartz and Shai Ben-David. *Understanding machine learning: From theory to algorithms*. Cambridge university press, 2014.
- [6] Camargo U. A., Ritschel P. S., Maia J. D. G., Quecini V., Machado C. A. E., Bosco D. D., Sinski I., Comachio V., and Zilio R. Banco ativo de germoplasma de uva. <https://www.embrapa.br/en/uva-e-vinho/banco-ativo-de-germoplasma-de-uva/pesquisa>, 2017 (accessed 25 June 2020).
- [7] Maul E., Tpferr R., Röckel F., Brühl U., Hundemer M., Mahler-Ries A., Walk M., Kecke S., Wolck A., and Ganesch A. *Vitis international variety catalogue (vivc)*. <http://www.vivc.de/>, 2020 (accessed 20 June 2020).
- [8] McKinney W. *Data Structures for Statistical Computing in Python*. In Stéfan van der Walt and Jarrod Millman, editors, *Proceedings of the 9th Python in Science Conference*, pages 56 – 61, 2010.
- [9] J. D. Hunter. *Matplotlib: A 2d graphics environment*. *Computing in Science & Engineering*, 9(3):90–95, 2007.
- [10] Michael Waskom and the seaborn development team. *mwaskom/ seaborn*, September 2020.

CAPÍTULO III

ANÁLISE DA DIVERSIDADE GENÉTICA DE PORTA-ENXERTOS DE VIDEIRA ATRAVÉS DE APRENDIZADO DE MÁQUINA

Análise da diversidade genética de porta-enxertos de videira através de aprendizado de máquina

Genetic diversity analysis of grapevine rootstocks through machine learning

Nina Iris Verslype^{a,*}, Rosimar dos Santos Musser^a, André Câmara Alves do Nascimento^b, Raphael Miller de Souza Caldas^a, Luiza Suely Semen Martins^d,
Patrícia Coelho de Souza Leão^c

^a Departamento de Agronomia, Universidade Federal Rural de Pernambuco, rua Dom Manoel de Medeiros, s/n - Dois Irmãos, Recife, 52171-900, Brasil

^b Departamento de Computação, Universidade Federal Rural de Pernambuco, rua Dom Manoel de Medeiros, s/n - Dois Irmãos, Recife, 52171-900, Brasil

^c Embrapa Semiárido, Zona Rural, s/n, Petrolina, 56302-970, Brasil.

^d Departamento de Biologia, Universidade Federal Rural de Pernambuco, rua Dom Manoel de Medeiros, s/n - Dois Irmãos, Recife, 52171-900, Brasil

* Autor correspondente.

E-mail: nina.iris.v@gmail.com (N. I. Verslype).

Artigo a ser enviado para publicação na “Scientia Horticulturae”.

RESUMO

A videira (*Vitis* spp.) é considerada uma fruteira de grande importância socioeconômica, cujas áreas cultivadas globalmente podem ser afetadas pela limitação de recursos hídricos e mudanças climáticas. Face a esta problemática tem havido um crescente investimento no desenvolvimento e uso de porta-enxertos tolerantes déficit hídrico. No entanto, a obtenção de novas cultivares trata-se de um processo lento e oneroso. Por conta disto, o uso de novas ferramentas, tais como o uso algoritmos de aprendizado de máquina podem facilitar as avaliações, devido a capacidade de gerenciar grandes quantidades de dados e identificar padrões relevantes, o que pode levar a resultados mais precisos e melhor interpretação do que os métodos estatísticos clássicos. Neste sentido, o objetivo deste trabalho foi avaliar a divergência genética de 45 cultivares de porta-enxertos de videira, usando-se informações coletadas, entre os meses de junho e novembro de 2020, de artigos, livros e sites. Foram aplicados algoritmos de agrupamento, com o intuito de formar grupos heteróticos e identificar os melhores cruzamentos. Os resultados obtidos indicaram a formação de cinco grupos heteróticos, mostrando-se uma ferramenta eficiente, que pode contribuir de forma significativa na seleção e indicação de cruzamentos entre porta-enxertos de videira com o objetivo de desenvolver novas cultivares tolerantes ao déficit hídrico.

Palavras- Chaves: *Vitis* spp., agrupamentos, tolerância à seca, ACP, K-means.

ABSTRACT

The grapevine (*Vitis* spp.) is a fruit of considered socioeconomic importance, whose globally cultivated areas may be affected by limited water resources and climate

change. Given this problem, there is a growing investment in the development and use of drought-tolerant rootstock cultivars. However, the obtention of new cultivars is a long time-consuming and costly process. Because of this, the use of new tools, such as the use of machine learning algorithms can facilitate evaluations, due to the ability to generate large amounts of data and identify patterns, which can lead to more accurate results and better interpretation than classical statistical methods. In this sense, this work aimed to assess the genetic divergence of 45 cultivars of grapevine rootstocks, using information collected, between June and November 2020, in articles, books, and websites. Which were submitted to clustering algorithms, intending to form heterotic groups and identify the best crosses. The results obtained indicated the formation of five heterotic groups, proving to be an efficient tool, which can significantly contribute to the selection and indication of crosses between grapevine rootstocks in order to develop new drought tolerant cultivars.

Key words: *Vitis* spp., clusters, drought tolerance, PCA, K-means.

1 Introdução

Entre as culturas agrícolas perenes exploradas pelo homem a nível mundial destaca-se a videira (*Vitis* spp.), por conta de sua grande versatilidade de produtos e importância socioeconômica. Esta cultura, vem sendo cultivada e moldada pelo homem desde a antiguidade, em regiões com condições ambientais muito diversas (Leão e Borges, 2009; Dougherty, 2012; Cantu e Walker, 2019). Os maiores produtores mundiais, na safra de 2019, encontravam-se localizados no continente Asiático seguido pelo continente Europeu, Americano, Africano e pela Oceania, onde foram colhidas 77,1 milhões de toneladas de uvas frescas (FAOSTAT 2020). O Brasil em 2019, foi considerado o décimo quinto maior produtor mundial de uva, contribuindo com cerca de 1,9% da produção mundial (IBGE, 2020; FAOSAT, 2020).

O sucesso da viticultura na maioria das regiões vitícolas pelo mundo, ocorre em parte por conta da técnica da enxertia, que agrega diversas vantagens (Hernandes *et al.* 2011). Isto possibilita às cultivares copas atingirem o seu potencial máximo de produção e qualidade das uvas, quando a associação entre a cultivar porta-enxerto e a copa, são usados de forma adequada (Leão *et al.*, 2009; Ollat *et al.*, 2016). Desta forma, a enxertia pode ser apontada como o fundamento da viticultura, sendo considerada relativamente recente no setor vitivinícola quando se leva em conta a sua extensa história de cultivo, tendo-se iniciado o seu uso no século XIX para os vinhedos europeus resistirem à infestação por filoxera (Hernandes *et al.*, 2011; Cantu e Walker, 2019; Riaz *et al.*, 2019). Apesar de sua importância ser reconhecida mundialmente,

quase todos os porta-enxertos de videira utilizados atualmente foram desenvolvidos há mais de 100 anos e apresentam uma estreita base genética (Riaz *et al.* 2019).

No mundo, tem-se havido uma crescente demanda por porta-enxertos de videira tolerantes ao déficit hídrico, face à previsão das mudanças climáticas e potencial limitação de recursos hídricos (Marguerit *et al.*, 2012; Ollat *et al.*, 2016; Riaz *et al.*, 2019). No entanto, a obtenção de novas cultivares tratar-se de um processo lento, que pode variar entre 7 e 25 anos através do melhoramento tradicional, além de ser oneroso e difícil por se tratar de uma espécie perene e de uma característica de herança complexa (Bruckner & Albuquerque, 2005; Dougherty, 2012; Marguerit *et al.*, 2012; Ollat *et al.*, 2016).

De acordo com Kanavi *et al.* (2020), o aprendizado de máquina vem sendo cada vez mais empregado em diversas áreas de estudo, como no melhoramento genético de plantas, por possibilitar identificar, prever e classificar genótipos conforme a necessidade do programa de melhoramento (Beiki, 2019; Etminan *et al.*, 2019). Neste contexto, algoritmos de agrupamento têm sido bastante utilizados, uma vez que trata-se de uma técnica muito poderosa e útil para identificar linhagens parentais promissoras em programas de melhoramento genético de plantas (Kanavi *et al.*, 2020; Nandini *et al.*, 2020). Bem como, para se desenvolver genótipos tolerantes ao déficit hídrico (Kanavi *et al.*, 2020).

Neste sentido, o objetivo deste trabalho é determinar a diversidade genética de 45 cultivares de porta-enxertos de videira, através do uso de algoritmos de agrupamento, para a orientação de cruzamentos entre genótipos divergentes visando o desenvolvimento de porta-enxertos de videira tolerantes ao déficit hídrico.

2 Material e Métodos

O estudo foi conduzido entre os meses de julho de 2020 e março de 2021, realizando-se previamente busca de informações disponíveis na literatura sobre quarenta e cinco cultivares de porta-enxertos de videira, que foram reunidas, organizadas e pré-processadas para finalmente serem submetidos ao algoritmo não supervisionado K-means, para a formação de grupos heteróticos, conforme pode ser visualizado na (Figura 1).



Figura 1 - Diagrama esquemático da análise de agrupamento do conjunto de dados. Fonte: Verslype *et al.* (2021).

2.1 Material vegetal

As quarenta e cinco cultivares de porta-enxertos de videira empregadas neste estudo, foram listadas na (Tabela 1) com informações sobre a sua origem genética e o sexo da flor. As cultivares foram selecionadas com base em três critérios: 1) cultivares comumente empregadas no Submédio do Vale do São Francisco; 2) disponibilidade de material no Banco Ativo de Germoplasma de videira da Embrapa (Camargo *et al.*, 2017); e 3) disponibilidade de informações na literatura sobre o grau de tolerância ao déficit hídrico de cada porta-enxerto.

Tabela 1 - Quarenta e cinco cultivares de porta-enxertos de videiras (*Vitis* spp.) avaliadas.

Cod	Cultivar	Origem genética	Flor
0	VR 039-16	<i>Vitis vinifera</i> L. x <i>Muscadinia rotundifolia</i> Michaux ¹	Feminina ¹
1	VR 043-43	<i>V. vinifera</i> L. x <i>M. rotundifolia</i> Michaux ¹	Masculina ³
2	101-14 MGt	<i>V. riparia</i> Michaux x <i>V. rupestris</i> Scheele ^{1;3}	Feminina ^{1;2;3;4}
3	106-8 MGt	<i>V. riparia</i> Michaux x <i>Cordifolia rupestris</i> de grasset n ^o 1 ¹	Feminina ^{1;2}
4	110 R	<i>V. berlandieri</i> Planchon x <i>V. rupestris</i> Scheele ^{1;5}	Masculina ^{1;2;4}
5	1103 P	<i>V. berlandieri</i> Planchon x <i>V. rupestris</i> Scheele ^{1;5}	Masculina ^{1;2;3;4}
6	1202 C	<i>V. vinifera</i> L. x <i>V. rupestris</i> Scheele ¹	Monóica ^{1;2}
7	125 AA	<i>V. berlandieri</i> Planchon x <i>V. riparia</i> Michaux ¹	Feminina ^{1;2;4}
8	140 Ru	<i>V. berlandieri</i> Planchon x <i>V. rupestris</i> Scheele ¹	Masculina ^{1;2;4}
9	154-11 C	<i>V. berlandieri</i> Planchon x <i>V. riparia</i> Michaux ¹	Feminina ^{1;2}
10	1613 C	Solonis x <i>Othello</i> ¹	Feminina ^{1;2}
11	161-49 C	<i>V. riparia</i> Michaux x <i>V. berlandieri</i> Planchon ¹	Feminina ^{1;2;4}
12	1616 C	Solonis x <i>V. riparia</i> Michaux ¹	Masculina ^{1;2}
13	196-17 Cl	1203C x <i>V. riparia</i> Michaux ¹	Masculina ^{1;2;4}
14	216-3 Cl	1616C x <i>V. rupestris</i> Scheele ¹	Masculina ^{1;2;4}
15	26 G	<i>V. vinifera</i> L. x <i>V. riparia</i> Michaux ¹	Feminina ^{1;2}
16	3306 C	<i>V. riparia</i> Michaux x <i>V. rupestris</i> Scheele ¹	Masculina ^{1;2}
17	3309 C	<i>V. riparia</i> Michaux x <i>V. rupestris</i> Scheele ¹	Masculina ^{1;2;4}
18	34 EM	<i>V. berlandieri</i> Planchon x <i>V. riparia</i> Michaux ¹	Masculina ^{1;2;4}
19	41 B MGt	<i>V. vinifera</i> L. x <i>V. berlandieri</i> Planchon ¹	Monóica ^{1;2}
20	420 A MGt	<i>V. berlandieri</i> Planchon x <i>V. riparia</i> Michaux ^{1;3}	Masculina ^{1;2;3;4}
21	44-53 M	<i>V. riparia</i> Michaux x Malegue 144 ¹	Masculina ^{1;2;4}
22	5 BB	<i>V. berlandieri</i> Planchon x <i>V. riparia</i> Michaux ¹	Feminina ^{1;2;4}
23	5 C	<i>V. berlandieri</i> Planchon x <i>V. riparia</i> Michaux ¹	Masculina ^{1;2;4}
24	62-66 C	<i>V. vinifera</i> L. x <i>V. cordifolia</i> Michaux ¹	Masculina ^{1;2}
25	8 B	<i>V. berlandieri</i> Planchon x <i>V. riparia</i> Michaux ¹	Masculina ^{1;2}
26	93-5 C	<i>V. vinifera</i> L. x <i>V. rupestris</i> Scheele ¹	Feminina ^{1;2}
27	99 R	<i>V. berlandieri</i> Planchon x <i>V. rupestris</i> Scheele ^{1;5}	Masculina ^{1;2;4}

28	Ganzin 1	<i>V. vinifera</i> L. x <i>V. rupestris</i> Scheele ¹	Masculina ^{1;2}
29	Dogridge	<i>V. rupestris</i> Scheele x <i>V. candicans</i> Engelman ¹	Feminina ^{1;2}
30	Fercal	<i>V. berlandieri</i> Planchon x 31R ¹	Feminina ^{1;2;4}
31	Freedom	Fresno 1613-59 x Dogridge ¹	Monóica ²
32	Golia	Castel 156-12 x <i>V. rupestris</i> Scheele ¹	Monóica ^{1;2}
33	Gravesac	161-49C x 3309C ¹	Masculina ^{1;2;4}
34	Harmony	1613C x Dogridge ¹	Feminina ^{1;2}
35	Riparia Gloire	<i>V. riparia</i> Michaux ¹	Masculina ^{1;2;4}
36	Rupestris du lot	<i>V. rupestris</i> Scheele ^{1;5}	Masculina ^{1;2;4}
37	Salt Creek	<i>V. champinii</i> Planchon ¹	Feminina ^{1;2}
38	Schwarzmann	<i>V. riparia</i> Michaux x <i>V. rupestris</i> Scheele ¹	Masculina ^{1;2}
39	SO4	<i>V. berlandieri</i> Planchon x <i>V. riparia</i> Michaux ¹	Masculina ^{1;2;3;4}
40	Sori	Solonis x <i>V. riparia</i> Michaux ¹	Feminina ¹
41	<i>Vitis champini</i>	<i>V. champinii</i> Planchon ¹	Feminina ²
42	IAC 313	Golia x <i>V. cinerea</i> Engelman ^{1;3}	Masculina ³
43	IAC 572	<i>V. caribaea</i> De Candolle x 101-14MGt ^{1;3}	Masculina ^{2;3}
44	IAC 766	106-8MGt x <i>V. caribaea</i> De Candolle ^{1;3}	Masculina ^{2;3}

Informações obtidas em: ¹Maul *et al.* (2020); ²Camargo *et al.* (2017); ³Embrapa Uva e Vinho (2016); ⁴Audeguin *et al.* (2020); ⁵Keller (2015).

2.2 Características avaliadas

Para a realização deste estudo foram coletados dados em revistas científicas, livros, anais de congressos, sites de órgãos de pesquisa e em viveiros de mudas, totalizando 59 variáveis listadas abaixo:

- Variáveis fisiológicas: condutância estomática sob condições ideais (g_s) e sob estresse hídrico ambas em $\mu\text{mol CO}_2 \cdot \text{m}^{-2} \cdot \text{s}^{-1}$; taxa de transpiração sob condições ideais (E) e sob estresse hídrico em $\text{mmol H}_2\text{O} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$; taxa fotossintética sob condições ideais (A) e sob estresse hídrico em $\mu\text{mol m}^{-2} \cdot \text{s}^{-1}$; eficiência no uso da água (EUA) sob condições ideais e sob estresse hídrico em $\mu\text{mol CO}_2 \text{ mol}^{-1} \text{ H}_2\text{O}$; eficiência instantânea do uso da água ($EiUA$) sob condições ideais e sob estresse hídrico em $\mu\text{mol CO}_2 \text{ mmol}^{-1} \text{ H}_2\text{O}$; potencial osmótico sob condições ideais (Ψ_{os}) e sob estresse hídrico em Mpa, condutividade hidráulica (KI) em kg/MPa^{-1} ;
- Variáveis bioquímicas: teor de ácido abscísico sob condições ideais (ABA), sob estresse hídrico e condição de reidratação da planta em $\text{ng} \cdot \text{g}^{-1}$; teor de prolina sob condições ideais e sob estresse hídrico em $\text{mg} \cdot \text{g}^{-1}$;
- Variáveis nutricionais: teor dos macronutrientes nitrogênio, fósforo e potássio presente no pecíolo da folha, em porcentagem; capacidade da planta em absorver nitrogênio, fósforo, potássio e magnésio, em uma escala de notas atribuída entre 1 e 5 (alta = 5, média-alta = 4, média = 3, média-baixa = 2 e baixa = 1);
- Variáveis morfoagronômicas: densidade estomática em estomato/ mm^2 ; sexo da flor (feminino, masculino, ou hermafrodita); início de quedas de folhas, brotação e

maturação em dias; tolerância ao calcário ativo em porcentagem; produção de estacas em número de estacas por plantas; profundidade do sistema radicular em uma escala de notas atribuída entre 1 e 6 (muito profundo = 6, profundo = 5, moderado-profundo = 4, moderado = 3, superficial = 2, muito superficial = 1); ângulo geotrópico em grau; distribuição de raízes até 60 cm e 100 cm em porcentagem; índice de poder clorosante (IPC) em porcentagem; tolerância ao déficit hídrico em uma escala de notas variando entre 1 e 3 (alto=3, médio=2 e baixo=1); ciclo vegetativo (precoce, intermediário e tardio); e nas variáveis listadas a seguir foram atribuídas uma escala de notas variando entre 1 e 5 (alta = 5, média-alta = 4, média = 3, média-baixa = 2 e baixa = 1), para resistência à antracnose, míldio, fusariose, filoxera e nematóide; tolerância ao calcário; pegamento na enxertia; tolerância à clorose férrica; facilidade de enraizamento de estacas; tolerância ao calcário total; vigor induzido a copa; tolerância a solo salino, encharcado, argiloso, arenoso, ácido, alcalino e compactado.

