

III - Artigo Científico

***Emergent Self-Organizing Maps* aplicados ao estudo da diversidade genética de acessos de uva de mesa do Banco Ativo de Germoplasma de videira (*Vitis* spp.) da Embrapa Semiárido**

Raphael Miller de Souza Caldas¹, Rosimar dos Santos Musser², André Câmara Alves do Nascimento² e Patrícia Coelho de Souza Leão³

Resumo

O objetivo deste trabalho foi aplicar a inteligência artificial, por meio do uso do algoritmo de agrupamento *Emergent self-organizing maps* (ESOM), ao estudo da diversidade genética de acessos de uvas de mesa do Banco Ativo de Germoplasma de videira da Embrapa Semiárido. O experimento foi realizado no Campo Experimental de Mandacaru, Juazeiro-BA, durante quatro safras consecutivas (2018 e 2019). O agrupamento realizado pelo algoritmo ESOM permitiu a formação de 10 grupos heteróticos. O grupo 0 apresentou os maiores valores máximos para as variáveis produção (8,44 kg por planta), peso do cacho (520,19 g), comprimento do cacho (22,30 cm), largura do cacho (15,10 cm), peso da baga (8,00 g), comprimento da baga (26,84 mm), diâmetro da baga (22,17 mm) e sólidos solúveis (23,90 °Brix). A presença de variabilidade genética para as variáveis analisadas foi evidenciada pela formação dos mapas de variabilidade genética. A matriz ESOM se mostrou eficiente na análise da diversidade genética e consequente formação de grupos heteróticos, além de indicar cruzamentos mais promissores. Futuras pesquisas sobre a validação dos ESOM como um método de agrupamento eficiente nos estudos de recursos genéticos e melhoramento de plantas são indicados.

Palavras-Chave: *Vitis* spp.; agrupamento; variabilidade genética; algoritmos inteligentes; ESOM.

Abstract

(Emergent Self-Organizing Maps applied to the study of the genetic diversity of table grape accessions from the Active Bank of Vine Germplasm (*Vitis* spp.) of Embrapa Semiárido) The objective of this work was to apply artificial intelligence, through the use of the clustering algorithm Emergent self-organizing maps (ESOM), to the study of the genetic diversity of accessions of table grapes from the Active Bank of Vine Germplasm of Embrapa Semiárido. The experiment was carried out at Campo Experimental de Mandacaru, Juazeiro-BA, during four consecutive harvests (2018 and 2019). The grouping performed by the ESOM algorithm allowed the formation of 10 heterotic groups. Group 0 had the highest maximum values for the variables production per vine (8.44 kg per vine), bunch weight (520.19 g), bunch length (22.30 cm), bunch width (15.10 cm), berry weight (8.00 g), berry length (26.84 mm), berry diameter (22.17 mm) and soluble solids (23.90 °Brix). The presence of genetic variability for the analyzed variables was evidenced by the formation of genetic variability maps. The ESOM matrix proved to be efficient in the analysis of genetic diversity and consequent formation of heterotic groups, in addition to indicating more promising crosses. Future studies on the validation of ESOM as an efficient clustering method in resources genetics studies and plant breeding are indicated.

Keywords: *Vitis* spp.; clustering; genetic variability; smart algorithms; ESOM.

¹Discente de Mestrado Universidade Federal Rural de Pernambuco – UFRPE, Rua Dom Manuel de Medeiros, s/n - Dois Irmãos, 52171-900, Recife - PE, Brasil, raphaelmillers@gmail.com

²Professor Universidade Federal Rural de Pernambuco – UFRPE, Rua Dom Manuel de Medeiros, s/n - Dois Irmãos, 52171-900, Recife - PE, rosimar.musser@ufrpe.br, andre.camara@ufrpe.br

³Pesquisador Embrapa Semiárido. Rodovia BR-428, Km 152, s/n - Zona Rural, Petrolina, PE, 56302-97, Brasil, patricia.leao@embrapa.br.

Introdução

Os estudos de diversidade genética são muito importantes no melhoramento genético, pois permitem a diferenciação e caracterização de acessos vegetais e auxiliam na identificação de genótipos contrastantes para a realização de cruzamentos promissores. Por meio de métodos clássicos tais como x e y, bem como ferramentas biotecnológicas ou novas abordagens, são caracterizadas populações com variabilidade genética e um ou mais indivíduos com caracteres agrônômicos de interesse são agrupados (LEÃO e BORGES 2009; CAMPOS et al. 2016).

Os Bancos Ativos de Germoplasma constituem-se em estratégias valiosas à preservação e à conservação de diversidade e variabilidade genética existente na natureza, que por sua vez, é muito importante para avançar nos programas de melhoramento, porque sem ela não pode haver ganho com a seleção (COSTA et al. 2012, BROWN et al. 2014). A diversidade genética pode ser estimada por meio de várias técnicas estatísticas e a escolha do método mais adequado será determinado de acordo com os objetivos da pesquisa, bem como pela facilidade da análise ou ainda pela forma como os dados são obtidos. Métodos multivariados, como a análise de componentes principais, variáveis canônicas e métodos de agrupamento (k-means, DBSCAN, agrupamento hierárquico da soma de quadrados do erro de Ward, entre outros) têm sido bastante utilizados em estudos de diversidade genética (SILVA et al. 2007; MURTAGH et al. 2011). Recentemente, vários são os estudos disponíveis na literatura que utilizam a inteligência artificial para fins de agrupamento ou classificação e análise de diversidade genética. Gutiérrez et al. (2015) propuseram em seu estudo um novo método de classificação para variedades de videiras através do uso de máquinas de vetor de suporte e redes, utilizando um espectrofotômetro infravermelho portátil nas folhas das plantas em situação de campo. Marques et al. (2019) propuseram um método automático para a classificação de variedades de videira com base nas características das folhas da planta utilizando o aprendizado de máquina. Barbosa et al. (2011) avaliaram a viabilidade de redes neurais artificiais como uma técnica de análise da diversidade genética de *Carica papaya* L., implementando uma rede neural artificial, conforme modelo proposto por Kohonen (1982), na tentativa de propor uma classificação e a formação de grupos divergentes dos acessos

Self-Organizing Maps (SOM) é um tipo de rede neural artificial, bidimensional, que organiza dados de um processo de aprendizagem não supervisionado por meio da noção de vizinhança e usando a distância euclidiana (SANTOS et al. 2019). O processo tem início

com a atribuição de pesos sinápticos, processo que gera competição e cada amostra de dados é alocada para o neurônio que melhor a representa, denominado *Best Matching Unit* (BMU) ou neurônio vencedor. Então, tem início a cooperação, na qual o neurônio vencedor determina a aproximação dos outros neurônios na ordem de proximidade. Finalmente, os neurônios que estabelecem sua vizinhança vão para a fase de adaptação, onde há ajustes de peso. Após todas as iterações, o mapa é organizado em uma estrutura topológica que reflete a proximidade dos elementos em estudo. (SANTOS et al. 2019).