2.3 Análises de correlação e agrupamento

A análise de correlação e divergência genética entre as 45 cultivares de porta-enxertos de videira foi realizada utilizando a linguagem Python na plataforma Google Colaboratory¹.

2.3.1 Análise de correlação de Pearson

As variáveis avaliadas foram inicialmente submetidas à estimativa do coeficiente de correlação de Pearson (r), empregando-se a função `corr()` disponível na biblioteca `pandas` (McKinney, 2010), com o intuito de examinar a correlação entre as variáveis e evitar a utilização de características fortemente correlacionadas. Na equação (01) podemos ver como são calculados os coeficientes de correlação de Pearson de acordo com Ferreira *et al.* (2011).

$$r = \left(\sum xy - \left(\sum x * \sum y / N \right) \right) / \sqrt{\left(\sum x^2 - \left(\left(\sum x \right)^2 / N \right) \right) * \left(\sum y^2 - \left(\left(\sum y \right)^2 / N \right) \right)} \quad (01)$$

onde: r corresponde ao coeficiente de correlação que pode variar entre -1 e +1; N é o número de observações; x e y são os valores de ambas as variáveis.

¹ <https://colab.research.google.com/>

Posteriormente, com o intuito de melhorar a visualização dos coeficientes de correlação de Pearson, foi gerado um mapa de calor disponível na biblioteca seaborn (Waskom *et al.*, 2020),

2.3.2 Pré-processamento do conjunto de dados

O pré-processamento do conjunto de dados consistiu inicialmente na adoção da média das cinco repetições de cada cultivar, e em seguida foi realizado o preenchimento dos dados faltantes com a média por meio da função *df.fillna()*, presente na biblioteca do pandas (McKinney, 2010), e a normalização através do *StandardScaler* (Figura 2), disponível na biblioteca do scikit-learn, que transforma os dados e os coloca na mesma escala, através da padronização das *features* removendo a média e escalonando a variância a uma unidade, conforme pode ser visualizado na equação (02) o cálculo da pontuação de uma amostra *x* (Pedregosa *et al.*, 2011).

$$Z = (x - u) / s \quad (02)$$

onde: *u* corresponde a média das amostras; *s* é o desvio padrão das amostras.

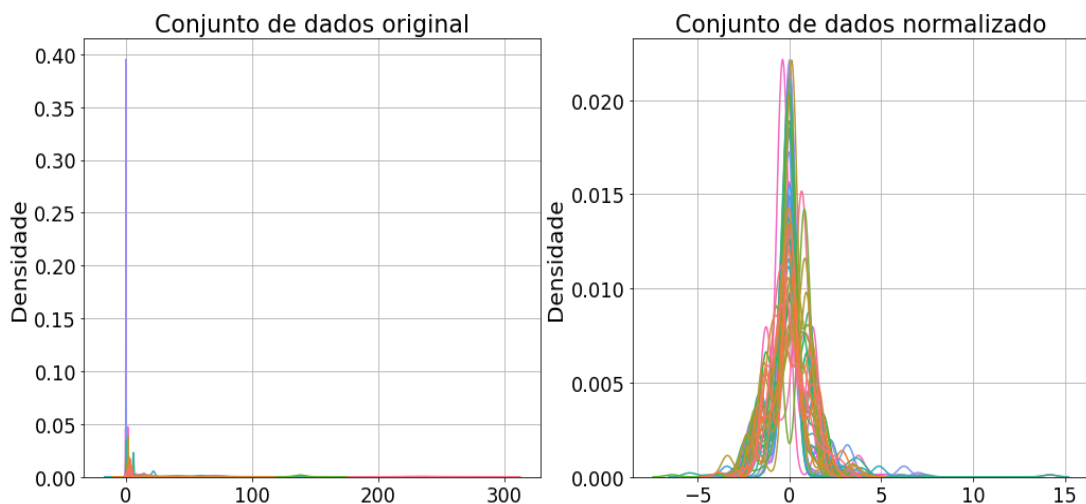


Figura 2 - Distribuição do conjunto de dados original e normalizado. Fonte: Verslype *et al.* (2021).

2.3.3 Determinação do número de agrupamentos

O número de agrupamentos (*K*), ou no inglês, *clusters*, foi determinado através do método do Cotovelo, ou no inglês, *Elbow Method*, que para Forsyth (2018) trata-se de uma estratégia interessante e útil por representar graficamente o valor de *K*, como uma função e, possibilitar observar o “cotovelo” da curva e escolher o melhor número de clusters. A distância das observações até o centro do cluster foi obtida através da

medida da soma dos quadrados intra-clusters, ou no inglês, *Within Cluster Sum of Squares* (WCSS), disponível na biblioteca do scikit-learn (Pedregosa *et al.*, 2011). Posteriormente foi aplicado o algoritmo Kneedle (Satopaa *et al.* 2011), baseado na noção de que os pontos de curvatura máxima em um conjunto de dados o ‘cotovelo’ da curva, com o intuito de preservar o comportamento geral dos dados, sendo útil para solucionar problemas de detecção do cotovelo da curva em questão.

2.3.4 Análise de K-means e Componentes Principais

As 45 cultivares de porta-enxerto de videira foram classificadas através da análise de K-means e a Análise de Componentes Principais, ou no inglês, *Principal Component Analysis* (PCA), conforme o modelo explicado por Skansi (2018). Ambas as análises encontram-se disponíveis na biblioteca do scikit-learn (Pedregosa *et al.*, 2011). A análise de PCA foi empregada para visualização da distribuição dos genótipos dentro dos agrupamentos, segundo Skansi (2018) trata-se de um algoritmo simples não supervisionado útil para construir representações distribuídas.

2.3.5 Diversidade genética

Para indicar os cruzamentos mais interessantes entre as 45 cultivares de porta-enxertos, foi adotado como medida de dissimilaridade a distância euclidiana, que pode ser calculada, segundo Ertel (2017), através da equação (04).

$$d(x, y) = |x - y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (04)$$

onde: d é a distância entre dois pontos x e y , i é o índice do vetor de n atributos dos objetos, x é o i ésimo atributo do objeto i e y é o i ésimo atributo do objeto j .

As distâncias foram obtidas através do módulo métrica disponível na biblioteca do scikit-learn (Pedregosa *et al.*, 2011), e em seguida para melhor visualização da distância euclidiana entre as 45 cultivares de porta-enxertos foi gerada uma matriz de similaridade também disponível no módulo métrica da biblioteca do scikit-learn (Pedregosa *et al.*, 2011).

3 Resultados e Discussão

Os coeficiente de correlação de Pearson observados na (Figura 3), indicam que houveram 39 casos com forte correlação entre as 61 variáveis avaliadas neste

estudo, sendo destas 13 com correlação negativa próximo a -1, e 26 com correlação positiva próxima a 1.

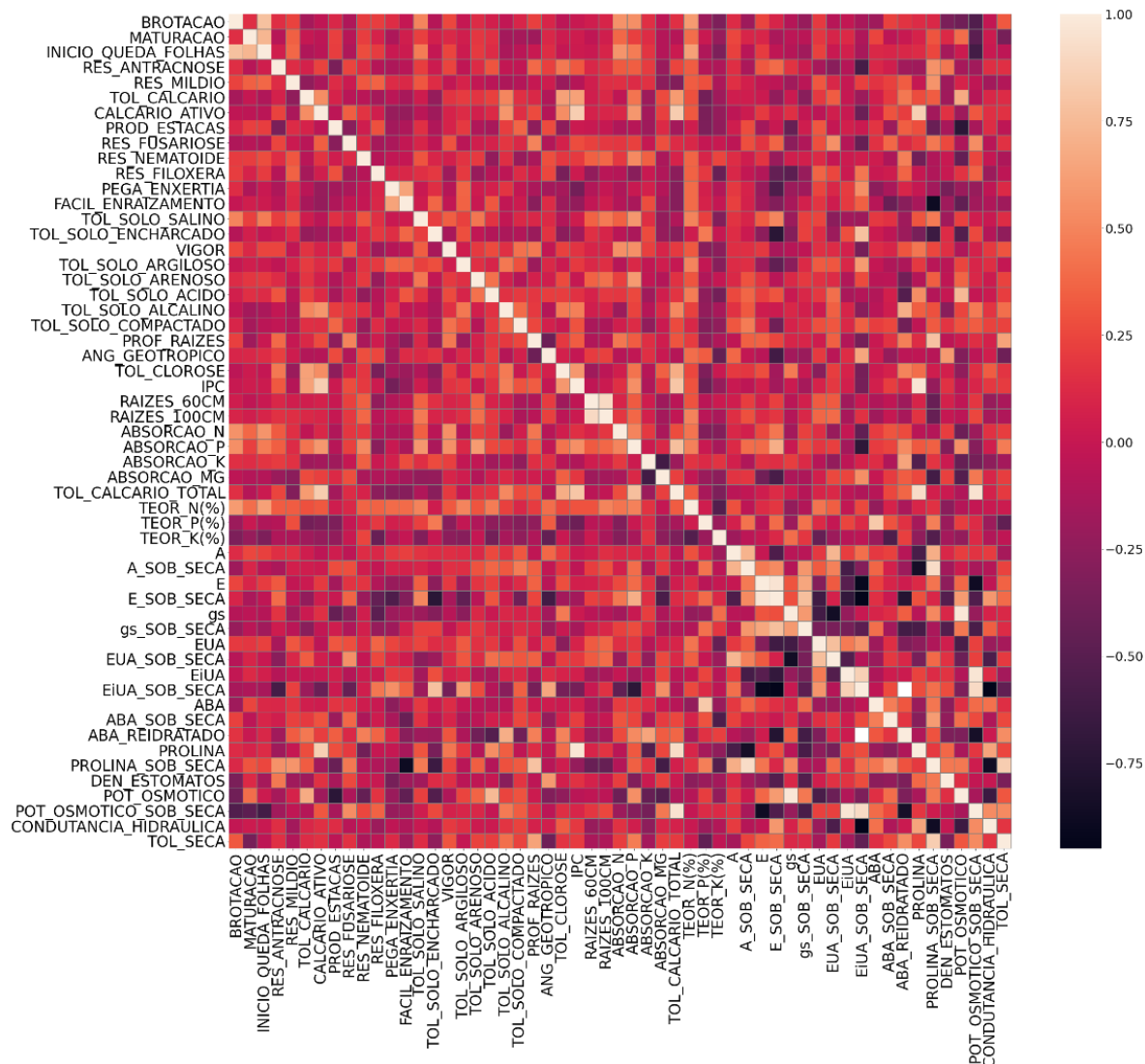


Figura 3 – Matriz de correlação entre as variáveis, onde as cores escuras representam a correlação negativa e as cores claras a correlação positiva. Fonte: Verslype *et al.* (2021).

A respeito disso, Ferreira *et al.* (2011) afirma que a análise de correlação é interessante por permitir explicar e determinar o grau de relação entre as variáveis. Desse modo, para Forsyth (2018), os coeficientes próximos a 1 e -1 possibilitam obter uma boa previsão de uma variável a partir da outra. Para tal afirmação podemos destacar entre as correlações fortes as variáveis distribuição de raízes até 100 cm e 60 cm, que apresentaram correlação positiva de ($r=0,91$), ou seja, a correlação alta pode indicar variáveis redundantes, indicando não haver necessidade de avaliar até a profundidade de 100 cm, o que pode evitar gastos dispendiosos e a redução significativa do tempo da análise.

Outra correlação que merece destaque, foi entre as variáveis brotação e maturação por apresentarem respectivamente ($r=0,78$) e ($r=0,73$) de correlação positiva com a variável início de queda de folhas, demonstrando que apenas a avaliação da variável início de queda de folhas seria suficiente.

Podemos destacar também a variável de teor de prolina sob estresse hídrico tem alta correlação com a tolerância à seca ($r=0,84$), pois a prolina poderia ser usada como critério de seleção indireta, quando se busca genótipos tolerantes à seca, visto que a tolerância à seca é uma característica poligênica de difícil identificação e seleção de materiais superiores em programas de melhoramento genético de plantas (Marguerit *et al.*, 2012; Serra *et al.*, 2014; Cantu e Walker, 2019). A alta correlação da prolina com a tolerância à seca, poderia ser explicada por ser a prolina considerada uma substância osmoticamente ativa, que é acumulada em níveis elevados no citoplasma quando a videira encontra-se em períodos de restrição hídrica, possibilitando assim manter em equilíbrio o potencial hídrico dentro da célula vegetal e a pressão de turgor (Taiz e Zieger, 2004; Serra *et al.*, 2014; Keller, 2015).

Desse modo, analisando-se as variáveis com correlação alta, (Figura 4), foram desconsideradas 14 variáveis para a análise de agrupamento, tais como, E sob estresse hídrico, A - sob estresse hídrico, EUA - sob estresse hídrico, EiUA - sob estresse hídrico, a prolina sob condições ideais e sob estresse hídrico, o teor de ABA sob condições ideais, o potencial osmótico sob condições ideais e sob estresse hídrico, o IPC, a tolerância ao calcário total, a distribuição de raízes até 100 cm, a brotação de gemas e a maturação do cacho.

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

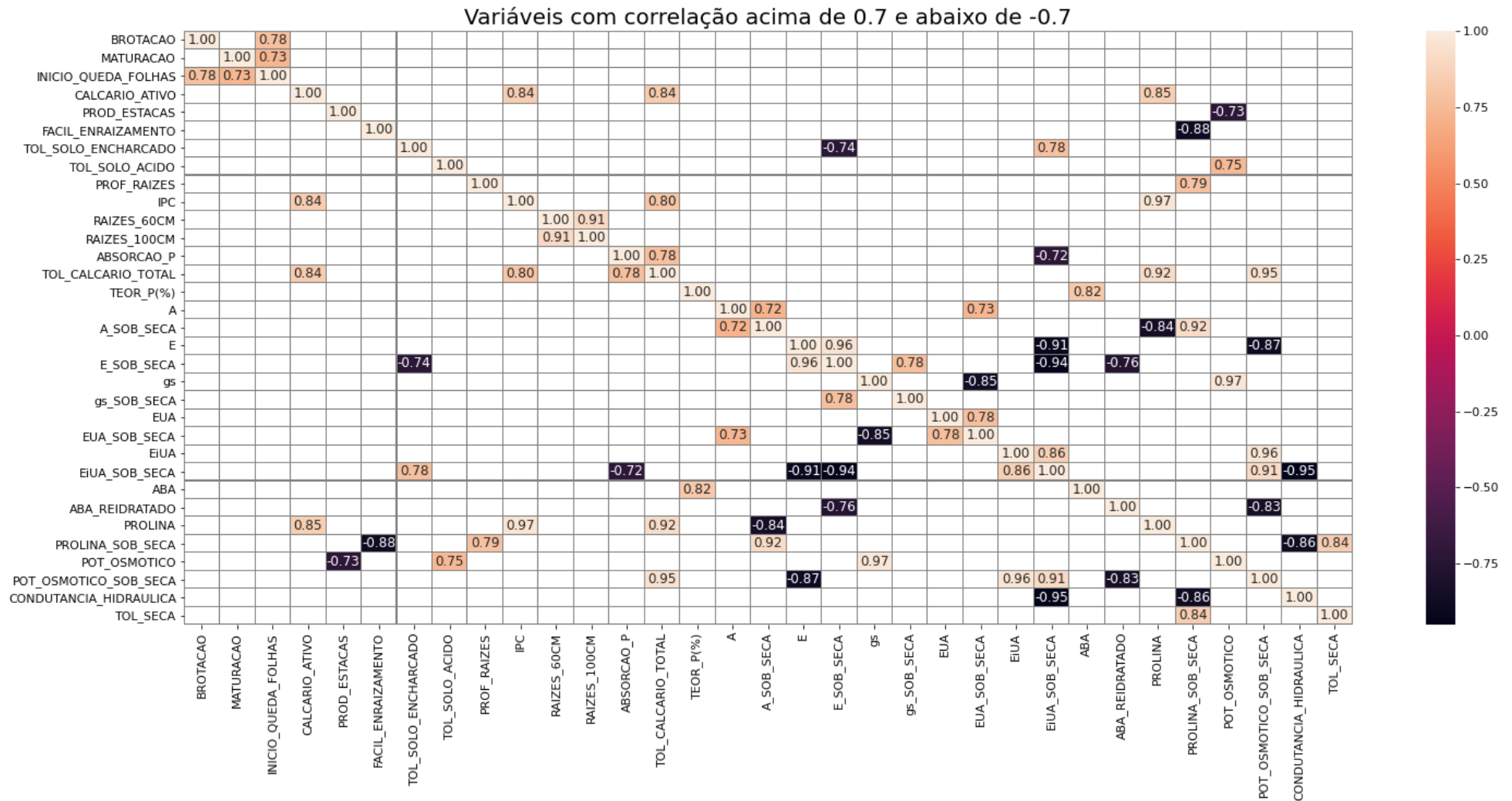


Figura 4 - Variáveis com forte correlação positiva e negativa. Fonte: Verslype *et al.* (2021).

A curva obtida pelo método do cotovelo indicou a adoção de k igual a cinco, ou seja, que sejam adotados cinco clusters na análise de agrupamento de k -means, o que foi confirmado através do algoritmo kneedle (Figura 5). Apesar dos algoritmos indicarem um valor de k , o especialista da área é quem tem o poder de escolha sobre o número mais adequado, porém em nosso estudo, optamos por manter o valor indicado pela curva obtida no método do cotovelo e o kneedle.