Os SOM são ferramentas de visualização que preservam a topologia dos dados originais, fornecendo uma representação visual de grupos de instâncias de dados semelhantes através da geração de mapas. SOM são compostos por elementos interconectados denominados neurônios e vizinhanças (ULTSCH e MÖRCHEN 2005; KOHONEN 2012). Seu uso como ferramenta para o estudo da diversidade genética tem se tornado cada vez mais frequente em diferentes espécies, como os estudos realizados por Mancuso (2001) que utilizou a rede SOM para realizar o agrupamento de 20 variedades de uvas de vinho com base em imagens das folhas. Estudos de diversidade genética utilizando as redes neurais SOM também foram realizados em arroz (SANTOS et al. 2019), mamão (BARBOSA et al. 2011) e goiaba (CAMPOS et al. 2016).

Por outro lado, *Emergent Self-Organizing Maps* (ESOM), também chamados de mapas auto-organizáveis emergentes, contém um número muito maior de neurônios do que uma rede SOM tradicional. Usando a cooperação de muitos neurônios, os ESOM são capazes de construir estruturas que permitem a visualização de dados de alta dimensão em mapas de baixa dimensão, que, de outra forma, seriam invisíveis. Esse grande número de neurônios pode representar clusters de dados individualmente, o que facilita sua detecção. Dessa forma, uma rede ESOM pode capturar a topologia das dimensões originais com maior acurácia (ULTSCH; KÄMPF 2004, ULTSCH 2007). Até o presente momento, não existem relatos na literatura sobre a utilização dos ESOM para estudos de diversidade genética em cultivos agrícolas.

O Banco Ativo de Germoplasma (BAG) de Videira da Embrapa Semiárido destaca-se por ser o único da região Nordeste do Brasil, constituindo-se um recurso estratégico fundamental para a sustentabilidade da vitivinicultura tropical no Submédio do Vale do São Francisco.

Diversos estudos foram realizados para estimar a diversidade genética presente neste BAG por diferentes métodos de análises estatísticas multivariadas (LEÃO et al. 2011; LEÃO; MOTOIKE 2011; BATISTA et al. 2015). Esses estudos foram desenvolvidos com o objetivo

de desenvolvimento de novas cultivares de uvas de mesa adaptadas ao semiárido brasileiro.

Em geral, os cruzamentos biparentais são os mais utilizados no melhoramento de videira e servem como base para os processos de seleção. A finalidade desses cruzamentos é a obtenção de híbridos através de polinização controlada. O sucesso dos cruzamentos biparentais está relacionado, entre outros fatores, à escolha dos genitores (LEÃO e BORGES 2009).

O objetivo deste estudo foi apresentar a técnica de ESOM como método alternativo para o estudo da diversidade genética em programas de melhoramento genético de plantas, através da sua aplicação para análise da diversidade genética de acessos de uva de mesa do BAG da Embrapa Semiárido.

Material e Métodos

O BAG de Videira da Embrapa Semiárido está localizado no Campo Experimental de Mandacaru, Juazeiro-BA (9°24'S, 40°26'O e 365,5 m de altitude), Brasil. As plantas são conduzidas em sistema de condução tipo espaldeira, em espaçamento de 3,0 metros entre filas de plantio e 2,0 metros entre plantas (adicione o número de plantas por ha), com irrigação localizada por gotejamento. Cada acesso é composto por quatro plantas, não havendo um delineamento experimental em cordão bilateral. As plantas são podadas em duas épocas ao longo do ano, sendo uma poda curta no 1º semestre e uma poda média com 6 a 8 gemas no 2º semestre. O manejo da copa consiste na poda, desbrota e desponte de ramos, além do controle de plantas daninhas com aplicação de herbicida, roçado mecanizado e controle químico de pragas e doenças.

O estudo foi realizado durante os anos de 2018 e 2019, utilizando as médias de quatro ciclos de produção (2018.1, 2018.2, 2019.1 e 2019.2) de 93 acessos de uva de mesa de *Vitis vinifera* L. e híbridos interespecíficos (Tabela 1).

Foram analisadas as seguintes variáveis agronômicas quantitativas:

a. Produção: avaliada na colheita por meio da pesagem de todos os cachos colhidos utilizando em balança eletrônica digital, expressa em quilogramas (kg); b. Número de cachos por planta: obtido pela contagem de todos os cachos; c. Peso dos cachos: determinado pela divisão do peso total dos cachos pelo número de cachos por videira, expresso em gramas (g); d. Comprimento e largura do cacho: medidos em uma amostra de cinco cachos por acesso, por meio de régua e expressos em centímetros (cm); e. Peso da baga: determinado numa amostra de 10 bagas colhidas em cada cacho, num total de 50 bagas por acesso, através de balança eletrônica digital, expresso em gramas (g); f. Comprimento e diâmetro das bagas: avaliados na mesma amostra de bagas do item

anterior, por meio de régua, expressos em milímetros (mm); g. Teor de sólidos solúveis totais: leituras feitas no mosto extraído de 50 bagas por planta em refratômetro digital com ajuste automático de temperatura e expresso em °Brix; h. Acidez titulável: através da diluição de 5 ml de polpa de uva em 50 mL de água destilada juntamente com solução de hidróxido de sódio (NaOH) 0,1 N utilizando um titulador automático (AOAC, 2010), e os resultados foram expressos em g de ácido tartárico.100 mL⁻¹ e i. Relação teor de sólidos solúveis totais/acidez titulável.

Para avaliar a diversidade genética de 93 acessos de uva de mesa do BAG de Videira da Embrapa Semiárido, foram gerados ESOM de configuração matricial de 16.000 neurônios (100 linhas x 160 colunas). Para a geração dos ESOM, os dados foram padronizados e foi utilizado o pacote Somoclu (WITTEK et al. 2013) em linguagem de programação *Python*.