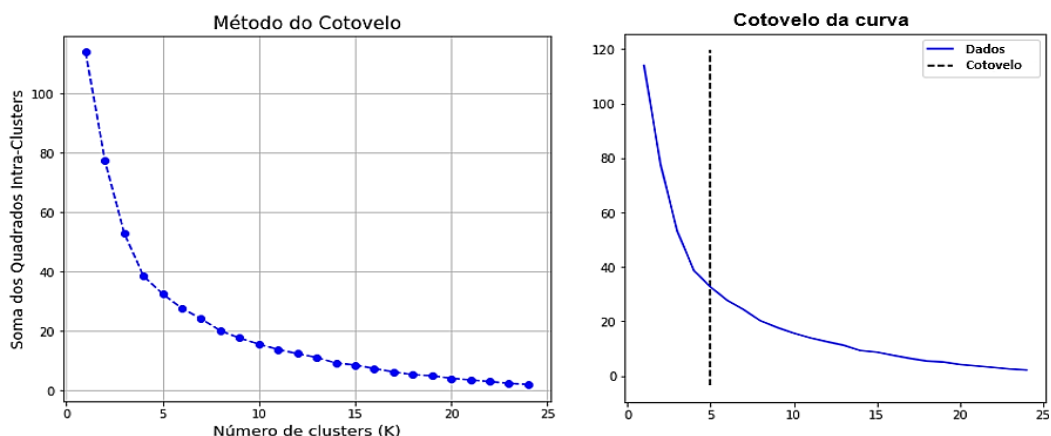


Figura 5 - Método do Cotovelo para determinar o número de clusters e o algoritmo Kneedle indicando o número ideal de clusters. Fonte: Verslype *et al.* (2021).

Aplicando-se a PCA e algoritmo de agrupamento K -means, (Figura 6), observou-se que todas as 45 cultivares de porta-enxertos de videira avaliadas foram agrupadas em cinco grupos heteróticos distintos, demonstrando que existe uma variação significativa entre as cultivares para todas as variáveis estudadas.

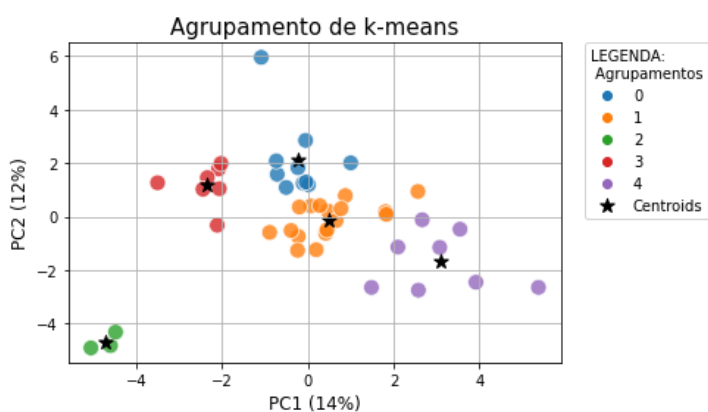


Figura 6 - Agrupamentos formados pelo k -means com base no primeiro e segundo componente principal da PCA. No qual os pontos representam os indivíduos dentro dos agrupamentos representados pelas diferentes cores, e em preto os centróides de cada cluster. Fonte: Verslype *et al.* (2021).

Com base nos resultados da análise de k -means, observamos na (Tabela 2), que o agrupamento 1 foi o maior com 17 cultivares, seguido pelo agrupamento 0 com 10 cultivares, agrupamento 4 com 8 cultivares, agrupamento 3 com 7 cultivares, e

finalmente o agrupamento 2 com apenas 3 cultivares. Os resultados corroboram em parte com outras pesquisas, como aqueles apresentados por Riaz *et al.* (2019), em que investigou a origem genética de 36 cultivares de porta-enxerto através de marcadores moleculares, e afirmam que a cultivar Schwarzmans foi agrupada intimamente com Riparia Gloire, assim como, foi observado em nosso estudo que ambas foram classificadas pelo k-means no agrupamento 4.

Tabela 2 - Distribuição das cultivares de porta-enxertos de videira (*Vitis* spp) dentro dos cinco agrupamentos.

Agrupamento	Número de cultivares	Lista de cultivares por agrupamento
0	10	110 R, 1103 P, 1202 C, 157-11 C, 1613 C, 26 G, 41 B MGt, 93-5 C, 99 R, Harmony
1	17	VR 039-16, 106-8 MGt, 125 AA, 161-49 C, 196-17 Castel, 216-3 Castel, 3306 C, 34 EM, 420 A MGt, 5 BB, 62-66 C, 8 B, Ganzin 1, Golia, Gravesac, Rupestris du Lot, Sori
2	3	IAC 313, IAC 572, IAC 766
3	7	VR 043-43, 140 Ru, Dogridge, Fercal, Freedom, Salt Creek, <i>Vitis Champini</i>
4	8	101-14 MGt, 1616 C, 3309 C, 44-53 M, 5 C, Riparia Gloire, Schwarzmans, SO4

Destacam-se também os agrupamentos que compreendem o maior número de cultivares com alta tolerância ao déficit hídrico, tais como o agrupamento 0 com cinco cultivares (110 R, 1103 P, 1202 C, 41 B MGt e 99 R), seguido pelo agrupamento 1 (106-8 MGt, 196-17 Castel, 62-66 C e Ganzin 1) e agrupamento 3 (140 Ru, Dogridge, Salt Creek, e *Vitis Champini*) ambas com quatro cultivares cada.

Com base nas características dos cinco agrupamentos, destacam-se na (Figura 7), as variáveis pegamento na enxertia, quantidade de dias para início de queda das folhas e o vigor induzido à copa, por apresentarem uma maior distinção e variabilidade entre os cinco agrupamentos. A respeito disso, observa-se que o agrupamento 2, formado pelas cultivares IAC 313, IAC 572 e IAC 766, apresentou os maiores valores e a menor variabilidade para as três variáveis em questão. Entre as possíveis justificativas pode-se afirmar que as três cultivares foram desenvolvidas pelo Instituto Agrônomo de Campinas, em São Paulo, Brasil, com o intuito de serem empregadas em regiões tropicais produtoras de uvas de mesa (Pommer, 2000; Nachtigal, 2003). Por conta disso, apresentam características similares, como indução de vigor elevado à copa, assim como, bons índices de pegamento na enxertia e enraizamento de estacas (Leão *et al.* 2009). Além disso, necessitam entre 261 e 293

dias para que haja o início da queda de folhas na região de Bento Gonçalves, Sul do Brasil, tratando-se de cultivares classificadas como precoce segundo Camargo *et al.* (2017). Nesse sentido, essas cultivares podem ser uma escolha interessante como genitor, quando buscam-se cultivares vigorosas com um ciclo vegetativo precoce.

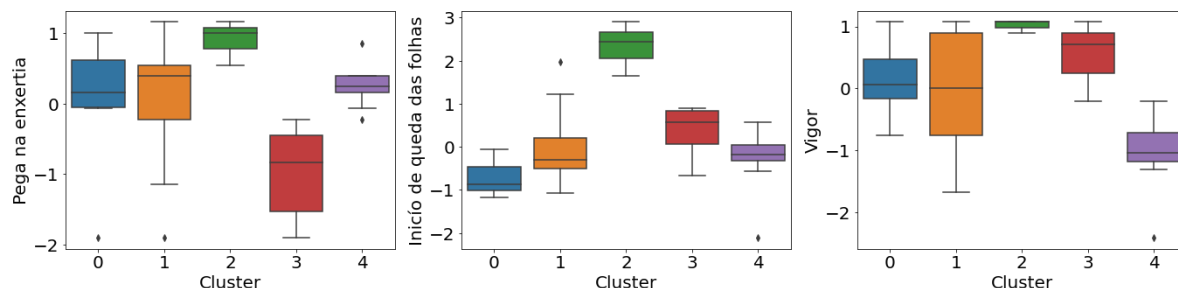


Figura 7 - Distribuição das características agrônômicas dentro de cada cluster. Fonte: Verslype *et al.* (2021).

Por sua vez, o agrupamento 1 formado por 17 cultivares, apresentou uma maior variabilidade para as três variáveis em questão, quando comparado com os demais agrupamentos. Esses resultados demonstram que as 17 cultivares apresentam uma maior diversidade genética, visto que tratam-se de cultivares originadas de seis espécies distintas de acordo com a Classificação do Catálogo Internacional de Variedades de *Vitis* – VIVC (Maul *et al.*, 2020). A variável vigor induzido à copa de distribuição simétrica, variou entre médio-baixo para 420 A MGt e alto nas cultivares VR 039-16, 216-3 Castel, Ganzin 1, Golia e Rupestris du Lot. Enquanto a variável pegamento na enxertia, de distribuição assimétrica positiva, variou entre alto para Ganzin 1 e baixo para uma única cultivar a VR 039-16. Esta baixa taxa de pegamento na enxertia pode ser explanada através da cultivar VR 039-16 ser um híbrido das espécies *V. vinifera* L. ($2n=38$) e *M. rotundifolia* Michaux ($2n=40$), que diferem no número de cromossomos e em algumas características morfológicas (Cantu e Walker, 2019). Segundo Albuquerque (2003), a região do câmbio desempenha um papel importante no êxito das enxertias. Neste sentido, possíveis diferenças no tamanho de seus vasos condutores, dificultam a soldadura entre a copa e o porta-enxerto. Além disso, foi observado que foram necessários entre 216 e 277 dias para iniciar a queda de folhas na região de Bento Gonçalves-RS, todas as cultivares apresentaram um ciclo vegetativo intermediário, com exceção da cultivar VR 039-16 que classificada como tardia (Camargo *et al.*, 2017).

No agrupamento zero, observamos uma distribuição assimétrica negativa para as três variáveis em questão, no qual o vigor variou entre alto para apenas quatro

cultivares 1202 C, 26 G, 41 B MGt e 99 R e baixo para 93-5 C, provocando a assimetria observada. Da mesma forma que foi observado para a variável pega na enxertia que variou entre alta para apenas duas cultivares 1202 C e 26 G e baixa para 157-11 C. Assim como, o início de queda de folhas variou entre 202 e 241 na região de Bento Gonçalves-RS, sendo classificadas com ciclo vegetativo intermediário (Camargo *et al.*, 2017).

No agrupamento 3, formado por 7 cultivares, as três variáveis em questão apresentaram distribuição assimétrica positiva. Para tal afirmação podemos destacar a variável vigor classificada como alta para as cultivares VR 043-43, 140 Ru, Dogridge, *Vitis Champini* e apenas três classificadas com vigor alto-médio (Fercal, Freedom, Salt Creek). Enquanto para variável pega na enxertia as cultivares VR 043-43 e Dogridge apresentaram baixa taxa de pegamento e as demais variaram entre médio e alta taxa de pegamento. Da mesma forma, foi observado para o início de queda de folhas variou entre 224 e 254 na região de Bento Gonçalves-RS, no qual apenas a cultivar VR 043-43 foi classificada como tardio e as demais como intermediário (Camargo *et al.*, 2017).

O agrupamento 4, composto por 8 cultivares, apresentou uma distribuição assimétrica negativa para as variáveis pegamento na enxertia e vigor, e simétrica para início de queda de folhas. Isto pode ser explicado, devido a apenas as cultivares 1616 C, 5 C e SO4 terem apresentado vigor alto-médio, e as demais terem variado entre médio e baixo vigor induzido à copa. Para a variável pegamento na enxertia apenas a cultivar 44-53 M foi classificada como alta e as demais foram classificadas entre média e alta-média. Todas as cultivares do agrupamento 4 apresentaram um ciclo vegetativo intermediário, entre 196 e 259 dias para iniciar a queda de folhas na região de Bento Gonçalves-RS, com exceção da cultivar Riparia Gloire classificada como precoce (Camargo *et al.*, 2017).

De acordo com Ertel (2017) a distância euclidiana indica o quanto dois genótipos são semelhantes, desta forma distâncias pequenas indicam maior similaridade genética e maiores distâncias maior divergência. Desse modo, Kanavi *et al.* (2020) e Nandini *et al.* (2020) comentam em seus estudos, a eficiência e a importância do algoritmo K-means como uma ferramenta útil na identificação de linhagens parentais promissoras em programas de melhoramento genético de plantas, através da semelhança de suas características e a determinação da distância

genéticas dos genótipos. De acordo com a matriz de similaridade obtida (Figura 8), e o sexo da flor das cultivares (Tabela 1), destacaram-se três cultivares de porta-enxerto, devido ao maior número de opções para se realizar cruzamentos, que foram as cultivares Harmony, 41 B MGt e VR 039-16, com respectivamente nove, oito e sete opções de cruzamento cada. Desse modo, na (Tabela 3), são indicados 37 cruzamentos viáveis, dentre as 45 cultivares de porta-enxertos de videira avaliadas, sendo as cultivares Harmony de média tolerância ao déficit hídrico (Sunridge, 2020) e 41 B MGt de alta tolerância ao déficit hídrico (ATVB, 2013) as cultivares com maior potencial para obter porta-enxertos tolerantes ao déficit hídrico, por conta do maior número de opções de se realizar cruzamentos, seguido pela cultivar Salt Creek de alta tolerância ao déficit hídrico (Wine Australia, 2016; Villa, 2018).

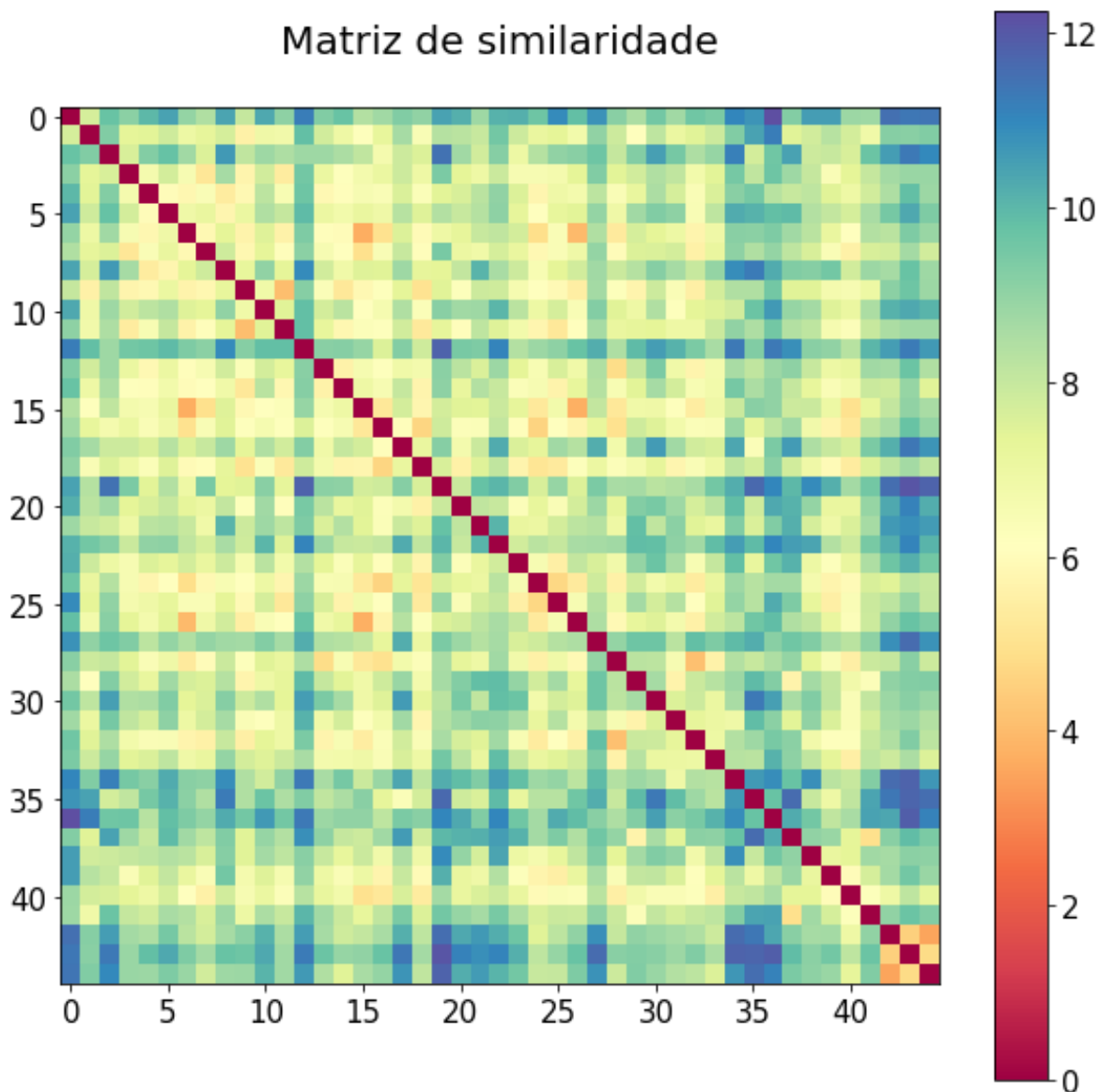


Figura 8 - Matriz de similaridade obtida através da distância euclidiana entre as 45 cultivares de porta-enxertos de videira, onde as cores nos tons de vermelho representam genótipos

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

geneticamente próximos e nos tons de azul genótipos geneticamente distantes. Fonte: Verslype *et al.* (2021).

Tabela 3 - Cruzamentos viáveis entre porta-enxertos sugeridos com base na matriz de similaridade genética.

Cultivar	Flor	Tolerância ao déficit hídrico	Cluster	Cultivar	Flor	Tolerância ao déficit hídrico	Cluster	
VR 039-16	Feminina	Baixa ¹	1	SO4	Masculina	Baixa ^{2;3;4}	4	
				1616 C	Masculina	Baixa ¹		
				Schwarzmann	Masculina	Baixa ^{3;4}		
				99 R	Masculina	Alta ⁴		0
				IAC 313	Masculina	- *		2
				IAC 572	Masculina	- *		
IAC 766	Masculina	- *						
41 B MGt	Monóica	Alta ²	0	101-14 MGt	Feminina	Baixa ^{1;3;4}	4	
				1616 C	Masculina	Baixa ¹		
				Riparia Gloire	Masculina	Baixa ^{1;2;3}		
				Schwarzmann	Masculina	Baixa ^{3;4}		
				Rupestris du lot	Masculina	Média ³		1
				IAC 313	Masculina	- *		2
IAC 572	Masculina	- *						
IAC 766	Masculina	- *						
Harmony	Feminina	Média ¹	0	3309 C	Masculina	Baixa ^{3;4}	4	
				Riparia Gloire	Masculina	Baixa ^{1;2;3}		
				Schwarzmann	Masculina	Baixa ^{3;4}		
				1616 C	Masculina	Baixa ¹		
				Rupestris du lot	Masculina	Média ³		1
				140 Ru	Masculina	Alta ^{1;3}		3
Fercal	Feminina	Média ^{1;3;4}	3	IAC 313	Masculina	- *	2	
				IAC 572	Masculina	- *		
				IAC 766	Masculina	- *		
				3309 C	Masculina	Baixa ^{3;4}		
161-49 C	Feminina	Baixa ^{1;3}	1	Riparia Gloire	Masculina	Baixa ^{1;2;3}	4	
				IAC 572	Masculina	- *		
IAC 572	Masculina	- *	2	101-14 MGt	Feminina	Baixa ^{1;3;4}	4	
Salt Creek	Feminina	Alta ^{3;4}	3	34 EM	Masculina	Baixa ³	1	
				3306 C	Masculina	Média ⁴		
				3309 C	Masculina	Baixa ^{3;4}		
				1616 C	Masculina	Baixa ¹		
5 BB	Feminina	Média ¹	1	Riparia Gloire	Masculina	Baixa ^{1;2;3}	4	
				IAC 572	Masculina	- *		
				3309 C	Masculina	Baixa ^{3;4}		
140 Ru	Masculina	Alta ^{1;3;4}	3	1616 C	Masculina	Baixa ¹	4	
				101-14 MGt	Feminina	Baixa ^{1;3;4}		

*Informação desconhecida; ¹Sunridge (2020); ²ATVB (2013); ³Villa (2018); ⁴Wine Australia (2016).

Nesse sentido, o melhoramento genético de porta-enxertos de videira para tolerância ao estresse hídrico são de extrema importância, no contexto das mudanças climáticas, por ser considerado o meio mais sustentável de adaptação e proteção da cultivar copa contra danos ocasionados pelo estresse hídrico à longo prazo e manutenção da produtividade (Marguerit *et al.*, 2012; Serra *et al.*, 2014; Ollat *et al.*,

2016; Zhang *et al.*, 2016). Em vista disso, os estudos de diversidade e formação de grupos heteróticos são fundamentais em programas de melhoramento para a obtenção de novas cultivares tolerantes aos estresses bióticos e abióticos, pois possibilitam selecionar os genitores contrastantes com maior potencial para determinada característica e consequente identificação dos melhores cruzamentos (Kanavi *et al.*, 2020; Nandini *et al.*, 2020).

4 Conclusão

O algoritmo K-means mostrou-se uma ferramenta eficiente na identificação de cinco grupos heteróticos entre as quarenta e cinco cultivares de porta-enxertos de videira avaliadas, com vista a identificar os melhores cruzamentos para obtenção de porta-enxertos tolerantes ao déficit hídrico. Dentre as 37 opções de cruzamentos viáveis as cultivares Harmony e 41 B MGt apresentaram maior potencial para obter porta-enxertos tolerantes ao déficit hídrico. Como estudos futuros na área, sugerimos a verificação experimental em campo dos resultados encontrados nesse trabalho.

As metodologias de análise de componentes principais e K-means são estratégias acessíveis e úteis ao melhorista, possibilitando alcançar um resultado de fácil entendimento, que podem contribuir no melhoramento genético de porta enxertos de videira tolerantes ao déficit hídrico.

Agradecimentos

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão de bolsa de mestrado à primeira autora.

Referências bibliográficas

Albuquerque TCS (2003) Videira (*Vitis* sp). In: Castro PRC and Kluge RA (Coord.). **Ecofisiologia de fruteiras:** abacateiro, aceroleira, macieira, pereira e videira. Piracicaba: Agronômica Ceres, p.93-119.

ATVB – Association Technique Viticole de Bourgogne (2013) **Caractéristiques des porte-greffes et greffons utilisés en côte-d’or.** Disponível em: <[http://www.atvb-bourgogne.com/sites/default/files/catalogue%20ATV Bsept2015%20\(2\).pdf](http://www.atvb-bourgogne.com/sites/default/files/catalogue%20ATV%20sept2015%20(2).pdf)>. Acesso em agosto 2020.

Audeguin L, Bécart V, Besnard E, Boursiquot JM, Ciccolini G, Dedet S, Delorme G, Follet J, Froelhy A, Garcin L, Gonin N, Lacombe T, Laffargue F, Marchal C, Pesato M,

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

Sereno C, Uriel G, Wentzel M and Yobregat O (2020) **PI@ntGrape**: le catalogue des vignes cultivées en France. Montpellier, France: IFV – INRA – Montpellier SupAgro, 2009-2020. Disponível em: <<http://plantgrape.plantnet-project.org/en/porte-greffe/>>. Acesso em: 23 set. 2020.

Beiki AH (2019) Molecular Predicting Drought Tolerance in Maize Inbred Lines by Machine Learning Approaches. **BioRxiv**.

Bruckner CH and Albuquerque AS (2005) Melhoramento de fruteiras. In: Borém A (ed.). **Melhoramento de espécies cultivadas**. Viçosa: Ed. UFV, p. 813-863.

Camargo UA, Ritschel PS, Maia JDG, Quecini V, Machado CAE, Bosco DD, Sinski I, Comachio V and Zilio R (2017) **Banco Ativo de Gesmoplasma de Uva**. Embrapa Uva e Vinho - Bento Gonçalves (RS). Disponível em: <<https://www.embrapa.br/en/uva-e-vinho/banco-ativo-de-germoplasma-de-uva/pesquisa>>. Acesso em: 25 jun. 2020.

Cantu D and Walker MA (eds.). **The grape genome**. Switzerland: Springer Nature Switzerland AG, p. 25-38.

Dougherty PH (2012) Introduction to the geographical study of viticulture and wine production. In: Dougherty, Percy H (ed.). **The Geography of Wine: Regions, Terroir and Technique**. Netherlands: Ed. Springer, p.3-36

Ertel E (2017) **Introduction to artificial intelligence**. Springer International Publishing AG, Switzerland, 365p.

Etminan A, Pour-Aboughadareh A, Mohammadi R, Shooshtari L, Yousefiazarkhanian M, and Moradkhani H (2019) Determining the best drought tolerance indices using artificial neural network (ANN): Insight into application of intelligent agriculture in agronomy and plant breeding. **Cereal Research Communications**, **47**: 170-181.

FAOSAT - **Food and Agriculture Organization Corporate Statistical Database** (2020). **Crops statistics**: data about production quantities of grapes in 2019. Disponível em: <<http://www.Fao.org/faostat/en/#data/QC/visualize>>. Acesso em: 25 dez. 2020.

Ferreira PV (2011) **Estatística experimental**. Maceió. Ed.UFAL. 559p.

Forsyth D (2018). **Probability and statistics for computer science** (pp. 36-42). Springer International Publishing.

Hernandes JL, Martins FP and Pedro Júnior MJ (2011) **Uso de porta-enxertos: Tecnologia simples e fundamental na cultura da videira**. Jundiaí, SP: Instituto Agrônomo de Campinas, 19p.

IBGE - **Instituto Brasileiro de Geografia e Estatística** (2020). Levantamento Sistemático da Produção Agrícola. Disponível em: <<https://sidra.ibge.gov.br/tabela/1618#resultado>>. Acesso em: 06 jan. 2020.

Kanavi MSP, Koler P, Somu G and Marappa N (2020) Genetic Diversity Study through K-means Clustering in Germplasm Accessions of Green gram [*Vigna radiata* (L.)] Under Drought Condition. **International Journal of Bio-resource and Stress Management** **18**: 138-147.

Keller M (2015) **The science of grapevines: anatomy and physiology**. 2nd ed., Academic press, San Diego, 509p.

Leão PCS and Borges RME (2009) **Melhoramento genético da videira**. Petrolina: Embrapa Semiárido, 61p.

Leão PCS, Soares JM and Rodrigues BL (2009) Principais Cultivares. In: Soares JM and Leão PCS (eds.). **A vitivinicultura no Semiárido Brasileiro**. Brasília, DF: Embrapa Informação Tecnológica, Petrolina: Embrapa Semi-Árido, p.151-214.

Marguerit E, Brendel O, Lebon E, Van Leeuwen C and Ollat N (2012) Rootstock control of scion transpiration and its acclimation to water deficit are controlled by different genes. **New Phytologist** **194**: 416-429.

Maul E, Töpfer R, Röckel F, Brühl U, Hundemer M, Mahler-Ries A, Walk M, Kecke S, Wolck A and Ganesch A (2020) **Vitis International Variety Catalogue (VIVC)**. Julius Kühn-Institut - Federal Research Centre for Cultivated Plants (JKI), Institute for Grapevine Breeding - Geilweilerhof (ZR). Disponível em: <<http://www.vivc.de/>>. Acesso em: 20 jul. 2020.

McKinney W. Data Structures for Statistical Computing in Python. In Stéfan van der Walt and Jarrod Millman, editors, **Proceedings of the 9th Python in Science Conference**, pages 56 – 61, 2010.

Nachtigal JC (2003) **Produção de mudas de videira em regiões tropicais e subtropicais do Brasil**. Embrapa Uva e Vinho, Bento Gonçalves, 22 p.

Nandini C, Bhat S, Saritha HS, Pandey CD, Sushil Pandey P, Bai L and Gowda J (2020) Characterization of barnyard millet (*Echinochloa frumentaceae* (Roxb.) Link) germplasm for quantitative traits to enhance its utilization. **Electronic Journal of Plant Breeding** **11**: 1066-1072.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M and Duchesnay E (2011) Scikit-learn: Machine Learning in Python. **Journal of Machine Learning Research** **12**: 2825-2830.

Pommer CV (2000) Cultivares de uva produzidos ou introduzidos pelo IAC. In: Boliani AC and Corrêa L de S (eds.) **Proceedings of the Simpósio Brasileiro sobre Uvas de Mesa 1**. Ilha Solteira, p. 51-67.

OIV - **Organización Internacional de la Viña y el Vino** (2020). Estadísticas: informe sobre la situación de la viticultura em 2019. Disponível em: <<http://www.oiv.int/es/statistiques/recherche>>. Acesso em: 25 dez. 2020.

Ollat N, Peccoux A, Papura D, Esmenjaud D, Marguerit E, Tandonnet JP, Bordenave L, Cookson SJ, Barrieu F, Rosssdeutsch L, Lecourt J, Lauvergeat V, Vivin P, Bert PF and Delrot S (2016) Rootstocks as a component of adaptation to environment. In Gerós H, Chaves MM, Gil HM and Delrot S (eds) **Grapevine in a changing environment: a molecular and ecophysiological perspective**. John Wiley & Sons, UK, p. 102-152.

Riaz S, Pap D, Uretsky J, Laucou V, Boursiquot JM, Kocsis L and Walker MA (2019) Genetic diversity and parentage analysis of grape rootstocks. **Theoretical and Applied Genetics** **132**: 1847–1860. Disponível em: <<https://doi.org/10.1007/s00122-019-03320-5>>. Acesso em: 29 mar. 2020.

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

Satopaa, V., Albrecht, J., Irwin, D., & Raghavan, B. (2011, June). Finding a "kneedle" in a haystack: Detecting knee points in system behavior. In 2011 31st **international conference on distributed computing systems workshops** (pp. 166-171). IEEE. Disponível em: <<https://raghavan.usc.edu/papers/kneedle-simplex11.pdf>>. Acesso em 16 mar. 2021.

Serra I, Strever A, Myburgh PA and Deloire A (2014) The interaction between rootstocks and cultivars (*Vitis vinifera* L.) to enhance drought tolerance in grapevine. **Australian Journal of Grape and Wine Research** 20: 1-14.

Skansi, S. (2018). **Introduction to Deep Learning**: from logical calculus to artificial intelligence. Springer. 196p.

Sunridge Nurseries (2020) **Rootstock chart**. Bakersfield, California. Disponível em <<https://www.sunridgenurseries.com/index.php/clonalselections/rootstock-chart>>. Acesso em: 25 Ago. 2020.

Taiz L and Zieger E (2004) **Fisiologia vegetal**. Artmed, Porto Alegre, 723 p.

Villa P (2018) **Cultivar la vid**. Editorial de Vecchi, USA, 160p.

Wine Australia Corporation (2016) **Grapevine rootstock selector tool**. Austrália. Disponível em: <<https://grapevine-rootstock.com/>>. Acesso em: 25 Ago. 2020.

Zhang L, Marguerit E, Rossdeutsch L, Ollat N and Gambetta GA (2016) The influence of grapevine rootstocks on scion growth and drought resistance. **Theoretical and Experimental Plant Physiology** 28: 143-157.

CAPÍTULO IV

CLASSIFICAÇÃO DE PORTA-ENXERTOS DE VIDEIRA TOLERANTES AO DÉFICIT HÍDRICO ATRAVÉS DE APRENDIZAGEM DE MÁQUINA

Classificação de porta-enxertos de videira tolerantes ao déficit hídrico através de aprendizagem de máquina

Drought tolerance classification of grapevine rootstock by machine learning

Nina Iris Verslype^{a*}; Rosimar dos Santos Musser^a; André Câmara Alves do Nascimento^b; Raphael Miller de Souza Caldas^a; Luiza Suely Semen Martins^d; Patrícia Coelho de Souza Leão^c

^a Departamento de Agronomia, Universidade Federal Rural de Pernambuco, rua Dom Manoel de Medeiros, s/n - Dois Irmãos, Recife, 52171-900, Brasil

^b Departamento de Computação, Universidade Federal Rural de Pernambuco, rua Dom Manoel de Medeiros, s/n - Dois Irmãos, Recife, 52171-900, Brasil

^c Embrapa Semiárido, Zona Rural, s/n, Petrolina, 56302-970, Brasil.

^d Departamento de Biologia, Universidade Federal Rural de Pernambuco, rua Dom Manoel de Medeiros, s/n - Dois Irmãos, Recife, 52171-900, Brasil

* Autor correspondente.

E-mail: nina.iris.v@gmail.com (N. I. Verslype).

Artigo a ser enviado para publicação na “Computers and Electronics in Agriculture”.

RESUMO

Os algoritmos de aprendizagem de máquina vem sendo cada vez mais empregado em diversas áreas de estudo, dentre as quais se inclui o melhoramento genético de plantas, pois podem auxiliar na descoberta de padrões significativos e/ou na predição de variáveis, facilitando assim as avaliações e trazendo resultados precisos. A identificação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico vem sendo cada vez mais demandadas especialmente diante dos cenários de mudanças climáticas globais, no entanto, são de difícil obtenção por ser a videira uma espécie perene e tratando-se de uma característica poligênica, de herança complexa. Nesse sentido, o objetivo deste trabalho foi comparar o desempenho de seis algoritmos de aprendizado supervisionado, com base em um conjunto de dados formado por 45 cultivares de porta-enxertos de videira. E em seguida predizer o grau de tolerância ao déficit hídrico dos porta-enxertos brasileiros IAC 313, IAC 572 e IAC 766, cuja informação na literatura é desconhecida, através do algoritmo com melhor desempenho. Os resultados obtidos indicaram que o melhor algoritmo foi o *Random Forest* e pode-se predizer que as cultivares IAC 313 e IAC 766 apresentam alta tolerância ao déficit hídrico e o IAC 572 baixa tolerância.

Palavras-chave: Estresse hídrico, *Vitis* spp., mudanças climáticas, aprendizado supervisionado, algoritmo.

ABSTRACT

Machine Learning algorithms are increasingly used in several areas of study, such as plant breeding, as they can assist in the identification of significant patterns or even in the prediction of variables, thus facilitating the evaluations and hence accurate results. The identification and selection of drought-tolerant grapevine rootstock (*Vitis* spp.) have been increasingly in demand because of global climate change scenarios. But, as a consequence of their polygenic characteristic of complex inheritance and grapevine is a perennial species, they are complicated to obtain. For this reason, this study's purpose was to compare the performance of six machine learning models on the prediction of drought tolerance class of grapevine rootstock cultivars. In addition, using the best performing model, it was predicted the drought tolerance class of IAC 313, IAC 572, and IAC 766 Brazilian grapevine rootstock cultivars, whose information still unknown in the literature. Our results indicated Random Forest as the best model, which predicted high tolerance for IAC 313 and IAC 766 cultivars, and a low tolerance for IAC 572.

Key words: Hydrical stress, *Vitis* spp., climatic changes, supervised learning, algorithm.

1 INTRODUÇÃO

O clima desempenha importante papel na vitivinicultura e encontra-se intimamente ligado a geografia do vinho, sendo considerado o aspecto mais crítico no amadurecimento e qualidade final das uvas, para produzir um determinado estilo de vinho, além da sua interferência direta na escolha de cultivares, local de plantio, potencial vegetativo, comportamento fitossanitário, rendimento, qualidade final das uvas e inclusive no manejo e práticas culturais adotadas (Moura *et al.*, 2009; Jones *et al.*, 2012; Fraga *et al.*, 2013; Jones, 2016).

As mudanças climáticas tornam a produção agrícola extremamente vulnerável (Jones, 2016), principalmente no quesito de disponibilidade hídrica, um dos constituintes mais importantes da vida (Kathpalia & Bhatla, 2018). Por conta disto, a seca é considerada um dos maiores estresses abióticos, ocasionando perdas consideráveis na produção agrícola mundial (Ashraf, 2010). Dessa forma, qualquer variação no clima pode vir a afetar significativamente o setor vitivinícola, tanto no rendimento quanto na qualidade das uvas, principalmente em condições futuras devido as mudanças climáticas (Jones & Webb, 2010; Marguerit *et al.*, 2011; Jones *et al.*, 2012; Fraga *et al.*, 2013). Portanto, a mitigação dos efeitos de mudanças climáticas tornou-se uma das principais demandas do setor vitivinícola, devido a videira ser uma

cultura de grande importância socioeconômica, considerada por Zhang *et al.* (2016), como uma das culturas perenes mais valiosas do mundo, devido ao alto valor agregado e versatilidade de seus produtos (Walker *et al.*, 2019).

As mudanças climáticas e a escassez crescente nos recursos hídricos enfatizam a importância do uso d'água mais eficiente no setor vitivinícola, pois a maioria das regiões vitivinícolas do mundo experimentam a seca em algum momento (Chaves *et al.*, 2010; Zarrouk *et al.*, 2016; Fort *et al.*, 2017). Embora várias abordagens possam ser empregadas para mitigar o problema da seca na viticultura mundial, o uso de porta-enxertos tolerantes à seca pode ser considerado como o elemento-chave de adaptação dos vinhedos (Marguerit *et al.*, 2011; Ollat *et al.*, 2016).