O agrupamento formado pela ESOM pode ser executado da seguinte forma: os *Best Matching Units* (BMU) são identificados e seus pontos de dados correspondentes podem ser agrupados manualmente em vários grupos. A associação dos grupos pode ser visualizada pela coloração dos BMU. Ou seja, os BMU que possuem cores iguais pertencem ao mesmo grupo (ULTSCH e MÖRCHEN, 2005).

Foram gerados onze *boxplots* a fim de facilitar a discussão ao permitir a visualização dos valores médios, máximos, mínimos para cada variável analisada, bem como os *outliers* de cada grupo.

Também foram gerados os mapas de variabilidade genética, que permitem visualizar a variabilidade genética existente entre os grupos formados para cada variável analisada. A interpretação dos mapas de variabilidade genética é bastante simples, sendo análoga à interpretação de um mapa de calor: o tom em vermelho mais escuro representa o valor de maior magnitude, enquanto o tom em azul mais escuro representa o valor de menor magnitude. Os valores de variação dentro desse intervalo são representados por cores intermediárias aos tons vermelho/azul escuro. Dentro das regiões coloridas, que representam os acessos, existem pontos de cor branca e seus respectivos rótulos numéricos de identificação. Como os mapas gerados têm a forma de um retângulo, deve-se perceber que o quadrilátero inferior é uma continuação do quadrilátero superior ou vice versa. O mesmo vale para o quadrilátero da direita e da esquerda.

Por fim, foi gerada a matriz ESOM de similaridade genética, que indica quais os cruzamentos mais promissores entre os grupos heteróticos formados pela rede ESOM com base na distância Euclidiana entre os grupos.

Resultados e Discussão

Agrupamento da rede ESOM

O mapa de agrupamento de xxx foi representado na Figura 1. Cada ponto colorido inserido no mapa corresponde a um genótipo que possui sua respectiva identificação numérica. Pontos da mesma cor representam os genótipos que pertencem ao mesmo grupo genético, ou seja, genótipos similares geneticamente de acordo com as variáveis analisadas e, portanto, é possível identificar os grupos formados pela rede ESOM (Figura 1).

A porcentagem de acessos alocados em cada grupo foi: 20,43 % no grupo 0; 11,82 % no grupo 1; 26,88 % no grupo 2; 11,82 % no grupo 3; 6,45 % no grupo 4; 5,37 % no grupo 5; 1,07 % no grupo 6; 6,45 % no grupo 7; 1,07 % no grupo 8 e 8,60 % no grupo 9.

Utilizando a análise de componentes principais para analisar a diversidade genética em 136 acessos de uva de mesa pertencentes ao BAG da Embrapa Semiárido, LEÃO (2008) não conseguiu identificar uma tendência na formação dos grupos, tais como uma característica comum, genealogia ou origem geográfica GUMBER (1998). De acordo com CAMPOS et al. (2016), a formação de grupos heteróticos facilita a seleção de genótipos divergentes para o melhoramento por meio da geração de híbridos, visto que permite a seleção de genótipos adequados para cruzamentos entre diferentes grupos heteróticos, aumentando as chances de obtenção de genótipos superiores.

Os grupos heteróticos têm forte impacto no melhoramento das culturas que apresentam o fenômeno da heterose; eles determinam o tipo de germoplasma que será usado em um programa de melhoramento genético para seleção de híbridos por um longo período de tempo (MELCHINGER e GUMBER 1998). LEÃO et al. (2011) afirmam que, em relação à videira, uma espécie altamente heterozigótica, espera-se obter o efeito heterótico máximo nas gerações segregantes a partir do cruzamento de pais divergentes.

Semelhante aos resultados obtidos por LEÃO (2008), foi observado a formação de um grupo composto por clones da cultivar Itália, de bagas vermelhas como 'Benitaka' e 'Brasil', que possuem teores de sólidos solúveis e sabor moscatel mais acentuado que a cultivar Itália, bem como as cultivares 'Itália Melhorada' e 'Itália Clone 1', todas inseridas no grupo 0. Nesse grupo também predominaram cultivares que se destacam pelo tamanho de suas bagas: 'Dona Maria', 'Queen' e 'Red Globe'. Também foram inseridas no grupo 0 as cultivares 'BRS Linda' e 'BRS Isis'; duas importantes cultivares comerciais desenvolvidas pela Embrapa.

O grupo 3, formado por dois acessos ('BRS Clara' e 'BRS Morena'), se destacou por possuir o maior valor para as variáveis produção por planta (3,97

No grupo 7 foram agrupadas duas importantes cultivares comerciais que representam o passado e presente do melhoramento genético das uvas de mesa: a cultivar 'Itália', uma das primeiras cultivares comerciais do Submédio Vale do São Francisco, e a cultivar 'BRS Vitória', uma das cultivares mais comercializadas no Submédio do Vale do São Francisco atualmente.

As cultivares 'Aurora' e 'IAC 77526', consideradas sinonímias, foram alocadas no grupo 2. LEÃO (2008) obteve resultado semelhante com relação a essas duas cultivares, porém, o mesmo resultado não foi obtido com relação a outras cultivares ('Dattier de Saint Vallier', 'Seyve Villard 20365', 'Emperatriz' e 'CG 28467'). Esses resultados indicam que a técnica multivariada de agrupamento utilizando variáveis morfoagronômicas não foi eficiente para agrupar genótipos idênticos. No presente estudo, não foi possível concluir se o agrupamento realizado pela rede ESOM foi eficiente ao agrupar sinonímias devido ao pequeno número de sinonímias utilizadas na análise (apenas duas).

Mapas de variabilidade genética

Os ESOM de variabilidade genética dos grupos formados também foram gerados e podem ser visualizados na Figura 3.

As variáveis que apresentam maior variabilidade foram a produção por planta, o número de cachos por planta, o peso e o comprimento da baga, o teor de sólidos solúveis e a acidez titulável (Figura 2).

Em um estudo de diversidade genética deste mesmo BAG, LEÃO (2008) também concluiu que o número de grupos formados e a distribuição satisfatória dos acessos nestes grupos possibilitou a identificação de genitores para a formação de populações segregantes com ampla base genética.

Matriz ESOM de similaridade genética

A matriz ESOM de similaridade genética (Figura 4) indica quais os cruzamentos mais promissores entre os 10 grupos heteróticos de uva de mesa com base na divergência genética. Características como apirenia (ausência de sementes), aparência do cacho, sabor da baga (neutro, moscatel, especial e foxado), baixo desgrane, consistência e textura da polpa, e resistência pós-colheita, são as de maior interesse (LEÃO e BORGES, 2009), sendo consideradas características que norteiam os programas de melhoramento genético da videira.

Segundo LEÃO (2008), a seleção de genótipos com base somente na divergência genética, sem considerar o desempenho produtivo e qualitativo de cada acesso individualmente, não é uma boa estratégia em um programa de melhoramento. Dessa forma, a recomendação de cruzamentos em programas de melhoramento entre genótipos divergentes, mas que

apresentem desempenho superior nas principais características de importância agrônômica é a mais indicada. O cruzamento de genótipos do grupo 0, característico por possuir genótipos com bagas de tamanho grande, com genótipos sem semente do grupo 1, são indicados. O genótipo ‘Flame Seedless’ pode ser utilizado como progenitor masculino e cruzado com o genótipo ‘Itália Melhorada’, utilizado como progenitor feminino. Também foram indicados cruzamentos entre os genótipos ‘Reliance’ e ‘Júpiter’ (grupo 1), utilizados como progenitores masculinos, com os genótipos ‘Itália Clone 1’ e ‘Michele Paglieri’ (grupo 0), utilizados como progenitores femininos.

O grupo 0 também apresentou divergência genética em relação aos grupos 2, 4 e 5. Os genótipos sem sementes e de bagas pequenas ‘Emerald Seedless’, ‘CG351’ e ‘Marroo Seedless’ (grupo 2), podem ser usados como progenitores masculinos e cruzados com os genótipos ‘Dona Maria’ e ‘Monte Serrat’ (grupo 0), que seriam os progenitores femininos. Os genótipos sem sementes do grupo 5, ‘Lakemont Seedless’, ‘Ruby Seedless’ e ‘Fiesta’ (progenitores masculinos) podem ser cruzados com ‘Itália Melhorada’, ‘Dona Maria’ e ‘Michele Paglieri’ (progenitores femininos) do grupo 0.

O grupo 7, no qual predominaram acessos de uvas sem sementes como a cultivar ‘BRS Vitória’, pode ser utilizado em cruzamentos com os grupos 2, 3 e 9.

Cruzamentos entre a ‘BRS Vitória’, progenitor masculino e ‘Moscatel de Hamburgo’ ou ‘Moscatel de Alexandria’ (grupo 3, progenitores femininos) são indicados, visto que pode-se aliar o sabor moscatel ao maior tamanho de bagas. Os grupos 0 e 9 são os mais distantes geneticamente. Porém, considerando que o objetivo principal do programa de melhoramento genético para uvas de mesa é a obtenção de novas cultivares apirênicas, seus cruzamentos não são indicados. Todos os genótipos alocados no grupo 9 possuem sementes e bagas de tamanho pequeno.

Os resultados obtidos demonstram a eficiência do método de redes neurais artificiais por meio de ESOM como alternativa para os estudos de divergência genética, sendo especialmente úteis para a análise de grandes quantidades de dados.

Conclusão

A matriz ESOM se mostrou eficiente na análise da diversidade genética e consequente formação de grupos heteróticos, além de indicar cruzamentos mais promissores no melhoramento genético de uvas de mesa. Estudos futuros sobre a validação dos ESOM como um método de agrupamento eficiente no melhoramento genético de plantas são indicados.