A tolerância ao estresse hídrico nas videiras, trata-se de uma característica poligênica, controlada por um grande número de genes (Ashraf, 2010; Marguerit *et al.*, 2012; Serra *et al.*, 2014; Dayer *et al.*, 2019), dificultando a identificação e seleção de genótipos mais promissores, por conta da interação do genótipo com o ambiente. Os objetivos de um programa de melhoramento de porta-enxertos devem ser bem definidos devido a necessidade de avaliação conjunta da tolerância à seca, com a facilidade de propagação, a compatibilidade com o enxerto, a longevidade da planta enxertada, a tolerância a fatores do meio, a resistência a pragas e doenças na região, além de outras características da planta e sua influência nos frutos.

O aprendizado de máquina é uma nova metodologia que vem sendo cada vez mais empregada em diversas áreas de estudo, como no melhoramento genético de plantas, por possibilitar identificar, prever e classificar genótipos conforme a necessidade do programa de melhoramento (Beiki, 2019; Etmnan *et al.*, 2019). Nesse sentido, o objetivo deste trabalho foi de prever o grau de tolerância à seca de três cultivares brasileiras de porta-enxertos de videira IAC 313, IAC 572 e IAC 766 utilizando algoritmos de aprendizado de máquina, com o intuito de indicar porta-enxertos tolerantes ao déficit hídrico para o cultivo da videira na região do Submédio do Vale do São Francisco, contribuindo economia da água de irrigação, e sustentabilidade da vitivinicultura no Nordeste brasileiro.

2 MATERIAL E MÉTODOS

2.1 Material vegetal

Foram avaliadas quarenta e cinco cultivares de porta-enxerto de videira, conforme pode ser visualizado na tabela 1 sua classe de tolerância ao déficit hídrico, nome comumente usado e a sua origem genética. A escolha destas cultivares foi realizada com base em três critérios, sendo o primeiro a inclusão das cultivares mais empregadas no Submédio do Vale do São Francisco, o segundo a sua disponibilidade entre os acessos no Banco Ativo de Germoplasma de videira da Embrapa Uva e Vinho (Camargo *et al.*, 2017) e o terceiro foi a disponibilidade de informação na literatura sobre o grau de tolerância ao déficit hídrico de cada porta-enxerto. Com a finalidade de treinar o algoritmo de aprendizado de máquina e possibilitar identificar o grau de tolerância ao déficit hídrico, dentre as 45 cultivares, estavam as cultivares IAC 313 (Tropical), IAC 572 (Jales) e IAC 766 (Campinas), para as quais não foram encontradas informações disponíveis na literatura.

Tabela 1 - Quarenta e cinco cultivares de porta-enxertos de videira, origem genética e grau de tolerância ao déficit hídrico.

Porta-enxerto	Origem genética	Tolerância ao déficit hídrico
VR 039-16	<i>Vitis vinifera</i> L. x <i>Muscadinia rotundifolia</i> Michaux ^{1;5}	Baixa ^{2;7}
VR 043-43	<i>Vitis vinifera</i> L. x <i>Muscadinia rotundifolia</i> Michaux ¹	Baixa ³
101-14 MGt	<i>Vitis riparia</i> Michaux x <i>Vitis rupestris</i> Scheele ^{1;3;5;6}	Baixa ^{2;3;4;5;8;10}
106-8 MGt	<i>Vitis riparia</i> Michaux x <i>Cordifolia rupestris</i> de grasset n ^o 1 ¹	Alta ⁶
110 R	<i>Vitis berlandieri</i> Planchon x <i>Vitis rupestris</i> Scheele ^{1;5;8}	Alta ^{2;4;5;6;7;8;13}
1103 P	<i>Vitis berlandieri</i> Planchon x <i>Vitis rupestris</i> Scheele ^{1;5;6;8}	Alta ^{2;3;5;7;8}
1202 C	<i>Vitis vinifera</i> L. x <i>Vitis rupestris</i> Scheele ¹	Alta ⁷
125 AA	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux ¹	Baixa ⁴
140 Ru	<i>Vitis berlandieri</i> Planchon x <i>Vitis rupestris</i> Scheele ^{1;6}	Alta ^{2;4;5;7;8;12;13}
157-11 C	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux ¹	Média ⁴
1613 C	Solonis (<i>Vitis riparia</i> Michaux x <i>Vitis longii</i> Prince) x Othello ^{1;5}	Média ⁹
161-49 C	<i>Vitis riparia</i> Michaux x <i>Vitis berlandieri</i> Planchon ^{1;6}	Baixa ⁵
1616 C	Solonis x <i>Vitis riparia</i> Michaux ¹	Baixa ²
196-17 CI	1203C x <i>Vitis riparia</i> Michaux ¹	Alta ^{6;7}
216-3 CI	1616C x <i>Vitis rupestris</i> Scheele ¹	Média ^{5;6;7}
26 G	<i>Vitis vinifera</i> L. x <i>Vitis riparia</i> Michaux ¹	Média ¹²
3306 C	<i>Vitis riparia</i> Michaux x <i>Vitis rupestris</i> Scheele ¹	Média ⁸
3309 C	<i>Vitis riparia</i> Michaux x <i>Vitis rupestris</i> Scheele ^{1;5;6}	Baixa ^{4;5;7;11;13}
34 EM	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux ¹	Baixa ^{4;5}
41 B MGt	<i>Vitis vinifera</i> L. x <i>Vitis berlandieri</i> Planchon ¹	Alta ^{5;10}
420 A MGt	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux ^{1;3;5;6;9}	Média ^{4;5;10;12}

44-53 M	<i>Vitis riparia</i> Michaux x Malegue 144 ¹	Alta ^{2;5;6;7}
5 BB	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux ^{1;5;6}	Média ^{2;4}
5 C	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux ^{1;5}	Baixa ^{5;8}
62-66 C	<i>Vitis vinifera</i> L. X <i>Vitis cordifolia</i> Michaux ¹	Alta ¹¹
8 B	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux ¹	Média ⁵
93-5 C	<i>Vitis vinifera</i> L. x <i>Vitis rupestris</i> Scheele ¹	Baixa ⁶
99 R	<i>Vitis berlandieri</i> Planchon x <i>Vitis rupestris</i> Scheele ^{1;5;6;8}	Alta ⁸
Ganzin 1	<i>Vitis vinifera</i> L. x <i>Vitis rupestris</i> Scheele ¹	Alta ⁷
Dogridge	<i>Vitis rupestris</i> Scheele x <i>Vitis candidans</i> Engelmann ^{1;5}	Alta ⁹
Fercal	<i>Vitis berlandieri</i> Planchon x 31R ¹	Média ^{2;5;7;8}
Freedom	Fresno 1613-59 x Dogridge ¹	Média ^{2;8}
Golia	Castel 156-12 x <i>Vitis rupestris</i> Scheele ¹	Baixa ⁴
Gravesac	161-49C x 3309C ¹	Média ^{6;7;8}
Harmony	1613C x Dogridge ¹	Média ²
Riparia Gloire	<i>Vitis riparia</i> Michaux ^{1;5;6}	Baixa ^{2;5;7;10;12;13}
Rupestris du lot	<i>Vitis rupestris</i> Scheele ^{1;5;6;8}	Média ^{5;7;9}
Salt Creek	<i>Vitis champinii</i> Planchon ^{1;5}	Alta ^{9;8}
Schwarzmann	<i>Vitis riparia</i> Michaux x <i>Vitis rupestris</i> Scheele ^{1;5}	Baixa ^{5;8}
SO4	<i>Vitis berlandieri</i> Planchon x <i>Vitis riparia</i> Michaux ^{1;5;6}	Baixa ^{3;5;8;10}
Sori	Solonis x <i>Vitis riparia</i> Michaux ¹	Média ^{7;12}
<i>Vitis champini</i>	<i>Vitis champinii</i> Planchon ¹	Alta ¹³
IAC 313	Golia x <i>Vitis cinerea</i> Engelmann ^{1;3}	-
IAC 572	<i>Vitis caribaea</i> De Candolle x 101-14MGt ^{1;3}	-
IAC 766	106-8MGt x <i>Vitis caribaea</i> De Candolle ^{1;3}	-

Informações obtidas em: ¹Maul *et al.* (2020); ²Sunridge (2020); ³Embrapa Uva e Vinho (2016); ⁴Vicopad (2020); ⁵Villa (2018) ⁶Storm & Krasokhina (2020); ⁷Audeguin *et al.* (2020); ⁸Wine Australia (2016); ⁹Satisha *et al.* (2006); ¹⁰ATVB (2013); ¹¹Chevalier (1925); ¹²Rebschule Mueller (2020), ¹³Carroll (2016).

2.2 Características avaliadas

Foram coletados dados em revistas científicas, livros, anais de congressos, teses, sites de órgãos de pesquisa e em viveiros de mudas, totalizando 30 variáveis fisiológicas, bioquímicas e morfoagronômicas listadas abaixo.

2.2.1 Variáveis fisiológicas

Para as variáveis fisiológicas foram consideradas a condutância estomática (g_s) sob condições ideais em $\mu\text{mol CO}_2 \cdot \text{m}^{-2} \cdot \text{s}^{-1}$; taxa de transpiração sob condições ideais (E) em $\text{mmol H}_2\text{O} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$; taxa fotossintética (A) sob condições ideais em $\mu\text{mol m}^{-2} \cdot \text{s}^{-1}$; eficiência no uso da água (EUA) sob condições ideais em $\mu\text{mol CO}_2 \text{ mol}^{-1} \text{ H}_2\text{O}$; e a eficiência instantânea do uso da água ($EiUA$) sob condições ideais em $\mu\text{mol CO}_2 \text{ mmol}^{-1} \text{ H}_2\text{O}$;

2.2.2 Variáveis nutricionais

Para as variáveis nutricionais foram considerados o teor dos macronutrientes nitrogênio, fósforo e potássio presente no pecíolo da folha, em porcentagem; assim como, a capacidade da planta em absorver nitrogênio, fósforo e potássio, em uma escala de notas atribuída entre 1 e 5 (alta = 5, média-alta = 4, média = 3, baixa- média = 2 e baixa = 1);

2.2.3 Variáveis morfoagronômicas

Para as variáveis morfoagronômicas foram considerados o início de quedas de folhas em dias; produção de estacas em número de estacas por plantas; ângulo geotrópico em grau; distribuição de raízes até 60 cm em porcentagem; tolerância ao déficit hídrico em uma escala de notas variando entre 1 e 3 (alto=3, médio=2 e baixo=1); e nas variáveis listadas a seguir foram atribuídas uma escala de notas variando entre 1 e 5 (alta = 5, média-alta = 4, média = 3, baixa- média = 2 e baixa = 1), para resistência à antracnose, míldio, fusariose, filoxera e nematóide; pegamento na enxertia; facilidade de enraizamento de estacas; vigor induzido a copa; tolerância a solo salino, encharcado, argiloso, arenoso, ácido e alcalino.

2.3 Análise de dados

Todas as análises foram realizadas utilizando a linguagem Python na plataforma Google Colaboratory².

2.3.1 Pré-processamento do conjunto de dados

O pré-processamento ocorreu da mesma forma para o conjunto de treinamento, com 42 cultivares, e o da predição, contendo três cultivares (IAC 313, IAC 572 e IAC 766). Para isso, foi realizado o preenchimento dos valores omissos pela média de cada variável, empregando-se a função *df.fillna()*, e foi verificado que não haviam duplicatas através da função *drop_duplicates()*, ambas disponíveis na biblioteca do pandas (McKinney 2010), e a normalização foi realizada através do *MinMaxScaler*, (Figura 1), cuja transformação é dada pela equação (01) e (02), disponível na biblioteca do scikit-learn (Pedregosa *et al.* 2011).

$$X_{std} = \frac{(x-x.min)}{(x.max - x.min)} \quad (01)$$

² <https://colab.research.google.com/>

$$X_{\text{scaled}} = X_{\text{std}} \times (\max - \min) + \min \quad (02)$$

onde: *min*, *max* corresponde ao intervalo da amostra.

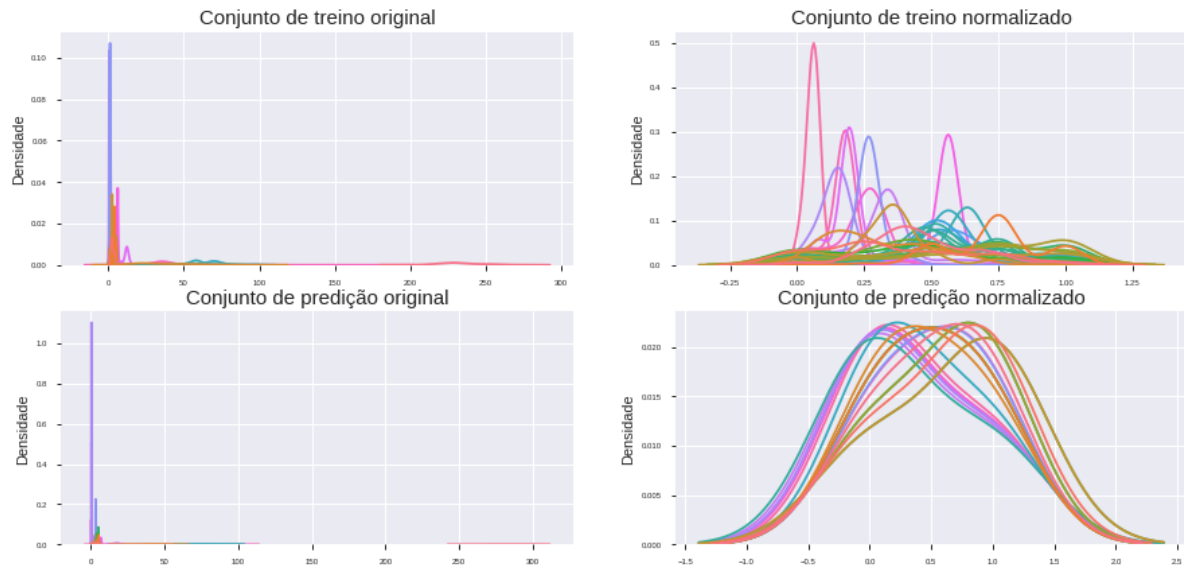


Figura 1 - Conjunto de dados antes e depois da normalização. Fonte: Verslype *et al.* (2021).

2.3.2 Comparação de algoritmos supervisionado

Com o intuito de avaliar o desempenho de diferentes algoritmos supervisionados para a predição de classes de tolerância ao déficit hídrico em porta-enxertos de videira, foram comparados seis algoritmos: *Decision Tree* (DT) (Breiman *et al.*, 1984), *Random Forest* (RF) (Breiman, 2001), *K-Nearest Neighbors* (KNN) (Fix *et al.*, 1951), *XGBoost* (XGB) (Chen & Guestrin, 2016), *Support Vector Machines* (SVM) (Cortes & Vapnik, 1995) e *Linear Discriminant Analysis* (LDA) (Rao, 1948). Esses algoritmos foram escolhidos por serem facilmente acessíveis pelo público, sendo que todos algoritmos encontram-se disponíveis na biblioteca scikit-learn (Pedregosa *et al.*, 2011), com exceção do XGB disponível na biblioteca xgboost (Chen & Guestrin, 2016). Para a determinação do melhor algoritmo foram realizados testes de validação cruzada e *grid search* ambos disponíveis na biblioteca do scikit-learn (Pedregosa *et al.*, 2011). Foram avaliadas as métricas taxa de precisão (*P*), *recall* (*R*), acurácia e *f1-score* de todos os modelos, representado respectivamente pelas equações (03), (04), (05) e (06) (Pedregosa *et al.* 2011; Shalev-Shwartz & Ben-David, 2014; Skansi, 2018).

$$P = VP / (VP + FP) \quad (03)$$

onde: *P* representa a precisão, *VP* indica o número de verdadeiros positivos e *FP* o número de falsos positivos.

$$R = VP / (VP + FN) \quad (04)$$

onde: R representa o *recall*, VP é o número de verdadeiros positivos e FN o número de falsos negativo.

$$acurácia = (VP + VN) / (VP + VN + FP + FN) \quad (05)$$

onde: VP é o número de verdadeiros positivo, VN o número de verdadeiros negativo, FP o número de falsos positivo e FN o número de falsos negativo.

$$f1 = 2 \times (precisão \times recall) / (precisão + recall) \quad (06)$$

onde: $f1$ representa o obtido pela média harmônica entre o *recall* e a precisão.

2.3.3 Predição de tolerância à seca

A análise de predição de tolerância ao déficit hídrico para as três cultivares de porta-enxertos de videira brasileiras IAC 313, IAC 572 e IAC 766, foi realizada empregando o algoritmo de melhor desempenho e a sua respectiva configuração de hiperparâmetros, indicada pelo *grid search* (Tabela 2).

Tabela 2 – Hiperparâmetros analisados pelo *grid search* em cada algoritmo.

Algoritmo	Hiperparâmetros
DT	Criterion: [gini, entropy]; Max_features [auto, log2, sqrt, None]
RF	N_estimators: [1, 5, 10, 100, 1000]; Max_features: [1, 2, 3]
XGB	N_estimators: [10, 100, 500]; Max_features: [auto, log2, sqrt, None]
SVM	C: [0.001, 0.1, 1, 10, 20, 100]; Kernel: [rbf, linear]
LDA	N_components: [10, 15, 20, 25, 30]
KNN	N_neighbors: [3, 7, 10]; Weights: [uniform, distance]

3 RESULTADOS E DISCUSSÃO

As diferenças no desempenho dos seis algoritmos sobre o conjunto de dados em estudo são observados na (Tabela 3). Entre as métricas avaliadas, destacaram-se a acurácia que variou entre 98,10% e 64,29%, a precisão entre 98% e 67% e o recall entre 98% e 64%. Para Skansi (2018), estas métricas de avaliação possibilitam comparar e obter uma visão significativa sobre o desempenho de um algoritmo, assim como, indicar qual é o melhor classificador quando são comparados dois ou mais algoritmos. Pois a acurácia permite identificar o quanto é bom o classificador, a

precisão determina a capacidade do modelo em evitar resultados errôneos, ou seja, de não rotular uma amostra positiva como negativa, evitando assim classificações falsas, e o recall aponta sucesso do modelo em questão para encontrar todas as amostras positivas (Pedregosa *et al.*, 2011; Skansi 2018).

Tabela 3 - Avaliação da acurácia, precisão (P), *recall* (R) e f1-score dos seis algoritmos obtidos pela validação cruzada.