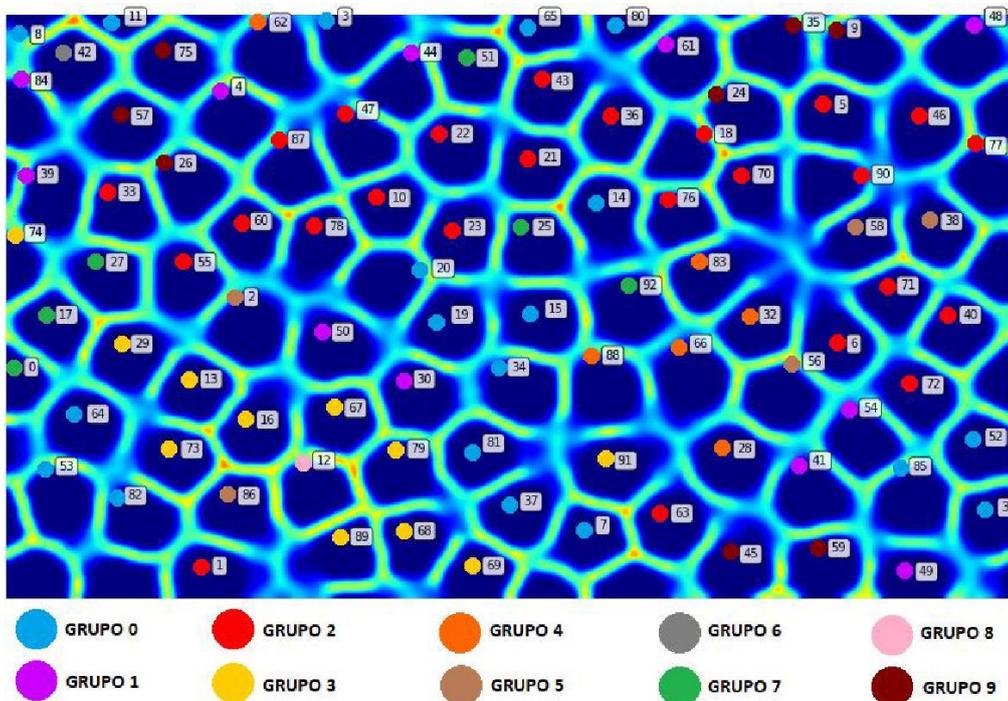
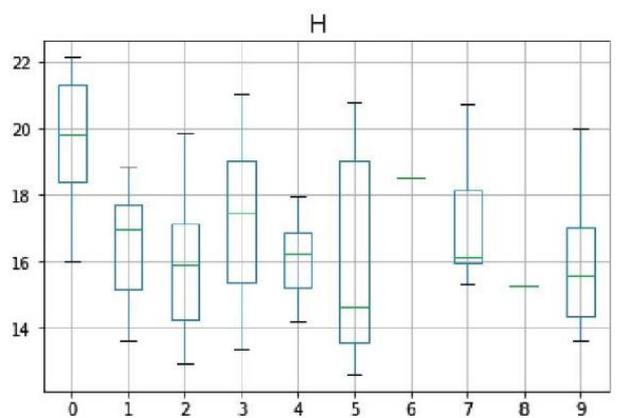
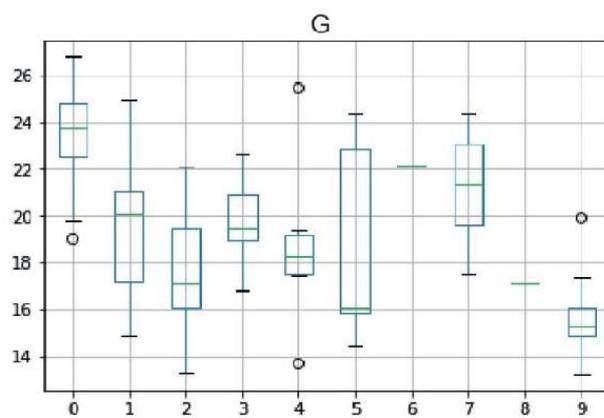
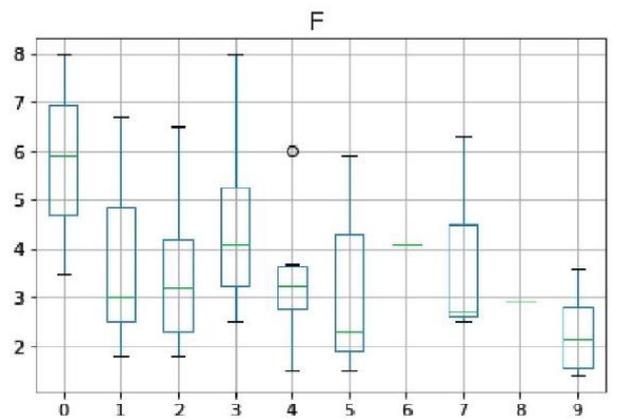
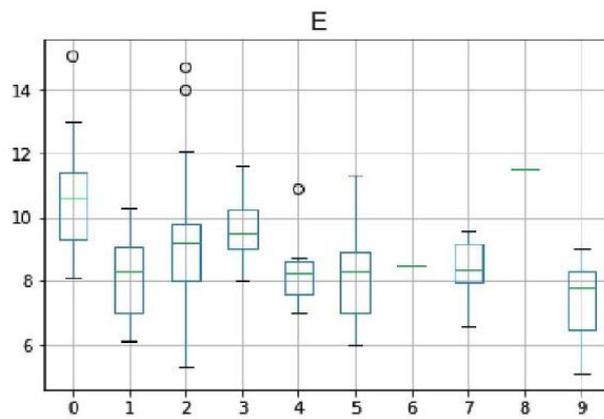
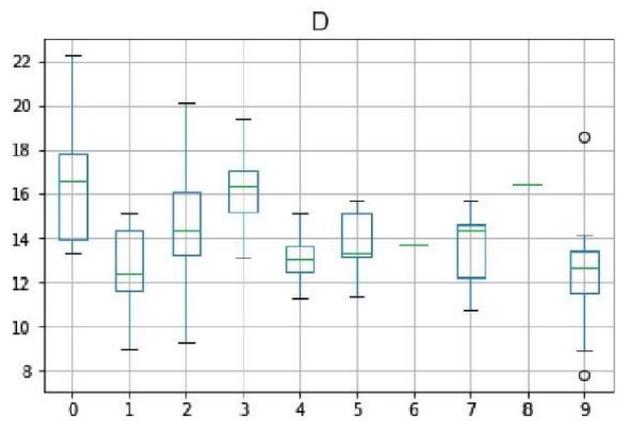
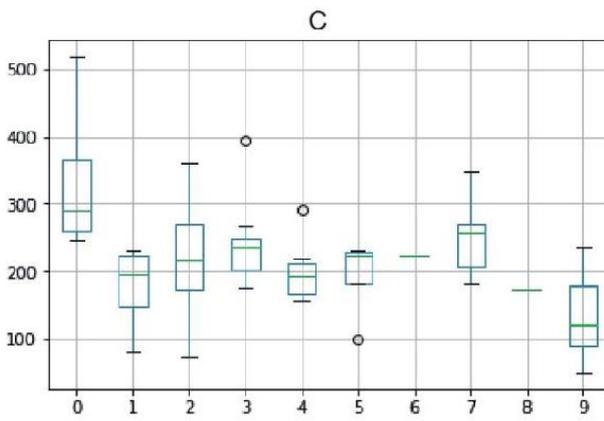
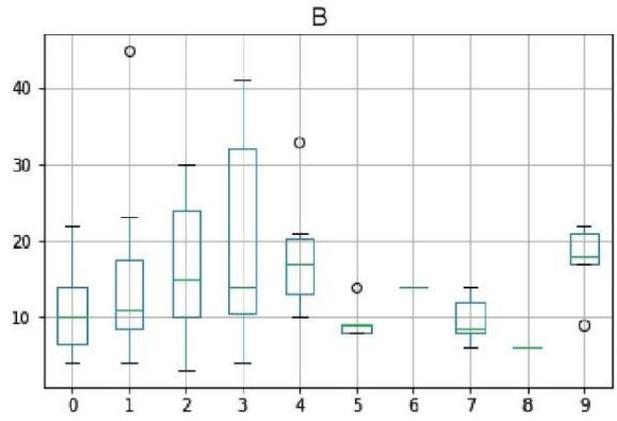
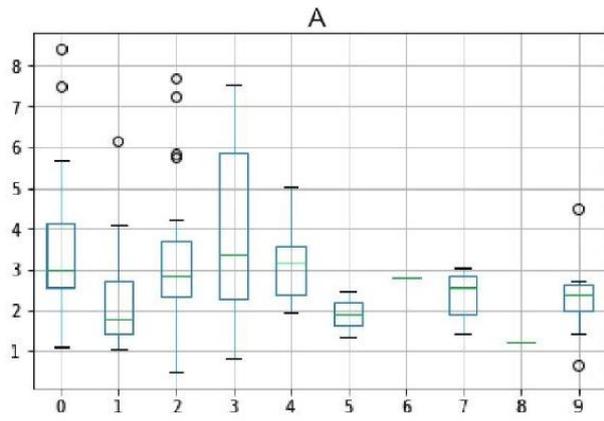


Figura 1. Emergent self-organizing map do agrupamento de acessos de videira realizado pela rede neural.

Tabela 1. Noventa e três genótipos de uvas de mesa (*Vitis vinifera* L. e híbridos interespecíficos) e códigos correspondentes utilizados no estudo.

Genótipo	Códig o	Genótipo	Códig o	Genótipo	Códig o	Genótipo	Códig o
A1105	0	Christmas Rose	24	IAC Madalena	47	Moscatel de Jundiaí	70
A1118	1	CNPUV 8	25	Igawa	48	Moscatel Grega	71
A1581	2	Concord Clone	26	Impero	49	Moscatel Nazareno	72
Ângelo Pirovano	3	Crimson Seedless	27	Isaura	50	Muscat Noir	73
August Giant	4	Dattier de Beirut	28	Itália	51	Neo Muscat	74
Aurora	5	Dawn Seedless	29	Itália Clone I	52	Niágara Rosada	75
Baresana	6	Don Mariano	30	Itália Melhorada	53	Patrícia	76
Benifuji	7	Dona Maria	31	Júpiter	54	Perlette	77
Benitaka	8	Early Muscat	32	Kagina	55	Perlona	78
Blue Lake	9	Emerald Sedless	33	Kyoho	56	Piratinunga	79
Branca Salitre	10	Estevão Marinho	34	Lake Emerald	57	Portuguesa Blanes	80
Brasil	11	Eumelan	35	Lakemont Seedless	58	Queen	81
Bronx Seedless	12	Feal	36	Liberty	59	Red Globe	82
BRS Clara	13	Ferral	37	Madeleine Royal	60	Regina Roma	83
BRS Isis	14	Fiesta	38	Marengo	61	Reliance	84
BRS Linda	15	Flame Seedless	39	Maria	62	Rosaky Rosada	85
BRS Morena	16	Flame Tokay	40	Marroo Seedless	63	Ruby Seedless	86
BRS Vitória	17	Frakenthal	41	Michele Paglieri	64	Seyve Villard 12327	87
Califórnia	18	Golden Queen	42	Monte Serrat	65	Soraya	88
Ceilad	19	H 449100	43	Moscatel Branca	66	Sovrana Pirovano	89
Centenial Seedless	20	Himoront	44	Moscatel Caillaba	67	Tardia de Caxias	90
CG 351	21	IAC 13822	45	Moscatel de Alexandria	68	Thompson Seedless	91
CG 38049	22	IAC 77526	46	Moscatel de Hamburgo	69	Vênus	92
CG 4113	23						



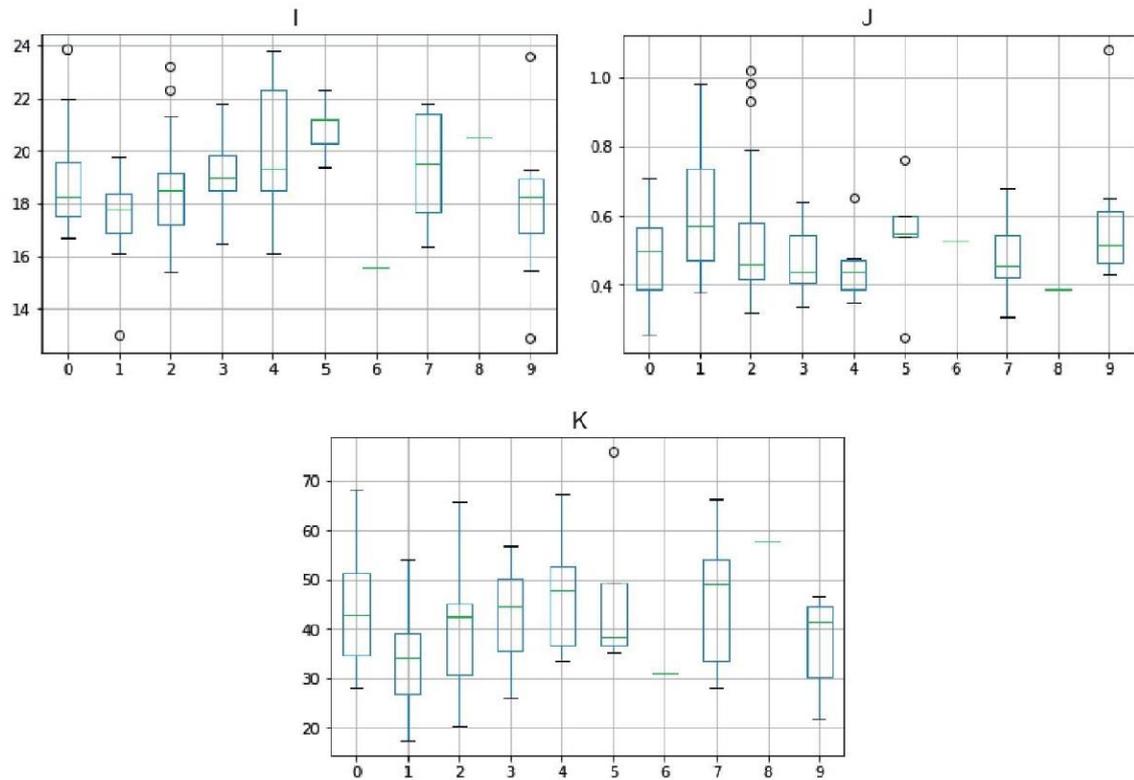
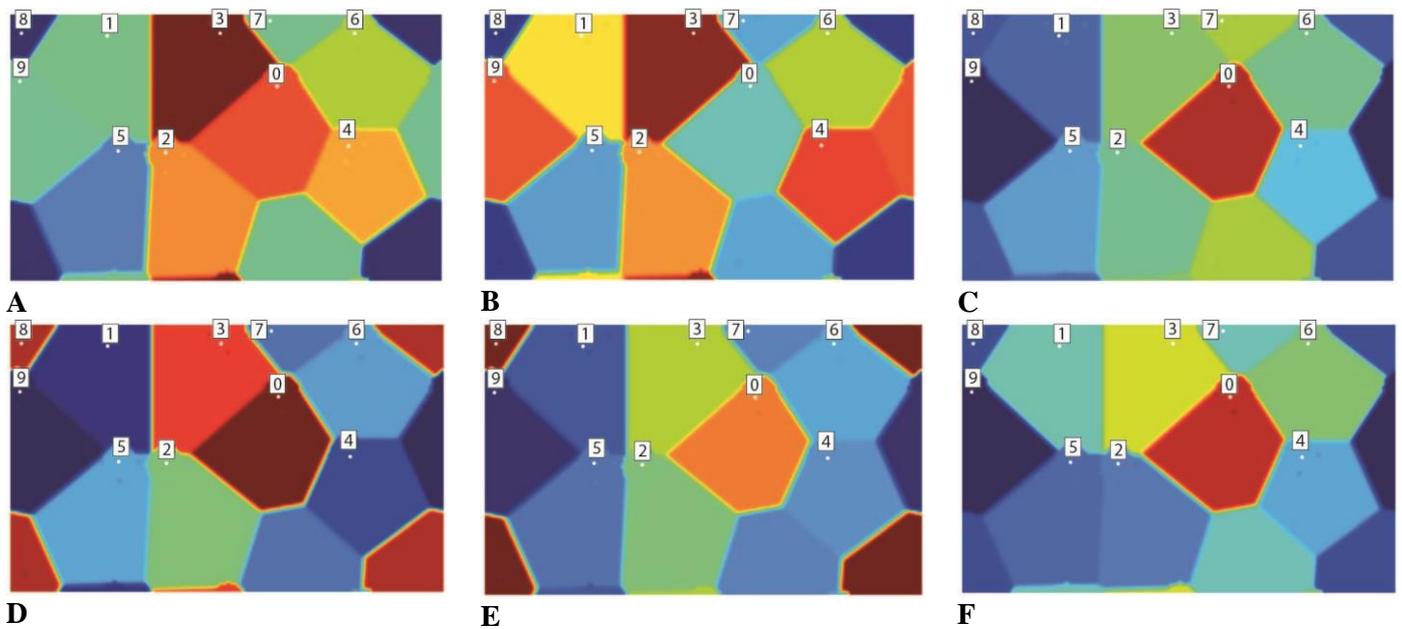