Algoritmo	Acurácia (%)	P (%)	R (%)	f1-score
DT	82,86	86	83	0,83
RF	98,10	98	98	0,98
XGB	96,67	97	97	0,97
SVM	87,14	89	87	0,87
LDA	64,29	67	64	0,65
KNN	85,71	89	86	0,86

Nesse sentido, a performance dos modelos com base na análise de validação cruzada, indicaram o algoritmo RF como o melhor classificador, devido aos maiores valores médios obtidos para as métricas acurácia (98,10%), precisão (98%) e recall (98%). Não obstante, o classificador LDA apresentou os menores valores de acurácia (64,29%), precisão (67%) e recall (64%) entre todos os modelos avaliados, indicando uma baixa taxa de acerto na indicação das três classes de tolerância ao déficit hídrico.

Para a indicação do melhor algoritmo, com a capacidade de prever classes de tolerância ao déficit hídrico, também foi utilizada a análise de *grid search* nos seis algoritmos em relação ao conjunto de dados (Tabela 3). Os resultados obtidos confirmam que o LDA apresentou o pior desempenho, obtendo apenas 67,14% de acerto como melhor resultado possível, para todas as configurações de hiperparâmetros testados sobre o conjunto de dados, e conseqüentemente a menor curva de aprendizagem (Figura 2). Enquanto, os algoritmos RF e XGB obtiveram os melhores resultados com respectivamente 98,57% e 96,19% de acerto para prever as três classes de tolerância ao déficit hídrico, (Tabela 4). O melhor parâmetro e a sua respectiva configuração de hiperparâmetros, e curva de aprendizagem plotada, na Figura 2, indicam uma maior taxa de aprendizado com o passar do tempo.

Tabela 4 - Avaliação dos melhores hiperparâmetros para cada modelo obtida pelo *Grid search*.

Algoritmo	Melhor resultado	Melhor parâmetro	Melhor configuração de hiperparâmetros
SVM	83,33%	{'C': 100, 'kernel': 'rbf'}	SVC(C=100, break_ties=False, cache_size=200, class_weight=None, coef0=0.0, decision_function_shape='ovr', degree=3, gamma='auto', kernel='rbf', max_iter=-1, probability=False, random_state=None, shrinking=True, tol=0.001, verbose=False)
RF	98,57%	{'max_features': 2, 'n_estimators': 1000}	RandomForestClassifier(bootstrap=True, ccp_alpha=0.0, class_weight=None, criterion='gini', max_depth=None, max_features=2, max_leaf_nodes=None, max_samples=None, min_impurity_decrease=0.0, min_impurity_split=None, min_samples_leaf=1, min_samples_split=2, min_weight_fraction_leaf=0.0, n_estimators=1000, n_jobs=None, oob_score=False, random_state=None, verbose=0, warm_start=False)
KNN	90,48%	{'n_neighbors': 3, 'weights': 'distance'}	KNeighborsClassifier(algorithm='auto', leaf_size=30, metric='minkowski', metric_params=None, n_jobs=None, n_neighbors=3, p=2, weights='distance')
XGB	96,19%	{'max_features': 'auto', 'n_estimators': 500}	XGBClassifier(base_score=0.5, booster='gbtree', colsample_bylevel=1, colsample_bynode=1, colsample_bytree=1, gamma=0, learning_rate=0.1, max_delta_step=0, max_depth=3, max_features='auto', min_child_weight=1, missing=None, n_estimators=500, n_jobs=1, nthread=None, objective='multi:softprob', random_state=0, reg_alpha=0, reg_lambda=1, scale_pos_weight=1, seed=None, silent=None, subsample=1, verbosity=1)
LDA	67,14%	{'n_components': 10}	LinearDiscriminantAnalysis(n_components=10, priors=None, shrinkage=None, solver='svd', store_covariance=False, tol=0.0001)
DT	86,19%	{'criterion': 'entropy', 'max_features': None}	DecisionTreeClassifier(ccp_alpha=0.0, class_weight=None, criterion='entropy', max_depth=None, max_features=None, max_leaf_nodes=None, min_impurity_decrease=0.0, min_impurity_split=None, min_samples_leaf=1, min_samples_split=2, min_weight_fraction_leaf=0.0, presort='deprecated', random_state=None, splitter='best')

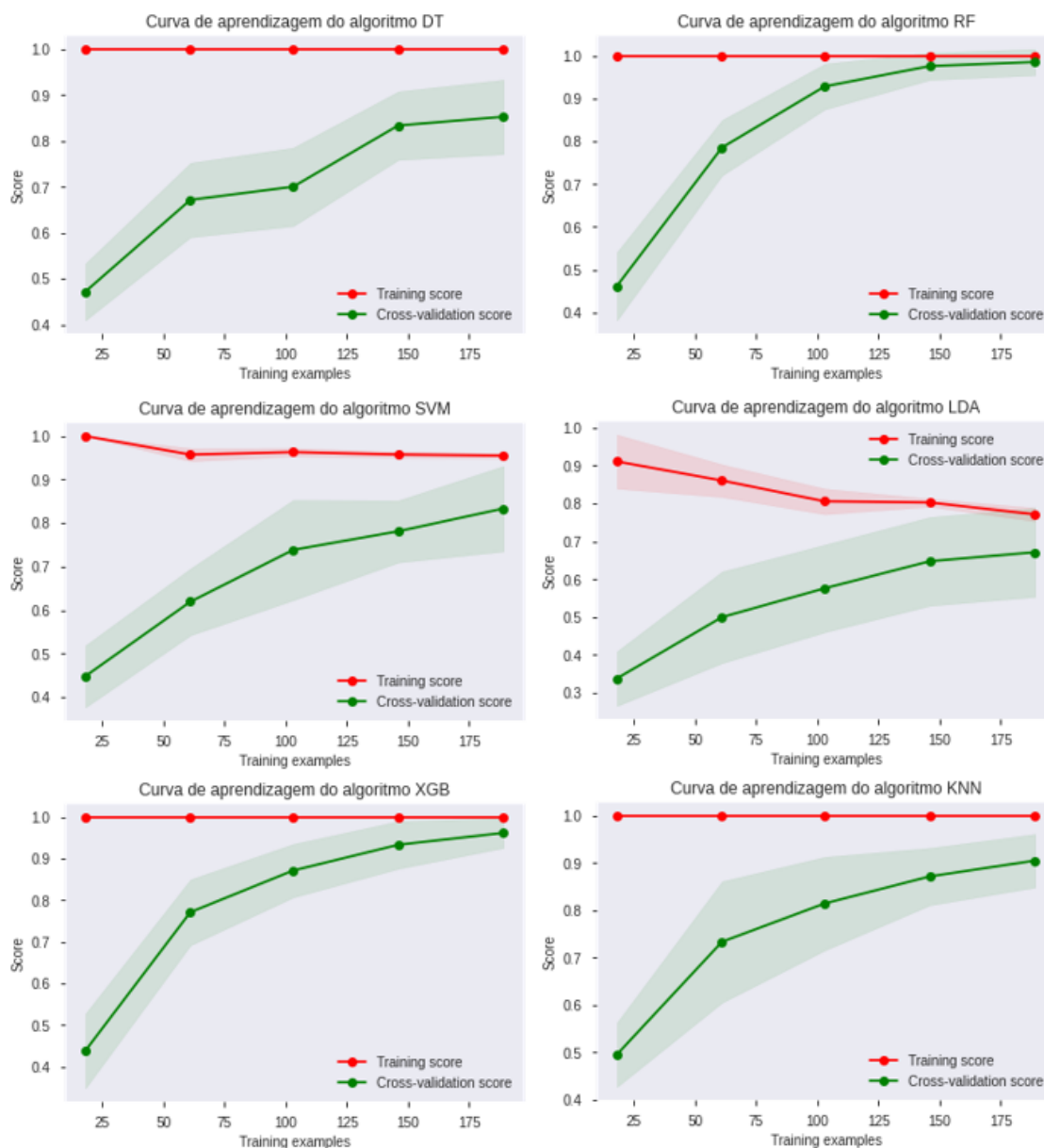


Figura 2 - Curva de aprendizagem dos seis algoritmos sobre o conjunto de dados, através da validação cruzada com K fold igual a 10 e métrica score (R^2). Sendo DT - *Decision Tree*, RF – *Random Forest*, XGB - *XGBoost*, SVM – *Support Vector Machines*, LDA - *Linear Discriminant Analysis* e KNN - *K-Nearest Neighbors*. Fonte: Verslype *et al.* (2021).

Para o melhor modelo indicado pela validação cruzada e configuração de hiperparâmetros fornecido pelo *grid search*, foi gerado o relatório de classificação, (Tabela 5). A partir destes resultados, observa-se que o RF apresentou um f1-score de 0,99 para classe de alta tolerância, de 0,98 para baixa e de 0,99 para média. Segundo Pedregosa *et al.* (2011), os valores de f1-score próximos a 1 indicam melhor desempenho e os próximos a zero a pior pontuação do modelo. Nesse sentido, estes resultados indicam que a taxa de acerto do algoritmo em cada uma das três classes de tolerância ao déficit hídrico foi alta.

Tabela 5 – Relatório de classificação da validação cruzada com 10 *flod* do algoritmo RF.

Classes	Precisão	Recall	f1-score	Support
Alta	1,00	0,99	0,99	70
Baixa	0,96	1,00	0,98	70
Media	1,00	0,97	0,99	70
Acurácia			0,99	210
Média macro	0,99	0,99	0,99	210
Média ponderada	0,99	0,99	0,99	210

Skansi (2018) relata que, o método da matriz de confusão é interessante por facilitar a visualização do número de resultados erroneamente rotulados, e corretamente rotulados representados na diagonal da matriz. A matriz de confusão obtida pode ser observada na (Figura 3), demonstrando que houve uma baixa taxa de resultados classificados erroneamente pelo algoritmo RF, o qual mostra-se como um bom classificador para o nosso problema em questão.

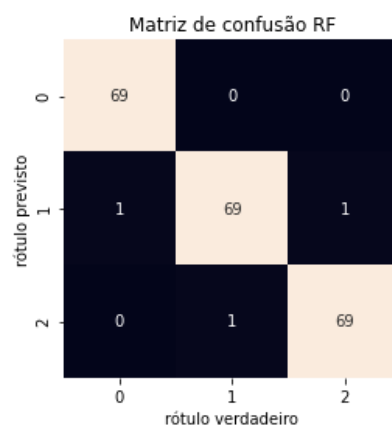


Figura 3 - Matriz de confusão do algoritmo RF. Onde 0 - representa a classe de alta tolerância ao déficit hídrico, 1 – baixa e 2 – média. Fonte: Verslype *et al.* (2021).

A predição de classes de tolerância ao déficit hídrico indicadas pelo algoritmo RF foi de alta tolerância para as cultivares IAC 313 e IAC 766, e de baixa tolerância para a cultivar IAC 572. Para esta predição, o algoritmo RF atribuiu uma escala de importância para todas as variáveis avaliadas no conjunto de dados. Entre estas, observamos que as variáveis produção de estacas, início de queda de folhas, resistência à antracnose, ângulo geotrópico, tolerância ao solo alcalino e a distribuição de raízes até 60 cm foram as primeiras variáveis com maior importância na sua tomada de decisão, conforme podemos visualizar na (Figura 4).

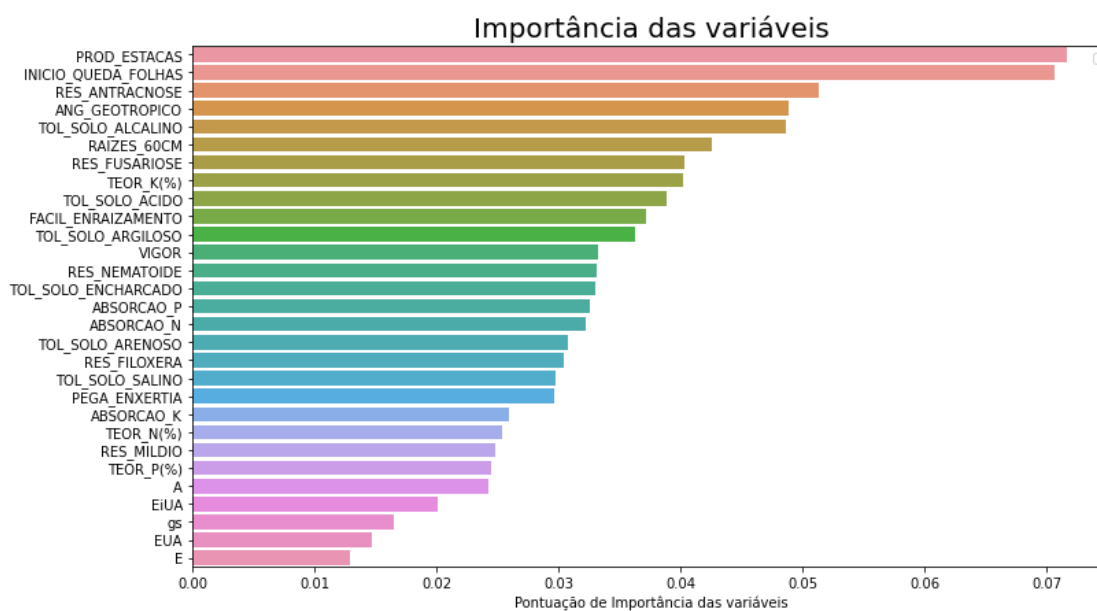


Figura 4 - Variáveis com maior importância na tomada de decisão do algoritmo RF. Fonte: Verslype *et al.* (2021).

A atribuição de maior importância à característica produção de estacas na tomada de decisão do RF, pode ser explicada através das plantas com maior eficiência no uso da água produzirem mais matéria seca por grama de água transpirada, em vista da necessidade de absorção média de 1 litro para cada 2 gramas de massa seca produzida nas videiras (Ojeda *et al.*, 2004; Serra *et al.*, 2014). Conseqüentemente essas plantas mais vigorosas em linhas gerais, toleram condições com menor disponibilidade hídrica (Ojeda *et al.*, 2004; Villa, 2018). Enquanto, a variável início de queda de folhas encontra-se relacionada com a tolerância ao déficit hídrico, pois nesta fase, as videiras direcionam os fotoassimilados para as raízes permitindo que aumente a superfície de absorção e se recupere do estresse hídrico (Ojeda *et al.*, 2004).

A maior importância atribuída as variáveis distribuição de raízes até 60 cm e ângulo geotrópico na tomada de decisão do RF, pode ser explicado por meio do ângulo geotrópico determinar a capacidade de aprofundamento do sistema radicular das videiras (Villa, 2018). No qual, a literatura tem ressaltado que, ângulos fechados propiciam o aprofundamento das raízes, e conseqüentemente possibilitam alcançar camadas mais profundas do solo que permanecem úmidas, assim como, de evitar a absorção de elementos prejudiciais as plantas em solos salinos, através da absorção seletiva (Smart *et al.*, 2006; Satisha *et al.*, 2006; Serra *et al.*, 2014; Villa, 2018). Estas características, conferem aos porta-enxertos uma melhor adaptação a condições de

estresse déficit hídrico, e consequentemente, maior tolerância ao déficit hídrico (Satisa *et al.*, 2006; Villa, 2018).

O RF na sua tomada de decisão atribuiu uma alta importância a variável resistência à antracnose. Segundo Lima *et al.* (2009), trata-se de uma doença de grande impacto na cultura da videira por acometer principalmente tecidos jovens e tenros na parte aérea da planta. Associação entre condições de estresse hídrico e predisposição a ocorrência de antracnose foram relatados por Erbaugh *et al.* (1995) que mencionam um aumento significativo na severidade da doença na espécie *Cornus florida* L. em plantas sombreadas e sob condição de estresse hídrico, e por Hsiang (2006), na poácea *Poa annua* L., que se torna incapaz de responder rapidamente à infecção da antracnose e superar a doença, quando submetida a estresses severos como estresse hídrico cíclico. Nesse sentido, deve existir algum grau de interferência entre as duas variáveis.

Como mencionado acima, após avaliar todas as métricas de acurácia, precisão, *recall* e *f1-score*, foi identificado o RF como o melhor modelo produzido, nos cenários e conjuntos de dados avaliados, para prever classes de tolerância ao déficit hídrico em porta-enxertos de videira. Apesar de ser um estudo inicial, que ainda precisa ser validado através de estudos em condições de campo sob condições de déficit hídrico, esta abordagem com algoritmos supervisionados, mostrou-se como uma estratégia útil e acessível ao melhorista, podendo contribuir no desenvolvimento dessas cultivares através da predição, identificação e seleção de genótipos promissores.

4 CONCLUSÃO

Os melhores classificadores foram o RF seguido pelo XGB, com respectivamente 98,57% e 96,19% de acerto, para a indicação de classes de tolerância ao déficit hídrico em porta-enxertos de videira.

Nos cenários e datasets avaliados observamos que a melhor configuração de hiperparâmetro para o RF foi *max_features* igual a dois e *n_estimators* igual a 1000. Enquanto o classificador LDA apresentou a menor eficiência, com uma taxa de acerto de 67,14% para prever as classes de tolerância ao déficit hídrico.

As cultivares de porta enxerto de videira IAC 313 e IAC 766, foram classificadas pelo melhor modelo produzido como genótipos de alta tolerância ao déficit hídrico, enquanto o IAC 572 apresenta baixa tolerância.

Neste sentido, os algoritmos de aprendizado de máquina, podem vir a contribuir na identificação e seleção de genótipos tolerantes ao déficit hídrico em programas melhoramento genético de videira, sugerindo-se fortalecer os resultados obtidos em ensaios de campo sob condições de déficit hídrico. Assim como, a seleção de características como possível extensão ao presente trabalho.

AGRADECIMENTOS

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão de bolsa de mestrado à primeira autora.

REFERÊNCIAS BIBLIOGRÁFICAS

ATVB – Association Technique Viticole de Bourgogne (2013) **Caractéristiques des porte-greffes et greffons utilisés en Côte-d’Or**. Disponível em: <<http://www.atvb-bourgogne.com/>>. Acesso em agosto 2020.

Audeguin L *et al.* (2020) **PI@ntGrape**: le catalogue des vignes cultivées en France. Montpellier, France: IFV – INRA – Montpellier SupAgro. Disponível em <<http://plantgrape.plantnet-project.org/>>. Acesso em setembro 2020.

Beiki AH (2019) Molecular Predicting Drought Tolerance in Maize Inbred Lines by Machine Learning Approaches. **BioRxiv** 10: 1-33.

Breiman L, Friedman J, Olshen RA e Stone CJ (1984) **Classification and Regression Trees**. Wadsworth, Belmont, CA: CRC press.

Breiman L (2001) Random forests. **Machine learning**, 45(1), 5-32.