Figura 2. *Boxplots* formados pelas variáveis produção por planta (kg) (A), número de cachos por planta (B), peso do cacho (g) (C), comprimento do cacho (cm) (D), largura do cacho (cm) (E), peso da baga (g) (F), comprimento da baga (mm) (G) diâmetro da baga (mm) (H), teor de sólidos solúveis (°Brix) (I), acidez titulável (g.100mL⁻¹) (J) e relação sólidos solúveis/acidez titulável (K) para os 10 grupos formados pelos *Emergent self-organizing maps* utilizando 93 genótipos de mesa.



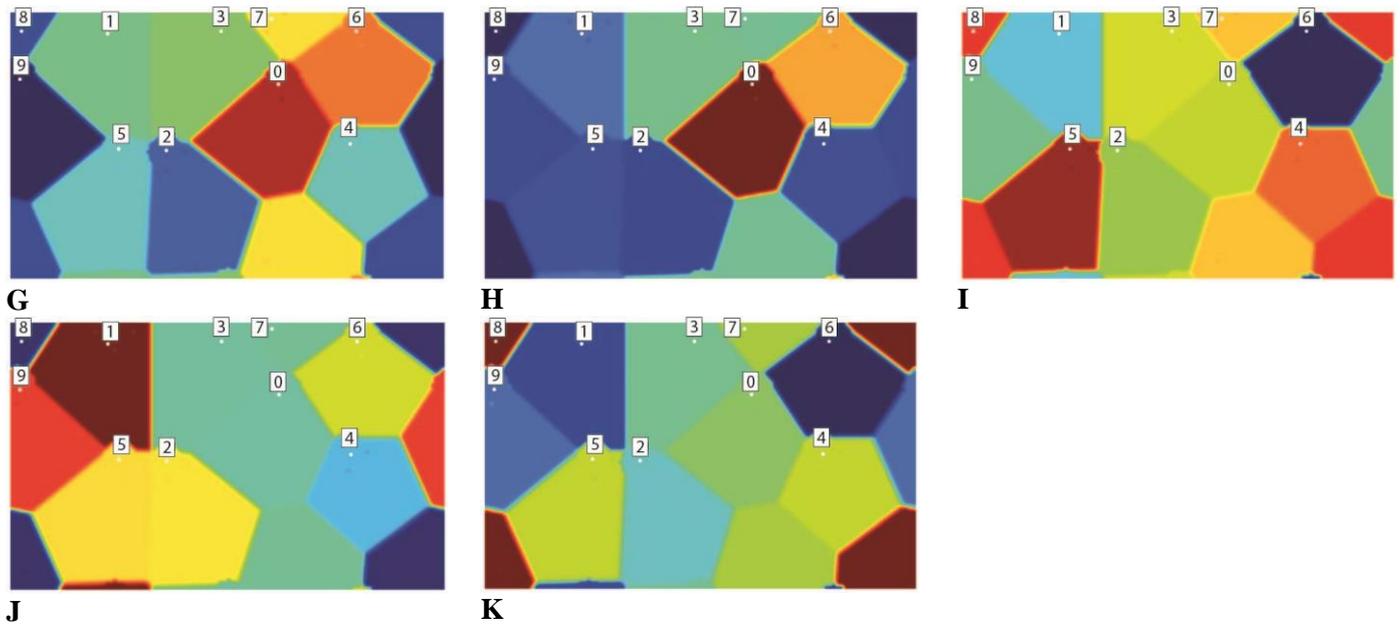


Figura 3. *Emergent self-organizing maps* das variáveis produção (A), número de cachos (B), peso do cacho (C) e comprimento do cacho (D), largura do cacho (E), peso da baga (F), diâmetro da baga (G), comprimento da baga (H), sólidos solúveis (I), acidez titulável (J) e relação sólidos solúveis / acidez titulável (K) para os 10 grupos heteróticos formados pela análise de 93 genótipos de uvas de mesa.

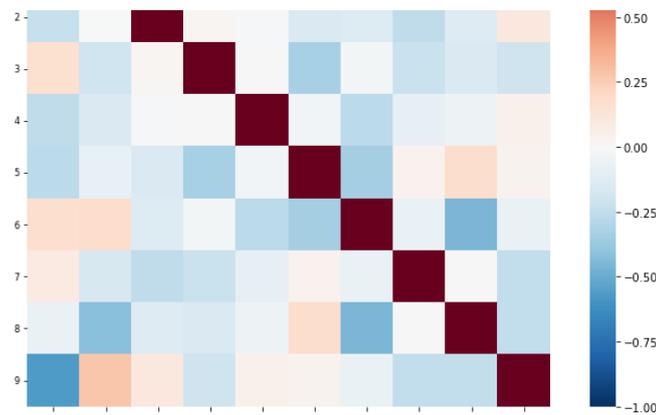


Figura 4. Matriz *Emergent self-organizing maps* de similaridade genética de 93 genótipos de uvas de mesa

Referências

AOAC. Association of Official Agricultural Chemists. **Official methods of analysis of the Association of the Agricultural Chemists**. 18. ed. Gaithersburg: AOAC. 2010, 1025p.

BARBOSA, C. D.; VIANA, A. P.; QUINTAL, S. S. R.; PEREIRA, M.G. Artificial neural network analysis of genetic diversity in *Carica papaya* L. **Crop Breeding and Applied Biotechnology**, v. 11, p. 224-231, 2011.

BORGES, R. M. E.; GONÇALVES, N. P. S.; GOMES, A. P. O.; ALVES, E. O. Divergência fenotípica entre acessos de uvas de mesa no Semi-Árido brasileiro. **Pesquisa Agropecuária Brasileira**, v. 43, p. 1025-1030, 2008.

BROWN, J.; CALIGARI, P.; CAMPOS, H. **Plant breeding**. Wiley-Blackwell, West Sussex, 2014, 296p.

- CAMPOS, B.; VIANA, A. P.; QUINTAL, S. S. R.; BARBOSA, C. D.; DAHER, R. F. Heterotic group formation in *Psidium guajava* L. by artificial neural network and discriminant analysis. **Revista Brasileira de Fruticultura**, v. 38, p. 151-157, 2016.
- COSTA, A. M.; SPEHAR, C. R., SERENO, J. R. **Conservação de recursos genéticos no Brasil**. Embrapa, Brasília, 2012, 628p.
- GUTIÉRREZ, S.; TARDAGUILA, J.; FERNÁNDEZ-NOVALES, J.; DIAGO, M. P. Support vector machine and artificial neural network models for the classification of grapevine varieties using a portable NIR spectrophotometer. **Plos One**, v. 10, e0143197, 2015.
- KOHONEN, T. **Self-Organizing Maps**. Springer Science & Business Media, Cham, 2012, 373p.
- LEÃO, P. C. S.; CRUZ, C. D.; MOTOIKE, S. Y. Genetic diversity of a Brazilian wine grape germplasm collection based on morphoagronomic traits. **Revista Brasileira de Fruticultura**, v. 32, p. 1164-1172, 2010.
- LEÃO, P.C. S. **Recursos Genéticos de Videira (*Vitis* spp.): análise da diversidade e caracterização da coleção de germoplasma da Embrapa Semiárido**. Tese de Doutorado, 126f. Universidade Federal de Viçosa. Genética e Melhoramento de Plantas, Viçosa-MG, 2008.
- LEÃO, P. C. S.; CRUZ, C. D.; MOTOIKE, S. Y. Genetic diversity of table grape based on morphoagronomic traits. **Scientia Agricola** v. 68, p. 42-49, 2011.
- LEÃO, P.C. S; Borges, R. M. E. **Melhoramento Genético da Videira**. Embrapa Semiárido-Documentos, 2009. Disponível em: < <https://www.infoteca.cnptia.embrapa.br/bitstream/doc/748878/1/SDC224.pdf>>. Acesso em: 15 abr. 2020.
- MANCUSO, S. Clustering of grapevine (*Vitis vinifera* L.) genotypes with Kohonen neural networks. **Vitis-Geilweilerhof**, v. 40, p. 59-64, 2001.
- MARQUES, P.; PÁDUA, L.; ADÃO, T.; HRUŠKA, J.; SOUSA, J.; PERES, E.; SOUSA, J. J.; MORAIS, R.; SOUSA, A. **Grapevine Varieties Classification Using Machine Learning**. EPIA Conference on Artificial Intelligence. Springer, Cham, p.186-199, 2019.
- MELCHINGER, A. E.; GUMBER, R. K. Overview of heterosis and heterotic groups in agronomic crops. **Concepts and Breeding of Heterosis in Crop Plants**, v. 25, p. 29-44, 1998.
- MURTAGH, F.; LEGENDRE, P. Ward's hierarchical clustering method: Clustering criterion and agglomerative algorithm. **arXiv**: v. 11, p. 62-85, 2011.
- NUNES, B. T. G.; REGO, J. I. S.; NASCIMENTO, J. H. B.; SOUZA, E. M. C.; LEÃO, P.C. S. Divergência fenotípica de uvas de mesa utilizando caracteres morfo-agronômicos no Vale do São Francisco. In: II Simpósio da Rede de Recursos Genéticos Vegetais do Nordeste, 2015, Fortaleza. Anais do II Simpósio da RGV Nordeste. Fortaleza, Embrapa Agroindústria Tropical., 2015.
- OLIVEIRA, M. S.; SANTOS, I. G., CRUZ, C. D. Self-organizing maps: a powerful tool for capturing genetic diversity patterns of populations. **Euphytica**, v. 216, p. 1-9, 2020.
- SANTOS, I. G. D.; CARNEIRO, V. Q.; SILVA JUNIOR, A. C. D.; CRUZ, C. D.; SOARES, P. C. Self-organizing maps in the study of genetic diversity among irrigated rice genotypes. **Acta Scientiarum Agronomy**, v. 41, p. 1-9, 2019.

SILVA, C. M.; KARASAWA, M. M. G.; VENCOVSKY, R.; VEASEY, E. A. Elevada diversidade genética interpopulacional em *Oryza glumaepatula* Steud. (Poaceae) avaliada com microssatélites. **Biota Neotropica**, v. 7, p. 165-171, 2007.

ULTSCH, A Emergence in self organizing feature maps. In: International Workshop on Self-Organizing Maps: Proceedings. Data Bionics Research Group of University of Marburg, 2007, 7p.

ULTSCH, A.; KAMPF, D. Knowledge discovery in DNA microarray data of cancer patients with emergent self organizing maps. In: ESANN. Belgium, Bruges, p. 501-506, 2004.

ULTSCH, A.; MÖRCHEN, F. ESOM-Maps: Tools for clustering, visualization, and classification with emergent SOM. Technical Report - Data Bionics Research Group of University of Marburg: 1-7, 2005.

WITTEK, P.; GAO, S. C.; LIM, I. S.; ZHAO, L. Somoclu: An efficient parallel library for self-organizing maps. **arXiv**, v. 13, p. 14-22, 2013.