Bruckner CH e Albuquerque AS (2005) Melhoramento de fruteiras. In: Borém A (ed.). **Melhoramento de espécies cultivadas**. Viçosa: Ed. UFV, p.813 – 863.

Carroll B (2020) **Characteristics of *Vitis* grape rootstocks and adaptability**. Oklahoma State University, Oklahoma, 4p.

Chaves MM *et al.* (2010) Grapevine under deficit irrigation: hints from physiological and molecular data. **Annals Appl. Biol.** 1: 661-676.

Chen T e Guestrin C (2016). XGBoost: A Scalable Tree Boosting System. In **Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining** (pp. 785–794). New York, USA: ACM.

Chevalier A (1925) L'amélioration de la vigne en France et les Travaux de G. Couderc sur l'Hybridation et le Greffage. **Revue de botanique appliquée et d'agriculture coloniale** 52: 926-945.

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

Cortes C e Vapnik V (1995) Support-vector networks. **Machine Learning**, 20(3), 273–297.

Embrapa Uva e Vinho. **Cultivares de porta-enxertos**. 2016. Disponível em: <<https://www.embrapa.br/en/uva-e-vinho/cultivares-e-porta-enxertos/>> . Acesso em: 11 abr. 2020.

Erbaugh DK, Windham MT, Stodola AJ e Augé RM (1995) Light intensity and drought stress as predisposition factors for dogwood anthracnose. **Journal of Environmental Horticulture**, 13(4), 186-189.

Etminan A *et al.* (2019) Determining the best drought tolerance indices using artificial neural network (ANN): Insight into application of intelligent agriculture in agronomy and plant breeding. **Cereal Research Communications**, 47: 170-181.

Fix E e Hodges JL (1951) **Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties**. Texas: USAF School of Aviation Medicine, Randolph Field.

Hsiang T (2006) **Anthracnose**: terrorizing turf grass. GCSAA: 2p. Disponível em: <<https://www.gcsaa.org/>>. Acesso em: 14 mai. 2020.

Leão PCS e Silva DJ (2014) Cultivo da videira no Semiárido brasileiro. In Pio R (eds) **Cultivo de fruteiras de clima temperado em regiões subtropicais e tropicais**. Lavras, UFLA, p. 578-618.

Leão PCS, Soares JM e Rodrigues BL (2009) Principais Cultivares. In Soares JM e Leão PCS (eds.) **A vitivinicultura no Semiárido Brasileiro**. Brasília, DF, Embrapa Informação Tecnológica, Petrolina, Embrapa Semi-Árido, p.151-214.

Lima MR, Lopes DB, Tavares SCCH, Tessmann DL e Melo NF (2009) Doenças e alternativas de controle. In Soares JM e Leão PCS (eds.) **A vitivinicultura no Semiárido Brasileiro**. Brasília, DF, Embrapa Informação Tecnológica, Petrolina, Embrapa Semi-Árido, p.543-596.

Maul E *et al.* (2020) **Vitis International Variety Catalogue**. Federal Research Centre for Cultivated Plants, Institute for Grapevine Breeding. Disponível em: <<https://www.vivc.de/>>. Acesso em: 30 Jul. 2020.

McKinney W (2010) **Data Structures for Statistical Computing in Python**. In Stéfan van der Walt and Jarrod Millman, editors, Proceedings of the 9th Python in Science Conference, pages 56 – 61.

Ojeda H, Deloire A, Wang Z and Carbonneau A (2004) Detereminación y control del estado hídrico de la vid: efectos morfológicos y fisiológicos de la restricción hídrica en vides. **Viticultura/Enología Profesional 90**: 27-43.

Pedregosa F *et al.* (2011) Scikit-learn: Machine Learning in Python. **Journal of Machine Learning Research** 12: 2825-2830.

Rao RC (1948) The Utilization of Multiple Measurements in Problems of Biological Classification. **Journal of the Royal Statistical Society 10**: 159-203.

Rebschule Mueller (2020) **Unterlagen für den weinbau**. Disponível em <<http://www.rebschulemueller.de/>>. Acesso em: 15 Ago. 2020.

VERSLYPE, N. I. Avaliação e seleção de porta-enxertos de videira (*Vitis* spp.) tolerantes ao déficit hídrico através de aprendizagem de máquina

Satisha J, Prakash GS, Murti GSR e Upreti KK (2006) Response of grape rootstocks to soil moisture stress. **Journal of Horticultural Sciences 1**: 19-23p.

Riaz S *et al.* (2019) Genetic diversity and parentage analysis of grape rootstocks. **Theoretical and Applied Genetics 132**: 1847–1860.

Serra I, Strever A, Myburgh PA e Deloire A (2014) The interaction between rootstocks and cultivars (*Vitis vinifera* L.) to enhance drought tolerance in grapevine. **Australian Journal of Grape and Wine Research 20**: 1-14.

Storm B e Krasokhina C (2020) **Rootstock grape varieties, gardening varieties, wild grapes, breeding varieties**. Disponível em <<https://vinograd.info/>>. Acesso em: 05 Ago. 2020.

Sunridge Nurseries (2020) **Rootstock chart**. Bakersfield, California. Disponível em <<https://www.sunridgenurseries.com/>>. Acesso em: 25 Ago. 2020.

VICOPAD - Vivai Cooperativi di Padergnone (2020) **Scelta del portinnesto**. Padergnone, Italy. Disponível em <<http://www.vicopad.it/>>. Acesso em: 19 Ago. 2020.

Villa P (2018) **Cultivar la vid**. Editorial de Vecchi, USA, 160p.

Waskom M (2021) Seaborn: statistical data visualization. **The Open Journal 60**: 3021p.

Wine Australia Corporation (2016) **Grapevine rootstock selector tool**. Austrália. Disponível em: <<https://grapevine rootstock.com/>>. Acesso em: 25 Ago. 2020.

ANEXO

Material suplementar do artigo intitulado: “Cultivares de porta-enxertos de videira: conjunto de dados descrevendo características fisiológicas, bioquímicas e agrônômicas”.

Tabela 1 - Parâmetros, descrição e referências bibliográficas.

Parâmetro	Descrição	Fonte
Genótipo	Nome comumente empregado para a cultivar	[1], [2], [3] e [4]
Espécie	Espécie da cultivar	[3]
Genitor 1	Nome do progenitor 1	[5], [3], [6] e [7]
Genitor 2	Nome do progenitor 2	[5], [3], [6] e [7]
Flor	Sexo da flor (feminino, masculino, ou hermafrodita)	[2] e [3]
Ciclo vegetativo	Ciclo vegetativo (precoce, intermediário e tardio)	[2]
Brotação	Dias para a brotação	[2]
Maturação	Dias para a maturação	[2]
Queda folhas	Dias para o início de queda de folhas	[2]
Resistência antracnose	Grau de resistência a antracnose (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[1], [2], [6], [8], [9] e [4]
Resistência míldio	Grau de resistência ao míldio (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[1], [2], [6], [8] e [4]
Produção de estacas	Número de estacas produzida por planta	[2]
Resistência à fusariose	Grau de resistência a fusariose (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[10], [6], [11] e [7]
Resistência a nematóides	Grau de resistência ao nematóide (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[1], [5], [12], [13], [6], [11], [7], [14], [15], [8], [16], [17], [18], [19], [20], [21], [22] e [23]
Resistência à filoxera	Grau de resistência a filoxera (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[1], [5], [12], [24], [13], [6], [11], [7], [8], [17], [25], [26], [27], [9], [28], [29], [30], [20], [31], [21], [4], [23], [22] e [32]
Pega na enxertia	Facilidade de pega na enxertia (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[1], [5], [33], [24], [7], [15], [8], [17], [25], [26], [19], [27], [9], [28], [34], [30], [31], [21] e [4]
Facilidade de enraizar	Facilidade de enraizar (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[1], [24], [6], [7], [8], [17], [25], [26], [19], [27], [9], [28], [34], [30] e [4]
Tolerância a salinidade	Grau de tolerância ao solo salino (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[5], [12], [6], [8], [17], [25], [18], [27], [9], [29], [20], [21], [35], [4], [23], [22] e [32]
Tolerância ao encharcamento	Grau de tolerância ao solo encharcado (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[36], [1], [5], [33], [12], [24], [7], [14], [8], [16], [17], [25], [18], [26], [19], [27], [9], [29], [34], [20], [21], [35], [37], [4] e [22]
Vigor	Vigor induzido à copa (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[1], [5], [12], [38], [24], [13], [6], [7], [14], [8], [17], [25], [18], [26], [19], [9], [28], [29],

		[30], [20], [31], [21], [35], [37], [4], [23], [22] e [32]
Tolerância ao solo argiloso	Grau de tolerância ao solo argiloso (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[36], [1], [5], [33], [38], [24], [6], [7], [14], [8], [16], [17], [25], [26], [19], [27], [9], [30], [20], [31], [21], [37] e [4]
Tolerância ao solo arenoso	Grau de tolerância ao solo arenoso (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[36], [1], [5], [38], [24], [6], [7], [14], [8], [16], [17], [25], [27], [30], [20], [31], [21], [37] e [4]
Tolerância ao solo ácido	Grau de tolerância ao solo ácido (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[1], [33], [38], [24], [6], [7], [15], [8], [17], [25], [26], [27], [29], [30], [20], [31], [35], [4] e [22]
Tolerância ao solo alcalino	Grau de tolerância ao solo alcalino (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[36], [1], [33], [12], [38], [24], [13], [7], [15], [8], [26], [29], [4] e [22]
Tolerância ao solo compactado	Grau de tolerância ao solo compactado (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[36], [1], [33], [7], [25], [18], [26], [9], [30], [35], [37] e [4]
Profundidade de raízes	Profundidade do sistema radicular (muito profundo=6, profundo=5, profundo-moderado=4, moderado=3, superficial=2, e muito superficial=1)	[1], [39], [12], [40], [15], [8], [34], [30], [35], [37] e [22]
Ângulo geotrópico	Ângulo geotrópico em grau	[39], [7] e [37]
Distribuição raízes 60cm (%)	Distribuição acumulada, em porcentagem, das raízes verticais até 60 cm de profundidade no solo	[41], [42], [43], [30], [44] e [45]
Distribuição raízes 100cm (%)	Distribuição acumulada, em porcentagem, das raízes verticais até 100 cm de profundidade no solo	[41], [42], [43], [30], [44] e [45]
Absorção de N	Capacidade de absorver nitrogênio (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[5] e [27]
Absorção de P	Capacidade de absorver fósforo (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[5] e [27]
Absorção de K	Capacidade de absorver potássio (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[36], [5], [33], [8], [27], [29] e [20]
Absorção de Mg	Capacidade de absorver magnésio (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[1], [5], [15], [27] e [29]
Teor de N (%)	Teor de nitrogênio no pecíolo em porcentagem	[46], [47], [48], [49], [50], [51], [52] e [53]
Teor de P (%)	Teor de fósforo no pecíolo em porcentagem	[47], [48], [49], [54], [52] e [53]
Teor de K (%)	Teor de potássio no pecíolo em porcentagem	[55], [47], [48], [54], [52] e [53]
Tolerância à clorose férrica	Grau de tolerância a clorose férrica (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[1], [8], [17], [25], [26], [9], [28], [34], [31] e [4]
IPC	Grau de índice de poder clorosante	[1], [8], [18], [28], [29], [21], [37] e [4]

Tolerância ao calcário	Grau de tolerância ao calcário (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[36], [1], [5], [33], [38], [7], [8], [17], [25], [26], [27], [34], [30], [20], [31], [37], [4] e [22]
Tolerância ao Calcário ativo (%)	Grau de tolerância ao calcário ativo em porcentagem	[1], [7], [8], [16],[25], [26], [28], [29],[34], [21], [35] [35],[37], [4] e [32]
Tolerância ao Calcário total	Grau de tolerância ao calcário total (alto=5, alto-médio=4, médio=3, médio-baixo=2, e baixo=1)	[1], [38], [25], [37] e [23]
A	Taxa de fotossíntese em $\mu\text{molm}^{-2}.\text{s}^{-1}$	[56], [57], [49], [58], [50], [59], [51], [50], [60], [61] e [52]
A sob estresse hídrico	A sob condição de déficit hídrico em $\mu\text{molm}^{-2}.\text{s}^{-1}$	[57], [49], [58], [59], [60] e [61]
E	Taxa de transpiração em m mole $\text{H}_2\text{O m}^{-2}.\text{s}^{-1}$	[56], [57], [58], [50], [59], [51], [50], [60], [62] e [52]
E sob estresse hídrico	E sob condição de déficit hídrico em m mole $\text{H}_2\text{O m}^{-2}.\text{s}^{-1}$	[63], [57], [59], [62] e [58]
gs	Condutância estomática em $\mu \text{CO}_2 \text{ m}^{-2}.\text{s}^{-1}$	[56], [40], [57], [50], [59], [61] e [52]
gs sob estresse hídrico	gs sob condição de déficit hídrico em $\mu \text{CO}_2 \text{ m}^{-2}.\text{s}^{-1}$	[63], [40], [57], [49], [58], [59] e [61]
EUA	Eficiência no uso da água	Calculado pela divisão entre A e gs
EUA sob estresse hídrico	EUA sob condição de déficit hídrico	Calculado pela divisão entre A sob estresse hídrico e gs sob estresse hídrico
EIUA	Eficiência instantânea do uso da água	Calculado pela divisão entre A e E
EIUA sob estresse hídrico	EIUA sob condição de déficit hídrico	Calculado pela divisão entre A sob estresse hídrico e E sob estresse hídrico
ABA	Teor de ácido abscísico em (ng/g)	[64], [59] e [58]
ABA sob estresse hídrico	Teor de ácido abscísico em condição de déficit hídrico (ng/g)	sob [64], [59] e [58]
ABA sob reidratação	Teor de ácido abscísico em condição de reidratação (ng/g)	sob [64]
Prolina	Prolina em	[60]
Prolina sob estresse hídrico	Prolina sob condição de déficit hídrico em	[60]
Densidade de estômatos	Densidade de estômatos em (estômato/ mm^2)	[65], [64], [66] e [67]
Potencial osmótico (Ψ_{os})	Potencial osmótico em (Mpa)	[60]
Ψ_{os} sob estresse hídrico	Potencial osmótico sob condição de déficit hídrico em (Mpa)	[60]
Condutividade hidráulica (K)	Condutividade hidráulica em $\text{kg MPa}^{-1}.\text{m}^{-2}.\text{s}^{-1} \times 104$	[50]
Tolerância à seca	Grau de tolerância à seca (alto=3, médio=2 e baixo=1)	[36], [1], [5], [55], [63], [12], [24], [38], [6], [7], [8], [17], [19], [9], [28], [59], [37], [35], [4] e [22]

Referências

- [1] L. Audeguin, V. Bécart, E. Besnard, J.M. Boursiquot, G. Ciccolini, S. Dedet, G. Delorme, J. Follet, A. Froelhy, L. Garcin, N. Gonin, T. Lacombe, F. Laffargue, C. Marchal, M. Pesato, C. Sereno, G. Uriel, M. Wentzel, and O. Yobregat. *PI@ntgrape: le catalogue des vignes cultivées en france*. <https://plantgrape.plantnet-project.org/>, 2020 (accessed 30 June 2020).
- [2] U.A. Camargo, P.S. Ritschel, J.D.G. Maia, V. Quecini, C.A.E. Machado, D.D. Bosco, I. Sinski, V. Comachio, and R. Zilio. Banco ativo de germoplasma de uva. <https://www.embrapa.br/en/uva-e-vinho/banco-ativo-degermoplasma-de-uva/pesquisa>, 2017 (accessed 25 June 2020).
- [3] E. Maul, R. Töpfer, F. Röckel, U. Brühl, M. Hundemer, A. Mahler-Ries, M. Walk, S. Kecke, A. Wolck, and A. Ganesch. Vitis international variety catalogue (vivc). <http://www.vivc.de/>, 2020 (accessed 20 June 2020).
- [4] B. Storm and C. Krasokhina. Rootstock grape varieties, gardening varieties, wild grapes, breeding varieties. <https://vinograd.info/sorta/ostalnoe/>, 2020 (accessed 05 August 2020).
- [5] L.J. Bettiga. *Wine grape varieties in California*, volume 3419. UCANR Publications, 2003.
- [6] Embrapa Uva e Vinho. Cultivares de uva e porta-enxertos de alta sanidades. <https://www.embrapa.br/en/uva-e-vinho/cultivares-eporta-enxertos/porta-enxertos>, 2020 (accessed 05 July 2020).
- [7] E. Giovannini. *Manual de viticultura: série Tekne*. Bookman Editora, 2014.
- [8] IFV Institut Français de la Vigne et du Vin. Catalogue des variétés et clones de vigne cultivés en france. <https://www.vignevin-occitanie.com/fichespratiques/cepages-et-porte-greffes-utilises-dans-le-sud-ouest/>, 2007. (accessed 05 September 2020).
- [9] M. Pecile, C. Zavaglia, and A. Ciardi. Registro nazionale delle varietà di vite. <http://catalogoviti.politicheagricole.it/catalogo.php>, 2020 (accessed 20 August 2020).
- [10] E.R. De Andrade, M.A. Dal Bo, E. Schuck, and G.J.M. Gallotti. Evaluation of grapevine (*Vitis* spp.) resistance to fusarium oxysporum f. sp. herbemontis In rio do peixe valley, santa catarina state, brazil. In *International Symposium on Viticulture and Enology* 388, pages 65–70, 1993.
- [11] T.V.M. Fajardo. *Uva para processamento: fitossanidade*. Embrapa Informação Tecnológica, 2003.
- [12] B. Carroll. Rootstocks for grape production. *Oklahoma Cooperative Extension Fact Sheets*, 1:1–4, 2016.
- [13] P. Cousins. Evolution, genetics, and breeding: viticultural applications of the origins of our rootstocks. *Grapevine Rootstocks: Current Use, Research, and Application*, page 1, 2005.
- [14] HerrickGrapevines. Grapevine rootstocks. <https://www.herrickgrapevines.com/rootstocks>, 2015. (Accessed 15 September 2020).
- [15] N. Hoskins, R. Bonfiglioli, and C. Wright. The riversun rootstock project: part 1. Information Sheet, riversun. co. nz, PO Box, 1199, 2003.
- [16] R.S. Jackson. *Wine science: principles and applications*. Academic press, 2008.
- [17] M. Keller. *The science of grapevines*. Academic press, 2020.
- [18] MC. Morisson Couderc. *Quality and services analysis*. Morisson Couderc, France, 2020.
- [19] L.T. Morton. The myth of the universal rootstock [grapes, varieties]. *Wines and Vines*, 1979.
- [20] Sunridge Nurseries. Rootstock chart. <https://www.sunridgenurseries.com/index.php/clonalselections/rootstock-chart>, 2020 (accessed 25 August 2020).
- [21] VCR Vivai Cooperativi Rauscedo. *Catalogue général des variétés et des clones à raisin de table*. Vivai Cooperativi Rauscedo sca, 2013.
- [22] Wine Australia Corporation. Grapevine rootstock selector tool. <https://grapevine-rootstock.com/>, 2016 (accessed 15 September 2020).
- [23] Vintage Nurseries. Rootstock guide. <http://www.dancagraphics.com/comps/vn/root.htm>, 2020(accessed 25 September 2020).

- [24] A.D. Clarke. Suggested rootstocks for New Zealand vineyards. Grovetown, Blenheim, 2004.
- [25] J.M.V.B Marfá. Los portainjertos en el viñedo. Anales de la Escuela de Peritos Agrícolas y de Especialidades Agropecuarias y de los Servicios Técnicos de Agricultura, pages 3–72, 1952.
- [26] Albino Morando *et al.* Vigna nuova. Vit. En, 2001.
- [27] Novavine. Characteristics of important rootstocks for California vineyards. Novavine, EUA, 2017.
- [28] Rebschule Mueller. Unterlagen für den weinbau. <http://www.rebschulemueller.de/unterlagen-info.html>, 2020 (accessed 15 August 2020).
- [29] A. Reynier. Manuel de viticulture: guide technique du viticulteur. Lavoisier, 2011.
- [30] D.R. Smart, E. Schwass, A. Lakso, and L. Morano. Grapevine rooting patterns: a comprehensive analysis and a review. *American Journal of Enology and Viticulture*, 57(1):89–104, 2006.
- [31] Brandon Thomas. Viticulture – characteristics of the vine rootstocks grafting. <http://www.thewalkinggrape.com/wpcontent/uploads/2015/08/rootstocks-grafting.pdf>, 2015 (accessed 01 August 2020).
- [32] Grapevine rootstock guide. <https://www.wonderfulnurseries.com/content/dam/web/sites/wonderfulnurseries/pdf/rootstock-guide.pdf>, 2020 (accessed 05 September 2020).
- [33] CA Chambre D’agriculture. Guide Technique: viticulture durable de Bourgogne. Bourgogne, France: Chambre d’agriculture Bourgogne-Franche Conté, 2006.
- [34] B.A. Samarakis. Bulletin de l’Office International du Vin N°140. Paris: Presses Universitaires de France, 1940.
- [35] VICOPAD Vivai Cooperative nurseries of Padergnone. Scelta del portinnesto. <http://www.vicopad.it/consigli-per-limpianto/>, 2020 (accessed 19 August 2020).
- [36] ATVB Association Technique Viticole de Bourgogne. Caractéristiques des porte-greffes et greffons utilisés en côte-d’or. ATVB: Beaune, France, 2020.
- [37] Pierluigi Villa. Cultivar la vid. Parkstone International, 2018.
- [38] A. Chevalier. L’amélioration de la vigne en France et les travaux de G. Couderc sur l’hybridation et le greffage. *Journal d’agriculture traditionnelle et de botanique appliquée*, 5(51):809–840, 1925.
- [39] E. Archer and D. Saayman. Vine Root. Institute for Grape and Wine Sciences (IGWS), Stellenbosch University, 2018.
- [40] K. Fort, J. Fraga, D. Grossi, and M.A. Walker. Early measures of drought tolerance in four grape rootstocks. *Journal of the American Society for Horticultural Science*, 142(1):36–46, 2017.
- [41] L.H. Bassoi, L.C. Grangeiro, J.A.M. E Silva, and E.E.G. da Silva. Root distribution of irrigated grapevine rootstocks in a coarse texture soil of the São Francisco Valley, Brazil. *Revista Brasileira de Fruticultura*, 24(1):35–38, 2002.
- [42] L.H. Bassoi and A.A. Miranda. Análise da distribuição radicular de videiras irrigadas em latossolo vermelho amarelo de petrolina, pe. i-massa seca de raízes. In Embrapa Semiárido - Resumo em anais de congresso (ALICE). In: CONGRESSO BRASILEIRO DE ENGENHARIA AGRÍCOLA, 26., 1997, Campina Grande ..., 1997.
- [43] M.A. Dalbó, M. da Veiga, and J.P.P. Gardin. Development of the vine root system in function of rootstocks and soil physical and chemical attributes. *Agropecuária Catarinense*, 24(3):59–63, 2011.
- [44] J.M. SOARES, A.L.C. COSTA, and M.S.B. de MOURA. Avaliação da distribuição do sistema radicular da videira na região do submédio São Francisco. In Embrapa Semiárido-Resumo em anais de congresso (ALICE). In: CONGRESSO LATINO-AMERICANO DE VITICULTURA E ENOLOGIA, 10.; CONGRESSO ..., 2005.
- [45] J.M. Southey. Root distribution of different grapevine rootstocks on a relatively saline soil. *South African Journal of Enology and Viticulture*, 13(1):1–9, 1992.
- [46] Aude Habran, Mauro Commisso, Pierre Helwi, Ghislaine Hilbert, Stefano Negri, Nathalie Ollat, Eric Gomés, Cornelis van Leeuwen, Flavia Guzzo, and Serge Delrot. Rootstocks/scion/nitrogen interactions affect secondary metabolism in the grape berry. *Frontiers in plant science*, 7:1134, 2016.

- [47] A.G. Ibacache and C.B. Sierra. Influence of rootstocks on nitrogen, phosphorus and potassium content in petioles of four table grape varieties. *Chilean Journal of Agricultural Research*, 69(4):503–508, 2009.
- [48] C.A.F do Nascimento, J.H.I. Martel, and C.G.P. Júnior. Comportamento de porta-enxertos cítricos submetidos em composições de diferentes substratos. *Estação Científica (UNIFAP)*, 8(2):47–56, 2018.
- [49] N. Nikolaou, K. Angelopoulos, and N. Karagiannidis. Effects of drought stress on mycorrhizal and non-mycorrhizal cabernet sauvignon grapevine, grafted onto various rootstocks. *Experimental Agriculture*, 39(3):241, 2003.
- [50] T.L.B. Sampaio. Using rootstocks to manipulate vine physiological performance and mediate changes in fruit and wine composition. 2007.
- [51] J. Satisha, S.D. Ramteke, and G.S. Karibasappa. Physiological and biochemical characterisation of grape rootstocks. *South African Journal of Enology and Viticulture*, 28(2):163–168, 2007.
- [52] R.G. Somkuwar, P.B. Taware, M.A. Bhange, J. Sharma, and I. Khan. Influence of different rootstocks on growth, photosynthesis, biochemical composition, and nutrient contents in ‘fantasy seedless’ grapes. *International Journal of Fruit Science*, 15(3):251–266, 2015.
- [53] D. Vijaya, V. Joshi, G.R. Reddy, D.A. Kumari, and B.S. Rao. Effect of rootstock-scion interaction on petiole nutrient content, bud break, and yield of three commercial grape cultivars. *Agricultural Research Journal*, 56(1):84–91, 2019.
- [54] M.S. Rizk-Alla, G.H. Sabry, and M.A. Abd-El-Wahab. Influence of some rootstocks on the performance of red globe grape cultivar. *The Journal of American Science*, 7(4):71–81, 2011.
- [55] L. Brancadoro, L. Valenti, A. Reina, and A. Scienza. Potassium content of grapevine during the vegetative period: The role of the rootstock. *Journal of Plant Nutrition*, 17(12):2165–2175, 1994.
- [56] A.R. de M. Chaves, P.C. de S. Leão, G.E. Pereira, S. de T. Aidar, L.M. Santos, and B.N.G. Caldas. Comportamento fotossintético em videiras syrah e chenin blanc cultivadas em lira e espaldeira no submédio do vale do são francisco. In *Embrapa Semiárido - Artigo em anais de congresso (ALICE)*. In: SIMPOSIO DE FRUTICULTURA DO VALE DO SÃO FRANCISCO, 1., 2015, Juazeiro ..., 2015.
- [57] S. Koundouras, I. T Tsialtas, E. Zioziou, and N. Nikolaou. Rootstock effects on the adaptive strategies of grapevine (*Vitis vinifera* L. cv. cabernet– sauvignon) under contrasting water status: leaf physiological and structural responses. *Agriculture, ecosystems & environment*, 128(1-2):86–96, 2008.
- [58] L. Rossdeutsch. Contribution du métabolisme de l’ABA et de la conductivité hydraulique à la réponse de la transpiration en situation de contrainte hydrique chez la Vigne: variabilité génétique et effets du greffage. PhD thesis, Université de Bordeaux, 2015.
- [59] J. Satisha, G.S. Prakash, G.S.R. Murti, and K.K. Upreti. Response of grape rootstocks to soil moisture stress. *Journal of Horticultural Sciences*, 1(1):19–23, 2006.
- [60] J. Satisha, S.D. Ramteke, J. Sharma, and A.K. Upadhyay. Moisture and salinity stress induced changes in biochemical constituents and water relations of different grape rootstock cultivars. *International Journal of Agronomy*, 2014, 2014.
- [61] R.M. Stevens, Joanne M. Pech, M.R. Gibberd, R.R. Walker, and P.R. Nicholas. Reduced irrigation and rootstock effects on vegetative growth, yield and its components, and leaf physiological responses of Shiraz. *Australian Journal of Grape and Wine Research*, 16(3):413–425, 2010.
- [62] C.R. de Souza, A.M. Soares, and M. de A. Regina. Trocas gasosas de mudas de videira, obtidas por dois porta-enxertos, submetidas à deficiência hídrica. *Pesquisa Agropecuária Brasileira*, 36(10):1221–1230, 2001.
- [63] A. Carbonneau. The early selection of grapevine rootstocks for resistance to drought conditions. *American Journal of Enology and Viticulture*, 36(3):195–198, 1985.

- [64] M. Fregoni, A. Scienza, and R. Miravalle. Evaluation précoce de la résistance des porte-greffes à la secheresse. In Proceedings of the IId Symposium of Genetics and Grape Breeding, volume 2, pages 287–96, 1978.
- [65] M. Borghezán, L.K.A. de Moraes, F.I.M. Moreira, and A.L. da Silva. Propagação in vitro e avaliação de parâmetros morfofisiológicos de porta enxertos de videira. Pesquisa Agropecuária Brasileira, 38(7):783–789, 2003.
- [66] S. Gargin. Determination of stomata densities of some american grape rootstocks in egirdir/isparta. In 2nd International Symposium on Sustainable Development, volume 2, pages 113–116, 2010.
- [67] S.I. Shiraishi and M. Shiraishi. Preliminary survey on stomatal density and length of grapevine. JOURNAL-FACULTY OF AGRICULTURE KYUSHU UNIVERSITY, 41:11–16, 1996.

Imagens complementares do artigo intitulado: “Análise da diversidade genética de porta-enxertos de videira através de aprendizado de máquina”.

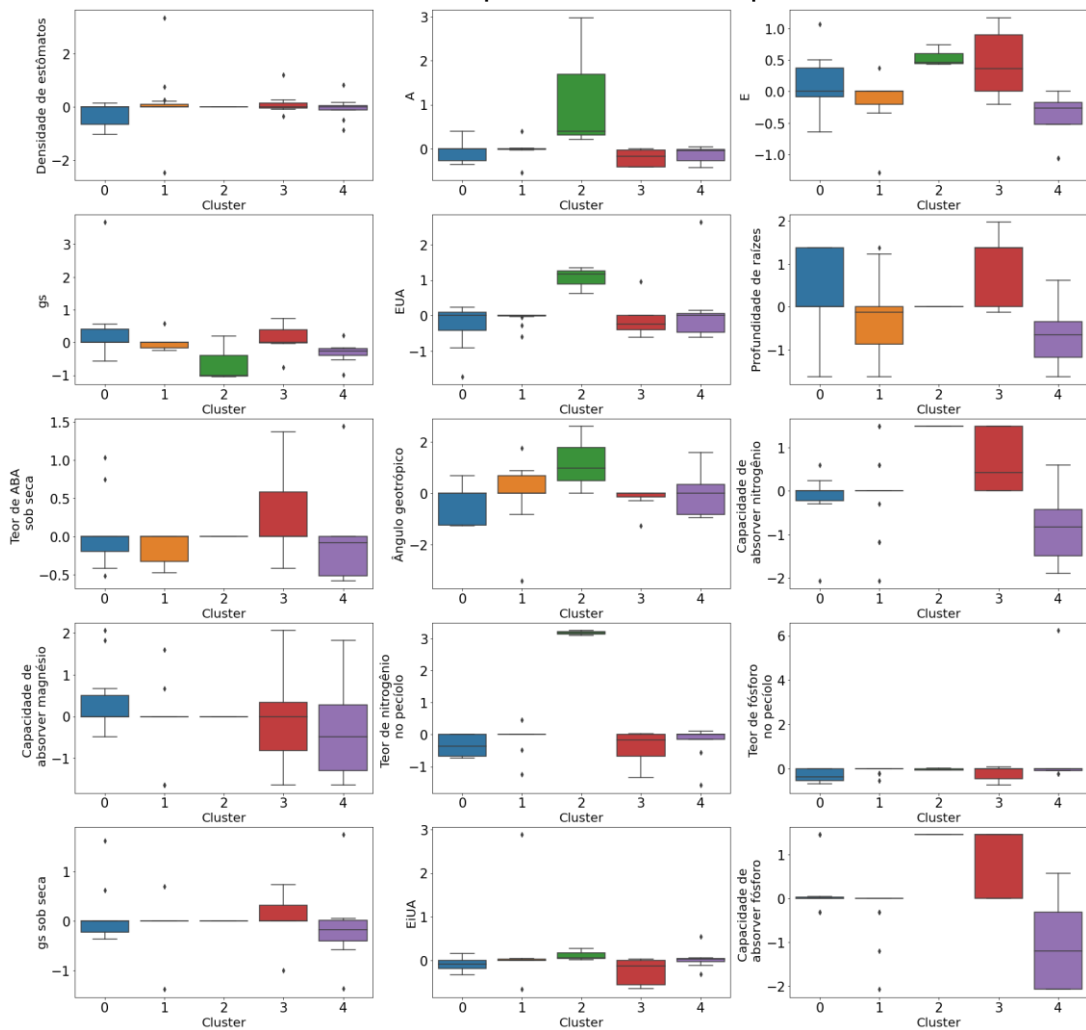


Figura 1 - Distribuição das características agrônômicas dentro de cada cluster. Fonte: Verslype *et al.* (2021).

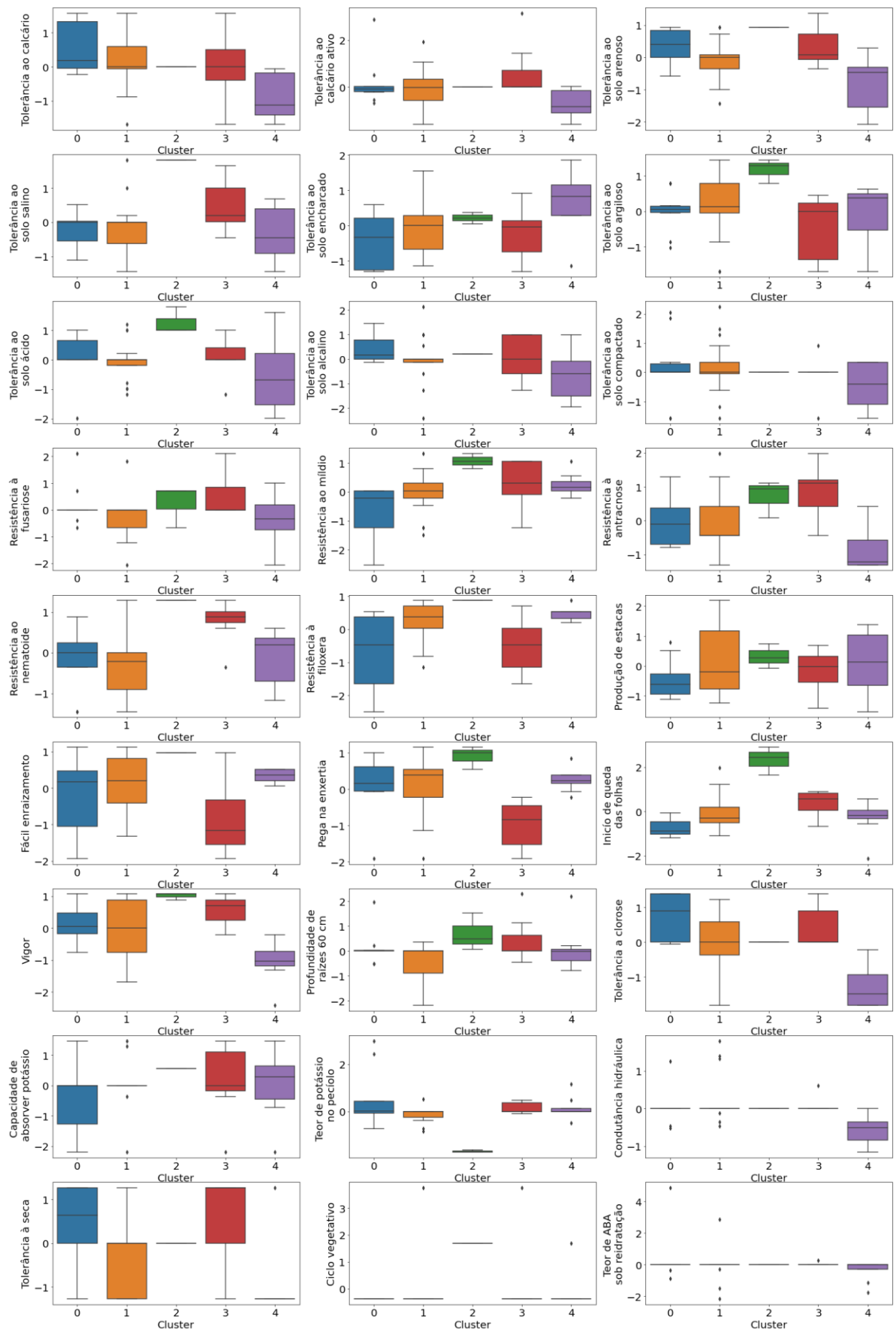


Figura 2 - Distribuição das características agrônômicas dentro de cada cluster. Fonte: Verslype *et al.* (2021).